Jørgensen's inequality for non-Archimedean metric spaces

J. Vernon Armitage and John R. Parker

Abstract. Jørgensen's inequality gives a necessary condition for a non-elementary group of Möbius transformations to be discrete. In this paper we generalise this to the case of groups of Möbius transformations of a non-Archimedean metric space. As an application, we give a version of Jørgensen's inequality for $\mathrm{SL}(2,\mathbb{Q}_p)$.

Dedicated to the memory of Alexander Reznikov

1. Introduction

In [6] Jørgensen proved a famous inequality giving a necessary condition for a non-elementary subgroup of $\mathrm{SL}(2,\mathbb{C})$ to be discrete. Intuitively, this inequality says that if two elements of $\mathrm{SL}(2,\mathbb{C})$ generate a non-elementary discrete group then they cannot both be very close to the identity. Jørgensen's theorem both makes this statement precise and gives explicit uniform bounds.

The methods used to prove this inequality have been generalised to a wide variety of different contexts but, generally, the statements look rather different from that given by Jørgensen. For example, a geometrical interpretation says there is always an embedded tubular neighbourhood of a very short geodesic in a hyperbolic manifold and that this neighbourhood, or "collar", has volume uniformly bounded away from zero. Hence handles in hyperbolic manifolds cannot be both short and thin.

In [7] Markham and Parker gave a general formulation of Jørgensen's inequality for Möbius transformations on a metric space which recovers many known versions as special cases. In these examples the one-point compactification of the metric space in question is the boundary of a rank one symmetric space of non-compact type, that is one of real, complex or quaternionic hyperbolic spaces or the octonionic hyperbolic plane. Additionally, this result applies when the metric space is a field, for example the p-adic numbers $\mathbb{Q}_p$ in which case $\mathrm{M\ddot{o}b}(\mathbb{Q}_p) = \mathrm{PSL}(2,\mathbb{Q}_p)$. In the main result of this paper, Theorem 3.1, we show that for non-Archimedean metric spaces one obtains a better inequality than Theorem 2.4 of [7]. In the case of $\mathbb{Q}_p$ this improved version of Jørgensen's inequality looks very similar to the original statement given by Jørgensen in [6]; see Theorem 4.2. We interpret this theorem geometrically in terms of the action of our group on an infinite, regular $p+1$ valent tree.

In the final section, we consider function field spaces. There is a strong analogy between these spaces and the p-adic numbers. It is possible to give a version of Theorem 4.2 in this case, but we leave details to the reader.

We would like to thank the referee for his/her valuable comments. Also, we would like to thank Guyan Robertson for his help, in particular for telling us about reference [3].

2. Non-Archimedean Möbius transformations

Let X be a non-empty set. A distance or *metric* on X is a function ρ from pairs of elements (x, y) to the real numbers satisfying:

(i) $\rho(x, y) \geq 0$ with equality if and only if $x = y$;
(ii) $\rho(x, y) = \rho(y, x)$;
(iii) $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ for all $z \in X$.

The inequality in (iii) is called the triangle inequality. A metric is said to be *non-Archimedean* if the triangle inequality is replaced with the following stronger inequality, called the *ultra-metric inequality*:

(iv) $\rho(x, y) \leq \max\{\rho(x, z),\ \rho(z, y)\}$ for all $z \in X$.

A simple consequence of the ultrametric inequality is the fact that every triangle in a non-Archimedean metric space is isosceles:

Lemma 2.1. *Suppose that ρ is a non-Archimedean metric on a space X. If x, y and z are points of X so that $\rho(x, y) < \rho(x, z)$ then $\rho(x, z) = \rho(y, z)$.*

Proof. We have

$$\rho(y, z) \leq \max\Big\{\rho(x, y),\ \rho(x, z)\Big\} = \rho(x, z)$$

by hypothesis. Likewise,

$$\rho(x, z) \leq \max\Big\{\rho(x, y),\ \rho(y, z)\Big\} = \rho(y, z)$$

since otherwise we would have $\rho(x, z) \leq \rho(x, y)$ which would be a contradiction. Therefore, we have

$$\rho(y, z) \leq \rho(x, z) \leq \rho(y, z)$$

and hence these quantities are equal. □

Many metrics arise from valuations on a ring. Let R denote a non-trivial ring. An *absolute value* (or *valuation* or *norm*) on R is a real valued function $x \longmapsto |x|$ on R satisfying:

(i) $|x| \geq 0$ with equality if and only if $x = 0$;
(ii) $|xy| = |x|\,|y|$;
(iii) $|x + y| \leq |x| + |y|$.

Once again, a valuation is said to be *non-Archimedean* if the inequality in (iii) is replaced with the stronger inequality:

(iv) $|x + y| \leq \max\{|x|,\ |y|\}$.

Given a valuation $|\ |$ on a ring R we may define a metric on R by:

$$\rho(x, y) = |x - y|.$$

Examples

(i) The standard absolute value on $\mathbb{R}$ or $\mathbb{C}$, which gives rise to the Euclidean metric.
(ii) Fix a prime number p and let $r \in \mathbb{Q}$ be non-zero. Write $r = p^f u/v$ where $f \in \mathbb{Z}$ and u, v are coprime integers both of which are also coprime to p. Then define a valuation $|\ |_p$ on $\mathbb{Q}$ by:

$$|r|_p = p^{-f}, \qquad |0|_p = 0. \tag{1}$$

One can then show that $|r + s|_p \leq \max\{|r|_p,\ |s|_p\}$. This valuation is called the *p-adic valuation*.

Let X be a complete non-Archimedean metric space with metric ρ. Following [7], we now define the Möbius transformations on X. Let $\mathrm{Aut}(X) \subset \mathrm{Isom}(X)$ be a group of isometries of X. This may be either the full isometry group or a sufficiently large subgroup that preserves some extra structure on X. We will suppose that $\mathrm{Aut}(X)$ acts transitively on X. The metric ρ induces a topology on X and we give $\mathrm{Aut}(X)$ the corresponding compact-open topology. Let o be a distinguished point of X. (Since $\mathrm{Aut}(X)$ acts transitively, in fact we may take o to be any point of X.) Suppose that the stabiliser of o in $\mathrm{Aut}(X)$ is compact. We make some more assumptions about X that allow us to extend $\mathrm{Aut}(X)$ to the group of Möbius transformations on X.

Given $d \in \mathbb{R}_+$, a *dilation* with *dilation factor* $d^2 \in \mathbb{R}_+$ is a map $D_d : X \longrightarrow X$ with $D_d o = o$ and for all z, $w \in X$ we have

$$\rho(D_d z, D_d w) = d^2 \rho(z, w). \tag{2}$$

(It may seem more natural to have taken d rather than d^2. However that would have introduced square roots into our formulae, such as (7) below.) Note that of $d \neq 1$ then D_d has a unique fixed point in X.

Let $X \cup \{\infty\}$ be the one point compactification of X. Suppose that there is an involution R interchanging o and ∞ and so that if z, $w \in X - \{o\}$ then

$$\rho(Rz, o) = \frac{1}{\rho(z, o)}, \tag{3}$$

$$\rho(Rz, Rw) = \frac{\rho(z, w)}{\rho(z, o)\rho(w, o)}. \tag{4}$$

We may think of R as reflection in the unit sphere of centre $o \in X$.

Let $\mathrm{M\ddot{o}b}(X)$ be the group generated by $\mathrm{Aut}(X)$, D_d and R for all d in some multiplicative subgroup of $\mathbb{R}_+$. We call $\mathrm{M\ddot{o}b}(X)$ the group of *Möbius transformations* of X. There is a natural topology on $X \cup \{\infty\}$ induced from the metric ρ (so neighbourhoods of ∞ are the exteriors of compact subsets of X). This enables us to define the compact-open topology for continuous functions from $X \cup \{\infty\}$ to itself. We will be interested in discrete subgroups of $\mathrm{M\ddot{o}b}(X)$ with respect to this topology.

Proposition 2.2 (Proposition 2.1 of [7]**).** *Let X be a metric space and* $\mathrm{M\ddot{o}b}(X)$ *be the group generated by* $\mathrm{Aut}(X)$, D_d *and R satisfying* (2), (3) *and* (4).

(i) *Let A be any element of* $\mathrm{M\ddot{o}b}(X)$ *for which $A\infty = \infty$. Then there exists a positive number d_A so that for all z, $w \in X$*

$$\rho(Az, Aw) = {d_A}^2 \rho(z, w).$$

(ii) *Let B be any element of* $\mathrm{M\ddot{o}b}(X)$ *for which $B\infty \neq \infty$. Then there exists a positive number r_B so that for all z, $w \in X - \{B^{-1}\infty\}$*

$$\rho(Bz, Bw) = \frac{{r_B}^2 \rho(z, w)}{\rho(z, B^{-1}\infty)\rho(w, B^{-1}\infty)},$$

$$\rho(Bz, B\infty) = \frac{{r_B}^2}{\rho(z, B^{-1}\infty)}.$$

The intuition behind Proposition 2.2(ii) is that B is like reflection in a sphere of radius r_B followed by an isometry. Also, we see that for all $B \in \mathrm{M\ddot{o}b}(X)$ with $B\infty \neq \infty$ we have

$$\frac{\rho(Bz, z)}{\rho(Bz, B\infty)} = \frac{\rho(z, B^{-1}z)}{\rho(B^{-1}z, B^{-1}\infty)}. \tag{5}$$

Lemma 2.3 (Lemma 2.2 of [7]**).** *Let X be a metric space. If* $\mathrm{Aut}(X)$ *acts transitively on X then* $\mathrm{M\ddot{o}b}(X)$ *acts 2-transitively on $X \cup \{\infty\}$. That is, given any two pairs x_1, y_1; x_2, y_2 of points in $X \cup \{\infty\}$ then there exists $B \in G$ so that $B(x_2) = x_1$ and $B(y_2) = y_1$.*

Lemma 2.4 (Lemma 2.3 of [7]). *Let X be a metric space. Suppose that $B \in \text{Möb}(X)$ fixes distinct points x, $y \in X \cup \{\infty\}$. Then B is conjugate to $A \in \text{Möb}(X)$ with fixed points o and ∞. Moreover, the dilation factor ${d_A}^2$ of A is independent of the conjugating map.*

Define the *cross-ratio* of quadruples of points in $X \cup \{\infty\}$ by

$$\begin{aligned} \mathbb{X}(z_1, z_2; w_1, w_2) &= \frac{\rho(w_1, z_1)\rho(w_2, z_2)}{\rho(w_2, z_1)\rho(w_1, z_2)}, \\ \mathbb{X}(z_1, \infty; w_1, w_2) &= \frac{\rho(w_1, z_1)}{\rho(w_2, z_1)}. \end{aligned}$$

Using Proposition 2.2 it is not hard to show that the cross-ratio of four points is preserved by the action of $\text{Möb}(X)$. Also, the cross-ratios satisfy the following property that resembles the ultrametric inequality

Proposition 2.5. *Let X be a non-Archimedean metric space. Let z_1, z_2, w_1, w_2 be four distinct points in $X \cup \{\infty\}$. Then*

$$\begin{aligned} \mathbb{X}(w_2, z_2; w_1, z_1) &\leq \max\Big\{1,\ \mathbb{X}(z_1, z_2; w_1, w_2)\Big\}, \\ 1 &\leq \max\Big\{\mathbb{X}(w_2, z_2; w_1, z_1),\ \mathbb{X}(z_1, z_2; w_1, w_2)\Big\}. \end{aligned}$$

Proof. When $z_2 = \infty$ we have

$$\mathbb{X}(z_1, \infty; w_1, w_2) = \frac{\rho(w_1, z_1)}{\rho(w_2, z_1)}, \quad \mathbb{X}(w_2, \infty; w_1, z_1) = \frac{\rho(w_1, w_2)}{\rho(w_2, z_1)}$$

and the result follows directly from

$$\rho(w_1, w_2) \leq \max\Big\{\rho(w_1, z_1),\ \rho(w_2, z_1)\Big\}, \quad \rho(w_2, z_1) \leq \max\Big\{\rho(w_1, z_1),\ \rho(w_1, w_2)\Big\}.$$

Now using the invariance of the cross-ratio under $\text{Möb}(X)$ we get the result for general quadruples of points. □

Let A be an element of $\text{Möb}(X)$ fixing x, $y \in X \cup \{\infty\}$ with dilation factor ${d_A}^2$ which may be 1 (see Lemma 2.4). Suppose that m_A is a positive number so that for all points $z \in X \cup \{\infty\} - \{x, y\}$ we have

$$\mathbb{X}(x, Az; y, z) \leq d_A m_A. \tag{6}$$

This is a conjugation invariant statement of the following inequality in the special case when $x = o$ and $y = \infty$:

$$\rho(z, Az) \leq d_A m_A \rho(o, z). \tag{7}$$

Observe that combining (7) with Proposition 2.2 gives

$$\rho(z, A^{-1}z) \leq {d_A}^{-1} m_A \rho(z, o)$$

and so $m_{A^{-1}} = m_A$. The number m_A gives a quantitative measure of how near A is to the identity: if A is close to the identity then the distance from z to Az should be small and hence m_A must be small. We remark that such an m_A always exists. For example using $Ao = o$ and the ultrametric inequality, we obtain

$$\rho(z, Az) \leq \max\big\{\rho(o, z),\ \rho(o, Az)\big\} = d_A \max\big\{d_A,\ 1/d_A\big\}\rho(o, z). \tag{8}$$

Thus one may always take $m_A = \max\{d_A,\ 1/d_A\} \geq 1$.

Lemma 2.6. *Let X be a non-Archimedean metric space. Suppose that $A \in \text{Möb}(X)$ is conjugate to a dilation with $d_A \neq 1$. If m_A is any positive number satisfying (6) then $m_A \geq \max\{d_A,\ 1/d_A\} > 1$.*

Proof. Assume A fixes o and ∞ and that m_A is any positive number satisfying (7). Since $d_A \neq 1$ then $\rho(z,o) \neq \rho(o,Az)$. Hence, using Lemma 2.1, we have equality in (8). In other words,

$$\rho(z,Az) = d_A \max\big\{d_A,\, 1/d_A\big\}\rho(o,z)$$

and so if m_A satisfies (7) we have $d_A m_A \rho(o,z) \geq \rho(z,Az) = d_A \max\big\{d_A,\, 1/d_A\big\}\rho(o,z)$ and $m_A \geq \max\{d_A,\, 1/d_A\} > 1$ as claimed. □

The intuition behind Lemma 2.6 is that, when $d_A \neq 1$, the map A is uniformly bounded away from the identity. For example, when A fixes o and ∞ we have

$$\rho(z,Az) = d_A \max\big\{d_A,\, 1/d_A\big\}\rho(z,o) \geq \rho(z,o)$$

and so $\rho(z,Az)$ is bounded below by a number depending on z but independent of A.

3. The main theorem

The main result of this paper is:

Theorem 3.1. *Let X be a complete non-Archimedean metric space and suppose that* $\mathrm{Aut}(X)$ *is a group of isometries of X that acts transitively on X with compact stabilisers. Suppose that* $\mathrm{Möb}(X)$*, the group of Möbius transformations on X, satisfies hypotheses* (2)*,* (3) *and* (4)*. Let A be an element of* $\mathrm{Möb}(X)$ *with exactly two fixed points, which we denote by x and y. Let m_A be a positive number satisfying* (6)*. If Γ is a discrete subgroup of* $\mathrm{Möb}(X)$ *containing A, then for all $B \in \Gamma$ so that $\{Bx,\, By\} \cap \{x,\, y\} = \emptyset$ we have*

$$m_A^2 \max\Big\{1,\, \mathbb{X}(Bx,y;x,By)\Big\} \geq 1. \tag{9}$$

Using Lemma 2.3, since $\mathrm{Aut}(X)$ acts transitively on X we see that $\mathrm{Möb}(X)$ acts 2-transitively on $X \cup \{\infty\}$. Thus, without loss of generality, in what follows we shall suppose that A fixes $x = o$ and $y = \infty$. Then the cross-ratio in (9) becomes:

$$\mathbb{X}(Bo,\infty;o,B\infty) = \frac{\rho(o,Bo)}{\rho(B\infty,Bo)}.$$

We now begin the proof of Theorem 3.1. This will broadly follow Section 2.3 of [7]. The main difference will come from the fact that we are working with a non-Archimedean metric. Our strategy is to assume that the hypothesis (9) fails. In particular, we must have $m_A < 1$ and so $d_A = 1$, using Lemma 2.6. (Recall, that as we saw above if $d_A \neq 1$ then A is uniformly bounded away from the identity in the sense that $\rho(z,Az) \geq \rho(z,o)$.) We construct a sequence B_n for $n = 0, 1, \ldots$ as follows. Let B_n be defined by $B_0 = B$ and $B_{n+1} = B_n A B_n^{-1}$. Let $x_n = B_n o$ and $y_n = B_n \infty$ be the fixed points of B_{n+1}. Let r_n denote r_{B_n}. We shall show that when the hypothesis (9) is not true then the B_n form a sequence of distinct elements of Γ that tend to the identity as n tends to infinity. This contradicts our hypothesis that Γ is discrete.

We begin by supposing that $x_n,\, y_n \notin \{o,\infty\}$ for all n. We then show that x_n tends to o and y_n tends to ∞ as n tends to infinity, Corollary 3.8. This immediately implies that the B_n are distinct.

Lemma 3.2. *Let A be a dilation fixing o and ∞ with $m_A < 1$. With the above notation*

$$\begin{aligned}
\rho(o,x_{n+1}) &\leq m_A \rho(x_{n+1},y_n)\,\frac{\rho(o,x_n)}{\rho(x_n,y_n)},\\
\frac{1}{\rho(x_{n+1},y_{n+1})} &\leq \frac{m_A}{\rho(x_{n+1},y_n)}\,\frac{\rho(o,y_n)}{\rho(x_n,y_n)}.
\end{aligned}$$

Proof. This follows the proof of Lemma 2.5 of [7]. Using Lemma 2.6, since $m_A < 1$ we have $d_A = 1$. Using Proposition 2.2 and (7) we have

$$\begin{aligned}
\rho(o, x_{n+1}) &= \rho(o, B_n A {B_n}^{-1} o) \\
&= \frac{{r_n}^2 \rho({B_n}^{-1} o, A{B_n}^{-1} o)}{\rho(A{B_n}^{-1} o, {B_n}^{-1}\infty)\rho({B_n}^{-1} o, {B_n}^{-1}\infty)} \\
&\leq \frac{m_A {r_n}^2 \rho(o, {B_n}^{-1} o)}{\rho(A{B_n}^{-1} o, {B_n}^{-1}\infty)\rho({B_n}^{-1} o, {B_n}^{-1}\infty)} \\
&= \frac{m_A \rho(B_n A B_n^{-1} o, B_n\infty)\rho(o, B_n o)}{\rho(B_n o, B_n \infty)} \\
&= m_A \rho(x_{n+1}, y_n) \frac{\rho(o, x_n)}{\rho(x_n, y_n)}.
\end{aligned}$$

We have used (5) on the penultimate line. Similarly, we have

$$\begin{aligned}
\frac{1}{\rho(x_{n+1}, y_{n+1})} &= \frac{1}{\rho(B_n A {B_n}^{-1} o, B_n A {B_n}^{-1}\infty)} \\
&= \frac{\rho(A{B_n}^{-1} o, {B_n}^{-1}\infty)\rho(A{B_n}^{-1}\infty, {B_n}^{-1}\infty)}{{r_n}^2 \rho(A{B_n}^{-1} o, A{B_n}^{-1}\infty)} \\
&\leq \frac{m_A \rho(A{B_n}^{-1} o, {B_n}^{-1}\infty)\rho(o, {B_n}^{-1}\infty)}{{r_n}^2 \rho({B_n}^{-1} o, {B_n}^{-1}\infty)} \\
&= \frac{m_A \rho(o, B_n\infty)}{\rho(B_n A B_n^{-1} o, B_n\infty)\rho(B_n o, B_n\infty)} \\
&= \frac{m_A}{\rho(x_{n+1}, y_n)} \frac{\rho(o, y_n)}{\rho(x_n, y_n)}.
\end{aligned}$$

□

Suppose that

$$\mathbb{X}_n = \mathbb{X}(B_n o, \infty; o, B_n\infty) = \frac{\rho(o, B_n o)}{\rho(B_n o, B_n\infty)} = \frac{\rho(o, x_n)}{\rho(x_n, y_n)}.$$

We may rewrite our hypothesis that (9) fails as

$$m_A < 1 \quad \text{and} \quad m_A^2 \mathbb{X}_0 < 1.$$

We shall show, first, that if the hypothesis (9) fails to hold then there is an $N \geq 1$ so that $\mathbb{X}_N \leq 1$ and, secondly, that this implies that $\mathbb{X}_n$ tends to zero as n tends to infinity.

Lemma 3.3. *Suppose that* $\mathbb{X}_n > 1$ *then* $\mathbb{X}_{n+1} \leq m_A^2 \mathbb{X}_n^2$.

Proof. Since $\mathbb{X}_n > 1$ we have $\rho(x_n, y_n) < \rho(o, x_n)$. Therefore, using Lemma 2.1, we see that $\rho(o, y_n) = \rho(o, x_n)$. This means that

$$\mathbb{X}_{n+1} = \frac{\rho(o, x_{n+1})}{\rho(x_{n+1}, y_{n+1})} \leq m_A^2 \frac{\rho(o, x_n)\rho(o, y_n)}{\rho(x_n, y_n)^2} = m_A^2 \left(\frac{\rho(o, x_n)}{\rho(x_n, y_n)} \right)^2 = m_A^2 \mathbb{X}_n^2.$$

□

Lemma 3.4. *Suppose that* $\mathbb{X}_n \leq 1$ *then* $\mathbb{X}_{n+1} \leq m_A^2 \mathbb{X}_n$.

Proof. If $\mathbb{X}_n \leq 1$ then $\rho(o,x_n) \leq \rho(x_n,y_n)$. Thus

$$\rho(o,y_n) \leq \max\Big\{\rho(o,x_n),\ \rho(x_n,y_n)\Big\} = \rho(x_n,y_n).$$

This means that

$$\mathbb{X}_{n+1} = \frac{\rho(o,x_{n+1})}{\rho(x_{n+1},y_{n+1})} \leq m_A^2 \frac{\rho(o,x_n)\rho(o,y_n)}{\rho(x_n,y_n)^2} \leq m_A^2 \frac{\rho(o,x_n)}{\rho(x_n,y_n)} = m_A^2 \mathbb{X}_n.$$

□

Lemma 3.5. *Suppose that $m_A^2\mathbb{X}_0 < 1$. Then there exists $N \geq 0$ so that $\mathbb{X}_N \leq 1$.*

Proof. If $\mathbb{X}_0 \leq 1$ then we choose $N = 0$. Suppose that $\mathbb{X}_k > 1$ for all $0 \leq k \leq n-1$. Then, using Lemma 3.3, we have

$$m_A^2\mathbb{X}_n \leq \big(m_A^2\mathbb{X}_{n-1}\big)^2 \leq \cdots \leq \big(m_A^2\mathbb{X}_0\big)^{2^n}.$$

Since $m_A^2\mathbb{X}_0 < 1$, this sequence is eventually at most m_A^2. Therefore we can only have $\mathbb{X}_n > 1$ for finitely many values of n. Hence there exists N with $\mathbb{X}_N \leq 1$. □

Lemma 3.6. *Suppose that $\mathbb{X}_N \leq 1$ then $\mathbb{X}_n \leq m_A^{2n-2N}$ for all $n \geq N$. In particular, if $m_A < 1$ then $\mathbb{X}_n$ tends to zero as n tends to infinity.*

Proof. We use induction. We have $\mathbb{X}_N \leq 1 = m_A^0$. Suppose that $\mathbb{X}_n \leq m_A^{2n-2N}$ for some $n \geq N$. Then, using Lemma 3.4, we have $\mathbb{X}_{n+1} \leq m_A^2\mathbb{X}_n \leq m_A^{2n+2-2N}$. The result follows. □

We now use the fact that $\mathbb{X}_n$ tends to zero as n tends to infinity to show that $\rho(o,x_n)$ tends to zero and $\rho(x_n,y_n)$ tends to infinity as n tends to infinity.

Lemma 3.7. *Suppose that $m_A < 1$ and $\mathbb{X}_N \leq 1$. Then for all $n \geq N$ we have*

$$\rho(o,x_n) \leq m_A^{n-N}\rho(o,x_N) \quad \text{and} \quad \frac{1}{\rho(x_n,y_n)} \leq \frac{m_A^{n-N}}{\rho(x_N,y_N)}.$$

In particular, $\rho(o,x_n)$ tends to zero and $\rho(x_n,y_n)$ tends to infinity as n tends to infinity.

Proof. Using Lemma 3.6, we see that $m_A\mathbb{X}_n \leq m_A^{2n+1-2N} \leq m_A < 1$ for all $n \geq N$. Thus

$$\rho(o,x_{n+1}) \leq m_A\mathbb{X}_n\rho(x_{n+1},y_n) < \rho(x_{n+1},y_n),$$

and so, using Lemma 2.1, we see that $\rho(o,y_n) = \rho(x_{n+1},y_n)$. As we already know that $\rho(o,y_n) \leq \rho(x_n,y_n)$, this means

$$\rho(x_{n+1},y_n) = \rho(o,y_n) \leq \rho(x_n,y_n).$$

Using Lemma 3.2, we have

$$\rho(o,x_{n+1}) \leq m_A\rho(o,x_n)\frac{\rho(x_{n+1},y_n)}{\rho(x_n,y_n)} \leq m_A\rho(o,x_n).$$

Using induction, we see that $\rho(o,x_n) \leq m_A^{n-N}\rho(o,x_N)$ as claimed.

Similarly, from the second part of Lemma 3.2, we have

$$\frac{1}{\rho(x_{n+1},y_{n+1})} \leq \frac{m_A}{\rho(x_{n+1},y_n)}\frac{\rho(o,y_n)}{\rho(x_n,y_n)} \leq \frac{m_A}{\rho(x_n,y_n)}.$$

Again, we use induction to get

$$\frac{1}{\rho(x_n,y_n)} \leq \frac{m_A^{n-N}}{\rho(x_N,y_N)}.$$

□

Corollary 3.8. *The points x_n tend to o and the points y_n tend to ∞ as n tends to infinity.*

We claim that the B_n lie in a compact subset of $\text{Möb}(X)$. Hence (a subsequence of) the B_n tend to the identity. Since the B_n are distinct, we see that $\langle A, B\rangle$ is not discrete. This will prove the main theorem in the case where x_n, $y_n \neq o, \infty$.

In order to verify the claim, observe that we may choose D_n lying in a compact subset of $\text{Möb}(X)$ so that $D_n B_n {D_n}^{-1}$ fixes both o and ∞. Secondly, since B_n is conjugate to A, using Lemma 2.4 we see that the dilation factor of $D_n B_n {D_n}^{-1}$ is 1. Thus for all z, $w \in X$ we have

$$\rho(D_n B_n {D_n}^{-1} A^{-1} z, D_n B_n {D_n}^{-1} A^{-1} w) = \rho(A^{-1} z, A^{-1} w) = \rho(z, w).$$

Hence $D_n B_n {D_n}^{-1} A^{-1}$ is in $\text{Aut}(X)$ and fixes o. By hypothesis the stabiliser of o in $\text{Aut}(X)$ is compact. Hence B_n lies in a compact subset of $\text{Möb}(X)$ as claimed.

We need to treat the case where there is an $N \geq 0$ for which either x_N or y_N is o or ∞, and so $x_{N+1} = o$ or $y_{N+1} = \infty$. Without loss of generality, suppose $y_{N+1} = \infty$ and hence $y_n = \infty$ for all $n \geq N + 1$.

Suppose $x_n \neq o$ for all n. We will not use (9) but only the fact that $\langle A, B\rangle$ is discrete. (Note that taking $N = 0$ this shows that if $\langle A, B\rangle$ is discrete then $\{Bo, B\infty\} \cap \{o, \infty\}$ cannot be just one point.) Consider the sequence B_n as defined above. By construction, B_n is conjugate to A and fixes ∞ for $n \geq N + 1$ and so $d_{B_n} = d_A = 1$. In other words, B_n is an isometry of X for $n \geq N + 1$. Hence for $n \geq N + 1$ we have

$$\begin{aligned}
\rho(x_{n+1}, o) &= \rho(B_n A {B_n}^{-1} o, o) \\
&= \rho(A {B_n}^{-1} o, {B_n}^{-1} o) \\
&\leq m_A \rho({B_n}^{-1} o, o) \\
&= m_A \rho(B_n o, o) \\
&= m_A \rho(x_n, o) \\
&\leq m_A^{n-N} \rho(x_{N+1}, o).
\end{aligned}$$

Therefore x_n tends to o as n tends to infinity and, arguing as above, B_n is a sequence of distinct elements of $\langle A, B\rangle$ converging to the identity. Again, $\langle A, B\rangle$ cannot be discrete.

Finally, suppose $x_{N+1} = o$ and $y_{N+1} = \infty$ for some $N \geq 0$. Thus B_{n+1} fixes both o and ∞ for all $n \geq N + 1$. Again we will not use (9), but this time we only use the fact that $\{Bo, B\infty\} \cap \{o, \infty\} = \emptyset$. Since A has precisely two fixed points, if $B_{n+1} = B_n A {B_n}^{-1}$ fixes both o and ∞ then B_n either fixes both o and ∞ or interchanges them. Without loss of generality, suppose that N is the smallest index for which $x_{N+1} = o$ and $y_{N+1} = \infty$. Since $\{B_0 o, B_0 \infty\} \cap \{o, \infty\} = \emptyset$, we may assume that $N \geq 1$. Then $B_N o = \infty$ and $B_N \infty = o$ and we see that B_N has an orbit of size 2. Thus B_N^2 fixes points that B_N does not. Since B_N is conjugate to A, this is a contradiction. This proves the theorem.

4. The p-adic numbers

In this section we consider the case where $X = \mathbb{Q}_p$, the p-adic numbers, that is, the completion of $\mathbb{Q}$ with respect to the p-adic valuation (1). We show that $\text{Möb}(X)$ is then the matrix group $\text{PSL}(2, \mathbb{Q}_p) = \text{SL}(2, \mathbb{Q}_p)/\{\pm I\}$ acting on $\mathbb{Q}_p \cup \{\infty\}$ by Möbius transformations. Discrete subgroups of $\text{SL}(2, \mathbb{Q}_p)$ have been considered by Ihara [5] and Serre in Chapter II.1 of [8], in particular page 84. Our main theorem gives a necessary condition for a subgroup of $\text{SL}(2, \mathbb{Q}_p)$ to be discrete, Theorem 4.2. This is very similar to the standard version of Jørgensen's inequality, [6]. In [4] Gromov and Schoen considered more general p-adic representations of lattices in non-compact, semisimple Lie groups. Our main result should apply in many of these cases.

The construction of the p-adic numbers and their properties in terms of non-Archimedean spaces is well known; see Artin [1], Cassels [2], and Serre [8], for example. We recall that a *p-adic integer* is any p-adic number α with $|\alpha|_p \le 1$. Thus, the ring of p-adic integers, denoted $\mathbb{Z}_p$, is the p-adic unit ball in $\mathbb{Q}_p$. Each p-adic integer α has an expansion

$$\alpha = \sum_{n=0}^{\infty} a_n p^n \tag{10}$$

where $a_n \in \{0, 1, \ldots, p-1\}$ and so $\mathbb{Z}_p$ is compact; see Lemma 2 on page 10 of Cassels [2]. Likewise, a *p-adic unit* is any $u \in \mathbb{Q}_p$ so that $u \in \mathbb{Z}_p$ and $u^{-1} \in \mathbb{Z}_p$. That is, u has the form (10) with $a_0 \neq 0$. Since the set of units is the intersection of two compact subsets of $\mathbb{Q}_p$, we see that it is compact.

We now show how to define a tree T whose boundary is $\mathbb{Q}_p \cup \{\infty\}$. This idea is due to Serre, [8], but our treatment will follow Figà-Talamanca [3]. The closed balls in $\mathbb{Q}_p$ are the vertices of T, that is

$$V = \left\{ x + p^k\mathbb{Z}_p \ : x \in \mathbb{Q}_p,\ k \in \mathbb{Z} \right\}.$$

Two vertices $x + p^k\mathbb{Z}_p$ and $y + p^j\mathbb{Z}_p$ are joined by an edge of T if and only if either $k = j+1$ and $x - y \in p^j\mathbb{Z}_p$ or else $j = k+1$ and $x - y \in p^k\mathbb{Z}_p$; see page 8 of [3]. In other words, $|j-k| = 1$ and one of the balls is contained in the other. Notice that each ball $x + p^k\mathbb{Z}_p$ of radius p^{-k} is contained in exactly one ball $x + p^{k-1}\mathbb{Z}_p$ of radius p^{-k+1} and contains exactly p balls $x + yp^k + p^{k+1}\mathbb{Z}_p$ of radius p^{-k-1} where $y = 0, 1, \ldots, p-1$. Hence each vertex has exactly $p+1$ edges emanating from it. Therefore the graph T we have just constructed is an infinite, regular $p+1$ tree.

We now find the boundary of T; see [3]. We consider geodesic paths through T. In other words, such a path is a (possibly infinite) sequence of vertices v_j so that for all j the vertices v_j, v_{j+1} are joined by an edge and $v_{j-1} \neq v_{j+1}$, that is there is no back tracking. The semi-infinite geodesic path $p^{-k}\mathbb{Z}_p$ for $k = 0, 1, 2, \ldots$ identifies a point of the boundary denoted by ∞. Every other semi-infinite geodesic path starting at the vertex $\mathbb{Z}_p$ eventually consists of a sequence of nested, decreasing balls $x + p^k\mathbb{Z}_p$ for $k = K, K+1, K+2, \ldots$. The limit of this sequence is the point x of $\mathbb{Q}_p$. Choosing a starting point other than $\mathbb{Z}_p$ makes only finitely many changes to these paths. Hence the boundary of T is $\mathbb{Q}_p \cup \{\infty\}$.

Any two distinct points z, w in $\mathbb{Q}_p \cup \{\infty\}$ are the end points of a unique doubly infinite geodesic path through T. We denote this path by $\gamma(z, w)$. The cross-ratio $\mathbb{X}(z_1, z_2; w_1, w_2)$ has the following interpretation in terms of T.

Lemma 4.1. *Suppose that z_1, z_2, w_1, w_2 are four distinct points of $\mathbb{Q}_p \cup \{\infty\}$. Let $\gamma(z_1, w_2)$ and $\gamma(z_2, w_1)$ be the geodesics joining z_1, w_2 and z_2, w_1. If $\mathbb{X}(z_1, z_2; w_1, w_2) = p^d > 1$. Then the shortest path in T from $\gamma(z_1, w_2)$ to $\gamma(z_2, w_1)$ has d edges. If $\mathbb{X}(z_1, z_2; w_1, w_2) \le 1$ then $\gamma(z_1, w_2)$ and $\gamma(z_2, w_1)$ intersect.*

Proof. Without loss of generality we suppose that $w_1 = o$ and $z_2 = \infty$. Then we have $\mathbb{X}(x, \infty; o, y) = \rho(o, x)/\rho(x, y)$. The geodesic $\gamma(o, \infty)$ passes through vertices $p^j\mathbb{Z}_p$ for $j \in \mathbb{Z}$.

Suppose first that the first few terms in the expansion of x and y are the same. In other words, we have $x = p^j(a + bp^k)$ and $y = p^j(a + cp^k)$ where $k > 0$ and a, b, c are units with $b \neq c$. Then $\rho(o, x) = p^{-j}$ and $\rho(x, y) = p^{-j-k}$ and thus we have $\mathbb{X}(x, \infty; o, y) = p^k > 1$. Every vertex on the geodesic $\gamma(x, y)$ has the form $p^j(a + bp^k + p^l\mathbb{Z}_p)$ or $p^j(a + cp^k + p^l\mathbb{Z}_p)$ where $l \ge k > 0$. The points of $\gamma(o, \infty)$ and $\gamma(x, y)$ closest to each other are $p^j\mathbb{Z}_p$ and $p^j(a + p^k\mathbb{Z}_p)$. The geodesic segment joining them has k edges and passes through the $k+1$ vertices $p^j(a + p^l\mathbb{Z}_p)$ for $l = 0, 1, \ldots, k$. This proves the first part of the lemma.

Suppose now that the first few terms of z_1 and w_2 are not the same. That is, we have $z_1 = ap^j$ and $w_2 = bp^k$ where a and b are units and either $j \neq k$ or, if $j = k$, then $a - b$ is a unit. Then $\rho(o, z_1) = p^{-j}$, $\rho(w_2, z_1) = p^{-\min(j,k)}$ and $\mathbb{X}(z_1, \infty; o, w_2) = p^{\min(0,k-j)} \le 1$.

Then the geodesic joining z_1 and z_2 passes through $p^j\mathbb{Z}_k$, which also lies on the geodesic joining o and ∞. $\square$

We claim that $\text{Möb}(\mathbb{Q}_p)$ is $\text{PSL}(2,\mathbb{Q}_p)$ acting on $\mathbb{Q}_p$ via Möbius transformations. Let $\text{Aut}(\mathbb{Q}_p)$ be the collection of maps $Ax = (ax+b)a$ where a is a unit in $\mathbb{Q}_p$ and b is any element of $\mathbb{Q}_p$. For $d = p^{-m}$ the dilation D_d is defined by $D_d x = p^{2m}x$ and satisfies (2):

$$\rho(D_d x, D_d y) = |p^{2m}z - p^{2m}w|_p = |p^{2m}|_p|z-w|_p = d^2|z-w|_p = d^2\rho(z,w).$$

The inversion R is given by $Rx = -1/x$ and clearly satisfies (3) and (4):

$$\begin{aligned}\rho(Rx,o) &= \left|\frac{-1}{x}\right|_p = \frac{1}{|x|_p} = \frac{1}{\rho(x,o)},\\ \rho(Rx,Ry) &= \left|\frac{-1}{x} - \frac{-1}{y}\right|_p = \frac{|x-y|_p}{|x|_p|y|_p} = \frac{\rho(x,y)}{\rho(x,o)\rho(y,o)}.\end{aligned}$$

As elements of $\text{SL}(2,\mathbb{Q}_p)$ these three maps are given by

$$A = \begin{pmatrix} a & b\\ 0 & a^{-1}\end{pmatrix}, \quad D_d = \begin{pmatrix} p^m & 0\\ 0 & p^{-m}\end{pmatrix}, \quad R = \begin{pmatrix} 0 & -1\\ 1 & 0\end{pmatrix}.$$

The maps A, D_d and R also act on T. Consider the vertex $v = p^j(x + p^k\mathbb{Z}_p)$ where x is a unit and $k \geq 0$. The action of isometries and dilations is straightforward; see pages 9 and 10 of [3]:

$$A\big(p^j(x+p^k\mathbb{Z}_p)\big) = p^j(a^2x + p^k\mathbb{Z}_p) + ba, \qquad D_d\big(p^j(x+p^k\mathbb{Z}_p)\big) = p^{j+2m}(x+p^k\mathbb{Z}_p).$$

The action of R is slightly more complicated. Let y be the unit with $xy = -1$. Then

$$R\big(p^j(x+p^k\mathbb{Z}_p)\big) = p^{-j}(y+p^k\mathbb{Z}_p).$$

if $k > 0$ and $R(p^j\mathbb{Z}_p) = p^{-j}\mathbb{Z}_p$. One may easily check that R preserves the structure of T.

Clearly $\text{Aut}(\mathbb{Q}_p)$ acts transitively on $\mathbb{Q}_p$: For any $b \in \mathbb{Q}_p$ the map $A(x) = x+b$ sends o to b. Notice that the stabiliser of o in $\text{Aut}(\mathbb{Q}_p)$ comprises those maps $A(x) = a^2x$ where a is a unit. Since the units form a compact subset of $\mathbb{Q}_p$, we see that $\text{Aut}(X)$ acts with compact stabilisers. This means that the hypotheses of Theorem 3.1 are satisfied in this case. In fact, we can restate Theorem 3.1 in a more familiar form:

Theorem 4.2. *Let A be a an element of $\text{SL}(2,\mathbb{Q}_p)$ conjugate to a diagonal matrix. Let B be any element of $\text{SL}(2,\mathbb{Q}_p)$ so that, when acting on $\mathbb{Q}_p \cup \{\infty\}$ via Möbius transformations, B neither fixes nor interchanges the fixed points of A. If $\Gamma = \langle A, B\rangle$ is discrete then*

$$\max\left\{\left|\text{tr}^2(A) - 4\right|_p, \ \left|\text{tr}[A,B] - 2\right|_p\right\} \geq 1.$$

Proof. Suppose that

$$A = \begin{pmatrix} \lambda & 0\\ 0 & \lambda^{-1}\end{pmatrix}, \qquad B = \begin{pmatrix} a & b\\ c & d\end{pmatrix},$$

where λ, a, b, c, $d \in \mathbb{Q}_p$ and $ad - bc = 1$. Then $m_A = |\lambda - \lambda^{-1}|_p$ and $Bo = b/d$, $B\infty = a/c$. By hypothesis neither Bo nor $B\infty$ is either o or ∞, so

$$\mathbb{X}(Bo,\infty;o,B\infty) = \frac{|b/d|_p}{|a/c - b/d|_p} = \frac{|b|_p|c|_p}{|ad-bc|_p} = |bc|_p.$$

We can then calculate

$$\begin{aligned}\left|\text{tr}^2(A) - 4\right|_p &= |\lambda - \lambda^{-1}|_p^2 = m_A^2,\\ \left|\text{tr}[A,B] - 2\right|_p &= |\lambda - \lambda^{-1}|_p^2\,|bc|_p = m_A^2\mathbb{X}(Bo,\infty;o,B\infty).\end{aligned}$$

The result follows directly from Theorem 3.1. $\square$

We can interpret this result geometrically in terms of the action of $\langle A,\, B\rangle$ on T as follows. Let A be as in the proof of Theorem 4.2 and write $\lambda = p^j a$ where $a = a_0 + a_1 p + \cdots$ is a unit. Then $\lambda^{-1} = p^{-j} b$ where $b = b_0 + b_1 p + \cdots$ is the unit with $ab = 1$.

Suppose that $m_A = |\lambda - \lambda^{-1}| < 1$. Then $j = 0$ and p divides $a - b$, that is $a_0 = b_0$. Since $ab = 1$ this means that $a_0 b_0 = a_0^2$ is congruent to 1 (mod p); that is λ^2 is congruent to 1 (mod p). In this case, $A(z) = \lambda^2 z$ fixes each vertex $p^j \mathbb{Z}_p$. In other words, A fixes $\gamma(o, \infty)$. For such maps, Theorem 3.1 states that $\mathbb{X}(Bo, \infty; o, B\infty) \geq 1/m_A > 1$. Geometrically this means that $\gamma(o, \infty)$ does not intersect its image under B.

On the other hand, if A has $d_A \neq 1$ then A maps the geodesic $\gamma(o, \infty)$ to itself shifting each vertex along by a fixed number of edges (see page 77 of Serre [8]). Recall that in this case $m_A \geq 1$ and $\rho(z, Az) \geq \rho(z, o)$. This corresponds to the fact that A must translate each vertex by a whole number of edges and so cannot have arbitrarily short translation length.

5. Function field spaces

We now explain how a function field can be thought of as resembling the p-adic numbers $\mathbb{Q}_p$ as developed in Section 4. We consider a field k and the field $k(t)$ of rational functions over k. The elements of $k(t)$ are quotients of two elements of the polynomial ring $k[t]$ over k. Then $k(t)$ is analogous to $\mathbb{Q}$ and $k[t]$ to $\mathbb{Z}$. We choose an irreducible polynomial $p(t)$ in $k[t]$ which plays the role analogous to the prime p in the definition of $\mathbb{Q}_p$. We consider an element $\phi(t) \in k(t)$ and we write

$$\phi(t) = p(t)^f \, \frac{u(t)}{v(t)}$$

where $f \in \mathbb{Z}$ and $u(t)$, $v(t)$ are polynomials in $k[t]$ without common factors and so that $p(t)$ does not divide either $u(t)$ or $v(t)$. Following (1) above, we define

$$\big|\phi(t)\big|_{p(t)} = c^{-f} \tag{11}$$

where $c > 1$ and we develop the theory in a manner resembling the p-adic case.

There is another approach, which we prefer in this section; see Artin [1] or Section II.1.6 of Serre [8]. If we replace the irreducible polynomial $p(t)$ with the rational function $1/t$ (which corresponds to ∞ at $t = 0$) then the valuation corresponding to (11) is

$$\left|\frac{u(t)}{v(t)}\right| = c^{\deg(u) - \deg(v)}. \tag{12}$$

Here the polynomials $u(t)$ and $v(t)$ have no common factors and have degree $\deg(u)$ and $\deg(v)$ respectively. This valuation corresponds to the standard absolute value in the case of $\mathbb{Q}$. The valuation (12) is non-Archimedean and leads to an ultrametric space.

In number theoretic applications (for example to the function fields of curves defined over finite fields) it is natural to define the number c in (11) to be q, where k is the field $\mathbb{F}_q$ of q elements.

We are led, accordingly, to consider $u(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_0$ in $k[t]$ with $a_n, a_{n-1}, \ldots, a_0 \in k$ and $n \geq 0$. We introduce the valuation

$$\big|u(t)\big| = \begin{cases} c^{\deg(u)} = c^n & \text{if } u(t) \neq 0, \\ 0 & \text{if } u(t) = 0. \end{cases} \tag{13}$$

It follows that

$$\begin{aligned} \bigl|u(t)\,v(t)\bigr| &= \bigl|u(t)\bigr|\,\bigl|v(t)\bigr|, \\ \bigl|u(t)+v(t)\bigr| &\le \max\Bigl\{\bigl|u(t)\bigr|\,\bigl|v(t)\bigr|\Bigr\}, \\ \bigl|u(t)+v(t)\bigr| &= \bigl|u(t)\bigr| \quad \text{if } \deg(v)<\deg(u). \end{aligned}$$

If we take $k=\mathbb{F}_q$ and $c=q$ then $\bigl|u(t)\bigr| = q^{\deg(u)}$ (for $u(t)\neq 0$) is the number of residue classes of polynomials in $\mathbb{F}_q[t]$ modulo $u(t)$, which is why $c=q$ is the natural choice. (Each residue class may be represented by a polynomial of degree less than $u(t)$. There are q choices for each of the $\deg(u)$ coefficients.)

If $k(t)$ denotes the quotient field of $k[t]$ then the the valuation defined by (13) extends in the obvious way to (12). The field $k\{t\}$ of formal Laurent series in $1/t$ consists of the series

$$\phi=\phi(t)=a_nt^n+a_{n-1}t^{n-1}+\cdots+a_0+a_{-1}t^{-1}+a_{-2}t^{-2}+\cdots$$

which is the completion of $k(t)$ with respect to the valuation (12) and is analogous to the completion of $\mathbb{Q}$ with respect to the Archimedean valuation. For such a ϕ we have

$$|\phi|=\bigl|\phi(t)\bigr|=c^n. \tag{14}$$

We may define $\text{Möb}\bigl(k\{t\}\bigr)$ in terms of $\text{SL}\bigl(2,k\{t\}\bigr)$ acting on $k\{t\}$ via Möbius transformations and, similarly, $\text{Aut}\bigl(k\{t\}\bigr)$. The dilations D_d are given by $D_d(\phi)=t^{2m}\phi$. We can prove the analogue of Theorem 4.2 with the valuation (14) in place of the p-adic valuation.

References

[1] E. Artin, Algebraic Numbers and Algebraic Functions, Nelson 1968.

[2] J.W.S. Cassels, Lectures on Elliptic Curves. London Mathematical Society Student Texts **18**, Cambridge University Press, 1991.

[3] A. Figà-Talamanca, *Local fields and trees.* Harmonic Functions on Trees and Buildings, ed A. Korányi, A.M.S. Contemporary Mathematics **206** (1997), 3–16.

[4] M. Gromov & R. Schoen, *Harmonic maps into singular spaces and p-adic superrigidity for lattices in groups of rank one*, I.H.E.S. Publ. Math. **76** (1992), 165–246.

[5] Y. Ihara, *On discrete subgroups of two by two projective linear groups over p-adic fields.* J. Mathematical Society of Japan **18** (1966) 219–235.

[6] T. Jørgensen, *On discrete groups of Möbius transformations*, American J. Maths. **98** (1976), 739–749.

[7] S. Markham & J.R. Parker, *Jørgensen's inequality for metric spaces with applications to the octonions*, Advances in Geometry **7** (2007), 19–38.

[8] J.-P. Serre, Trees, Springer 1980.

J. Vernon Armitage
Department of Mathematical Sciences,
University of Durham,
Durham DH1 3LE, England
e-mail: `j.v.armitage@durham.ac.uk`

John R. Parker
Department of Mathematical Sciences,
University of Durham,
Durham DH1 3LE, England
e-mail: `j.r.parker@durham.ac.uk`