CONVEX HULLS IN COARSE MEDIAN SPACES

BRIAN H. BOWDITCH

Abstract. We describe the geometry coarse convex hulls in coarse median spaces. The main results apply when the space comes equipped with a family of projection maps to hyperbolic spaces satisfying certain natural conditions. We show that coarse hulls in such spaces are quasi-isometric to CAT(0) cube complexes. From this we deduce a distance formula, and show the existence of monotone paths connecting pairs of points. In particular, we recover a number of results about hierarchically hyperbolic spaces. Among the main motivating examples are the mapping class groups and Teichmüller spaces of compact surfaces.

1. Introduction

The aim of this paper is to describe the geometry of coarse convex hulls in coarse median spaces equipped with a family of maps to hyperbolic spaces satisfying certain hypotheses. From this one can derive a number of consequences, such as a "distance formula" and the existence of "monotone paths" between two points. Related statements have been obtained for "hierarchically hyperbolic" spaces in [BeHS2] and [BeHS3]. The arguments of the present paper are somewhat different. (In particular, our description of hulls does not rely on the distance formula a-priori.) We will reduce the results to more combinatorial statements about cube complexes, which may have some intrinsic interest in that context. In the course of doing this, we give an account of coarse hulls in general finite-rank coarse median spaces (Section 6).

It is well known that hyperbolic spaces as originally defined by Gromov can be characterised as having a certain "treelike structure". There are several ways in which this principle can be expressed. Probably the most useful formulation features in the original paper [G] where finite subsets are approximated by simplicial trees. (See Lemma 5.1 here for a statement of this.)

Since then, various authors have observed that many naturally occurring spaces admit a more general kind of "cubical structure". In this case, a simplicial tree is generalised to a CAT(0) cube complex, typically with its dimension bounded by some finite "rank". (The hyperbolic case therefore corresponds to "rank 1".) Again there are several formulations of this principle. A starting point for the present discussion is the paper [BeM] where it is shown, using the theory developed in [MaM], that the mapping class group of a surface admits a natural ternary

Date: First draft: 13th August 2018. Revised: 19th October 2020.

operation, called the "centroid map" defined up to bounded distance. In [Bo1], we abstracted some its key properties into the notion of a "coarse median space". The "median" here corresponds to the centroid in the case of the mapping class group, or the centre of a geodesic triangle in the case of a hyperbolic space. More recently the notion of a "hierarchically hyperbolic space" was formulated in [BeHS1]. (A slightly different formulation of this notion is given in [BeHS2] which the authors show to be equivalent.) This property implies coarse median [BeHS2, Bo3]. In addition to the mapping class groups and hyperbolic spaces, these notions apply to Teichmüller space in either the Teichmüller or the Weil-Petersson metric, the separating curve graph, right-angled Artin groups, various classes of relatively hyperbolic groups and a large class of 3-manifold groups, etc. See [Bo4] for some general exposition about coarse median spaces.

In this paper, we will use a collection of axioms, (B1)–(B10), which are implied by those of a hierarchically hyperbolic space, and which imply those of a coarse median space (see Section 7). From these we derive a number of results, most of which are already known in some form for a hierarchically hyperbolic space. The results here are, in some ways more general, and their conclusions a little stronger. The main point however is that the statements are quickly reduced to combinatorial statements about cube complexes (or equivalently discrete median algebras). We also deduce from these results some basic facts about such spaces, such as the existence of monotone paths and the "distance formula".

We briefly summarise the main results as follows. These are all readily deduced from the main result (Theorem 1.3) though the first two are easier to state. In each of these results, it is implicit that the constants of the conclusion depend only on the constants introduced in the hypotheses.

Let (Λ, ρ) be a geodesic metric space. We assume that we have a collection of maps, $\theta_X : \Lambda \longrightarrow \Theta(X)$, to a family of hyperbolic spaces, $((\Theta(X), \sigma_X))_{X \in \mathcal{X}}$, indexed by a set $\mathcal{X}$. We suppose that these satisfy the axioms (B1)–(B10) given in Section 7. In particular, (B10) asserts that there is a "median" operation, $\mu : \Lambda^3 \longrightarrow \Lambda$. It is a consequence of the other axioms that (Λ, ρ, μ) is a coarse median space of rank at most ν (the constant featuring in axiom (B1)), as observed in Lemma 7.1.

Definition. A monotone path is a coarsely lipschitz map $\zeta : I \longrightarrow \Lambda$ from an interval $I \subseteq \mathbb{R}$, such that whenever $t, u, v \in I$ with $t \leq u \leq v$, we have $\rho(\zeta(u), \mu(\zeta(t), \zeta(u), \zeta(v))) \leq k$, for some fixed $k \geq 0$.

Recall that "coarsely lipschitz" means that there are constants $k_1, k_2 \geq 0$ such that for all $t, u \in I$, $\rho(\zeta(t), \zeta(u)) \leq k_1 |t - u| + k_2$. Loosely speaking, the monotone property means that $\zeta(u)$ lies "between" $\zeta(t)$ and $\zeta(v)$ whenever $t \leq u \leq v$.

We do not necessarily assume ζ to be continuous, though it can always be approximated by a continuous path, as we will discuss below.

We will show:

Theorem 1.1. *Given any $a, b \in \Lambda$, there is a monotone path $\zeta : I \longrightarrow \Lambda$ from a to b. Moreover, any monotone path is a quasigeodesic up to reparameterisation of the domain.*

By "reparameterisation" we mean precomposition with an orientation-preserving homeomorphism of intervals. Here we take a "quasigeodesic" to mean a quasi-isometric embedding of a real interval into Λ. In a geodesic space, such a map can always be approximated up to bounded distance by a rectifiable (indeed piecewise geodesic) path such that the length of any subpath is bounded above by a linear function of the distance between its endpoints: in other words it is a "quasigeodesic" in the more traditional sense.

Next, we formulate the "distance formula". Given $t, r \geq 0$, define $\{\{t\}\}_r$ to be equal to t if $t \geq r$ and to be equal to 0 if $t < r$. Given $x, y \in \Lambda$, write $D_k(a, b) = \sum_{X \in \mathcal{X}} \{\{\sigma_X(\theta_X a, \theta_X b\}\}_k$. Clearly, $D_{k'}(a, b) \leq D_k(a, b)$ if $k' \geq k$.

Theorem 1.2. $(\exists k_0 \geq 0)(\forall k \geq k_0)(\exists k_1, k_2 \geq 0)(\forall a, b \in \Lambda)$ $D_k(x, y) \leq k_1\rho(x, y) + k_2$ *and* $\rho(x, y) \leq k_1 D_k(x, y) + k_2$.

Implicit in the conclusion is the fact that $D_k(a, b) < \infty$ for all $k \geq k_0$.

Given any subset, $A \subseteq \Lambda$, one can define a notion of "coarse hull", $H(A)$, of A. This can be formulated in a number of ways, and more discussion will be given in Section 6 for general finite-rank coarse median spaces. Briefly, $A \subseteq H(A)$, and $H(A)$ is "coarsely convex" in the sense that if $x, y \in H(A)$ and $z \in \Lambda$, then $\mu(x, y, z)$ lies a bounded distance from $H(A)$. Moreover, $H(A)$ is, in some sense, the "smallest" set with this property. (Of course, one needs to properly quantify this: see Proposition 6.2.)

We remark that in the case where $A = \{a, b\}$, then $H(A)$ is the "coarse interval" from a to b. This is in turn the union of all monotone paths from a to b (again, up to bounded distance).

In this paper, cube complexes are given the l^1 metric with unit edge-lengths (though the usual l^2 CAT(0) metric is bilipschitz equivalent, and would hence be equivalent for the following statement). The standard definitions of quasi-isometry etc. will be summarised in Section 5. Cube complexes are discussed in Section 2.

Theorem 1.3. *Suppose that $n \in \mathbb{N}$ and $A \subseteq \Lambda$ with $|A| \leq n$. There is a CAT(0) cube complex, Δ, of dimension at most ν, a subset A_0 of vertices of Δ, and a map $f : \Delta \longrightarrow \Lambda$ such that $f|A_0$ maps A_0 bijectively to A, f is a quasi-isometric embedding, and $f(\Delta)$ is a bounded Hausdorff distance from $H(A)$. Moreover, we can assume that Δ is the combinatorial convex hull (in the median sense) of A_0. The map f preserves the respective median structures on Δ and Λ up to bounded distance.*

The statement of the last result calls for some elaboration which we supply in Section 9.

In each of the above results, the constants involved in the conclusions depend only on those of the hypotheses; that is to say, axioms (B1)–(B10) together with k in Theorem 1.2 and n in Theorem 1.3. Moreover, all the arguments of the present paper are constructive, and so give rise to computable bounds. However, we won't generally give explicit formulae.

We note that a monotone path is essentially the same as a "hierarchy path", in the sense that the composition $\theta_X \circ \zeta$ is a uniform unparameterised quasigeodesic in each of the hyperbolic spaces, $\Theta(X)$. Such paths were shown to exist for the mapping class group in [MaM] (as part of their theory of "hierarchies") and more generally for hierarchically hyperbolic spaces in [BeHS2].

The result of Theorem 1.2 is the "distance formula", proven for the mapping class group, as well as the Weil-Petersson metric, in [MaM], for the Teichmüller metric in [Ra] (see also [Du]) and more generally for hierarchically hyperbolic spaces in [BeHS2]. (It was taken as an axiom for a hierarchically hyperbolic space in [BeHS1].) It was used in [BeHS3] in the proof of their description of convex hulls. Here the logic is reversed, in that we derive it as a consequence of Theorem 1.3. (For this, we only need to consider coarse intervals, i.e. when $|A| = 2$, which means one could simplify the argument somewhat.) The distance formula is key to many applications of the general theory.

The notion of coarse convex hull is equivalent, in the context of the mapping class group, to the notion of a "Σ-hull" which was central to the paper [BeKMM]. (This equivalence is a consequence of Lemmas 7.2 and 7.3 here.) For hierarchically hyperbolic spaces a version of Theorem 1.3 was proven in [BeHS3] and was apparently new to both the mapping class group and to Teichmüller space. (In their version it is only asserted that Δ is the convex hull of some finite set of bounded cardinality.) We also note that various aspects of convexity in hierarchically hyperbolic spaces are further explored in [RuST]. A brief account of applications to the mapping class group and Teichmüller space will be given in Section 10 here.

Another recent application of the cubical nature of coarse hulls (in particular coarse intervals) is given in [HHP]. There the authors axiomatise this in terms of a coarse median space with "quasicubical intervals" (which of course, applies in the situations mentioned above). They show that such a space admits a natural quasi-isometrically equivalent metric in which metric balls are median convex. Under the additional assumptions given in [BeHS1], this implies that the space is coarse Helly, which in turn has various applications. (We suspect this would also apply under the more general axioms (B1)–(B10) of the present paper, but we have not checked this.) The (coarse) Helly property is in many ways analogous to (or an extension of) the median property. It has some interesting applications in geometric group theory. A recent paper on the subject is [ChalCGHO].

As another consequence of Theorem 1.3, we have the following:

Theorem 1.4. *If Λ has bounded geometry, then coarse median intervals in Λ have at most uniform polynomial growth of degree ν.*

The relevant definitions can be found at the end of Section 9. We note that this is a key property of the "rapid decay" criterion of [ChatR], as used in [BeM] in the case of the mapping class group. A similar result is given in [Bo2] under much more general hypotheses, though with a slightly weaker bound on degree, and by a non-constructive argument. Further discussion of the growth rate of intervals in general finite-rank coarse median spaces can be found in [NWZ2].

We remark that some amount of work is required to obtain the dimension bound of ν in Theorem 1.3, and the consequent bound of the polynomial degree in Proposition 1.4. If we were to be satisfied with a weaker bound (such as $\lambda = \nu\kappa$, where κ is the constant of Axiom (B2)), then one could bypass some of more technical arguments (including most of Section 3). This would be sufficient for the proofs of Theorems 1.1 and 1.2.

Much of the discussion is set in a more general context. In particular, in Section 6 we give some general results about coarse hulls in coarse median spaces of finite rank.

We should make a few comments about our hypotheses.

First, we have assumed that (Λ, ρ) is a geodesic space; that is, any two points are connected by a geodesic. We could weaken this by demanding only that they be connected by a uniform quasigeodesic. It can be seen that this would make no essential difference to our arguments. Indeed any such "quasigeodesic space" is quasi-isometric to a graph, hence to a genuine geodesic space. Moreover, the various hypotheses only really require Λ to be defined up to quasi-isometry. For simplicity of exposition, we will stick with geodesic spaces in this paper. This is directly applicable to most situations.

Our hypotheses (B1)–(B9) are all standard properties of projection maps (originating in [MaM]). They can easily be seen to be consequences, either of the hypotheses of a hierarchically hyperbolic space, or of the Axioms (A1)–(A10) listed in Section 7 of [Bo3]. In contrast to those accounts, we have taken the existence of a median as an axiom, namely (B10). This is also a consequence of either set of axioms referred to above (see [BeHS2] and [Bo3]). Both those formulations included instead a "(partial) realisation" axiom (cf. Axiom 8 of [BeHS2] or Axiom (A10) of [Bo3]) which we have omitted here. In this respect, the spaces we consider are more general.

As alluded to above, our aim will be to interpret the above statements in combinatorial terms, using cube complexes. We will implicitly describe much of this in terms of median algebras, though we will not need to get too involved in the general theory of these structures here. Some standard references to this are [BaH, I, Ro]. Some further discussion, relevant to present paper, can be found in [Bo1, Bo3, Bo4]. We will begin with a general discussion of these in the next section.

I thank Jason Behrstock and Alessandro Sisto for their interest and comments.

2. Cube complexes

In this section, we prove some general statements about CAT(0) cube complexes, which we mostly view combinatorially.

Let Δ be a cube complex (a cell complex built out of cubes). Write Δ^0 for its set of vertices, and Δ^1 for its 1-skeleton. We will assume that Δ is (combinatorially) CAT(0). Here we take this to mean that it is connected and simply connected, and that the link of every cell is a flag simplicial complex. We give each cell the structure of a unit cube with the l^1 metric, and write $d = d_\Delta$ for the induced path metric. In this way, Δ^1 is an isometrically embedded graph with all side-lengths equal to 1. (The more usual CAT(0) metric is obtained by using the l^2 or "euclidean" metric instead. If Δ is finite dimensional then this is bilipschitz equivalent.)

Given $x, y \in \Delta^0$, let $[x, y]_d = \{z \in \Delta^0 \mid d(x, y) = d(x, z) + d(z, y)\}$. This is always finite. Also it turns out that if $x, y, z \in \Delta^0$, then $[x, y]_d \cap [y, z]_d \cap [z, x]_d$ consists of a single point, denoted $\mu(x, y, z)$. Moreover, (Δ^0, μ) is a discrete median algebra. Indeed every discrete median algebra arises canonically in this way [Che]. (Here "discrete" means that all intervals are finite.) For our purposes, this can serve as a definition of the term "discrete median algebra" — the more standard axioms can be found in [BaH, I, Ro, Bo1]. One can check that μ is symmetric in x, y, z and that $\mu(x, x, y) = x$. We write $[x, y]_\mu = \{z \in \Delta^0 \mid \mu(x, y, z) = z\}$, for the *median interval* from x to y. It turns out that $[x, y]_d = [x, y]_\mu$, and we will generally abbreviate this to $[x, y]$. We also have $[x, y] = \{\mu(x, y, z) \mid z \in \Delta^0\}$.

Let Π be a discrete median algebra, and write $\Delta = \Delta(\Pi)$ for the associated cube complex with $\Pi = \Delta^0$. We write $d = d_\Pi = d_\Delta$. A subset, $M \subseteq \Pi$, is a *subalgebra* if it is closed under μ. It is *convex* if $[x, y] \subseteq M$ for all $x, y \in M$. Any convex subset is a subalgebra. Given any $A \subseteq \Pi$, we write $\langle A \rangle \subseteq \Pi$ for the subalgebra generated by A, and $\operatorname{hull}(A)$ for the convex hull of A: that is the smallest convex subset containing A. We note that $\operatorname{hull}(\{x, y\}) = [x, y]$. A useful property is that $d(\mu(x, y, z), \mu(x, y, w)) \le d(z, w)$ for all $x, y, z, w \in \Pi$. This implies that the median map is 1-lipschitz with respect to l^1 metric on Π^3.

Any two-point set admits a unique median structure. By a *cube* we mean a product, $\Omega(\mathcal{A}) = \prod_{\alpha \in \mathcal{A}} E(\alpha)$, of two-point median algebras, $E(\alpha)$, indexed by a set $\mathcal{A}$. If $|\mathcal{A}| = n < \infty$, we refer to $\Omega(\mathcal{A})$ as an *n-cube*. A *square* is a 2-cube. The *rank* of Π is the maximal n such that Π contains a subalgebra isomorphic to an n-cube. (This will be finite in all cases of interest here.) It turns out that $\operatorname{rank}(\Pi) = \dim(\Delta(\Pi))$. One can also show that any discrete median algebra, Π, can be embedded in a cube, which we can take to be finite if Π is finite.

The following definition is perhaps less standard. Let $r \in \mathbb{N}$, and $F \subseteq \Pi$. By an *r-path* in F, we mean a sequence, $x_0, x_1, \ldots, x_n$ in F with $d_\Pi(x_i, x_{i+1}) \le r$ for all i. We say that F is *r-connected* if given any $x, y \in F$, there is an r-path, $x = x_0, \ldots, x_n = y$ in F from x to y.

Lemma 2.1. *If F is r-connected, then so is $\langle F \rangle$.*

For the proof, we define $L(F) = L^0(F) = \{\mu(x,y,z) \mid x,y,z \in F\} \supseteq F$; and inductively, $L^{i+1}(F) = L(L^i(F))$ with $L^0(F) = F$. Thus $\langle F \rangle = \bigcup_{i=0}^{\infty} L^i(F)$.

Proof. It suffices to show that $L(F)$ is r-connected. In fact, if $w = \mu(x,y,z) \in L(F)$, we $x = x_0, \ldots, x_n = y$ be a path from x to y in F. Let $y_i = \mu(x, x_i, z)$. Then $x = y_0, \ldots, y_n = w$ is a path from x to w in $L(F)$. $\square$

By a similar argument, we see that if $M \subseteq \Pi$ is an r-connected subalgebra, and $C \subseteq \Pi$ is convex, then $C \cap M$ is r-connected.

If $M \subseteq \Pi$ is a subalgebra, then M is 1-connected if and only if the full subcomplex on M in $\Delta(\Pi)$ is connected. In this case, we can naturally identify $\Delta(M)$ with this subcomplex. Moreover, the inclusion of $\Delta(M)$ into $\Delta(\Pi)$ is isometric with respect to the l^1 metrics, $d_{\Delta(M)}$ and $d_{\Delta(\Pi)}$.

Before continuing, we make the following general definition.

Definition. If $\mathcal{S}$ is any set, and $R \subseteq \mathcal{S} \times \mathcal{S}$ is any relation, we define the *width*, $\mathrm{width}(\mathcal{S}, R)$, of R as the maximal cardinality of a subset $B \subseteq \mathcal{S}$ such that for all distinct $a, b \in B$, either aRb or bRa (or both).

For example, if R is an equivalence relation, then $\mathrm{width}(\mathcal{S}, R)$ is the maximal cardinality of an equivalence class. If R is a partial order, then $\mathrm{width}(\mathcal{S}, R)$ is the maximal length of a chain in $\mathcal{S}$.

We move on to consider subalgebras of cubes. We have already noted that every discrete median algebra can be embedded in a cube. In fact, the ones we meet below will all arise directly in that way.

Let $\Omega(\mathcal{A}) = \prod_{\alpha \in \mathcal{A}} E(\alpha)$ be a cube. We will write $x_\alpha = \pi_\alpha x$ for the α-coordinate of $x \in \Omega(\mathcal{A})$. Given any $\mathcal{E} \subseteq \mathcal{A}$, there is a natural projection map, $\pi_{\mathcal{E}} : \Omega(\mathcal{A}) \longrightarrow \Omega(\mathcal{E})$. If $\alpha, \beta \in \mathcal{A}$ are distinct, we will write $\Omega_{\alpha\beta} = E(\alpha) \times E(\beta) = \Omega(\{\alpha, \beta\})$. We will abbreviate $\pi_{\{\alpha,\beta\}}$ to $\pi_{\alpha,\beta}$. We similarly write $\Omega_{\alpha\beta\gamma} = E(\alpha) \times E(\beta) \times E(\gamma)$. Note that any convex subset of a cube is a subcube. If it is finite, it is the convex hull of any pair of opposite corners.

Definition. We say that a subset $F \subseteq \Omega(\mathcal{A})$ is *filling* if $\Omega(\mathcal{A}) = \mathrm{hull}(F)$.

Clearly this implies that $\pi_{\mathcal{E}}(F) \subseteq \Omega(\mathcal{E})$ is filling for all $\mathcal{E} \subseteq \mathcal{A}$. Indeed it is equivalent to saying that $\pi_\alpha(F) = E(\alpha)$ for all $\alpha \in \mathcal{A}$.

Now any subset, $G \subseteq E(\alpha) \times E(\beta) = \Omega_{\alpha\beta}$, of a square is a subalgebra. If G is filling, then there are three possibilities up to isomorphism.

Maybe $|G| = 2$, in which case, G consists of two opposite corners of the square $\Omega_{\alpha\beta}$. Maybe $|G| = 4$, that is $G = \Omega_{\alpha\beta}$. Otherwise, $|G| = 3$. In this case, we write $G = \Omega_{\alpha\beta} \setminus \{(\psi^C_\alpha \beta, \psi^C_\beta \alpha)\}$, where $\psi^C_\alpha \beta \in E(\alpha)$ and $\psi^C_\beta \alpha \in E(\beta)$. We write $\psi_\alpha \beta$ and $\psi_\beta \alpha$ respectively for the other points of $E(\alpha)$ and $E(\beta)$. So, $G = \{(\psi_\alpha \beta, \psi^C_\beta \alpha), (\psi_\alpha \beta, \psi_\beta \alpha), (\psi^C_\alpha \beta, \psi_\beta \alpha)\}$. (Thus, intrinsically, G is isomorphic to $\{-1, 0, 1\} \subseteq \mathbb{R}$, with $(\psi_\alpha \beta, \psi_\beta \alpha)$ corresponding to the midpoint, 0.)

Now suppose that $F \subseteq \Omega(\mathcal{A})$ is filling. We define relations $\sim$ and $\approx$ on $\mathcal{A}$ by writing $\alpha \sim \beta$ if $\alpha = \beta$ or $|\pi_{\alpha\beta}(F)| = 2$, and writing $\alpha \approx \beta$ if $|\pi_{\alpha\beta}(F)| = 4$.

Clearly these are symmetric, and (since any subset of a square is a subalgebra) they agree with the relations on F, similarly defined with $\langle F\rangle$ replacing F. In particular, $\mathrm{width}(\mathcal{A},\sim)$ and $\mathrm{width}(\mathcal{A},\approx)$ are both well defined whether we use F or $\langle F\rangle$ to define these relations.

In fact, $\sim$ is an equivalence relation. Indeed, if $\mathcal{E}$ is finite and its elements are pairwise $\sim$-related, then $\pi_{\mathcal{E}}(F)$ consists of opposite corners of the cube, $\Omega(\mathcal{E})$.

The relation $\approx$ need not be transitive. However, we note that if $\mathcal{E}$ is finite and its elements are pairwise $\approx$-related, then $\pi_{\mathcal{E}}(\langle F\rangle)=\Omega(\mathcal{E})$. (This can be seen by induction on $|\mathcal{E}|$, projecting to codimension-1 cubes, similarly as in the proof of Lemma 2.2 below.)

Note that if $F\subseteq\Omega(\mathcal{A})$ is 1-connected, then $\mathrm{width}(\mathcal{A},\sim)=1$. In other words, each $\sim$-class is a singleton, or equivalently, $|\pi_{\alpha\beta}(F)|\geq 3$ whenever $\alpha\neq\beta$. For subalgebras, we have the following converse.

Lemma 2.2. *Let $M\subseteq\Omega(\mathcal{A})$ be a filling subalgebra, and suppose* $\mathrm{width}(\mathcal{A},\sim)=1$. *Then M is 1-connected.*

Proof. Let $x,y\in\Omega(\mathcal{A})$. Passing to the subcube, $[x,y]\subseteq\Omega(\mathcal{A})$, we can assume that $\mathcal{A}$ is finite. We proceed by induction on $n=|\mathcal{A}|$. If $n=2$, this is clear. In fact, if $Q\subseteq\Omega(\mathcal{A})$ is a square face with opposite corners, a,b in $Q\cap M$, then there is a third point, $c\neq a,b$ in $Q\cap M$. (To see this, let $\alpha,\beta\in\mathcal{A}$ be the labels corresponding to the edges of Q. Since $\alpha\not\sim\beta$, $|\pi_{\alpha\beta}(M)|\geq 3$, so we can find some $z\in M$ with $\pi_{\alpha\beta}z\neq\pi_{\alpha\beta}a,\pi_{\alpha\beta}b$. This gives a point, $c\in\mu(a,b,z)\in Q\cap M$, as required.) Now let $\alpha\in\mathcal{A}$ be any element, and let $\mathcal{E}=\mathcal{A}\setminus\{\alpha\}$. Clearly each $\sim$-class in $\mathcal{E}$ corresponding to $\pi_{\mathcal{E}}(M)\subseteq\mathcal{E}$ is also a singleton. Therefore, by the inductive hypothesis, $\pi_{\mathcal{E}}(M)$ is 1-connected in $\Omega(\mathcal{E})$. We can therefore connect $\pi_{\mathcal{E}}x$ to $\pi_{\mathcal{E}}y$ by a 1-path in $\pi_{\mathcal{E}}(M)$. By the above observation about square faces, we can interpolate points as appropriate to give us a 1-path from x to y in M. $\square$

Now let $F\subseteq\Omega(\mathcal{A})$ be any filling subset. Let $\mathcal{F}=\mathcal{A}/\sim$. There is a natural projection $\pi_{\mathcal{F}}:\Omega(\mathcal{A})\longrightarrow\Omega(\mathcal{F})$.

Lemma 2.3. *$\pi_{\mathcal{F}}|F$ is injective.*

In fact, let $\mathcal{E}$ be any $\sim$-transversal in $\mathcal{A}$. We can naturally identify $\mathcal{F}$ with $\mathcal{E}$ and $\Omega(\mathcal{F})$ with $\Omega(\mathcal{E})$.

If $x,y\in F\subseteq\Omega(\mathcal{A})$ with $\pi_{\mathcal{E}}x=\pi_{\mathcal{E}}y$, then $x_\alpha=y_\alpha$ for all $\alpha\in\mathcal{E}$, and so by the definition of $\sim$, we see that $x_\beta=y_\beta$ for all $\beta\in\mathcal{A}$. In other words, $x=y$. This proves Lemma 2.3.

We note that if $\approx$ is the relation on $\mathcal{A}$ defined with respect to a 1-connected median algebra, M, then $\mathrm{rank}(M)=\dim(\Delta(M))=\mathrm{width}(\mathcal{A},\approx)$.

Returning to the earlier set-up, if F is r-connected, then $\mathrm{width}(\mathcal{A},\sim)\leq r$. (For if $\mathcal{E}\subseteq\mathcal{A}$ lies in an equivalence class then $\pi_{\mathcal{E}}(F)$ consists of a pair, a,b, of opposite corners of $\Omega(\mathcal{E})$. If $|\mathcal{E}|>r$, there is no r-path from $\pi_{\mathcal{E}}^{-1}a$ to $\pi_{\mathcal{E}}^{-1}b$.) The map $\pi_{\mathcal{A}/\sim}:\Omega(\mathcal{A})\longrightarrow\Omega(\mathcal{A}/\sim)$ restricted to F reduces distances by a factor of at most r.

Now suppose that $F \subseteq \Omega(\mathcal{A})$ is filling, and write $M = \langle F \rangle$. Let $\mathcal{B}$ be any $\sim$-transversal. We have embeddings $F \subseteq M \subseteq \Omega(\mathcal{B})$, and M is 1-connected in $\Omega(\mathcal{B})$. (So we can identify $\Delta(M)$ as a full subcomplex of $\Delta(\Omega(\mathcal{B}))$). From this point on, we could forget about $\mathcal{A}$.

3. More cubes

This section is a continuation of the last, though it is specifically geared towards the applications in Section 7. It will not be needed again until then, but fits more logically into to the present discussion. Some of the definitions here are rather formal. The motivation behind them in terms of model spaces can be found in Section 11. As alluded to in the Introduction, the discussion beyond Lemma 3.1 can be ignored if we were prepared to weaken the dimension bound in the conclusion of Theorem 1.3. This would be sufficient for proving Theorems 1.1 and 1.2.

Recall that we have a cube, $\Omega = \Omega(\mathcal{B})$, and a 1-connected filling median subalgebra, $M \subseteq \Omega(\mathcal{B})$. This gives rise to a relation, $\approx$, on $\mathcal{B}$. In other words, if $\alpha, \beta \in \mathcal{B}$, then either $\alpha \approx \beta$ and $\pi_{\alpha\beta} M = \Omega_{\alpha\beta}$, or $\alpha \not\approx \beta$ and $|\pi_{\alpha\beta} M| = 3$. In the latter case, $\psi_\alpha \beta$ and $\psi_\beta \alpha$ are defined as in Section 2.

In addition, we suppose there are symmetric relations, $\pitchfork$ and $\boxminus$ on $\mathcal{B}$ such that for all $\alpha, \beta \in \mathcal{B}$ exactly one of the relations $\alpha = \beta$, $\alpha \pitchfork \beta$ or $\alpha \boxminus \beta$ holds. We suppose that $\mathrm{width}(\mathcal{B}, \boxminus) \leq \lambda < \infty$. (In the case of the mapping class group, where the indexing set consists of subsurfaces of a given surface, $\pitchfork$ means "transverse" and $\boxminus$ means "disjoint or nested", see Sections 10 and 11.)

We note the following particular case of Ramsey's Theorem:

Lemma 3.1. *Suppose that $\mathcal{E} \subseteq \mathcal{B}$ with $\mathrm{width}(\mathcal{E}, \pitchfork) \leq 2$. Then $|\mathcal{E}| \leq 3^\lambda$.*

Proof. Recall that $\mathrm{width}(\mathcal{E}, \pitchfork) \leq 2$ means that no three elements of $\mathcal{E}$ are pairwise $\pitchfork$-related. We can suppose that there exist $\alpha, \beta \in \mathcal{E}$ with $\alpha \pitchfork \beta$ (otherwise all distinct pairs in $\mathcal{E}$ are $\boxminus$-related, and so $|\mathcal{E}| \leq \lambda \leq 3^\lambda$). Then $\mathcal{E} = \{\alpha, \beta\} \cup \mathcal{E}_\alpha \cup \mathcal{E}_\beta$, where $\mathcal{E}_\alpha = \{\gamma \in \mathcal{E} \mid \gamma \boxminus \alpha\}$ and $\mathcal{E}_\beta = \{\gamma \in \mathcal{E} \mid \gamma \boxminus \beta\}$. Now $\mathrm{width}(\mathcal{E}_\alpha, \boxminus) < \lambda$ and $\mathrm{width}(\mathcal{E}_\beta, \boxminus) < \lambda$. By induction, we can assume that $|\mathcal{E}_\alpha|, |\mathcal{E}_\beta| \leq 3^\lambda$ (since the statement clearly holds for $\lambda = 1$). Therefore, $|\mathcal{E}| \leq 2 + 2.3^{\lambda-1} \leq 3^\lambda$. $\square$

We will also assume that $\alpha \approx \beta$ implies $\alpha \boxminus \beta$, and so $\alpha \pitchfork \beta$ implies $\alpha \approx \beta$. (Thus $\mathrm{width}(\mathcal{B}, \approx) \leq \lambda$, and so the rank of $\Omega(\mathcal{B})$ is at most λ.)

The remainder of this section will only be relevant to obtaining the dimension bound of ν in Theorem 1.1.

We assume, in addition, that $\boxplus$ is another symmetric relation on $\mathcal{B}$ such that $\alpha \boxplus \beta$ implies $\alpha \boxminus \beta$. (For the mapping class group, $\boxplus$ can be interpreted as a certain "nesting" property: see Section 11.) We will assume:

$(*)$: If $\alpha \boxplus \beta$ there is a subset $D_{\alpha\beta} \subseteq \Omega_{\alpha\beta} = E(\alpha) \times E(\beta)$, consisting of a pair of opposite corners of $\Omega_{\alpha\beta}$ such that if $\gamma \in \mathcal{B}$ with $\gamma \pitchfork \alpha$ and $\gamma \pitchfork \beta$, then $\psi_\gamma \alpha = \psi_\gamma \beta$

and $(\psi_\alpha\gamma, \psi_\beta\gamma) \in D_{\alpha\beta}$.

(A simple illustration of this hypothesis in terms of intersections of real intervals is given in Section 11.)

Note that $\psi_\alpha\gamma$ and $\psi_\gamma\alpha$ are defined since $\alpha \pitchfork \gamma$ implies $\alpha \not\approx \gamma$. (That is, $|\pi_{\alpha\gamma}M| = 3$.) Similarly for $\psi_\beta\gamma$ and $\psi_\gamma\beta$. In fact, we see that $\pi_{\alpha\beta\gamma}M \subseteq (\Omega_{\alpha\beta} \times \{\psi_\gamma\alpha\}) \cup \{(\psi_\alpha\gamma, \psi_\beta\gamma, \psi_\gamma\alpha)\} \subseteq \Omega_{\alpha\beta\gamma}$.

Recall that $M \subseteq \Omega(\mathcal{B})$ is 1-connected. A 1-path, in $x = x_0, \dots, x_n = y$, in M gives us a sequence of edges, $e_1, \dots, e_n$, of the 1-skeleton, $\Delta^1(M)$, where e_i connects x_{i-1} to x_i. Let $\alpha_i \in \mathcal{B}$ be the "label" associated with e_i. In other words, x_{i-1} and x_i differ precisely in the α_i coordinate. If this is a shortest 1-path from x to y (that is, $d_M(x, y) = n$), then the α_i are all distinct. (For suppose $\alpha_i = \alpha_j$ with $i < j$ and $\alpha_k \neq \alpha_i$ for $i < k < j$. We set $y_k = \mu(x_{i-1}, x_i, x_k)$, so that $y_i = x_{i-1}$ and $y_{j-1} = x_j$. Replacing $x_{i-1}, \dots, x_j$ by $y_i, \dots, y_{j-1}$ would then give us a shorter 1-path from x to y in M contrary to our assumption.)

We now suppose that $\mathcal{C} \subseteq \mathcal{B}$ is a subset with the property that for all $\alpha \in \mathcal{B}$ there is some $\beta \in \mathcal{C}$ with $\beta \boxplus \alpha$. (The motivation for this can be found in the proof of Theorem 1.3 in Section 9.)

Let $\pi_{\mathcal{C}} : \Omega(\mathcal{B}) \longrightarrow \Omega(\mathcal{C})$ be the quotient map. If $p \in \Omega(\mathcal{C})$, then $\pi_{\mathcal{C}}^{-1}p$ is a face (i.e. convex subcube) of $\Omega(\mathcal{B})$, and $M \cap \pi_{\mathcal{C}}^{-1}p$ is 1-connected. In fact, we have a projection, $\pi : \Delta(\Omega(\mathcal{B})) \longrightarrow \Delta(\Omega(\mathcal{C}))$, and we can identify $\Delta(M \cap \pi_{\mathcal{C}}^{-1}p)$ with $\Delta(M) \cap \pi^{-1}p$, which is isometrically embedded in $\Delta(M)$.

We claim:

Lemma 3.2. *If $p \in \Omega(\mathcal{C})$, then $M \cap \pi_{\mathcal{C}}^{-1}p$ has d_M-diameter at most 3^λ.*

Proof. Let $x, y \in M \cap \pi_{\mathcal{C}}^{-1}p$. Let $x = x_0, \dots, x_n = y$ be a shortest 1-path in $M \cap \pi_{\mathcal{C}}^{-1}p$ from x to y. Let $e_1, \dots, e_n$ be its sequence of edges in $\Delta^1(M)$, and let $\alpha_1, \dots, \alpha_n$ be their labels which must lie in $\mathcal{B} \setminus \mathcal{C}$. As observed above, the α_i are all distinct.

Suppose that $i < j < k$. We claim that either $\alpha_j \boxminus \alpha_i$ or $\alpha_j \boxminus \alpha_k$ (or both). For suppose, to the contrary, that $\alpha_j \pitchfork \alpha_i$ and $\alpha_j \pitchfork \alpha_k$. By the assumption on $\mathcal{C}$, there is some $\beta \in \mathcal{C}$ with $\beta \boxplus \alpha_j$. Write $\Psi = \Omega_{\beta\alpha_j} = E(\beta) \times E(\alpha_j)$ and $D = D_{\beta\alpha_j} \subseteq \Psi$, as given by $(*)$. Thus D is a pair of opposite corners of Ψ. Let $u = (\psi_\beta\alpha_i, \psi_{\alpha_j}\alpha_i) \in \Psi$, and $v = (\psi_\beta\alpha_k, \psi_{\alpha_j}\alpha_k) \in \Psi$. By $(*)$, we have $u, v \in D$. In fact, $\pi_{\beta\alpha_j}e_i = \{u\}$, and $\pi_{\beta\alpha_j}e_k = \{v\}$. Since the subpath from e_i to e_k crosses e_j, but no other edge labelled α_j, we must have $u \neq v$. In other words, u, v are opposite corners of Ψ, and so in order to get from u to v, our subpath must also cross some edge labelled β. This is a contradiction, since $\beta \in \mathcal{C}$, and we have observed that all edges in our path have labels in $\mathcal{B} \setminus \mathcal{C}$. This proves the claim.

In particular, we see that no three of the α_i are all pairwise $\pitchfork$-related. In other words, the width of the relation, $\pitchfork$, restricted to $\{\alpha_1, \dots, \alpha_n\}$ is at most 2. It follows by Lemma 3.1 that $n \leq 3^\lambda$. $\square$

As an immediate corollary, we have:

Corollary 3.3. *If $x, y \in M \subseteq \Omega(\mathcal{B})$, then $d_M(x, y) \leq 3^\lambda d_{\pi_\mathcal{C}(M)}(\pi_\mathcal{C} x, \pi_\mathcal{C} y) + 3^\lambda$.*

In particular, we see that the projection map, $\pi : \Delta(\Omega(\mathcal{B})) \longrightarrow \Delta(\Omega(\mathcal{C}))$ restricted to $\Delta(M) \subseteq \Delta(\Omega(\mathcal{B}))$ gives a quasi-isometry from $\Delta(M)$ to $\pi(\Delta(M)) = \Delta(\pi_\mathcal{C}(M))$.

The eventual point of this discussion will be to reduce dimension in the following sense. We will show, under certain hypotheses, that $\mathrm{width}(\mathcal{C}, \approx) \leq \nu$. It will then follow that the dimension of $\Delta(\pi_\mathcal{C}(M))$ is at most ν.

4. Subalgebras

In this section, we prove a general result about discrete median algebras of finite rank (or equivalently, finite dimensional CAT(0) cube complexes). Essentially this says that a subset which is a subalgebra up to bounded distance lies a bounded distance from a genuine subalgebra.

Let Π be a discrete median algebra, and let $\Delta(\Pi)$ be the associated cube complex, with $\Pi = \Delta^0(\Pi)$. We suppose that $\mathrm{rank}(\Pi) \leq \nu < \infty$. (Recall that this is the same as the dimension of $\Delta(\Pi)$.)

Definition. We say that a subset $F \subseteq \Pi$ is *r-median* if for all $x, y, z \in F$ we have $d_\Pi(\mu(x, y, z), F) \leq r$.

Proposition 4.1. *Given $r, \nu \in \mathbb{N}$, there is some $k \in \mathbb{N}$ such that if* $\mathrm{rank}(\Pi) \leq \nu$ *and if $F \subseteq \Pi$ is 1-connected and r-median, then F is k-dense in $\langle F \rangle$.*

In other words, $\langle F \rangle$ is contained in the r-neighbourhood of F in Π.

In fact, we could weaken "1-connected" to "r-connected", by a simple adaptation of the argument, but we won't be needing that.

Recall from Section 2 that $L(F) = \{\mu(x, y, z) \mid x, y, z \in F\}$. The r-median hypothesis tells us that $L(F) \subseteq N(F; r)$, where $N(.; r)$ denotes r-neighbourhood. Since the median map is 1-lipschitz, it follows that $L^2(F) = L(L(F)) \subseteq L(N(F; r)) \subseteq N(L(F); 3r) \subseteq N(L; 4r)$. We see inductively, that $L^n(F) \subseteq N(F; 4^n r)$ for all n.

For the proof of the proposition, we will use the "binary subdivision", $\hat{\Pi}$, of Π. We first consider the case of cubes.

The binary subdivision of the 1-cube, $\{-1, 1\}$, is the set $\{-1, 0, 1\}$, with the obvious median of "betweenness". The binary subdivision of $Q = \{-1, 1\}^n$ is then $\hat{Q} = \{-1, 0, 1\}^n$. We write $o(Q)$ for the centre, $(0, 0, \ldots, 0)$, of $\hat{Q}$.

Note that we can naturally identify $\Delta(\hat{Q})$ with $\Delta(Q)$ with l^1 metric scaled by a factor of $1/2$, so that each edge of $\Delta^1(\hat{Q})$ has length $1/2$. (For the purposes of the current argument, the real interval $[-1, 1]$ is deemed to have length 1.)

In general, we can subdivide Π in this way to give us $\hat{\Pi}$. We can similarly identify, $\Delta(\hat{\Pi})$, in $\Delta(\Pi)$. Given any cell, Q, of Π, we write $o(Q) \in \Delta^0(\hat{\Pi})$ for its centre.

We write $C(\Pi) \subseteq \hat{\Pi}$ for the set of centres (or "midpoints") of 1-cells of Π. Thus, the set of vertices of the 1-skeleton of $\Delta(\hat{\Pi})$ is $\Delta^0(\Pi) \sqcup C(\Pi)$.

We begin with a lemma about cubes.

Lemma 4.2. *Let Q be a finite-dimensional cube, and suppose that $\Gamma \subseteq \Delta^1(Q)$ is a connected subgraph of the 1-skeleton of $\Delta(Q)$. Suppose that $Q = \langle \Gamma \cap Q \rangle$. Then $o(Q) \in \langle \Gamma \cap C(Q) \rangle$ in $\hat{Q}$.*

Proof. Write $Q = \{-1, 1\}^n$. We proceed by induction on n. If $n = 1$, then $\Gamma = \Delta(Q)$, so $o(Q) = \Gamma \cap C(Q) = \{o(Q)\}$. If $n = 2$, then $Q \subseteq \Gamma$, and Γ omits at most one edge of $\Delta(Q)$. Now $o(Q)$ is the median of the midpoints of any three edges, and so again we have $o(Q) \in \langle \Gamma \cap C(Q) \rangle$. We therefore assume that $n \geq 3$.

Given $i \in \{1, \ldots, n\}$, write $\hat{Q}_i \subseteq \hat{Q}$ for the set of points with ith coordinate 0. We can identify $\hat{Q}_i$ with the binary subdivision of an $(n-1)$-cube Q_i. Write $\hat{\pi}_i : \hat{Q} \longrightarrow \hat{Q}_i$ for the natural projection, and write $\pi_i : Q \longrightarrow Q_i$ for its restriction to Q. (In other words, we can think of Q_i and $\hat{Q}_i$ as quotients of Q and $\hat{Q}$ respectively.) We can extend π_i to a map $\pi_i : \Delta(Q) \longrightarrow \Delta(Q_i)$. Write $C = C(Q)$ and write $C_i = C(Q_i)$ which we can identify as a subset of $\pi_i C$. Now $\Gamma_i = \pi_i \Gamma$ is a connected subgraph of the 1-skeleton of $\Delta(Q_i)$. Also, $\Gamma_i \cap Q_i = \pi_i(\Gamma \cap Q)$ generates Q_i as a median algebra. Therefore, by our inductive assumption, $o(Q_i) \in \langle \Gamma_i \cap C_i \rangle$ in $\hat{Q}_i$. Now, $\langle \Gamma_i \cap C_i \rangle \subseteq \hat{\pi}_i(\langle \Gamma \cap C \rangle)$. Since $\hat{\pi}_i : \hat{Q} \longrightarrow \hat{Q}_i$ is a median epimorphism, it follows that either $o(Q_i^+)$ or $o(Q_i^-)$ lies in $\langle \Gamma \cap C \rangle$, where $Q_i^{\pm}$ is the subcube of Q with ith coordinate ± 1. Without loss of generality, we can assume that $o(Q_i^+) \in \langle \Gamma \cap C \rangle$. Note that $o(Q_i^+)$ has ith coordinate 1, and all other coordinates 0. Thus, $o(Q) = \mu(o(Q_1^+), o(Q_2^+), o(Q_3^+)) \in \langle \Gamma \cap C \rangle$, and the statement follows by induction. $\square$

If Q is any cell of Π there is a nearest-point projection, $\omega : \Pi \longrightarrow Q$. (It is an instance of the more general notion of a "gate map" of a median algebra to a convex subset.) This extends to a nearest-point projection, $\hat{\omega} : \hat{\Pi} \longrightarrow \hat{Q}$. Both ω and $\hat{\omega}$ are median epimorphisms. Note that Q is a maximal cell (i.e. not contained in any strictly larger cell) if and only if $\hat{\omega}^{-1}(o(Q)) = \{o(Q)\}$.

Proof of Proposition 4.1. Replacing Π by $\langle F \rangle$, we may as well assume that F generates Π. We want to show that F is cobounded in Π. Let $G \subseteq \Delta^1(\Pi)$ be the full subgraph with vertex set $F \subseteq \Pi$. Since F is 1-connected, G is connected. Let $\hat{F} = G \cap \hat{\Pi}$. In other words, $\hat{F} = F \sqcup C(F)$, where F is the set of vertices of G and $C(F)$ is the set of midpoints of edges of G. Let Q be any maximal subcube of Π. Projecting G to $\Delta(Q)$, we get a connected graph, Γ, in the 1-skeleton of $\Delta(Q)$, with $\omega(F) = \Gamma \cap Q$ and with $\hat{\omega}(C(F)) = \Gamma \cap \hat{Q}$. Since $\Pi = \langle F \rangle$ and ω is an epimorphism, we have have $Q = \langle \Gamma \cap Q \rangle$. Therefore, by Lemma 4.2, $o(Q) \in \langle \Gamma \cap C(Q) \rangle$. It follows that $o(Q)$ can be written as a median expression of bounded complexity involving elements, $\omega(x)$, in $\omega(C(F))$, with the bound just depending on $\dim(Q) \leq \nu$. Since ω is a homomorphism, applying the same expression to the elements, x, in $C(F)$, we arrive at $y \in \Pi$, with $\omega(y) = o(Q)$. Therefore, since Q is maximal, we have $y = o(Q)$. By definition of $C(F)$, each such x is the midpoint of an edge with both vertices in F. Choose one such vertex $x' \in F$. We now apply the same expression to these x' to give us a point $y' \in F$. One can check that $y' \in Q$. (In fact, it would be sufficient for the proof to show that $d_\Pi(y, y')$ is bounded, which is clear given that the median operation is lipschitz.) This shows that $Q \cap L^n(F) \neq \varnothing$, where n just depends on the dimension of Q, which is at most $\operatorname{rank}(\Pi) \leq \nu$. In other words, Π lies in a 1-neighbourhood of $L^n(F)$. Therefore, as observed above, we have $\Pi \subseteq N(F; k)$, where $k = 4^n r + 1$. $\square$

We remark that (an equivalent of) the result could be expressed in terms of cube complexes as follows.

Let Δ be a CAT(0) cube complex of dimension ν, and let $S \subseteq \Delta$ be a connected subset. Suppose that the median in Δ of any three points of S lies in $N(S; r)$ for some $r \geq 0$. Then there is a subcomplex $\Delta' \subseteq \Delta$ which is a subalgebra of Δ with respect to the standard median, and with $S \subseteq \Delta' \subseteq N(S; k)$. In particular, Δ' is isometrically embedded with respect to the l^1 metrics. Moreover, with respect to the l^2 (euclidean) metrics, Δ' is intrinsically CAT(0) and quasi-isometrically embedded in Δ.

5. Coarse median spaces

We now move on to coarse geometry. We begin by recalling some general definitions.

Let (Υ, d) be a metric space. Given a subset, $A \subseteq \Upsilon$, we write $N(A; r)$ for the r-neighbourhood of A. We say that A is *r-dense* in Υ if $\Upsilon = N(A; r)$. We say A is *cobounded* if it is r-dense for some r. Given $A, B \subseteq \Upsilon$, write $\operatorname{hd}(A, B)$ for the Hausdorff distance from A to B.

A *geodesic* in Υ is a path, $\alpha : [a, b] \longrightarrow \Upsilon$, with $\operatorname{length}(\alpha) = d(\alpha(a), \alpha(b))$. We say that a path, α, is a (k_1, k_2)-*quasigeodesic* if for all $t, u \in [a, b]$, $\operatorname{length}(\alpha|[t, u]) \leq k_1 d(\alpha(t), \alpha(u)) + k_2$. We say that Υ is a geodesic space if every pair of points are connected by a geodesic. We say that a subset $A \subseteq \Upsilon$ is *r-connected* if any two points $x, y \in A$ are connected by an *r-path* in A; that is, a sequence, $x = x_0, x_1, \dots, x_n = y$, with $x_i \in A$, and with $d(x_i, x_{i+1}) \leq r$ for all i. If Υ is a geodesic space, this is equivalent to saying that $N(A; r/2)$ is connected.

We say that a map $f : (\Upsilon, d) \longrightarrow (\Upsilon', d')$ between metric spaces is *coarsely lipschitz* if $(\exists k_1, k_2 \geq 0)(\forall x, y \in \Upsilon)(d'(fx, fy) \leq k_1 d(x, y) + k_2)$. It is a *quasi-isometric embedding* if, in addition, $(\exists k_3, k_4 \geq 0)(\forall x, y \in \Upsilon)(d(x, y) \leq k_3 d'(fx, fy) + k_4)$. We say that f is a *quasi-isometry* if it is a quasi-isometric embedding and $f(\Upsilon)$ is k_5-dense in Υ' for some fixed $k_5 \geq 0$.

A map $f : \Upsilon \longrightarrow \Upsilon'$ is a *coarse embedding* if it is coarsely lipschitz and there is some function, $F : [0, \infty) \longrightarrow [0, \infty)$, such that for all $x, y \in \Upsilon$, $d(x, y) \leq F(d'(fx, fy))$. (So a quasi-isometric embedding corresponds to the case where F

is linear.) Note that a coarse embedding with cobounded image is necessarily a quasi-isometry.

A *coarse quasigeodesic* is a quasi-isometric embedding of a real interval into Υ. If Υ is a geodesic space, then a coarse quasigeodesic can always be approximated up to bounded distance by a quasigeodesic, and we will generally deal with the latter notion.

Recall that a geodesic space, Θ, is *k-hyperbolic* if every geodesic triangle has a *k-centre*; that is a point a distance at most k from each of its three sides. Given $x, y, z \in \Theta$, we choose a k-centre, $\mu(x, y, z)$, for some geodesic triangle with vertices, x, y, z. It is well defined up to bounded distance.

A key fact about hyperbolic spaces is their treelike structure [G]. This can be expressed as follows.

Lemma 5.1. *There is a function, $h : \mathbb{N} \longrightarrow [0, \infty)$ such that if Θ is k-hyperbolic and $A \subseteq \Theta$ with $|A| \leq n < \infty$, then there is an embedded tree, $T \subseteq \Theta$, with $A \subseteq T$ such that for all $a, b \in A$, $d_T(a, b) \leq d(a, b) + kh(n)$, where d_T is the induced path metric on T.*

Here, of course, d_T denotes the intrinsic path-metric induced on T.

Note that we can assume that each edge of T is a geodesic, and that the extreme points of T all lie in $\theta_X A$. We write T_A for some such choice of T.

Given a subset, $A \subseteq \Theta$, we write $\operatorname{join}(A)$ for the union of all geodesics with endpoints in A. It is not hard to see that $\operatorname{hd}(T_A, \operatorname{join}(A))$ is bounded in terms of k and n.

We can view T_A combinatorially as a simplicial tree, and as such, it comes equipped with a median map, $\mu_{T_A} : T_A^3 \longrightarrow T_A$. In this way, (T_A, μ_A) is a rank-1 median algebra, with its vertex set as a subalgebra. One can check easily that for all $x, y, z \in T_A$, $d(\mu(x, y, z), \mu_{T_A}(x, y, z))$ is bounded above in terms of k and n.

This leads naturally to the definition of a "coarse median space" as defined in [Bo1], where a simplicial tree is generalised to (the vertex set of) a CAT(0) cube complex, or equivalently, a discrete median algebra.

Let (Λ, ρ) be a geodesic space, and suppose that $\mu : \Lambda^3 \longrightarrow \Lambda$ is a ternary operation.

Definition. We say that (Λ, ρ, μ) is a *coarse median space* (of rank at most ν) if it satisfies:

(C1): There are constants, $k, h(0)$, such that for all $a, b, c, a', b', c' \in \Lambda$,

$$\rho(\mu(a, b, c), \mu(a', b', c')) \leq k(\rho(a, a') + \rho(b, b') + \rho(c, c')) + h(0).$$

(C2): There is a function $h : \mathbb{N} \longrightarrow [0, \infty)$ such that if $A \subseteq \Lambda$ with $1 \leq |A| \leq n < \infty$, then there is a finite median algebra (of median rank at most ν) and an $h(n)$-quasimorphism, $\lambda : \Pi \longrightarrow \Lambda$ such that for all $a \in A$, $\rho(a, \lambda\pi a) \leq h(n)$.

To say that λ is an *h-quasimorphism* means that $\rho(\lambda\mu_\Pi(x, y, z), \mu(\lambda x, \lambda y, \lambda z)) \leq h$, for all $x, y, z \in \Pi$. (We will apply this terminology more generally, when the domain is any space equipped with a ternary operation. In particular, the inclusion $T_A \hookrightarrow \Theta$ referred to earlier is a quasimorphism.)

There is no loss in assuming that μ is symmetric in a, b, c and that $\mu(a, a, b) = a$ for all $a, b \in \Lambda$ (since these conditions necessarily hold up to bounded distance).

One can show that for geodesic spaces, coarse median of rank 1 is equivalent to hyperbolic (see [Bo1, NWZ1]).

To simplify the exposition, we will suppress mention of the map, λ, and identify Π with its image in Λ. (There is no essential loss in assuming λ to be injective.) We will also assume that $A \subseteq \Pi \subseteq \Lambda$. (Again, this can be achieved after some modification of Π.) This means that the median operation on Π agrees up to bounded distance (depending on n) with the coarse median operation on Λ.

A simple consequence of the axioms is that any tautological median identity (i.e. one that holds exactly in any median algebra) holds up to bounded distance in a coarse median space. The bound depends only on the complexity of the expression and the parameters of the coarse median space. A more formal expression of this principle can be found in [Z], [Bo3] and in [NWZ1].

As an illustration, the identity, $\mu(a, b, \mu(a, b, c)) = \mu(a, b, c)$ holds in any median algebra. Therefore in any coarse median space, Λ, we have that $\rho(\mu(a, b, \mu(a, b, c)), \mu(a, b, c))$ is bounded.

There is also a generalisation of this principle to conditional identities [Z, Bo3, NWZ1].

Given $a, b \in \Lambda$, write $[a, b] = \{\mu(a, b, x) \mid x \in \Lambda\}$ for the *coarse (median) interval* from a to b. Up to bounded Hausdorff distance, it be can described in a number of equivalent ways. For example, given $r \geq 0$, write $[a, b]_r = \{x \in \Lambda \mid \mu(x, \mu(a, b, x)) \leq r\}$. For all sufficiently large r, $\mathrm{hd}([a, b], [a, b]_r)$ is bounded above in terms of r. (Note that in a median algebra, these two definitions correspond to two equivalent ways of defining an interval, as in our earlier illustration.) For more discussion of coarse intervals, see [NWZ2, Bo5].

We remark that any monotone path from a to b lies in (a bounded neighbourhood of) $[a, b]$. If Λ happens to satisfy (B1)–(B10), then one can see from Theorem 1.1 (or Theorem 1.3) that the converse holds, so that up to bounded Hausdorff distance, $[a, b]$ is the union of all monotone paths. (It seems likely that this should hold in greater generality, though we won't explore that issue here.)

We also note that if $a, a', b, b' \in \Lambda$, then $\mathrm{hd}([a, b], [a', b'])$ is bounded above in terms of $\max(\rho(a, a'), \rho(b, b'))$.

The notion of a coarse interval can be thought of as a special case of a "coarse (convex) hull" which we describe in the next section.

6. Coarse hulls

Let Λ be a coarse median space of rank $\nu < \infty$. Our aim is to describe the coarse hull of a subset of Λ. The conclusion of Proposition 6.2 below can be thought of as characterising what we mean by this. We will say a bit more about the structure of hulls under additional assumptions in Section 7. A related construction in the context of hierarchically hyperbolic spaces, by iterating hierarchy paths, has been given independently in [RuST].

Definition. We say that a subset $C \subseteq \Lambda$ is *r-convex* if for all $a, b \in C$, $[a, b] \subseteq N(C; r)$. We say that C is *coarsely convex* if it is r-convex for some r.

In other words, if $a, b \in C$ and $x \in \Lambda$, then $\rho(\mu(a, b, x), C) \leq r$.

It turns out that any interval, $[a, b]$, is uniformly coarsely convex. (Again, this is a consequence of the general principle discussed above.) Of course, this property is not closed under intersection, so it does not cleanly give us a notion of convex hull. However, at least with a bound on the rank, this is possible as we now describe.

Given $A \subseteq \Lambda$, write $J(A) = \bigcup_{a,b \in A} [a, b] \subseteq \Lambda$ for the *coarse median join* of A. We define $J^n(A)$ inductively, by $J^0(A) = A$ and $J^{n+1}(A) = J(J^n(A))$. We will write $J_\Lambda(A)$ and $J^n_\Lambda(A)$ if we need to specify Λ.

Lemma 6.1. *Let Λ be a coarse median space of rank at most ν. Then there is a constant, $k \geq 0$, depending only on the parameters of Λ (including ν) such that if $A \subseteq \Lambda$, then $J^{\nu+1}(A) \subseteq N(J^\nu(A); k)$.*

Proof. In fact the corresponding statement holds tautologically for any median algebra, Π, of rank at most ν, where we define the "median join" similarly in terms of unions of median intervals. In this case, we necessarily have $J^{\nu+1}_\Pi(B) = J^\nu_\Pi(B)$ for any $B \subseteq \Pi$ (which is therefore the median convex hull of B), see Lemma 6.4 of [Bo1].

For Λ, we want to show that if $y \in \Lambda$ and $x_1, x_2 \in J^\nu_\Lambda(A)$ then $z = \mu(y, x_1, x_2)$ is a bounded distance from some point of $J^\nu_\Lambda(A)$. Now the fact that $x_i \in J^\nu_\Lambda(A)$ can be expressed as a median formula of bounded complexity involving elements in some subset $B_i \subseteq A$, with $|B_i| \leq 3^\nu$. Let $B = B_1 \cup B_2 \cup \{y\}$. Since $|B|$ is bounded, we have a finite median algebra, $\Pi \subseteq \Lambda$, with $B \subseteq \Pi$ such that the median operation, μ_Π, on Π agrees with that on Λ up to bounded distance.

We now apply the same formula in Π to the elements of B_i in Π to give us points, $x'_i \in \Pi \subseteq \Lambda$, with $\rho(x_i, x'_i)$ bounded. By definition, x'_i lies in the νth iterated join, $J^\nu_\Pi(B_i)$ of $B_i \subseteq B$ in Π. Let $z' = \mu_\Pi(y, x'_1, x'_2) \in J^{\nu+1}_\Pi(B)$. By the observation of the first paragraph, $z' \in J^\nu_\Pi(B)$. This fact can be expressed by another formula of bounded complexity involving elements of B. Applying the same formula to these elements in Λ with μ in place of μ_Π, we arrive at some $z'' \in \Lambda$, with $\rho(z'', z')$ bounded. By the definition of $J^\nu_\Lambda(B)$, we must have $z'' \in J^\nu_\Lambda(B) \subseteq J^\nu_\Lambda(A)$. Now, $\rho(z, z')$ is also bounded, so $\rho(z, z'')$ is bounded as required. $\square$

Note that the conclusion Lemma 6.1 is equivalent to asserting that $J^\nu(A)$ is k-convex in Λ. In fact, we get the following.

Proposition 6.2. *Let Λ be a coarse median space of rank ν, and let $A \subseteq \Lambda$. There is some $r \geq 0$ depending only on the parameters of Λ (including ν) such* *that there is an r-convex subset $H(A) \subseteq \Lambda$, with $A \subseteq H(A)$ and with the property that if $H' \subseteq \Lambda$ is an r'-convex subset containing A, then $H(A) \subseteq N(H'; r'')$, where r'' depends only on r' and the parameters of Λ.*

Proof. We set $H(A) = J^\nu(A)$. By hypothesis, $J(A) \subseteq N(H'; r')$. We have observed that if $a, b, a', b' \in \Lambda$ with $\rho(a, a') \leq r$ and $\rho(b, b') \leq r$ then $\mathrm{hd}([a, b], [a', b']) \leq r'''$, where r''' depends only on r and the parameters of Λ. Therefore, $J^2(A) \subseteq N(H'; r' + r''')$. The statement follows by induction up to ν. □

Note that the conclusion of Proposition 6.2 determines $H(A)$ up to bounded Hausdorff distance. For definiteness, in what follows, we will set $H(A) = J^\nu(A)$.

One can check that if $a, b \in \Lambda$, then $\mathrm{hd}([a, b], H(\{a, b\}))$ is bounded. In other words, any coarse interval is the coarse hull of its endpoints.

Another immediate consequence is that, up to bounded Hausdorff distance, $H(A)$ equals $\bigcup\{H(B) \mid B \subseteq A, |B| \leq 3^\nu\}$.

Here is another way of describing coarse hulls (cf. [SW, NWZ1, NWZ2]).

First, let (Π, μ_Π) be a median algebra (finite, in the cases of interest here). Given $a_1, \ldots, a_n, x \in \Pi$, with $n \geq 2$, we define y_i inductively for $i \geq 2$ by $y_2 = \mu_\Pi(a_1, a_2, x)$ and $y_{i+1} = \mu_\Pi(a_{i+1}, y_i, x)$. We write $\mu_\Pi(a_1, \ldots, a_n; x) = y_n$. One can show that this is symmetric under any permutation of the a_i (see Lemma 5.2 of [SW]). Thus, given any $A \subseteq \Pi$ with $2 \leq |A| \leq n < \infty$, we can write $\mu_\Pi(A; x) = \mu_\Pi(a_1, \ldots, a_n; x)$ where $A = \{a_1, \ldots, a_n\}$. We also set $\mu_\Pi(\{a\}; x) = a$. (The a_i need not be distinct. If $a_i = a_j$ for $i \neq j$, then the value of the expression does not change on deleting a_j.) One can check that $\mathrm{hull}(A) = \{x \in \Lambda \mid \mu_\Pi(A; x) = x\}$, cf. [SW]. In fact, $\mu_\Pi(A; x) \in \mathrm{hull}(A)$ for all $x \in \Lambda$. In particular, we see that $\mu_\Pi(\mu_\Pi(A; x); x) = \mu_\Pi(A; x)$. (This also follows directly from the fact that $\mu_\Pi(\mu_\Pi(a_1, a_2, x), a_1, x) = \mu_\Pi(a_1, a_2, x)$.)

One can adapt this to a coarse median space, Λ (cf. [NWZ1, NWZ2].) Given $a_1, \ldots, a_n, x \in \Lambda$, one can define $\mu(a_1, \ldots, a_n; x)$ inductively in the same way. As observed in [NWZ2] (Lemma 2.15, thereof), permuting the a_i moves this point at most a bounded distance. We can therefore define $\mu(A; x)$ in a similar way. This is well defined up to bounded distance (depending on $n = |A|$).

Note that from the definition of $H(A)$ it is clear that for all $x \in \Lambda$, $\mu_\Pi(A; x)$ is a bounded distance from $H(A)$.

Lemma 6.3. *For all $x \in \Lambda$, $\rho(\mu(\mu(A; x); x), \mu(A; x))$ is bounded above in terms of n and the parameters of Λ.*

Proof. Given that this corresponds to a tautological identity in a median algebra, the statement follows from the general principle described above.

More explicitly, we can argue as follows. We have a finite median algebra, $\Pi \subseteq \Lambda$, with $A \cup \{x\} \subseteq \Pi$, such that the median, $\mu = \mu_\Lambda$, agrees with μ_Π, on Π up to bounded distance. Now $\mu_\Pi(\mu_\Pi(A; x); x) = \mu_\Pi(A; x)$. Thus, the corresponding statement holds up to bounded distance in Λ. □

Given a finite nonempty subset $A \subseteq \Lambda$ and $r \geq 0$, define $H'_r(A) = \{x \in \Lambda \mid \rho(x, \mu(A; x)) \leq r\}$.

Lemma 6.4. *There is some $r_0 \geq 0$ depending only on n, ν and the parameters of Λ, such that if $r \geq r_0$ and $A \subseteq \Lambda$ with $|A| \leq n < \infty$, then $\mathrm{hd}(H(A), H'_r(A))$ is bounded above in terms of r, n, ν and the parameters of Λ.*

Proof. If $x \in H'_r(A)$, then it follows immediately from Lemma 6.3 that $\rho(x, H(A))$ is bounded.

Conversely, suppose that $x \in H(A) = J^\nu_\Lambda(A)$. This fact can be expressed by a median expression of bounded complexity with arguments in A, and consequently in some subset $B \subseteq A$ with $|B|$ bounded. We assume can $x \in B$. Now there is a finite median algebra $\Pi \subseteq \Lambda$, of rank at most ν and with $B \subseteq \Pi$, such that μ and μ_Π agree up to bounded distance. Reinterpreting this expression with μ_Π in place of μ, we get some $y \in \Pi$, with $\rho(x, y)$ bounded. This expression tells us that $y \in J^\nu_\Pi(B)$. (One can see, inductively on i that if $z \in J^(_\Pi B)$, then $\mu_\Pi(B; y) = y$.) Therefore, $\mu_\Pi(B; y) = y$. It follows that $\rho(\mu(A; x), x) \leq \rho(\mu(B; x), x)$ is bounded as required. □

In view of an earlier observation, it follows that, up to a distance bounded in terms of ν, and for any sufficiently large r, given any $A \subseteq \Lambda$, we have that $H(A)$ is equal to the union of $H'_r(B)$ as B ranges over all subsets of A with $|B| \leq 3^\nu$.

Before concluding this section, we make the general observation that any r-convex subset, C, of a coarse median space, Λ, is r'-connected, where r' depends on r and the parameters of Λ. In fact, if $a, b \in \Lambda$, we can connect a, b by a geodesic, to give us a sequence, $a = x_0, \dots, x_n = b$, with each $\rho(x_i, x_{i+1})$ as small as we like. Setting $y_i = \mu(a, b, x_i)$ we get a sequence $y_0, \dots, y_n$ in $[a, b]$ with $\rho(y_i, y_{i+1})$ bounded. If $a, b \in C$, then we can find points, $z_i \in C$ with $\rho(y_i, z_i) \leq r$. We also have $\rho(z_i, z_{i+1})$ bounded, and the statement now follows easily. Note also that we can choose the x_i so that n is bounded by some linear function of $\rho(a, b)$.

In particular, if C is r-convex, then $N(C; r')$ is path-connected, and the induced path metric is geodesic. Indeed the inclusion of $N(C; r')$ into Λ is a quasi-isometric embedding.

Finally, we note that in the hyperbolic case (i.e. where $\nu = 1$) all of the above becomes much simpler. In particular, the median interval $[a, b]$ is a bounded Hausdorff distance from any geodesic from a to b. It follows that $\mathrm{hd}(H(A), \mathrm{join}(A))$ is bounded above in terms of the hyperbolicity constant. Similarly, if $|A| \leq n < \infty$ then $\mathrm{hd}(H(A), T_A)$ is bounded, where T_A is the tree given by Lemma 5.1, though this bound will also depend on n.

7. Projection maps

In this section, we introduce projection maps to hyperbolic spaces, as mentioned in the introduction.

Let $\mathcal{X}$ be a set with binary relations, $\perp$, $\pitchfork$ and $\prec$, with $\perp$ and $\pitchfork$ both symmetric, and with $\prec$ a strict partial order. (The relation $\pitchfork$ will eventually coincide with that used in Section 3.) We assume that for all $X, Y \in \mathcal{X}$, exactly one of the relations $X = Y$, $X \pitchfork Y$, $X \perp Y$, $X \prec Y$ or $Y \prec X$ holds. Also, if $X, Y, Z \in \mathcal{X}$ with $Y \prec X$ and $X \perp Z$, then $Y \perp Z$.

(As motivation, in the case of the mapping class group or Teichmüller space, $\mathcal{X}$ is a set of subsurfaces of a compact surface. In this case, $\pitchfork$ denotes transversality, $\perp$ denotes disjointness, and $\prec$ denotes strict inclusion. In this context, $\perp$ was denoted by "$\wedge$" in [Bo1, Bo3, Bo5]. See Section 10 for further discussion of this.)

We will assume:

(B1) "disjointness bound": There is some $\nu \in \mathbb{N}$, such that if $\mathcal{Y} \subseteq \mathcal{X}$ with $X \perp Y$ for all $X, Y \in \mathcal{Y}$, then $|\mathcal{Y}| \leq \nu$.

(B2) "nesting bound": There is some $\kappa \in \mathbb{N}$, such that any chain $X_1 \prec X_2 \prec \cdots \prec X_m$ has length, m, at most κ.

In other words, we have $\mathrm{width}(\mathcal{X}, \perp) \leq \nu$ and $\mathrm{width}(\mathcal{X}, \prec) \leq \kappa$.

Now suppose that (Λ, ρ) is a geodesic metric space. Suppose that to each $X \in \mathcal{X}$, we have associated a geodesic metric space $(\Theta(X), \sigma_X)$ and a map $\theta_X : \Lambda \longrightarrow \Theta(X)$. Also, if $Y \in \mathcal{X}$, with $Y \prec X$ or $Y \pitchfork X$, we suppose that we have associated a point, $\theta_X Y \in \Theta(X)$. In (B8) below, $\langle ., . : . \rangle$ denotes the Gromov product, as we discuss afterwards. We will make the following assumptions (motivations for which can be found in [MaM, Bo3, BeHS1] etc.):

(B3) "hyperbolic": $(\exists r_0 \geq 0)(\forall X \in \mathcal{X})$ $\Theta(X)$ is r_0-hyperbolic.

(B4) "disjoint projection": There is some $r_1 \geq 0$ such that if $X, Y, Z \in \mathcal{X}$ with ($Y \pitchfork X$ or $Y \prec X$) and ($Z \pitchfork X$ or $Z \prec X$) and ($Y \perp Z$ or $Y \prec Z$), then $\sigma_X(\theta_X Y, \theta_X Z) \leq r_1$.

(B5) "lipschitz projections": $(\exists k_1, k_2 \geq 0)(\forall x, y \in \Lambda)$ $\sigma_X(\theta_X x, \theta_X y) \leq k_1 \rho(x, y) + k_2$.

(B6) "finiteness": $(\exists r_2 \geq 0)$ $(\forall x, y \in \Lambda)$ the set of $X \in \mathcal{X}$ with $\sigma_X(\theta_X x, \theta_X y) \geq r_2$ is finite.

(B7) "distance bound": $(\forall r \geq 0)(\exists r' \geq 0)(\forall x, y \in \Lambda)$ if $\sigma_X(\theta_X y, \theta_X y) \leq r$ for all $X \in \mathcal{X}$, then $\rho(x, y) \leq r'$.

(B8) "bounded image": $(\exists r_3 \geq 0)(\forall X, Y \in \mathcal{X}$ with $Y \prec X)$ $(\forall x, y \in \Lambda)$ if $\langle \theta_X x, \theta_X y : \theta_X Y \rangle \geq r_3$ then $\sigma_Y(\theta_Y x, \theta_Y y) \leq r_3$. Moreover, if $Z \in \mathcal{X}$ with $Z \pitchfork X$ and $Z \pitchfork Y$ and $\langle \theta_X x, \theta_X Z : \theta_X Y \rangle \geq r_3$, then $\sigma_Y(\theta_Y x, \theta_Y Z) \leq r_3$.

(B9) "transverse projections": $(\exists r_4 \geq 0)(\forall X, Y \in \mathcal{X})$ with $X \pitchfork Y$ if $x \in \Lambda$, then $\min\{\sigma_X(\theta_X x, \theta_X Y), \sigma_Y(\theta_Y x, \theta_Y X)\} \leq r_4$. Moreover if $Z \in \mathcal{X}$ with ($Z \pitchfork X$ or

$Z \prec X$) and ($Z \pitchfork Y$ or $Z \prec Y$) then $\min\{\sigma_X(\theta_X Z, \theta_X Y), \sigma_Y(\theta_Y Z, \theta_Y X)\} \leq r_4$.

(B10) "medians": There is some $r_5 \geq 0$ and a ternary operation, $\mu : \Lambda^3 \longrightarrow \Lambda$ such that $(\forall X \in \mathcal{X})(\forall x, y, z \in \Lambda)$, we have $\sigma_X(\theta_X \mu(x, y, z), \mu_X(\theta_X x, \theta_X y, \theta_X z)) \leq r_5$. Here, μ_X is the standard centroid operation on the hyperbolic space $\Theta(X)$ (which is well defined up to bounded distance depending only on the constant r_0 of (B3)).

For (B8) above, we recall the definition of the "Gromov product" in a metric space, (Θ, σ), as $\langle x, y : z \rangle = (\sigma(x, z) + \sigma(y, z) - \sigma(x, y))/2$. If Θ is hyperbolic, this can be thought of, up to an additive constant, as the distance between z and any geodesic from x to y.

Henceforth, we will abbreviate $\sigma_X(\theta_X x, \theta_X y)$ to $\sigma_X(x, y)$ for $x, y \in \Lambda$ and $X \in \mathcal{X}$. (In this way, we can view σ_X as a pseudometric on Λ.)

Note that (B7) implies that the median operation, μ, described by (B10) is unique up to bounded distance. In other words, it is characterised by the fact that the projection maps θ_X are all uniform quasimorphisms. We also note:

Lemma 7.1. (Λ, ρ, μ) *is a coarse median space of rank at most* ν.

Proof. We just note that hypotheses (P1)–(P4) of [Bo1] are satisfied: (P1) is (B7), (P2) is (B3), (P3) is (B1), and (P4) is a simple consequence of (B8) and (B9) (see the proof of Lemma 11.7 of [Bo1]). The statement now follows from Proposition 10.2 of [Bo1]. □

As noted in the introduction, these hypotheses apply to various naturally occurring spaces. In particular, the axioms (B1)–(B10) are implied by (A1)–(A10) in Section 7 of [Bo3], as well as by the axioms of a hierarchically hyperbolic space given in [BeHS1] or [BeHS2].

Remark. The fact that Axiom 8 of [BeHS2] (namely "partial realisation") is stronger than property (B10) here can be illustrated as follows. Suppose $\mathcal{X} = \{1, 2\}$ with $1 \perp 2$ and $\Theta(1) = \Theta(2) = \mathbb{R}$. Suppose that Λ is an isometrically embedded subset of $\mathbb{R}^2$ with the l^1 metric (for example the diagonal) and let θ_1, θ_2 be projection to the respective coordinates. Then Λ satisfies all the axioms (B1)–(B10), whereas "partial realisation" would require Λ to be cobounded in $\mathbb{R}^2$.

We next interpret the notions of coarse intervals and coarse hulls in these terms.

Recall that in a hyperbolic space such as $\Theta(X)$, there is only one sensible way of defining an "interval" up to bounded distance. To be precise, we write $[a, b]_X$ for the median interval in $\Theta(X)$, as defined in Section 5. It coarsely agrees with any geodesic from a to b. Similarly, writing $H_X(A)$, for the coarse hull of $A \subseteq \Theta(X)$, we see that this is the same, up to bounded Hausdorff distance, as the geodesic join, $\operatorname{join}(A)$. If $|A| \leq n < \infty$, then it also agrees with the tree, T_A, except that the distance bound may then depend on n.

Note that using (B7) and (B10), we see that a set $C \subseteq \Lambda$ is coarsely convex if and only if $\theta_X C$ is coarsely convex in $\Theta(X)$ for all X.

Given $A \subseteq \Lambda$, and $r \geq 0$, let $H''_r(A)$ be the set of $x \in \Lambda$ such that $\sigma_X(\theta_X x, H_X(\theta_X A)) \leq r$ for all $X \in \mathcal{X}$.

The following two lemmas only require the fact that Λ is coarse median of rank at most ν, together with properties (B3), (B5), (B7) and (B10).

Lemma 7.2. *For any $A \subseteq \Lambda$, we have $H(A) \subseteq H''_r(A)$, where r depends only on and the parameters of the hypotheses.*

Proof. Let $x \in H(A)$. Then, as observed in Section 6, we have $x \in H(B)$ for some $B \subseteq A$, with $|B| \leq 3^\nu$. Now, by Lemma 6.3, $\rho(x, \mu(B; x))$ is bounded. By (B5) and (B10), it follows that $\sigma_X(\theta_X x, \mu_X(\theta_X B; \theta_X x))$ is bounded for all $X \in \mathcal{X}$. Therefore, $\theta_X x$ lies a bounded distance from $H_X(B) \subseteq H_X(A)$. In other words, $x \in H''_r(A)$ for some $r \geq 0$ depending only on the parameters of the hypotheses. □

Lemma 7.3. *Suppose $r \geq 0$ and $A \subseteq \Lambda$ with $|A| \leq n < \infty$. Then $H''_r(A) \subseteq N(H(A); r')$, where r' depends only on r, n and the parameters of the hypotheses.*

Proof. Let $x \in H''_r(A)$. So $\theta_X x$ is a bounded distance from $H_X(\theta_X A)$ for all $X \in \mathcal{X}$. Therefore, $\sigma_X(\theta_X x, \mu_X(\theta_X A; \theta_X x))$ is bounded. Up to bounded distance, this is $\sigma_X(\theta_X x, \theta_X \mu(A; x))$, which is therefore also bounded. (These facts use the bound on $|A|$.) Now (B7) tells us that $\rho(x, \mu(A; x))$ is bounded, and so by Lemma 6.4, x is a bounded distance from $H(A)$ as required. □

In particular, this shows that if $|A| \leq n < \infty$ for all sufficiently large r, $H(A)$ and $H''_r(A)$ agree up to bounded distance depending on r and n. (It is not clear whether this holds for arbitrary A, regardless of cardinality.)

This fits in with the discussion in [BeKMM] regarding the mapping class group. The set $H''_r(A)$ is a "Σ-hull" for A in their terminology. This is also essentially the notion of hull that is used in [BeHS3]. In other words, these are equivalent up to bounded distance for finite sets in finite-rank coarse median spaces.

8. Hulls of finite sets

The aim of this section is to show that coarse hulls of finite sets can be coarsely embedded into cubes (see Lemma 8.1).

Let Λ satisfy the conditions (B1)–(B10) laid out in Section 7. Suppose $A \subseteq \Lambda$ with $|A| \leq n < \infty$.

We observed in Section 6 that $H(A)$ is coarsely connected, and there is some $r_0 \geq 0$, depending only on the parameters, such that $H = N(H(A); r_0)$, is intrinsically geodesic, with the inclusion into Λ a uniform quasi-isometric embedding. Moreover, as described in Section 7, $\theta_X(H(A))$, hence also $\theta_X(H)$, is a bounded Hausdorff distance in $\Theta(X)$, from $T(X)$, where $T(X) = T_{\theta_X A}$ is the tree given by Lemma 5.1 and subsequent remarks.

Now the properties (B1)–(B10) are all invariant under quasi-isometry. For most of this section we will simplify notation by assuming that $\Lambda = H$, and that

$\Theta(X) = T(X)$ for all $X \in \mathcal{X}$. We can restrict the original maps, θ_X, to H, and move them a bounded amount, so that their respective images lie in $T(X)$. We can then also assume that μ_X (as in (B10)) is precisely the standard median operation on a tree. (It must be equal to this up to bounded distance anyway.)

We will write $V(T(X))$ for the vertex set of $T(X)$. By construction, $\theta_X A \subseteq V(T(X))$, and every extreme point of $T(X)$ lies in A. Note that $|V(T(X))| \leq 2n-2$.

The next step is to reduce further to a collection of real intervals, $(I(\alpha))_{\alpha\in\mathcal{A}}$, indexed by a set $\mathcal{A}$, together with maps, $\theta_\alpha : \Lambda \longrightarrow I(\alpha)$. It is not hard to see that these again will satisfy properties (B1)–(B10) with appropriate modifications of constants. We will only make explicit those properties which we use subsequently.

To do this, we choose $L \geq 0$ sufficiently large as determined below (see Lemma 8.1). Given $X \in \mathcal{X}$, let $\mathcal{A}(X)$ be an indexing set, and let $(I(\alpha))_{\alpha\in\mathcal{A}(X)}$, be a family of closed real intervals each of length L, with disjoint interiors, and whose interiors do not meet $V(T(X))$. We assume that $|\mathcal{A}(X)|$ is maximal subject to these conditions. (Possibly $\mathcal{A}(X) = \varnothing$.) Note that the total length of $T(X) \setminus \bigcup_{\alpha\in\mathcal{A}(X)} I(\alpha)$ is at most $(2n-2)L$ (since each edge of $T(X)$ contributes at most L). We write $E(\alpha)$ for the boundary of $I(\alpha)$, thought of as a 2-point median algebra. We write σ_α for the metric on $I(\alpha)$.

We write $\tau_\alpha : T(X) \longrightarrow I(\alpha)$ for the nearest-point projection. Clearly, this is 1-lipschitz. Moreover, $\tau_\alpha(\theta_X A) \subseteq E(\alpha)$. Note that if $x \in T(X)$, then $\tau_\alpha x \notin E(\alpha)$ for at most one $\alpha \in \mathcal{A}(X)$.

Let $\Omega(X) = \prod_{\alpha\in\mathcal{A}(X)} E(\alpha)$ be the $|\mathcal{A}(X)|$-cube. Let $\Delta(X) = \prod_{\alpha\in\mathcal{A}(X)} I(\alpha)$. We can think of this as the realisation, $\Delta(\Omega(X))$, of $\Omega(X)$, rescaled by a factor of L. We have a map from $T(X)$ to $\Delta(X)$ obtained by sending x to $(\tau_\alpha(x))_\alpha$. This maps to the 1-skeleton, $\Delta^1(X)$, of $\Delta(X)$. In fact, the image of $T(X)$ in $\Delta(X)$ can be described by collapsing each component of $T(X) \setminus \bigcup_{\alpha\in\mathcal{A}(X)} I(\alpha)$ to a point, so as to give another tree. This collapsing map is a median homomorphism. (If $\mathcal{A}(X) = \varnothing$, then $T(X)$ has bounded diameter, and $\Omega(X)$ is a singleton, so we can effectively ignore such X.)

Given distinct $\alpha, \beta \in \mathcal{A}(X)$, $\tau_\alpha(I(\beta))$ consists of a single point of $E(\alpha)$, which we denote by $\psi_\alpha\beta$. Note that, if $x \in T(X)$, then at least one of $\tau_\alpha x = \psi_\alpha\beta$ or $\tau_\beta x = \psi_\beta\alpha$ holds.

Now set $\mathcal{A} = \bigsqcup_{X\in\mathcal{X}} \mathcal{A}(X)$. Given $\alpha \in \mathcal{A}$, write $X(\alpha) = X$, where $X \in \mathcal{X}$ is such that $\alpha \in \mathcal{A}(X)$.

Write $\Omega(\mathcal{A}) = \prod_{\alpha\in\mathcal{A}} E(\alpha) = \prod_{X\in\mathcal{X}} \Omega(X)$, and $\Delta(\mathcal{A}) = \prod_{\alpha\in\mathcal{A}} I(\alpha) = \prod_{X\in\mathcal{X}} \Delta(X)$. Again, we can think of $\Delta(\mathcal{A})$ as the realisation, $\Delta(\Omega(A))$, of $\Omega(\mathcal{A})$, rescaled by a factor of L. Given $x \in \Lambda$ and $\alpha \in \mathcal{A}$, we have a point, $\tau_\alpha(\theta_{X(\alpha)}x) \in I(\alpha)$, which we will simply denote by $\theta_\alpha(x)$. This gives a map, $\theta_\alpha : \Lambda \longrightarrow E(\alpha)$. Note that $\theta_\alpha(A) = E(\alpha)$. We also get a map $\theta : \Lambda \longrightarrow \Delta(\mathcal{A})$ by setting $\theta(x) = (\theta_\alpha(x))_\alpha$. Note that $\theta(A) \subseteq \Omega(\mathcal{A})$.

We will also want a discrete approximation of this. Given any $\alpha \in \mathcal{A}$, define $\omega_\alpha : I(\alpha) \longrightarrow E(\alpha)$ to be the nearest-point projection (defined arbitrarily on the midpoint of $I(\alpha)$). We write $\phi_\alpha = \omega_\alpha \circ \theta_\alpha : \Lambda \longrightarrow E(\alpha)$, and set $\phi(x) = (\phi_\alpha(x))_\alpha$, so that $\phi : \Lambda \longrightarrow \Omega(\mathcal{A})$. Note that $\phi|A = \theta|A$. Since $\phi_\alpha A = E(\alpha)$ for all $\alpha \in \mathcal{A}$, we see that $\phi(A) \subseteq \Omega(\mathcal{A})$ is filling (as defined in Section 2).

Given any distinct $a, b \in A$, the set $\{\alpha \in \mathcal{A} \mid \phi_\alpha a \neq \phi_\alpha b\}$ is finite, provided L is chosen bigger than the constant, r_2, of (B6). Since $\phi_\alpha(A) = E(\alpha)$ for all α, it follows that $\mathcal{A}$ must be finite. (In particular, for all but finitely many $X \in \mathcal{X}$, we have $\mathcal{A}(X) = \varnothing$, which implies that $T(X)$ has bounded diameter.)

So far, we have defined $\phi : \Lambda \longrightarrow \Omega(\mathcal{A})$. We set $F = \phi\Lambda \subseteq \Omega(\mathcal{A})$. We have already observed that F is filling. We therefore have relations, $\sim$ and $\approx$ on $\mathcal{A}$, as defined in Section 2. We will forget about $\sim$ for the moment. Recall that $\alpha \approx \beta$ means that $\pi_{\alpha\beta} F = \Omega_{\alpha\beta} = E(\alpha) \times E(\beta)$. Therefore $\alpha \not\approx \beta$ means that $|\pi_{\alpha\beta} F| \leq 3$.

We next set about defining the relations, $\pitchfork$, $\perp$, and $\prec$ on $\mathcal{A}$.

Given $\alpha, \beta \in \mathcal{A}$, let $X = X(\alpha)$ and $Y = X(\beta)$, so that $X, Y \in \mathcal{X}$. We split into four cases. We will make some assumptions about L as we go along. (These will retrospectively be used in defining L.)

(S1): $X = Y$.
Write $\alpha \pitchfork \beta$ whenever $\alpha \neq \beta$.
We have already defined $\psi_\alpha\beta$, $\psi_\beta\alpha$ and noted that if $x \in \Lambda$, then either $\theta_\alpha x = \psi_\alpha\beta$ or $\theta_\beta x = \psi_\beta\alpha$. In particular, $|\pi_{\alpha\beta} F| \leq 3$, so $\alpha \not\approx \beta$.

(S2): $X \perp Y$.
We set $\alpha \perp \beta$. (In this case, $\psi_\alpha\beta$ and $\psi_\beta\alpha$ are undefined.)

(S3): $X \pitchfork Y$.
We set $\alpha \pitchfork \beta$.

We assume that $L > 12 r_4$, where r_4 is the constant of (B9). We write $\theta_\alpha\beta = \tau_\alpha(\theta_X Y) \in I(\alpha)$. We write $\psi_\alpha\beta = \omega_\alpha(\theta_\alpha\beta) \in E(\alpha)$ for the nearest point in $E(\alpha)$ to $\theta_\alpha\beta$. In other words, $\sigma_\alpha(\theta_\alpha\beta, \psi_\alpha\beta) \leq L/2$.

In fact, we claim that $\sigma_\alpha(\theta_\alpha\beta, \psi_\alpha\beta) \leq r_4 < L/12$. For if not, let any $a \in A$ be any point of A. By definition, $\theta_\alpha a = \tau_\alpha(\theta_X a) \in E(\alpha)$, and $\theta_\alpha\beta = \tau_\alpha(\theta_X Y)$. Since τ_α is 1-lipschitz, we have $\sigma_X(\theta_X a, \theta_X Y) \geq \sigma_\alpha(\theta_\alpha a, \theta_\alpha\beta) \geq r_4$, and so by (B9), we have $\sigma_\beta(\theta_\beta a, \theta_\beta\alpha) \leq \sigma_Y(\theta_Y a, \theta_Y X) \leq r_4$. Since this holds for all $a \in A$, this gives the contradiction that the diameter of $E(\beta) = \theta_\beta A$ in $I(\beta)$ is at most $2r_4 < L$.

Swapping α and β, we similarly have elements $\theta_\beta\alpha$ and $\psi_\beta\alpha$ in $I(\beta)$, which satisfy $\sigma_\beta(\theta_\beta\alpha, \psi_\beta\alpha) \leq r_4 < L/12$.

If $x \in \Lambda$, then $\min\{\sigma_\alpha(\theta_\alpha x, \psi_\alpha\beta), \sigma_\beta(\theta_\beta x, \psi_\beta\alpha)\} \leq \min\{\sigma_\alpha(\theta_\alpha x, \theta_\alpha\beta), \sigma_\beta(\theta_\beta x, \theta_\beta\alpha)\} + r_4 \leq 2r_4 < L/6$, again, by (B9).

In particular, we must have $\phi_\alpha x = \psi_\alpha\beta$ or $\phi_\beta x = \psi_\beta\alpha$. As in (S1), we get $\alpha \not\approx \beta$.

(S4): $Y \prec X$.
We set $\beta \prec \alpha$.

Here we assume that $L > 6(2r_1 + r_3 + r_4)$, where r_1, r_3, r_4 are respectively the constants of (B4), (B8) and (B9).

First, write $\theta_\alpha \beta = \theta_\alpha(\theta_X Y)$, and let $p_\alpha = \omega_\alpha(\theta_\alpha \beta) \in E(\alpha)$, so that $\sigma_\alpha(\theta_\alpha\beta, p_\alpha) \leq L/2$. Let q_α be the other point of $E(\alpha)$. Thus $\sigma_\alpha(\theta_\alpha\beta, q_\alpha) \geq (L/2) > r_3$. Suppose $a, b \in A$ with $\theta_\alpha a = \theta_\alpha b = q_\alpha$. Since $\theta_\beta a, \theta_\beta b \in E(\beta)$, it follows by (B8) that $\theta_\beta a = \theta_\beta b$. In other words, there is a unique $q_\beta \in E(\beta)$ such that $A \cap \theta_\alpha^{-1} q_\alpha \subseteq A \cap \theta_\beta^{-1} q_\beta$.

This holds whenever $\beta \prec \alpha$. We now consider a number of specific cases.

Suppose first that $\sigma_\alpha(\theta_\alpha\beta, p_\alpha) \geq L/4 > r_3 + r_4 > r_3$, then the same argument as above gives us some $p_\beta \in E(\beta)$ with $A \cap \theta_\alpha^{-1} p_\alpha \subseteq A \cap \theta_\beta^{-1} p_\beta$. Therefore, $A \cap \theta_\alpha^{-1} p_\alpha = A \cap \theta_\beta^{-1} p_\beta$ and $A \cap \theta_\alpha^{-1} q_\alpha = A \cap \theta_\beta^{-1} q_\beta$. Clearly, $p_\beta \neq q_\beta$. In other words, we can write $E(\alpha) = \{p_\alpha, q_\alpha\}$, $E(\beta) = \{p_\beta, q_\beta\}$, and $A = A_p \sqcup A_q$ with $\theta_\alpha A_p = \{p_\alpha\}$, $\theta_\beta A_p = \{p_\beta\}$, $\theta_\alpha A_q = \{q_\alpha\}$ and $\theta_\beta A_q = \{q_\beta\}$. We set $D_{\alpha\beta} = \{(p_\alpha, p_\beta), (q_\alpha, q_\beta)\} \subseteq \Omega_{\alpha\beta} = E(\alpha) \times E(\beta)$. (In other words, $D_{\alpha\beta}$ is the image of A under the map $(\theta_\alpha, \theta_\beta)$ to $\Omega_{\alpha\beta}$.)

Continuing under the assumption that $\sigma_\alpha(\theta_\alpha\beta, p_\alpha) \geq L/4$, suppose that $\gamma \in \mathcal{A}$ with $\gamma \pitchfork \alpha$ and $\gamma \pitchfork \beta$. Let $Z = X(\gamma)$. We cannot have $Z = X$ (otherwise $Y \prec Z$ so $\beta \prec \gamma$) nor $Z = Y$ (otherwise $Z \prec X$ so $\gamma \prec \alpha$). So we must have $Z \pitchfork X$ and $Z \pitchfork Y$. Applying case (S3) above, we see that $\sigma_\alpha(\theta_\alpha\gamma, \psi_\alpha\gamma) \leq r_4$ and $\sigma_\beta(\theta_\beta\gamma, \psi_\beta\gamma) \leq r_4$. We claim that if $\psi_\alpha\gamma = p_\alpha$, then $\psi_\beta\gamma = p_\beta$. To see this, choose any $a \in A_p$, so that $\theta_\alpha a = p_\alpha$. Now $\sigma_\alpha(p_\alpha, \theta_\alpha\gamma) = \sigma_\alpha(\psi_\alpha\gamma, \theta_\alpha\gamma) \leq r_4$ (by (S3)). In other words, $\theta_\alpha a = p_\alpha$, $\sigma_\alpha(p_\alpha, \theta_\alpha\gamma) \leq r_4$, and by assumption $\sigma_\alpha(p_\alpha, \theta_\alpha\beta) > r_3 + r_4$. Therefore $\langle \theta_\alpha a, \theta_\alpha\gamma : \theta_\alpha\beta \rangle > r_3$, and so $\sigma_\beta(\theta_\beta a, \theta_\beta\gamma) \leq r_3$ (by (B8)). Since $\theta_\beta a = p_\beta$, and $r_3 < L/2$, we get $\psi_\beta\gamma = p_\beta$ as claimed. Similarly, if $\psi_\alpha\gamma = q_\alpha$, then $\psi_\beta\gamma = q_\beta$. In other words, we have shown that $(\psi_\alpha\gamma, \psi_\beta\gamma) \in D_{\alpha\beta}$.

Now suppose that $\sigma_\alpha(\theta_\alpha\beta, p_\alpha) \leq L/3$. (Of course, this overlaps with the case of the previous two paragraphs.) We claim that if $x \in \Lambda$, then $\phi_\alpha x = p_\alpha$ or $\phi_\beta x = q_\beta$. For if $\phi_\alpha x = q_\alpha$, then $\sigma_\alpha(\theta_\alpha x, \theta_\alpha\beta) \geq (L/2) - (L/3) \geq r_3$. If $a \in A$ is any point with $\theta_\alpha a = q_\alpha$, then $\langle \theta_\alpha a, \theta_\alpha x : \theta_\alpha\beta \rangle \geq r_3$, so by (B8), we see that $\sigma_\beta(\theta_\beta x, \theta_\beta a) \leq r_3$, so $\phi_\beta x = \theta_\beta a = q_\beta$. This proves the claim. In particular (as in (S1) or (S3)), this shows that $\alpha \not\approx \beta$ in this case.

(The following will be applied in Section 9 with $\mathcal{A}$ replaced by the $\sim$-transversal, $\mathcal{B}$, but that makes no difference to the present discussion, so we retain the same notation.)

Given $\beta \prec \alpha$, we write $\beta \ll \alpha$ if $\sigma_\alpha(\theta_\alpha\beta, E(\alpha)) \geq L/4$ and $\beta \lll \alpha$ if $\sigma_\alpha(\theta_\alpha\beta, E(\alpha)) \geq L/3$. Clearly, $\beta \lll \alpha$ implies $\beta \ll \alpha$. (If we think of $\beta \prec \alpha$ to mean that β is "nested in" in α, then $\beta \ll \alpha$ means that it is "deeply nested", and $\beta \lll \alpha$ means that it is "very deeply nested". This has an interpretation in terms of model spaces, as we discuss in Section 11.) Also note that if $\beta \prec \alpha$ and $\beta \approx \alpha$, then $\beta \lll \alpha$.

Now these relations need not be transitive. However, $\beta \prec \gamma \lll \alpha$ implies $\beta \ll \alpha$. (Since, by (B4), we have $\sigma_\alpha(\theta_\alpha \beta, \theta_\alpha \gamma) \le r_1 \le L/12 = (L/3) - (L/4)$.)

Note that if $\beta \ll \alpha$, then we have defined a pair of opposite corners, $D_{\alpha\beta}$, of the square $\Omega_{\alpha\beta}$ with the property that if $\gamma \pitchfork \alpha$ and $\gamma \pitchfork \beta$, then $(\psi_\alpha \gamma, \psi_\beta \gamma) \in D_{\alpha\beta}$ (cf. Property $(*)$ of Section 3).

In summary, we have the following pentachotomy: given any $\alpha, \beta \in \mathcal{A}$, exactly one of the relations $\alpha = \beta$, $\alpha \perp \beta$, $\alpha \pitchfork \beta$, $\alpha \prec \beta$ or $\beta \prec \alpha$ holds. The first three relations are symmetric, and $\prec$ is transitive. Moreover, $\alpha \prec \beta$ and $\beta \perp \gamma$ implies $\alpha \perp \gamma$. We also have relations $\ll$ and $\lll$ with $\alpha \lll \beta \Rightarrow \alpha \ll \beta \Rightarrow \alpha \prec \beta$. Also, $\alpha \prec \beta \lll \gamma \Rightarrow \alpha \ll \gamma$.

We set $\alpha \boxplus \beta$ to mean $\alpha \ll \beta$ or $\beta \ll \alpha$. From the above, we see that Property $(*)$ of Section 3 is satisfied.

Write $\mathcal{C}$ for the set of $\alpha \in \mathcal{A}$ such that there is no $\beta \in \mathcal{A}$ with $\beta \lll \alpha$. Given any $\alpha \in \mathcal{A} \setminus \mathcal{C}$, there is some $\beta \in \mathcal{C}$ with $\beta \ll \alpha$. (To see this, choose $\gamma \in \mathcal{A}$ with $\gamma \lll \alpha$. If $\gamma \in \mathcal{C}$, set $\beta = \gamma$. If not, since $\mathcal{A}$ is finite, we can find some $\beta \in \mathcal{C}$ with $\beta \prec \gamma$. By the earlier observation, $\beta \ll \alpha$.)

We now relate this to the discussion in Section 3. To this end, we set $\alpha \boxminus \beta$ to mean that $\alpha \ne \beta$ and not $\alpha \pitchfork \beta$. In other words, $\alpha \boxminus \beta$, means that one of $\alpha \perp \beta$ or $\alpha \prec \beta$ or $\beta \prec \alpha$ holds. Setting $\lambda = \nu\kappa$ (the constants from (B1) and (B2)), we see that $\mathrm{width}(\mathcal{A}, \boxminus) \le \lambda$ (as was assumed in Section 3).

Suppose $\alpha, \beta \in \mathcal{C}$ with $\alpha \approx \beta$. We claim that $\alpha \perp \beta$. To see this, note first we cannot have $\alpha \pitchfork \beta$ (by (S3)). Therefore, if not $\alpha \perp \beta$, then we must have either $\beta \prec \alpha$ or $\alpha \prec \beta$. We suppose $\beta \prec \alpha$. In other words, we are in case (S4) above. As noted there, $\beta \prec \alpha$ and $\beta \approx \alpha$ together imply $\beta \lll \alpha$. But this contradicts the definition of $\mathcal{C}$, and so proves the claim. In particular, we see that $\mathrm{width}(\mathcal{C}, \approx) \le \mathrm{width}(\mathcal{C}, \perp) \le \nu$, by (B1).

Next, we investigate properties of the map, $\phi : \Lambda \longrightarrow \Omega(\mathcal{A})$.

To this end, we set $\mathcal{A}(x) = \{\alpha \in \mathcal{A} \mid \sigma_\alpha(\theta_\alpha x, \phi_\alpha x) \ge L/6\}$ for $x \in \Lambda$. In other words, $\theta_\alpha x$ is at least $L/6$ away from $E(\alpha) = \partial I(\alpha)$. If $\alpha, \beta \in \mathcal{A}(x)$ are distinct, then we must have $\alpha \boxminus \beta$. (Otherwise $\alpha \pitchfork \beta$, so by (S3) above, after swapping α and β, we can assume that $\sigma_\alpha(\theta_\alpha x, \psi_\alpha \beta) < L/6$ so $\phi_\alpha x = \psi_\alpha \beta$, contradicting $\alpha \in \mathcal{A}(x)$.) Since $\mathrm{width}(\mathcal{A}, \boxminus) \le \lambda$, it follows that $|\mathcal{A}(x)| \le \lambda$.

If $x, y \in \Lambda$, set $\mathcal{A}(x, y) = \{\alpha \in \mathcal{A} \mid \phi_\alpha x \ne \phi_\alpha y\}$. From the definition of the l^1 metric, $d_{\Omega(\mathcal{A})}$, on $\Omega(\mathcal{A})$, we see that $d_{\Omega(\mathcal{A})}(\phi x, \phi y) = |\mathcal{A}(x, y)|$.

We claim that ϕ is uniformly coarsely lipschitz. Since Λ is a geodesic space, it's enough to bound $d_{\Omega(\mathcal{A})}(\phi x, \phi y)$ in terms of $\rho(x, y)$. Suppose that $\rho(x, y)$ is less than some fixed constant. Since θ_α is coarsely lipschitz, this bounds $\sigma_\alpha(\theta_\alpha x, \theta_\alpha y)$. We assume that L is at least 6 times this bound. If $\alpha \in \mathcal{A}(x, y)$, then $\phi_\alpha x \ne$

$\phi_\beta x$, so at least one of $\sigma_\alpha(\phi_\alpha x, \theta_\alpha x)$ or $\sigma_\alpha(\phi_\alpha y, \theta_\alpha y)$ is at least $L/3$. In other words, $\mathcal{A}(x,y) \subseteq \mathcal{A}(x) \cup \mathcal{A}(y)$, so $|\mathcal{A}(x,y)| \leq |\mathcal{A}(x) \cup \mathcal{A}(y)| \leq 2\lambda$. We see that $d_{\Omega(\mathcal{A})}(\phi x, \phi y) = |\mathcal{A}(x,y)| \leq 2\lambda$ is bounded as required.

We next claim that ϕ is a "coarse embedding" (or "uniform embedding"). That is, for any $x, y \in \Lambda$, $\rho(x,y)$ is bounded above as a function of $d_{\Omega(\mathcal{A})}(\phi x, \phi y)$. Now for any $X \in \mathcal{X}$, $\sigma_X(x,y)$ is bounded above in terms of the number of $\alpha \in \mathcal{A}(X)$ for which $\phi_\alpha x \neq \phi_\alpha y$, that is to say, $|\mathcal{A}(X) \cap \mathcal{A}(x,y)|$. This in turn is at most $d_{\Omega(\mathcal{A})}(\phi x, \phi y)$ hence bounded. By (B7), this bounds $\rho(x,y)$ as required.

Again provided that L is large enough, we claim that ϕ is a quasimorphism. In other words, there is a bound on $d_{\Omega(\mathcal{A})}(\phi\mu(x,y,z), \mu_{\Omega(\mathcal{A})}(\phi x, \phi y, \phi z))$ for $x,y,z \in \Lambda$. We write $m = \mu(x,y,z)$. Write μ_α for the median on $I(\alpha)$. By (B10), $\sigma_X(\theta_X m, \mu_X(\theta_X x, \theta_X y, \theta_X z))$ is bounded for all $X \in \mathcal{X}$, and so in particular, for $X = X(\alpha)$. Therefore, $\sigma_\alpha(\theta_\alpha m, \mu_\alpha(\theta_\alpha x, \theta_\alpha y, \theta_\alpha z)) \leq L_0$, where L_0 is some fixed constant. We assume that $L > 6L_0$. Write $\mathcal{A}(x,y,z) = \{\alpha \in \mathcal{A} \mid \phi_\alpha m \neq \mu_\alpha(\phi_\alpha x, \phi_\alpha y, \phi_\alpha z)\}$, so that $d_{\Omega(\mathcal{A})}(\theta_\alpha m, \mu_{\Omega(\mathcal{A})}(\phi x, \phi y, \phi z)) = |\mathcal{A}(x,y,z)|$. If $\alpha \in \mathcal{A}(x,y,z)$, then at least one of $\theta_\alpha x, \theta_\alpha y, \theta_\alpha z$ lies at least L_0 from $E(\alpha)$. (To see this, note that without loss of generality $\phi_\alpha x = \phi_\alpha y = \mu_\alpha(\phi_\alpha x, \phi_\alpha y, \phi_\alpha z)$, and so if $\sigma_\alpha(\phi_\alpha x, \theta_\alpha x), \sigma_\alpha(\theta_\alpha y, \phi_\alpha y) \leq L_0$, then $\sigma_\alpha(\theta_\alpha m, \phi_\alpha m) \leq 2L_0$, so $\phi_\alpha m = \phi_\alpha x = \phi_\alpha y$.) In other words, $\mathcal{A}(x,y,z) \subseteq \mathcal{A}(x) \cup \mathcal{A}(y) \cup \mathcal{A}(z)$ and so $|\mathcal{A}(x,y,z)| \leq 3\lambda$ as required.

In summary we have shown the following lemma.

Let Λ be a geodesic space satisfying (B1)–(B10). Let $A \subseteq \Lambda$ with $|A| \leq n < \infty$. Let $H = H(A)$ be its coarse hull. Let L be the constant used to define the sets, $\mathcal{A}$, $\Omega(\mathcal{A})$ etc. earlier in this section.

Lemma 8.1. *Provided L is chosen large enough in relation to the parameters of the hypotheses, the map $\phi : H \longrightarrow \Omega(\mathcal{A})$ is coarsely lipschitz, a coarse embedding, and a median quasimorphism, where all the constants of the conclusion depend only on the parameters of the hypotheses.*

Of course, we could also take $H = N(H(A); k)$ for any fixed $k \geq 0$. In fact, we can choose k so that H is intrinsically geodesic, and quasi-isometrically embedded in Λ, as discussed that the end of Section 8.

This will be the starting point of the proof of Theorem 1.3 in the next section.

9. Proofs of the main results

In this section we assemble the earlier constructions to give a proof of Theorem 1.3. We will see that Theorems 1.1 and 1.2 follow easily from this.

Let (Λ, ρ) satisfy (B1)–(B10) of Section 7. Implicit in what follows is the assumption that all constants involved depend ultimately only on those introduced in the hypotheses.

Proof of Theorem 1.3. Suppose that $A \subseteq \Lambda$ with $|A| \leq n < \infty$, and let $H(A)$ be its coarse hull as defined in Section 6. We want to construct a CAT(0) cube complex, Δ, and a map $f : \Delta \longrightarrow \Lambda$ satisfying the conclusion of Theorem 1.3. In particular, f will be a quasi-isometric embedding with $\mathrm{hd}(f(\Delta), H(A))$ bounded.

We have already noted in Section 6 that, for some fixed $k_0 \geq 0$, the set $H = N(H(A); k_0)$ is intrinsically geodesic, and the inclusion of H into Λ is a quasi-isometric embedding. We can therefore equivalently define a map in the opposite direction, namely a quasi-isometry from H to a cube complex, Δ.

We will do this in a series of steps. We will start with a map from H to a cube $\Omega(\mathcal{A})$. We then postcompose with projection to smaller cubes, first to $\Omega(\mathcal{B})$ and then to $\Omega(\mathcal{C})$. We finally take a certain subcomplex, Δ, of $\Delta(\Omega(\mathcal{C}))$ containing the image of H.

To begin this procedure, let $\mathcal{A}$, $\Omega(\mathcal{A})$ and $\phi : H \longrightarrow \Omega(\mathcal{A})$ be as constructed in Section 8. Then $\phi A \subseteq \Omega(\mathcal{A})$ is filling. Moreover, by Lemma 8.1, ϕ is a coarse embedding and a median quasimorphism. In particular, ϕH is k_1-connected for some fixed $k_1 \geq 0$. Let $\sim$ be the equivalence relation on $\mathcal{A}$ defined in Section 2, and let $\mathcal{B} \subseteq \mathcal{A}$ be a $\sim$-transversal. Let $\phi_1 : H \longrightarrow \Omega(\mathcal{B})$ be the composition of ϕ with the projection $\Omega(\mathcal{A}) \longrightarrow \Omega(\mathcal{B})$. Since the latter map is a median homomorphism, ϕ_1 is also a median quasimorphism. Moreover, $\phi_1 H \subseteq \Omega(\mathcal{B})$ is filling. Since $\mathrm{width}(\phi_1 H, \sim) \leq k_1$, the projection from $\Omega(\mathcal{A})$ to $\Omega(\mathcal{B})$ contracts distances by at most some fixed factor. Therefore, ϕ_1 is also a coarse embedding.

Now write $M = \langle \phi_1 H \rangle \subseteq \Omega(\mathcal{B})$ for the median algebra generated by $\phi_1 H$. By Lemma 2.2, M is 1-connected. Therefore, we can identify $\Delta(M)$ as the full subcomplex of $\Delta(\Omega(\mathcal{B}))$ with vertex set M. The inclusion of $\Delta(M)$ into $\Delta(\Omega(\mathcal{B}))$ isometric with respect to the respective l^1 metrics.

We now have relations on $\mathcal{B} \subseteq \mathcal{A}$ as defined in Section 3. Let $\mathcal{C} \subseteq \mathcal{B}$ be as described there (with $\mathcal{B}$ replacing $\mathcal{A}$).

We need to check hypothesis $(*)$ of Section 3. For this, note that $\alpha \boxplus \beta$ was defined in Section 8 to mean that $\alpha \ll \beta$ or $\beta \ll \alpha$ (so that $X(\alpha) \prec X(\beta)$ or $X(\beta) \prec X(\alpha)$). In this case, we constructed $D_{\alpha\beta} \subseteq \Omega_{\alpha\beta}$ and showed there that $(*)$ is satisfied.

Let $\phi_2 : H \longrightarrow \Omega(\mathcal{C})$ be the map ϕ_1 postcomposed with the projection map from $\Omega(\mathcal{B})$ to $\Omega(\mathcal{C})$. Let $\Pi = \langle \phi_2 H \rangle \subseteq \Omega(\mathcal{C})$. This is also the image of M under the projection to $\Omega(\mathcal{C})$. By the construction of $\mathcal{C}$, for any $\beta \in \mathcal{B}$, there is some $\gamma \in \mathcal{C}$ with $\gamma \ll \beta$, hence $\gamma \boxplus \beta$. Therefore the hypotheses of Lemma 3.2 and Corollary 3.3 are satisfied. By Corollary 3.3 and the subsequent remark, the quotient map from $\Delta(M)$ to $\Delta(\Pi)$ is a quasi-isometry. Therefore, ϕ_2 is also a coarse embedding. We observed in Section 8 that $\mathrm{width}(\mathcal{C}, \approx) \leq \mathrm{width}(\mathcal{C}, \perp) \leq \nu$, where $\approx$ is the relation on $\mathcal{C}$ defined by M. Therefore, the dimension of $\Delta(\Pi)$ is at most ν. (Obtaining this dimension bound is the whole point of projecting everything to $\Omega(\mathcal{C})$. The current paragraph can be ignored if we don't require this. Instead, we just recall from Section 3 that $\dim(\Delta(M)) \leq \mathrm{width}(\mathcal{B}, \approx) \leq \lambda = \nu\kappa$ is bounded.)

Now since ϕ_2 is a quasimorphism, we see that the hypotheses of Proposition 4.1 are satisfied with $F = \phi_2 H$. Therefore, $\phi_2 H$ is cobounded in $\Delta = \Delta(\Pi)$. It follows that $\phi_2 : H \longrightarrow \Delta$ is a quasi-isometry.

We now set f to be a quasi-inverse of ϕ_2. This satisfies the conclusion of Theorem 1.3, except that $\phi_2 | A$ might not be injective. Nevertheless $\rho(a, f(\phi_2(a)))$ is bounded for all $a \in A$. One can fix this in a number of simple (if artificial) ways. For example, let A_0 be a copy of A, and adjoin to Δ a free edge from each $b \in A_0$ to $\phi_2(b)$ in Δ. We then adjust f so that it maps each such edge to b. □

Proof of Theorem 1.1. First, we show the existence of monotone paths. Let $a, b \in \Lambda$, and let $[a, b] \subseteq \Lambda$ be the coarse interval. Up to bounded Hausdorff distance, this agrees with $H(\{a, b\})$, and so Theorem 1.3 gives a quasi-isometric embedding $f : \Delta \longrightarrow \Lambda$ with $\mathrm{hd}(f(\Delta), [a, b])$ bounded in terms of the parameters of (B1)–(B10). Let $a_0, b_0 \in \Delta$ be points with $f(a_0) = a$ and $f(b_0) = b$.

Set $t_0 = d_\Delta(a_0, b_0)$, and let $\xi : [0, t_0] \longrightarrow \Delta$ be any l^1 geodesic with $\xi(0) = a_0$ and $\xi(t_0) = b_0$. (For example, take any geodesic from a_0 to b_0 in the 1-skeleton of Δ.) This is a median homomorphism with respect to the standard median on the real interval $[0, t_0]$. Let $\zeta = f \circ \xi : [0, t_0] \longrightarrow [a, b] \subseteq \Lambda$. This is a uniform median quasimorphism, as required by the first statement of Theorem 1.1.

It remains to show that monotone paths can be reparameterised as quasi-geodesics. Let $\zeta : I \longrightarrow \Lambda$ be any monotone path from a to b, where $I \subseteq \mathbb{R}$ is an interval. Up to bounded distance, we can assume that $\zeta(I) \subseteq [a, b]$. After postcomposing with the inverse of the quasi-isometry given by Theorem 1.3, we get a map $\xi : I \longrightarrow \Delta$ which is also monotone. We are therefore reduced to considering monotone paths in cube complexes. In this case, such a path will be a bounded distance from a geodesic in the 1-skeleton of Δ. □

Proof of Theorem 1.2. Although this can be viewed as a direct consequence of Theorem 1.3, we can simplify the argument by assuming that, in the construction of Section 7, each tree $T(X)$ is a geodesic from $\theta_X a$ to $\theta_X b$. Therefore, in the notation used at the end of Section 7, we have $|\mathcal{A}(X) \cap \mathcal{A}(a, b)| = \lfloor \sigma_X(a, b)/L \rfloor$. Provided we assume that $r \geq L$, then $\{\{\sigma_X(a, b)\}\}_r$ agrees with this up to linear bounds. Now $d_\Delta(a_0, b_0) = |\mathcal{A}(a, b)| = \sum_{X \in \mathcal{X}} |\mathcal{A}(X) \cap \mathcal{A}(a, b)|$. Moreover, since f is a quasi-isometric embedding, $d_\Delta(a_0, b_0)$ agrees with $\rho(a, b)$ to within linear bounds. In other other words, we have shown that $\rho(a, b)$ and $\sum_{X \in \mathcal{X}} \{\{\sigma_X(a, b)\}\}_r$ agree to within linear bounds depending on r and the parameters of (B1)–(B10). □

It remains to prove Theorem 1.4. First we need to give a more formal statement of the result.

There are several equivalent ways of formulating the notion of "coarse bounded geometry". Here we will take it to mean that Λ is quasi-isometric to a connected graph of bounded valence and with unit edge-lengths. Since our constructions are quasi-isometry invariant, we may as well assume that Λ is such a graph. We write V for its vertex set. Up to bounded distance, we need only consider subsets which are subgraphs of Λ. We say that a collection, $\mathcal{G}$, of subgraphs of Λ has "at most uniform polynomial growth of degree ν", if there is a polynomial, p, of degree at most ν such that for all $G \in \mathcal{G}$, all $x \in V$ and all $n \in \mathbb{N}$ we have $|G \cap V \cap N(x; n)| \leq p(n)$.

Here will show that, for the collection of intervals in Λ, we can choose such a polynomial to depend only on the parameters of the hypotheses.

Proof of Theorem 1.4. Let $a, b \in \Lambda$. Let Δ be as given by Theorem 1.3. Since it is intrinsically the interval between a and b, it follows via Dilworth's Lemma [Di] that Δ isometrically embeds into $\mathbb{R}^\nu$ in the l^1 metric. (This arises from the fact that "walls" of Δ admit a "ν-colouring" — see for example, [Bo1, Bo2].) The statement now follows easily from the fact that $\mathbb{R}^\nu$ has polynomial growth of degree ν. □

We remark polynomial growth of intervals is a key condition in the criterion for rapid decay given in [ChatR]. See [BeM], [Bo2] and [NWZ2] for further discussion in the present context.

10. Mapping class groups and Teichmüller space

We briefly describe how these results apply to the mapping class group and Teichmüller space.

We begin with the mapping class group. Here the key definitions, and many of the results, can be found in [MaM]. Some further discussion of this case can be found in [BeM, BeKMM, Bo3].

Let Σ be a compact orientable surface of complexity $\xi = \xi(\Sigma) \geq 2$. (That is, the number of boundary components plus 3 times the genus minus 3.) Let $\Lambda = \mathbb{M}(\Sigma)$ be the marking graph. (There are several ways of formulating this, but they are all equivalent up to equivariant quasi-isometry. Alternatively, we could take it to be the Cayley graph of the mapping class group with respect to any finite generating set.) Let $\mathcal{X}$ be the set of subsurface π_1-injective subsurfaces defined up to homotopy, disallowing discs, 3-holed spheres and peripheral annuli. In this context, the relations $\prec$ and $\perp$ refer respectively to inclusion and disjointness up to homotopy. (For clarification, see [Bo1, Bo4] or Section 11 here).

Given $X \in \mathcal{X}$, let $\Theta(X) = \mathbb{G}(X)$ be the curve graph of X. (This needs to appropriately defined when X is an annulus, and in such a case, it is quasi-isometric to $\mathbb{R}$.) Let $\theta_X : \mathbb{M}(\Sigma) \longrightarrow \mathbb{G}(X)$ be subsurface projection (well defined up to bounded distance). In this case, all of the properties (B1)–(B10) of Section 7 are satisfied. In (B1) and (B2) we can take $\nu = \kappa = \xi$.

In this context, the existence of median was established in [BeM], and the fact that $\mathbb{M}(\Sigma)$ is a coarse median space of rank ξ follows from [Bo1]. Theorem 1.1 follows from the resolution of "hierarchies", and Theorem 1.2 is the "distance formula", both established in [MaM]. A version of Proposition 1.4 appears in [BeM]. A version of the coarse hull construction in this context is given in [BeKMM] —

in fact, their "Σ-hulls" are essentially the same as the sets $H''_r(A)$ described in Section 7 here. In this case, Theorem 1.3 follows from the account in [BeHS3].

Remark. It should be pointed out that the indexing set corresponding to $\mathcal{X}$ used in [BeHS2, BeHS3] (there denoted $\mathfrak{S}$) is larger than ours in that it includes disconnected subsurfaces. This is necessary in order to satisfy their "orthogonality" axiom. However, if X is disconnected, then $\Theta(X)$ has bounded diameter. In particular, the corresponding terms do not feature in the distance formula.

The case of Weil-Petersson space, $\mathbb{W}(\Sigma)$, that is Teichmüller space equipped with the Weil-Petersson metric, is similar. In this case, we simply omit all annular subsurfaces from $\mathcal{X}$. We take $\nu = \lfloor (\xi+1)/2 \rfloor$. The distance formula for $\mathbb{W}(\Sigma)$ was also described in [MaM]. (In the account in [Bo5] we included annular subsurfaces in $\mathcal{X}$, but defined $\mathbb{G}(X)$ to be a singleton in that case. This amounts to the same thing.)

If $\Lambda = \mathbb{T}(\Sigma)$ is Teichmüller space in the Teichmüller metric, then $\mathcal{X}$ is the same as for $\mathbb{M}(\Sigma)$ and $\nu = \xi$. However, we now need to modify $\Theta(X)$ when X is an annulus. In this case, it is quasi-isometric to a horodisc in the hyperbolic plane, and we can think of $\mathbb{G}(X)$ as being identified with the bounding horocircle. A distance formula in this case was proven in [Ra], and another proof can be found in [Du]. In particular, if $a, b \in \mathbb{T}(\Sigma)$ lie in the thick part of Teichmüller space, then the distance formula is the same as that for $\mathbb{M}(\Sigma)$ except that we replace each summand $\sigma_{\mathbb{G}(X)}(a, b)$ by $\log(\sigma_{\mathbb{G}(X)}(a, b))$ when X is an annulus. (Here of course, $\sigma_{\mathbb{G}(X)}$ denotes distance in $\mathbb{G}(X)$.) This stems from the fact that a horocircle is exponentially distorted in the hyperbolic plane: therefore to measure distances in the modified curve graph of X we need to introduce a logarithm.

Remark. We remark that the relevant description of $\mathbb{T}(\Sigma)$ makes use of the combinatorial model described in [Du], which in turn makes use of the distance formula to show that it is quasi-isometric. Therefore the result given here cannot really be considered an independent proof of this formula. It is natural to ask if one can show that the combinatorial model is quasi-isometric to $\mathbb{T}(\Sigma)$ without explicit use of this formula.

11. The relation to model spaces

In this section, we motivate various constructions of the paper in terms of "model spaces" which were introduced in [Mi] as a key step in proof the Ending Lamination Conjecture by Brock, Canary and Minsky. These ideas can be formulated in a number of essentially equivalent ways.

For simplicity we will focus on the case where $\Lambda = \mathbb{P}(\Sigma)$ is the pants graph of a compact surface, Σ. (This is quasi-isometric to Weil-Petersson space, $\mathbb{W}(\Sigma)$, [Br]. A very similar discussion would apply to the mapping class group of Σ, though this is complicated by having to deal also with projection to annular subsurfaces.) More about model spaces can be found in [MaM, Mi, Bo7]. Our main focus here will be to give an informal description of an interval in $\mathbb{P}(\Sigma)$, and explain why it is quasi-isometric to a CAT(0) cube complex. This section is logically independent of previous sections. We will use similar notation, though the interpretations are a little different. One can retrospectively see that they are compatible with the earlier definitions.

11.1. **A construction of cube complexes.**

Before introducing surfaces and model spaces, we give an abstract construction of cube complexes. This will serve to motivate the relations $\boxminus$ and $\boxplus$ which have been used throughout this paper.

Let $\mathcal{A}$ be a finite set. To each $\alpha \in \mathcal{A}$, we associate a two-point median algebra, $E(\alpha) = \{e_-(\alpha), e_+(\alpha)\}$. Let $\Omega = \prod_{\alpha\in\mathcal{A}} E(\alpha)$. This is a cube with opposite corners, $c_- = (e_-(\alpha))_\alpha$ and $c_+ = (e_+(\alpha))_\alpha$. Suppose we have relations, $\boxminus$ and $<$, on $\mathcal{A}$, with the property that for all $\alpha, \beta \in \mathcal{A}$, exactly one of the relations $\alpha = \beta$, $\alpha \boxminus \beta$, $\alpha < \beta$ or $\beta < \alpha$ holds. Moreover, we assume that the relation $<$ is transitive. Note that $\boxminus$ is symmetric. We write $\alpha \pitchfork \beta$ to mean $\alpha < \beta$ or $\beta < \alpha$. (Note that $<$ has a different interpretation from $\prec$, as we shall see.)

Suppose $v = (v_\alpha)_\alpha \in \Omega$. We can think of v as determining a direction on each $\alpha \in \mathcal{A}$, which is "positive" if $v_\alpha = e_+(\alpha)$ and "negative" if $v_\beta = e_-(\alpha)$. Let $M \subseteq \Omega$ be the set of $v \in \Omega$ with the property that if $\alpha < \beta \in \mathcal{A}$, then either $v_\alpha = e_+(\alpha)$ or $v_\beta = e_-(\beta)$ (or both). One readily checks that M is a median subalgebra of Ω. Moreover, M is 1-connected. (To see this, let $v, w \in M$. Let $\mathcal{A}(v, w) = \{\alpha \in \mathcal{A} \mid v_\alpha \neq w_\alpha\}$. We proceed by induction on $|\mathcal{A}(v, w)|$. Let $\alpha \in \mathcal{A}(v, w)$ be minimal with respect to $<$. We can suppose that $v_\alpha = e_-(\alpha)$. We reverse the direction on v_α to give $v' \in M$ adjacent to v. Replacing v by v', this reduces $|\mathcal{A}(v, w)|$ and we continue inductively.) It follows that the full subcomplex of Ω with vertex set M is intrinsically CAT(0), and we can identify it with $\Delta(M)$. Note that if two walls of M corresponding to α and β cross (that is there is a square in $\Delta(M)$ with these labels) then $\alpha \boxminus \beta$. It follows that $\dim(\Delta(M)) = \operatorname{rank}(M) \leq \operatorname{width}(\mathcal{A}, \boxminus)$.

One situation in which such a set-up arises occurs when we have associated a real interval to each index α in $\mathcal{A}$ as we now describe. Suppose $J \subseteq \mathbb{R}$ is a non-empty open interval. We write $J = (\partial_- J, \partial_+ J)$ where $\partial_- J < \partial_+ J$. Write $\partial J = \{\partial_- J, \partial_+ J\}$. If J, J' are intervals we write $J < J'$ to mean $\partial_+ J \leq \partial_- J'$. Clearly this implies that $J \cap J' = \varnothing$. Now suppose that $(J(\alpha))_\alpha$ is a family of open intervals indexed by $\mathcal{A}$. We set $E(\alpha) = \partial J(\alpha)$ and define Ω as above. We set $e_\pm(\alpha) = \partial_\pm J(\alpha)$. Suppose that $\alpha < \beta$ implies $J(\alpha) < J(\beta)$. In this case, we write $\psi_\alpha \beta = e_+(\beta)$ and $\psi_\beta \alpha = e_-(\alpha)$. We can think of a vertex $v \in \Omega$ as determining a direction on each interval $J(\alpha)$ in the sense defined above. By definition, v lies in M if, whenever $\alpha \pitchfork \beta$, the directions on $J(\alpha)$ and $J(\beta)$ do not both point away from each other. If $\alpha \pitchfork \beta$, then $J(\alpha) \cap J(\beta) \neq \varnothing$. Note that in this setting, $\alpha \boxminus \beta$ means that $J(\alpha)$ and $J(\beta)$ are distinct and intersect.

We write $D_{\alpha\beta} = \{(e_-(\alpha), e_-(\beta)), (e_+(\alpha), e_+(\beta))\} \subseteq \Omega_{\alpha\beta} \subseteq E(\alpha) \times E(\beta)$. It is easily checked that if $\gamma \in \mathcal{A}$ with $\gamma \pitchfork \alpha$ and $\gamma \pitchfork \beta$, then $\psi_\gamma \alpha = \psi_\gamma \beta$ and $(\psi_\alpha \gamma, \psi_\beta \gamma) \in D_{\alpha\beta}$. If we have any symmetric relation, $\boxplus$, on $\mathcal{A}$ with the property that $\alpha \boxplus \beta$ implies $\alpha \boxminus \beta$, then this gives us property $(*)$ of Section 3.

We can realise $\Delta(\Omega)$ as $\prod_{\alpha \in \mathcal{A}} \bar{J}(\alpha)$. Given $t \in \mathbb{R}$, let $\zeta_\alpha(t)$ be the nearest point of $\bar{J}(\alpha)$ to t, and let $\zeta(t) = (\zeta_\alpha(t))_\alpha \in \Delta(\Omega)$. If $\bar{J} \subseteq \mathbb{R}$ is any compact interval containing $\bigcup_{\alpha \in \mathcal{A}} \bar{J}(\alpha)$, then we get a continuous monotone path, $\zeta : \bar{J} \longrightarrow \Delta(\Omega)$, connecting the opposite corners, c_- and c_+, of $\Delta(\Omega)$. Note that the image of ζ lies in the subcomplex, $\Delta(M) \subseteq \Delta(\Omega)$, described above.

11.2. **Subsurfaces and bands.**

Let S be a finite type orientable surface. By this we mean that it is the interior of a compact surface, $\bar{S}$, with (possibly empty) boundary, $\partial \bar{S}$. (We should note that most accounts elsewhere refer directly to compact surfaces. Here we will deal with open surfaces and subsurfaces. This will simplify notation and terminology for the purposes of our discussion. To be compatible with other accounts, one may need to take closures or metric completions of these surfaces. In particular $\bar{S}$ corresponds to the compact surface, Σ, of Section 10.) We will assume that the complexity, $\xi = \xi(S)$, of S is at least 2. Let $\mathbb{G}(S) = \mathbb{G}(\bar{S})$ be the curve graph of S. Its vertex set, $\mathbb{G}^0(S)$, consists of curves which we can realise simultaneously so that any pair have minimal intersection. (For example, take closed geodesics in any complete finite-area hyperbolic structure on S.) A *multicurve*, $a \subseteq S$, is a non-empty disjoint union of such curves. It is *complete* (or a *pants decomposition*) if each component of $S \setminus a$ is a 3-holed sphere. (This is equivalent to saying that a has exactly ξ components.)

By a *subsurface* of S, we mean an open connected subset, $X \subseteq S$, whose topological boundary, $\partial X \subseteq S$, is either empty (if $X = S$) or a multicurve. Note that X is intrinsically a finite-type surface, and there is a natural map from $\bar{X}$ to $\bar{S}$. Each component of $\partial \bar{X}$ maps either to a component of $\partial \bar{S}$ or to a curve in $\partial X \subseteq S$. It is possible that two components of $\partial \bar{X}$ might get identified to a single curve in ∂X. Note that a subsurface of X is also a subsurface of S.

We write $\mathcal{X}$ for the set of subsurfaces of S which are not 3-holed spheres. (Here we are not including annuli in $\mathcal{X}$, contrary to the notation used in Section 10 here or in [Bo3, Bo5]. Moreover, we have chosen preferred realisations of these subsurfaces. As we have noted, the subsurfaces here are open sets.) Given $X, Y \in \mathcal{X}$, we write $X \prec Y$ to mean $X \subseteq Y$ and $X \neq Y$. We write $X \perp Y$ to mean $X \cap Y = \varnothing$. We write $X \pitchfork Y$ to mean none of $X = Y$, $X \prec Y$, $Y \prec X$ or $X \perp Y$. This gives the usual pentachotomy on $\mathcal{X}$.

We fix an open interval, $J(S) \subseteq \mathbb{R}$, and write $\Psi = S \times J(S)$. We think of the first and last factors as being "horizontal" and "vertical" respectively. A *band* in Ψ is an (open) subset, $B \subseteq \Psi$, of the form $X \times J$, where $X \in \mathcal{X}$ and $J \subseteq J(S)$ is an open interval. We write $\partial_\pm B = X \times \partial_\pm J$, $\partial_H B = \partial_- B \cup \partial_+ B$ and $\partial_V B = \partial X \times J$. Note that Ψ is itself a band (with $\partial_V \Psi = \varnothing$).

Let $\mathcal{A}$ be a finite indexing set, and let $[\alpha \mapsto X(\alpha)] : \mathcal{A} \longrightarrow \mathcal{X}$ be a map. Given $X \in \mathcal{X}$, let $\mathcal{A}(X) = \{\alpha \in \mathcal{A} \mid X(\alpha) = X\}$. Thus $\mathcal{A} = \bigsqcup_{X \in \mathcal{X}} \mathcal{A}(X)$ is a partition of $\mathcal{A}$ into subsets $\mathcal{A}(X) \subseteq \mathcal{A}$, all but finitely many of which are empty. Let $\alpha, \beta \in \mathcal{A}$. If $X(\alpha) = X(\beta)$, we write $\alpha \pitchfork \beta$ to mean that $\alpha \neq \beta$. Otherwise, we write $\alpha \pitchfork \beta$, $\alpha \perp \beta$ and $\alpha \prec \beta$ to mean respectively, $X(\alpha) \pitchfork X(\beta)$, $X(\alpha) \perp X(\beta)$ and $X(\alpha) \prec X(\beta)$. (This accords with the definitions in Section 8.)

We suppose that to each $\alpha \in \mathcal{A}$, we have associated an open interval $J(\alpha) \subseteq J(S)$. Given $X \in \mathcal{X}$, we suppose that the intervals $J(\alpha)$ for $\alpha \in \mathcal{A}(X)$ are disjoint, and that the union of their closures, $\bigcup_{\alpha \in X(\alpha)} \bar{J}(\alpha)$, is a closed interval $\bar{J}(X)$, with interior, $J(X) \subseteq J(S)$. We write $B(\alpha) = X \times J(\alpha)$ and $B(X) = X \times J(X)$. These are bands in Ψ. We say that the family of bands, $\{B(\alpha)\}_{\alpha \in \mathcal{A}}$ is *nested* if for all $\alpha, \beta \in \mathcal{A}$, one of $B(\alpha) \subseteq B(\beta)$, $B(\beta) \subseteq B(\alpha)$ or $B(\alpha) \cap B(\beta) = \varnothing$ holds.

We write $\alpha \boxminus \beta$ to mean $\alpha \neq \beta$ and not $\alpha \pitchfork \beta$. In this case we write $\alpha \ll \beta$ to mean $B(\alpha) \subseteq B(\beta)$ (so that $J(\alpha) \subseteq J(\beta)$). This implies $\alpha \prec \beta$. We write $\alpha \boxplus \beta$ to mean $\alpha \neq \beta$ and either $\alpha \ll \beta$ or $\beta \ll \alpha$. Note that $\alpha \boxplus \beta$ implies $\alpha \boxminus \beta$. This is all consistent with the assumptions we made in Subsection 11.1.

If $\alpha \pitchfork \beta$, we define $\psi_\alpha \beta$ and $\psi_\beta \alpha$ as in Subsection 11.1. Note that, exactly as before, Property $(*)$ of Section 3 now holds.

We now write $\alpha < \beta$ to mean that $\alpha \pitchfork \beta$ and $J(\alpha) < J(\beta)$. Thus, $\alpha \pitchfork \beta$ is equivalent to ($\alpha < \beta$ or $\beta < \alpha$). In general $<$ might not be transitive. However, in the cases of interest to us later, it will be. (Note that for transitivity, it is sufficient that $\alpha < \beta < \gamma$ should imply that $\alpha \pitchfork \gamma$.) If $<$ is indeed transitive, then we can construct the complex, $\Delta(M) \subseteq \Delta(\Omega)$ as above.

11.3. **The pants graph.**

We now relate this to the pants graph. Let $\mathbb{P}(S)$ be the pants graph of S. Its vertex set $\mathbb{P}^0(S)$ is the set of complete multicurves in S. Two multicurves, $a, b \in \mathbb{P}^0(S)$ are deemed adjacent if they can be written in the form $a = c \sqcup d$ and $b = c \sqcup e$ where d, e are components of a, b respectively, and moreover that d, e have minimal possible intersection number (1 or 2) in the component of $S \setminus c$ containing $c \cup d$ (namely a 1-holed torus or a 4-holed sphere). We write $c = c(a, b)$. Note that if $X \in \mathcal{X}$, we can define the pants graph $\mathbb{P}(X)$ intrinsically to X. (If X is complexity-1, then in the above, c will be empty.)

Let $a_0, a_1, \ldots, a_n$ be the vertex set of a path, $\underline{a}$ in $\mathbb{P}(S)$. Let $c_i = c(a_{i-1}, a_i)$. This gives us a sequence of multicurves, $a_0, c_1, a_1, c_2, a_2, \ldots, a_{n-1}, c_n, a_n$, adding or deleting a component at each step. Let $t_0 < u_0 < t_1 < u_1 < \cdots < u_{n-1} < t_n < u_n \in \mathbb{R}$. Let $J(S) = (t_0, u_n)$ and $\Psi = S \times J(S)$. Let $W = \bigcup_{i=0}^n (a_i \times [t_i, u_i]) \cup \bigcup_{i=1}^n (c_i \times [u_{i-1}, t_i]) \subseteq \bar{S} \times \bar{J}(S)$. Thus, W is a disjoint union of "vertical" annuli. It is well defined up to vertical reparameterisation (that is, a map of the form $[(x,t) \mapsto (x, f(t))]$, where f is an orientation preserving self-homeomorphism of

$J(S)$). Note that W also determines the path $\underline{a}$ in $\mathbb{P}(S)$. We can view $\underline{a}$ as being parameterised by $J(S)$. More precisely, we set $\underline{a}(t) = a_i$ if $t \in [t_i, t_{i+1})$, where $t_{n+1} = u_n$.

We say that a band $B = X \times J$ is *compatible* with $\underline{a}$ (or with W) if $\partial_V B \subseteq W$ and if $W \cap \partial_\pm B$ is a complete multicurve in $\partial_\pm B$, which we can identify with an element $a_\pm \in \mathbb{P}(X)$, on identifying $\partial_\pm B$ with X. We say that a family of bands, $(B(\alpha))_{\alpha\in\mathcal{A}}$, is *compatible* with W if it is nested, and $B(\alpha)$ is compatible with W for all $\alpha \in \mathcal{A}$. Note that we get multicurves, $a_\pm(\alpha)$, in $X(\alpha)$. Moreover, $W \cap B(\alpha)$ determines a path, $\underline{a}(\alpha)$, from $a_-(\alpha)$ to $a_+(\alpha)$ in $\mathbb{P}(X(\alpha))$.

So far, the discussion has been combinatorial. We now start on coarse geometry.

Suppose now that $\underline{a}$ is a (coarsely) monotone path in $\mathbb{P}(S)$. By definition, this means that if $i < j < k$, then $\mu(a_i, a_j, a_k)$ is a bounded distance from a_j. Here μ is the median on $\mathbb{P}(S)$. By the characterising property of μ (given as (B10) here), if $X \in \mathcal{X}$, then $\theta_X \underline{a} = (\theta_X a_i)_i$ is a monotone path in the curve graph, $\mathbb{G}(X)$, where $\theta_X : \mathbb{P}(S) \longrightarrow \mathbb{G}(X)$ is subsurface projection. Since $\mathbb{G}(X)$ is hyperbolic, $\theta_X \underline{a}$ fellow travels a geodesic, $I(X)$, from $\theta_X a_-$ to $\theta_X a_+$. In other words, up to reparametisation, the paths remain a bounded distance apart.

Suppose that $(B(\alpha))_{\alpha\in\mathcal{A}}$ is a compatible band system in Ψ. Given $\alpha \in \mathcal{A}(X)$, we have noted that this defines a path, $\underline{a}(\alpha)$ in $\mathbb{P}(X)$, which projects to a path in $\mathbb{G}(X)$ (defined up to bounded distance). By the definition of subsurface projection, $\theta_X \underline{a}(\alpha)$ agrees with a subpath of $\theta_X \underline{a}$, and therefore fellow travels a subpath, $I(\alpha) \subseteq I(X)$. We can think of $I(\alpha)$ as being parameterised by the interval $J(\alpha)$. Up to bounded distance, we can take this parameterisation to be a homeomorphism. In particular, we can identify $\partial I(\alpha)$ with $E(\alpha) = \partial J(\alpha)$.

We now introduce the model space [Mi]. An account using the present terminology can be found in [Bo6]. By the general construction of model spaces, we can find a band system, $(B(X))_{X\in\mathcal{X}}$ with the following property. Suppose $X \in \mathcal{X}$. If $\sigma_X(\theta_X a_-, \theta_X a_+)$ is at most some fixed bound, then $B(X) = \varnothing$. Otherwise, $B(X)$ is a band with $a_\pm(X) = \theta_X a_\pm$, so that $I(X)$ is a geodesic from $\theta_X a_-$ to $\theta_X a_+$. (This corresponds to the tree, $T(X)$, described in Section 8 with $A = \{a_-, a_+\}$.) We can subdivide $I(X)$ into intervals $I(\alpha)$ indexed by some set $\mathcal{A}(X)$, so that each $I(X)$ has length approximately equal to some sufficiently large constant, $L \geq 0$. Writing $\mathcal{A} = \bigsqcup_{X\in\mathcal{X}} \mathcal{A}(X)$, we get a band system $(B(\alpha))_{\alpha\in\mathcal{A}}$ compatible with $\underline{a}$. A feature of the construction of the band system for a monotone path is that the relation, $<$, (defined via the intervals $J(\alpha)$ as in Subsection 11.2) is transitive.

Let $\Delta = \prod_{\alpha\in\mathcal{A}} \bar{J}(\alpha)$. We can identify Δ with $\Delta(\Omega) = \prod_{\alpha\in\mathcal{A}} I(\alpha)$, via the parameterisations, $\bar{J}(\alpha) \longrightarrow I(\alpha)$. As we discussed in Subsection 11.1, a monotone path, $\zeta : \bar{J}(S) \longrightarrow \Delta$ between opposite corners of Δ. Up to bounded distance, this can be described as follows. Given $x \in \mathbb{P}(S)$ and $X \in \mathcal{X}$, we can suppose, up to bounded distance, that $\theta_X x \in I(X)$. If $\alpha \in \mathcal{A}(X)$, let $\zeta'_\alpha(x)$ be the nearest point in the interval $I(\alpha) \subseteq I(X)$. Let $\zeta'(x) = (\zeta'_\alpha(x))_\alpha \in \Delta$. Note that our monotone path is parameterised by a map, $J(S) \longrightarrow \mathbb{P}(S)$. Postcomposing with

ζ gives us a map $J(S) \longrightarrow \Delta$. One can check from the various definitions that ζ and ζ' agree up to bounded distance.

We would like to understand the coarse interval, $[a_-, a_+] \subseteq \mathbb{P}(S)$. We have seen that, up to bounded distance, this is the union of all monotone paths from a_- to a_+. Note that the family of surfaces $(X(\alpha))_\alpha$ is determined (from a_- and a_+) by subsurface projection. Therefore, changing the monotone path does not affect this. However, the family of intervals, $(J(\alpha))_\alpha$ may change. We can think of this as sliding the bands $B(\alpha)$ in the vertical direction in Ψ. In general, the bands may move past or through each other (if the base surfaces are disjoint or nested, respectively). However, the underlying topological structure does not change insofar as the relation $<$ on $\mathcal{A}$ remains constant. (If $\alpha \pitchfork \beta$, then we cannot push $B(\alpha)$ past $B(\beta)$.) Therefore all the monotone paths, ζ, lie in the same subcomplex, $\Delta(M) \subseteq \Delta$, constructed above. Note that this has bounded dimension. In fact, on projecting to $\Delta(\mathcal{C})$, it has dimension bounded by $\nu = \lfloor(\xi + 1)/2\rfloor$. This loosely explains why $[a_-, a_+]$ is quasi-isometric to a CAT(0) cube complex of at most this dimension.

To make this more precise, we needed proper bookkeeping of constants. In particular, we used to relations, $\alpha \ll \beta$ and $\alpha \lll \beta$ which might be interpreted to mean respectively that $B(\alpha)$ is "deeply nested" in $B(\beta)$ and that $B(\alpha)$ is "very deeply nested" in $B(\beta)$. In this way, $\alpha \lll \beta \Rightarrow \alpha \ll \beta \Rightarrow \alpha \prec \beta$. Of course, our proofs made no reference to models or bands. Instead, these properties were described entirely in terms of subsurface projections.

In this section, we have restricted attention to the case where $A = \{a_-, a_+\}$ is a 2-point set. The formal description using subsurface projections generalises fairly readily to any finite set A, though the interpretation in terms of models does not generalise as cleanly.

References

[BaH] H.-J.Bandelt, J.Hedlíková, *Median algebras* : Discrete Math. **45** (1983) 1–30.

[BeHS1] J.A.Behrstock, M.F.Hagen, A.Sisto, *Hierarchically hyperbolic spaces I: curve complexes for cubical groups* : Geom. Topol. **21** (2017) 1731–1804.

[BeHS2] J.A.Behrstock, M.F.Hagen, A.Sisto, *Hierarchically hyperbolic spaces II: Combination theorems and the distance formula* : Pacific J. Math. **299** (2019) 257–338.

[BeHS3] J.A.Behrstock, M.F.Hagen, A.Sisto, *Quasiflats in hierarchically hyperbolic spaces* : preprint, 2017, posted at arXiv:1704.04271.

[BeKMM] J.A.Behrstock, B.Kleiner, Y.Minsky, L.Mosher, *Geometry and rigidity of mapping class groups* : Geom. Topol. **16** (2012) 781–888.

[BeM] J.A.Behrstock, Y.N.Minsky, *Centroids and the rapid decay property in mapping class groups* : J. London Math. Soc. **84** (2011) 765–784.

[Bo1] B.H.Bowditch, *Coarse median spaces and groups* : Pacific J. Math. **261** (2013) 53–93.

[Bo2] B.H.Bowditch, *Embedding median algebras in products of trees* : Geom. Dedicata **170** (2014) 157–176.

[Bo3] B.H.Bowditch, *Large-scale rigidity properties of the mapping class groups* : Pacific J. Math. **293** (2018) 1–73.

[Bo4] B.H.Bowditch, *Notes on coarse median spaces* : in "Beyond hyperbolicity", LMS Lecture Note Series, Vol. 454 (ed. M.Hagen, R.C.H.Webb, H.Wilton) Cambridge University Press (2019) 3–24.

[Bo5] B.H.Bowditch, *Large-scale rank and rigidity of the Weil-Petersson metric* : Groups Geom. Dyn. **14** (2020) 607–652.

[Bo6] B.H.Bowditch, *The ending lamination theorem* : preprint, Warwick 2011.

[Bo7] B.H.Bowditch, *Quasiflats in coarse median spaces* : preprint, Warwick 2019.

[Br] J.F.Brock, *The Weil-Petersson metric and volumes of 3-dimensional hyperbolic convex cores* : J. Amer. Math. Soc. **16** (2003) 495–535.

[ChalCGHO] J.Chalopin, V.Chepoi, A.Genevois, H.Hirai, D.Osajda, *Helly groups* : preprint, 2020, posted at arXiv:2002.06895.

[ChatR] I.Chatterji, K.Ruane, *Some geometric groups with rapid decay* : Geom. Funct. Anal. **15** (2005) 311–339.

[Che] V.Chepoi, *Graphs of some CAT(0) complexes* : Adv. in Appl. Math. **24** (2000) 125–179.

[Di] R.P.Dilworth, *A decomposition theorem for partially ordered sets* : Ann. Math. **51** (1950) 161–166.

[Du] M.G.Durham, *The augmented marking complex of a surface* : J. Lond. Math. Soc. **94** (2016) 933–969.

[G] M.Gromov, *Hyperbolic groups* : in "Essays in group theory", Math. Sci. Res. Inst. Publ. No. 8, Springer (1987) 75–263.

[HHP] T.Haettel, N.Hoda, H.Petyt, *The coarse Helly property, hierarchical hyperbolicity and semihyperbolicity* : preprint, 2020, posted at arXiv:2009.14053.

[I] J.R.Isbell, *Median algebra* : Trans. Amer. Math. Soc. **260** (1980) 319–362.

[MaM] H.A.Masur, Y.N.Minsky, *Geometry of the complex of curves II: hierarchical structure* : Geom. Funct. Anal. **10** (2000) 902–974.

[Mi] Y.Minsky, *The classification of Kleinian surface groups. I. Models and bounds* : Ann. of Math. **171** (2010) 1–107.

[NWZ1] G.A.Niblo, N.Wright, J.Zhang, *A four point characterisation for coarse median spaces* : Groups Geom. Dyn. **13** (2019) 939–980.

[NWZ2] G.A.Niblo, N.Wright, J.Zhang, *The intrinsic geometry of coarse median spaces and their intervals* : preprint, 2018, posted at: arXiv:1802.02499.

[Ra] K.Rafi, *A combinatorial model for the Teichmüller metric* : Geom. Funct. Anal. **17** (2007) 936–959.

[Ro] M.A.Roller, *Poc-sets, median algebras and group actions, an extended study of Dunwoody's construction and Sageev's theorem* : Habilitationsschrift, Regensberg, 1998.

[RuST] J.Russell, D.Spriano, H.C.Tran, *Convexity in hierarchically hyperbolic spaces* : preprint 2018, posted at: arXiv:1809.09303.

[SW] J.Špakula, N.Wright, *Coarse medians and property A* : Algebr. Geom. Topol. **17** (2017) 2481–2498.

[Z] R.Zeidler, *Coarse median structures on groups* : Masterarbeit, University of Vienna, 2013.

Mathematics Institute, University of Warwick, Coventry, CV4 7AL, Great Britain