ON A CLASS OF MARKOV SEMIGROUPS ON DISCRETE ULTRA-METRIC SPACES

ALEXANDER BENDIKOV, ALEXANDER GRIGOR'YAN, AND CHRISTOPHE PITTET

Abstract. We consider a discrete ultra-metric space (X, d) with a measure m and construct in a natural way a symmetric Markov semigroup $\{P_t\}_{t\geq 0}$ in $L^2(X, m)$ and the corresponding Markov process $\{\mathcal{X}_t\}$. We prove upper and lower bounds of its transition function and its Green function, give a criterion for the transience, and estimate its moments.

Contents

1. Introduction

We study here certain Markov semigroups on ultra-metric measure spaces. An ultra-metric space is a metric space (X, d) where the distance function satisfies a stronger triangle inequality

$$d(x, y) \leq \max\{d(x, z); d(z, y)\},$$

Date: November 25, 2012.
2000 *Mathematics Subject Classification.* Primary 60J05, Secondary 60J27, 60J35, 60B15.
Key words and phrases. Markov chain, Markov semigroup, Markov generator, ultra-metric space, heat kernel, transition probability.
AB was supported by the Polish Government Scientific Research Fund, Grant No. N201 371736. AG was supported by SFB 701 of German Research Council. CP was supported by the CNRS.

for all $x, y, z \in X$, that is called the *ultra-metric* (or ultra-triangle) inequality. We say that an ultra-metric space (X, d) is *discrete* if X is an infinite set, all metric balls of (X, d) are finite sets, and the distance function d takes only integer values[1] (various equivalent descriptions of such spaces are given in Section 2). In this paper we consider only discrete ultra-metric spaces postponing general ultra-metric spaces to a sequel paper.

Various aspects of potential analysis on ultra-metric spaces have been studied before. In the important case of non-archimedean local fields, the theory has been developed in [22], and even in a more general setting in [20]. The Riesz potentials as defined in the above two references are related (as usual) to Laplacians defined in the present paper. Also the Sobolev embedding (see [20, VI.4]) is as usual related to properties of heat kernels studied in the present paper.

A well-known example of an ultra-metric space is the field $\mathbb{Q}_p$ of p-adic numbers with the metric $\|x-y\|_p$ where $\|\cdot\|_p$ is the p-adic norm. This space is not discrete but its subset

$$\Lambda_p = \left\{ \frac{a}{p^k} : a < p^k, \quad a, k \in \mathbb{Z}_+, \right\}$$

is a discrete ultra-metric space. By representing p-adic fractions in the numeral base p one easily identifies Λ_p with the inductive limit

$$\mathbb{Z}(p)^{(\infty)} = \varinjlim \mathbb{Z}\left(p\right)^n ,$$

where the ultra-metric distance between two elements $x, y \in \mathbb{Z}(p)^{(\infty)}$ is defined as the minimal n so that x, y belong to the same coset of $\mathbb{Z}\left(p\right)^n$ in $\mathbb{Z}(p)^{(\infty)}$.

There is a huge literature devoted to random walks on finitely (or compactly) generated groups, based on the study of the geometry of such groups (see e.g. [4], [7], [12], [18], [19], [21], [23], [25], and references therein). The group $\mathbb{Z}\left(p\right)^{(\infty)}$ is not finitely generated. A basic notion of geometry of finitely generated groups – the word metric, does not apply here. However, the group $\mathbb{Z}\left(p\right)^{(\infty)}$ is a member of the class of *locally finite* groups, where the notion of an ultra-metric can be used instead of the word metric.

Having defined random walks on $\mathbb{Z}\left(p\right)^n$ for each n let us select a stochastic sequence $\{c_n\}_{n=0}^{\infty}$ and define a random walk in $\mathbb{Z}(p)^{(\infty)}$ as follows: with probability c_n we choose n and then move according to the law in the subgroup $\mathbb{Z}\left(p\right)^n$. Similarly one defines random walks on other locally finite groups. Certain properties of such random walks, including recurrence, have been studied by a number of authors – see, for example, [3], [6], [8], [9], [10], [13], [14], [15], [16], [17].

Let us now define a Markov chain on a discrete ultra-metric space. Fix a measure m on a discrete ultra-metric space (X, d) such that $0 < m\left(x\right) < \infty$ for any $x \in X$ and $m\left(X\right) = \infty$. For example, m can be a counting measure. Denote by $B_r\left(x\right)$ the closed d-ball of radius r centered at x. Consider the following Markov operator, which is defined for all functions $f \in L^2\left(X, m\right)$:

$$Pf\left(x\right) = \sum_{k=0}^{\infty} \frac{c_k}{m\left(B_k\left(x\right)\right)} \int_{B_k(x)} f dm,$$

where $\{c_k\}_{k=0}^{\infty}$ is a given sequence of positive numbers such that $\sum_{k=0}^{\infty} c_k = 1$. The operator P is associated with two types of Markov chains.

1. A discrete time Markov chain $\{\mathcal{X}_n\}$, $n = 0, 1, ...$ with the following transition rule: $\mathcal{X}_{n+1}$ is uniformly distributed in $B_k\left(\mathcal{X}_n\right)$ with probability c_k. This can be viewed as

[1] One can slightly relax the latter assumption by assuming that the set of values of $d\left(x, y\right)$ is contained in a sequence $\{v_k\}_{k=0}^{\infty}$ such that $v_k \to \infty$ as $k \to \infty$. Since the composition of an ultra-metric with a strictly monotone increasing function with a fixed point 0 is again an ultra metric, one can use such a change of metric to reduce a general sequence $\{v_k\}$ to the sequence $v_k = k$.

follows: in each ball $B_k(x)$ we define a Markov chain that gets uniformly distributed in one step, and then take a convex combination of the corresponding Markov kernels with coefficients c_k.

2. A continuous time Markov process $\{\mathcal{X}_t\}_{t\geq 0}$ with transition density $p(t,x,y)$, which is the integral kernel of the operator P^t with respect to measure m. The Markov semigroup $\left\{P^t\right\}_{t\geq 0}$ is defined using the functional calculus of self-adjoint operators and the fact that P is symmetric and non-negative definite (note, that a symmetric Markov operator does not have to be non-negative definite in general).

It follows that the discrete time Markov chain coincides with the restriction of the continuous time process $\{\mathcal{X}_t\}$ to integer values of t, which allows us to focus on the study of $\{\mathcal{X}_t\}$. One of the main purposes of this paper is to obtain estimates of the transition density (=heat kernel) $p(t,x,y)$ from the given data: the sequences $\{c_k\}$ and $\{m(B_k(x))\}$.

We should emphasize that, to the best of our knowledge, this is the first time that this problem has been considered without assuming a group structure on X, which excludes the powerful methods of harmonic analysis such as the Fourier transform. Our approach is based on the simple observation that the building blocks of the operator P, namely, the operators

$$\mathcal{P}_k f(x) = \frac{1}{m(B_k(x))}\int_{B_k(x)} f dm,$$

are not only Markov operators, but also orthogonal projectors in $L^2(X,m)$. The latter is generally not true on an arbitrary metric measure space, but is a specific property of an ultra-metric. Indeed, it follows easily from the ultra-metric inequality that any two balls of the same radius are either disjoint or identical, which implies that $\mathcal{P}_k f$ belongs to the subspace $\mathcal{V}_k$ of $L^2(X,m)$ of all functions that are constant on all balls of radius k, and $\mathcal{P}_k$ is an orthoprojector onto $\mathcal{V}_k$. Moreover, the sequence $\{\mathcal{V}_k\}$ is decreasing in k, so that the family $\{\mathcal{P}_k\}$ is a spectral resolution of the identity, up to reparametrization.

Hence, the ultra-metric property allows one to obtain immediately a spectral resolution of P, where the spectral projectors are also Markov operators. This enables us to engage at an early stage the methods of the spectral theory and functional calculus. Our results are stated in terms of the spectral density function $N(x,\lambda)$ that is defined in Section 3.2 using the sequences $\{c_k\}$ and $\{m(B_k(x))\}$. It is an increasing staircase function on $[0,1]$ that changes from 0 to $\frac{1}{m(x)}$. Its behavior as $\lambda\to 0$ is intimately related to the behavior of the heat kernel $p(t,x,y)$ as $t\to\infty$. By Theorem 3.5, we have the following explicit identity

$$p(t,x,y) = t\int_0^{\frac{1}{1+d_\sigma(x,y)}} N(x,\lambda)(1-\lambda)^{t-1}d\lambda, \tag{1.1}$$

where d_σ is another ultra-metric on X that is defined in (3.21) using the original ultra-metric d and the sequence $\{c_k\}$.

Under certain additional assumptions about the function $N(x,\lambda)$ we obtain in Section 3 various estimates for the heat kernel. For example, if the function $\lambda\mapsto N(x,\lambda)$ satisfies the doubling property then, by Corollary 3.17, for all $t\geq 1$,

$$p(t,x,y)\simeq \frac{t}{t+d_\sigma(x,y)}N\left(x,\frac{1}{t+d_\sigma(x,y)}\right)$$

(where the sign $\simeq$ means that the ratio of the left and right hand sides is bounded from above and below by positive constants). In particular, if $N(x,\lambda)\simeq\lambda^\alpha$ then

$$p(t,x,y)\simeq\frac{t}{(t+d_\sigma(x,y))^{1+\alpha}},$$

that is, $p(t,x,y)$ behaves like the Cauchy distribution in "α-dimensional" space.

In Section 4 we obtain estimates of the Green function and conditions for the transience of the Markov process $\{\mathcal{X}_t\}$. Theorem 4.1 says that the process is transient if and only if

$$\int_0 N(x,\lambda)\frac{d\lambda}{\lambda^2} < \infty.$$

For example, for the function $N(x,\lambda) \simeq \lambda^\alpha$ the transience is equivalent to $\alpha > 1$.

In Section 5 we give examples of applications of the above results to specific sequences $\{c_k\}$ and $\{m(B_k(x))\}$.

In Section 6 we estimate the moments

$$M_\gamma(x,t) = \mathbb{E}_x\left(d_\sigma(x,\mathcal{X}_t)^\gamma\right)$$

of the process $\{\mathcal{X}_t\}$. Theorem 6.2 provides a general upper bound: if $0<\gamma<1$ then, for large enough t,

$$M_\gamma(x,t) \le \frac{t^\gamma}{1-\gamma}.$$

In particular, the γ-moment is finite for all $0<\gamma<1$. If $N(x,\lambda)$ satisfies the reverse doubling condition then there is a matching lower bound (Theorem 6.3). In this case the γ-moment is finite if and only if $0<\gamma<1$.

In Section 7 we describe the generator $\mathcal{L}$ of the semigroup $\left\{P^t\right\}$ and show that, for any continuous, strictly monotone increasing function $\phi:[0,\infty)\to[0,\infty)$ such that $\phi(0)=0$, the operator $\phi(\mathcal{L})$ is also a generator of a similar semigroup (Corollary 7.2). In particular, $\mathcal{L}^\alpha$ is also a Markov generator for any $\alpha>0$. Recall for comparison that, for a general symmetric non-negative definite Markov generator $\mathcal{L}$, the operator $\mathcal{L}^\alpha$ generates a Markov semigroup only for $0<\alpha\le 1$. We also prove the independence of p of the spectrum of $\mathcal{L}$ in $L^p(X,m)$ as well as a strong Liouville property of $\mathcal{L}$.

Notation. Given two functions f and g of the same argument, the relation

$$f \simeq g$$

means that $C_1 g \le f \le C_2 g$ for some positive constants C_1, C_2 and for a specified range of the argument of f, g.

If $f(t)$ and $g(t)$ are two functions of a real variable t then the relation

$$f(t) \asymp g(t)$$

means that there are positive constants C_1, C_1', C_2, C_2' such that

$$C_1 g\left(C_1' t\right) \le f(t) \le C_2 g\left(C_2' t\right)$$

for all t from a specified range.

2. Explicit description of discrete ultra-metric spaces

In this section we provide alternative descriptions of discrete ultra-metric spaces that allow to construct many examples of such spaces. This material is not used in the main part of the paper, though.

2.1. **Downward trees.** Let Γ be a countable connected graph that is constructed in the following way. The set of vertices of Γ consists of disjoint union of subsets $\{\Gamma_k\}_{k=0}^\infty$ with the following properties:

- from each vertex $v\in\Gamma_k$ there is exactly one edge to Γ_{k+1};
- for each vertex $v\in\Gamma_k$ the number of edges connecting v to Γ_{k-1} is finite and positive, provided $k\ge 1$;
- if $|k-l|\neq 1$ then there is no edges between vertices of Γ_k and Γ_l.

We will call such a graph Γ a *downward tree* (it is clear that Γ is a tree – see Fig. 1). The set Γ_0 that plays a special role and will be called the bottom of Γ.

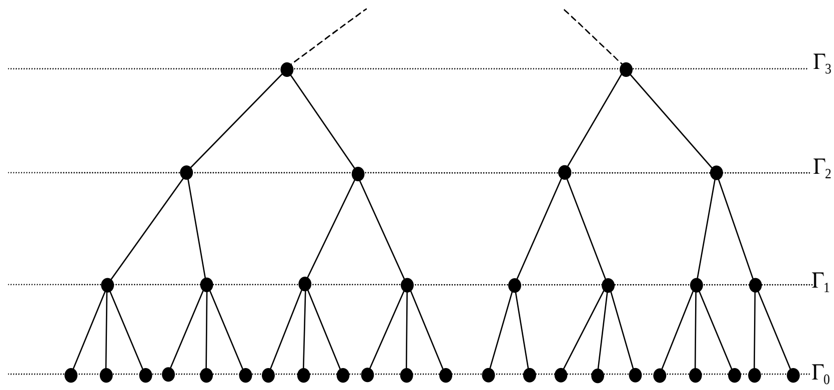

FIGURE 1. Graph Γ

Let d_Γ denote the graph distance between the vertices of Γ, that is, the smallest number of edges in an edge path connecting two vertices.

Lemma 2.1. (Γ_0, d_Γ) *is a discrete ultra-metric space.*

Proof. Observe that the shortest path between vertices $x, y \in \Gamma_0$ goes through the nearest common ancestor: a vertex $a \in \Gamma_k$ with the minimal value of k, that is connected to x and y by downward paths. It follows that $d_\Gamma(x, y) = 2k$. Let z be another vertex on Γ_0 and $b \in \Gamma_l$ be the nearest common ancestor of y and z, so that $d_\Gamma(y, z) = 2l$. Then y is connected to both a and b by upward paths. Since there is only one upward path emanating from y, the vertices a and b lie on the same upward path. Without loss of generality, assume that $k \le l$. Then we obtain a path from b to x that goes through a and the number of edges in this path is l. Since $d_\Gamma(b, z) = l$, we obtain that

$$d_\Gamma(x, z) \le d_\Gamma(x, b) + d_\Gamma(b, z) \le 2l = \max(d_\Gamma(x, y), d_\Gamma(y, z)),$$

which proves the ultra-metric inequality. ■

As it is clear from the proof, d_Γ on Γ_0 takes only even values, so it would be more natural to consider the distance $\frac{1}{2}d_\Gamma$ whose range is $\mathbb{Z}_+$.

Let us now show that any discrete ultra-metric space (X, d) admits a representation as the bottom of a downward tree Γ. Define the vertices of Γ to be all distinct balls $\{B_k(x)\}$ where $x \in X$ and $k \in \mathbb{Z}_+$. Two balls $B_k(x)$ and $B_l(y)$ are connected by an edge in Γ if $|k - l| = 1$ and one of them is a subset of the other. The graph Γ is naturally split into levels Γ_k, $k \in \mathbb{Z}_+$, where Γ_k is the set of all vertices that correspond to the balls of radii k. That is, Γ_0 coincides with the set X, Γ_1 consists of balls of radii 1, etc. Clearly, edges exist only between the vertices of Γ_k and Γ_{k+1}.

Lemma 2.2. *The graph* Γ *is a downward tree. Furthermore, for all* $x, y \in X$, $d_\Gamma(x, y) = 2d(x, y)$.

Proof. Any vertex of Γ_k has exactly one edge to Γ_{k+1}, because any ball of radius k is contained exactly in one ball of radius $k + 1$ (indeed, any two intersecting $(k + 1)$-balls coincide due to the property of an ultra-metric). If $k \ge 1$ then any ball of radius k contains at least one ball of radius $k - 1$, and the number of such balls is finite. Hence, Γ is a downward tree.

To prove the second claim, set $n = d(x, y)$ and observe that $B_n(x) = B_n(y)$. Then the following path between x and y consists of $2n$ edges:

$$x \sim B_1(x) \sim ... \sim B_n(x) \sim B_{n-1}(y) \sim ... \sim B_1(y) \sim y,$$

which implies that $d_\Gamma(x, y) \leq 2n$. To prove the opposite inequality set $k = \frac{1}{2} d_\Gamma(x, y)$ and observe that the nearest common ancestor of x an y in Γ is at the level k. Let it be a ball $B_k(z)$. Then $x, y \in B_k(z)$, and the ultra-metric inequality implies

$$d(x, y) \leq \max(d(x, z), d(y, z)) \leq k = \frac{1}{2} d_\Gamma(x, y),$$

whence the identity $d_\Gamma(x, y) = 2d(x, y)$ follows. ■

2.2. **Partitions.** The graph Γ that is associated with a discrete ultra-metric space can be also described as follows. All balls of a given radius k provide a partition of X, so that Γ_k consists of the elements of this partition. Each of the balls of radius k is partitioned into finitely many smaller balls of radius $k - 1$, and is contained in exactly one ball of radius $k + 1$ (see Fig. 2).

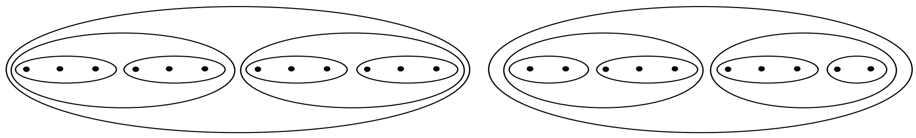

FIGURE 2. Ultra-metric balls matching the tree on Fig. 1

Conversely, assume that X is a countable set that admits for any $k \in \mathbb{Z}_+$ a partition Γ_k into finite disjoint sets with the following properties:

- elements of Γ_0 are single points of X;
- partition Γ_k is a refinement of Γ_{k+1};
- for each $x \in X$ and $k \in \mathbb{Z}_+$ denote by $B_k(x)$ the unique element of Γ_k that contains x; then

$$\bigcup_{k=0}^{\infty} B_k(x) = X. \tag{2.1}$$

Given that much, define

$$d(x, y) = \min\{k : y \in B_k(x)\}. \tag{2.2}$$

Lemma 2.3. *Under the above hypotheses, (X, d) is a discrete ultra-metric space and $B_k(x)$ are its ultra-metric balls.*

Proof. The family Γ of all elements of all partitions Γ_k has an obvious structure of a downward tree: if $u \in \Gamma_k$ and $v \in \Gamma_l$ then u and v are connected by an edge in Γ if $|k - l| = 1$ and one of the sets u, v is contained in the other. The hypothesis (2.1) implies the connectedness of Γ. The definition (2.2) of $d(x, y)$ means that $d(x, y)$ is equal to the level of the nearest common ancestor of x, y. By the proof of Lemma 2.1 we obtain $d(x, y) = \frac{1}{2} d_\Gamma(x, y)$. Hence, (X, d) coincides with $\left(\Gamma_0, \frac{1}{2} d_\Gamma(x, y)\right)$ and, hence, is a discrete ultra-metric space by Lemma 2.1. ■

2.3. $\{n_k\}$**-adic sequences.** Fix a sequence $\{n_k\}_{k=1}^{\infty}$ of positive integers. We say that a sequence $\{a_k\}_{k=1}^{\infty}$ is $\{n_k\}$-adic, if it satisfies the following conditions:

- $a_k \in \{0, 1, ..., n_k - 1\}$ for any $k \in \mathbb{N}$
- $a_k = 0$ for all large enough k.

Denote by $\mathcal{A} = \mathcal{A}(\{n_k\})$ the class of all $\{n_k\}$-adic sequences. For any two sequences $\{a_k\}, \{b_k\} \in \mathcal{A}$, define $N(\{a_k\}, \{b_k\})$ as the non-negative integer N such that $a_k = b_k$ for all $k > N$ but $a_N \neq b_N$. Such N is unique by definition and exists because $a_k = b_k = 0$ for large enough k. If $\{a_k\}, \{b_k\}$ are not identical then one obtains

$$N(\{a_k\}, \{b_k\}) = \max\{k \geq 1 : a_k \neq b_k\} \geq 1. \tag{2.3}$$

If $\{a_k\} \equiv \{b_k\}$ then $N = 0$.

Lemma 2.4. $(\mathcal{A}, N)$ *is a discrete ultra-metric space.*

Proof. Let us prove the ultra-metric inequality. Set $n = N(\{a_k\}, \{b_k\})$ and $m = N(\{b_k\}, \{c_k\})$. Then for $k > \max(n, m)$ we have $a_k = b_k$ and $b_k = c_k$ whence it follows that $a_k = c_k$ and, hence,

$$N(\{a_k\}, \{c_k\}) \leq \max(n, m).$$

The set $\mathcal{A}$ is countable due to the hypothesis that $a_k = 0$ for all large k. Finally, any ball $B_r(\{a_k\})$ with respect to the ultra-metric N is finite because if $\{x_k\}$ is a point in this ball then $x_k = a_k$ for all $k > r$ so that only $x_1, ..., x_r$ are variable. ■

Obviously one can interpret the component a_k as an element of $\mathbb{Z}(n_k) := \mathbb{Z}/(n_k\mathbb{Z})$. Then the set $\mathcal{A}$ is identified with the additive group

$$G(\{n_k\}) = \mathbb{Z}(n_1) \oplus \mathbb{Z}(n_2) \oplus ... = \bigoplus_{k=1}^{\infty} \mathbb{Z}(n_k),$$

where the infinite direct sum is defined as an inductive limit of finite sums, that is, as the union of $\bigoplus_{k=1}^{m} \mathbb{Z}(n_k)$, $m \in \mathbb{N}$.

2.4. **Radially homogeneous ultra-metric spaces.** We say that a discrete ultra-metric space (X, d) is *radially homogeneous* if, for any $k \in \mathbb{N}$, there is a positive integer n_k such that every ball $B_k(x)$ of radius k contains exactly n_k distinct balls of radii $k - 1$. Fix a reference point $o \in X$. Let us enumerate all balls of radius $k - 1$ that are contained in $B_k(x)$ by integers $0, 1, ..., n_k - 1$. If one of these balls is $B_{k-1}(o)$ then it receives the ordinal number 0. Then any point $x \in X$ can be associated with a sequence $\{x_k\}_{k=1}^{\infty}$ where x_k is the ordinal number of the ball $B_{k-1}(x)$ in $B_k(x)$; in particular, x_k takes values $0, 1, ..., n_k - 1$ (see Fig. 3). Furthermore, if $n = d(x, o)$ then $B_k(x) = B_k(o)$ for all $k \geq n$, which means that $x_k = 0$ for all $k > n$.

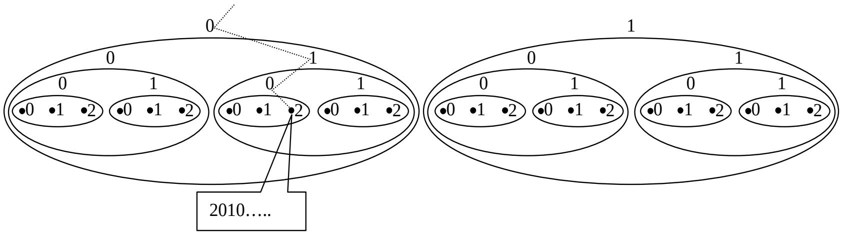

FIGURE 3. Associating a sequence $\{x_k\}$ to a point x, using the sequence $\{n_k\}_{k=1}^{\infty} = \{3, 2, 2, ...\}$

Lemma 2.5. *The mapping $x \mapsto \{x_k\}$ is a bijection between X and the set $\mathcal{A}(\{n_k\})$ of all $\{n_k\}$-adic sequences. Furthermore, for any two points $x, y \in X$,*

$$d(x,y) = N(\{x_k\}, \{y_k\}). \tag{2.4}$$

Hence, (X,d) and $(\mathcal{A}, N)$ are isometric.

Proof. Let $\{a_k\}$ be a $\{n_k\}$-adic sequence and let l be so big that $a_k = 0$ for all $k \geq l$. Consider a sequence of balls

$$B_l(o) \supset B_{l-1} \supset B_{l-2} \supset ... \supset B_1 \supset B_0$$

that is constructed inductively as follows: B_{k-1} is the ball of radius $k-1$ that has the ordinal number a_k in B_k. The ball B_0 consists of a single point x, and for this point we have $x_k = a_k$ for all k. The uniqueness follows from the same construction.

Given two points $x, y \in X$ and their sequences $\{x_k\}, \{y_k\}$, let $n = d(x,y)$. Then $B_{k-1}(x) = B_{k-1}(y)$ for any $k > n$ which implies $x_k = y_k$. On the contrary, $B_{n-1}(x) \neq B_{n-1}(y)$, which implies that the balls $B_{n-1}(x)$ and $B_{n-1}(y)$ have different ordinal numbers in the ball $B_n(x) = B_n(y)$, whence $x_n \neq y_n$. Hence, $n = N(\{x_k\}, \{y_k\})$, which proves (2.4). ■

2.5. **p-adic net.** Let all n_k in the sequence $\{n_k\}$ be equal to the same value p. We refer to $\{n_k\}$-adic sequences as p-adic and set $\mathcal{A}_p \equiv \mathcal{A}(\{n_k\})$. By the above results, $\mathcal{A}_p$ can be identified with a downward p-regular tree as well as with $\mathbb{Z}(p)^{(\infty)}$. On the other hand, if p is a prime then one can identify $\mathcal{A}_p$ with a subset of p-adic numbers as follows.

Recall that, for a given prime p, the p-adic norm $\|\cdot\|_p$ of any rational $x \in \mathbb{Q}$ is defined by $\|x\|_p = p^n$ provided x can be written in the form $x = p^{-n}\frac{a}{b}$, where a, b, n are integers such that neither a nor b is divisible by p; also set $\|0\|_p = 0$. Then $\|\cdot\|_p$ is an ultra-norm, that is, it satisfies the ultra-metric inequality

$$\|x+y\|_p \leq \max\left(\|x\|_p, \|y\|_p\right).$$

Consequently, $\|x-y\|_p$ is an ultra-metric on $\mathbb{Q}$. Consider now the set

$$\Lambda_p = \left\{\frac{a}{p^n} : n, a \in \mathbb{Z}_+, \quad a < p^n\right\},$$

which is an 1-net in $\mathbb{Q}_p$. Define $\log^+ r$ as the positive part of $\log r$; that is, $\log^+ r = \log r$ if $r \geq 1$ and $\log^+ r = 0$ if $r < 1$.

Lemma 2.6. $\left(\Lambda_p, \log_p^+ \|\cdot\|_p\right)$ *is a discrete ultra-metric space that is isometric to $(\mathcal{A}_p, N)$.*

Proof. Any integer $a \in [0, p^n)$ can be expanded in a base p as $a = \sum_{k=0}^{n-1} a_k p^k$ where $a_k \in \{0, 1, .., p-1\}$, whence, for $x = \frac{a}{p^n}$, we obtain an expansion

$$x = \frac{x_1}{p} + \frac{x_2}{p^2} + ... + \frac{x_n}{p^n},$$

where $x_k = a_{k-1} \in \{0, 1, ..., p-1\}$. Defining $x_k = 0$ for all $k > n$, we obtain a p-adic sequence $\{x_k\}_{k=1}^{\infty}$. Hence, we have constructed a bijection $x \mapsto \{x_k\}$ from Λ_p onto $\mathcal{A}_p$.

Fix $x \in \Lambda_p \setminus \{0\}$ and set

$$n = \max\{k : x_k \neq 0\}.$$

Then we have by (2.3) $N(x,0) = n$, while $\|x\|_p = p^n$, whence it follows that

$$\log_p \|x\|_p = N(x,0).$$

If $x = 0$ then

$$\log_p^+ \|x\|_p = 0 = N(x,0).$$

It follows that, for all $x, y \in \Lambda_p$,

$$\log_p^+ \|x-y\|_p = N(x,y),$$

which finishes the proof. ■

This example can be slightly generalized by considering the additive group $\mathbb{F}_q[T]$ of all polynomials over the field $\mathbb{F}_q$, where $q = p^s$. Indeed, any non-zero polynomial $f \in \mathbb{F}_q[T]$ has the form

$$f = \sum_{k=0}^{n} a_k T^k,$$

where $n \in \mathbb{Z}_+$, $a_k \in \mathbb{F}_q$, $a_n \neq 0$, and its norm in $\mathbb{F}_q[T]$ is defined by $\|f\|_T := q^n$ (see [24]). Then the associated metric $\|f-g\|_T$ makes $\mathbb{F}_q[T]$ into a discrete ultra-metric space. Moreover, $\left(\mathbb{F}_q[T], \log_q^+(q\|\cdot\|_T)\right)$ is isometric to $(\mathcal{A}_q, N)$.

3. Heat semigroup and heat kernel

3.1. **Markov chains on ultra-metric spaces.** A distance function d on a set X is called an *ultra-metric* if d satisfies the ultra-metric inequality

$$d(x,y) \leq \max\{d(x,z); d(z,y)\}, \forall x,y,z \in X.$$

A metric space (X,d) with an ultra-metric d is called an ultra-metric space. The ultra-metric balls

$$B_r(x) = \{y \in X : d(x,y) \leq r\}$$

have the property that any two balls of the same radius are either identical or disjoint. In other words, any point inside a ball B is its center.

An ultra-metric space (X,d) is called *discrete* if the set X is countable, all balls $B_r(x)$ are finite, and the distance function d takes only integer values. In this paper we treat only discrete ultra-metric spaces postponing more general cases to a follow-up paper.

From now on (X,d) is a discrete ultra-metric space. Let us fix any measure m on 2^X such that $0 < m(x) < \infty$ for any $x \in X$ and $m(X) = \infty$. For example, m can be a counting measure. For any non-negative integer k, define the operator $\mathcal{P}_k$ acting on all functions f on X as follows:

$$\mathcal{P}_k f(x) = \frac{1}{m(B_k(x))} \int_{B_k(x)} f dm. \tag{3.1}$$

By the property of ultra-metric balls, $\mathcal{P}_k f$ is constant on any ball of radius k. Clearly, $\mathcal{P}_k$ can be written in the form

$$\mathcal{P}_k f(x) = \int K_k(x,y) f(y) dm(y)$$

where $K_k(x,y)$ is the integral kernel of $\mathcal{P}_k$ given by

$$K_k(x,y) = \frac{1}{m(B_k(x))} \mathbf{1}_{B_k(x)}(y). \tag{3.2}$$

By the ultra-metric property, if $y \in B_k(x)$ then $B_k(x) = B_k(y)$, which implies that K_k is symmetric in x, y, that is,

$$K_k(x,y) = K_k(y,x).$$

Since $\mathcal{P}_k$ preserves positivity and $\mathcal{P}_k 1 = 1$, it follows that $\mathcal{P}_k$ is a symmetric Markov operator with stationary measure m. By a standard argument, $\mathcal{P}_k$ is a bounded self-adjoint operator in $L^2 = L^2(X,m)$ with $\|\mathcal{P}_k\|_{L^2 \to L^2} \leq 1$.

Choose now a sequence $\{c_k\}_{k=0}^{\infty}$ of strictly positive reals such that

$$\sum_{k=0}^{\infty} c_k = 1.$$

Then we define the following operator

$$P = \sum_{k=0}^{\infty} c_k \mathcal{P}_k. \tag{3.3}$$

Clearly, the series converges in the operator norm of L^2. Then P is also a symmetric Markov operator in L^2 and, hence, it determines a reversible Markov chain $\{\mathcal{X}_n\}$ on X whose transition operator at time $n \in \mathbb{Z}_+$ is P^n. If $\mathcal{X}_n = x$ then $\mathcal{X}_{n+1}$ is uniformly distributed in $B_k(x)$ with probability c_k, for any $k \in \mathbb{Z}_+$. One of our main purposes is to obtain explicit estimates for transitions probabilities of this Markov chain.

Denote by $\mathcal{V}_k$ the subspace of L^2 that consists of functions that are constants on all balls of radii k. It is clear that $\mathcal{V}_k$ is a closed subspace and

$$L^2 = \mathcal{V}_0 \supset ... \supset \mathcal{V}_k \supset \mathcal{V}_{k+1} \supset ... .$$

By the hypothesis $m(X) = \infty$, constants are not in L^2 and, hence, the intersection of all subspaces $\{\mathcal{V}_k\}$ is a trivial subspace $\{0\}$.

A major observation that takes full advantage of an ultra-metric structure is as follows.

Lemma 3.1. *$\mathcal{P}_k$ is the orthoprojector of L^2 onto $\mathcal{V}_k$.*

Proof. It is obvious from (3.1) that $\mathcal{P}_k^2 = \mathcal{P}_k$ and that the image of $\mathcal{P}_k$ is $\mathcal{V}_k$. Since $\mathcal{P}_k$ is a bounded self-adjoint operator in L^2, we obtain that $\mathcal{P}_k$ is an orthoprojector onto its image, that is, onto $\mathcal{V}_k$. ■

Consequently, $\{\mathcal{P}_k\}$ is a decreasing sequence of orthoprojectors such that $\mathcal{P}_0 = \mathrm{id}$ and $\mathcal{P}_k \to 0$ as $k \to \infty$ in the strong operator topology.

Therefore, $\{\mathcal{P}_k - \mathcal{P}_{k+1}\}_{k=0}^{\infty}$ is a sequence of orthoprojectors with mutually orthogonal images, and the identity (3.3) implies by the Abel transformation[2] that

$$P = \sum_{k=0}^{\infty} s_k \left(\mathcal{P}_k - \mathcal{P}_{k+1}\right), \tag{3.4}$$

where

$$s_k = c_0 + ... + c_k.$$

Set also $s_{-1} = 0$. Note that $0 < s_k < 1$ for $k = 0, 1, ...$ and $s_k \uparrow 1$ as $k \to \infty$. The series (3.4) converges in the strong operator topology, due to the Bessel inequality and the boundedness of $\{s_k\}$.

Hence, the identity (3.4) is the spectral resolution in L^2 of the operator P. By functional calculus of self-adjoint operators, we have, for any $t \geq 0$,

$$P^t = \sum_{k=0}^{\infty} s_k^t \left(\mathcal{P}_k - \mathcal{P}_{k+1}\right) = \sum_{k=0}^{\infty} \left(s_k^t - s_{k-1}^t\right) \mathcal{P}_k, \tag{3.5}$$

where we use in the second identity the Abel transformation. Note that the first series in (3.5) converges in the strong operator topology while the second one converges in the operator norm.

[2]We use the Abel transformation in the form

$$\sum_{k=0}^{\infty} a_k \left(b_k - b_{k+1}\right) = a_{-1} b_0 + \sum_{k=0}^{\infty} \left(a_k - a_{k-1}\right) b_k$$

provided one of the above series converges and $a_n b_{n+1} \to 0$ as $n \to \infty$.

Lemma 3.2. *The family $\left\{P^t\right\}_{t\geq 0}$ is a strongly continuous Markov semigroup in $L^2(X,m)$.*

Proof. Since $\left\{s_k^t - s_{k-1}^t\right\}_{k=0}^{\infty}$ is a stochastic sequence, the operator P^t is Markov. The semigroup identity $P^{t+s} = P^t P^s$ follows from the functional calculus. The strong continuity of P^t, that is, $s\text{-}\lim_{t\to 0} P^t = \mathrm{id}$, follows from the first identity of (3.5) by the bounded convergence theorem (cf. [11, Lemma 4.8]), because $s_k^t \to 1$ as $t \to 0$ and the sequence $\left\{s_k^t\right\}$ is uniformly bounded by 1. ■

Hence, the semigroup $\left\{P^t\right\}_{t\geq 0}$ determines a continuous time Markov chain $\{\mathcal{X}_t\}_{t\geq 0}$ that extends the above mentioned discrete time Markov chain $\{\mathcal{X}_n\}$ to the real-valued time. That the Markov chain $\{\mathcal{X}_n\}$ can be embedded into a continuous time Markov chain is a very specific property of $\{\mathcal{X}_n\}$ that is a consequence of the ultra-metric structure (for example, a simple random walk on $\mathbb{Z}$ cannot be embedded into a continuous time Markov chain).

Theorem 3.3. *The operator P^t has the integral kernel $p(t,x,y)$, that is*

$$P^t f(x) = \int_X p(t,x,y) f(y)\, dm(y),$$

where $p(t,x,y)$ is given explicitly by the identity

$$p(t,x,y) = \sum_{k=d(x,y)}^{\infty} \left(s_k^t - s_{k-1}^t\right) \frac{1}{m(B_k(x))}. \tag{3.6}$$

Proof. Using (3.5) and the integral kernel K_k of $\mathcal{P}_k$ that is given by (3.2), we obtain that the integral kernel of P^t is

$$\begin{aligned} p(t,x,y) &= \sum_{k=0}^{\infty} \left(s_k^t - s_{k-1}^t\right) K_k(x,y) \\ &= \sum_{k=0}^{\infty} \left(s_k^t - s_{k-1}^t\right) \frac{\mathbf{1}_{B_k(x)}(y)}{m(B_k(x))} \\ &= \sum_{k=d(x,y)}^{\infty} \left(s_k^t - s_{k-1}^t\right) \frac{1}{m(B_k(x))}. \end{aligned}$$

■

Corollary 3.4. *For all $x \in X$ and $t > 0$, the function $y \mapsto p(t,x,y)$ depends only on $d(x,y)$. If in addition the sequence $\{m(B_k(x))\}$ is independent of x then function $x, y \mapsto p(t,x,y)$ depends only on $d(x,y)$.*

Proof. This is clear from (3.6). ■

3.2. **The spectral density.** Using (3.6) we obtain another convenient expression for $p(t,x,y)$. Set

$$\sigma_k = 1 - s_k = \sum_{l>k} c_l$$

for any non-negative integer k. In what follows we use the sequence $\{\sigma_k\}$ as the main input data instead of $\{c_k\}$. Clearly, $\{\sigma_k\}_{k=0}^{\infty}$ can be any sequence of positive reals that satisfies the following conditions:

$$\sigma_{k+1} < \sigma_k < 1, \quad k = 0, 1, \ldots \quad \text{and} \quad \sigma_k \to 0 \text{ as } k \to \infty.$$

Put for convenience $\sigma_{-1} = 1$.

Our main tool is the function $N(x,\lambda)$ defined for all $x \in X$ and $\lambda \in [0,+\infty)$ as follows:

$$N(x,\lambda) = \begin{cases} 0, & \text{if } \lambda = 0, \\ \dfrac{1}{m(B_k(x))}, & \text{if } \lambda \in [\sigma_k, \sigma_{k-1}) \quad k = 0,1,2,..., \\ \dfrac{1}{m(x)}, & \text{if } \lambda \geq 1, \end{cases} \tag{3.7}$$

see Fig. 4. The function $N(x,\lambda)$ is nothing other than the spectral density of the associated discrete Laplace operator id $-P$. This function encodes all necessary information about the Markov kernel P: the sequence $\{\sigma_k\}$ and the measures $m(B_k(x))$ of balls. Observe that the function $\lambda \mapsto N(x,\lambda)$ is monotone increasing and right continuous on $[0,+\infty)$.

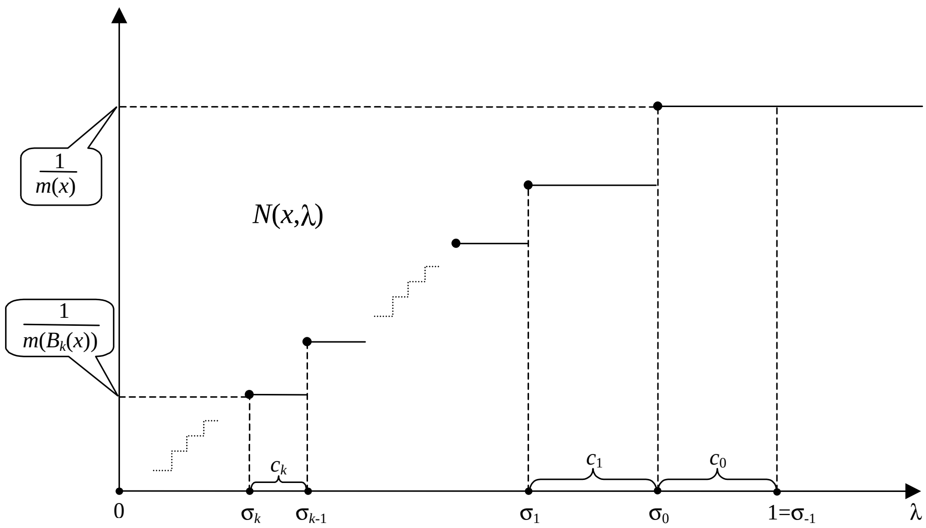

FIGURE 4. Function $\lambda \mapsto N(x,\lambda)$

The significance of the spectral density is determined by the following identities.

Theorem 3.5. *The following identity holds for all $x,y \in X$ and $t > 0$*

$$p(t,x,y) = t\int_0^{\sigma_{n-1}} N(x,\lambda)(1-\lambda)^{t-1}d\lambda, \tag{3.8}$$

where $n = d(x,y)$, and

$$p(t,x,x) = \int_{[0,1]} (1-\lambda)^t \, dN(x,\lambda). \tag{3.9}$$

Proof. Using (3.6) and (3.7), we obtain

$$\begin{aligned} p(t,x,y) &= \sum_{k=n}^{\infty}\left(s_k^t - s_{k-1}^t\right)\frac{1}{m(B_k(x))} \\ &= \sum_{k=n}^{\infty}\int_{s_{k-1}}^{s_k} t\xi^{t-1}d\xi \frac{1}{m(B_k(x))} \\ &= \sum_{k=n}^{\infty}\int_{\sigma_k}^{\sigma_{k-1}} t(1-\lambda)^{t-1}d\lambda\frac{1}{m(B_k(x))} \\ &= \sum_{k=n}^{\infty}\int_{\sigma_k}^{\sigma_{k-1}} t(1-\lambda)^{t-1}N(x,\lambda)\,d\lambda \\ &= \int_0^{\sigma_{n-1}} t(1-\lambda)^{t-1}N(x,\lambda)\,d\lambda. \end{aligned}$$

For the case $n=0$ it follows from (3.5) and (3.7) that

$$\begin{aligned} p(t,x,x) &= \sum_{k=0}^{\infty} s_k^t\left(\frac{1}{m(B_k(x))}-\frac{1}{m(B_{k+1}(x))}\right) \\ &= \sum_{k=0}^{\infty}(1-\sigma_k)^t\left(N(x,\sigma_k)-N(x,\sigma_{k+1})\right) \\ &= \sum_{k=0}^{\infty}\int_{\{\sigma_k\}}(1-\lambda)^t\,dN(x,\lambda) \\ &= \int_{[0,1]}(1-\lambda)^t\,dN(x,\lambda). \end{aligned}$$

■

Corollary 3.6. *The heat kernel satisfies the following properties.*

(a) *For any fixed $x\in X$ and $t>0$, the function $y\mapsto p(t,x,y)$ is strictly monotone decreasing in $d(x,y)$* (recall that by Corollary 3.4 the function $y\mapsto p(t,x,y)$ depends on $d(x,y)$ only).

(b) *For any $x\in X$, the function $t\mapsto p(t,x,x)$ is strictly monotone decreasing in t and*

$$\lim_{t\to 0} p(t,x,x) = \frac{1}{m(x)}.$$

(c) *For all distinct $x,y\in X$ and $t>0$,*

$$p(t,x,y) < p(t,x,x) < \frac{1}{m(x)}. \tag{3.10}$$

Proof. (a) The right hand side of (3.8) is strictly monotone increasing in $\sigma_{n-1}\in(0,1]$ while σ_{n-1} is strictly monotone decreasing in $n=d(x,y)$, whence it follows that $p(t,x,y)$ is strictly monotone decreasing in n.

(b) That $p(t,x,x)$ is strictly monotone decreasing in t follows obviously from (3.9). By the monotone convergence theorem we obtain from (3.9)

$$\lim_{t\to 0} p(t,x,x) = \int_{[0,1]} dN(x,\lambda) = \frac{1}{m(x)}.$$

(c) This follows from (a) and (b). ■

Corollary 3.7. *If $x\neq y$ then $p(t,x,y)\to 0$ as $t\to 0$.*

Proof. Let $n = d(x,y) \geq 1$. Then $\sigma := \sigma_{n-1} < 1$ and we obtain by (3.8)

$$\begin{aligned} p(t,x,y) &\leq N(x,\sigma)\int_0^\sigma t(1-\lambda)^{t-1}d\lambda \\ &= N(x,\sigma)\int_0^\sigma -d(1-\lambda)^t \\ &= N(x,\sigma)\left(1-(1-\sigma)^t\right) \underset{t\to 0}{\longrightarrow} 0. \end{aligned}$$

■

Corollary 3.8. *For all $x, y \in X$, we have*

$$p(t,x,y) \sim p(t,x,x) \quad \text{as } t \to \infty.$$

Proof. Denoting $\sigma = \sigma_{n-1}$ where $n = d(x,y)$, we obtain by Theorem 3.5,

$$\begin{aligned} p(t,x,x) - p(t,x,y) &= t\int_\sigma^1 N(x,\lambda)(1-\lambda)^{t-1}d\lambda \\ &\leq N(x,1)\int_\sigma^1 -d(1-\lambda)^t \\ &= N(x,1)(1-\sigma)^t. \end{aligned}$$

On the other hand, by (3.9)

$$p(t,x,x) \geq \int_{[0,\sigma/2]} (1-\lambda)^t dN(x,\lambda) \geq (1-\sigma/2)^t N(x,\sigma/2).$$

Using (3.10), $\sigma > 0$ and $N(x,\sigma/2) > 0$, we obtain

$$\frac{|p(t,x,x) - p(t,x,y)|}{p(t,x,x)} \leq \left(\frac{1-\sigma}{1-\sigma/2}\right)^t \frac{N(x,1)}{N(x,\sigma/2)} \to 0 \quad \text{as } t \to \infty,$$

whence the claim follows. ■

Corollary 3.9. *There is $\delta = \delta(\sigma_0) > 0$ such that the following inequalities are true for any $x \in X$ and $t > 0$:*

$$\int_0^\infty e^{-\lambda\delta t}dN(x,\lambda) \leq p(t,x,x) \leq \int_0^\infty e^{-\lambda t}dN(x,\lambda) \tag{3.11}$$

Proof. By (3.9) we have

$$p(t,x,x) = \int\limits_{[0,\sigma_0]} (1-\lambda)^t dN(x,\lambda), \tag{3.12}$$

because the measure $dN(x,\lambda)$ vanishes in $(\sigma_0, +\infty)$. Define δ by

$$\delta = -\frac{\log(1-\sigma_0)}{\sigma_0},$$

which is equivalent to

$$e^{-\sigma_0\delta} = 1 - \sigma_0.$$

Then the following inequality is true for all $\lambda \in [0, \sigma_0]$:

$$e^{-\delta\lambda} \leq 1 - \lambda \leq e^{-\lambda}. \tag{3.13}$$

Substituting into (3.12) and noticing that the integration can be extended to $[0,\infty)$, we obtain (3.11). ■

3.3. **Basic estimates of the heat kernel.** The purpose of this and the next section is to provide the heat kernel estimates for large values of t.

Theorem 3.10. *Put $n = d(x, y)$ and $\tau = \frac{1}{\sigma_{n-1}}$.*

(a) *For all $1 \le t \le \tau$,*

$$\frac{1}{2e}\frac{t}{\tau}N\left(x, \frac{1}{2\tau}\right) \le p(t, x, y) \le \frac{t}{\tau}N\left(x, \frac{1}{\tau}\right). \tag{3.14}$$

(b) *For all $t \ge \tau$,*

$$p(t, x, y) \ge \frac{1}{2e}N\left(x, \frac{1}{2t}\right). \tag{3.15}$$

In particular, for all $t \ge 1$,

$$p(t, x, x) \ge \frac{1}{2e}N\left(x, \frac{1}{2t}\right). \tag{3.16}$$

Proof. (a) We have by (3.8)

$$p(t, x, y) = t\int_0^{1/\tau} N(x, \lambda)(1-\lambda)^{t-1}d\lambda. \tag{3.17}$$

Since $(1-\lambda)^{t-1} \le 1$ and $N(x, \lambda)$ is monotone increasing in λ, we obtain

$$p(t, x, y) \le t\int_0^{1/\tau} N(x, \frac{1}{\tau})d\lambda = \frac{t}{\tau}N\left(x, \frac{1}{\tau}\right).$$

To prove the lower bound in (3.14), consider the two cases. If $\tau = 1$ then necessarily $t = 1$. Using the monotonicity of $N(x, \lambda)$ in λ, we obtain from (3.17)

$$p(t, x, y) \ge \int_{1/2}^{1} N(x, \lambda)\, d\lambda \ge \frac{1}{2}N\left(x, \frac{1}{2}\right).$$

If $\tau > 1$ then we have by (3.17)

$$\begin{aligned} p(t, x, y) &\ge t\int_{1/(2\tau)}^{1/\tau} N(x, \lambda)(1-\lambda)^{t-1}d\lambda \\ &\ge \frac{t}{2\tau}N\left(x, \frac{1}{2\tau}\right)(1-\frac{1}{\tau})^{t-1} \\ &\ge \frac{t}{2\tau}N\left(x, \frac{1}{2\tau}\right)\left(1-\frac{1}{\tau}\right)^{\tau-1}, \end{aligned}$$

where we have used that $t \le \tau$. We are left to notice that

$$\inf_{\tau>1}\left(1-\frac{1}{\tau}\right)^{\tau-1} = e^{-1}.$$

(b) It follows from (3.17) and $1/\tau \ge 1/t$ that

$$\begin{aligned} p(t, x, y) &\ge t\int_{1/(2t)}^{1/t} N(x, \lambda)(1-\lambda)^{t-1}d\lambda \\ &\ge N(x, \frac{1}{2t})t\int_{1/(2t)}^{1/t}(1-\frac{1}{t})^{t-1}d\lambda \\ &\ge \frac{1}{2}N\left(x, \frac{1}{2t}\right)\left(1-\frac{1}{t}\right)^{t-1} \\ &\ge \frac{1}{2e}N\left(x, \frac{1}{2t}\right), \end{aligned}$$

which proves (3.15). Then (3.16) follows from (3.15) because in the case $x = y$ we have $n = 0$ and $\tau = \frac{1}{\sigma_{-1}} = 1$. ∎

A non-decreasing function $F : \mathbb{R}_+ \to \mathbb{R}_+$ is said to be doubling if there exists a constant $D > 0$ such that

$$F(2s) \leq DF(s) \text{ for all } s > 0.$$

If F is doubling then, for all $0 < s_1 < s_2$,

$$F(s_2) \leq D \left(\frac{s_2}{s_1}\right)^\delta F(s_1), \tag{3.18}$$

where $\delta = \log_2 D$.

Theorem 3.11. *If the function $\lambda \mapsto N(x, \lambda)$ is doubling then, for all $t \geq 1$,*

$$cN\left(x, \frac{1}{t}\right) \leq p(t, x, x) \leq CN\left(x, \frac{1}{t}\right) \tag{3.19}$$

where $C, c > 0$ depend on the doubling constant.

Proof. The lower bound in (3.19) follows from (3.16) and

$$N\left(x, \frac{1}{2t}\right) \geq D^{-1} N\left(x, \frac{1}{t}\right).$$

To prove the upper bound in (3.19), consider two cases. If $t < 2$ then by (3.10) and the doubling property

$$p(t, x, x) \leq \frac{1}{m(x)} = N(x, 1) \leq CN\left(x, \frac{1}{t}\right).$$

Let now $t \geq 2$. Making change $u = \lambda t$ in the integral (3.8), we obtain

$$\begin{aligned}
p(t, x, x) &= t \int_0^1 N(x, \lambda)(1 - \lambda)^{t-1} d\lambda \\
&= \int_0^t N\left(x, \frac{u}{t}\right)\left(1 - \frac{u}{t}\right)^{t-1} du \\
&\leq \int_0^t D \max(1, u)^\delta N\left(x, \frac{1}{t}\right)\left(1 - \frac{u}{t}\right)^{t-1} du \\
&\leq DN\left(x, \frac{1}{t}\right) \int_0^\infty \max(1, u)^\delta e^{-\frac{1}{2}u} du \\
&= CN\left(x, \frac{1}{t}\right),
\end{aligned}$$

where we have used the doubling property (3.18) and the inequality

$$\left(1 - \frac{u}{t}\right)^{t-1} \leq e^{-\frac{u}{t}(t-1)} \leq e^{-\frac{1}{2}u},$$

that is true for all $t \geq 2$ and $u \geq 0$. ∎

Let us introduce the following two conditions:

Condition 3.12. *There exists a constant $0 < \kappa_+ < 1$ such that for all $k \geq 0$,*

$$\sigma_{k+1} \leq \kappa_+ \sigma_k.$$

Condition 3.13. *There exists a constant $\nu_+ > 1$ such that for all $k \geq 0$,*

$$m(B_{k+1}(x)) \leq \nu_+ m(B_k(x)).$$

The point x is fixed here.

Proposition 3.14. *Assume that the conditions 3.12 and 3.13 are satisfied for some x. Then the function $\lambda \to N(x,\lambda)$ is doubling. Consequently, the heat kernel satisfies the estimate*

$$p\,(t,x,x) \simeq N\left(x,\frac{1}{t}\right) \tag{3.20}$$

for all $t \geq 1$.

Proof. Assume first that $0 < 2\lambda < 1$. Then there is a non-negative integer k such that $\sigma_k \leq 2\lambda < \sigma_{k-1}$. Hence, for any $l \geq 1$, we have

$$\begin{aligned} N(x,2\lambda) &= \frac{1}{m(B_k\,(x))} \leq \frac{\nu_+^l}{m(B_{k+l}\,(x))} \\ &= \nu_+^l N(x,\sigma_{k+l}) \\ &\leq \nu_+^l N\left(x,\kappa_+^l\sigma_k\right). \end{aligned}$$

Choose l big enough so that $\kappa_+^l < 1/2$. Then letting $C = \nu_+^l$ we obtain

$$N(x,2\lambda) \leq CN(x,\frac{1}{2}\sigma_k) \leq CN(x,\lambda).$$

If $2\lambda \geq 1$ then

$$N\,(x,2\lambda) = N\,(x,\sigma_0) \leq CN\left(x,\frac{1}{2}\sigma_0\right) \leq CN\,(x,\lambda)\,.$$

Finally, the estimate (3.20) follows from (3.19). ■

3.4. **Change of metric and off-diagonal estimates.** For all $x,y \in X$ set

$$d_\sigma\,(x,y) = \frac{1}{\sigma_{n-1}} - 1, \tag{3.21}$$

where $n = d\,(x,y)$. Note that the right hand side of (3.21) is monotone increasing in n because σ_{n-1} is decreasing, and vanishes at $n=0$ since $\sigma_{-1} = 1$. It follows that $d_\sigma\,(x,y)$ satisfies the ultra-metric inequality. It is clear that $d_\sigma\,(x,y) = d_\sigma\,(y,x)$. Also, $d_\sigma\,(x,y) = 0$ if and only if $\sigma_{n-1} = 1$ which is equivalent to $n = 0$, that is, to $x = y$. Hence, $d_\sigma\,(x,y)$ is an ultra-metric.

The heat kernel estimates can be conveniently stated in terms of the metric d_σ. In the next statements, we write d_σ for $d_\sigma\,(x,y)$.

Theorem 3.15. *The following estimates hold for all $x,y \in X$ and $t \geq 1$:*

$$p(t,x,y) \geq \frac{1}{2e}\frac{t}{t+d_\sigma}N\left(x,\frac{1}{2\,(t+d_\sigma)}\right). \tag{3.22}$$

and

$$p(t,x,y) \leq 4e\frac{t}{t+d_\sigma}p\left(\frac{t+d_\sigma}{4},x,x\right). \tag{3.23}$$

Proof. Let us use the notation from Theorem 3.10: $\tau = \frac{1}{\sigma_{n-1}}$ where $n = d\,(x,y)\,,$ so that $d_\sigma = \tau - 1$. Consider two cases.

If $t \leq \tau$ then we use the lower bound in (3.14) and

$$\tau = d_\sigma + 1 \leq d_\sigma + t$$

to obtain

$$p(t,x,y) \geq \frac{1}{2e}\frac{t}{\tau}N\left(x,\frac{1}{2\tau}\right) \geq \frac{1}{2e}\frac{t}{t+d_\sigma}N\left(x,\frac{1}{2(t+d_\sigma)}\right),$$

which proves (3.22) in this case. To prove (3.23), we use $t+d_\sigma \le 2\tau$, (3.16), and the upper bound in (3.14):

$$\frac{t}{t+d_\sigma}p\left(\frac{t+d_\sigma}{4},x,x\right) \ge \frac{t}{2\tau}p\left(\frac{\tau}{2},x,x\right) \ge \frac{t}{4e\tau}N\left(x,\frac{1}{\tau}\right) \ge \frac{1}{4e}p\left(t,x,y\right).$$

If $t>\tau$, then by (3.15)

$$p(t,x,y) \ge \frac{1}{2e}N\left(x,\frac{1}{2t}\right) \ge \frac{t}{2e\left(t+d_\sigma\right)}N\left(x,\frac{1}{2\left(t+d_\sigma\right)}\right),$$

which proves (3.22). To prove (3.23) observe that

$$t+d_\sigma \le 2t$$

whence

$$\frac{t}{t+d_\sigma}p\left(\frac{t+d_\sigma}{4},x,x\right) \ge \frac{1}{2}p\left(t,x,x\right) \ge \frac{1}{2}p(t,x,y),$$

which finishes the proof. ∎

Corollary 3.16. *Assume that,*

$$p\left(t,x,x\right) \le C_1 N\left(x,\frac{C_2}{t}\right) \tag{3.24}$$

for some $x \in X$ and all $t \ge 1$. Then, for all $y \in X$ and $t \ge 1$,

$$\frac{1}{2e}\frac{t}{t+d_\sigma}N\left(x,\frac{1}{2\left(t+d_\sigma\right)}\right) \le p(t,x,y) \le \frac{C_1' t}{t+d_\sigma}N\left(x,\frac{C_2'}{t+d_\sigma}\right) \tag{3.25}$$

where $C_1' = 4eC_1$ and $C_2' = 4C_2$.

Proof. The lower bound in (3.25) coincides with (3.22). The upper bound follows from (3.23) and (3.24) as follows:

$$\begin{aligned} p(t,x,y) &\le 4e\frac{t}{t+d_\sigma}p\left(\frac{t+d_\sigma}{4},x,x\right) \\ &\le 4eC_1\frac{t}{t+d_\sigma}N\left(x,\frac{4C_2}{t+d_\sigma}\right) \end{aligned}$$

∎

Corollary 3.17. *If the function $\lambda \mapsto N\left(x,\lambda\right)$ is doubling for some $x \in X$ then, for all $y \in X$ and $t \ge 1$,*

$$p\left(t,x,y\right) \simeq \frac{t}{t+d_\sigma}N\left(x,\frac{1}{t+d_\sigma}\right) \tag{3.26}$$

where the constants that bound the ratio of the two sides in (3.26) depend only on the doubling constant.

Proof. Indeed, this is a combination of Corollary 3.16 with Theorem 3.11. ∎

4. Green function and transience

The potential operator R associated with the semigroup $\left\{P^t\right\}_{t\ge0}$ is defined as follows

$$Rf = \int_0^\infty P^t f\ dt,$$

for any non-negative function f on X. The semigroup $\left\{P^t\right\}_{t\ge0}$ is called *transient* if $Rf<\infty$ for any non-negative function f with finite support, and *recurrent* otherwise.

Define the Green function $r(x,y)$ by

$$r(x,y)=\int_0^{\infty} p(t,x,y)\,dt.$$

Theorem 4.1. *The following conditions are equivalent.*

(i) The semigroup $\{P^t\}_{t\geq 0}$ *is transient.*

(ii) $r(x,y)<\infty$ *for all* $x,y\in X$

(iii) For all/some $x\in X$,

$$\int_0 N(x,\lambda)\frac{d\lambda}{\lambda^2}<\infty. \tag{4.1}$$

In this case the Green function is the integral kernel of R *and it is given by*

$$r(x,y)=\int_0^{\sigma_{n-1}}\frac{N(x,\lambda)d\lambda}{(1-\lambda)\left(\log\frac{1}{1-\lambda}\right)^2}, \tag{4.2}$$

where $n=d(x,y)$.

Proof. Using the definitions of R and $r(x,y)$, we obtain

$$Rf(x)=\int_0^{\infty}\int_X p(t,x,y)\,f(y)\,dm(y)\,dt=\int_X r(x,y)\,f(y)\,dm(y)\,,$$

so that $Rf<\infty$ if and only if $r(x,y)<\infty$ for all x,y, and in the latter case $r(x,y)$ is indeed the integral kernel of R. Integrating the identity (3.8) in t, we obtain

$$\begin{aligned} r(x,y) &= \int_0^{\infty} t\int_0^{\sigma_{n-1}} N(x,\lambda)(1-\lambda)^{t-1}d\lambda dt \\ &= \int_0^{\sigma_{n-1}} N(x,\lambda)\left(\int_0^{\infty} t\,(1-\lambda)^{t-1}\,dt\right)d\lambda \\ &= \int_0^{\sigma_{n-1}}\frac{N(x,\lambda)d\lambda}{(1-\lambda)\left(\log\frac{1}{1-\lambda}\right)^2}, \end{aligned}$$

which proves (4.2). The finiteness of $r(x,y)$ for all x,y is equivalent to

$$\int_0^1\frac{N(x,\lambda)d\lambda}{(1-\lambda)\left(\log\frac{1}{1-\lambda}\right)^2}<\infty,$$

for all $x\in X$. Clearly, this integral converges always at 1 while the convergence at 0 is equivalent to (4.1).

Finally, let us show that the convergence of the integral (4.1) for some x implies that for all x. It suffices to show that, for all $x,y\in X$,

$$N(x,\lambda)=N(y,\lambda)$$

for small enough λ. Indeed, by (3.7), we have

$$N(x,\lambda)=\frac{1}{m(B_k(x))}\quad \text{if}\quad \lambda\in[\sigma_k,\sigma_{k-1}).$$

If λ is small enough then $k>d(x,y)$ and $B_k(x)=B_k(y)$ whence the claim follows. ■

Corollary 4.2. *If the Green function is finite then it satisfies the following properties.*

(a) For all distinct $x,y\in X$,

$$0<r(x,y)<r(x,x)\,.$$

(b) The function $x,y\mapsto\frac{1}{r(x,y)}$ *satisfies the ultra-metric inequality.*

(c) *The following asymptotic holds*

$$r(x,y) \sim \int_0^{\frac{1}{1+d_\sigma(x,y)}} N(x,\lambda)\frac{d\lambda}{\lambda^2} \tag{4.3}$$

as $d_\sigma(x,y) \to \infty$.

Proof. (a) This follows from (4.2) and $0 < \sigma_{n-1} < 1$ where $n = d(x,y) \geq 1$.

(b) By 4.2, $r(x,y)$ is an increasing function of σ_{n-1} and, hence, is a decreasing function of $n = d(x,y)$. Therefore, $\frac{1}{r(x,y)}$ is an increasing function of $d(x,y)$ and, hence, satisfies the ultra-metric inequality.

(c) This is a consequence of (4.2) since $\sigma_{n-1} = \frac{1}{1+d_\sigma(x,y)} \to 0$. ■

Corollary 4.3. *Assume that there exist constants* $0 < \varepsilon'' < \varepsilon' < \varepsilon < 1$ *such that*

$$\varepsilon'' \leq \frac{N(x,\varepsilon\lambda)}{N(x,\lambda)} \leq \varepsilon' \tag{4.4}$$

for some $x \in X$ *and all* $\lambda \in (0,1)$. *Then the semigroup* $\{P^t\}_{t\geq 0}$ *is transient and*

$$r(x,y) \simeq d_\sigma(x,y) N\left(x, \frac{1}{d_\sigma(x,y)}\right) \tag{4.5}$$

for all $y \in X$ *with large enough* $d_\sigma(x,y)$.

Proof. For any $\Lambda \in (0,1]$ we have

$$\begin{aligned}
\int_0^\Lambda N(x,\lambda)\frac{d\lambda}{\lambda^2} &= \sum_{k=0}^\infty \int_{\varepsilon^{k+1}\Lambda}^{\varepsilon^k\Lambda} N(x,\lambda)\frac{d\lambda}{\lambda^2} \quad [\text{change } u = \varepsilon^{-k}\lambda] \\
&= \sum_{k=0}^\infty \int_{\varepsilon\Lambda}^{\Lambda} N(x,\varepsilon^k u)\varepsilon^{-k}\frac{du}{u^2} \\
&\leq \sum_{k=0}^\infty \int_{\varepsilon\Lambda}^{\Lambda} N(x,u)\left(\varepsilon'\right)^k \varepsilon^{-k}\frac{du}{u^2} \\
&\leq \frac{1}{1-\varepsilon'/\varepsilon}(1-\varepsilon)\Lambda N(x,\Lambda)\frac{1}{(\varepsilon\Lambda)^2} \\
&= \frac{1-\varepsilon}{\varepsilon^2(1-\varepsilon'/\varepsilon)}\frac{N(x,\Lambda)}{\Lambda}
\end{aligned}$$

and

$$\int_0^\Lambda N(x,\lambda)\frac{d\lambda}{\lambda^2} \geq \int_{\varepsilon\Lambda}^\Lambda N(x,\lambda)\frac{d\lambda}{\lambda^2} \geq (1-\varepsilon)\Lambda N(x,\varepsilon\Lambda)\frac{1}{\Lambda^2} \geq \varepsilon''(1-\varepsilon)\frac{N(x,\Lambda)}{\Lambda},$$

whence

$$\int_0^\Lambda N(x,\lambda)\frac{d\lambda}{\lambda^2} \simeq \frac{N(x,\Lambda)}{\Lambda}.$$

Setting here $\Lambda = \frac{1}{1+d_\sigma(x,y)}$, using (4.3) and noticing that $\Lambda \sim \frac{1}{d_\sigma(x,y)}$ for large enough $d_\sigma(x,y)$ we obtain (4.5). ■

5. Examples

Here we present some examples of sequences $\{\sigma_k\}$ and $\{m(B_k(x))\}$ where the above results yield explicit estimates of the heat kernel and the Green function.

Example 5.1. Assume that, for some $x \in X$ and all $k \in \mathbb{Z}_+$

$$m(B_k(x)) \simeq p^k, \tag{5.1}$$

where $p > 1$. For example, this is the case when $X = \mathbb{Z}(p)^{(\infty)}$ or $X = \Lambda_p$ (the p-adic net) with counting measure m. Clearly, (5.1) implies that Condition 3.13 is satisfied. Assume also that the sequence $\{\sigma_k\}_{k=0}^{\infty}$ is satisfied the estimate

$$\sigma_k \simeq \sigma(k)$$

where $\sigma(k)$ is a continuous, strictly monotone decreasing function of $k \in [0, +\infty)$ such that $\sigma(k) \to 0$ as $k \to \infty$. It follows from (3.7) that

$$N(x,\lambda) \asymp p^{-\sigma^{-1}(\lambda)}, \tag{5.2}$$

for all $0 < \lambda \le \sigma_0$. Note that if the function $\lambda \mapsto N(x,\lambda)$ is doubling then the sign $\asymp$ in (5.2) can be replaced by $\simeq$. Consider some specific examples of $\sigma(k)$.

1. Let $\sigma(k) = a^{-k}$, for some $a > 1$. Condition 3.12 holds so that the function $\lambda \to N(x,\lambda)$ is doubling. Then by (5.2)

$$N(x,\lambda) \simeq \lambda^{\alpha}$$

for small λ, where $\alpha = \frac{\log p}{\log a}$. It follows from (3.20) that

$$p(t,x,x) \simeq N\left(x, \frac{1}{t}\right) \simeq t^{-\alpha},$$

for large enough t. Then Corollary 3.16 yields, for all $y \in X$ and large enough t,

$$p(t,x,y) \simeq \frac{t}{(t + d_\sigma(x,y))^{\alpha+1}}.$$

In connection with this example recall that if $h_\beta(t,x,y)$ is the transition function of the symmetric stable process in $\mathbb{R}^d$ of index $0 < \beta < 2$, then according to [5] (see also [2]),

$$h_\beta(t,x,y) \simeq \frac{t}{(t^{1/\beta} + |x-y|)^{d+\beta}}.$$

Our result shows that $p(t,x,y)$ has a shape similar to that of the 1-stable law of "dimension α".

The transience test (4.1) is satisfied if and only if $\alpha > 1$ (that is, $p > a$). In this case we obtain from (4.3) or (4.5) the following estimate of the Green function

$$r(x,y) \simeq d_\sigma(x,y)^{1-\alpha}$$

for large $d_\sigma(x,y)$.

2. Let

$$\sigma(k) = \exp(-a^k),$$

for some constant $a > 1$. Condition 3.12 holds and, hence, the function $\lambda \to N(x,\lambda)$ is doubling. It follows from (5.2) that, for small enough λ,

$$N(x,\lambda) \simeq \left(\log \frac{1}{\lambda}\right)^{-\alpha},$$

where $\alpha = \frac{\log p}{\log a}$, which implies by (3.20)

$$p\left(t,x,x\right) \simeq N\left(x,\frac{1}{t}\right) \simeq \left(\log t\right)^{-\alpha},$$

for large enough t. By Corollary 3.16 we obtain

$$p\left(t,x,y\right) \simeq \frac{t}{\left(t+d_{\sigma}\left(x,y\right)\right)\log^{\alpha}\left(t+d_{\sigma}\left(x,y\right)\right)}$$

for all $y \in X$ and large enough t.

The transience test (4.1) fails so that in this case the semigroup $\left\{P^{t}\right\}_{t\geq 0}$ is recurrent.

3. Let $\sigma\left(k\right) = k^{-a}$, $k \geq 1$, for some $a > 0$. In this case Condition 3.12 fails. By (5.2) we obtain

$$N\left(x,\lambda\right) \asymp \exp\left\{-\left(\frac{1}{\lambda}\right)^{\frac{1}{a}}\right\}$$

for small λ. A direct estimate of the integrals in (3.11) yields

$$p\left(t,x,x\right) \asymp \exp\left(-t^{\frac{1}{a+1}}\right),$$

for large t. Then Theorem 3.15 gives the following results

$$p\left(t,x,y\right) \geq \frac{C_1 t}{t+d_{\sigma}(x,y)}\exp\{-C_1'(t+d_{\sigma}(x,y))^{\frac{1}{a}}\}$$

and

$$p(t,x,y) \leq \frac{C_2 t}{t+d_{\sigma}(x,y)}\exp\{-C_2'(t+d_{\sigma}(x,y))^{\frac{1}{a+1}}\},$$

that are not as precise as in the previous two cases.

The transience test (4.1) is clearly satisfied, and we obtain from (4.3)

$$r\left(x,y\right) \asymp \exp\left(-d_{\sigma}\left(x,y\right)^{\frac{1}{a}}\right)$$

for large $d_{\sigma}\left(x,y\right)$.

4. Let $\sigma(k) = (\log_{(n)} k)^{-a}$ for large enough k,where $a > 0$ and

$$\log_{(n)} = \underbrace{\log\ldots\log}_{n \text{ times}}$$

Condition 3.12 does not hold. By (5.2) we obtain

$$N(x,\lambda) \asymp \exp\left\{-\exp_{(n)}\left(\frac{1}{\lambda}\right)^{\frac{1}{a}}\right\},$$

where

$$\exp_{(n)} = \underbrace{\exp\exp\ldots\exp}_{n \text{ times}}.$$

Using (3.11) it is possible to show that

$$p\left(t,x,x\right) \asymp \exp\left(-\frac{t}{(\log_{(n)} t)^{a}}\right),$$

for large enough t (see [3] for the details of the computation).

Example 5.2. Let m_0 be a counting measure on a discrete ultra-metric space (X,d) that satisfies

$$m_0\left(B_k\left(x\right)\right)=p^k$$

for all $x\in X$ and $k\in\mathbb{Z}_+$, where $p>1$ is a constant. Fix a reference point $o\in X$ and introduce the notation

$$|x|=d\left(x,o\right).$$

Let m be another measure on X that satisfy the relation

$$m\left(x\right)\simeq q^{|x|}, \tag{5.3}$$

where $q>1$ is a constant. For example, in the case of the p-adic net Λ_p we have

$$|x|=\log_p^+\|x\|_p$$

and (5.3) means that

$$m\left(x\right)\simeq\|x\|_p^{\gamma}$$

where $\gamma=\frac{\log q}{\log p}$.

We claim that

$$m\left(B_k\left(x\right)\right)\simeq p^kq^{\max(|x|,k)}=\left\{\begin{array}{ll}(pq)^k, & |x|\leq k,\\ p^kq^{|x|}, & |x|>k.\end{array}\right. \tag{5.4}$$

Indeed, if the balls $B_k\left(x\right)$ and $B_k\left(o\right)$ are disjoint then, for any $y\in B_k\left(x\right)$, we have $|y|>k$. It follows that

$$|x|\leq\max\left(|y|,d\left(x,y\right)\right)=|y|$$

and in the same way $|y|\leq|x|$ so that $|y|=|x|$. Therefore, $m\left(y\right)\simeq q^{|x|}$ and

$$m\left(B_k\left(x\right)\right)\simeq q^{|x|}m_0\left(B_k\left(x\right)\right)=p^kq^{|x|}.$$

If the balls $B_k\left(x\right)$ and $B_k\left(o\right)$ intersect then they coincide and we obtain

$$\begin{aligned} m\left(B_k\left(x\right)\right) &= m\left(B_k\left(o\right)\right)\\ &\simeq 1+\sum_{i=1}^{k}\left(p^i-p^{i-1}\right)q^i\\ &= 1+q\left(p-1\right)\frac{(pq)^k-1}{pq-1}\\ &\simeq p^kq^k,\end{aligned}$$

which finishes the proof of (5.4).

Note that, for any fixed x the function (5.4) satisfies Condition 3.13 with a constant ν_+ that is the same for all $x\in X$. Now let us consider two examples of the sequence $\{\sigma_k\}$ as in Example 5.1, that satisfy Condition 3.12. Hence, the function $\lambda\mapsto N\left(x,\lambda\right)$ is doubling and we can obtain the heat kernel bounds by Theorem 3.11 and Corollary 3.17.

1. Let

$$\sigma_k=a^{-(k+1)}$$

where $a>1$. Then by (3.7)

$$N\left(x,\lambda\right)\simeq p^{-k}q^{-\max(|x|,k)}\ \text{ if }\ a^{-(k+1)}\leq\lambda<a^{-k},$$

where the latter condition implies

$$\lambda\simeq a^{-k}.$$

It follows that, for small enough λ,

$$N\left(x,\lambda\right)\simeq\lambda^{\alpha}q^{-\max\left(|x|,\log_a\frac{1}{\lambda}\right)}=\left\{\begin{array}{ll}\lambda^{\alpha+\beta}, & \lambda\leq a^{-|x|},\\ \lambda^{\alpha}q^{-|x|}, & \lambda>a^{-|x|},\end{array}\right. \tag{5.5}$$

where

$$\alpha=\frac{\log p}{\log a}\quad\text{and}\quad\beta=\frac{\log q}{\log a}. \tag{5.6}$$

By (3.20) we obtain, for large enough t,

$$p\left(t,x,x\right)\simeq t^{-\alpha}q^{-\max\left(|x|,\log_a t\right)}=\begin{cases}t^{-(\alpha+\beta)}, & t\geq a^{|x|},\\ t^{-\alpha}q^{-|x|}, & t<a^{|x|}.\end{cases}$$

In terms of the metric d_σ, we have

$$d_\sigma\left(x,o\right)=\frac{1}{\sigma_{|x|-1}}-1=a^{|x|}-1$$

whence

$$|x|=\log_a\left(1+d_\sigma\left(x,o\right)\right).$$

By Corollary 3.17 we obtain

$$p\left(t,x,y\right)\simeq t\lambda N\left(x,\lambda\right) \tag{5.7}$$

where

$$\lambda=\frac{1}{t+d_\sigma\left(x,y\right)}.$$

Substituting this value of λ into (5.5) and (5.7) and noticing that

$$\begin{aligned}\max\left(|x|,\log_a\frac{1}{\lambda}\right) &= \log_a\max\left(1+d_\sigma\left(x,o\right),t+d_\sigma\left(x,y\right)\right)\\ &\simeq \log_a\left(t+d_\sigma\left(x,y\right)+d_\sigma\left(x,o\right)\right)\end{aligned}$$

we obtain

$$p\left(t,x,y\right)\simeq\frac{t}{\left(t+d_\sigma\left(x,y\right)\right)^{1+\alpha}\left(t+d_\sigma\left(x,y\right)+d_\sigma\left(x,o\right)\right)^{\beta}}.$$

In particular, for $x=o$ we have

$$p\left(t,x,y\right)\simeq\frac{t}{\left(t+d_\sigma\left(x,y\right)\right)^{1+\alpha+\beta}}.$$

The transience test (4.1) is satisfied if and only if $\alpha+\beta>1$, that is, if $pq>a$. In this case we obtain by (4.3) and (5.5) that

$$r\left(x,y\right)\simeq d_\sigma\left(x,y\right)^{1-(\alpha+\beta)}$$

for large enough $d_\sigma\left(x,y\right)$.

2. Put now

$$\sigma_k=\exp\left(1-a^{k+1}\right),$$

where $a>1$. Then by (3.7)

$$N\left(x,\lambda\right)\simeq p^{-k}q^{-\max(|x|,k)}\ \text{ if }\ e^{1-a^{k+1}}\leq\lambda<e^{1-a^k},$$

where the latter condition implies

$$\log\frac{1}{\lambda}\simeq a^k.$$

It follows that, for small enough λ,

$$N\left(x,\lambda\right)\ \simeq\ \left(\log\frac{1}{\lambda}\right)^{-\alpha}q^{-\max\left(|x|,\log_a\log\frac{1}{\lambda}\right)} \tag{5.8}$$

$$=\begin{cases}\left(\log\frac{1}{\lambda}\right)^{-(\alpha+\beta)}, & \lambda\leq\exp\left(-a^{|x|}\right),\\ \left(\log\frac{1}{\lambda}\right)^{-\alpha}q^{-|x|}, & \lambda>\exp\left(-a^{|x|}\right).\end{cases} \tag{5.9}$$

where α and β are given by (5.6). By (3.20) we obtain, for large enough t,

$$\begin{aligned} p(t,x,x) &\simeq (\log t)^{-\alpha} q^{-\max(|x|,\log_a \log t)} \\ &= \begin{cases} (\log t)^{-(\alpha+\beta)}, & t \geq \exp\left(a^{|x|}\right), \\ (\log t)^{-\alpha} q^{-|x|}, & t < \exp\left(a^{|x|}\right). \end{cases} \end{aligned}$$

In terms of the metric d_σ, we obtain

$$d_\sigma(x,o) = \frac{1}{\sigma_{|x|-1}} - 1 = \exp\left(a^{|x|} - 1\right) - 1$$

whence

$$|x| = \log_a\left(1 + \log\left(1 + d_\sigma(x,o)\right)\right)$$

For

$$\lambda = \frac{1}{t + d_\sigma(x,y)},$$

we obtain from (5.8) and (5.7)

$$p(t,x,y) \simeq \frac{t}{(t + d_\sigma(x,y)) \log^\alpha (t + d_\sigma(x,y)) \log^\beta (t + d_\sigma(x,y) + d_\sigma(x,o))}.$$

In particular, for $x = o$, we have

$$p(t,x,y) \simeq \frac{t}{(t + d_\sigma(x,y)) \log^{\alpha+\beta}(t + d_\sigma(x,y))}.$$

The transience test (4.1) fails so that the semigroup $\left\{P^t\right\}_{t\geq 0}$ is recurrent.

6. The moments of the Markov process

Let $\mathcal{X} = \left(\{\mathcal{X}_t\}_{t\geq 0}, \{\mathbb{P}_x\}_{x\in X}\right)$ be the Markov process associated with the semigroup $\left\{P^t\right\}_{t\geq 0}$. For any $\gamma > 0$, the γ-moment of $\mathcal{X}$ is defined by

$$M_\gamma(x,t) := \mathbb{E}_x\left(d_\sigma(x,\mathcal{X}_t)^\gamma\right), \tag{6.1}$$

where $\mathbb{E}_x$ is expectation associated with the probability measure $\mathbb{P}_x$. In terms of the heat kernel, we have

$$M_\gamma(x,t) = \int_X d_\sigma(x,y)^\gamma\, p(t,x,y)\, dm(y) \tag{6.2}$$

The aim of this section is to estimate the moment $M_\gamma(x,t)$ as a function of t. We precede the main result by a lemma. Consider open balls with respect to the metric d_σ:

$$B_r^\sigma(x) = \{y \in X : d_\sigma(x,y) < r\}.$$

Lemma 6.1. *For all $x \in X$ and $r > 0$,*

$$m\left(B_r^\sigma(x)\right) = \frac{1}{N\left(x, \frac{1}{1+r}\right)}. \tag{6.3}$$

Proof. Set $r_n = \frac{1}{\sigma_n} - 1$. By (3.21), the equation $d(x,y) = n+1$ is equivalent to $d_\sigma(x,y) = r_n$, whence it follows that

$$d_\sigma(x,y) < r_n \Leftrightarrow d(x,y) \leq n.$$

For any $r > 0$ there is $n \geq 0$ such that

$$r_{n-1} < r \leq r_n.$$

Then we have

$$d_\sigma(x,y) < r \Leftrightarrow d_\sigma(x,y) < r_n$$

whence it follows that

$$B_r^\sigma(x) = B_{r_n}^\sigma(x) = B_n(x). \tag{6.4}$$

On the other hand, we have

$$\sigma_n = \frac{1}{1+r_n} \le \frac{1}{1+r} < \frac{1}{1+r_{n-1}} = \sigma_{n-1}$$

whence by (3.7)

$$N\left(x, \frac{1}{1+r}\right) = \frac{1}{m(B_n(x))}.$$

Comparing with (6.4) we obtain (6.3). ■

Theorem 6.2. *For any $\gamma \in (0,1)$ and all $x \in X$, $t > 0$, the following estimate is true:*

$$M_\gamma(x,t) \le \frac{1}{1-\gamma} \min\left(t^\gamma, \frac{t}{\gamma}\right). \tag{6.5}$$

Proof. Set $V(r) = m(B_r^\sigma(x))$ so that by Lemma 6.1

$$V(r) = \frac{1}{N\left(x, \frac{1}{1+r}\right)}.$$

By (3.8) we have

$$p(t,x,y) = t\int_0^{\frac{1}{1+d_\sigma(x,y)}} N(x,\lambda)(1-\lambda)^{t-1}d\lambda,$$

so that $p(t,x,y)$ is a function of t and $r = d_\sigma(x,y)$. Therefore, we obtain from (6.2)

$$\begin{aligned}
M_\gamma(x,t) &= \int_0^\infty r^\gamma p(t,x,y)\,dV(r) \\
&= t\int_0^\infty r^\gamma \left(\int_0^{\frac{1}{1+r}} N(x,\lambda)(1-\lambda)^{t-1}d\lambda\right) dV(r) \\
&= t\int\int_{\left\{0<\lambda<\frac{1}{1+r}, r>0\right\}} r^\gamma N(x,\lambda)(1-\lambda)^{t-1}d\lambda dV(r) \\
&= t\int\int_{\left\{0<r<\frac{1}{\lambda}-1, 0<\lambda<1\right\}} r^\gamma N(x,\lambda)(1-\lambda)^{t-1}d\lambda dV(r) \\
&= t\int_0^1 \left(\int_{(0,\frac{1}{\lambda}-1)} r^\gamma dV(r)\right) N(x,\lambda)(1-\lambda)^{t-1}d\lambda \qquad (6.6)\\
&\le t\int_0^1 \left(\frac{1}{\lambda}-1\right)^\gamma V\left(\frac{1}{\lambda}-1\right) N(x,\lambda)(1-\lambda)^{t-1}d\lambda \\
&= t\int_0^1 \lambda^{-\gamma}(1-\lambda)^{\gamma+t-1}\frac{1}{N(x,\lambda)}N(x,\lambda)d\lambda \\
&= tB(1-\gamma, \gamma+t) \\
&= t\frac{\Gamma(1-\gamma)\Gamma(t+\gamma)}{\Gamma(t+1)},
\end{aligned}$$

where $B(\cdot,\cdot)$ is the beta function and $\Gamma(\cdot)$ is the gamma function.

Let us use the following properties of the gamma function:

(1) For all $t > 0$ and $\gamma \in (0,1)$

$$\frac{\Gamma(t+\gamma)}{\Gamma(t)} \le t^\gamma$$

(cf. Lemma 8.1 in Appendix).

(2) $\Gamma(z) \le \frac{1}{z}$ for all $z \in (0,1)$. Indeed, $z\Gamma(z) = \Gamma(z+1)$, while the gamma function is bounded by 1 on the interval $[1,2]$.

Using these properties, we obtain, for all $t>0$ and $\gamma \in (0,1)$,

$$M_\gamma(x,t) \le t\Gamma(1-\gamma)\frac{\Gamma(t+\gamma)}{\Gamma(t+1)} = \Gamma(1-\gamma)\,t\frac{\Gamma(t+\gamma)}{t\Gamma(t)} \le \frac{1}{1-\gamma}t^\gamma.$$

The function

$$t \mapsto \frac{\Gamma(t+\gamma)}{\Gamma(t+1)}$$

is monotone decreasing in $t \ge 0$ (cf. Lemma 8.1), which implies that

$$\frac{\Gamma(1-\gamma)\Gamma(t+\gamma)}{\Gamma(t+1)} \le \frac{\Gamma(1-\gamma)\Gamma(\gamma)}{\Gamma(1)} = \frac{\pi}{\sin\pi\gamma} \le \frac{1}{\gamma(1-\gamma)},$$

whence

$$M_\gamma(x,t) \le \frac{t}{\gamma(1-\gamma)}.$$

Combining the above two upper bounds of $M_\gamma(x,t)$, we obtain (6.5). ■

Theorem 6.3. *Assume that, for some $x \in X$, the function $\lambda \mapsto N(x,\lambda)$ satisfies the reverse doubling property:*

$$N(x,\lambda) \ge (1+\eta)\,N(x,\delta\lambda) \tag{6.7}$$

for all $\lambda \in (0,1)$ and some $\delta,\eta \in (0,1)$. Then, for all $t>0$ and $\gamma \in (0,1)$,

$$M_\gamma(x,t) \ge \frac{c}{1-\gamma}\min(t^\gamma, t)$$

where $c = c(\delta,\eta) > 0$.

Proof. For any $R>0$ and $\varepsilon = \delta/2$, we have

$$\begin{aligned}\int_{(0,R)} r^\gamma dV(r) &\ge \int_{[\varepsilon R,R)} r^\gamma dV(r) \ge (\varepsilon R)^\gamma (V(R) - V(\varepsilon R)) \\ &= (\varepsilon R)^\gamma \left(\frac{1}{N\left(x,\frac{1}{1+R}\right)} - \frac{1}{N\left(x,\frac{1}{1+\varepsilon R}\right)}\right).\end{aligned}$$

Note that

$$\frac{\frac{1}{1+R}}{\frac{1}{1+\varepsilon R}} = \frac{1+\varepsilon R}{1+R} \le \delta$$

provided

$$R \ge \frac{1-\delta}{\delta-\varepsilon}.$$

In this case, we obtain by (6.7)

$$\frac{1}{N\left(x,\frac{1}{1+R}\right)} - \frac{1}{N\left(x,\frac{1}{1+\varepsilon R}\right)} \ge \frac{1}{N\left(x,\frac{1}{1+R}\right)}\left(1 - \frac{1}{1+\eta}\right)$$

whence

$$\int_{(0,R)} r^\gamma dV(r) \ge \frac{\varepsilon^\gamma \eta}{1+\eta}\frac{R^\gamma}{N\left(x,\frac{1}{1+R}\right)}$$

Setting $R = \frac{1}{\lambda} - 1$ we obtain that, if

$$\lambda \le \lambda_0 := \frac{\delta-\varepsilon}{1-\varepsilon},$$

then

$$
\begin{aligned}
\int_{(0,\frac{1}{\lambda}-1)} r^{\gamma} dV(r) &\geq \frac{\varepsilon^{\gamma}\eta}{1+\eta}\frac{(1-\lambda)^{\gamma}\lambda^{-\gamma}}{N(x,\lambda)} \\
&\geq \frac{\left(\varepsilon\frac{1-\delta}{1-\varepsilon}\right)^{\gamma}\eta}{1+\eta}\frac{\lambda^{-\gamma}}{N(x,\lambda)} \\
&\geq c\frac{\lambda^{-\gamma}}{N(x,\lambda)},
\end{aligned}
$$

where $c = \frac{\varepsilon\frac{1-\delta}{1-\varepsilon}\eta}{1+\eta}$ (we use that $\gamma < 1$). Substituting into (6.6) and assuming that $t \geq \frac{1}{\lambda_0}$ we obtain

$$
\begin{aligned}
M_{\gamma}(x,t) &\geq ct\int_0^{\lambda_0}\lambda^{-\gamma}(1-\lambda)^{t-1}d\lambda \\
&\geq ct\int_0^{1/t}\lambda^{-\gamma}(1-\lambda)^{t-1}d\lambda \\
&\geq ct\left(1-\frac{1}{t}\right)^{t-1}\int_0^{1/t}\lambda^{-\gamma}d\lambda \\
&\geq \frac{c}{e}t\frac{t^{\gamma-1}}{1-\gamma} \\
&= \frac{c}{e}\frac{t^{\gamma}}{1-\gamma}.
\end{aligned}
$$

If $t < \frac{1}{\lambda_0}$ then similarly

$$
\begin{aligned}
M_{\gamma}(x,t) &\geq ct\int_0^{\lambda_0}\lambda^{-\gamma}(1-\lambda)^{t-1}d\lambda \\
&\geq ct\int_0^{\lambda_0}\lambda^{-\gamma}(1-\lambda_0)^{\frac{1}{\lambda_0}-1}d\lambda \\
&\geq \frac{c}{e}t\frac{\lambda_0^{1-\gamma}}{1-\gamma} \\
&\geq \frac{c}{e}\frac{\lambda_0}{1-\gamma}t.
\end{aligned}
$$

■

Let us introduce the following two conditions (cf. Conditions 3.12 and 3.13).

Condition 6.4. *There exists a constant* $0 < \kappa_- < 1$ *such that for all* $k \geq 0$,

$$\sigma_{k+1} \geq \kappa_- \sigma_k.$$

Condition 6.5. *For a fixed* $x \in X$, *there exists a constant* $\nu_- > 1$ *such that for all* $k \geq 0$,

$$m(B_{k+1}(x)) \geq \nu_- m(B_k(x)).$$

Proposition 6.6. *If Conditions 6.4 and 6.5 are satisfied then the function* $N(x,\lambda)$ *satisfies the reverse doubling condition* (6.7). *Consequently, the conclusion of Theorem 6.3 is satisfied.*

Proof. If $\sigma_k \leq \lambda < \sigma_{k-1}$ then by Condition 6.4

$$\kappa_- \lambda < \kappa_- \sigma_{k-1} \leq \sigma_k.$$

Therefore, by (3.7) and Condition 6.5,

$$N(x,\lambda)=\frac{1}{m(B_k(x))}\geq\frac{\nu_-}{m(B_{k+1}(x))}\geq\nu_-N(x,\kappa_-\lambda),$$

so that (6.7) is satisfied with $\delta=\kappa_-$ and $1+\eta=\nu_-$. ■

The case $\gamma\geq 1$ is treated in the following theorem.

Theorem 6.7. *If $\gamma\geq 1$ and Condition 6.5 holds, then $M_\gamma(x,t)=\infty$ for all $x\in X$ and $t>0$.*

Proof. Since

$$\int_X p(t,x,y)\,dm(y)=P^t1(x)=1,$$

the divergence of the integral (6.2) can occur only for large $d_\sigma(x,y)$. Therefore, the divergence of (6.2) for $\gamma=1$ implies that for $\gamma>1$. Hence, we can assume in the sequel that $\gamma=1$. We use notation

$$p(t,x,A):=\int_A p(t,x,y)\,dm(y)$$

where A is a subset of A. Set

$$r_k=\frac{1}{\sigma_k}-1,\quad k\geq -1, \tag{6.8}$$

so that $d_\sigma(x,y)=r_k$ if $d(x,y)=k+1$. It follows from (6.2) that

$$\begin{aligned}M_1(x,t)&=\sum_{k\geq 0}\int_{B_{k+1}(x)\setminus B_k(x)}d_\sigma(x,y)p(t,x,y)dm(y)\\&=\sum_{k\geq 0}r_kp(t,x,B_{k+1}(x)\setminus B_k(x))\\&=\sum_{k\geq 0}r_k(p(t,x,B_k(x)^c)-p(t,x,B_{k+1}(x)^c)).\end{aligned}\tag{6.9}$$

Next, by (3.5) and (3.1) we have

$$p(t,x,B_k(x)^c)=\sum_{l\geq 0}\left(s_l^t-s_{l-1}^t\right)\mathcal{P}_l1_{B_k(x)^c}(x)=\sum_{l>k}\left(s_l^t-s_{l-1}^t\right)\left(1-\frac{m(B_k(x))}{m(B_l(x))}\right).\tag{6.10}$$

It follows from (6.10) and Condition 6.5 that

$$\begin{aligned}p(t,x,B_k(x)^c)-p(t,x,B_{k+1}(x)^c)&=\left(s_{k+1}^t-s_k^t\right)\left(1-\frac{m(B_k(x))}{m(B_{k+1}(x))}\right)\\&\geq c\left(s_{k+1}^t-s_k^t\right),\end{aligned}\tag{6.11}$$

where $c=1-\frac{1}{\nu_-}$. Therefore, we obtain by (6.9)

$$M_1(x,t)\geq c\sum_{k\geq 0}r_k(s_{k+1}^t-s_k^t).$$

Next, we use the inequality

$$a^t-b^t\geq t(a-b)\min\left(a^{t-1},b^{t-1}\right)$$

that is true for all $0<b<a$ and $t>0$, which follows from the identity of the mean-value theorem

$$a^t-b^t=t\xi^{t-1}(a-b),$$

where $\xi \in [a,b]$. It follows that

$$M_1(x,t) \geq ct \sum_{k\geq 0} r_k (s_{k+1}-s_k) \min\left(s_k^{t-1}, s_{k+1}^{t-1}\right).$$

Since $s_k^{t-1} \to 1$ as $k \to \infty$, it suffices to prove that

$$\sum_{k\geq N} r_k (s_{k+1}-s_k) = \infty$$

for some N. Choose N so large that $r_N \geq 1$. By (6.8) we have

$$s_k = 1 - \sigma_k = \frac{r_k}{r_k+1},$$

whence it follows that

$$\begin{aligned}\sum_{k\geq N} r_k (s_{k+1}-s_k) &= \sum_{k\geq N} r_k \frac{r_{k+1}-r_k}{(r_{k+1}+1)(r_k+1)} \\ &\geq \frac{1}{4}\sum_{k\geq N} r_k \frac{r_{k+1}-r_k}{r_{k+1}r_k} \\ &= \frac{1}{4}\sum_{k\geq N}\left(1-\frac{r_k}{r_{k+1}}\right).\end{aligned}$$

Set $\alpha_k = 1 - \frac{r_k}{r_{k+1}}$. Since

$$\prod_{k\geq N}(1-\alpha_k) = \prod_{k\geq N}\frac{r_k}{r_{k+1}} = 0,$$

it follows that $\sum_{k\geq N}\alpha_k = \infty$, whence $M_1(x,t) = \infty$. ∎

Example 6.8. As in example 5.1, assume that, for some $x \in X$ and for all $k = 0, 1, ...$

$$m(B_k(x)) \simeq p^k.$$

for some $p > 1$. Clearly, Condition 6.5 is satisfied. Therefore, by Theorems 6.2 and 6.7, $M_\gamma(x,t)$ is finite if and only if $\gamma < 1$; for such γ and for large enough t, we have

$$M_\gamma(x,t) \leq \frac{t^\gamma}{1-\gamma}.$$

Furthermore, if Condition 6.4 is also satisfied then we have by Theorem 6.3 a matching lower bound, so that

$$M_\gamma(t,x) \simeq \frac{t^\gamma}{1-\gamma}. \tag{6.12}$$

Let us show that without Condition 6.4 the lower bound of Theorem 6.3 fails. Assume that

$$\sigma_k \simeq \exp\left(-a^k\right),$$

where $a > 1$, and show that in this case M_γ admits the following upper bound

$$M_\gamma(t,x) \leq Ct^\gamma + o(t^\gamma) \quad \text{as } t \to \infty, \tag{6.13}$$

where $\alpha = \frac{\log p}{\log a}$ and the constant C does not depend on γ (but may depend on α). The difference between the estimates (6.12) and (6.13) is that in the latter the leading coefficient in front of t^γ remains uniformly bounded for all $\gamma \in (0,1)$.

As was shown in Example 5.1, we have

$$N(x,\lambda) \simeq \left(\log\left(1+\frac{1}{\lambda}\right)\right)^{-\alpha} \tag{6.14}$$

for all $\lambda \in (0,1)$, where $\alpha = \frac{\log p}{\log a}$. By Lemma 6.1, we have

$$V(r) := m(B_r^\sigma(x)) = \frac{1}{N\left(x, \frac{1}{1+r}\right)} \simeq \log^\alpha(2+r)$$

for all $r > 0$. By Corollary 3.17, we have

$$p(t,x,y) \simeq \frac{t}{t + d_\sigma(x,y)} N\left(x, \frac{1}{t + d_\sigma(x,y)}\right)$$

for all $y \in X$ and $t \geq 1$. Therefore, we obtain by (6.2), for $t \geq 1$,

$$M_\gamma(x,t) \simeq \int_0^\infty r^\gamma \frac{t}{t+r} N\left(x, \frac{1}{t+r}\right) dV(r).$$

Making change $u = r/t$ and using the monotonicity of $N(x,\cdot)$, we obtain

$$\begin{aligned} M_\gamma(x,t) &\simeq \int_0^\infty (ut)^\gamma \frac{1}{1+u} N\left(x, \frac{1}{t+ut}\right) dV(ut) \\ &\leq t^\gamma \int_0^\infty \frac{u^\gamma}{1+u} N\left(x, \frac{1}{t}\right) dV(ut). \end{aligned}$$

Let us define the function

$$F(u) = \frac{u^\gamma}{1+u}.$$

Using the estimate (6.14), we can write, for large t,

$$\begin{aligned} M_\gamma(x,t) &\leq C \frac{t^\gamma}{\log^\alpha t} \int_0^\infty F(u)\, dV(ut) \\ &= C \frac{t^\gamma}{\log^\alpha t} \int_0^\infty V(ut) \left(-F'(u)\right) du, \end{aligned} \tag{6.15}$$

where we have integrated by parts and used the fact that the function

$$V(ut) F(u) \simeq \log^\alpha(2+ut) \frac{u^\gamma}{1+u}$$

vanishes at $u = 0$ and $u \to \infty$. Noticing that

$$-F'(u) = \frac{u^{\gamma-1}}{(1+u)^2} \left((1-\gamma) u - \gamma\right)$$

and that $-F'(u) \geq 0$ for $u \geq u_0 := \frac{\gamma}{1-\gamma}$ we obtain

$$\begin{aligned} \int_0^\infty V(ut) \left(-F'(u)\right) du &\leq \int_{u_0}^\infty V(ut) \left(-F'(u)\right) du \\ &\leq C \int_{u_0}^\infty \left(\log^\alpha t + \log^\alpha(2+u)\right) \left(-dF(u)\right) \end{aligned}$$

Clearly, we have

$$\int_{u_0}^\infty \log^\alpha t \left(-dF(u)\right) = \left(\log^\alpha t\right) F(u_0) = \left(\log^\alpha t\right) \gamma^\gamma (1-\gamma)^{1-\gamma} \leq \log^\alpha t.$$

Using the estimate

$$-F'(u) \leq \frac{u^{\gamma-1} u}{(1+u)^2} \leq u^{\gamma-2},$$

we obtain

$$\begin{aligned}\int_{u_0}^{\infty} \log^{\alpha}(2+u)(-dF(u)) &\leq \int_{u_0}^{\infty} \log^{\alpha}(2+u)\, u^{\gamma-2} du \\ &\leq C_\gamma \int_{u_0}^{\infty} u^{\frac{1-\gamma}{2}+(\gamma-2)} du \\ &\leq C_\gamma u_0^{-\frac{1-\gamma}{2}}.\end{aligned}$$

Since u_0 depends only on γ, we can write

$$\int_0^{\infty} V(ut)\left(-F'(u)\right) du \leq C \log^{\alpha} t + C_\gamma$$

where the constant C does not depend on γ (but can depend on α). Substituting this estimate into (6.15), we obtain

$$M_\gamma(x,t) \leq C \frac{t^\gamma}{\log^\alpha t}\left(C \log^\alpha t + C_\gamma\right) = Ct^\gamma + C_\gamma \frac{t^\gamma}{\log^\alpha t},$$

whence (6.13) follows.

7. The Laplace operator

It is known that any strongly continuous contraction semigroup $\{P_t\}_{t\geq 0}$ in a Hilbert space H has the generator

$$\mathcal{L} = s\text{-}\lim_{t\to 0} \frac{\operatorname{id} - P_t}{t},$$

that is a densely defined operator in H. Moreover, if the operators P_t are symmetric then $\mathcal{L}$ is a self-adjoint operator and $P_t = \exp(-t\mathcal{L})$. The purpose of the next theorem is to evaluate the generator $\mathcal{L}$ of the semigroup $\left\{P^t\right\}_{t\geq 0}$ in $L^2(X,m)$ defined by (3.5). We will refer to $\mathcal{L}$ as the *Laplace operator* of $\left\{P^t\right\}_{t\geq 0}$.

Theorem 7.1. *Let $\mathcal{L}$ be the generator of $\left\{P^t\right\}_{t\geq 0}$. Then the following identity holds*

$$\mathcal{L} = \sum_{k=0}^{\infty} \left(\log \frac{1}{s_k}\right)\left(\mathcal{P}_k - \mathcal{P}_{k+1}\right), \tag{7.1}$$

where the series converges in the strong operator topology of $L^2(X,m)$. Consequently, $\mathcal{L}$ is a bounded, non-negative definite, self-adjoint operator in $L^2(X,m)$, and the spectrum of $\mathcal{L}$ is given by

$$\operatorname{spec}_{L^2} \mathcal{L} = \left\{\log \frac{1}{s_k}\right\}_{k=0}^{\infty} \bigcup \{0\}. \tag{7.2}$$

Each value $\log \frac{1}{s_k}$ is an eigenvalue of $\mathcal{L}$ with infinite multiplicity.

Proof. The defining identity $P^t = e^{-t\mathcal{L}}$ is equivalent to $P = e^{-\mathcal{L}}$ whence $\mathcal{L} = \log \frac{1}{P}$. Using the spectral resolution (3.4) of P, we obtain (7.1). It follows from (7.1) that $\mathcal{L}$ has the spectrum (7.2), where 0 is added as the only accumulation point of the sequence $\left\{\log \frac{1}{s_k}\right\}$. Since this sequence is non-negative and bounded, the operator $\mathcal{L}$ is non-negative definite and bounded.

It follows from (7.1) that $\log \frac{1}{s_k}$ is an eigenvalue of $\mathcal{L}$ with the eigenspace $\left(\mathcal{P}_k - \mathcal{P}_{k+1}\right) L^2$. Let us show that this space is infinite dimensional. For any $a \in X$ and $k \in \mathbb{Z}_+$ define a function

$$f_{k,a} = (\mathcal{P}_k - \mathcal{P}_{k+1})\mathbf{1}_{\{a\}}$$

so that $f_{k,a}$ is an eigenfunction of $\mathcal{L}$ with the eigenvalue $\log \frac{1}{s_k}$. Since

$$\mathcal{P}_l \mathbf{1}_{\{a\}}(x) = \frac{1}{m\left(B_l\left(x\right)\right)} \int_{B_l(x)} \mathbf{1}_{\{a\}} dm = \frac{m(\{a\})\mathbf{1}_{B_l(a)}(x)}{m(B_l(x))},$$

we see that $\operatorname{supp} \mathcal{P}_l \mathbf{1}_{\{a\}} \subset B_l\left(a\right)$. It follows that $\operatorname{supp} f_{k,a} \subset B_{k+1}(a)$. In particular, if the balls $B_{k+1}(a)$ and $B_{k+1}(b)$ are disjoint then the functions $f_{k,a}$ and $f_{k,b}$ have disjoint supports and therefore are orthogonal in L^2. Since there are infinitely many disjoint balls of radius $k+1$, we obtain that the eigenspace of $\mathcal{L}$ with the eigenvalue $\log \frac{1}{s_k}$ is infinitely dimensional. ■

Corollary 7.2. *Let $\mathcal{L}$ be the Laplace operator of the semigroup $\left\{P^t\right\}_{t\geq 0}$. Let $\phi : [0,\infty) \to [0,\infty)$ be a continuous, strictly monotone increasing function such that $\phi\left(0\right) = 0$. Then the operator $\phi(\mathcal{L})$ is the Laplace operator of the semigroup $\left\{P_\phi^t\right\}_{t\geq 0}$ that is defined by*

$$P_\phi = \sum_{k=0}^{\infty} c_k^\phi \mathcal{P}_k,$$

where $\left\{c_k^\phi\right\}_{k=0}^{\infty}$ is a stochastic sequence given by

$$c_k^\phi = \exp\left(-\phi\left(\log \frac{1}{s_k}\right)\right) - \exp\left(-\phi\left(\log \frac{1}{s_{k-1}}\right)\right),\ k \geq 1, \tag{7.3}$$

$$c_0^\phi = \exp\left(-\phi\left(\log \frac{1}{s_0}\right)\right). \tag{7.4}$$

Proof. Note that the sequence $\left\{\exp\left(-\phi\left(\log \frac{1}{s_k}\right)\right)\right\}_{k=0}^{\infty}$ is strictly monotone increasing in k, which implies $c_k^\phi > 0$. It follows from (7.3)-(7.4) that

$$s_k^\phi := c_0^\phi + ... + c_k^\phi = \exp\left(-\phi\left(\log \frac{1}{s_k}\right)\right).$$

Hence,

$$\sum_{k=0}^{\infty} c_k^\phi = \lim_{k\to\infty} s_k^\phi = \exp\left(-\phi\left(0\right)\right) = 1$$

so that the sequence $\left\{c_k^\phi\right\}_{k=0}^{\infty}$ is stochastic. By Lemma 3.2, $\left\{P_\phi^t\right\}_{t\geq 0}$ is a strongly continuous Markov semigroup in $L^2\left(X,m\right)$, and by Theorem 7.1, the generator $\mathcal{L}_\phi$ of this semigroup is given by

$$\mathcal{L}_\varphi = \sum_{k=0}^{\infty}\left(\log \frac{1}{s_k^\phi}\right)\left(\mathcal{P}_k - \mathcal{P}_{k+1}\right).$$

On the other hand, we obtain from (7.1) by the functional calculus that

$$\phi\left(\mathcal{L}\right) = \sum_{k=0}^{\infty} \phi\left(\log \frac{1}{s_k}\right)\left(\mathcal{P}_k - \mathcal{P}_{k+1}\right).$$

Then the identity $\mathcal{L}_\phi = \phi\left(\mathcal{L}\right)$ follows from

$$\log \frac{1}{s_k^\phi} = \phi\left(\log \frac{1}{s_k}\right).$$

■

If A is a (non-negative definite) symmetric operator such that the semigroup $\left\{e^{-tA}\right\}_{t\geq 0}$ is Markov then we refer to A as a Laplace operator. By a theorem of Bochner, for any Laplace operator A, the operator $\phi(A)$ is again a Laplace operator, provided ϕ is a Bernstein function. For example, $\phi(\lambda) = \lambda^\alpha$ is a Bernstein function if $0 < \alpha \le 1$ and is not if $\alpha > 1$. In general, it is not true that A^α with $\alpha > 1$ is a Laplace operator. However, in the specific case of the semigroup $\{P^t\}_{t\ge 0}$ given by (3.5), the operator $\mathcal{L}^\alpha$ is by Corollary 7.2 a Laplace operator for any $\alpha > 0$.

Theorem 7.3. *For any $p \in [1, +\infty]$, the operator $\mathcal{L}$ can be extended as a bounded operator acting on $L^p = L^p(X, \mathcal{F}, m)$. Moreover, we have*

$$\operatorname{spec}_{L^p} \mathcal{L} = \operatorname{spec}_{L^2} \mathcal{L}, \quad \textit{for any } p \in [1, +\infty]. \tag{7.5}$$

Proof. Denoting $l_k = \log \frac{1}{s_k}$ we have by (7.1)

$$\mathcal{L}f = \sum_{k=0}^{\infty} l_k (\mathcal{P}_k f - \mathcal{P}_{k+1} f), \tag{7.6}$$

for any $f \in L^2$. Applying the Abel transformation to (7.6), we obtain

$$\mathcal{L}f = l_0 f - \sum_{k=1}^{\infty} (l_{k-1} - l_k)\, \mathcal{P}_k f. \tag{7.7}$$

We use (7.7) to define $\mathcal{L}f$ for $f \in L^p$ for any $p \in [1, +\infty]$. Indeed, since $\|\mathcal{P}_k\|_{L^p \to L^p} \le 1$ and $l_k \downarrow 0$ as $k \to \infty$, we have

$$\sum_{k=1}^{\infty} \|(l_{k-1} - l_k)\, \mathcal{P}_k\|_{L^p \to L^p} \le \sum_{k=1}^{\infty} (l_{k-1} - l_k) = l_0 < \infty$$

so that the series in (7.7) converges in L^p. Hence, (7.7) defines $\mathcal{L}$ as a bounded operator in L^p, which proves the first claim.

For any $f \in L^2$ and for any $\lambda \notin \operatorname{spec}_{L^2} \mathcal{L}$, we have from (7.6)

$$(\mathcal{L} - \lambda\, \mathrm{id})^{-1} f = \sum_{k=0}^{\infty} \frac{1}{l_k - \lambda} (\mathcal{P}_k f - \mathcal{P}_{k+1} f).$$

Applying again the Abel transformation with $l_{-1} := 0$ we obtain

$$\begin{aligned}(\mathcal{L} - \lambda\, \mathrm{id})^{-1} f &= -\frac{1}{\lambda} f + \sum_{k=0}^{\infty} \left(\frac{1}{l_k - \lambda} - \frac{1}{l_{k-1} - \lambda} \right) \mathcal{P}_k f \\ &= -\frac{1}{\lambda} f + \sum_{k=0}^{\infty} \left(\frac{l_{k-1} - l_k}{(l_k - \lambda)\,(l_{k-1} - \lambda)} \right) \mathcal{P}_k f. \end{aligned} \tag{7.8}$$

Since by (7.2)

$$\inf_{k \ge 0} |(l_k - \lambda)\,(l_{k-1} - \lambda)| > 0,$$

we have

$$\sum_{k=0}^{\infty} \left| \frac{l_{k-1} - l_k}{(l_k - \lambda)\,(l_{k-1} - \lambda)} \right| < \infty,$$

whence it follows that the series (7.8) converges in all L^p. Clearly, (7.8) defines $(\mathcal{L} - \lambda\, \mathrm{id})^{-1}$ as a bounded operator in L^p, which proves that $\lambda \notin \operatorname{spec}_{L^p} \mathcal{L}$, that is,

$$\operatorname{spec}_{L^p} \mathcal{L} \subset \operatorname{spec}_{L^2} \mathcal{L}.$$

To prove the opposite inclusion, observe that, for any $k \in \mathbb{Z}_+$ and $a \in X$, the function $f_{k,a} = (\mathcal{P}_k - \mathcal{P}_{k+1})\delta_a$ belongs to all spaces L^p and, by (7.1), is an eigenfunction of the operator $\mathcal{L}$ with the eigenvalue l_k. Therefore, $\operatorname{spec}_{L^p} \mathcal{L}$ contains all l_k and, hence, also 0 as an accumulation point of $\{l_k\}$, which together with (7.2) finishes the proof of (7.5). ∎

To state the next theorem observe that, for any non-negative function f on X, $\mathcal{P}_k f$ is also a non-negative function on X (cf. (3.1)). Hence, the identity (7.7) defines $\mathcal{L}f$ as a function on X with values in $[-\infty, +\infty)$. For example, $\mathcal{L}\,\mathrm{const} \equiv 0$.

Theorem 7.4. (A strong Liouville property) *If f is a non-negative function on X such that $\mathcal{L}f \equiv 0$ then $f = \mathrm{const}$. In particular, the point 0 of $\mathrm{spec}_{L^\infty} \mathcal{L}$ is an eigenvalue of multiplicity 1.*

Proof. Since $\mathcal{P}_k\mathcal{P}_m = \mathcal{P}_m\mathcal{P}_k = \mathcal{P}_{\max(k,m)}$, we obtain from (7.7) that, for any $m \in \mathbb{Z}_+$,

$$\mathcal{L}(\mathcal{P}_m f) = \mathcal{P}_m(\mathcal{L}f) = 0, \tag{7.9}$$

and

$$\begin{aligned}\mathcal{L}(\mathcal{P}_m f) &= l_0\mathcal{P}_m f - \sum_{k\le m}(l_{k-1}-l_k)\,\mathcal{P}_m f - \sum_{k\ge m+1}(l_{k-1}-l_k)\,\mathcal{P}_k f \\ &= l_m\mathcal{P}_m f - \sum_{k=m+1}^{\infty}(l_{k-1}-l_k)\,\mathcal{P}_k f. \end{aligned}\tag{7.10}$$

Applying (7.10) to $m+1$ instead of m and subtracting the result from (7.10) we obtain

$$\mathcal{L}(\mathcal{P}_m f) - \mathcal{L}(\mathcal{P}_{m+1} f) = l_m(\mathcal{P}_m f - \mathcal{P}_{m+1} f).$$

Since $l_m \neq 0$, it follows from (7.9) that

$$\mathcal{P}_m f - \mathcal{P}_{m+1} f = 0.$$

It follows by induction that $f = \mathcal{P}_m f$ for all $m \in \mathbb{Z}_+$, which implies that f must be constant on all balls and, hence, $f = \mathrm{const}$. ■

Corollary 7.5. *Let ϕ be a function as in* Corollary 7.2. *If f is a non-negative function on X such that $\phi(\mathcal{L}) f \equiv 0$ then $f = \mathrm{const}$.*

Proof. Indeed, by Corollary 7.2 the operator $\phi(\mathcal{L})$ is a Laplace operator. Applying to this operator Theorem 7.4, we finish the proof. ■

Example 7.6. Taking $\phi(t) = t^\alpha$ where $\alpha > 0$, we obtain that $\mathcal{L}^\alpha f = 0$ implies $f = \mathrm{const}$ (assuming a priori that $f \ge 0$). Take now $\phi(t) = 1 - e^{-\alpha t}$ and observe that by (7.1) and (3.5)

$$\phi(\mathcal{L}) = \sum_{k=0}^{\infty}(1 - s_k^\alpha)(\mathcal{P}_k - \mathcal{P}_{k+1}) = \mathrm{id} - P^\alpha.$$

Hence, $P^\alpha f = f$ for some $\alpha > 0$ implies $f = \mathrm{const}$.

8. Appendix: some properties of the gamma function

Lemma 8.1. *Fix $\gamma \in (0,1)$.*

(a) *The function $t \mapsto \frac{\Gamma(t+\gamma)}{\Gamma(t+1)}$ is monotone decreasing in $t \ge 0$.*
(b) *For all $t > 0$, we have*

$$\Gamma(t+\gamma) \le t^\gamma \Gamma(t). \tag{8.1}$$

Proof. It is known that, for all $t > 0$,

$$\varphi(t) := \frac{d}{dt}\ln\Gamma(t) = -C + \sum_{n=0}^{\infty}\left(\frac{1}{n+1} - \frac{1}{n+t}\right), \tag{8.2}$$

where C is the Euler constant (see [1, eq. 6.3.16]).

(a) It is clear from (8.2) that $\varphi(t)$ is monotone increasing in t. Therefore, we have

$$\frac{d}{dt}\left(\ln\Gamma(t+\gamma)-\ln\Gamma(t+1)\right)=\varphi(t+\gamma)-\varphi(t+1)\leq 0,$$

whence the claim follows.

(b) It follows from (8.2) that

$$\begin{aligned}\varphi(t+\gamma)-\varphi(t) &= \sum_{n=0}^{\infty}\left(\frac{1}{n+t}-\frac{1}{n+t+\gamma}\right)\\ &= \sum_{n=0}^{\infty}\frac{\gamma}{(n+t)(n+t+\gamma)}\\ &\geq \gamma\sum_{n=0}^{\infty}\frac{1}{(n+t)(n+1+t)}\\ &= \frac{\gamma}{t}.\end{aligned}$$

Therefore, setting

$$F(t):=\ln\Gamma(t+\gamma)-\ln\Gamma(t)-\gamma\ln t,$$

we obtain that

$$\frac{d}{dt}F(t)=\varphi(t+\gamma)-\varphi(t)-\frac{\gamma}{t}\geq 0,$$

that is, the function $F(t)$ is monotone increasing. Using Stirling's formula in the form

$$\ln\Gamma(t)=\left(t-\frac{1}{2}\right)\ln t-t+\ln\sqrt{2\pi}+o(1)\quad\text{as } t\to\infty,$$

we obtain

$$\begin{aligned}F(t) &= \left(t+\gamma-\frac{1}{2}\right)\ln(t+\gamma)-(t+\gamma)-\left(t-\frac{1}{2}\right)\ln t+t-\gamma\ln t+o(1)\\ &= \left(t+\gamma-\frac{1}{2}\right)\ln\left(\frac{t+\gamma}{t}\right)-\gamma+o(1)\\ &= t\frac{\gamma}{t}-\gamma+o(1)\\ &= o(1).\end{aligned}$$

Hence, $\lim_{t\to\infty}F(t)=0$. Since the function $F(t)$ is monotone increasing, it follows that $F(t)\leq 0$ for all $t>0$, which is equivalent to (8.1). ■

Acknowledgement. This paper was written during a visit of the first-named author to the University of Bielefeld. He would like to express his gratitude to the colleagues from SFB 701 for the opportunity to work on the project. The authors thank Wolfhard Hansen, Laurent Saloff-Coste and Wolfgang Woess for fruitful discussions and valuable comments.

References

[1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables,*. New York: Dover, 1972.

[2] A. Bendikov. Asymptotic formulas for symmetric stable semigroups. *Exposition. Math.*, **12**, no. 4:381–384, 1994.

[3] A. Bendikov, B. Bobikau, and Ch. Pittet. Spectral properties of a class of random walks on locally finite groups. *Preprint*, , 2010.

[4] A. Bendikov, Ch. Pittet, and R. Sauer. Spectral distribution and l^2-isoperimetric profile of laplace operators on groups. *to appear in Math. Ann.*, , 2011.

[5] R. M. Blumenthal and R. K. Getoor. Some theorems on stable processes. *Trans. Amer. Math. Soc.*, **95**:263–273, 1960.

[6] S. Brofferio and W. Woess. On transience of card shuffling. *Proc. Amer. Math. Soc.*, **129**, no. 5:1513–1519, 2001.

[7] T. Coulhon, A. Grigor'yan, and Ch. Pittet. A geometric approach to on-diagonal heat kernel low bounds on groups. *Ann. Inst. Fourier (Grenoble).*, **51**, no. 6:1763–1827, 2001.

[8] D. A. Darling and P. Erdös. On the recurrence of a certain chain. *Proc. Amer. Math. Soc.*, **13**:447–450, 1962.

[9] N. Fereig and S. A. Molchanov. Random walks on abelian groups with an infinite number of generators. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, **5** :22–29, 1978.

[10] L. Flatto and J. Pitt. Recurrence criteria for random walks on countable abelian groups. *Illinois J. Math.*, **18**:1–19, 1974.

[11] A. Grigor'yan. *Heat Kernel and Analysis on Manifolds*, volume **47**. AMS/IP Studies in Advanced Mathematics, 2009.

[12] M. Gromov. *Asymptotic invariants of infinite groups*, volume **182** of *Geometric group theory, Vol. 2 (Sussex, 1991), London Math. Soc. Lecture Notes Ser.*, pages 1–295. Cambridge Univ. Press, Cambridge, 1993.

[13] M. A. Kasymdzhanova. Recurrence of invariant markov chains on a class of abelian groups. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, **3**:3–7, 1981.

[14] H. Kesten and F. Spitzer. Random walks on countably infinite abelian groups. *Acta. Math.*, **114**:237–265, 1965.

[15] G. F. Lawler. Recurrence and transience for a card shuffling model. *Combin. Probab. Comput.*, **4**, no. 2:133–142, 1995.

[16] M. Del Muto and A. Figa-Talamanca. Diffusion on locally compact ultrametric spaces. *Expo. Math.*, **22**:197–211, 2004.

[17] M. Del Muto and A. Figa-Talamanca. Anisotropic diffusion on totally disconnected abelian groups. *Pacific J. Math.*, **225**:221–229, 2006.

[18] Ch. Pittet and L. Saloff-Coste. Amenable groups, isoperimetric profiles and random walks. *in: Geometric group Theory (Canberra, 1996)*, pages 293–316, 1999.

[19] Ch. Pittet and L. Saloff-Coste. On random walks on wreath products. *Ann. Probab.*, 30(2):948–977, 2002.

[20] L. Saloff-Coste. Opérateurs pseudo-différentiels sur certains groupes totalement discontinus. *Studia Mathematica*, 83:205–228, 1986.

[21] L. Saloff-Coste. Probability on groups: random walks and invariant diffusions. *Notices Amer. Math. Soc.*, 48(9):968–977, 2001.

[22] M. H. Taibleson. *Fourier Analysis on Local Fields*. Princeton Univ. Press, 1975.

[23] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon. *Analysis and geometry on groups*. Cambridge Tracts in Mathematics, vol. 100, Cambridge Univ. Press, Cambridge, 1992.

[24] André Weil. *Basic number theory.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the second (1973) edition.

[25] W. Woess. *Random Walks on Infinite Graphs and Groups*, volume **138**. Cambridge University Press, 2000.

Institute of Mathematics, Wrocław University, 50-384 Wrocław, Pl. Grundwaldzki 2/4, Poland
E-mail address: `bendikov@math.uni.wroc.pl`

Department of Mathematics, University of Bielefeld, 33501 Bielefeld, Germany
E-mail address: `grigor@math.uni-bielefeld.de`

CMI, 39 rue Joliot-Curie, 13453 Marseille Cedex 13, Université d'Aix-Marseille I, France
E-mail address: `pittet@cmi.univ-mrs.fr`