Thus, as $i \to \infty$, we see that x'_i, y'_i, z'_i converge to the same point, say z', on $[x, y]$. Thus $d(x_i, y'_i) + d(y_i, x'_i) - d(x_i, y_i)$ converges to zero. Since $d(x_i, z_i) + d(y_i, z_i) - d(x_i, y_i)$ also converges to zero, we have that $d(y'_i, z_i) + d(z_i, x'_i)$ converges to zero. Since $d(z_i, z'_i) \leqslant d(z'_i, y'_i) + d(y'_i, z_i) \leqslant 4\delta_i + d(y'_i, z_i)$ we see that the z_i converge to the point z' on our original geodesic segment $[x, y]$. Thus z, the midpoint of our arbitrary geodesic from x to y, coincides with the midpoint of our fixed geodesic. Repeating the argument we see that these geodesics must agree at a dense set of points, and hence everywhere. Since geodesic triangles in C_i are δ_i-slim, and geodesics in C all arise as limits of geodesics in C_i, we see that geodesic triangles in C must be 0-slim, and hence C is an $\mathbf{R}$-tree. □

Remark. If one has a sequence of δ_i-hyperbolic spaces C_i, with $C_i \to C$ and $\delta_i \to \delta > 0$, then one can extend the preceding argument to show that C is δ'-hyperbolic (with $\delta' = 19\delta$, for example).

SECTION 2: THE PROOF OF PAULIN'S THEOREM

In this section we shall prove the following theorem of F. Paulin [P4].

2.1 THEOREM (Paulin). *If Γ is a word hyperbolic group and $Out(\Gamma)$ is infinite, then Γ acts by isometries on an* $\mathbf{R}$*-tree with virtually cyclic segment stabilizers and no global fixed points.*

In its outline, the proof given below is very similar to Paulin's original proof, except that we use Hausdorff-Gromov convergence instead of the equivariant Gromov convergence used by Paulin. In particular, this allows us to avoid the difficulties discussed in the next section.

Let S be a finite set of generators for Γ and let $X = X(\Gamma, S)$ denote the Cayley graph of Γ with respect to S, as defined in the introduction. Γ is the vertex set of X and receives the induced metric. The hypothesis that Γ is word hyperbolic means precisely that there exists $\delta > 0$ such that X is a δ-hyperbolic geodesic metric space. Note that with our definition of a Cayley graph, the endpoints of each edge are distinct, and there is at most one edge joining each pair of vertices; hence the action of Γ on itself be left multiplication can be extended linearly across edges in a unique way to give an isometric action of Γ on X.

The proof of Theorem 2.1 will be broken into a number of smaller results. We begin by noting that, because $Out(\Gamma)$ is infinite, we can choose a sequence of automorphisms $\{\phi_i\}_{i \in \mathbf{N}}$ such that none of the ϕ_i is an inner automorphism and no two of the ϕ_i have the same image in $Out(\Gamma)$. For each $i \in \mathbf{N}$ we consider the function $f_i : X \to [0, \infty)$ defined by:

$$f_i(x) = \max_{s \in S} d(x, \phi_i(s)x) . \tag{2.2}$$

This function has been used by Bestvina in his study of degeneration of real hyperbolic structures [B], and our use of this function is similar to his. (A similar idea was used earlier in a different context by Thurston [T, Prop. 1.1].)

Note that f_i takes on integer values at vertices and midpoints of edges in X, and its restriction to half-edges is linear. It follows that f_i attains its infimum (which is an integer) at some point, $x_i \in X$ say. (In the case where Γ is not virtually cyclic one can also see this by showing that f_i is a proper map, i.e., a map with the property that the inverse image of a compact set is compact.)

Let

$$\begin{aligned} \lambda_i &= \max_{s \in S} d(x_i, \phi_i(s)x_i) \\ &= \inf_{x \in X} \max_{s \in S} d(x, \phi_i(s)x) . \end{aligned} \tag{2.3}$$

We fix a definite choice of points x_i with the above property.

For future reference, we note that by passing to a subsequence of the ϕ_i we may assume there is a single element $s_0 \in S$ such that $\lambda_i = d(x_i, \phi_i(s_0)x_i)$ for all $i \in \mathbf{N}$. We also note that with the above choice of x_i, the triangle inequality yields:

$$d(x_i, \phi_i(\gamma)x_i) \leqslant \lambda_i d(e, \gamma) . \tag{2.4}$$

Following Paulin, we next note that because $Out(\Gamma)$ is infinite, the sequence λ_i must be unbounded. For suppose that there were a uniform bound, ρ say, on the value of λ_i. Then for any vertex $y_i \in X$ closest to x_i, we would have $d(e, y_i^{-1}\phi_i(s)y_i) = d(y_i, \phi_i(s)y_i) \leqslant \rho + 2$ for all $s \in S$, $i \in \mathbf{N}$. But there are only finitely many vertices in the ball of radius $\rho + 2$ about e, so this bound would imply the existence of integers $n \neq m$ such that $y_n^{-1}\phi_n(s)y_n = y_m^{-1}\phi_m(s)y_m$ for all $s \in S$. Whence ϕ_n and ϕ_m would be equal in $Out(\Gamma)$, contrary to hypothesis. Thus we have shown that the sequence of numbers $\{\lambda_i\}_{i \in \mathbf{N}}$ is unbounded, so we may pass to a subsequence $\{\lambda_n\}_{n \in \mathbf{N}}$ which is *strictly increasing* and assume that $\lambda_n \to \infty$ as $n \to \infty$.

Consider the sequence of metric spaces $X_k = (X, d_k)$, where $d_k := d/\lambda_k$ is the original metric on X scaled down by λ_k. In what follows we shall intermittently use both the original metric d and the scaled metric d_k, specifying which on each occasion and, where appropriate, using the formal notation (Y, d) for a metric space which consists of the set Y together with a distance function d. But for the moment, the most important distinction between the X_k will be that we shall regard Γ as acting on X_k via ϕ_k, and think of our chosen point x_k, at which the minimax λ_k is attained, as a *basepoint* in X_k. More precisely, we consider the sequence of pointed Γ-spaces (X_k, x_k), where the action of $\gamma \in \Gamma$ on X_k is $x \to \phi_k(\gamma)x$.

We wish to use the hyperbolic nature of X_k to approximate it by a sequence of star-like compact subsets $X_k(P_i)$ centred at x_k. To this end, we fix a sequence of finite subsets $\{e\} = P_0 \subseteq P_1 \subseteq P_2 \cdots \subseteq P_i \subseteq \cdots$ which exhaust Γ. Let $n_i = |P_i|$ denote the cardinality of P_i. The desired subsets of X_k are defined inductively as follows: $X_k(P_0) = \{x_k\}$, and $X_k(P_i)$ is the union of $n_i - 1$ geodesic segments, those in $X_k(P_{i-1})$ together with a choice of geodesic segment from x_k to each element of $\{\phi_k(\gamma)x_k \mid \gamma \in P_i - P_{i-1}\}$.

We next 'fatten-up' each of the sets $X_k(P_i)$ by taking its closed δ-neighbourhood in the metric d. Henceforth we shall denote this neighbourhood V_k^i. Let $d_{i,k}$ be the induced *path metric* on V_k^i. As we discussed in Section 1, $(V_k^i, d_{i,k})$ is a geodesic metric space. It is also important to notice that the induced path metric which V_k^i receives from d_k is $d_{i,k}/\lambda_k$. The following lemma is suggested by an argument of B. Bowditch [Bo].

2.5 LEMMA. *With the above notation, for all* $x, y \in V_k^i$ *we have:*

$$d(x, y) \leqslant d_{i,k}(x, y) \leqslant d(x, y) + 4\delta\ .$$

Proof. The left-most inequality comes from the general fact that for any subspace of a geodesic metric space the induced metric is dominated by the induced path metric. In order to establish the other inequality, we first note that $X_k(P_i)$ is δ-convex in (X_k, d), in the sense that if a geodesic segment in X_k joins a pair of points $x, y \in X_k(P_i)$, then this geodesic segment lies entirely within the closed δ-neighbourhood V_k^i of $X_k(P_i)$.

Given $x, y \in V_k^i$, we fix points $z, w \in X_k(P_i)$ closest to x and y respectively. (Such points are not unique in general.) Let $[x, z]$, $[z, w]$ and $[w, y]$ be choices of geodesic segments joining x to z, z to w and w to y, respectively. Each is contained in V_k^i, and hence so is the broken geodesic $[x, z, w, y]$ obtained by concatenating them. The length of this broken geodesic is at most $d(z, w) + 2\delta \leqslant d(x, y) + 4\delta$. Hence $d_{i,k}(x, y) \leqslant d(x, y) + 4\delta$. □

The subspace V^i_k forms a good substitute for the notion of a convex hull for $\phi_k(P_i)x_i$ in X_k. According to the above lemma, geodesics in $(V^i_k, d_{i,k})$ are $(1, 4\delta)$-quasigeodesics in (X_k, d), and hence by [GH, p. 82] there exists a constant $\eta = \eta(\delta)$ (independent of k, i) such that geodesic triangles in $(V^i_k, d_{i,k})$ are η-slim. Thus we have proved the first part of:

2.6 LEMMA. *There exists a constant* $\eta = \eta(\delta)$ *such that, for all* $k \in \mathbf{N}$, *with respect to the path metric* $d_{i,k}$ *on* V^i_k, *geodesic triangles in* V^i_k *are* η*-slim. Moreover, for fixed* i, *with respect to the (scaled) path metrics* $d_{i,k}/\lambda_k$, *the metric spaces* $\{V^i_k\}_{k \in \mathrm{N}}$ *are uniformly compact.*

Proof. It remains to prove the assertion of the second sentence. We follow an argument of Bestvina [B]. Until further notice we work with the metric d. Let μ_i be the maximum of the integers $\{d(e, \gamma) \mid \gamma \in P_i\}$. Each of the geodesic segments used to define $X_k(P_i)$ has length at most $\mu_i\lambda_k$ (by (2.4)). Therefore, given $\varepsilon > 0$, we can cover $X_k(P_i)$ by $2n_i\mu_i/\varepsilon$ segments of length at most $\lambda_k\varepsilon/2$. (Recall that $n_i = |P_i|$.) Hence, if $\lambda_k\varepsilon > 2\delta$, then in order to cover V^i_k we need at most $2n_i\mu_i/\varepsilon$ balls of radius $\lambda_k\varepsilon$. But we arranged that $\lambda_k \to \infty$ as $k \to \infty$, so this is true for large k.

Now we change viewpoints and work with the scaled metric d_k on X_k, and the induced path metric on V^i_k. In this setting, the preceding argument shows that for large k one needs only $2n_i\mu_i/\varepsilon$ balls of radius ε to cover V^i_k. Since the path metric on V^i_k and the restriction to V^i_k of d_k differ by at most an additive constant of $4\delta/\lambda_k$, we have thus established the existence of a uniform ε-count for the $\{V^i_k\}_{k \in \mathrm{N}}$ both when equipped with the restriction of the metrics d_k and when equipped with the induced path metrics. Because they are *path* metric spaces, a uniform ε-count also yields a bound on the diameter of the V^i_k. □

Continuing with the proof of Paulin's theorem, we fix an integer j and suppose that we are given a positive constant ε. According to the preceding lemma, we can choose ε-nets $N_\varepsilon(k, j)$ for V^j_k on whose cardinalities there is a bound independent of k. We may also assume that the set $N_\varepsilon(k, j)$ includes $\phi_k(P_j)x_k$. Since, for fixed j, the $N_\varepsilon(k, j)$ are finite metric spaces of uniformly bounded cardinality and diameter, we can pass to a subsequence (using a diagonal type argument, as in Section 1) so as assume that, for all $\gamma, \gamma' \in P_j$, the sequence of numbers $d_{j,k}(\phi_k(\gamma)x_k, \phi_k(\gamma')x_k)$ converges as $k \to \infty$. Passing to a further subsequence which is convergent in the Hausdorff-Gromov topology we obtain a limit metric space $L_{\varepsilon,j}$ (whose cardinality will be no greater than that of the $N_\varepsilon(k, j)$). As a basepoint in the limit space we choose the limit of the sequence x_k, and we christen this point x_∞. For each $\gamma \in P_j$, we denote the limit of the sequence $\phi_k(\gamma)x_k$ by γx_∞.

We next take an $\varepsilon/2$-net for V^j_k which is constructed so as to include the previously chosen ε-net. Passing to a subsequence if necessary, we obtain a finite limit metric space $L_{\varepsilon/2,j}$. We proceed in this manner, taking finer ε-nets, and at each stage including the previous (coarser) ones and extracting convergent subsequences to obtain finite limit metric spaces. The natural inclusions of each ε-net into its refinements gives a natural identification of points in the limit, so it is not too abusive a notation to write:

$$L_{\varepsilon,j} \subset L_{\varepsilon/2,j} \cdots \subset L_{\varepsilon/2^n,j} \subset \cdots$$

We define L_j to be the direct limit of this sequence, that is, $L_j = \bigcup\{L_{\varepsilon/2^n,j} \mid n \in \mathbf{N}\}$. We denote by $\hat{L}_j$ the metric completion of L_j. Since the diameters of the V^j_k are uniformly bounded in the scaled metrics, we see that $\hat{L}_j$ is a complete space of finite diameter, and hence is compact.

By choosing a diagonal type subsequence and renumbering, we obtain the following array of spaces with convergence in both the horizontal and vertical directions:

$$\begin{array}{ccccccccccc}
N_\varepsilon(1,j) & \subseteq & N_{\varepsilon/2}(2,j) & \subseteq & \cdots\cdots & \subseteq & N_{\varepsilon/2^n}(1,j) & \subseteq & \cdots & \subseteq & V^j_1 \subseteq X_1 \\
N_\varepsilon(2,j) & \subseteq & N_{\varepsilon/2}(2,j) & \subseteq & \cdots\cdots & \subseteq & N_{\varepsilon/2^n}(2,j) & \subseteq & \cdots & \subseteq & V^j_2 \subseteq X_2 \\
\vdots & & \vdots & & & & \vdots & & & & \vdots \\
N_\varepsilon(m,j) & \subseteq & N_{\varepsilon/2}(m,j) & \subseteq & \cdots\cdots & \subseteq & N_{\varepsilon/2^n}(m,j) & \subseteq & \cdots & \subseteq & V^j_m \subseteq X_m \\
\vdots & & \vdots & & & & \vdots & & & & \vdots \\
L_{\varepsilon,j} & \subseteq & L_{\varepsilon/2,j} & \subseteq & \cdots\cdots & \subseteq & L_{\varepsilon/2^n,j} & \subseteq & \cdots & \subseteq & \hat{L}_j
\end{array}$$

Our next goal is to show that as $k \to \infty$ the V^j_k actually converge to $\hat{L}_j$ in the Hausdorff-Gromov topology. We have that $N_{\varepsilon/2^n}(m,j)$ is $\varepsilon/2^{n-1}$ close to V^j_m for all m. After passing to yet another diagonal type subsequence, we may assume that $N_{\varepsilon/2^n}(m,j)$ is $\varepsilon/2^{m-1}$ close to $L_{\varepsilon/2^n,j}$ for all $m \geqslant n$. Thus V^j_m and $L_{\varepsilon/2^n,j}$ are $\varepsilon/2^{n-2}$ close for $m \geqslant n$. On the other hand, $L_{\varepsilon/2^n,j}$ and $L_{\varepsilon/2^{n+1},j}$ are $\varepsilon/2^{n+1}$ close (since any choice of $\varepsilon/2^n$ and $\varepsilon/2^{n+1}$ nets of V^j_k are $\varepsilon/2^{n+1}$ close). Thus $L_{\varepsilon/2^n,j}$ is $\Sigma_{i \geqslant n}\varepsilon/2^i$ close to L_j and $\hat{L}_j$. Hence V^j_n and $\hat{L}_j$ are $\varepsilon/2^{n-3}$ close, so V^j_n converges to $\hat{L}_j$, in the Hausdorff-Gromov topology, as $n \to \infty$.

Notice that, by (1.9) and (2.6), the spaces $\hat{L}_j$ are **R**-trees of finite diameter, because V_k^j is η/λ_k-hyperbolic and $\lambda_k \to \infty$. It is also useful to observe that $\hat{L}_j$ is spanned by γx_∞, with $\gamma \in P_j$. Furthermore, the $X_k(P_j)$ themselves converge to $\hat{L}_j$ because $X_k(P_j)$ and V_k^j are $4\delta/\lambda_k$-close and $\lambda_k \to \infty$. However, in what follows it is most convenient to still work with V_k^j rather than $X_k(P_j)$ when we need to take a choice of geodesic between two points of $X_k(P_j)$. Also, because the scaled path metric on V_k^j and the induced metric d_k/λ_k differ only by $4\delta/\lambda_k$, which tends to 0 as $k \to \infty$, henceforth it is not important to keep track of the difference between these two metrics.

By construction, all of our $\varepsilon/2^n$-nets include the set $\{\phi(\gamma)x_k \mid \gamma \in P_j\}$ and each of the sequences $d_k(\phi(\gamma)x_k, \phi(\gamma')x_k)$ converges. Thus, if we denote by $x_\infty \in \hat{L}_j$ the 'limit' of the x_k, and by γx_∞ the limit of the $\phi(\gamma)x_k$, then we see that $d(\gamma x_\infty, \gamma' x_\infty)$ (distance in $\hat{L}_j$) is independent of j. Since the tree $\hat{L}_j$ is the convex hull of the points γx_∞, we can define an isometric embedding of $\hat{L}_j$ into $\hat{L}_{j+1}$ for all j and hence obtain an **R**-tree by taking the direct limit of the resulting system of inclusions. We denote the direct limit metric space with basepoint (which as the limit of **R**-trees is itself an **R**-tree) by $(X_\infty; x_\infty)$. The final important observation to make is that Γ acts isometrically on X_∞, because it acts isometrically on the subset $\{\gamma x_\infty\}_{\gamma \in \Gamma}$ (by left translation), and the convex hull of this subset is the whole of X_∞.

Let us now examine the nature of the action of Γ on X_∞. We claim that it has the following properties:

(1) There is no point of X_∞ whose stabilizer is the whole of Γ.

(2) The stabilizer of every non-trivial segment in X_∞ is virtually cyclic.

To see that (1) is true, let us see what would happen if it were to fail. Suppose that Γ were to stabilize a point $z_\infty \in X_\infty$. We fix a segment $z_\infty \in [\gamma x_\infty, \gamma' x_\infty] \subseteq \hat{L}_j$. Up to the taking of subsequences, we have that the closures in $\hat{L}_j$ of the images of the geodesic segments $[\gamma x_k, \gamma' x_k] \subseteq V_k^j$ converge (in the Hausdorff metric) to $[\gamma x_\infty, \gamma' x_\infty]$, and we fix points $z_k \in [\gamma x_k, \gamma' x_k]$ which converge to z_∞. We then choose j large enough to ensure that $S \subset P_j$ (recall that S is our fixed finite generating set for Γ), and l large enough to ensure that $P_j P_j \subset P_l$.

We have, for every $s \in S$, geodesics $[s\gamma x_k, s\gamma' x_k] := s \cdot [\gamma x_k, \gamma' x_k]$ in V_k^l, and (by definition of the action on X_∞) the closures of their images in $\hat{L}_l \subseteq X_\infty$ converge to $[s\gamma x_\infty, s\gamma' x_\infty]$. Moreover, $\{sz_k\}_{k \in \mathbf{N}}$ converges to $s \cdot z_\infty = z_\infty$, so for large k we have that $d_k(s \cdot z_k, z_k) < 1/4$ in the *scaled* metric of X_k. Hence $d(s \cdot z_k, z_k) < \lambda_k/4$, for large k, in the original metric on X_k. But this contradicts the definition of λ_k.

Remark. The preceding argument actually shows that for every finite set $P \subseteq \Gamma$ which fixes z_∞, given any $\varepsilon > 0$ one has that for k sufficiently large z_k and γz_k are ε-close, in the scaled metric d_k, for every $\gamma \in P$.

We next need to show that segment stabilizers are virtually cyclic. This seems to be the place where some sort of discreteness assumption on Γ is needed. In the classical real-hyperbolic case, Margulis' Lemma implies the result for discrete actions (see [B] and [P2]). Since we are using Cayley graphs and the group actions are (almost) free there is still some sort of discreteness and Paulin gives a delicate argument to show that segment stabilizers are virtually cyclic. The following algebraic lemma is taken from [P4]:

2.7 LEMMA. *Let G be a finitely generated group. If the set of commutators $\{aba^{-1}b^{-1} \mid a, b \in G\}$ is finite, then G is virtually abelian.*

Proof. The action of G on itself by conjugation determines a map $G \to Aut(\Gamma)$, whose image is $Inn(G)$ and whose kernel is the centre of G; it suffices to prove that $Inn(G)$ is finite. If A is a finite generating set for G, then the action of $g \in G$ by conjugation is determined by its action on the elements $a \in A$. But $g^{-1}ag = (g^{-1}aga^{-1})a$, and by hypothesis there are only finitely many possibilities, M say, for the commutator $g^{-1}aga^{-1}$. Hence the cardinality of $Inn(G)$ is at most $M^{|A|}$. $\square$

We proceed with the proof of assertion (2) on segment stabilizers. We call a subgroup *large* if it contains a non-abelian free subgroup (for hyperbolic groups this is equivalent to not having a cyclic subgroup of finite index). Suppose that a large subgroup G of Γ stabilizes a non-trivial segment $e \subseteq X_\infty$. If e is finite, then a subgroup of index 2 in G fixes e pointwise. If e is infinite, a subgroup of index 2 in G acts as translations on a ray in e and thus a large subgroup of G, obtained by taking commutators, fixes a segment of positive length in e pointwise. Thus, in any case, if a large subgroup of Γ stabilizes a segment, then a (perhaps smaller) large subgroup of Γ fixes a segment e of positive length pointwise. Therefore, in order to complete the proof of Paulin's theorem, it suffices to show that if a subgroup of Γ fixes a segment of X_∞ pointwise, then that subgroup is virtually cyclic. Let D denote the length of such a segment which is fixed pointwise by the subgroup $G \subset \Gamma$, and let z and z' denote the endpoints of the segment.

We fix $\varepsilon > 0$ small (to be estimated later) compared to D, and k so large that if $z_k, z'_k \in X_k$ correspond to $z, z' \in X$ then $|d(z_k, z'_k) - D| < \varepsilon$. We fix a geodesic segment $[z_k, z'_k]$ from z_k to z'_k *in* X_k. Given any finite subset $P \subseteq G$, we choose a finite subset $Q \subseteq G$ which contains all products of length $\leqslant 4$ in s, t, s^{-1}, t^{-1}, as s and t vary over P. We choose k large enough so that z_k, z'_k are moved by less than ε by each $\gamma \in Q$ with respect to the scaled metric $d_k = d/\lambda_k$. If $D > 3\varepsilon + (24\delta/\lambda_k)$ then if we omit segments of d-length $\lambda_k\varepsilon + 12\delta$ from the ends of $[z_k, z'_k]$, the remaining sub-segment is non-empty; call this segment C_k. We assume that ε is small enough to satisfy the above inequality; we shall place further restrictions on ε later.

Now we use the original metric d on X_k. From the proof 'slim $\Rightarrow$ thin' (see [Sho] p. 17), if $x \in C_k$ then γx is within 12δ of $[z_k, z'_k]$. We denote by $\gamma_* x$ the projection of γx on $[z_k, z'_k]$. Of course, the 'projection' is not uniquely defined, but the preceding sentence is true no matter which closest point on $[z_k, z'_k]$ one chooses — we fix a definite choice for each $x \in C_k$, thus defining a map $\gamma_* : C_k \to [z_k, z'_k]$ for each γ. Next, we omit segments of length $5(\lambda_k\varepsilon + 12\delta)$ from the ends of $[z_k, z'_k]$ and denote the remaining long segment by $E_k \subseteq C_k$. The map $C_k \to [z_k, z'_k]$ just defined restricts to a map $E_k \to [z_k, z'_k]$; we continue to denote this map by γ_*. Notice that this map is a 24δ-isometry, that is to say, it distorts distances by at most an additive constant of 24δ; in fact it is 24δ close to a translation of E_k along $[z_k, z'_k]$. (Here, and in what follows, the terminology η-*close* is used to describe functions f, g with the same domain such that $d(f(x), g(x)) < \eta$ for all points in their common domain.)

Note that on E_k the maps s_*, s_*t_*, $s_*t_*(s^{-1})_*$, $s_*t_*(s^{-1})_*(t^{-1})_*$ etc. are well-defined and uniformly close to translations. Choose $M = \mathrm{Max}\{5(\lambda_k\varepsilon + 12\delta), 600\delta\}$. We will denote by e_k the segment obtained from $[z_k, z'_k]$ by omitting segments of length M from the ends. We have $e_k \subset E_k$. To make sure that $e_k \neq \emptyset$ we assume $D - \varepsilon > 5\varepsilon + (60\delta/\lambda_k)$, we also assume $D - \varepsilon > (600\delta/\lambda_k)$. Since $\lambda_k \to \infty$, we can choose large enough k and small enough ε so that the above conditions are satisfied.

We shall consider the restrictions $\gamma_* : e_k \to C_k$ to e_k of the maps γ_* defined above; we retain the notation γ_* for these restricted maps. Our goal is to obtain a bound (independent of $|Q|$) on the number commutators $tst^{-1}s^{-1}$ in Q by estimating how close the action of such a commutator on e_k is to the identity map. We first compare $t_*s_*(t^{-1})_*(s^{-1})_*$ to $tst^{-1}s^{-1}$. Observe that, since the maps s and s_* are 12δ close, ts and $t(s_*)$ are 12δ close (the left-action of Γ on X_k is by isometries in the metric d). Hence, $(ts)_*$ and t_*s_* are 36δ close. Comparing successively $tst^{-1}s^{-1}$, $(tst^{-1}s^{-1})_*$, $(ts)_*(t^{-1}s^{-1})_*$, $t_*s_*(t^{-1})_*(s^{-1})_*$ shows that $tst^{-1}s^{-1}$ and $t_*s_*(t^{-1})_*(s^{-1})_*$ are $(12 + 36 + 108)\delta$ close.

Next, we compare $t_* s_* (t^{-1})_* (s^{-1})_*$ to the identity map on e_k. Since $s_* (t^{-1})_*$ and $(t^{-1})_* s_*$ are 72δ close to the same translation, and translations commute, we have that $t_* s_* (t^{-1})_* (s^{-1})_*$ and $t_* (t^{-1})_* s_* (s^{-1})_*$ are $(144 + 24)\delta$ close. Moreover, $t_* (t^{-1})_*$ and $s_* (s^{-1})_*$ are 36δ close to the identity. Thus $t_* (t^{-1})_* s_* (s^{-1})_*$ is 108δ close to the identity. Hence $t_* s_* (t^{-1})_* (s^{-1})_*$ is 276δ close to the identity. Combining this with the estimate in the previous paragraph we have that the restriction of $(tst^{-1}s^{-1})$ to e_k is 532δ close to the identity on e_k. Therefore, a vertex close to the midpoint of e_k is moved by less than $532\delta + 2$ by $tst^{-1}s^{-1}$. Thus $tst^{-1}s^{-1}$ lies in the ball of radius $532\delta + 2$ about the identity in Γ, and we have the desired bound on the number of commutators in the arbitrary finite subset $P \subset G$.

Now Lemma 2.7 implies that G is virtually abelian. But every abelian subgroup of a hyperbolic group is virtually cyclic. Hence the segment stabilizers for the action of Γ on X_∞ must be virtually cyclic. This completes the proof of Paulin's theorem. $\square$

SECTION 3: CONVEX HULLS

A subset Σ of a geodesic metric space X is said to be *geodesically convex* if for all $p, q \in \Sigma$ every geodesic segment from p to q is completely contained in Σ. Given a bounded set $Y \subset X$, perhaps the most natural way to define its convex hull is as the intersection of all geodesically convex sets containing Y.

If X is simply connected and non-positively curved then round balls are geodesically convex and hence the convex hull of a bounded set is bounded. However, for more general geodesic metric spaces, even δ-hyperbolic spaces, it may happen that the convex hull of a finite set is the whole of the ambient space X. The following example illustrates how general this problem is.

3.1 PROPOSITION. *Given any finitely generated group Γ there exists a finite generating set S and a finite subset $Y \subset \Gamma$ such that the convex hull of Y in the Cayley graph $X(\Gamma, S)$ is the whole of $X(\Gamma, S)$.*

Proof. Let A be any finite generating set for Γ, and take S to be the set of those elements of Γ which are a distance 1 or 4 from the identity in the Cayley graph of Γ with respect to S. Let Y be the set of elements of Γ which are a distance at most 3 away from the identity in the Cayley graph associated to A.