Ergodic theorems for noncommuting random products

Anders Karlsson

Abstract. Outline of lectures given at the meeting Probabilistic and Dynamical Properties of (semi)-group Actions, held at Universidad de Santiago de Chile, January 2008, and at Wrocław University as part of the EU Marie Curie Host Fellowship program Transfer of Knowledge, June 2008. The author is grateful to Andrés Navas and to Ewa Damek for their kind invitations to Santiago and to Wrocław, respectively.

Contents

1. Lecture 1: Introduction and Outline

1.1. The classical law of large numbers asserts that for i.i.d. random variables X_i with $E[|X_i|] < \infty$,

$$\frac{1}{n}(X_1 + X_2 + ... + X_n) \to E[X_1] \text{ a.s.}$$

(the strong version). The first version was proved by J. Bernoulli in his Ars Conjectandi from around 1700: X_i took values 0 or 1 and convergence in weak sense (convergence in probability). The modern version is due to efforts of Chebyshev, Markov, Borel, Cantelli, and Kolmogorov.

1.2. In the early 1930s von Neumann proved the first ergodic theorem (after Poincaré's recurrence theorem) and immediately afterwards Birkhoff proved his pointwise ergodic theorem. The set-up is: Let (Ω, μ) be a standard Borel measure space with $\mu(\Omega) = 1$ (from now on called a *probability space*), and $L : \Omega \to \Omega$ a measure preserving (i.e. $\mu(L^{-1}A) = \mu(A)$), ergodic (i.e. $\mu(L^{-1}A\Delta A) = 0$ implies $\mu(A) = 0$ or 1), measurable transformation (an *ergodic m.p.t).*

Birkhoff's theorem reads, given $g : \Omega \to \mathbb{R}$, L^1-integrable, i.e. $\int_\Omega |g|\, d\mu < \infty$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(L^k\omega) = \int_\Omega g d\mu \text{ a.e.}$$

(The mean ergodic theorem of von Neumann will be discussed later.)

1.3. The strong law of large number is a special case: $(\Omega, \mu) = (\mathbb{R}, \nu)^{\mathbb{N}}$ and transformation $L =$the shift, which is ergodic by Kolmogorov's 0-1 law. The map g is the projection onto the first coordinate, so that $X_i(\omega) = g(L^{i-1}\omega) = \omega_i$ where $\omega = (\omega_1, \omega_2, ...)$.

1.4. We are after a noncommutative version of the ergodic theorem, meaning that the function f now takes values in a more general group G instead of $\mathbb{R}$. We use the following notations. Let G be a topological group. Given a measurable map $g : \Omega \to G$, we let

$$Z_n(\omega) = g(\omega)g(L\omega)...g(L^{n-1}\omega)$$

which is the extension of $\sum g(L^k\omega)$ to a general, possibly noncommutative, group. Assuming some appropriate integrability, what can we say about Z_n as n tends to infinity? The first people to consider this type of questions were Bellman, Kesten, and Furstenberg from the 1950s and onward. As Grenander noted in his book from the time, it is not even clear how to formulate a possible extention. We will see one possible answer here. To fix terminology we call Z_n as above an *ergodic cocycle* in the general case, and in the special case of i.i.d variables, we call Z_n a *random walk* (defined by some probability measure ν on G).

1.5. One of the first studies for products of random matrices is a paper of Furstenberg-Kesten from 1960, where they in particular prove that if Z_n is an ergodic cocycle, integrable in the sense that $\int \log \|g\|\, d\mu < \infty$ for some matrix norm $\|\cdot\|$, then the following limit exists a.e.

$$\lim_{n\to\infty} \frac{1}{n} \log \|Z_n\| = A.$$

This was later generalized in two important ways: by Oseledec's multiplicative ergodic theorem, to be discussed later, and by Kingman's subadditive ergodic theorem.

1.6. In a paper published in 1968, Kingman in response to a question of Hammersley proved a subadditive ergodic theorem. This extension of the Birkhoff theorem is of basic importance in several contexts. Let $a(n,\omega)$ be a sequence of measurable functions such that a.e. for almost every $n, m \geq 0$

$$a(n+m,\omega) \leq a(n,\omega) + a(m, L^n\omega).$$

(For example, $a(n,\omega) = \sum_{k=0}^{n-1} f(L^k\omega)$, the additive case.) Then, provided that

$$\int_\Omega a(1,\omega)d\mu(\omega) < \infty,$$

the following limit exists a.e.

$$\lim_{n\to\infty} \frac{1}{n} a(n,\omega) = A.$$

The Furstenberg-Kesten theorem above follows when this is applied to

$$a(n,\omega) = \log \|Z_n(\omega)\|$$

in view of the submultiplicative property of matrix norms.

1.7. Another related implication of importance for us is the definition of the linear drift of Z_n. Fix a base point x_0 in X. Let

$$a(n,\omega) = d(Z_n(\omega)x_0, x_0).$$

From the triangle inequality and since G acts by isometry the subadditivity condition is easily verified:

$$\begin{aligned} d(Z_{n+m}(\omega)x_0, x_0) &\leq d(Z_{n+m}(\omega)x_0, Z_n(\omega)x_0) + d(Z_n(\omega)x_0, x_0) \\ &= d(Z_m(L^n\omega)x_0, x_0) + d(Z_n(\omega)x_0, x_0). \end{aligned}$$

(Note that for the subadditivity to hold the order in which the random elements are multiplied, from the right or from the left, does not matter.) Therefore, we have that if $\int_\Omega d(gx_0, x_0)d\mu < \infty$, then the *linear drift*

$$l = \lim_{n\to\infty} \frac{1}{n} d(Z_n(\omega)x_0, x_0)$$

exists a.e. and is a constant (because of the ergodicity assumption) depending on μ and g. Note that nothing depends on the choice of x_0. If we consider a random walk defined by ν, we sometimes write $l = l(\nu)$.

Example: Consider simple symmetric random walk on a regular tree. Calculate the drift and show the convergence in direction to an end. A main point here is to generalize this to a much more general setting.

1.8. Here is an outline of the main results. The first can be viewed as a noncommutative ergodic theorem (not to be confused with the subject of noncommutative geometry):

THEOREM 1 ([**KL1**]). *Let X be a proper metric space and Z_n an integrable ergodic cocycle taking values in a topological group G which acts by isometry on X. Then, for almost every ω, there is a horofunction $h = h_\omega$ depending measurably on ω such that*

$$\lim_{n\to\infty} -\frac{1}{n}h(Z_n x_0) = l$$

where $l := \lim_{n\to\infty} \frac{1}{n} d(Z_n x_0, x_0)$.

Remarks at this point (a more complete discussion later):

- When $G = GL(n,\mathbb{R})$ and $X = Pos(n,\mathbb{R}) = G/K$, this specializes to Oseledec's theorem, which is fundamental in the subject of differentiable dynamics, and hence extends Birkhoff's ergodic theorem and the law of large numbers.
- Previous work of Kaimanovich: symmetric spaces of nonpositive curvature and Gromov hyperbolic spaces (even for nonproper X).
- When X is uniformly convex, Busemann NPC (e.g. CAT(0)) this was proved by K.-Margulis in [**KM99**] (even for nonproper X)
- Interesting even when applied to $G = X = \mathbb{R}$ with various invariant metrics d.

1.9. From the proof of Theorem 1 with some additional arguments one can prove (note that horofunctions appear as a crucial concept in the proof)

THEOREM 2 ([**KL2**]). *Let G be a locally compact group with a left invariant proper metric d and ν a nondegenerate probability measure on G with finite first moment. Then, if the Poisson boundary is trivial, there is a 1-Lipschitz homomorphism $T : G \to \mathbb{R}$ such that for almost every trajectory Z_n of the random walk defined by ν, we have:*

$$\lim_{n\to\infty} \frac{1}{n}T(Z_n) = \int_G T(g)d\nu(g) = l(\nu).$$

This existence of a *drift homorphism* seemed to have been mostly an unnoticed phenomenon, except that in the work of Guivarc'h on random walks on amenable Lie groups, the existence of such homomorphism in this case was in fact established in the proofs, as he has informed us.

Immediately, one has

COROLLARY 1. *Let G be a locally compact group with a left invariant proper metric d and ν a nondegenerate centered probability measure on G with finite first moment. Then, if the drift $l(\nu) > 0$, there exist nonconstant bounded ν-harmonic functions.*

The special case of finitely supported and symmetric measures was proved by Varopoulos. His proof rests on estimates for n-step transition probabilities of symmetric Markov chains. To extend this to measures of infinite support was a problem of some significance because of the Furstenberg-Lyons-Sullivan's discretization procedure for Brownian motion. Indeed, using Corollary 1 we obtained:

THEOREM 3 ([**KL3**]). *Assume that (M, g) is a regular covering of a Riemannian manifold which has finite volume and bounded sectional curvatures. Then M is Liouville if, and only if,*

$$\lim_{t\to\infty} \frac{1}{t} d(x_0, B_t) = 0 \text{ a.s.}$$

where B_t is Brownian motion and d the Riemannian distance.

This will not be discussed in detail. In short, the "if" part was proved by Kaimanovich. The "only if" part uses Ballmann-Ledrappier's version of the FLS-discretization procedure and the observation of Babillot that the measure so obtained is symmetric.

Another consequence of Theorem 2:

COROLLARY 2. *Let G be a finitely generated torsion group of subexponential growth (e.g. Grigorchuk groups) and ν any probability measure of finite first moment. Then $l(\nu) = 0$.*

2. Lectures 2-3: Horofunctions

2.1. Horofunctions and horospheres first appeared in noneuclidean geometry and complex analysis in the unit disk or upper half plane. Consider the unit disk D in the complex plane. The metric with constant Gaussian curvature -1 is

$$ds = \frac{2\,|dz|}{1-|z|^2} \text{ or } d(0,z) = \log\frac{1+|z|}{1-|z|}.$$

It is called the Poincaré metric and for a point $\zeta \in \partial D$ one has the horofunction

$$h_\zeta(z) = \log\frac{|\zeta - z|^2}{1-|z|^2}.$$

These appear explicitly or implicitly in for example the Poisson formula, Eisenstein series, and the Wolff-Denjoy theorem.

An abstract definition was later introduced by Busemann who defines the *Busemann function* associated to a geodesic ray γ to be the function

$$h_\gamma(z) = \lim_{t\to\infty} d(\gamma(t), z) - t.$$

Note here that the limit indeed exists for any metric space, since the triangle inequality implies that the expression on the right is monotonically decreasing and bounded from below by $-d(z, \gamma(0))$. The convergence is moreover uniform if X is proper as can be seen from a 3ε-proof using the compactness of closed balls. Horoballs are sublevel sets of horofunctions $h(\cdot) \leq C$. In euclidean geometry horoballs are halfspaces and in the disk model of the hyperbolic plane horoballs, or horodisks in this case, are euclidean disks tangent to the boundary circle.

Busemann functions appear for example as a curcial ingredient in the proof the theorem of Cheeger-Gromoll that any manifold M of semipositive curvature which has a geodesic line must split isometrically as $M = M' \times \mathbb{R}$.

2.2. There is a more general definition probably first considered by Gromov around 1980. Namely, let X be a complete metric space and let $C(X)$ denote the space of continuous real functions on X equipped with topology of uniform convergence on bounded subsets. Fix a base point $x_0 \in X$. Consider now the map $\Phi : X \to C(X)$ defined by

$$\Phi : z \mapsto d(z, \cdot) - d(z, x_0).$$

(A related map was considered by Kuratowski and independently K. Kunugui in the 1930s.) We will sometimes denote by h_z the function $\Phi(z)$. Note that every h_z is 1-Lipschitz because

$$|h_z(x) - h_z(y)| = |d(z, x) - d(z, x_0) - d(z, y) + d(z, x_0)| \leq d(x, y)$$

which applied to $y = x_0$ gives

$$|h_z(x)| \leq d(x, x_0).$$

We have:

Proposition 1. *The map Φ is a continuous injection.*

Proof. For injectivity and partial continuity of the inverse, say $d(x, x_0) \geq d(y, x_0)$ and note that

$$h_y(x) - h_x(x) = d(y, x) - d(y, x_0) - d(x, x) + d(x, x_0) \geq d(y, x).$$

For continuity of Φ it suffices to see that

$$\begin{aligned} |h_x(\cdot) - h_y(\cdot)| &\leq |d(x, \cdot) - d(x, x_0) - d(y, \cdot) + d(y, x_0)| \\ &\leq |d(x, \cdot) - d(y, \cdot)| + |-d(x, x_0) + d(y, x_0)| \leq 2d(x, y). \end{aligned}$$

□

2.3. An exercise in [**BH99**] asks you to prove that Φ is a homeomorphism onto its image. In other words it remains to show that the map is an embedding. It may however happen that $x_n \to \infty$ in X but $h_{x_n} \to h_{x_0}$. Indeed, I learnt a counterexample from Bader: consider a graph with one central vertex x_0 from which countably many edges e_n are attached of corresponding length n. If we take as metric space just the (end) vertices x_n then this space is a proper metric space with exactly the nonembedding property. So for $n \neq m$, $d(x_n, x_m) = n + m$. (If we instead add also the edges to the metric space, then we have a geodesic, although nonproper space. In this case h_{x_n} does not converge in the topology chosen,pointwise the convergence is the same). On the other hand, if X is a proper geodesic space, then Φ is an embedding, see Ballmann's book. Indeed, given a sequence $x_n \to \infty$

(properness) and $y \in X$. Select z_n on the geodesic from x_0 to x_n for every large n on distance R from x, where $R >> d(y, x_0)$. Then $h_y(z) > 0$ but $h_{x_n}(z_n) < -R$. Taking a limit point z of z_n shows that any limit of h_{x_n} is distinct from h_y.

2.4. We go on to define, following Rieffel's terminology, the *metric bordification* of X by taking the closure of the image in $C(X)$. We use the notation

$$X \cup \partial X = \overline{X} := \overline{\Phi(X)}.$$

The points, or functions, in ∂X are called *horofunctions* with $h(x_0) = 0$. The sublevel sets $\{z : h_\xi(z) \leq C\}$ are called *horoballs centered at* ξ. The Lipschitz estimates above passes to the limits and hence holds for all functions in $\overline{\Phi(X)}$.

If the metric space is proper, $\overline{X}$ is (sequentially) compact by the Arzela-Ascoli theorem, in which case it is called the *metric compactification.* Note that $\overline{X}$ is metrizable, since X is proper (take supmetric on each ball and build a convergent sum of metric). Following [**BH99**] we prove:

LEMMA 1. *Let X be a complete metric space. Suppose $x_n \to h \in \partial X$. Then $x_n \to \infty$.*

PROOF. If not, we may assume that $d(x_n, x_0) \to R$. Given $\varepsilon > 0$ there is an N such that

$$|h(x) - d(x_n, x) + d(x_n, x_0)| < \varepsilon$$

for every $x \in B_{R+1}(x_0)$. Hence

$$|h(x_m) - d(x_n, x_m) + d(x_n, x_0)| < \varepsilon$$

for all n and m sufficiently large, and in particular

$$|h(x_n) - d(x_n, x_n) + d(x_n, x_0)| < \varepsilon$$

which means that $|h(x_n) - R| < 2\varepsilon$ for all large n. It follows from the previous inequality then that $d(x_n, x_m) < 3\varepsilon$. This means $x_n \to x \in X$ and so $h_x = h$ which contradicts $h \in \partial X$. □

The action of $\mathrm{Isom}(X)$ on X extends continuously to an action by homeomorphisms to the whole of $\overline{\Phi(X)}$ and is given by

$$g \cdot h(z) = h(g^{-1}z) - h(g^{-1}x_0).$$

(Write out what h_{gx} is for $x \in X$ and g an isometry.) A related remark is the fact that the metric bordification is independent of the base point x_0 up to homeomorphism, as can be seen from

$$h_x^{x_0}(\cdot) = d(x, \cdot) - d(x, x_0) - d(x, y_0) + d(x, y_0) = h_x^{y_0}(\cdot) - h_x^{y_0}(x_0).$$

(An alternative definition of the bordification is to map X into $C(X)/\mathbb{R}$ induced by $x \mapsto d(x, \cdot)$.)

2.5. We will apply the following observation.

PROPOSITION 2. *Assume that a group* Γ *acts on a proper metric space* X *and assume that* λ *is a* Γ*-inavriant measure on* ∂X. *Then*

$$T(g) := \int_{\partial G} h(gx_0) d\lambda(h)$$

is a homomorphism $G \to \mathbb{R}$.

Recall that the action of G on H is given by:

$$\gamma_1 \cdot h(\gamma_2) \ = \ h(\gamma_1^{-1}\gamma_2) - h(\gamma_1^{-1}),$$

so that:

$$\begin{aligned} T(gg') &= -\int h(g'^{-1}g^{-1})dm(h) \\ &= -\int g' \cdot h(g^{-1})dm - \int h(g'^{-1})dm \\ &= T(g) + T(g'), \end{aligned}$$

where we used the invariance of m in the last equation. Hence the map T defines a homomorphism

$$T : G \to \mathbb{R}.$$

Finally we note that T moreover is a 1-Lipschitz map:

$$\left|T(g) - T(g')\right| \le \int \left|h(g^{-1}) - h(g'^{-1})\right| dm(h) \le \int d(g,g')dm(h) = d(g,g').$$

2.6. We also record some related propositions.

LEMMA 2. *Let* G *be an infinite, finitely generated group and* d *a proper word metric. Then the function* $h \equiv 0$ *cannot belong to* ∂G *(nor to* $\overline{\Phi(G)}$*). In fact, any* $h \in \overline{\Phi(G)}$ *must be unbounded.*

PROOF. For any g the function $d(g,\cdot) - d(g,e)$ is clearly unbounded since G is infinite.

Given $h \in \partial G$. Suppose $h(\cdot) = \lim_{n\to\infty} d(g_n,\cdot) - d(g_n,e)$. Take a ball $B_r(e)$ of radius $r > 1$ around e. For large enough n, it holds that $d(g_n,\cdot) - d(g_n,e) = h(\cdot)$ on this ball in view of the convergence and that all functions in $\overline{\Phi(G)}$ are integer-valued. Connect e to g_n by a geodesic. This geodesic must intersect the ball in a point s of distance r from e. We have $h(s) = -r$. Note that r was arbitrary. □

By the same proof as above we have:

PROPOSITION 3. *Assume* m *is a probability measure on* ∂G *which is invariant under a subgroup* H *of* G. *Then*

$$T(g) := \int_{\partial G} h(g) dm(h)$$

is a homomorphism $H \to \mathbb{R}$.

Corollary 3. *Assume that G fixes a point h in ∂G. Then G surjects onto $\mathbb{Z}$.*

Proof. The nontriviality of the homomorphism comes from the lemma above that no h is identically 0. □

Corollary 4. *Assume that G has a finite orbit in ∂G. Then G has a finite index subgroup H which surjects onto $\mathbb{Z}$.*

Proof. Take as H the stabilizer of a point h on the finite orbit. We need to show that h does not vanish identically on H. Since the orbit is finite, the index of H is finite and there are hence G is partition in a finite set of coset Hg_i. Since all horofunctions on G are unbounded and the distance $d(Hg_i, H)$ is finite, the assertion follows. □

Corollary 5. *Suppose ∂G is countable. Then G has a finite index subgroup which surjects onto $\mathbb{Z}$.*

Proof. Take e.g. a uniform probability measure ν on the generators of G. Take a ν-stationary probability measure μ. Since the measure is finite it can only have a finite number of atoms of maximal mass. These must constitute a finite set in ∂G invariant under G. □

Walsh [**Wa**] has proved that for any finitely generated nilpotent group G and any word metric there is a finite G-orbit in ∂G. What about virtually nilpotent groups? What about polynomial growth? If the same statement would be true, it could provide an alternative approach to Gromov's celebrated polynomial growth theorem. (Lemma: Let G be a group of polynomial growth d and G' an infinite index finitely generated subgroup. Then the polynomial growth of $G\ '$ is at most $d-1$.)

2.7. The following is a basic general result on the iteration of semicontractions (i.e. 1-Lipschitz maps) in proper spaces. Let

$$l = \lim_{n\to\infty} \frac{1}{n} d(f^n(x_0), x_0)$$

which exists by subadditivity, similar to the linear drift of Z_n above.

Theorem 4 ([**Ka02**]). *Let (X, d) be a proper metric space and let f be a semicontraction. Then there is a function $h \in \overline{\Phi(X)}$ such that for all $k \geq 0$,*

$$h(f^k(x_0)) \leq -lk$$

and for any $x \in X$

$$\lim_{n\to\infty} -\frac{1}{n} h(f^n(x)) = l.$$

Proof. Define $a_n = d(x_0, f^n(x_0))$ and let ε_i be a sequence of positive numbers decreasing to 0. Let $b_n^i = a_n - (l - \varepsilon_i)n$. Then since b_n^i is unbounded in n we can find n_i such that $b_{n_i}^i > b_m^i$ for all $m < n_i$. By taking a subsequence we can in view of compactness assume that $f^{n_i}(x_0)$ converges to some point $h \in \overline{X}$. Then we have

$$\begin{aligned}-a_k \leq h(f^k(x_0)) &= \lim_{i\to\infty} d(f^{n_i}(x_0), f^k(x_0)) - d(f^{k_i}(x_0), x_0) \\ &\leq \liminf_{i\to\infty} a_{n_i-k} - a_{n_i} \\ &= \liminf_{i\to\infty} b^i_{n_i-k} + (l-\varepsilon_i)(n_i-k) - b^i_{n_i} - (l-\varepsilon_i)n_i \\ &\leq \liminf_{i\to\infty} -(l-\varepsilon_i)k = -lk.\end{aligned}$$

The proves the first part of the proposition and that the limit exists for $y = x_0$ in the second part. But since $d(f^k(x_0), f^k(y)) \leq d(x_0, y)$ and $|h(z) - h(w)| \leq d(z, w)$, the limit is the same for any y and the proposition is proved. □

Keep in mind that the base point x_0 is arbitrary. A nonserious, but still somewhat illuminating example: a rotation of the circle, rational or irrational. Holomorphic maps in one or several variables constitute a source of semicontractions.

2.8. Let X be a complete geodesic space. Consider the set $\partial_{ray}X$ of all geodesic rays starting from a point x_0. Quite generally, this set (or the equivalence classes of divergent rays) can be viewed as a boundary at infinity. The topology is given by uniform convergence on closed balls. For example, let γ_i be any sequence of rays starting at x_0 and assume that $\{\gamma_i(R)\}_{i=1}^{\infty}$ is a Cauchy sequence for every R. By the completeness of (X, d), we can for each R define $\gamma(R) = \lim \gamma_i(R)$. It is then immediate that γ is a ray starting at y and we say that γ_i converges to γ.

A complete metric space X is a *CAT(0)-space* or *nonpositively curved* if for any $x, y \in X$ there exists a point z such that

$$d(x,y)^2 + 4d(z,w)^2 \leq 2d(x,w)^2 + 2d(y,w)^2$$

holds for every $w \in X$. This inequality is called the *semiparallelogram law* motivated by the fact that in case of equality it is the usual parallelogram law for Hilbert spaces. Apart from euclidean spaces, other main examples are the classical hyperbolic spaces and $\mathrm{Pos}_N(\mathbb{R})$.

The latter space is the space of all positive definite symmetric $N \times N$ real matrices. It is an open set in the vector space of symmetric matrices Sym_N. The Riemannian metric is

$$< v, w >_p = tr(p^{-1}vp^{-1}w)$$

for $v, w \in Sym_N \simeq T_pPos_N$. This yields a distance in the usual way which becomes

$$d(I, p) = \sqrt{\sum (\log \lambda_i)^2}.$$

Analogous to the situation in Banach space theory one has for nice spaces that horofunctions (linear functionals) are in bijective correspondence with geodesic rays (vectors). It will turn out that two properties, uniform convexity (UC) and uniform smoothness (US) guarantees this "reflexivity".

The space X is *uniformly convex* if it is convex in the sense that any two points can be joined by a geodesic, and that there is a strictly decreasing continuous function g on [0,1] with $g(0) = 1$ and $g(1) = 0$, such that for any $x, y, w \in X$ and midpoint z of x and y,

$$\frac{d(z,w)}{R} \leq g\left(\frac{d(x,y)}{2R}\right),$$

where $R := \max\{d(x,w), d(y,w)\}$.

The space X is *uniformly smooth* if for any $\varepsilon > 0$ and r, there is a number $R > 0$ sucht that for any geodesic $\gamma(t)$ and $z \in B_r(\gamma(0))$ it holds that

$$d(z,\gamma(s)) + d(\gamma(s),\gamma(t)) \leq d(z,\gamma(t)) + \varepsilon$$

for all $s, t \geq R$. This condition immitates the corresponding standard notion for Banach spaces, namely that for any $\varepsilon > 0$ there should exists $\delta > 0$ such that for any unit vector x and y with $\|y\| < \delta$ that

$$\|x + y\| + \|x - y\| \leq 2 + \varepsilon \|y\|.$$

It is proved in [**BH99**] that every CAT(0)-space is uniformly smooth.

The purpose of this section is to prove:

THEOREM 5. *Assume that X is UC and US. Then the Busemann map*

$$h : \partial_{ray}X \to \partial X$$

$\gamma \mapsto h_\gamma$ is a homeomorphism and the corresponding bordifications are homeomorphic as well.

The proof will consist of two lemmas of independent interest as they clearly display the roles of UC and US.

LEMMA 3. *Assume that X is US and let x_n be a sequence such that for some geodesics $[x_0, x_n] \to \gamma \in \partial_{ray}X$. Then for some $h \in \partial X$ we have $h_{x_n} \to h$ in $C(X)$.*

PROOF. Given $z \in B_r(x_0)$ and $\varepsilon > 0$. Select R such that the inequality in the definition of US holds. Let $y_n = [x_0, x_n](R)$ pick N so that $d(y_n, \gamma(R)) < \varepsilon$ for all $n > N$. We have

$$0 \leq d(z, y_n) + d(y_n, x_n) - d(z, x_n) < \varepsilon$$

which means

$$|h_{y_n}(z) - h_{x_n}(z)| = |d(z, y_n) - R + d(y_n, x_n) - d(z, x_n) + R| < \varepsilon.$$

Since for all $n, m > N$ we have $d(y_n, y_m) < 2\varepsilon$ we finally get

$$|h_{x_m}(z) - h_{x_n}(z)| < 4\varepsilon.$$

This proves the lemma, since this shows that $h_{x_n} \to h$ in $C(X)$. □

Note that this in particular shows that the Busemann map is well defined (the only issue being the uniform convergence on compact sets). Next we have:

LEMMA 4. *Assume that X is UC and let x_n be a sequence such that $h_{x_n} \to h \in \partial X$. Then for some geodesic ray γ we have the geodesic segments $[x_0, x_n] \to \gamma$.*

PROOF. We know by a previous lemma that $d(x_0, x_n) \to \infty$. Given $\varepsilon > 0$ pick N such that

$$|h_{x_n}(z) - h(z)| < \varepsilon$$

for all $z \in B_R(x_0)$ and $n > N$. Let $y_n = [x_0, x_n](R)$ and note that $h_{x_n}(y_n) = -R$ for every n (sufficiently large). This implies that

$$|h_{x_m}(y_n) + R| < 2\varepsilon$$

which means

$$|d(x_m, y_n) - d(x_m, x_0) + R| = |d(x_m, y_n) - d(x_m, y_m)| < 2\varepsilon.$$

Now let z be the midpoint of y_n and y_m. Note first that by uniform convexity

$$d(x_0, x_m) \leq d(x_0, z) + d(z, x_m) \leq d(x_0, z) + d(y_m, x_m) + 2\varepsilon$$

so we have $d(x_0, z) \geq R - 2\varepsilon$. Therefore, by uniform convexity we get

$$R - 2\varepsilon < d(z, x_0) \leq g\left(\frac{d(y_m, y_n)}{2R}\right) R$$

which implies that

$$d(y_m, y_n) \leq 2Rg^{-1}(1 - 2\varepsilon/R).$$

Hence y_m converges and in turn that $[x_0, x_n](t)$ converges for every $t \leq R$. The limit is a geodesic ray γ and the lemma is proven. □

To prove the theorem, we first note that both the ray topology and the topology on the metric bordification are first countable: for each point we can pick consider neighborhoods picking rational ε and R. Therefore we can argue with sequences. The map is surjective becuase let $h_{x_n} \to h$ then by the second lemma $x_n \to \gamma$. Applying the frist lemma to $y_n = x_n$ if n odd and $\gamma(n)$ if n even, so y_n converges to γ. It follwos that $h = h_\gamma$. For injectivity assume $x_n \to \gamma$ and $y_n \to \xi$ with $h_{x_n} \to h$ and $h_{y_n} \to h$. Then by considering $z_n = x_n$ if n odd and y_n if n even we get from the second lemma that $\gamma = \xi$. The continuity of the Busemann map and its inverse is immediate from the lemmas.

2.9. In view of this we have that every horofunction h based at $x_0 = 0$ in $\mathbb{R}^N$ with the euclidean metric can be written as $h(\cdot) = - < v, \cdot >$ for some unit vector v.

In the prototypical symmetric space $Pos_N(\mathbb{R})$ every horofunction is given as follows (Proposition II.10.69 in [**BH99**]). Let $c(t) = \exp(tX)$ be the geodesic line defined by symmetric matrix X of norm 1. Let $F(c)$ be the union of the geodesic lines parallel to c (if c is regular this is the unique maximal flat containing c). There is a diffeomorphism $N_{[c]} \times F(c) \to Pos_N$ by $(v, p) \mapsto vpv^t$, where $N_{[c]}$ is the horospherical subgroup associated to $[c]$, which is the group which leaves the Busemann function defined by c invariant. (Similar to the Iwasawa decomposition.) The Busemann function defined by $c(t)$, $t \to \infty$, is given by

$$h(z) = -tr(XY),$$

where $z = vpv^t$ and $p = \exp(Y) \in F(c)$.

2.10. Let X be a Gromov hyperbolic space. In the case X is Gromov hyperbolic, it is known that for any two horofunctions h_1 and h_2 associated to sequences converging to ξ in the Gromov boundary there is a constant C such that $|h_1(z) - h_2(z)| < C$ for all $z \in X$, see [**BH99**], and for a boundary point ξ there may indeed be several such associated horofunctions. This shows that the h in the theorem is not necessarily unique (only up to suitable equivalence it is unique).Then there is a natural continuous surjection of ∂X onto the usual Gromov boundary $\partial_{hyp} X$ ($=\partial_{ray} X$ if X is geodesic), see [**BH99**]. Example: X=an infinite ladder.

2.11. There are studies of horofunctions in various other classes of metric spaces, such as abelian groups, Heisenberg groups, Banach spaces, and Hilbert's metric on convex sets, by e.g. Rieffel, Webster-Winchester, Walsh, Ledrappier-Lim, Karlsson-Metz-Noskov.

3. Lecture 4: Proof of the noncommutative ergodic theorem

3.1. Let (Ω, μ) be a pobability space and $L : \Omega \to \Omega$ an ergodic m.p.t. Let (X, d) be a proper metric space with a base point x_0. Assume that G is a group which act on X by isometry, $G \to Isom(X, d)$ (we supress the notation of this homomorphism). Given a measurable map $g : \Omega \to G$ (where the measurable structure on G comes from the action on X). We assume the integrability condition

$$\int_\Omega d(gx_0, x_0) d\mu < \infty.$$

Therefore the *linear drift*

$$l := \lim_{n\to\infty} \frac{1}{n} d(Z_n(\omega)x_0, x_0) = \inf_{n>0} \frac{1}{n} \int_\Omega d(Z_n(\omega)x_0, x_0) d\mu$$

exists a.e. and is a constant by the subadditive ergodic theorem. We consider the metric compactification $\overline{X} := \overline{\Phi(X)} \subset C(X)$ and the corresponding boundary $\partial X = \overline{X} - \Phi(X)$.

Recall also the statement of the noncommutative ergodic theorem in the present context:

THEOREM 6 ([**KL1**]). *Under the above assumptions there is an a.e. defined measurable map $\omega \mapsto h_\omega \in \partial X$ such that*

$$\lim_{n\to\infty} -\frac{1}{n} h(Z_n x_0) = l.$$

In this lecture we give a proof of this theorem. Note that we may assume $l > 0$ since otherwise the statement is trivial: any horofunction h would do. Indeed, when $l = 0$, for any h we have

$$\left| \frac{1}{n} h(Z_n x_0) \right| \leq \frac{1}{n} d(Z_n x_0, x_0) \to 0.$$

3.2. The cocycle F. Define for $g \in G$ and $h \in \overline{X}$, $F(g,h) = -h(g^{-1}x_0)$ and note the following cocycle relation

$$\begin{aligned} F(g_1, g_2 h) + F(g_2, h) &= -(g_2.h)(g_1^{-1}x_0) - h(g_2^{-1}x_0) \\ &= -h(g_2^{-1}g_1^{-1}x_0) + h(g_2^{-1}x_0) - h(g_2^{-1}x_0) \\ &= F(g_1 g_2, h). \end{aligned}$$

Note also that for any $g \in G$,

$$d(x_0, g x_0) = \max_{h \in \overline{X}} F(g, h),$$

since $F(g, \Phi(g^{-1}x_0)) = -d(g^{-1}x_0, g^{-1}x_0) + d(g^{-1}x_0, x_0) = d(g^{-1}x_0, x_0)$ and $-h(g^{-1}x_0) \leq d(g^{-1}x_0, x_0)$ for any $h \in \overline{X}$.

3.3. A skew product. Let $Z_n(\omega)$ be an integrable cocycle taking values in G defined by a map $g : \Omega \to G$. We define the skew product $\overline{L} : \Omega \times \overline{X} \to \Omega \times \overline{X}$ via

$$\overline{L}(\omega, h) = (L\omega, g(\omega)^{-1}h).$$

Let $\overline{F}(\omega, h) = F(g(\omega)^{-1}, h)$. For detailed information about skew-products that we will need here, see [**L98**, Ch. 1].

Consider the space $L^1(\Omega, C(\overline{X}))$ which is the space of (equivalence classes of) measurable maps $f : \Omega \to \overline{X}$ such that

$$\int_\Omega \sup_{h \in \overline{X}} |f(\omega)(h)| \, d\mu(\omega) < \infty.$$

Note that $\overline{F}$ is in $L^1(\Omega, C(\overline{X}))$ in view of $\left|F(g(\omega)^{-1}, h)\right| \leq d(x_0, g(\omega)x_0)$ and the basic integrability assumption on g.

3.4. A Birkhoff sum. Using the cocycle relation we have

$$\begin{aligned}\overline{F_n}(\omega,h) &:= \sum_{i=0}^{n-1}\overline{F}(\overline{L}^i(\omega,h))\\ &= F(g(\omega)^{-1},h)+F(g(L\omega)^{-1},g(\omega)^{-1}.h)+\\ &\dots+F(g(L^{n-1}(\omega)^{-1},g(L^{n-2}\omega)^{-1}\dots g(\omega)^{-1}.h)\\ &= F(Z_n(\omega)^{-1},h).\end{aligned}$$

By the subadditive ergodic theorem and from the remarks above,

$$\begin{aligned}0<l &= \inf_{n>0}\frac{1}{n}\int_\Omega d(Z_n(\omega)x_0,x_0)d\mu(\omega)\\ &= \inf_{n>0}\frac{1}{n}\int_\Omega \max_{h\in\overline{X}}\overline{F_n}(\omega,h)d\mu(\omega).\end{aligned}$$

3.5. A certain space of measures. Consider the space $\mathcal{P}_\mu(\Omega\times\overline{X})$ of probability measures ν on $\Omega\times\overline{X}$ which projects onto μ on Ω, i.e. $\nu(B\times\overline{X})=\mu(B)$ for any measurable set $B\subset\Omega$. The topology is the weak topology coming from the duality with $L^1(\Omega,C(\overline{X}))$ (see [**L98**, p. 27]). (See also this reference for a discussion of measurability of maps like $\overline{F}$.) In other words, $\mu_n\to\lambda$ if for every $f\in L^1(\Omega,C(\overline{X}))$,

$$\mu_n(f):=\int_{\Omega\times\overline{X}}f(\omega,h)d\mu_n\to\lambda(f).$$

Note that this space of measures is weakly sequential compact ([**L98**, p. 27]).

3.6. Construction of a good measure. Now, for each n choose a probability measure μ_n in $\mathcal{P}_\mu(\Omega\times\overline{X})$ such that

$$\frac{1}{n}\int_{\Omega\times\overline{X}}\overline{F_n}(\omega,h)d\mu_n(\omega,h)\geq l.$$

For example, the measures defined by $\mu_{n,\omega}=\delta_{\Phi(Z_n(\omega)x_0)}$ in the terminology of [**L98**, p. 22-25] (disintegration of measures) would do, so

$$\mu_n(A)=\int_\Omega \delta_{\Phi(Z_n(\omega)x_0)}(A_\omega)d\mu(\omega),$$

where the section $A_\omega=\{h:(\omega,h)\in A\}$.

Let

$$\eta_n=\frac{1}{n}\sum_{i=0}^{n-1}(\overline{L}^i)_*\mu_n$$

and let η be a weak limit of these measures, which is possible by the weak sequential compactness.

The measure η is an $\overline{L}$-invariant probability measure projecting onto μ and satisfying $\int \overline{F} d\eta = l$. Indeed, it is clearly a probability measure, projecting onto μ since each μ_n does, and the invariance is simple to check:

$$\begin{aligned}\eta(\overline{L}^{-1}A) &= \lim_{k\to\infty} \frac{1}{n_k} \sum_{i=0}^{n_k-1} \left(\overline{L}^i\right)_* \mu_{n_k}(\overline{L}^{-1}A) \\ &= \lim_{k\to\infty} \left(\frac{1}{n_k} \sum_{i=0}^{n_k-1} \left(\overline{L}^i\right)_* \mu_{n_k}(A) - \frac{1}{n_k}\left(\mu_{n_k}(A) - \mu_{n_k}(\overline{L}^{-n_k}A)\right)\right) \\ &= \eta(A).\end{aligned}$$

Moreover, it is set up by construction so that $\int \overline{F} d\eta_n \geq l$:

$$\begin{aligned}\int \overline{F} d\eta_n &= \frac{1}{n}\int \overline{F}(\omega,h)\sum_{i=0}^{n-1}(\overline{L}^i)_*\mu_n = \frac{1}{n}\int \sum_{i=0}^{n-1} \overline{F}(\overline{L}^i(\omega,h))d\mu_n \\ &= \frac{1}{n}\int_{\Omega\times\overline{X}} \overline{F_n}(\omega,h)d\mu_n(\omega,h) \geq l.\end{aligned}$$

and by the definition of weak limits this property passes to η as well. On the other hand

$$\begin{aligned}\frac{1}{n}\int_{\Omega\times\overline{X}} \overline{F_n}(\omega,h)d\mu_n(\omega,h) &\leq \frac{1}{n}\int_{\Omega\times\overline{X}} d(Z_n(\omega)x_0,x_0)d\mu_n(\omega,h) \\ &= \frac{1}{n}\int_{\Omega} d(Z_n(\omega)x_0,x_0)d\mu(\omega) \to l.\end{aligned}$$

Hence

$$\int \overline{F} d\eta = l.$$

3.7. Applying Birkhoff's theorem. The Birkhoff ergodic theorem implies that for (ω,h) in a set $E\subset\Omega\times\overline{X}$ of η-measure 1 it holds that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\overline{F}(\overline{L}^i(\omega,h)) = \int_{\Omega\times\overline{X}} \overline{F} d\eta = l.$$

Since η projects onto μ, we have that for μ-almost every ω that there is a non-empty set of h with the desired property.

3.8. Applying a measurable selection theorem. Finally we will appeal to a measurable selection theorem (of von Neumann in a version due to Aumann) to get a measurable section. There is a Polish topology on Ω compatible with the standard Borel structure and such that the projection $f:\Omega\times\overline{X}\to\Omega$ maps open sets to Borel sets, and the inverse image of each point in Ω is a closed subset. By regularity of η, we can find closed subsets of P with arbitrarily large measure. These subsets are Polish spaces for the induced topology and the restriction of f still satisfies the hypotheses of Theorem 3.4.1 in [**Av**], which then gives a (partially-defined) cross section. Putting them together yields a measurable, a.e. defined cross section $\omega\mapsto$

(ω, h_ω) with h_ω having the desired property and the theorem is proved. QED.

3.9. Questions: Does limits of this type exists also for other (or all) horofunction? (It is true for CAT(0)-spaces and Gromov hyperbolic spaces by geometric arguments.) What about general nonproper spaces? What about nonhomogeneuous spaces? The map $\omega \to h_\omega$ obtained is not guaranteed to be equivariant. Is it possible to achieve that by possible consider a quotient of ∂X and in such a way get a ν-boundary? When is this the Poisson boundary? (Most of these questions can be answered in the two good classes of spaces: CAT(0)-space and Gromov hyperbolic spaces.)

4. Lecture 5: Consequences and ray approximation

4.1. Let us begin by carefully see the case of Birkhoff's theorem. Here $G = \mathbb{R}$ acting by translations on $X = (\mathbb{R}, |\cdot|)$ via the map $g : \Omega \to \mathbb{R}$. There are two horofunctions $h_{+\infty}(x) = -x$ and $h_{-\infty}(x) = x$. Theorem 1 states that there is a horofunction h such that

$$\lim_{n\to\infty} -\frac{1}{n} h\left(\sum_{k=0}^{n-1} g(L^k\omega)\right) = l.$$

This is the pointwise convergence statement and the limit can be seen to be given by $l = \left|\int_\Omega g(x)d\mu(x)\right|$ and $h = h_{+\infty}$ if $\int_\Omega g(x)d\mu(x) > 0$ and $h = h_{-\infty}$ if $\int_\Omega g(x)d\mu(x) < 0$. (If $l = 0$, then either one works and the statement is trivial.)

4.2. The following follows [**KMo**]. Let $D : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be an increasing function, $D(t) \to \infty$ such that $D(0) = 0$ and $D(t)/t \to 0$ monotonically. From the inequality

$$\frac{1}{t+s} D(t+s) \leq \frac{1}{t} D(t)$$

we get the following subadditivity property

$$D(t+s) \leq D(t) + \frac{s}{t} D(t) = D(t) + \frac{D(t)/t}{D(s)/s} D(s) \leq D(t) + D(s).$$

From all these properties of D, it follows that $(\mathbb{R}, D(|\cdot|))$ is a proper metric space, and clearly invariant under translations.

Now we determine $\partial\mathbb{R}$ with respect to this metric. Wlog we may assume that $x_n \to \infty$. We claim that for any z

$$h(z) = \lim_{n\to\infty} D(x_n - z) - D(x_n) = 0.$$

Assume not. Then for some $s > 0$ and an infinite sequence of $t \to \infty$ that $D(t+s) - D(t) > c > 0$ (wlog). For such s, t we have and t large so that $D(t)/t < c$

$$\frac{D(t+s)}{t+s} \geq \frac{D(t) + sc}{t+s} \geq \frac{D(t) + \frac{D(t)}{t}s}{t+s} = \frac{D(t)}{t}$$

but this contradicts that $D(t)/t$ is strictly decreasing. Hence $\partial\mathbb{R} = \{h \equiv 0\}$.

4.3. Let us now apply Theorem 1, which turns out to have been proved earlier by Aaronson with a very different argument.

THEOREM 7. *Let $f : \Omega \to \mathbb{R}$ which is D-integrable, i.e.*

$$\int_\Omega D(|f|)d\mu < \infty.$$

Then

$$\lim_{n\to\infty} \frac{1}{n} D\left(\left|\sum_{k=0}^{n-1} f(L^k\omega)\right|\right) = 0.$$

PROOF. Since the only horofunction is the zero function, it is impossible in view of the noncommutative ergodic theorem that the drift l with respect to the D-metric is strictly positive. □

4.4. One can relax the conditions on D, not having to have $D(0) = 0$, and the condition $D(t)/t \to 0$ can be weakened to $d(t) = o(t)$ and $d(t+s) \leq d(t) + d(t)$. To see this define

$$D(t) = \sup\{d(ut)/u : u \geq 1\}.$$

Note that this has the required properties. Moreover

$$d(t) \leq D(t) \leq 2d(t),$$

since if $D(t) = d(tu)/u$, set $n = [u] + 1$ and then $D(u) \leq d(nt)/u \leq nd(t)/u \leq 2d(t)$. (See Aaronson-Weiss.) In this way we can apply the argument for any metric on $\mathbb{R}$.

4.5. From this one obtains as special case classical results like the one of Marcinkiewics-Zygmund:

COROLLARY 6. *Let $0 < p < 1$. If $f \in L^p$, then*

$$\lim_{n\to\infty} \frac{1}{n^{1/p}} \sum_{k=0}^{n-1} f(L^k\omega) = 0.$$

Such moment conditions arise naturally in probability theory. These results are essentially best possible (e.g. Sawyer). For the iid case converses also hold (M-Z, Feller). Another example

COROLLARY 7. *If f is log-integrable, then*

$$\lim_{n\to\infty} \left|\sum_{k=0}^{n-1} f(L^k\omega)\right|^{1/n} = 1.$$

4.6. By approvimating a subadditive by a additive cocycle in the obvious way, we can reduce the question for the drift on groups to the theorem above. For example: Let G be a finitely generated group with word metric $\|\cdot\|$acting on a metric space X such that the action is distorted in such away that

$$d(gx_0, x_0) = o(\|g\|).$$

(For example, certain groups acting properly by (affine) isometries on a Hilbert space have such properties, pull back metric to X.) Let Z_n be an integrable cocycle (wrt d as in Theorem 1), then

$$l = \lim_{n\to\infty} \frac{1}{n} d(Z_n x_0, x_0) = 0.$$

4.7. To further relate Theorem 1 we should first discuss the following notion introduced and studied by Kaimanovich. A sequence of points $\{x_n\}$ in a metric space is approximated by a geodesic ray γ if there is a constant $l \geq 0$ such that

$$\frac{1}{n} d(x_n, \gamma(nl)) \to 0 \text{ as } n \to \infty.$$

For any metric space and sequence, to be approximated by rays is a priori stronger than to be "horofunction regular" in the sense of Theorem 1:

Proposition 4. *Let x_n be a sequence of points in X and $l \geq 0$. Assume that there is a geodesic ray γ such that $d(x_n, \gamma(nl))/n \to 0$. Then*

$$\lim_{n\to\infty} -\frac{1}{n} b_\gamma(x_n) = l.$$

Proof. For any horofunction h it is true that $|h(x_n)| \leq d(x_n, x_0)$ from the triangle inequality. Since $d(x_n, x_0)/n \to l$, this implies that

$$\liminf_{n\to\infty} \frac{1}{n} b_\gamma(x_n) \geq -l.$$

On the other hand note that for $t > \ln$ we have $d(\gamma(t), x_n) \leq t - nl + d(\gamma(nl), x_n)$. Hence

$$b_\gamma(x_n) \leq -nl + d(\gamma(nl), x_n)$$

and the proposition follows upon dividing by n and taking the limit as $n \to \infty$. □

The converse is true for e.g. CAT(0)-spaces, see the next proposition, and Gromov hyperbolic spaces (exercise, or see [**KL4**]).

Proposition 5. *Let X be a CAT(0)-space and $\{x_n\}$ be a sequence of point such that for some horofunction $h = b_\gamma$ it holds $-h(x_n)/n \to l \geq 0$. Then*

$$\lim_{n\to\infty} \frac{1}{n} d(x_n, \gamma(nl)) = 0.$$

PROOF. Recall first that every horofunction is a Busemann function defined by a geodesic ray γ with $\gamma(0) = x_0$. Let $\hat{x}_n$ be the projection of x_n (i.e. the point closest) on the convex set being the image of γ. By the cosine law,

$$d(x_0, x_n)^2 \geq d(x_0, \hat{x}_n) + d(\hat{x}_n, x_n)^2 - 2d(x_0, \hat{x}_n)d(\hat{x}_n, x_n)\cos\alpha,$$

a property of projections ($\alpha \geq \pi/2$), the fact that horoballs are convex:

$$d(x_0, x_n)^2 \geq d(x_0, \hat{x}_n)^2 + d(\hat{x}_n, x_n)^2 \geq b_\gamma(x_n)^2 + d(\hat{x}_n, x_n)^2.$$

This implies that $d(\hat{x}_n, x) = o(n)$ and by the triangle inequality that

$$d(\gamma(nl), x_n) = o(n)$$

as desired. □

Kaimanovich (see [**Kai89**]) characterized sequences $\{x_n\}$ in symmetric spaces of nonpositive curvature (e.g. classical hyperbolic spaces and Pos_N) which are regular in the sense of ray approximation. In rank 1, such $\{x_n\}$ is approximated by a geodesic ray iff

$$d(x_n, x_{n+1}) = o(n) \text{ and } d(x_0, x_n) \sim nl.$$

(The conditions are easily seen to be necessary, for the sufficieny draw triangles and use that angles are exponentially small.) In the higher rank, one must in addition have that the A-component must stabilize. Note that these conditions hold a.e. for any integrable ergodic cocycle $x_n := Z_n x_0$ in view of Birkhoff's theorem. This implies Theorem 1 for these spaces in view of Proposition 4, and as pointed out in [**Kai89**] this is moreover equivalent to Oseledec's theorem, see the next paragraph. For general CAT(0)-spaces the theorem was established by Margulis and the author in [**KM99**], or more precisely an equivalent version of it in view of the above propositions. Note also that with the help of an idea of Delzant, Kaimanovich established ray approximation (and hence also Theorem 1 in this case) for Gromov hyperbolic spaces, see [**Kai00**].

4.8. Oseledec's multiplicative ergodic theorem which appeared in 1968 (a different form of this was also proved by Millionshchikov) asserts that for μ-a.e. ω the sequence $A(n,\omega) := Z_n(\omega)^{-1}$, an integrable ergodic cocycle of $N \times N$ invertible matrices, is *Lyapunov regular*, which by definition means that there is a constant s, a filtration of subspaces

$$\{0\} = V_0^\omega \subsetneq V_1^\omega \subsetneq \dots \subsetneq V_s^\omega = \mathbb{R}^N$$

and numbers $\lambda_1 < \lambda_2 < \dots < \lambda_s$ such that for any $v \in V_i^\omega \setminus V_{i-1}^\omega$,

$$\lim_{n\to\infty} \frac{1}{n} \log \|A(n,\omega)v\| = \lambda_i$$

and

$$\lim_{n\to\infty} \frac{1}{n} \log |\det A(n,\omega)| = \sum_{i=1}^{s} \lambda_i (\dim V_i^\omega - \dim V_{i-1}^\omega).$$

This is equivalent to the existence of a positive symmetric matrix $\Lambda = \Lambda(\omega)$ for which

$$\lim_{n\to\infty} \frac{1}{n} \log \left\| (\Lambda^n Z_n)^{\pm 1} \right\| \to 0$$

as was observed in [**Kai89**]. In words, this says that there exists an "average" matrix $\Lambda = \Lambda(\omega)$ whose powers approximate the random product $Z_n(\omega)$ similar to the classical law of large numbers except that Z_n is written multiplicatively.

This in turn is equivalent to the ray approximation property of $Z_n(\omega)x_0$ for a.e. ω.

4.9. Note that most of the discussed special cases of Theorem 1 are in fact ray approximation, which has a more intuitive appeal. One could ask whether one could prove ray approximation more generally. On the other hand Theorem 1 with its horofunction works when there are no geodesics, and its cocycle property is crucial also in the proof of Theorem 2.

5. Lecture 6: An application to random walks on groups

5.1. Let G be a locally compact, second countable group and ν a probability measure on G. It is natural to assume that the support of ν generates G as a group, in which case we refer to ν as *nondegenerate*. Let $f : G \to \mathbb{R}$ be a function such that $f(g\cdot)$ is in $L^1(G,\nu)$ for every $g \in G$. Then f is *ν-harmonic* if

$$f(g) = \int_G f(gh)d\nu(h)$$

for any $g \in G$. Constant functions are obviously ν-harmonic. The pair (G,ν) is called *Liouville* (or has trivial Poisson boundary) if every bounded ν-harmonic function is constant. For example abelian groups are Liouville for any measure.

5.2. Let G be a locally compact group and d a left invariant proper metric on G (it is assumed throughout that the topology generated by d coincides with the given one). When G is second countable such a metric always exists, see [**HP**]. Let ν be a probability measure on G of finite first moment, which means that

$$\int_G d(e,g)d\nu(g) < \infty.$$

Let Z_n denote trajectories of the corresponding random walk, that is,

$$Z_n = g_0 g_1 ... g_{n-1}$$

where g_i are independent random variables taking values in G with distribution ν.

Let ν be a probability measure on a topological group G. Assume G acts on a space K with measure η. The convolution measure on K is defined by

$$\nu * \eta(A) = \int_{g\in G} \eta(g^{-1}A) d\nu(g).$$

The measure η is called ν-*stationary* if $\nu * \eta = \eta$. Assume that there are no nonconstant bounded ν-harmonic functions. Then, as is well-known, η is in fact G-invariant. Indeed, given a continuous function f on K it follows from the stationarity relation that

$$F(g) := \int_K f(gz) d\eta(z) = \int_K f(z) d(g_*\eta)(z)$$

is a bounded ν-harmonic function, hence constant. Since this holds for all continuous functions f, we have that η must be invariant.

The probability distribution ν^{*n}, defined as the n-times convolution $\nu * \nu * ... * \nu$ of ν, is the distribution of Z_n, and the linear drift is, as above, a consequence of Kingman's theorem:

$$l(\nu) := \lim_{n\to\infty} \frac{1}{n} d(e, Z_n) = \inf_n \frac{1}{n} \int_G d(e,g) \nu^{*n}(g).$$

If f is a bounded harmonic function, then $f(Z_n)$ is a bounded martingale and therefore converges almost surely.

5.3. Assume that G is a finitely generated group. Let S be a symmetric finite generating set. The distance on G is the corresponding left invariant word metric $|\cdot|$. It is clearly a proper metric space. Let ν be a probability measure on G of finite first moment, which means that

$$\sum_{g\in G} |g| \, \nu(g) < \infty.$$

Define the *entropy* of ν by

$$H(\nu) := -\sum_{g\in G} \nu(g) \log(\nu(g)).$$

Recall that we have

$$H(\nu) := -\sum_{g\in G} \nu(g) \log(\nu(g)) \leq \log(2|S|) \sum_{g\in G} |g| \, \nu(g) + \log 2. \tag{5.1}$$

Indeed, let a_n be the number of group elements of wordlength n. Then $a_n \leq |S|^n$. Define a probability measure ν' on G by $\nu'(g) = 1/(2^{|g|+1} a_{|g|})$. Then,

$$H(\nu) - \sum_{g\in G} \nu(g) \log(2^{|g|+1} a_{|g|}) = -\sum_{g\in G} \nu(g) \log \frac{\nu(g)}{\nu'(g)} \leq 0,$$

where the inequality comes from Jensen's inequality (or $-\log t \leq 1/t - 1$) keeping in mind that both measures ν and ν' are probability measures.

The estimate (5.1) follows. By Kingman's subadditive ergodic theorem,the *entropy of the random walk*

$$h(\nu) := \lim_{n\to\infty} -\frac{1}{n}\log \nu^{*n}(Z_n) = \inf_n \frac{1}{n} H(\nu^{*n})$$

exists and is constant almost surely. The vanishing of the entropy is related to bounded harmonic functions; it is proved in [**Av**], [**D**] and [**KV**] that

$$h(\nu) = 0 \text{ if,and only if, } (G,\nu) \text{ is Liouville.}$$

. By applying (5.1) to ν^{*n}, dividing by n, and letting $n \to \infty$, we get:

$$h(\nu) \le \log(2|S|)l(\nu).$$

Finally, taking as generators the set S_k of elements with word length smaller than k, the new drift is not bigger than the old one divided by k. This shows that, for all k, $h(\nu) \le \log(2|S_k|)l(\nu)/k$. Letting $k \to \infty$ yields $h \le vl$, which is called *the fundamental inequality* in [**Ve**], and explains the role of subexponential growth in Corollary 11.

5.4. Our main result here is the following:

Theorem 8. *[Furstenberg-Khasminskii formula for the linear drift]. Let (G,ν) be a topological group with a nondegenerate probability measure of finite first moment, d a left-invariant proper metric, and let $\overline{G}$ be the metric compactification of (G,d). Then there exists a measure λ on $\overline{G}$ with the following properties:*

- *λ is ν-stationary, i.e. $\lambda = \int (g_*\lambda)d\nu(g)$ and*
- *$\ell(\nu) = \int h(g^{-1})d\lambda(h)d\nu(g)$.*

Moreover, if $\ell(\nu) > 0$, then λ is supported on ∂G.

Proof. In the proof of Theorem 1, we constructed a measure m on $\Omega \times \overline{G}$. The measure λ can be seen as the projection on $\overline{G}$ of m, but it turns out that the measure λ can be directly constructed. Let $(\Omega^+, \mathcal{A}^+, \mathbb{P})$ be the space of sequences $\{g_0, g_1, \dots\}$ with product topology, σ-algebra and measure $\mathbb{P} = \nu^{\otimes \mathbb{N}}$. For $n \ge 0$, let ν_n be the distribution of $Z_n(\omega)$ in $\overline{G}$. In other words, define, for any continuous function f on $\overline{G}$:

$$\int f d\nu_n \;=\; \int f(g_0 g_1 \cdots g_{n-1}) d\nu(g_0) d\nu(g_1) \cdots d\nu(g_{n-1}), \quad \nu_0 = \delta_e.$$

We claim that any weak* limit λ of the measures $\frac{1}{n}\sum_{i=0}^{n-1} \nu_i$ satisfies the conclusions of Theorem 8. Clearly, the measure λ is stationary: for any continuous function f on $\overline{G}$, we have (note that when $h \in \overline{G} - \partial G$ then it is just a group element)

$$\int f(g.h)d\mu(h)d\nu(g)$$

$$= \lim_{k\to\infty}\frac{1}{n_k}\sum_{i=0}^{n_k-1}\int f(gg_0g_1\cdots g_{i-1})d\nu(g_0)d\nu(g_1)\cdots d\nu(g_{i-1})d\nu(g)$$

$$= \lim_{k\to\infty}\frac{1}{n_k}\sum_{i=0}^{n_k-1}\int fd\nu_{i+1}$$

$$= \int fd\lambda + \lim_{k\to\infty}\frac{1}{n_k}[\int fd\nu_{n_k} - f(e)] \;=\; \int fd\lambda.$$

In the same way, we get:

$$\int h(g^{-1})d\lambda(h)d\nu(g)$$

$$= \lim_{k\to\infty}\frac{1}{n_k}\sum_{i=0}^{n_k-1}\int[d(Z_i(\omega),g^{-1})-d(Z_i(\omega),e)]d\mathbb{P}(\omega)d\nu(g)$$

Now note that $\int d(gZ_i,e)d\mathbb{P}d\nu \;=\; \int d(Z_{i+1},e)d\mathbb{P}$ because of i.i.d). This makes the sum into a telescoping sum and we obtain that

$$\int h(g^{-1})d\lambda(h)d\nu(g) = \lim_{k\to\infty}\frac{1}{n_k}\int d(Z_{n_k},e)-d(Z_0,e)d\mathbb{P}(\omega) \;=\; \ell(\nu).$$

This shows that the measure λ has the desired properties. Moreover, the measure $\mathbb{P}\times\lambda$ on the space $\Omega^+\times\overline{G}$ is $\overline{T}$-invariant. There is a unique $\overline{T}$-invariant measure m on $\Omega\times\overline{G}$ that extends $\mathbb{P}\times\lambda$. The measure m satisfies all the properties we needed in the proof of Theorem 1. In particular, if $\ell(\nu)$ is positive,

$$\mu(\partial G) \;=\; (\mathbb{P}\times\lambda)(\Omega\times\partial G) \;=\; m(\Omega\times\partial G) \;=\; 1.$$

For this notice that for $x\in G$, $h_x(Z_n) = d(x,Z_n)-d(x,x_0)\to+\infty$ (and not to $-\infty$) if $Z_n\to\infty$. □

6. Lecture 7: The Liouville property and drift homomorphisms

6.1. Varopoulos proved in 1985 that if a finitely generated group admits a symmetric (i.e. $\nu(g^{-1}) = \nu(g)$) finitely supported probability measure ν with positive drift $l(\nu) > 0$, then the group admits nontrivial bounded ν-harmonic functions. The symmetry condition cannot be removed: any measure on $\mathbb{Z}$ with finite $\sum_{x\in\mathbb{Z}} x\nu(x)\neq 0$ has positive drift. However, in a certain sense this is the only counterexample.

Theorem 9. *Let G be a locally compact group with a left invariant proper metric and ν be a nondegenerate probability measure on G with first moment. Then, if the Poisson boundary is trivial, there is a 1-Lipschitz* *homomorphism $T : G \to \mathbb{R}$ such that for almost every trajectory Z_n of the corresponding random walk, we have:*

$$\lim_{n\to\infty} \frac{1}{n} T(Z_n) \;=\; \int_G T(g) d\nu(g) \;=\; l(\nu).$$

PROOF. This follows from the Furstenberg-Khasminskii type formula above, the remark above that if (G, ν) is Liouville any ν-stationary measure must be G-invariant, and the proposition on construction homomorphism from an invariant measure on the metric boundary. □

A measure ν is *symmetric* if $d\nu(g^{-1}) = d\nu(g)$ for every $g \in G$. A measure is *centered* if every homomorphism of G into $\mathbb{R}$ is centered, meaning that the ν-weighted mean value of the image is 0 (cf. [**G**]). Every symmetric measure ν is centered, since for any homomorphism $T : G \to \mathbb{R}$, the mean value, which is

$$\int_G T(g) d\nu(g) = \int_G T(g^{-1}) d\nu(g) = -\int_G T(g) d\nu(g),$$

must hence equal 0. By simple contraposition, we get:

COROLLARY 8. *Let G be a locally compact group with a left invariant proper metric and ν be a nondegenerate centered probability measure on G with first moment. Then, if $l(\nu) > 0$, there exist nonconstant bounded ν-harmonic functions.*

Corollary 8 was proved by Varopoulos ([**Va**]) in the case ν is symmetric and of finite support on a finitely generated group. His proof rests on estimates for n-step transition probabilities of symmetric Markov chains. A simpler proof of the crucial estimate was given by Carne [**C**]. See also [**Al**] and [**M**] for interesting extensions. Note however that so far these estimates do not work for measures of infinite support. Measures with infinite support and finite first moment occur for example in the Furstenberg-Lyons-Sullivan discretization procedure of the Brownian motion, see [**KL3**].

Corollary 8 may also be compared with one of the main theorems in the paper [**G**] of Guivarc'h which states that for any connected amenable Lie group and any nondegenerate, centered measure ν with finite moments of all orders, the linear drift vanishes. The proof goes via a reduction to the case of connected, simply connected, nilpotent Lie groups. Guivarc'h pointed out to us that it is in fact proved in [**G**] that in the case of a connected amenable Lie group all the drift comes from an additive character (similar to Theorem 9 above). For finitely generated amenable groups this is no longer true: consider a simple symmetric random walk on the wreath product of $\mathbb{Z}^3$ with $\mathbb{Z}/2\mathbb{Z}$. This example has nontrivial bounded harmonic functions, hence the drift is positive, but all additive characters factor through $\mathbb{Z}^3$ and there the random walk moves sublinearly. In this discrete case, one should also mention the result of Kaimanovich ([**K**]) that when the group G is polycyclic and ν is centered, then the linear drift vanishes.

6.2. In the case when G is a finitely generated group, entropy theory (see below) yields a kind of converse to Corollary 8.

COROLLARY 9. *Let G be a finitely generated group and ν be a nondegenerate centered probability measure on G with first moment. Then $l(\nu) > 0$ if, and only if, there are nonconstant bounded ν-harmonic functions on G.*

PROOF. Any measure ν with first moment on G and with $l(\nu) = 0$ has $h(\nu) = 0$ and therefore only constant bounded ν-harmonic functions. So if there are nonconstant bounded harmonic functions, $l(\nu) > 0$. Otherwise, since ν is centered, $l(\nu) = 0$ by Corollary 8. This proves Corollary 9. □

Note that a measure may be centered for the simple reason that there are no nontrivial homomorphisms into $\mathbb{R}$, in this case Corollary 8 gives:

COROLLARY 10. *Let G be a locally compact group with a left invariant proper metric and ν be a nondegenerate probability measure on G with first moment. Assume that the only bounded ν-harmonic functions are the constants and that $H^1(G, \mathbb{R}) = 0$. Then $l(\nu) = 0$.*

The point here is that ν is not necessarily symmetric. Example: $\mathrm{Isom}(\mathbb{Z})$, the infinite dihedral group, vs $\mathbb{Z}$. It is remarkable that for a whole general class of groups, the nonexistence of homomorphisms can have such a strong influence on the drift; this is in great contrast with the case of nonamenable groups where no matter what, any nondegenerate measure of first moment must have positive drift (see [**G**]).

6.3. Recall that, if v be the volume growth rate of a finitely generated group G, $v \leq \log |S| < +\infty$, then it is a fact (see [**Av**], [**G**], [**Ve**] and below) that if $v = 0$, then the Poisson boundary is trivial. Examples of groups with subexponential growth ($v = 0$) and no nontrivial homomorphisms into the reals include the torsion groups with subexponential (but superpolynomial) growth constructed by Grigorchuk. We may formulate:

COROLLARY 11. *Let G be a group of subexponential growth generated by torsion elements, and ν any nondegenerate measure with first moment. Then $l(\nu) = 0$.*

Indeed, since a set of generators has finite order, any homomorphism $G \to \mathbb{R}$ vanishes and the statement follows from the volume criterion and Theorem 9.

See [**E**] for properties of measures without first moment, but with finite entropy, on such groups of subexponential growth.

6.4. Question: Let us call such a homomorphism as in the theorem for a *drift homomorphism.* When is the following converse true: assume (G, ν) admits such a homomorphism, then is it Liouville? Erschler indicated an example to me which seems to show that the converse is not always true.

7. Lecture 8: Extensions of von Neumann's ergodic theorem

7.1. Let (Ω, μ) be a probability space, L an ergodic m.p.t., and $f \in L^1(\Omega, \mathbb{R})$. One of the first basic questions in ergodic theory was to establish the convergence of the time averages to the space averages

$$\frac{1}{n}\sum_{k=0}^{n-1} f(L^k\omega) \to \int_\Omega f d\mu$$

which had come up in statistical mechanics. Around 1930, Koopman and later independently Weil suggested to von Neumann that it might be a useful to view the ergodic average as

$$\frac{1}{n}\sum_{k=0}^{n-1} U^k f$$

where U is the unitary operator of $L^2(\Omega)$ defined by $(Uf)(\omega) = f(L\omega)$. Challenged and inspired by this von Neumann indeed proved that the above convergence in L^2. Shortly afterwards Birkhoff (with a remark of Khintchine) proved the deeper convergence almost everywhere of the ergodic average. One further generalization in the von Neumann setting was that the isometry U can be replaced by any linear semicontraction, $\|U\| \leq 1$.

7.2. Notice that we can take one step further: Let ϕ be the affine semicontraction of $L^2 \to L^2$ defined by $\phi(g) = Ug + f$. Then

$$\frac{1}{n}\sum_{k=0}^{n-1} U^k f = \frac{1}{n}\phi^n(f).$$

Thus von Neumann's theorem can be formulated in terms of the behaviour of one semicontractions. This generalization was proven by Pazy (without commenting on that it generalizes von Neumann's theorem), namely that for any semicontraction ϕ of a Hilbert space and any vector f,

$$\frac{1}{n}\phi^n(f) \to v$$

in L^2 to some vector v (which Pazy gives a description of; it is the infimal vectorial displacement.) This can be generalized as follows (Pazy's case being $\Omega = \{\omega\}$ and X=a Hilbert space) in a theorem which already has been mentioned:

THEOREM 10 ([**KM99**]). *Assume that X is a complete metric space which is uniformly convex and Busemann nonpositively curved (e.g. CAT(0)-space or uniformly convex Banach space). Let $Z_n(\omega)$ be an integrable ergodic cocycle of semicontractions. Denote the drift by l. If $l > 0$ then there is a unique geodesic ray $\gamma_\omega(\cdot)$ in X with $\gamma_\omega(0) = x_0$ such that*

$$\lim_{n\to\infty} \frac{1}{n} d(\gamma_\omega(nl), Z_n(\omega)x_0) = 0.$$

7.3. Applications of Theorem 10:

- By the same remark as in the previous paragraphs on von Neumann's theorem, we recover a random mean ergodic theorem of Beck-Schwarz: Let U_ω be linear operators of a Hilbert space with $\|U_\omega\| \leq 1$. Then for any vector v, there is a vector $\hat{v}(\omega)$ a.e. such that
$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} U_\omega U_{L\omega}...U_{L^{k-1}\omega}v = \hat{v}(\omega)$$
strongly. (Note however that the original theorem was proved for Banach spaces more general than uniform convexity.)
- The special case of X a CAT(0)-space and $Z_n(\omega) = g^n$ was not known even for proper spaces. It gives convergence to a boundary point which then is a canonical fixed point for g.
- Another application is a version of Oseledec's theorem to infinite dimensions related to results of Ruelle.
- If $l > 0$, the map $\omega \mapsto \gamma_\omega$ induces a hitting meaure defined on ∂X. This is a ν-boundary which under a mild condition is isomorphic to the Poisson boundary. (Previous results by Kaimanovich and Ballmann-Ledrappier). Example: G a Coxeter group and X the associated Moussong-Davis complex.

7.4. One can also prove (see e.g. [**KL4**] for a discussion andreferences):

Theorem 11 (K). *Let X be a reflexive Banach space and $Z_n(\omega)$ an integrable ergodic cocycle taking values in semicontractions of X. Then for a.e. ω there is a linear functional f_ω of norm 1 such that*
$$\lim_{n\to\infty} \frac{1}{n} f_\omega(Z_n(\omega)0) = l.$$

Question: Is the theorem true without the assumption of X being reflexive?

Here is another question which I do not know if it is known and/or trivially true/false.

Question: Beardon showed that for any semicontraction ϕ of a Euclidean space there is either a fixed point or a linear functional such that $f(\phi^n 0) \geq 0$ for all $n \geq 1$. (A more general version was proved in a previous lecture). Is the same true for Hilbert spaces?

References

[L98] Arnold, L., *Random Dynamical Systems*, Springer Monographs in Mathematics, Springer-Verlag 1998

[B89] Ballmann, W., *On the Dirichlet problem at infinity for manifolds of nonpositive curvature*, Forum Math. 1 (1989) 201-213

[BGS] Ballmann, W., Gromov, M., Schroeder, V., *Manifolds of Nonpositive Curvature,* Progress in Math. vol. 61, Birkhäuser, Boston, 1985

[BH99] Bridson, M., Haefliger, A., *Metric spaces of non-positive curvature*, Grundlehren 319, Springer Verlag 1999

[CKW94] Cartwright, D. I., Kaimanovich, V. A., Woess, W., *Random walks on the affine group of local fields and of homogeneous trees*, Ann. de l'institut Fourier, 44 (1994), 1243-1288

[Gu 80] Guivarc'h, Y. *Sur la loi des grands nombres et le rayon spectral d'une marche aléatoire*, Astérisque 74 (1980) 47-98

[Kai89] Kaimanovich, V. A., *Lyapunov exponents, symmetric spaces and a multiplicative ergodic theorem for semisimple Lie groups*, J. Soviet Math. 47 (1989), no. 2, 2387–2398

[Kai00] Kaimanovich, V., *The Poisson formula for groups with hyperbolic properties*, Ann. Math. **152** (2000), no. 3, 659-692

[KM99] Karlsson, A., Margulis, G.A., *A multiplicative ergodic theorem and nonpositively curved spaces*, Comm. Math. Phys. **208** (1999) 107-123

[Ka04] Karlsson, A., *Linear rate of escape and convergence in directions*, In: Proceedings of a Workshop at the Schrödinger Institute, Vienna 2001, (Ed. by V.A. Kaimanovich, in collab. with K. Schmidt, W. Woess) de Gruyter, 2004

[KMN04] Karlsson, A., Metz, V., Noskov, G., *Horoballs in simplices and Minkowski spaces*, FG-Preprint, Bielefeld, 2004

[NW02] Nagnibeda, T., Woess, W., Random walks on trees with finitely many cone types. J. Theoret. Probab. 15 (2002), no. 2, 383–422

[R02] Rieffel, M., *Group C*-algebras as compact quantum metric spaces,* Doc. Math. 7 (2002), 605–651

[WW03] Webster, C., Winchester, A., *Busemann Points of Infinite Graphs,*Trans. Am. Math. Soc., vol. 358, no9, (2006), 4209-4224

[Al] G. Alexopoulos, On the mean distance of random walks on groups, *Bull. Sci. Math.* 111 (1987) 189–199

[Av] A. Avez, Entropie des groupes de type fini, *C.R.Acad.Sci. Paris Sér.A-B* 275 (1972) A1363–A1366

[C] T. K. Carne, A transmutation formula for Markov chains, *Bull. Sci. Math.* 109 (1985) 399–405

[D] Y. Derrienic, Quelques applications du théorème ergodique sous-additif, *Astérisque,* 74 (1980) 183–201

[E] A. Erschler, Boundary behaviour for groups of subexponential growth, *Ann. Math.* 160 (2004) 1183–1210

[G] Y. Guivarc'h, Sur la loi des grands nombres et le rayon spectral d'une marche aléatoire, *Astérisque,* 74 (1980) 47–98

[HP] U. Haagerup, A. Przybyszewska, Proper metrics on locally compact groups and proper affine isometric actions on Banach spaces, *Preprint* (2006)

[K] V. Kaimanovich, Poisson boundaries of random walks on discrete solvable groups, *in* Probability measures on groups, X (Oberwolfach, 1990), Plenum, New York, (1991) 205–238

[KV] V. Kaimanovich, A. Vershik, Random walks on discrete groups: boundary and entropy, *Ann. Prob.* 11 (1983) 457–490

[Ka02] Karlsson, A., Nonexpanding maps and Busemann functions, *Erg. Th. & Dyn. Sys.* 21 (2001) 1447-1457

[M] P. Mathieu, Carne-Varopoulos bounds for centered random walks, *Ann. Prob.* 34 (2006) 987–1011

[Va] N. Th. Varopoulos, Long range estimates for Markov chains, *Bull. Sci. Math.* 109 (1985) 225–252

[Ve] A. Vershik, Dynamic theory of growth in groups: entropy, boundary, examples, *Russian Math. Surveys* 55 (2000) 667–733.

[KL1] Karlsson, A., Ledrappier, F., On laws of large numbers for random walks, *Ann. of Prob.*, 34 (2006) 1693-1706

[KL2] Karlsson, A., Ledrappier, F., Linear drift and Poisson boundary for random walks, *Pure Applied Math. Quart.* 3 (2007) 1027-1036

[KL3] Karlsson, A., Ledrappier, F., Propriété de Liouville et vitesse de fuite du mouvement Brownien, *C. R. Acad. Sci. Paris Ser. I*, 344 (2007) 685-690

[KL4] Karlsson, A., Ledrappier, F., Noncommutative ergodic theorems, Preprint.

[KMo] Karlsson, A., Monod, N., Strong law of large numbers for concave moments, *Unpublished manuscript.*

[Wa] Walsh, C., The action of a nilpotent group on its horofunction boundary has finite orbits, preprint 2008.

KTH Stockholm, 100 44 Stockholm, Sweden

Exercises for "Ergodic theory of noncommuting random products"

1. Given a sequence of numbers a_n and a constant c such that

$$a_{n+m} \leq a_n + a_m + c$$

for all $n, m > 0$. Show that

$$\lim_{n\to\infty} \frac{1}{n} a_n = \inf_n \frac{1}{n} a_n.$$

2. Let (X, d) be a metric space and let $f : X \to X$ be 1-Lipschitz. Prove that

$$l := \lim_{n\to\infty} \frac{1}{n} d(f^n(x), x)$$

exists and is independent of x. Can you think of some other numbers, naturally and nontrivially, associated to f ?

3. Let G be the free group on two generators and consider a probability measure supported on the generators and their inverses. Compute the drift of the associated random walk.

4. Let X be a proper metric space with base point $x_0 \in X$, and consider the map

$$\Phi : x \longmapsto d(x, \cdot) - d(x, x_0)$$

of X into the space of continuous maps $C(X)$ equipped with the uniform convergence on compact sets. Take the closure $H := \overline{\Phi(X)}$. What is H if $X = \mathbb{R}$ with the standard metric, if $X = \mathbb{Z}^2$ with the word metric ("Manhattan metric"), and if $X = \mathbb{R}^2$ with the euclidean metric?

5. (Busemann) Let $\gamma : \mathbb{R}_{\geq 0} \to X$ be geodesic ray (i.e. γ is an isometry) in the proper metric space X. Show that

$$h_\gamma(y) = \lim_{t\to\infty} d(y, \gamma(t)) - t$$

is an element of $H := \overline{\Phi(X)}$.

6. Consider the additive group $\mathbb{Z}$ and a probability measure

$$\nu(x) = p\delta_{-1}(x) + (1-p)\delta_{+1}.$$

Determine all ν-harmonic functions on $\mathbb{Z}$.

7. Construct a bounded ν-harmonic function on the free group of two generators with ν being the uniform measure on the two generators and its inverses.

8. Let G be the infinite dihedral group, which can be realized as all automorphisms/isometries (not necessarily orientation preserving) of the simplicial graph $\mathbb{Z}$ (the standard Cayley graph of $\mathbb{Z}$). Show that there are no nontrivial homomorphisms of G into $\mathbb{R}$.

9. (Kuratowski; Kunugui, 1930s) Show that every metric space (X, d) is isometric to a subset of a Banach space (Hint: consider a close variant of the map Φ in the lectures: interchanging x and x_0, and $C(X)$ with sup-norm.)

10. A complete metric space X satisfies the *semiparallelogram law* if for any x, y there is a z such that

$$d(x,y)^2 + 4d(z,w)^2 \leq 2d(x,w)^2 + 2d(y,w)^2$$

holds for every $w \in X$.(This in fact equivalent to the usual CAT(0) definition.) Let M^n denote a Riemannian manifold. Assume that M is simply connected and has everywhere nonpositive sectional curvature $K \leq 0$; this is equivalent to that the exponential map

$$\exp_p : T_pM \to M,$$

which by definition maps lines to geodesics, semi-increases distances, i.e.

$$d(\exp_p(v), \exp_p(w)) \geq \|v - w\|_p$$

and preserves distances on the lines. Show that M satisfies the semiparallelogram law. (Hint: given a configuration x, y, z, w make a clever choice for p.)

11. In 1948, Busemann introduced and studied a weaker form of nonpositive curvature for metric spaces. Denote by $\frac{x+y}{2}$ a *midpoint* of x and y, i.e.

$$d(\frac{x+y}{2},x) = d(\frac{x+y}{2},y) = \frac{1}{2}d(x,y).$$

(The notation makes most sense when midpoints are unique, which they will be in the context of nonpositive curvature). A geodesic metric space X is said to be *Busemann NPC* if for any three points x, y and z the following inequality holds

$$d(\frac{x+y}{2},\frac{x+z}{2}) \leq \frac{1}{2}d(y,z).$$

(Note that this is intimately related to the property of $\exp_p$ in the previous exercise.) Prove that the semiparallelogram law implies Busemann NPC. (In view of the previous exercise, this means that we have more or less showed that for Riemannian manifolds that ($K \leq 0$ & simply connected) $\Longleftrightarrow$ (the semiparallelogram law)$\Longleftrightarrow$(Busemann NPC).)

12. Give an example of a space which is Busemann NPC but does not satisfy the semi-parallelogram law.

13. Prove Oseledec's multiplicative ergodic theorem in the deterministic case of itereates of one single invertible real $N \times N$ matrix A (surely known before Oseledec), i.e.that the sequence A^n is *Lyapunov regular*, which by definition means that there is a constant s, a filtration of subspaces

$$\{0\} = V_0 \subsetneq V_1 \subsetneq ... \subsetneq V_s = \mathbb{R}^N$$

and numbers $\lambda_1 < \lambda_2 < ... < \lambda_s$ such that for any $v \in V_i \setminus V_{i-1}$,

$$\lim_{n\to\infty} \frac{1}{n} \log \|A^n v\| = \lambda_i$$

and

$$\lim_{n\to\infty} \frac{1}{n} \log |\det A^n| = \sum_{i=1}^{s} \lambda_i (\dim V_i - \dim V_{i-1}).$$

(Hint: Jordan normal form).

14. Let x_n be a sequence of points in the hyperbolic plane such that $d(x_n, x_{n+1}) < C$ for a constant C and $d(x_0, x_n) = nl$ for some constant $l > 0$. Prove that the points converge to a point in the boundary circle (Hint: consider angles.) Is the same true in the euclidean plane?

15. Let (X, d) be a metric space. Assume that x_n is a sequence of points in X such that for some geodesic ray $\gamma : \mathbb{R}_{\geq 0} \to X$ with $\gamma(0) = x_0$ one has that
$$\frac{1}{n} d(x_n, \gamma(nl)) \to 0 \text{ as } n \to \infty.$$
Let h be the Busemann function associated to γ, so $h = b_\gamma$. Show that
$$-\frac{1}{n} h(x_n) \to l.$$

16. Let (X, d) be the Euclidean space of dimension N. Verify that for sequences of points x_n such that $d(x_n, x_0)/n \to l$, the following two statements are equivalent: 1. there exists a geodesic ray $\gamma : \mathbb{R}_{\geq 0} \to X$ with $\gamma(0) = x_0$ such that
$$\frac{1}{n} d(x_n, \gamma(nl)) \to 0 \text{ as } n \to \infty,$$
and 2. there is a horofunction h such that
$$-\frac{1}{n} h(x_n) \to l.$$

17. Given a group G with a probability measure ν. Assume that G acts on a countable space B with ν-stationary probability measure λ. Show that λ has finite support. Is λ necessarily G-invariant?

18. Given a finitely generated group G with a word metric d and a nondegenerate probability measure ν with $\nu(g) \in \mathbb{Q}$ for every $g \in G$. Assume that the metric boundary ∂G of (G, d) is countable (examples?). Prove that the linear drift $l(\nu)$ is a rational number.

19. Let G be the infinite dihedral group and d a word metric. Show that for any nondegenerate measure ν of finite first moment that the drift $l(\nu) = 0$.