THE NEAREST NEIGHBOR RANDOM WALK ON SUBSPACES OF A VECTOR SPACE AND TIME TO STATIONARITY

BY

ANTHONY J. D'ARISTOTILE

TECHNICAL REPORT No. 418
FEBRUARY 1993

PREPARED UNDER THE AUSPICES OF
NATIONAL SCIENCE FOUNDATION GRANT DMS92-04864

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4065
THE NEAREST NEIGHBOR RANDOM WALK
ON SUBSPACES OF A VECTOR SPACE
AND TIME TO STATIONARITY

BY
ANTHONY J. D’ARISTOTILE

TECHNICAL REPORT No. 418
FEBRUARY 1993

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT DMS92-04864

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4065
The Nearest Neighbor Random Walk
On Subspaces of a Vector Space
and Time To Stationarity

by

Anthony J. D'Aristotile

Abstract

Let $X$ be the collection of $k$-dimensional subspaces of an $n$-dimensional vector space $V_n$ over $GF(q)$. A metric may be defined on $X$ by letting

$$d(W_k, V_k) = k \cdot \dim(W_k \cap V_k) \quad \text{for} \quad W_k, V_k \in X.$$ 

The nearest neighbor random walk on $X$ starts at an initial fixed point $Z_k$ of $X$ and, from wherever it finds itself, moves with the same probability to any of the points of $X$ at distance one from it. We analyze the rate of convergence of the nearest neighbor random walk on $X$ to its stationary distribution. The argument involves lifting the process on $X$ to a random walk on $GL_n(GF(q))$ and using the Fourier transform and $q$-Hahn polynomials.

Key words: Markov chains, $q$-binomial coefficients, Gelfand pairs, $q$-Hahn polynomials.

AMS(MOS) Subject Classifications: Primary 60B15; secondary 60J20.
The Nearest Neighbor Random Walk  
On Subspaces of a Vector Space  
and Time To Stationarity  

by  
Anthony J. D'Aristotile  

1. Introduction

In recent years there has been great interest in the use of Markov chains and random walk on groups as a part of algorithms in computer science. This paper presents a careful analysis of one such algorithm and compares the best current deterministic algorithm with a suitable random walk.

The problem involves choosing a random subspace over a finite field. Thus, let $V_n$ be an $n$-dimensional vector space over $GF(q)$, and for $0 \leq k \leq \frac{n}{2}$ let $X = X_{n,k,q}$ denote the collection of $k$-dimensional subspaces of $V_n$. Calabi and Wilf [CW] provide an algorithm for selecting a member of $X$ in such a way that all subspaces have equal a priori probabilities of being chosen. They further show that this can be accomplished in time $Cnk$ which is minimal for the problem.

Here, we study the “nearest neighbor random walk” on $X$. Specifically, a metric is defined on $X$ by letting $d(W_k, V_k) = k \cdot \dim(W_k \cap V_k)$ for $W_k, V_k \in X$. The nearest neighbor random walk on $X$ starts at an initial fixed point $Z_k$ of $X$ and at each step moves from an arbitrary point with the same probability to any of the points of $X$ at distance one from it. Using the result of Calabi and Wilf, there is a fairly simple algorithm to describe the process. If $W_k \in X$, choose at random a (k-1)-dimensional subspace $W_{k-1}$ of $W_k$. This is possible by the procedure given in Calabi and Wilf [CW] where in fact linearly independent vectors $w_1, \ldots, w_{k-1}$ are produced which generate $W_{k-1}$. Extend $\{w_1, w_1, \ldots, w_{k-1}\}$ to a basis $\{w_1, \ldots, w_{k-1}, w_k, z_1, z_2, \ldots, z_{n-k}\}$ of $V_n$ where $w_k \in W_k$. Again appealing to [CW], select randomly any 1-dimensional subspace $J_1$ of $Sp\{w_k, z_1, \ldots, z_{n-k}\}$ different from the one generated by $w_k$. From $W_k$ we move to $W_{k-1} \oplus J_1$. This procedure is possible be-
cause there is a 1-1 correspondence between the k-dimensional extensions of $W_{k-1}$ and the 1-dimensional subspaces of $Sp\{w_k, z_1, ..., z_{n-k}\}$ given by

$$V_k \rightarrow V_k \cap Sp\{w_k, z_1, ..., z_{n-k}\}.$$ 

In this article we analyze and bound the rate of convergence of the nearest neighbor random walk on $X$ to its stationary distribution. This problem was posed by Diaconis and Shahshahani [DS]. Work on it was initiated by Greenhalgh [Gr].

In section 5 we study situations where convergence to stationarity is independent of the initial point, and in section 6 we use this result to show that the operating time for our algorithm is $C(nk + k^3)$ (versus $Cnk$ for the Calabi-Wilf algorithm).

The present study is hopefully a contribution to understanding what has come to be called the “cutoff phenomenon”. Intuition says that a Markov chain should approach its stationary distribution monotonically. While this is formally true, detailed study of many chains shows that there is often a sharp cutoff in this approach: the variation distance stays at essentially 1 for a while and then in a relatively short time interval cuts down to values near 0 and thereafter tends to 0 exponentially fast at a rate determined by the second largest eigenvalue. This phenomenon is pervasive and difficult to understand. The problem under study offers domains where cutoffs occur and others where they don’t.

The two contrasting behaviors can be seen in considering $X_{n,k,q}$ for fixed $k$. As $q$ tends to infinity, with $n$ fixed or free, there is a sharp cutoff. As $n$ tends to infinity, with $q$ fixed, there is proveably no cutoff. On the other hand, if $k$ is allowed to vary with $n$, say $k = \frac{n}{2}$, then for $X_{n,\frac{n}{2},q}$, with $q$ fixed and $n$ tending to infinity, there is again a sharp cutoff.

This behavior can be contrasted with previous results. In many cases, the asymptotics do not show a cutoff as the field size increases. A simple example is random walk on $Z(p)$, the integers mod $p$. There, $c(p)p^2$ steps are required to approach stationarity with $c(p)$ tending to infinity as $p$ increases. See [D]. On the other hand, there is a cutoff for simple random walk on the “hypercube” $Z_2^n$ (a step is change of a single coordinate or no change at all). Note that this example may be considered as a much different type of random walk.
on 1-dimensional subspaces of $V_n$ over $GF(2)$. In [D], Diaconis shows that $\frac{1}{4} n \log n + cn$ steps are necessary and sufficient to approach stationarity. However, with respect to the random walk discussed in this paper, randomness would be achieved after a single step.

Stong [Sto] bounded the rate of convergence of Broder’s algorithm for generating a random spanning subtree of a graph (See [Br]). It was one of the first situations where the Upper Bound Lemma of [D] failed in a certain range of parameters. The cases $k$ finite with either $q$ or $n$ tending to infinity provide two additional examples for which the Upper Bound Lemma gives the wrong answer.

2. Preliminaries

With respect to the nearest neighbor random walk on $X = X_{n,k,q}$, let $P_m$ be the law of the process after $m$ steps. Observe that $P_m$ is constant on d-circles centered at $Z_k$, and so $P_m$ may be considered the law after $n$ steps of the induced process on $Y = \{0, 1, 2, ..., j, ..., k\}$. Here, $j$ corresponds to the subset, call it $A_j$, of those numbers of $X$ at distance $j$ from $Z_k$. The stationary distribution $\pi_n$ is given by

$$\pi_n(j) = \frac{|A_j|}{|X|}, \quad 1 \leq j \leq k,$$

where $|\cdot|$ denotes the cardinality of a set. Later we shall derive the formula

$$\pi_n(j) = \frac{|A_j|}{|X|} = \frac{\binom{k}{j} q^{n-j} \binom{n-j}{k-j} q^{j^2}}{\binom{n}{k} q} \quad (1.1)$$

where $\binom{y}{x}_q$ is the so-called q-binomial coefficient [Du] which we define below. Distance to stationarity is measured by variation distance (see [D])

$$||P_m - \pi_n|| = \frac{1}{2} \sum_{j \in Y} |P_m(j) - \pi_n(j)| = \sup_{A \subseteq Y} |P_m(A) - \pi(A)|. \quad (1.2)$$

In many concrete examples of a random walk, the underlying space $Y$ may depend on one or more parameters, and we may wish to study the processes when one of the parameters tends to infinity. (For example, the set $X$ above depends on $n, k,$ and $q$.) If, for instance, $Y = Y_\alpha$ and we write $d_\alpha(m) = ||P_\alpha^m - \pi^\alpha||$, then quite frequently $d_\alpha$ stays
near its maximum value of 1 for a certain number of steps, say $\tau_d(\alpha)$, and then very quickly cuts down to almost zero. This illustrates the so-called "threshold" or "cutoff" phenomenon. More formally [AD 2], a sequence $\tau_d(\alpha)$ is a variation threshold for $Y_\alpha$ if for each $\varepsilon > 0$

$$d_\alpha((1 - \varepsilon)\tau_d(\alpha)) \to 1 \quad \text{as } \alpha \to \infty$$

$$d_\alpha((1 + \varepsilon)\tau_d(\alpha)) \to 0 \quad \text{as } \alpha \to \infty.$$  \hfill (1.3)

The definition of threshold makes explicit the notion of "the number of steps of the random walk required to approach the stationary distribution."

The q-binomial coefficient $[\binom{n}{k}]_q$ is defined for integers $k$, $n$ with $0 \leq k \leq n$ by

$$[\binom{n}{k}]_q = \frac{(q^n - 1)(q^{n-1} - 1)\cdots(q^{n-(k-1)} - 1)}{(q - 1)(q^2 - 1)\cdots(q^k - 1)}.$$  

It is analogous to the binomial coefficient in the sense that it represents the number of $k$-dimensional subspaces of an $n$-dimensional vector space $V_n$ over $GF(q)$. It is also the generating function for restricted partitions of a positive integer $n$, a fact which will be explained and exploited in section 4. See [An], [GR], and [S] for more details.

The following proposition for counting subspaces is due to Dunkel [Du] and shall prove to be extremely useful.

**Proposition 2.1.** Let $\zeta, \eta$ be subspace of $V_n$ with $\dim \zeta = a$, $\dim \eta = b$, $\dim (\zeta \cap \eta) = x$. For integers $c, y$ with $x \leq y \leq a$, $b \leq c \leq n - a + y$, the number of subspaces $\omega \subset V_n$ with $\omega \supset \eta$, $\dim \omega = c$, $\dim (\omega \cap \zeta) = y$ is

$$[\binom{a-x}{y-x}]_q \cdot [\binom{n-b-a+x}{c-b-y+x}]_q q^{(a-y)(c-b-y+x)}.$$  

If $A_j$ as above denotes the collection of those members of $X$ at distance $j$ from $Z_k$, we can use the above proposition to calculate the cardinality of $A_j$:

**Proposition 2.2:** The cardinality of the set $A_j$ is given by

$$|A_j| = [\binom{k}{k-j}]_q \cdot [\binom{n-k}{j}]_q q^{j^2} = [\binom{k}{j}]_q \cdot [\binom{n-k}{j}]_q q^{j^2}.$$
Proof. In the above setting, \( \delta, \eta \) correspond respectively to \( Z_k \) and the zero subspace. Thus \( a = k, \ b = 0, \) and \( x = 0. \) By the definition of \( A_j, \) we take \( y = k - j \) and the result follows.

The formula for the stationary distribution \( \pi_n \) which we stated above is now clear. We now complete the description of the induced random walk.

Proposition 2.3. The transition probabilities on \( Y = \{0, 1, 2, \ldots, k\} \) relative to the nearest neighbor random walk are given by

\[
Q_{i,i-1} = \frac{[i]^2}{\left[ \begin{array}{c} k \\ 1 \end{array} \right]_q \left[ \begin{array}{c} n-k \\ 1 \end{array} \right]_q q} \quad 1 \leq i \leq k,
\]

\[
Q_{i,i} = \frac{[i]_q \left( \left[ \begin{array}{c} n-k-i+1 \\ 1 \end{array} \right]_q q^i - 1 \right) + \left[ \begin{array}{c} k-1 \\ 1 \end{array} \right]_q \left( \left[ \begin{array}{c} i+1 \\ 1 \end{array} \right]_q - 1 \right)}{\left[ \begin{array}{c} k \\ 1 \end{array} \right]_q \left[ \begin{array}{c} n-k \\ 1 \end{array} \right]_q q} \quad 1 \leq i < k,
\]

\[
Q_{k,k} = \frac{\left[ \begin{array}{c} k \\ 1 \end{array} \right]_q \left( \left[ \begin{array}{c} n-2k+1 \\ 1 \end{array} \right]_q q^k - 1 \right)}{\left[ \begin{array}{c} k \\ 1 \end{array} \right]_q \left[ \begin{array}{c} n-k \\ 1 \end{array} \right]_q q},
\]

\[
Q_{i,i+1} = \frac{\left[ \begin{array}{c} k-i \\ 1 \end{array} \right]_q \left[ \begin{array}{c} n-k-i \\ 1 \end{array} \right]_q q^2i+1}{\left[ \begin{array}{c} k \\ 1 \end{array} \right]_q \left[ \begin{array}{c} n-k \\ 1 \end{array} \right]_q q} \quad 0 \leq i \leq k-1.
\]

Proof. We derive the formula for \( Q_{i,i+1}. \) The proofs of the others are established similarly.

In the language of (2.1), we are given \( k \)-dimensional subspaces \( \zeta, \beta \) of \( V_n \) with \( d(\zeta, \beta) = i. \) Since the denominator in the above expression equals the number of elements of \( X \) at distance 1 from \( \beta, \) it suffices to show that the number of subspaces \( \omega \) of \( X \) at distance 1 from \( \beta \) and at distance \( i+1 \) from \( \zeta \) is precisely the numerator. (Here, \( \zeta \) corresponds to the initial point \( Z_k \) of the walk.)

First, fix a \((k-1)\)-dimensional subspace \( n \subseteq \beta, \) and we calculate the number of subspaces \( \omega \) of \( X \) with \( \eta \subseteq \omega \) and satisfying \( d(\beta, \omega) = 1 \) and \( d(\zeta, \omega) = i+1. \) Note
dim(η ∩ ζ) = k - i or k - i - 1. If dim(η ∩ ζ) = k - i, then there is no such ω for, to the contrary, dim(ω ∩ ζ) ≥ k - i which is not compatible with the above conditions. Therefore we assume dim(η ∩ ζ) = k - i - 1. In the context of (2.1), a = k, b = k - 1, c = k, x = k - i - 1, and y = k - i - 1. It is immediate that the number of such ω is

\[
\left[\begin{array}{c} n-k-i \\ 1 \end{array}\right]_q q^{i+1}.
\]

It remains to find the number of (k - 1)-dimensional subspaces γ with γ ⊆ β and dim(γ ∩ ζ) = k - i - 1. Let η be a (k - i - 1)-dimensional subspace of β ∩ ζ. By (2.1) there are \( q^i \) subspaces ω ⊆ β with η ⊆ ω, dim ω = k - 1, and dim(ω ∩ ζ) = k - i - 1. To see this, note that our β ∩ ζ, β corresponds respectively to ζ, \( V_n \) of (2.1) and ω ∩ β ∩ ζ = η. Thus

\[
a = \dim β ∩ ζ = k - i,
\]
\[
b = \dim η = k - i - 1,
\]
\[
c = \dim ω = k - 1,
\]
\[
x = \dim(η ∩ β ∩ ζ) = \dim η = k - i - 1,
\]
\[
y = \dim(ω ∩ β ∩ ζ) = k - i - 1.
\]

Since there are \( \left[\begin{array}{c} k-i \\ 1 \end{array}\right]_q \) (k - i - 1)-dimensional subspaces η of β ∩ ζ, we have \( \left[\begin{array}{c} k-i \\ 1 \end{array}\right]_q q^i \) spaces γ. Combining this with the result of the above paragraph, our assertion follows.

For 0 ≤ x ≤ \( \frac{1}{2} \), the inequality \( e^{2x} < 1 - x \) is well known ([B], [DS]). We establish and employ a tighter lower bound for 1 - x over a more restricted interval. This is provided in the following

**Lemma 2.4.** If

\[
0 < \alpha \leq \sqrt{2} - 1 \quad \text{and} \quad 0 \leq x \leq \alpha
\]

then

\[
e^{-x(1+\alpha)} < 1 - x.
\]
**Proof.** We have for all \( x \) that
\[
e^{-x(1+\alpha)} = 1 - (1 + \alpha)x + \frac{(1 + \alpha)x^2}{2!} - \frac{(1 + \alpha)x^3}{3!} + \frac{(1 + \alpha)x^4}{4!} - \cdots .
\]
Thus it suffices to show that
\[
-\alpha x + \frac{(1 + \alpha)x^2}{2!} \leq 0
\]
and that
\[
- \frac{(1 + \alpha)^{2n+1}x^{2n+1}}{(2n + 1)!} + \frac{(1 + \alpha)^{2n+2}x^{2n+2}}{(2n + 2)!} < 0
\]
for positive integers \( n \) and for \( x \) and \( \alpha \) of the hypothesis. These calculations are straightforward.

The above result yields immediately that
\[
e^{-x-x^2} < 1 - x \quad \text{for} \quad 0 \leq x \leq \sqrt{2} - 1 .\quad (2.5)
\]

The proof of the following elementary lemma is left as an exercise.

**Lemma 2.6.** If \( P^r(i,j) \) is the \( r \)-step transition matrix of a Markov chain with finite state space \( S \) and stationary distribution \( \pi \), then for positive integers \( r \) greater than 1 and all \( i \in S \) we have
\[
\|P^r(i, \cdot) - \pi\| \leq \|P^{r-1}(i, \cdot) - \pi\| .
\]
In treating certain cases of our problem, it will be necessary to utilize a variation of the well known coupling method. The following lemma is similar to a result appearing on page 197 of Bhattacharya and Waymire [BW] but requires a slightly different proof. I thank Jeff Rosenthal for pointing out the potential usefulness of this fact.

**Lemma 2.7.** Let \( P(i,j) \) be the transition matrix of a Markov chain with finite state space \( S \) and stationary distribution \( \pi \). If there exists \( j_0 \in S \), a positive integer \( m_0 \), and a real number \( c > 0 \) such that \( P^{m_0}(i,j_0) \geq c \) for all \( i \in S \), then for each \( i_0 \in S \)
\[
\|P^n(i_0, \cdot) - \pi\| \leq (1 - c)^{[m/m_0]} \quad \text{for} \quad m \geq m_0 .
\]
Proof. Suppose first that $m_0 = 1$ so that $P(i, j_0) \geq c$ for all $i \in S$. Set $X_0 = i_0$ and choose $Y_0 \sim \pi$. For each step, flip a coin with probability of heads $c$. Suppose $X_1, X_2, \ldots, X_t$ and $Y_1, Y_2, \ldots, Y_t$ have already been chosen. If heads results, set $X_{t+1} = Y_{t+1} = j_0$ (i.e., choose $X_{t+1}, Y_{t+1} \sim \delta_{j_0}$). If the coin comes up tails, choose $X_{t+1}, Y_{t+1}$ independently according to

$$X_{t+1} \sim \frac{P(X_t, \cdot) - c\delta_{j_0}}{1 - c}$$

and

$$Y_{t+1} \sim \frac{P(Y_t, \cdot) - c\delta_{j_0}}{1 - c}.$$

Note that $X_t, Y_t$ are independent Markov chains with transition law $P$ and respective initial distributions $\delta_{i_0}$ and $\pi$. (Observe, for instance, that

$$P(X_t, \cdot) = (1 - c)\frac{P(X_t, \cdot) - c\delta_{j_0}}{1 - c} + c\delta_{j_0}$$

and thus $X_{t+1}$ is chosen overall according to $P(X_t, \cdot)$.) From Diaconis [D], or Bhattacharya and Waymire [BW], we have

$$||P_m(i_{0, \cdot}) - \pi|| \leq Q(S \times S \setminus \Delta)$$

where $Q$ is the distribution of $(X_m, Y_m)$ and $\Delta$ is the diagonal of $S \times S$.

Let $T$ be the first time that the coin comes up heads. It is plain that

$$Q(S \times S \setminus \{j_0, j_0\}) \leq P(T > m)$$

and thus

$$||P_m(i_{0, \cdot}) - \pi|| \leq P(T > m).$$

Now

$$P(T > m) = \sum_{s=m+1}^{\infty} P(T = s) = \sum_{s=m+1}^{\infty} (1 - c)^{s-1} c = (1 - c)^m$$

and our claim follows for $m_0 = 1$. 8
For $m_0 > 1$, apply the argument to the transition matrix $P^{m_0}$. In particular, for positive integers $\ell$

$$
||P^{m_0 \ell}(i_0,.) - \pi|| \leq (1 - c)^\ell .
$$

Take $\ell = [m/m_0]$, note that $m_0 \ell \leq m$ and so by Lemma 2.6

$$
||P^m(i_0,.) - \pi|| \leq ||P^{m_0 \ell}(i_0,.) - \pi|| \leq (1 - c)^{[m/m_0]} .
$$


Here, we are concerned with the case of fixed $k$ and either $q$ or $n$ tending to infinity. Expanding $Q_{i,i+1}$ of Proposition 2.3, we have

$$
Q_{i,i+1} = \frac{q^n - q^{k+i} - q^{n-k+i} + q^{2i}}{q^n - q^k - q^{n-k} + 1} \tag{3.1}
$$

$$
= 1 - \frac{q^{n-k+i} + q^{k+i} + 1 - q^k - q^{n-k} - q^{2i}}{q^n - q^k - q^{n-k} + 1} .
$$

We claim that

$$
\frac{q^{n-k+i} + q^{k+i} + 1 - q^k - q^{n-k} - q^{2i}}{q^n - q^k - q^{n-k} + 1} \leq \frac{2}{q^{k-i}}
$$

for $q \geq 3$. This is equivalent to

$$
q^{2k} + q^{k-i} + 2q^k + 2q^{n-k} \leq q^n + 2 + q^{2k-i} + q^{n-i} .
$$

Without loss, take $k \leq n/2$. We thus observe, assuming $1 \leq i \leq k - 1$

$$
q^{2k} \leq q^n \quad \text{for} \quad q \geq 2 ,
$$

$$
2q^{n-k} \leq q^{n-i} \quad \text{for} \quad q \geq 2 ,
$$

$$
q^{n-k} + 2q^k \leq q^{k+1} \leq q^{2k-i} \quad \text{for} \quad q \geq 3 .
$$

Our claim follows and therefore we obtain the following fact:

**Lemma 3.2.** With respect to the above,

$$
Q_{i,i+1} > 1 - \frac{2}{q^{k-i}} \quad \text{for} \quad q \geq 3 \quad \text{and all} \quad n .
$$
This inequality implies that the nearest neighbor random walk on $Y = \{0, 1, 2, \ldots, k\}$ moves very rapidly from 0 to $k$ for large $q$. This suggests that the variation threshold for $X = X_q$, $q \to \infty$, is $k$. We need a few preliminary results.

**Lemma 3.3.** Suppose $a, j$ are positive integers with $j - 1 < a$. Then

$$ (q^a - 1) (q^{a-1} - 1) \cdots (q^{a-(j-1)} - 1) \geq q^{\sum_{i=0}^{j-1} b_i} - \sum_{\ell=0}^{j-1} q^{c_\ell} $$

where

$$ b_i = a - i $$

$$ 0 \leq i \leq j - 1 $$

$$ c_\ell = \left( \sum_{i=0}^{\ell-1} b_i \right) - b_\ell $$

$$ 0 \leq \ell \leq j - 1. $$

**Proof.** This is proved by induction on $j$. For any positive integer $N$ one proves easily that

$$ 1 + q + q^2 + \cdots + q^N \leq 2q^N. \quad (3.4) $$

Since the constants $c_\ell$ are distinct and increasing, combining (3.3) and (3.4) we have

$$ (q^a - 1) (q^{a-1} - 1) \cdots (q^{a-(j-1)} - 1) \geq q^{\sum_{i=0}^{j-1} b_i} - 2 q^{c_{j-1}} \quad (3.5) $$

**Lemma 3.6.**

$$ \pi(k) > 1 - \frac{2}{q} \quad (q \geq 2 \text{ and all } n). $$

**Proof.**

$$ \pi(k) = \frac{(q^{n-k} - 1) (q^{n-k-1} - 1) \cdots (q^{n-k-(k-1)} - 1) q^{k^2}}{(q^n - 1) (q^{n-1} - 1) \cdots (q^{n-(k-1)} - 1)} $$

by (1.1). We observe that

$$ (q^{n-k} - 1) (q^{n-k-1} - 1) \cdots (q^{n-k-(k-1)} - 1) q^{k^2} = q^s - q^t + \cdots $$

$$ (q^n - 1) (q^{n-1} - 1) \cdots (q^{n-(k-1)} - 1) = q^s - q^u + \cdots, $$
where
\[s = kn - \frac{(k-1)k}{2}\]
\[t = kn - n + \frac{5}{2} k - \frac{1}{2} k^2 - 1\]
\[u = kn - n + \frac{3}{2} k - \frac{1}{2} k^2 - 1\,.

Note that
\[s - t = n - 2k + 1 \geq 1\quad \text{since} \quad k \leq \frac{n}{2} \quad \text{and} \quad t - u = k\,.

By applying (3.5) to the numerator of \(\pi(k)\),
\[
\pi(k) > \frac{q^s - 2q^t}{q^s - y(q)} \quad \text{(where \(\deg y = u\))}
\]
\[
= \frac{q^s - y(q) - (2q^t - y(q))}{q^s - y(q)}
\]
\[
= 1 - \frac{2q^t - y(q)}{q^s - y(q)}
\]

But
\[
\frac{2q^t - y(q)}{q^s - y(q)} \leq \frac{2}{q}
\]

and our assertion follows. (That \(y(q) > 0\) may be seen by employing induction on \(k\) in the denominator of \(\pi(k)\).)

**Proposition 3.7.** The variation threshold for the nearest neighbor random walk on \(X = X_q\), \(q \to \infty\), is \(k\).

**Proof.** For \(\varepsilon > 0\),
\[
d_q ((1 - \varepsilon)k) = \left| |P_{(1-\varepsilon)k} - \pi|\right| \geq |P_{(1-\varepsilon)k} (k) - \pi(k)|
\]

However,
\[P_{(1-\varepsilon)k} (k) = 0\]
and by Lemma 3.6 we have

\[ 1 - \frac{2}{q} < \pi(k) \leq \| P_{(1-\epsilon)k} - \pi \|. \]

Clearly,

\[ \lim_{q \to \infty} d_q((1 - \epsilon)k) = 1. \]

(For any probability measures \( P, Q \) it is plain that \( \| P - Q \| \leq 1 \).)

It remains to show that

\[ \lim_{q \to \infty} d_q((1 + \epsilon)k) = 0. \]

We appeal to Lemma 2.7. Here, we take \( m_0 = j_0 = k, \ i_0 = 0 \), and

\[ c = c(q) = \prod_{i=1}^{k-1} \left( 1 - \frac{2}{q^{k-i}} \right). \]

Thus

\[ d_q((1 + \epsilon)k) \leq (1 - c(q))^{\left\lceil \frac{(1+\epsilon)k}{k} \right\rceil} \leq 1 - c(q) \]

since the latter is between 0 and 1. But plainly, \( \lim_{q \to \infty} c(q) = 1 \) and we are done.

**Remarks.** Note that by Lemma 2.4

\[ c \geq \left( 1 - \frac{2}{q} \right) \prod_{i=2}^{k-1} e^{\frac{2}{q^i} \left( -1 - \frac{2}{q^i} \right)} = \left( 1 - \frac{2}{q} \right) e^{-2 \left( \sum_{j=2}^{k-1} \left( \frac{1}{q^j} + \frac{2}{(q^j)^2} \right) \right)} . \]

However,

\[ e^{-2 \left( \sum_{j=2}^{k-1} \left( \frac{1}{q^j} + \frac{2}{(q^j)^2} \right) \right)} \geq e^{-2 \left( \sum_{j=2}^{\infty} \left( \frac{1}{q^j} + \frac{2}{(q^j)^2} \right) \right)} = e^{-\frac{2}{q(q-1)} - \frac{4}{q^2(q^2-1)}} . \]

From here and from Lemma 3.6, we get

**Proposition 3.6.** Consider the nearest neighbor random walk on \( X = X_q \). For all \( \epsilon > 0 \) and all \( q \) (a power of a prime)

\[ \left( 1 - \frac{2}{q} \right) \leq \| P_{(1-\epsilon)k} - \pi \|. \]
\[ \| P_{(1+\varepsilon)k} - \pi \| \leq 1 - \left( 1 - \frac{2}{q} \right) e^{-\frac{2}{q(q-1)}} - \frac{4}{q^2(q-1)} \].

The inequalities of (3.8) imply immediately Proposition 3.7 but they say much more. The existence of the cutoff phenomenon is a statement about large \( q \). But these inequalities are noteworthy for say, \( q = 11 \). Here,

\[ \frac{2}{q(q-1)} - \frac{4}{q^2(q-1)} \]

is negligible and so if we take \( \varepsilon = 1/k \), then practically,

\[ \frac{9}{11} \leq \| P_{k-1} - \pi \|, \]

\[ \| P_{k+1} - \pi \| \leq \frac{2}{11}. \]

Thus the variation threshold begins to take shape for small \( q \) independent of \( k \).

We turn our attention to the process where \( n \) goes to infinity and \( q, k \) are fixed. It is plain from (3.1) that

\[ \lim_{n \to \infty} Q_{i,i+1} = \frac{1 - \frac{1}{q^{i+1}}}{1 - \frac{1}{q^k}} < 1 \]  \hspace{1cm} (3.9)

for \( i \leq i \leq k-1 \), and so, roughly, for large \( n \) our walk starts at 0 but does some zig-zagging before reaching state \( k \). This suggests

**Proposition 3.10.** There is no cutoff phenomenon for the nearest neighbor random walk on \( X = X_n, q \geq 2 \), as \( n \) tends to infinity.

**Proof.** Assume first of all that \( q \geq 3 \). Suppose it is false so that there is a sequence \( \tau_d(n) \) satisfying (1.3). We claim that \( \lim_{n \to \infty} \tau_d(n) = \infty \) for otherwise \( \tau_d(n) \) has a bounded subsequence and hence a convergent subsequence \( \tau_d(n_\varepsilon) \), say, \( \lim_{\varepsilon \to \infty} \tau_d(n_\varepsilon) = z \). We show there is no such \( z \). If \( z < k + 2 \), consider

\[ |P_{(1+\varepsilon)\tau_d(n_\varepsilon)}(k) - \pi^{n_\varepsilon}(k)|. \]
One verifies easily that \( \lim_{n \to \infty} \pi(k) = 1 \). However, for sufficiently small \( \varepsilon \) and sufficiently large \( \ell \), \( (1 + \varepsilon)\tau_d(n_\ell) \) is less than \( k + 2 \). By (3.9) \( P_{(1+\varepsilon)\tau_d(n_\ell)}(k) \) does not approach 1 as \( n \) tends to infinity, and so

\[
\lim_{\ell \to \infty} d_{n_\ell}((1 + \varepsilon)\tau_d(n_\ell)) \neq 0
\]

which is a contradiction. If \( z \geq k + 2 \) then \((1 - \varepsilon)\tau_d(n_\ell)\) may be considered greater than \( k + 1 \). By Lemma 2.7,

\[
\left\| P_{(1-\varepsilon)\tau_d(n_\ell)} - \pi \right\| \leq 1 - c ,
\]

where

\[
c = \prod_{i=1}^{k-1} \left( 1 - \frac{2}{q^{k-i}} \right).
\]

Since \( c > 0 \),

\[
\lim_{\ell \to \infty} \left\| P_{(1-\varepsilon)\tau_d(n_\ell)} - \pi \right\| \neq 1
\]

which is not possible. Thus

\[
\lim_{n \to \infty} \tau_d(n) = \infty.
\]

It follows that

\[
\lim_{n \to \infty} (1 - \varepsilon)\tau_d(n) = \infty
\]

and so

\[
\lim_{n \to \infty} \left\| P_{(1-\varepsilon)\tau_d(n)} - \pi \right\| = 0
\]

by Lemma 2.7. This is impossible and the proof is complete for \( q \geq 3 \).

If \( q = 2 \), by (3.9), there is an \( \eta > 0 \) and an \( M > 0 \) such that \( P_k(k) > \eta \) for \( n > M \) (\( \eta \) corresponds to \( c \) above). The proof now proceeds as in the case of \( q \geq 3 \).

Although there is no variation threshold on \( X_n, n \to \infty \), Lemma 2.7 provides good information on the number of steps required to become uniform:

\[
\left\| P_m - \pi \right\| \leq (1 - c)^{\left\lfloor \frac{m}{q} \right\rfloor} = c^{\left\lfloor \frac{m}{q} \right\rfloor \log(1-c)}
\]  

(3.11)
where $c$ can be taken to be

$$
\left( 1 - \frac{2}{q} \right) e^{\frac{2}{q^2(q^2 - 1)} - \frac{2}{q^2(q^2 - 1)}} .
$$

If $q = 3$ (resp. 11), then $m = 6k$ (resp. $2k$) steps would make the right side of (3.11) small. If $q \geq 27$, then $k$ steps would suffice.

On the other hand, if $m < k$, then $||P_m - \pi||$ is close to 1 even for small $n$. This may be seen as follows:

$$
\begin{align*}
||P_m - \pi|| &\geq \pi(k) \\
&= \frac{(q^{n-k} - 1)(q^{n-2k-1} - 1) \cdots (q^{n-k-(k-1)} - 1)}{(q^n - 1)(q^{n-1} - 1) \cdots (q^{n-(k-1)} - 1)} q^{k^2} \\
&= \left( \frac{q^n - q^k}{q^n - 1} \right) \left( \frac{q^{n-1} - q^k}{q^{n-1} - 1} \right) \cdots \left( \frac{q^{n-(k-1)} - q^k}{q^{n-(k-1)} - 1} \right) \\
&> \left( 1 - \frac{1}{q^{n-k}} \right) \left( 1 - \frac{1}{q^{n-k-1}} \right) \cdots \left( 1 - \frac{1}{q^{n-k-(k-1)}} \right) \\
&> e^{-\frac{1}{q^n-k} \left( 1 + \frac{1}{q^n-k-1} \right) \cdots \left( 1 + \frac{1}{q^n-k-(k-1)} \right)} - \frac{1}{q^n-k-(k-1)} \left( 1 + \frac{1}{q^n-k-(k-1)} \right)^2 \\
&> e^{-\frac{1}{q^n-k-k+1} \left( 1 + \frac{1}{q^n-k-k+1} \right)^2}.
\end{align*}
$$

Thus, for $n$ slightly bigger than $2k$ and $m < k$, $||P_m - \pi||$ is close to 1. This is, of course, not much of a restriction on $n$ since $n \geq 2k$ to begin with. We summarize these considerations with

**Proposition 3.11.** With respect to the nearest neighbor random walk on $X = X_n$, $n \to \infty$, $c(q)k$ steps are both necessary and sufficient for convergence to stationarity.

4. Subspaces of Varying Dimension.

We now consider the case where $k$ grows big with $n$. For simplicity, we take $k = \frac{n}{2}$. Since $k$ varies, we can no longer appeal to Lemma 2.7 which proved so useful in the previous
section. Instead, we shall make use of Fourier Analysis on the group \( G = GL_n(F_q) \) and the Upper Bound Lemma of Diaconis [D]. We now give some background for this approach. In doing so, we follow the setup of Diaconis & Shahshahani [DS] where more details can be found. See also Belsley [B] for an even more complete explanation.

The group \( G \) acts transitively on \( X = X_n \) in a natural way. If \( Z_k \in X \) let

\[
K = \{ A \in G: AZ_k = Z_k \}
\]

be the isotropy subgroup of \( G \) relative to \( Z_k \). Of course \( AK \to AZ_k \) defines a one to one correspondence from \( G/K \) to \( X_n \) which respects the action of \( G \). The original probability on \( X_n \) is uniform on the distance one neighbors of \( Z_k \) and can be lifted in an obvious way to a \( K \)-bi-invariant probability \( P \) on \( G \) (i.e., \( P(ABC) = P(B) \) for all \( B \in G, A, C \in K \)). Now \( (G, K) \) is a Gelfand pair and let

\[
L(X) = V_0 \oplus V_1 \oplus \cdots \oplus V_k
\]

be a decomposition into distinct irreducible representations of \( G \). We can assume

\[
\dim V_i = \begin{bmatrix} m \\ i \end{bmatrix}_q - \begin{bmatrix} m \\ i - 1 \end{bmatrix}_q, \quad \text{for } i \geq 1 \text{ and } \dim V_0 = 1.
\]

Here, \( L(X) \) is the set of all functions from \( X \) to the complex numbers, and \( V_0 \) is the subspace of constant functions. Each \( V_i \) has a unique one-dimensional subspace of left \( K \)-invariant functions. If we take an element \( s_i \) of this subspace of \( V_i \) which is normalized so that \( s_i(Z_k) = 1 \), then \( s_i \) is the so-called \( i \)th spherical function. The functions \( s_i \) depend only on the distance \( d(W_k, Z_k) \) and are polynomials in \( d \) given by

\[
s_i(d) = \sum_{m=0}^i \frac{(q^{-i}; q)_m (q^{i-n-1}; q)_m (q^{-d}; q)_m q^m}{(q; q)_m (q^{k-n}; q)_m (q^{-k}; q)_m}, \quad 0 \leq i \leq k,
\]

where

\[
(a; q)_m = (1 - a) (1 - aq) \cdots (1 - aq^{m-1}) \quad \text{for } m > 0 \text{ and } (a; q)_0 = 1.
\]

These functions are known as \( q \)-Hahn polynomials. See Stanton [St] for more information.

The following result is proved in Diaconis [D].
**Proposition 4.1.** In the above setup,

\[
||P_m - \pi||^2 \leq \frac{1}{4} \sum_{i=1}^{m} d_i |s_i(1)|^{2m}
\]

where \(d_i = \dim V_i\).

We presently show that \(\frac{n}{2}\) is a variation threshold for this process. But among other things, we first need to bound \(d_i\) and \(s_i(1)\). With regard to \(s_i(1)\), there isn't much of a problem:

\[
s_i(1) = 1 - \frac{(1 - q^{-i}) (1 - q^{i-n-1})}{(1 - q^{k-n}) (1 - q^{-k})}
\]

\[
= \frac{(1 - q^{k-n}) (1 - q^{-k}) - (1 - q^{-i}) (1 - q^{i-n-i})}{(1 - q^{k-n}) (1 - q^{-k})}
\]

\[
\leq \frac{(1 - q^{i-n-1}) [(1 - q^{-k}) - (1 - q^{-i})]}{(1 - q^{k-n}) (1 - q^{-k})}
\]

\[
\leq \frac{q^{-i} - q^{-k}}{(1 - q^{k-n}) (1 - q^{-k})}
\]

\[
\leq \frac{q^{-i}}{1 - q^{k-n}}.
\]

Since

\[
\frac{1}{1-x} \leq 1 + 2x \leq e^{2x} \quad \text{for} \quad 0 \leq x \leq \frac{1}{2}
\]

it follows that

\[
s_i(1) \leq \frac{\exp(2q^{k-n})}{q^i}.
\]

This bound was obtained by Greenhalgh [Gr]. The situation for \(d_i\) (of course it suffices to bound \(\left[n\atop k\right]_q\)) is a bit more complicated. Here, a suggestion of George Andrews was very useful and led to Lemma 4.2 below.

A partition of a positive integer \(n\) is a finite nonincreasing sequence of positive integers \(\lambda_1, \lambda_2, ..., \lambda_r\) such that \(\Sigma_{i=1}^r \lambda_i = n\). The partition function \(p(n)\) is the number of partitions of \(n\). Let \(p(N, M, n)\) denote the number of partitions of \(n\) into at most \(M\)
parts, each of size less than or equal to $N$. Clearly,

$$p(N, M, n) = 0 \quad \text{if} \quad n > MN ,$$

$$p(N, M, NM) = 1 .$$

Thus the generating function

$$G(N, M; q) = \sum_{n \geq 0} p(N, M, n) q^n$$

is a polynomial in $q$ of degree $NM$. It is well known (See Th. 3.1 of [An]) that

$$G(N, M; q) = {M + N \choose N}_q .$$

If

$$A(x) = \prod_{m=1}^{\infty} \frac{1}{1 - x^m} ,$$

it is also known (5.2.1 of [An]) that

$$\sum_{m \geq 0} p(m)x^m = A(x) .$$

Here $p(0) = 1$ and $A(x)$ converges for $|x| < 1$. A polynomial

$$p(x) = \sum_{i=0}^{m} a_i x^i$$

is called reciprocal if for each $i, a_i = a_{m-i}$ or equivalently

$$x^m p\left(\frac{1}{x}\right) = p(x) .$$

**Lemma 4.2.** If $q > 1$, then

$$q^{k(n-k)} \leq {n \choose k}_q \leq A\left(\frac{1}{q}\right) q^{k(n-k)} .$$
Proof. In Theorem 3.10 of Andrews [An], it is shown that \( \left[ \begin{array}{c} n \\ k \end{array} \right]_x \) is a reciprocal polynomial in \( x \). Hence
\[
\left[ \begin{array}{c} n \\ k \end{array} \right]_q = q^{k(n-k)} \left[ \begin{array}{c} n \\ k \end{array} \right]_{\frac{1}{q}}
\]
\[
= q^{k(n-k)} \sum_{r \geq 0} p(k, n-k, r) \left( \frac{1}{q} \right)^r
\]
\[
\leq q^{k(n-k)} \sum_{r \geq 0} p(r) \left( \frac{1}{q} \right)^r
\]
\[
= q^{k(n-k)} A \left( \frac{1}{q} \right).
\]
Since
\[
\left[ \begin{array}{c} n \\ k \end{array} \right]_{\frac{1}{q}} > 1,
\]
the first equality above implies that
\[
q^{k(n-k)} \leq \left[ \begin{array}{c} n \\ k \end{array} \right]_q
\]
and so we are done.

For our purposes, we can be a little more specific:

Lemma 4.3 If \( q \geq 3 \), then
\[
q^{k(n-k)} \leq \left[ \begin{array}{c} n \\ k \end{array} \right]_q \leq e^{\frac{1}{q^2-1}} q^{k(n-k)}.
\]

Proof. In view of Lemma 4.2, it suffices to show that
\[
A \left( \frac{1}{q} \right) < e^{\frac{1}{q^2-1}}.
\]
Now
\[
A \left( \frac{1}{q} \right) = \prod_{m=1}^{\infty} \frac{1}{1 - \left( \frac{1}{q} \right)^m}
\]
and from Lemma 2.5
\[
e^{-\left( \frac{1}{4} \right)^m (1+\left( \frac{1}{4} \right)^m)} < 1 - \left( \frac{1}{q} \right)^m
\]
19
which implies
\[
\frac{1}{1 - \left(\frac{1}{q}\right)^m} < \left(\frac{1}{q}\right)^m \left(1 + \left(\frac{1}{q}\right)^m\right).
\]

The result is now clear. For \(q = 3, 5, 9, 25\), \(e^{\frac{1}{q} + \frac{1}{q^2 - 1}}\) is respectively \(1.87, 1.37, 1.14\), and 1.04. Thus Lemma 4.2 provides a very sharp bound for \([n \atop k]_x\).

We now give an upper bound for the variation distance.

**Proposition 4.4.** Let \(P_m\) be the law of the process on

\[
X = X_{n,k,q} = X_{n,\frac{n}{2},q} \ (q \geq 3).
\]

If

\[
m = \frac{n}{2} + c
\]

and

\[
C = C(q) = \sqrt{\frac{1}{8}} \left(\frac{\pi}{\log q}\right)^{\frac{1}{4}} e^{\frac{2}{\log q}} a, \quad D_n(q) = \log q - \frac{2}{q^{\frac{1}{2}}},
\]

then for all \(n \geq 3\) we have

\[
||P_m - \pi_n|| \leq C e^{-D_n c}.
\]

**Proof.** Appealing to Proposition 4.1 and the preceding discussion, we have

\[
||P_m - \pi_n||^2 \leq \frac{1}{4} e^{\frac{1}{q^2 - 1} + \frac{1}{q^2 - 1}} \sum_{i=1}^{m/2} q^{ni-i^2} q^{-2mi} \exp(4m q^{-n/2}).
\]

Here we have used the simple bound \(d_i \leq \left[\begin{array}{c} n \\ i \end{array}\right]_q\). To bound the sum, consider first the term corresponding to \(i = 1:\)

\[
q^{n-1} q^{-2m} \exp(4m q^{-\frac{m}{2}}).
\]
To drive $q^{n-1} q^{-2m}$ to zero, $m$ must be at least $\frac{n}{2} + c$. With $m$ of this form, the $i^{th}$ term is
\[
q^{ni-i^2} q^{-2(\frac{n}{2} + c)} e^{i(\frac{n}{2} + c)/q^{\frac{3}{2}}}
\]
\[
= q^{ni-i^2} q^{-ni-2ci} e^{2n/q^{\frac{3}{2}}}
\]
\[
= q^{-i^2} e^{2n/q^{\frac{3}{2}}} q^{-2ci} e^{4c/q^{\frac{3}{2}}}
\]
\[
\leq q^{-2c} e^{4c/q^{\frac{3}{2}}} q^{-i^2} e^{2n/q^{\frac{3}{2}}}.
\]
Thus it suffices to prove
\[
\sum_{i=1}^{n/2} q^{-i^2} e^{2n/q^{\frac{3}{2}}} \leq M \text{ independent of } n.
\]
However, note that
\[
\sum_{i=1}^{\infty} \frac{1}{q^{i^2}} \leq \int_{0}^{\infty} \frac{1}{q^{x^2}} = \frac{1}{2} \sqrt{\frac{\pi}{\log q}}.
\]
Also, $2x/q^{\frac{3}{2}}$ achieves its unique maximum value at
\[
x = \frac{2}{\log q}.
\]
Thus
\[
\frac{2n}{q^{\frac{n}{2}}} \leq \frac{\frac{4}{\log q}}{q^{\frac{1}{2}\log q}} = \frac{4}{e \log q}
\]
and
\[
e^{2n/(q^{\frac{3}{2}})} \leq e^{\left(\frac{4}{e \log q}\right)}.
\]
We have
\[
\|P_m - \pi_n\|^2 \leq \frac{1}{4} \frac{1}{2} \sqrt{\frac{\pi}{\log q}} e^{\left(\frac{4}{e \log q}\right)} e^{2(\log q - \frac{2c}{q^{\frac{3}{2}}})(-c)}.
\]
Thus
\[
\|P_m - \pi_n\| \leq \sqrt{\frac{1}{8}} \left(\frac{\pi}{\log q}\right)^{\frac{1}{4}} e^{\left(\frac{2}{e \log q}\right)} e^{\left((\log q - \frac{2c}{q^{\frac{3}{2}}})(-c)\right)}.
\]
Put
\[ C(q) = \sqrt{\frac{1}{8}} \left( \frac{\pi}{\log q} \right)^{\frac{1}{4}} e^{\left( \frac{2}{\log q} \right)} , \quad D_n(q) = \log q - \frac{2}{q^\frac{3}{2}} . \]

Comment. Note that \( C(q) \) decreases with increasing \( q \), \( C(3) = .8984 \), and \( \lim_{q \to \infty} C(q) = 0 \) with \( C(q) \) approaching 0 fairly fast. Also, \( d_n(q) \) increases with increasing \( q \), and
\[ |D_n(q) - \log q| = \frac{2}{q^\frac{3}{2}} . \]
Thus, the bound is quite good for small \( q \) and can only improve for increasing \( q \). It is extremely sharp for large \( q \).

Concerning a lower bound for the variation distance, we have the following.

Proposition 4.5. Let \( P_m \) be the law of the process on
\[ X = X_{n,k,q} = X_{n,\frac{n}{2},q} \quad (q \geq 3) \]
and let
\[ E(q) = e^{\frac{1}{q-1} + \frac{1}{q^2-1}} . \]
If
\[ m = \frac{n}{2} - c , \]
then
\[ \|P_m - \pi\| \geq 1 - E^2(q) \left( \frac{1}{q^{c/2}} + \frac{1}{q^{(c+1)/2}-1} \right) . \]

Proof.
\[ P_m(j) = 0 \quad \text{for} \quad j > \frac{n}{2} - c \]
and so
\[ \|P_m - \pi_n\| \geq \sum_{j=(n/2) - c+1}^{n/2} \pi_n(j) = 1 - \sum_{j=0}^{(n/2)-c} \pi_n(j) . \]

From (1.1) and Lemma 4.3,
\[
\sum_{j=0}^{(n/2)-c} \pi_n(j) \leq \frac{E^2(q) \sum_{j=0}^{(n/2)-c} q^{2((n/2)-j)j} q^{j^2}}{q^{n^2/4}}.
\]

Note that
\[
q^{2((n/2)-j)j} q^{j^2} = q^{n-j^2}
\]

and after completing the square, we have
\[
\sum_{j=0}^{(n/2)-c} \pi_n(j) \leq E^2(q) \sum_{0}^{(n/2)-c} q^{-(j-(n/2))^2}
\]
\[
\leq E^2(q) \sum_{\ell=c}^{\infty} q^{-\ell^2}
\]
\[
\leq E^2(q) \left( \frac{1}{qc^2} + \frac{1}{q(c+1)^2 - 1} \right).
\]

Propositions 4.4 and 4.5 demonstrate that there is an extremely sharp cutoff for the process at \( \frac{n}{2} \) steps. Somewhat fewer steps yield a variation very close to 1 while slightly more steps drives this distance to 0 exponentially fast.

5. Independence of the Starting Point.

In the above situation, it is intuitively clear that the rate of convergence to the stationary distribution does not depend on the initial fixed point \( Z_k \) of \( X \). This fact is typical of random walks on finite groups but is not true in general for Markov chains. To see this, consider a stack of cards numbers 1 to \( n \) from top to bottom. The top card is removed and then inserted into the deck at random. The procedure is repeated. Let \( X_m \) be the position of the ♠ ace at time \( m \). This defines a Markov chain on \( \{1, 2, ..., n\} \). If \( X_0 = 1 \), then one step is required to achieve the stationary distribution. On the other hand, if \( X_0 = n \), then \( n \log n \) steps are needed. See Aldous and Diaconis [AD1] for details.

We now provide a class of random walks (which include random walks on groups) where convergence to stationarity is independent of the starting point. Specifically, this is the case for walks on homogeneous spaces induced by a Gelfand pair \((G, K)\) and hence on 2-point homogeneous graphs. A list of pertinent examples may be found in [DS] and
Stanton [St] and include random walks on certain sets of matrices, vector subspaces (the problem considered in this paper), the n-cube $Z_2^n$, and partitions of a set with an even number of elements.

In what follows, $(G, K)$ is a Gelfand pair. If $P$ is a probability on $G$, let $\tilde{P}$ be the induced probability on $X$. We need a few lemmas.

**Lemma 5.1.** If $Q$ is a bi-$K$-invariant probability on $G$, then $\delta_x \ast \tilde{Q}^k$ is the law of the induced process (right random walk) starting at $x \in X$, i.e.,

$$\delta_x \ast \tilde{Q}^k(y) = \delta_x \ast Q^k(yK)$$

$$= |K| \delta_x \ast Q^k(y) \quad (Q \text{ is bi-$K$-invariant})$$

$$= |K| \sum_{t \in G} \delta_x(t) \ Q^k(t^{-1}y)$$

$$= |K| \ Q^k(x^{-1}y)$$

$$= Q^k(x^{-1}yK)$$

$$= Q^k(x, y). \quad \text{(See [DS])}$$

**Lemma 5.2.** Let $P, Q, U$ be probabilities on $G$ with $Q$ right-$K$-invariant and $U$ the uniform probability on $G$. It follows that

$$||P \ast Q - \tilde{U}|| \leq ||\tilde{Q} - \tilde{U}||.$$

**Proof.** Suppose

$$G = \{t_0, t_1, \cdots\}$$

where $t_0 = x_0 = e$ and $X = \{x_0, x_1, \cdots, x_m\}$. Let $B$ be any subset of $X$, say, $B = \{y_1, y_2, \cdots, y_i\}$.

$$P \ast Q(x_i) = P \ast Q(x_iK) = |K|P \ast Q(x_i) \quad (Q \text{ is right-$K$-invariant})$$

24
Thus
\[ P \tilde{\star} Q (B) = |K| \left[ P(t_0) Q(t_0^{-1}y_1) + P(t_1) Q(t_1^{-1}y_1) + \cdots \right] \]
\[ + |K| \left[ P(t_0) Q(t_0^{-1}y_2) + P(t_1) Q(t_1^{-1}y_2) + \cdots \right] \]
\[ + \cdots \]
\[ + |K| \left[ P(t_0) Q(t_0^{-1}y_\ell) + \cdots \right] \]
\[ = \sum_{t_j \in G} P(t_j) \tilde{Q} (t_j^{-1}B) . \]

Similarly,
\[ \tilde{U}(B) = \sum_{t_j \in G} P(t_j) \tilde{U} (t_j^{-1}B) . \]

Thus,
\[ ||P \tilde{\star} Q(B) - \tilde{U}(B)|| \leq \sum_{t_j \in G} P(t_j) |\tilde{Q} (t_j^{-1}B) - \tilde{U} (t_j^{-1}B)| \]
\[ \leq \sum_{t_j \in G} P(t_j) ||\tilde{Q} - \tilde{U}|| \]
\[ = ||\tilde{Q} - \tilde{U}|| . \]

Taking the supremum overall, \( B \subseteq X \), our assertion is proved.

We now can prove

**Proposition 5.3.** The rate of convergence of a right random walk with respect to the homogeneous space of a Gelfand pair is invariant of the initial fixed point.

**Proof.** By Lemma 5.2
\[ ||\delta_x \tilde{\star} Q^* - \tilde{U}|| \leq ||\tilde{Q}^* - \tilde{U}|| . \]

However,
\[ Q^* = \delta_x^{-1} \ast \delta_x \ast Q^* . \]
Thus
\[ \| Q^* - \tilde{U} \| \leq \| \delta_x Q^* - \tilde{U} \| \]
and
\[ \| Q^* - \tilde{U} \| = \| \delta_x Q^* - \tilde{U} \|. \]

6. Comparison With the Calabi-Wilf Algorithm.

We commented in the Introduction section that the algorithm of Calabi and Wilf [CW] requires $C kn$ steps to generate a random subspace of dimension $k$ over $V_n$. We would like to compare this with the time required for our algorithm to operate.

To this end, identify $V_n$ with n-tuples having entries from $GF(q)$. By Proposition 5.3, convergence to stationarity is independent of the starting point and so we take

\[ W_k = Sp\{e_1, e_2, \ldots, e_k\} \]

where $e_j \in V_n$ is the $j^{th}$ unit vector. By [CW], choose linearly independent vectors $w_1, \ldots, w_{k-1}$ of $W_k$ in $C(k-1)k$ steps. A vector $w_k$ of $W_k$ which is not in $Sp\{w_1, \ldots, w_{k-1}\}$ can be found in time $Ck^2$. Note that

\[ \{w_1, w_2, \ldots, w_k, e_{k+1}, \ldots, e_n\} \]

is a basis for $V_n$ and that $C \cdot 1 \cdot (n-k+1)$ steps are required to produce a 1-dimensional subspace of

\[ Sp\{w_1, e_{k+1}, \ldots, e_n\} \].

We are now at distance 1 from $W_k$ and so this procedure must be repeated $Ck$ times. Thus our algorithm operates in $C(nk + k^3)$ steps versus $Cnk$ for the procedure of Calabi and Wilf. When $k = \frac{n}{2}$, this becomes $Cn^3$ versus $Cn^2$.

Acknowledgments

The author would like to thank George Andrews, Persi Diaconis, Jeff Rosenthal, and Zeev Rudnick for a number of indispensable conversations. Thanks is also extended to the Department of Statistics of Stanford University for their hospitality during the summers of 1991 and 1992.
References


