Discrete homotopy theory and critical values of metric spaces

by

Jim Conant (Knoxville, TN),
Victoria Curnutte (Sedalia, CO),
Corey Jones (Nashville, TN), **Conrad Plaut** (Knoxville, TN),
Kristen Pueschel (Ithaca, NY), **Maria Lusby** (Hicksville, NY)
and **Jay Wilkins** (Pembroke, NC)

Abstract. Utilizing the discrete homotopy methods developed for uniform spaces by Berestovskii–Plaut, we define the critical spectrum $\mathrm{Cr}(X)$ of a metric space, generalizing to the non-geodesic case the covering spectrum defined by Sormani–Wei and the homotopy critical spectrum defined by Plaut–Wilkins. If X is geodesic, $\mathrm{Cr}(X)$ is the same as the homotopy critical spectrum, which differs from the covering spectrum by a factor of 3/2. The latter two spectra are known to be discrete for compact geodesic spaces, and correspond to the values at which certain special covering maps, called δ-covers (Sormani–Wei) or ε-covers (Plaut–Wilkins), change equivalence type. In this paper we initiate the study of these ideas for non-geodesic spaces, motivated by the need to understand the extent to which the accompanying covering maps are topological invariants. We show that discreteness of the critical spectrum for general metric spaces can fail in several ways, which we classify. The "newcomer" critical values for compact non-geodesic spaces are completely determined by the homotopy critical values and the *refinement* critical values, the latter of which can, in many cases, be removed by changing the metric in a bi-Lipschitz way.

1. Introduction. The *covering spectrum* of a geodesic space, introduced by Sormani and Wei ([14], [15]), provides a way to understand the topology of the space at a given scale. Their construction uses a classical method of Spanier ([16]) that produces a covering map for a given open cover; in particular, they use the covering map $\pi^\delta : \widetilde{X}^\delta \to X$ of a geodesic space X corresponding to the open cover by δ-balls. Such a cover is constructed for every $\delta > 0$, yielding a parameterized collection $\{\widetilde{X}^\delta\}_{\delta > 0}$. The covering spectrum consists of those $\delta > 0$ such that for all $\delta' > \delta$, the cover-

[97]

ing maps $\pi^{\delta'}$ and π^δ are not equivalent (see also Remark 16 below). Sormani and Wei use these covering maps, and the covering spectrum, to obtain interesting results concerning limits of compact Riemannian manifolds with Ricci curvature uniformly bounded below.

Berestovskii and Plaut independently developed a method to construct covering maps determined by geometric information—first for topological groups ([2], [4]) and later for uniform spaces ([3], [9]), hence metric spaces. Unlike Spanier's construction, which uses classical (continuous) paths and homotopies, the Berestovskii–Plaut construction uses discrete chains and homotopies, which we will discuss in more detail in the following section. One consequence is that the construction works for an arbitrary connected metric space and yields, for each $\varepsilon > 0$, a covering map $\varphi_\varepsilon : X_\varepsilon \to X$ and corresponding deck group $\pi_\varepsilon(X)$, called the ε-cover and ε-group, respectively. Plaut and Wilkins show in [11] that for compact geodesic spaces the Sormani–Wei cover π^δ is equivalent to φ_ε when $\varepsilon = 2\delta/3$. At the same time, in [10], [11], and [19], they give several topological and geometric applications of the discrete methods.

For example, in [10] the authors strengthen a theorem of M. Gromov ([6]) that the fundamental group of any compact Riemannian manifold of diameter at most D has a set of generators $g_1, \ldots, g_k$ of length at most $2D$ and relators of the form $g_i g_m = g_j$. In fact, they obtain an explicit bound for the number k of generators in terms of the number of equivalence classes of "short loops" at every point and the number of balls required to cover a given semilocally simply connected geodesic space. As a corollary they obtain a "curvature free" fundamental group finiteness theorem (new even for Riemannian manifolds) that generalizes the finiteness theorems of Anderson ([1]) and Shen–Wei ([13]). In [11], they investigate the covers of Gromov–Hausdorff convergent sequences, where discrete methods are particularly useful due to the well-established principle that many properties of the limit of a convergent sequence of compact metric spaces X_n can be formulated in terms of the limits of finite point sets (cf. Chapter 7 of [5]). They establish a type of weak closure result for a collection of covers called *circle covers* that generalize ε-covers, allowing them to complete the investigation of Sormani-Wei into the structure of limits of δ-covers. In [19], Wilkins employs these discrete methods to relate the existence of universal covers of Peano continua to the geometry and topology of various generalized fundamental groups. For example, he shows that such a space has a universal cover if and only if the revised and uniform fundamental groups are finitely presented, which, in turn, holds if and only both groups are discrete in a specific topology called the *covering topology*.

In [11], Plaut and Wilkins ask to what degree these concepts are topological invariants. The present paper represents a start on that program by

moving beyond the geodesic case. There are many reasons why one would want to pursue such a program, including the existence of new topological invariants. As an example, given a Peano continuum X, one can consider the collection of all covering maps of X (with connected domain) that are ε-covers for some geodesic metric (or more general metric) on X. For instance, up to rescaling, there is only one geodesic metric on a topological circle, and the only ε-covers for that metric are the trivial cover and the universal cover; the n-fold covers of the circle are not ε-covers for the geodesic metric. We conjecture that the same is true for any metric on the circle (assuming the covers have connected domain). On the other hand, $\mathbb{RP}^2$ with an open disc removed is homeomorphic to the Möbius band, and so there are geodesic metrics on the Möbius band that Gromov–Hausdorff approximate $\mathbb{RP}^2$. It follows from the convergence results of [11] that the 2-fold cover of the Möbius band occurs as an ε-cover for some geodesic metrics. In other words, one can distinguish between the circle and the Möbius band on this basis, despite the fact that the circle is a deformation retraction of the Möbius band. However, other than directly classifying all of the ε-covers in examples like the circle, at this point we have few good methods to show that a given cover is *not* an ε-cover for some geodesic metric, and it may be easier to consider more broadly the collection of covering maps that are ε-covers for some compatible metric in general. In this case one has to consider the possibility that X_ε is not connected, and also that the analog of the covering spectrum may not be discrete in $\mathbb{R}_+$. The latter possibility is one of the main complications of the non-geodesic case, and understanding it is a main thrust of this paper.

Note also that these discrete methods—unlike geodesic methods—can be applied to fractals with no rectifiable curves like the Sierpiński Gasket with its "resistance metric," which is not geodesic and is different from the metric induced by the plane (cf. [7] and [17]). One may also consider more exotic spaces such as metrized solenoids, and—as was already observed in [3]—the Topologist's Sine Curve. Finally, one should recall that Gromov–Hausdorff convergence of non-compact proper geodesic spaces is generally reduced to consideration of metric balls, which, even in geodesic spaces, need not themselves be geodesic spaces when equipped with the subspace metric.

The results of this paper are built around a natural defintion of the *critical spectrum* for an arbitrary metric space. The precise definition (Definition 19) requires some technical background given in Section 2, but the critical spectrum of a metric space X—denoted by $\mathrm{Cr}(X)$—can be intuitively thought of as the set of all positive real numbers ε at which the equivalence type of ε-covers changes as ε decreases toward 0. This spectrum extends both the covering spectrum of Sormani–Wei and the homotopy crit-

ical spectrum of Plaut–Wilkins ([10]) to the non-geodesic case. The latter spectra are known to be discrete in $\mathbb{R}_+$ for any compact geodesic metric (see Section 3), but for non-geodesic spaces the situation is more complicated.

First of all, we have

THEOREM 1. *There exist compact connected metric spaces with critical spectra that are not discrete in* $\mathbb{R}_+$.

In Section 3 we give a formal classification of the ways in which discreteness may fail. In Section 4, we give examples of many of the types of critical values deduced in Section 3 using spaces we call Rapunzel's Combs. To begin with, one may identify two types of fundamental critical values: homotopy and refinement critical values. The resulting spectra are denoted by $H(X)$ and $R(X)$, respectively. Homotopy critical values are defined for metric spaces in the same way that they are defined for geodesic spaces, and are more topologically significant than refinement critical values, although the latter do indicate a lack of connectivity of certain metric balls (Proposition 32). In particular, $R(X)$ is empty for geodesic spaces, and refinement critical values can often be removed by changing the metric:

THEOREM 2. *Let* (X, d) *be a compact connected metric space. Then for every* $\varepsilon > 0$ *there is a metric* D_ε *on* X *such that*

(1) $d \le D_\varepsilon$, *and* $d(x, y) = D_\varepsilon(x, y)$ *if and only if* $d(x, y) < \varepsilon$.
(2) d *and* D_ε *are bi-Lipschitz equivalent.*
(3) D_ε *has no refinement critical values greater than* ε.

The metric D_ε is called the *ε-intrinsic metric induced by* d. Note that the only obstruction in Theorem 2 to possibly removing *all* refinement critical values is that one may have $\inf R(X) = 0$ (see Theorem 6).

Rapunzel's Combs show that not all critical values are refinement or homotopy critical values, but these two types determine the entire spectrum:

THEOREM 3. *Let* X *be a compact connected metric space. If* $\overline{H(X)}$ *denotes the closure of the homotopy critical spectrum* $H(X)$ *in* $\mathbb{R}_+ = (0, \infty)$, *then*

$$\mathrm{Cr}(X) = \overline{H(X)} \cup R(X).$$

Since a geodesic space X has no refinement critical values, the discreteness of $H(X)$ in $\mathbb{R}_+$ when X is compact immediately yields the following.

COROLLARY 4. *If* X *is geodesic, then* $\mathrm{Cr}(X) = \overline{H(X)}$. *If* X *is also compact, then* $\mathrm{Cr}(X) = H(X)$.

We are also able to classify isolated critical values:

THEOREM 5. *If X is a compact metric space, and if $\varepsilon > 0$ is an isolated critical value of X, then ε is a homotopy or a refinement critical value.*

Finally, it is interesting to consider the limit of the metrics D_ε described in Theorem 2, which are monotone increasing as $\varepsilon \searrow 0$. In general the limit need not be finite, or if it is finite, topologically equivalent to the original metric. However, we do have:

THEOREM 6. *Let (X, d) be a compact connected metric space. If D_0 is finite, then (X, D_0) has no refinement critical values. If, in addition, (X, d) and (X, D_0) are topologically equivalent then D_0 is a geodesic metric—and hence X is locally path connected.*

REMARK 7. Some of the research contained in this paper was carried out by the undergraduate authors during an REU at the University of Tennessee in summer 2009 and during the subsequent year. The quality of contributions by undergraduate co-authors in published papers is known to vary from paper to paper. In the case of the present work, the contributions of the undergraduate authors are substantial, including the discovery of Rapunzel's Combs and insights into the distinct types of critical values in general metric spaces.

2. Background on discrete homotopy theory. Omitted proofs of statements in this section may be found in [3] or [18]. Given a metric space (X, d), we will denote the open metric balls of radius ε centered at $x \in X$ by $B(x, \varepsilon) := \{y \in X : d(x, y) < \varepsilon\}$. If multiple metrics on the same space are being considered, we will distinguish them by subscripts: $B_{d_1}(x, r)$, $B_{d_2}(x, r)$, etc. A map $f : Y \to X$ between metric spaces is an *isometry* if f is surjective and preserves distances (i.e. $d_X(f(y_1), f(y_2)) = d_Y(y_1, y_2)$ for all $y_1, y_2 \in Y$). If $f : Y \to X$ is surjective and for some $\varepsilon > 0$ the restriction of f to any $B(y, \varepsilon) \subset Y$ is an isometry onto $B(f(y), \varepsilon) \subset X$, we call f an *ε-local isometry*. It follows easily that an ε-local isometry is in fact a covering map such that ε-balls are evenly covered; when ε is unimportant we may simply call f a *metric covering map*. Recall that X is a *geodesic space* if every pair of points in X is joined by a *geodesic*, an arclength parameterized curve having length equal to the distance between its endpoints. Note that when X is a geodesic space, Y is a connected topological space, and $f : Y \to X$ is any covering map, then the well-known "lifted metric" on Y is the unique, topologically equivalent, geodesic metric on Y such that f is a metric covering map.

An *ε-chain* α in X is a finite sequence of points $\alpha = \{x_0, x_1, \ldots, x_{n-1}, x_n\}$ such that $d(x_{i-1}, x_i) < \varepsilon$ for $i = 1, \ldots, n$. For any ε-chain $\alpha = \{x_0, \ldots, x_n\}$, the *reversal* of α is denoted by $\alpha^{-1} := \{x_n, x_{n-1}, \ldots, x_1, x_0\}$. A *basic move* on an ε-chain is the addition or removal of a single point with the conditions

that the endpoints remain fixed and the resulting chain is still an ε-chain. If two ε-chains α and β have the same initial and terminal points, then we say they are ε-*homotopic* if there is a finite sequence of ε-chains, $H = \{\alpha = \gamma_0, \gamma_1, \ldots, \gamma_{k-1}, \gamma_k = \beta\}$—called an ε-*homotopy*—such that each γ_i differs from γ_{i-1} by a basic move. For a fixed basepoint $* \in X$, let X_ε be the set of all ε-equivalence classes $[\{* = x_0, \ldots, x_n\}]_\varepsilon$ of ε-chains in X beginning at $*$, which has a natural metric (Definition 9). One may then define $\varphi_\varepsilon : X_\varepsilon \to X$ to be the "endpoint map" taking $[\{* = x_0, \ldots, x_n\}]_\varepsilon$ to x_n. If one assumes that X is *chain connected* in the sense that every pair of points is joined by ε-chains in X for all $\varepsilon > 0$, then $\varphi_\varepsilon : X_\varepsilon \to X$ is surjective, and in fact is a covering map and local isometry. Given $0 < \delta < \varepsilon$, there is a well-defined map (cf. [3]) $\varphi_{\varepsilon\delta} : X_\delta \to X_\varepsilon$ that simply treats a δ-chain like an ε-chain, i.e. $\varphi_{\varepsilon\delta}([\alpha]_\delta) = [\alpha]_\varepsilon$. These maps satisfy the composition relation $\varphi_{\varepsilon\delta} = \varphi_{\varepsilon\lambda} \circ \varphi_{\lambda\delta}$ for $\delta < \lambda < \varepsilon$.

Basic arguments show that connected spaces are chain connected, and compact chain connected spaces are connected—although, for example, the rationals with the subspace metric from $\mathbb{R}$ are chain connected but not connected. To ensure that the maps φ_ε are always covering maps, we assume from now on that all metric spaces are connected unless otherwise noted. We are concerned primarily with compact connected metric spaces.

Although it is important to keep in mind that all chains technically have both initial and terminal points, for simplicity we will often use the notation $\{x\}$ instead of $\{x, x\}$ for constant or trivial chains. The notation $\alpha \sim_\varepsilon \beta$ will denote the (equivalence) relation "α is ε-homotopic to β." We will often write $[x_0, \ldots, x_n]_\varepsilon$ rather than the technically correct $[\{x_0, \ldots, x_n\}]_\varepsilon$. There is a well-defined, associative concatenation operation on equivalence classes of ε-chains: if the initial point of β equals the terminal point of α, then we can define $[\alpha]_\varepsilon[\beta]_\varepsilon = [\alpha\beta]_\varepsilon$. The following properties are easily checked:

LEMMA 8. *Let α and β be ε-chains. If the initial point of β is the terminal point of α, and if $\alpha' \sim_\varepsilon \alpha$ and $\beta' \sim_\varepsilon \beta$, then $\alpha\beta \sim_\varepsilon \alpha'\beta'$. Furthermore, $\alpha\alpha^{-1}$ is ε-null, and if $\alpha \sim_\varepsilon \beta$, then $\alpha^{-1} \sim_\varepsilon \beta^{-1}$.*

An ε-*loop* is an ε-chain with equal initial and terminal points. If an ε-loop $\alpha = \{* = x_0, \ldots, x_n = *\}$ is ε-homotopic to the trivial loop $\{*\}$ then we say that α is ε-*null*. If α is an ε-chain, $\delta < \varepsilon$, and α is ε-homotopic to a δ-chain β, then we say that α *can be ε-refined to β*. We call such a homotopy an ε-*refinement*, and refer to β as a δ-*refinement* of α. When no confusion will result, we will often drop the ε's and δ's in our notation. Note that if α is an ε-chain in a *geodesic* space then we can refine α to any desired degree of fineness by successively adding midpoints of geodesics joining each pair of consecutive points in the chain. Refinements of this sort play important

roles in [10] and [11], and the ability (or inability) to refine chains is an important possible feature of non-geodesic spaces.

The set of classes $[\lambda]_\varepsilon$, where λ is an ε-loop at $*$, forms a group $\pi_\varepsilon(X)$ with operation induced by concatenation. This group acts on X_ε by preconcatenation: for any ε-loop λ, define $h_\lambda([\alpha]_\varepsilon) := [\lambda\alpha]_\varepsilon$. Then h_λ is a bijection that depends only on the ε-equivalence class of λ and, as was shown in [3], φ_ε is a well-defined regular covering map with deck group naturally identified with $\pi_\varepsilon(X)$. One can think of $\pi_\varepsilon(X)$, roughly, as a fundamental group at a specific metric scale. For instance, for a geodesic space X there is always a natural surjective homomorphism $h : \pi_1(X) \to \pi_\varepsilon(X)$, the kernel of which is generated by classes of loops that are, roughly, "small" on the scale of ε.

We define a metric d_ε on X_ε as follows:

DEFINITION 9. The *length* of an ε-chain, $\alpha = \{x_0, \dots, x_n\}$, in (X, d) is

$$L(\alpha) := \sum_{i=1}^{n} d(x_{i-1}, x_i).$$

The length of the equivalence class $[\alpha]_\varepsilon \in X_\varepsilon$ is $L([\alpha]_\varepsilon) := \inf\{L(\beta) : \beta \in [\alpha]_\varepsilon\}$, and for $[\alpha]_\varepsilon, [\beta]_\varepsilon \in X_\varepsilon$, we define $d_\varepsilon([\alpha]_\varepsilon, [\beta]_\varepsilon) = L([\alpha^{-1}\beta]_\varepsilon)$.

Note that if the initial point of β is the terminal point of α, then $L(\alpha\beta) = L(\alpha) + L(\beta)$. We also have $L(\alpha) = L(\alpha^{-1})$, and hence $L([\alpha^{-1}]_\varepsilon) = L([\alpha]_\varepsilon)$.

It is straightforward to check that $L([\alpha\beta]_\varepsilon) \leq L([\alpha]_\varepsilon) + L([\beta]_\varepsilon)$, from which the triangle inequality for d_ε follows. Symmetry is obvious, and positive definiteness is implied by the otherwise useful fact that if $L([\alpha]_\varepsilon) < \varepsilon$, where $\alpha = \{x_0, \dots, x_n\}$ is an ε-chain, then $d(x_0, x_n) < \varepsilon$, α is ε-homotopic to the two-point ε-chain $\{x_0, x_n\}$, and $L([\alpha]_\varepsilon) = d(x_0, x_n)$. More generally one also sees that for $0 < \delta \leq \varepsilon$ and an ε-chain $\alpha = \{* = x_0, \dots, x_n\}$, $B([\alpha]_\varepsilon, \delta) \subset X_\varepsilon$ consists precisely of those $[\beta]_\varepsilon \in X_\varepsilon$ such that β is ε-homotopic to an ε-chain of the form $\{* = x_0, \dots, x_n, y\}$, where $d(y, x_n) < \delta$. Furthermore, φ_ε is 1-Lipschitz (or distance non-increasing) and an $\varepsilon/2$-local isometry. In fact, every $\varepsilon/2$-ball in X is evenly covered by a union of $\varepsilon/2$-balls in X_ε. As with the case of classical covering space theory, change of basepoint in a chain connected space induces a natural equivalence of covering spaces, which with the present metric is an isometry. Therefore we will treat base points very informally, but assuming, when needed, that maps are base point preserving. In particular, if $*$ is the base point in X we will always use $[*]_\varepsilon$ as the base point in X_ε. For brevity we will denote $[*]_\varepsilon$ by $\tilde{*}$. Finally, the maps $h_\lambda : X_\varepsilon \to X_\varepsilon$ defining the action of $\pi_\varepsilon(X)$ on X_ε are isometries.

There is an identification of fundamental importance, $\iota_{\delta\varepsilon} : X_\delta \to (X_\varepsilon)_\delta$, that is defined by

$$\iota_{\delta\varepsilon}\big([* = x_0, x_1, \dots, x_n]_\delta\big) = \big[[*]_\varepsilon, [*, x_1]_\varepsilon, \dots, [*, x_1, \dots, x_n]_\varepsilon\big]_\delta$$

and was shown in [3] to be a well-defined uniform homeomorphism. According to the metrics we have defined in the spaces X_ε and $(X_\varepsilon)_\delta$,

$$d_\varepsilon\big([*,\ldots,x_i]_\varepsilon, [*,\ldots,x_i,x_{i+1}]_\varepsilon\big) = d(x_i,x_{i+1}),$$

and therefore

$$L\big(\{* = x_0,\ldots,x_n\}\big) = L\big(\{[*]_\varepsilon, [*,x_1]_\varepsilon,\ldots,[*,x_1,\ldots,x_n]_\varepsilon\}\big).$$

This implies that $\iota_{\delta\varepsilon}$ is in fact an isometry in the case of metric spaces. For one thing, this fact implies that the basic results that we have proved about φ_δ hold equally well for the function $\varphi_{\varepsilon\delta}$, including the properties of being Lipschitz and a $\delta/2$-local isometry when it is surjective.

We will denote the restriction of $\varphi_{\varepsilon\delta}$ to $\pi_\delta(X)$ by $\Phi_{\varepsilon\delta} : \pi_\delta(X) \to \pi_\varepsilon(X)$. These maps are homomorphisms (again, invoke $\iota_{\delta\varepsilon}$), and $\varphi_{\varepsilon\delta}$ is injective (respectively, surjective) if and only if $\Phi_{\varepsilon\delta}$ is injective (respectively, surjective). In the case of geodesic spaces, the functions $\varphi_{\varepsilon\delta}$ are always surjective. In fact, when X is geodesic, the metric d_ε coincides with the lifted geodesic metric (Proposition 23 of [10]), so that X_ε is path connected, and hence chain connected. The identification $\iota_{\delta\varepsilon}$ then tells us that $\varphi_{\varepsilon\delta}$ is surjective.

An ε-loop of the form $\{x,y,z,x\}$ will be called an ε-*triangle*. Note that an ε-triangle is necessarily ε-null. If $f : Y \to X$ is a map between metric spaces, then a lift of a chain, $\alpha = \{x_0,\ldots,x_n\}$, from X to a point $y \in f^{-1}(x_0)$ is a chain $\tilde\alpha = \{y = \tilde x_0,\ldots,\tilde x_n\}$ such that $f(\tilde x_i) = x_i$ for each $i = 0,1,\ldots,n$.

The next lemma shows that X_ε largely inherits the local topology and metric properties of X. The first part follows directly from the fact that $\varphi_\varepsilon : X_\varepsilon \to X$ is an isometry from any $\varepsilon/2$-ball onto its image. The last parts were proved in [3] in the more general setting of uniform spaces.

LEMMA 10. *Let X be a connected metric space. If X is locally compact (respectively, complete), then X_ε is locally compact (respectively, complete). Furthermore, suppose the ε-balls in X have any one of the following properties: connected, chain connected, path connected. Then the whole space X_ε is, respectively, connected, chain connected, path connected.*

LEMMA 11 (Chain and homotopy lifting). *Let $f : Y \to X$ be a surjective map between metric spaces that is a bijection from ε-balls in Y onto ε-balls in X. Let $\alpha = \{x_0,x_1,\ldots,x_n\}$ be an ε-chain in X, and let $\tilde x_0$ be any point in $f^{-1}(x_0)$. Then α lifts uniquely to an ε-chain $\tilde\alpha$ beginning at $\tilde x_0$. If, in addition, f has the property that the lift of any ε-triangle in X is an ε-triangle in Y, and if β is an ε-chain that begins at x_0 and is ε-homotopic to α, then the lifts of α and β to $\tilde x_0$ end at the same point and are ε-homotopic.*

Proof. The first part is proved by induction on the number of points in α. If α contains one point, the result is trivial. So, assume the result holds for

all chains with n or fewer points for some $n \geq 1$. Let $\alpha = \{x_0, \ldots, x_n\}$ be an ε-chain with $n + 1$ points, and let $\lambda = \{x_0, \ldots, x_{n-1}\}$. Using the inductive hypothesis, let $\tilde{\lambda} = \{\tilde{x}_0, \ldots, \tilde{x}_{n-1}\}$ be the unique lift of λ to $\tilde{x}_0$. Then $f(\tilde{x}_{n-1}) = x_{n-1}$, f is a bijection from $B(\tilde{x}_{n-1}, \varepsilon)$ onto $B(x_{n-1}, \varepsilon)$, and $x_n \in B(x_{n-1}, \varepsilon)$. Let $\tilde{x}_n$ be the unique point in $B(\tilde{x}_{n-1}, \varepsilon)$ mapping to x_n under f, and let $\tilde{\alpha} = \tilde{\lambda}\{\tilde{x}_n\} = \{\tilde{x}_0, \ldots, \tilde{x}_{n-1}, \tilde{x}_n\}$. Then $\tilde{\alpha}$ is an ε-chain and $f(\tilde{\alpha}) = \alpha$, proving existence. If there were another ε-chain, $\bar{\alpha} = \{\tilde{x}_0 = \bar{x}_0, \ldots, \bar{x}_n\}$, beginning at $\tilde{x}_0$ and projecting to α, then by the uniqueness part of the inductive hypothesis, the first n points of $\bar{\alpha}$ and $\tilde{\lambda}$ must coincide: $\bar{x}_i = \tilde{x}_i$ for $0 \leq i \leq n - 1$. The bijectivity of f on $B(\tilde{x}_{n-1}, \varepsilon)$ then implies that $\tilde{x}_n$ and $\bar{x}_n$ must be the same, proving uniqueness.

To prove the second part, it suffices to consider the case in which α and β differ by only a basic move, since an ε-homotopy is just a finite sequence of basic moves. Suppose β is obtained by removing a point from α, say $\alpha = \{x_0, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n\}$ and $\beta = \{x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n\}$. Let $\tilde{\alpha}$ and $\tilde{\beta}$ denote the unique lifts of α and β to $\tilde{x}_0$. By uniqueness, $\tilde{\alpha}$ and $\tilde{\beta}$ must agree for their first i points. Denote $\tilde{\alpha}$ by $\{\tilde{x}_0, \ldots, \tilde{x}_{i-1}, \tilde{x}_i, \tilde{x}_{i+1}, \ldots, \tilde{x}_n\}$ and $\tilde{\beta}$ by $\{\tilde{x}_0, \ldots, \tilde{x}_{i-1}, \tilde{y}_{i+1}, \ldots, \tilde{y}_n\}$. Since we can remove x_i from α, the loop $\{x_{i-1}, x_i, x_{i+1}, x_{i-1}\}$ is an ε-triangle. The lift of this ε-triangle to $\tilde{x}_{i-1}$ is an ε-triangle by hypothesis, and, by uniqueness of lifts, the first three points of that triangle must be $\tilde{x}_{i-1}$, $\tilde{x}_i$, and $\tilde{x}_{i+1}$. Since a triangle is a loop, the fourth point of this lift must be $\tilde{x}_{i-1}$. In other words, we have $d_Y(\tilde{x}_{i-1}, \tilde{x}_{i+1}) < \varepsilon$. So, $\tilde{x}_{i+1}$ and $\tilde{y}_{i+1}$ both lie in $B(\tilde{x}_{i-1}, \varepsilon)$ and project under f to x_{i+1}, implying that $\tilde{x}_{i+1} = \tilde{y}_{i+1}$. Finally, by uniqueness of lifts, the rest of $\tilde{\beta}$ agrees with $\tilde{\alpha}$. Thus, $\tilde{\beta}$ is obtained by removing a point from $\tilde{\alpha}$, so $\tilde{\alpha} \sim_\varepsilon \tilde{\beta}$. The reasoning for the basic move is similar. ∎

COROLLARY 12. *Let X be a connected metric space, and let α and β be ε-chains beginning at a common point $x \in X$ that are ε-homotopic. Then their lifts, $\tilde{\alpha}$ and $\tilde{\beta}$, to any $\tilde{x} \in \varphi_\varepsilon^{-1}(x)$ are ε-homotopic in X_ε.*

Proof. Let $\{z_0, z_1, z_2, z_0\}$ be an ε-triangle in X, and let $[\alpha]_\varepsilon$ be any point in $\varphi_\varepsilon^{-1}(z_0)$, where $\alpha = \{* = x_0, \ldots, x_n = z_0\}$. Let $\beta = \{* = x_0, \ldots, x_n = z_0, z_1\}$ and $\lambda = \{* = x_0, \ldots, x_n = z_0, z_2\}$. Then $d_\varepsilon([\alpha]_\varepsilon, [\beta]_\varepsilon)$ and $d_\varepsilon([\alpha]_\varepsilon, [\lambda]_\varepsilon)$ are less than ε. Moreover, $\beta^{-1}\lambda \sim_\varepsilon \{z_1, z_0, z_2\}$, and since $d(z_1, z_2) < \varepsilon$, we can remove z_0 from this chain to conclude that $\beta^{-1}\lambda \sim_\varepsilon \{z_1, z_2\}$. Thus, $d_\varepsilon([\beta]_\varepsilon, [\lambda]_\varepsilon) < \varepsilon$. It follows that $\{[\alpha]_\varepsilon, [\beta]_\varepsilon, [\lambda]_\varepsilon, [\alpha]_\varepsilon\}$ is an ε-triangle, and it projects under φ_ε to $\{z_0, z_1, z_2, z_0\}$. Uniqueness of lifts now implies that ε-triangles lift to ε-triangles, so Lemma 11 then applies. ∎

For ε-covers, we can precisely characterize the lifts of ε-chains.

LEMMA 13. *Let X be a connected metric space, and let $\varepsilon > 0$ be given. If $\alpha = \{* = x_0, \ldots, x_n\}$ is an ε-chain beginning at the base point $* \in X$ then*

the unique lift of α to $\tilde{*} = [*]_\varepsilon \in X_\varepsilon$ is given by

$$\tilde{\alpha} = \{[*]_\varepsilon, [x_0, x_1]_\varepsilon, \ldots, [x_0, \ldots, x_{n-1}]_\varepsilon, [x_0, \ldots, x_n]_\varepsilon = [\alpha]_\varepsilon\}.$$

In particular, the endpoint of the lift of α is $[\alpha]_\varepsilon$ and the distances between consecutive points, as well as the chain length, are preserved in the lift.

Proof. First, note that, for $i = 1, \ldots, n$,

$$d_\varepsilon([x_0, \ldots, x_{i-1}]_\varepsilon, [x_0, \ldots, x_{i-1}, x_i]_\varepsilon)$$
$$= L([x_{i-1}, \ldots, x_1, x_0, x_1, \ldots, x_{i-1}, x_i]_\varepsilon)$$
$$= L([x_{i-1}, x_i]_\varepsilon) = d(x_{i-1}, x_i) < \varepsilon,$$

since we can successively remove from $\{x_{i-1}, \ldots, x_1, x_0, x_1, \ldots, x_{i-1}, x_i\}$ the point x_0, then each x_1, and so on via an ε-homotopy. Thus, $\tilde{\alpha}$ is an ε-chain, and clearly $\varphi_\varepsilon(\tilde{\alpha}) = \alpha$. ∎

The previous lemma immediately yields the following.

COROLLARY 14. *If X is a connected metric space and $\varepsilon > 0$, then an ε-loop γ based at $*$ lifts to an ε-loop at $\tilde{*} \in X_\varepsilon$ if and only if γ is ε-null. Thus, any representative of a non-trivial element of $\pi_\varepsilon(X)$ lifts open (i.e. to a non-loop).*

LEMMA 15. *For a connected metric space X and any $\varepsilon > 0$, X_ε is ε-connected and ε-simply connected, i.e. every ε-loop based at $\tilde{*} \in X_\varepsilon$ is ε-null, or equivalently, $\pi_\varepsilon(X_\varepsilon)$ is trivial.*

Proof. The ε-connectivity follows from Lemma 13. Given an ε-loop $\tilde{\gamma}$ at $\tilde{*} \in X_\varepsilon$, it will project to an ε-loop, $\gamma := \varphi_\varepsilon(\tilde{\gamma})$, at $*$. Since γ lifts to a closed loop, it is ε-null. By Corollary 12, this ε-nullhomotopy will lift to X_ε. ∎

We will postpone some concrete examples until the next section when we introduce the critical spectrum.

REMARK 16. In their definition of the covering spectrum, Sormani–Wei use the condition $\widetilde{X}^\delta \neq \widetilde{X}^{\delta'}$ for all $\delta' > \delta$, but from their proofs it is clear that they take this to mean non-equivalence of the corresponding covering maps π^σ and π^δ. In fact, it seems to be an interesting open question (in our terminology) whether it is possible for X_ε and X_δ to be homeomorphic when φ_ε and φ_δ are not equivalent. Recall that it is possible in general for non-equivalent covers to involve homeomorphic spaces—for example with n-fold covers of the circle. However, these specific covers may not be ε-covers of the geodesic circle with any compatible metric, as we conjectured in the introduction.

3. The critical spectrum. We begin by defining the two aforementioned fundamental types of critical values, the first of which was originally introduced in [10].

DEFINITION 17. A number $\varepsilon > 0$ is a *homotopy critical value* of X if there is an ε-loop γ in X that is not ε-null but is δ-null for all $\delta > \varepsilon$ (when γ is considered as a δ-chain). Such a γ is an *essential ε-loop*. The *homotopy critical spectrum* of X is the set $H(X)$ of all homotopy critial values.

DEFINITION 18. Let $\alpha = \{x, y\}$ be a two-point chain in a metric space X such that

(1) $d(x, y) = \varepsilon > 0$;
(2) for all δ greater than but sufficiently close to ε, α is not δ-homotopic to an ε-chain.

Then we say that ε is a *refinement critical value* of X and that $\{x, y\}$ is an *essential ε-gap*, or just an essential gap when ε is clear. The set of all refinement critical values of X is denoted by $R(X)$.

Of course, since $d(x, y) = \varepsilon$, an essential ε-gap $\{x, y\}$ is not an ε-chain, and the point of the above definition is that, in a sense, we cannot "refine" it to one either. The existence (or non-existence) of essential gaps is closely connected to the surjectivity of the maps $\varphi_{\varepsilon\delta}$; see Lemma 24.

Even adding refinement critical values to the mix, however, does not tell the full story, and to completely understand the behavior of the ε-covers we need the following general definition of a critical value.

DEFINITION 19. Let X be a chain connected metric space. A *non-critical interval* of X is a non-empty open interval $I \subset \mathbb{R}_+$ such that, for each $\delta < \varepsilon$ in I, the map $\varphi_{\varepsilon\delta} : X_\delta \to X_\varepsilon$ is bijective. We call $\varepsilon > 0$ a *critical value* of X if it does not lie in a non-critical interval. The set $\mathrm{Cr}(X)$ of all critical values of X is called the *critical spectrum* of X.

There are two immediate consequences of the definition. First, $\mathrm{Cr}(X)$ is bounded above by $\mathrm{diam}(X)$ when X is compact (or simply bounded). Second, the set of non-critical values, $\mathbb{R}_+ \setminus \mathrm{Cr}(X)$, is open in $\mathbb{R}_+$. This means that $\mathrm{Cr}(X)$ is closed in $\mathbb{R}_+$, although 0 may be a limit point of the critical spectrum.

Even though we are interested largely in non-geodesic spaces, it is useful to sketch out what happens in some simple but illustrative instances by starting with the geodesic case. Two non-geodesic examples follow the first two, and more are given in Section 4.

EXAMPLE 20. Consider $X = S^1$, the geodesic circle of circumference 1. The details of this example are given in [18]. If $\varepsilon > 1/3$, then all ε-loops are ε-null, meaning that the ε-cover X_ε is isometric to X and is, therefore, the trivial cover. For any $0 < \varepsilon \leq 1/3$, however, the ε-loop that wraps around the circle one time (in either direction) is no longer ε-null. The ε-cover for any $\varepsilon \leq 1/3$ is now $\mathbb{R}$, the universal cover, and these covers "unravel" the

hole in X. Since there is a 1/3-loop in X that is not 1/3-null but is δ-null for all $\delta > 1/3$, it follows that $\frac{1}{3}$ is a homotopy critical value of X. It is, in fact, the only critical value of X, since X_ε is universal for $0 < \varepsilon \leq 1/3$. ∎

While basic, this first example illustrates some important phenomena. First, as noted in Corollary 4, the critical spectrum of a compact geodesic space X is simply its homotopy critical spectrum. Furthermore, as we mentioned previously, one of the fundamental results concerning the covering/homotopy critical spectra of a compact geodesic space X is that this set is discrete in $\mathbb{R}_+ := (0, \infty)$ when X is compact. These properties do not hold for general metric spaces. The discreteness was first proved by Sormani and Wei for their covering spectrum in [14], and a stronger result was later established by Plaut and Wilkins ([10]) for the homotopy critical spectrum. Thus, the situation in the previous example is, by extension, typical of compact geodesic spaces. One can imagine $\varepsilon > 0$ sliding continuously along the positive real axis from the diameter of X towards 0 and consider the corresponding covers $\varphi_\varepsilon : X_\varepsilon \to X$. For $\varepsilon \geq \mathrm{diam}(X)$, the covering map φ_ε is trivial. As ε decreases to 0, the covering map remains trivial for a while until ε is the first critical value of X, if one exists. At that point, the equivalence class of $\varphi_\varepsilon : X_\varepsilon \to X$ and the ε-group change from the trivial cover and group, respectively, to the next cover and group in the sequence. The covering map φ_ε remains the same for some non-trivial interval, then changes again at the next critical value—again, if one exists.

The common discreteness of both the critical and covering spectra in the compact geodesic case is actually a consequence of a more general result. Plaut and Wilkins also showed in [11] that despite the apparent difference in methods, the Berestovskii–Plaut construction, when applied to geodesic spaces, yields the same covers as the δ-covers used by Sormani–Wei, for $\varepsilon = 2\delta/3$. It follows that the covering spectrum and homotopy critical spectrum of a compact geodesic space differ only by a multiplicative factor of 2/3. However, the requirement of chain connectedness is in general weaker than connectedness, and is even satisfied by some totally disconnected spaces. Therefore, the requirement of local path connectedness in the Spanier construction used by Sormani–Wei may be dropped when using discrete chains and homotopies. As was noted in the introduction, this opens the way to investigate all of these topics for non-geodesic spaces.

EXAMPLE 21. For the geodesic Hawaiian Earring determined by circles with distinct circumferences $d_i \to 0$, the homotopy critical spectrum is $\{d_1/3, d_2/3, \dots\}$. At each of the critical values, another circle is "unrolled," and X_ε consists of an infinite tree with successively smaller Hawaiian Earrings attached to each vertex. ∎

The previous example shows that for a compact geodesic space that is not semilocally simply connected, the critical spectrum is still discrete but may not have a positive lower bound. This is not true in general, though; there are compact geodesic spaces that are not semilocally simply connected but have finite or even empty critical spectra (cf. [19]). In general, for a compact, semilocally simply connected geodesic space, X_ε is the universal cover of X for all sufficiently small $\varepsilon > 0$; in fact, a stronger statement holds in the more general setting of uniform spaces (Theorem 88 in [3]). Note that in this case it is interesting to consider the so-called *uniform universal cover of X* ([3]), which is the inverse limit of the covers $\varphi_\varepsilon : X_\varepsilon \to X$ as $\varepsilon \to 0$. The bonding maps are the functions $\varphi_{\varepsilon\delta}$ mentioned earlier. The uniform universal cover—which exists for all compact geodesic spaces, even if they are not semilocally simply connected—has the same lifting, regularity, and categorical universality properties as the traditional universal cover, though it is not generally a classical covering space defined by evenly covered neighborhoods. In this paper, however, we are only interested in the maps φ_ε and $\varphi_{\delta\varepsilon}$ and not the inverse limit space.

The next two examples are non-geodesic, and Example 23 gives a simple instance of a refinement critical value.

EXAMPLE 22. Let X be the square of side-length s with its (non-geodesic!) Euclidean metric inherited from $\mathbb{R}^2$. It is not hard to see that, for $\varepsilon > s$, X_ε is the trivial cover—while for $0 < \varepsilon \le s$, X_ε is the universal cover, which is homeomorphic to $\mathbb{R}$. ∎

EXAMPLE 23. Let S^1 be the geodesic circle of circumference 1. Fix $a, b \in S^1$ so that $d(a, b) = 1/4$, and remove the open geodesic segment from a to b. Let $X \subset S^1$ be the resulting set with the inherited subspace metric, not the induced geodesic metric, which would just make X a line segment. Choose as the base point $*$ the midpoint of the longer segment from a to b. See Figure 1 (showing the full sequence of ε-covers). As in the case of S^1, X_ε is the trivial cover for $\varepsilon > 1/3$. Since ε is larger than the gap we created by removing the segment, an ε-loop can cross over the gap; informally speaking, X_ε does not "see" the gap. For $1/4 < \varepsilon \le 1/3$, one can show that $\pi_\varepsilon(X) \cong \mathbb{Z}$. In this
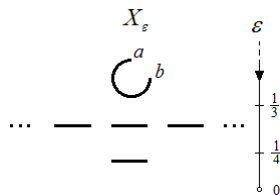

Fig. 1. The ε-covers of the geodesic circle with a gap

case, X_ε is the subspace of $\mathbb{R}$ with the open segments $(n/2-1/8, n/2+1/8)$, $n \in \mathbb{Z}$, removed. The ε-group, $\pi_\varepsilon(X)$, acts by shifts, as in the case of S^1 itself. Intuitively, we still unravel the circle as in the standard geodesic case, but the missing segment of the circle gets unraveled along with it. Note that X_ε is not connected in this case.

Now, suppose $0 < \varepsilon \leq 1/4$. Since the removed segment has length $1/4$, it becomes the case at $\varepsilon = 1/4$ that no ε-loop can cross the gap. Thus, the ε-covers now recognize the gap, and it is now impossible to travel around the circle via an ε-loop. In other words, all ε-loops are necessarily trivial, and $\pi_\varepsilon(X)$ is the trivial group. The pair $\{a, b\}$ is, in fact, an essential gap, and $1/4$ is a refinement critical value. In this case, $H(X) = \{1/3\}$, $R(X) = \{1/4\}$, and $\mathrm{Cr}(X) = \{1/4, 1/3\}$.

A subtle point deserves mention: $\varphi_\varepsilon : X_\varepsilon \to X$ is topologically trivial for $\varepsilon > 1/3$ and $0 < \varepsilon \leq 1/4$, but it is only an isometry for $\varepsilon > 1/3$. The reason for this is that, for $\varepsilon \leq 1/4$, the metric that we have defined makes X_ε isometric to a Euclidean segment. In fact, the changes in the topology of the covering space are measuring geometric features of the space rather than topological features. This example illustrates two things that can occur in non-geodesic spaces but not in geodesic ones: the covering space X_ε may not be connected even when X is path connected and locally path connected, and $\pi_\varepsilon(X)$ may be non-trivial even when $\pi_1(X)$ is trivial. ∎

We now continue to analyze the general critical spectrum.

LEMMA 24. *Let X be a metric space, and let $0 < \delta < \varepsilon$ be given. The following are equivalent:*

(1) *$\varphi_{\varepsilon\delta} : X_\delta \to X_\varepsilon$ is surjective;*
(2) *every ε-chain in X can be ε-refined to a δ-chain;*
(3) *every two-point ε-chain in X can be ε-refined to a δ-chain.*

Proof. The equivalence of (2) and (3) is obvious, since an ε-chain is the concatenation of two-point ε-chains. The equivalence of (1) and parts (2) and (3) follows from the definition of $\varphi_{\varepsilon\delta}$. Specifically, given $[\alpha]_\varepsilon \in X_\varepsilon$, the existence of an element $[\beta]_\delta \in X_\delta$ such that $[\beta]_\varepsilon = \varphi_{\varepsilon\delta}([\beta]_\delta) = [\alpha]_\varepsilon$ means precisely that α is ε-homotopic to the δ-chain β. ∎

In the following, we may occasionally need to refer to the bonding map between two covers X_δ and X_ε when either $\delta < \varepsilon$ or vice versa is possible. In this case, we will simply refer to "the map between X_δ and X_ε," and this will mean $\varphi_{\varepsilon\delta}$ if $\delta < \varepsilon$ and $\varphi_{\delta\varepsilon}$ if $\varepsilon < \delta$.

LEMMA 25. *A positive number ε is in $\mathrm{Cr}(X)$ if and only if there is a sequence $\{\varepsilon_n\}$ such that $\varepsilon_n \neq \varepsilon$ for all n, $\varepsilon_n \to \varepsilon$, and the map between X_ε and X_{ε_n} is not bijective for all n.*

Proof. If such a sequence exists, then ε cannot lie in a non-critical interval, for if it did then ε_n would also lie in that interval for all large n, contradicting that the map between X_ε and X_{ε_n} is not bijective.

Conversely, suppose $\varepsilon \in \mathrm{Cr}(X)$. Then there is no non-critical interval containing ε. So there are two positive numbers, r_1 and r_2, in $(\varepsilon - 1, \varepsilon + 1)$ such that $r_1 < r_2$ and the map $\varphi_{r_2 r_1} : X_{r_1} \to X_{r_2}$ is not bijective. We will show that at least one of the maps between X_ε and X_{r_1}, X_{r_2} must be non-bijective. In other words, we can choose ε_1 to be r_1 or r_2, and the proof is complete by iterating this process to construct $\{\varepsilon_n\}$. If ε equals r_1 or r_2, we are done. If $r_1 < r_2 < \varepsilon$, then $\varphi_{\varepsilon r_1} = \varphi_{\varepsilon r_2} \circ \varphi_{r_2 r_1}$. If both $\varphi_{\varepsilon r_1}$ and $\varphi_{\varepsilon r_2}$ were bijective, then $\varphi_{r_2 r_1}$ would be also, a contradiction. The proofs for the other two possible orderings are similar. ∎

In light of this result, there are four possible ways in which a critical value might occur. That is, $\varepsilon \in \mathrm{Cr}(X)$ if and only if one or more of the following hold (the sloped arrows indicate strictly increasing or decreasing sequences):

(1) There is a sequence, $\varepsilon_n \searrow \varepsilon$, such that the map $\varphi_{\varepsilon_n \varepsilon} : X_\varepsilon \to X_{\varepsilon_n}$ is non-injective (resp. non-surjective) for all n. In these cases, we say that X is *ε-upper non-injective* (resp. *ε-upper non-surjective*), and we call ε an *upper non-injective critical value* (resp. *upper non-surjective critical value*).

(2) There is a sequence, $\varepsilon_n \nearrow \varepsilon$, such that the map $\varphi_{\varepsilon \varepsilon_n} : X_{\varepsilon_n} \to X_\varepsilon$ is non-injective (resp. non-surjective) for all n. In these cases, we say that X is *ε-lower non-injective* (resp. *ε-lower non-surjective*), and we call ε a *lower non-injective critical value* (resp. *lower non-surjective critical value*).

As we will see in the last section, each of these possibilities may occur in a compact metric space, and in fact two or more of these cases may simultaneously hold for a given ε.

REMARK 26. Spaces for which all maps $\varphi_{\varepsilon\delta}$ are surjective exhibit a much stronger connection between critical values and topology. Such spaces—which include, for example, all geodesic spaces, all convex subsets of Euclidean space, and all Peano continua—are called *refinable*. By Definition 18, such spaces have no esssential gaps, and thus no refinement critical values.

If ε is a homotopy critical value of X, then the maps $X_\varepsilon \to X_{\varepsilon+t}$ are non-injective for all $t > 0$, meaning that X is ε-upper non-injective. The converse of this statement is true when X is compact geodesic; that is, every upper non-injective critical value is a homotopy critical value. If ε is a refinement critical value of X then the maps $X_\varepsilon \to X_{\varepsilon+t}$ are non-surjective

for all sufficiently small t, and ε is an upper non-surjective critical value. The converse of this statement is not true in general. Example 23 shows that the two notions may coincide, but it may also be the case that ε is an upper non-surjective critical value when there is no essential ε-gap, as illustrated in Examples 31 and 42. These examples also show that the second part of Lemma 27 below cannot be improved to the existence of a refinement critical value.

LEMMA 27. *Let $0 < \delta < \varepsilon$ be given. If $\varphi_{\varepsilon\delta} : X_\delta \to X_\varepsilon$ is not injective, then there is a homotopy critical value of X in the interval $[\delta, \varepsilon)$. If $\varphi_{\varepsilon\delta} : X_\delta \to X_\varepsilon$ is not surjective, then there is an upper non-surjective critical value in $[\delta, \varepsilon)$.*

Proof. If $\varphi_{\varepsilon\delta} : X_\delta \to X_\varepsilon$ is not injective, then there is a non-trivial δ-loop γ that is ε-null. Since an ε-chain is also an $(\varepsilon-t)$-chain for sufficiently small t, and since an ε-homotopy is just a finite sequence of ε-chains, it follows that an ε-homotopy is also an $(\varepsilon-t)$-homotopy for sufficiently small $t > 0$. Hence, γ is also $(\varepsilon - t)$-null for small t. Let $\varepsilon^* = \inf\{\tau \in (\delta, \varepsilon) : \gamma \text{ is } \tau\text{-null}\}$. Note that ε^* cannot be in this set, for if γ were ε^*-null, it would be $(\varepsilon^* - t)$-null for small t, contradicting that ε^* is the infimum. Thus γ is ε^*-non-trivial but $(\varepsilon^* + t)$-null for all $t > 0$, making ε^* a homotopy critical value.

If $\varphi_{\varepsilon\delta} : X_\delta \to X_\varepsilon$ is not surjective then there is a two-point ε-chain $\gamma = \{x, y\}$ that cannot be ε-refined to a δ-chain. Let $l = d(x, y) < \varepsilon$ and ε^* be the infimum of the set R of all $\tau \in (\delta, \varepsilon)$ such that γ can be ε-refined to a τ-chain. Clearly we have $\delta \le \varepsilon^* \le l < \varepsilon$. Moreover, γ cannot be ε-refined to an ε^*-chain, for such a chain would also be an $(\varepsilon^* - t)$-chain for sufficiently small $t > 0$, contradicting $\varepsilon^* = \inf R$. Suppose ε^* were not an upper non-surjective critical value. Then there would be some interval $[\varepsilon^*, \varepsilon^* + t^*)$ such that $\varepsilon^* + t^* < \varepsilon$ and, for every $0 < t < t^*$, every $(\varepsilon^* + t)$-chain can be $(\varepsilon^* + t)$-refined to an ε^*-chain. But by definition of ε^*, we can ε-refine γ to an $(\varepsilon^* + t)$-chain for some $t < t^*$. Then we could $(\varepsilon^* + t)$-refine that chain to an ε^*-chain, yielding an ε-refinement of γ to an ε^*-chain, a contradiction. ∎

LEMMA 28. *For a connected metric space X the following hold:*

(1) *ε is a lower non-injective critical value of X if and only if it is the upper limit of a strictly increasing sequence of homotopy critical values;*

(2) *if ε is an upper non-injective critical value, then it is either a homotopy critical value or the limit of a strictly decreasing sequence of homotopy critical values.*

Proof. If ε is a lower non-injective critical value of X then there exists a sequence $\varepsilon_n \nearrow \varepsilon$ such that each map $X_{\varepsilon_n} \to X_\varepsilon$ is non-injective. By Lemma 27, for each n there is a homotopy critical value ε_n^* in the inter-

val $[\varepsilon_n, \varepsilon)$. From the sequence $\{\varepsilon_n^*\}$ we may choose any strictly increasing subsequence.

Conversely, suppose there is a sequence of homotopy critical values, $\{\varepsilon_n\}$, such that $\varepsilon_n \nearrow \varepsilon$. Then for each n, there is some $t > 0$ such that $\varepsilon_n < \varepsilon_n + t < \varepsilon$ and the map $\varphi_{\varepsilon_n + t, \varepsilon_n} : X_{\varepsilon_n} \to X_{\varepsilon_n + t}$ is not injective. The map $\varphi_{\varepsilon \varepsilon_n} : X_{\varepsilon_n} \to X_\varepsilon$ is equal to the composition $\varphi_{\varepsilon, \varepsilon_n + t} \circ \varphi_{\varepsilon_n + t, \varepsilon_n}$. Since $\varphi_{\varepsilon_n + t, \varepsilon_n}$ is not injective, neither is $\varphi_{\varepsilon \varepsilon_n}$, completing the proof of the first statement.

For the second statement suppose ε is not a homotopy critical value and consider a sequence $\varepsilon_n \searrow \varepsilon$ such that each map $\varphi_{\varepsilon_n \varepsilon} : X_\varepsilon \to X_{\varepsilon_n}$ is non-injective. Then there is a homotopy critical value ε_n^* in $[\varepsilon, \varepsilon_n)$, and since $\varepsilon_n^* \neq \varepsilon$, we may choose any strictly decreasing subsequence of $\{\varepsilon_n^*\}$. ∎

LEMMA 29. *Let X be a metric space and $\varepsilon > 0$ a lower non-surjective critical value. Then there exists a sequence of upper non-surjective critical values converging up to ε. If in addition X is compact, then ε is also an upper non-injective critical value.*

Proof. The proof of the first part is an application of Lemma 27, similar to part (1) of Lemma 28. Going further, by hypothesis, there is an ε-chain $\alpha_1 = \{x_1, y_1\}$ that cannot be ε-refined to an $\varepsilon/2$-chain. Let $\varepsilon_1 = d(x_1, y_1)$, and note that $\varepsilon > \varepsilon_1 \geq \varepsilon/2$. Now, let $\tau_1 = \max\{\varepsilon_1, \varepsilon - \varepsilon/3\}$ and $\tau_1^* = (\tau_1 + \varepsilon)/2$. Then let $\alpha_2 = \{x_2, y_2\}$ be an ε-chain that cannot be ε-refined to a τ_1^*-chain. Let $\varepsilon_2 = d(x_2, y_2)$, noting that $\varepsilon_2 \geq \tau_1^* > \tau_1 \geq \varepsilon_1$ and $\varepsilon > \varepsilon_2 > \varepsilon - \varepsilon/3$. Continuing this process inductively, we obtain a sequence of ε-chains, $\alpha_n = \{x_n, y_n\}$, with $\varepsilon_n := d(x_n, y_n)$ such that (a) $\varepsilon - \frac{\varepsilon}{n+1} < \varepsilon_n < \varepsilon$ for each $n \geq 2$, and (b) for each $n \geq 1$, there is some τ strictly between $\varepsilon_n < \varepsilon_{n+1}$ such that α_{n+1} cannot be ε-refined to a τ-chain. In particular, $\varepsilon_n \nearrow \varepsilon$.

If X is compact, by choosing subsequences if necessary we may suppose that $x_n \to x$ and $y_n \to y$ for some x, y such that $d(x, y) = \varepsilon$, and $d(x_n, x_{n-1})$, $d(y_n, y_{n-1}) < \varepsilon/2$ for all n.

For each n, let $\beta_n := \{x_n, x_{n-1}, y_{n-1}, y_n\}$, and let γ_n denote the loop $\beta_n \alpha_n^{-1} = \{x_n, x_{n-1}, y_{n-1}, y_n, x_n\}$. Note that β_n is an ε-chain and γ_n is an ε-loop. For any $\lambda > \varepsilon_{n-1}$, β_n is a λ-chain, since $d(x_n, x_{n-1}), d(y_n, y_{n-1}) < \varepsilon/2 \leq \varepsilon_{n-1}$ and $d(x_{n-1}, y_{n-1}) = \varepsilon_{n-1}$. In addition, the chain α_n cannot be ε-homotopic to β_n; if this were true, then α_n would be ε-homotopic to a chain in which all distances are at most ε_{n-1}, contradicting the fact that there is some τ strictly between ε_{n-1} and ε_n such that α_n is not ε-homotopic to a τ-chain. It follows that γ_n is not ε-null. We next observe that

$$\varepsilon_{n-1} + d(x_n, x_{n-1}) \geq \varepsilon \quad \text{and} \quad \varepsilon_{n-1} + d(y_n, y_{n-1}) \geq \varepsilon.$$

In fact, if the first inequality did not hold, we would have the following

ε-nullhomotopy between γ_n and $\{x_n\}$:

(1) $\quad \gamma_n = \{x_n, x_{n-1}, y_{n-1}, y_n, x_n\} \sim \{x_n, y_{n-1}, y_n, x_n\} \sim \{x_n, y_n, x_n\}$
$$\sim \{x_n, x_n\},$$

where the first step is allowed because

$$d(x_n, y_{n-1}) \leq d(x_n, x_{n-1}) + d(x_{n-1}, y_{n-1}) = d(x_n, x_{n-1}) + \varepsilon_{n-1} < \varepsilon.$$

A similar conclusion follows if the other inequality does not hold. Finally, define

$$\delta_n := \max\{\varepsilon_{n-1} + d(x_n, x_{n-1}), \varepsilon_{n-1} + d(y_n, y_{n-1})\} \geq \varepsilon.$$

Then γ_n is $(\delta_n + 1/n)$-null, via the same sequence of steps in homotopy (1). This means that the map $X_\varepsilon \to X_{\delta_n + 1/n}$ is not injective. Since $\delta_n + 1/n \to \varepsilon$, it now follows that ε is an upper non-injective critical value. ∎

LEMMA 30. *Suppose that X is a compact metric space, ε is an upper non-surjective critical value, and there exists $\delta > \varepsilon$ such that whenever $\varepsilon < \delta_1 < \delta_2 < \delta$, the map $\varphi_{\delta_2 \delta_1}$ is injective. Then ε is a refinement critical value.*

Proof. By assumption, there is a sequence $\varepsilon_n \searrow \varepsilon$, with $\varepsilon_n < \delta$ for all n, such that each map $X_\varepsilon \to X_{\varepsilon_n}$ is not surjective. That is, for each n there is an ε_n-chain $\gamma_n = \{x_n, y_n\}$ that cannot be ε_n-refined to an ε-chain. We claim that for any $n \geq 1$, γ_n cannot be ε_1-refined to an ε-chain either. If it could be, then, equivalently, $[\gamma_n]_{\varepsilon_1}$ would lie in the image of $\varphi_{\varepsilon_1 \varepsilon}$. But by assumption, $\varphi_{\varepsilon_1 \varepsilon_n}$ is an injection, and since $\varphi_{\varepsilon_1 \varepsilon} = \varphi_{\varepsilon_1 \varepsilon_n} \circ \varphi_{\varepsilon_n \varepsilon}$, that would place $[\gamma_n]_{\varepsilon_n}$ in the image of $\varphi_{\varepsilon_n \varepsilon}$, a contradiction.

Since X is compact, by choosing subsequences if necessary we may suppose $x_n \to x$ and $y_n \to y$ for some x, y with $d(x, y) = \varepsilon$. We finish by proving that $\gamma := \{x, y\}$ is an essential gap. Suppose, to the contrary, that we can find values δ strictly greater than but arbitrarily close to ε such that γ can be δ-refined to an ε-chain. We may choose $\delta < \varepsilon_1$ and n large enough that $d(x_n, y_n) + d(x_n, x) + d(y_n, y) < \delta$ and $d(x_n, x), d(y_n, y) < \varepsilon$. Noting that $d(x, y_n) \leq d(x, x_n) + d(x_n, y_n) < \delta$, we have a δ-homotopy

$$\gamma_n = \{x_n, y_n\} \to \{x_n, x, y_n\} \to \{x_n, x, y, y_n\}.$$

Combining this with a δ-refinement of $\{x, y\}$ to an ε-chain, we obtain a δ-refinement of γ_n to an ε-chain. Since $\delta < \varepsilon_1$, this δ-refinement would also be an ε_1-refinement, which we showed previously does not exist. ∎

Proof of Theorem 5. Lemmas 28 and 29 together imply that ε must be upper non-injective or upper non-surjective. If the former holds, then Lemma 28 and the fact that this critical value is isolated imply that ε is a homotopy critical value. If ε is upper non-surjective, then since ε is an isolated critical value, the hypotheses of Lemma 30 are satisfied and ε is a refinement critical value. ∎

Compactness is only required for the second part of the proof of Theorem 5, but the theorem is false without it:

EXAMPLE 31. Let X be the union of the graphs of $f(x) = 1 + e^x$ and $g(x) = -1 - e^x$, for $x \leq 0$, and the vertical segment connecting the right endpoints of each graph. Give X the subspace metric inherited from $\mathbb{R}^2$. Pictured as a subset of $\mathbb{R}^2$, as $x \to -\infty$ the tails of this space asymptotically approach a distance of 2 from each other. In fact, 2 is an upper non-surjective critical value. To see this, take any pair of vertically aligned points, $\{(x, -1 - e^x), (x, 1 + e^x)\}$, for any $x \ll 0$. The distance between these points is $2 + 2e^x$. For all ε greater than but sufficiently close to this value, this chain can be ε-refined to a $(2 + \tau)$-chain for any $\tau > 0$, but it cannot be refined to a 2-chain. (If ε is large enough we can simply refine it around the other end.) However, there is no essential 2-gap in this space. ∎

Proof of Theorem 3. The containment $\overline{H(X)} \cup R(X) \subset \mathrm{Cr}(X)$ is clear since $\mathrm{Cr}(X)$ is closed in $\mathbb{R}_+$. Conversely, let $\varepsilon > 0$ be a critical value of X. Lemma 28 immediately handles the cases when ε is upper or lower non-injective. If ε is lower non-surjective then we may use Lemma 29 followed by Lemma 28. If ε is upper non-surjective then by Lemma 30 we need only consider the following case: there exist, for each $n \geq 1$, τ_n and δ_n such that $\varepsilon < \delta_n < \tau_n < \varepsilon + 1/n$ and the map $\varphi_{\tau_n \delta_n} : X_{\delta_n} \to X_{\tau_n}$ is non-injective. But then Lemma 27 implies there is a homotopy critical value in each interval $[\delta_n, \tau_n)$, and the proof is finished. ∎

The next few results further examine the utility and topological significance of refinement critical values.

PROPOSITION 32. *If $\{x, y\}$ is an essential ε-gap in a metric space X, then the balls $B(x, \delta)$ and $B(y, \delta)$ are disconnected for all δ greater than but sufficiently close to ε.*

Proof. Let $\alpha := \{x, y\}$ be an essential ε-gap in X. There is a $\tau > \varepsilon$ such that for all $\delta \in (\varepsilon, \tau)$, α is not δ-homotopic to an ε-chain. We claim that $B(x, \delta)$ is not connected for all such δ; a parallel argument holds for $B(y, \delta)$. In fact, $B(x, \delta)$ is not even ε-connected. If $B(x, \delta)$ were ε-connected, there would be an ε-chain $\beta = \{x = x_0, x_1, \ldots, x_n = y\}$ lying in $B(x, \delta)$. Then we could construct a δ-homotopy between β and α by just successively removing $x_1, x_2, \ldots, x_{n-1}$. But this contradicts that $\{x, y\}$ is an essential ε-gap. ∎

The Topologist's Sine Curve with its Euclidean metric has no refinement critical values but has a continuum of points at which small balls are not connected. Thus, the converse of the previous proposition is not true.

DEFINITION 33. Let (X, d) be a connected metric space with $\varepsilon > 0$. The ε-intrinsic metric determined or induced by d is the metric D_ε defined by

$$D_\varepsilon(x, y) = \inf\{L(\alpha) : \alpha \text{ is an } \varepsilon\text{-chain in } (X, d) \text{ from } x \text{ to } y\}.$$

The fact that D_ε is a metric follows from the same type of argument that shows d_ε is a metric on the ε-cover, and it follows immediately from the triangle inequality that $d \leq D_\varepsilon$. This inequality and the fact that $\{x, y\}$ is the shortest ε-chain between x and y when $d(x, y) < \varepsilon$ further imply that $D_\varepsilon(x, y) < \varepsilon$ if and only if $d(x, y) < \varepsilon$, in which case the two metrics agree. This moreover shows that not only are (X, d) and (X, D_ε) locally isometric, but also the topologies induced by d and D_ε are equivalent. Finally, for fixed $x, y \in X$, it is easy to see that $D_\varepsilon(x, y) \leq D_\delta(x, y)$ for $\delta < \varepsilon$, since any δ-chain from x to y is also an ε-chain. Thus, the function $(\varepsilon, x, y) \mapsto D_\varepsilon(x, y)$ is decreasing in ε, or increasing as $\varepsilon \to 0$.

The ε-intrinsic metric induced by a given metric can be thought of as a discretized or coarse analog of an induced length metric. If (X, d) is already a length or geodesic space, then D_ε is equal to d for every ε. Thus, one can roughly think of the difference between a given metric d and the induced metric D_ε as a measure—at a particular metric scale—of how far d is from being intrinsic in some sense. It is also interesting to note that the ε-cover, X_ε, with its natural metric d_ε, is *always* ε-intrinsic, whether (X, d) is geodesic or not.

One of the results of Theorem 2 is that, for each $\varepsilon > 0$, one can replace the given metric d on a compact connected metric space X with a topologically equivalent metric that eliminates all refinement critical values greater than ε. In particular, we have the following immediate corollary.

COROLLARY 34. *If (X, d) is a compact connected metric space such that $\inf R(X) > 0$, then there is a metric $\bar{d}$ on X that is topologically equivalent to d and is such that $(X, \bar{d})$ has no refinement critical values.*

This result also supports the idea put forth in the introduction that refinement critical values arise not as much as a result of the underlying topology of the given space as they do as a result of the particular metric one imposes on the space.

Proof of Theorem 2. Part (1) has already been noted. Consider the metric space (X, D_ε). We claim that for any $\delta > \varepsilon$, any two-point δ-chain can be δ-refined to an ε-chain, showing that there are no refinement critical values greater than ε. But this essentially follows from the definition of D_ε. If $D_\varepsilon(x, y) < \delta$, then there is—with respect to d—an ε-chain $\alpha = \{x = x_0, x_1, \ldots, x_n = y\}$ such that $L(\alpha) = \sum_{i=1}^{n} d(x_{i-1}, x_i) < \delta$. Since the metrics agree on ε-balls, α is also an ε-chain in (X, D_ε). To see that α is δ-homotopic to $\{x, y\}$ in (X, D_ε), we construct a δ-homotopy as follows. Since $\{x_0, x_1, x_2\}$

is an ε-chain with respect to d, we have $D_\varepsilon(x_0, x_2) \le d(x_0, x_1) + d(x_1, x_2) \le \sum_{i=1}^n d(x_{i-1}, x_i) < \delta$. Thus, we can remove x_1 from α via δ-homotopy to obtain $\{x_0, x_2, \ldots, x_n\}$. Similarly, $\{x_0, x_1, x_2, x_3\}$ is an ε-chain in (X, d) from x_0 to x_3, and $D_\varepsilon(x_0, x_3) \le \sum_{i=1}^3 d(x_{i-1}, x_i) \le \sum_{i=1}^n d(x_{i-1}, x_i) < \delta$. So, we can then remove x_2 from the previous chain to obtain $\{x_0, x_3, \ldots, x_n\}$. Continuing in the obvious way, we obtain a δ-homotopy in (X, D_ε) from α to $\{x = x_0, \, x_n = y\}$.

Now, assume (X, d) is compact, which implies that there is a natural number M such that any two points in (X, d) can be joined by an ε-chain having at most M points. If $x, y \in X$ and $d(x, y) < \varepsilon$, then $D_\varepsilon(x, y) = d(x, y)$ and we trivially have $D_\varepsilon(x, y) \le M d(x, y)$. Suppose $d(x, y) \ge \varepsilon$, and let α be an ε-chain in (X, d) from x to y with M or fewer points. It follows that $D_\varepsilon(x, y) \le L(\alpha) \le M\varepsilon \le M d(x, y)$. Thus, $D_\varepsilon \le Md$, and the two metrics are bi-Lipschitz equivalent. ∎

Since $D_\varepsilon(x, y)$ is increasing as ε decreases, it is natural to consider what happens when we let ε go to 0. If we set

$$D_0(x, y) = \sup_\varepsilon D_\varepsilon(x, y),$$

then it is easy to see that D_0 is also a metric on X when $D_0(x, y)$ is finite for all x and y. Furthermore, we obviously have $d(x, y) \le D_\varepsilon(x, y) \le D_0(x, y)$ for all $x, y \in X$ and $\varepsilon > 0$. In general, however, D_0 need not be finite even for a compact connected metric space. For instance, if we take the usual Koch snowflake with its subspace metric in $\mathbb{R}^2$, then $D_0(x, y)$ will be infinite for all pairs of points. Moreover, even if D_0 is finite and uniformly bounded, the resulting metric space (X, D_0) need not be topologically equivalent to the original space (X, d). If X is the boundary of the unit square in $\mathbb{R}^2$ with the vertical segments from the top boundary to the bottom attached at the points $(1/2^n, 0)$, then X with the subspace metric is compact and path connected, though not locally connected. In this case, $D_0(x, y)$ is finite for every (x, y), but the resulting space (X, D_0) is not compact. The issue that arises here is that the while D_0 is bounded, the relative distortion $D_0(x, y)/d(x, y)$ is unbounded.

Proof of Theorem 6. The proof that (X, D_0) has no refinement critical values follows as in the proof of Theorem 2. In fact, if $\varepsilon < \delta$ and $\{x, y\}$ is a δ-chain with respect to D_0, then $D_\varepsilon(x, y) \le D_0(x, y) < \delta$, and the preceding proof goes through without change.

For the second part, it suffices to prove that for every $x, y \in X$ there is a midpoint between x and y (cf. [5] or [8]), i.e. a point $m \in X$ such that $D_0(x, m) = D_0(y, m) \le D_0(x, y)/2$. By definition of D_0 for every natural number i there is a $1/i$-chain $\alpha := \{x = x_0, \ldots, x_n = y\}$ such that $L(\alpha) < D_0(x, y) + 1/i$. Let j_i be the largest index such that $L(\{x_0, \ldots, x_{j_i}\}) \le$

$D_0(x, y)/2$. By the triangle inequality,

$$L\left(\{x_0, \ldots, x_{j_i}\}\right) \geq L\left(\{x_0, \ldots, x_{j_i+1}\}\right) - \frac{1}{i} > \frac{D_0(x, y)}{2} - \frac{1}{i}.$$

Setting $m_i := x_{j_i}$ we have by definition $D_{1/i}(x, m_i) \leq D_0(x, y)/2$. On the other hand,

$$D_{1/i}(y, m_i) \leq L(\{m_i = x_{j_i}, \ldots, y\}) = L(\alpha) - L(\{x_0, \ldots, x_{j_i}\})$$
$$< D_0(x, y) + \frac{1}{i} - \left(\frac{D_0(x, y)}{2} - \frac{1}{i}\right) = \frac{D_0(x, y)}{2} + \frac{2}{i}.$$

Combining these inequalites we obtain

$$D_{1/i}(x, m_i) \leq \frac{D_0(x, y)}{2} \quad \text{and} \quad D_{1/i}(y, m_i) \leq \frac{D_0(x, y)}{2} + \frac{2}{i}, \quad i \geq 1.$$

By choosing a subsequence if necessary we may assume that $m_i \to m \in X$. Since X is compact, it follows from Dini's Theorem (cf. [12]) that $D_{1/i}$ actually converges to D_0 uniformly. Thus, given any $k > 0$, we have $D_0(x, m_i) - 1/k < D_{1/i}(x, m_i) \leq D_0(x, m_i)$ and $D_0(y, m_i) - 1/k < D_{1/i}(y, m_i) \leq D_0(y, m_i)$ for all large i. For a fixed k, it follows that

$$D_0(x, m_i) - \frac{1}{k} < \frac{D_0(x, y)}{2} \quad \text{and} \quad D_0(y, m_i) - \frac{1}{k} < \frac{D_0(x, y)}{2} + \frac{2}{i}$$

for all sufficiently large i. Letting $i \to \infty$ in these inequalities, we obtain

$$D_0(x, m) - \frac{1}{k} \leq \frac{D_0(x, y)}{2} \quad \text{and} \quad D_0(y, m) - \frac{1}{k} \leq \frac{D_0(x, y)}{2}.$$

Finally, we let $k \to \infty$, and this is the desired result. ∎

Note that an immediate corollary of the above proof is that (X, D_ε) converges in the uniform sense (cf. [5])—hence, in the Gromov–Hausdorff sense—to (X, D_0) as $\varepsilon \to 0$.

4. Examples of non-discrete spectra. In this section we will present several examples illustrating some of the phenomena that can occur concerning the critical spectrum of a general compact metric space. We will first prove some technical results that facilitate identifying refinement critical values. Roughly speaking, the following construction yields a method for detecting or constructing essential gaps. Indeed, the following discussion and definition should make it clear why the name 'essential gap' is appropriate for the structure that induces a refinement critical value.

Let X be a connected metric space. Assume there are points, $x, y \in X$, with $d(x, y) = l > 0$, and a number $\varepsilon^* > l$ such that the following holds: for each ε in the interval $(l, \varepsilon^*]$, if we let $B_x = B(x, \varepsilon - l)$ and $B_y = B(y, \varepsilon - l)$, then we can express X as a disjoint union, $X = Z \cup Y$, such that

(1) $B_x \subset Z$ and $B_y \subset Y$ (hence $B_x \cap B_y = \emptyset$),
(2) the only points in Z that are strictly within ε of a point in B_y lie in B_x, and the only points of Y that are strictly within ε of a point in B_x lie in B_y.

If these conditions hold, we call $\{x, y\}$ a *pre-essential gap*. A pre-essential gap need *not* be an essential gap. Note also that though it behaves like one locally around x and y, $\{Z, Y\}$ need not be a disconnection of X. As we have already seen, connected spaces can have refinement critical values.

Given a pre-essential gap, $\{x, y\}$, with $d(x, y) = l < \varepsilon \leq \varepsilon^*$ as above, let $\gamma = \{x_0, \ldots, x_n\}$ be any ε-chain in X. A pair of consecutive points, (x_{i-1}, x_i), $1 \leq i \leq n$, will be said to *contain* or *cross the x, y-gap* if x_{i-1} lies in either B_x or B_y and x_i lies in the other ball. Assign to each pair of consecutive points a value $|x_{i-1}, x_i|$ as follows:

$$|x_{i-1}, x_i| = \begin{cases} 0, & (x_{i-1}, x_i) \text{ does not contain the } x, y\text{-gap}, \\ 1, & x_{i-1} \in B_x, \ x_i \in B_y, \\ -1, & x_{i-1} \in B_y, \ x_i \in B_x. \end{cases}$$

Note that the order of the points in the notation $|x_{i-1}, x_i|$ does matter; the second case in this definition, for instance, occurs when the first point of the pair lies in B_x and the second point lies in B_y, while the third case occurs when the opposite holds. Now, define $\mathscr{G}(\gamma; x, y, \varepsilon) := \sum_{i=1}^{n} |x_{i-1}, x_i|$. We call this the *$(x, y, \varepsilon)$-gap number of γ*; it measures the net number of times γ crosses the x, y-gap.

LEMMA 35. *Assume that the above conditions hold for some $\varepsilon^* > l = d(x, y)$, so that $\{x, y\}$ is a pre-essential gap. Given ε such that $l < \varepsilon \leq \varepsilon^*$, the integer $\mathscr{G}(\gamma; x, y, \varepsilon)$ is an ε-homotopy invariant. That is, for fixed $\varepsilon \in (l, \varepsilon^*]$, if α and γ are ε-chains such that $\alpha \sim_\varepsilon \gamma$, then $\mathscr{G}(\gamma; x, y, \varepsilon) = \mathscr{G}(\alpha; x, y, \varepsilon)$.*

Proof. Since any ε-homotopy taking γ to α will consist of a finite sequence of basic moves, it suffices to prove the result in the case where α is obtained by adding or removing a single point to/from γ. The proof is not difficult, but it is a tedious process in working through all the possible cases. We will prove one case to illustrate the reasoning used. The rest of the cases follow in exactly the same manner.

Let $\gamma = \{x_0, x_1, \ldots, x_n\}$, and assume that α is obtained by adding z between x_{i-1} and x_i. Since this basic move only affects three different pairs of points in the sums defining the (x, y, ε)-gap numbers of γ and α, we only need to show that $|x_{i-1}, x_i| = |x_{i-1}, z| + |z, x_i|$.

Assume that $|x_{i-1}, x_i| = 0$. If $|x_{i-1}, z| = |z, x_i| = 0$, then the result is clear. If $|x_{i-1}, z| = 1$ and $|z, x_i| = -1$ (or $|x_{i-1}, z| = -1$ and $|z, x_i| = 1$), the result is also clear. The subcase $|x_{i-1}, z| = 1 = |z, x_i|$ cannot occur,

for the first equality would imply that $x_{i-1} \in B_x$ and $z \in B_y$, while the second would imply that $z \in B_x$ and $x_i \in B_y$, which would further imply that $z \in B_x \cap B_y$, a contradiction. The case $|x_{i-1}, z| = -1 = |z, x_i|$ cannot occur either, for the first equality would imply $x_{i-1} \in B_y$ and $z \in B_x$, while the second would imply that $z \in B_y$ and $x_i \in B_x$, another contradiction. Suppose $|x_{i-1}, z| = 1$ and $|z, x_i| = 0$. Then $x_{i-1} \in B_x$, $z \in B_y$, and x_i cannot be in B_x or B_y (or else we would have $|x_{i-1}, x_i| = 1$ in the latter case and $|z, x_i| = -1$ in the former). But x_i must lie in Z or Y, and x_i is strictly within ε of x_{i-1}, a point in B_x, and strictly within ε of z, a point in B_y. If $x_i \in Z$, then, since $d(x_i, z) < \varepsilon$, condition (2) above implies that $x_i \in B_x$, a contradiction. If $x_i \in Y$, then $d(x_i, x_{i-1}) < \varepsilon$ implies that $x_i \in B_y$, another contradiction. Similar reasoning applies to the cases $|x_{i-1}, z| = -1$ and $|z, x_i| = 0$, $|x_{i-1}, z| = 0$ and $|z, x_i| = 1$, and $|x_{i-1}, z| = 0$ and $|z, x_i| = -1$. Thus, given that $|x_{i-1}, x_i| = 0$, the only possible cases that can occur result in the equality $|x_{i-1}, x_i| = |x_{i-1}, z| + |z, x_i|$. Proceeding, one would argue simliarly for the cases $|x_{i-1}, x_i| = \pm 1$, and then work through the same procedure in the case where a point is removed from γ to obtain α. All cases that can occur lead to the desired equality, thus proving the result. ∎

LEMMA 36 (Essential Gap Lemma). *Let X be a chain connected metric space, and suppose $\{x, y\}$ is a pre-essential gap with $d(x, y) = l$. If $\mathrm{dist}(B(x, r), B(y, r)) = d(x, y)$ for all sufficiently small r, then $\{x, y\}$ is an essential l-gap.*

Proof. Let $\varepsilon^* > l$ be as in the definition of a pre-essential gap; we may assume that $\varepsilon^* - l$ is small enough that $\mathrm{dist}(B(x, r), B(y, r)) = d(x, y)$ for all $r \leq \varepsilon^* - l$. Fix ε so that $l < \varepsilon \leq \varepsilon^*$. Then $\gamma := \{x, y\}$ is an ε-chain, and $\mathscr{G}(\gamma; x, y, \varepsilon) = 1$. No l-chain can cross the x, y-gap. In fact, if $\{z_0, \ldots, z_n\}$ is an l-chain, and if we had $z_{i-1} \in B_x = B(x, \varepsilon - l)$ and $z_{i-1} \in B_y = B(y, \varepsilon - l)$, then we would have $\mathrm{dist}(B_x, B_y) \leq d(z_{i-1}, z_i) < l$, contradicting the fact that $\mathrm{dist}(B_x, B_y) = d(x, y) = l$. Thus, the (x, y, ε)-gap number of any l-chain must be 0. The ε-homotopy invariance of this value then implies that $\gamma = \{x, y\}$ is not ε-homotopic to an l-chain. Since $\varepsilon \in (l, \varepsilon^*)$ was arbitrary, it follows that $\{x, y\}$ is an essential l-gap. ∎

A partial converse to the Essential Gap Lemma also holds. However, this condition alone is not sufficient to ensure that $\{x, y\}$ is an essential gap.

LEMMA 37. *If $\{x, y\}$ is an essential gap in X, then $\mathrm{dist}(B(x, r), B(y, r)) = d(x, y)$ for sufficiently small r.*

Proof. Suppose the conclusion does not hold. Let $l = d(x, y)$, and let $\varepsilon > l$ be given. Choose $r < \min\{l, (\varepsilon - l)/2\}$ such that there are points $u \in B(x, r)$, $v \in B(y, r)$ satisfying $\mathrm{dist}(B(x, r), B(y, r)) \leq d(u, v) < d(x, y) = l$. We can transform $\{x, y\}$ via ε-homotopy as follows: $\{x, y\} \to \{x, u, y\} \to$

$\{x, u, v, y\}$. The second step is valid since $d(u, y) \le d(u, v) + d(v, y) < l + r < (\varepsilon + l)/2 < \varepsilon$. Moreover, this last chain is an l-chain, since $d(u, v) < d(x, y) = l$ and $r < l$. In other words, $\{x, y\}$ can be ε-refined to an l-chain for all ε sufficiently close to l, contradicting the hypothesis. ∎

This is a rather technical set-up, and it may not yet be visually clear what this structure looks like. The following can be taken as a sort of canonical example illustrating this concept.

EXAMPLE 38. Let L, l_1, l_2, and h be positive real numbers such that L is significantly larger than l_1 (say, $L > 3l_1$), $l_2 \le l_1$, and $h^2 + (l_1 + l_2)^2/4 > l_1^2$. Let X be the metric subspace of $\mathbb{R}^2$ shown in Figure 2, and let x, y, u, and v be the points $\left(\frac{L-l_1}{2}, h\right)$, $\left(\frac{L+l_1}{2}, h\right)$, $\left(\frac{L-l_2}{2}, 0\right)$, and $\left(\frac{L+l_2}{2}, 0\right)$, respectively. Let d be the diagonal distance from x to v (or y to u, by symmetry). The condition $h^2 + (l_1 + l_2)^2/4 > l_1^2$ implies that $d > l_1$. Let ε^* be such that $l_1 < \varepsilon^* < \min\{d, 2l_1, (L - l_1)/2\}$. Now, fix any ε with $l_1 < \varepsilon \le \varepsilon^*$. Let Z be the left half of X, and let Y be the right half. The conditions $\varepsilon < 2l_1$ and $\varepsilon < (L - l_2)/2$ ensure that the balls $B_x := B(x, \varepsilon - l_1)$ and $B_y := B(y, \varepsilon - l_1)$ do not intersect or extend to the vertical sides of X.
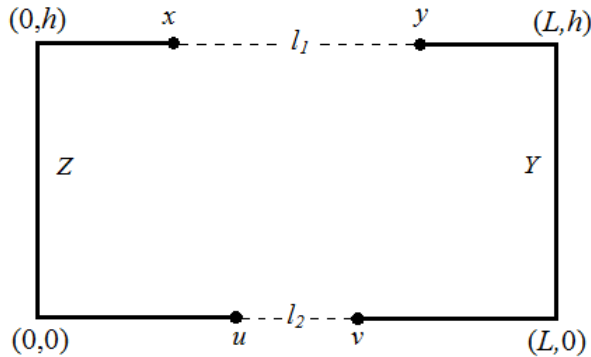

Fig. 2. An essential gap

Suppose that $z \in Z$ and is strictly within ε of a point in B_y. Clearly, z cannot lie on the vertical segment of Z. If z lies on the lower boundary of Z, then the closest it could be to any point in B_y is d, which occurs when $z = u$. But $d > \varepsilon^* \ge \varepsilon$, so z cannot, in fact, lie on this lower boundary. If z lies on the upper boundary of Z but outside of B_x, then it is at least $\varepsilon - l_1 + l_1 = \varepsilon$ away from any point of B_y. Thus, it must be the case that z lies in B_x. Likewise, by symmetry, if $z \in Y$ and is strictly within ε of a point in B_x, then $z \in B_y$. Moreover, we clearly have $\text{dist}(B(x, r), B(y, r)) = d(x, y)$ for sufficiently small r. Therefore, $\{x, y\}$ is an essential l-gap, and l is a refinement critical value.

To see what makes this essential gap phenomenon occur, consider the trapezoid $\{x, u, v, y\}$. The diagonals of this trapezoid are longer than the longest base of the trapezoid. This, essentially, is why $\{x, y\}$ cannot be ε-refined to an l_1-chain for $\varepsilon \in (l_1, \varepsilon^*)$. Adding v in between x and y is not allowed, because $\varepsilon \leq \varepsilon^* < d = d(x, v)$. Likewise, one cannot jump from x to u to y for the same reason. In other words, because the diagonal is too long, one cannot overcome the $\{x, y\}$-gap by going around it, at least via "hops" that are sufficiently close to l_1 in length. Note that if $d \leq l_1$ then we could, in fact, go around the $\{x, y\}$-gap. Indeed, for any $\varepsilon > l_1$, we could then transform the ε-chain, $\{x, y\}$, via ε-homotopy by adding v and then u. So, it is the diagonal length that makes this gap essential.

Finally, X is not connected, but we could attach a long joining curve to X to make it path connected, without affecting the critical value. ∎

Now, we will use the Essential Gap Lemma to produce examples of compact metric spaces having critical spectra with positive limit points. Moreover, these examples will show that critical values of one type can converge to critical values of the other type. They will also illustrate some of the other properties mentioned in the previous section.

EXAMPLE 39. We define the following sets:

(1) For $n \geq 0$, $A_n = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, y = 1/2^n\} \cup \{(x, y) \in \mathbb{R}^2 : 2 \leq x \leq 3, y = 1/2^n\}$.

(2) $A_\infty = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, y = 0\} \cup \{(x, y) \in \mathbb{R}^2 : 2 \leq x \leq 3, y = 0\}$.

(3) $B_1 = \{(x, y) \in \mathbb{R}^2 : x = 0, 0 \leq y \leq 2\}$, $B_2 = \{(x, y) \in \mathbb{R}^2 : x = 3, 0 \leq y \leq 2\}$.

(4) $C = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 3, y = 2\}$.

Define a metric subspace of $\mathbb{R}^2$ by $X = (\bigcup_{n=0}^\infty A_n) \cup A_\infty \cup B_1 \cup B_2 \cup C$. For $n \geq 0$, let $x_n = (1, 1/2^n)$ and $y_n = (2, 1/2^n)$, and let $x_\infty = (1, 0)$, $y_\infty = (2, 0)$, $z_0 = (3/2, 2)$. Let $d_0 = d(x_0, z_0)$ and, for $n \geq 1$, let $d_n = d(x_{n-1}, y_n)$. Note that $d_0 = d_1$. For $m > n \geq 0$, let $d_m^n = d(x_n, y_m)$, and note that $d_n^{n-1} = d_n$ for $n \geq 1$. See Figure 3 below. We call X a "Rapunzel's Comb."

The following results can be easily verified: (a) $1 < d_n < d_{n-1}$ for all $n \geq 1$, and $d_n \searrow 1$ as $n \to \infty$; (b) $d_m^n > 1$ for all $m > n \geq 0$, and for fixed n, d_m^n is minimized when $m = n + 1$. Now, fix $n \geq 1$. It is evident that $\operatorname{dist}(B(x_n, r), B(y_n, r)) = d(x_n, y_n)$ for sufficiently small r. Fix any ε such that $d(x_n, y_n) = 1 < \varepsilon \leq d_{n+1}$. Let Z be the left half of X, including z_0 (so Z is closed), and let Y be the rest of the space. Let $B_{x_n} = B(x_n, \varepsilon - 1) \subset Z$ and $B_{y_n} = B(y_n, \varepsilon - 1) \subset Y$. Suppose $z \in Z$ is strictly within ε of a point in B_{y_n}. Clearly $z \notin B_1$, and $z \notin C$. In fact, the closest any point of $C \cap Z$ can be to B_{y_n} is the distance from z_0 to y_n, which is greater than d_0. But
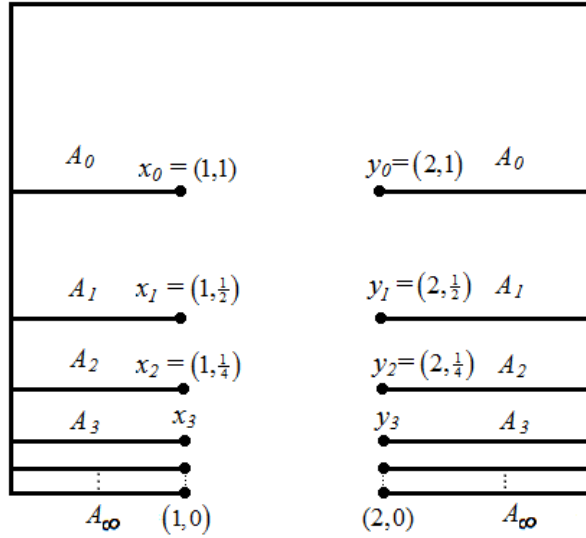
Fig. 3. Rapunzel's Comb

$d_0 = d_1 > d_{n+1} \geq \varepsilon$, so $z \notin C$. Thus, z is in the left half of one of the sets, A_k. However, z cannot be in A_k for $0 \leq k < n$, since in that case the distance between z and B_{y_n} would be at least d_n^k, which in turn is at least as great as $d_n^{n-1} = d_n$. Since $d_n > d_{n+1} \geq \varepsilon$, this shows that this case cannot occur. Hence, z must lie in A_m for some $m \geq n$. If z were in A_m for $m > n$, the distance between z and B_{y_n} would be at least d_m^n. For fixed n, d_m^n is minimized when $m = n + 1$, so the distance between z and any point of B_{y_n} is at least $d_{n+1}^n = d_{n+1} \geq \varepsilon$. This contradicts that z is strictly within ε of a point of B_{y_n}. Therefore, $z \in A_n$, and since it is within ε of a point of B_{y_n}, it must lie in B_{x_n}. By symmetry, if $z \in Y$ and is strictly within ε of a point of B_{x_n}, then $z \in B_{y_n}$. Hence, $\{x_n, y_n\}$ is an essential 1-gap, and 1 is a refinement critical value.

Fix $n \geq 1$, and let $\gamma_n = \{x_n, x_{n+1}, y_{n+1}, y_n, x_n\}$. For $1 < \varepsilon \leq d_{n+1}$, γ_n is an ε-loop with (x_n, y_n, ε)-gap number -1. An ε-null loop has (x_n, y_n, ε)-gap number 0, implying that γ_n is ε-nontrivial. This holds for all $1 < \varepsilon \leq d_{n+1}$. However, for $\varepsilon > d_{n+1}$, γ_n is easily seen to be trivial. Thus, $d_{n+1} \in H(X)$, and $d_n \searrow 1$, showing that $\mathrm{Cr}(X)$ is not discrete. Note that X is compact, path connected, and even simply connected.

This example illustrates some other interesting properties. There are infinitely many essential 1-gaps, but there are no non-trivial 1-loops in X. Thus, $\pi_1(X)$ is trivial, and the map $\varphi_{\varepsilon,1} : X_1 \to X_\varepsilon$ is injective for all ε sufficiently close to 1. So, there is a sequence, d_n, of critical values of one type converging to a critical value, 1, of an entirely different type. ∎

There are many different variations on Rapunzel's Comb that one can use to illustrate critical value limiting behavior. All of them use the Essential Gap Lemma in some form, and the details follow much as before.

EXAMPLE 40 (Rapunzel's Comb, variation 1). For $n \geq 1$, let $h_n = 2^{-n/2}$ and $H = \sum_{n=1}^{\infty} h_n = 1 + \sqrt{2}$. Let X be the subspace of $\mathbb{R}^2$ shown in Figure 4 below (purposefully not drawn exactly to scale to show detail). Here, the gaps increase in length to a limiting gap of length 1.

We also define the following: $z_0 = (3/2, H + 2)$, $d_0 = d(x_\infty, z_0) = d(y_\infty, z_0)$, $d_n = d(x_n, y_{n+1})$ for all $n \geq 1$, $x_\infty = (1, H)$, $y_\infty = (2, H)$, and

$$ x_n = \left(1 + \frac{1}{2^{n+1}}, \sum_{i=1}^{n-1} h_i\right), \quad y_n = \left(2 - \frac{1}{2^{n+1}}, \sum_{i=1}^{n-1} h_i\right), \quad \forall n \geq 1. $$

Reasoning as in the previous example, one can show that each $\{x_n, y_n\}$, for $n \geq 1$, is an essential $(1 - 1/2^n)$-gap. These values converge up to 1, but 1 is not a refinement critical value or even an upper non-surjective critical value. In fact, for every $\varepsilon > 1$, every ε-chain can be ε-refined to a 1-chain, even the chain $\{x_\infty, y_\infty\}$. Given any $\varepsilon > 1$, the distances $d(x_\infty, y_n)$ and

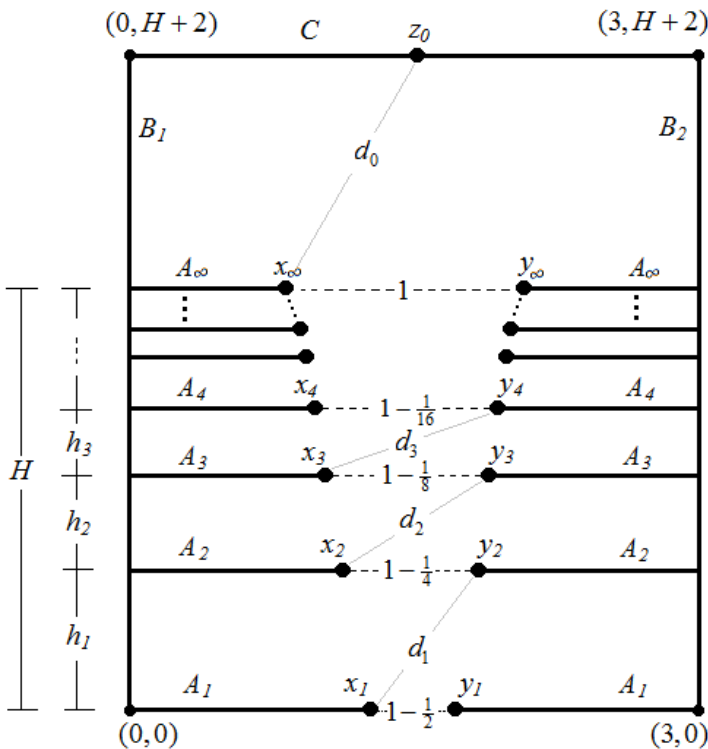Fig. 4. Rapunzel's Comb, variation 1

$d(y_\infty, x_n)$ are eventually less than ε, thus making this refinement possible. So, what type of critical value is 1?

Fix $n \geq 2$, and let γ_n be the loop $\{x_n, x_{n-1}, y_{n-1}, y_n, x_n\}$. For all ε greater than $1 - 1/2^n$, γ_n is an ε-loop. Since $\{x_n, y_n\}$ is an essential gap, we also know that, for each ε greater than but sufficiently close to $1 - 1/2^n$, the (x_n, y_n, ε)-gap number is an ε-homotopy invariant. Fixing any such ε, we see that the ε-chain $\alpha_n := \{x_n, y_n\}$ has non-zero (x_n, y_n, ε)-gap number, while the ε-chain $\beta_n := \{x_n, x_{n-1}, y_{n-1}, y_n\}$ does not cross the x_n, y_n-gap at all. Hence, $\gamma_n = \beta_n \alpha_n^{-1}$ cannot be ε-null for such an ε. So, γ_n is ε-non-trivial for all ε greater than but sufficiently close to $1 - 1/2^n$. On the other hand, γ_n is 1-null. In fact, since the diagonals between x_n and y_{n-1} are less than 1 in length (this is easily verified), we can successively remove y_{n-1}, x_{n-1}, and y_n from γ_n, giving us the trivial chain. Therefore, 1 is a lower non-injective critical value that is an upper limit of sequences of refinement critical values and homotopy critical values. Note that, as before, X is compact, path connected, and simply connected. ∎

EXAMPLE 41 (Rapunzel's Comb, variation 2). The construction of this example is very similar to the previous case. In fact, the lengths of the gaps will be the same. The key difference will be changing the heights between the gaps in the comb. We increase them just enough so that the diagonal lengths between x_n and y_{n+1} are greater than 1 for each n, but still small enough so that the sum of the heights is finite.

So, for $n \geq 1$, let $h_n = \sqrt{3}/(\sqrt{2})^n$, and let $H = \sum_{n=1}^\infty h_n = \sqrt{3} + \sqrt{6}$. We define X exactly as in the previous example except for the different values h_n, and we similarly define the points x_n, y_n, x_∞, y_∞, and z_0. Also as before, we let $d_n = d(x_n, y_{n+1})$, so that d_n is the length of the diagonal between x_n and y_{n+1}. Since the construction is the same, Figure 4 holds equally well for this example. We just need to keep in mind that the heights, h_n, and therefore the diagonals, d_n, are larger in this case. One can verify by direct computation the following:

$$d_n^2 = 1 + \frac{3}{2^{n+1}} + \frac{9}{2^{2n+4}}, \quad d_{n+1} < d_n \quad \text{for all } n, \quad d_n \searrow 1,$$
$$d(x_n, y_m) > 1 \quad \text{for all } 1 \leq n < m.$$

In addition, for fixed n and $m > n$, the diagonal lengths, $d(x_n, y_m)$, increase as m increases.

Now, fix $n \geq 2$, and, recalling that $d(x_n, y_n) = 1 - 1/2^n$, let ε be such that

$$1 - 1/2^n < \varepsilon \leq \min\{d(x_1, y_n), \ldots, d(x_{n-1}, y_n), d(x_{n+1}, y_n), 1 + h_n - 1/2^n\}.$$

The condition that ε be less than or equal to $1 + h_n - 1/2^n$ is to ensure that the ball of radius $\varepsilon - (1 - 1/2^n)$ centered at x_n (or y_n) does not intersect any

nearby teeth of the comb or either of the vertical sides of X. That is, these balls are just segments of the teeth of the comb formed by A_n. As before, we let Z be the left half of X, and we let Y be the right half. Suppose $z \in Z$ and lies within ε of a point of $B_{y_n} := B(y_n, \varepsilon - (1 - 1/2^n))$. Clearly, z cannot lie in C or B_1. If z were in A_m for some $m \leq n-1$, the distance between z and any point of B_{y_n} would be at least $d(x_m, y_n)$. But $\varepsilon \leq d(x_m, y_n)$ for such m, so this cannot occur. If z were in A_{n+1}, then the closest z could be to any point of B_{y_n} is $d_n = d(x_{n+1}, y_n)$, but again, we have $\varepsilon \leq d(x_{n+1}, y_n)$. So, this cannot occur either. Neither can z be in A_m for $m > n+1$, since the diagonal lengths, $d(x_m, y_n)$, are greater than d_n for $m > n+1$. Hence, z must lie in A_n, and in fact it must lie in $B(x_n, \varepsilon - (1 - 1/2^n))$. By symmetry, the same result holds if $z \in Y$ and lies within ε of a point in $B(x_n, \varepsilon - (1 - 1/2^n))$. We also have $\operatorname{dist}(B(x_n, r), B(y_n, r)) = d(x_n, y_n)$ for sufficiently small r, so $\{x_n, y_n\}$ is an essential $(1 - 1/2^n)$-gap. It follows that $1 - 1/2^n = d(x_n, y_n)$ is an upper non-surjective critical value; for all ε greater than but sufficiently close to $1 - 1/2^n$, $\{x_n, y_n\}$ is an ε-chain that cannot be ε-refined to a $(1 - 1/2^n)$-chain.

Finally, since $1 - 1/2^n \nearrow 1$, we know that 1 is a critical value. Note that

$$1 < \min\{d(x_1, y_n), \ldots, d(x_{n-1}, y_n), d(x_{n+1}, y_n), 1 + h_n - 1/2^n\},$$

because all diagonals have length greater than 1 and

$$h_n > \frac{1}{(\sqrt{2})^n} > \frac{1}{2^n}, \quad \text{so} \quad 1 + h_n - \frac{1}{2^n} > 1.$$

Thus, for $\varepsilon = 1$, the (x_n, y_n, ε)-gap number of an ε-chain is an ε-homotopy invariant. Now, $\{x_n, y_n\}$ is a 1-chain, and its $(x_n, y_n, 1)$-gap number is 1. However, no $(1 - 1/2^n)$-chain can cross the x_n, y_n-gap. Thus, $\{x_n, y_n\}$ cannot be 1-homotopic to a $(1 - 1/2^n)$-chain. In other words, the map $\varphi_{1, 1-1/2^n} : X_{1-1/2^n} \to X_1$ is not surjective, and this holds for all $n \geq 1$. Hence, 1 is a lower non-surjective critical value. As we have shown, such a critical value can only occur as the upper limit of upper non-surjective critical values. ∎

EXAMPLE 42 (Rapunzel's Comb, variation 3). Working as in the previous examples, let X be the variation of Rapunzel's Comb shown in Figure 5. Here, we have a sequence of gaps of length $1 + 1/2^n$ converging down to a gap of length 1. Moreover, we have attached an extra single set of teeth below the limiting gap, which adds a gap of length $l < 1$.

We choose the heights, h_n, to be $h_n = 1/(\sqrt{2})^n$, and we choose h so that $d(x_\infty, b) = d(y_\infty, a) = 1$ and $d(x_\infty, a), d(y_\infty, b) < 1$. A straightforward computation shows that this can be done. It can then be seen as in the previous examples that $\{x_n, y_n\}$ is an essential $(1 + 1/2^n)$-gap for each n. Thus, for each n and all ε greater than but sufficiently close to $1 + 1/2^n$, $\{x_n, y_n\}$ cannot be ε-refined to a $(1 + 1/2^n)$-chain, and thus to a 1-chain either. Hence, 1 is an upper non-surjective critical value. However, there is
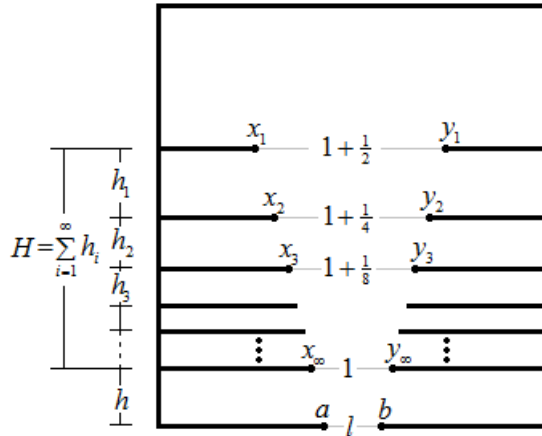
Fig. 5. Rapunzel's Comb, variation 3

no essential 1-gap in X. Because we have added the gap $\{a, b\}$ below the gap $\{x_\infty, y_\infty\}$, and chosen h so that $d(x_\infty, b) = d(y_\infty, a) = 1$, the chain $\{x_\infty, y_\infty\}$ can be ε-refined to a 1-chain for all ε greater than 1. The homotopy itself is simply

$$\{x_\infty, y_\infty\} \rightarrow \{x_\infty, a, y_\infty\} \rightarrow \{x_\infty, a, b, y_\infty\}.$$

Since $l < 1$, this is a 1-chain. This example shows that—even in a compact space—an upper non-surjective critical value need not correspond to an essential gap. As we have already shown, the two only correspond for certain when the critical value in question is isolated. ∎

References

[1]   M. Anderson, *Short geodesics and gravitational instantons*, J. Differential Geom. 31 (1990), 265–275.

[2]   V. Berestovski and C. Plaut, *Covering group theory for topological groups*, Topology Appl. 114 (2001), 141–186.

[3]   V. Berestovski and C. Plaut, *Uniform universal covers of uniform spaces*, Topology Appl. 154 (2007), 1748–1777.

[4]   V. Berestovskii, C. Plaut and C. Stallman, *Geometric groups. I*, Trans. Amer. Math. Soc. 351 (1999), 1403–1422.

[5]   D. Burago, Y. Burago and S. Ivanov, *A Course in Metric Geometry*, Grad. Stud. Math. 33, Amer. Math. Soc., 2001.

[6]   M. Gromov, *Metric Structures for Riemannian and Non-Riemannian Spaces*, Progr. Math. 152, Birkhäuser Boston, 1999.

[7]   J. Kigami, *Analysis on Fractals*, Cambridge Univ. Press, 2008.

[8]   C. Plaut, *Metric spaces of curvature* $\geq k$, in: Handbook of Geometric Topology, North-Holland, Amsterdam, 2002, 819–898.

[9]   C. Plaut, *Quotients of uniform spaces*, Topology Appl. 153 (2006), 2430–2444.

[10] C. Plaut and J. Wilkins, *Discrete homotopies and the fundamental group*, Adv. Math. 232 (2013), 271–294.

[11] C. Plaut and J. Wilkins, *Essential circles and Gromov–Hausdorff convergence of covers*, arXiv:1304.0808.

[12] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill, 1976.

[13] Z. Shen and G. Wei, *On Riemannian manifolds of almost nonnegative Ricci curvature*, Indiana Univ. Math. J. 40 (1991), 551–565.

[14] C. Sormani and G. Wei, *The covering spectrum of a compact length space*, J. Differential Geom. 67 (2004), 35–77.

[15] C. Sormani and G. Wei, *The cut-off covering spectrum*, Trans. Amer. Math. Soc. 362 (2010), 2339–2391.

[16] E. Spanier, *Algebraic Topology*, Springer, 1981.

[17] R. Strichartz, *Differential Equations on Fractals*, Princeton Univ. Press, 2006.

[18] J. Wilkins, *Discrete geometric homotopy theory and critical values of metric spaces*, PhD thesis, Univ. of Tennessee, 2011; http://trace.tennessee.edu/utk_graddiss/1039.

[19] J. Wilkins, *The revised and uniform fundamental groups and universal covers of geodesic spaces*, Topology Appl. 160 (2013), 812–835.

Jim Conant, Conrad Plaut
Department of Mathematics
University of Tennessee
Knoxville, TN 37996, U.S.A.
E-mail: jconant@math.utk.edu
          cplaut@math.utk.edu

Corey Jones
Department of Mathematics
Vanderbilt University
Nashville, TN 37240, U.S.A.
E-mail: corey.m.jones@vanderbilt.edu

Maria Lusby
17 Sutherland Rd.
Hicksville, NY 11801, U.S.A.
E-mail: mariawalpole@gmail.com

Victoria Curnutte
7078 W. Rainbow Rd.
Sedalia, CO 80135, U.S.A.
E-mail: vcurnutte@gmail.com

Kristen Pueschel
Department of Mathematics
Cornell University
Ithaca, NY 14853-4201, U.S.A.
E-mail: kp338@cornell.edu

Jay Wilkins
Department of Mathematics and
Computer Science
University of North Carolina at Pembroke
Pembroke, NC 28372, U.S.A.
E-mail: leonard.wilkins@uncp.edu