UNIVERSITY OF LJUBLJANA
FACULTY OF MATHEMATICS AND PHYSICS
DEPARTMENT OF MATHEMATICS

Dejan Govc

UNIMODAL CATEGORY

Doctoral thesis

Advisor: prof. dr. Dušan Repovš
Co-Advisor: doc. dr. Primož Škraba

Ljubljana, 2017

UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO
ODDELEK ZA MATEMATIKO

Dejan Govc

UNIMODALNA KATEGORIJA

Doktorska disertacija

Mentor: prof. dr. Dušan Repovš
Somentor: doc. dr. Primož Škraba

Ljubljana, 2017

Abstract

In this thesis, we completely characterize the unimodal category for functions $f : \mathbb{R} \to [0, \infty)$ using a decomposition theorem obtained by generalizing the sweeping algorithm of Baryshnikov and Ghrist. We also give a characterization of the unimodal category for functions $f : S^1 \to [0, \infty)$ and provide an algorithm to compute the unimodal category of such a function in the case of finitely many critical points.

We then turn to the monotonicity conjecture of Baryshnikov and Ghrist. We show that this conjecture is true for functions on $\mathbb{R}$ and S^1 using the above characterizations and that it is false on certain graphs and on the Euclidean plane by providing explicit counterexamples. We also show that it holds for functions on the Euclidean plane whose Morse-Smale graph is a tree using a result of Hickok, Villatoro and Wang. We then present several open questions indicating promising research directions.

After this, we prove an approximate nerve theorem, which is a generalization of the nerve theorem from classical algebraic topology to the context of persistent homology. This is done by introducing the notion of an ε-acyclic cover of a filtered space. We use spectral sequences to relate the persistent homologies of the various spaces involved. The approximation is stated in terms of the interleaving distance between persistence modules. To obtain a tight bound, the technical notions of left and right interleavings are introduced. Finally, examples are provided, which realize the bound and thus prove the tightness of the result.

Math. Subj. Class. (2010): 55, 55M30, 55M99, 55T, 18
Keywords: unimodal category, monotonicity, counterexample, bounded variation, persistence module, approximation, Mayer-Vietoris, spectral sequence

Povzetek

V tej disertaciji popolnoma karakteriziramo unimodalno kategorijo funkcij $f : \mathbb{R} \to [0, \infty)$ s pomočjo izreka o dekompoziciji, ki ga dobimo kot posplošitev algoritma s pometanjem, ki sta ga vpeljala Baryshnikov in Ghrist. Podamo tudi karakterizacijo unimodalne kategorije za funkcije $f : S^1 \to [0, \infty)$ in od tod dobimo algoritem za izračun unimodalne kategorije take funkcije v primeru, ko ima le končno mnogo kritičnih točk.

Nato obravnavamo domnevo Baryshnikova in Ghrista o monotonosti. Pokažemo, da ta domneva drži za funkcije na $\mathbb{R}$ in S^1 s pomočjo zgornjih karakterizacij, in da ne drži za funkcije na določenih grafih in na evklidski ravnini, tako da konstruiramo eksplicitne protiprimere. Poleg tega pokažemo, da drži za funkcije na evklidski ravnini, katerih Morse-Smaleov graf je drevo, z uporabo rezultata, ki so ga dokazali Hickok, Villatoro in Wang. Nato predstavimo nekaj odprtih vprašanj, ki nakazujejo obetavne smeri raziskovanja.

Potem dokažemo še aproksimativni izrek o živcu, ki je posplošitev izreka o živcu iz klasične algebraične topologije v kontekst vztrajne homologije. To storimo z vpeljavo pojma ε-acikličnega pokritja filtriranega prostora. Z uporabo spektralnih zaporedij povežemo vztrajne homologije raznih prostorov, na katere pri tem naletimo. Aproksimacija je podana v jeziku prepletne razdalje med vztrajnostnimi moduli. Da dobimo optimalne meje, vpeljemo tehnična pojma levih in desnih prepletanj. Nazadnje podamo še primere, kjer so meje realizirane in s tem dokažemo optimalnost rezultata.

Math. Subj. Class. (2010): 55, 55M30, 55M99, 55T, 18
Ključne besede: unimodalna kategorija, monotonost, protiprimer, omejena variacija, vztrajnostni modul, aproksimacija, Mayer-Vietoris, spektralno zaporedje

Contents

Zahvala/Acknowledgments

Rad bi se zahvalil vsem, ki so kakorkoli pripomogli pri nastajanju te doktorske disertacije.

Pri tem bi rad posebej izpostavil mentorja prof. dr. Dušana Repovša in somentorja doc. dr. Primoža Škrabo.

Mentorju bi se rad zahvalil za pozicijo mladega raziskovalca, ki mi jo je omogočil, za hitro in natančno branje predhodnih verzij disertacije in za vse konstruktivne pripombe in nasvete, ki jih je ob tem podal, ter za spremljanje mojih predavanj na Seminarju za Geometrijsko Topologijo.

Somentorju bi se rad zahvalil za problem, ki mi ga je predlagal in ki je prerasel v pričujočo doktorsko disertacijo, ki jo je prav tako hitro in natančno prebral. Še bolj pa bi se mu rad zahvalil za prijateljski odnos in za vsa leta tesnega sodelovanja, med katerimi mi je ves čas stal ob strani in bil vedno pripravljen priskočiti na pomoč v trenutkih težav ali dvomov. Brez njegove pozitivne energije, širokega poznavanja matematičnih pojmov in obilice idej, ki jih je bil vedno pripravljen deliti, bi bil moj matematični razvoj precej počasnejši, moje življenje pa dolgočasnejše.

Nadalje bi se rad zahvalil vsem sodelavcem Fakultete za Matematiko in Fiziko, s katerimi sem bil v teh letih v stiku. Med njimi sem še posebej hvaležen vsem tistim, ki so mi izkazali podporo, ko sem sprejemal odločitev, da se prijavim na postdoc pozicijo v Aberdeenu.

Poimensko bi se rad zahvalil Sari Kališnik Verovšek, ki mi je kot prva omenila možnosti zaposlitve v tujini in mi predstavila Pawła Dłotka. Hvaležen sem ji tudi za zanimive matematične pogovore. Zahvala gre tudi Jaku Smrekarju, ki mi je že na začetku doktorskega študija omenil, da bi bilo dobro, da razmislim o možnostih zaposlitve v tujini. Tudi njemu se zahvaljujem, da mi je bil vedno pripravljen pomagati in odgovarjati na vprašanja matematične in nematematične narave, in seveda prav tako za pripravljenost prebrati disertacijo. Zahvaljujem se Petru Pavešiću, ki je bil tudi vedno pripravljen pomagati, hkrati pa je bil tudi vedno pripravljen na zanimive pogovore. Hvaležen sem mu tudi, da mi je omenil možnost udeležbe na zelo zanimivi konferenci Applied Topology in Będlewo. Nazadnje bi rad izkazal še hvaležnost Barbari Drinovec Drnovšek in Matjažu Konvalinki.

Rad bi se zahvalil tudi vsem sodelavcem na IJS, IMFM in FMF, vsem prijateljem, sorodnikom, odjemalcem in hišnim ljubljenčkom, ki so mi stali ob strani in vnesli tudi nekaj nematematičnega veselja v moje življenje, ter obema svojima družinama. Nazadnje bi se posebej zahvalil še Maruši za vsa leta skupnega življenja, moralne podpore in za vse, česar sem se naučil od nje.

Kot vedno bi se rad zahvalil tudi tistim, ki sem jih morda na tem mestu pozabil omeniti, čeprav bi si to zaslužili.

I would like to thank Katharine Turner for her constructive comments regarding an earlier version of this dissertation and for being prepared to read it quickly and carefully. I would like to thank Ran Levi, who offered me a postdoc position in Aberdeen and had an illuminating talk with me in Będlewo. I would also like to thank Paweł Dłotko, who introduced me to Ran Levi and supported my decision to apply to the postdoc position in Aberdeen.

Regarding the approximate nerve theorem, I would like to thank Don Sheehy, who suggested the problem, and Petar Pavešić for suggesting the proof of Proposition 6.4. During my PhD studies I was supported by the Slovenian Research Agency Grant P1-0292-0101 and by the EU projects TOPOSYS (FP7-ICT-318493-STREP) and ProaSense (FP7-ICT-612329).

1 Introduction

For science to function successfully, it is important that we are able to statistically analyze the data from experiments. In this process, we come across various probability distributions, among which the normal distribution is especially important and well understood. One of its notable features is unimodality, which informally means that its probability density function has a unique local maximum. Most other classical probability distributions are unimodal as well. On the other hand, experimentally collected data often results in distributions with more than one mode, where the data is accumulated around more than one value. Understanding this phenomenon is especially important because it can indicate the presence of more than one underlying effect influencing the values of the data. As a somewhat simplified example, consider the density of traffic. The traffic is denser in the morning and in the afternoon, and the two underlying effects that explain this are the facts that people are driving to work in the morning and back home in the afternoon. Another example would be the study of height in a population consisting of two different types of individuals. Here we expect to obtain two peaks, corresponding to the average heights of each type of individual.

Hence, given a distribution that describes a certain phenomenon of interest, it is especially important to be able determine the minimal number of underlying effects[1] explaining this phenomenon. In other words, we would like to decompose a given distribution with more than one mode as a sum of unimodal summands. In statistics, this problem is especially well-studied in the special case when the underlying effects are normally distributed [7, 33, 56], and to a certain extent also for more general distributions [45, 46]. Another interesting phenomenon is the appearance of ghost modes, described in [28]: a mixture of k isotropic normal distributions can have more than k modes. In practice, the distributions of different effects may vary and we might be dealing with non-numerical data, in which case any analitical description of the distibution would be just an artifact of our choice of coordinate system, not the problem itself. Despite this, we usually have a natural concept of closeness/similarity. Baryshnikov and Ghrist have addressed these issues by introducing the concept of unimodal category [4], which is a topological abstraction of such problems. Here, instead of a probability density function $\mathbb{R}^m \to [0,\infty)$, we study a nonnegative function $f : X \to [0,\infty)$ on a topological space X. We say that the function $u : X \to [0,\infty)$ is unimodal if there is an $M > 0$ such that the superlevel sets $u^{-1}[c,\infty)$ are contractible for $c \in (0, M]$ and empty for $c > M$. The unimodal category of a function $f : X \to [0,\infty)$ is then defined as the smallest $n \in \mathbb{N}_0$ for which there exist unimodal functions $u_1, \dots, u_n : X \to [0,\infty)$, such that $f = \sum_{i=1}^{n} u_i$ (where the summation is pointwise). In this case, we write $n = \mathbf{ucat}(f)$. Therefore, the unimodal category is a lower bound for the number of summands, regardless of the unimodal distributions of the underlying effects we might be interested in. This concept generalizes naturally if instead of decomposing functions into sums of unimodals, we decompose them into ℓ^p-combinations of unimodals, $p \in (0,\infty]$. In this way, we obtain the concept of unimodal p-category $\mathbf{ucat}^p(f)$. Baryshnikov and Ghrist [4] also suggested $\mathbf{gcat}(\mathrm{supp}(f))$, the geometric category of the support of f,

[1] Baryshnikov and Ghrist like to reference Ockham's razor here, which provides a philosophical motivation for such endeavor.

as a natural candidate for the case $p = 0$. In light of this, the unimodal p-category can also be understood as a deformation of the geometric category of the support of f. (The geometric category is a variant of the classical Lusternik-Schnirelmann category [24]. These invariants count the minimal number of pieces of a certain type a topological space can be decomposed into. For the geometric category the pieces are required to be contractible, whereas for the Lusternik-Schnirelmann category they are only required to be contractible within the ambient space.) In other words, the unimodal category can be understood as a lift of geometric category, an invariant of spaces, to an invariant of functions on these spaces. Recently, such lifting of invariants has proven to be a particularly successful idea, consider e.g. Euler calculus [25], where the Euler characteristic is interpreted as a measure and integrating a constructible function against this measure yields the concept of its Euler integral. The very successful concept of persistent homology [30] can be understood as a similar lift; in the basic variant, this is a lift of homology, an invariant of spaces (and maps between them), to an invariant of filtrations of these spaces.

Not much is known about the unimodal category, even in the basic case of $X = \mathbb{R}^m$, which is the most interesting case from the statistical point of view. For the case $m = 1$, Baryshnikov and Ghrist provided a simple algorithm [4], which allows us to compute the unimodal category of any function with only finitely many critical points. Their paper treats the case $m = 2$ only partially and concludes with a monotonicity conjecture, which they suggest will play a key role in providing precise bounds for the unimodal category in higher dimensions – the conjecture is that for a fixed function $f : X \to [0, \infty)$ and $0 < p_1 < p_2 \leq \infty$, we always have $\mathbf{ucat}^{p_1}(f) \leq \mathbf{ucat}^{p_2}(f)$. Computation of $\mathbf{ucat}$ in the case $m = 2$ is treated in some more detail by Hickok, Villatoro and Wang in [43], which is focused on those Morse distributions on the plane whose Morse-Smale graphs are trees. For these, the unimodal category is almost completely characterized.

The thesis consists of two conceptual parts. The first part is mostly centered around the monotonicity conjecture and is the subject of the paper [40], which is currently being prepared for publication, whereas the second part regards the approximate nerve theorem [41], which is related to persistent homology and can thus be regarded as a continuation of the research [39] which spawned the author's diploma thesis.

In the first part, we begin by showing that the decomposition provided by the algorithm for $X = \mathbb{R}$ can be generalized to arbitrary functions $f : \mathbb{R} \to [0, \infty)$ as a variant of the Jordan decomposition [58] of functions with bounded variation (each such function can be written as the difference of two monotonically increasing functions). Then, we generalize these results to obtain a generalization to $X = S^1$ suitable for the study of monotonicity, which also yields a simple algorithm in the case of finitely many critical points. The results obtained are general enough to allow for proving the monotonicity conjecture for arbitrary functions on $X = \mathbb{R}$ and $X = S^1$. Next, we show that monotonicity does not hold for certain more general spaces. Namely, we construct two counterexamples on graphs, leading us to conclude that the conjecture is false for most graphs, and more importantly, we also construct two counterexamples on $X = \mathbb{R}^2$. Finally, we show that, despite this, the conjecture is true for $X = \mathbb{R}^2$ in the case of functions whose Morse-Smale graph is a tree.

1.1 Approximate Nerve Theorem

The second part of the thesis, concentrated in Section 6, is related to more developed areas of computational topology, such as persistent homology. For this reason it requires more background and deserves a separate introduction.

To motivate this part of the thesis, note that every notion of category is related to a certain type of cover of the underlying space. For instance the Lusternik-Schnirelmann category is concerned with categorical covers, i.e. such that the cover elements are contractible within the space, and the geometric category is concerned with covers consisting of contractible sets. The unimodal category is related to the concept of cover as well, however, the relation is more complicated. The simplest case is that of the unimodal ∞-category, where the function $f : X \to [0, \infty)$ is decomposed as $f = \min_{1 \leq i \leq n} u_i$ where each $u_i : X \to [0, \infty)$ is unimodal. This means that at each level $c > 0$ we have $f^{-1}[c, \infty) = \bigcup_{i=1}^{n} u_i^{-1}[c, \infty)$, which means that at each level the superlevel sets of the unimodal functions in the decomposition form a cover of the corresponding superlevel set of the original function and this cover consists of contractible sets.

Inferring global properties of the space from local properties, for instance the homology of a space from the homology of cover elements in an appropriate cover of the space, is a common theme in algebraic topology. It is therefore reasonable to expect that such techniques will prove to be fruitful in the context of unimodal category as well, once the area is sufficiently developed.

From the point of view of persistent homology, a function $f : X \to [0, \infty)$ provides us with a superlevel set filtration of the space X and the unimodal ∞-decomposition $f = \min_{1 \leq i \leq n} u_i$ yields a filtered cover of this space, whose cover elements have trivial persistent homology. So far, it is unclear what the general connection between persistence and the unimodal category is. However, in the case when the cover is particularly simple, namely, a good cover, there are classical results for the unfiltered case. Using spectral sequences, we have been able to extend these results into the setting of persistent homology. It seems likely that similar techniques can be used to work more generally with filtered covers whose elements are contractible at each level, as in the case of the unimodal ∞-category.

It is also known that for functions $f : \mathbb{R} \to [0, \infty)$ the concept of persistence is related to the concept of total variation [6]. By the first part of the thesis, this means that in the one-dimensional case, **ucat** is intimately related to persistence.

The classical result alluded to above is the nerve theorem, which relates a sufficiently nice cover of a topological space with the nerve of that cover, and goes back to Alexandroff [2].

Theorem 1.1 (Corollary 4G.3 [42])**.** *If $\mathcal{U}$ is an open cover of a paracompact space X such that every non-empty intersection of finitely many sets in $\mathcal{U}$ is contractible, then X is homotopy equivalent to the nerve $\mathcal{N}(X)$.*

One more recent application is in the area of *topological data analysis* [37, 14, 63]. The goal is to obtain information about the topology of a space, often given a discrete sample of the space. There has been a large body of work proving results in different contexts, including [21, 9, 26] just to name a few. A common point is the use of the nerve theorem, either explicitly or implicitly, through constructions such as the Čech complex.

The main idea of *persistent homology*, a powerful tool in topological data analysis, is to study the homology of a filtration rather than a single space. Applying the homology functor yields a *persistence module*. If we compute homology with field coefficients, then we can obtain a complete topological invariant called *persistence barcode* or *persistence diagram*. One useful source of filtrations are sublevel (resp. superlevel) set filtrations – given a space endowed with a real-valued continuous function, $f : X \to \mathbb{R}$, the sublevel sets of the function form a filtration yielding a persistence diagram denoted by $\mathrm{Dgm}(X, f)$.

One important example of a function is the distance to a compact set. When the compact set consists of sample points, this function relates to a notion of scale and is equivalent to the Čech filtration. Recall that the Čech complex on a point set P is the nerve of the union of balls of radius r. The points are usually embedded in Euclidean space, allowing the nerve theorem to be applied via convexity. By varying the radius r, we obtain the Čech filtration. Other filtrations which are often considered are: superlevel sets of probability density functions [9], sublevel set filtration of a sampled function [21], and the elevation function on 2-manifolds [1]. Persistence diagrams have proven interesting because they are stable [23] – meaning a small change in the filtration bounds the change in the invariant. One way of measuring the magnitude of the change is the *bottleneck distance*. Stability enables us to prove theorems about approximating the persistent homology of a filtration using an alternate filtration constructed from a discrete sample, i.e. that the bottleneck distance is small.

An important technique in proving such an approximation is *interleaving* [19], which provides an algebraic condition for approximation (Section 2.5). A common theme is to construct an interleaving with a good cover, providing an approximation guarantee. In some cases, such as for distance filtration, an interleaving with a good cover can often be shown directly. In more general settings, it can often be more difficult to directly prove an interleaving. The main goal in Section 6 is to prove an approximation bound using the stability of persistent homology to relax the need for a good cover. Importantly, we only make assumptions on the *local properties* of the space and function, which make it useful in a variety of applications.

As we deal with persistent homology, we concentrate on a homological version of Theorem 1.1.

Theorem 1.2 (Theorem 4.4 [12]). *Suppose X is the union of subcomplexes U_i such that every non-empty intersection $U_{i_0} \cap \cdots \cap U_{i_p}$ for $p \geq 0$ is acyclic. Then $\mathsf{H}_*(X) \cong \mathsf{H}_*(\mathcal{N}(\mathcal{U}))$ where $\mathcal{N}(\mathcal{U})$ is the nerve of the cover.*

The main result of Section 6 is to provide an *approximate* version of the above theorem in the context of persistent homology. Given a space and function, we first define the notion of an ε-acyclic cover. Note that we do not restrict ourselves to induced functions on a fixed cover, but consider a covering by filtrations. This notion is less intuitive but is applicable in a wider range of settings. We follow the formalization of covers by filtrations by Sheehy [60]. Informally, our main result is:

Result 1.3. *Given a space X endowed with a function f and a (filtered) cover $\mathcal{U}$, if every non-empty finite intersection of cover elements is ε-acyclic, then there exists a function on the nerve $g : \mathcal{N}(\mathcal{U}) \to \mathbb{R}$, such that the bottleneck distance $d_B(\cdot)$ is bounded by*

$$d_B(\mathrm{Dgm}(X, f), \mathrm{Dgm}(\mathcal{N}(\mathcal{U}), g)) \leq 2(Q + 1)\varepsilon,$$

where

$$Q = \min\{\dim(X), \dim(\mathcal{N}(\mathcal{U}))\}.$$

The construction of the function on the nerve is given explicitly and agrees with what is currently done in practice when computing persistent homology.

We do not use persistence diagrams and bottleneck distance, but find it more convenient to work directly with the corresponding persistence modules and interleavings. Therefore, we do not explicitly define bottleneck distance or diagrams as they are not required for the statement of our results, but do allude to them to give intuition for readers who are familiar with persistence. In cases where the diagrams are well defined, results of bottleneck distance type follow automatically.

We prove this result by using the Mayer-Vietoris spectral sequence to glue together the ε-acyclic pieces into the global persistent homology. To obtain a tight bound, we introduce the notion of *left* and *right interleavings* (Section 6.2), which have additional structure. This refinement of interleavings captures similar phenomena as the results of [5], but works at the level of modules rather than barcodes. Hence, it does not require modules to be decomposable and so we believe these notions are of independent interest.

Approximation results of this type have received significant attention in the community. In addition to the applications mentioned above, the persistent nerve lemma in [22] showed that the homotopy equivalence between the a good cover its nerve commutes with inclusions allowing it to be applied to a persistent setting. More recently, this was extended in [10], who approximate the persistent homology of a Čech complex in Euclidean space using non-good covers. Also recently a comparable result has been reported in [15]. Their result was more general than an earlier version of paper [41] which did not cover the case of *filtrations of filtrations* (also called a multicover) - which is now addressed in the paper, as well as this thesis. Furthermore, the proofs are different as the [15] requires cover elements (and finite intersections) to be "ε-nullhomotopic". They show that under this assumption, it is possible to construct an explicit chain map interleaving (with the same constant that we report). Our approach, however, is purely algebraic and so only requires structure at the level of homology rather than homotopy.

We prove the result in two steps: first we show how the approximation bound evolves through the computation of the spectral sequence (Section 6.3), then resolve the extension problem to relate the result of the spectral sequence with the persistent homology of the underlying space (Section 6.4). While we have tried to make this Section 6 contained, we do assume some familiarity with spectral sequences, but we try and provide intuition and references whenever possible.

1.2 Structure of the Thesis

In Section 2 we introduce the relevant concepts and some useful results. In Section 3, we generalize the sweeping algorithm of Baryshnikov and Ghrist to obtain a general decomposition theorem for functions on $X = \mathbb{R}$, thereby removing the assumption that the function only has finitely many critical points. This is achieved by a systematic use of the concepts of positive and negative variation. Then we show how to compute the unimodal category in the case $X = S^1$. In Section 4, we treat the monotonicity conjecture. We prove it in the case $X = \mathbb{R}$ and give various counterexamples to demonstrate the ways in which it may fail in general. Specifically, we construct two simple counterexamples on graphs, the first of which shows that the conjecture is false in the case $p_1 < p_2 < \infty$, whereas the second shows that the conjecture is also false for $p_1 < p_2 = \infty$. Then we construct two counterexamples on the plane $X = \mathbb{R}^2$ that are to a certain extent analogous to the counterexamples on graphs. We also prove that the conjecture is nevertheless true for $X = \mathbb{R}^2$ if the Morse-Smale graph is a tree. In Section 5 we present some investigations regarding higher dimensions and some ideas for future research. Section 6 gives an account of the approximate nerve theorem, as proved in [41].

1.3 Original Contributions

All results from Section 3, Section 4 and Section 5, except where otherwise noted, are original results of the author. The results of Section 6 are original as well and have been obtained by the joint work [41] of the author and his co-advisor Primož Škraba.

1.4 Conventions Used in the Text

Unless otherwise noted, in Sections 2–5 all functions $f : X \to [0, \infty)$ are assumed to be **continuous**. In addition to this, all functions are assumed to be of **bounded variation** in Section 3. Compact support is not assumed. When we say $f : X \to [0, \infty)$ is a **Morse function** we mean that f is a function that has finitely many critical points and is Morse on $f^{-1}(0, \infty)$. In Section 6, the symbol H_* is used to denote both ordinary as well as **persistent homology**, depending on whether its argument is just a space or a space equipped with a filtration.

2 Preliminaries and Previously Known Results

2.1 Unimodal Category

From the point of view of algebraic topology, the simplest spaces are contractible ones. Intuitively, these are the spaces which can be contracted to a point within themselves. A related notion is that of a space contractible within a larger space. Such spaces can be contracted to a point within this larger space. Formally, these are defined as follows.

Definition. A space X is said to be *contractible* if the identity map $\mathrm{id}_X : X \to X$ is nullhomotopic. A subspace $A \subseteq X$ is said to be *contractible within* X if the inclusion map $A \hookrightarrow X$ is nullhomotopic.

For instance, the space $\mathbb{R}^2$ is contractible, whereas the unit circle S^1 is not. However, the unit circle is contractible within $\mathbb{R}^2$, since we can contract it to a point by restricting the homotopy which contracts $\mathbb{R}^2$ to a point within itself. These notions are the starting point of the theory of Lusternik-Schnirelmann category, which measures the complexity of a space X by how many sets contractible within X are required to cover the space. There is also a related notion of geometric category. Another related notion by the name of *topological complexity* has been defined by Farber [35].

Definition. The **Lusternik-Schnirelmann category** $\mathbf{cat}(X)$ of a topological space X is the minimum number n of open sets $U_1, \ldots, U_n$ contractible within X such that $X = \bigcup_{i=1}^n U_i$. The **geometric category** $\mathbf{gcat}(X)$ of a topological space X is the minimum number n of contractible open sets $U_1, \ldots, U_n$ in X such that $X = \bigcup_{i=1}^n U_i$.

These categories have been well-studied, see for instance [24] for a nice introduction into the topic. It is known for instance that the Lusternik-Schnirelmann category is a homotopy invariant, whereas the geometric category is not. The standard example illustrating this is due to Fox [36], namely let $X_1 = S^2 \vee S^1 \vee S^1$ and $X_2 = S^2/\{x, y, z\}$. Then $X_1 \simeq X_2$ and $\mathbf{cat}(X_1) = \mathbf{cat}(X_2) = 2$, whereas $\mathbf{gcat}(X_1) = 2$ and $\mathbf{gcat}(X_2) = 3$. For a proof of this, see [24, Proposition 3.11]. This also illustrates that **cat** and **gcat** are not the same.

Following Baryshnikov and Ghrist [4], we are interested in a notion of category applicable to the study of continuous functions $f : X \to [0, \infty)$. To define it, we need a concept analogous to contractibility, applicable to functions:

Definition. A continuous function $u : X \to [0, \infty)$ is *unimodal* if there is an $M > 0$ such that the superlevel sets $u^{-1}[c, \infty)$ are contractible for $0 < c \leq M$ and empty for $c > M$.

It should be noted that in probability, unimodality strictly means that the relevant probability distribution has a single mode, which corresponds to a unique global maximum, however, in practice, the word "unimodality" is also commonly used to refer to functions with a single local maximum or maximal region. The above definition is a formalization of the latter convention. We can now state the main definition.

Definition. Let $p \in (0,\infty)$. The *unimodal p-category* $\mathbf{ucat}^p(f)$ of a function $f : X \to [0,\infty)$ is the minimum number n of unimodal functions $u_1,\dots,u_n : X \to [0,\infty)$ such that pointwise, $f = (\sum_{i=1}^n u_i^p)^{\frac{1}{p}}$. Similarly, the *unimodal ∞-category* $\mathbf{ucat}^\infty(f)$ of a function $f : X \to [0,\infty)$ is the minimum number n of unimodal functions $u_1,\dots,u_n : X \to [0,\infty)$ such that pointwise, $f = \max_{1\le i\le n} u_i$. In place of $\mathbf{ucat}^1(f)$ we usually write $\mathbf{ucat}(f)$.

Since the definition of unimodality does not say anything about $u^{-1}[0,\infty)$, we can safely ignore any part of the space where the function is zero. In particular, if $A = f^{-1}(0,\infty)$, we have

$$\mathbf{ucat}^p(f) = \mathbf{ucat}^p(f|_A).$$

Baryshnikov and Ghrist [4] also give a more general notion of unimodal ν-category, corresponding to any norm ν on an appropriate space of real-valued sequences. We formalize this as follows.

Definition. Let $\mathbb{R}^{(\mathbb{N})} \equiv \bigoplus_{n\in\mathbb{N}} \mathbb{R}$ denote the vector space of all eventually zero sequences of real numbers, i.e. sequences with at most finitely many nonzero terms. Suppose ν is a norm on $\mathbb{R}^{(\mathbb{N})}$. The *unimodal ν-category* $\mathbf{ucat}^\nu(f)$ of a function $f : X \to [0,\infty)$ is the minimum number n of unimodal functions $u_1,\dots,u_n : X \to [0,\infty)$ such that for each $x \in X$, $f(x) = \nu(u(x))$, where $u(x) = (u_i(x))_{i=1}^\infty$ and we take $u_i := 0$ for $i > n$. We refer to the sequence $u_1,\dots,u_n$ as a *unimodal ν-decomposition* of f and to n as the *length* of this decomposition.

Remark 2.1. It can happen that a function $f : X \to [0,\infty)$ does not have a unimodal ν-decomposition $u_1,\dots,u_n$ for any $n \in \mathbb{N}$, so that technically $\mathbf{ucat}^\nu(f)$ is undefined. We express this fact by writing $\mathbf{ucat}^\nu(f) = \infty$. We will see in Section 3, for instance, that $\mathbf{ucat}^p(f) = \infty$ if f is not of bounded variation.

The unimodal p-category $\mathbf{ucat}^p$ corresponds to the p-norm, $p \in [1,\infty]$. For $p < 1$, we do not have a corresponding norm, as the triangle inequality fails, but the definition of $\mathbf{ucat}^p$ is still of interest. The more general concept of $\mathbf{ucat}^\nu$ does not seem as well-behaved as $\mathbf{ucat}^p$ since changing the order of functions $u_1,\dots,u_n$ may change the value of $\nu(u(x))$. For this reason, we limit our further investigations to the case of p-norms.

Remark 2.2. Baryshnikov and Ghrist also consider $\mathbf{gcat}(f^{-1}(0,\infty))$ as a natural candidate[2] for the notion of $\mathbf{ucat}^0(f)$ and claim that $\lim_{p\downarrow 0} \mathbf{ucat}^p(f) = \mathbf{ucat}^0(f)$. We note that while this may be a natural candidate, the equation does actually not hold as stated. For instance, if $f : \mathbb{R} \to [0,\infty)$ is a function with two peaks such that $f^{-1}(0,\infty)$ is connected, e.g. $f(x) = \max\{0, 2 - ||x| - 1|\}$, then $\mathbf{ucat}^p(f) = 2$ for all $p \in (0,\infty]$ whereas $\mathbf{gcat}(f^{-1}(0,\infty)) = 1$.

We also note that there is a typo in the statement of their Lemma 9 in [4], which relates the various notions of $\mathbf{ucat}^p$ for $p < \infty$. The proof as stated there, remains valid. For convenience, we restate the lemma in its correct form.

Lemma 2.1 ([4], Lemma 9)**.** *If $f : X \to [0,\infty)$ is any continuous function, then*

$$\mathbf{ucat}^p(f) = \mathbf{ucat}(f^p).$$

[2] Note that they seem to be using the notation $\operatorname{supp} f$ to denote $f^{-1}(0,\infty)$, i.e. the set-theoretic support of f.

2.1.1 Discussion

Note that $\mathbf{ucat}^p(f)$ can only be finite if $f : X \to [0, \infty)$ is continuous, since continuity is assumed in the definition of unimodality. This motivates the following convention.

Convention. Unless explicitly noted otherwise, all functions $f : X \to [0, \infty)$ will hereafter be assumed to be continuous.

Baryshnikov and Ghrist further assume that the functions they study are compactly supported. We do **not** make this assumption. Because of this, our notion of unimodal category slightly differs from theirs. Note that the two notions of $\mathbf{ucat}^p(f)$ agree whenever f is compactly supported. However, if f is not compactly supported, their notion is undefined, while ours may still give a finite answer. Hence, the notion we use is a slight generalization of theirs. (Note that all the main examples and counterexamples considered in the thesis will be compactly supported, so they are still valid under the original definition.)

While the definition of $\mathbf{ucat}^p$ makes sense for general topological spaces, it seems to be the most interesting for spaces that are not too pathological. To illustrate this point, we make a simple observation.

Proposition 2.2. *Suppose $f : X \to [0, \infty)$ is a function such that $\mathbf{ucat}^p(f) < \infty$. Then each point $x \in X$ with $f(x) \neq 0$ has a contractible neighborhood.*

Proof. Since $\mathbf{ucat}^p(f) < \infty$ and $f(x) \neq 0$ there is a unimodal $u : X \to [0, \infty)$ such that $u(x) = a > 0$. Hence, $u^{-1}[\frac{a}{2}, \infty)$ is a contractible neighborhood of a. □

For instance, if X is the Hawaiian earring and $\mathbf{ucat}^p(f) < \infty$ for some $f : X \to [0, \infty)$, this f must necessarily have a zero at the point where the circles intersect. Similarly, if X is the topologist's sine curve, defined as the closure of

$$A = \{(x, y) \in (0, 1) \times \mathbb{R} \mid y = \sin \tfrac{\pi}{x}\}$$

in $\mathbb{R}^2$, any function $f : X \to [0, \infty)$ whose unimodal p-category is finite must vanish at the points of $\{0\} \times [-1, 1]$ as none of these has a contractible neighborhood.

However, even if there is a function $f : X \to [0, \infty)$ with $\mathbf{ucat}^p(f) < \infty$ and no zeros, the local structure of X can still be pathological. For instance, if X is contractible, but not locally contractible, the constant function $f : X \to [0, \infty)$, $f(x) = 1$, is nonetheless unimodal, so $\mathbf{ucat}^p(f) = 1$.

Ideally, the space should be at least locally contractible to allow for the existence of unimodal functions supported in a neighborhood of any given point. As there are plenty of open questions already in the case when X is a manifold or a CW complex, we will restrict our attention to those.

2.2 Total Variation and Jordan Decomposition

The concept of total variation (see e.g. [3, Chapter 6]) for functions $f : \mathbb{R} \to [0, \infty)$ will be useful, so we recall the basic definitions and results here.

Definition. The *total variation* of $f : \mathbb{R} \to \mathbb{R}$ on the interval $[a, b]$ is defined by the formula

$$V(f; [a,b]) = \sup \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|,$$

where the supremum is taken over all partitions $a = x_0 < x_1 < \ldots < x_n = b$ of the interval $[a, b]$. Similarly, we define the *positive variation* of f on the interval $[a, b]$ as

$$V^+(f; [a,b]) = \sup \sum_{i=1}^{n} \max\{0, f(x_i) - f(x_{i-1})\},$$

and the *negative variation* of f on the interval $[a, b]$ as

$$V^-(f; [a,b]) = \sup \sum_{i=1}^{n} \max\{0, f(x_{i-1}) - f(x_i)\}.$$

We will use the following basic facts:

Theorem 2.3. *Let V^* stand for V, V^+ or V^- and let $f : [a, b] \to \mathbb{R}$. Then the following hold:*

- *If f is increasing on $[a, b]$, then $V(f; [a, b]) = V^+(f; [a, b]) = f(b) - f(a)$ and $V^-(f; [a, b]) = 0$.*
- *If f is decreasing on $[a, b]$, then $V(f; [a, b]) = V^-(f; [a, b]) = f(a) - f(b)$ and $V^+(f; [a, b]) = 0$.*
- *If $f = \sum_{i=1}^{n} f_i$, then $V^*(f; [a, b]) \leq \sum_{i=1}^{n} V^*(f_i; [a, b])$.*
- *If $a = x_0 < x_1 < \ldots < x_n = b$, then $V^*(f; [a, b]) = \sum_{i=1}^{n} V^*(f; [x_{i-1}, x_i])$.*
- *$V^+(f; [a, b]) + V^-(f; [a, b]) = V(f; [a, b])$ and $V^+(f; [a, b]) - V^-(f; [a, b]) = f(b) - f(a)$.*

Proof. The first two facts are obvious. For the third and the fourth fact in the case of V, see Theorems 6.9 and 6.11 in [3]. For V^+ and V^-, the idea is completely analogous. The fifth fact is a standard exercise, see [3], Exercise 6.7. □

By undergraduate measure theory, every right continuous increasing function on $\mathbb{R}$ determines a Borel measure [27, Section IV.8], so V, V^+ and V^- can be extended to define measures on $\mathbb{R}$. We are therefore also allowed to compute these variations over open intervals or even Borel sets, but note that continuity of f implies that $V^*(f; [a, b]) = V^*(f; (a, b))$.

These concepts also make sense for $f : J \to \mathbb{R}$, where $J \subseteq \mathbb{R}$ is an interval. Functions of bounded variation, i.e. such that $V(f; J) < \infty$, are of particular interest to us. Note that the first two bullet points of Theorem 2.3 imply that monotone functions over finite intervals have bounded variation. More generally, bounded monotone functions also have bounded variation. Using the fourth bullet point, this also implies that unimodal functions $u : J \to \mathbb{R}$ have bounded variation, since the domain of any such function can be split into two intervals where the function is monotone and bounded.

Any function $f : J \to \mathbb{R}$ of bounded variation can be split as the difference of two monotone functions, i.e. it satisfies the following Jordan decomposition theorem:

Theorem 2.4. *Suppose $f : J \to \mathbb{R}$ is of bounded variation. Then f can be expressed as the difference $f = g - h$ of two increasing functions $g, h : J \to \mathbb{R}$.*

Proof. See [3], Theorem 6.15. The result is stated there for closed intervals, but it actually holds in general. The idea is to take $g(x) = V(f; J \cap (-\infty, x))$ and $h(x) = g(x) - f(x)$. $\square$

In Section 3.1, we use a slightly different decomposition, namely $g(x) = V^+(f; J \cap (-\infty, x))$ and $h(x) = V^-(f; J \cap (-\infty, x))$, assuming $\lim_{x \to -\infty} f(x) = 0$. The limit at $-\infty$ can always be subtracted from f, so this is not really a restriction. To see that this is indeed a decomposition, use the fifth property of Theorem 2.3.

2.3 Morse Functions

For convenience, we remind the reader of some basic notions of Morse theory. First, recall that a *Morse function* on a manifold M is a smooth function $f : M \to \mathbb{R}$ which only has nondegenerate critical points. A nondegenerate point is always isolated, so if M is a closed manifold, there are only finitely many such points. If M is not closed, the situation is slightly more complicated, so to keep things simple, we adopt the following convention.

Convention. Suppose M is a noncompact manifold and $f : M \to [0, \infty)$ a continuous function which is smooth on $f^{-1}(0, \infty)$. We say that f is a *Morse function* if its restriction to $f^{-1}(0, \infty)$ is a Morse function with finitely many critical points.

In other words, we are not interested in the case with infinitely many critical points and we are also not interested in what happens in the zero set.

Definition. A Morse function all of whose critical values are distinct is called *nonresonant.*

This terminology seems to be due to Nicolaescu [53], who also mentions that Thom calls such functions *excellent*. One of the main points of Morse theory is that it enables us to describe the topology of a manifold M using a Morse function on M:

Theorem 2.5 ([51])**.** *Suppose $f : M \to \mathbb{R}$ is a Morse function, $p \in M$ is a critical point of index i and $f(p) = a$ the corresponding critical value. For every $x \in \mathbb{R}$ let $M_x = f^{-1}(-\infty, x]$. Then if $[a - \epsilon, a + \epsilon]$ contains no other critical values of f, the sublevel set $M_{a+\epsilon}$ is obtained (up to homotopy) from $M_{a-\epsilon}$ by attaching an i-handle.*

Remark 2.3. For our purposes, the relevant fact here is that attaching an i-handle either destroys a homology class in H_{i-1} or creates a homology class in H_i.

2.4 Filtered Simplicial Complexes

To understand the part of the thesis pertaining to the approximate nerve theorem, the reader is presumed to be familiar with persistent homology. We refer the reader to [29] and [64] for complete introductions. The relevant preliminaries are given below – as much as possible we have tried to avoid technical complications but we try to point out where generalizations are possible.

To minimize technical complications, we work primarily with $\mathbb{Z}$-filtered simplicial complexes (see the definition below), denoted X and $\mathbb{Z}$-filtered covers $\mathcal{U}$ of such complexes by subcomplexes (where each subcomplex itself has a specified filtration). However, our proofs work directly on the algebraic level and hence should be extendable to much more general settings than those presented here without changing the bounds. Note that already the results for $\mathbb{Z}$-filtered simplicial complexes are widely applicable. It is known, for instance, that each smooth manifold or, more generally, Whitney-stratified space can be triangulated. Hence, such a space Y equipped with a sublevel set filtration induced by some function $f : Y \to \mathbb{R}$ can be approximated arbitrarily well by a piecewise linear (PL) function on a simplicial complex and therefore by an $\mathbb{R}$-filtered simplicial complex.

Definition. Let $J \subseteq \mathbb{R}$. A *J-filtered simplicial complex* is a pair $(X, \mathcal{F})$, where X is an abstract simplicial complex and $\mathcal{F} = (X^j)_{j\in J}$ is a family of subcomplexes such that $j_1 \leq j_2$ implies $X^{j_1} \subseteq X^{j_2}$, $X^{-\infty} := \bigcap_{j\in J} X^j = \emptyset$ and $X^{\infty} := \bigcup_{j\in J} X^j = X$.

A *J-filtered cover*[3] *by subcomplexes* of a J-filtered simplicial complex $(X, \mathcal{F})$ is an indexed family $\mathcal{U} = (U_i, \mathcal{F}_i)_{i\in\Lambda}$, where each U_i is a subcomplex of X and $\mathcal{F}_i$ is a filtration of this subcomplex, such that the filtrations $\mathcal{F}$ and $\mathcal{F}_i$ satisfy a compatibility requirement, namely that $X^j = \bigcup_{i\in\Lambda} U_i^j$ holds for each $j \in J$. Note that whenever $I \subseteq \Lambda$, the intersection $U_I := \bigcap_{i\in I} U_i$ has a natural filtration $\mathcal{F}_I$ given by $U_I^j := \bigcap_{i\in I} U_i^j$.

Note that the requirements on $X^{-\infty}$ and X^{∞} are sometimes dropped. If $J \subseteq \mathbb{R}$ is a discrete subset, for instance $J = \mathbb{Z}$, the filtration $\mathcal{F}$ may also be given as a function $f : X \to \mathbb{Z}$ whose sublevel sets are $f^{-1}(-\infty, j] = X^j$. For this reason, a $\mathbb{Z}$-filtered simplicial complex is sometimes written as (X, f). Since the filtration is regarded as part of the structure, we often suppress it from notation and simply write X.

When the filtrations are given as functions, the compatibility requirement in the definition of the filtered cover $\mathcal{U} = (U_i, f_i)_{i\in\Lambda}$ of the filtered simplicial complex (X, f) can be stated[4] as $f = \min_{i\in\Lambda} f_i$.

Remark 2.4. This definition of J-filtered cover is the one given by Sheehy [60], which allows for the extension of Theorem 1.2 to the persistent setting via the persistent nerve lemma of Chazal and Oudot [22].

There is also a natural way to assign a filtered cover to an unfiltered cover of a filtered complex. Namely, if $(X, \mathcal{F})$ is a J-filtered simplicial complex and $\overline{\mathcal{U}} = (U_i)_{i\in\Lambda}$ is a cover of the underlying complex X by subcomplexes, $\overline{\mathcal{U}}$ can naturally be given the structure of a J-filtered cover $\mathcal{U} = (U_i, \mathcal{F}_i)_{i\in\Lambda}$ by defining $U_i^j = U_i \cap X^j$. We call $\mathcal{U}$ the *induced J-filtered cover of* $(X, \mathcal{F})$ associated to $\overline{\mathcal{U}}$. In this case, if the filtrations are given by functions $f : X \to \mathbb{Z}$ and $f_i : U_i \to \mathbb{Z}$, the functions f_i are simply restrictions $f_i = f|_{U_i}$.

Our results also make sense in the setting of triangulable spaces, which we now recall.

[3] All covers are assumed to be indexed.

[4] To make sense of the minimum, we may consider f_i to be extended to the whole X by defining it to be ∞ outside of U_i.

Definition. A topological space Y is said to be *triangulable* if there exists a simplicial complex X and a homeomorphism $h : |X| \to Y$, where $|X|$ denotes the carrier of X. The pair (X, h) is said to be a *triangulation* of Y.

In the persistent setting, we also consider filtered triangulable spaces. To do this, start with a space Y and a continuous function $f : Y \to \mathbb{R}$. The pair (Y, f) is then regarded to be an $\mathbb{R}$-filtered topological space. The filtration is defined by $Y^j = f^{-1}(-\infty, j]$ and is known as *the sublevel set filtration of Y induced by f.*

We sometimes need to replace the function $f : Y \to \mathbb{R}$ by a piecewise linear approximation.

Definition. Suppose (X, h) is a triangulation of Y and $f : Y \to \mathbb{R}$ a continuous function. The *piecewise linear approximation of f associated to* (X, h) is the function $\hat{f} : |X| \to \mathbb{R}$ defined on the vertices of X by $\hat{f}(v) = f(h(v))$ and extended affinely over the simplices.

Definition. Suppose X is a simplicial complex and $\hat{f} : |X| \to \mathbb{R}$ a piecewise linear function (w.r.t. the triangulation). Then X can be given the structure of an $\mathbb{R}$-filtered simplicial complex $(X, \hat{f})$ by defining X^j to consist of all simplices contained in $\hat{f}^{-1}(-\infty, j]$. We call this filtration *the lower star filtration of $\hat{f}$.*

Note that lower star filtrations are usually considered only for finite simplicial complexes and the function values on the vertices are assumed to be distinct. (See, for instance, [29].)

It is a standard fact (for finite simplicial complexes, this is explained in [29]) that the sublevel set filtration of $|X|$ and the lower star filtration of X induced by the same function $\hat{f}$ are related by the fact that $|X^j|$ is a deformation retract of $|X|^j$. Of more interest to us, however, is comparing the persistence modules of these two filtrations.

Finally, we recall the standard construction for the nerve given a cover is:

Definition. Given a cover $(U_i)_{i\in\Lambda}$ of X, the nerve $\mathcal{N}$ is the set of finite subsets of Λ defined as follows: a finite set $I \subseteq \Lambda$ belongs to $\mathcal{N}$ if and only if the intersection of the U_i whose indices are in I, is non-empty, or equivalently

$$U_I = \bigcap_{i\in I} U_i \neq \emptyset.$$

If I belongs to $\mathcal{N}$, then so do all of its subsets making $\mathcal{N}$ an abstract simplicial complex.

2.5 Modules and Interleavings

Let $\Bbbk$ be a field. Both the graded and the non-graded ring of polynomials with coefficients in $\Bbbk$ are commonly denoted by $\Bbbk[t]$ in the literature. We mostly work with the former. For this reason, we reserve the notation $\Bbbk[t] = \Bbbk[t]_{(\mathrm{Gr})}$ for the graded version and the non-graded version is always explicitly denoted as such by $\Bbbk[t]_{(\mathrm{NGr})}$.

Here $\Bbbk[t]$ is graded by degree, namely $\Bbbk[t] = \bigoplus_{i\in\mathbb{N}_0} \Bbbk[t]_i$, where $\Bbbk[t]_i = \Bbbk \cdot t^i$ consists of the homogeneous polynomials of degree i, i.e. scalar multiples of t^i. This decomposition is regarded as part of the structure of $\Bbbk[t]$ and has to be taken into account when defining $\Bbbk[t]$-modules and their morphisms, whereas $\Bbbk[t]_{(\mathrm{NGr})}$ is simply a ring without any additional structure, so $\Bbbk[t]_{(\mathrm{NGr})}$-modules and their morphisms are not required to respect any such grading.

Definition. [34, p. 42] A $\Bbbk[t]$-*module* is a $\Bbbk[t]_{(\mathrm{NGr})}$-module M together with a decomposition (also called grading) into abelian subgroups $M = \bigoplus_{j\in\mathbb{Z}} M^j$ such that $\Bbbk[t]_i \cdot M^j \subseteq M^{i+j}$ holds for all $i \in \mathbb{N}_0$ and $j \in \mathbb{Z}$. Let $\varepsilon \in \mathbb{N}_0$. An ε-*morphism of* $\Bbbk[t]$-*modules* M and N is a morphism $f : M \to N$ of the underlying $\Bbbk[t]_{(\mathrm{NGr})}$-modules such that for all $j \in \mathbb{Z}$ we have $f(M^j) \subseteq N^{j+\varepsilon}$. A 0-morphism is also called *a morphism.*

Example. There is a distinguished ε-morphism $\mathrm{id}_\varepsilon : M \to M$ given by $\mathrm{id}_\varepsilon(m) = t^\varepsilon m$.

Since $\Bbbk[t]_0 = \Bbbk$, any such M is also a $\mathbb{Z}$-graded $\Bbbk$-module. Consequently, some authors [34] call this a *graded* $\Bbbk[t]$-*module.* For us, "graded $\Bbbk[t]$-module" means something else (see Section 2.6).

Definition. Let $\varepsilon \in \mathbb{N}_0$. An ε-*interleaving of* $\Bbbk[t]$-*modules* M and N is a pair (ϕ, ψ) of ε-morphisms $\phi : M \to N$ and $\psi : N \to M$ such that $\phi\psi = \mathrm{id}_{2\varepsilon}$ and $\psi\phi = \mathrm{id}_{2\varepsilon}$. A 0-interleaving is the same as an isomorphism. If there is an ε-interleaving between M and N, we say that M and N are ε-*interleaved* and write $M \overset{\varepsilon}{\sim} N$.

Remark 2.5. We also work with interleavings of graded modules and chain complexes. These are defined by components, and for the latter, we additionally require that the interleaving maps are chain maps, i.e. that they commute with the differentials.

The notion of ε-interleaving defines an extended[5] metric between isomorphism classes of $\Bbbk[t]$-modules, i.e. it satisfies the following basic properties.

Proposition 2.6. *Suppose M, N and P are $\Bbbk[t]$-modules. Then the following properties hold.*

1. *Positive definiteness: $M \overset{0}{\sim} N$ holds if and only if $M \cong N$.*
2. *Symmetry: $M \overset{\varepsilon}{\sim} N$ implies $N \overset{\varepsilon}{\sim} M$.*
3. *Triangle inequality: $M \overset{\varepsilon_1}{\sim} N$ and $N \overset{\varepsilon_2}{\sim} P$ imply $M \overset{\varepsilon_1+\varepsilon_2}{\sim} P$.*

Proof. The first two properties are immediate. To show the third, let (ϕ, ψ) be an ε_1-interleaving of M and N and (η, θ) an ε_2-interleaving of N and P. Then, $(\eta\phi, \theta\psi)$ is an $(\varepsilon_1 + \varepsilon_2)$-interleaving of M and P. $\square$

Definition. The *interleaving distance* between $\Bbbk[t]$-modules M and N is defined by the formula

$$d_I(M, N) = \min\{\varepsilon \in \mathbb{N}_0 \mid M \overset{\varepsilon}{\sim} N\}.$$

[5] This means that we allow it to take the value ∞.

Therefore, the notion of interleaving provides a means to quantify how close two modules are to each other. Modules ε-interleaved with 0 are of particular importance, as they may be regarded as small, and are therefore useful as a model of experimental error. Alternatively, they are characterized as follows.

Proposition 2.7. *A $\Bbbk[t]$-module M is ε-interleaved with 0 if and only if $t^{2\varepsilon}M = 0$.*

Proof. If $M \overset{\varepsilon}{\sim} 0$, let (ϕ, ψ) be the interleaving. This means that $\mathrm{id}_{2\varepsilon} = \psi\phi = 0$. Conversely, if $t^{2\varepsilon}M = 0$, the interleaving is given by $(\phi, \psi) = (0, 0)$. □

This immediately implies that subquotients of small modules are small.

Corollary 2.8. *Let M be a $\Bbbk[t]$-module and P its subquotient. Then $M \overset{\varepsilon}{\sim} 0$ implies $P \overset{\varepsilon}{\sim} 0$.*

Proof. Let $P = N/_\sim$ for some $N \leq M$. Since $t^{2\varepsilon}M = 0$, we have $t^{2\varepsilon}N = 0$ and therefore $t^{2\varepsilon}P = 0$. □

In the context of persistence modules, it is useful to define the notions of interleavings categorically. Let **Vect** be the category of vector spaces over $\Bbbk$ and let $I \subseteq \mathbb{R}$ be closed under addition. Being a poset, I may be viewed as a category in the usual way. For each $\varepsilon \in I$ and $\varepsilon \geq 0$, there is a functor $T_\varepsilon : (I, \leq) \to (I, \leq)$ given by $T_\varepsilon(a) = a + \varepsilon$ and a natural transformation $\eta_\varepsilon : \mathrm{id} \Rightarrow T_\varepsilon$ given by $\eta_\varepsilon(a) : a \to a + \varepsilon$. These observations are due to Bubenik and Scott [13]. This leads to the following definition.

Definition. A *persistence module* is a functor $F : (I, \leq) \to \mathbf{Vect}$. For $\varepsilon \geq 0$, an *ε-morphism* $\phi : F \overset{\varepsilon}{\to} G$ is a natural transformation $\phi : F \Rightarrow G \circ T_\varepsilon$. A *morphism* is a 0-morphism. An ε_1-morphism and an ε_2-morphism can be composed in the natural way to yield a $(\varepsilon_1 + \varepsilon_2)$-morphism. An *$\varepsilon$-interleaving* is a pair (ϕ, ψ) of ε-morphisms $\phi : F \overset{\varepsilon}{\to} G$ and $\psi : F \overset{\varepsilon}{\to} G$ such that $\psi\phi = F\eta_{2\varepsilon}$ and $\phi\psi = G\eta_{2\varepsilon}$. We say F and G are *ε-interleaved*, $F \overset{\varepsilon}{\sim} G$.

We denote the corresponding functor category by $\mathbf{Vect}^{(I,\leq)}$. The notion of *interleaving distance* also makes sense in the setting of persistence modules and is defined by the analogous formula

$$d_I(F, G) = \inf\{\varepsilon \in I \cap [0, \infty) \mid F \overset{\varepsilon}{\sim} G\}.$$

Note, however, that the infimum is not necessarily attained in this case (see [18]).

A standard fact about persistence modules over $I = \mathbb{Z}$ is that they correspond in a natural way to $\Bbbk[t]$-modules. Let $\mathbf{Mod}_{\Bbbk[t]}$ denote the category of modules over $\Bbbk[t]$ (in a graded sense). Then, the following holds.

Theorem 2.9. *The categories $\mathbf{Vect}^{(\mathbb{Z},\leq)}$ and $\mathbf{Mod}_{\Bbbk[t]}$ are isomorphic.*

Proof. Inverse functors $\Phi : \mathbf{Vect}^{(\mathbb{Z},\leq)} \to \mathbf{Mod}_{\Bbbk[t]}$ and $\Psi : \mathbf{Mod}_{\Bbbk[t]} \to \mathbf{Vect}^{(\mathbb{Z},\leq)}$ can be defined explicitly. On objects, these are defined as $\Phi(F) = \bigoplus_{j\in\mathbb{Z}} F(j)$ and $\Psi(M)(j) = M^j$. On morphisms, we have $\Phi(\eta) = (\eta_j)_{j\in\mathbb{Z}}$ and $\Psi(f)_j = f_j$, where $f_j : M^j \to N^j$ is the restriction of f to the j-th step of the filtration. □

Note that this also holds for $I = \varepsilon\mathbb{Z}$, $\varepsilon > 0$. The correspondence between persistence modules over $I = \mathbb{Z}$ and $\Bbbk[t]$-modules was first noted in [64]. We use this extensively in this paper. There is a similar correspondence between persistence modules over $I = \mathbb{R}$ and modules over the monoid algebra over $[0, \infty)$, as noted by Lesnick [47]. However, for our application, as we shall see, this is unnecessary.

In particular, we can show that each persistence module over $\mathbb{R}$ can be approximated by a persistence module over $\varepsilon\mathbb{Z}$ up to ε. To make sense of this, first observe that there is a natural inclusion functor $i_\varepsilon : (\varepsilon\mathbb{Z}, \leq) \to (\mathbb{R}, \leq)$. This functor has a left inverse $p_\varepsilon : (\mathbb{R}, \leq) \to (\varepsilon\mathbb{Z}, \leq)$ given by $p_\varepsilon(a) = \lfloor \frac{a}{\varepsilon} \rfloor \varepsilon$. Note that this left inverse is not unique. In a sense, however, it is the most natural choice in our situation.

These two functors give rise to the (natural) restriction functor $I_\varepsilon : \mathbf{Vect}^{(\mathbb{R},\leq)} \to \mathbf{Vect}^{(\varepsilon\mathbb{Z},\leq)}$ given by $I_\varepsilon(F) = F i_\varepsilon$ and an extension functor $P_\varepsilon : \mathbf{Vect}^{(\varepsilon\mathbb{Z},\leq)} \to \mathbf{Vect}^{(\mathbb{R},\leq)}$ given by $P_\varepsilon(F) = F p_\varepsilon$. Under our choice of p_ε, when defined, the persistence diagrams of $F : (\varepsilon\mathbb{Z}, \leq) \to \mathbf{Vect}$ and $P_\varepsilon(F) : (\mathbb{R}, \leq) \to \mathbf{Vect}$ agree as multisets (except perhaps on the diagonal, depending on the convention used).

The functors I_ε and P_ε have various useful properties. Since $p_\varepsilon i_\varepsilon = \mathrm{id}$, we have $I_\varepsilon P_\varepsilon = \mathrm{id}$. The composition $P_\varepsilon I_\varepsilon$ is, in a certain sense, also not far from the identity. Furthermore, P_ε is an isometric embedding, and I_ε is an almost isometry.

Proposition 2.10. *Let $F : (\mathbb{R}, \leq) \to \mathbf{Vect}$ be a persistence module. Then F and $P_\varepsilon I_\varepsilon(F)$ are ε-interleaved.*

Proof. An ε-interleaving (ϕ, ψ) is given by $\phi_x : F(x) \to F(p_\varepsilon(x) + \varepsilon)$ and $\psi_x : F(p_\varepsilon(x)) \to F(x + \varepsilon)$, given by the shifting morphisms $\phi_x = \mathrm{id}_{p_\varepsilon(x)+\varepsilon-x}$ and $\psi_x = \mathrm{id}_{x+\varepsilon-p_\varepsilon(x)}$. □

Proposition 2.11. *The functor $P_\varepsilon : \mathbf{Vect}^{(\varepsilon\mathbb{Z},\leq)} \to \mathbf{Vect}^{(\mathbb{R},\leq)}$ is an isometric embedding.*

Proof. Suppose $F, G \in \mathbf{Vect}^{(\varepsilon\mathbb{Z},\leq)}$ and suppose $F \overset{\eta}{\sim} G$ and let (ϕ, ψ) be the relevant interleaving. Then, $(\phi p_\varepsilon, \psi p_\varepsilon)$ is an η-interleaving of $P_\varepsilon(F)$ and $P_\varepsilon(G)$. This implies that

$$\begin{aligned} d_I(F, G) &= \min\{\eta \in \varepsilon\mathbb{N}_0 \mid F \overset{\eta}{\sim} G\} \geq \inf\{\eta \in [0, \infty) \mid P_\varepsilon(F) \overset{\eta}{\sim} P_\varepsilon(G)\} \\ &= d_I(P_\varepsilon(F), P_\varepsilon(G)). \end{aligned}$$

To prove the converse inequality, suppose $P_\varepsilon(F)$ and $P_\varepsilon(G)$ are η-interleaved and let (ϕ, ψ) be the interleaving. We claim that this implies F and G are $p_\varepsilon(\eta)$-interleaved. To define an interleaving, note that for $k \in \mathbb{Z}$, $\phi_{k\varepsilon} : F(k\varepsilon) \to G(p_\varepsilon(k\varepsilon + \eta)) = G(k\varepsilon + p_\varepsilon(\eta))$ and $\psi_{k\varepsilon} : G(k\varepsilon) \to F(p_\varepsilon(k\varepsilon + \eta)) = F(k\varepsilon + p_\varepsilon(\eta))$, so the maps $\tilde{\phi}_{k\varepsilon} = \phi_{k\varepsilon}$ and $\tilde{\psi}_{k\varepsilon} = \psi_{k\varepsilon}$ are components of a $p_\varepsilon(\eta)$-interleaving $(\tilde{\phi}, \tilde{\psi})$, showing that $d_I(F, G) \leq d_I(P_\varepsilon(F), P_\varepsilon(G))$. □

Proposition 2.12. *Given persistence modules $F, G : (\mathbb{R}, \leq) \to \mathbf{Vect}$, we have*

$$d_I(F, G) - 2\varepsilon \leq d_I(I_\varepsilon(F), I_\varepsilon(G)) \leq d_I(F, G) + \varepsilon.$$

Proof. Let $A = \{\eta \in \varepsilon\mathbb{N}_0 \mid I_\varepsilon(F) \overset{\eta}{\sim} I_\varepsilon(G)\}$ and $B = \{\eta \in [0, \infty) \mid F \overset{\eta}{\sim} G\}$. Note that if two modules are η-interleaved, they are also θ-interleaved for all $\theta \geq \eta$, so these sets are upward closed in $\varepsilon\mathbb{N}_0$ and $[0, \infty)$, respectively. By definition, we have

$$d_I(I_\varepsilon(F), I_\varepsilon(G)) = \min A \qquad \text{and} \qquad d_I(F, G) = \inf B.$$

Suppose $\eta \in B \cap \varepsilon\mathbb{N}_0$ and let (ϕ, ψ) be the relevant η-interleaving. Then $(\phi i_\varepsilon, \psi i_\varepsilon)$ is an η-interleaving of $I_\varepsilon(F)$ and $I_\varepsilon(G)$. Therefore, $B \cap \varepsilon\mathbb{N}_0 \subseteq A$. Since B is upward closed, this immediately implies $\inf B \geq \min A - \varepsilon$ and therefore

$$d_I(I_\varepsilon(F), I_\varepsilon(G)) \leq d_I(F, G) + \varepsilon.$$

The other inequality follows from Propositions 2.11 and 2.10:

$$\begin{aligned} d_I(F, G) &\leq d_I(F, P_\varepsilon(I_\varepsilon(F))) + d_I(P_\varepsilon(I_\varepsilon(F)), P_\varepsilon(I_\varepsilon(G))) + d_I(P_\varepsilon(I_\varepsilon(G)), G) \\ &\leq d_I(P_\varepsilon(I_\varepsilon(F)), P_\varepsilon(I_\varepsilon(G))) + 2\varepsilon = d_I(I_\varepsilon(F), I_\varepsilon(G)) + 2\varepsilon. \end{aligned}$$

□

These observations allow us to compare persistence modules over $\varepsilon\mathbb{Z}$ and persistence modules over $\mathbb{R}$. Namely, since P_ε is an isometric embedding, $\varepsilon\mathbb{Z}$-persistence modules can be understood as a special case of $\mathbb{R}$-persistence modules, namely those satisfying the property $F(a \to b) = \mathrm{id}$ for any pair of points $a \leq b$ lying the same interval $[k\varepsilon, (k+1)\varepsilon)$. Therefore, we regard persistence modules $F : (\varepsilon\mathbb{Z}, \leq) \to \mathbf{Vect}$ and $G : (\mathbb{R}, \leq) \to \mathbf{Vect}$ as ε-close if $P_\varepsilon(F)$ and G are ε-interleaved. With this understanding, we may state:

Corollary 2.13. *For any $\varepsilon > 0$, any continuous-valued persistence module $F : (\mathbb{R}, \leq) \to \mathbf{Vect}$ can be ε-approximated by a $\Bbbk[t]$-module $F_\varepsilon : (\varepsilon\mathbb{Z}, \leq) \to \mathbf{Vect}$, namely $F_\varepsilon = I_\varepsilon(F)$.*

We concern ourselves with strictly positive ε. The connection between discrete and continuous parameter persistence was first exploited in the first algebraic persistence stability result [16] and has been studied in [62]. The related notion of *observable structure* was further introduced in [18]. In principle, this discretization is technically unnecessary but desirable in algorithmic applications (see Discussion).

As mentioned at the end of the preceding section, we would like to compare the persistence modules of a sublevel set filtration and a lower star filtration associated to the same piecewise linear function $\hat{f}$. The functorial approach is fruitful here, as the two filtrations may also be regarded as functors $S_{\hat{f}} : (\mathbb{R}, \leq) \to (\mathbf{Top}, \subseteq)$ and $L_{\hat{f}} : (\mathbb{R}, \leq) \to (\mathbf{SCx}, \subseteq)$, respectively.

Let H_n denote the n-th simplicial homology functor and H^s_n the n-th singular homology functor. Note that these are related by $\mathsf{H}_n \cong \mathsf{H}^s_n G$, where $G : (\mathbf{SCx}, \subseteq) \to (\mathbf{Top}, \subseteq)$ is the geometric realization functor.

Proposition 2.14. *Suppose X is a simplicial complex and $\hat{f} : |X| \to \mathbb{R}$ a piecewise linear function (w.r.t. the triangulation). Then the persistence modules $\mathsf{H}^s_n S_{\hat{f}} : (\mathbb{R}, \leq) \to \mathbf{Vect}$ and $\mathsf{H}_n L_{\hat{f}} : (\mathbb{R}, \leq) \to \mathbf{Vect}$ are isomorphic.*

Proof. There is a natural transformation $\eta : G \circ L_{\hat{f}} \Rightarrow S_{\hat{f}}$ given componentwise by the inclusions $|X^j| \to |X|^j$. However, since $|X^j|$ and $|X|^j$ are homotopy equivalent, the components of the natural transformation $\mathsf{H}^s_n \eta : \mathsf{H}_n L_{\hat{f}} \to \mathsf{H}^s_n S_{\hat{f}}$ are isomorphisms; therefore, it is a natural isomorphism. □

Another important fact about sublevel set filtrations is that the persistent homologies associated to a pair of ε-close functions on the same space are ε-interleaved. We recall a classical result.

Proposition 2.15. *Suppose Y is a topological space and $f, g : Y \to \mathbb{R}$ are functions satisfying $\|f - g\|_\infty \leq \varepsilon$. Then the persistence modules $\mathsf{H}_*(Y, f)$ and $\mathsf{H}_*(Y, g)$ are ε-interleaved.*

Proof. For each $x \in \mathbb{R}$, there are inclusions $f^{-1}(-\infty, x] \to g^{-1}(-\infty, x + \varepsilon]$ and $g^{-1}(-\infty, x] \to f^{-1}(-\infty, x + \varepsilon]$. Upon taking their homology, we obtain the desired ε-interleaving. □

Remark 2.6. This also holds for lower star filtrations, with completely analogous proof.

As we have seen, continuous persistence modules can be approximated by discrete ones. For this reason, we mostly work with $\Bbbk[t]$-modules in the remainder of the paper. To avoid notational clutter, we also adopt the following convention.

Convention. Both ordinary and persistent homology are denoted by the same symbol H_*. In case the filtration is explicitly mentioned, as in $\mathsf{H}_*(X, \mathcal{F})$ or $\mathsf{H}_*(X, f)$, the meaning is unambiguous. However, when suppressing the filtration, $\mathsf{H}_*(X)$ could in principle mean either the persistent homology of the filtered simplicial complex $(X, \mathcal{F})$ or the ordinary homology of its underlying space X. Whenever X has the structure of a filtered simplicial complex, $\mathsf{H}_*(X)$ will always mean persistent homology and $\mathsf{H}_*(X^j)$, with the filtration step explicitly specified (possibly $j = \infty$), will always mean ordinary homology.

2.6 Spectral Sequences

In this section, we introduce the concept of a *spectral sequence* and examine its various basic properties. Then, we focus our attention on the Mayer-Vietoris spectral sequence which is the one most suitable for our needs. Many spectral sequences arise from double complexes. A description of these spectral sequences can be found in [57, Chapter 10] and [50]. Versions of the Mayer-Vietoris spectral sequence can also be found in [11] and [12] among numerous others.

Definition. *A graded* $\Bbbk[t]$*-module* is a $\mathbb{Z}$-indexed family $M = M_* = (M_p)_{p\in\mathbb{Z}}$ of $\Bbbk[t]$-modules.

Definition. *A (chain) complex of* $\Bbbk[t]$*-modules* is a pair (C, ∂) where C is a graded $\Bbbk[t]$-module and $\partial = (\partial_p)_{p\in\mathbb{Z}}$ is a family of morphisms $\partial_p : \mathsf{C}_p \to \mathsf{C}_{p-1}$ of $\Bbbk[t]$-modules such that $\partial_{p-1}\partial_p = 0$ for each $p \in \mathbb{Z}$.

It is often convenient to view a graded $\Bbbk[t]$-module as a genuine $\Bbbk[t]$-module by identifying it with the direct sum of its components $M \equiv \bigoplus_{p\in\mathbb{Z}} M_p$. The decomposition is regarded as part of the structure. Similarly, we often view a chain complex as a *differential graded module*, i.e. the $\Bbbk[t]$-module $\mathsf{C} \equiv \bigoplus_{p\in\mathbb{Z}} \mathsf{C}_p$ equipped with a $\Bbbk[t]$-module homomorphism ∂ such that $\partial(\mathsf{C}_p) \subseteq \mathsf{C}_{p-1}$ and $\partial \circ \partial = 0$.

Definition. *A bigraded* $\Bbbk[t]$*-module* is a $\mathbb{Z}\times\mathbb{Z}$-indexed family $M = M_{*,*} = (M_{p,q})_{p,q\in\mathbb{Z}}$ of $\Bbbk[t]$-modules.

Definition. *A double complex (bicomplex) of* $\Bbbk[t]$*-modules* is a triple $(M, \partial^0, \partial^1)$ where M is a bigraded $\Bbbk[t]$-module and $\partial^0 = (\partial^0_{p,q})_{p,q\in\mathbb{Z}}$ and $\partial^1 = (\partial^1_{p,q})_{p,q\in\mathbb{Z}}$ are two families of morphisms $\partial^0_{p,q} : M_{p,q} \to M_{p,q-1}$ and $\partial^1_{p,q} : M_{p,q} \to M_{p-1,q}$ such that $\partial^0_{p,q-1}\partial^0_{p,q} = 0, \partial^1_{p-1,q}\partial^1_{p,q} = 0$ and $\partial^1_{p,q-1}\partial^0_{p,q} + \partial^0_{p-1,q}\partial^1_{p,q} = 0$ for $p, q \in \mathbb{Z}$.

Note that the notions of ε-morphisms and interleavings make sense for bigraded modules and therefore for spectral sequences. We may define them by components and, in the case of double complexes, additionally assume that they commute with both differentials.

As with graded modules and complexes, we often view bigraded $\Bbbk[t]$-modules as genuine $\Bbbk[t]$-modules with additional structure, namely via the identification $M = \bigoplus_{p,q\in\mathbb{Z}} M_{p,q}$. We can also view M as a graded $\Bbbk[t]$-module in (at least) two ways, namely by summing over all p or by summing over all q. Using this view, a double complex can be seen as a bigraded module M that is a differential module with respect to ∂^0 as well as with respect to ∂^1, and the two structures are related by the equation $\partial^0\partial^1 + \partial^1\partial^0 = 0$.

The relevance of this anticommutativity property is that combining the two differentials by summing them also yields a differential $\partial^0 + \partial^1$. In fact, we may equivalently work with *commutative double complexes* $(M, \partial^0, \partial^1)$ with the only difference that ∂^0 and ∂^1 commute instead of anticommute, i.e. $\partial^0\partial^1 = \partial^1\partial^0$. Note that such M becomes an anticommutative double complex upon replacing ∂^0 by $(-1)^p\partial^0$. The advantage of the anticommutative case is that we do not have to keep track of signs in the combined differential.

$$
\begin{tikzcd}
E^0_{0,3} \arrow[d, "\partial^0"'] & E^0_{1,3} \arrow[l, "\partial^1"'] \arrow[d, "\partial^0"'] & E^0_{2,3} \arrow[l, "\partial^1"'] \arrow[d, "\partial^0"'] & E^0_{3,3} \arrow[l, "\partial^1"'] \arrow[d, "\partial^0"'] \\
E^0_{0,2} \arrow[d, "\partial^0"'] & E^0_{1,2} \arrow[l, "\partial^1"'] \arrow[d, "\partial^0"'] & E^0_{2,2} \arrow[l, "\partial^1"'] \arrow[d, "\partial^0"'] & E^0_{3,2} \arrow[l, "\partial^1"'] \arrow[d, "\partial^0"'] \\
E^0_{0,1} \arrow[d, "\partial^0"'] & E^0_{1,1} \arrow[l, "\partial^1"'] \arrow[d, "\partial^0"'] & E^0_{2,1} \arrow[l, "\partial^1"'] \arrow[d, "\partial^0"'] & E^0_{3,1} \arrow[l, "\partial^1"'] \arrow[d, "\partial^0"'] \\
E^0_{0,0} & E^0_{1,0} \arrow[l, "\partial^1"'] & E^0_{2,0} \arrow[l, "\partial^1"'] & E^0_{3,0} \arrow[l, "\partial^1"']
\end{tikzcd}
$$

Figure 1: A double complex comes equipped with two differentials ∂^0 and ∂^1. Considering the antidiagonals, we also obtain a chain complex, called the total complex with $\partial^0\partial^1 + \partial^1\partial^0 = 0$ by anticommutativity.

To each double complex, one may associate a total complex by summing over the antidiagonals and combining the two boundary operators into a total boundary operator. Note that the n-th antidiagonal is the direct sum of all entries in the double complex such that $p + q = n$. This leads to the following definition.

Definition. Let M be a double complex. The *total complex* $(\mathrm{Tot}(M), D)$ associated to M is the chain complex defined by $\mathrm{Tot}_n(M) = \bigoplus_{p+q=n} M_{p,q}$ and $D = \partial^0 + \partial^1$.

Spectral sequences are a tool that allows us to compute the homology of this total complex. This is a very common situation in practice. Suppose we are given a chain complex (C, ∂) whose homology we would like to compute. It is often possible to find a natural filtration of such a complex. By taking successive quotients, one then obtains a double complex M, whose total complex is isomorphic to the original chain complex. In particular, their homologies agree:

$$\mathsf{H}_*(\mathrm{Tot}(M), D) \cong \mathsf{H}_*(\mathsf{C}, \partial).$$

The homology of such a complex (C, ∂) can therefore be computed systematically using the associated spectral sequence. In fact, this is precisely what happens in our case. The associated *spectral sequence* consists of *pages*, where each page E^r, $r = 0, 1, \ldots$, is a differential bigraded module, computed successively by taking the homology with respect to the differential on the previous page. On the r-th page, the differential is given by

$$d^r : E^r_{p,q} \to E^r_{p-r,q+r-1}.$$

It may happen that there is a R such that for $r > R$ all differentials beginning or ending at $E^r_{p,q}$ are zero maps. In this case the (p, q)-th component stabilizes in the sense that all these modules $E^r_{p,q}$ are isomorphic. If such a R exists for each pair (p, q), the spectral sequence is said to *converge* and the stabilized modules $E^r_{p,q}$ are denoted by $E^\infty_{p,q}$. In this case, the bigraded module with components $E^\infty_{p,q}$ is called the ∞-page of the spectral sequence. If $E^N = E^\infty$ for some finite N, the spectral sequence is said to *collapse* on the N-th page.

Each successive page of the spectral sequence provides a successively better approximation of the homology of the total complex, so if the spectral sequence of a double complex M converges, it is said to converge to $\mathsf{H}_*(\mathrm{Tot}(M))$. In practice, this means that $\mathsf{H}_*(\mathrm{Tot}(M))$ can be reconstructed from the E^∞ page. In particular, if E^∞ consists of free modules, $\mathsf{H}_n(\mathrm{Tot}(M))$ is isomorphic to $\bigoplus_{p+q=n} E^\infty_{p,q}$. Generally, however, the relation between $\mathsf{H}_*(\mathrm{Tot}(M))$ and E^∞ is slightly more complicated. If a spectral sequence converges to $\mathsf{H}_*(\mathrm{Tot}(M))$, then there exists a filtration

$$\mathsf{H}_{p+q}(\mathrm{Tot}(M))^0 \subseteq \mathsf{H}_{p+q}(\mathrm{Tot}(M))^1 \subseteq \ldots \subseteq \mathsf{H}_{p+q}(\mathrm{Tot}(M))^p \subseteq \ldots \subseteq \mathsf{H}_*(\mathrm{Tot}(M)) \tag{1}$$

and the E^∞ consists of successive quotients of various steps of the filtration of $\mathsf{H}_*(\mathrm{Tot}(M))$ arising from the structure of the double complex:

$$E^\infty_{p,q} \cong \frac{\mathsf{H}_{p+q}(\mathrm{Tot}(M))^p}{\mathsf{H}_{p+q}(\mathrm{Tot}(M))^{p-1}}.$$

Note that the p here denotes the position in the filtration which coincides with the column of the double complex. It is straightforward to check that in our case, the spectral sequences are convergent.

Hence, reconstructing $\mathsf{H}_*(\mathrm{Tot}(M))$ up to isomorphism from E^∞ in general requires us to solve a series of extension problems over each antidiagonal $p + q = n$. For our particular spectral sequence (an upper quadrant spectral sequence of a double complex), it is known that the filtration has two additional properties which follow from the explicit description in the Appendix (see also [50]), namely

$$\mathsf{H}_n(\mathrm{Tot}(M))^{-1} = 0 \qquad \text{and} \qquad \mathsf{H}_n(\mathrm{Tot}(M))^n = \mathsf{H}_n(X).$$

The first three steps in a spectral sequence are shown in Figure 2. The spectral sequence relevant to our needs is called the Mayer-Vietoris spectral sequence. It is a first quadrant spectral sequence, meaning that $E^r_{p,q} = 0$ if either $p < 0$ or $q < 0$. Note that first quadrant spectral sequences always converge, since eventually all differentials beginning or ending at a particular (p, q) in the first quadrant will point outside this quadrant. The Mayer-Vietoris spectral sequence is defined as the spectral sequence of a particular double complex arising from a cover of the space whose homology we are interested in.

$$\begin{tikzcd}
E^0_{0,3} \arrow[d, "d^0"'] & E^0_{1,3} \arrow[d, "d^0"'] & E^0_{2,3} \arrow[d, "d^0"'] & E^0_{3,3} \arrow[d, "d^0"'] \\
E^0_{0,2} \arrow[d, "d^0"'] & E^0_{1,2} \arrow[d, "d^0"'] & E^0_{2,2} \arrow[d, "d^0"'] & E^0_{3,2} \arrow[d, "d^0"'] \\
E^0_{0,1} \arrow[d, "d^0"'] & E^0_{1,1} \arrow[d, "d^0"'] & E^0_{2,1} \arrow[d, "d^0"'] & E^0_{3,1} \arrow[d, "d^0"'] \\
E^0_{0,0} & E^0_{1,0} & E^0_{2,0} & E^0_{3,0}
\end{tikzcd}
\quad
\begin{tikzcd}
E^1_{0,3} & E^1_{1,3} \arrow[l, "d^1"'] & E^1_{2,3} \arrow[l, "d^1"'] & E^1_{3,3} \arrow[l, "d^1"'] \\
E^1_{0,2} & E^1_{1,2} \arrow[l, "d^1"'] & E^1_{2,2} \arrow[l, "d^1"'] & E^1_{3,2} \arrow[l, "d^1"'] \\
E^1_{0,1} & E^1_{1,1} \arrow[l, "d^1"'] & E^1_{2,1} \arrow[l, "d^1"'] & E^1_{3,1} \arrow[l, "d^1"'] \\
E^1_{0,0} & E^1_{1,0} \arrow[l, "d^1"'] & E^1_{2,0} \arrow[l, "d^1"'] & E^1_{3,0} \arrow[l, "d^1"']
\end{tikzcd}
\quad
\begin{tikzcd}
E^2_{0,3} & E^2_{1,3} & E^2_{2,3} & E^2_{3,3} \\
E^2_{0,2} & E^2_{1,2} & E^2_{2,2} \arrow[ull, "d^2"] & E^2_{3,2} \arrow[ull, "d^2"] \\
E^2_{0,1} & E^2_{1,1} & E^2_{2,1} \arrow[ull, "d^2"] & E^2_{3,1} \arrow[ull, "d^2"] \\
E^2_{0,0} & E^2_{1,0} & E^2_{2,0} \arrow[ull, "d^2"] & E^2_{3,0} \arrow[ull, "d^2"]
\end{tikzcd}$$

Figure 2: The differentials for the first three pages of the spectral sequence. In each case, to compute the next page we take homology with respect to the differential on the current page. We set $d^0 = \partial^0$ and d^1 is the homomorphism induced by ∂^1.

Now, suppose we are given a pair $(X, \mathcal{U})$, where X is a filtered simplicial complex and $\mathcal{U} = (U_i, \mathcal{F}_i)_{i \in \Lambda}$ is a filtered cover of X by subcomplexes. To any such pair, we may associate a commutative double complex $(E^0, \partial^0, \partial^1)$ where the underlying bigraded module is given by (recall that U_I is given the filtration $\mathcal{F}_I$)

$$E^0_{p,q} = \bigoplus_{|I|=p+1} \mathsf{C}_q(U_I)$$

and the boundary maps $\partial^0_{p,q} : E^0_{p,q} \to E^0_{p,q-1}$ and $\partial^1_{p,q} : E^0_{p,q} \to E^0_{p-1,q}$ are defined on the simplices by

$$\partial^0_{p,q}(\sigma, I) = \sum_{k=0}^{q} (-1)^k t^{\deg(\sigma,I) - \deg(\sigma_k, I)} (\sigma_k, I)$$

and

$$\partial^1_{p,q}(\sigma, I) = \sum_{l=0}^{p} (-1)^l t^{\deg(\sigma,I) - \deg(\sigma, I_l)} (\sigma, I_l).$$

These formulae require some explanation. To simplify things, we choose total orderings on the set V of vertices of X and the index set Λ of the cover $\mathcal{U}$. Note that Λ is the set of vertices of the nerve $\mathcal{N}$ of $\mathcal{U}$. These total orders of V and Λ allow us to speak unambiguously of "the k-th vertex of σ" and "l-th vertex of I." The simplices are denoted as pairs (σ, I) to distinguish between two copies of σ corresponding to different summands in $E^0_{p,q}$. Each simplex $\sigma = \{v_0, \ldots, v_q\} \in U_I$ has a birth time $\deg(\sigma, I)$ in the filtration of U_I. As usual, $\sigma_k := \sigma - \{v_k\}$ are the faces of codimension 1 in σ. Since $I = \{i_0, \ldots, i_p\}$ is a p-simplex in the nerve, it also makes sense to think of $I_l := I - \{i_l\}$ as the faces of codimension 1 in I.

It is a standard fact that E^0 is indeed a chain complex with respect to ∂^0 and ∂^1. Furthermore, ∂^0 and ∂^1 commute, since the first only operates on the chains of X, whereas the second operates on the chains of the nerve $\mathcal{N}$. Hence, replacing ∂^0 by $(-1)^p\partial^0$ yields a double complex. The spectral sequence (E^r, d^r) associated to this double complex is called *the Mayer-Vietoris spectral sequence of* $(X, \mathcal{U})$.

This double complex is designed so that its homology is precisely the homology of $(X, \mathcal{F})$, implying the following fact, which is the main reason for the importance of the Mayer-Vietoris spectral sequence.

Theorem 2.16. *The Mayer-Vietoris spectral sequence of* $(X, \mathcal{U})$ *converges to* $\mathsf{H}_*(X)$.

This result can be found in [48] and variations can be found in [11, 12, 38]. For completeness, we include the idea of proof in the Appendix. The first page of the Mayer-Vietoris spectral sequence can be expressed as follows. Note that d^0 is simply ∂^0, which acts on each summand as the simplicial boundary operator, therefore

$$E^1_{p,q} = \bigoplus_{|I|=p+1} \mathsf{H}_q(U_I). \tag{2}$$

The boundary map d^1 is induced by ∂^1. Explicitly, representing homology classes by cycles, we have:

$$d^1_{p,q}\left(\left[\sum_{n=0}^{N} \lambda_n t^{\mu_n}\sigma_n\right], I\right) = \sum_{l=0}^{p}(-1)^l \left(\left[\sum_{n=0}^{N} \lambda_n t^{\mu_n+\deg(\sigma_n,I)-\deg(\sigma_n,I_l)}\sigma_n\right], I_l\right).$$

The only case we really need is $q = 0$. In this case, the explicit formula can be simplified and has the same form as that of $\partial^1_{p,0}$, the only difference being that it is defined on homology classes instead of simplices:

$$d^1_{p,0}([v], I) = \sum_{l=0}^{p}(-1)^l t^{\deg(v,I)-\deg(v,I_l)}([v], I_l).$$

In the case of induced covers (see Remark 2.4), the explicit formula of $d^1_{p,q}$ has the same form as that of $\partial^1_{p,q}$ for all q. Computing E^2 is also straightforward, simply take the homology with respect to d^1. The higher pages require us to compute the higher differentials, which usually requires more work.

For illustrative purposes, we now prove the persistent nerves theorem of Sheehy [60, Theorem 6] using spectral sequences. In [60], this is proved by using the persistent nerve lemma of Chazal and Oudot [22, Lemma 3.4.]. Our proof also uses the idea of Chazal and Oudot, namely the fact that the Mayer-Vietoris blowup complex associated to $(X, \mathcal{U})$ is homotopy equivalent to X is used to establish Theorem 2.16 (see Appendix). Our theorems are motivated by and can be thought of as a generalization of this proof (recall that $\mathsf{H}_*(\cdot)$ denotes persistent homology). We begin with a preliminary lemma.

Lemma 2.17. *Suppose the chain complexes* (C', ∂') *and* $(\mathsf{C}'', \partial'')$ *are* ε*-interleaved as chain complexes. Then their homologies* $\mathsf{H}'_* = \mathsf{H}_*(\mathsf{C}')$ *and* $\mathsf{H}''_* = \mathsf{H}_*(\mathsf{C}'')$ *are* ε*-interleaved as graded modules.*

Proof. Let $\phi : \mathsf{C}' \to \mathsf{C}''$ and $\psi : \mathsf{C}'' \to \mathsf{C}'$ be the interleaving maps. Since (ϕ, ψ) is an interleaving of chain complexes, ϕ and ψ preserve cycles and boundaries. Therefore, the restrictions $\phi_{\mathsf{Z}} : \mathsf{Z}' \to \mathsf{Z}''$ and $\psi_{\mathsf{Z}} : \mathsf{Z}'' \to \mathsf{Z}'$ of the interleaving maps define an ε-interleaving $(\phi_{\mathsf{Z}}, \psi_{\mathsf{Z}})$ of the cycle modules, and the restrictions $\phi_{\mathsf{B}} : \mathsf{B}' \to \mathsf{B}''$ and $\psi_{\mathsf{B}} : \mathsf{B}'' \to \mathsf{B}'$ provide an ε-interleaving $(\phi_{\mathsf{B}}, \psi_{\mathsf{B}})$ of the boundary modules. These also descend to the level of quotients, i.e. we may define an ε-interleaving $(\phi_{\mathsf{H}}, \psi_{\mathsf{H}})$ of H'_* and H''_* by the formulae

$$\phi_{\mathsf{H}}([x]) = [\phi_{\mathsf{Z}}(x)] \qquad \text{and} \qquad \psi_{\mathsf{H}}([x]) = [\psi_{\mathsf{Z}}(x)].$$

It is readily verified that these maps are well defined and provide the appropriate interleaving. □

This leads us immediately to the persistent nerves theorem. Here, the notion of acyclicity is analogous to the usual one, i.e. each non-empty intersection has the persistent homology of a point (see Section 6.1).

Theorem 2.18. *Suppose X is a filtered simplicial complex and $\mathcal{U}$ a persistently acyclic filtered cover of X. Then, $\mathsf{H}_*(X) \cong \mathsf{H}_*(\mathcal{N}(\mathcal{U}))$.*

Proof. We use the Mayer-Vietoris spectral sequence E associated to $(X, \mathcal{U})$. Let (C, ∂) be the simplicial chain complex of the nerve $\mathcal{N}$ of $\mathcal{U}$. The boundary operator is given by the explicit formula

$$\partial_p(I) = \sum_{l=0}^{p} (-1)^l t^{\deg I - \deg I_l} I_l.$$

This has the same form as the boundary operators d^1 in the bottom row of E^1 and ∂^1 in the bottom row of the double complex. In particular,

$$d^1_{p,0}([v], I) = \sum_{l=0}^{p} (-1)^l t^{\deg(v,I) - \deg(v,I_l)} ([v], I_l).$$

In fact, (C, ∂) and $(E^1_{*,0}, d^1_{*,0})$ are isomorphic as chain complexes, the inverse isomorphisms $\phi_p : E^1_{p,0} \to \mathsf{C}_p$ and $\psi_p : \mathsf{C}_p \to E^1_{p,0}$ being given by

$$\phi_p([v], I) = t^{\deg v - \deg I} I \qquad \text{and} \qquad \psi_p(I) = ([v_I], I),$$

where we choose a vertex $v_I \in V$ with the property $\deg v_I = \deg I$. Note that ψ is well defined, because U_I is acyclic: if $v \neq v_I$ is another vertex with $\deg v = \deg I$, it belongs to the same homology class as v_I. That ϕ and ψ are inverse to each other follows by direct computation.

By Lemma 2.17, this implies that $E^2_{*,0} \cong \mathsf{H}_*(\mathcal{N})$. Using the fact that all U_I are acyclic, we have that $E^1_{p,q} = 0$ for $q > 0$, so the higher differentials d^r for $r > 1$ are all trivial and therefore $E^2 \cong E^\infty$. As all modules above the bottom row are zero, there are also no extension problems, so the conclusion follows.

In more detail, by Theorem 2.16, there is a filtration $(\mathsf{H}_*(X)^p)_{p \in \mathbb{Z}}$, defined on $\mathsf{H}_*(X)$, such that

$$E^\infty_{p,q} \cong \frac{\mathsf{H}_{p+q}(X)^p}{\mathsf{H}_{p+q}(X)^{p-1}} = \begin{cases} 0; & q \neq 0, \\ \mathsf{H}_{p+q}(\mathcal{N}); & q = 0. \end{cases}$$

Therefore, applying the third isomorphism theorem to the modules $\mathsf{H}_n(X)^{p-1} \subseteq \mathsf{H}_n(X)^p \subseteq \mathsf{H}_n(X)^n$ for each $p = 0, \ldots, n-1$ (see Equation (1)) and recalling the fact that $\mathsf{H}_n(X)^{-1} = 0$ and $\mathsf{H}_n(X)^n = \mathsf{H}_n(X)$, we have

$$\mathsf{H}_n(X) \cong \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{-1}} \cong \ldots \cong \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{n-2}} \cong \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{n-1}} \cong E^{\infty}_{n,0} \cong \mathsf{H}_n(\mathcal{N}),$$

as desired. $\square$

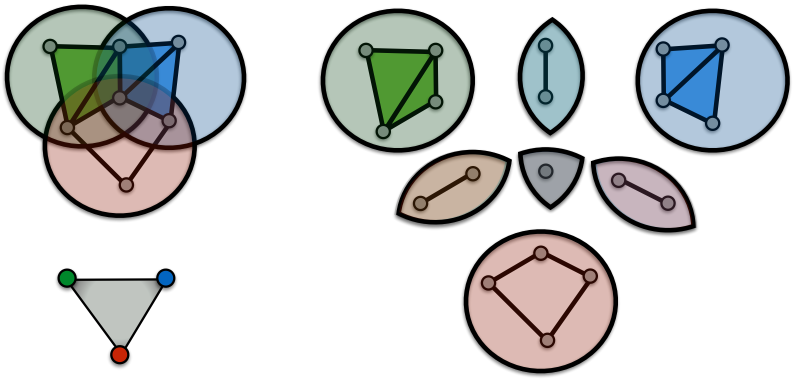


Figure 3: An example construction. On the left, we have a simplicial complex which is covered by a cover with three elements. The corresponding nerve is shown below it as it is a triangle. This is **not** an example of a good cover. The zeroth column of the double complex consists of a direct sum of the subcomplexes which lie in each individual element, the first column, the subcomplexes which lie in the pairwise intersections and the finally the second column contains the triple intersection. The row index represents the dimension grading from the underlying complex, i.e. vertices in the zeroth row, edges in the first row and triangles in the second. Note that the total complex has potentially multiple copies of a simplex and is much larger than the original complex.

3 Computing the Unimodal Category

Given a function $f : X \to [0, \infty)$, the most basic question we are interested in is how to compute its unimodal category $\mathbf{ucat}(f)$. This has been answered for functions $f : \mathbb{R} \to [0, \infty)$ with finitely many critical points by Baryshnikov and Ghrist [4, Theorem 11], using a simple sweeping algorithm. We show that this sweeping algorithm can be seen as arising from the Jordan decomposition theorem for functions of bounded variation. This gives a complete answer to the question in the case of $X = \mathbb{R}$. We can use the same idea to answer the question when $X = S^1$.

3.1 Real Line and Intervals

We show that the method to obtain a minimal unimodal decomposition for a compactly supported continuous function $f : \mathbb{R} \to [0, \infty)$ described in [4] in the case of finitely many critical points actually works for an arbitrary $f : \mathbb{R} \to [0, \infty)$. First, observe that the problem is only interesting for functions of bounded variation:

Proposition 3.1. *Suppose $f : \mathbb{R} \to [0, \infty)$ has a unimodal decomposition $f = \sum_{i=1}^n u_i$. Then f is of bounded variation on each interval $[a, b] \subseteq \mathbb{R}$.*

Proof. Let $[a, b]$ be an arbitrary interval. Then

$$V(f; [a, b]) = V\left(\sum_{i=1}^n u_i; [a, b]\right) \leq \sum_{i=1}^n V(u_i; [a, b]) < \infty,$$

since unimodal functions are of bounded variation. □

Remark 3.1. This also holds if the function only has a locally finite decomposition into unimodal summands, even though we may have $\mathbf{ucat}(f) = \infty$. An example of this kind is given by $f(x) = 1 + \sin x$. Note, however, that the converse of the above statement is not true even if f is compactly supported. Consider for instance

$$f(x) = \begin{cases} x^2(1 + \sin \frac{\pi}{x}); & x \in [-\frac{2}{5}, \frac{2}{3}], \\ 0; & \text{otherwise}, \end{cases}$$

which is compactly supported, has bounded variation, but $\mathbf{ucat}(f) = \infty$ as the support of any unimodal function $u \leq f$ must be contained between two zeros of f. For a complete characterization of functions $f : \mathbb{R} \to [0, \infty)$ with finite unimodal category, see Theorem 3.7 below.

We may therefore restrict our attention to functions of bounded variation:

Convention. In Section 3.1, in addition to continuity, we shall also assume that all functions $f : \mathbb{R} \to [0, \infty)$ are of bounded variation (but not necessarily compactly supported).

Remark 3.2. The fact that unimodal functions have bounded variation is specific to $\mathbb{R}$. It is not difficult to construct a unimodal function $u : \mathbb{R}^2 \to [0, \infty)$ which does not have bounded variation. Take any differentiable function $f : (0, 1) \to \mathbb{R}$ whose total variation is unbounded and $0 \le f(x) \le 1$ for all x. Further assume that the limits $\lim_{x\to 0} f(x)$ and $\lim_{x\to 1} f(x)$ exist. For instance, one might take

$$f(x) = x(1 + \sin \tfrac{\pi}{x}).$$

Define $u : [0,1] \times [0,1] \to [0,\infty)$ on $(0,1)\times(0,1)$ by

$$u(x,t) = tf(x) + (1-t)$$

and by its unique continuous extension elsewhere. This is a unimodal function but does not have bounded variation. Recall that the total variation of a differentiable function can be computed as the integral of the norm of its gradient:

$$\begin{aligned} V(u;[0,1]\times[0,1]) &= \int_0^1\int_0^1 \sqrt{u_x^2+u_t^2}\,\mathrm{d}x\mathrm{d}t \\ &\ge \int_0^1\int_0^1 |u_x|\mathrm{d}x\mathrm{d}t \\ &= \int_0^1 t\mathrm{d}t \int_0^1 |f'(x)|\mathrm{d}x \\ &= \tfrac{1}{2}V(f;(0,1)) = \infty. \end{aligned}$$

If desired, any such function u can also be continuously extended to a compactly supported unimodal function $\overline{u} : \mathbb{R}^2 \to [0,\infty)$. (This can be shown, for instance, by using Lemma 5.2 and the fact that $\mathbf{ucat}(f)$ is invariant under homeomorphisms of the domain.)

To see that the method of Ghrist and Baryshnikov generalizes, we first suitably modify Proposition 10 of [4]. First, recall Definition 7 of [4]:

Definition. Let $\mathcal{D} = (u_i)_{i=1}^n$ be a unimodal decomposition of $f : \mathbb{R} \to [0,\infty)$. An open interval (x,y) is *$\mathcal{D}$-max-free* if it contains no local maxima of any of u_i.

In [4], the authors only use this concept for intervals bounded by local minima of f. However, it is more generally applicable:

Proposition 3.2. *If an* arbitrary *open interval (x,y) is $\mathcal{D}$-max-free, where $\mathcal{D} = (u_i)_{i=1}^n$ is a unimodal decomposition of $f : \mathbb{R} \to [0,\infty)$, then*

$$V^-(f;(x,y)) \le f(x).$$

Proof. The idea is the same as in [4]:

$$V^-(f;(x,y)) \le \sum_{i=1}^n V^-(u_i;(x,y)) = \sum_{i=1}^n \max\{0, u_i(x) - u_i(y)\} \le \sum_{i=1}^n u_i(x) = f(x),$$

where the first equality uses the max-free condition. □

So, the concept of **forced-max** interval from [4] makes sense even if the interval is not bounded by local minima:

Definition. An interval (x, y) is called **forced-max** (with respect to f) if

$$V^{-}(f;(x,y)) > f(x).$$

In addition to this, we will call an interval (x, y) **almost forced-max** if $(x, y+\delta)$ is forced-max for each $\delta > 0$. If $f : \mathbb{R} \to [0, \infty)$ is compactly supported, we define $M(f)$ to be the maximum number of pairwise disjoint open intervals that are forced-max with respect to f, and $\widetilde{M}(f)$ to be the maximum number of pairwise disjoint open intervals that are almost forced-max with respect to f. (It is important to use open intervals here, as closed intervals with the same endpoints could intersect.)

More precisely, we should separate two cases here: if there is a finite bound to the number of such intervals, the maximum indeed exists and can be realized by a collection of such intervals; however, if there is no upper bound, this does not by itself imply the existence of an infinite collection of such intervals. We resolve this issue in Proposition 3.3, where we show that such a collection does indeed always exist.

The observation from [4] that (almost) forced-max intervals form an ideal in the sense that an interval containing an (almost) forced-max interval is itself (almost) forced-max, remains valid in this context. We now show that the numbers $M(f)$ and $\widetilde{M}(f)$ always agree. For this reason we only speak of $M(f)$ in the rest of the text.

Proposition 3.3. *If $f : \mathbb{R} \to [0, \infty)$ is a compactly supported function, we have*

$$\widetilde{M}(f) = M(f).$$

Proof. The inequality $\widetilde{M}(f) \geq M(f)$ trivially follows from the definitions. It remains to show that $\widetilde{M}(f) \leq M(f)$. We do this by an inductive construction. Suppose $(x_0, y_0), ..., (x_{n-1}, y_{n-1})$ is an arbitrary family of pairwise disjoint almost forced-max intervals. We may assume without loss of generality that $x_0 < x_1 < \ldots < x_n := y_{n-1}$ and that $y_i = x_{i+1}$ for $i = 1, \ldots, n-1$ and $x_0 = \inf \operatorname{supp} f$, $x_n = \sup \operatorname{supp} f$. We claim that it is possible to choose $\delta_1, \ldots, \delta_{n-1} > 0$ such that the intervals $(x_0, x_1 + \delta_1), (x_1 + \delta_1, x_2 + \delta_2), \ldots, (x_{n-1} + \delta_{n-1}, x_n)$ are forced-max and clearly these are still disjoint.

We prove this by induction from n downwards. For the base case, we do not actually change anything, we just note that (x_{n-1}, x_n) is already forced-max. This is because $V^{-}(f; [x_{n-1}, x_n]) = V^{-}(f; [x_{n-1}, x_n + \delta])$, since f is zero on $[x_n, \infty)$. Assuming that $\delta_k, \ldots, \delta_{n-1}$ (possibly none of them in the case $k = n$) have already been chosen such that $(x_{k-1}, x_k + \delta_k), \ldots, (x_{n-1} + \delta_{n-1}, x_n)$ are forced-max, we observe that since $V^{-}(f; [x_{k-1}, x_k + \delta_k]) > f(x_{k-1})$ and f is continuous, a $\delta_{k-1} > 0$ exists such that $V^{-}(f; [x_{k-1} + \delta_{k-1}, x_k + \delta_k]) > f(x_{k-1} + \delta_{k-1})$. The interval $(x_{k-2}, x_{k-1} + \delta_{k-1})$ thus becomes genuinely forced-max. $\square$

Furthermore, Theorem 11 in [4] states that $\mathbf{ucat}(f) = M(f)$ for functions that are nice enough. This can be also generalized to our context. In fact, we are going to explicitly construct a unimodal decomposition of f. This is achieved by a recursive construction, which generalizes the sweeping algorithm from [4]. For convenience, we first construct a minimal unimodal decomposition in the compactly supported case. The general case is then treated in Proposition 3.8 below.

Theorem 3.4. *If $f : \mathbb{R} \to [0, \infty)$ is compactly supported, then*

$$\mathbf{ucat}(f) = M(f)$$

holds. If $M(f) = n < \infty$, then an explicit minimal unimodal decomposition $(u_i)_{i=1}^n$ is obtained by the following procedure. First, extend f to $\overline{\mathbb{R}} = [-\infty, \infty]$ and define $g, h : \overline{\mathbb{R}} \to [0, \infty)$ by

$$g(x) = V^+(f; (-\infty, x]) \qquad \text{and} \qquad h(x) = V^-(f; (-\infty, x]).$$

Recursively define a finite sequence[6] $(x_i)_{i=0}^n$:

$$\begin{aligned} x_0 &= -\infty, \\ x_i &= \inf\{x \mid V^-(f; (x_{i-1}, x)) > f(x_{i-1})\}, \qquad i = 1, \ldots, n, \\ x_{n+1} &= \infty. \end{aligned}$$

Finally, define $u_i : \mathbb{R} \to [0, \infty)$ by

$$u_i(x) = \begin{cases} 0; & x \leq x_{i-1}, \\ g(x) - g(x_{i-1}); & x \in [x_{i-1}, x_i], \\ h(x_{i+1}) - h(x); & x \in [x_i, x_{i+1}], \\ 0; & x \geq x_{i+1}. \end{cases}$$

Proof. The inequality $\mathbf{ucat}(f) \geq M(f)$ follows directly from the definitions. This is because given any unimodal decomposition of f, each forced-max interval must contain a local maximum of some unimodal summand in this decomposition, hence there are at least as many summands as there are disjoint forced-max intervals. In particular, if $M(f) = \infty$, we immediately have $\mathbf{ucat}(f) = \infty$.

Now, assume $M(f) < \infty$ and let $M(f) = n \in \mathbb{N}_0$. Note that $f = g - h$ and that g and h are increasing functions. This latter fact implies that each function u_i is increasing on $(-\infty, x_i]$ and decreasing on $[x_i, \infty)$ and therefore unimodal.

To see that the functions u_i are well defined, we have to show that $g(x_i) - g(x_{i-1}) = h(x_{i+1}) - h(x_i)$ holds for $1 \leq i \leq n$. In fact, we can show more, i.e. that $g(x_{i-1}) = h(x_i)$. First, observe that $V^-(f; [x_{i-1}, x_i]) = f(x_{i-1})$. If $x_i < \infty$, this is true by definition of x_i, since $(x_{i-1}, x_i + \delta)$ is forced-max if $\delta > 0$ and is not forced-max if $\delta < 0$. If $x_i = \infty$, this is trivial in the case $x_{i-1} = \infty$, otherwise it follows by the definition of x_i that (x_{i-1}, x_i) is not forced-max, so $V^-(f; [x_{i-1}, x_i]) \leq f(x_{i-1})$, and $V^-(f; [x_{i-1}, x_i]) \geq f(x_{i-1}) - f(x_i) = f(x_{i-1})$ is true by definition of V^-. We therefore have:

$$\begin{aligned} g(x_{i-1}) &= V^+(f; (-\infty, x_{i-1}]) \\ &= f(x_{i-1}) + V^-(f; (-\infty, x_{i-1}]) \\ &= V^-(f; [x_{i-1}, x_i]) + V^-(f; (-\infty, x_{i-1}]) \\ &= V^-(f; (-\infty, x_i]) \\ &= h(x_i). \end{aligned}$$

[6] We use the standard convention that $\inf \emptyset = \infty$ and $(a, b) = \emptyset$ if $a \geq b$.

This establishes that the functions u_i are well defined. Clearly, they are also unimodal. Finally, we also have $f = \sum_{i=1}^n u_i$. To see this, observe that for $x \in [x_{i-1}, x_i]$, we have:

$$\sum_{i=1}^n u_i(x) = g(x) - g(x_{i-1}) + h(x_i) - h(x) = f(x).$$

This concludes the proof. □

A corollary of this is that in the case $M(f) = \infty$, we can actually find an infinite set of pairwise disjoint almost forced-max intervals.

Corollary 3.5. *Suppose $f : \mathbb{R} \to [0, \infty)$ is compactly supported. If $M(f) = \infty$, there is an infinite set of disjoint almost forced-max intervals with respect to f.*

Proof. The fact that $M(f) = \infty$ implies that the recursively defined sequence

$$\begin{aligned} x_0 &= -\infty, \\ x_i &= \inf\{x \mid V^-(f; (x_{i-1}, x)) > f(x_{i-1})\}, \qquad i \in \mathbb{N} \end{aligned}$$

is strictly increasing (otherwise, using the theorem we would have a finite unimodal decomposition). This gives us infinitely many almost-forced max intervals. □

From the explicit construction above, it also follows that a function with finite unimodal category must be well-behaved near the boundary of its support:

Corollary 3.6. *Suppose $f : \mathbb{R} \to [0, \infty)$ has finite unimodal category and* $\operatorname{supp} f = [a, b] \subseteq \mathbb{R}$. *Then there is an $\epsilon > 0$ such that f is increasing on $[a, a+\epsilon]$ and decreasing on $[b - \epsilon, b]$.*

This allows us to completely characterize the functions with finite unimodal category.

Theorem 3.7. *A compactly supported function $f : \mathbb{R} \to [0, \infty)$ has finite unimodal category if and only if:*

- *$f^{-1}(0, \infty) = \bigcup_{j=1}^m (a_j, b_j)$ for some $-\infty \leq a_1 < b_1 \leq a_2 < b_2 \leq \ldots \leq a_m < b_n \leq \infty$, and*
- *there is an $\epsilon > 0$ such that f is increasing on $[a_j, a_j + \epsilon]$ and decreasing on $[b_j - \epsilon, b_j]$ for each $j = 1, \ldots, m$.*

Proof. The forward implication is obvious from Theorem 3.4 and Corollary 3.5. Conversely, suppose the two bullet points are satisfied. We can decompose f uniquely as $f = \sum_{j=1}^m f_j$, where f_j is supported in the interval $[a_j, b_j]$. It suffices to show that each of these functions f_j has a finite unimodal decomposition, so without loss of generality, we can assume that $m = 1$ and $f_1 = f$. Let $a = a_1$ and $b = b_1$. We can assume that ϵ is small enough so that the interval $[a + \epsilon, b - \epsilon]$ is nonempty. This interval is also compact, so f achieves a minimum on it, say $f(x) \geq \eta > 0$ for all $x \in [a + \epsilon, b - \epsilon]$. Since f is of bounded variation, there is a $q \in \mathbb{N}$ such that $V^-(f; [a+\epsilon, b-\epsilon]) < q\eta$. This allows us to show that f has finite unimodal category. Suppose $(x_1, y_1), \ldots, (x_k, y_k)$ is a set of disjoint intervals. At most one of these intervals can contain the point $a+\epsilon$ and at most one of them can contain the point $b+\epsilon$.

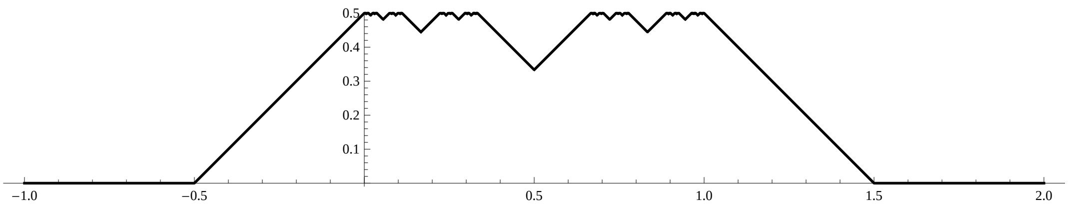
Therefore, all the other intervals are contained in one of the intervals $(-\infty, a+\epsilon)$, $(a+\epsilon, b-\epsilon)$, $(b-\epsilon, \infty)$. Now observe that any interval $(x, y) \subseteq (-\infty, a+\epsilon)$ or $(x, y) \subseteq (b-\epsilon, \infty)$ cannot be forced-max (since f is increasing on the first interval and decreasing on the second one). On the other hand, among the aforementioned intervals, there are less than q disjoint forced-max intervals $(x'_l, y'_l) \subseteq (a+\epsilon, b-\epsilon)$, otherwise we would have:

$$V^-(f; (x, y)) \geq \sum_{l=1}^{q} V^-(f; (x'_l, y'_l)) > \sum_{l=1}^{q} f(x'_l) \geq q\eta,$$

which is not the case. Therefore f has at most $q+2$ forced-max intervals and the proof is complete. $\square$

Example. To illustrate the theorems on a function with an infinite set of local maxima, let $C \subseteq [0, 1]$ be the usual middle-thirds Cantor set and define

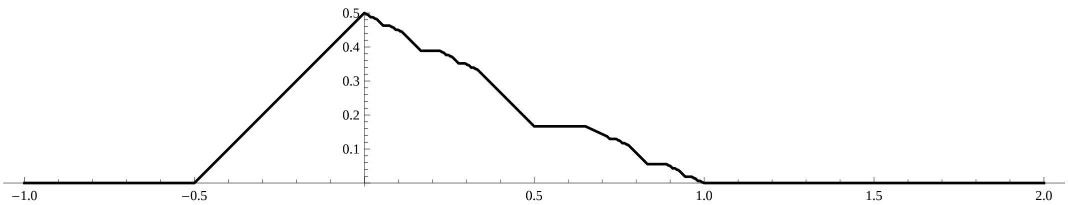
$$f(x) = \max\{0, \tfrac{1}{2} - d(x, C)\}.$$

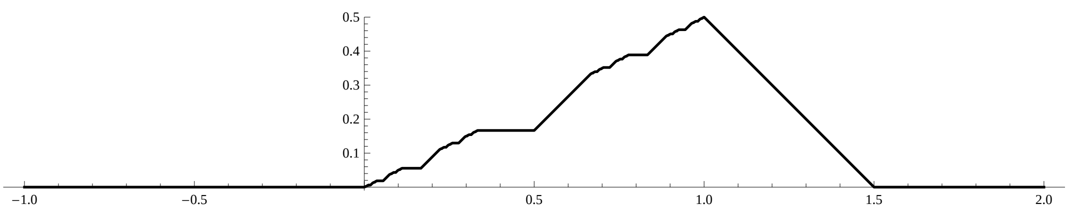


Observe that $f^{-1}(0, \infty) = (-\frac{1}{2}, \frac{3}{2})$ and that f is increasing on $[-\frac{1}{2}, 0]$ and decreasing on $[1, \frac{3}{2}]$. This already implies that f has finite unimodal category. In fact, using the explicit construction of Theorem 3.4, we can show that $\mathbf{ucat}(f) = 2$. To see this, observe that $(-\infty, x)$ is forced-max if and only if $x > 0$, so in the construction, we must take $x_1 = 0$. Next, observe that

$$V^-(f; (0, 1)) = \frac{1}{6} + \frac{2}{3 \times 6} + \frac{4}{3^2 \times 6} + \ldots = \frac{1}{2}.$$

Since f is strictly decreasing on $[1, \frac{3}{2}]$, for any $x > 1$, we have $V^-(f; (0, x)) > \frac{1}{2}$. Therefore, we must take $x_2 = 1$. Finally, $(1, \infty)$ is not forced-max. Hence, the decomposition obtained by sweeping has precisely two unimodal summands u_1 and u_2. These can be graphed as follows:

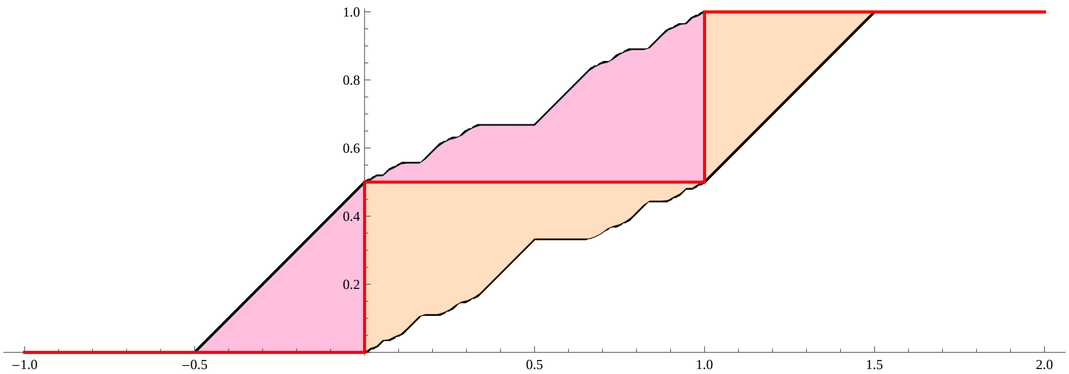




The construction used in the theorem has a nice geometric interpretation:

- Plot the positive variation g and the negative variation h.
- Draw a broken line, as follows. For each $i \in \{1, 2, \ldots, n\}$, draw a horizontal segment (or infinite ray) from $(x_{i-1}, g(x_{i-1}))$ to $(x_i, h(x_i))$ and a vertical segment from $(x_i, h(x_i))$ to $(x_i, g(x_i))$. Conclude with a horizontal ray from $(x_n, h(x_n))$ to $(x_{n+1}, g(x_{n+1}))$.
- This divides the area between g and h into $2n$ pieces. For each $i = 1, \ldots, n$, there are two pieces between the $(i-1)$-th and the i-th horizontal line (where the line at the bottom is indexed by 0), and they are divided by a vertical line. Flipping the second piece and translating both pieces downward so that their bottom edge is on the x-axis yields a set bounded by two curves: the x-axis and the graph of u_i.

Note that in the second step, the horizontal segment drawn for each i is the longest horizontal segment beginning at the endpoint of the $(i-1)$-th vertical segment that does not cross the graph of h. Similarly, each vertical segment is the longest vertical segment beginning at the given point that does not cross the graph of g. In the given example, the first two steps are pictured as follows:



The third step consists of assembling the first two pieces to obtain u_1 and the second two pieces to obtain u_2, whose graphs are pictured above. (The orange pieces are flipped in the process.)

Now we turn to functions $f : J \to [0, \infty)$, where $J \subseteq \mathbb{R}$ is an interval. In this case, we define $M(f)$ as the maximum number of forced-max intervals contained in J. Note that again, this agrees with the maximum number of almost forced-max intervals $\widetilde{M}(f)$. The proof is the same as for Proposition 3.3, but note that the notion of "almost forced-max" depends on the ambient, as the enlarged intervals must be contained in J. The construction of Theorem 3.4 can be used to construct a minimal unimodal decomposition[7] of f. Using an appropriate homeomorphism h, first map this interval to an interval with endpoints 0 and 1 (note that this does not affect unimodality and hence does not change **ucat**). Now, since $f \circ h^{-1}$ is

[7]This actually allows us to compute $\mathbf{ucat}(f)$ of $f : X \to [0, \infty)$ for any set $X \subseteq \mathbb{R}$: if $f^{-1}(0, \infty)$ has infinitely many components, $\mathbf{ucat}(f)$ is infinite, otherwise $f^{-1}(0, \infty)$ is a finite union of intervals, each of which can be treated separately.

of bounded variation, it has limits at 0 and 1, so we may extend it uniquely to a function $\bar{f} : [0,1] \to \mathbb{R}$. Finally, we can extend this function to $\hat{f} : \mathbb{R} \to [0,\infty)$ by

$$\hat{f}(x) = \begin{cases} (x+1)\bar{f}(0); & x \in [-1,0], \\ \bar{f}(x); & x \in [0,1], \\ (2-x)\bar{f}(1); & x \in [1,2], \\ 0; & x \notin [-1,2]. \end{cases} \tag{3}$$

We claim that using the construction of Theorem 3.4 on $\hat{f}$ gives us a minimal unimodal decomposition of f:

Proposition 3.8. *Let $J \subseteq \mathbb{R}$ be an interval and $f : J \to [0,\infty)$ a function. Then*

$$\mathbf{ucat}(f) = \mathbf{ucat}(\hat{f}),$$

where $\hat{f}$ is as defined in Equation (3)*. A minimal unimodal decomposition of f is obtained by using the construction of Theorem 3.4 on $\hat{f}$, restricting the obtained unimodal summands to $h(J)$ and composing with h.*

Furthermore, we either have $M(\hat{f}|_{(-\infty,1]}) = M(\hat{f}) - 1$ or $M(\hat{f}|_{(-\infty,1]}) = M(\hat{f})$. The last summand (i.e. the one whose maximum appears at the largest x) in this minimal unimodal decomposition of f is increasing if and only if $M(\hat{f}|_{(-\infty,1]}) = M(\hat{f}) - 1$.

Proof. Restriction of a unimodal function to an interval is still unimodal, therefore $\mathbf{ucat}(f) \leq \mathbf{ucat}(\hat{f})$. Conversely, suppose $u_1, \dots, u_n$ is a unimodal decomposition of f. Then construct $\hat{u}_i$ from u_i using the same procedure used to construct $\hat{f}$ from f. This yields a unimodal decomposition of $\hat{f}$, so $\mathbf{ucat}(f) \geq \mathbf{ucat}(\hat{f})$. Note that the output of the "sweeping algorithm" for f on $[0,1]$ does not depend on left-most interval where the input function $\hat{f}$ increases, so we can sweep directly over $[0,1]$. (The reader is warned that $M(f) = \mathbf{ucat}(f)$ may no longer be true in this case.)

Next we establish the bounds for $M(\hat{f}|_{(-\infty,1]})$. It is clear that $M(\hat{f}|_{(-\infty,1]}) \leq M(\hat{f})$.

In the other direction, let $(x_1,y_1), \dots, (x_n,y_n)$ be a sequence of forced-max intervals that realizes $M(\hat{f})$, ordered by the size of the endpoints. Observe that $x_n < 1$, otherwise $(1,2)$ would be a forced-max interval, which cannot be the case, since $\hat{f}$ is decreasing there. Therefore $(x_1,y_1), \dots, (x_{n-1},y_{n-1})$ are forced-max intervals for $\hat{f}|_{(-\infty,1]}$. This proves that $M(\hat{f}|_{(-\infty,1]}) \geq M(\hat{f}) - 1$.

Now, let $u_1, u_2, \dots, u_n$ be the decomposition of $\hat{f}$ obtained by sweeping, where the summands are ordered as in the construction, and suppose $u_n|_{(-\infty,1]}$ is increasing. Then $\mathbf{ucat}(\hat{f}) = M(\hat{f}) = n$. We claim that $M(\hat{f}|_{(-\infty,1]}) = n - 1$. Suppose that this does not hold. Then there are forced-max intervals $(x_1,y_1), \dots, (x_n,y_n)$ for $\hat{f}|_{(-\infty,1]}$. By Proposition 3.2 this means that for each unimodal decomposition of $\hat{f}|_{(-\infty,1]}$ and for each i, there is a summand that achieves its maximum in (x_i,y_i). This contradicts the fact that $u_n|_{(-\infty,1]}$ is increasing.

Conversely, suppose that $M(\hat{f}|_{(-\infty,1]}) = n - 1$. Define $x_0, x_1, \dots, x_{n+1}$ as in Theorem 3.4. The intervals $(x_0,x_1), \dots, (x_{n-1},x_n)$ are then almost forced-max with respect to $\hat{f}$. Since there are n of them, at least one is not almost-forced max with respect to $\hat{f}|_{(-\infty,1]}$. In fact, this is true for any interval whose closure intersects $[1,\infty)$. Therefore, (x_{n-1}, x_n) is not almost-forced max with respect to $\hat{f}|_{(-\infty,1]}$. Furthermore, there is at most one such interval, since $\mathbf{ucat}(f) = n$. We can conclude that $x_{n-1} < 1 \leq x_n$. Now, simply recall that u_n is increasing on (x_{n-1}, x_n) by construction, and the proof is complete. $\square$

The following property of unimodal decompositions is sometimes useful.

Proposition 3.9. *Suppose $f : [0,1] \to [0,\infty)$, $\mathbf{ucat}(f) \geq 2$ and $u_1, \ldots, u_n$ is the unimodal decomposition obtained by sweeping. Then $u_1(1) = \ldots = u_{n-2}(1) = 0$ and there exist unimodal functions u'_{n-1} and u'_n such that $u_1, \ldots, u_{n-2}, u'_{n-1}, u'_n$ is a unimodal decomposition of f and $u'_{n-1}(1) = 0$. Let*

$$a = \max\{x \mid u_{n-1}|_{[0,x]} \textit{ is increasing}\}.$$

It is possible to choose u'_{n-1} so that it only differs from u_{n-1} in the interval $(a, 1]$ and if u_n was increasing, u'_n is again increasing.[8]

Proof. The fact that $u_1(1) = \ldots = u_{n-2}(1) = 0$ follows directly from the construction in the proof of Theorem 3.4. Let $c := u_{n-1}(1)$ and let $m = u_{n-1}(a)$. If $c = 0$, there is nothing to do, so we shall assume that $c > 0$. The sweeping construction then implies that u_n is increasing. It also implies that $c < m$, so for $x \in [a, 1]$ we may define

$$u'_{n-1}(x) = m\frac{u_{n-1}(x) - c}{m - c}$$

and

$$u'_n(x) = u_n(x) + c\frac{m - u_{n-1}(x)}{m - c}.$$

Note that $u'_{n-1}(1) = 0$, $u'_{n-1}(a) = m$ and since $u_{n-1} \leq m$, we also have $u'_{n-1} \leq u_{n-1}$ and u'_{n-1} is decreasing in $[a, 1]$. Therefore, u'_{n-1} is unimodal. The function u'_n is defined as the sum of two increasing functions, so it is also unimodal. This completes the proof. $\square$

3.2 Circle

As in the previous section, we are only interested in functions $f : S^1 \to [0,\infty)$ that are continuous and of bounded variation, in the sense that the function $[0,1] \to [0,\infty)$ defined by $t \mapsto f(\exp(2\pi i t))$ has bounded variation. Note that this is again a necessary condition for finiteness of **ucat**. If f has a zero, the problem immediately reduces to the case of $X = \mathbb{R}$, so we only deal with functions without zeros. Note that this already implies $\mathbf{ucat}(f) \geq \mathbf{gcat}(S^1) = 2$. We now introduce some notations.

Notation. Suppose $a \in S^1$. In the rest of this section, $\phi_a^+ : [0,1] \to S^1$ is defined by $\phi_a^+(t) = a\exp(2\pi i t)$ and $\phi_a^- : [0,1] \to S^1$ by $\phi_a^-(t) = a\exp(-2\pi i t)$. We also define $f_a^+ = f \circ \phi_a^+$ and $f_a^- = f \circ \phi_a^-$. If $g : [0,1] \to [0,\infty)$ is any function, its extension $\hat{g} : \mathbb{R} \to [0,\infty)$ is defined as in Equation (3). For easier readability we sometimes omit the $+$ sign and e.g. write f_a instead of f_a^+.

[8]Note that u_n is automatically increasing if $u_{n-1}(1) \neq 0$.

We are also going to make use of the following numbers associated to f:

$$M_a^+(f) = M(\hat{f}_a^+|_{(-\infty,1]}), \qquad M^+(f) = \min_{a\in S^1} M_a^+(f),$$
$$M_a^-(f) = M(\hat{f}_a^-|_{(-\infty,1]}), \qquad M^-(f) = \min_{a\in S^1} M_a^-(f).$$

We proceed to prove a lemma.

Lemma 3.10. *Suppose $f : S^1 \to [0,\infty)$ has no zeros and $\mathbf{ucat}(f) = n$. Then it is possible to choose $a \in S^1$ so that $\mathbf{ucat}(f_a) \le \mathbf{ucat}(f)+1$. Furthermore, f_a admits a unimodal decomposition $u_1,\dots,u_{n+1}$ where u_1 is decreasing and u_{n+1} is increasing.*

Proof. Let $v_1,\dots,v_n$ be a minimal unimodal decomposition of f and choose $a \in S^1$ such that $v_1(x) \le v_1(a)$ holds for all $x \in S^1$. Define functions $\tilde{v}_i : [0,1] \to [0,\infty)$ by $\tilde{v}_i = v_i \circ \phi_a$. By unimodality, each "open support" $\tilde{v}_i^{-1}(0,\infty)$ has either one or two components. Reorder the indices $i > 1$ and choose k so that the open supports of $\tilde{v}_1,\dots,\tilde{v}_k$ have two components and the open supports of $\tilde{v}_{k+1},\dots,\tilde{v}_n$ have one component.

Each $\tilde{v}_i$ for $1 \le i \le k$ can uniquely be split as a sum of two unimodal functions $\tilde{v}_i = \tilde{v}_i^{(1)} + \tilde{v}_i^{(2)}$ so that $0 \in \operatorname{supp}\tilde{v}_i^{(1)}$ and $1 \in \operatorname{supp}\tilde{v}_i^{(2)}$. By unimodality, for each i either $\tilde{v}_i^{(1)}$ is decreasing or $\tilde{v}_i^{(2)}$ is increasing and by our choice of a both of these hold for $i = 1$. After possibly reordering the indices $i \in \{2,\dots,k\}$ again, we can choose an l such that $\tilde{v}_i^{(1)}$ is decreasing for $i = 2,...,l$ and $\tilde{v}_i^{(2)}$ is increasing for $i = l+1,\dots,k$. (For $i = 1$, both of these hold.)

For $i = 1,\dots,n+1$, we may therefore define $u_i : [0,1] \to [0,\infty)$ as follows:

$$u_i = \begin{cases} \sum_{i=1}^{l} \tilde{v}_i^{(1)}; & \text{for } i = 1, \\ \tilde{v}_i^{(2)}; & \text{for } i = 2,\dots,l, \\ \tilde{v}_i^{(1)}; & \text{for } i = l+1,\dots,k, \\ \tilde{v}_i; & \text{for } i = k+1,\dots,n, \\ \tilde{v}_1^{(2)} + \sum_{i=l+1}^{k} \tilde{v}_i^{(2)}; & \text{for } i = n+1. \end{cases}$$

The function u_1 is decreasing since it is the sum of a sequence of decreasing functions and u_{n+1} is increasing since it is the sum of a sequence of increasing functions. The unimodality of the other functions is clear. □

Proposition 3.11. *Suppose $f : S^1 \to [0,\infty)$ has no zeros and $\mathbf{ucat}(f_a) = 2$ for some $a \in S^1$. Then $\mathbf{ucat}(f) = 2$.*

Proof. Suppose $f_a = u_1 + u_2$ is the unimodal decomposition obtained by sweeping. Using Proposition 3.9, we can modify it so that $u_1(1) = 0$. If u_1 is increasing on some small interval $[0,\epsilon]$, let

$$v(x) = \max\{0, (1 - \tfrac{x}{\epsilon}) f_a(0)\}$$

and define functions $\tilde{u}_1, \tilde{u}_2 : S^1 \to [0,\infty)$ by $\tilde{u}_1 \circ \phi_a = u_1 - v$ and $\tilde{u}_1 + \tilde{u}_2 = f$. It can be directly verified that these are unimodal. An entirely analogous construction works if u_2 is decreasing on some interval $[1-\epsilon, 1]$.

Finally, if u_1 is decreasing everywhere and u_2 is increasing, choose a point x_0 such that $u_1(x_0) = u_2(x_0)$ and define

$$u(x) = \begin{cases} 2u_2(x); & x \leq x_0, \\ 2u_1(x); & x \geq x_0. \end{cases}$$

Define functions $\tilde{u}_1, \tilde{u}_2 : S^1 \to [0,\infty)$ by $\tilde{u}_1 \circ \phi_a = u$ and $\tilde{u}_1 + \tilde{u}_2 = f$. A direct verification shows that these are unimodal. □

This allows us to characterize $\mathbf{ucat}(f)$ in the following way.

Theorem 3.12. *The unimodal category of $f : S^1 \to [0,\infty)$ without zeros is characterized as follows:*

$$\mathbf{ucat}(f) = \max\{2, M^+(f)\} = \max\{2, M^-(f)\}.$$

Proof. Clearly, $\mathbf{ucat}(f) \geq \mathbf{gcat}(S^1) = 2$, since f has no zeros. Notice that is suffices to prove the first equality. The second equality then follows by a simple observation that composing f with a reflection of the circle does not affect the unimodal category, since a reflection is a homeomorphism.

Next, we show that $M^+(f) \leq \mathbf{ucat}(f) =: n$. By Lemma 3.10, there is a point $a \in S^1$ and a unimodal decomposition $f_a = \sum_{i=1}^{n+1} u_i$ such that u_{n+1} is increasing. By Proposition 3.8 this means that $M_a^+(f) \leq n$, which proves $M^+(f) \leq \mathbf{ucat}(f)$.

Note that this already completes the proof in the case $\mathbf{ucat}(f) = 2$. We may therefore assume that $n = \mathbf{ucat}(f) \geq 3$, in which case it remains to establish that $\mathbf{ucat}(f) \leq M^+(f)$. By the definition of $M^+(f)$, it suffices to show that $\mathbf{ucat}(f) \leq M_a^+(f)$ holds for all $a \in S^1$. So let $a \in S^1$. By Proposition 3.8, $M_a^+(f)$ is either $\mathbf{ucat}(f_a)$ or $\mathbf{ucat}(f_a) - 1$.

Case 1: $M_a^+(f) = \mathbf{ucat}(f_a)$. In this case, it is sufficient to see that $\mathbf{ucat}(f_a) \geq n$. Suppose that this does not hold. Then the sweeping algorithm returns a unimodal decomposition of f_a with $k < n$ summands and by Proposition 3.9 we may modify the last two so that $u_i(1) = 0$ for $i < k$, which also implies $u_k(1) = u_1(0)$. If $k = 2$, Proposition 3.11 yields a contradiction. Otherwise note that for $i = 2, \ldots, k-1$, we have $u_i^{-1}(0,\infty) \subseteq (0,1)$, so these summands may be transported to S^1 by defining $\tilde{u}_i \circ \phi_a = u_i$. Since $k > 2$, the sets $u_1^{-1}(0,\infty)$ and $u_k^{-1}(0,\infty)$ are disjoint. The summands u_1 and u_k may be glued together, i.e. we may define $g : S^1 \to [0,\infty)$ by $g \circ \phi_a = u_1 + u_k$. If g is unimodal, we have $\mathbf{ucat}(f) \leq k - 1$, a contradiction. If not, there exist $B > A > 0$ such that for $x > B$, $g^{-1}[x,\infty)$ is empty, for $x \in (A, B]$ it has two contractible components and for $x \in (0, A]$ it has one contractible component. This implies that $\mathbf{ucat}(g) = 2$ and therefore $\mathbf{ucat}(f) \leq k$, which is again a contradiction.

Case 2: $M_a^+(f) = \mathbf{ucat}(f_a) - 1$. By Proposition 3.8, this means that f_a has a unimodal decomposition into $k = \mathbf{ucat}(f_a)$ summands $u_1, \ldots, u_k$ such that u_k is increasing. It is sufficient to prove that $k \geq n + 1$. Suppose not: then $k < n + 1$. If $k = 2$, we have $\mathbf{ucat}(f) = 2$ by Proposition 3.11, so we may assume that $k \geq 3$. By Proposition 3.9, we may again modify the last two summands so that $u_{k-1}(1) = 0$. This allows us to use the same construction as in Case 1: define $\tilde{u}_i$ by $\tilde{u}_i \circ \phi_a = u_i$ for $i = 2, \ldots, k-1$ and $\tilde{u}_1 \circ \phi_a = u_1 + u_k$. Since u_k is increasing, $\tilde{u}_1$ is unimodal by definition. But this implies that $\mathbf{ucat}(f) \leq k - 1 < n$, yet another contradiction. □

Remark 3.3. As a consequence of this proof we can see that $M_a^+(f)$ can be computed in a greedy manner as in Theorem 3.4, i.e. by recursively defining

$$\begin{aligned} x_0 &= -\infty, \\ x_i &= \inf\{x \mid V^-(\hat{f}_a; (x_{i-1}, x)) > \hat{f}_a(x_{i-1})\}, \qquad i \in \mathbb{N}, \end{aligned}$$

however, in this case, we must stop once x_i exceeds 1.

3.2.1 Algorithm in the Case of Finitely Many Critical Points

One possible interpretation of these results is that $\mathbf{ucat}(f)$ of $f : S^1 \to [0, \infty)$ can be computed by sweeping if we know where to start. In general, it is not entirely clear how to find the starting point. However, if the function only has finitely many critical points, it suffices to check these critical points. Therefore the unimodal category of any such function can be computed in a completely algorithmic manner. This is justified by the following result which shows that $M_a^+(f)$ achieves its minimum at a critical point. An explicit description of the algorithm is available in Appendix B.

For convenience we introduce some further notation. If $a, a' \in S^1$ let $[a, a']$ denote the closed arc between a and a', i.e. the set of points obtained by starting at a and traversing the circle in the positive direction until reaching a'. Let $(a, a'), (a, a']$ and $[a, a')$ denote the corresponding open and half-open arcs.

Proposition 3.13. *Suppose $f : S^1 \to [0, \infty)$ is a function and $a_1, a_2 \in S^1$ are points such that (a_1, a_2) does not contain a critical point of f. Then*

$$M_{a_1}^+(f) \leq M_a^+(f)$$

for all $a \in (a_1, a_2)$.

Proof. Let $0 \leq t_1 < t < t_2 \leq 1$, $a_1 = \exp(2\pi i t_1)$, $a_2 = \exp(2\pi i t_2)$ and $a = \exp(2\pi i t)$. Also let $t' = t - t_1$. Suppose $M_{a_1}^+(f) = n$ and let $(-\infty, x_1), (x_1, x_2), \ldots, (x_{n-1}, x_n)$ be a collection of forced-max intervals for $\hat{f}_{a_1}|_{(-\infty,1]}$. There are now two cases to treat.

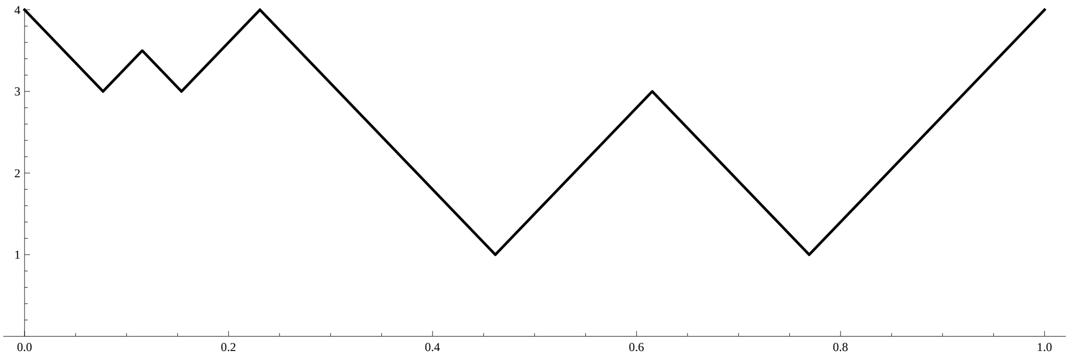
Case 1: $\hat{f}_{a_1}$ is increasing on $[0, t']$. In this case, we have $x_1 \geq t'$ and the intervals $(-\infty, x_1 - t'), (x_1 - t', x_2 - t'), \ldots, (x_{n-1} - t', x_n - t')$ is obviously a collection of forced-max intervals for $\hat{f}_a|_{(-\infty,1]}$, so $M_a^+(f) \geq n$.

Case 2: $\hat{f}_{a_1}$ is decreasing on $[0, t']$. Without loss of generality, we can assume $\hat{f}_{a_1}$ is non-constant on $[0, t']$, otherwise the first case applies. We also have $t' \leq x_2$. Without loss of generality we can further assume that $x_1 \leq t'$. (If this is not the case, we can replace x_1 by $x := \min\{x_1, t'\}$. Upon doing this, $(-\infty, x)$ is still forced-max, because $\hat{f}_{a_1}$ is decreasing on $[0, t']$, and (x, x_2) is forced-max because forced-max intervals form an ideal.) There is an $\epsilon > 0$ such that $\hat{f}_{a_1}$ is still decreasing on $[0, t' + \epsilon]$. The intervals $(-\infty, \epsilon), (\epsilon, x_2 - t'), (x_2 - t', x_3 - t'), \ldots, (x_{n-1} - t', x_n - t')$ are a collection of forced-max intervals for $\hat{f}_a|_{(-\infty,1]}$, so $M_a^+(f) \geq n$. For $(\epsilon, x_2 - t')$, this is true because the interval $(t' + \epsilon, x_2)$ is still forced-max for $\hat{f}_{a_1}|_{(-\infty,1]}$ since it is decreasing on $[x_1, t' + \epsilon]$ and therefore

$$V^-(\hat{f}_a; (\epsilon, x_2 - t')) = V^-(\hat{f}_{a_1}; (t' + \epsilon, x_2)) > \hat{f}_{a_1}(t' + \epsilon) = \hat{f}_a(\epsilon).$$

For all the other intervals, this is obvious. $\square$

Example. Consider the function $f : S^1 \to [0, \infty)$ obtained by choosing eight points on the circle, for instance $a_j = \exp(2\pi i t_j) \in S^1$, $j = 0, 1, \ldots, 7$, where $0 = t_0 < t_1 < \ldots < t_7 < 1$, taking the values there to be $4, 3, \frac{7}{2}, 3, 4, 1, 3, 1$ respectively, and interpolating linearly in between. Slicing the circle at $a = 0$, we obtain the following graph:



By the above proposition, the calculation of $M^+_{a_i}(f)$ for $i = 0, 1, \ldots, 7$ suffices to determine $\mathbf{ucat}(f)$. However, it is illustrative to calculate $M^+_a(f)$ for all $a \in S^1$. By a direct calculation, we obtain:

$$M^+_a(f) = \begin{cases} 2; & a \in [a_1, a_2] \cup [a_5, a_6], \\ 3; & a \in (a_6, a_1) \cup (a_2, a_5). \end{cases}$$

We conclude that $\mathbf{ucat}(f) = 2$. An explicit unimodal decomposition can also be described. The first summand is obtained by mapping the points a_j, $j = 0, 1, \ldots, 7$, to the sequence $3, 3, \frac{7}{2}, 3, 3, 0, 0, 0$, and interpolating linearly. The second summand is obtained by mapping the points a_j, $j = 0, 1, \ldots, 7$, to the sequence $1, 0, 0, 0, 1, 1, 3, 1$ and interpolating linearly.

4 The Monotonicity Conjecture

Baryshnikov and Ghrist conclude their paper with the following conjecture [4, Conjecture 18], which they believe could play an important role in obtaining various bounds for $\mathbf{ucat}^p$.

Conjecture. Suppose $f : X \to [0, \infty)$ and $0 < p_1 < p_2 \leq \infty$. Then $\mathbf{ucat}^{p_1}(f) \leq \mathbf{ucat}^{p_2}(f)$.

In other words, they conjecture that $\mathbf{ucat}^p$ is monotone in p. The aim of this section is to investigate this conjecture for various spaces and functions. We show that the conjecture is true for $X = \mathbb{R}$ and $X = S^1$. However, there are counterexamples if X is a certain type of graph and if $X = \mathbb{R}^2$ is the Euclidean plane.

4.1 Proof for the Real Line and the Circle

In the case of $X = \mathbb{R}$, the conjecture is true. As in Section 3.1, we assume throughout Section 4.1 that $f : \mathbb{R} \to [0, \infty)$ is of bounded variation.

Definition. Let $k \in \mathbb{N}_0$. We call a sequence $a = (a_0, a_1, \ldots, a_{2k})$ such that $a_0 \leq a_1 \geq a_2 \leq \ldots \geq a_{2k}$ an **up-down** sequence of length $2k$. For such a sequence, define its **negative variation** as follows:

$$V^-(a) = \sum_{i=1}^{k} (a_{2i-1} - a_{2i}).$$

Note that this is a sum of non-negative numbers. We also define its p-th power for $p > 0$ as $a^p = (a_0^p, a_1^p, \ldots, a_{2k}^p)$. Note that this is again an up-down sequence.

Now, recall the following inequality of Karamata:

Theorem 4.1 ([44])**.** *Suppose we are given two finite sequences $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ with terms in (α, β) satisfying the following conditions*[9]*:*

- $a_1 \geq a_2 \geq \ldots \geq a_n$ *and* $b_1 \geq b_2 \geq \ldots \geq b_n$*,*
- $a_1 + a_2 + \ldots + a_k \geq b_1 + b_2 + \ldots + b_k$ *for* $1 \leq k \leq n-1$*,*
- $a_1 + a_2 + \ldots + a_n = b_1 + b_2 + \ldots + b_n$*.*

Further suppose $\phi : (\alpha, \beta) \to \mathbb{R}$ is a convex function. Then we have

$$\phi(a_1) + \phi(a_2) + \ldots + \phi(a_n) \geq \phi(b_1) + \phi(b_2) + \ldots + \phi(b_n).$$

Note that the inequality is reversed if ϕ is concave, because in that case $-\phi$ is convex. We only need the following special case:

Lemma 4.2. *Suppose $0 < q < 1$ and let $x, y, z \geq 0$ such that $\max\{x, z\} \leq y \leq x + z$. Then*

$$(x - y + z)^q \leq x^q - y^q + z^q.$$

[9] If these conditions are satisfied, we say that a majorizes b.

Proof. By continuity, it is sufficient to treat the case when $x, y, z > 0$ and $y < x+z$. The inequality is equivalent to

$$(x-y+z)^q + y^q \leq x^q + z^q,$$

which is just a special case of Theorem 4.1 in the case of $n = 2$, $a_1 = y, a_2 = x-y+z$, $b_1 = \max\{x, z\}, b_2 = \min\{x, z\}$ and $\phi(t) = t^q$, which is concave. $\square$

Lemma 4.3. *Suppose $0 < q < 1$, a is an up-down sequence of length $2k$ and $V^-(a) \leq a_0$. Then $V^-(a^q) \leq a_0^q$.*

Proof. We prove this by induction on k. For $k = 0$, this is trivial, as both negative variations are zero. Suppose the lemma holds for all sequences of length at most $2k-2$ and a is a sequence of length $2k$. Then $\tilde{a} = (a_0 - a_1 + a_2, a_3, a_4, \ldots, a_{2k-1}, a_{2k})$ is an up-down sequence of length $2k-2$. Observe that

$$V^-(\tilde{a}) = V^-(a) - (a_1 - a_2) \leq a_0 - a_1 + a_2.$$

By the inductive hypothesis and the previous lemma, we have

$$V^-(\tilde{a}^q) \leq (a_0 - a_1 + a_2)^q \leq a_0^q - a_1^q + a_2^q.$$

But

$$V^-(\tilde{a}^q) + a_1^q - a_2^q = V^-(a^q),$$

so this is precisely the conclusion we wanted. $\square$

We can now prove

Theorem 4.4. *Suppose $f : \mathbb{R} \to [0, \infty)$ and $0 < p_1 < p_2 \leq \infty$. Then $\mathbf{ucat}^{p_1}(f) \leq \mathbf{ucat}^{p_2}(f)$.*

Proof. In the case of $p_2 = \infty$ this follows trivially from Lemma 6, Lemma 9 (our Lemma 2.1) and Lemma 16 of [4], so it suffices to treat the case where $p_1, p_2 < \infty$. In this case, observe that $\mathbf{ucat}^{p_1}(f) \leq \mathbf{ucat}^{p_2}(f)$ is equivalent to $\mathbf{ucat}(f^{p_1}) \leq \mathbf{ucat}^{\frac{p_2}{p_1}}(f^{p_1})$ by Lemma 2.1. Therefore, we may assume without loss of generality that $p_1 = 1$ and $p_2 = p > 1$. Under this reduction, we want to prove that $\mathbf{ucat}(f) \leq \mathbf{ucat}(f^p)$.

To prove this, observe that it is sufficient to prove the following statement:

$$\text{if } V^-(f; [x, y]) > f(x), \text{ then } V^-(f^p; [x, y]) > f(x)^p.$$

This means precisely that each forced-max interval of f is also a forced interval of f^p, which then establishes the claim by Theorem 3.4. We are going to prove the contrapositive. Assume therefore that $V^-(f^p; [x, y]) \leq f(x)^p$. By the definition of negative variation, this means precisely that

$$\sup \left[\sum_{i=1}^{n} \max\{0, f(a_{i-1})^p - f(a_i)^p\} \right] \leq f(x)^p,$$

where the supremum is taken over all partitions P of the form $x = a_0 \leq a_1 \leq a_2 \leq \ldots \leq a_n = y$ of the interval $[x, y]$. Fix an arbitrary partition P of this kind. Note that this time, we allow duplications $a_i = a_{i+1}$ in P. This does not change the relevant supremum and allows us to assume that $n = 2k$ for some $k \in \mathbb{N}$ and that

$$f(a_0) \leq f(a_1) \geq f(a_2) \leq \ldots \geq f(a_{2k}).$$

Let c be the up-down sequence of length $2k$ defined by $c_i = f(a_i)^q$. We have

$$V^-(c) = \sum_{i=1}^{n} \max\{0, f(a_{i-1})^p - f(a_i)^p\} \leq f(x)^p = c_0.$$

By Lemma 4.3 with $q = \frac{1}{p}$, we have

$$V^-\left(c^{\frac{1}{p}}\right) \leq c_0^{\frac{1}{p}}.$$

This means precisely that

$$\sum_{i=1}^{n} \max\{0, f(a_{i-1}) - f(a_i)\} \leq f(x).$$

Since the partition P was arbitrary, the same holds for the supremum over all partitions:

$$V^-(f; [x,y]) \leq f(x).$$

This concludes the proof. $\square$

This immediately proves the monotonocity conjecture for the circle S^1.

Corollary 4.5. *Suppose $f : S^1 \to [0, \infty)$ and $0 < p_1 < p_2 \leq \infty$. Then $\mathbf{ucat}^{p_1}(f) \leq \mathbf{ucat}^{p_2}(f)$.*

Proof. In the case of $p_2 = \infty$ this follows from Lemma 6, Lemma 9 (our Lemma 2.1) and Lemma 16 of [4] (note that Lemma 16 still holds for S^1), so it suffices to treat the case $p_1 < p_2 < \infty$. In this case, we use Lemma 2.1 and Theorem 3.12:

$$\mathbf{ucat}^{p_1}(f) = \min\{2, M^+(f^{p_1})\} \leq \min\{2, M^+(f^{p_2})\} = \mathbf{ucat}^{p_2}(f).$$

The inequality follows from Theorem 4.4. $\square$

4.2 Graphs: First Counterexample

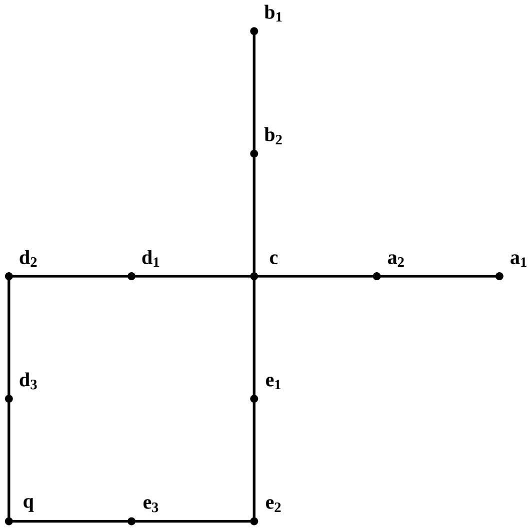
For general spaces, the monotonicity conjecture is false. The simplest counterexamples can be constructed on graphs. Similar ideas can then be exploited to yield counterexamples on Euclidean spaces.

Let G be the graph (abstract simplicial complex of dimension 1) whose vertices and edges are given by[10]

$$\begin{aligned}
V &= \{a_1, a_2, b_1, b_2, c, d_1, d_2, d_3, e_1, e_2, e_3, q\}, \\
E &= \{a_1a_2, a_2c, b_1b_2, b_2c, cd_1, ce_1, d_1d_2, d_2d_3, d_3f, e_1e_2, e_2e_3, e_3q\},
\end{aligned}$$

[10]Here, xy is considered as shorthand for $\{x, y\}$.

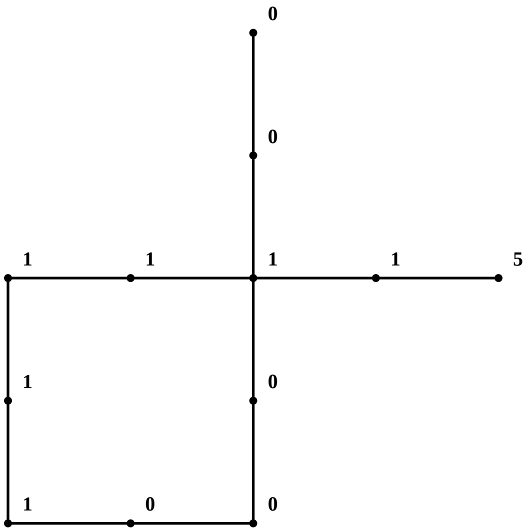
and let X be the polytope of its geometric realization, for instance as in the picture below, where the vertices c and q are realized as $(0,0)$ and $(-2,-2)$, a_1, a_2, d_1, d_2, d_3 are realized as $(2,0), (1,0), (-1,0), (-2,0)$ and $(-2,-1)$ and b_1, b_2, e_1, e_2, e_3 as their reflections across $y = x$. For simplicity, we identify the vertices of G with their corresponding points in X.

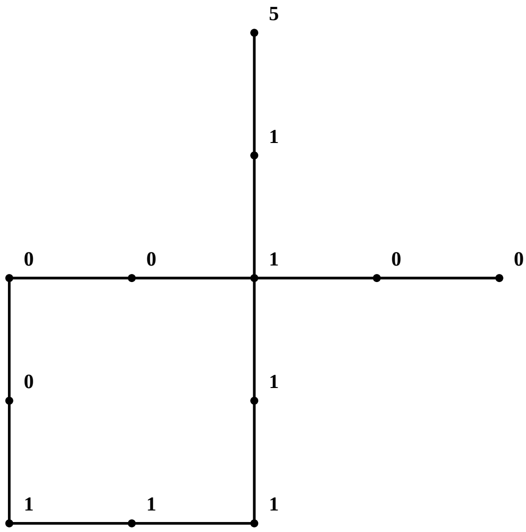


We define three piecewise linear (with respect to the graph structure above) functions $f, u_1, u_2 : X \to [0, \infty)$. The functions u_1 and u_2 are defined by specifying their values on the vertices ($i = 1, 2, 3$):

$$u_1(a_1) = 5, \quad u_1(a_2) = u_1(c) = u_1(d_i) = u_1(q) = 1, \quad u_1(b_1) = u_1(b_2) = u_1(e_i) = 0,$$
$$u_2(b_1) = 5, \quad u_2(b_2) = u_2(c) = u_2(e_i) = u_2(q) = 1, \quad u_2(a_1) = u_2(a_2) = u_2(d_i) = 0.$$

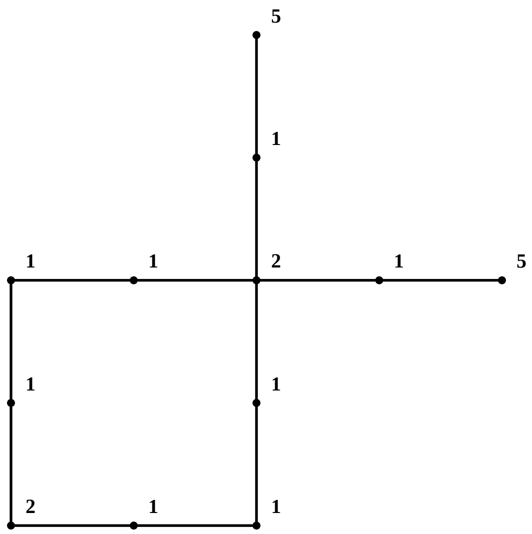
See picture (where u_2 is u_1 reflected across the axis of symmetry):





Finally, we define $f = u_1 + u_2$. Note that f may also be given by its values at the vertices ($i = 1, 2, 3$):

$$f(a_1) = f(b_1) = 5, \quad f(c) = f(q) = 2, \quad f(a_2) = f(b_2) = f(d_i) = f(e_i) = 1.$$



The following follows directly from the definitions.

Observation 4.6. *The functions u_1 and u_2 are unimodal, f is not. So* $\mathbf{ucat}(f) = 2$.

This f yields a simple counterexample to the monotonicity conjecture.

Proposition 4.7. *The unimodal $\frac{1}{2}$-category of f is* $\mathbf{ucat}(\sqrt{f}) = \mathbf{ucat}^{\frac{1}{2}}(f) = 3$.

Proof. A unimodal decomposition of length 3 can easily be constructed explicitly. Note that $\sqrt{f}$ is not piecewise linear, but can be turned into such by using an appropriate homeomorphism on the domain X, all the while retaining the function values at the vertices. The resulting piecewise linear function can be decomposed into three piecewise linear unimodal summands v_1, v_2 and v_3, which we define by their values at the vertices. The nonzero values are given by ($i = 1, 2, 3$):

$$\begin{aligned} v_1(a_1) &= \sqrt{5}, & v_1(a_2) &= 1, & v_1(c) &= \tfrac{\sqrt{2}}{2}, \\ v_2(a_1) &= \sqrt{5}, & v_2(a_2) &= 1, & v_2(c) &= \tfrac{\sqrt{2}}{2}, \\ v_3(q) &= \sqrt{2}, & v_3(d_i) &= 1, & v_3(e_i) &= 1. \end{aligned}$$

The function values at the remaining vertices are zero.

To complete the proof, it therefore remains to show that $\sqrt{f}$ cannot be decomposed into two unimodal summands. We argue by contradiction. Suppose $\sqrt{f} = v_a + v_b$, where v_a and v_b are unimodal. If $r > 0$, define superlevel sets

$$Q_a(r) = v_a^{-1}[r, \infty) \qquad \text{and} \qquad Q_b(r) = v_b^{-1}[r, \infty).$$

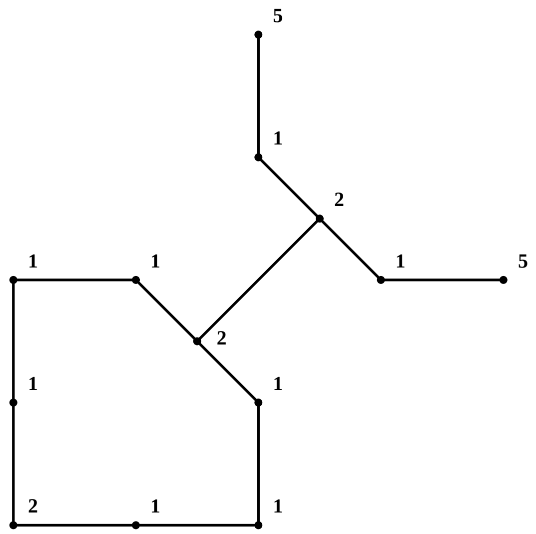
By unimodality, these are all contractible. Note that, since $v_a(a_1)+v_b(a_1) = \sqrt{f(a_1)}$, we either have $v_a(a_1) \geq \frac{\sqrt{5}}{2}$ or $v_b(a_1) \geq \frac{\sqrt{5}}{2}$. Without loss of generality, assume that the first possibility holds. (This is also the reason behind the choice of notation for v_a and v_b.) Since $\sqrt{f(a_2)} = 1$, we have $v_a(a_2) \leq 1$. This immediately implies that $v_a(b_1) \leq 1$, since otherwise we would have $v_a(b_1) = r > 1$ and $Q_a(\min\{r, \frac{\sqrt{5}}{2}\})$ would not be connected (since it is a subspace of X containing a_1 and b_1 but not a_2). This means that $v_b(b_1) \geq \sqrt{5} - 1$. By a symmetric argument, we also have $v_a(a_1) \geq \sqrt{5} - 1$, but we do not use this fact.

Now, let K be the subspace of X consisting of the vertices c, d_i, e_i, q $(i = 1, 2, 3)$ and all the edges between these vertices. Note that K is homeomorphic to a circle. By unimodality, there are points $x, y \in K$ such that $v_a(y) = 0$ and $v_b(x) = 0$. Otherwise we would have $K \subseteq Q_a(r)$ or $K \subseteq Q_b(r)$ for some $r > 0$. Therefore $v_a(x) = v_b(y) = 1$. Since $Q_a(1)$ and $Q_b(1)$ are contractible, there is a path from a_1 to x in $Q_a(1)$, implying that $c \in Q_a(1)$, and a path from b_1 to y in $Q_b(1)$, implying that $c \in Q_b(1)$. This implies that $v_a(c) + v_b(c) \geq 2$, contradicting the fact that $v_a(c) + v_b(c) = \sqrt{f(c)} = \sqrt{2}$ and concluding the proof. $\square$

To sum up, we have found a space X, a function $f : X \to [0, \infty)$ and values $0 < p_1 < p_2 < \infty$ such that $\mathbf{ucat}^{p_1}(f) > \mathbf{ucat}^{p_2}(f)$. Hence, the monotonicity conjecture is not generally true.

Remark 4.1. In fact, by changing the function values of u_1 and u_2 appropriately, the same example can be modified to show that the monotonicity conjecture is not generally true for any pair of exponents $0 < p_1 < p_2 < \infty$.

Further note that this example implies the failure of monotonicity for a very general class of graphs: namely, whenever the graph contains a cycle which contains a point of valence 4, monotonicity cannot hold in general. In fact, monotonicity fails for an even larger class of graphs: the point of valence 4 can be replaced by two points of valence 3 as in the picture below, yielding another counterexample. The proof is very similar to the one above, so we omit it. So, if a connected graph contains a cycle and a point of valence 3 or more somewhere outside this cycle, monotonicity cannot hold in general. Note that this severely limits the collection of CW complexes for which monotonicity can possibly hold.

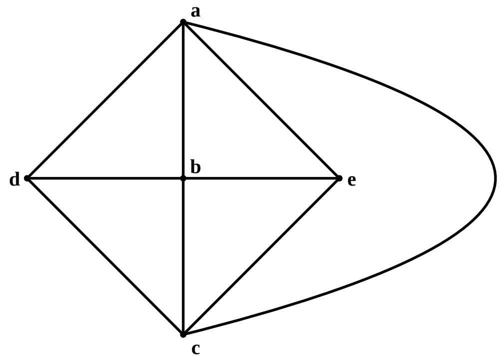


4.3 Graphs: Second Counterexample

The previous example does not preclude the possibility that the monotonicity conjecture holds in the case $0 < p_1 < p_2 = \infty$. We must therefore construct a different example to show that it also fails here. Let G be the graph whose vertices and edges are given by

$$V = \{a, b, c, d, e\},$$
$$E = \{ab, ac, ad, ae, bc, bd, be, cd, ce\}.$$

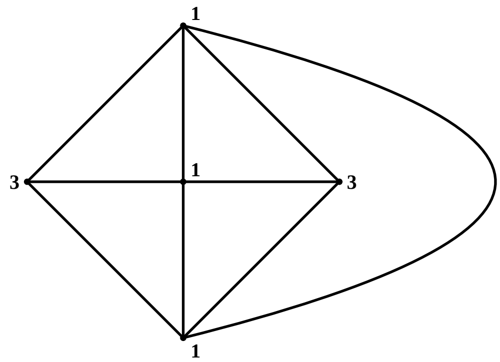
We can realize G geometrically in $\mathbb{R}^3$ as the 1-skeleton of a triangular bipyramid. Let X be the polytope of its geometric realization. At the cost of losing some symmetry, but simplifying the illustrations, we prefer to picture X as embedded into the plane $\mathbb{R}^2$:



We define a piecewise linear function $f : X \to [0, \infty)$ by specifying its values at the vertices:

$$f(d) = f(e) = 3 \qquad \text{and} \qquad f(a) = f(b) = f(c) = 1.$$

We now calculate **ucat** and $\mathbf{ucat}^\infty$ for this function.



Proposition 4.8. *The unimodal* ∞*-category of* f *is* $\mathbf{ucat}^\infty(f) = 2$.

Proof. Clearly, f is not unimodal. However, it has a unimodal ∞-decomposition of length 2. To construct it, we first subdivide the edges ab, ac and bc, by adding three points on each of them. Let i, j, k be the vertices added on ab, so that it is now replaced by four edges ai, ij, jk, kb. Similarly, add vertices p, q, r on bc to subdivide it into bp, pq, qr, rc. Finally, add x, y, z on the edge ac to subdivide it into cx, xy, yz, za.

Now the decomposition u_1, u_2 can be defined by piecewise linear functions defined by the values on the vertices of this subdivision. Namely, take

$$u_1(d) = 3, \qquad u_1|_{\{a,b,c,j,k,q,r,y,z\}} \equiv 1 \qquad \text{and} \qquad u_1|_{\{e,i,p,x\}} \equiv 0,$$
$$u_2(e) = 3, \qquad u_2|_{\{a,b,c,i,j,p,q,x,y\}} \equiv 1 \qquad \text{and} \qquad u_2|_{\{d,k,r,z\}} \equiv 0.$$

These have the desired properties: they are unimodal and $f = \max\{u_1, u_2\}$. $\square$

Proposition 4.9. *The unimodal category of f is* $\mathbf{ucat}(f) = 3$.

Proof. A unimodal decomposition of length 3 can easily be constructed explicitly so this is left to the reader.

It remains to show that there is no unimodal decomposition of length 2. Again, we argue by contradiction. Suppose $f = u_1 + u_2$ is such a decomposition. If $r > 0$, define superlevel sets

$$Q_1(r) = u_1^{-1}[r, \infty) \qquad \text{and} \qquad Q_2(r) = u_2^{-1}[r, \infty).$$

Without loss of generality, we may assume that $u_1(d) \geq u_2(d)$, in other words, $u_1(d) \geq \frac{3}{2}$. By unimodality, it follows that $u_1(e) \leq 1$ (otherwise $u_1(e) = r > 1$ and the points d and e would have to lie in separate components of $Q_1(\min\{\frac{3}{2}, r\})$). It follows that $u_2(e) \geq 2$. Using unimodality again, we have $u_2(d) \leq 1$ (otherwise $u_2(d) = r > 1$ and the points d and e would have to lie in separate components of $Q_2(\min\{2, r\})$). It follows that $u_1(d) \geq 2$. Let K be the subspace of X consisting of the edges ab, ac and bc.

Since $u_1(a)+u_2(a) = u_1(b)+u_2(b) = u_1(c)+u_2(c) = 1$, at least three of the values $u_1(a), u_2(a), u_1(b), u_2(b), u_1(c), u_2(c)$ are $\geq \frac{1}{2}$. Two of these three values necessarily correspond to the same function u_i, $i = 1, 2$. Therefore, without loss of generality (renaming the vertices and functions if necessary), we may assume that $u_1(b) \geq \frac{1}{2}$ and $u_1(c) \geq \frac{1}{2}$. Observe that this implies that $u_1(x) \geq \frac{1}{2}$ for any point x on the edge bc (otherwise, we would have $u_2(b) \leq \frac{1}{2}, u_2(c) \leq \frac{1}{2}, u_2(x) = r > \frac{1}{2}$ and $u_2(e) \geq 2$, contradicting unimodality as points x and e would lie in different components of $Q_2(r)$). But then, u_1 has at least one zero z somewhere in the interior of the union of segments bd and cd. (Otherwise $Q_1(r)$ would contain the whole cycle bc, bd, cd for some $r > 0$, contradicting unimodality.) This implies that $u_2(z) = r > 1$, contradicting unimodality, as this means that z and e lie in different components of $Q_2(\min\{r, 2\})$. This concludes the proof. □

Therefore, the monotonicity conjecture fails for $0 < p_1 < p_2 = \infty$ as well.

Remark 4.2. Note that the basic idea underlying the proof is the fact that a cycle of odd length has chromatic number 3. It remains an open question what exactly the connection between chromatic numbers and **ucat** is for general graphs.

4.4 Euclidean Plane: First Counterexample

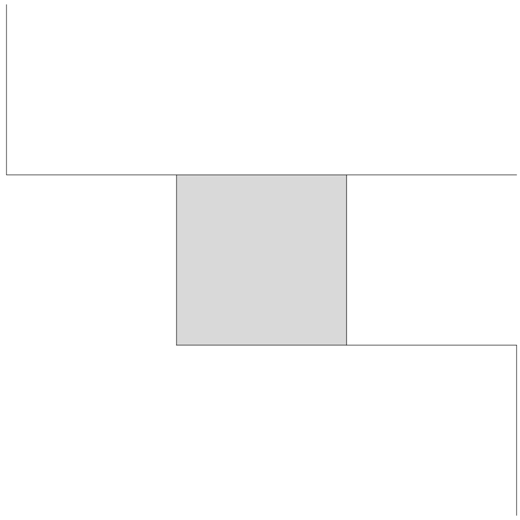
In our construction of the counterexamples to the monotonicity conjecture on the two graphs above, we have exploited the fact that these graphs contain cycles. As we have also seen, the unimodal p-category is indeed monotone in p for $X = \mathbb{R}$. The question then arises whether it is essential that the space X has non-trivial homology for such counterexamples to exist. We show that the answer to this question is also negative by constructing two counterexamples to the monotonicity conjecture in the Euclidean plane $X = \mathbb{R}^2$.

4.4.1 Concise Description

The first counterexample we give is motivated by the first counterexample in the case when X is a graph, so it bears some resemblance to it. For simplicity, whenever

$p = (x, y) \in \mathbb{R}^2$, we write $p^* = (y, x)$. We also adopt the convention that $\times$ binds more strongly than $\cup$. Let $a = (3, 1)$ and define the following compact subset of $\mathbb{R}^2$:

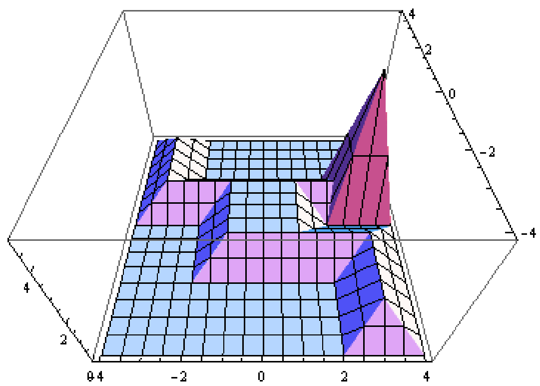
$$K = [-1,1]\times[-1,1] \cup [-3,3]\times\{1\} \cup [1,3]\times\{-1\} \cup \{-3\}\times[1,3] \cup \{3\}\times[-3,-1].$$

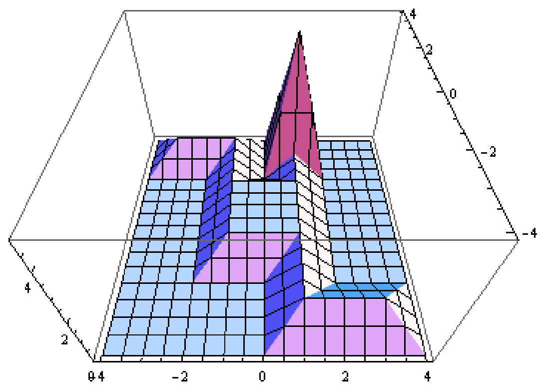


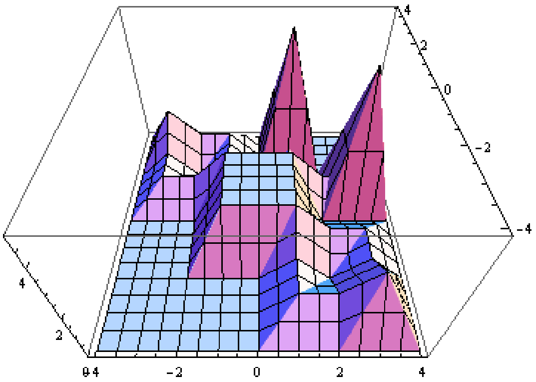
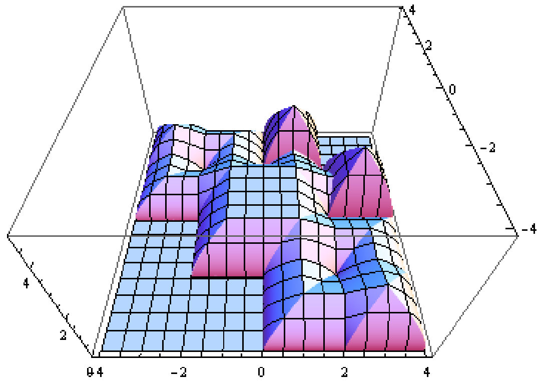
Our first counterexample can be concisely described using the ∞-distance to the set K. We define the following functions $u_1, u_2, F, f : \mathbb{R}^2 \to [0, \infty)$:

$$\begin{aligned} u_1(p) &= \max\big\{0, 1 - d_\infty\big(p, K\big), 5 - 5d_\infty\big(p, a\big)\big\}, \\ u_2(p) &= u_1(p^*), \\ F(p) &= u_1(p) + u_2(p), \\ f(p) &= \sqrt{u_1(p) + u_2(p)}. \end{aligned}$$

For convenience, the following are the graphs of u_1, u_2, F and f:





Note that each of these is continuous and compactly supported. Our main claim is the following:

Proposition 4.10. *The function f is a counterexample to the conjecture. Concretely,* $\mathbf{ucat}^2(f) = \mathbf{ucat}(F) = 2$ *and* $\mathbf{ucat}(f) \geq 3$.

The rest of this section is devoted to proving this proposition.

4.4.2 Direct Description

We can make a straightforward observation regarding the nature of the functions we have defined.

Observation 4.11. *The functions $u_1, u_2 : \mathbb{R}^2 \to [0, \infty)$ are piecewise linear and unimodal. The function $F : \mathbb{R}^2 \to [0, \infty)$ is piecewise linear. It is not unimodal, but is by definition the sum of two unimodal functions.*

Piecewise linearity follows from the properties of the ∞-distance. For concreteness, we describe the decomposition of the plane, with respect to which the functions are piecewise linear, explicitly. The main advantage of u and F being piecewise linear is that other facts about these functions can be verified completely computationally.

Before we begin, we need some notation. Two points $p_1, p_2 \in \mathbb{R}^2$ determine a segment

$$p_1p_2 = \{p \in \mathbb{R}^2 \mid \exists t \in [0, 1] : p = (1 - t)p_1 + tp_2\}.$$

We write $p_1p_2 \ldots p_n$ for the union of segments $p_1p_2, p_2p_3, \ldots, p_{n-1}p_n$. If $p_1p_2 \ldots p_np_1$ is a topological circle in $\mathbb{R}^2$, it is the boundary of a uniquely determined compact set in $\mathbb{R}^2$, which we denote by $\overline{p_1p_2 \ldots p_n}$.

We can describe u_1, u_2 and F as piecewise linear functions defined by their values on the vertices of a polygonal decomposition of $S = \operatorname{supp} f$ consisting of 44 vertices, 95 edges and 52 faces, namely triangles, trapezoids and two non-convex quadrilaterals. Note that we do not count the "face at infinity" and we consider parallelograms to be a special case of trapezoids. Some care must be taken as not every choice of values at the vertices of a quadrilateral can be extended to a linear function. First, we list the vertices (indexed lexicographically):

$$\begin{array}{llll}
x_1 = (-4,0), & x_{12} = (-2,4), & x_{23} = (1,-1), & x_{34} = (2,2), \\
x_2 = (-4,2), & x_{13} = (-1,-1), & x_{24} = (1,1), & x_{35} = (2,4), \\
x_3 = (-4,4), & x_{14} = (-1,1), & x_{25} = (1,2), & x_{36} = (\frac{11}{5},3), \\
x_4 = (-3,1), & x_{15} = (-1,2), & x_{26} = (1,\frac{11}{5}), & x_{37} = (3,-3), \\
x_5 = (-3,2), & x_{16} = (-1,3), & x_{27} = (1,3), & x_{38} = (3,-2), \\
x_6 = (-3,3), & x_{17} = (0,-4), & x_{28} = (2,-4), & x_{39} = (3,-1), \\
x_7 = (-2,-2), & x_{18} = (0,-2), & x_{29} = (2,-3), & x_{40} = (3,1), \\
x_8 = (-2,0), & x_{19} = (0,2), & x_{30} = (2,-2), & x_{41} = (4,-4), \\
x_9 = (-2,1), & x_{20} = (0,4), & x_{31} = (2,-1), & x_{42} = (4,-2), \\
x_{10} = (-2,2), & x_{21} = (1,-3), & x_{32} = (2,0), & x_{43} = (4,0), \\
x_{11} = (-2,3), & x_{22} = (1,-2), & x_{33} = (2,1), & x_{44} = (4,2).
\end{array}$$

We omit listing the edges $h_1, h_2, \dots, h_{95}$ (ordered lexicographically by the indices of the vertices) as they are simply the edges of the 52 polygons in the decomposition. Finally, we list the faces, using the notation defined above (ordered lexicographically by the corresponding sets of vertices):

$$\begin{array}{llll}
f_1 = \overline{x_1x_4x_5x_2}, & f_{14} = \overline{x_{10}x_{16}x_{11}}, & f_{27} = \overline{x_{19}x_{26}x_{34}x_{27}}, & f_{40} = \overline{x_{29}x_{37}x_{38}x_{30}}, \\
f_2 = \overline{x_1x_8x_9x_4}, & f_{15} = \overline{x_{10}x_{15}x_{16}}, & f_{28} = \overline{x_{20}x_{27}x_{35}}, & f_{41} = \overline{x_{30}x_{39}x_{31}}, \\
f_3 = \overline{x_2x_6x_3}, & f_{16} = \overline{x_{11}x_{16}x_{20}x_{12}}, & f_{29} = \overline{x_{21}x_{30}x_{22}}, & f_{42} = \overline{x_{30}x_{38}x_{39}}, \\
f_4 = \overline{x_2x_5x_6}, & f_{17} = \overline{x_{13}x_{23}x_{24}x_{14}}, & f_{30} = \overline{x_{21}x_{29}x_{30}}, & f_{43} = \overline{x_{31}x_{39}x_{43}x_{32}}, \\
f_5 = \overline{x_3x_6x_{12}}, & f_{18} = \overline{x_{14}x_{19}x_{15}}, & f_{31} = \overline{x_{22}x_{30}x_{31}x_{23}}, & f_{44} = \overline{x_{32}x_{36}x_{33}}, \\
f_6 = \overline{x_4x_{10}x_5}, & f_{19} = \overline{x_{14}x_{24}x_{19}}, & f_{32} = \overline{x_{23}x_{32}x_{24}}, & f_{45} = \overline{x_{32}x_{40}x_{34}x_{36}}, \\
f_7 = \overline{x_4x_9x_{10}}, & f_{20} = \overline{x_{15}x_{19}x_{20}x_{16}}, & f_{33} = \overline{x_{23}x_{31}x_{32}}, & f_{46} = \overline{x_{32}x_{43}x_{40}}, \\
f_8 = \overline{x_5x_{10}x_{11}x_6}, & f_{21} = \overline{x_{17}x_{21}x_{22}x_{18}}, & f_{34} = \overline{x_{24}x_{33}x_{34}x_{25}}, & f_{47} = \overline{x_{33}x_{36}x_{34}}, \\
f_9 = \overline{x_6x_{11}x_{12}}, & f_{22} = \overline{x_{17}x_{28}x_{29}x_{21}}, & f_{35} = \overline{x_{24}x_{32}x_{33}}, & f_{48} = \overline{x_{34}x_{40}x_{44}}, \\
f_{10} = \overline{x_7x_{13}x_{14}x_8}, & f_{23} = \overline{x_{18}x_{22}x_{23}}, & f_{36} = \overline{x_{25}x_{34}x_{26}}, & f_{49} = \overline{x_{37}x_{42}x_{38}}, \\
f_{11} = \overline{x_7x_{18}x_{23}x_{13}}, & f_{24} = \overline{x_{19}x_{27}x_{20}}, & f_{37} = \overline{x_{27}x_{34}x_{35}}, & f_{50} = \overline{x_{37}x_{41}x_{42}}, \\
f_{12} = \overline{x_8x_{14}x_9}, & f_{25} = \overline{x_{19}x_{24}x_{25}}, & f_{38} = \overline{x_{28}x_{37}x_{29}}, & f_{51} = \overline{x_{38}x_{42}x_{43}x_{39}}, \\
f_{13} = \overline{x_9x_{14}x_{15}x_{10}}, & f_{26} = \overline{x_{19}x_{25}x_{26}}, & f_{39} = \overline{x_{28}x_{41}x_{37}}, & f_{52} = \overline{x_{40}x_{43}x_{44}}.
\end{array}$$

We can now state the alternative descriptions of u_1, u_2 and F. The function u_1 can be defined on the vertices of the above decomposition:

$$u_1(x_i) = \begin{cases} 5; & i = 40, \\ 1; & i = 4,5,6,9,13,14,23,24,31,33,36,37,38,39, \\ 0; & \text{elsewhere.} \end{cases}$$

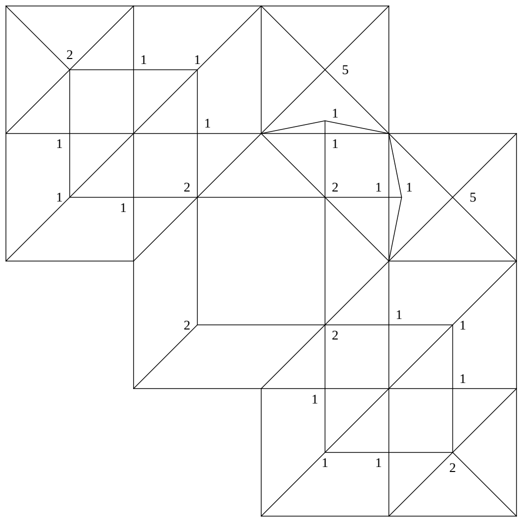
Similarly, we have

$$u_2(x_i) = \begin{cases} 5; & i = 27, \\ 1; & i = 6,11,13,14,15,16,21,22,23,24,25,26,29,37, \\ 0; & \text{elsewhere.} \end{cases}$$

Summing these, we obtain

$$F(x_i) = \begin{cases} 5; & i = 27, 40, \\ 2; & i = 6, 13, 14, 23, 24, 37, \\ 1; & i = 4, 5, 9, 11, 15, 16, 21, 22, 25, 26, 29, 31, 33, 36, 38, 39, \\ 0; & \text{elsewhere.} \end{cases}$$

The decomposition (and the function F) can be pictured as follows:



In the proofs we will also make use of the following points (see the picture below):

$$\begin{aligned} a &= (3,1) \\ a_1 &= a + (-\tfrac{4}{5}, -\tfrac{4}{5}) \\ a_2 &= a + (\tfrac{4}{5}, -\tfrac{4}{5}) \\ a_3 &= a + (\tfrac{4}{5}, \tfrac{4}{5}) \\ a_4 &= a + (-\tfrac{4}{5}, \tfrac{4}{5}) \\ a_5 &= a + (-\tfrac{4}{5}, 0) \\ b &= a^* \\ b_1 &= a_1^* \\ b_2 &= a_2^* \\ b_3 &= a_3^* \\ b_4 &= a_4^* \\ b_5 &= a_5^* \end{aligned}$$

$$\begin{aligned} c_1 &= (2,1) \\ c_2 &= (\tfrac{3}{2}, \tfrac{1}{2}) \\ c_3 &= (\tfrac{3}{2}, -\tfrac{1}{2}) \\ c_4 &= (2,-1) \\ c_5 &= (1,-2) \\ c_6 &= (\tfrac{1}{2}, -\tfrac{3}{2}) \\ c_7 &= (-\tfrac{3}{2}, -\tfrac{3}{2}) \\ c_8 &= c_6^* \\ c_9 &= c_5^* \\ c_{10} &= c_4^* \\ c_{11} &= c_3^* \\ c_{12} &= c_2^* \\ c_{13} &= c_1^* \end{aligned}$$

$$\begin{aligned} d &= (3,-3) \\ d_1 &= (2,-3) \\ d_2 &= (\tfrac{5}{2}, -\tfrac{7}{2}) \\ d_3 &= (\tfrac{7}{2}, -\tfrac{7}{2}) \\ d_4 &= (\tfrac{7}{2}, -\tfrac{5}{2}) \\ d_5 &= (3,-2) \\ e &= (-3,3) \\ e_1 &= d_1^* \\ e_2 &= d_2^* \\ e_3 &= d_3^* \\ e_4 &= d_4^* \\ e_5 &= d_5^* \end{aligned}$$

$$\begin{aligned} z_0 &= (2,-2) \\ z_1 &= (3,-1) \\ z_2 &= (1,-3) \\ w_0 &= z_0^* \\ w_1 &= z_1^* \\ w_2 &= z_2^* \\ q_1 &= (-1,-1) \\ q_2 &= (1,-1) \\ q_3 &= (1,1) \\ q_4 &= (-1,1) \end{aligned}$$

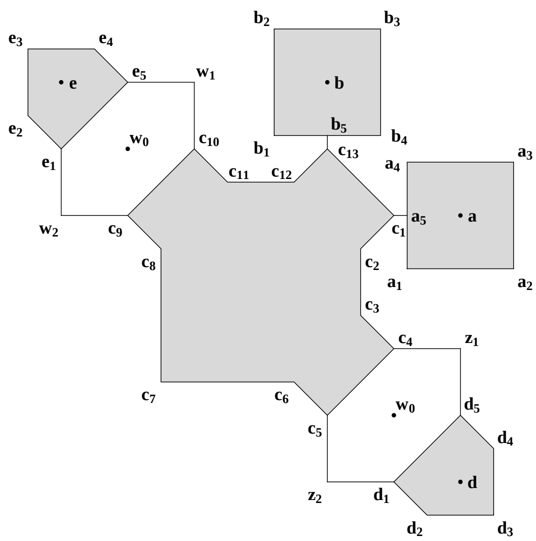
Finally, we also need the following sets:

$$
\begin{aligned}
A &= \overline{a_1a_2a_3a_4}, & P_a &= a_5c_1,\\
B &= \overline{b_1b_2b_3b_4}, & P_b &= b_5c_{13},\\
C &= \overline{c_1c_2c_3c_4c_5c_6c_7c_8c_9c_{10}c_{11}c_{12}c_{13}}, & Z_1 &= c_5z_1d_1,\\
D &= \overline{d_1d_2d_3d_4d_5}, & Z_2 &= c_4z_2d_5,\\
E &= \overline{e_1e_2e_3e_4e_5}, & W_1 &= c_9w_1e_1,\\
K &= aw_2e \cup q_1z_1d \cup \overline{q_1q_2q_3q_4}, & W_2 &= c_{10}w_2e_5.
\end{aligned}
$$

The following two observations are straightforward, so we omit their proofs.

Observation 4.12. *The superlevel set $Q = f^{-1}[1,\infty) = F^{-1}[1,\infty)$ can be expressed as follows:*

$$Q = A \cup B \cup C \cup D \cup E \cup P_a \cup P_b \cup Z_1 \cup Z_2 \cup W_1 \cup W_2.$$



Observation 4.13. *The function f has the following properties:*

- $f(z_0) = f(w_0) = 0$,
- $f(a) = f(b) = \sqrt{5}$,
- $f(p) = 1$ *for all* $p \in \partial Q$,[11]
- $f(p) \le \sqrt{2}$ *for all* $p \notin A \cup B$.

[11] Here ∂Q means the topological boundary of Q.

4.4.3 Proof

Before proceeding to the main proof, we require the following lemma.

Lemma 4.14. *Let $I = [0,1]$. Suppose $\gamma_1, \gamma_2 : I \to I \times I$ are paths such that $\gamma_1(0) = (0,0)$, $\gamma_1(1) = (1,1)$, $\gamma_2(0) = (0,1)$ and $\gamma_2(1) = (1,0)$. Then $\gamma_1(I) \cap \gamma_2(I) \neq \emptyset$.*

Proof. This follows directly from [49, Lemma 2] by taking $a = c = 0$, $b = d = 1$, $h = \gamma_1$ and $v = \gamma_2$. $\square$

We can now prove our proposition.

Proof of Proposition 4.10. The first part of the claim, that $\mathbf{ucat}^2(f) = 2$, follows directly from Observation 4.11.

To prove the second part, suppose for the sake of contradiction that there exists a decomposition $f = v_a + v_b$, where v_a and v_b are unimodal.[12] If $r > 0$, define superlevel sets

$$Q_a(r) = v_a^{-1}[r, \infty), \qquad \text{and} \qquad Q_b(r) = v_b^{-1}[r, \infty).$$

We are especially interested in

$$Q_a = Q_a(1) \qquad \text{and} \qquad Q_b = Q_b(1).$$

Since v_a and v_b are unimodal, the sets $Q_a(r)$ and $Q_b(r)$ (where $r > 0$) are all contractible. In particular, they are path-connected. If $r_1 \leq r_2$, the inclusions

$$Q_a(r_1) \supseteq Q_a(r_2), \qquad \text{and} \qquad Q_b(r_1) \supseteq Q_b(r_2)$$

hold. Since $v_a, v_b \leq f$, we also have

$$Q_a, Q_b \subseteq Q.$$

By Observation 4.13, $f(p) \leq \sqrt{2}$ for $p \notin A \cup B$. If $v_a(p) \geq 1$ and $v_b(p) \geq 1$, we have $f(p) \geq 2$, so $p \in A \cup B$. This shows that $Q_a \cap Q_b \subseteq A \cup B$.

Observe that $v_a(a) + v_b(a) = \sqrt{5}$ holds. By interchanging v_a and v_b, if necessary, we may assume without loss of generality that $v_a(a) \geq \frac{\sqrt{5}}{2}$. (Which is in fact the reason why we chose this notation for v_a and v_b.)

Step 1. *The following inequalities are satisfied:*

- $v_a(p) \leq 1$ *for all* $p \in Q \setminus A$,
- $v_b(p) \leq 1$ *for all* $p \in Q \setminus B$,
- $v_a(a) \geq \sqrt{5} - 1$ *and*
- $v_b(b) \geq \sqrt{5} - 1$.

[12] The notation is not meant to imply any relation with a and b at this point.

Proof. To prove the first of these, it suffices to show that $Q_a(r) \subseteq A$ for all $r > 1$. We know that $Q_a(r) \subseteq Q_a \subseteq Q$ holds. Note that $Q \setminus \{a_5\}$ is not path connected. The path component of a in $Q \setminus \{a_5\}$ is a subset of A. Since $v_a(a_5) \leq f(a_5) = 1$, the point a_5 is not contained in $Q_a(r)$. Therefore, the path component of a in $Q_a(r)$ is also a subset of A and we are done.

Since $v_a(b)+v_b(b) = \sqrt{5}$, this means that $v_b(b) \geq \sqrt{5}-1$. The fact that $v_b(p) \leq 1$ holds for all $p \in Q \setminus B$ is obtained by a symmetric argument. This also implies that $v(a) \geq \sqrt{5}-1$ holds. □

Step 2. *Let $S = \{c_9, c_{10}\}$ or $S = \{c_4, c_5\}$. Then there exist points $p, q \in S$ such that $v_a(p) = 1$ and $v_b(q) = 1$.*

Proof. We show this for $S = \{c_4, c_5\}$, the other proof is symmetric. Note that $f(z_0) = 0$, so $v_a(z_0) = v_b(z_0) = 0$. For each $t \in (0, 1)$ define $c_4(t) = (1-t)c_4 + tc_3$ and $c_5(t) = (1-t)c_5 + tc_6$. Define

$$L(t) = D \cup Z_1 \cup Z_2 \cup c_4c_4(t) \cup c_4(t)c_5(t) \cup c_5(t)c_5.$$

Note that $L(t)$ is compact, so v_b attains its minimum in $L(t)$, say $v_b(p_t) = m_t$. But m_t cannot be positive: if $m_t > 0$, we would have $L(t) \subseteq Q_b(m_t)$ and $z_0 \notin Q_b(m_t)$, implying the existence of a retraction $Q_b(m_t) \to L(t)$, which is impossible, since $Q_b(m_t)$ is contractible.

Therefore, $v_b(p_t) = 0$, or since f is at least 1 on $L(t)$, $v_a(p_t) \geq 1$. But we know that $v_a(p) \leq 1$ holds for $p \notin A$, so we must have $v_a(p_t) = 1$. This also implies that $p \in \partial Q$.

There are now two possibilities: if for some t we obtain $p_t \in D \cup Z_1 \cup Z_2$, we also have a path in Q_a from p_t to a, which must necessarily pass either through c_4 or c_5, since p_t and a lie in different path components of $Q \setminus \{c_4, c_5\}$. The only remaining possibility is that $p_t \in c_4c_4(t) \cup c_4(t)c_5(t) \cup c_5(t)c_5$ for all $t \in (0, 1)$. Since $p_t \in \partial Q$, this means that $p_t \in c_4c_4(t) \cup c_5c_5(t)$ holds for all t. Therefore we may choose a convergent subsequence $(p_{t_n})_n$ of $(p_{\frac{1}{k}})_k$ such that $v_a(p_{t_n}) = 1$ for all n. This sequence converges either to c_4 or c_5, so one of $v_a(c_4) = 1$, $v_a(c_5) = 1$ must hold.

The proof that one of $v_b(c_4) = 1$, $v_b(c_5) = 1$ holds is symmetric. □

Step 3. *The equalities $v_a(c_1) = 1$ and $v_b(c_{13}) = 1$ hold.*

Proof. This follows from the previous step. Let S be any of the two sets from the previous step. Let $p, q \in S$ be points such that $v_a(p) = 1$ and $v_b(q) = 1$. There are paths in Q_a and Q_b from a to p and from b to q, respectively. The first path must cross c_1, since a and p lie in different path components of $Q \setminus \{c_1\}$ and the second one must cross c_{13}, since b and q lie in different path components of $Q \setminus \{c_{13}\}$. □

Step 4. *The results of the previous steps contradict each other.*

Proof. Let $p \in \{c_9, c_{10}\}$ have the property that $v_a(p) = 1$ and let q be the unique element of $\{c_9, c_{10}\} \setminus \{p\}$. Let $q' \in \{c_4, c_5\}$ have the property that $v_b(q') = 1$ and let $p' \in \{c_4, c_5\} \setminus \{q'\}$. These points exist by step 2. There is a path $\gamma_a : I \to Q_a$ from p to c_1 and a path $\gamma_b : I \to Q_b$ from q' to c_{13}.

Observe that $\gamma_a(I) \cap (B \cup P_b) = \emptyset$ and $\gamma_b(I) \cap (A \cup P_a) = \emptyset$, since $c_1 \notin Q_b$ and $c_{13} \notin Q_a$. We can also ensure that $\gamma_a(I) \cap (E \cup W_1 \cup W_2) = \{p\}$ and $\gamma_a(I) \cap (A \cup P_a) = \{c_1\}$

by cropping the path at both ends (formally, take $t_1 = \sup\{t \in [0,1] \mid \gamma_a(t) = p\}$ and $t_2 = \inf\{t \in [t_1,1] \mid \gamma_a(t) = c_1\}$ and reparametrize the restriction of γ_a to $[t_1,t_2]$). In the same way we can ensure that $\gamma_b(I) \cap (D \cup Z_1 \cup Z_2) = \{q'\}$ and $\gamma_b(I) \cap (B \cup P_b) = \{c_{13}\}$. Assume, therefore, without loss of generality that γ_a and γ_b have these properties.

Now, let $R_a : Q_a \to Q_a \setminus \big((D \cup Z_1 \cup Z_2) \setminus \{p'\}\big)$ be the retraction that takes the points of $(D \cup Z_1 \cup Z_2) \setminus \{p'\}$ to p' and let $R_b : Q_b \to Q_b \setminus \big((E \cup W_1 \cup W_2) \setminus \{q\}\big)$ be the retraction that takes the points of $(E \cup W_1 \cup W_2) \setminus \{q\}$ to q. These retractions are well defined, since $q' \notin Q_a$ and $p \notin Q_b$. Note that $R_a \circ \gamma_a$ is a path in $Q_a \cap C$ and $R_b \circ \gamma_b$ is a path in $Q_b \cap C$ and these two paths have the same endpoints as γ_a and γ_b, respectively.

Note that ∂C is a Jordan curve and that $\partial C \setminus \{c_1, p\}$ has two path components, each of which contains exactly one of the points c_{13}, q'. Using the Jordan-Schönflies theorem [8, Chapter III], we obtain the situation in which Lemma 4.14 applies and we can conclude that the paths $R_a \circ \gamma_a$ and $R_b \circ \gamma_b$ intersect somewhere in C. But this is a contradiction, as we have already established that $Q_a \cap Q_b \subseteq A \cup B$. $\square$

This contradiction shows that $\mathbf{ucat}(f) \geq 3$, which concludes the proof of our proposition. $\square$

Remark 4.3. The homology of the space is trivial, but the first homology of the superlevel sets is not. This seems to be an essential feature of the counterexample, as it enables us to force certain values upon the functions in an unimodal decomposition. An explicit unimodal decomposition of length 3 can be constructed for f, but we do not describe it here, as it has no bearing on the validity of the counterexample. Furthermore, by modifying the function values at the vertices, we can obtain such counterexamples for any pair $0 < p_1 < p_2 < \infty$. We expect such counterexamples to exist on $\mathbb{R}^m$ for any $m \geq 2$.

4.5 Euclidean Plane: Second Counterexample

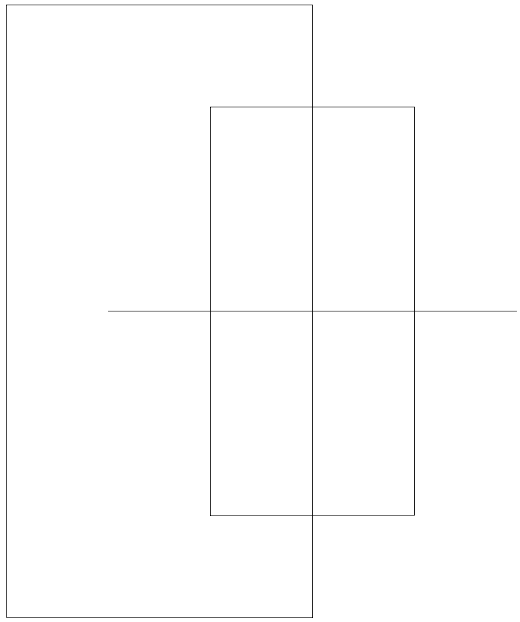
We need a different idea to show that the monotonicity conjecture fails for $X = \mathbb{R}^2$ also in the case of $0 < p_1 < p_2 = \infty$. The example we give is completely analogous to the second example in the case of graphs. As with the first example on $\mathbb{R}^2$, we also describe it in two ways: using the ∞-distance and as a piecewise linear function.

4.5.1 Concise Description

Let $d_0 = (-4, 0)$ and $e_0 = (4, 0)$. Define the following compact subset of $\mathbb{R}^2$:

$$K = [-4,4]\times\{0\}\cup\{-6,0\}\times[-6,6]\cup[-6,0]\times\{-6,6\}\cup\{-2,2\}\times[-4,4]\cup[-2,2]\times\{-4,4\}.$$

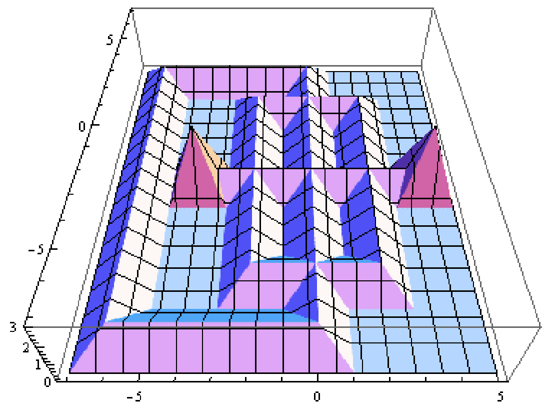
This can be pictured as follows:



Define $f : \mathbb{R}^2 \to [0, \infty)$:

$$f(p) = \max\big\{0, 1 - d_\infty\big(p, K\big), 3 - 3d_\infty\big(p, d_0\big), 3 - 3d_\infty\big(p, e_0\big)\big\}.$$

Note that f is continuous and compactly supported. Here is the graph of f:



Our main claim is the following:

Proposition 4.15. *The function f is a counterexample to the conjecture. Concretely,* $\mathbf{ucat}^\infty(f) = 2$ *and* $\mathbf{ucat}(f) = 3$.

The rest of this section is devoted to proving this proposition.

4.5.2 Direct Description

For the same reason as our first example, $f : \mathbb{R}^2 \to [0, \infty)$ is actually a piecewise linear function, so it can be described by its function values at the vertices of a polygonal decomposition of its support $S = \operatorname{supp} f$. This decomposition consists of 47 vertices, 94 edges and 46 faces, namely triangles, trapezoids and two non-convex quadrilaterals. Again note that not every choice of values at the vertices of a trapezoid can be extended to a linear function, but in our case such issues do not arise as the two function values on each of two parallel sides agree. We first list the vertices (indexed lexicographically):

$$
\begin{aligned}
x_1 &= (-7,-7), & x_{13} &= (-3,1), & x_{25} &= (0,-4), & x_{37} &= (2,-4),\\
x_2 &= (-7,7), & x_{14} &= (-3,5), & x_{26} &= (0,0), & x_{38} &= (2,0),\\
x_3 &= (-6,-6), & x_{15} &= (-2,-4), & x_{27} &= (0,4), & x_{39} &= (2,4),\\
x_4 &= (-6,6), & x_{16} &= (-2,0), & x_{28} &= (0,6), & x_{40} &= (3,-5),\\
x_5 &= (-5,-5), & x_{17} &= (-2,4), & x_{29} &= (1,-7), & x_{41} &= (3,-1),\\
x_6 &= (-5,-1), & x_{18} &= (-1,-5), & x_{30} &= (1,-5), & x_{42} &= (3,1),\\
x_7 &= (-5,1), & x_{19} &= (-1,-3), & x_{31} &= (1,-3), & x_{43} &= (3,5),\\
x_8 &= (-5,5), & x_{20} &= (-1,-1), & x_{32} &= (1,-1), & x_{44} &= (\tfrac{10}{3},0),\\
x_9 &= (-4,0), & x_{21} &= (-1,1), & x_{33} &= (1,1), & x_{45} &= (4,0),\\
x_{10} &= (-\tfrac{10}{3},0), & x_{22} &= (-1,3), & x_{34} &= (1,3), & x_{46} &= (5,-1),\\
x_{11} &= (-3,-5), & x_{23} &= (-1,5), & x_{35} &= (1,5), & x_{47} &= (5,1).\\
x_{12} &= (-3,-1), & x_{24} &= (0,-6), & x_{36} &= (1,7), & &
\end{aligned}
$$

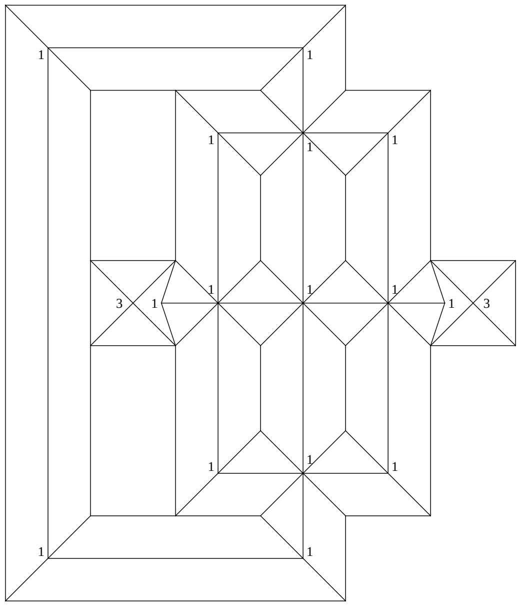
We again omit listing the edges $h_1, h_2, \dots, h_{94}$ and proceed to the faces:

$$
\begin{aligned}
f_1 &= \overline{x_1x_3x_4x_2}, & f_{13} &= \overline{x_{11}x_{15}x_{16}x_{12}}, & f_{25} &= \overline{x_{21}x_{26}x_{27}x_{22}}, & f_{36} &= \overline{x_{27}x_{39}x_{43}x_{35}},\\
f_2 &= \overline{x_1x_{29}x_{24}x_3}, & f_{14} &= \overline{x_{11}x_{18}x_{25}x_{15}}, & f_{26} &= \overline{x_{23}x_{27}x_{28}}, & f_{37} &= \overline{x_{31}x_{37}x_{38}x_{32}},\\
f_3 &= \overline{x_2x_4x_{28}x_{36}}, & f_{15} &= \overline{x_{13}x_{16}x_{17}x_{14}}, & f_{27} &= \overline{x_{24}x_{29}x_{30}x_{25}}, & f_{38} &= \overline{x_{33}x_{38}x_{39}x_{34}},\\
f_4 &= \overline{x_3x_5x_8x_4}, & f_{16} &= \overline{x_{14}x_{17}x_{27}x_{23}}, & f_{28} &= \overline{x_{25}x_{31}x_{32}x_{26}}, & f_{39} &= \overline{x_{37}x_{40}x_{41}x_{38}},\\
f_5 &= \overline{x_3x_{24}x_{18}x_5}, & f_{17} &= \overline{x_{15}x_{19}x_{20}x_{16}}, & f_{29} &= \overline{x_{25}x_{30}x_{40}x_{37}}, & f_{40} &= \overline{x_{38}x_{41}x_{44}},\\
f_6 &= \overline{x_4x_8x_{23}x_{28}}, & f_{18} &= \overline{x_{15}x_{25}x_{19}}, & f_{30} &= \overline{x_{25}x_{37}x_{31}}, & f_{41} &= \overline{x_{38}x_{42}x_{43}x_{39}},\\
f_7 &= \overline{x_6x_9x_7}, & f_{19} &= \overline{x_{16}x_{21}x_{22}x_{17}}, & f_{31} &= \overline{x_{26}x_{33}x_{34}x_{27}}, & f_{42} &= \overline{x_{38}x_{44}x_{42}},\\
f_8 &= \overline{x_6x_{12}x_9}, & f_{20} &= \overline{x_{16}x_{20}x_{26}}, & f_{32} &= \overline{x_{26}x_{32}x_{38}}, & f_{43} &= \overline{x_{41}x_{45}x_{42}x_{44}},\\
f_9 &= \overline{x_7x_9x_{13}}, & f_{21} &= \overline{x_{16}x_{26}x_{21}}, & f_{33} &= \overline{x_{26}x_{38}x_{33}}, & f_{44} &= \overline{x_{41}x_{46}x_{45}},\\
f_{10} &= \overline{x_9x_{12}x_{10}x_{13}}, & f_{22} &= \overline{x_{17}x_{22}x_{27}}, & f_{34} &= \overline{x_{27}x_{35}x_{36}x_{28}}, & f_{45} &= \overline{x_{42}x_{45}x_{47}},\\
f_{11} &= \overline{x_{10}x_{12}x_{16}}, & f_{23} &= \overline{x_{18}x_{24}x_{25}}, & f_{35} &= \overline{x_{27}x_{34}x_{39}}, & f_{46} &= \overline{x_{45}x_{46}x_{47}}.\\
f_{12} &= \overline{x_{10}x_{16}x_{13}}, & f_{24} &= \overline{x_{19}x_{25}x_{26}x_{20}}, & & & &
\end{aligned}
$$

An alternative definition of the piecewise linear function $f : \mathbb{R}^2 \to [0,\infty)$ is then given by specifying it on the vertices as follows:

$$
f(x_i) = \begin{cases} 3; & i = 9, 45,\\ 1; & i = 3,4,10,15,16,17,24,25,26,27,28,37,38,39,44,\\ 0; & \text{otherwise.} \end{cases}
$$

The decomposition (and the function) can be pictured as follows:



4.5.3 Proof

The idea of computing $\mathbf{ucat}^\infty(f)$ is completely the same as for the second counterexample in the case of graphs.

Proposition 4.16. *The unimodal ∞-category of f is $\mathbf{ucat}^\infty(f) = 2$.*

Proof. Clearly, f is not unimodal, so it suffices to find an ∞-decomposition of length 2. We define two piecewise linear functions $v_1, v_2 : \mathbb{R}^2 \to [0, \infty)$ on a polygonal decomposition of S by specifying their values at the vertices. Start with the decomposition of S defined above. Further subdivide it[13] by adding the edges $x_{15}x_{18}, x_{17}x_{23}, x_{30}x_{37}, x_{35}x_{39}$. The function v_1 assumes the value 1 at the vertices $x_3, x_4, x_{10}, x_{15}, x_{16}, x_{17}, x_{24}, x_{25}, x_{26}, x_{27}, x_{28}$; the value 3 at the vertex x_9; and the value 0 at all the other vertices. The function v_2 assumes the value 1 at the vertices $x_3, x_4, x_{24}, x_{25}, x_{26}, x_{27}, x_{28}, x_{37}, x_{38}, x_{39}, x_{44}$; the value 3 at the vertex x_{45}; and the value 0 at all the other vertices. It is clear that $\max\{v_1, v_2\} = f$, however, these functions are not unimodal, as $x_3x_{24}x_{28}x_4x_3$ yields a non-trivial cycle in some of the superlevel sets. However, we can modify v_1 and v_2 to obtain unimodal functions $u_1, u_2 : \mathbb{R}^2 \to [0, \infty)$. To retain the property $\max\{u_1, u_2\} = f$, we modify them on sets with disjoint interiors, namely, u_1 is modified on the sets $R_1 = [-1, 1] \times [2, 3]$, $R_2 = [-1, 1] \times [-2, -1]$ and $R_3 = [-7, -5] \times [-2, -1]$, whereas u_2 is modified on the sets $R_4 = [-1, 1] \times [1, 2]$, $R_5 = [-1, 1] \times [-3, -2]$ and $R_6 = [-7, -5] \times [1, 2]$. In fact, they are modified in the same way on each of these. Namely, if R is of one of these rectangles, subdivide it into four triangles using the center point of R. Then define a piecewise linear function on R that takes the value 1 at the center point of R and value 0 at the vertices of the rectangle, and extend it by 0 to the whole plane to obtain a function $\varphi_R : \mathbb{R}^2 \to [0, \infty)$. Now, u_1 and u_2 are defined by putting

[13] This is necessary as the functions we define are not linear when restricted to some of the trapezoids of the original decomposition.

$u_1 = v_1 - \varphi_{R_1} - \varphi_{R_2} - \varphi_{R_3}$ and $u_2 = v_2 - \varphi_{R_4} - \varphi_{R_5} - \varphi_{R_6}$. (These functions are piecewise linear w.r.t. a further subdivision of S that takes into account the six rectangles.) A straightforward verification shows that u_1 and u_2 are unimodal[14], so the proof is complete. $\square$

Proving that $\mathbf{ucat}(f) = 3$ is also very similar as in the graph case. Most of the action takes place in the sublevel set $Q = f^{-1}[1, \infty)$, so before beginning the proof, we describe it explicitly. The notation we use is similar as in the first counterexample in $\mathbb{R}^2$. In addition to $d_0 = (-4, 0)$ and $e_0 = (4, 0)$, we define the following points:

$$\begin{aligned}
d_1 &= (-\tfrac{14}{3}, -\tfrac{2}{3}), & e_1 &= (\tfrac{14}{3}, -\tfrac{2}{3}), & a &= (0, 4),\\
d_2 &= (-\tfrac{10}{3}, -\tfrac{2}{3}), & e_2 &= (\tfrac{10}{3}, -\tfrac{2}{3}), & b &= (0, 0),\\
d_3 &= (-\tfrac{10}{3}, \tfrac{2}{3}), & e_3 &= (\tfrac{10}{3}, \tfrac{2}{3}), & c &= (0, -4),\\
d_4 &= (-\tfrac{14}{3}, \tfrac{2}{3}), & e_4 &= (\tfrac{14}{3}, \tfrac{2}{3}), & k_1 &= (0, -6),\\
d_5 &= (-2, -4), & e_5 &= (2, -4), & k_2 &= (-6, -6),\\
d_6 &= (-2, 0), & e_6 &= (2, 0), & k_3 &= (-6, 6),\\
d_7 &= (-2, 4), & e_7 &= (2, 4), & k_4 &= (0, 6).
\end{aligned}$$

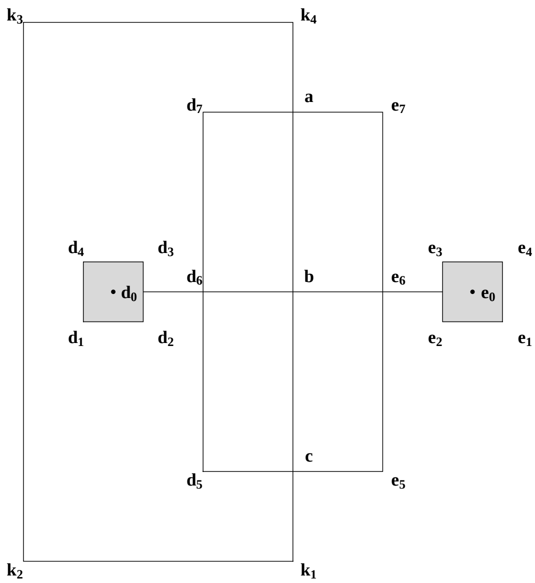
Next, we define the following sets:

$$D = \overline{d_1 d_2 d_3 d_4}, \quad E = \overline{e_1 e_2 e_3 e_4}, \quad P = d_5 e_5 e_7 d_7 d_5, \quad R = d_0 e_0, \quad K = k_1 k_2 k_3 k_4 k_1.$$

We can now give a simple description of Q.

Observation 4.17. *The superlevel set $Q = f^{-1}[1, \infty)$ can be expressed as follows:*

$$Q = D \cup E \cup P \cup R \cup K.$$



[14] We omit the unenlightening formal proof of this fact. Instead, the graphs of these unimodal functions are included in Appendix A.

We are now ready to compute the unimodal category of f. The proof in the graph case relied on the fact that a, b, c are local cut points of the graph and the various superlevel sets they appear in. In the case of $\mathbb{R}^2$, which has no local cut points, a more careful case by case analysis is required.

Proposition 4.18. *The unimodal category of f is* $\mathbf{ucat}(f) = 3$.

Proof. A unimodal decomposition of length 3 can easily be constructed explicitly: take the polygonal decomposition of S as defined in the beginning. Further subdivide it[15] by adding the edges $x_1x_4, x_4x_5, x_{15}x_{18}, x_{20}x_{25}, x_{21}x_{27}, x_{25}x_{32}, x_{27}x_{33}, x_{35}x_{39}$. Now, define piecewise linear functions $u_1, u_2, u_3 : \mathbb{R}^2 \to [0, \infty)$ by specifying their values at the vertices of this decomposition. For the sake of brevity, we only specify the nonzero values. The function u_1 is defined by taking the value 3 at the vertex x_9 and the value 1 at the vertices $x_4, x_{10}, x_{15}, x_{16}, x_{17}, x_{27}, x_{28}$. The function u_2 is defined by taking the value 3 at the vertex x_{45} and the value 1 at the vertices $x_3, x_{24}, x_{25}, x_{37}, x_{38}, x_{39}, x_{44}$. Finally, u_3 takes the value 1 at the vertex x_{26}. It is straightforward to verify that these are all unimodal[16] and that $f = u_1 + u_2 + u_3$.

To show that there is no unimodal decomposition of length 2, we argue by contradiction. Suppose $f = u_1 + u_2$ is such a decomposition. If $r > 0$, define superlevel sets

$$Q_1(r) = u_1^{-1}[r, \infty) \qquad \text{and} \qquad Q_2(r) = u_2^{-1}[r, \infty).$$

Without loss of generality, we may assume that $u_1(d_0) \geq u_2(d_0)$, in other words, $u_1(d_0) \geq \frac{3}{2}$. By unimodality, it follows that $u_1(e_0) \leq 1$ (otherwise $u_1(e_0) = r > 1$ and the points d_0 and e_0 would have to lie in separate components of $Q_1(\min\{\frac{3}{2}, r\})$, separated by K). It follows that $u_2(e_0) \geq 2$. Using unimodality again, we have $u_2(d_0) \leq 1$ (otherwise $u_2(d_0) = r > 1$ and the points d_0 and e_0 would have to lie in separate components of $Q_2(\min\{2, r\})$, separated by K). It follows that $u_1(d_0) \geq 2$.

Observe that for $i = 1, 2$, u_i must have a zero somewhere in K (if $u_i(x) \geq r > 0$ for all $x \in K$, the inclusion $K \hookrightarrow Q_1(r)$ is non-trivial on H_1), say $u_i(z_i) = 0$. It follows that $u_i(z_{3-i}) = 1$, so $Q_i(1) \cap K \neq \emptyset$. Since $Q_1(1)$ is contractible, there is a path in $Q \supseteq Q_1(1)$ from d_0 to z_2. Therefore, there is a point $q_1 \in \{a, b, c\}$ such that $u_1(q_1) = 1$ (there is no path from d_0 to $K \setminus \{a, b, c\}$ in $Q \setminus \{a, b, c\}$). Similarly, there is a point $q_2 \in \{a, b, c\}$ such that $u_2(q_2) = 1$. Since there are paths in $Q_1(1)$ from d_6 to q_1 and in $Q_2(1)$ from e_6 to q_2, we also have $u_1(d_6) = 1$ and $u_2(e_6) = 1$. Note that q_1 and q_2 are distinct and let q_3 be the third point in $\{a, b, c\}$.

There are now six cases, each of which leads to a contradiction. To avoid treating these cases separately, we proceed as follows. Let Z_1 be the unique subspace of Q which is homeomorphic to the circle, contains d_6 and q_1, but does not contain e_6 and q_2. Similarly, let Z_2 be the unique subspace of Q which is homeomorphic to the circle, contains e_6 and q_2, but does not contain d_6 and q_1. For instance, if $q_2 = c$, we have $Z_1 = abd_6d_7a$ and if $q_2 = b$ we have $Z_1 = ad_7d_6d_5ck_1k_2k_3k_4a$.

Now note that u_1 must have a zero z_1 somewhere in Z_1 (otherwise $Z_1 \hookrightarrow Q_1(r)$ would be non-trivial on H_1 for some $r > 0$). Therefore $u_2(z_1) = 1$ and by unimodality

[15] This is necessary for the same reason as in Proposition 4.16.

[16] The graphs of these unimodal functions can also be found in Appendix A.

there is a path in $Q_2(1)$ from q_2 to z_1. This path must contain the point q_3 (there is no path from q_2 to $Z_1 \setminus \{q_1, q_3, d_6\}$ in $Q \setminus \{q_1, q_3, d_6\}$), therefore $u_2(q_3) = 1$. Similarly, u_2 must have a zero z_2 somewhere in Z_2. Therefore $u_1(z_2) = 1$ and by unimodality, there is a path in $Q_1(1)$ from q_1 to z_2 which must contain the point q_3. We conclude that $u_1(q_3) = 1$. As $u_1(q_3)+u_2(q_3) = 2 \neq 1 = f(q_3)$, we have reached a contradiction, thus concluding the proof. □

Remark 4.4. Note that this counterexample can be modified to work in any $\mathbb{R}^m$, $m \geq 2$.

4.6 Proof in $\mathbb{R}^2$ if the Morse-Smale Graph is a Tree

Hickok, Villatoro and Wang describe in [43] how to compute the unimodal category of a nonresonant function $f : \mathbb{R}^2 \to [0, \infty)$ whose Morse-Smale graph is a tree. We show that their results in fact also imply that the monotonicity conjecture is true for such functions, which appears to have gone unnoticed, even though it follows from their result almost immediately. So at least in the case of Morse functions, the presence of cycles is an essential feature of counterexamples to the monotonicity conjecture.

The concept of Morse-Smale graph used in [43] seems somewhat nonstandard, and seems to be something akin to “the upper half of a splittable quasi-Morse-Smale complex” in the language of [32], which is a nice exposition of Morse-Smale complexes. This means that their structure only takes into account local maxima (as vertices of the graph) and saddles (as the edges). Local minima seem to be disregarded completely. Their definition can be phrased as follows.

Definition. A *Morse-Smale graph* associated to a Morse function $f : \mathbb{R}^2 \to [0, \infty)$ is a weighted graph, embedded in $\mathbb{R}^2$, whose vertices are the local maxima of f and whose edges are associated to the saddles of f in the following way: corresponding to each saddle, the graph has precisely one edge, which is realized as a path connecting two local maxima and passing through the saddle, so that the function values are decreasing on the portion of the path between each maximum and the saddle. The weight corresponding to a maximum $m \in \mathbb{R}^2$ is given by $f(m)$ and the weight corresponding to the edge associated to the saddle $s \in \mathbb{R}^2$ is given by $f(s)$.

Note that a Morse-Smale graph is not uniquely determined by the function.

4.6.1 Path values

Following [43], we are going to use the concept of *the path value*, however, we phrase it in a slightly different way, which we feel should be more amenable to generalization. In [43], this concept is defined using the concept of the Morse-Smale graph. We prefer to bypass this using a somewhat more general definition, that applies to general topological spaces. The benefit of this approach is that we obtain new lower bounds for general topological spaces.

Definition. Let $f : X \to [0, \infty)$ and $x_1, x \in X$. Then the *path value* from x_1 to x is the number

$$\mathrm{PV}(x_1, x) = \sup_{\gamma \in \Gamma(x_1, x)} \min_{t \in [0,1]} f(\gamma(t)),$$

where $\Gamma(x_1, x)$ is the set of all paths $\gamma : (I, 0, 1) \to (X, x_1, x)$. If $x_1, x_2, \ldots, x_n, x \in X$, we also define the *total path value* from $x_1, x_2, \ldots, x_n$ to x:

$$\mathrm{PV}(x_1, \ldots, x_n; x) = \sum_{i=1}^{n} \mathrm{PV}(x_i, x).$$

The concept of path value can be used to obtain lower bounds for **ucat**:

Proposition 4.19. *Suppose $f : X \to [0, \infty)$ is such that $\mathbf{ucat}(f) \leq n \in \mathbb{N}$. Then there exist points $x_1, x_2, \ldots, x_n$ such that*

$$\mathrm{PV}(x_1, \ldots, x_n; x) \geq f(x)$$

holds for each $x \in X$.

Proof. Let $f = \sum_{i=1}^{n} u_i$ be a unimodal decomposition and choose points $x_1, \ldots, x_n$ so that for each i, x_i is a maximum of u_i. Now, observe that for each $x \in X$ we have $u_i(x) \leq \mathrm{PV}(x_i, x)$. This is because $u_i^{-1}[u_i(x), \infty)$ is path connected, so there exists a path γ from x_i to x such that $u_i(\gamma(t)) \geq u_i(x)$ holds for all t. Therefore

$$u_i(x) = \min_{t \in [0,1]} u_i(\gamma(t)) \leq \min_{t \in [0,1]} f(\gamma(t)) \leq \mathrm{PV}(x_i, x).$$

This implies

$$f(x) = \sum_{i=1}^{n} u_i(x) \leq \sum_{i=1}^{n} \mathrm{PV}(x_i, x) = \mathrm{PV}(x_1, \ldots, x_n; x).$$

□

The converse of this proposition is not generally true, however, as the authors of [43] observe, it is almost true in the case $X = \mathbb{R}^2$ under their definition of path value. Namely, they prove the following result:

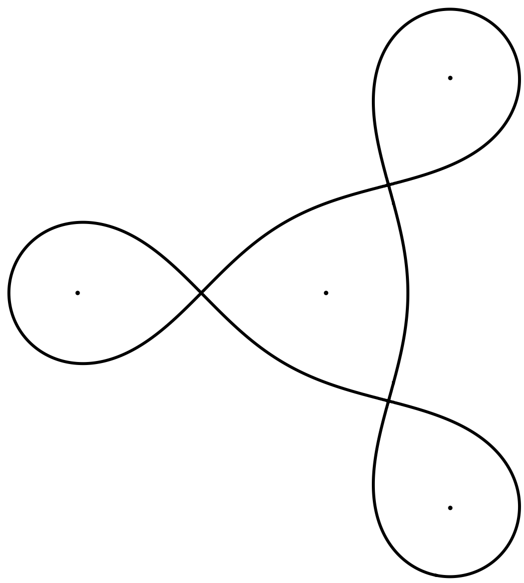
Theorem 4.20 ([43], Proposition 4.3)**.** *Suppose $f : \mathbb{R}^2 \to [0, \infty)$ is a nonresonant function whose Morse-Smale graph is a tree and there are local maxima $x_1, \ldots, x_n$ (not necessarily distinct) such that*

$$\mathrm{PV}(x_1, \ldots, x_n; x) > f(x)$$

holds for each local maximum $x \neq x_i$ $(i = 1, 2, \ldots, n)$. Then $\mathbf{ucat}(f) \leq n$.

This result relies on the fact that a nonresonant function whose Morse-Smale graph is a tree always has a Morse-Smale graph which is a path. This allows the authors to describe a general function of this kind in terms simple enough to allow for the construction of an explicit unimodal decomposition, which is what they proceed to do.

Remark 4.5. We note that the assumption of nonresonance which seems to have been overlooked by the authors of [43] is crucial here, otherwise it could happen that the Morse-Smale graph cannot be converted into a path. For instance, a function whose critical sublevel sets are as depicted in the following picture, has a Morse-Smale graph which is a tree, but which cannot be converted into a path by the procedure described in [43].



4.6.2 Monotonicity

Using the results of [43], we can now prove that the monotonicity conjecture holds for any nonresonant function $f : \mathbb{R}^2 \to [0, \infty)$ whose Morse-Smale graph is a tree. This follows almost immediately from the characterization using path values.

Theorem 4.21. *Suppose $f : \mathbb{R}^2 \to [0, \infty)$ is a nonresonant function whose Morse-Smale graph is a tree and $0 < p_1 < p_2 \leq \infty$. Then*

$$\mathbf{ucat}^{p_1}(f) \leq \mathbf{ucat}^{p_2}(f).$$

Proof. Let $g = f^{p_1}$ and $p = \frac{p_2}{p_1}$. Note that g is again nonresonant and its Morse-Smale graph is still a tree. By Lemma 2.1, the statement we wish to prove is equivalent to

$$\mathbf{ucat}(g) \leq \mathbf{ucat}(g^p).$$

Suppose $\mathbf{ucat}(g) \leq n$. Then by Theorem 4.19, there exist points $x_1, \ldots, x_n$ such that

$$\sum_{i=1}^{n} \mathrm{PV}(x_i, x) \geq g(x).$$

Now, if x is a local maximum of g distinct from all x_i, $i = 1, 2, \ldots, n$, at least two path values $\mathrm{PV}(x_i, x)$ must be nonzero, since otherwise they cannot sum to $\geq g(x)$. By the usual norm inequalities, this immediately implies

$$\sum_{i=1}^{n} \mathrm{PV}(x_i, x)^p > g(x)^p,$$

which, using Theorem 4.20, yields

$$\mathbf{ucat}(g^p) \leq n,$$

as desired. $\square$

5 Miscellanea

5.1 Higher Dimensions

5.1.1 Multimodal Functions

Given the successful application of the concept of Morse-Smale graphs which are trees in the case of $\mathbb{R}^2$, it would be desirable to have a similar concept in $\mathbb{R}^m$ for $m > 2$. In fact, such a graph can be defined if the function $f : \mathbb{R}^m \to [0, \infty)$ only has critical points of indices m and $m - 1$.

However, we may be able to generalize this a bit. The main problem which we are trying to avoid using the requirement that the Morse-Smale graph is a tree, is the presence of cycles in the superlevel sets. As a more general notion, akin to unimodality, that captures this, we propose the following:

Definition. A function $f : X \to [0, \infty)$ is *multimodal* if there is a $M > 0$ such that each superlevel set $f^{-1}[c, \infty)$ is homotopy equivalent to a finite set of points for $0 < c \leq M$ and empty for $c > M$.

Such a function cannot have any cycles that would allow us to force certain values upon the unimodal summand in the decomposition as we did with the counterexamples in the plane, so it seems more likely that the following question could admit a positive answer:

Question 1. Suppose X is a sufficiently nice space (for instance a manifold) and $f : X \to [0, \infty)$ is a multimodal function. Does this imply that $\mathbf{ucat}^p(f)$ is monotone in p?

To demonstrate that this is indeed a generalization of the case studied in [43], we now prove the following result.

Proposition 5.1. *Suppose $f : \mathbb{R}^m \to [0, \infty)$ is a multimodal nonresonant function with compact support. Then f (restricted to $f^{-1}(0, \infty)$, as per Convention in Section 2.3) has only critical points of index m and $m - 1$.*

Proof. Consider instead the function $-f$ and let $M_x = (-f)^{-1}(-\infty, x]$ for each $x < 0$. By Morse theory, it suffices to prove that this function only has critical points of index 0 and 1. Suppose $-f$ has a critical point p of index $i > 1$ and choose it so that the corresponding critical value $a = -f(p)$ is minimal. Suppose $[a - \epsilon, a + \epsilon]$ contains no other critical values. By Theorem 2.5, up to homotopy, passing a critical point of index i corresponds to attaching an i-handle, so $M_{a+\epsilon}$ is homotopy equivalent to $M_{a-\epsilon}$ with an i-handle attached. However, attaching an i-handle must either kill a homology class in H_{i-1} or create a homology class in H_i. In both cases, this is a contradiction: by multimodality, $M_{a-\epsilon}$ and $M_{a+\epsilon}$ are both finite unions of contractible sets, so such homology classes cannot exist. $\square$

5.1.2 Extension Lemma

In Section 3 we have proved Proposition 3.8 which shows that computing the unimodal category of functions on $X = [0, 1]$ and on $X = \mathbb{R}$ is basically the same thing. This fact has a straightforward generalization to the closed unit ball B^m in $\mathbb{R}^m$.

Lemma 5.2. *Suppose $f : B^m \to [0, \infty)$ is unimodal. Then we can extend it to a compactly supported unimodal function $\tilde{f} : \mathbb{R}^m \to [0, \infty)$ by defining*

$$\tilde{f}(x) = \begin{cases} f(x); & \|x\| \le 1, \\ (2 - \|x\|) f(\frac{x}{\|x\|}); & \|x\| \in [1, 2], \\ 0; & \|x\| \ge 2. \end{cases}$$

Proof. Suppose $c > 0$. Observe that $B^m \cap \tilde{f}^{-1}[c, \infty) = f^{-1}[c, \infty)$. This is contractible because f is unimodal. Therefore we only need to construct a deformation retraction $H : \tilde{f}^{-1}[c, \infty) \times I \to \tilde{f}^{-1}[c, \infty)$ from $\tilde{f}^{-1}[c, \infty)$ to $B^m \cap \tilde{f}^{-1}[c, \infty)$. Define it as follows:

$$H(x, t) = \begin{cases} x; & \|x\| \le 1, \\ (1 - t)x + t\frac{x}{\|x\|} & \|x\| \ge 1. \end{cases}$$

If well defined, this is clearly a deformation retraction. We only have to verify that it is indeed well defined. This is obvious for $\|x\| \le 1$. For $\|x\| \ge 1$, we have to show that $x \in \tilde{f}^{-1}[c, \infty)$ implies $v := (1-t)x + t\frac{x}{\|x\|} \in \tilde{f}^{-1}[c, \infty)$. It is enough to establish that

$$\tilde{f}(v) \ge \tilde{f}(x).$$

Observe that $\|v\| = (1-t)\|x\| + t$, so by definition of $\tilde{f}$ this reduces to proving that

$$2 - ((1 - t)\|x\| + t) \ge 2 - \|x\|,$$

which is equivalent to $1 \le \|x\|$ and therefore true. □

This has the following straightforward corollary.

Corollary 5.3. *Suppose $f : B^m \to [0, \infty)$ is a continuous function. Associate to it a function $\tilde{f} : \mathbb{R}^m \to [0, \infty)$ as in the above lemma. Then*

$$\mathbf{ucat}(f) = \mathbf{ucat}(\tilde{f}).$$

5.2 Open Questions

In this section we list several questions that we have not been able to resolve, to outline some directions for future research that seem to be promising.

5.2.1 Other Notions of Category for Functions

We propose an alternative notions of category suitable for the study of distributions. In the case of $X = \mathbb{R}$ it coincides with the unimodal category. In higher dimensions, however, we can expect the behavior to be quite different. We believe that the notion is closely related to the notion of unimodal category and may prove useful in establishing various bounds.

If in the condition of unimodality, we use path-connectedness instead of contractibility, we obtain the following definition:

Definition. A continuous function $u : X \to [0, \infty)$ is π_0-*unimodal* if there is a $M > 0$ such that the superlevel sets $u^{-1}[c, \infty)$ are path-connected for $0 < c \le M$ and empty for $c > M$.

This yields the following notion:

Definition. Let $p \in (0, \infty)$. The *π_0-unimodal p-category* $\mathbf{ucat}^p_{\pi_0}(f)$ of a function $f : X \to [0, \infty)$ is the minimum number n of π_0-unimodal functions $u_1, \ldots, u_n : X \to [0, \infty)$ such that pointwise, $f = (\sum_{i=1}^n u_i^p)^{\frac{1}{p}}$. The *$\pi_0$-unimodal ∞-category* is defined analogously, using the ∞-norm instead.

Similarly, a notion of *$\widetilde{\pi}_0$-unimodal p-category* may be defined by using connectedness instead of path-connectedness. (Here $\widetilde{\pi}_0(X)$ denotes the set of connected components of X.) This notion is coarser than the notion of $\mathbf{ucat}^p$, so we expect that it will turn out to be easier to compute. In that case, one particular use for these concepts is that they provide lower bounds for $\mathbf{ucat}^p(f)$. A possible approach to computation of $\mathbf{ucat}^p_{\pi_0}(f)$ might be via Reeb graphs.

Question 2. Can $\mathbf{ucat}^p_{\pi_0}(f)$ be reconstructed from the Reeb graph of f?

Given that in our approach to constructing counterexamples to monotonicity, cycles have been of fundamental importance, whereas the concept of $\mathbf{ucat}^p_{\pi_0}$ allows cycles of all kinds, it seems more plausible to expect that such a concept could be monotone in p.

Question 3. Is $\mathbf{ucat}^p_{\pi_0}(f)$ monotone in p? What about $\mathbf{ucat}^p_{\widetilde{\pi}_0}(f)$?

Note that many other variations of the concept are possible.

Question 4. Are there other notions with interesting properties that can be obtained by replacing contractibility by some other property in the definition of unimodality? Such properties might be for instance: convex, homeomorphic to a ball, etc.

5.2.2 Continuity

Question 5. Suppose we do not assume continuity in the definition of unimodality. Does this change the minimum number of summands a continuous function $f : X \to [0, \infty)$ can be decomposed into?

If X is pathological enough, the answer to this question is positive. For instance, if X is the topologist's sine curve as in Section 2 and $f : X \to [0, \infty)$ is a constant function, for instance $f \equiv 1$, then $\mathbf{ucat}(f) = \infty$. However, f can be decomposed as the sum of the two characteristic functions of the path components of X, which are discontinuous unimodal functions. It is unclear, however, whether this phenomenon can occur for nice spaces such as manifolds or CW complexes.

5.2.3 Morse-Smale Graphs

As mentioned in Section 4.6, the concept of Morse-Smale graph used by Hickok, Villatoro and Wang disregards the local minima of the function completely. One might speculate that this could be the reason why their approach is insufficient to treat the case where the graph contains cycles. This leads us to the following open-ended question:

Question 6. Can a more general and systematic treatment of the unimodal category of Morse functions $f : \mathbb{R}^2 \to [0, \infty)$ be given by taking into account the complete structure of the Morse-Smale complex (in the sense of [32]) of the function f? Morse-Smale complexes also make sense in $\mathbb{R}^3$, see [31]. What can be said in that case?

5.2.4 Graphs

Computation of **ucat** as well as its monotonicity for $\mathbb{R}$ and S^1 follows by using appropriate notions of sweeping. We have also seen that monotonicity fails for most graphs.

Question 7. Which kinds of graphs admit sweeping algorithms? Can these algorithms be used to establish mononicity in the case of trees? What about $S^1 \vee I$?

5.2.5 Cohomological Approach

In the study of Lusternik-Schnirelmann category [24], cohomological methods have been very successful. For instance, one of the basic bounds that is used is

$$\mathrm{cup}_R(X) < \mathbf{cat}(X).$$

A natural question is whether such cohomological methods can be developed for the study of unimodal category. Of course, first an appropriate cohomology theory is needed. If we expect such a theory to be functorial, it must be defined on a suitable category of functions. One natural candidate seems to be the category whose objects are functions $X \to [0, \infty)$ and a morphism between two such functions $X_1 \to [0, \infty)$ and $X_2 \to [0, \infty)$ is an appropriate commutative triangle. To be useful, one property such a cohomology theory should have is that the cohomology ring of a unimodal function is trivial. Furthermore, it should be additive with respect to functions with disjoint supports. The other properties are not immediately clear. One issue that arises is that the notion of unimodal category does not seem to be homotopy invariant in any sense. This is analogous to the notion of **gcat**, which is also not a homotopy invariant.

Question 8. Can such a cohomology theory be constructed?

In both counterexamples on the plane, we studied various paths in the superlevel sets of the functions to establish that the unimodal category cannot be less than three.

Question 9. Is it possible to establish some kind of "path calculus" to automatize this process?

Note that "path calculus" in this case might turn out to be just another term for "cohomology theory". Speculating further, the fact that certain paths intersect might be amenable to a description using cup products. Next, we note that paths also play a crucial role in the approach used by [43]. Namely, their notion of path value is basically the largest value such that the corresponding superlevel set still contains a path between the two points we are interested in.

Question 10. Should the cohomology theory in question be based on the concept of path value?

6 Approximate Nerve Theorem

6.1 ε-Acyclic Covers

Here we introduce the notion of an ε-acyclic cover. For convenience, we find it easier to work with the notion of interleaving and modules rather than persistence diagrams. However, we also include the diagrams for the definitions, for completeness and to help with intuition.

In the classical setting of Theorem 1.2, we assume that each non-empty finite intersection U_I has the homology of a point. In our case, we wish to assume that the homology of each non-empty intersection U_I is ε-close to the homology of a point; specifically, we require that the two homologies are ε-interleaved.

To be more precise, for each $a \in \mathbb{Z}$, we define pt_a to be the $\mathbb{Z}$-filtered simplicial complex consisting of a single point, with the filtration defined by the requirement that $\mathrm{pt}_a^j = \emptyset$ for $j < a$ and $\mathrm{pt}_a^j = \{*\}$ for $j \geq a$.

Definition. A non-empty $\mathbb{Z}$-filtered simplicial complex X is *(persistently) acyclic* if it has the persistent homology of a point, i.e. $\mathsf{H}_*(X) \cong \mathsf{H}_*(\mathrm{pt}_a)$ for some $a \in \mathbb{Z}$. It is *ε-acyclic* if its persistent homology is ε-interleaved with the persistent homology of a point, i.e. $\mathsf{H}_*(X) \overset{\varepsilon}{\sim} \mathsf{H}_*(\mathrm{pt}_a)$ for some $a \in \mathbb{Z}$.

In other words, ε-acyclicity means that $\mathsf{H}_q(X) \overset{\varepsilon}{\sim} 0$ for $q \neq 0$ and $\mathsf{H}_0(X) \overset{\varepsilon}{\sim} t^a\Bbbk[t]$ for some $a \in \mathbb{Z}$. A persistence module M that is ε-close to the trivial module 0, i.e. $M \overset{\varepsilon}{\sim} 0$ is said to be *ε-trivial*. The same understanding applies to persistence diagrams.

The persistence diagram of an acyclic complex consists of only the diagonal in degrees other than 0, and a single point of the form (a, ∞) in degree 0 representing the essential class (corresponding to the first component that appears), while the persistence diagram of an ε-acyclic complex consists of points which are at most ε-away from the diagonal (see Figure 4) and a single point (a, ∞) in degree 0.

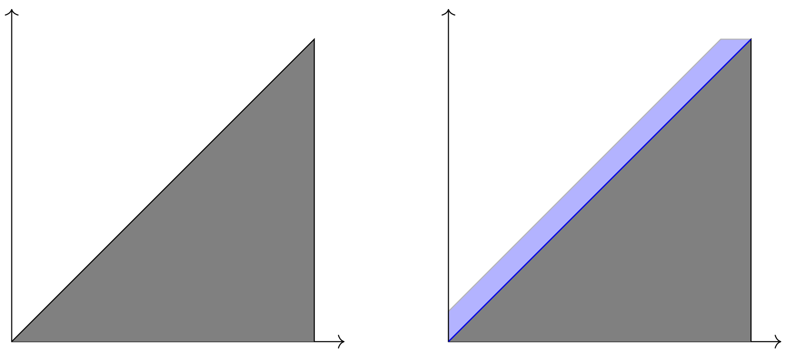


Figure 4: On the left, we have a trivial persistence diagram and on the right an ε-trivial persistence diagram, where points can occur with any multiplicity within the shaded region.

We can now define an ε-acyclic cover.

Definition. Let $\varepsilon \in \mathbb{N}_0$. We say that the filtered cover $\mathcal{U}$ of X is an *ε-acyclic cover* if for each $I \in \mathcal{N}(\mathcal{U})$ there is an $a \in \mathbb{Z}$ such that $\mathsf{H}_*(U_I) \overset{\varepsilon}{\sim} \mathsf{H}_*(\mathrm{pt}_a)$.

Assuming $\mathcal{U}$ is an ε-acyclic cover of X, our aim is to prove that $\mathsf{H}_*(X)$ and $\mathsf{H}_*(\mathcal{N})$ are η-interleaved, where η is bounded above in terms of ε and possibly some other parameter. To help with intuition, we now relate the double complex we use in the spectral sequence with the notion of an ε-acyclic cover. This is best expressed in terms of the E^1 pages. The E^1 page of an acyclic cover and an ε-acyclic cover are shown in Figure 5. For $q > 0$, the elements are 0 or ε-interleaved with 0 respectively. For $q = 0$, each element or non-empty intersection yields one essential class. Since

$$E^1_{p,0} = \bigoplus_{|I|=p+1} \mathsf{H}_0(U_I),$$

the meaning of ε-acyclicity is that each element of the E^1-page is either an essential class corresponding to some $I \in \mathcal{N}(\mathcal{U})$ or ε-trivial.

0	0	0
0	0	0
$E^1_{0,0}$	$E^1_{1,0}$	$E^1_{2,0}$

$E^1_{0,2} \overset{\varepsilon}{\sim} 0$	$E^1_{1,2} \overset{\varepsilon}{\sim} 0$	$E^1_{2,2} \overset{\varepsilon}{\sim} 0$
$E^1_{0,1} \overset{\varepsilon}{\sim} 0$	$E^1_{1,1} \overset{\varepsilon}{\sim} 0$	$E^1_{2,1} \overset{\varepsilon}{\sim} 0$
$E^1_{0,0}$	$E^1_{1,0}$	$E^1_{2,0}$

Figure 5: The E^1 page of an acyclic cover (left) and an ε-acyclic cover (right). In the case of an acyclic cover, the spectral sequence degenerates on the E^2 page because the non-trivial terms are concentrated in the first row. For the ε-acyclic cover, the terms above the first row are only required to be ε-trivial.

The notion of ε-acyclicity need only hold at the level of homology, or equivalently, the interleaving is defined on the E^1-page of the spectral sequence. Consider the corresponding condition at the chain level, i.e. the cover is interleaved with an acyclic cover at the chain level. This implies that the terms on the E^0-page are ε-interleaved. It is straightforward to check that an interleaving on the E^0-page induces an interleaving on the total complex and hence on the persistent homology. This observation combined with the lower bounds presented in Section 6.7 illustrates that ε-acyclicity is a strictly weaker requirement than requiring chain level interleaving as well as that in certain natural cases, chain level interleavings do not exist.

6.1.1 Construction

Here we describe an explicit construction of the filtration for the nerve. Recall the standard construction for the nerve, as given in Section 2.4. In our case, however, the cover elements U_i are filtered by functions f_i, and the space X has a filtration as well, given by $f = \min_{i\in\Lambda} f_i$. Therefore, we must also describe a function g on the nerve. One natural construction is the following. For $I \in \mathcal{N}$, define

$$g(I) = \min\{j \mid U_I^j \neq \emptyset\}. \tag{4}$$

That is, we place a simplex in the filtration, the first time the intersection is not empty. Note there are numerous other constructions, such as taking the average or maximum value which may make more sense in certain cases. It is clear that the sublevel sets of g define a filtration on the nerve.

6.2 Left and Right Interleavings

We extend the usual notion of interleaving to left and right interleaving. This is a refinement of interleaving and certain structural properties will be useful for proving our main result. Readers may skip this section and replace the notions of left and right interleaving in Section 6.5 simply by interleaving, since this is all that is needed for the easy result (Theorem 6.23). The main result in this section is Proposition 6.11. We note that a similar result could be obtained using the techniques in [5] by considering matchings between barcodes. One drawback of using matchings is that it requires the persistence module to be pointwise finite dimensional [5] or at least have an interval decomposition. Our alternative approach has no such requirement, as it applies to modules where no such decomposition exists.

This represents a new viewpoint on interleavings since left and right interleavings are asymmetric leading to several different types of composition (addressed in Proposition 6.11). Though we only use one type of composition in Section 6.5, the others are included for completeness as well as to highlight an interesting phenomenon. We show that for most types of composition of right and left interleavings, the factors are not additive but rather take the maximum of the two component interleavings. Except for one specific case, this holds for more general persistence theories such as persistence over $\mathbb{Z}$ [55] and with appropriate modification, to multidimensional persistence [47]. The exception is the fourth case in Proposition 6.11, which has the additional requirement of having projective dimension 1. Unfortunately, this is precisely the case used in Section 6.5. We conjecture that this is not an artifact of the proof technique but that the statement does not hold for this type of composition in the case of more general persistence modules. If so, we believe this asymmetry is of independent interest. Finally, we show an equivalence between a general interleaving and a sequence of right and left interleavings. This decomposition can be interpreted as "shortening" and "lengthening" bars, but holds even when a barcode does not exist.

In order to prove our result, we must work with approximations of persistence modules efficiently. In particular, we must be able to estimate kernels and cokernels of maps. Intuitively, given a map whose codomain is approximately zero, the kernel should be approximately equal to the domain. The following proposition justifies this intuition. We remind the reader that "morphism" always means 0-morphism, i.e. it is assumed that degrees are preserved.

Proposition 6.1. *Let $g : N \to P$ be a morphism of $\Bbbk[t]$-modules and $P \overset{\varepsilon}{\sim} 0$. Then, $N \overset{2\varepsilon}{\sim} \ker g$. In fact, $\phi : N \to \ker g$ and $\psi : \ker g \to N$ defined by $\phi(n) = t^{2\varepsilon}n$ and $\psi(m) = m$ satisfy $\phi\psi = \mathrm{id}_{2\varepsilon}$ and $\psi\phi = \mathrm{id}_{2\varepsilon}$.*

Proof. The equalities follow directly from the definitions of ϕ and ψ. Therefore, $(\phi, \mathrm{id}_{2\varepsilon}\,\psi)$ is a 2ε-interleaving. We only have to verify that ϕ is well defined. To see this, note that $P \overset{\varepsilon}{\sim} 0$, so multiplication by $t^{2\varepsilon}$ is the zero map on P. This means that for any $n \in N$, we have $t^{2\varepsilon}n \in \ker g$, because $g(t^{2\varepsilon}n) = t^{2\varepsilon}g(n) = 0$. $\square$

The analogous statement for cokernels is also true by the dual argument.

Proposition 6.2. *Let $f : M \to N$ be a morphism of $\Bbbk[t]$-modules and $M \overset{\varepsilon}{\sim} 0$. Then, $N \overset{2\varepsilon}{\sim} \operatorname{coker} f$. In fact, $\eta : N \to \operatorname{coker} f$ and $\theta : \operatorname{coker} f \to N$ defined by $\eta(n) = [n]$ and $\theta([n]) = t^{2\varepsilon} n$ satisfy $\eta\theta = \mathrm{id}_{2\varepsilon}$ and $\theta\eta = \mathrm{id}_{2\varepsilon}$.*

Proof. Again, the two equalities follow directly from the definitions and $(\mathrm{id}_{2\varepsilon} \eta, \theta)$ is a 2ε-interleaving. We only have to verify that θ is well defined. To see this, note that $M \overset{\varepsilon}{\sim} 0$, so $t^{2\varepsilon} M = 0$ and thus $t^{2\varepsilon} \operatorname{im} f = 0$. Now suppose $[n_1] = [n_2]$. It follows that $n_1 - n_2 \in \operatorname{im} f$, so $t^{2\varepsilon}(n_1 - n_2) = 0$, concluding the proof. $\square$

As described in Section 2.5, interleavings define a metric between modules. It turns out, however, that the interleavings arising in these two situations have somewhat special properties, so they deserve separate definitions to distinguish them from ordinary interleavings. We will exploit the properties of such interleavings to obtain tight bounds in the approximate nerve theorem.

Definition. Suppose M and N are $\Bbbk[t]$-modules. We say that M and N are 2ε*-left interleaved* and write $M \overset{2\varepsilon}{\sim}_L N$ if there is a $\Bbbk[t]$-module $P \overset{\varepsilon}{\sim} 0$ and a short exact sequence of the form $0 \to M \to N \to P \to 0$.

Definition. Suppose N and P are $\Bbbk[t]$-modules. We say that N and P are 2ε*-right interleaved* and write $N \overset{2\varepsilon}{\sim}_R P$ if there is a $\Bbbk[t]$-module $M \overset{\varepsilon}{\sim} 0$ and a short exact sequence of the form $0 \to M \to N \to P \to 0$.

Remark 6.1. Note that these definitions are not symmetric, i.e. $M \overset{2\varepsilon}{\sim}_L N$ does not imply $N \overset{2\varepsilon}{\sim}_L M$ and $N \overset{2\varepsilon}{\sim}_R P$ does not imply $P \overset{2\varepsilon}{\sim}_R N$. To see the asymmetry, let M consist of one generator born at $j = 0$ with one relation at $j = a + 2\varepsilon$, and N have one generator born at $j = 0$ with one relation at $j = a$. The kernel of the obvious map is ε-interleaved with 0; hence M and N are ε-left interleaved. In fact, there exists no 0-morphism from $N \to M$, hence their right or left interleaving distance is infinite.

We now prove some properties of left and right interleavings. As mentioned above, left and right interleavings are not symmetric, but they do still satisfy the triangle inequality. Positive definiteness also holds, but this is easy to see by definition – simply take the 0 morphism.

Before continuing, we establish a basic proposition, similar in spirit to Propositions 6.1 and 6.2.

Proposition 6.3. *Suppose we are given an exact sequence*

$$0 \to M \xrightarrow{i} N \xrightarrow{f} P \to 0$$

where $M \overset{\varepsilon_1}{\sim} 0$ and $P \overset{\varepsilon_2}{\sim} 0$. Then $N \overset{\varepsilon_1+\varepsilon_2}{\sim} 0$.

Proof. We need to show that $t^{2(\varepsilon_1+\varepsilon_2)} N = 0$. Let $n \in N$. Note that $f(t^{2\varepsilon_2} n) = t^{2\varepsilon_2} f(n) = 0$, since $t^{2\varepsilon_2} P = 0$, so $m = t^{2\varepsilon_2} n \in \ker f = M$. Therefore $t^{2(\varepsilon_1+\varepsilon_2)} n = t^{2\varepsilon_1} m = 0$, since $t^{2\varepsilon_1} M = 0$. $\square$

Our main motivation for introducing left and right interleavings is to study how the metrics between modules act with respect to composition. We show that

- the approximation factors are additive under composition of the same types of interleaving (i.e. left with left or right with right),
- only the maximum of the approximation factors is relevant when composing different types of interleaving (i.e. left with right or right with left).

We first require some basic structural propositions.

Proposition 6.4. *Suppose $f : M \to N$ and $g : N \to P$ are morphisms of modules. Then there is exact sequence of the form*

$$0 \to \ker f \to \ker gf \to \ker g \to \operatorname{coker} f \to \operatorname{coker} gf \to \operatorname{coker} g \to 0$$

Proof. Note that the diagrams

$$\begin{tikzcd} & 0 \arrow[r] \arrow[d] & M \arrow[r, "\mathrm{id}"] \arrow[d, "f"'] & M \arrow[r] \arrow[d, "gf"] & 0 \\ 0 \arrow[r] & \ker g \arrow[r] & N \arrow[r, "g"'] & P & \end{tikzcd} \qquad \begin{tikzcd} & M \arrow[r, "f"] \arrow[d, "gf"'] & N \arrow[r] \arrow[d, "g"] & \operatorname{coker} f \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & P \arrow[r, "\mathrm{id}"'] & P \arrow[r] & 0 & \end{tikzcd}$$

have exact rows. Applying the snake lemma to each of these diagrams, we obtain exact sequences

$$0 \to \ker f \to \ker gf \to \ker g \to \operatorname{coker} f \to \operatorname{coker} gf$$

and

$$\ker gf \to \ker g \to \operatorname{coker} f \to \operatorname{coker} gf \to \operatorname{coker} g \to 0$$

By construction, the two maps $\ker g \to \operatorname{coker} f$ are actually the same, so splicing the two sequences yields

$$0 \to \ker f \to \ker gf \to \ker g \to \operatorname{coker} f \to \operatorname{coker} gf \to \operatorname{coker} g \to 0$$

as desired. □

This immediately yields two useful corollaries, dual to each other.

Corollary 6.5. *Suppose $f : M \to N$ and $g : N \to P$ are morphisms of modules with g injective. Then the sequence*

$$0 \to \operatorname{coker} f \to \operatorname{coker} gf \to \operatorname{coker} g \to 0$$

is exact.

Corollary 6.6. *Suppose $f : M \to N$ and $g : N \to P$ are morphisms of modules with f surjective. Then the sequence*

$$0 \to \ker f \to \ker gf \to \ker g \to 0$$

is exact.

Using these, we may now prove the triangle inequality for left interleavings.

Proposition 6.7. *Suppose $M \overset{2\varepsilon_1}{\sim}_L N$ and $N \overset{2\varepsilon_2}{\sim}_L P$. Then $M \overset{2(\varepsilon_1+\varepsilon_2)}{\sim}_L P$.*

Proof. The assumptions mean that we have exact sequences

$$0 \to M \xrightarrow{i} N \xrightarrow{f} \operatorname{coker} i \to 0 \qquad \text{and} \qquad 0 \to N \xrightarrow{j} P \xrightarrow{g} \operatorname{coker} j \to 0$$

with $\operatorname{coker} i \overset{\varepsilon_1}{\sim} 0$ and $\operatorname{coker} j \overset{\varepsilon_2}{\sim} 0$. Since j is injective, we have

$$0 \to \operatorname{coker} i \to \operatorname{coker} ji \to \operatorname{coker} j \to 0$$

by Corollary 6.5, so $\operatorname{coker} ji \overset{\varepsilon_1+\varepsilon_2}{\sim} 0$ by Proposition 6.3. Observing that the sequence

$$0 \to M \xrightarrow{ji} P \to \operatorname{coker} ji \to 0$$

is exact completes the proof. □

The same result holds for right interleavings.

Proposition 6.8. *Suppose $M \overset{2\varepsilon_1}{\sim}_R N$ and $N \overset{2\varepsilon_2}{\sim}_R P$. Then $M \overset{2(\varepsilon_1+\varepsilon_2)}{\sim}_R P$.*

Proof. By the assumptions, there are exact sequences

$$0 \to \ker f \xrightarrow{i} M \xrightarrow{f} N \to 0 \qquad \text{and} \qquad 0 \to \ker g \xrightarrow{j} N \xrightarrow{g} P \to 0$$

with $\ker f \overset{\varepsilon_1}{\sim} 0$ and $\ker g \overset{\varepsilon_2}{\sim} 0$. Since f is surjective, we have

$$0 \to \ker f \to \ker gf \to \ker g \to 0$$

by Corollary 6.6, so $\ker gf \overset{\varepsilon_1+\varepsilon_2}{\sim} 0$ by Proposition 6.3. Observing that the sequence

$$0 \to \ker gf \to M \xrightarrow{gf} P \to 0$$

is exact completes the proof. □

The previous results are required to show that the interleavings are in a sense closed under composition, e.g. composing two left interleavings (with a suitable ordering of terms), yields a left interleaving (with an additive approximation factor). Now, we show the more interesting property: most combinations of the different notions of interleavings do not interact, i.e. composition yields the maximum of the two rather than an additive factor. First, we show that if composition is **not** in the natural order as in Propositions 6.8 and 6.7, the interleavings do not yield an additive factor.

Proposition 6.9. *Suppose one of the following two possibilities holds,*

$$M \overset{2\varepsilon}{\sim}_L N \quad \text{and} \quad P \overset{2\varepsilon}{\sim}_L N, \qquad \text{or} \qquad M \overset{2\varepsilon}{\sim}_R N \quad \text{and} \quad P \overset{2\varepsilon}{\sim}_R N,$$

then $M \overset{2\varepsilon}{\sim} P$.

Proof. In the first case, we have the following two exact sequences:

$$0 \to M \xrightarrow{i} N \xrightarrow{f} X \to 0 \qquad \text{and} \qquad 0 \to P \xrightarrow{g} N \xrightarrow{j} Y \to 0$$

Similarly in the second case, we have:

$$0 \to X \xrightarrow{f} M \xrightarrow{i} N \to 0 \qquad \text{and} \qquad 0 \to Y \xrightarrow{j} P \xrightarrow{g} N \to 0$$

In both cases, by assumption $X \overset{\varepsilon}{\sim} 0$ and $Y \overset{\varepsilon}{\sim} 0$. By Propositions 6.1 and 6.2, for both cases there are 2ε-interleavings (ϕ, ψ) of M and N and (η, θ) of P and N. These fit into the following commutative diagram, where the horizontal arrows are ordinary morphisms and all other arrows are 2ε-morphisms.

$$\begin{tikzcd}
M \arrow[r, "i"] \arrow[d, "t^{2\varepsilon}"'] & N \arrow[dl, "\psi"] \arrow[d, "t^{2\varepsilon}"] \arrow[dr, "\theta"] & P \arrow[l, "g"'] \arrow[d, "t^{2\varepsilon}"] \\
M \arrow[r, "i"'] \arrow[d, "t^{2\varepsilon}"'] & N \arrow[dl, "\psi"] \arrow[d, "t^{2\varepsilon}"] \arrow[dr, "\theta"] & P \arrow[l, "g"] \arrow[d, "t^{2\varepsilon}"] \\
M \arrow[r, "i"'] & N & P \arrow[l, "g"]
\end{tikzcd}$$

By inspection of this diagram, we see that $(\theta i, \psi g)$ is a 2ε-interleaving of M and P. □

Proposition 6.10. *Suppose one of the following two possibilities holds,*

$$N \overset{2\varepsilon}{\sim}_L M \quad \textit{and} \quad N \overset{2\varepsilon}{\sim}_L P, \qquad \textit{or} \qquad N \overset{2\varepsilon}{\sim}_R M \quad \textit{and} \quad N \overset{2\varepsilon}{\sim}_R P,$$

then $M \overset{2\varepsilon}{\sim} P$.

Proof. The proof is similar as above. For each case, we get two pairs of exact sequences

$$0 \to N \xrightarrow{i} M \xrightarrow{f} X \to 0 \qquad \text{and} \qquad 0 \to N \xrightarrow{g} P \xrightarrow{j} Y \to 0$$

and

$$0 \to X \xrightarrow{f} N \xrightarrow{i} M \to 0 \qquad \text{and} \qquad 0 \to Y \xrightarrow{j} N \xrightarrow{g} P \to 0$$

with $X \overset{\varepsilon}{\sim} 0$ and $Y \overset{\varepsilon}{\sim} 0$. Again, by Propositions 6.1 and 6.2, we have 2ε-interleavings (ϕ, ψ) of N and M and (η, θ) of N and P, which fit into the following commutative diagram, where the horizontal arrows are ordinary morphisms and all other arrows are 2ε-morphisms.

$$\begin{tikzcd}
M \arrow[d, "t^{2\varepsilon}"'] \arrow[dr, "\psi"'] & N \arrow[l, "i"'] \arrow[r, "g"] \arrow[d, "t^{2\varepsilon}"] & P \arrow[d, "t^{2\varepsilon}"] \arrow[dl, "\theta"] \\
M \arrow[d, "t^{2\varepsilon}"'] \arrow[dr, "\psi"'] & N \arrow[l, "i"] \arrow[r, "g"'] \arrow[d, "t^{2\varepsilon}"] & P \arrow[d, "t^{2\varepsilon}"] \arrow[dl, "\theta"] \\
M & N \arrow[l, "i"] \arrow[r, "g"'] & P
\end{tikzcd}$$

By inspection of this diagram, we see that $(g\psi, i\theta)$ is a 2ε-interleaving of M and P. $\square$

Finally, we show that all other combinations of left and right interleavings do not interact, i.e. composing a 2ε-left interleaving followed by a 2ε-right interleaving still yields a 2ε-interleaving. As the two notions are not symmetric, there are four such possible cases to treat. It turns out that three of the four cases can be handled directly, while the fourth is more involved.

Proposition 6.11. *Suppose one of the following four possibilities holds:*

- $M \stackrel{2\varepsilon}{\sim}_L N$ *and* $N \stackrel{2\varepsilon}{\sim}_R P$*,*
- $N \stackrel{2\varepsilon}{\sim}_L M$ *and* $N \stackrel{2\varepsilon}{\sim}_R P$*,*
- $M \stackrel{2\varepsilon}{\sim}_L N$ *and* $P \stackrel{2\varepsilon}{\sim}_R N$ *or*
- $N \stackrel{2\varepsilon}{\sim}_L M$ *and* $P \stackrel{2\varepsilon}{\sim}_R N$*.*

Then $M \stackrel{2\varepsilon}{\sim} P$*.*

Proof. We treat each possibility separately.

First case. We give a direct argument. There are exact sequences

$$0 \to M \xrightarrow{i} N \xrightarrow{f} X \to 0 \qquad \text{and} \qquad 0 \to Y \xrightarrow{j} N \xrightarrow{g} P \to 0$$

that is $M = \ker f$ and $P = \operatorname{coker} j$ with $X \stackrel{\varepsilon}{\sim} 0$ and $Y \stackrel{\varepsilon}{\sim} 0$. The interleaving maps $\phi : M \to P$ and $\psi : P \to M$ may be defined explicitly by the formulae $\phi(m) = t^{2\varepsilon}g(i(m))$ and $\psi([n]) = t^{2\varepsilon}n$. Here, $[n] = g(n)$ is the class in $\operatorname{coker} j$ represented by $n \in N$. Note that $t^{2\varepsilon}n \in M$, since $X \stackrel{\varepsilon}{\sim} 0$. It is clear that ϕ is well defined. To show that ψ is well defined, observe that $Y \stackrel{\varepsilon}{\sim} 0$ implies $t^{2\varepsilon}Y = 0$ and therefore, $t^{2\varepsilon} \operatorname{im} j = 0$, so if $[n_1] = [n_2]$, we have $t^{2\varepsilon}n_1 = t^{2\varepsilon}n_2$.

We remark that shifting by 2ε is not necessary for the first map to be well defined and is only done to adhere to the definition of interleaving. In fact, without this shifting we already have that $(gi) \circ \psi = \mathrm{id}_{2\varepsilon}$ and $\psi \circ (gi) = \mathrm{id}_{2\varepsilon}$, which is important, as it is used in the proof of Proposition 6.12.

Second case. There are exact sequences

$$0 \to N \xrightarrow{i} M \xrightarrow{f} X \to 0 \qquad \text{and} \qquad 0 \to Y \xrightarrow{j} N \xrightarrow{g} P \to 0$$

with $X \overset{\varepsilon}{\sim} 0$ and $Y \overset{\varepsilon}{\sim} 0$. There are 2ε-interleavings (ϕ, ψ) of N and M and (η, θ) of N and P, which fit into the same commutative diagram as in the proof of Proposition 6.10. Similarly, we conclude that $(g\psi, i\theta)$ is a 2ε-interleaving of M and P.

Third case. There are exact sequences

$$0 \to M \xrightarrow{i} N \xrightarrow{f} X \to 0 \qquad \text{and} \qquad 0 \to Y \xrightarrow{j} P \xrightarrow{g} N \to 0$$

with $X \overset{\varepsilon}{\sim} 0$ and $Y \overset{\varepsilon}{\sim} 0$. There are 2ε-interleavings (ϕ, ψ) of N and M and (η, θ) of N and P, which fit into the same commutative diagram as in the proof of Proposition 6.9. Similarly, we conclude that $(\theta i, \psi g)$ is a 2ε-interleaving of M and P.

Fourth case. There are exact sequences

$$0 \to N \xrightarrow{i} M \xrightarrow{f} X \to 0 \qquad \text{and} \qquad 0 \to Y \xrightarrow{j} P \xrightarrow{g} N \to 0$$

with $X \overset{\varepsilon}{\sim} 0$ and $Y \overset{\varepsilon}{\sim} 0$. To the latter, we associate the following long exact sequence of Ext-modules[17]:

$$\begin{aligned} 0 \longrightarrow \mathrm{Hom}(X, Y) \longrightarrow \mathrm{Hom}(X, P) \longrightarrow \mathrm{Hom}(X, N) \longrightarrow \\ \longrightarrow \mathrm{Ext}(X, Y) \longrightarrow \mathrm{Ext}(X, P) \longrightarrow \mathrm{Ext}(X, N) \longrightarrow 0 \end{aligned}$$

Note that all higher Ext-modules are 0. To see this, recall that the projective dimension $\mathrm{projdim}(X)$ of a $\Bbbk[t]$-module X is the smallest $n \in \mathbb{N}_0$ such that $\mathrm{Ext}^{n+1}(X, M)$ vanishes for all $\Bbbk[t]$-modules M (see [57, Proposition 8.6]). It is known that any module over a principal ideal domain has projective dimension at most 1, so in particular $\mathrm{Ext}^2(X, Y) = 0$, as desired. (There is a slight subtlety here that the number $\mathrm{projdim}(X)$ could in principle depend on whether X is regarded as a $\Bbbk[t]$-module or a $\Bbbk[t]_{(\mathrm{NGr})}$-module. That this is not the case follows from [52, Corollary 3.3.7].)

In particular, $\mathrm{Ext}(X, P) \to \mathrm{Ext}(X, N)$ is an epimorphism. Using the classical interpretation of elements of Ext-modules as (equivalence classes of) extensions of modules and maps between them as morphisms of such extensions implies that there is a map of extensions

$$\begin{tikzcd} 0 \arrow[r] & P \arrow[r] \arrow[d, "g"'] & Q \arrow[r] \arrow[d, "h"'] & X \arrow[r] \arrow[d, "\mathrm{id}"'] & 0 \\ 0 \arrow[r] & N \arrow[r] & M \arrow[r] & X \arrow[r] & 0 \end{tikzcd}$$

Now, using the snake lemma on this diagram, we see that the sequence

$$0 \to \ker g \to \ker h \to \ker \mathrm{id} \to \operatorname{coker} g \to \operatorname{coker} h \to \operatorname{coker} \mathrm{id} \to 0$$

is exact. Since $\ker g = Y$ and $\operatorname{coker} g = \ker \mathrm{id} = \operatorname{coker} \mathrm{id} = 0$, the sequence

$$0 \to Y \to Q \to M \to 0$$

is exact. Therefore $Q \overset{2\varepsilon}{\sim}_R M$ and $P \overset{2\varepsilon}{\sim}_L Q$, so the fourth case reduces to the first case. □

[17]Note that Hom-modules consist of morphisms of $\Bbbk[t]$-modules. These are degree-preserving. The appropriate notion of Ext-module needs to reflect this. In particular, the maps used in the relevant projective resolutions must also be degree-preserving.

Intuitively, the notion of left and right interleavings corresponds to the notion of shortening (respectively lengthening) bars by changing birth and death times. This was described in [5] using matchings. The main advantage of using short exact sequences is that the independence between modifying birth and death times can be captured without a decomposition existing. It also gives an alternative algebraic characterization of when this holds, namely that the projective dimension is one. To make this connection concrete, we prove that every interleaving admits a decomposition into left and right interleavings. We also show the converse, giving a characterization of an interleaving given a decomposition. We first require one additional definition.

Definition. If S is a persistence module, there is an *ε-shifted* module $S(\varepsilon)$ which is a reparameterization of S by

$$S^\alpha(\varepsilon) = S^{\alpha+\varepsilon}.$$

Proposition 6.12. *There exists an interleaving $M \overset{2\varepsilon}{\sim} S$ if and only if $\exists N, P, Q$ such that*

$$M \overset{2\varepsilon}{\sim}_R N, \qquad N \overset{2\varepsilon}{\sim}_L P,$$
$$Q \overset{2\varepsilon}{\sim}_L P, \qquad S \overset{2\varepsilon}{\sim}_R Q.$$

Proof. We first show if M is 2ε-interleaved with S then N, P, and Q exist. First, we construct an interpolation. Let Z be such that $M \overset{\varepsilon}{\sim} Z$ with the interleaving maps (ξ, η) and $S \overset{\varepsilon}{\sim} Z$ with the interleaving maps (ζ, ν). For the construction of the interpolated module, see [16]. Now we set

$$P = Z(\varepsilon)$$

the shifted version of Z. Then let

$$f : M \to Z(\varepsilon) \qquad \text{and} \qquad g : S \to Z(\varepsilon)$$

where f and g are the interleaving maps ξ and ζ respectively. Note that as morphisms into $Z(\varepsilon)$, f and g are 0-morphisms, that is, they are ungraded morphisms. Setting

$$N = \operatorname{im} f \qquad \text{and} \qquad Q = \operatorname{im} g$$

we have the following set of short exact sequences:

$$0 \longrightarrow \ker f \longrightarrow M \longrightarrow \operatorname{im} f \longrightarrow 0$$

$$0 \longrightarrow \operatorname{im} f \longrightarrow Z(\varepsilon) \longrightarrow \operatorname{coker} f \longrightarrow 0$$

$$0 \longrightarrow \operatorname{im} g \longrightarrow Z(\varepsilon) \longrightarrow \operatorname{coker} g \longrightarrow 0$$

$$0 \longrightarrow \ker g \longrightarrow S \longrightarrow \operatorname{im} g \longrightarrow 0$$

We can directly verify that $\ker f$, $\operatorname{coker} f$, $\operatorname{coker} g$ and $\ker g$ are ε-interleaved with 0, hence completing the proof. In the other direction, assume N, P, and Q exist. This gives rise to the following short exact sequences, where the ε denote (possibly distinct) modules ε-interleaved with 0.

$$
\begin{tikzcd}
0 \arrow[r] & \varepsilon \arrow[r] & M \arrow[r, "i"] \arrow[d, "j\circ i"] & N \arrow[r] & 0 \\
0 \arrow[r] & N \arrow[r, "j"] & P \arrow[r] \arrow[d, leftrightarrow, "\cong"] & \varepsilon \arrow[r] & 0 \\
0 \arrow[r] & Q \arrow[r, "k"] & P \arrow[r] & \varepsilon \arrow[r] & 0 \\
0 \arrow[r] & \varepsilon \arrow[r] & S \arrow[r, "\ell"] \arrow[u, "k\circ\ell"'] & Q \arrow[r] & 0
\end{tikzcd}
$$

If we consider the composition of the first two exact sequences and the last two, we are in the fourth case of Proposition 6.11. This implies that there exists a 2ε-morphism $\varphi : P \to M$ and $\psi : P \to S$, each of which is the component of an appropriate interleaving. Hence, we can consider the following commutative diagram:

$$
\begin{tikzcd}
M \arrow[r, "j\circ i"] \arrow[d, "t^{2\varepsilon}"'] & P \arrow[dl, "\varphi"] \arrow[d, "t^{2\varepsilon}"] \arrow[dr, "\psi"] & S \arrow[l, "k\circ\ell"'] \arrow[d, "t^{2\varepsilon}"] \\
M \arrow[r, "j\circ i"'] \arrow[d, "t^{2\varepsilon}"'] & P \arrow[dl, "\varphi"] \arrow[d, "t^{2\varepsilon}"] \arrow[dr, "\psi"] & S \arrow[l, "k\circ\ell"] \arrow[d, "t^{2\varepsilon}"] \\
M \arrow[r, "j\circ i"'] & P & S \arrow[l, "k\circ\ell"]
\end{tikzcd}
$$

This diagram commutes, since $t^{2\varepsilon} = \varphi \circ j \circ i$ and $t^{2\varepsilon} = \psi \circ k \circ \ell$ by the remark in the proof of the first case of Proposition 6.11. Hence, we have the required 2ε-interleaving given by $(\psi \circ j \circ i, \varphi \circ k \circ \ell)$. $\square$

This decomposition helps give an interpretation to right and left interleaving in the case where the barcode exists. The first short exact sequence is a right interleaving which shortens bars by changing the death time of a bar; the second sequence is a left interleaving, which lengthens the bars by changing the birth time of a bar; the third sequence is again a left interleaving which now shortens the bars by changing the birth time; finally the last sequence is a right interleaving which lengths bars by changing the death time. This interpretation of shortening and lengthening bars leads us to the following conjecture.

Conjecture. For each $i = 1, 2, 3, 4$, let $X \overset{2\varepsilon}{\sim}_i Y$ denote one of the following four types of interleavings, so that each of these interleaving types occurs precisely once:

$$X \overset{2\varepsilon}{\sim}_R Y, \qquad Y \overset{2\varepsilon}{\sim}_R X, \qquad X \overset{2\varepsilon}{\sim}_L Y \qquad \text{or} \qquad Y \overset{2\varepsilon}{\sim}_L X.$$

Then, there exists an interleaving $M \overset{2\varepsilon}{\sim} S$ if and only if $\exists N, P, Q$ such that

$$M \overset{2\varepsilon}{\sim}_1 N \overset{2\varepsilon}{\sim}_2 P \overset{2\varepsilon}{\sim}_3 Q \overset{2\varepsilon}{\sim}_4 S.$$

Essentially, we should be able to shorten and lengthen bars (when these notions are well defined) in any order, rather than just the order we list in Proposition 6.12. Note since we use the fourth case of Proposition 6.11, the results do not hold for general persistence modules, but rather require projective dimension one. We believe this approach may help highlight what results hold for more general modules.

Remark 6.2. We note that it is likely that Example 6.8 in [5] may be modified to provide a counter-example to Proposition 6.12 for multi-parameter (or multidimensional) persistence modules. This alternate characterization of the decomposition of interleavings may provide a different perspective on how interleavings can interact in the case of more general modules.

We conclude this section with the analysis of a special case: when a module is ε-interleaved with the trivial module. This was studied extensively in [59] for more complicated modules. Unfortunately, the results were not applicable directly; however, the connection of left and right interleavings with [59] remains open. We conclude with a lemma that further illustrates that interleaving with the trivial module has special structure.

Lemma 6.13. *If a module is ε-interleaved with 0, then it is both right and left 2ε-interleaved with 0.*

Proof. To prove the result, we consider the following short exact sequences illustrating left and right interleaving respectively:

$$0 \to 0 \xrightarrow{t^{2\varepsilon}} A \xrightarrow{\cong} \mathrm{coker}(t^{2\varepsilon}) \to 0$$

$$0 \to \ker(t^{2\varepsilon}) \xrightarrow{\cong} A \xrightarrow{t^{2\varepsilon}} 0 \to 0$$

It follows directly that $\ker(t^{2\varepsilon})$ and $\mathrm{coker}(t^{2\varepsilon})$ are ε-interleaved with 0, fufilling the definitions of right and left interleaving and hence A is both 2ε-left and right interleaved with 0. □

6.3 Approximating Higher Pages

The main work in the proof is to track the approximation factors through the spectral sequence. Let E be the Mayer-Vietoris spectral sequence associated to $(X, \mathcal{U})$. In the acyclic case, as in the case for many spectral sequences, the sequence collapses on the second page. Furthermore, the special structure of the second page, i.e. $E^2_{p,q} = 0$ for $q > 0$, eliminates the possibility of extension problems. This allows for the homology of the space to be read off from the bottom row, and hence corresponding with the homology of the nerve (Theorem 1.2). The extension problems which arise in our setting are further discussed in Section 6.4.

Therefore, a natural first step is to compare the bottom row of the E^2 page with the homology of the nerve.

Proposition 6.14. *If $\mathcal{U}$ is an ε-acyclic cover of X, $(E^1_{*,0}, d^1_{*,0})$ and $(\mathsf{C}_*(\mathcal{N}), \partial)$ are 2ε-interleaved as chain complexes.*

Proof. The interleaving maps $\phi_p : E^1_{p,0} \to \mathsf{C}_p(\mathcal{N})$ and $\psi_p : \mathsf{C}_p(\mathcal{N}) \to E^1_{p,0}$ are defined by the formulae

$$\phi_p([v], I) = t^{\deg(v)-\deg(I)+2\varepsilon} I \qquad \text{and} \quad \psi_p(I) = t^{2\varepsilon}([v_I], I),$$

where $v_I \in V$ is any vertex such that $\deg v_I = \deg I$. Note that the definition of ψ requires a choice of v_I, but since $\mathcal{U}$ is an ε-acyclic cover, $t^{2\varepsilon}[v_I]$ is independent of this choice, so ψ is well defined.

A straightforward calculation now shows that (ϕ, ψ) is a 2ε-interleaving and that ϕ and ψ commute with the differentials ∂ and d^1. (For the latter note that the differentials only really act on the information coming from the nerve, i.e. I, while the interleaving maps preserve this information.) □

Using Lemma 2.17, this immediately yields:

Corollary 6.15. *If $\mathcal{U}$ is an ε-acyclic cover of X, then $E^2_{*,0}$ and $\mathsf{H}_*(\mathcal{N})$ are 2ε-interleaved as graded modules.*

Note that setting $\varepsilon = 0$ recovers Theorem 1.2. We now observe that in the nerve construction, the dimension of the nerve is $D = \dim \mathcal{N}$, all $(D+1)$-intersections are empty and hence 0. In this case, Corollary 6.15 can be sharpened:

Remark 6.3. For $d \geq D+1$, $E^2_{d,0}$ and $\mathsf{H}_d(\mathcal{N})$ are both trivial and hence isomorphic.

The next step is to establish a relation between E^2 and E^∞.

Proposition 6.16. *If $\mathcal{U}$ is an ε-acyclic cover of X, then $E^r_{p,q} \overset{\varepsilon}{\sim} 0$ holds for all $p \in \mathbb{Z}$ and $q \neq 0$ and all $r \geq 1$.*

Proof. Using

$$E^1_{p,q} = \bigoplus_{|I|=p+1} \mathsf{H}_q(U_I)$$

(see Equation (2)) and the definition of ε-acyclic cover, we obtain the claim for $E^1_{p,q}$ with $q > 0$. Since all the $E^r_{p,q}$ with $r > 1$ are subquotients of $E^1_{p,q}$, the claim is now a direct consequence of Corollary 2.8. □

We can now prove the following proposition:

Proposition 6.17. *If $\mathcal{U}$ is an ε-acyclic cover of X, then $E^{r+1}_{*,0} \overset{2\varepsilon}{\sim}_L E^r_{*,0}$ as graded modules for all $r \geq 2$.*

Proof. Notice that $E^{r+1}_{p,0} = \ker d^r_{p,0}$, since the domain of $d^r_{p+r,-r+1}$ is 0. We conclude that $\ker d^r_{p,0} \overset{2\varepsilon}{\sim}_L E^r_{p,0}$ is true by the definition of left interleaving, since $E^r_{p-r,r-1} \overset{\varepsilon}{\sim} 0$ by Proposition 6.16. □

If the spectral sequence collapses after finitely many steps, E^2 may already give a good approximation to E^∞. This happens, for instance, if dimension of the nerve or underlying space are finite. We define $D := \dim \mathcal{N}$, the maximum dimension of any simplex in $\mathcal{N}$. Since simplices in $\mathcal{N}$ correspond to non-empty intersections of cover elements, D is also the smallest number such that any intersection of more than $D+1$ distinct cover elements is empty. Note that in the following the number of pages required until the spectral sequence collapses may be bounded by the dimension of the underlying space.

Theorem 6.18. *If $\mathcal{U}$ is an ε-acyclic cover of X and $0 < D < \infty$, then $E^\infty_{*,0} \overset{2(D-1)\varepsilon}{\sim_L} E^2_{*,0}$ as graded modules. For $D = 0, 1$ we have $E^\infty_{*,0} \cong E^2_{*,0}$.*

Proof. Since the intersections of more than $D + 1$ cover elements are necessarily empty, $E^r_{p,q} = 0$ holds for all $p > D$. Therefore, for $r > D$, we have $d^r = 0$, since either the domain or codomain of each $d^r_{p,q}$ is zero. This immediately implies that the spectral sequence has collapsed by the $(D+1)$-th page, i.e. $E^{D+1} = E^{D+2} = \ldots$. This concludes the proof for $D = 0$. For $D > 0$, using Proposition 6.17, this shows that

$$E^\infty_{*,0} = E^{D+1}_{*,0} \overset{2\varepsilon}{\sim_L} E^D_{*,0} \overset{2\varepsilon}{\sim_L} \ldots \overset{2\varepsilon}{\sim_L} E^3_{*,0} \overset{2\varepsilon}{\sim_L} E^2_{*,0}$$

and therefore $E^\infty_{*,0} \overset{2(D-1)\varepsilon}{\sim_L} E^2_{*,0}$ by the triangle inequality for left interleavings. □

Remark 6.4. For dimension $d > D$, since all the modules are trivial, it follows that $E^\infty_{d,0} \cong \mathsf{H}_d(\mathcal{N})$.

A similar argument shows a weaker property without any assumptions on the dimension of the nerve.

Theorem 6.19. *If $\mathcal{U}$ is an ε-acyclic cover of X and $n > 0$, we have $E^\infty_{n,0} \overset{2(n-1)\varepsilon}{\sim_L} E^2_{n,0}$. For $n = 0$ we have $E^\infty_{n,0} \cong E^2_{n,0}$.*

Proof. Observe that for $r > n > 0$, we have $d^r_{n,0} = 0$ and $d^r_{n+r,-r+1} = 0$, since $E^r_{n-r,r-1}$ and $E^r_{n+r,-r+1}$ are zero. Therefore, $E^{n+1}_{n,0} = E^{n+2}_{n,0} = \ldots$. Combined with Proposition 6.17, this shows that

$$E^\infty_{n,0} = E^{n+1}_{n,0} \overset{2\varepsilon}{\sim_L} E^n_{n,0} \overset{2\varepsilon}{\sim_L} \ldots \overset{2\varepsilon}{\sim_L} E^3_{n,0} \overset{2\varepsilon}{\sim_L} E^2_{n,0}$$

and therefore $E^\infty_{n,0} \overset{2(n-1)\varepsilon}{\sim_L} E^2_{n,0}$ by the triangle inequality for left interleavings.

The case $n = 0$ holds since for $r > 1$ all differentials to and from $E^r_{0,0}$ are zero. □

6.4 From E^∞ to Homology

If there were no extension problems, the direct sum of the antidiagonals on the E^∞ page of the spectral sequence would be isomorphic to the homology of the space, and completing the proof would be straightforward. However, when dealing with persistence modules, we **do** have to worry about extensions. As noted before, in the acyclic case, $E^2_{p,q} = 0$ for all $q > 0$, so the only possible extension is the trivial one. If we replace the ε-modules below by 0, we see that each step becomes an isomorphism. We now show how to infer an approximate nerve theorem from these results. For technical reasons, we have to distinguish between several cases depending on the dimension of the nerve and beyond that dimension.

Proposition 6.20. *If $\mathcal{U}$ is an ε-acyclic cover of X and $D < \infty$, $\mathsf{H}_d(X) \overset{2d\varepsilon}{\sim_R} E^\infty_{d,0}$ holds for $0 \le d \le D$.*

Proof. By Theorem 2.16, we already know that E converges to $\mathsf{H}_*(X)$. Explicitly, this means that a filtration $(\mathsf{H}_*(X)^p)_{p\in\mathbb{Z}}$ is defined on $\mathsf{H}_*(X)$ such that

$$E^\infty_{p,q} \cong \frac{\mathsf{H}_{p+q}(X)^p}{\mathsf{H}_{p+q}(X)^{p-1}}.$$

In the process of reconstructing $\mathsf{H}_n(X) = \mathsf{H}_n(X)^n$ from $E^\infty_{p,q}$ with $p+q=n$, we therefore encounter a series of extension problems. The effect of each of these extension problems in our case, however, is simply to add an error of 2ε to our approximation of $\mathsf{H}_n(X)$. Specifically, we have

$$\frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{p-1}} \overset{2\varepsilon}{\sim}_R \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^p} \tag{5}$$

for each $p \neq n$ (equivalently $q \neq 0$). To see this, observe that the sequence

$$0 \to \frac{\mathsf{H}_n(X)^p}{\mathsf{H}_n(X)^{p-1}} \to \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{p-1}} \to \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^p} \to 0 \tag{6}$$

is exact and

$$\frac{\mathsf{H}_n(X)^p}{\mathsf{H}_n(X)^{p-1}} = E^\infty_{p,q} \overset{\varepsilon}{\sim} 0 \tag{7}$$

holds by Proposition 6.16 if $q \neq 0$. Since the left most term is ε-interleaved with 0, (5) then follows by the definition of right interleaving. The claim now follows inductively. For $0 \leq n \leq D$, we have

$$\mathsf{H}_n(X) \cong \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{-1}} \overset{2\varepsilon}{\sim}_R \ldots \overset{2\varepsilon}{\sim}_R \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{n-2}} \overset{2\varepsilon}{\sim}_R \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{n-1}} \cong E^\infty_{n,0},$$

Since $n \leq D$, there are at most D 2ε-right interleavings, proving the result by Proposition 6.8. $\square$

Note that in the case where $\varepsilon = 0$, the extensions become trivial as the maps in the filtration are isomorphisms by exactness. The second case is for $\mathsf{H}_d(X)$ when $d > D$.

Proposition 6.21. *If $\mathcal{U}$ is an ε-acyclic cover of X and $D < \infty$, $\mathsf{H}_d(X) \overset{2(D+1)\varepsilon}{\sim}_R E^\infty_{d,0}$ holds for $d > D$.*

Proof. For $n > D$, we use the fact that $E^\infty_{p,q} \cong 0$ holds for all $p > D$ (equivalently $q < n - D$). The short exact sequence (6) for these p implies that

$$\frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^D} \cong \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{D+1}} \cong \ldots \cong \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{n-1}}.$$

Using (6) and (7) we obtain the following sequence of right interleavings

$$\mathsf{H}_n(X) \cong \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{-1}} \overset{2\varepsilon}{\sim}_R \ldots \overset{2\varepsilon}{\sim}_R \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{D-1}} \overset{2\varepsilon}{\sim}_R \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^D} \cong \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{n-1}} \cong E^\infty_{n,0}.$$

By counting that there are $(D+1)$ 2ε-right interleavings, we obtain the result. $\square$

For completeness, we add one further case: where the dimension of the space is lower than the dimension of the nerve. For example, the nerve of a cubical cover of k-dimensional Euclidean space has $D = 2^k$. We could redo much of our work for cubical complexes; however, the following result shows this is unnecessary. Let $\Delta := \dim X$. For the case, $D > \Delta$ we show the approximation constant depends on Δ instead of D.

Proposition 6.22. *If $\mathcal{U}$ is an ε-acyclic cover of X and $\Delta < \infty$, $\mathsf{H}_d(X) \overset{2\Delta\varepsilon}{\sim}_R E^{\infty}_{d,0}$ holds for all d.*

Proof. The proof follows as in the above propositions. However, since Δ is the dimension of the space

$$\frac{\mathsf{H}_n(X)^p}{\mathsf{H}_n(X)^{p-1}} = 0, \qquad p \leq n - \Delta - 1,$$

Therefore using (6) and (7) we obtain the sequence

$$\mathsf{H}_n(X) \cong \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{-1}} \cong \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{n-\Delta-1}} \overset{2\varepsilon}{\sim}_R \dots \overset{2\varepsilon}{\sim}_R \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{n-2}} \overset{2\varepsilon}{\sim}_R \frac{\mathsf{H}_n(X)^n}{\mathsf{H}_n(X)^{n-1}} \cong E^{\infty}_{n,0}.$$

There are Δ 2ε-right interleavings, proving the result. $\square$

6.5 Main Theorems

Here we connect the results of the previous two sections to obtain our main result. The idea is to consider the chain of approximations. Unfortunately there are several cases we have to consider depending on the dimension of the nerve and the space. The basic idea, however, is to consider the relationships in the sequence

$$\mathsf{H}_*(\mathcal{N}) \sim E^2_{*,0} \sim E^{\infty}_{*,0} \sim \mathsf{H}_*(X).$$

where we recall that X is a filtered simplicial complexes and $\mathcal{N}$ is another filtered complex given by the nerve of a cover on X. Before stating the result with the tight constant, we consider an easy case of the result which does not use the specific properties of left and right interleavings. Recall that a 2ε-left or right interleaving implies a 2ε-interleaving.

Theorem 6.23. *If $\mathcal{U}$ is an ε-acyclic cover of X and $D < \infty$, we have $\mathsf{H}_*(X) \overset{(4D+2)\varepsilon}{\sim} \mathsf{H}_*(\mathcal{N})$.*

Proof. Assuming $D > 0$ and composing interleavings with constants, we obtain

$$\mathsf{H}_*(\mathcal{N}) \overset{2\varepsilon}{\sim} E^2_{*,0} \overset{2(D-1)\varepsilon}{\sim} E^{\infty}_{*,0} \overset{2(D+1)\varepsilon}{\sim} \mathsf{H}_*(X).$$

The first interleaving is from Corollary 6.15 and the second follows from Theorem 6.18. Finally the last interleaving follows from Proposition 6.20 for $0 \leq d \leq D$ and Proposition 6.21 for $d > D$. Adding the terms, we obtain the result. The case $D = 0$ is straightforward. $\square$

Theorem 6.24. *Let $Q = \min(D, \Delta)$. If $\mathcal{U}$ is an ε-acyclic cover of X and $Q < \infty$, we have $\mathsf{H}_*(X) \overset{2(Q+1)\varepsilon}{\sim} \mathsf{H}_*(\mathcal{N})$.*

Proof. Observe that in the proof of the previous theorem, for $0 \leq d \leq D$ and $\Delta \geq D$, the precise relationship is

$$\mathsf{H}_d(X) \overset{2D\varepsilon}{\sim}_R E^{\infty}_{d,0} \overset{2D\varepsilon}{\sim}_L E^2_{d,0} \overset{2\varepsilon}{\sim} \mathsf{H}_d(\mathcal{N}).$$

The first interleaving follows from Proposition 6.20, the second from Theorem 6.18 and the last one from Corollary 6.15. However, the interleaving obtained from Theorem 6.18 is a left interleaving, whereas the one from Proposition 6.20 is a right interleaving. By Proposition 6.11, together these imply

$$\mathsf{H}_d(\mathcal{N}) \overset{2\varepsilon}{\sim} E^2_{d,0} \overset{2D\varepsilon}{\sim} \mathsf{H}_d(X).$$

For $d > D$ and $\Delta \geq D$,

$$\mathsf{H}_d(X) \overset{2(D+1)\varepsilon}{\sim_R} E^\infty_{d,0} \cong \mathsf{H}_d(\mathcal{N}),$$

where the isomorphism follows from Remark 6.4 and the interleaving follows from Proposition 6.21. As a right interleaving implies interleaving, this proves this case. Finally, for $\Delta < D$, we note the spectral sequence stabilizes after $\Delta + 1$ steps; therefore, the relationship is

$$\mathsf{H}_*(X) \overset{2\Delta\varepsilon}{\sim_R} E^\infty_{*,0} \overset{2(\Delta-1)\varepsilon}{\sim_L} E^2_{*,0} \overset{2\varepsilon}{\sim} \mathsf{H}_*(\mathcal{N}),$$

where the right interleaving is due to Proposition 6.22. Again noting that right and left interleavings do not interact, we obtain

$$\mathsf{H}_*(\mathcal{N}) \overset{2\varepsilon}{\sim} E^2_{*,0} \overset{2\Delta\varepsilon}{\sim} \mathsf{H}_*(X).$$

We can now directly verify that the approximation is bounded by $2(\min(D, \Delta)+1)\varepsilon$, concluding the proof. □

Using an analogous argument without any assumptions on D or Δ, we obtain

Theorem 6.25. *If $\mathcal{U}$ is an ε-acyclic cover of X, $\mathsf{H}_n(X) \overset{2(n+1)\varepsilon}{\sim} \mathsf{H}_n(\mathcal{N})$.*

Proof. The key observation is that since we have a first quadrant spectral sequence, $E^{n+1}_{p,q} \cong E^\infty_{p,q}$ for $0 \leq p+q \leq n$. Applying Propositions 6.20 and 6.21, yields

$$\mathsf{H}_n(X) \overset{2n\varepsilon}{\sim_R} E^\infty_{n,0} \cong E^{n+2}_{n,0} \overset{2n\varepsilon}{\sim_L} E^2_{n,0} \overset{2\varepsilon}{\sim} \mathsf{H}_n(\mathcal{N}).$$

As in Theorem 6.24, combining the interleavings yields the result. □

6.6 Applications

We prove a simple result of a possible application of our main result. While the result is not new, the proof is an immediate consequence of our result. There are many related approximation results in the literature (e.g., [9, 20, 54, 17, 22, 60]). We do not provide a comprehensive account of these approximation results but provide two example applications to illustrate the approximate nerve theorem.

Throughout this section, we use the function g on the nerve which was defined in Section 6.1, which inserts a simplex into nerve as soon as the corresponding intersection is non-empty.

Theorem 6.26. *Given a c-Lipschitz function f on a D-dimensional manifold X embedded in Euclidean space with positive reach ρ, given an ε-sample of the space with $\varepsilon < \rho$, consider the cover of balls of radius ε centered at the sample points. Let $h : \mathcal{N} \to \mathbb{R}$ be the function defined by the formula*

$$h(I) = \max_{i \in I} f(x_i)$$

where x_i is the corresponding sample point. Then,

$$d_I(\mathsf{H}_*(X, f), \mathsf{H}_*(\mathcal{N}, h)) \leq (4D + 3)c\varepsilon.$$

Proof. Ignoring the function for the time being, since we have an ε-sample, balls of radius ε centered at the sample points form a cover of the manifold X. Since the reach is larger than ε, it follows that these balls form a good cover of X. Now, we show that for a c-Lipschitz function, this is a $2c\varepsilon$-acyclic cover. Using the construction in Section 6.1, we note that the maximum value attained in any cover element is

$$f(U_I) \leq g(I) + 2c\varepsilon.$$

where g is defined in Equation 4. Hence, after $2c\varepsilon$, the sublevel set fills the entire cover element, so it is a $2c\varepsilon$-good cover. In this construction, we also note that if the elements are $2c\varepsilon$-interleaved with the trivial diagram so are all intersections. This gives an approximation of $(2D + 1)2c\varepsilon$. Finally, we note that since the cover elements are bounded in size by ε and the definition of g, $|h - g| \leq c\varepsilon$. Adding these constants together yields an interleaving which implies the result. □

The bound above is not meant to be tight as a slightly longer argument would remove a $c\varepsilon$, and many similar results have been proven. Importantly it illustrates that we can approximate the sublevel set persistence with a single filtration rather than an image between two cover elements as in [9] without requiring any one sublevel set to have a good cover. We do note that in this instance, it is possible but cumbersome to construct an explicit functional interleaving. An almost identical result can also be stated replacing reach with other measures such as convexity radius, homotopy feature size, etc.

We also wish to derive an approximate nerve theorem for ε-acyclic covers of triangulable spaces directly from the one for simplicial complexes. However, covers of triangulable spaces by triangulable subsets are too general for this, as their triangulations may not interact well. To circumvent this issue, we introduce the following technical notion.

Definition. Suppose $\overline{\mathcal{V}} = (V_i)_{i \in \Lambda}$ is a cover of a locally compact triangulable space Y. We say $\overline{\mathcal{V}}$ is a *triangulable cover* if there exists some triangulation $(\widetilde{X}, h)$ of Y such that each cover element V_i is the image of a subcomplex of $\widetilde{X}$ under h.

Such covers are very common in practical applications. The notion of ε-acyclic cover is analogous to the one for simplicial complexes; however, continuous persistence modules must be used. A triangulable cover by itself is not filtered, but we will impose a filtration on it by specifying a function on each cover element. We do *not* require that the triangulable cover condition holds at the intermediate stages of the filtration.

First we prove a preliminary lemma to establish that a filtered cover of a triangulable space can be approximated arbitrarily well by one whose filtration is given by piecewise linear functions.

Lemma 6.27. *Let Y be a locally compact triangulable space and $\overline{\mathcal{V}} = (V_i)_{i\in\Lambda}$ a locally finite cover of Y. Suppose $\overline{\mathcal{V}}$ is triangulable, with triangulation $(\widetilde{X}, h)$. Let $\varepsilon > 0$. Given continuous functions $f : Y \to \mathbb{R}$ and $f_i : V_i \to \mathbb{R}$, $i \in \Lambda$, there exists a subdivision X of $\widetilde{X}$ such that for each simplex σ of X we have*

$$\max_{x\in|\sigma|} f(h(x)) - \min_{x\in|\sigma|} f(h(x)) < \varepsilon$$

and

$$\max_{x\in|\sigma|} f_i(h(x)) - \min_{x\in|\sigma|} f_i(h(x)) < \varepsilon \quad \textit{for all } i \in \Lambda.$$

Proof. Let $(\widetilde{X}, h)$ be a triangulation of Y. By local compactness, $\widetilde{X}$ is locally finite, so $|\widetilde{X}|$ is metrizable. Choose a metric d on $|\widetilde{X}|$. Since fh is uniformly continuous on each simplex $\tilde{\sigma}$, there exists a $\delta(\tilde{\sigma}) > 0$ such that $d(x_1, x_2) < \delta(\tilde{\sigma})$ implies $|f(h(x_1)) - f(h(x_2))| < \varepsilon$ for all $x_1, x_2 \in |\tilde{\sigma}|$. Since f_ih, $i \in \Lambda$, is uniformly continuous on each simplex $\tilde{\sigma}$, there exists a $\delta_i(\tilde{\sigma}) > 0$ such that $d(x_1, x_2) < \delta_i(\tilde{\sigma})$ implies $|f_i(h(x_1)) - f_i(h(x_2))| < \varepsilon$ for all $x_1, x_2 \in |\tilde{\sigma}|$. Since the cover $\overline{\mathcal{V}}$ is locally finite, each simplex $\tilde{\sigma} \in \widetilde{X}$ is only contained in finitely many cover elements $V_{i_1}, \ldots, V_{i_k}$. Let $\delta'(\tilde{\sigma}) = \min\{\delta(\tilde{\sigma}), \delta_{i_1}(\tilde{\sigma}), \ldots, \delta_{i_k}(\tilde{\sigma})\}$. Using iterated barycentric subdivision on each simplex $\tilde{\sigma}$, we can now construct a subdivision X of $\widetilde{X}$ such that the diameter of each simplex in $|\tilde{\sigma}|$ is less than $\delta'(\tilde{\sigma})$ and so X has the desired property. $\square$

Corollary 6.28. *Under the assumptions of Lemma 6.27, the piecewise linear functions $\hat{f} : |X| \to \mathbb{R}$ and $\hat{f}_i : |U_i| \to \mathbb{R}$ defined on the vertices by $\hat{f}(v) = f(h(v))$ and $\hat{f}_i(v) = f_i(h(v))$ and extended affinely over the simplices satisfy $\|\hat{f} - fh\|_\infty \leq \varepsilon$ and $\|\hat{f}_i - f_ih\|_\infty \leq \varepsilon$, respectively. Consequently, $\|\min_{i\in\Lambda} \hat{f}_i - \hat{f}\| \leq 2\varepsilon$.*

The final inequality means that upon replacing the functions f and f_i by piecewise linear approximations, the compatibility condition $f = \min_{i\in\Lambda} f_i$ remains approximately true. This is important, because the compatibility condition is needed to invoke the approximate nerve theorem for filtered simplicial complexes. We now have the necessary tools to prove an approximate nerve theorem for triangulable spaces.

Proposition 6.29. *Let Y be a locally compact triangulable space and $\overline{\mathcal{V}} = (V_i)_{i\in\Lambda}$ a locally finite triangulable cover of Y. Let $f : Y \to \mathbb{R}$ and $f_i : V_i \to \mathbb{R}$, $i \in \Lambda$, be continuous functions such that $f = \min_{i\in\Lambda} f_i$. Let $\mathcal{N}(\mathcal{V}) = (\mathcal{N}, g)$ be the nerve of the filtered cover $\mathcal{V} = (V_i, f_i)_{i\in\Lambda}$ of (Y, f). Let $D = \dim \mathcal{N}$, $\Delta = \dim Y$ and $Q = \min(D, \Delta) < \infty$. If $\mathcal{V}$ is ε-acyclic, $\mathsf{H}_*(Y, f) \overset{2(Q+1)\varepsilon+\eta}{\sim} \mathsf{H}_*(\mathcal{N}, g)$ holds for any $\eta > 0$. In particular,*

$$d_I(\mathsf{H}_*(Y, f), \mathsf{H}_*(\mathcal{N}, g)) \leq 2(Q+1)\varepsilon.$$

Proof. Let $(\widetilde{X}, h)$ be the triangulation from the definition of triangulable cover. By Lemma 6.27 and its corollary, there is a subdivision X of $\widetilde{X}$ and a corresponding cover $\overline{\mathcal{U}} = (U_i)_{i\in\Lambda}$ of X, satisfying $h(|U_i|) = V_i$, such that the piecewise linear functions $\hat{f} : |X| \to \mathbb{R}$ associated to fh and $\hat{f}_i : |X| \to \mathbb{R}$ associated to $f_i h$ satisfy $\|\hat{f} - fh\|_\infty < \delta$ and $\|\hat{f}_i - f_i h\|_\infty < \delta$, where $\delta > 0$ is to be chosen later.

Recall from Section 2.5 that there are two functors: the (natural) restriction functor $I_\delta : \mathbf{Vect}^{(\mathbb{R},\leq)} \to \mathbf{Vect}^{(\delta\mathbb{Z},\leq)}$ given by $I_\delta(F) = Fi_\delta$ and an extension functor $P_\delta : \mathbf{Vect}^{(\delta\mathbb{Z},\leq)} \to \mathbf{Vect}^{(\mathbb{R},\leq)}$ given by $P_\delta(F) = Fp_\delta$. Next, observe that defining $u^\delta(x) := \lceil \frac{u(x)}{\delta} \rceil \delta$, whenever u is a real-valued function, we have

$$P_\delta(I_\delta(\mathsf{H}_*(U_I, \hat{f}_I))) = \mathsf{H}_*(U_I, \hat{f}_I^\delta) \qquad \text{and} \qquad P_\delta(I_\delta(\mathsf{H}_*(X, \hat{f}))) = \mathsf{H}_*(X, \hat{f}^\delta). \tag{8}$$

Using the interleavings/isomorphisms provided by Propositions 2.15, 2.14, 2.10 and Equation (8), we obtain in turn

$$\mathsf{H}_*(V_I, f_I) \overset{\delta}{\sim} \mathsf{H}_*(V_I, \hat{f}_I h^{-1}) \cong \mathsf{H}_*(U_I, \hat{f}_I) \overset{\delta}{\sim} P_\delta(I_\delta(\mathsf{H}_*(U_I, \hat{f}_I))) = \mathsf{H}_*(U_I, \hat{f}_I^\delta).$$

By the same logic and using Corollary 6.28 to obtain the additional 2δ-interleaving in the middle, we have[18]

$$\begin{aligned}\mathsf{H}_*(Y, f) \overset{\delta}{\sim} \mathsf{H}_*(Y, \hat{f}h^{-1}) &\cong \mathsf{H}_*(X, \hat{f}) \\ &\overset{2\delta}{\sim} \mathsf{H}_*(X, \min_{i\in\Lambda} \hat{f}_i) \overset{\delta}{\sim} P_\delta(I_\delta(\mathsf{H}_*(X, \min_{i\in\Lambda} \hat{f}_i))) = \mathsf{H}_*(X, \min_{i\in\Lambda} \hat{f}_i^\delta).\end{aligned}$$

Since $\mathcal{V}$ is ε-acyclic and $\mathsf{H}_*(V_I, f_I) \overset{2\delta}{\sim} \mathsf{H}_*(U_I, \hat{f}_I^\delta)$ for all I, $\mathcal{U}$ is a $(\varepsilon + 2\delta)$-acyclic cover of $(X, \min_{i\in\Lambda} \hat{f}_i^\delta)$. In fact, it is $(p_\delta(\varepsilon) + 2\delta)$-acyclic. To see this, note that the $\mathbb{R}$-persistence modules $\mathsf{H}_*(U_I, \hat{f}_I^\delta)$ and $\mathsf{H}_*(X, \hat{f}^\delta)$ may be represented as $\delta\mathbb{Z}$-persistence modules, since their filtrations only change at $\delta\mathbb{Z}$ (see discussion following Proposition 2.12). This means that they lie in the image of the isometry P_δ. In particular, interleaving distances between such modules must be multiples of δ. Therefore, Theorem 6.24 applies to the pair $(X, \mathcal{U})$, where $\mathcal{U} = (U_i, \hat{f}_i^\delta)_{i\in\Lambda}$. Taking into account that $\mathcal{N}(\overline{\mathcal{U}}) = \mathcal{N}(\overline{\mathcal{V}}) = \mathcal{N}$, this means that

$$\mathsf{H}_*(X, \min_{i\in\Lambda} \hat{f}_i^\delta) \overset{2(Q+1)(p_\delta(\varepsilon)+2\delta)}{\sim} \mathsf{H}_*(\mathcal{N}, g_\delta),$$

where g_δ is the function on the nerve corresponding to the family of filtrations $(\hat{f}_i^\delta)_{i\in\Lambda}$. Using Proposition 2.11 we may now once again regard these as $\mathbb{R}$-persistence modules. It remains to compare g_δ with the function g corresponding to the family $(f_i)_{i\in\Lambda}$. Note that replacing each $f_i h$ by $\hat{f}_i^\delta$ changes the function values by at most 2δ, therefore we have $\|g - g_\delta\|_\infty \leq 2\delta$. Using Remark 2.6 we conclude that $\mathsf{H}_*(\mathcal{N}, g_\delta) \overset{2\delta}{\sim} \mathsf{H}_*(\mathcal{N}, g)$. Combining all these observations, we have

$$\mathsf{H}_*(Y, f) \overset{4\delta}{\sim} \mathsf{H}_*(X, \min_{i\in\Lambda} \hat{f}_i^\delta) \overset{2(Q+1)(p_\delta(\varepsilon)+2\delta)}{\sim} \mathsf{H}_*(\mathcal{N}, g_\delta) \overset{2\delta}{\sim} \mathsf{H}_*(\mathcal{N}, g),$$

so $\mathsf{H}_*(Y, f)$ and $\mathsf{H}_*(\mathcal{N}, g)$ are $(2(Q+1)\varepsilon + (4Q+10)\delta)$-interleaved, using $p_\delta(\varepsilon) \leq \varepsilon$. Choosing $\delta := \frac{\eta}{4Q+10}$ completes the proof. $\square$

[18]Note that $(\min_{i\in\Lambda} \hat{f}_i)^\delta = \min_{i\in\Lambda}(\hat{f}_i^\delta)$, so we can drop the parentheses in the final expression.

6.7 Lower Bounds

Here we construct simple examples to show that the bounds in Corollary 6.15, Theorem 6.18 and Proposition 6.20 are sharp. For each example, we compute the homology of the nerve, the homology of the filtered simplicial complex and the E^1, E^2 and E^∞ pages of the spectral sequence (up to isomorphism). For better readability, we use the notations

$$[a, b] = \{k \in \mathbb{Z} \mid a \leq k \leq b\} \qquad \text{and} \qquad [a, b) = [a, b] \setminus \{b\}.$$

Without loss of generality, we work with $\epsilon = 1$, otherwise simply multiply each time in the filtration by ϵ. To simplify the exposition, the pages of the spectral sequence are not computed directly, but rather inferred from the homology of the space and various intersections of its cover elements.

In each of the two examples provided, the filtration is defined on the total space X. The cover elements U_i are assumed to be equipped with the induced filtrations (see Remark 2.4). Each example illustrates the tightness of each step of our approximation proof. To construct a topological example which achieves all three, we can simply take the direct sum of the three examples.

6.7.1 First Example

Our first example realizes the bounds in Corollary 6.15 and Theorem 6.18. Let X be the D-sphere, realized as the boundary of the $(D+1)$-simplex with vertex set $[0, D+1]$. A cover $\mathcal{U}$ of X is given by its set of maximal faces, i.e. $\mathcal{U} = \{U_i \mid i \in [0, D+1]\}$, where U_i is the D-simplex spanned by $[0, D+1] \setminus \{i\}$.

We also define a filtration $X^0 \leq X^2 \leq \ldots \leq X^{2D+2}$ by adding one cover element at a time, i.e.

$$X^{2j} = U_0 \cup U_1 \cup \ldots \cup U_j.$$

Proposition 6.30. *The homology of the nerve of $\mathcal{U}$ is given by*

$$\mathsf{H}_q(\mathcal{N}) \cong \begin{cases} \mathbb{k}[t]; & q = 0, \\ t^2\mathbb{k}[t]; & q = D, \\ 0; & \textit{otherwise.} \end{cases}$$

Proof. At time 0, the vertices $1, \ldots, D+1$ are born in X. For $I \subseteq [0, D+1]$ each U_I except $U_{[1,D+1]}$ and $U_{[0,D+1]}$ contains one of these vertices, so the nerve at time 0 consists of all $I \subseteq [0, D+1]$, except for $[1, D+1]$ and $[0, D+1]$. At time 2, the vertex 0 is born, which corresponds to the birth of $[1, D+1]$ in the nerve. Since $U_{[0,D+1]}$ is always empty, $\mathcal{N}^j$ is contractible for $j = 0, 1$ and homeomorphic to a D-sphere for $j \geq 2$. $\square$

Proposition 6.31. *The homology of the filtered simplicial complex X is given by*

$$\mathsf{H}_q(X) \cong \begin{cases} \mathbb{k}[t]; & q = 0, \\ t^{2D+2}\mathbb{k}[t]; & q = D, \\ 0; & \textit{otherwise.} \end{cases}$$

Proof. X^j is contractible at the times $j = 0, \ldots, 2D+1$. For $j \geq 2D+2$ it is homeomorphic to a D-sphere. $\square$

Computing the E^1 page requires some preparation, namely simplifying U_I^{2j}.

Proposition 6.32. *Suppose that $\emptyset \neq I \subseteq [0, D+1]$ and let $j \in [0, D+1]$. If $j \geq \min I$, U_I^{2j} is a $(D+1-|I|)$-simplex. If $j < \min I$, U_I^{2j} is the join of a $(j-1)$-sphere, realized as the boundary of a j-simplex, and a $(D-|I|-j)$-simplex. We allow $D-|I|-j = -1$ and interpret "(-1)-simplex" as the empty set.*

Proof. Observe that

$$U_I^{2j} = U_I \cap X^{2j} = U_I \cap (U_0 \cup \ldots \cup U_j) = U_{I\cup\{0\}} \cup U_{I\cup\{1\}} \cup \ldots \cup U_{I\cup\{j\}}.$$

For $j \geq \min I$, one of the terms is $U_{I\cup\{\min I\}} = U_I$, so $U_I^{2j} = U_I$ is the $(D+1-|I|)$-simplex with vertices $[0, D+1] \setminus I$. (Intersecting with U_i corresponds to removing the vertex i.)

For $j < \min I$, $U_{I\cup\{k\}}$ (where $k \in [0, j]$) is the $(D-|I|)$-simplex with vertices $[0, D+1] \setminus (I \cup \{k\})$. So, U_I^{2j} is the complex spanned by all the simplices of the form $J \cup ([j+1, D+1] \setminus I)$ where J is a j-element subset of $[0, j]$. But this means precisely that U_I^{2j} is the simplicial join of the $(D-|I|-j)$-simplex $[j+1, D+1] \setminus I$ and the $(j-1)$-sphere, realized as the boundary of the j-simplex $[0, j]$. $\square$

Proposition 6.33. *Suppose that $\emptyset \neq I \subseteq [0, D+1]$. Then*

$$\mathsf{H}_q(U_I) \cong \begin{cases} \frac{t^{2q+2}\Bbbk[t]}{t^{2q+4}\Bbbk[t]}; & q = D-|I| > 0, I = [q+2, D+1], \\ \Bbbk[t] \oplus \frac{t^2\Bbbk[t]}{t^4\Bbbk[t]}; & q = 0, I = [2, D+1], \\ \Bbbk[t]; & q = 0 < D-|I| \text{ or } q = 0, |I| = D+1, I \neq [1, D+1], \\ t^2\Bbbk[t]; & q = 0, I = [1, D+1], \\ 0; & \textit{otherwise.} \end{cases}$$

Proof. The join of a sphere and a non-empty simplex is contractible, so U_I^{2j} can only be non-acyclic is if it is the join of a sphere and an empty simplex. By the previous proposition, this occurs precisely if $D = |I| + j - 1$ and $I = [j+1, D+1]$ (the latter is required so that $\min I > j$) in which case U_I^{2j} is a $(j-1)$-sphere. If $j-1 > 0$, this means that $\mathsf{H}_{j-1}(U_I^{2j}) \cong \Bbbk$ and $\mathsf{H}_0(U_I^{2j}) \cong \Bbbk$, if $j-1 = 0$, it means that $\mathsf{H}_{j-1}(U_I^{2j}) \cong \Bbbk^2$, and for $j = 0$ all homology groups (corresponding to $U_{[1,D+1]}$) are trivial. In all other cases, U_I^{2j} is contractible, so $\mathsf{H}_0(U_I^{2j}) \cong \Bbbk$. The remaining homology groups are 0. Setting $q = j-1$ completes the proof. $\square$

Note that we have computed persistent homology slicewise, i.e. by computing the simplicial homology at each step of the filtration. To infer the correct $\Bbbk[t]$-module from this, we have used the facts that the filtration only changes at even times and that once it is born, the first class appearing in dimension 0 lives forever. One immediate consequence of these computations is the following.

Corollary 6.34. *The cover $\mathcal{U}$ is 1-acyclic.*

Since we already know that

$$E^1_{p,q} = \bigoplus_{|I|=p+1} \mathsf{H}_q(U_I),$$

the previous proposition also immediately yields the E^1 page.

Corollary 6.35. *The E^1 page of the Mayer-Vietoris spectral sequence of $(X,\mathcal{U})$ is given by:*

$$\begin{array}{|llllll}
\frac{t^{2D}\Bbbk[t]}{t^{2D+2}\Bbbk[t]} & & & & & \\
& \frac{t^{2D-2}\Bbbk[t]}{t^{2D}\Bbbk[t]} & & & & \\
& & \ddots & & & \\
& & & \frac{t^{4}\Bbbk[t]}{t^{6}\Bbbk[t]} & & \\
\Bbbk[t]^{\binom{D+2}{1}} & \Bbbk[t]^{\binom{D+2}{2}} & \cdots & \Bbbk[t]^{\binom{D+2}{D-1}} & \Bbbk[t]^{\binom{D+2}{D}} \oplus \frac{t^2\Bbbk[t]}{t^4\Bbbk[t]} & \Bbbk[t]^{\binom{D+2}{D+1}-1} \oplus t^2\Bbbk[t] \\
\hline
\end{array}$$

The E^2 page can be inferred from this.

Corollary 6.36. *The E^2 page of the Mayer-Vietoris spectral sequence of $(X,\mathcal{U})$ is given by:*

$$\begin{array}{|llllll}
\frac{t^{2D}\Bbbk[t]}{t^{2D+2}\Bbbk[t]} & & & & & \\
& \frac{t^{2D-2}\Bbbk[t]}{t^{2D}\Bbbk[t]} & & & & \\
& & \ddots & & & \\
& & & \frac{t^{4}\Bbbk[t]}{t^{6}\Bbbk[t]} & & \\
\Bbbk[t] & & & & 0 & t^4\Bbbk[t] \\
\hline
\end{array}$$

Proof. We have already seen that all persistent homology groups $\mathsf{H}_q(X)$ for $q \neq 0, D$ are trivial. This means that the corresponding antidiagonals on the E^∞ page must consist of trivial modules. As there are no non-trivial differentials to and from $E^r_{p,0}$ for $p \neq D$ for $r \geq 2$, these modules stabilize already on E^2. Hence, these are all trivial, except for $E^2_{0,0} \cong \mathsf{H}_0(X) \cong \Bbbk[t]$. The modules $E^2_{p,q}$ for $q > 0$ are isomorphic to $E^1_{p,q}$ since d^1 is trivial above the bottom row. Finally, $E^2_{D,0} = \ker d^1_{D,0}$. This can be computed explicitly from the generators, or inductively, as follows. We already know most of E^2, so we can use this to our advantage. Namely, we know that

$$\frac{\ker d^1_{0,0}}{\operatorname{im} d^1_{1,0}} \cong \Bbbk[t]$$

and for $p = 1, \dots, D-1$ we have

$$\ker d^1_{p,0} \cong \operatorname{im} d^1_{p+1,0}.$$

From this, using the first isomorphism theorem, we can inductively infer that for $p = 0, \dots, D-2$

$$\ker d^1_{p,0} \cong \mathbb{k}[t]^{\sum_{k=0}^{p+1}(-1)^k\binom{D+2}{p+1-k}}$$

and using the binomial theorem

$$\ker d^1_{D-1,0} \cong \frac{t^2\mathbb{k}[t]}{t^4\mathbb{k}[t]} \oplus \mathbb{k}[t]^{\sum_{l=2}^{D+2}(-1)^l\binom{D+2}{D+2-l}} \cong \frac{t^2\mathbb{k}[t]}{t^4\mathbb{k}[t]} \oplus \mathbb{k}[t]^{D+1}.$$

Since

$$E^1_{D,0} \cong t^2\mathbb{k}[t] \oplus \mathbb{k}[t]^{D+1}$$

and

$$\frac{t^2\mathbb{k}[t]}{t^4\mathbb{k}[t]} \oplus \mathbb{k}[t]^{D+1} \cong \operatorname{im} d^1_{D,0} \cong \frac{E^1_{D,0}}{\ker d^1_{D,0}} \cong \frac{t^2\mathbb{k}[t] \oplus \mathbb{k}[t]^{D+1}}{\ker d^1_{D,0}},$$

we finally infer that $\ker d^1_{D,0} \cong t^4\mathbb{k}[t]$ and thus conclude the proof. □

The E^∞ page can be inferred in a similar fashion.

Corollary 6.37. *The E^∞ page of the Mayer-Vietoris spectral sequence of $(X, \mathcal{U})$ is given by:*

$\mathbb{k}[t]$	$t^{2D+2}\mathbb{k}[t]$

Proof. Note that the only non-trivial differential on the r-th page, $2 \leq r \leq D$, is $d^r_{D,0}$. Note that $E^R_{D-r,r-1}$ has already stabilized for $R > r$, as there are no more non-trivial differentials to and from this module. Since $\mathsf{H}_{D-1}(X) = 0$, we can infer that $E^{r+1}_{D-r,r-1} = 0$ and that $d^r_{D,0}$ is surjective. A simple inductive argument shows that $E^r_{D,0} \cong t^{2r}\mathbb{k}[t]$. The spectral sequence collapses at $r = D+1$ where $E^r_{D,0} \cong t^{2D+2}\mathbb{k}[t]$. □

From these considerations, it follows that this example has the following properties:

- $E^2_{*,0} \overset{\eta}{\sim} \mathsf{H}_*(\mathcal{N})$ holds for $\eta = 2$ but not for $\eta < 2$,
- $E^2_{*,0} \overset{\eta}{\sim} E^\infty_{*,0}$ holds for $\eta = 2(D-1)$ but not for $\eta < 2(D-1)$,

and therefore, it attains the bounds from Corollary 6.15 and Theorem 6.18, so these bounds are in fact sharp.

6.7.2 Second Example

Our second example shows that the bound in Proposition 6.20 is also sharp. Let $D \geq 1$ and let X be the simplicial complex with vertex set $[0, D+3]$ consisting of all simplices $\sigma \subseteq [0, D+3]$ such that $[1, D+1] \nsubseteq \sigma$ and $\{D+2, D+3\} \nsubseteq \sigma$.

We may visualize X geometrically as a bipyramid consisting of two $(D+1)$-simplices, each of which is subdivided into $D+1$ smaller $(D+1)$-simplices. More specifically, consider the subdivision of the D-simplex $[1, D+1]$ into $D+1$ smaller D-simplices obtained by adding the point 0 at the barycenter and connecting it to the vertices (note that this is *not* the barycentric subdivision). Then X can be understood as the union of two cones over this subdivision, whose apices are $D+2$ and $D+3$.

A cover $\mathcal{U}$ of X is given by the cone with apex $D+3$ and the $D+1$ small $(D+1)$-simplices the cone with apex $D+2$ is subdivided into. Specifically, the 0-th cover element U_0 is the full subcomplex of X spanned by $[0, D+1] \cup \{D+3\}$ and for each $i \in [1, D+1]$, the i-th cover element U_i is defined as the full subcomplex of X spanned by $[0, D+2] \setminus \{i\}$. In the geometric interpretation mentioned above, the intersection U_I with $0 \notin I$ is the cone with apex $D+2$ over the corresponding $(D+1-|I|)$-simplex occurring in the subdivision of the base D-simplex $[1, D+1]$, and $U_{I \cup \{0\}}$ is this base $(D+1-|I|)$-simplex.

Next, we define a filtration. The idea is to start with the boundary of the bipyramid X and fill in the U_i one at a time. Let A be the subcomplex of X obtained by removing all simplices $\sigma \subseteq [0, D+3]$ such that $0 \in \sigma$. Geometrically, A corresponds to the boundary of the bipyramid X. A filtration $X^{-2D} \leq X^0 \leq X^2 \leq X^4 \leq \ldots \leq X^{2D+2}$ of X is defined by

$$X^{-2D} = A \qquad \text{and} \qquad X^{2j} = A \cup U_0 \cup \ldots \cup U_j \qquad \text{for } j \geq 0.$$

We claim that with this filtration, X achieves the relevant bound of $2(D+1)$. To see this, we compute the E^1 page of the spectral sequence directly. This corresponds to computing the persistent homology of the $|I|$-fold intersections U_I, equipped with the naturally induced filtrations $U_I^{2j} = U_I \cap X^{2j}$.

First, we compute homology of the nerve of the cover of X.

Proposition 6.38. *The persistent homology of the nerve of $\mathcal{U}$ is given by*

$$\mathsf{H}_q(\mathcal{N}) = \begin{cases} t^{-2D}\Bbbk[t]; & q = 0, \\ \frac{t^{-2D}\Bbbk[t]}{\Bbbk[t]}; & q = D, \\ 0; & \textit{otherwise}. \end{cases}$$

Proof. First recall the definition of a cover element U_I. For $I \subseteq [0, D+2]$, U_I consists of all simplices in the space spanned by $[0, D+2] \setminus I$. The nerve becomes non-empty at time $-2D$, since at this time all D-simplices $I \subseteq [0, D+1]$, $|I| = D+1$, are born. This is because for the simplex $I = [0, D+1] \setminus \{i\}$ in the nerve, where $i > 0$, by definition U_I^{-2D} contains the point i. On the other hand, if $I = [1, D+1]$, the point $D+2$ is contained in U_I^{-2D}. However, the top simplex $[0, D+1]$ only contains the point 0, which is born at time 0. It follows from all this that $\mathcal{N}^j$ is a $(D+1)$-simplex for $j \in [0, \infty)$ and the boundary of this $(D+1)$-simplex for $j \in [-2D, 0)$. ☐

Next, we compute the persistent homology of the union.

Proposition 6.39. *The persistent homology of X is given by*

$$\mathsf{H}_q(X) = \begin{cases} t^{-2D}\Bbbk[t]; & q = 0, \\ \frac{t^{-2D}\Bbbk[t]}{t^{2D+2}\Bbbk[t]}; & q = D, \\ 0; & \textit{otherwise.} \end{cases}$$

Proof. For each $j \in [0, D]$, there is a collapse of X^{2j} to X^{-2D}. Observing that $X^{2D+2} = X$ is a bipyramid and $X^{-2D} = A$ is its boundary completes the proof. □

In order to compute E^1, we describe U_I^{2j} in more familiar terms.

Proposition 6.40. *Suppose that $\emptyset \neq I \subseteq [1, D+1]$ and let $j \in [0, D+1]$. Then the following hold:*

- *if $j \geq \min I$, U_I^{2j} is a $(D+2-|I|)$-simplex,*
- *if $j < \min I$, U_I^{2j} is the join of the boundary j-sphere of a $(j+1)$-simplex and a $(D-|I|-j)$-simplex,*
- *U_I^{-2D} is a $(D+1-|I|)$-simplex,*
- *$U_{I\cup\{0\}}^{2j}$ is a $(D+1-|I|)$-simplex and $U_{I\cup\{0\}}^{-2D}$ is a $(D-|I|)$-simplex,*
- *U_0^{2j} is a subdivided $(D+1)$-simplex and U_0^{-2D} is the cone over the boundary of a D-simplex.*

In the second and fourth bullet points, we allow $D-|I|-j = -1$ resp. $D-|I| = -1$ and interpret "(-1)-simplex" as the empty set.

Proof. We begin by proving the first three bullet points. Observe that

$$U_I^{2j} = U_I \cap X^{2j} = U_I \cap ((A \cup U_0) \cup U_1 \cup \ldots \cup U_j) = (U_I \cap (A \cup U_0)) \cup U_{I\cup\{1\}} \cup \ldots \cup U_{I\cup\{j\}}.$$

For $j \geq \min I$, one of the terms is $U_{I\cup\{\min I\}} = U_I$, so $U_I^{2j} = U_I$ is the $(D+2-|I|)$-simplex with vertices $[0, D+2] \setminus I$. (Intersecting with $U_i, i > 0$, corresponds to removing the vertex i.)

For $j < \min I$, $U_{I\cup\{k\}}$ (where $k \in [1, j]$) is the $(D+1-|I|)$-simplex with vertices $[0, D+2] \setminus (I \cup \{k\})$. The first term, $U_I \cap (A \cup U_0)$, consists of two $(D+1-|I|)$-simplices, $[0, D+1] \setminus I$ and $[1, D+2] \setminus I$. So, U_I^{2j} is the complex spanned by all the simplices of the form $J \cup ([j+1, D+1] \setminus I)$ where J is a $(j+1)$-element subset of $[0, j] \cup \{D+2\}$. But this means precisely that U_I^{2j} is the simplicial join of the $(D-|I|-j)$-simplex $[j+1, D+1] \setminus I$ and the j-sphere, realized as the boundary of the $(j+1)$-simplex $[0, j] \cup \{D+2\}$.

By definition, $U_I^{-2D} = U_I \cap A$ is the $(D+1-|I|)$-simplex spanned by $[0, D+2] \setminus (I \cup \{0\})$.

To prove the fourth bullet point, note that

$$\begin{aligned} U_{I\cup\{0\}}^{2j} = U_{I\cup\{0\}} \cap X^{2j} &= U_{I\cup\{0\}} \cap ((A \cup U_0) \cup U_1 \cup \ldots \cup U_j) \\ &= U_{I\cup\{0\}} \cup U_{I\cup\{0,1\}} \cup \ldots \cup U_{I\cup\{0,j\}} = U_{I\cup\{0\}} \end{aligned}$$

is the $(D+1-|I|)$-simplex spanned by $[0,D+1]\setminus I$ and $U_{I\cup\{0\}}^{-2D} = U_{I\cup\{0\}}\cap A$ is the $(D-|I|)$-simplex spanned by $[1,D+1]\setminus I$.

The last bullet point follows by definition of U_0, namely $U_0^{2j} = U_0 \cap X^{2j} = U_0$ is the half of the bipyramid with apex $D+3$, so it is a subdivided $(D+1)$-simplex, and $U_0^{-2D} = U_0 \cap X^{-2D} = U_0 \cap A$ is the cone with apex $D+3$ over the boundary $(D-1)$-sphere of the base D-simplex of the bipyramid. □

Using this fact, we can compute the persistent homology of the intersections U_I.

Proposition 6.41. *Suppose that $\emptyset \neq I \subseteq [1,D+1]$. Then*

$$\mathsf{H}_q(U_I) = \begin{cases} \frac{t^{2q}\Bbbk[t]}{t^{2q+2}\Bbbk[t]}; & q = D+1-|I| > 0, I = [q+1,D+1], \\ t^{-2D}\Bbbk[t] \oplus \frac{\Bbbk[t]}{t^2\Bbbk[t]}; & q = D+1-|I| = 0, \\ t^{-2D}\Bbbk[t]; & q = 0 < D+1-|I|, \\ 0; & \textit{otherwise.} \end{cases}$$

and for any $I \subseteq [1,D+1]$ we have

$$\mathsf{H}_q(U_{I\cup\{0\}}) = \begin{cases} t^{-2D}\Bbbk[t]; & q = 0 \textit{ and } I \neq [1,D+1], \\ \Bbbk[t]; & q = 0 \textit{ and } I = [1,D+1], \\ 0; & \textit{otherwise.} \end{cases}$$

Proof. The join of a sphere and a non-empty simplex is contractible, so U_I^{2q} can only be non-acyclic if it is the join of a sphere and an empty simplex. The previous proposition shows that this occurs precisely if $D = |I|+q-1$ and $I = [q+1,D+1]$ (the latter is required so that $\min I > q$) in which case U_I^{2q} is a q-sphere. If $q > 0$, this means that $\mathsf{H}_q(U_I^{2q}) \cong \Bbbk$ and $\mathsf{H}_0(U_I^{2q}) \cong \Bbbk$ and if $q = 0$, it means that $\mathsf{H}_q(U_I^{2q}) \cong \Bbbk^2$. In all other cases, including $q = -D$, U_I^{2q} is contractible, so $\mathsf{H}_0(U_I^{2q}) \cong \Bbbk$. All other homology groups of U_I^{2q} are 0.

The second part holds because, once born, $U_{I\cup\{0\}}^{2q}$ is contractible, so we have $\mathsf{H}_0(U_{I\cup\{0\}}^{2q}) = \Bbbk$ for $q \geq -D$ if $I \neq [1,D+1]$ and for $q \geq 0$ otherwise. □

Again, persistent homology has been computed slicewise, so the remark from the first example applies.

Corollary 6.42. *The cover $\mathcal{U}$ is 1-acyclic.*

As in the previous example, this immediately yields the E^1 page.

Corollary 6.43. *The E^1 page of the Mayer-Vietoris spectral sequence of $(X,\mathcal{U})$ is given by:*

$$
\begin{array}{|lllll}
\frac{t^{2D}\Bbbk[t]}{t^{2D+2}\Bbbk[t]} & & & & \\
& \frac{t^{2D-2}\Bbbk[t]}{t^{2D}\Bbbk[t]} & & & \\
& & \ddots & & \\
& & & \frac{t^{2}\Bbbk[t]}{t^{4}\Bbbk[t]} & \\
(t^{-2D}\Bbbk[t])^{\binom{D+2}{1}} & (t^{-2D}\Bbbk[t])^{\binom{D+2}{2}} & \cdots & (t^{-2D}\Bbbk[t])^{\binom{D+2}{D}}\frac{\Bbbk[t]}{t^2\Bbbk[t]} \oplus (t^{-2D}\Bbbk[t])^{\binom{D+2}{D+1}} & \Bbbk[t]^{\binom{D+2}{D+2}} \\
\hline
\end{array}
$$

This allows us to infer the E^2 page.

Corollary 6.44. *The $E^2 = E^\infty$ page of the Mayer-Vietoris spectral sequence of $(X, \mathcal{U})$ is given by:*

$$
\begin{array}{|lllll}
\frac{t^{2D}\Bbbk[t]}{t^{2D+2}\Bbbk[t]} & & & & \\
& \frac{t^{2D-2}\Bbbk[t]}{t^{2D}\Bbbk[t]} & & & \\
& & \ddots & & \\
& & & \frac{t^{2}\Bbbk[t]}{t^{4}\Bbbk[t]} & \\
t^{-2D}\Bbbk[t] & & & & \frac{t^{-2D}\Bbbk[t]}{t^2\Bbbk[t]} \\
\hline
\end{array}
$$

Proof. We claim that the map $d^1_{D+1,0} : E^1_{D+1,0} \to E^1_{D,0}$ is injective. To see this, note that $([0], [0, D+1])$ is a generator of E^1_{D+1} (see Section 2.6 for notation) and its image

$$d^1([0], [0, D+1]) = \sum_{l=0}^{D+1} (-1)^l ([0], [0, D+1] \setminus \{l\})$$

generates a free submodule of $E^1_{D,0}$, since $([0], [0, D+1] \setminus \{l\})$ generates a free submodule of $\mathsf{H}_0(U_{[0,D+1]\setminus\{l\}})$. This implies that $E^2_{D+1,0} = 0$, so there are no non-trivial differentials to or from any $E^r_{p,q}$ for $r \geq 2$. If $q > 0$, this is true also for $r = 1$. Therefore, the spectral sequence collapses on E^2 and $E^1_{p,q} \cong E^2_{p,q}$ for $q > 0$. As we have seen, the persistent homology groups $\mathsf{H}_q(X)$ for $q \neq 0, D$ are trivial. This means that the corresponding antidiagonals on the $E^2 = E^\infty$ page must consist of trivial modules. Furthermore, $E^2_{0,0} \cong \mathsf{H}_0(X) \cong t^{-2D}\Bbbk[t]$. Finally, we shall compute $E^2_{D,0}$ explicitly. We already know $\operatorname{im} d^1_{D+1,0}$, so it remains to compute $\ker d^1_{D,0}$ and the corresponding quotient. Using a similar inductive argument as in the proof of Corollary 6.36, we have

$$
\begin{aligned}
\frac{\Bbbk[t]}{t^2\Bbbk[t]} \oplus (t^{-2D}\Bbbk[t])^{D+2} \cong E^1_{D,0} &\cong \operatorname{im} d^1_{D,0} \oplus \ker d^1_{D,0} \\
&\cong \ker d^1_{D-1,0} \oplus \ker d^1_{D,0} \cong (t^{-2D}\Bbbk[t])^{D+1} \oplus \ker d^1_{D,0}.
\end{aligned}
$$

Since these modules are finitely generated, we may conclude that

$$\ker d^1_{D,0} \cong \frac{\Bbbk[t]}{t^2\Bbbk[t]} \oplus t^{-2D}\Bbbk[t].$$

In fact, the generators may be deduced from the explicit description of the intersections of the cover elements. Namely, they are given by

$$a := (t^{2D}[D+2] - [0], [1, D+1])$$

and

$$b := ([D+2], [1, D+1]) + \sum_{l=1}^{D+1} (-1)^l ([l], [0, D+1] \setminus \{l\}),$$

subject to the single relation $t^2 a = 0$. Note that since $([0], [0, D+1] \setminus \{l\}) = t^{2D}([l], [0, D+1] \setminus \{l\})$, the generator of $\operatorname{im} d^1_{D+1,0}$ may be written as $t^{2D}b - a$. Therefore, letting x and y be the generators of $\Bbbk[t] \oplus t^{-2D}\Bbbk[t]$, the quotient may be computed as follows:

$$E^2_{D,0} \cong \frac{\langle x, y\rangle}{\langle t^2 x, t^{2D} y - x\rangle} = \frac{\langle t^{2D} y - x, y\rangle}{\langle t^{2D+2} y, t^{2D} y - x\rangle} \cong \frac{\langle y\rangle}{\langle t^{2D+2} y\rangle} \cong \frac{t^{-2D}\Bbbk[t]}{t^2\Bbbk[t]}.$$

□

The modules $\mathsf{H}_D(X) = \frac{t^{-2D}\Bbbk[t]}{t^{2D+2}\Bbbk[t]}$ and $E^\infty_{D,0} = \frac{t^{-2D}\Bbbk[t]}{t^2\Bbbk[t]}$ are η-interleaved for $\eta = 2D$, but not so for any $\eta < 2D$. Therefore, this example attains the bound of Proposition 6.20, so this bound is also sharp.

7 Conclusions

The concept of unimodal category, as envisioned by Baryshnikov and Ghrist in [4], has its origins in statistics. However, the definition of the concept requires much less than this, namely a nonnegative function $f : X \to [0, \infty)$. We feel that the question of how many functions with a unique local maximum are needed to express such an f is very natural from the point of view of mathematical analysis. For this reason, we find it somewhat surprising that this question has received so little attention thus far. One of the main aims of our work was to demonstrate that this area of mathematics admits many interesting questions and has the potential to become a vibrant area of research.

Our work is mainly centered around the monotonicity conjecture of [4], which has turned out to be a more interesting question than initially thought, especially since it turned out to be false. This leaves us with many open questions. For instance, the constructions we provide rely on the existence of cycles in the superlevel sets of the functions. This leads one to wonder if there is a more conceptual reason explaining this failure of monotonicity in the presence of cycles and what are the precise conditions a function should satisfy for monotonicity to hold. It would be interesting to construct topological invariants measuring the extent to which monotonicity can fail.

We have reformulated the original results of Baryshnikov and Ghrist for functions $f : \mathbb{R} \to [0, \infty)$ in a language that we feel is more natural than that of the original article, using the concepts of total, positive and negative variation. This has led us to a general decomposition theorem for such functions, as well as a characterization of functions $f : S^1 \to [0, \infty)$. The question of what the natural context for a general treatment of **ucat** for continuous functions $f : \mathbb{R}^n \to [0, \infty)$ might be, remains widely open. We speculate that the answer might lie in a new kind of (co)homology theory, designed especially to treat such problems in general. This could in turn provide a definite connection to the methods exploited in Section 6.

In fact, our work opens many more questions than it answers. This is also the reason for the inclusion of Section 5 in the dissertation. Our hope is that the various open questions posed there might stimulate other researchers from various areas of mathematics to work on such problems.

The initial motivation for the work regarding the approximate nerve theorem was algorithmic - given a filtered simplicial complex, it would be computationally desirable to construct a coarser simplicial complex via a cover such that the persistent homology was preserved. This has been done for metric spaces [61] but not for more general filtrations. An alternate spectral sequence approach is used for computation of persistence, but it does not allow for passing to a coarser representation. Our results suggest a natural approximation algorithm, where a coarse cover is constructed and the condition of ε-acyclicity is checked locally for each finite intersection. Conversely, the maximum ε overall finite non-empty intersections could provide the bound. We would then have an explicit error bound relating the persistent homology of the input simplicial complex and the coarser (and presumably smaller) nerve.

Beyond the initial motivation, our setting of $\Bbbk[t]$-modules and simplicial complexes may seem restrictive. However, these were chosen to make the constructions as explicit as possible and to avoid technical complications. We believe the bounds hold in much greater generality. For example, a natural direction is to consider a sheaf of q-tame persistence modules and to use the Leray spectral sequence, of which the Mayer-Vietoris spectral sequence is a special case. We believe the error analysis goes through identically and plan to address this in a separate note. The main technical obstacles are in setting up the spectral sequence so that the differentials are well defined.

Likewise, the restriction to simplicial complexes is mainly to avoid complications and should hold for CW complexes or perhaps even suitably nice singular spaces. In general, our results should simplify proving approximation results. It does not require individual sublevel sets of a function to have a good cover at any particular level. In particular, this removes the need to consider the image of a pair of covers. Finally, the ε-acyclicity is a local condition, making it easier to verify in a number of applications.

Finally, we note that this work can also be extended to multidimensional persistence modules. The weaker bound using only interleaving applies directly. The tighter bound does not however, as in our proof that left and right interleavings do not interact (Proposition 6.11), the last case uses the fact that persistence modules have projective dimension of 1, which does not hold for multidimensional persistence modules.

References

[1] P. Agarwal, H. Edelsbrunner, J. Harer, Y. Wang, *Extreme elevation on a 2-manifold*, Discrete and Computational Geometry, **36**(4), 553–572 (2006).

[2] P. Alexandroff, *Über den allgemeinen dimensionsbegriff und seine beziehungen zur elementaren geometrischen anschauung*, Mathematische Annalen, **98**(1), 617–635 (1928).

[3] T. Apostol, *Mathematical Analysis*, 2nd edition, Addison-Wesley Publishing Company, Reading, Massachusetts (1974).

[4] Y. Baryshnikov, R. Ghrist, *Unimodal Category and Topological Statistics*, Proc. NOLTA (2011).

[5] U. Bauer, M. Lesnick, *Induced matchings of barcodes and the algebraic stability of persistence*, Journal of Computational Geometry **6**(2), 162–191 (2015).

[6] U. Bauer, C. B. Schönlieb, M, Wardetzky, *Total Variation Meets Topological Persistence: A First Encounter*, Proceedings of ICNAAM, 1022–1026 (2010).

[7] J. Behboodian, *On the modes of a mixture of two normal distributions*, Technometrics **12**, 131–139 (1970).

[8] R. H. Bing, *The Geometric Topology of 3-Manifolds*, American Mathematical Society, Providence, Rhode Island (1983).

[9] O. Bobrowski, S. Mukherjee, *The topology of probability distributions on manifolds*, Probability Theory and Related Fields, **161**(3-4), 651–686 (2015).

[10] M. B. Botnan, G. Spreemann, *Approximating persistent homology in euclidean space through collapses,* arXiv preprint, arXiv:1403.0533 (2014).

[11] R. Bott, L. W. Tu, *Differential forms in algebraic topology*, Springer, New York (1982).

[12] K. S. Brown, *Cohomology of groups*, Springer, New York (1982).

[13] P. Bubenik, J. A. Scott, *Categorification of persistent homology*, Discrete and Computational Geometry, **51**(3), 600–627 (2014).

[14] G. Carlsson, *Topology and data*, Bulletin of the American Mathematical Society, **46**, 255–308, (2009).

[15] N. J. Cavanna, D. R. Sheehy, *Persistent nerves revisited*, URL: http://matthewpjohnson.org/fwcg2016/FWCG_2016_paper_17.pdf.

[16] F. Chazal, D. Cohen-Steiner, M. Glisse, L. Guibas, S. Oudot, *Proximity of persistence modules and their diagrams*, Proceedings 25th Annual Symposium on Computational Geometry, 237–246 (2009).

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

[18] F. Chazal, W. Crawley-Boevey, V. De Silva, *The observable structure of persistence modules*, Homology, Homotopy and Applications, **18**(2), 247–265 (2016).

[19] F. Chazal, V. De Silva, M. Glisse, S. Oudot, *The structure and stability of persistence modules*, Springer, USA (2016).

[20] F. Chazal, L. Guibas, S. Oudot, P. Škraba, *Analysis of scalar fields over point cloud data*, Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, 1021–1030 (2009).

[21] F. Chazal, L. Guibas, S. Oudot, P. Škraba, *Scalar field analysis over point cloud data*, Discrete and Computational Geometry, **46**(4), 743–775 (2011).

[22] F. Chazal, S. Oudot, *Towards persistence-based reconstruction in euclidean spaces*, Proceedings 24th Annual Symposium on Computational Geometry, 232–241 (2008).

[23] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, *Stability of persistence diagrams*, Discrete and Computational Geometry, **37**, 103–120 (2007).

[24] O. Cornea, G. Lupton, J. Oprea, D. Tanré, *Lusternik-Schnirelmann Category*, American Mathematical Society, Providence, Rhode Island (2003).

[25] J. Curry, R. Ghrist, M. Robinson, *Euler Calculus with Applications to Signals and Sensing*, Proc. Sympos. Appl. Math., American Mathematical Society (2012).

[26] T. Dey, F. Fan, Y. Wang, *Graph induced complex on point data*, Computational Geometry, **48**(8), 575–588 (2015).

[27] J. L. Doob, *Measure Theory*, Springer-Verlag, New York (1994).

[28] H. Edelsbrunner, B. T. Fasy, G. Rote, *Add Isotropic Gaussian Kernels at Own Risk: More and More Resilient Modes in Higher Dimensions*, Discrete Comput. Geom. **49** (4), 797–822 (2013).

[29] H. Edelsbrunner, J. Harer, *Computational Topology*, American Mathematical Society, USA (2010).

[30] H. Edelsbrunner, J. Harer, *Persistent homology – a survey*, Contemporary Mathematics **453**, 257–282 (2008).

[31] H. Edelsbrunner, J. Harer, V. Natarajan, V. Pascucci, *Morse-Smale Complexes for Piecewise Linear 3-Manifolds*, Proc. Symp. on Computational Geometry, 361–370 (2003).

[32] H. Edelsbrunner, J. Harer, A. Zomorodian, *Hierarchical Morse-Smale Complexes for Piecewise Linear 2-Manifolds*, Discrete Comput. Geom. **30** (1), 87–107 (2003).

[33] I. Eisenberger, *Genesis of bimodal distributions*, Technometrics **6**, 357–363 (1964).

[34] D. Eisenbud, *Commutative Algebra: with a view toward algebraic geometry*, Springer-Verlag, New York (1985).

[35] M. Farber, *Topological complexity of motion planning*, arXiv preprint, arXiv:math/0111197v1 (2001).

[36] R. H. Fox, *On the Lusternik-Schnirelmann category*, Ann. of Math **42**, 333–370 (1941).

[37] R. Ghrist, *Barcodes: the persistent topology of data,* Bulletin of the American Mathematical Society **45**(1), 61–75 (2008).

[38] R. Godement, *Topologie algébrique et théorie des faisceaux*, Hermann Paris (1958).

[39] D. Govc, *On the Definition of Homological Critical Value*, J. Homotopy Relat. Struct. **11** (1), 143–151 (2016).

[40] D. Govc, *Unimodal Category*, in preparation.

[41] D. Govc, P. Škraba, *An Approximate Nerve Theorem*, J. Foundations of Comput. Math., in press.

[42] A. Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge (2002).

[43] L. Hickok, J. Villatoro, X. Wang, *Unimodal Category of 2-Dimensional Distributions*, URL: http://www.math.uiuc.edu/ xwang105/unimodal.pdf.

[44] Z. Kadelburg, D. Đukić, M. Lukić, I. Matić, *Inequalities of Karamata, Schur and Muirhead, and Some Applications*, The Teaching of Mathematics **8**, 31–45 (2005).

[45] I. Kakiuchi, *Unimodality conditions of the distribution of a mixture of two distributions*, Math. Sem. Notes Kobe Univ. **9**, 315–325 (1981).

[46] J. Kemperman, *Mixtures with a limited number of modal intervals*, Ann. Statist. **19**, 2120–2144 (1991).

[47] M. Lesnick, *The theory of the interleaving distance on multidimensional persistence modules*, Foundations of Computational Mathematics, **15**(3), 613–650 (2015).

[48] D. Lipsky, P. Škraba, M. Vejdemo-Johansson, *A spectral sequence for parallelized persistence*, arXiv preprint, arXiv:1112.1245 (2011).

[49] R. Maehara, *The Jordan Curve Theorem via the Brouwer Fixed Point Theorem*, Amer. Math. Month. **91**, 641–643 (1984).

[50] J. McCleary, *A user's guide to spectral sequences*, Cambridge University Press, Cambridge (2001).

[51] J. Milnor, *Morse Theory*, Princeton University Press, Princeton (1963).

[52] C. Nastasescu, F. Van Oystaeyen, *Graded and filtered rings and modules*, North-Holland Publishing Company, Netherlands (1979).

[53] L. Nicolaescu, *An Invitation to Morse Theory*, 2^{nd} edition, Springer Science & Business Media, New York (2011).

[54] P. Niyogi, S. Smale, S. Weinberger, *Finding the homology of submanifolds with high confidence from random samples*, Discrete and Computational Geometry, **39**(1-3), 419–441 (2008).

[55] A. Patel, *Semicontinuity of persistence diagrams*, arXiv preprint, arXiv:1601.03107 (2016).

[56] C. Robertson, J. Fryer, *Some descriptive properties of normal mixtures*, Skand. Aktuarietidskr. **69**, 137–146 (1969).

[57] J. Rotman, *An introduction to homological algebra*, Springer, New York (2008).

[58] H. L. Royden, P. M. Fitzpatrick, *Real Analysis*, 4^{th} edition, Prentice Hall, USA (2010).

[59] M. Scolamiero, W. Chachólski, A. Lundman, R. Ramanujam, S. Öberg, *Multidimensional persistence and noise*, Foundations of Computational Mathematics, 1–40 (2016). URL: https://doi.org/10.1007/s10208-016-9323-y

[60] D. Sheehy, *A multicover nerve for geometric inference*, Proceedings of the Canadian Conference of Computational Geometry, 309–314 (2012).

[61] D. Sheehy, *Linear-size approximations to the Vietoris–Rips filtration*, Discrete and Computational Geometry, **49**(4), 778–796 (2013).

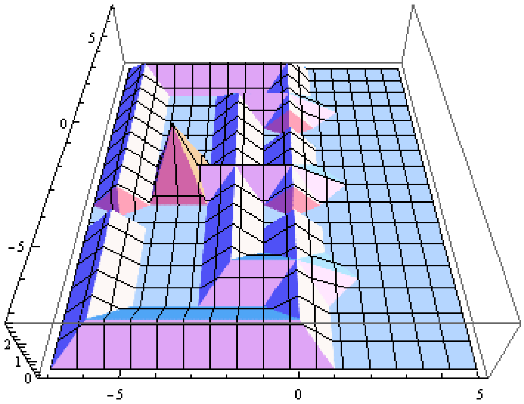
[62] M. Vejdemo-Johansson, *Interleaved equivalence of categories of persistence modules*, arXiv preprint, arXiv:1210.7913 (2012).

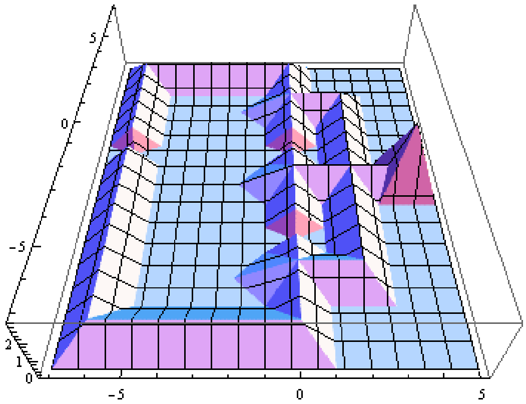
[63] S. Weinberger, *What is... persistent homology?*, Notices of the American Mathematical Society, **58**(1), 36–39 (2011).

[64] A. Zomorodian, G. Carlsson, *Computing persistent homology*, Discrete and Computational Geometry, **33**(2), 249–274 (2005).

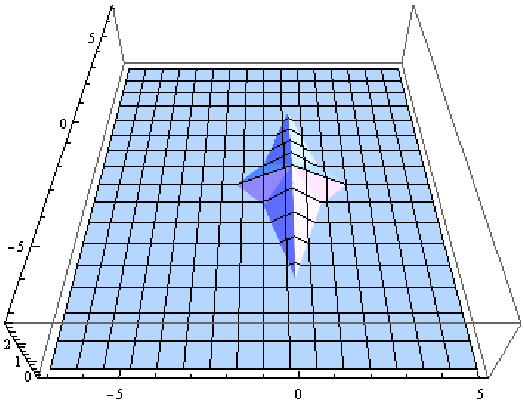
A Graphs

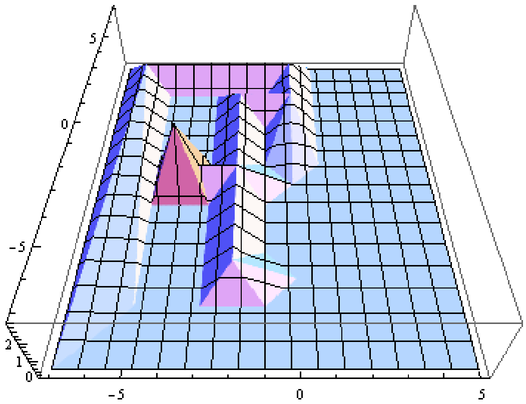
The following are the graphs of $u_1, u_2 : \mathbb{R}^2 \to [0, \infty)$ from Proposition 4.16:

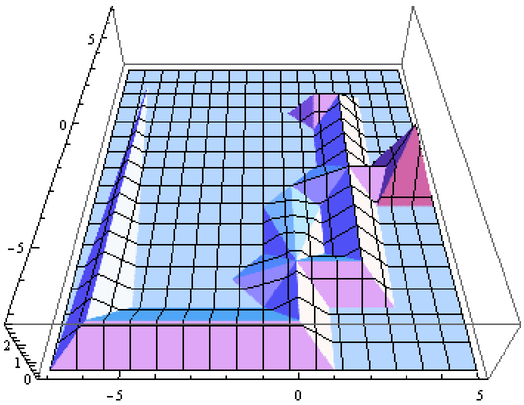




These are the graphs of $u_1, u_2, u_3 : \mathbb{R}^2 \to [0, \infty)$ from Proposition 4.18:







B Algorithm for the Circle

An algorithm computing $\mathbf{ucat}(f)$ for $f : \mathbb{R} \to [0, \infty)$ with finitely many critical points is given in [4]. We do not reproduce it here, but we do note that Theorem 3.4 provides a generalization of it. We do describe an algorithm to compute $\mathbf{ucat}(f)$ for $f : S^1 \to [0, \infty)$ in the case of finitely many critical points. Although this is possible, as the proofs of the theorems leading to the algorithm are constructive in nature, we do not compute the explicit decomposition.

Algorithm 1 Computing **ucat** for the Circle

Require: $f : S^1 \to [0, \infty)$ without zeros and with critical points $x_1, \ldots, x_k \in S^1$.

  for $i = 1, \ldots, k$ **do**

$$f_i(t) := \begin{cases} (t+1)f(x_i); & t \in [-1, 0], \\ f(x_i e^{2\pi i t}); & t \in [0, 1], \\ (2-t)f(x_i); & t \in [1, 2], \\ 0; & t \notin [-1, 2]. \end{cases}$$

  for $i = 1, \ldots, k$ **do**

    $j = 0$

    $t_0 = -\infty$

    while $t_j < 1$ **do**

      $j := j + 1$

      $t_j = \min\{(t_{j-1}, t) \mid V^-(f_i; (t_{j-1}, t)) > f_i(t_{j-1})\}$

    $\alpha_i := j - 1$

  return $\min_{1 \le i \le k} \alpha_i$.

C Convergence of the Mayer-Vietoris spectral sequence

The aim of this Appendix is to briefly describe the basic idea of the proof of Theorem 2.16. Let $(M, \partial^0, \partial^1)$ be the double complex associated to a filtered cover of a filtered simplicial complex, i.e. a pair $(X,\mathcal{U})$, and (E^r, d^r) the spectral sequence associated to this double complex, as defined in Section 2.6. As mentioned there, the spectral sequence associated to a double complex $(M, \partial^0, \partial^1)$ is just a tool to compute the homology of the associated total complex $(\mathrm{Tot}(M), D)$, namely (E^r, d^r) will converge to $\mathsf{H}_*(\mathrm{Tot}(M))$. This standard fact can be established by a series of elementary but tedious computations, so we do not replicate the proof here, see, for instance, [57]. To prove Theorem 2.16, it is therefore sufficient to show that the homology of the total complex $\mathrm{Tot}(M)$ is isomorphic to the homology of (X, f).

In fact, the double complex $(M, \partial^0, \partial^1)$ has a geometric counterpart, namely, the *filtered Mayer-Vietoris blowup complex* $X_{\mathcal{U}}$ associated to $(X,\mathcal{U})$. The total complex $\mathrm{Tot}(M)$ arises as the chain complex associated to $X_{\mathcal{U}}$ and M arises from a filtration on $\mathrm{Tot}(M)$ which in turn is induced by a natural filtration of $X_{\mathcal{U}}$.

So far, we have been working mostly with filtered simplicial complexes; however, in the case at hand, it is slightly more convenient to work with filtered CW complexes and cellular homology. Any filtered simplicial complex X gives rise to a filtered CW complex X^{CW} whose cellular homology is isomorphic to the simplicial homology of X. In fact, the corresponding chain complexes are isomorphic. Each simplex σ in X is assigned a cell e_σ in X^{CW}. The cartesian product $X_1^{\mathrm{CW}} \times X_2^{\mathrm{CW}}$ of two such complexes again has the structure of a CW complex whose cells are given as $e_{\sigma_1} \times e_{\sigma_2}$ for each pair of simplices σ_1 in X_1 and σ_2 in X_2. The blowup complex is a subcomplex of such a product.

Definition. The *filtered Mayer-Vietoris blowup complex* associated to $(X,\mathcal{U})$ is the filtered CW complex $(X_{\mathcal{U}}, \mathcal{F}_{\mathcal{U}})$, where $X_{\mathcal{U}} \leq X \times \mathcal{N}$ is given by

$$X_{\mathcal{U}} = \bigcup_{\sigma \in U_I} e_\sigma \times e_I$$

and the filtration $\mathcal{F}_{\mathcal{U}} = (X_{\mathcal{U}}^j)_{j\in\mathbb{Z}}$ is given by

$$X_{\mathcal{U}}^j = \bigcup_{\sigma \in U_I^j} e_\sigma \times e_I.$$

Let $(\mathsf{C}_*^X, \partial^X)$ be the (persistent) cellular chain complex associated to X^{CW} and let $(\mathsf{C}_*^{\mathcal{N}}, \partial^{\mathcal{N}})$ be the cellular chain complex associated to $\mathcal{N}^{\mathrm{CW}}$. Let (C_*, ∂) be the cellular chain complex associated to the blowup complex $X_{\mathcal{U}}$. Explicitly, each C_n is the free $\mathbb{k}[t]$-module, generated by the graded set of all cells $e_\sigma \times e_I$ with $\dim\sigma + \dim I = n$, where the grading is given by $\deg(e_\sigma \times e_I)$, i.e. the birth times of the cells in the filtration of the blowup complex. Since the blowup complex is a subcomplex of $X^{\mathrm{CW}} \times \mathcal{N}^{\mathrm{CW}}$, the boundary homomorphisms ∂_n are simply restrictions of the boundary homomorphisms of the chain complex associated to this product. These satisfy the following relation:

$$\partial_n(e_\sigma \times e_I) = \partial_n^0(e_\sigma) \times e_I + (-1)^{\dim\sigma} e_\sigma \times \partial_n^1(e_I).$$

Taking into account the isomorphisms between $(\mathsf{C}^X_*, \partial^X)$ and $(\mathsf{C}^{\mathcal{N}}_*, \partial^{\mathcal{N}})$ and the corresponding simplicial chain complexes, it follows that $(\mathsf{C}_*, \partial) \cong (\mathrm{Tot}_*(M), D)$. For comparison, here is the boundary formula for the latter chain complex, written out in full. If (σ, I) is a pair with $\dim \sigma = q$ and $\dim I = p$ such that $p+q=n$, we have

$$D_n(\sigma, I) = \sum_{k=0}^{q} (-1)^k t^{\deg(\sigma,I)-\deg(\sigma_k,I)}(\sigma_k, I) + (-1)^q \sum_{l=0}^{p} (-1)^l t^{\deg(\sigma,I)-\deg(\sigma,I_l)}(\sigma, I_l).$$

This also explains the grading from which the double complex structure of M arises. Namely for each pair p, q with $p+q=n$, let $N_{p,q} \leq \mathsf{C}_n$ be the submodule freely generated by all cells $e_\sigma \times e_I$ with $\dim \sigma = q$ and $\dim I = p$. Then, we have $\mathsf{C}_n = \bigoplus_{p+q=n} N_{p,q}$ and $\partial^X \times \mathrm{id}$ and $\mathrm{id} \times \partial^{\mathcal{N}}$ respect this grading. The aforementioned isomorphism of $(\mathsf{C}_*, \partial) \cong (\mathrm{Tot}_*(M), D)$ isomorphically maps the double complex structure of N into that of M. Therefore, the homology of the total complex $(\mathrm{Tot}(M), D)$ is precisely the (persistent) cellular homology of the blowup complex. In other words, we have:

Proposition C.1. *The homology of the total complex is isomorphic to the homology of the filtered blowup complex:*

$$\mathsf{H}_*(\mathrm{Tot}(M), D) \cong \mathsf{H}^{\mathrm{CW}}_*(X_{\mathcal{U}}, \mathcal{F}_{\mathcal{U}}).$$

It only remains to check that $\mathsf{H}^{\mathrm{CW}}_*(X_{\mathcal{U}}, \mathcal{F}_{\mathcal{U}}) \cong \mathsf{H}_*(X, \mathcal{F})$. To see this, it suffices to construct a homotopy equivalence of the two spaces, in the filtered sense. Let $\pi : X_{\mathcal{U}} \to X^{\mathrm{CW}}$ be the natural projection (to the first component) and let $\pi^j : X^j_{\mathcal{U}} \to (X^j)^{\mathrm{CW}}$ be the appropriate restriction. It is a standard fact [42, Proposition 4G.2] that these projections are homotopy equivalences. Finally, these maps obviously also respect the filtration, i.e. for each $j_1 \leq j_2$, the diagram

$$\begin{tikzcd} X^{j_1}_{\mathcal{U}} \arrow[r] \arrow[d, "\pi^{j_1}"'] & X^{j_2}_{\mathcal{U}} \arrow[d, "\pi^{j_2}"] \\ (X^{j_1})^{\mathrm{CW}} \arrow[r] & (X^{j_2})^{\mathrm{CW}} \end{tikzcd}$$

commutes, because the projections simply forget the second component, whereas the information from the first component remains unchanged. Therefore, as claimed in Theorem 2.16, the Mayer-Vietoris spectral sequence of $(X, \mathcal{U})$ converges to the cellular persistent homology of X^{CW} and therefore to the simplicial persistent homology of X.

D Slovenski povzetek

Za uspešno delovanje znanosti je pomembno, da znamo eksperimentalno zbrane podatke statistično analizirati. Pri tem naletimo na različne verjetnostne porazdelitve, med katerimi je še posebej pomembna in dobro razumljena normalna porazdelitev. Ena od opaznejših značilnosti te porazdelitve je unimodalnost, ki pomeni, da ima njena gostota porazdelitve enolično določen lokalni maksimum. Večina drugih porazdelitev, ki so znane iz klasičnega verjetnostnega računa, je prav tako unimodalnih. Po drugi strani v eksperimentalno izmerjenih porazdelitvah dostikrat opazimo porazdelitve z več kot enim modusom, pri katerih se podatki ne zgoščajo okrog ene same vrednosti, pač pa imajo več opaznih zgostitev. Razumevanje tega pojava je še posebej pomembno, saj običajno nakazuje na prisotnost več kot enega pomembnega vpliva na vrednosti podatkov. Kot nekoliko poenostavljen primer lahko omenimo zgoščevanje prometa. Ta je bolj zgoščen v jutranjih in popoldanskih urah, vpliva, ki pojasnjujeta ti zgostitvi, pa sta dejstvi, da se zjutraj ljudje vozijo v službo, popoldne pa se vračajo domov.

Posebej pomembno je torej znati iz porazdelitve, ki opisuje neki pojav, ki nas zanima, določiti minimalno število vplivov, ki ta pojav pojasnjujejo. Z drugimi besedami, dano porazdelitev z več kot enim modusom bi radi razcepili na unimodalne sumande. V statistiki je ta problem podrobno študiran v posebnem primeru, ko so posamezni vplivi normalno porazdeljeni [7, 33, 56], in do neke mere tudi za splošnejše porazdelitve [45, 46]. Zanimiv je tudi pojav fantomskih modusov (ang. *ghost modes*), opisan v [28]: mešanica n izotropnih normalnih porazdelitev ima lahko več kot n modusov. V praksi so porazdelitve posameznih vplivov lahko različne, poleg tega pa morda podatki sploh niso numerični in zato vsakršen analitičen opis porazdelitve izhaja iz umetno izbranega koordinatnega sistema, ne pa iz problema samega. Kljub temu imamo na podatkih običajno naraven pojem bližine oziroma podobnosti. Baryshnikov in Ghrist sta zato leta 2007 problem topološko abstrahirala in vpeljala pojem unimodalne kategorije, ki je natančneje opisan v [4]. V tem kontekstu namesto gostote verjetnosti $\mathbb{R}^n \to [0, \infty)$ študiramo poljubno funkcijo $f : X \to [0, \infty)$ na topološkem prostoru X. Pravimo, da je funkcija $u : X \to [0, \infty)$ unimodalna, če obstaja število $M > 0$, da so nadnivojnice $u^{-1}[a, \infty)$ kontraktibilne za $a \in (0, M]$ in prazne za $a > M$. Unimodalna kategorija funkcije $f : X \to [0, \infty)$ je najmanjše naravno število k, za katerega obstajajo unimodalne funkcije $u_1, \ldots, u_k : X \to [0, \infty)$, da velja $f = \sum_{i=1}^{k} u_i$ (kjer je vsota definirana po točkah). V tem primeru pišemo $k = \mathbf{ucat}(f)$. Unimodalna kategorija je torej spodnja meja za število sumandov, ne glede na to, kakšne unimodalne porazdelitve imajo posamezni vplivi, ki nas zanimajo. Pojem lahko naravno posplošimo, če namesto razcepov funkcije f na vsote unimodalnih funkcij opazujemo razcepe na ℓ^p-kombinacije unimodalnih funkcij, $p \in (0, \infty]$, kar nam da pojem unimodalne p-kategorije $\mathbf{ucat}^p(f)$. Baryshnikov in Ghrist [4] kot naravnega kandidata za primer $p = 0$ predlagata $\mathbf{gcat}(\text{supp}(f))$, geometrijsko kategorijo nosilca funkcije f. Unimodalno p-kategorijo funkcije f torej lahko razumemo tudi kot nekakšno deformacijo geometrijske kategorije njenega nosilca. (Geometrijska kategorija je različica slavne Lusternik-Schnirelmannove kategorije [24], obe pa merita minimalno število "enostavnih kosov", na katere je mogoče razcepiti prostor.) Drugače rečeno, unimodalno kategorijo lahko razumemo kot dvig geometrijske kategorije, ki je topološka invarianta prostorov, do topološke invariante funkcij na teh prostorih. Tako dviganje invariant se je v zadnjem času izkazalo kot posebej uspešno, primer je npr. Eulerjev račun [25] (ang. *Euler calculus*), kjer Eulerjevo karakteristiko razumemo kot mero, integral konstruktibilne funkcije po tej meri pa imenujemo Eulerjev integral. Tudi zelo uspešen pojem vztrajne homologije [30] lahko razumemo kot podoben dvig; v najosnovnejši različici je to dvig homologije, ki je invarianta prostorov (in preslikav med njimi), do invariante filtracij teh prostorov.

O unimodalni kategoriji ni veliko znanega, niti v osnovnem primeru $X = \mathbb{R}^n$, ki je najbolj zanimiv s stališča statistike. Za primer $n = 1$ sta Baryshnikov in Ghrist zasnovala enostaven algoritem [4], s pomočjo katerega je mogoče unimodalno kategorijo učinkovito izračunati za vse funkcije, ki imajo končno mnogo kritičnih točk. Primer $n = 2$ v članku obravnavata le delno, zaključita pa z domnevo o monotonosti, za katero menita, da bo ključnega pomena pri izračunu natančnejših mej za unimodalno kategorijo v višjih dimenzijah – domnevata namreč, da za fiksno funkcijo $f : X \to [0, \infty)$ in $0 < p < q \leq \infty$ vselej velja $\mathbf{ucat}^p(f) \leq \mathbf{ucat}^q(f)$. Primer $n = 2$ nekoliko podrobneje obravnavajo Hickok, Villatoro in Wang v [43], in sicer se omejijo na Morsove distribucije na ravnini, katerih Morse-Smaleov graf je drevo. V tem primeru unimodalno kategorijo skoraj popolnoma karakterizirajo.

Disertacija sestoji iz dveh konceptualnih delov. Prvi del je večinoma osredotočen na domnevo o monotonosti, medtem ko drugi del zadeva aproksimativni izrek o živcu [41].

V prvem delu najprej pokažemo, da lahko dekompozicijo, ki sledi iz algoritma za $X = \mathbb{R}$, posplošimo na poljubne funkcije $f : \mathbb{R} \to [0, \infty)$, in sicer kot različico Jordanove dekompozicije [58] za funkcije z omejeno variacijo (vsako tako funkcijo lahko zapišemo kot razliko dveh monotono naraščajočih funkcij). Potem te rezultate posplošimo na primer $X = S^1$. Dobljena posplošitev je primerna za študij domneve o monotonosti, poleg tega pa porodi preprost algoritem za izračun unimodalne kategorije funkcije $f : S^1 \to [0, \infty)$, če ima ta le končno mnogo kritičnih točk. Rezultati so dovolj splošni, da nam omogočijo dokaz domneve o monotonosti za poljubne funkcije na $X = \mathbb{R}$ in $X = S^1$. Nato pokažemo, da monotonost ne drži za nekatere splošnejše prostore, in sicer konstruiramo dva protiprimera na grafih, ki pokažeta, da domneva za večino grafov ne drži. Pomembneje, konstruiramo tudi dva protiprimera na prostoru $X = \mathbb{R}^2$. Nazadnje pokažemo, da domneva kljub temu velja za $X = \mathbb{R}^2$ v primeru funkcij, katerih Morse-Smaleov graf je drevo.

Drugi del disertacije, zgoščen v razdelku 6, je povezan z bolj razvitimi področji računske topologije, kot je npr. vztrajna homologija. Zato zanj potrebujemo nekoliko več ozadja in si zasluži poseben uvod.

Da bi motivirali ta del disertacije, opazimo, da je vsak pojem kategorije povezan z določenim tipom pokritja danega prostora. Npr. Lusternik-Schnirelmannova kategorija zadeva kategorična pokritja, tj. pokritja, kjer so elementi kontraktibilni znotraj prostora, geometrijska kategorija pa zadeva pokritja s kontraktibilnimi množicami. Unimodalna kategorija je prav tako povezana s konceptom pokritja, le da je povezava zapletenejša. Najenostavnejši primer je primer unimodalne ∞-kategorije, kjer funkcijo $f : X \to [0, \infty)$ razcepimo kot $f = \min_{1 \leq i \leq n} u_i$, pri čemer so funkcije $u_i : X \to [0, \infty)$ unimodalne. To pomeni, da na vsakem nivoju $c > 0$ velja $f^{-1}[c, \infty) = \bigcup_{i=1}^n u_i^{-1}[c, \infty)$, tj. da na vsakem nivoju nadnivojnice unimodalnih funkcij v dekompoziciji tvorijo pokritje ustreznih nadnivojnic originalne funkcije in da to pokritje sestoji iz kontraktibilnih množic.

Sklepanje na globalne lastnosti prostora iz lokalnih lastnosti, npr. na homologijo prostora iz homologije elementov primernega pokritja prostora, je pogosta tema v algebraični topologiji. Zato lahko pričakujemo, da se bodo take metode izkazale kot plodovite tudi v kontekstu unimodalne kategorije, ko se to področje dovolj razvije.

Z vidika vztrajne homologije funkcija $f : X \to [0, \infty)$ podaja nadnivojsko filtracijo prostora X in unimodalna ∞-dekompozicija $f = \min_{1 \leq i \leq n} u_i$ poraja filtrirano pokritje tega prostora, pri čemer je vztrajna homologija elementov tega pokritja trivialna. Zaenkrat ni jasno, kakšna je splošna povezava med vztrajnostjo in unimodalno kategorijo. V primeru, ko je pokritje posebej enostavno, in sicer dobro pokritje, pa obstajajo klasični rezultati za nefiltrirani primer. Z uporabo spektralnih zaporedij smo te rezultate posplošili v kontekst vztrajne homologije. Zdi se verjetno, da je mogoče podobne metode uporabiti splošneje za delo s filtriranimi pokritji, katerih elementi so na vsakem nivoju kontraktibilni, kot v primeru unimodalne ∞-kategorije.

Znano je tudi, da je za funkcije $f : \mathbb{R} \to [0, \infty)$ koncept vztrajnosti povezan s konceptom totalne variacije [6]. Če upoštevamo prvi del disertacije, to pomeni, da je v enodimenzionalnem primeru pojem **ucat** tesno povezan s pojmom vztrajnosti.

D.1 Definicije in predhodno znani rezultati

Doktorska disertacija se začne z omembo nekaterih splošno znanih rezultatov, ki jih potrebujemo za nadaljnje delo. Vpeljan je pojem unimodalne kategorije in podana primerjava z nekaterimi drugimi klasičnimi pojmi kategorije. Medtem ko npr. geometrijska kategorija podaja minimalno število elementov v pokritju danega prostora s kontraktibilnimi podmnožicami, pojem unimodalne kategorije zadeva razcep funkcij. Pojem, analogen kontraktibilnosti, ki ima smisel za funkcije, je unimodalnost.

Definicija. Zvezna funkcija $u : X \to [0, \infty)$ je *unimodalna*, če obstaja $M > 0$, tako da so nadnivojnice $u^{-1}[c, \infty)$ kontraktibilne za $0 < c \leq M$ in prazne za $c > M$.

To nas pripelje direktno do definicije unimodalne kategorije, kakršno sta podala Baryshnikov in Ghrist v članku [4].

Definicija. Naj bo $p \in (0, \infty)$. *Unimodalna p-kategorija* $\mathbf{ucat}^p(f)$ funkcije $f : X \to [0, \infty)$ je minimalno število n unimodalnih funkcij $u_1, \ldots, u_n : X \to [0, \infty)$, tako da po točkah velja $f = (\sum_{i=1}^n u_i^p)^{\frac{1}{p}}$. Podobno je *unimodalna ∞-kategorija* $\mathbf{ucat}^\infty(f)$ funkcije $f : X \to [0, \infty)$ minimalno število n unimodalnih funkcij $u_1, \ldots, u_n : X \to [0, \infty)$, tako da po točkah velja $f = \max_{1 \leq i \leq n} u_i$. Namesto $\mathbf{ucat}^1(f)$ običajno pišemo $\mathbf{ucat}(f)$.

Baryshnikov in Ghrist pojem še nekoliko posplošita in poljubni normi ν na prostoru $\mathbb{R}^{(\mathbb{N})}$ priredita ustrezen pojem unimodalne ν-kategorije. Pomembna lastnost unimodalne kategorije je naslednje dejstvo [4, Lemma 9], iz katerega lahko razberemo, da metode za računanje unimodalne 1-kategorije lahko uporabimo tudi za računanje unimodalne p-kategorije za $p \in (0, \infty)$.

Lema D.1 ([4], Lemma 9)**.** *Če je $f : X \to [0, \infty)$ poljubna zvezna funkcija, velja*

$$\mathbf{ucat}^p(f) = \mathbf{ucat}(f^p).$$

Pojem unimodalne kategorije je zanimiv predvsem na topoloških prostorih, ki niso preveč patološki, recimo na mnogoterostih ali CW kompleksih.

Pri izračunu unimodalne kategorije za splošne zvezne funkcije $f : \mathbb{R} \to [0, \infty)$ je koristno poznavanje klasičnega pojma totalne variacije (glej npr. [3, Chapter 6]) in nekaterih sorodnih pojmov.

Definicija. *Totalna variacija* funkcije $f : \mathbb{R} \to \mathbb{R}$ na intervalu $[a, b]$ je definirana s formulo

$$V(f; [a, b]) = \sup \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|,$$

kjer supremum teče po vseh delitvah $a = x_0 < x_1 < \ldots < x_n = b$ intervala $[a, b]$. Podobno definiramo *pozitivno variacijo* funkcije f na intervalu $[a, b]$ kot

$$V^+(f; [a, b]) = \sup \sum_{i=1}^{n} \max\{0, f(x_i) - f(x_{i-1})\},$$

in *negativno variacijo* funkcije f na intervalu $[a, b]$ kot

$$V^-(f; [a, b]) = \sup \sum_{i=1}^{n} \max\{0, f(x_{i-1}) - f(x_i)\}.$$

Ti pojmi imajo naslednje lastnosti.

Izrek D.2. *Naj V^* označuje V, V^+ ali V^- in naj bo $f : [a, b] \to \mathbb{R}$. Potem velja:*

- *Če je f naraščajoča na $[a, b]$, potem je $V(f; [a, b]) = V^+(f; [a, b]) = f(b) - f(a)$ in $V^-(f; [a, b]) = 0$.*
- *Če je f naraščajoča na $[a, b]$, potem je $V(f; [a, b]) = V^-(f; [a, b]) = f(a) - f(b)$ in $V^+(f; [a, b]) = 0$.*
- *Če je $f = \sum_{i=1}^{n} f_i$, potem je $V^*(f; [a, b]) \leq \sum_{i=1}^{n} V^*(f_i; [a, b])$.*
- *Če je $a = x_0 < \ldots < x_n = b$, potem je $V^*(f; [a, b]) = \sum_{i=1}^{n} V^*(f; [x_{i-1}, x_i])$.*
- *$V^+(f; [a, b]) + V^-(f; [a, b]) = V(f; [a, b])$ in $V^+(f; [a, b]) - V^-(f; [a, b]) = f(b) - f(a)$.*

Vse tri pojme lahko razširimo tudi do pozitivnih Borelovih mer na $\mathbb{R}$, smisel pa imajo tudi za funkcije $f : J \to \mathbb{R}$, kjer je $J \subseteq \mathbb{R}$ poljuben interval. Najzanimivejše za nas so funkcije z omejeno variacijo, tj. take, za katere velja $V(f; J) < \infty$. Take so npr. monotone funkcije na končnih intervalih, pa tudi omejene monotone funkcije. Od tod sledi, da imajo tudi unimodalne funkcije $u : \mathbb{R} \to [0, \infty)$ omejeno variacijo.

Če definiramo npr. $g(x) = V^+(f; J \cap (-\infty, x))$ in $h(x) = V^-(f; J \cap (-\infty, x))$, lahko dokažemo, da za funkcije z omejeno variacijo velja Jordanov izrek o dekompoziciji.

Izrek D.3. *Naj bo $f : J \to \mathbb{R}$ funkcija z omejeno variacijo. Potem lahko f izrazimo kot razliko $f = g - h$ dveh monotono naraščajočih funkcij $g, h : J \to \mathbb{R}$.*

Za obravnavo funkcij na $\mathbb{R}^n$, kjer je $n > 1$, je koristno poznati tudi nekaj osnov Morsove teorije.

Izrek D.4 ([51]). *Naj bo $f : M \to \mathbb{R}$ Morsova funkcija na mnogoterosti M, $p \in M$ kritična točka indeksa i in $f(p) = a$ ustrezna kritična vrednost. Za vsak $x \in \mathbb{R}$ naj bo $M_x = f^{-1}(-\infty, x]$. Če $[a-\epsilon, a+\epsilon]$ ne vsebuje nobene druge kritične vrednosti funkcije f, potem podnivojnico $M_{a+\epsilon}$ dobimo (do homotopije natančno) iz podnivojnice $M_{a-\epsilon}$ z lepljenjem i-ročaja.*

Za naše namene je pomembno dejstvo, da lepljenje i-ročaja ali uniči homološki razred v H_{i-1} ali pa ustvari homološki razred v H_i.

Za obravnavo aproksimativnega izreka o živcu prav tako potrebujemo nekaj predhodnih definicij in znanih izrekov s področja vztrajne homologije. Delamo predvsem z $\mathbb{Z}$-filtriranimi simplicialnimi kompleksi in indeksiranimi $\mathbb{Z}$-filtriranimi pokritji.

Definicija. Naj bo $J \subseteq \mathbb{R}$. *J-filtriran simplicialni kompleks* je par $(X, \mathcal{F})$, kjer je X abstrakten simplicialni kompleks in $\mathcal{F} = (X^j)_{j \in J}$ družina podkompleksov, tako da iz $j_1 \leq j_2$ sledi $X^{j_1} \subseteq X^{j_2}$, $X^{-\infty} := \bigcap_{j \in J} X^j = \emptyset$ in $X^{\infty} := \bigcup_{j \in J} X^j = X$.

J-filtrirano pokritje s podkompleksi J-filtriranega simplicialnega kompleksa $(X, \mathcal{F})$ je indeksirana družina $\mathcal{U} = (U_i, \mathcal{F}_i)_{i \in \Lambda}$, kjer je vsak U_i podkompleks v X in je $\mathcal{F}_i$ filtracija tega podkompleksa, tako da filtracije $\mathcal{F}$ in $\mathcal{F}_i$ zadoščajo kompatibilnostnemu pogoju, in sicer, da velja $X^j = \bigcup_{i \in \Lambda} U_i^j$ za vse $j \in J$. Pripomnimo, da ima v primeru, da je $I \subseteq \Lambda$, presek $U_I := \bigcap_{i \in I} U_i$ naravno filtracijo $\mathcal{F}_I$, in sicer $U_I^j := \bigcap_{i \in I} U_i^j$.

Filtracijo lahko podamo tudi s podnivojnicami funkcije $f : X \to \mathbb{Z}$. Če je na danem simplicialnem kompleksu podana filtracija, ji lahko na poljubnem pokritju tega simplicialnega kompleksa s podkompleksi priredimo inducirano filtracijo.

Smiseln je tudi pojem filtriranega triangulabilnega prostora. V tem primeru filtracijo na prostoru Y običajno podamo s podnivojnicami zvezne funkcije $f : Y \to \mathbb{R}$.

Spomnimo se še standardne konstrukcije živca danega pokritja:

Definicija. Danemu pokritju $(U_i)_{i \in \Lambda}$ prostora X priredimo *živec* $\mathcal{N}$, in sicer kot množico vseh končnih podmnožic indeksne množice Λ, ki ustrezajo naslednjemu pogoju: končna množica $I \subseteq \Lambda$ je element $\mathcal{N}$ natanko tedaj, ko velja

$$U_I := \bigcap_{i \in I} U_i \neq \emptyset.$$

Če I pripada $\mathcal{N}$, potem to velja tudi za vse podmnožice množice I. Torej je $\mathcal{N}$ abstrakten simplicialni kompleks.

Rezultati v razdelku 6 so zapisani v jeziku $\Bbbk[t]$-modulov (glej npr. [34], kjer so ti pojmi opisani v nekoliko drugačnem jeziku). Tu $\Bbbk$ označuje polje koeficientov, $\Bbbk[t]$ pa stopničasti kolobar polinomov v eni spremenljivki s koeficienti v $\Bbbk$. Pojem $\Bbbk[t]$-modula razumemo v primernem stopničastem smislu, del strukture $\Bbbk[t]$-modula M je torej tudi razcep $M = \bigoplus_{j \in \mathbb{Z}} M^j$, kjer so M^j ustrezni $\Bbbk$-podmoduli v M, delovanje kolobarja $\Bbbk[t]$ pa upošteva ta razcep.

Če je $\varepsilon \in \mathbb{N}_0$, vpeljemo pojem *$\varepsilon$-morfizma* $\Bbbk[t]$-*modulov* M in N, to je homomorfizem $f : M \to N$, ki upošteva omenjeni razcep modulov M in N, tj. za vse $j \in \mathbb{Z}$ velja $f(M^j) \subseteq N^{j+\varepsilon}$. Namesto izraza 0-morfizem raje uporabljamo izraz *morfizem.*

To nam omogoča vpeljavo pojma prepletanja, ki je formalizacija intuitivne ideje "aproksimativnega izomorfizma", tj. para morfizmov, ki sta si skoraj inverzna, v smislu, da je njun kompozitum v poljubnem vrstnem redu enak $\mathrm{id}_{2\varepsilon}$. Tu je $\mathrm{id}_{2\varepsilon}$ 2ε-morfizem, ki je karseda blizu identiteti, in sicer $\mathrm{id}_{2\varepsilon}(x) = t^{2\varepsilon}x$.

Definicija. Par (ϕ, ψ) ε-morfizmov $\phi : M \to N$ and $\psi : N \to M$ se imenuje *ε-prepletanje* $\Bbbk[t]$-*modulov* M in N, če velja $\phi\psi = \mathrm{id}_{2\varepsilon}$ in $\psi\phi = \mathrm{id}_{2\varepsilon}$. Pojem 0-prepletanja se ujema s pojmom izomorfizma. Če obstaja ε-prepletanje modulov M in N, pravimo, da sta M in N *ε-prepletena* in pišemo $M \overset{\varepsilon}{\sim} N$.

Pojem prepletanja nam omogoči vpeljavo pojma *prepletne razdalje* med moduloma M in N, tj. najmanjša vrednost parametra ε, za katero velja $M \overset{\varepsilon}{\sim} N$. Tako formaliziramo idejo aproksimacije modulov. Module, ki so ε-prepleteni s trivialnim modulom 0, razumemo kot v nekem smislu majhne, in so tako uporabni za modeliranje eksperimentalnih napak v podatkih.

Pri študiju vztrajne homologije običajno vpeljemo še pojem *vztrajnostnega modula*, ki je funktor $F : (I, \leq) \to \mathbf{Vect}$, kjer je $(I, \leq)$ ustrezna delno urejena množica, ki jo razumemo kot kategorijo. V tem pogledu $\Bbbk[t]$-modulom ustrezajo ravno vztrajnostni moduli $F : (\mathbb{Z}, \leq) \to \mathbf{Vect}$. V praksi običajno vztrajnostni modul $F : (\mathbb{Z}, \leq) \to \mathbf{Vect}$ dobimo z diskretizacijo ustreznega modula $G : (\mathbb{R}, \leq) \to \mathbf{Vect}$, ki je idealizacija danega naravnega procesa, ki ga želimo opisati. Vztrajnostni modul slednje oblike npr. lahko dobimo tako, da na podnivojski filtraciji, ki ustreza dani funkciji $f : X \to \mathbb{R}$, uporabimo homološki funktor.

O opisanem procesu diskretizacije si lahko bralec več prebere npr. v [16], [62] in [18].

Da je računanje s $\Bbbk[t]$-moduli karseda učinkovito, nazadnje potrebujemo še pojem spektralnega zaporedja. Spektralna zaporedja običajno dobimo iz dvojnih kompleksov. Opis teh lahko najdemo v [57, Chapter 10] in [50]. Uporabljamo Mayer-Vietorisovo spektralno zaporedje. Nekatere njegove različice so opisane npr. v [11] in [12].

Definicija. *Dvojni kompleks (bikompleks)* $\Bbbk[t]$-*modulov* je trojica $(M, \partial^0, \partial^1)$, kjer je M dvojno stopničast $\Bbbk[t]$-modul in sta $\partial^0 = (\partial^0_{p,q})_{p,q\in\mathbb{Z}}$ in $\partial^1 = (\partial^1_{p,q})_{p,q\in\mathbb{Z}}$ dve družini morfizmov $\partial^0_{p,q} : M_{p,q} \to M_{p,q-1}$ in $\partial^1_{p,q} : M_{p,q} \to M_{p-1,q}$, tako da velja $\partial^0_{p,q-1}\partial^0_{p,q} = 0, \partial^1_{p-1,q}\partial^1_{p,q} = 0$ in $\partial^1_{p,q-1}\partial^0_{p,q} + \partial^0_{p-1,q}\partial^1_{p,q} = 0$ za $p, q \in \mathbb{Z}$.

Dvojnemu kompleksu lahko priredimo *totalni kompleks.*

Definicija. Naj bo M dvojni kompleks. *Totalni kompleks* $(\mathrm{Tot}(M), D)$, prirejen M je verižni kompleks $\Bbbk[t]$-modulov $\mathrm{Tot}_n(M) = \bigoplus_{p+q=n} M_{p,q}$ z robnim operatorjem $D = \partial^0 + \partial^1$.

Spektralna zaporedja so orodje za izračun homologije takega totalnega kompleksa. V praksi je taka situacija precej pogosta, saj npr. iz filtracij topoloških prostorov lahko na naraven način dobimo dvojne komplekse, homologija ustreznih totalnih kompleksov pa se ujema s homologijo originalnega prostora, ki jo želimo izračunati. Prav to se zgodi tudi v našem primeru.

Spektralno zaporedje sestoji iz *strani*, kjer je vsaka stran $E^r, r = 0, 1, \ldots$, diferencialni bistopničasti modul, iz dane strani pa se da izračunati naslednjo, tako da izračunamo njeno homologijo glede na dani diferencialni operator. Diferencialni operator na r-ti strani je oblike

$$d^r : E^r_{p,q} \to E^r_{p-r,q+r-1}.$$

Lahko se zgodi, da obstaja R, tako da so za vse $r > R$ diferenciali z začetkom ali koncem v $E^r_{p,q}$ ničelne preslikave. V tem primeru se (p, q)-ta komponenta stabilizira v smislu, da so vsi moduli $E^r_{p,q}$ za $r > R$ izomorfni. Če se to zgodi za vsak par (p, q), rečemo, da spektralno zaporedje *konvergira* in stabilne module $E^r_{p,q}$ označimo $E^\infty_{p,q}$. Dvojno stopničasti modul s komponentami $E^\infty_{p,q}$ imenujemo ∞-stran spektralnega zaporedja. Če velja $E^N = E^\infty$ za neki končni N, rečemo, da spektralno zaporedje *kolabira* na N-ti strani.

Vsaka naslednja stran spektralnega zaporedja podaja boljšo aproksimacijo homologije totalnega kompleksa. Če spektralno zaporedje bikompleksa M konvergira, zato rečemo, da konvergira k $\mathsf{H}_*(\mathrm{Tot}(M))$. V praksi to pomeni, da $\mathsf{H}_*(\mathrm{Tot}(M))$ lahko rekonstruiramo iz strani E^∞, vendar je pri tem morda treba razrešiti določene razširitvene probleme.

Paru $(X,\mathcal{U})$, kjer je X filtrirani simplicialni kompleks in $\mathcal{U} = (U_i, \mathcal{F}_i)_{i\in\Lambda}$ njegovo filtrirano pokritje, lahko priredimo (komutativni) dvojni kompleks $(E^0, \partial^0, \partial^1)$. Ustrezne module definiramo kot

$$E^0_{p,q} = \bigoplus_{|I|=p+1} \mathsf{C}_q(U_I),$$

robne preslikave $\partial^0_{p,q} : E^0_{p,q} \to E^0_{p,q-1}$ in $\partial^1_{p,q} : E^0_{p,q} \to E^0_{p-1,q}$ pa s predpisoma

$$\partial^0_{p,q}(\sigma, I) = \sum_{k=0}^{q} (-1)^k t^{\deg(\sigma,I)-\deg(\sigma_k,I)}(\sigma_k, I)$$

in

$$\partial^1_{p,q}(\sigma, I) = \sum_{l=0}^{p} (-1)^l t^{\deg(\sigma,I)-\deg(\sigma,I_l)}(\sigma, I_l).$$

Spektralno zaporedje (E^r, d^r), prirejeno temu dvojnemu kompleksu, se imenuje *Mayer-Vietorisovo spektralno zaporedje para* $(X,\mathcal{U})$, pomembno pa je zato, ker zadošča naslednjemu izreku (glej npr. [48, 11, 12, 38]).

Izrek D.5. *Mayer-Vietorisovo spektralno zaporedje para* $(X,\mathcal{U})$ *konvergira k vztrajni homologiji* $\mathsf{H}_*(X)$.

Od tod je mogoče razmeroma enostavno izpeljati Sheehyjev izrek o vztrajnih živcih.

Izrek D.6. *Recimo, da je* X *filtriran simplicialni kompleks in* $\mathcal{U}$ *vztrajno aciklično pokritje za* X. *Potem velja* $\mathsf{H}_*(X) \cong \mathsf{H}_*(\mathcal{N}(\mathcal{U}))$.

Ta rezultat vzamemo za osnovo nadaljnjega dela, aproksimativni izrek o živcu je namreč njegova posplošitev na primer, ko je pokritje le približno aciklično, kar je pogosta situacija v praksi, kjer je nemogoče zagotoviti popolno natančnost v podatkih.

D.2 Računanje unimodalne kategorije

Najosnovnejše vprašanje, ki nas zanima za dano funkcijo $f : X \to [0, \infty)$, je kako izračunati njeno unimodalno kategorijo $\mathbf{ucat}(f)$. V primeru funkcij $f : \mathbb{R} \to [0, \infty)$ s končnim številom kritičnih točk sta nanj odgovorila že Ghrist in Baryshnikov [4, Theorem 11] z uporabo enostavnega algoritma s pometanjem.

V razdelku 3 pokažemo, da ta algoritem lahko razumemo kot posledico Jordanovega izreka o dekompoziciji za funkcije z omejeno variacijo. Ta opazka nam omogoča popolno karakterizacijo unimodalne kategorije za funkcije na $X = \mathbb{R}$. Isto idejo lahko uporabimo za izračun unimodalne kategorije v primeru $X = S^1$.

Najprej opazimo, da je vprašanje zanimivo le za funkcije z omejeno variacijo, saj imajo unimodalne funkcije $u : \mathbb{R} \to [0, \infty)$ omejeno variacijo. (To ne drži za unimodalne funkcije $u : \mathbb{R}^n \to [0, \infty)$, če je $n > 1$.) Pojem intervala izsiljenega maksimuma iz dela [4] lahko nato z uporabo negativne variacije posplošimo na poljubne intervale in poljubne funkcije.

Definicija. Interval (x, y) je **interval izsiljenega maksimuma** (glede na f), če velja

$$V^-(f; (x, y)) > f(x).$$

Rečemo še, da je (x, y) **skoraj interval izsiljenega maksimuma**, če je $(x, y + \delta)$ interval izsiljenega maksimuma za vsak $\delta > 0$. Če ima $f : \mathbb{R} \to [0, \infty)$ kompakten nosilec, definiramo $M(f)$ kot največje možno število paroma disjunktnih odprtih intervalov, ki so intervali izsiljenega maksimuma za f, $\widetilde{M}(f)$ pa je analogno število za skoraj intervale izsiljenega maksimuma.

Izkaže se, da se ti dve števili ujemata. Z uporabo Jordanove dekompozicije lahko dani funkciji $f : \mathbb{R} \to [0, \infty)$ priredimo delitev realne osi na skoraj intervale izsiljenega maksimuma, tem pa priredimo unimodalno dekompozicijo funkcije f in tako posplošimo algoritem iz [4].

Izrek D.7. *Če ima $f : \mathbb{R} \to [0, \infty)$ kompakten nosilec, potem velja*

$$\mathbf{ucat}(f) = M(f).$$

Če je $M(f) = n < \infty$, lahko konstruiramo eksplicitno minimalno unimodalno dekompozicijo $(u_i)_{i=1}^n$ funkcije f po naslednjem postopku. Najprej razširimo f na $\overline{\mathbb{R}} = [-\infty, \infty]$ in definiramo $g, h : \overline{\mathbb{R}} \to [0, \infty)$ s predpisoma

$$g(x) = V^+(f; (-\infty, x]) \qquad \text{in} \qquad h(x) = V^-(f; (-\infty, x]).$$

Rekurzivno definiramo končno zaporedje $(x_i)_{i=0}^n$:

$$\begin{aligned} x_0 &= -\infty, \\ x_i &= \inf\{x \mid V^-(f; (x_{i-1}, x)) > f(x_i - 1)\}, \qquad i = 1, \ldots, n, \\ x_{n+1} &= \infty. \end{aligned}$$

Nazadnje definiramo $u_i : \mathbb{R} \to [0, \infty)$ s predpisom

$$u_i(x) = \begin{cases} 0; & x \le x_{i-1}, \\ g(x) - g(x_{i-1}); & x \in [x_{i-1}, x_i], \\ h(x_{i+1}) - h(x); & x \in [x_i, x_{i+1}], \\ 0; & x \ge x_{i+1}. \end{cases}$$

To nam omogoči tudi popolno karakterizacijo funkcij s končno unimodalno kategorijo ter izračun unimodalne kategorije za funkcije na intervalih $J \subseteq \mathbb{R}$.

Podoben postopek deluje za funkcije $f : S^1 \to [0, \infty)$. Tudi v tem primeru so zanimive le funkcije z omejeno variacijo. Privzamemo pa lahko tudi, da funkcija nima ničel, saj se sicer problem enostavno prevede na primer $X = \mathbb{R}$. Nato vpeljemo nekaj oznak.

Oznake. Če je $a \in S^1$, definiramo $\phi_a^+ : [0,1] \to S^1$ s predpisom $\phi_a^+(t) = a\exp(2\pi i t)$ in $\phi_a^- : [0,1] \to S^1$ s predpisom $\phi_a^-(t) = a\exp(-2\pi i t)$. Vpeljemo $f_a^+ = f \circ \phi_a^+$ in $f_a^- = f \circ \phi_a^-$. Poljubno funkcijo $g : [0,1] \to [0,\infty)$ razširimo do funkcije $\hat{g} : \mathbb{R} \to [0,\infty)$ kot v enačbi (3). Namesto $\widehat{f_a^+}$ včasih pišemo $\hat{f}_a^+$.

Funkciji f priredimo naslednja števila:

$$M_a^+(f) = M(\hat{f}_a^+|_{(-\infty,1]}), \qquad M^+(f) = \min_{a \in S^1} M_a^+(f),$$
$$M_a^-(f) = M(\hat{f}_a^-|_{(-\infty,1]}), \qquad M^-(f) = \min_{a \in S^1} M_a^-(f).$$

Unimodalno kategorijo funkcije $f : S^1 \to [0,\infty)$ je zdaj mogoče karakterizirati takole.

Izrek D.8. *Če funkcija $f : S^1 \to [0,\infty)$ nima ničel, potem velja:*

$$\mathbf{ucat}(f) = \max\{2, M^+(f)\} = \max\{2, M^-(f)\}.$$

Rezultat pomeni, da lahko unimodalno kategorijo funkcije $f : S^1 \to [0,\infty)$ izračunamo s konstrukcijo družine skoraj intervalov izsiljenega maksimuma podobno kot v primeru $f : \mathbb{R} \to [0,\infty)$, če le vemo, v kateri točki krožnice začeti. Izkaže pa se, da je v primeru, ko je kritičnih točk le končno mnogo, dovolj preizkusiti te kritične točke, torej je v tem primeru mogoč popolnoma algoritmičen izračun.

D.3 Domneva o monotonosti

Baryshnikov in Ghrist svoj članek [4] zaključita z naslednjo domnevo.

Domneva. Recimo, da je $f : X \to [0,\infty)$ in $0 < p_1 < p_2 \leq \infty$. Potem velja $\mathbf{ucat}^{p_1}(f) \leq \mathbf{ucat}^{p_2}(f)$.

Z drugimi besedami, domnevata, da je število $\mathbf{ucat}^p(f)$ pri fiksni funkciji f monotona funkcija parametra p. V razdelku 4 dokažemo, da ta domneva drži za $X = \mathbb{R}$ in $X = S^1$ in da odpove, če je X primerno izbran graf, pa tudi v primeru, ko je $X = \mathbb{R}^2$ evklidska ravnina.

V primeru $X = \mathbb{R}$ oziroma $X = S^1$ domnevo dokažemo s pomočjo naslednje leme, ki je posledica Karamatove neenakosti [44].

Lema D.9. *Naj bo $0 < q < 1$ in $x, y, z \geq 0$, tako da velja $\max\{x,z\} \leq y \leq x+z$. Potem je*

$$(x - y + z)^q \leq x^q - y^q + z^q.$$

Od tod lahko izpeljemo, da mora biti v primeru $p_1 < p_2$ vsak interval izsiljenega maksimuma za funkcijo f^{p_1} tudi interval izsiljenega maksimuma za funkcijo f^{p_2}, od tod pa lahko potem dokažemo želeni izrek.

Izrek D.10. *Naj bo $f : \mathbb{R} \to [0, \infty)$ in $0 < p_1 < p_2 \leq \infty$. Potem velja* $\mathbf{ucat}^{p_1}(f) \leq \mathbf{ucat}^{p_2}(f)$.

Kot posledico tega izreka in karakterizacije unimodalne kategorije za funkcije $f : S^1 \to [0, \infty)$ iz razdelka 3 od tod dobimo, da je unimodalna kategorija monotona tudi za funkcije na S^1.

Posledica D.11. *Naj bo $f : S^1 \to [0, \infty)$ in $0 < p_1 < p_2 \leq \infty$. Potem velja* $\mathbf{ucat}^{p_1}(f) \leq \mathbf{ucat}^{p_2}(f)$.

Za splošne prostore pa domneva o monotonosti ne drži. V razdelku 4.2 je podan primer funkcije $f : X \to [0, \infty)$ na grafu X, za katero v nasprotju z domnevo o monotonosti velja $\mathbf{ucat}(f) = 2$ in $\mathbf{ucat}^{\frac{1}{2}}(f) = 3$. Graf je dovolj enostaven, da se da ta protiprimer posplošiti na precej širok razred grafov. Te lastnosti primera dokažemo neposredno, ideja konstrukcije pa je v tem, da lahko s pomočjo ciklov v nadnivojnicah funkcije f do neke mere nadzorujemo vrednosti unimodalnih sumandov.

V razdelku 4.3 konstruiramo nekoliko drugačen protiprimer, in sicer funkcijo $f : X \to [0, \infty)$ na nekem drugem grafu X, za katero velja $\mathbf{ucat}(f) = 3$ in $\mathbf{ucat}^{\infty}(f) = 2$. Konstrukcija protiprimera v tem primeru sloni na dejstvu, da ima cikel dolžine 3 kromatično število 3.

Podobne primere z uporabo podobnih idej nato konstruiramo na ravnini. V razdelku 4.4 je podan primer funkcije $f : \mathbb{R}^2 \to [0, \infty)$, ki zadošča pogojema $\mathbf{ucat}(f) = 3$ in $\mathbf{ucat}^2(f) = 2$, v razdelku 4.5 pa primer funkcije $f : \mathbb{R}^2 \to [0, \infty)$, ki zadošča pogojema $\mathbf{ucat}(f) = 3$ in $\mathbf{ucat}^{\infty}(f) = 2$.

Nazadnje z uporabo trditve Proposition 4.3 iz [43] dokažemo še monotonost za nekatere dovolj enostavne funkcije $f : \mathbb{R}^2 \to [0, \infty)$.

Izrek D.12. *Naj bo $f : \mathbb{R}^2 \to [0, \infty)$ Morsova funkcija, ki ima v različnih kritičnih točkah različne kritične vrednosti. Recimo, da je njen Morse-Smaleov graf drevo in $0 < p_1 < p_2 \leq \infty$. Potem velja*

$$\mathbf{ucat}^{p_1}(f) \leq \mathbf{ucat}^{p_2}(f).$$

D.4 Razno

V razdelku 5 je predstavljenih nekaj možnih smeri raziskovanja, ki se avtorju zdijo obetavne.

Najprej predstavimo nekaj idej za študij funkcij $f : \mathbb{R}^m \to [0, \infty)$. To sta pojem multimodalne funkcije in lema o razširitvi. Namen prvega je posplošiti pojem funkcije, katere Morse-Smaleov graf je drevo, ne da bi se pri tem sklicevali na gladkost. Namesto tega za osnovo vzamemo povsem topološki pogoj, soroden unimodalnosti.

Definicija. Funkcija $f : X \to [0, \infty)$ je *multimodalna*, če obstaja $M > 0$ da je vsaka nadnivojnica $f^{-1}[c, \infty)$ za $0 < c \leq M$ homotopsko ekvivalentna končni množici točk in prazna za $c > M$.

Zanimivo je vprašanje, ali je unimodalna p-kategorija take funkcije monotona v p. V primeru, ko poleg multimodalnosti privzamemo, da je funkcija $f : \mathbb{R}^m \to$

$[0,\infty)$ Morsova in ima v različnih kritičnih točkah različne kritične vrednosti, lahko pokažemo, da ima le kritične točke indeksov m in $m-1$, kar pomeni, da ji lahko priredimo Morse-Smaleov graf, kar za take funkcije odpira možnost podobnega pristopa, kot so ga uporabili avtorji [43] v primeru $\mathbb{R}^2$.

Nato si ogledamo nekaj odprtih vprašanj. Vprašamo se lahko npr., kakšen pojem kategorije dobimo, če v definiciji unimodalnosti besedo "kontraktibilnost" zamenjamo za "povezanost s potmi".

Definicija. Zvezna funkcija $u : X \to [0,\infty)$ je *π_0-unimodalna* če obstaja $M > 0$, da so nadnivojnice $u^{-1}[c,\infty)$ povezane s potmi za $0 < c \leq M$ in prazne za $c > M$.

Od tod dobimo naslednji pojem kategorije, katerega lastnosti bi bilo zanimivo podrobneje študirati. Prav tako bi si bilo zanimivo ogledati različice, ki jih dobimo, če namesto kontraktibilnosti za osnovo vzamemo npr. povezanost, konveksnost ali pa če za unimodalne funkcije ne privzemamo zveznosti.

Definicija. Naj bo $p \in (0,\infty)$. Vpeljimo *π_0-unimodalno p-kategorijo* $\mathbf{ucat}^p_{\pi_0}(f)$ funkcije $f : X \to [0,\infty)$ kot minimalno število n π_0-unimodalnih funkcij $u_1,\dots,u_n : X \to [0,\infty)$, tako da po točkah velja $f = (\sum_{i=1}^n u_i^p)^{\frac{1}{p}}$. Pojem *$\pi_0$-unimodalne ∞-kategorije* definirajmo analogno z uporabo ∞-norme.

Opazimo lahko, da pojem Morse-Smaleovega grafa v [43] nikjer ne upošteva lokalnih minimumov. Ti nakazujejo prisotnost ciklov in si po avtorjevem mnenju zaslužijo večje pozornosti. Še eno nerešeno vprašanje je, na kakšnih grafih lahko upamo na algoritme, sorodne algoritmu s pometanjem iz [4].

Nazadnje se vprašamo še, ali je pojem unimodalne kategorije mogoče študirati s pomočjo kohomološkega pristopa. Ta se je izkazal kot precej plodovit v primeru Lusternik-Schnirelmannove kategorije. V primeru unimodalne kategorije za začetek ni jasno niti, kakšno homološko teorijo bi sploh lahko uporabili, niti če je tako kohomološko teorijo sploh mogoče konstruirati.

D.5 Aproksimativni izrek o živcu

Izrek o živcu je eden od klasičnih rezultatov algebraične topologije, ki govori o odnosu med dovolj lepim pokritjem topološkega prostora in živcem tega pokritja, in ima korenine v delu Aleksandrova [2].

Izrek D.13 (Corollary 4G.3 [42])**.** *Če je $\mathcal{U}$ odprto pokritje parakompaktnega prostora X, tako da je vsak neprazen presek končnega števila množic v $\mathcal{U}$ kontraktibilen, potem je X homotopsko ekvivalenten živcu $\mathcal{N}(\mathcal{U})$.*

Ena od novejših aplikacij tega izreka je na področju *topološke analize podatkov* [37, 14, 63]. Cilj je pridobiti informacije o topologiji prostora, pri čemer pogosto poznamo le diskreten vzorec tega prostora. Na to temo obstaja precej člankov, kjer so dokazani rezultati v različnih kontekstih, kar vključuje [21, 9, 26], če jih naštejemo le nekaj. Skupna točka vseh teh rezultatov je uporaba izreka o živcu, eksplicitno ali pa implicitno, z uporabo konstrukcij kot je npr. Čechov kompleks.

Glavna ideja *vztrajne homologije*, ki je močno orodje na področju topološke analize podatkov, je da namesto homologije enega samega prostora študiramo homologijo filtracije. Uporaba homološkega funktorja na taki filtraciji porodi *vztrajnostni modul*. Če računamo homologijo s koeficienti v polju, lahko dobimo popolno topološko invarianto, ki se imenuje *vztrajnostna črtna koda* ali *vztrajnostni diagram*. Uporaben primer filtracij so npr. podnivojske (oziroma nadnivojske) filtracije – za dani prostor, opremljen z realno zvezno funkcijo $f : X \to \mathbb{R}$ podnivojnice te funkcije tvorijo filtracijo, ta pa porodi vztrajnostni diagram, ki ga označimo $\mathrm{Dgm}(X, f)$.

Pomemben primer take funkcije je razdalja do dane kompaktne množice. Kadar ta kompaktna množica sestoji iz točk vzorca, je funkcija povezana s pojmom merila in je ekvivalentna Čechovi filtraciji. Spomnimo se, da je Čechov kompleks na množici točk P definiran kot živec družine krogel z radijem r. Točke običajno ležijo v evklidskem prostoru, kar omogoča uporabo izreka o živcu s pomočjo konveksnosti. Spreminjanje radija r nam da Čechovo filtracijo. Druge filtracije, ki se pogosto obravnavajo so: nadnivojska filtracija prirejena dani funkciji gostote verjetnosti [9], podnivojska filtracija prirejena vzorčeni funkciji [21] in filtracije prijene višinskim funkcijam na ploskvah [1]. Vztrajnostni diagrami so se izkazali kot zanimivi, ker so stabilni [23], kar pomeni, da majhna sprememba v filtraciji povzroči enako majhno spremembo v invarianti. En način za merjenje te spremembe je *ozkogrlna razdalja*. Stabilnost nam omogoča, da dokažemo izreke o aproksimaciji vztrajne homologije filtracije s pomočjo druge filtracije, ki jo dobimo iz diskretnega vzorca, in sicer nam pove, da je ozkogrlna razdalja majhna.

Pomembna metoda v dokazovanju take aproksimacije je *prepletanje* [19], ki podaja algebraični pogoj za aproksimacijo (razdelek 2.5). Pogosta tema je konstrukcija prepletanja z dobrim pokritjem, kjer je aproksimacija teoretično zagotovljena. V nekaterih primerih, kot npr. za distančno filtracijo, lahko prepletanje z dobrim pokritjem pogosto pokažemo direktno. V splošnejših kontekstih je direkten dokaz včasih težji. Glavni cilj razdelka 6 je, da dokažemo zgornjo mejo za aproksimacijo s pomočjo stabilnosti vztrajne homologije, ki nam omogoči nekoliko omiliti zahtevo, da je pokritje dobro. Pomembno je dejstvo, da rezultate izpeljemo zgolj na podlagi predpostavk o *lokalnih lastnostih* prostora in funkcije, kar omogoča uporabo v zelo različnih aplikacijah.

Delamo z vztrajno homologijo, zato izhajamo iz homološke verzije izreka 1.1.

Izrek D.14 (Theorem 4.4 [12])**.** *Recimo, da je X unija podkompleksov U_i, tako da je vsak neprazen presek $U_{i_0} \cap \cdots \cap U_{i_p}$ za $p \geq 0$ acikličen. Potem velja $\mathsf{H}_*(X) \cong \mathsf{H}_*(\mathcal{N}(\mathcal{U}))$, kjer je $\mathcal{N}(\mathcal{U})$ živec pokritja.*

Glavni rezultat razdelka 6 je *aproksimativna* verzija zgornjega izreka v kontekstu vztrajne homologije. Za dani prostor in funkcijo najprej definiramo pojem ε-acikličnega pokritja. Pripomnimo, da se ne omejimo na inducirane funkcije na fiksnem pokritju, ampak študiramo pokritje s filtriranimi prostori. Ta pojem je nekoliko manj intuitiven, je pa zato precej širše aplikativen. Uporabljamo Sheehyjevo formalizacijo pokritja s filtriranimi prostori [60]. Naš glavni rezultat se v neformalnem jeziku glasi:

Rezultat D.15. *Naj bo dan prostor X, opremljen s funkcijo f in (filtriranim) pokritjem $\mathcal{U}$. Če je vsak neprazen presek elementov pokritja ε-acikličen, potem obstaja funkcija na živcu $g : \mathcal{N}(\mathcal{U}) \to \mathbb{R}$, tako da ozkogrlna razdalja $d_B(\cdot)$ zadošča oceni*

$$d_B(\mathrm{Dgm}(X, f), \mathrm{Dgm}(\mathcal{N}(\mathcal{U}), g)) \leq 2(Q + 1)\varepsilon,$$

kjer je
$$Q = \min\{\dim(X), \dim(\mathcal{N}(\mathcal{U}))\}.$$

Konstrukcija funkcije na živcu je podana eksplicitno in se ujema z metodami, ki se za izračun vztrajne homologije trenutno uporabljajo v praksi.

V dokazih ne uporabljamo vztrajnostnih diagramov in ozkogrlne razdalje, saj menimo, da je priročneje delati neposredno z ustreznimi vztrajnostnimi moduli in prepletanji. Zato ozkogrlne razdalje in vztrajnostnih diagramov ne definiramo eksplicitno, saj niso nujni za formulacijo naših rezultatov. Kljub temu aludiramo na te pojme, kadar bi to utegnilo biti v pomoč bralcem, seznanjenim z vztrajnostjo. V primeru, ko so diagrami dobro definirani, ocene na podlagi ozkogrlne razdalje sledijo avtomatično.

Glavni rezultat dokažemo z uporabo Mayer-Vietorisovega spektralnega zaporedja, ki nam omogoča zlepiti informacije o ε-acikličnih elementih pokritja v globalno informacijo o vztrajni homologiji prostora. Da dobimo tesno mejo, vpeljemo pojma *levih* in *desnih prepletanj* (razdelek 6.2), ki imajo dodatno strukturo. Na ta način lahko zajamemo podobne fenomene kot rezultati v [5], s to razliko, da delamo neposredno na nivoju modulov, in ne črtnih kod. Zato ne zahtevamo, da so moduli razcepni, za vpeljane pojme pa menimo, da so neodvisno zanimivi.

Ta tip aproksimacijskih rezultatov je požel precej pozornosti na področju računske topologije. Poleg zgoraj omenjenih aplikacij vztrajna lema o živcu v [22] pokaže, da homotopska ekvivalenca med dobrim pokritjem in njegovim živcem komutira z inkluzijami, kar omogoča uporabo v vztrajnem kontekstu. V zadnjem času je bilo to uporabljeno v [10], kjer avtorji aproksimirajo vztrajno homologijo Čechovega kompleksa v evklidskem prostoru z uporabo pokritij, ki niso dobra. Primerljiv rezultat se je pred kratkim pojavil tudi v [15]. Ta rezultat je bil splošnejši od predhodne različice članka [41], v kateri niso bile obravnavane *filtracije filtracij* (oziroma multipokritja) – v trenutni verziji članka in v pričujoči disertaciji pa zdaj obravnavamo tudi ta primer. Različen je tudi pristop, saj [15] zahteva, da so elementi pokritja (in njihovi končni preseki) "ε-nulhomotopni". Pokažejo, da je pri tej predpostavki mogoče zgraditi eksplicitno verižno preslikavo, ki je hkrati prepletanje (in ustreza isti aproksimacijski konstanti kot v našem primeru). Naš pristop pa je povsem algebraičen in tako zanj potrebujemo le strukturo na nivoju homologije, ne pa tudi homotopije.

Rezultat dokažemo v dveh korakih: najprej pokažemo, kako se aproksimacijska meja spreminja skozi izračun spektralnega zaporedja (razdelek 6.3), potem pa razrešimo razširitveni problem in tako povežemo rezultat spektralnega zaporedja z vztrajno homologijo danega prostora (razdelek 6.4). Razdelek 6 smo poskusili napisati tako, da bi bil samostojno berljiv, a na nekaterih mestih še vedno privzemamo nekaj poznavanja spektralnih zaporedij, pri tem pa kjer je le mogoče, podajamo tudi ustrezno intuicijo in vire.

Izjava

Podpisani Dejan Govc izjavljam, da je disertacija z naslovom *Unimodalna kategorija* oziroma *Unimodal category* plod mojega lastnega študija in raziskovalnega dela.

Ljubljana, 2017 Dejan Govc