ADDENDUM TO "RANDOM WALK ON RANDOM GROUPS" BY M. GROMOV

LIOR SILBERMAN

ABSTRACT. Written at the request of GAFA's Editorial Board, with the goal of explicating some aspects of M. Gromov's paper [2]. Section 1 recalls the construction of the random group, while section 2 contains a proof that the random group has property (T) based on the ideas from the preprint. The appendix defines CAT(0) spaces and works out in detail some geometric propositions from the preprint used in the proof.

1. RANDOM GROUPS

In this application of the probabilistic method, a "random group" will be a quotient of a given group. We fix the set of generators (and possibly some initial relations) and pick extra relations "at random" to get the group:

Let Γ be a finitely generated group, specifically generated by the finite symmetric set S (of size $2k$). Let $G = (V, E)$ be a (locally finite) *undirected* graph. $\vec{E}$ will denote the (multi)set of *oriented edges* of G, i.e. $\vec{E} = \{(u, v), (v, u) \mid \{u, v\} \in E\}$. Given a map ("$S$-coloring") $\alpha : \vec{E} \to S$ and an oriented path $\vec{p} = (\vec{e}_1, \ldots, \vec{e}_r)$ in G, we set $\alpha(\vec{p}) = \alpha(\vec{e}_1) \cdot \ldots \cdot \alpha(\vec{e}_r)$. We will only consider the case of *symmetric* α (i.e. $\alpha(u, v) = \alpha(v, u)^{-1}$ for all $(u, v) \in \vec{E}$). Then connected components of the labelled graph (G, α) look like pieces of the Cayley graph of a group generated by S. Such a group can result from "patching together" labelled copies of G, starting from the observation that in a Cayley graph the cycles correspond precisely to the relations defining the group. Following this idea let $R_\alpha = \{\alpha(\vec{c}) \mid \vec{c} \text{ a cycle in } G\}$, $W_\alpha = \langle R_\alpha \rangle^N$ (normal closure in Γ) and

$$\Gamma_\alpha = \Gamma / W_\alpha.$$

The Γ_α will be our random groups, with $\alpha(\vec{e})$ chosen independently uniformly at random from S, subject to the symmetry condition. Properties of Γ_α then become random variables, which can be studied using the techniques of the probabilistic method (e.g. concentration of measure). We prove here that subject to certain conditions on G (depending on k), the groups Γ_α furnish examples of Kazhdan groups with high probability as $|V| \to \infty$.

We remark that Γ_α is presented by the k generators subject to the relations corresponding to the labelled cycles in the graph, together with the relations already present in Γ (unless Γ is a free group). In particular if G is finite and Γ is finitely presented then so is Γ_α.

Remark 1.1. This can be done in greater generality: Let Γ be any group, let ρ be a symmetric probability measure on Γ (i.e. $\rho(A) = \rho(A^{-1})$ for any measurable $A \subset \Gamma$), and denote by ρ^E the (probability) product measure on the space $\mathcal{A}$ of symmetric maps $\vec{E} \to \Gamma$.

Department of Mathematics, Princeton University, Princeton, NJ 08544, USA
homepage: `http://www.math.princeton.edu/~lior`.

This is well defined since ρ is symmetric and justifies the power notation. Now proceed as before. The case considered in this paper is that of the *standard measure*, assigning probabilities $\frac{1}{2k}$ to the generators and their reciprocals.

Remark 1.2. Assume that $\Gamma = \langle S|R\rangle$ is a presentation of Γ w.r.t. S. Then $\Gamma_\alpha = \langle S|R \cup R_\alpha\rangle$ is a quotient of $\langle S|R_\alpha\rangle$. Since a quotient of a (T)-group is also a (T)-group, it suffices to prove that the latter groups has property (T) with high probability. In other words, we can assume for the purposes of the next section that $\Gamma = F_k$ is the free group on k generators ($k \geq 2$).

2. Geometry, Random Walks on Trees and Property (T).

2.1. **Overview.** In this section we prove that with high probability Γ_α has property (T). The idea is to start with a vector in a representation, and consider the average of its translates by the generators. Typically iterating the averaging produces a sequence rapidly converging to a fixed point. The proof of this breaks down in the following parts:

(1) Property (T) can be understood in geometric language by examining random walks on the group Γ_α.
(2) A general analysis of random walks on trees gives some technical results.
(3) The spectral gap of G can be expressed as a bound on the long-range variation of functions on G in terms of their short-range variation.
(4) ("Transfer of the spectral gap"): With high probability (w.r.t the random choice of α), a similar bound holds on the variation of certain equivariant functions on Γ (these are Γ-translates of vectors in a representation of Γ_α).
(5) By a geometric inequality and an estimate on the random walk on the tree Γ, averaging such a function over the action of the the generators n times produces a function whose (short-range) variation is bounded by the long-range variation of the original function.
(6) Combining (3), (4) and (5) shows that repeated averaging over the action of the generators converges to a fixed point. The rate of convergence gives a Kazhdan constant.

2.2. **Property (T).** Let Γ be a locally compact group generated by the compact subset S (not to be confused with the Γ of the previous section).

Definition 2.1. Let $\pi : \Gamma \to \mathrm{Isom}(Y)$ be an isometric action of Γ on the metric space Y. For $y \in Y$ set $\mathrm{di}_S(y) = \sup_{\gamma \in S} d_Y(\pi(\gamma)y, y)$.

Say that $y \in Y$ is *ε-almost-fixed* if $\mathrm{di}_S(y) \leq \varepsilon$. Say that $\{y_n\}_{n=1}^{\infty} \subseteq Y$ *represents an almost-fixed-point* (a.f.p.) if $\lim_{n\to\infty} \mathrm{di}_S(y_n) = 0$.

Definition 2.2. (Kazhdan, [3]) Say that Γ has *property (T)* if there exists $\varepsilon > 0$ such that every unitary representation of Γ which has ε-almost fixed unit vectors is the trivial representation. Such an ε is called a *Kazhdan constant* for Γ w.r.t S. The largest such ε is called *the* Kazhdan constant of Γ w.r.t. S.

Remark 2.3. It is easy to see that the question of whether Γ has property (T) is independent of the choice of S. Different choices may result in different constants though.

An alternative definition considers "affine" representations. For the purpose of most of the discussion, the choice of origin in the representation space is immaterial, as we consider an action of the group through the entire isometry group of the Hilbert space, rather than through the unitary subgroup fixing a particular point (informally, we allow action by translations in addition to rotations). In this case we will say the Hilbert space is “affine”. One can always declare an arbitrary point of the space to be the “origin”, letting the norm denote distances from that point and vector addition and scalar multiplication work “around” that point. However, the results will always be independent of such a choice.

Theorem 2.4. *(Guichardet-Delorme, see* [1, Ch. 4]*)* Γ *has property (T) iff every affine (=isometric) action of* Γ *on a Hilbert space* Y *has a fixed point.*

We now return to the group Γ of the previous section and introduce the geometric language used in the remainder of the discussion. As explained above we specialize to the case of Γ being a free group. Let Y be a metric space, $\pi : \Gamma_\alpha \to \mathrm{Isom}(Y)$. Since Γ_α is a quotient of Γ we can think of π as a representation of Γ as well. Setting $X = \mathrm{Cay}(\Gamma, S)$ (a $2k$-regular tree) allows us to separate the geometric object X from the group Γ acting on it (by the usual Cayley left-regular action). We can now identify Y with the space[1] $B^\Gamma(X,Y)$ of Γ-equivariant functions $f : X \to Y$ (e.g. by taking the value of f at 1).

We are interested in bounding the distances $d_Y(sy, y)$ for $s \in S$. More precisely we will bound

$$\frac{1}{2|S|}\sum_{\gamma\in S} d_Y^2(\gamma y, y).$$

Under the identification of Y with $B^\Gamma(X,Y)$ this is:

$$\sum_{x'\in X} \mu_X(x \to x') d_Y^2(f(x'), f(x)),$$

where μ_X is the standard random walk on X (i.e. $\mu_X(x \to x') = \frac{1}{2k}$ if $x' = x\gamma$ for some generator $\gamma \in S$ and $\mu_X(x \to x') = 0$ otherwise). We note that since f and μ_X are Γ-equivariant, this “energy” is independent of x, and we can set:

$$E_{\mu_X}(f) = \frac{1}{2}\sum_{x'} \mu(x \to x') d_Y^2(f(x'), f(x))$$

and call it the μ_X*-energy* of f. To conform with the notation of section B.4 in the appendix, one should formally integrate w.r.t. a measure $\bar\nu$ on $\Gamma\backslash X$, but that space is trivial. In this language f is constant (i.e. is a fixed point) iff $E_{\mu_X}(f) = 0$ and $\{f_n\}_{n=1}^\infty$ represent an a.f.p. iff $\lim_{n\to\infty} E_{\mu_X}(f_n) = 0$.

In much the same way we can also consider longer-range variations in f, using the n-step random walk instead. $\mu_X^n(x \to x')$ will denote the probability of going from x to x' in n steps of the standard random walk on X, $E_{\mu_X^n}(f)$ the respective energy. Secondly, we can apply the same notion of energy to functions on a graph G as well (no equivariance here: we consider all functions $f : V \to Y$). Here μ_G^n will denote the usual random walk on the graph, ν_G will be the standard measure on G ($\nu_G(u) = \frac{1}{2|E|}\deg u$) and $E_{\mu_G^n}(f) = \frac{1}{2}\sum_{u,v} \nu_G(u)\mu_G^n(v) d_Y^2(f(u), f(v))$. The “spectral gap” property of G can then be written as the inequality (Lemma 2.10, and note that the RHS does not depend on n!)

$$E_{\mu_G^n}(f) \le \frac{1}{1-\lambda^1(G)} E_{\mu_G}(f),$$

where $\lambda^r(G) = \max\{|\lambda_i|^r \mid \lambda_i \text{ is an e.v. of } G \text{ and } \lambda_i^r \neq 1\}$.

The core of the proof, section 2.4, carries this over to X with a worse constant. There is one caveat: we prove that with high probability, for every equivariant f coming from a

[1] For the motivation for this notation see appendix B.

representation of Γ_α, the inequality $E_{\mu_X^{2l}}(f) \le \frac{10.5}{1-\lambda^2(G)} E_{\mu_X^2}(f)$ holds for *some* value of l, large enough. We no longer claim this holds for *every* l.

Leveraging this bound to produce an a.f.p. is straightforward: we use the iterated averages $H_{\mu_X}^l f(x)$, which are simply

$$(H_{\mu_X})^l f(x) = H_{\mu_X^l} f(x) = \sum_{x'} \mu_X^l(x \to x') f(x').$$

Geometric considerations show that if l is large enough then (approximately, see Proposition 2.15 for the accurate result) $E_{\mu_X}(H_{\mu_X^l} f) \ll E_{\mu_X^l}(f)$, and together with the spectral gap property this shows that continued averaging gives an a.f.p. which converges to a fixed point. Moreover, if the representation is unitary (i.e. the action of Γ on Y fixed a point $0 \in Y$), if f represents a unit vector (i.e. the values of f are unit vectors) and if $E_{\mu_X}(f)$ is small enough to start with, then this fixed point will be nonzero. This gives an explicit Kazhdan constant (for details see section 2.6).

One technical problem complicates matters: the tree X is a bipartite graph. It is thus more convenient to consider the random walk μ_X^2 and its powers instead, since they are all supported on the same half of the tree. Then the above considerations actually produce a Γ^2-f.p. where Γ^2 is the subgroup of Γ of index 2 consisting of the words of even length. If W_α (i.e. G) contains a cycle of odd order then $\Gamma_\alpha^2 = \Gamma_\alpha$ and we're done. Otherwise $[\Gamma_\alpha : \Gamma_\alpha^2] = 2$ and $\Gamma_\alpha^2 \lhd \Gamma_\alpha$. Now averaging w.r.t $\Gamma_\alpha/\Gamma_\alpha^2$ produces a Γ_α-f.p. out of a Γ_α^2-f.p.

2.3. **Random walks on trees.** Let T_d be a d-regular tree, rooted at $x_0 \in T_d$. Consider the distance from x_0 of the random walk on T_d starting at x_0. At each step this distance increases by 1 with probability $p_d = \frac{d-1}{d}$ and decreases by 1 with probability $q_d = \frac{1}{d}$, except if the walk happens to be at x_0 when it must go away. Except for this small edge effect, the distance of the walk from x_0 looks very much like a binomial random variable.

Formally let $\Omega_1 = \{+1, -1\}$ with measure p_d on $+1$, q_d on -1 and let $\Omega = \Omega_1^{\mathbb{N}}$ with product measure. We define two sequences of random variables on Ω: the usual Bernoulli walk:

$$X_n(\omega) = \sum_{i=1}^n \omega_i$$

as well as the "distance from x_0" one by setting $Y_0(\omega) = 0$ and:

$$Y_{n+1}(\omega) = \begin{cases} Y_n(\omega) + \omega_{n+1} & Y_n(\omega) \ge 1 \\ 1 & Y_n(\omega) = 0 \end{cases}.$$

Let $\mu_d^n(r) = \mathcal{P}(Y_n = r)$ and $b_d^n(r) = \mathcal{P}(X_n = r)$. If μ_T is the standard random walk on the tree then by spherical symmetry $\mu_d^n(r) = \delta_d(r)\mu_T^n(x_0 \to y_0)$ where $\delta_d(r)$ is the size of the r-sphere in T and $d_T(x_0, y_0) = r$. In similar fashion $b_d^n(r)$ is the probability that the skewed random walk on $\mathbb{Z}$ with probabilities p_d, q_d goes from 0 to r in n steps. We also add $\eta_d = p_d - q_d$ and $\sigma_d^2 = 4p_d q_d$ (the expectation and variance of ω_n). Of course $\mu_d^n(r) = b_d^n(r) = 0$ if $r \not\equiv n \ (2)$ and otherwise

$$b_d^n(r) = \binom{n}{\frac{n+r}{2}} p_d^{\frac{n+r}{2}} q_d^{\frac{n-r}{2}}.$$

We drop the subscript $'d'$ for the remainder of the section. The following Proposition collects some facts about the Bernoulli walk:

Proposition 2.5. $\mathbb{E}(X_n) = n\eta$, $\sigma^2(X_n) = n\sigma^2$ *and:*

(1) *(Large deviation inequality, e.g. see* [4]*)*

$$\mathcal{P}\left(X_n > \eta n + \varepsilon n\right), \mathcal{P}\left(X_n < \eta n - \varepsilon n\right) \leq e^{-2n\varepsilon^2}$$

so that

$$\mathcal{P}\left(|X_n - \eta n| > \varepsilon n\right) \leq 2e^{-2n\varepsilon^2}.$$

(2) *(Non-recurrence) Let* $r \geq 0$ *and assume* $p > q$*. Then*

$$\mathcal{P}\left(\forall_n X_n > -r\right) = 1 - \left(\frac{q}{p}\right)^r.$$

The trivial observation that $X_n(\omega) \leq Y_n(\omega)$ implies $\mathcal{P}(Y_n \leq r) \leq \mathcal{P}(X_n \leq r)$ leading to the (one-sided) deviation estimate

$$\mathcal{P}\left(Y_n \leq r\right) \leq e^{-2n(\eta - \frac{r}{n})^2}. \tag{2.1}$$

In the other direction the recurrence relation

$$\mu^{n+1}(r) = \begin{cases} p \cdot \mu^n(r-1) + q \cdot \mu^n(r+1) & r \geq 2 \\ \mu^n(0) + q \cdot \mu^n(2) & r = 1 \\ q \cdot \mu^n(1) & r = 0 \end{cases}$$

implies $\mu^n(r) \leq \frac{1}{p} b^n(r)$, with equality for $r = n$ (proof by induction, also carrying the stronger assertion $\mu^n(0) \leq b^n(0)$).

Corollary 2.6. $\mathcal{P}\left(|Y_n - \eta n| > \theta\sqrt{n \log n}\right) \leq \frac{2}{p} n^{-2\theta}$.

This is a crucial point, since this will allow us to analyze expressions like $\sum_{x'} \mu_X^n(x, x') f(x')$ only when $d_X(x, x') \sim \eta n$, making a trivial estimate otherwise. In fact, from now on we will only consider the range $|r - \eta n| \leq 2\sqrt{n \log n}$.

Lemma 2.7. *(Reduction to Bernoulli walks) Let* $t \leq n$ *. Then*

$$\left|\mathcal{P}(Y_n = r) - \sum_{\frac{1}{2}\eta t \leq j \leq t} \mathcal{P}(Y_t = j)\mathcal{P}(X_{n-t} = r - j)\right| \leq e^{-\eta^2 t/2} + \left(\frac{q}{p}\right)^{\frac{1}{2}\eta t}.$$

Proof. As has been remarked before, $\mathcal{P}(Y_t < \frac{1}{2}\eta t) \leq \mathcal{P}(X_t < \frac{1}{2}\eta t) \leq e^{-\eta^2 t/2}$. Furthermore, defining $\tilde{X}_n = Y_t + X_n - X_t$ we clearly have:

$$\left\{\omega \in \Omega \mid Y_t(\omega) \geq j_0 \wedge Y_n \neq \tilde{X}_n\right\} \subseteq \left\{\omega \in \Omega \mid Y_t(\omega) \geq j_0 \wedge \exists u > t : \tilde{X}_u = 0\right\}.$$

However, by Proposition 2.5(2) $\mathcal{P}(Y_t \geq j_0 \wedge \exists u > t : \tilde{X}_u = 0) \leq \left(\frac{q}{p}\right)^{j_0}$. Also by the time translation-invariance of the usual random talk,

$$\mathcal{P}(Y_t = j \mid \tilde{X}_n = r) = \mathcal{P}(X_n - X_t = r) = \mathcal{P}(X_{n-t} = r - j).$$

□

Proposition 2.8. *Let* $|r - \eta n| \leq 2\sqrt{n \log n}$*. Then for some constants* $c_1(d), c_2(d)$ *independent of* n*,*

$$\left|\mu_d^{n+2}(r) - \mu_d^n(r)\right| \leq \frac{\sqrt{\log n}}{\sqrt{n}} c_1(d) \cdot \mu_d^n(r) + c_2(d)\left(e^{-\eta^2\sqrt{n}/2} + \left(\frac{q}{p}\right)^{\frac{1}{2}\eta\sqrt{n}}\right)$$

Proof. Set $t = \lfloor \sqrt{n} \rfloor$. By the Lemma,

$$|\mathcal{P}(Y_{n+2} = r) - \mathcal{P}(Y_n = r)| \leq 2e^{-\eta^2 t/2} + 2\left(\frac{q}{p}\right)^{\frac{1}{2}\eta t} +$$
$$+ \sum_{\frac{1}{2}\eta t \leq j \leq t} \mathcal{P}(Y_t = j) \left|\mathcal{P}(X_{n-t+2} = r - j) - \mathcal{P}(X_{n-t} = r - j)\right|.$$

An explicit computation shows:

$$|\mathcal{P}(X_{n-t+2} = r - j) - \mathcal{P}(X_{n-t} = r - j)| \leq c_1(d) \frac{\sqrt{\log n}}{\sqrt{n}} \mathcal{P}(X_{n-t} = r - j).$$

Since $\sqrt{\log n/n} \leq 1$ the result follows with $c_2(d) = (2 + c_1(d))(e^{\eta^2/2} + (p/q)^{\eta/2})$. $\square$

We also need a result about general trees. Let T be a tree rooted at $x_0 \in T$ such that the degree of any vertex in T is at least 3. We would like to prove that the random walk on T tends to go further away from x_0 than the random walk on T_3 due to the higher branching rate.

Proposition 2.9. *Let $r < \frac{1}{3}n$. Then $\mu_T^n(x_0 \to B_T(x_0, r)) \leq e^{-2n(\frac{1}{3} - \frac{r}{n})^2}$.*

Proof. Label the (directed) edges exiting each vertex $x \in T$ with the integers $0, \ldots, \deg(x) - 1$ such that the edge leading toward x_0 is labelled 0 (any labelling if $x = x_0$). Assume that the maximal degree occurring in $B_T(x_0, n)$ is d, and let $\Omega = \{0, \ldots, d! - 1\}$ with uniform measure. Again we define two random variables on Ω^n. First, let $X_i(\omega) = \sum_{j=1}^{i} X(\omega_j)$ where $X(\omega_j) = -1$ if $\frac{3}{d!}\omega_j < 1$, $X(\omega_j) = +1$ if not. Secondly define $Y_i : \Omega^n \to T$ by $Y_0 \equiv x_0$ and $Y_i(\omega)$ follows the edge labelled $\left\lfloor \frac{\deg(Y_{i-1})}{d!}\omega_i \right\rfloor$ leaving $Y_{i-1}(\omega)$.

It is easy to see that the Y_i give a model for the standard random walk on T, while the X_i are a Bernoulli walk on $\mathbb{Z}$ with probability p_3 of going to the right. Moreover, by induction it is clear that $d_T(x_0, T_i(\omega)) \geq X_i(\omega)$. The deviation inequality 2.5 now concludes the proof. $\square$

2.4. **The spectral gap and its transfer.**

Lemma 2.10. *Let $G = (V, E)$ be a finite graph, let ν_G be the standard probability measure on V ($\nu_G(u) = \frac{\deg(u)}{2|E|}$), and let μ_G be the standard random walk on G ($\mu_G(u \to v) = \frac{1}{\deg(u)}$ if $(u, v) \in \vec{E}$). Set $\lambda^2(G) = \max\left\{\lambda^2 \mid \lambda \neq \pm 1 \text{ is an eigenvalue of } H_{\mu_G}\right\}$ (note that $\lambda^2(G) < 1$). Let Y be an affine Hilbert space, and let $f : V \to Y$. Then $E_{\mu_G^{2n}}(f) \leq \frac{1}{1 - \lambda^2(G)} E_{\mu_G^2}(f)$ for all $n \in \mathbb{N}$.*

Proof. Choose an arbitrary origin in Y, making it into a Hilbert space. Then

$$E_{\mu_G^q}(f) = \frac{1}{2} \sum_{u \in V} \nu_G(u) \sum_{v \in V} \mu_G^q(u \to v) \|f(u) - f(v)\|_Y^2$$
$$= \frac{1}{2} \sum_{u, v \in V} \nu_G(u) \mu_G^q(u \to v) \left[\|f(u)\|_Y^2 + \|f(v)\|_Y^2 - 2\langle f(u), f(v) \rangle_Y\right].$$

μ is ν-symmetric, i.e. $\nu(x)\mu^q(x \to y) = \nu(y)\mu^q(y \to x)$ and therefore
(2.2)
$$E_{\mu_G^q}(f) = \sum_{u \in V} \nu_G(u) \left\langle f(u), f(/u) - \sum_{v} \mu_G^q(u \to v) f(v) \right\rangle = \left\langle f, (I - H_{\mu_G}^q) f \right\rangle_{L^2(\nu_G)},$$

where we used $H_{\mu_G^q} = (H_{\mu_G})^q$. Just like the case $Y = \mathbb{C}$ we can now decompose (the vector-valued) f in terms of the (scalar!) eigenfunctions of H_{μ_G} (the normalized adjacency matrix). Assume these are $\psi_0, \ldots, \psi_{n-1}$ corresponding to the eigenvalues $\lambda_0 = 1 \geq \lambda_1 \geq \cdots \geq \lambda_{|V|-1} \geq -1$, such that λ_{i_0} is the first smaller than 1 and λ_{i_1} is the last larger than -1. Writing $f = \sum a_i \psi_i$ ($a_i \in Y$) and specializing to the case $q = 2n$ we get:

$$E_{\mu_G^{2n}}(f) = \sum_{i=i_0}^{i_1} (1 - \lambda_i^{2n}) |a_i|^2 .$$

For $i_0 \leq i \leq i_1$,

$$1 - \lambda_i^{2t} = (1 - \lambda_i^2) \sum_{j=0}^{t-1} \lambda_i^{2j} \leq (1 - \lambda_i^2) \sum_{j=0}^{\infty} \lambda(G)^{2j} = \frac{1 - \lambda_i^2}{1 - \lambda^2(G)},$$

and thus

$$E_{\mu_G^{2n}}(f) = \sum_{i=i_0}^{i_1} (1 - \lambda_i^{2n}) |a_i|^2 \leq \frac{1}{1 - \lambda^2(G)} \sum_{i=i_0}^{i_1} (1 - \lambda_i^2) |a_i|^2 = \frac{1}{1 - \lambda^2(G)} E_{\mu_G^2}(f).$$

□

Now recall the definition of the factor group Γ_α in Section 1. We would like to transfer the "boundedness of energy" property to the random walk on the random group Γ_α. Let $X_\alpha = \mathrm{Cay}(\Gamma_\alpha; S)$ with the standard S-labelling and Γ-action. Given a vertex $u \in V$ and an element $x \in \Gamma_\alpha$ there exists a unique homomorphism of labelled graphs,

$$\alpha_{u \to x} : G_u \to X_\alpha$$

(G_u is the component of G containing u) which takes u to x and maps a directed edge of $\vec{e} \in \vec{E}$ to an edge $(y, y \cdot \alpha(\vec{e}))$ of the Cayley graph $X_\alpha = \mathrm{Cay}(\Gamma_\alpha; S)$. Given the measure $\mu_G(u \to)$ on V_u we thus have its pushforward measure $\alpha^*_{u \to x}(\mu_G(u \to))$ on Γ_α. We can then define a pushforward random walk on Γ_α by averaging over all choices of $u \in V$:

$$\bar{\mu}_{X,\alpha}(x \to A) = \sum_u \nu(u) \alpha^*_{u \to x}(\mu_G(u \to A)) = \sum_u \nu(u) \mu_G(\alpha^{-1}_{u \to x}(A)).$$

It is clear that $\bar{\mu}_{X,\alpha}$ is Γ-invariant. The $\bar{\mu}_{X,\alpha}$-odds of going from x to y are the average over u of the odds of going from a vertex u to a vertex v such that the path from u to v is α-mapped to $x^{-1}y$.

We now fix $u_0 \in V, x_0 \in \Gamma_\alpha$ (e.g. $x_0 = 1$). Let Γ_α act by isometries on the metric space Y. Then we can identify an element of Y with a Γ_α-equivariant map $f : X_\alpha \to Y$ through the value $f(x_0)$. Any such function can be clearly pulled back to a function $f \circ \alpha_{u_0 \to x_0} : V \to Y$. Moreover,

$$E_{\mu_G^q}(f \circ \alpha_{u_0 \to x_0}) = \frac{1}{2} \sum_{u,v \in V} \nu_G(u) \mu_G^q(u \to v) d_Y^2 \left(f(\alpha_{u_0 \to x_0}(u)), f(\alpha_{u_0 \to x_0}(v)) \right)$$

and by the equivariance of f this equals:

$$\frac{1}{2} \sum_{x \in X} d_Y^2(f(e), f(x)) \cdot \sum_{\alpha_{u \to e}(v) = x} \nu_G(u) \mu_G^q(u \to v) = \frac{1}{2} |df|^2_{\tilde{\mu}^q_{X,\alpha}}(e) = E_{\tilde{\mu}^q_{X,\alpha}}(f),$$

pushing forward any walk μ_G^q to a random walk $\tilde{\mu}^q_{X,\alpha}$ on X_α. We can now rewrite the Lemma as $E_{\tilde{\mu}^{2n}_{X,\alpha}}(f) \leq \frac{1}{1 - \lambda^2(G)} E_{\tilde{\mu}^2_{X,\alpha}}(f)$.

The space X_α, however, is difficult to analyze – in particular it varies with α. We would rather consider the tree $X = \mathrm{Cay}(\Gamma; S)$, which we also take with the S-labelling and Γ-action. Fixing $x \in X$, every path in G has a unique α-pushforward to an path on X starting at x and preserving the labelling. By averaging over all paths of length q in G we get a Γ-invariant random walk on X, which we will denote by $\bar{\mu}^q_{X,\alpha}$:

$$\bar{\mu}^q_{X,\alpha}(x \to x') = \sum_{|\vec{p}|=q; \alpha_{p_0 \to x}(\vec{p})=x'} \nu_G(p_0)\mu^q_G(\vec{p})$$

(this notation is acceptable since the pushforward of this walk by the quotient map $X \to X_\alpha$ will give the walk $\tilde{\mu}^q_{X,\alpha}$on X_α). The relevant space of functions is now all the Γ-equivariant functions $f : X \to Y$ such that $wf = f$ for all $w \in W_\alpha$. It is clear that averaging such a function f w.r.t. this walk on X or on X_α will give the same answer, allowing us to only consider walks on the regular tree X. The final form of the Lemma is then

$$E_{\bar{\mu}^{2n}_{X,\alpha}}(f) \leq \frac{1}{1-\lambda^2(G)} E_{\bar{\mu}^2_{X,\alpha}}(f). \tag{2.3}$$

We now use the results of section 2.3 to obtain (with high probability) a similar inequality about the variation of functions w.r.t. the standard walk μ^q_X:

For fixed q, x and x', we think of the transition probability $\bar{\mu}^q_{X,\alpha}(x \to x')$ as a function of α, in other words a random variable. It expectation will be denoted by $\bar{\mu}^q_{X,G}(x \to x') \stackrel{\text{def}}{=} \mathbb{E}\bar{\mu}^q_{X,\alpha}(x \to x')$. We show that $\bar{\mu}^{2n}_{X,G}(x \to x')$ is a weighted sum of the $\mu^q_X(x \to x')$, where small values of q give small contributions. Thus any bound on $E_{\bar{\mu}^{2n}_{X,G}}(f)$ can be used to bound some $E_{\mu^{2l}_X}(f)$. We then show that with high probability $\bar{\mu}^q_{X,\alpha}(x \to x')$ is close to its expectation, so that equation (2.3) essentially applies to $E_{\bar{\mu}^{2n}_{X,G}}(f)$ as well.

We recall that the *girth* of a graph G, denoted $g(G)$, is the length of the shortest non-trivial closed cycle in G. If $q < \frac{1}{2}g(G)$ then any ball of radius q in G is a tree. We also denote the *minimal vertex degree* of G by $\delta(G) = \min\{\deg(v) \mid v \in V\}$.

Lemma 2.11. *Let $2n < \frac{1}{2}g(G)$. Then there exist nonnegative weights $P^{2n}_G(2l)$ such that $\sum_{l=0}^n P^{2n}_G(2l) = 1$, and*

$$\bar{\mu}^{2n}_{X,G}(x \to x') \stackrel{\text{def}}{=} \mathbb{E}\bar{\mu}^{2n}_{X,\alpha}(x \to x') = \sum_{l=0}^{n} P^{2n}_G(2l)\mu^{2l}_X(x \to x').$$

Moreover if $\delta(G) \geq 3$ then

$$Q^{2n}_G \stackrel{\text{def}}{=} \sum_{l \leq n/6} P^{2n}_G(2l) \leq e^{-n/9}.$$

Proof. Since

$$\bar{\mu}^{2n}_{X,\alpha}(x \to x') \stackrel{\text{def}}{=} \sum_{|\vec{p}|=2n; \alpha_{p_0 \to x}(\vec{p})=x'} \nu_G(p_0)\mu^{2n}_G(\vec{p}),$$

the expectation is

$$\bar{\mu}^{2n}_{X,G}(x \to x') = \sum_{|\vec{p}|=2n} \nu_G(p_0)\mu^{2n}_G(\vec{p})\mathcal{P}(\alpha_{p_0 \to x}(p_{2n}) = x').$$

Where the probability is w.r.t. the choice of α.

Let $\vec{p}$ be a path of length $2n$ in G. By the girth assumption, the ball of radius $2n$ around $p(0)$ is a tree. Thus there is a unique shortest path $\vec{p}'$ in it from p_0 to p_{2n}, and

$$\alpha_{p_0\to x}(\vec{p}) = \alpha_{p_0\to x}(\vec{p}') = x\cdot\alpha(\vec{e}'_1)\cdot\alpha(\vec{e}'_2)\cdot\cdots\cdot\alpha(\vec{e}'_{|\vec{p}'|})$$

(just remove the backtracks by the symmetry of α). Moreover, the $\alpha(\vec{e}'_i)$ are independent random variables since those edges are distinct. We conclude that the probability that $\alpha_{p_0\to x}(\vec{p}) = x'$ is equal to the probability of the $|\vec{p}'|$-step random walk on X getting from x to x'. Thus:

$$\bar{\mu}^{2n}_{X,G}(x\to x') = \sum_{l=0}^{n} P^{2n}_G(2l)\mu^{2l}_X(x\to x')$$

where

$$P^{2n}_G(2l) = \sum_{|\vec{p}|=2n;|\vec{p}'|=2l} \nu_G(p_0)\mu^{2n}_G(\vec{p}).$$

For any $u\in V$ Let $T_u = B_G(u,2n)$. This is a tree, and if $P^{2n}_{G,u}(2l)$ is the probability of the $2n$-step random walk on T_u starting at u reaching the distance $2l$ from u then clearly $P^{2n}_G(2l) = \sum_u \nu_G(u)P^{2n}_{G,u}(2l)$. In particular it is clear that $\sum_l P^{2n}_G(2l) = 1$. Moreover since we assume that the minimal degree in G is 3, we have $\sum_{l\le n/6} P^{2n}_{G,u} \le e^{-2n/18}$ by Proposition 2.9. Averaging over u gives the bound on Q^{2n}_G. □

Lemma 2.12. *In addition to the assumptions of the previous Lemma, let* $\deg(u)\le d$ *for all* $u\in V$*, where* d *is independent of* $|V|$*. Then with probability tending exponentially to* 1 *with* $|V|$*,*

$$\bar{\mu}^{2n}_{X,\alpha}(x\to x') \ge \frac{1}{2}\bar{\mu}^{2n}_{X,G}(x\to x')$$

for all $x,x'\in X$ *and*

$$\bar{\mu}^{2}_{X,\alpha}(x\to x') \le \mu^{2}_X(x\to x')$$

for all $x\ne x'\in X$.

Proof. The random variable $\bar{\mu}^{q}_{X,\alpha}(x\to x')$ is a Lipschitz function on a product measure space: return to the definition of $\bar{\mu}^{q}_{X,\alpha}(x\to x')$ as the average of the pushforwards of random walks centered at the various vertices of G. Changing the value of α on one edge only affects random walks starting at vertices with distance at most $q-1$ from the endpoints of the edge. Since each such contribution can change by at most 1 the average can change by at most

$$\tau \le \frac{2(d-1)^{q-1}}{|V|},$$

which is therefore the Lipschitz constant[2]. By the concentration of measure inequality (see [4, Corollary 1.17]) the probability of $\bar{\mu}^{q}_{X,\alpha}(x\to x')$ deviating from its mean by at least ε (in one direction) is at most

$$e^{-\frac{\varepsilon^2}{2\tau^2|E|}} \le e^{-\frac{\varepsilon^2}{4d(d-1)^{2q-2}}|V|}$$

since $2|E|\le d|V|$.

Fixing q, x we now consider all the $\bar{\mu}^{q}_{X,\alpha}(x\to x')$ (different x') together. Let $N_k(q)$ be the number of random variables:

$$N_k(q) = \left|\{x'\in X \mid d_X(x,x')\le q \,\text{and}\, d_X(x,x')\equiv q \pmod 2\}\right|,$$

[2] w.r.t. the Hamming metric on the product space.

and let $\varepsilon(k,d,q)$ be the infimum over such x' (essentially over their distances – the walk $\bar{\mu}^q_{X,G}$ is spherically symmetric) and over rooted trees T of depth q and degrees bounded between 3 and d of the expression[3]

$$\bar{\mu}^q_{X,T}(x \to x') = \sum_{l \le q} \sum_{|\vec{p}|=q; |\vec{p}'|=l} \mu^q_T(\vec{p}) \mu^l_X(x \to x'),$$

the probability of *some* $\bar{\mu}^q_{X,\alpha}(x \to x')$ being less than half its mean is then at most

$$N_k(2n) \cdot e^{-\frac{\varepsilon(k,d,2n)^2}{16d(d-1)^{4n-2}}|V|}.$$

To see this note that if the girth of G is larger than $2q$ the $\bar{\mu}^q_{X,G}$ is the ν_G-average of $\bar{\mu}^q_{X,T_u}$ where T_u is the ball of radius q around u in G, which is a tree rooted at u.

As to $\bar{\mu}^2_{X,\alpha}(x \to x')$, note that if $x \ne x'$ but $d_X(x,x') = 2$ then $\bar{\mu}^2_{X,G}(x \to x') = P^2_G(2)\mu^2_X(x \to x')$ (to get from x to x' we can only consider the paths of length 2 in G) so that the probability of $\bar{\mu}^2_{X,\alpha}(x \to x') > \mu^2_X(x \to x')$ for some such x' is at most:

$$N_k(2) \cdot e^{-\frac{(1-P^2_G(2))^2\left(\mu^2_X(x \to x')\right)^2}{4d(d-1)^2}|V|}.$$

□

Combining the deviation estimates with the spectral gap of the graph, we obtain the main result:

Proposition 2.13. *Assume $4 \le 2n < \frac{1}{2}g(G)$ and that for every $u \in V$, $3 \le \deg(u) \le d$. Then with probability at least*

$$1 - N_k(2n) \cdot e^{-\frac{\varepsilon(k,d,2n)^2}{16d(d-1)^{4n-2}}|V|} - N_k(2) \cdot e^{-\frac{(1-P^2_G(2))^2}{16k^2(2k-1)^2 d(d-1)^2}|V|}$$

for every Y, $\pi : \Gamma_\alpha \to \mathrm{Isom}(Y)$ and every equivariant $f : \Gamma_\alpha \to Y$ there exists an l (depending on f) such that $\frac{1}{2}\eta_d n < l \le n$ and

$$E_{\mu^{2l}_X}(f) \le \frac{2}{1-e^{-2/9}} \cdot \frac{1}{1-\lambda^2(G)} E_{\mu^2_X}(f).$$

Proof. Let $l_0 = \frac{1}{2}\eta_d n$. By Lemma 2.12 we know that with high probability (as in the statement of this Proposition),

$$E_{\bar{\mu}^{2n}_{X,\alpha}}(f) \ge \frac{1}{2} E_{\bar{\mu}^{2n}_{X,G}}(f),$$

and

$$E_{\bar{\mu}^2_{X,\alpha}}(f) \le E_{\mu^2_X}(f)$$

(The second inequality follows from the assumed bound on the random walk $\bar{\mu}^2_{X,\alpha}$ since terms with $x = x'$ don't contribute on either side). Combining these two inequalities with the graph spectral gap (equation 2.3) we get:

$$E_{\bar{\mu}^{2n}_{X,G}}(f) \le \frac{2}{1-\lambda^2(G)} E_{\mu^2_X}(f)$$

We now use Lemma 2.11 in the form:

$$\frac{1}{1-e^{-2/9}} E_{\bar{\mu}^{2n}_{X,G}}(f) \ge \frac{1}{1-Q^{2n}_G} E_{\bar{\mu}^{2n}_{X,G}}(f) \ge \sum_{l_0 < l \le n} \frac{P^{2n}_G(2l)}{1-Q^{2n}_G} E_{\mu^{2l}_X}(f)$$

[3]The sum is over paths $\vec{p}$ starting at the *marked root* of the tree.

where $\sum_{l_0<l\leq n}\frac{P_G^{2n}(2l)}{1-Q_G^{2n}}=1$ by the definition of Q_G^{2n}, to get:

$$\sum_{l_0<l\leq n}\frac{P_G^{2n}(2l)}{1-Q_G^{2n}}E_{\mu_X^{2l}}(f)\leq\frac{2}{1-e^{-2/9}}\frac{1}{1-\lambda^2(G)}E_{\mu_X^2}(f).$$

Finally, the smallest of the $E_{\mu^{2l}}(f)$ is at most equal to their average, so the desired conclusion holds for that particular l. □

Remark 2.14. Modifying the second Lemma and assuming n large enough, the factor $\frac{2}{1-Q_G^{2n}}$ could be replaced with a bound arbitrarily close to 1. For brevity we also note $\frac{2}{1-e^{-2/9}}\leq 10.5$.

2.5. **Geometry.** We begin with some motivation from the case of a unitary representation. The derivation of equation (2.2) then shows that $E_{\mu_X^q}(f)=\left\langle f,(I-H^q_{\mu X})f\right\rangle$ (this inner product is Γ-invariant since the origin is now assumed to be Γ-invariant). H_μ is self-adjoint w.r.t. this inner product, so

$$E_{\mu_X^2}(H_{\mu_X^{2n}}f)=\left\langle f,H_{\mu_X^{2n}}\left(I-H_{\mu_X^2}\right)H_{\mu_X^{2n}}f\right\rangle=E_{\mu_X^{4n}}(f)-E_{\mu_X^{4n+2}}(f).$$

Since $\mu_X^{4n}(x\to x')$, $\mu_X^{4n+2}(x\to x')$ are in some sense close for large n, this should imply that averages of f have small energy (see the rigorous discussion below). This is precisely what we need in order to prove the existence of fixed points.

This analysis is insufficient for our purposes, however. We would like to analyze affine (isometric) actions when inner products are no longer invariant, and even actions on CAT(0) metric spaces, where the equation wouldn’t even make sense. Indeed, it turns out that the the non-positive curvature of Hilbert space is all that is needed here: an analogue of the above formula is proved in the appendix (Proposition B.25, for the random walk $\mu=\mu_X^2$) in two parts, reading:

(2.4)
$$E_{\mu_X^2}(H_{\mu_X^{2n}}f)\leq\frac{1}{2}\int_{\Gamma\backslash X}d\bar{\nu}_X(x)\int_X\left[d\mu_X^{2n+2}(x\to x')-d\mu_X^{2n}(x\to x')\right]d_Y^2(H_{\mu_X^{2n}}f(x),f(x')),$$

and

$$\frac{1}{2}\int_{\Gamma\backslash X}d\bar{\nu}_X(x)\int_X d\mu_X^{2n}(x\to x')d_Y^2(H_{\mu_X^{2n}}f(x),f(x'))\leq E_{\mu_X^{2n}}(f).$$

In the present case $\Gamma\backslash X$ is a single point, and the outer integrals can be ignored (any $x\in X$ is a “fundamental domain” for $\Gamma\backslash X$).

As just indicated, we would like to prove that averaging indeed reduces the variation of f by producing an inequality of the form[4]:

$$E_{\mu_X^2}(H_{\mu_X^{2n}}f)\leq o(1)\cdot E_{\mu_X^{2n}}(f).$$

If $\mu_X^{2n+2}(x\to x')$ were all close to the respective $\mu_X^{2n}(x\to x')$ (e.g. $\mu_X^{2n+2}(x\to x')\leq(1+o(1))\mu_X^{2n}(x\to x')$) we would be done immediately. Unfortunately, such an inequality does not hold for all x'. Fortunately, such an inequality does hold for most x'. The exceptions lie in the “tails” of the distribution: x' which are very far or very close to x. Moreover, in these cases both $\mu_X^{2n+2}(x\to x')$ and $\mu_X^{2n}(x\to x')$ are extremely small and a simple estimate for $d_Y^2(H_{\mu_X^{2n}}f(x),f(x'))$ suffices, leading to an inequality of the form:

$$E_{\mu_X^2}(H_{\mu_X^{2n}}f)\leq o(1)\cdot E_{\mu_X^{2n}}(f)+o(1)\cdot E_{\mu_X^2}(f).$$

[4]Here lies the main motivation for only looking at walks of even length: one of $\mu_X^n(x\to x')$, $\mu_X^{n+1}(x\to x')$ is always be zero since the tree X is bipartite, meaning such an argument could not work.

To be precise let $D = \max\{d_Y(f(x), f(x')) \mid d_X(x,x') = 2\}$ (compare the displacement di_S of the introduction). Then $D^2 \leq 2(2k)(2k-1)E_{\mu_X^2}(f) \leq 8k^2 E_{\mu_X^2}(f)$. Also if $d_X(x,x') \leq 2n$ and has even parity then $d_Y^2(f(x), f(x')) \leq nD^2$ by the triangle inequality and the inequality of the means. By the convexity of the ball of radius $\sqrt{n}D$ around $f(x)$ we then have $d_Y^2(f(x), H_{\mu_X^{2n}} f(x)) \leq nD^2$ ($H_{\mu_X^{2n}} f(x)$ is an average of such $f(x')$). Hence if $d(x,x') \leq 2n$ and is even:

$$d_Y^2(H_{\mu_X^{2n}} f(x), f(x')) \leq 2 \cdot nD^2 \leq 16nk^2 E_{\mu_X^2}(f). \tag{2.5}$$

We can now split the integration in equation (2.4) into the region where $|d_X(x,x') - 2\eta_{2k} n| \leq 2\sqrt{(2n)\log(2n)}$ and its complement. In the first region the difference

$$\left[d\mu_X^{2n+2}(x \to x') - d\mu_X^{2n}(x \to x')\right]$$

is small by Proposition 2.8. The measure of the second region is small by Corollary 2.6, and we can use there the simple bound (2.5) on the integrand. All-in-all this gives:

$$E_{\mu_X^2}(H_{\mu_X^{2n}} f) \leq c_1(2k)\frac{\sqrt{\log(2n)}}{\sqrt{2n}} E_{\mu_X^{2n}}(f) + \frac{32k^2}{p_{2k}} n^{-3} E_{\mu_X^2}(f) +$$
$$+16k^2 c_2(2k)\left(e^{-\eta_{2k}^2\sqrt{n/2}} + \left(\frac{q_{2k}}{p_{2k}}\right)^{\eta_{2k}\sqrt{n/2}}\right) n^2 E_{\mu_X^2}(f).$$

The first term is the main term for the first region. The second is the bound on the contribution from the second region. The third is the contribution from the error term in Proposition 2.8 (again the first region) using simple bound and multiplying by n to account for the possible (even) values of r. This last term goes to zero quickly as $n \to \infty$ so we can conclude:

Proposition 2.15. *There exists constants $c_3(k)$, $c_4(k)$ depending only on k such that for all n:*

$$E_{\mu_X^2}(H_{\mu_X^{2n}} f) \leq c_3(k)\frac{\sqrt{\log n}}{\sqrt{n}} E_{\mu_X^{2n}}(f) + c_3(k)\frac{1}{n^3} E_{\mu_X^2}(f).$$

2.6. **Conclusion.** We first prove the main theorem.

Theorem 2.16. *If G is an expander, $3 \leq \deg(u) \leq d$ for all $u \in V$ and the girth of G is large enough then Γ_α has property (T) with high probability.*

Remark 2.17. Formally we claim: given $k \geq 2$, $d \geq 3$ and $\lambda_0 < 1$ there exists an explicit $g_0 = g(k, \lambda_0)$ such that if the girth of G is at least g_0, $\lambda^2(G) \leq \lambda_0^2$ and the degree of every vertex in G is between 3 and d, then the probability of Γ_α having property (T) is at least $1 - ae^{-b|V|}$ where a, b are explicit and only depend on the parameters k, d and λ_0.

Proof. Choose n large enough such that for some $l_0 \leq \frac{1}{6}n$ we have $r = c_3(k)\frac{\sqrt{\log l_0}}{\sqrt{l_0}}\frac{10.5}{1-\lambda^2(G)} + c_3(k)\frac{1}{l_0^3} < 1$.

By Proposition 2.13 if $g(G) \geq 4n$ (note that this minimum girth essentially only depends on λ,k and the desired smallness of r) then with high probability (going to 1 at least as fast as $1 - ae^{-b|V|}$ for some a, b depending only on n, d, k) for any affine representation Y of Γ_α and any equivariant $f : X \to Y$ we can find l in the range $l_0 < l \leq n$ such that

$$E_{\mu_X^{2l}}(f) \leq \frac{10.5}{1 - \lambda^2(G)} E_{\mu_X^2}(f).$$

By Proposition 2.15 and the choice of n we thus have

$$E_{\mu_X^2}(H_{\mu_X^{2l}}(f)) \leq rE_{\mu_X^2}(f)$$

for all $f \in B^{\Gamma_\alpha}(X_\alpha, Y)$. This means that the sequence of functions defined by $f_{q+1} = H_{\mu_X^{2l}} f_q$ (the choice of l of course differs in each iteration) represents an almost-fixed point for the action of the subgroup Γ_α^2 on Y where Γ_α^2 is the subgroup of Γ_α generated by all words of even length. In fact, $E_{\mu_X^2}(f_q) \leq r^q E_{\mu_X^2}(f_0)$. Using $d_Y^2(H_{\mu_X^{2n}} f(x), f(x)) \leq 8nk^2 E_{\mu_X^2}(f)$ (derived in section 2.5), we see that $d_Y^2(f_{q+1}, f_q) \leq 8nk^2 r^q E_{\mu_X^2}(f_0)$. Since $r < 1$ this makes $\{f_q\}$ into a Cauchy sequence, which thus converges to a fixed-point f_∞. If f_∞ is not Γ_α-fixed, then the midpoint of the interval $[f_\infty, \gamma f_\infty]$ is for any $\gamma \in \Gamma_\alpha \setminus \Gamma_\alpha^2$.

Moreover,

$$d_Y(f_\infty, f_0) \leq \frac{1}{1-\sqrt{r}} \sqrt{8nk^2 E_{\mu_X^2}(f_0)}.$$

This means that

$$\varepsilon = \frac{1-\sqrt{r}}{2\sqrt{16nk^2}}$$

is a Kazhdan constant for Γ_α w.r.t S: Let $\pi : \Gamma_\alpha \to U(Y)$ be a unitary representation, and let f_0 be a unit vector, ε-almost invariant for S. Then $d_Y(\gamma_1\gamma_2 f_0, f_0) < \varepsilon$ for all $\gamma_1, \gamma_2 \in S$ so $E_{\mu_X^2}(f_0) < \frac{1}{2}\varepsilon^2$ which means $d_Y(f_\infty, f_0) < \frac{1}{2}$, in particular $f_\infty \neq 0$. Note that f_∞ is Γ_α^2-invariant, so that the closed subspace $Y_0 \subset Y$ of Γ_α^2-invariant vectors is nonempty. Pick $\gamma \in S$, $y \in Y_0$ and consider $y + \pi(\gamma)y$. This vector is clearly Γ_α-invariant and we are thus done unless $\pi(\gamma)y = -y$ for all $y \in Y_0$. In that case we have $\gamma f_\infty = -f_\infty$, and since $\varepsilon \leq \frac{1}{8}$ we obtain the contradiction:

$$\begin{aligned} 2 = d_Y(f_\infty, \pi(\gamma) f_\infty) &\leq d_Y(f_\infty, f_0) + d_Y(f_0, \gamma f_0) + d_Y(\gamma f_0, \gamma f_\infty) \\ &< \frac{1}{2} + \frac{1}{8} + \frac{1}{2}. \end{aligned}$$

□

A slightly different version is actually needed for the result in [2]. For a fixed integer j, let G_j be the graph obtained from G by subdividing each edge of G into a path of length j, adding $j-1$ vertices in the process. On large scales, the new graph resembles the original one (e.g. every ball of radius $< \frac{1}{2} g(G) \cdot j$ is a tree), but the minimum degree is no longer 3: most vertices, in fact, now have degree 2. We now indicate how to adapt the proof above to this case.

We first remark that the spectral gap of G_j can be bounded in terms of the spectral gap of G. In fact, if $f : V(G_j) \mapsto \mathbb{R}$ is an eigenfunction on G_j with eigenvalue $\cos\varphi = \lambda$, then $f \restriction_V$ is an eigenfunction on G with eigenvalue $\cos j\varphi$. To see this observe first that if $(u, v) \in E$ then $f(u), f(v)$ determine the values of f along the subdivided edge since $\lambda f(w_i)$ is the average of the neighbouring values for any internal vertex w_i of the subdivided edge. Plugging this in the expression of $\lambda f(u)$ as an average over the neighbours in G_j gives the desired result. Now λ cannot be too close to ± 1since that that would imply φ too close to 0 or π, making $j\varphi$ close (but not equal) to a multiple of π, so that $\cos j\varphi$ would be too close to ± 1, contradicting the spectral gap of G. Write this bound as $\frac{1}{1-\lambda^2(G_j)} \leq c(\lambda^2(G), j)$.

We next adjust Proposition 2.13. Define $\varepsilon(k, d, j, 2n)$ to be the maximum of $\mu_T^{2n}(x \to x')$ over all trees of depth $2n$ rooted at x, which are obtained by subdividing edges in trees with degrees in the interval $[3, d]$. We then have:

Proposition. *Assume that* $200j \leq 2n < \frac{1}{2}g(G_j)$ *and that for every* $u \in V$, $3 \leq \deg(u) \leq d$. *Then with probability at least*

$$1 - N_k(2n) \cdot e^{-\frac{\varepsilon(k,d,j,2n)^2}{16d(d-1)^{4n-2}}|V|} - N_k(2) \cdot e^{-\frac{(1-P^2_{G_j}(2))^2}{8kd(2k-1)(d-1)^2}|V|}$$

for every Y, $\pi : \Gamma_\alpha \to \mathrm{Isom}(Y)$ *and every equivariant* $f : \Gamma_\alpha \to Y$ *there exists an* l *(depending on* f*) such that* $\frac{1}{2}\eta_d \frac{n}{8j\log n} < l \leq n$ *and*

$$E_{\mu_X^{2l}}(f) \leq \frac{11}{1-\lambda^2(G_j)} E_{\mu_X^2}(f).$$

Proof. Also note that $P^2_{G_j}(2)$ can be readily bounded since we know all possible balls of radius 2 in G_j. The only other change needed in the proof is a reevaluation of the bound on the inverse of the probability that the $2n$-step random walk on the graph G_j travels a distance at least $\frac{1}{2}\eta_d \frac{n}{8j\log n}$ from its starting point (the factor 2 used to compare $\mu^{2n}_{X,\alpha}(x \to x')$ to $\mu^{2n}_{X,G}(x \to x')$ remains). The idea of this is to think of the random walk on G_j as a series of 'macro-steps', each consisting of a random walk on the 'star' of radius j centered at a vertex of G until a neighbouring vertex in G is reached. In this context we will term 'micro-steps' the steps of this last random walk, i.e. the usual steps from before. The sequence of 'macro-steps' is a random walk on G (every neighbour is clearly reach with equal probability), which we know to travel away from the origin with high probability, assuming enough 'macro-steps' are taken. The expected number of 'micro-steps' until reaching an 'endpoint' is j^2, so on first approximation we can think of the $2n$-step walk on G_j as a variant of the $\frac{2n}{j^2}$-step walk on G (up to a correction of length j at the first and last steps). Of course, there can deviations. It clearly suffices to estimate the probability that all 'macro-steps' take less than $8j^2\log n$ 'micro-steps' to complete, since if that happens then we must have made at least $j^{-2}\frac{2n}{8\log n}$ 'macro-steps'. Then the probability of the final distance from the origin being less than $\frac{1}{2}\eta_d \frac{2n}{j^2 8\log n} \cdot j$ is at most $e^{-2/9}$ for the same reason as in the original proposition. A bound for an individual 'macro-step' follows from a large deviation estimate. Noting that there are at most $2n$ macro-steps in the process completes the bound. □

Corollary 2.18. *Given* k, d, λ_0 *and* j *there exists an explicit* $g_0 = g(k, \lambda_0, j)$ *such that if the girth of* G *is at least* g_0, $\lambda^2(G) \leq \lambda_0^2$ *and the degree of every vertex in* G *is between* 3 *and* d, *then the probability of a random group* Γ_α *resulting from a labelling of* G_j *having property (T) is at least* $1 - ae^{-b|V|}$ *where* a, b *are explicit and only depend on the parameters* k, d, λ_0 *and* j.

Proof. Identical to the main theorem, except we now choose n large enough so that $l_0 < \frac{1}{2}\eta_d \frac{n}{8j\log n}$ satisfies $r = c_3(k)\frac{\sqrt{\log l_0}}{\sqrt{l_0}} 11 \cdot c(\lambda^2(G), j) + c_3(k)\frac{1}{l_0^3} < 1$. □

APPENDIX A. CAT(0) SPACES AND CONVEXITY

Let (Y, d) be a metric space.

Definition A.1. A *geodesic path* in Y is a rectifiable path $\gamma : [0, l] \to Y$ such that $d(\gamma(a), \gamma(b)) = |a - b|$ for all $a, b \in [0, l]$. The space Y is called *geodesic* if every two points of Y are connected by a geodesic path. We say that Y is *uniquely geodesic* if that path is unique.

Definition A.2. Let Y be a geodesic space, $\lambda > 0$. Say that $f : Y \to \mathbb{R}$ is *(λ-)convex* if $f \circ \gamma$ is (λ-)convex for every geodesic path γ. We say $f : [a, b] \to \mathbb{R}$ is λ-convex if it is convex, and furthermore $(D^+f)(y) - (D^-f)(x) \geq \lambda(y - x)$ for every $a < x < y < b$. (informally, the second derivative of the function is bounded away from zero).

We need the following property of λ-convex functions on intervals:

Lemma A.3. *Let $f : [a, b] \to \mathbb{R}$ be λ-convex. Then* $\max f([a, b]) - \min f([a, b]) \geq \frac{1}{8}\lambda(b - a)^2$.

Proof. Assume first that f is monotone nondecreasing. By convexity, the convergence of $\frac{f(x+h)-f(x)}{h}$ to $(D^+f)(x)$ as $h \to 0^+$ is monotone. The limit is nonnegative. By the monotone convergence Theorem,

$$\int_a^{b-\delta} (D^+f)(x)dx = \lim_{h\to 0^+} \frac{1}{h}\int_{b-\delta}^{b-\delta+h} f(x)dx - \lim_{h\to 0^+} \frac{1}{h}\int_a^{a+h} f(x)dx$$

and by continuity of f in the interior of the interval we get:

$$\int_a^{b-\delta} (D^+f)(x)dx = f(b - \delta) - f(a)$$

letting $\delta \to 0$ and using the monotone convergence Theorem again (D^+f is nonnegative by assumption) we obtain:

$$\int_a^{b} (D^+f)(x)dx \leq f(b) - f(a)$$

now by the λ-convexity assumption, $(D^+f)(x) \geq \lambda(x-a)$ (since $(D^+f)(x) \geq (D^-f)(x) \geq (D^+f)(a) + \lambda(x - a)$ and $(D^+f)(a) \geq 0$). Thus

$$f(b) - f(a) \geq \int_a^b \lambda(x - a)dx = \lambda\frac{(b - a)^2}{2} \tag{A.1}$$

In the general case, let f obtain its minimum on $[a, b]$ at $c \in (a, b)$. Then f is monotone non-increasing on $[a, c]$ and nondecreasing on $[c, b]$. It follows that $f(a)-f(c) \geq \frac{\lambda}{2}(a-c)^2$ and $f(b) - f(c) \geq \frac{\lambda}{2}(b - c)^2$. Since either $c - a \geq \frac{1}{2}(b - a)$ or $b - c \geq \frac{1}{2}(b - a)$ we are done. □

Lemma A.4. *Let f be a λ-convex function, bounded below, on a complete geodesic metric space Y. Then f has a unique global minimum.*

Proof. Let $m = \inf\{f(y) \mid y \in Y\}$, and for $\varepsilon > 0$ consider the closed set

$$Y_\varepsilon = \{y \in Y \mid f(y) \leq m + \varepsilon\}.$$

We claim that $\lim_{\varepsilon\to 0} \operatorname{diam}(Y_\varepsilon) = 0$ and therefore that their intersection is nonempty. Let $x, y \in Y_\varepsilon$, and consider a geodesic $\gamma : [0, d(x, y)] \to Y$ connecting x, y. $g = f \circ \gamma$ is a 2-convex function on $[0, d(x, y)]$ and since $x, y \in Y_\varepsilon$ we have $g(0), g(d(x, y)) \leq m + \varepsilon$, and thus $m \leq g(t) \leq m + \varepsilon$ for all $t \in [0, d(x, y)]$. By the previous Lemma we get $d(x, y)^2 \leq \frac{8\varepsilon}{\lambda}$ and therefore $\operatorname{diam}(Y_\varepsilon) \leq \sqrt{\frac{8\varepsilon}{\lambda}}$ as promised. □

Definition A.5. A geodesic space (Y, d) will be called a CAT(0) space if for every three points $p, q, r \in Y$, and every point s on a geodesic connecting p, q, it is true that $d(s, r) \leq |SR|$ where $P, Q, R \in \mathbb{E}^2$ (Euclidean 2-space) form a triangle with sides $|PQ| = d(p, q)$, $|QR| = d(q, r)$, $|RP| = d(r, p)$ and $S \in PQ$ satisfies $|PS| = d(p, s)$.

Remark A.6. This clearly implies the usual CAT(0) property: if $s \in [p,q]$, $t \in [p,r]$ and $S \in PQ$, $T \in PR$ such that $|PS| = d(p,s)$, $|PT| = d(p,t)$ then $d(s,t) \leq |ST|$.

From now on let Y be a complete CAT(0) space.

Lemma A.7. *(Explicit CAT(0) inequality) Let P, A, B be a triangle in $\mathbb{E}^2$ with sides $|PA| = a, |PB| = b, |AB| = c + d$, and let $Q \in AB$ satisfy $|Q - A| = c$, $|Q - B| = d$. Let $l = |PQ|$. Then:*

$$l^2 = \frac{d}{c+d}a^2 + \frac{c}{c+d}b^2 - cd$$

Moreover, let $p, q, r, s \in Y$ where s lies on the geodesic containing p, q. Then

$$d^2(r,s) \leq \frac{d(p,s)}{d(p,q)}d^2(q,r) + \frac{d(s,q)}{d(p,q)}d^2(p,r) - d(p,s)d(s,q)$$

Proof. $\cos(\angle BAP) = \frac{c^2+a^2-l^2}{2ac} = \frac{(c+d)^2+a^2-b^2}{2(c+d)a}$ and therefore

$$l^2 = c^2 + a^2 - c(c+d) - \frac{c}{c+d}a^2 + \frac{c}{c+d}b^2$$

as desired. The second part is a restatement of the definition of a CAT(0) space. □

Corollary A.8. *Fix $y_0 \in Y$. Then the function $f(y) = d^2(y, y_0)$ is strictly convex along geodesics. In fact, it is 2-convex.*

Proof. Let $y_1, y_2, y_3 \in Y$ be distinct and lie along a geodesic in that order. By Lemma A.7

$$f(y_2) \leq \frac{d(y_2,y_3)}{d(y_1,y_3)}f(y_1) + \frac{d(y_1,y_2)}{d(y_1,y_3)}f(y_3) - d(y_1,y_3)d(y_2,y_3),$$

and since $d(y_1,y_3)d(y_2,y_3) > 0$ we have strict convexity. In particular, we obtain the two inequalities

$$\frac{f(y_2)-f(y_1)}{d(y_2,y_1)} \leq \frac{f(y_3)-f(y_1)}{d(y_3,y_1)} - d(y_2,y_3)$$

and

$$\frac{f(y_3)-f(y_2)}{d(y_3,y_2)} \geq \frac{f(y_3)-f(y_1)}{d(y_3,y_1)} + d(y_1,y_2)$$

which together imply:

$$\frac{f(y_3)-f(y_2)}{d(y_3,y_2)} - \frac{f(y_2)-f(y_1)}{d(y_2,y_1)} \geq d(y_1,y_2) + d(y_2,y_3) = d(y_1,y_3)$$

As to the second part, let $y_1 < y_4$ lie along a geodesic path γ. Let $y_1 < y_2 < y_3 < y_4$. Then applying the last inequality twice, for the triplets (y_1,y_2,y_3) and (y_2,y_3,y_4) we obtain:

$$\frac{f(y_4)-f(y_3)}{d(y_4,y_3)} - \frac{f(y_2)-f(y_1)}{d(y_2,y_1)} \geq d(y_1,y_3) + d(y_2,y_4)$$

letting $y_2 \to y_1$ and $y_3 \to y_4$ the LHS converge to the difference of the right- and left-derivatives of $f \circ \gamma$ at y_1 and y_4 respectively, while the RHS converges to $2d(y_1,y_4)$ as desired. □

Lemma A.9. *The metric $d : Y \times Y \to \mathbb{R}$ is convex.*

Proof. Let $a_1, a_2, b_1, b_2 \in Y$ and let $\gamma_i : [0,1] \to Y$ be uniformly parametrized geodesics from a_i to b_i. We need to prove:

$$d(\gamma_1(t), \gamma_2(t)) \leq (1-t)d(a_1, a_2) + td(b_1, b_2).$$

Consider first the case $\gamma_1(0) = \gamma_2(0)$ (i.e. $a_1 = a_2$). Then $\gamma_i(t)$ are two points along two edges of the geodesic triangle a_1, b_1, b_2. By similarity of triangles in $\mathbb{E}^2$ and the strong CAT(0) condition we are done.

In the general case, let $p_i = \gamma_i(t)$. Let $\gamma_0 : [0,1] \to Y$ be the uniformly parametrized geodesic from a_1 to b_2, and let $r = \gamma'(t)$. Then by the special case for γ_1, γ_0 which begin at a_1 we have $d(p_1, r) \leq td(b_1, b_2)$. Similarly by the special case for $\gamma_2^{-1}, \gamma_0^{-1}$ which begin at b_2 are have $d(p_2, r) \leq (1-t)d(a_2, a_1)$. By the triangle inequality we are done. □

Appendix B. Random walks on metric spaces

This appendix follows quite closely section 3 of [2] supplying proofs of the results. The target is Proposition B.25, the geometric result needed for the property (T) proof.

B.1. **Random Walks and the Center of Mass.** Let X be a topological space, and let $\mathcal{M}_X$ be the set of regular Borel probability measures on X (if X is countable & discrete this is the set of non-negative norm 1 elements of $l_1(X)$). Topologize $\mathcal{M}_X$ as a subset of the space of finite regular Borel measures on X (with the total variation norm). This makes $\mathcal{M}_X$ into a closed convex subset of a Banach space.

Definition B.1. A *random walk (*or a *diffusion)* on X is a continuous map $\mu : X \to \mathcal{M}_X$, whose value at x we write as $\mu(x \to)$. The set of random walks will be denoted by $\mathcal{W}_X$.

(1) The *composition* of a measure $\nu \in \mathcal{M}_X$ with the random walk $\mu \in \mathcal{W}_X$ is the measure:

$$(\nu \cdot \mu)(A) = \int_X d\nu(x)\mu(x \to A).$$

This is well defined since μ is continuous and bounded. It is clearly a probability measure. It seems natural to also think of this definition in terms of a vector-valued integral.

(2) The *composition* (or *convolution*) of two random walks μ, μ' is the random walk:

$$(\mu * \mu')(x \to) = \int_X d\mu(x \to x')\mu'(x' \to).$$

The integral is continuous so this is, indeed, a random walk.

(3) We will also write

$$\mu^n \stackrel{\text{def}}{=} \underbrace{\mu * \cdots * \mu}_{n},$$

and also $d\mu^n(x \to y)$ for the probability (density) of going from x to y in n independent steps.

If X is discrete, we think of $\mu(x \to y) = \mu(x \to)(y)$ as the transition probability for a Markov process on X. In this case the stationary Markov process

$$\delta(x \to A) = \delta_x(A) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

is a random walk under the above definition. It isn't in the non-discrete case since the map $x \to \delta_x$ is weakly continuous but not strongly continuous. It might be possible to define random walks as *weakly* continuous maps $X \to \mathcal{M}_X$ but this would require more analysis.

Definition B.2. Let ν be a (possibly infinite) regular Borel measure on X, $\mu \in \mathcal{W}_X$. Say that μ is *ν-symmetric*, or that ν is a *stationary measure for* μ if for all measurable $A, B \subseteq X$
$$\int_A \mu(x \to B) d\nu(x) = \int_B \mu(y \to A) d\nu(y).$$
In other words, μ is ν-symmetric iff for all measurable $\varphi : X \times X \to \mathbb{R}_{\geq 0}$
$$\int_{X\times X} \varphi(x,x') d\nu(x) d\mu(x \to x') = \int_{X\times X} \varphi(x,x') d\nu(x') d\mu(x' \to x).$$
If X is discrete, we say that μ is *symmetric* if this holds where ν is the counting measure (i.e. if $\mu(x \to \{y\}) = \mu(y \to \{x\})$ for all x, y).

Example B.3. Let $G = (V, E)$ be a graph where the degree of each vertex is finite. We can then define a random walk by having $\mu(u \to v) = \frac{1}{d_u} N_u(v)$, where N_u is the number of edges between u and v (i.e. we take account of multiple and self- edges if they exist). This walk is ν-symmetric for the measure $\nu(u) = d_u$.

In particular, if G is the Cayley graph of a group Γ w.r.t. a finite symmetric generating set S we obtain the *standard random walk* on Γ. This is a symmetric random walk since the associated measure $\bar{\nu}_G$ on Γ is Γ-invariant.

Definition B.4. A *codiffusion* on a topological space Y is a continuous map $c : M \to Y$ defined on a convex subset $M \subseteq \mathcal{M}_Y$ containing all the delta-measures δ_y such that $c(\delta_y) = y$ for every $y \in Y$ and such that the pullback $c^{-1}(y)$ is *convex* for every $y \in Y$.

Example B.5. In an affine space we have the "*centre of mass*":
$$c(\sigma) = \int\limits_Y \vec{y} \cdot d\sigma(\vec{y})$$
defined on the set the measures for which the co-ordinate functions y_i are integrable.

In the case where Y is an affine inner product space, we can characterize $c(\sigma)$ for some[5] $\sigma \in \mathcal{M}_Y$ in a different fashion: consider the function ("Mean-Square distance from y")
$$d_\sigma^2(y) = \int\limits_Y \|y - y'\|_Y^2 \, d\sigma(y'),$$
and note that
$$d_\sigma^2(y) = \int\limits_Y \|(y - c(\sigma)) - (y' - c(\sigma))\|_Y^2 \, d\sigma(y') =$$
$$= \|(y - c(\sigma))\|_Y^2 + d_\sigma^2(c(\sigma)) + 2 \left\langle y - c(\sigma), c(\sigma) - \int\limits_Y y' d\sigma(y') \right\rangle.$$
Since $c(\sigma) - \int\limits_Y y' d\sigma(y') = 0$ by definition, we find:
$$d_\sigma^2(y) = \|(y - c(\sigma))\|_Y^2 + d_\sigma^2(c(\sigma)). \tag{B.1}$$
In other words, $c(\sigma)$ is the *unique* point of Y where $d_\sigma^2(y)$ achieves its minimum. More generally if Y be a metric space and $\sigma \in \mathcal{M}_Y$, we set:
$$d_\sigma^2(y) \stackrel{\text{def}}{=} \int\limits_Y d^2(y, y') \cdot d\sigma(y').$$

[5] σ must be such that the following integral converges for all $y \in Y$.

Lemma B.6. $d_\sigma^2(y) \leq d_\sigma^2(y_1) + d_Y^2(y, y_1) + 2d_Y(y, y_1)d_\sigma(y_1) = (d_Y(y, y_1) + d_\sigma(y_1))^2$.

Proof. By the triangle inequality, $d_Y(y, y') \leq d_Y(y, y_1) + d_Y(y_1, y')$. Squaring and integrating $d\sigma(y')$ gives:

$$d_\sigma^2(y_1) \leq d_Y^2(y, y_1) + d_\sigma^2(y_1) + 2d_Y(y, y_1) \int_Y d_Y(y_1, y')d\sigma(y').$$

Using Cauchy-Schwarz (which, for probability measures σ, reads $(\int f d\sigma)^2 \leq \int f^2 d\sigma$) completes the proof. □

Corollary B.7. *If $d_\sigma(y)$ is finite for some $y \in Y$ then it is finite everywhere.*

Corollary B.8. *If $d_\sigma(y)$ is finite, then it is σ-integrable. Furthermore, for any $y_1 \in Y$ we have:*

$$4d_\sigma^2(y_1) \geq \int_Y d_\sigma^2(y)d\sigma(y) = \left\langle d_Y^2(y, y')\right\rangle_{\sigma \times \sigma}.$$

Proof. This follows from Lemma B.6 immediately by integrating $d\sigma(y)$ and using Cauchy-Schwarz again. □

We therefore let M_0 be the set of (regular Borel) probability measures on Y such that $d_\sigma^2(y)$ is finite. This is clearly a convex set (though it is not closed if d is unbounded).

Definition B.9. Let $\sigma \in M_0$. If $d_\sigma(y)$ has a unique minimum on Y, it is called the (*Riemannian*) *centre of mass of* σ, denoted again by $c(\sigma)$.

If Y is a CAT(0)-space, then Corollary A.8 states that, as a function of y, $d_Y^2(y, y')$ is 2-convex. This property clearly also holds to $d_\sigma^2(y)$ (seen e.g. by differentiating under the integral sign and using the monotone convergence Theorem). The existence of $c(\sigma)$ then follows from Lemma A.4. In this case we can take $M = M_0$.

Remark B.10. The 2-convexity of $d_\sigma^2(y)$ on the segment $[y, c(\sigma)]$ implies the following crucial inequality[6]:

$$d_\sigma^2(y) \geq d_\sigma^2(c(\sigma)) + d_Y^2(c(\sigma), y) \tag{B.2}$$

(c.f. equation (A.1)).

We can also integrate this inequality $d\sigma(y)$ to obtain the bound $d_\sigma^2(c(\sigma)) \leq \frac{1}{2}\left\langle d_Y^2(y, y')\right\rangle_{\sigma \times \sigma}$.

B.2. **The Heat operator.** From now on let X be a topological space, (Y, d) a complete CAT(0) space. We also fix a measure $\nu \in \mathcal{M}_X$, a ν-symmetric random walk $\mu \in \mathcal{W}_X$, and let c be the center-of-mass codiffusion on Y, defined on the convex set $M_0 \subseteq \mathcal{M}_Y$. Our arena of play will be two subspaces of $M(X, Y)$, the space of measurable functions from X to Y. The first one is a generalization of the usual L^2 spaces:

Definition B.11. Let $f, g : X \to Y$ be measurable. The L^2 *distance* between f, g (denoted $d_{L^2(\nu)}(f, g)$) is given by

$$d_{L^2(\nu)}(f, g) = \left(\int_X d_Y^2(f(x), g(x))d\nu(x)\right)^{1/2}.$$

[6]This formula (and the next one) are actually equalities in the Hilbertian case – see equation (B.1).

This is a (possibly infinite) pseudometric. Moreover, $d_{L^2(\nu)}(f,g)=0$ iff $f=g$ ν-a.e. as usual. Fixing $f_0\in M(X,Y)$, we set:

$$L^2_\nu(X,Y)=\left\{f\in M(X,Y)\middle| d_{L^2(\nu)}(f,f_0)<\infty\right\}.$$

As usual, the completeness of Y implies the completeness of $L^2_\nu(X,Y)$.

Remark B.12. That $L^2_\nu(X,Y)$ is a CAT(0) space follows from Y having the property.

Proof. We start with some notation. Let $y_1,y_2\in Y$, $t\in[0,1]$. Set $[y_1,y_2]_t$ to be the point at distance $td(y_1,y_2)$ from y_1 along the segment. In a CAT(0) space this is a continuous function on $Y\times Y\times[0,1]$. Now let $f,g\in M(X,Y)$ and define $[f,g]_t(x)=[f(x),g(x)]_t$. Then $[f,g]_t$ is measurable as well, and clearly $d^2_{L^2(\nu)}(f,[f,g]_t)=t^2d^2_{L^2(\nu)}(f,g)$ since $d_Y(f(x),[f,g]_t(x))=t^2d_Y(f(x),g(x))$ holds pointwise. Thus $L^2_\nu(X,Y)$ is a geodesic space. Next, let $f,g,h\in L^2_\nu(X,Y)$ and let $u=[f,g]_t$. The explicit CAT(0) inequality for $h(x),f(x),g(x),u(x)$ reads:

$$d^2_Y(h(x),u(x))\le td^2_Y(g(x),h(x))+(1-t)d^2_Y(f(x),h(x))-t(1-t)d^2_Y(f(x),g(x),$$

and integration w.r.t. $d\nu(x)$ gives the explicit CAT(0) inequality of $L^2_\nu(X,Y)$. This is immediately equivalent to the general CAT(0) condition. □

Definition. Let $f\in M(X,Y)$, and let $\tau\in\mathcal{M}_X$. The *pushforward measure* $f_*\tau\in\mathcal{M}_Y$ is the Borel measure on Y defined by $(f_*\tau)(E)=\tau(f^{-1}(E))$.

Definition B.13. Let $\varepsilon\in[0,1]$, $f\in M(X,Y)$.

(1) The *Heat operators* H^ε_μ are

$$(H^\varepsilon_\mu f)(x)=c\left(f_*\left(\varepsilon\mu\left(x\rightarrow\right)+\left(1-\varepsilon\right)\delta_x\right)\right)$$

Whenever this makes sense. In particular $H=H^1_\mu=H^1$ is called the *heat operator*. Note that by the convexity of M, $H^\varepsilon f$ are defined whenever H^1f is.

(2) Say that a function $f\in M(X,Y)$ is *(μ-) harmonic* if $Hf(x)$ is defined for *all* $x\in X$ and equal to $f(x)$.

Remark B.14. If $Hf(x)=f(x)$ then $d^2_{f_*\mu(x\rightarrow)}(y)$ and $d^2_Y(y,f(x))$ achieve their minimum at the same point, $y=f(x)$. Thus $H^\varepsilon f(x)=f(x)$ for all $0\le\varepsilon\le1$. In other words, f is harmonic iff $H^\varepsilon f=f$ for all ε.

Example B.15. Consider a graph $G=(V,E)$ with the graph metric, and let μ be the standard random walk on G. Let $f:V\to\mathbb{R}$ by any function. It is then clear that H_1 is the "local average" operator (the normalized adjacency matrix).

By the following Lemma (since $f_*(\varepsilon\mu(x\rightarrow)+(1-\varepsilon)\delta_x)=\varepsilon f_*(\mu(x\rightarrow))+(1-\varepsilon)\delta_{f(x)}$), $H^\varepsilon f$ is well defined for some $L^2_\nu(X,Y)$ spaces.

Lemma B.16. *Let $f_0\in M(X,Y)$ be a μ-harmonic, and let $f\in L^2_\nu(X,Y)$ (around f_0). Then $f_*\mu(x\rightarrow)\in M$ for ν-a.e. $x\in X$. In particular, $H^\varepsilon f$ is defined ν-a.e.*

Proof. Since $\frac{a+b}{2}\le\sqrt{\frac{a^2+b^2}{2}}$ we have for any $y\in Y$, $x'\in X$:

$$\int_X d^2_Y\left(y,f(x')\right)d\mu(x\rightarrow x')\le 2\int_X d^2_Y\left(y,f_0(x')\right)d\mu(x\rightarrow x')+2\int_X d^2_Y\left(f_0(x'),f(x')\right)d\mu(x\rightarrow x').$$

The first integral is finite by μ-harmonicity and Lemma B.6. Integrating the second one $d\nu(x)$ and using the fact that μ is ν-symmetric, we find:

$$\int\limits_{X\times X} d\nu(x)d\mu(x\to x')d_Y^2\left(f_0(x'),f(x')\right)d\mu(x\to x')$$

$$=\int_{X\times X} d_Y^2\left(f_0(x'),f(x')\right)d\nu(x')d\mu(x'\to x)=\int\limits_X d_Y^2\left(f_0(x'),f(x')\right)d\nu(x')$$

$$=d^2_{L^2(\nu)}(f,f_0)<\infty.$$

It must thus be finite ν-a.e. □

Actually the same proof shows that if Hf_0 is well-defined ν-a.e. then so is Hf for all $f\in L^2_\nu(X,Y)$ – but we should know more:

Claim B.17. Let $\sigma_1,\sigma_2\in M_0$, and let τ be a probability measure on $Y\times Y$ such that $\tau(A\times Y)=\sigma_1(A)$ and $\tau(Y\times A)=\sigma_2(A)$ for all measurable $A\subseteq Y$. Then

$$d_Y^2\left(c(\sigma_1),c(\sigma_2)\right)\le\int\limits_{Y\times Y} d_Y^2(y,y')d\tau(y,y').$$

Proof. I have only managed to prove the Hilbertian case. This should hold for CAT(0) spaces in general.

$$c(\sigma_1)=\int_Y yd\sigma_1(y)=\int_{Y\times Y} yd\tau(y,y')$$

and

$$c(\sigma_2)=\int_{Y\times Y} y'd\tau(y,y'),$$

we have by Cauchy-Schwarz:

$$\|c(\sigma_1)-c(\sigma_1)\|_Y^2\le\int_{Y\times Y}\|y-y'\|_Y^2\,d\tau(y,y')\cdot\int_{Y\times Y}d\tau(y,y').$$

□

Proposition B.18. *Let $f_1,f_2\in M(X,Y)$ satisfy $f_{i*}\mu(x\to)\in M_0$ for ν-a.e. $x\in X$. Then $d_{L^2(\nu)}(H^\varepsilon f_1,H^\varepsilon f_2)\le d_{L^2(\nu)}(f_1,f_2)$. In particular, $H^\varepsilon:L^2_\nu(X,Y)\to L^2_\nu(X,Y)$ is Lipschitz continuous.*

Proof. Let $\sigma_i=f_{i*}\mu_\varepsilon(x\to)$, and let $\tau=(f_1\times f_2)_*(\mu_\varepsilon(x\to))$ where $f_1\times f_2:X\to Y\times Y$ is the product map. By the claim

$$d_Y^2\left(H^\varepsilon f_1(x),H^\varepsilon f_2(x)\right)\le\int_{X\times X} d_Y^2\left(f_1(x'),f_2(x')\right)d\mu_\varepsilon(x\to x').$$

The result now follows by integrating $d\nu(x)$ and using the symmetry of μ_ε. In particular, we note that

$$d_{L^2(\nu)}(H^\varepsilon f,H^\varepsilon f_0)\le d_{L^2(\nu)}(f,f_0)$$

and if, in addition, $H^\varepsilon f_0\in L^2_\nu(X,Y)$ (e.g. if f_0 is harmonic) then $H^\varepsilon f\in L^2_\nu(X,Y)$ for all $f\in L^2_\nu(X,Y)$. □

Continuing in a different direction, in order for $Hf(x)$ to be defined, it suffices to know that $\int_X d_Y^2(y, f(x'))d\mu(x \to x')$ is finite for some y. Thinking "locally" we make the natural choice $y = f(x)$, and set:

$$|df|_\mu^2(x) \stackrel{\text{def}}{=} \int_X d_Y^2\left(f(x), f(x')\right) d\mu(x \to x').$$

The second (and more important) space of functions to be considered is

$$B_\nu(X, Y) = \left\{ f \in M(X, Y) \middle| |df|_\mu \in L^2(\nu) \right\}$$

($L^2(\nu) = L^2_\nu(X, \mathbb{R})$ is the usual space of square-integrable real valued-functions on X). By the preceding discussion if $f \in B_\nu(X, Y)$ then $Hf(x)$ is defined ν-a.e. Anticipating the following section we call $B_\nu(X, Y)$ the space of functions of *finite energy*.

We return now to the formula $d_\sigma^2(y) \geq d_\sigma^2(c(\sigma)) + d_Y^2(y, c(\sigma))$ which followed from the 2-convexity of $d_\sigma^2(y)$. Setting $\sigma = f_*(\mu(x \to))$ (so that $c(\sigma) = H_\mu f(x)$) and writing out $d_\sigma^2(H_\mu f(x))$ in full we define:

$$|d'f|_\mu^2(x) \stackrel{\text{def}}{=} d_\sigma^2(c(\sigma)) = \int_X d_Y^2\left(H_\mu f(x), f(x')\right) d\mu(x \to x')$$

(finite for any $f \in B_\nu(X, Y)$). Then for any measurable f

$$\int_X d_Y^2(y, f(x'))d\mu(x \to x') \geq d_Y^2(y, H_\mu f(x)) + |d'f|_\mu^2(x) \tag{B.3}$$

(Note that $H_\mu f$ is undefined iff the LHS is infinite). We now derive an important pair of inequalities which are a basic ingredient of the proof that our random groups have property (T):

Lemma B.19.

$$|d'f|_{\mu^n}^2(x) \leq |df|_{\mu^n}^2(x).$$

Proof. Follows from equation (B.3) by replacing μ with μ^n (which is also ν-symmetric), setting $y = f(x)$ and ignoring the $d_Y^2(y, Hf(x))$ term. □

Proposition B.20. *Let*[7] $H_n = H_{\mu^n}$. *Then*

$$\int_X d\nu(x) \cdot |d(H_n f)|_\mu^2(x) \leq \int_{X\times X} d\nu(x) \left[d\mu^{n+1}(x \to x') - d\mu^n(x \to x')\right] d_Y^2(H_n f(x), f(x')).$$

Proof. Set $y = H_n f(x'')$, and integrate (B.3) w.r.t. $d\nu(x'')d\mu(x'' \to x) = d\nu(x)d\mu(x \to x'')$ on the RHS and LHS respectively to get:

$$\int_{X\times X} d\nu(x'')d\mu(x'' \to x) \int_X d\mu^n(x \to x')d_Y^2(H_n f(x''), f(x')) \geq \int_{X\times X} d\nu(x)d\mu(x \to x'')d_Y^2(H_n f(x), H_n f(x''))+$$

$$+ \int_{X\times X} d\nu(x)d\mu(x \to x'') \int_X d\mu^n(x \to x')d_Y^2(H_n f(x), f(x')).$$

Now on the LHS, $\int_X d\mu(x'' \to x)\mu^n(x \to A) \stackrel{\text{def}}{=} \mu^{n+1}(x'' \to A)$. On the LHS the inner integral does not depend on x'' and $\int_X d\mu(x \to x'') = 1$. In other words:

$$\int_{X\times X} d\nu(x'')d\mu^{n+1}(x'' \to x')d_Y^2(H_n f(x''), f(x')) \geq \int_{X\times X} d\nu(x)d\mu(x \to x'')d_Y^2(H_n f(x), H_n f(x''))+$$

[7] In the affine case (but not in general) this equals the n-th power (iterate) of H_μ.

$$+\int_{X\times X} d\nu(x)d\mu^n(x\to x')d_Y^2(H_nf(x),f(x')).$$

On the LHS we can now rename the variable x'' to be x. We also move the second term on the RHS to the left. Finally we note that $\int_X d\mu(x \to x'')d_Y^2(H_nf(x), H_nf(x'')) \stackrel{\text{def}}{=} |d(H_nf)|_\mu^2(x)$. □

B.3. **The Energy.** We can rewrite the results of the previous section more concisely by introducing one more notion.

Definition B.21. Let $f \in M(X,Y)$. The *energy of f* is:

$$E(f) = \frac{1}{2}\left\||df|_\mu\right\|^2_{L^2(\nu)} = \frac{1}{2}\int_X d\nu(x)\int_X d\mu(x\to x')d_Y^2(f(x),f(x')).$$

Lemma B.22. *Let $f_0, f \in M(X,Y)$ such that $d_{L^2(\nu)}(f,f_0) < \infty$. If f_0 has finite energy then so does f.*

Proof. We recall that $\int d\nu(x)d\mu(x\to x')d_Y^2(f(x'),f_0(x')) = d^2_{L^2(\nu)}(f,f_0)$. Using the inequality of the means and the triangle inequality we get:

$$d_Y^2(f(x),f(x')) \le 3d_Y^2(f(x),f_0(x)) + 3d_Y^2(f_0(x),f_0(x')) + 3d_Y^2(f_0(x'),f(x')).$$

Multiplying by $\frac{1}{2}$ and integrating $d\nu(x)d\mu(x\to x')$ gives:

$$E(f) \le 3E(f_0) + 3d^2_{L^2(\nu)}(f_0,f).$$

□

Proposition B.23. *$B_\nu(X,Y)$ is a convex subset of $M(X,Y)$. $E(f)$ is a convex function on $B_\nu(X,Y)$.*

Proof. We prove the stronger assertion that $\sqrt{E}$ is convex. Let $f,g \in B_\nu(X,Y)$ and let $t\in[0,1]$. Let $u_t = [f,g]_t$. By the convexity of the metric (Lemma A.9), we have

$$d_Y(u_t(x),u_t(x')) \le (1-t)\cdot d_Y(f(x),f(x')) + t\cdot d_Y(g(x),g(x')).$$

Now since

$$\sqrt{2E(f)} = \|d_Y(f(x),f(x'))\|_{L^2(\nu\cdot\mu)},$$

the triangle inequality of $L^2(\nu\cdot\mu)$ reads:

$$\sqrt{E(u_t)} \le (1-t)\sqrt{E(f)} + t\sqrt{E(g)}.$$

This implies both that $E(u_t) < \infty$ (i.e. $u_t \in B_\nu(X,Y)$) and the convexity of $\sqrt{E}$. □

Definition B.24. Let (X,μ,ν) and (Y,d) be as usual, and let f range over $B_\nu(X,Y)$. We define the *Poincaré constants:*

$$\pi_n(X,Y) = \sup_{E_\mu(f)>0}\frac{E_{\mu^n}(f)}{E_\mu(f)}.$$

For example, in section 2.4 we saw that for a graph G with the usual random walk and a Hilbert space Y one has $\pi_n(G,Y) \le \frac{1}{1-\lambda(G)}$ where $1-\lambda(G)$ is the (one-sided) spectral gap of G).

B.4. **Group actions.** Let Γ be a locally compact group acting on the topological space X, and assume $\nu \in \mathcal{M}_X$ is Γ-invariant. Assume that there exists a measure $\bar{\nu}$ on $\bar{X} = \Gamma\backslash X$ such that if $E \subset X$ is measurable and $\gamma E \cap E = \emptyset$ for all $\gamma \in \Gamma \setminus \{e\}$ then $\nu(E) = \bar{\nu}(\bar{E})$. Also assume that $\Gamma\backslash X$ can be covered by countably many such $\bar{E}$ of finite measure. In this setup every Γ-invariant measurable function $f : X \to \mathbb{R}$ descends to a measurable function $\bar{f} : \bar{X} \to \mathbb{R}$, and we can set:

$$\|f : \Gamma\|^2_{L^2(\nu)} = \int_{\bar{X}} |\bar{f}(\bar{x})|^2 d\bar{\nu}(\bar{x})$$

For example, if Γ is a finite group acting freely on X then $\|f : \Gamma\|^2_{L^2(\nu)} = \frac{1}{|\Gamma|} \|f\|^2_{L^2(\nu)}$ where $\|f\|^2_{L^2(\nu)}$ is the usual L^2 norm on (X, ν).

We now throw in an equivariant diffusion $\mu \in \mathcal{W}_X$, and a metric Γ-space Y (i.e. Γ acts on Y by isometries). If $f \in M^\Gamma(X, Y)$ is equivariant then $H_\mu f$ (wherever defined) is also equivariant, $|df|_\mu(x)$ is invariant and the energy is properly defined by

$$E_\mu(f) = \frac{1}{2} \||df|_\mu : \Gamma\|^2_{L^2(\nu)}$$

We then consider the space $B^\Gamma_\nu(X, Y)$ of equivariant functions of finite energy. In this context Lemma B.19and Proposition B.20 above read:

Proposition B.25. *Let*[8] $H_n = H_{\mu^n}$. *Then*

$$E_\mu(H_n f) \le \frac{1}{2} \int_{\bar{X}} d\bar{\nu}(x) \int_X \left[d\mu^{n+1}(x \to x') - d\mu^n(x \to x')\right] d^2_Y(H_n f(x), f(x')),$$

and

$$\frac{1}{2} \int_{\bar{X}} d\bar{\nu}(x) \int_X d\mu^n(x \to x') d^2_Y(H_n f(x), f(x')) \le E_{\mu^n}(f).$$

Proof. Essentially the same as before, integrating $d\bar{\nu}$ instead of $d\nu$ since the integrands are Γ-invariant in all cases. □

Equation (B.3) has two more important implications:

Proposition B.26. *Let* $f \in B^\Gamma(X, Y)$. *Then*

$$d^2_{L^2(\nu)}(f, H_\mu f) \le 2E_\mu(f).$$

Proof. Set $y = f(x)$in the equation and ignore the second term on the RHS to get:

$$|df|^2_\mu(x) = \int_X d^2_Y(f(x), f(x')) d\mu(x \to x') \ge d^2_Y(f(x), H_\mu f(x)).$$

Integrating this $d\bar{\nu}(x)$ we obtain the desired result. □

Proposition B.27. *Let* $f \in B^\Gamma(X, Y)$. *Then* $E(H_\mu f) \le 2E(f)$.

Proof. By the triangle inequality and the inequality of the means,

$$\frac{1}{2} d^2_Y(H_\mu f(x), H_\mu f(x')) \le d^2_Y(H_\mu f(x), f(x')) + d^2_Y(f(x'), H_\mu f(x')).$$

Integrating $d\mu(x \to x')$ gives:

$$\frac{1}{2} |dH_\mu f|^2_\mu (x) \le |d' f|^2_\mu (x) + \int_X d_\mu(x \to x') d^2_Y(f(x'), H_\mu f(x')).$$

[8] In the Hilbertian case (but not in general) this equals the n-th power (iterate) of H_μ.

We now integrate $d\bar{\nu}(x)$. We evaluate the second term first using the ν-symmetry of μ:

$$\int_{\bar{X}} d\bar{\nu}(x) \int_X d_\mu(x \to x') d_Y^2(f(x'), H_\mu f(x')) = \int d\bar{\nu}(x') d_Y^2(f(x'), H_\mu f(x')),$$

so that:

$$E_\mu(H_\mu f) \leq \int_{\bar{X}} d\bar{\nu}(x) \left[\left| d' f \right|_\mu^2 (x) + d_Y^2(f(x), H_\mu f(x)) \right].$$

By Equation (B.3) we get:

$$E_\mu(H_\mu f) \leq \int_{\bar{X}} d\bar{\nu}(x) \int_X d\mu(x \to x') d_Y^2(f(x), f(x')) = 2E_\mu(f).$$

□

References

1. P. de la Harpe and A. Valette, *La propriété (T) de Kazhdan pour les groupes localement compacts*, Astérisque, no. 175, Société mathémathique de France, 1989.
2. M. Gromov, *Random walk in random groups*, Geom. Funct. Anal. **13** (2003), no. 1, 73–146.
3. D. Kazhdan, *Connection of the dual space of a group with the structure of its closed subgroups*, Func. Anal. Appl. **1** (1967), 63–65.
4. M. Ledoux, *Concentration of measure phenomenon*, Mathematical Surveys and Monographs, vol. 89, American Mathematical Society, 2001.