# A note on an alternating upper bound for random walks on semigroups

Fan Chung* and Jacob Hughes
University of California, San Diego

August 9, 2012

**Abstract**

We consider random walks on idempotent semigroups, called Left Regular Bands, satisfying the relation $xyx = xy$ for any two elements $x$ and $y$ of the semigroup. We give an alternating upper bound for the total variation distance of a random walk on a Left Regular Band semigroup, improving the previous bound by Brown and Diaconis.

**Keywords:** Random Walk, Semigroup, Left Regular Band

## 1 Introduction

In this note we consider random walks on semigroups. In particular, we restrict our attention to a class of semigroups known as Left Regular Bands (or LRB's, for short) which are idempotent with the additional relation $xyx = xy$ (see [7, 8, 9]). Many problems can be interpreted as random walks on LRB's, such as the move-to-front self-organizing schemes [2, 6], hyperplane arrangements [1] and graph coloring games [5]. It is known that the random walks on Left Regular Bands have many amazing properties, including having real eigenvalues which can be expressed in elegant formula [2, 3]. In addition, Diaconis and Brown [4] gave a variation of the Plancherel formula for bounding the total variation distance $\Delta_{\mathrm{TV}}(t)$ of a LRB random walk after $t$ steps:

$$\Delta_{\mathrm{TV}}(t) \le \sum_{\{l \in L^* | l \text{ is co-maximal}\}} \lambda_l^t \tag{1}$$

where the eigenvalues are indexed by the co-maximal elements in the semilattice $L$ associated with $S$ and the random walk under consideration is on the ideal of chambers in $S$. (Detailed definitions will be given later.) This is in contrast with the Plancherel formula for random walks on groups (or, on vertex transitive graphs), which states that

$$\Delta_{\mathrm{TV}}(t) \le \frac{1}{2}\Big(\sum_{\lambda_i < 1} \lambda_i^{2s}\Big)^{1/2}$$

where $\lambda_i$'s are eigenvalues of the transition probability matrix.

In this paper we will give a slightly improved formula of (1). The proof combines the techniques of Diaconis-Brown in [4] and the methods of Bidigare, Hanlon, and Rockmore [1] for random walks on chambers of hyperplane arrangements. A similar result was independently obtained by Benjamin Steinberg [10].

*Research supported in part by ONR MURI N000140810747, and FA9550-09-1-0900.

**Theorem 1.** *For a random walk on chambers of an LRB semigroup, the total deviation distance after t steps is bounded by:*

$$\Delta_{TV}(t) \leq \sum_{l \in L^*} -\mu(l, \hat{1})\lambda_l^t,$$

*where $\mu(l, \hat{1})$ is the Möbius function on the support lattice $L$, and $L^*$ denotes the lattice $L$ with its maximal element, $\hat{1}$, removed.*

To show that Theorem 1 is an improvement of (1), we will show the following.

**Corollary 1.**

$$\sum_{l \in L^*} -\mu(l, \hat{1})\lambda_l^t \leq \sum_{\{l \in L^* | l \text{ is co-maximal}\}} \lambda_l^t.$$

The proofs of Theorem 1 and Corollary 1 are included in Section 3.

When $L$ is a Boolean lattice of a set V, then we can index elements of the lattice by subsets of $V$. For a subset $Y \subset V$, the Möbius function $\mu(Y, \hat{1}) = (-1)^{|V \setminus Y|}$.

**Corollary 2.** *If $L$ is a Boolean lattice of a set $V$, then for a subset $Y \subset V$ $\mu(Y, \hat{1}) = (-1)^{|V \setminus Y|}$, and therefore*

$$\Delta_{TV}(t) \leq \sum_{Y \subset V} (-1)^{|V \setminus Y|+1} \lambda_Y^t$$

## 2 The eigenvalues of a LRB walk

Before we give the proof of Theorem 1, we state some definitions concerning an LRB semigroup $S$ as in [2].

**(a)** In $S$, there is a natural partial order $<$, defined by:

$$x \leq y \Leftrightarrow xy = y.$$

For $x$ in $S$, we define

$$S_{\geq x} = \{y \in S : y \geq x\}.$$

**(b)** A semilattice $L$ can be derived from $S$ as follows: First we define a relation $\preceq$ on $S$ :

$$y \preceq x \Leftrightarrow xy = x.$$

The equivalence class under $\preceq$ containing $x$ is said to be the support of $x$, denoted by supp $x$. For $x, y$ in $S$, we have

$$\text{supp } xy = \text{supp } x \vee \text{supp } y$$

and

$$x \leq y \Rightarrow \text{supp } x \preceq \text{supp } y.$$

Elements in $L$ are called *flats* (following the terminology for semigroups associated with matroids [2]). A flat $l$ is *co-maximal* if $x \succ l \Rightarrow x = \hat{1}$, where $\hat{1}$ denotes the maximal element of $L$

**(c)** An element $c \in S$ is said to be a *chamber* if $cx = c$ for all $x \in S$. Therefore supp $c$ is maximal in the semilattice $L$. The set of all chambers forms an ideal of $S$.

The eigenvalues of a random walk on chambers of semigroups have an elegant form (see [2]). For each $X \in L$, there is an eigenvalue

$$\lambda_X = \sum_{\text{supp } x \preceq X} w_x$$

with multiplicity $m_X$ satisfying

$$\sum_{Y \succeq X} m_Y = c_X$$

where $c_Y$ is the cardinality of $S_{\succeq Y} = S_{\geq y} = \{z \in S : z \geq y\}$ where $y$ is any element with support $Y$, (this is independent of the choice of $y$). Alternatively,

$$m_X = \sum_{Y \succeq X} \mu(X,Y) c_Y$$

where $\mu$ is the Möbius function of the lattice $L$.

Examples of random walks on chambers of various LRB's will be given after the proof of the Theorem.

# 3 The proof of the main theorem

*Proof of Theorem 1*:
Let $\{p_s\}$ be a probability distribution on $S$, so that $p_s \geq 0$ and $\sum_s p_s = 1$. The transition probability matrix of the associated random walk is denoted by

$$P(u,v) = \sum_{\substack{s \\ su=v}} p_s.$$

Let $x_1, x_2, \cdots$ be an i.i.d. sequence of random elements of $S$. We consider $x^{(t)} = x_1 x_2 \cdots x_t$ and $x_{t,s} = x^{(t)} s$, which is the location of a random walk after $t$ steps starting at $s$.

Note that if $x^{(t)}$ is a chamber then $x^{(t)} s_1 = x^{(t)} s_2$ for any $s_1, s_2 \in S$. We define $\pi_{t,s}$ to be the distribution of $x_{t,s}$, that is $\pi_{t,s}(u) = P(x_{t,s} = u)$, for any $u \in S$. Let $\pi$ denote the stationary distribution of the random walk. Note that if $x^{(t_0)}$ is a chamber for a fixed time $t_0$, then $x^{(N)} = x^{(t)}$ for all $N \geq t_0$. Thus $\pi(u) = P(Cs = u)$ where $C$ is a the random chamber first reached by $x^{(t)}$ as $t$ increases.

We consider the total variation distance

$$\Delta_{TV}(t) := \max_{s \in S} \max_{A \subset S} |\pi_{t,s}(A) - \pi(A)|. \tag{2}$$

where $\pi_{t,s}(A) = P(x_{t,s} \in A) = \sum_{u \in A} \pi_{t,s}(u)$ and $\pi(A) = P(Cs \in A) = \sum_{u \in A} \pi(u)$. We split up both events according to whether or not $x^{(t)}$ is a chamber. Let $B_t$ be the event that $x^{(t)}$ is a chamber. Then we have

$$\pi_{t,s}(A) = P(B_t \text{ and } x^{(t)} s \in A) + P(\neg B_t \text{ and } x^{(t)} s \in A)$$
$$\pi(A) = P(B_t \text{ and } Cs \in A) + P(\neg B_t \text{ and } Cs \in A)$$

If $B_t$ occurs then $x^{(t)}$ is a chamber, and thus the first term of each expression is the same. This follows from the known fact that the stationary distribution in the original (unbounded) process (with replacement) is the same as the stationary distribution for the process without replacement. For detailed discussions, the reader is referred to Section 4 in [5].

Therefore we have

$$|\pi_{t,s}(A) - \pi(A)| = |P(\neg B_t \text{ and } x^{(t)}s \in A) - P(\neg B_t \text{ and } Cs \in A)|$$

which is at most $P(\neg B_t)$, as both terms in the difference are between 0 and $P(\neg B_t)$. Thus we have

$$\Delta_{TV}(t) \leq P(\neg B_t) = P(x^{(t)} \text{ is not chamber}).$$

By definition, the only way for $x^{(t)}$ not be a chamber is for some $m \in L, m \neq \hat{1}$, supp $x^{(t)} = m$. Therefore,

$$\Delta_{TV}(t) \leq \sum_{m \in L^*} P(\text{supp } x^{(t)} = m). \tag{3}$$

Let us denote $P(\text{supp } x^{(t)} = m)$ by $\beta_{t,m}$, then equation (3) becomes

$$\Delta_{TV}(t) \leq \sum_{m \in L^*} \beta_{t,m}. \tag{4}$$

We will evaluate $\beta_{t,m}$ using a Möbius inversion on the lattice L. From [2], we have

$$\lambda_m = \sum_{\text{supp } c \subseteq m} p_c = P(\text{supp } x \preceq m)$$

where $P(x = c) = p_c$. Since $x_1, \cdots, x_t$ are chosen independently, we have that

$$\lambda_m^t = P(\text{supp } x_1 \preceq m) P(\text{supp } x_2 \preceq m) \cdots P(\text{supp } x_t \preceq m) = P(\text{supp } x^{(t)} \preceq m). \tag{5}$$

Note that supp $x^{(t)} \preceq m$ if and only if supp $x^{(t)} = l$ for some $l \preceq m$. Thus

$$P(\text{supp } x^{(t)} \preceq m) = \sum_{l \preceq m} \beta_{t,l}.$$

By (5) we have

$$\lambda_m^t = \sum_{l \preceq m} \beta_{t,l}.$$

Therefore we can use a Möbius inversion to derive

$$\beta_{t,m} = \sum_{l \preceq m} \mu(l,m) \lambda_l^t.$$

Plugging this into (4) we obtain

$$\begin{aligned}
\Delta_{TV}(t) &\leq \sum_{m \in L^*} \sum_{\{l | l \preceq m\}} \mu(l,m) \lambda_l^t \\
&= \sum_{l \in L^*} \sum_{\{m | l \preceq m, m \neq \hat{1}\}} \mu(l,m) \lambda_l^t \\
&= \sum_{l \in L^*} \left( \sum_{\{m | l \preceq m,, m \neq \hat{1}\}} \mu(l,m) \right) \lambda_l^t \\
&= \sum_{l \in L^*} \left( -\mu(l, \hat{1}) \right) \lambda_l^t
\end{aligned}$$

as desired. $\square$

*Proof of Corollary 1*:
We note that the bound given in (1) can be compared to the bound of Theorem 1 as follows:

$$\begin{aligned}\sum_{\{l\in L^*|l \text{ is co-maximal}\}} \lambda_l^t &= \sum_{\{l\in L^*|l \text{ is co-maximal}\}} P(\text{supp } x^{(t)} \preceq l) \\ &= \sum_{m\in L^*} |\{l|m \preceq l \text{ and } l \text{ is co-maximal}\}| P(\text{supp } x^{(t)} = m) \\ &\geq \sum_{m\in L^*} P(\text{supp } x^{(t)} = m)\end{aligned}$$

This last term is the right hand side of (3) above, which is equal to the bound given in Theorem 1. Thus Theorem 1 is an improvement over (1). $\square$

# 4 Several examples of LRB's using the improved convergence bounds

**The Tsetlin Library**: Let $T$ be the set of permutations of the elements of $C = \{c_1, \cdots, c_m\}$. We define the action of $C$ on $T$ as follows. For $c_j$ in $C$, and $t = c_{i_1}c_{i_2}\cdots c_{i_m}$, then $c_jt = c_jc_{i_1}c_{i_2}\cdots \widehat{c_j}\cdots c_{i_m}$, where the $\widehat{c_j}$ means to delete $c_j$ where it appears later in the string. It is not hard to check that $T$ is the ideal of chambers of the free semigroup generated by actions of $C$. This is known as the Tsetlin library, a well studied LRB, which has many applications including the "move-to-front" self-organizing search.

Let us consider the random walk on $T$ with respect to the distribution $\{p_c\}$ where $p_c = \frac{1}{m}$ for all $c \in C$. The support lattice $L$ is the boolean lattice $\{0,1\}^C$. The eigenvalues for the random walk on $T$ were first determined by Phatarfod [8] in 1991. For each subset $X \subset C$, there is an eigenvalue $\lambda_X = |X|/m$, with multiplicity equal to the so-called derangement number $d_k$ where $k = m - |X|$. It is known that

$$d_k = k!\sum_{j=0}^{k}\frac{(-1)^j}{j!} = \left\lfloor \frac{k!}{e} + \frac{1}{2}\right\rfloor.$$

Note that the derangement numbers $d_k$ satisfy

$$\sum_{Y\supseteq X} d_{m-|Y|} = c_X = (m-|X|)!.$$

However, one of the advantages of both the bounds in (1) and Theorem 1 is that they do not depend on the multiplicity of each eigenvalue. Equation (1) yields the bound

$$\Delta_{\text{TV}}(t) \leq m\left(1-\frac{1}{m}\right)^t.$$

Theorem 1 improves this to

$$\Delta_{\text{TV}}(t) \leq \sum_{k=1}^{m-1}(-1)^{k+1}\binom{m}{k}\left(\frac{m-k}{m}\right)^t.$$

Since the old bound is simply the first term of this alternating series, it has been improved by $\sum\limits_{k=2}^{m-1}(-1)^k\binom{m}{k}\left(\frac{m-k}{m}\right)^t$.

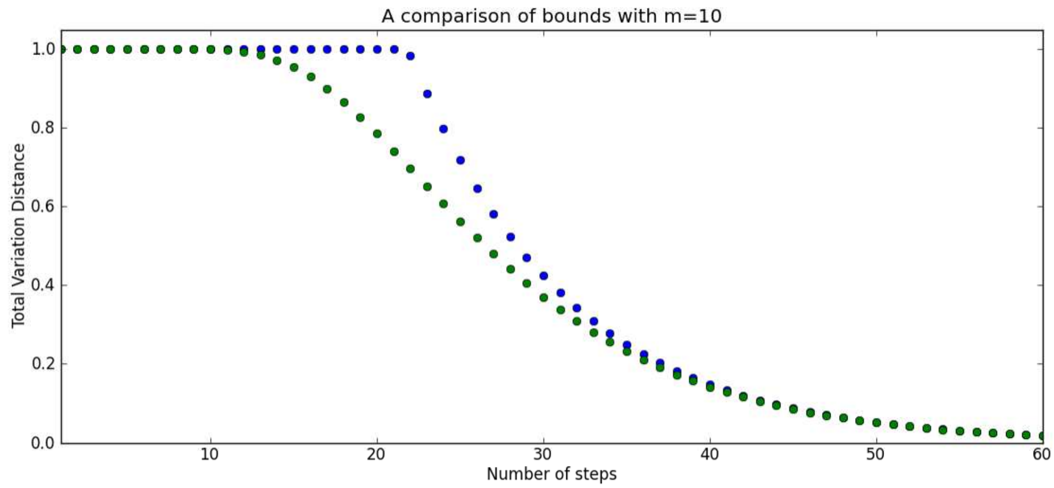


Figure 1: The bounds given by Equation (1) and Theorem 1 for the Tsetlin Library on 10 elements

To illustrate the improvement, Figure 4 is a plot of the two bounds for the case when $m = 10$.

**Strings of Colored Numbers**: For given positive integers $m, k$, let $C = \{(i,j)\}_{i=1,j=1}^{m,k}$. It is convenient to think of $(i,j)$ as being the number $i$ with color $j$. We define the strings of colored numbers, $SCN(m,k)$, to be the LRB consisting of strings of elements of $C$ with respect to the relation $(i,j_0)x(i,j_1) = (i,j_0)x$ for all $x \in SCN(m,k)$. The support lattice $L$ is the boolean lattice of the numbers $1, ..., m$ with supp $(i,j) = i \in L$. Thus the maximal ideal $I$ consists of strings of length $m$, with each number represented exactly once.

Consider a random walk on the LRB $SCN(m,k)$ where each color from $1, ..., k$ is assigned a probability $p_j$ of being chosen, where $\sum_j p_j = 1, p_j > 0$, and each number has equal probability of being chosen. More explicitly, for $(i,j) \in C$, $p_{(i,j)} = \frac{p_j}{m}$. Note that if $k = 1$, then we have the Tsetlin Library considered above. The support lattice $L$ is the boolean lattice $\{0,1\}^m$, and thus elements are indexed by subsets $Y \subset \{1, ..., m\}$. The eigenvalues are thus

$$\lambda_Y = \sum_{\text{supp } c \subset Y} p_c = \sum_{\substack{(i,j) \\ i \in Y}} \frac{p_j}{m} = \sum_{i \in Y} \sum_j \frac{p_j}{m} = \sum_{i \in Y} \frac{1}{m} \sum_j p_j = \frac{|Y|}{m}.$$

Thus the eigenvalues and support lattice are the same as in the Tsetlin library, and so we have the same bound of

$$\Delta_{\text{TV}}(t) \leq \sum_{k=1}^{m-1} (-1)^{k+1} \binom{m}{k} \left(\frac{m-k}{m}\right)^t.$$

It is interesting to note that the bound does not depend at all on the choice of probabilities for a color $p_j$, so long as each number has equal probability of being chosen.

**Edge Flipping in graphs:** In [5], the following random process was studied: Initially each vertex of a graph $G$ is colored red or blue. At each step in the process, we select a random edge of $G$ and (re-)color both its endpoints blue with probability $p$, or red with probability $q = 1 - p$. This process is then repeated some large number of times. The color configuration of $G$ changes at each step. This edge-flipping process corresponds to a random walk on the associated state graph in which each coloring configuration is a node. It can be shown that the actions (each of which is associated with picking an edge $e$ and changing the colors of its endpoints ) form a LRB and the random walk on the state graph is exactly the chamber random walk on this LRB. It was shown in [5] that, for example, for the uniform case of $p = 1/2$, the random walk on the state graph has, for each subset $T$ of the vertex set $V$ of $G$, the eigenvalue $\lambda_T$ (with multiplicity 1) being the ratio of the number of edges in the induced subgraph of $T$ divided by the total number of edges in $G$.

To bound the total variation distance, we can use Theorem 1 to improve previous bounds since the negative terms in the Möbius function lead to cancelations. For the example of the path $P_5$, the bound given in [5] yields $\Delta_{\mathrm{TV}}(t) \leq 2(\frac{3}{4})^t + 6(\frac{1}{2})^t + 10(\frac{1}{4})^t$, whereas Theorem 1 and Corollary 1 give the improved bound $\Delta_{\mathrm{TV}}(t) \leq 2(\frac{3}{4})^t - 2(\frac{1}{4})^t$.

## References


[1] T. P. Bidigare, P. Hanlon, and D. Rockmore, A Combinatorial Description of the Spectrum for the Tsetlin library and its generalization to hyperplane arrangements, *Duke Math. J.* **99** (1), (1999) 135–174.

[2] K. S. Brown, Semigroups, rings, and Markov chains, *Journal of Theoretical Proabability,* **13** (3), (2000), 871–938.

[3] K. S. Brown, Semigroup and ring theoretical methods in probability, *Representations of Finite Dimensional Algebras and Related Topics in Lie Theory and Geometry, Fields Institute Communications* , **40**, (2004), 3–26.

[4] K. S. Brown and P. Diaconis, Random walks and hyperplane arrangements, *Annals of Probability*, **26** (4), (1998), 1813–1854.

[5] F. Chung and R. Graham., Edge flipping in graphs, *Advances in Applied Mathematics*, **48**, (2012), 37–63.

[6] J. A. Fill, An exact formula for the move-to-front rule for self-organizing lists, *J. Theoret. Probab.* **9** (1), (1996), 113–160.

[7] F. Klein-Barmen, Über eine weitere Versallgemeinerung des Verbandsbegriffes, *Math. Z.* **46**, (1940), 472–480.

[8] R. M. Phatarfod, On the matrix occurring in a linear search problem, *J. Appl. Probab.* **28** (2), (1991), 336–346.

[9] M.-P. Schützenberger, Sur certains treillis gauches, *C. R. Acad. Sci. Paris* **224**, (1947), 776–778.

[10] B. Steinberg, An algebraic approach to the Brown-Diaconis upper bound lemma for left regular band walks, preprint.