Random Rotations: Characters and Random Walks on SO(N)

(Appeared in *Annals of Probability* **22** (1994), 398–423.)

by

Jeffrey S. Rosenthal

University of Minnesota

Summary. We analyze a random walk on the orthogonal group $SO(N)$ given by repeatedly rotating by a fixed angle through randomly chosen planes of $\mathbf{R}^N$. We derive estimates of the rate at which this random walk will converge to Haar measure on $SO(N)$, using character theory and the Upper Bound Lemma of Diaconis and Shashahani. In some cases we are able to establish the existence of a "cut-off phenomenon" for the random walk. This is the first interesting such result on a non-finite group.

Key words and phrases: Random walk, Haar measure, Rate of convergence, Upper bound lemma, Cut-off phenomenon.

MSC 1991 classification numbers: 60J05, 60B15, 43A75.

Running head: Random Rotations.

1. Introduction.

When studying convergence to a stationary distribution, Markov chain theory has traditionally concerned itself with asymptotic issues, including the asymptotic exponential rate of convergence. More recently, *non*-asymptotic convergence rates have become a topic of interest. The question becomes, given a Markov chain and an initial distribution, how many iterations are required before the chain is "close" to its stationary distribution? This question has been motivated by such diverse areas as card shuffling ("How many times do you have to shuffle a deck of cards to make it random?"; see [D] for background) and stochastic algorithms ("How long do you have to run the algorithm until the answers are satisfactory?"; see e.g. [GS] and [R]). In each case, it is desired to know how long a Markov chain should be run until it has converged to the desired stationary distribution.

The study of non-asymptotic convergence rates often yields interesting results. The best known of these is the "cut-off phenomonon" of Diaconis and Shashahani [DS1] (see also [AD] and [D]) in which the variation distance to stationarity decreases sharply from 1 to 0 over a relatively short length of time. This phenomenon has been observed in a number of random walks on finite groups, using such techniques as Coupling, Strong Stopping Times, and Fourier Analysis. See Diaconis [D] for an excellent, extensive survey of known examples and methods. See Hildebrand [H] for some recent results.

On a compact Lie group, many random walks converge in total variation distance to (normalized) Haar measure. Again, one may ask how quickly the convergence occurs. The finite-group methods mentioned above would appear to be applicable to compact groups. In this paper, we present an analysis of a process of "random rotations" on the group $SO(N)$ of real, orthogonal, $N \times N$ matrices with unit determinant. Roughly, this process involves repeatedly picking a "random plane" in $\mathbf{R}^N$, and *rotating* the $(N-1)$-sphere S^{N-1} in that plane through an angle θ. Here θ is a fixed, pre-chosen angle of rotation. We show that $\frac{1}{2(1-\cos\theta)} N \log N$ such random rotations are necessary to get close to Haar measure in total variation distance, to first order in N. We further show that in the case when θ is 180 degrees, $\frac{1}{4} N \log N$ rotations are also sufficient. This allows us to conclude a cut-off phenomenon in this case; this is the first interesting such result on a non-finite group.

The method employed in proving these results is again Fourier Analysis. The necessity of doing $\frac{1}{2(1-\cos\theta)} N \log N$ rotations is proved using the standard technique of showing that if fewer rotations are done, then a certain character of $SO(N)$ will have large expectation value, while it should have expectation value 1 under Haar measure. Calculation of variances, and an appeal to Chebychev's Inequality, then show that the variation distance to Haar measure is large. The sufficiency of doing $\frac{1}{4} N \log N$ rotations when θ is 180 degrees is proved using the "Upper Bound Lemma" of Diaconis and Shashahani (see [DS1], [DS2], [D]). This involves summing the squares of the expected values of *all* of the irreducible characters of $SO(N)$, and showing that this sum is small. To do this, a careful description of the irreducible characters of $SO(N)$ (evaluated at certain group elements) is required. This is developed in Section 3 by means of the Weyl Character Formula.

This paper is organized as follows. Section 2 provides precise definitions of the process being studied, as well as a precise statement of the results obtained. Section 3 describes the irreducible characters of $SO(N)$ in sufficient detail to be able to apply the Upper Bound Lemma. Section 4 discusses Fourier analysis. It describes the Upper Bound Lemma of Diaconis and Shashahani, and includes a proof for the case of conjugate-invariant measures on a compact Lie group (the case at hand). Section 5 proves the necessity of doing $\frac{1}{2(1-\cos\theta)} N \log N$ rotations, by considering the expectation values of a particular character (namely, the ordinary trace). Section 6 applies the Upper Bound Lemma to prove the sufficiency of doing $\frac{1}{4} N \log N$ rotations when θ is 180 degrees.

2. Definitions and Results.

Fix a non-zero angle θ. Let R_θ be the element of $SO(N)$ defined by

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & \dots & 0 \\ \sin\theta & \cos\theta & 0 & \dots & 0 \\ 0 & 0 & & & \\ \vdots & \vdots & & I_{N-2} & \\ 0 & 0 & & & \end{pmatrix},$$

where I_{N-2} is the $(N-2) \times (N-2)$ identity matrix. We define the following random process on $SO(N)$. Let X_0 be the $N \times N$ identity matrix. At each step $k \geq 1$, choose at random an element B_k of $SO(N)$ in the conjugacy class of R_θ, and set $X_k = B_k X_{k-1}$. (Here "at random" means according to the unique probability measure concentrated on the conjugacy class of R_θ which is invariant under conjugation by $SO(N)$. Equivalently, we may set $B_k = C_k^{-1} R_\theta C_k$, where C_k is an element of $SO(N)$ chosen at random according to Haar measure.)

Note that the process is equivalent to acting on the $(N-1)$-sphere S^{N-1} by means of rotating it, at each step, by an angle θ through a random two-dimensional subspace of $\mathbf{R}^N$.

If we write Q_k for the probability measure on $SO(N)$ for the random variable X_k (and suppress the dependence on θ and N), we have that Q_0 is the measure concentrated at the identity matrix, that Q_1 is the measure which is uniformly concentrated on the conjugacy class of R_θ, and that for each k, Q_k is the *convolution* of Q_1 with Q_{k-1}:

$$Q_k(A) = (Q_1 * Q_{k-1})(A) \equiv \int\limits_{SO(N)} Q_{k-1}(s^{-1}A)\, dQ_1(s)$$

for each measureable set $A \subseteq SO(N)$. This convolution defines a measure on the Borel subsets of $SO(N)$ – see Hewitt and Ross [HR], Theorems 19.6 and 19.11. Proceeding by induction, we see that

$$Q_k = Q_1 * Q_1 * \ldots * Q_1 \equiv Q_1^{*k}\ , \tag{2-1}$$

the k-fold convolution of the measure Q_1 with itself.

Given two Borel probability measures μ and ν on a set X, we define their *variation distance* $\|\mu - \nu\|$ by

$$\|\mu - \nu\| = \sup_{A \subseteq X} |\mu(A) - \nu(A)| \ ,$$

where the supremum is over all Borel subsets A of X. The variation distance is always between 0 and 1. In terms of this variation distance, we may ask very precise questions about our random processes above. Writing λ for normalized Haar measure on $SO(N)$, we may ask how large k has to be to make $\|Q_k - \lambda\|$ small. In this paper, we prove the following results.

Theorem 2-1. *For any non-zero angle θ, there exist constants Γ and Δ (where Δ may depend on θ but Γ does not), such that for any positive integer N, and any positive real number c, if $k = \frac{1}{2(1-\cos\theta)}(N \log N - cN)$, then*

$$\|Q_k - \lambda\| \geq 1 - \Gamma e^{-2c} - \Delta\left(\frac{\log N}{N}\right) \ .$$

Theorem 2-2. *For the case $\theta = \pi$, there are positive constants Λ and γ such that for any integer $N \geq 3$ and any positive real number c, if $k = \frac{1}{4}N \log N + cN$, then*

$$\|Q_k - \lambda\| \leq \Lambda e^{-\gamma c} \ .$$

Remark: Roughly speaking, Theorem 2-1 states that we need to do at least

$\frac{1}{2(1-\cos\theta)} N \log N$ random rotations through an angle θ before we can possibly get close to Haar measure, while Theorem 2-2 states that $\frac{1}{4} N \log N$ random rotations through an angle of 180 degrees is approximately enough to get close to Haar measure.

The representation theory needed to prove the above two theorems is developed in the next two sections. Theorem 2-1 is proved in Section 5, and Theorem 2-2 is proved in Section 6.

3. Character Theory of SO(N).

We present here the required character values of $SO(N)$. For each irreducible character, we require only its dimension and its value at the matrix R_θ defined above. We quote without proof the Weyl Character Formula for compact Lie groups, as applied to $SO(N)$.

The Weyl Character Formula for SO(N).

(A) Let $N = 2n+1$ be odd. Then the irreducible characters of $SO(N)$ can be indexed by n "half-integers" $a_1, a_2, \ldots, a_n$ such that for each j, $a_j - \frac{1}{2}$ is an integer, $a_{j+1} \geq a_j + 1$, and $a_1 \geq \frac{1}{2}$. The value of the corresponding character on an element of $SO(N)$ of the form

$$\begin{pmatrix} \cos(2\pi x_1) & -\sin(2\pi x_1) & \ldots & 0 & 0 & 0 \\ \sin(2\pi x_1) & \cos(2\pi x_1) & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \cos(2\pi x_n) & -\sin(2\pi x_n) & 0 \\ 0 & 0 & \ldots & \sin(2\pi x_n) & \cos(2\pi x_n) & 0 \\ 0 & 0 & \ldots & 0 & 0 & 1 \end{pmatrix}$$

is

$$\frac{\sum\limits_{\tau \in S_n} \sum\limits_{\substack{\epsilon_m = \pm 1 \\ (1 \leq m \leq n)}} \operatorname{sgn}(\tau) \left(\prod\limits_{j=1}^{n} \epsilon_j \right) \exp\left(2\pi i \sum\limits_{q=1}^{n} \epsilon_q a_{\tau(q)} x_q \right)}{\prod\limits_{1 \leq s < r \leq n} \left(\exp\left(\pi i (x_r + x_s)\right) - \exp\left(-\pi i (x_r + x_s)\right) \right)} \times$$

$$\times \frac{1}{\prod\limits_{1 \leq s < r \leq n} \left(\exp\left(\pi i (x_r - x_s)\right) - \exp\left(-\pi i (x_r - x_s)\right) \right)} \times$$

$$\times \frac{1}{\prod\limits_{1 \leq r \leq n} \left(\exp\left(\pi i x_r\right) - \exp\left(-\pi i x_r\right) \right)} .$$

(B) Let $N = 2n$ be even. Then the irreducible characters of $SO(N)$ can be indexed by n integers $a_1, a_2, \ldots, a_n$ such that for each j, $a_{j+1} \geq a_j + 1$, and $a_1 \geq -a_2 + 1$. The value of the corresponding character on an element of $SO(N)$ of the form

$$\begin{pmatrix} \cos(2\pi x_1) & -\sin(2\pi x_1) & \ldots & 0 & 0 \\ \sin(2\pi x_1) & \cos(2\pi x_1) & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & \cos(2\pi x_n) & -\sin(2\pi x_n) \\ 0 & 0 & \ldots & \sin(2\pi x_n) & \cos(2\pi x_n) \end{pmatrix}$$

is

$$\frac{\sum\limits_{\tau \in S_n} \sum\limits^{*}_{\substack{\epsilon_m = \pm 1 \\ (1 \leq m \leq n)}} \operatorname{sgn}(\tau) \exp\left(2\pi i \sum\limits_{q=1}^{n} \epsilon_q a_{\tau(q)} x_q\right)}{\prod\limits_{1 \leq s < r \leq n} \left(\exp\left(\pi i(x_r + x_s)\right) - \exp\left(-\pi i(x_r + x_s)\right)\right)} \times$$
$$\times \frac{1}{\prod\limits_{1 \leq s < r \leq n} \left(\exp\left(\pi i(x_r - x_s)\right) - \exp\left(-\pi i(x_r - x_s)\right)\right)},$$

where $\sum^{*}$ means we sum only over those choices of the ϵ_m for which $\prod_j \epsilon_j = 1$.

Remarks: In the above expressions, the sums in the numerators are over all τ in the permutation group S_n, and over n different choices of $+1$ or -1 (except that when N is even, we must choose an *even* number of -1). Thus, the sums are over $2^n n!$ (or $2^{n-1}n!$) different terms. These sums correspond to summing over the "Weyl group" of $SO(N)$. The products in the denominators correspond to products over the "positive roots" of $SO(N)$. The values $a_1, \ldots, a_n$ are related to the "highest weights". The interested reader is refered to Adams's book [Ad] for background; see especially his Propositions 6.16 and 6.17 and Lemma 5.58.

The Weyl Character Formula expresses the irreducible characters of $SO(N)$, evaluated on elements of the "maximal torus" of $SO(N)$, as a quotient of two functions of the arguments $x_1, \ldots, x_n$. The functions vanish to high order as the x_j approach 0, which is where we need to evaluate them. However, since the characters themselves are continuous functions on $SO(N)$, we may evaluate the quotient at these singular points by taking limits. The rest of this Section is devoted to evaluating these limits; the results are given in Proposition 3-1 below.

To evaluate the dimensions of the characters, we must evaluate the characters themselves at the identity matrix, i.e. when all the x_j are 0. We do this as follows, say for SO(2n+1). We set $x_j = b^j t$ for each j, where $b > 1$ is some fixed number, and t is a non-zero real number. We then let $t \to 0$. Thus, writing $d_{\vec{a}}$ for the dimension of the character indexed by $\vec{a} = (a_1, \ldots, a_n)$, we have

$$d_{\vec{a}} = \lim_{t\to 0} \frac{\sum_{\tau\in S_n} \operatorname{sgn}(\tau) \sum_{\epsilon_m} \left(\prod_j \epsilon_j\right) \exp\left(2\pi i \sum_{q=1}^n \epsilon_q b^q t a_{\tau(q)}\right)}{(2i)^{n^2} \prod_{r>s} \sin(\pi(b^r - b^s)t)\sin(\pi(b^r + b^s)t) \prod_r \sin(\pi b^r t)} .$$

Since $t \to 0$, we need only evaluate the numerator and denominator to lowest order in t. It is easily computed that the denominator is

$$(2i)^{n^2} t^{n^2} \pi^{n^2} \left(\prod_{r>s} \left(b^{2r} - b^{2s}\right)\right) \left(\prod_r b^r\right) + \text{ higher order terms },$$

and that the coefficient of t^{n^2} in the numerator is

$$\sum_{\tau\in S_n} \operatorname{sgn}(\tau) \sum_{\epsilon_m} \left(\prod_j \epsilon_j\right) \frac{1}{(n^2)!} \left(2\pi i \sum_{q=1}^n \epsilon_q b^q a_{\tau(q)}\right)^{n^2} .$$

Now, in the sum over the ϵ_m in this coefficient, any term with an *odd* power of some ϵ_q will sum to zero. Hence, the only surviving terms in the expansion of

$$\sum_{\epsilon_m} \left(\prod_j \epsilon_j\right) \left(\sum_{q=1}^n \epsilon_q b^q a_{\tau(q)}\right)^{n^2}$$

are those in which each factor $\epsilon_q b^q a_{\tau(q)}$ occurs an odd number of times. Furthermore, any term in which two such factors occur exactly the same number of times cancels out when we sum over τ. This means that the only surviving terms are those in which one such factor occurs exactly once, another exactly three times, ..., and one exactly $2n-1$ times. This argument also shows that the coefficient of t^p in the numerator is 0 for $p < n^2$. Hence, to lowest order in t, the numerator is

$$(2\pi i t)^{n^2} \frac{1}{(n^2)!} \sum_{\tau\in S_n} \operatorname{sgn}(\tau) \sum_{\epsilon_m} \left(\prod_j \epsilon_j\right) \binom{n^2}{1\ 3\ \ldots\ 2n-1} \sum_{\sigma\in S_n} \prod_{q=1}^n \left(\epsilon_q b^q a_{\tau(q)}\right)^{2\sigma(q)-1} ,$$

where we have introduced a second sum over S_n to take into account the different ways in which the powers $1, 3, \ldots, 2n-1$ can be distributed over the factors $\epsilon_q b^q a_{\tau(q)}$, and where $\binom{n^2}{1\ 3\ \ldots\ 2n-1}$ is the appropriate multinomial coefficient. In this last equation, each ϵ_q is raised to an even power, and hence the sum over the ϵ_m is now a sum of 2^n identical terms. Hence, to lowest order in t, the numerator is

$$(2\pi i t)^{n^2} 2^n \frac{1}{(n^2)!} \binom{n^2}{1\ 3\ \ldots\ 2n-1} \sum_{\tau,\sigma \in S_n} \mathrm{sgn}(\tau) \prod_{q=1}^{n} \left(b^q a_{\tau(q)}\right)^{2\sigma(q)-1}$$

$$= \frac{(2\pi i t)^{n^2} 2^n}{1!3!\ldots(2n-1)!} \sum_{\tau,\sigma \in S_n} b^{\left(\sum\limits_{q=1}^{n} q(2\sigma(q)-1)\right)} \mathrm{sgn}(\tau) \prod_{q=1}^{n} \left(a_{\tau(q)}\right)^{2\sigma(q)-1} .$$

Hence,

$$d_{\vec{a}} = \frac{2^n}{1!3!\ldots(2n-1)!} \frac{\sum\limits_{\tau,\sigma \in S_n} b^{\left(\sum\limits_{q=1}^{n} q(2\sigma(q)-1)\right)} \mathrm{sgn}(\tau) \prod\limits_{q=1}^{n} \left(a_{\tau(q)}\right)^{2\sigma(q)-1}}{\prod\limits_{r>s} \left(b^{2r} - b^{2s}\right) \prod\limits_{r} b^r}$$

$$= \frac{2^n}{1!3!\ldots(2n-1)!} \frac{\sum\limits_{\tau,\sigma \in S_n} b^{\left(\sum\limits_{q=1}^{n} q(2\sigma(q)-1)\right)} \mathrm{sgn}(\tau) \prod\limits_{q=1}^{n} \left(a_{\tau(q)}\right)^{2\sigma(q)-1}}{\prod\limits_{r>s} \left(b^{2r} - b^{2s}\right) \prod\limits_{r} b^r} .$$

Now, this is true for any $b > 1$. Hence, we can let $b \to \infty$. To evaluate this limit, we need only the highest power of b in both the numerator and denominator. In the numerator, this power arises precisely when σ is the identity permutation, and one then obtains the coefficient of $b^{\left(\sum_{q=1}^{n} q(2q-1)\right)}$. The denominator has the same highest power of b, with unit coefficient. We conclude that

$$d_{\vec{a}} = \frac{2^n}{1!3!\ldots(2n-1)!} \sum_{\tau \in S_n} \mathrm{sgn}(\tau) \prod_{q=1}^{n} \left(a_{\tau(q)}\right)^{2q-1} .$$

We claim that this is the same as writing

$$d_{\vec{a}} = \frac{2^n}{1!3!\ldots(2n-1)!} \left(\prod_{q=1}^{n} a_q\right) \prod_{r>s} \left(a_r^2 - a_s^2\right) .$$

Indeed, all terms occuring in the first expression occur in the second. There are some extra terms in the second expression, corresponding to two or more of the a_q being raised to the same power, but these terms precisely cancel out.

The computation of the dimensions of the representations for *even* N is very similar. The main point of departure is that while above, each term $\left(\epsilon_q b^q a_{\tau(q)}\right)$ had to be raised to an *odd* power in a surviving term of the sum over the ϵ_m, in the even N case we only require that either each ϵ_q be raised to an odd power, *or* that each ϵ_q be raised to an even power. The lowest-order term in t in the numerator arises when one term $\left(\epsilon_q b^q a_{\tau(q)}\right)$ is raised to the power 0, another to the power 2, ..., and another to the power $2n-2$. After letting $t \to 0$, we let $b \to \infty$ as above, and finally conclude that

$$d_{\vec{a}} \;=\; \frac{2^{n-1}}{0!2!4!\dots(2n{-}2)!}\prod_{r>s}\left(a_r^2 - a_s^2\right) \;.$$

In addition to the dimensions, we also require the values of the characters at the conjugacy class of the matrix R_θ defined above. This corresponds to letting each x_q approach 0 with the exception of x_1, which we will eventually set to $\theta/2\pi$. The computation is somewhat similar to the above dimensions computations, but there are some differences. We proceed as follows, say for $SO(2n+1)$. We set $x_j = b^j t$, for $2 \le j \le n$, where again $b > 1$ and $t \to 0$. In the Weyl Character Formula, to lowest order in t the denominator is

$$(-1)^{n-1}\left(2i\sin(\pi x_1)\right)^{2n-1}(2\pi i t)^{(n-1)^2}\prod_{r>s>1}\left(b^{2r}-b^{2s}\right)\prod_{r>1} b^r \;.$$

The coefficient of $t^{(n-1)^2}$ in the numerator is

$$\sum_\tau \mathrm{sgn}(\tau)(2i)\sin(2\pi a_{\tau(1)}x_1)\sum_{\substack{\epsilon_m\\ m=2,3,\dots,n}}\left(\prod_{j=2}^n \epsilon_j\right)\frac{1}{(n-1)^2!}\left(2\pi i\sum_{q=2}^n \epsilon_q b^q a_{\tau(q)}\right)^{(n-1)^2}$$

$$= \frac{1}{(n-1)^2!}\sum_\tau \mathrm{sgn}(\tau)(2i)\sin(2\pi a_{\tau(1)}x_1)\sum_{\epsilon_m}(2\pi i)^{(n-1)^2}\sum_{\sigma\in S_{n-1}}\prod_{q=2}^n\left(b^q a_{\tau(q)}\right)^{2\sigma(q-1)-1}$$

$$= \frac{2^n i(2\pi i)^{(n-1)^2}}{1!3!\dots(2n-3)!}\sum_\tau \mathrm{sgn}(\tau)\sin(2\pi a_{\tau(1)}x_1)\sum_{\sigma\in S_{n-1}}\prod_{q=2}^n\left(b^q a_{\tau(q)}\right)^{2\sigma(q-1)-1} \;,$$

where we have again considered only terms in which each term $\left(\epsilon_q b^q a_{\tau(q)}\right)$ is raised to a distinct odd power. We conclude, writing $c_{\vec{a}}$ for the value of the character indexed by $\vec{a}$ evaluated at the matrix R_θ, that

$$c_{\vec{a}} \;=\; \frac{1}{2^{n-1}1!3!\ldots(2n-3)!(\sin(\pi x_1))^{2n-1}} \frac{\sum_\tau \operatorname{sgn}(\tau)\sin(2\pi a_{\tau(1)}) \sum_{\sigma\in S_{n-1}} \prod_{q=2}^{n}\left(b^q a_{\tau(q)}\right)^{2\sigma(q-1)-1}}{\prod_{r>s>1}(b^2 r - b^2 s)\prod_{r>1} b^r} \;.$$

As before, we let $b \to \infty$ and consider only the highest power of b in both numerator (i.e. when σ is the identity permutation) and denominator, and conclude that

$$c_{\vec{a}} \;=\; \frac{1}{2^{n-1}1!3!\ldots(2n-3)!(\sin(\pi x_1))^{2n-1}} \sum_\tau \operatorname{sgn}(\tau)\sin(2\pi a_{\tau(1)}x_1)\prod_{q=2}^{n}\left(a_{\tau(q)}\right)^{2q-3} \;.$$

We now claim that this is the same as

$$c_{\vec{a}} \;=\; \frac{1}{2^{n-1}1!3!\ldots(2n-3)!(\sin(\pi x_1))^{2n-1}} \sum_{j=1}^{n}(-1)^{j-1}\sin(2\pi a_j x_1)\left(\prod_{r\neq j} a_r\right)\prod_{\substack{s<r\\ s,r\neq j}}(a_r^2 - a_s^2) \;.$$

We leave the verification to the reader.

The computation of $c_{\vec{a}}$ for $N = 2n$ is similar, with the points of departure from the above being similar to those for $d_{\vec{a}}$. One difference is that while for $N = 2n+1$, we could factor out in the numerator an expression of the form

$$\exp(2\pi i a_{\tau(1)}x_1) - \exp(-2\pi i a_{\tau(1)}x_1) \;=\; 2i\sin(2\pi a_{\tau(1)}x_1) \;,$$

for $N = 2n$ we want to factor out an expression of the form

$$\exp(2\pi i a_{\tau(1)}x_1) + \exp(-2\pi i a_{\tau(1)}x_1) \;=\; 2\cos(2\pi a_{\tau(1)}x_1) \;.$$

This last factorization cannot be done directly because the condition $\prod_j \epsilon_j = 1$ is affected, but this does not affect the result provided $n > 1$. We leave the details to the reader.

We summarize the results here (writing θ for $2\pi x_1$).

Proposition 3-1. *Let $d_{\vec{a}}$ and $c_{\vec{a}}$ be, respectively, the dimension and the value at the matrix R_θ defined above, of the irreducible representation of $SO(N)$ corresponding to the index $\vec{a} = (a_1,\ldots,a_n)$. Then for $N = 2n+1$,*

$$d_{\vec{a}} \;=\; \frac{2^n}{1!3!\ldots(2n-1)!}\left(\prod_{q=1}^{n} a_q\right)\prod_{r>s}(a_r^2 - a_s^2)$$

and

$$c_{\vec{a}} \;=\; \frac{1}{2^{n-1}1!3!\ldots(2n-3)!(\sin(\theta/2))^{2n-1}}\sum_{j=1}^{n}(-1)^{j-1}\sin(a_j\theta)\left(\prod_{r\neq j}a_r\right)\prod_{\substack{s<r\\ s,r\neq j}}\left(a_r^2-a_s^2\right)\,.$$

For $N = 2n$*,*

$$d_{\vec{a}} \;=\; \frac{2^{n-1}}{0!2!4!\ldots(2n-2)!}\prod_{r>s}\left(a_r^2-a_s^2\right)$$

and (for $n > 1$*)*

$$c_{\vec{a}} \;=\; \frac{1}{2^{n-1}0!2!4!\ldots(2n-4)!\,(\sin(\theta/2))^{2n-2}}\sum_{j=1}^{n}(-1)^{j-1}\cos(a_j\theta)\prod_{\substack{s<r\\ s,r\neq j}}\left(a_r^2-a_s^2\right)\,.$$

To make these characters more concrete, we mention the following. For $N = 2n+1$, the smallest possible values of $\vec{a}$ are $a_1 = \frac{1}{2}, a_2 = \frac{3}{2}, \ldots, a_n = n - \frac{1}{2}$. This corresponds to the *trivial* representation of $SO(N)$, and so both $d_{\vec{a}}$ and $c_{\vec{a}}$ are 1. The next lowest values of $\vec{a}$ are when we change a_n to $n+\frac{1}{2}$ and leave the other a_j the same. This corresponds to the *natural* representation of $SO(N)$ (i.e. the representation in which each element of $SO(N)$ is sent to itself as a matrix), and so $d_{\vec{a}} = N$ and $c_{\vec{a}} = (N-2) + 2\cos(\theta)$. Similarly, for $N = 2n$, the trivial representation corresponds to setting $a_1 = 0, a_2 = 1, \ldots, a_n = n-1$, and the natural representation corresponds to changing a_n to n and leaving the other a_j the same.

4. Fourier Analysis and the Upper Bound Lemma.

Let G be a compact Lie group with irreducible representations $\rho_1, \rho_2, \ldots$, and corresponding characters $\chi_1, \chi_2, \ldots$. Recall that these characters are orthonormal with respect to the inner product

$$\langle \chi_i, \chi_j \rangle = \int_G \chi_i \overline{\chi_j} \, d\lambda \ ,$$

where λ is normalized Haar measure. Given a finite measure μ on G, the *Fourier transform* of μ is defined by

$$\widehat{\mu}(\rho_j) = \int_G \rho_j \, d\mu \ ,$$

and the *Fourier coefficients* by

$$\widehat{\mu}(\chi_j) = \int_G \chi_j \, d\mu = \operatorname{trace} \widehat{\mu}(\rho_j) \ .$$

(This terminology may not be completely standard.) It is easily verified that

$$\widehat{\mu * \nu}(\rho_j) = \widehat{\mu}(\rho_j) \widehat{\nu}(\rho_j) \ ,$$

i.e. that Fourier transforms change convolutions into ordinary matrix products.

Now, if μ is conjugate-invariant (in the sense that $\mu(g^{-1}Ag) = \mu(A)$ for all $g \in G$ and all measurable $A \subseteq G$), then it is easily verified that $\widehat{\mu}(\rho_j)$ commutes with $\rho_j(g)$ for all $g \in G$, so that Schur's Lemma implies $\widehat{\mu}(\rho_j)$ is a scalar multiple of the identity matrix. In this case, the scalar is easily seen (by taking traces) to be $\frac{\operatorname{trace}(\widehat{\mu}(\rho_j))}{\dim(\rho_j)}$, so

$$\widehat{\mu}(\rho_j) = \left(\frac{\operatorname{trace}(\widehat{\mu}(\rho_j))}{\dim(\rho_j)} \right) I \ ,$$

where I is the appropriately sized identity matrix. Hence,

$$\widehat{\mu^{*k}}(\rho_j) = \left(\frac{\operatorname{trace}(\widehat{\mu}(\rho_j))}{\dim(\rho_j)} \right)^k I \ .$$

In particular, for the measures Q_k defined on $SO(N)$ in Section 2, with $\rho_{\vec{a}}$ the representation corresponding to the index $\vec{a}$, with $c_{\vec{a}}$ and $d_{\vec{a}}$ as in Proposition 3-1, and using equation (2-1),

$$\widehat{Q_k}(\rho_{\vec{a}}) = \widehat{Q_1^{*k}}(\rho_{\vec{a}}) = \left(\frac{\operatorname{trace}(\widehat{Q_1}(\rho_{\vec{a}}))}{\dim(\rho_{\vec{a}})} \right)^k I = \left(\frac{\operatorname{trace}(\rho_{\vec{a}}(R_\theta))}{\dim(\rho_{\vec{a}})} \right)^k I = \left(\frac{c_{\vec{a}}}{d_{\vec{a}}} \right)^k I \ ,$$

and so

$$\widehat{Q_k}(\chi_{\vec{a}}) = \left(\frac{c_{\vec{a}}}{d_{\vec{a}}}\right)^k d_{\vec{a}} \ . \tag{4-1}$$

We require the following standard result from harmonic analysis.

Lemma 4-1 (Fourier Inversion Theorem). *Let G be any compact Lie group, with μ any conjugate-invariant Borel measure on G. Let $\ell_j = \widehat{\mu}(\chi_j)$, and suppose $\sum_j |\ell_j|^2$ is finite. Then μ is absolutely continuous with respect to Haar measure λ, and in fact $d\mu = f\, d\lambda$ where f is the L^2 function on G defined by*

$$f(g) = \sum_j \ell_j \, \overline{\chi_j}(g) \ .$$

Proof. Let ν be the measure defined by $d\nu = f\, d\lambda$. Then ν and μ are two conjugate-invariant measures which are easily seen to have the same Fourier coefficients. The Fourier Uniqueness Theorem (cf. [HR], Theorem 31.5) then implies that $\mu = \nu$. ∎

As an illustration of a use of this Lemma, we have

Remark 4-2. *Let $g \in G$, with G an (infinite) compact Lie group. Then $\sum_j |\chi_j(g)|^2$ is infinite, and so $\sum_j |\chi_j(g)|$ is infinite.*

Proof. Let μ be the unique conjugate-invariant Borel probability measure on G which is concentrated on the conjugacy class of g. Then $\widehat{\mu}(\chi_j) = \chi_j(g)$. If $\sum_j |\chi_j(g)|^2$ were finite, the measure μ would be absolutely continuous with respect to Haar measure, which is clearly false since the conjugacy class of g has μ-measure 1 but Haar-measure 0. ∎

The following Lemma was developed and used extensively by Diaconis and Shashahani (see [DS1], [DS2], [D]). We include a statement and proof here for the case of a conjugate-invariant measure on a compact Lie group.

Lemma 4-3 (Upper Bound Lemma). *Let G be a compact Lie group, with normalized Haar measure λ, and let μ be any conjugate- invariant probability measure on G. Let $\ell_j = \widehat{\mu}(\chi_j) = \int\limits_G \chi_j \, d\mu$. Then*

$$\|\mu - \lambda\|^2 \le \frac{1}{4}\left(\sum_j |\ell_j|^2 \; - 1\right) ,$$

where $\|\cdot\|$ denotes the variation distance defined in Section 2.

Proof. The statement is vacuous if $\sum\limits_j |\ell_j|^2$ is infinite, so we assume it is finite. Then, with f as in Lemma 4-1,

$$\begin{aligned} 4\|\mu - \lambda\|^2 &= \left(\int\limits_G |f - 1| \, d\lambda\right)^2 \\ &\le \int\limits_G |f - 1|^2 \, d\lambda \\ &= \int\limits_G |f|^2 d\lambda \; - 1 \\ &= \sum_j |\ell_j|^2 \; - 1 \end{aligned}$$

where the inequality is Cauchy-Schwarz, and the final equality uses the fact that the irreducible characters form an orthonormal basis for the L^2 class functions on G. ∎

Remark. If χ_1 is the character corresponding to the trivial representation, then $\ell_1 = 1$, so the Upper Bound Lemma can be written $\|\mu - \lambda\|^2 \le \frac{1}{4} \sum\limits_{j \ne 1} |\ell_j|^2$.

We shall apply the Upper Bound Lemma to the problem at hand in Section 6.

5. Proof of Theorem 2-1.

We fix $N \geq 5$ (smaller N can be easily handled by modifying the constant Δ as necessary). We let χ_1 be the irreducible character of $SO(N)$ corresponding to the natural representation (see the end of Section 2). We shall compute the expectation and variance of χ_1 under the measures λ and Q_k (with $k = \frac{1}{2(1-\cos\theta)}(N \log N - cN)$). We have that

$$E_{\lambda}(\chi_1) = \int\limits_{SO(N)} \chi_1 \, d\lambda = \langle \chi_1, 1 \rangle = 0$$

and

$$Var_{\lambda}(\chi_1) = \int\limits_{SO(N)} |\chi_1|^2 \, d\lambda = \langle \chi_1, \chi_1 \rangle = 1 \ ,$$

using orthonormality. Also, from equation (4-1),

$$\begin{aligned} E_{Q_k}(\chi_1) &= \left(\frac{\widehat{Q_1}(\chi_1)}{N} \right)^k N \\ &= \left(\frac{N - 2(1-\cos\theta)}{N} \right)^k N \\ &= N \left(1 - \frac{t}{N} \right)^{\frac{1}{t} N \log N} \left(1 - \frac{t}{N} \right)^{-\frac{1}{t} cN} \ , \end{aligned}$$

where we have written t for $2(1-\cos\theta)$. Now, it is easily checked that

$$\left(1 - \frac{t}{N} \right)^{-\frac{1}{t} cN} > e^c \ .$$

Also, using the fact that

$$\begin{aligned} \log\left(1 - \frac{t}{N}\right) &= -\frac{t}{N} - \frac{t^2}{2N^2} - \frac{t^3}{3N^3} - \dots \\ &\geq -\frac{t}{N} - \frac{t^2}{2N^2} - \frac{t^3}{2N^3} - \dots \\ &= -\frac{t}{N} - \frac{t^2}{2N^2} \left(\frac{1}{1 - \frac{t}{N}} \right) \end{aligned}$$

for $-1 < \frac{t}{N} < 1$, it is not hard to show that

$$N \left(1 - \frac{t}{N} \right)^{\frac{1}{t} N \log N} \geq \frac{1}{5} \ ,$$

for $N \geq 5$ and $-4 \leq t \leq 4$. We conclude that

$$E_{Q_k}(\chi_1) > \frac{1}{5} e^{-c} \ .$$

It remains only to compute the variance of χ_1 under the measure Q_k. We have the following.

Lemma 5-1.

$$Var_{Q_k}(\chi_1) \leq 1 + \frac{16}{1-\cos\theta} \exp\left(\frac{2}{1-\cos\theta}\right)\ e^{2c}\left(\frac{\log N}{N}\right) \ .$$

Assuming the Lemma, we complete the proof as follows. We have that

$$E_{Q_k}(\chi_1) \geq \frac{1}{5} e^{c} \ .$$

Hence, from Chebychev's Inequality,

$$\mathrm{Prob}_{Q_k}\left(\chi_1 \leq \frac{1}{10} e^{c}\right) \leq 100 e^{-2c} + \frac{1600}{1-\cos\theta} \exp\left(\frac{2}{1-\cos\theta}\right) \frac{\log N}{N} \ .$$

Also

$$\mathrm{Prob}_{\lambda}\left(\chi_1 \geq \frac{1}{10} e^{c}\right) \leq 100 e^{-2c} \ .$$

Hence,

$$\begin{aligned}
\|Q_k - \lambda\| &\geq \mathrm{Prob}_{\lambda}\left(\chi_1 < \frac{1}{10} e^{c}\right) - \mathrm{Prob}_{Q_k}\left(\chi_1 < \frac{1}{10} e^{c}\right) \\
&= 1 - \mathrm{Prob}_{\lambda}\left(\chi_1 \geq 1 - \frac{1}{10} e^{c}\right) - \mathrm{Prob}_{Q_k}\left(\chi_1 < \frac{1}{10} e^{c}\right) \\
&\geq 1 - 200 e^{-2c} - \frac{1600}{1-\cos\theta} \exp\left(\frac{2}{1-\cos\theta}\right) \frac{\log N}{N}
\end{aligned}$$

as required.

Proof of Lemma 5-1. Let ρ_1 be the natural representation of $SO(N)$ (so $\chi_1 = \operatorname{trace}\rho_1$). Then $\chi_1^2 = \operatorname{trace}\rho_1 \otimes \rho_1$. Now, $\rho_1 \otimes \rho_1$ acts on $\mathbf{R}^N \otimes \mathbf{R}^N$. Under the natural identification of this space with the space $M_N(\mathbf{R})$ of real, $N \times N$ matrices, the action of $\rho_1 \otimes \rho_1$ becomes the following. For $g \in SO(N)$ and $B \in M_N(\mathbf{R})$,

$$(\rho_1 \otimes \rho_1)(g) B = g^t B g \ ,$$

where g^t denotes the transpose of g as a matrix. It immediately follows that $M_N(\mathbf{R})$ decomposes into the direct sum of three subspaces invariant under $\rho_1 \otimes \rho_1$: the one-dimensional subspace M_N^0 consisting of scalar multiples of the identity matrix, the $\frac{N(N-1)}{2}$-dimensional subspace M_N^- consisting of the skew-symmetric matrices, and the $\left(\frac{N(N+1)}{2} - 1\right)$ -dimensional subspace M_N^+ consisting of the symmetric matrices with trace 0. We claim that each of these subspaces is irreducible under $\rho_1 \otimes \rho_1$. Indeed, it is obvious that the action of $\rho_1 \otimes \rho_1$ on M_N^0 is the trivial representation. It is a well-known fact that M_N^- is invariant, and in fact the action of $\rho_1 \otimes \rho_1$ on M_N^- is isomorphic to the representation corresponding to the index $\vec{a}_-$ given by $a_1 = \frac{1}{2}, a_2 = \frac{3}{2}, \ldots, a_{n-2} = n - \frac{5}{2}, a_{n-1} = n - \frac{1}{2}, a_n = n + \frac{1}{2}$ for $N = 2n+1$, and given by $a_1 = 0, a_2 = 1, \ldots, a_{n-2} = n-3, a_{n-1} = n-1, a_n = n$ for $N = 2n$. We further claim that the action of $\rho_1 \otimes \rho_1$ on M_N^+ is isomorphic to the representation corresponding to the index $\vec{a}_+$ given by $a_1 = \frac{1}{2}, a_2 = \frac{3}{2}, \ldots, a_{n-2} = n - \frac{5}{2}, a_{n-1} = n - \frac{3}{2}, a_n = n + \frac{3}{2}$ for $N = 2n+1$, and given by $a_1 = 0, a_2 = 1, \ldots, a_{n-2} = n-3, a_{n-1} = n-2, a_n = n+1$ for $N = 2n$. Indeed, the dimensions are equal, so it is enough to show that M_N^+ contains the approprate "highest weight". Specifically, we must show that there exists a matrix B_* in the complexification of M_N^+, such that if $g \in SO(N)$ is of the form

$$g = \begin{pmatrix} & & & 0 & 0 \\ & * & & \vdots & \vdots \\ & & & 0 & 0 \\ 0 & \ldots & 0 & \cos\alpha & -\sin\alpha \\ 0 & \ldots & 0 & \sin\alpha & \cos\alpha \end{pmatrix}$$

where α is any angle, and $*$ indicates an arbitrary $(N-2) \times (N-2)$ matrix, then

$$g^t B_* g = e^{i(2\alpha)} B_* \ .$$

The matrix

$$B_* = \begin{pmatrix} 0 & \ldots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & 0 \\ 0 & \ldots & 0 & 1 & i \\ 0 & \ldots & 0 & i & -1 \end{pmatrix}$$

is easily seen to satisfy this requirement.

We conclude that

$$\rho_1 \otimes \rho_1 \simeq (\text{trivial rep.}) \oplus \rho_{\vec{a}_-} \oplus \rho_{\vec{a}_+} \ ,$$

and so

$$E_{Q_k}(\chi_1^2) = E_{Q_k}(1 + \chi_{\vec{a}_-} + \chi_{\vec{a}_+})$$
$$= 1 + \left(\frac{c_{\vec{a}_-}}{d_{\vec{a}_-}}\right)^k d_{\vec{a}_-} + \left(\frac{c_{\vec{a}_-}}{d_{\vec{a}_-}}\right)^k d_{\vec{a}_-} \quad .$$

Thus, writing $\vec{a}_1$ for the index corresponding to χ_1,

$$Var_{Q_k}(\chi_1) = E_{Q_k}(\chi_1^2) - (E_{Q_k}(\chi_1))^2$$
$$= 1 + \left(\frac{c_{\vec{a}_-}}{d_{\vec{a}_-}}\right)^k d_{\vec{a}_-} + \left(\frac{c_{\vec{a}_-}}{d_{\vec{a}_-}}\right)^k d_{\vec{a}_-} - \left(\left(\frac{c_{\vec{a}_1}}{d_{\vec{a}_1}}\right)^k d_{\vec{a}_1}\right)^2$$

(5-1)
$$= \left(1 - \left(\frac{c_{\vec{a}_1}}{d_{\vec{a}_1}}\right)^{2k}\right) + \left(\left(\frac{c_{\vec{a}_-}}{d_{\vec{a}_-}}\right)^k - \left(\frac{c_{\vec{a}_1}}{d_{\vec{a}_1}}\right)^{2k}\right) d_{\vec{a}_-} +$$
$$\left(\left(\frac{c_{\vec{a}_+}}{d_{\vec{a}_+}}\right)^k - \left(\frac{c_{\vec{a}_1}}{d_{\vec{a}_1}}\right)^{2k}\right) d_{\vec{a}_+}$$

(using the fact that $(d_{\vec{a}_1})^2 = 1 + d_{\vec{a}_-} + d_{\vec{a}_+}$). Now, we have that $d_{\vec{a}_1} = N$ and $c_{\vec{a}_1} = N - 2 + 2\cos\theta$. Using the basis

$$\{e_i \otimes e_j - e_j \otimes e_i | i \neq j\}$$

for M_N^-, one sees that $d_{\vec{a}_-} = \frac{N(N-1)}{2}$ and $c_{\vec{a}_-} = \frac{(N-2)(N-3)}{2} + 2(N-2)\cos\theta + 1$. Similarly, using the basis

$$\{e_i \otimes e_j + e_j \otimes e_i | i \neq j\} \cup \{e_j \otimes e_j + e_n \otimes e_n | 1 \leq j < n\}$$

for M_N^+, one sees that $d_{\vec{a}_+} = \frac{N(N+1)}{2} - 1$ and

$$c_{\vec{a}_+} = \left(\frac{(N-2)(N-3)}{2} + 2(N-2)\cos\theta + \cos 2\theta\right) + \left((N-3) + 2\cos^2\theta\right) \quad .$$

From these facts, one easily obtains for $N \geq 5$ the bounds

$$\left|\frac{c_{\vec{a}_-}}{d_{\vec{a}_-}} - \left(\frac{c_{\vec{a}_1}}{d_{\vec{a}_1}}\right)^2\right| \leq \frac{30}{N^2}$$

$$\left|\frac{c_{\vec{a}_+}}{d_{\vec{a}_+}} - \left(\frac{c_{\vec{a}_1}}{d_{\vec{a}_1}}\right)^2\right| \leq \frac{26}{N^2}$$

$$\left|\frac{c_{\vec{a}_-}}{d_{\vec{a}_-}}\right| \leq 1 - \frac{4(1-\cos\theta) - \frac{4}{N}}{N-1}$$

$$\text{and } \left|\frac{c_{\vec{a}_+}}{d_{\vec{a}_+}}\right| \le 1 - \frac{4(1-\cos\theta) - \frac{6}{N}}{N-1} \ .$$

If we then use the general fact

$$\left|x^k - y^k\right| = \left|(x-y)(x^{k-1} + x^{k-2}y + \ldots + y^{k-1})\right| \ \le \ |(x-y)| \ k \ (\max(|x|,|y|))^k$$

(with $y = \left(\frac{c_{\vec{a}_1}}{d_{\vec{a}_1}}\right)^2$) and recall that $k = \frac{1}{2(1-\cos\theta)}(N\log N - cN)$, we obtain that

$$\left|\left(\frac{c_{\vec{a}_-}}{d_{\vec{a}_-}}\right)^k - \left(\frac{c_{\vec{a}_1}}{d_{\vec{a}_1}}\right)^{2k}\right| d_{\vec{a}_-} \le \frac{8}{1-\cos\theta}\exp\left(\frac{2}{1-\cos\theta}\right) \ e^{2c} \left(\frac{\log N}{N}\right)$$

and

$$\left|\left(\frac{c_{\vec{a}_+}}{d_{\vec{a}_+}}\right)^k - \left(\frac{c_{\vec{a}_1}}{d_{\vec{a}_1}}\right)^{2k}\right| d_{\vec{a}_+} \le \frac{8}{1-\cos\theta}\exp\left(\frac{2}{1-\cos\theta}\right) \ e^{2c} \left(\frac{\log N}{N}\right) .$$

Also

$$\left(1 - \left(\frac{c_{\vec{a}_1}}{d_{\vec{a}_1}}\right)^{2k}\right) \le 1 \ .$$

Hence, from equation (5-1),

$$Var_{Q_k}(\chi_1) \le 1 + \frac{16}{1-\cos\theta}\exp\left(\frac{2}{1-\cos\theta}\right) \ e^{2c} \ \left(\frac{\log N}{N}\right) \ ,$$

as required. ∎

Remarks.

1. We have been a bit sloppy with the constants above. The key point, however, is that for large N, the variance of χ_1 under the measure Q_k is not much larger than 1.
2. It is a fact that in many of the examples studied, including those on finite groups, the variance as above is small; see Diaconis [D]. It is an open problem to explain why this is so.

6. Proof of Theorem 2-2.

We have from the Upper Bound Lemma (Lemma 4-3) and equation (4-1) that

$$4\|Q_k - \lambda\|^2 \leq \sum_{\vec{a}} \left(\frac{c_{\vec{a}}}{d_{\vec{a}}}\right)^{2k} d_{\vec{a}}^2 \quad - 1 \ ,$$

where the sum is over all values of $\vec{a} = (a_1, \ldots, a_n)$ allowed by the Weyl Character Formula. Hence, we need only show that

$$\sum_{\vec{a}} \left(\frac{c_{\vec{a}}}{d_{\vec{a}}}\right)^{2k} d_{\vec{a}}^2 \quad - 1 \ \leq 4\left(\Lambda e^{-\gamma c}\right)^2 \ .$$

We restrict ourselves to the odd N case; the even N case is very similar. Thus $N = 2n+1$, and $k \geq \frac{1}{2} n \log n + 2cn$.

From Proposition 3-1,

$$d_{\vec{a}} \ = \ \frac{2^n}{1!3!\ldots(2n{-}1)!} \left(\prod_{q=1}^{n} a_q\right) \prod_{r>s} \left(a_r^2 - a_s^2\right)$$

and

$$\frac{c_{\vec{a}}}{d_{\vec{a}}} = \frac{(2n-1)!}{(2\sin(\theta/2))^{2n-1}} \sum_{j=1}^{n} \frac{(-1)^{j-1}\sin(a_j\theta)}{a_j \prod_{r=j+1}^{n} \left(a_r^2 - a_j^2\right) \prod_{s=1}^{j-1} \left(a_j^2 - a_s^2\right)} \ .$$

Now, when θ is 180 degrees, $\sin(\theta/2) = 1$, and $\sin(a_j\theta) = (-1)^{a_j - \frac{1}{2}}$ (for a_j a half-integer). Hence

$$\frac{c_{\vec{a}}}{d_{\vec{a}}} = \frac{(2n-1)!}{(2)^{2n-1}} \sum_{j=1}^{n} \frac{(-1)^{a_j - j + \frac{1}{2}}}{a_j \prod_{r=j+1}^{n} \left(a_r^2 - a_j^2\right) \prod_{s=1}^{j-1} \left(a_j^2 - a_s^2\right)} \ .$$

Thus, for the *trivial* representation, we have

$$1 = \frac{c_{\vec{a}}}{d_{\vec{a}}} = \frac{(2n-1)!}{2^{2n-1}} \sum_{j=1}^{n} \left(\frac{1}{a_j \prod_{r=j+1}^{n} \left(a_r^2 - a_j^2\right) \prod_{s=1}^{j-1} \left(a_j^2 - a_s^2\right)}\right) \ .$$

Let

$$r_{\vec{a}} = \frac{(2n-1)!}{2^{2n-1}} \sum_{j=1}^{n} \left(\frac{1}{a_j \prod_{r=j+1}^{n} \left(a_r^2 - a_j^2\right) \prod_{s=1}^{j-1} \left(a_j^2 - a_s^2\right)}\right) \ .$$

Then $|\frac{c_{\vec{a}}}{d_{\vec{a}}}| \leq r_{\vec{a}}$ for all $\vec{a}$ by the triangle inequality, with $r_{\vec{a}} = 1$ for the trivial representation. We shall use $r_{\vec{a}}$ as an upper bound for $|\frac{c_{\vec{a}}}{d_{\vec{a}}}|$, and we shall approximate it by considering by what fraction it changes as the index $\vec{a}$ changes from the index for the trivial representation to other values.

We begin by noting that it is sufficient to prove the Theorem for sufficiently large values of N and c; the constant Γ can be adjusted to make the Theorem vacuous for smaller values of N and c.

The key computational fact we shall require is the following.

Proposition 6-1. *There is a constant K such that for sufficiently large n, and for any allowable index $\vec{a} = (a_1, \ldots, a_n)$ with $a_n < 8n$,*

$$d_{\vec{a}} r_{\vec{a}}^k \leq \left(K e^{-c/8}\right)^{b_1+b_2+\ldots+b_n} ,$$

where $b_j = a_j - (j - \frac{1}{2})$ is the amount that a_j differs from its value for the trivial representation. (In fact, we can take $K = 252$.)

Assuming Proposition 6-1, and writing q for $\left(Ke^{-c/8}\right)^2$, we have that

$$\begin{aligned}
\sum_{\substack{\vec{a} \\ (a_n < 8n)}} d_{\vec{a}}^2 r_{\vec{a}}^{2k} &\leq \sum_{b_1=0}^{8n} \sum_{b_2=0}^{b_1} \cdots \sum_{b_n=0}^{b_{n-1}} q^{b_1+b_2+\ldots+b_n} \\
&\leq \sum_{b_1=0}^{\infty} \sum_{b_2=0}^{b_1} \cdots \sum_{b_n=0}^{b_{n-1}} q^{b_1+b_2+\ldots+b_n} \\
&= \sum_{g_1=0}^{\infty} \sum_{g_2=0}^{\infty} \cdots \sum_{g_n=0}^{\infty} q^{ng_1+(n-1)g_2+\ldots+g_n}
\end{aligned}$$

where we have set $g_n = b_n$, and $g_j = b_j - b_{j+1}$ for $1 \leq j \leq n-1$. Hence,

$$\begin{aligned}
\sum_{\substack{\vec{a} \\ (a_n < 8n)}} d_{\vec{a}}^2 r_{\vec{a}}^{2k} &\leq \prod_{j=1}^{n} \left(\frac{1}{1-q^j}\right) \\
&\leq \frac{1}{\prod_{j=1}^{\infty}(1-q^j)} \\
&= \frac{1}{1-q-q^2+q^5+\ldots} \\
&= 1+q+2q^2+\ldots \\
&\leq 1+2q
\end{aligned}$$

for q sufficiently small, i.e. for c larger than some universal constant c_0. (The key point here is that $\prod_{j=1}^{\infty}(1-q^j)$ is analytic in q for $|q|<1$, so the power series makes sense.) Hence

$$\sum_{\substack{\vec{a} \\ (a_n<8n)}} d_{\vec{a}}^2 r_{\vec{a}}^{2k} \; - 1 \;\leq\; 2q \;=\; 2K^2 e^{-2c/8} \; .$$

This is almost exactly what we need; once Proposition 6-1 is proved, we need only worry about those $\vec{a}$ with $a_n > 8n$. That is done at the end of this Section.

To prove Proposition 6-1, we make use of the following lemmas.

Lemma 6-2. *Let $d(T)$ stand for the dimension of the character corresponding to the index in which $a_j = j - \frac{1}{2}$ for $1 \leq j \leq n-1$, and in which $a_n = T$. Then*

$$\frac{d(T+1)}{d(T)} \;=\; \left(1+\frac{1}{T}\right)\left(1+\frac{2n-2}{T-n+\frac{3}{2}}\right) \; .$$

Proof. This is a straightforward computation. If we change a_n from T to $T+1$, the only factors of $d_{\vec{a}}$ which are affected are those of the form

$$a_n \prod_{s=1}^{n-1}\left(a_n^2 - a_s^2\right) = a_n \prod_{s=1}^{n-1}(a_n - a_s) \prod_{s=1}^{n-1}(a_n + a_s) \; .$$

Hence

$$\frac{d(T+1)}{d(T)} = \frac{T+1}{T} \prod_{s=1}^{n-1}\left(\frac{T+1-a_s}{T-a_s}\right) \prod_{s=1}^{n-1}\left(\frac{T+1+a_s}{T+a_s}\right) \; .$$

Each of these two products is a "collapsing product" which can easily be evaluated. ∎

Lemma 6-3. *Choose any ℓ with $0 \leq \ell \leq n-1$. Let $\vec{a} = (a_1, \ldots, a_n)$ be any allowable index with $a_j = T + n - j$, for $\ell + 1 \leq j \leq n$. If we increase each of $a_{\ell+1}, a_{\ell+2}, \ldots, a_n$ by 1, then $d_{\vec{a}}$ is multiplied by an amount which is less than or equal to the binomial coefficient $\binom{2n+1}{n-\ell}$.*

Proof. A computation similar to that for Lemma 6-2 above shows that $d_{\vec{a}}$ is multiplied by exactly $\binom{2n+1}{n-\ell}$ in the case when $T = n - \frac{1}{2}$ (i.e. if we started at the trivial representation). Now, the only factors in $d_{\vec{a}}$ which are affected are those of the form

$$a_{\ell+1} \ldots a_n \prod_{r>s>\ell}(a_r + a_s) \prod_{r>\ell\geq s}(a_r + a_s) \prod_{r>\ell\geq s}(a_r - a_s) \; .$$

Now, if T is larger than $n - \frac{1}{2}$, and if the values of $a_1, \ldots, a_\ell$ get more "spread out", then it is easily verified that each factor of $d_{\vec{a}}$ is multiplied by *less* than it was for the trivial representation. This gives the inequality. ∎

Corollary 6-4. *Under the hypothesis of Lemma 6-3 above, with $n \geq 2$, $d_{\vec{a}}$ is multiplied by an amount which is less than or equal to $8e^{2n}$.*

Proof. This follows immediately from Lemma 6-3 and a weak form of Sterling's Approximation. Indeed, for any integer p,

$$e\left(\frac{p}{e}\right)^p < p! < pe\left(\frac{p}{e}\right)^p$$

(cf. [Ar], p. 20). Hence,

$$\begin{aligned}\binom{2n+1}{n-\ell} \leq \binom{2n+1}{n} &= \frac{(2n+1)!}{n!(n+1)!} \\ &< \frac{e(2n+1)\left(\frac{2n+1}{e}\right)^{2n+1}}{e\left(\frac{n}{e}\right)^n e\left(\frac{n+1}{e}\right)^{n+1}},\end{aligned}$$

and this last expression is easily seen to be less than $8e^{2n}$ for $n \geq 2$. ∎

Proof of Proposition 6-1. Note that $r_{\vec{a}}$ is a positive linear combination of terms of the form

$$t_{\vec{a}}^{(j)} = \frac{1}{a_j \prod_{r>j}(a_r^2 - a_j^2) \prod_{s<j}(a_j^2 - a_s^2)},$$

and recall that $r_{\vec{a}_0} = 1$ where $\vec{a}_0$ is the index corresponding to the trivial representation. Hence, for any permissible index $\vec{a}$,

$$r_{\vec{a}} = \frac{r_{\vec{a}}}{r_{\vec{a}_0}} \leq \max_{1 \leq j \leq n} \frac{t_{\vec{a}}^{(j)}}{t_{\vec{a}_0}^{(j)}},$$

so it suffices to show that

$$d_{\vec{a}}\, r_{\vec{a}}^{(j)k} \leq \left(Ke^{-c/8}\right)^{b_1+\ldots+b_n}, \qquad j = 1, 2, \ldots, n,$$

where $r_{\vec{a}}^{(j)} = \frac{t_{\vec{a}}^{(j)}}{t_{\vec{a}_0}^{(j)}}$.

To this end, we fix $j \in \{1, 2, \ldots, n\}$, and fix an index $\vec{a}^* = (a_1^*, a_2^*, \ldots, a_n^*)$. We shall proceed by *moving* the values of $\vec{a}$ from their initial values of $\vec{a}_0$ to the final values of $\vec{a}^*$.

The outline is as follows. We shall first increase each of the indices $a_{j+1}, \ldots, a_n$ enough to get the differences $a_r - a_j$ right for $r > j$. In other words, we shall increase a_n from its value of $n - \frac{1}{2}$ for the trivial representation to the value $a_n^* - a_j^* + (j - \frac{1}{2})$. We shall show that the value of $d_{\vec{a}}\, r_{\vec{a}}^{(j)k}$ decreases by a factor of at least $(Ke^{-c/8})$ each time we increase a_n by 1. We shall do the same for each of $a_{j+1}, \ldots, a_{n-1}$. We shall then increase *all* of the indices $a_1, \ldots, a_n$ *simultaneously* until a_1 takes on its correct value, and we shall show that $d_{\vec{a}}\, r_{\vec{a}}^{(j)k}$ decreases by a factor of at least $(Ke^{-c/8})^n$ each time we increase each of $a_1, \ldots, a_n$ by 1. Once a_1 has its correct value, we shall increase each of $a_2, \ldots, a_n$ simultaneously until a_2 takes on its correct value, and we shall again get a decrease in $d_{\vec{a}}\, r_{\vec{a}}^{(j)k}$ by a factor of at least $(Ke^{-c/8})^n$. We continue in this manner. When finally a_j takes on its correct value, we are done. The rest of this proof is merely an ellaboration of these ideas.

We begin with a_n (assuming $j < n$). We move it, one step at a time, from its initial value of $n - \frac{1}{2}$ to the value $(j - \frac{1}{2}) + a_n^* - a_j^*$ (to get the value of $a_n - a_j$ right). We claim that each such "move" decreases $d_{\vec{a}}\, r_{\vec{a}}^{(j)k}$ by a factor of $\left(Ke^{-c/8}\right)$ or better. Indeed, from Lemma 6-2, $d_{\vec{a}}$ is multiplied each time by an amount

$$\left(1 + \frac{1}{a_n}\right)\left(1 + \frac{2n-2}{a_n - n + \frac{3}{2}}\right) < 2\left(1 + \frac{2n}{a_n - n + \frac{3}{2}}\right) .$$

Also the only term in $r_{\vec{a}}^{(j)}$ which is affected is the term $\frac{1}{(a_n^2 - a_j^2)}$ which is multiplied by an amount

$$\begin{aligned}\frac{a_n^2 - a_j^2}{(a_n+1)^2 - a_j^2} &\le \frac{a_n^2}{(a_n+1)^2} \\ &= \frac{1}{(1 + \frac{1}{a_n})^2} .\end{aligned}$$

Now, it is easily verified that for $0 \le x \le 1$,

$$(1+x)e^{x^2/2} \ge e^x .$$

Hence,

$$\frac{1}{1+x} \le e^{-x}e^{x^2/2} .$$

Thus $r_{\vec{a}}^{(j)k}$ is multiplied by an amount which is less than

$$\left(\frac{1}{\left(1+\frac{1}{a_n}\right)^2}\right)^k \leq \left(\frac{1}{1+\frac{1}{a_n}}\right)^{(n\log n+4cn)}$$

$$\leq \left(e^{-1/a_n}e^{1/2(a_n)^2}\right)^{(n\log n+4cn)}$$

For $n \geq 20$, this is less than

$$2e^{c/9}\exp(-(n\log n+4cn)/a_n)\ .$$

Hence, writing $b_n = a_n - (n-\frac{1}{2})$, we have that $d_{\vec{a}}\, r_{\vec{a}}^{(j)k}$ is multiplied by less than

$$\begin{aligned} &4e^{c/9}\left(1+\frac{2n}{a_n-n+\frac{3}{2}}\right)\exp(-(n\log n+4cn)/a_n) \\ &\leq 4e^{c/9}\left(1+\frac{2n}{b_n+1}\right)\exp(-(n\log n+4cn)/(n+b_n)) \\ &= 4e^{c/9}\left(1+\frac{2n}{b_n+1}\right)\exp((-(n+b_n)\log n+b_n\log n-4cn)/(n+b_n)) \\ &= 4e^{c/9}\left(\frac{1}{n}+\frac{2}{b_n+1}\right)\exp\left((b_n\log n-4cn)/(n+b_n)\right)\ . \end{aligned}$$

Now, if $b_n < \frac{n}{\log n}$, this amount is less than

$$12e^{c/9}\exp\left((n-4cn)/2n\right) = 12e^{c/9}\exp\left(\frac{1-4c}{2}\right) < 10e^{-c}$$

which is smaller than is required. If $\frac{n}{\log n} < b_n < 0.1n$, this amount is less than

$$\frac{12\,e^{c/9}}{\left(\frac{n}{\log n}\right)}\exp\left((0.1n\log n-4cn)/1.1n\right)\ \leq\ \frac{12\log n}{n^{0.9}}e^{-c}\ \leq\ 12e^{-c}\ ,$$

which is again smaller than required. Finally, if $0.1n \leq b_n \leq 8n$, then writing t for b_n/n, so that $0.1 \leq t \leq 8$, we have that this amount is less than

$$\begin{aligned} 12\,e^{c/9}\left(\frac{1}{n}\right)\left(1+\frac{2}{t}\right)\exp\left((tn\log n-4cn)/((1+t)n)\right) &= 12\,e^{c/9}\left(\frac{1}{n}\right)\left(1+\frac{2}{t}\right)n^{\frac{t}{1+t}}e^{-\frac{4c}{1+t}} \\ &= 12\,e^{c/9}\left(\frac{t+2}{t}\right)n^{\frac{-1}{1+t}}e^{-\frac{c}{1+t}} \\ &< 252\,e^{-c/3}\ , \end{aligned}$$

which is smaller than is required.

To summarize, we have moved a_n from its initial value of $n - \frac{1}{2}$ to the value $a_n^* - a_j^* + (j - \frac{1}{2})$, by increasing the value of a_n in steps of 1 in such a way that the value of $d_{\vec{a}}\, r_{\vec{a}}^{(j)k}$ was multiplied by a factor smaller than $Ke^{-c/8}$ each time.

We now move a_{n-1} from its initial value of $n - \frac{3}{2}$ to the value $a_{n-1}^* - a_j^* + (j - \frac{1}{2})$, in steps of 1. The process is the same as the above, and the same argument shows that $d_{\vec{a}}\, r_{\vec{a}}^{(j)k}$ is multiplied by a factor smaller than $Ke^{-c/8}$ each time. In fact, the argument is a bit *less* delicate, since $d_{\vec{a}}$ is actually multiplied by a bit *less* than it was when we moved a_n, while $r_{\vec{a}}$ is actually decreased by a bit more.

We continue in this manner, moving each of $a_{n-2}, a_{n-3}, \ldots, a_{j+1}$ in turn, by steps of 1. When we are done, we have that $a_m - a_j = a_m^* - a_j^*$ for $j+1 \le m \le n$, and furthermore that $d_{\vec{a}}\, r_{\vec{a}}^{(j)k}$ has been multiplied by an amount less than $\left(Ke^{-c/8}\right)^{b_{j+1}+\ldots+b_n}$, where b_m represents the amount a_m has been increased *so far.*

To complete the proof of Proposition 6-1, we proceed as follows. We increase *each* of $a_1, a_2, \ldots, a_n$ by 1, simultaneously. Under such a move, $d_{\vec{a}}$ is multiplied by less than $8e^{2n}$, by Corollary 6- 4. Also, the terms in $r_{\vec{a}}^{(j)}$ which are affected are those of the form $\frac{1}{a_j \prod\limits_{r\neq j}(a_r+a_j)}$. They are multiplied by

$$\left(\frac{a_j}{a_j+1}\right)\prod_{r\neq j}\left(\frac{a_r+a_j}{a_r+a_j+2}\right) \le \left(\frac{1}{1+\frac{1}{a_n}}\right)^n .$$

Reasoning as before, this is less than

$$\left(2e^{c/9}e^{1/a_n}\right)^n .$$

Hence, $d_{\vec{a}}\, r_{\vec{a}}^{(j)k}$ is multiplied by an amount which is less than

$$\begin{aligned} 8(2e^{c/9})^n e^{2n}\left(e^{-n/a_n}\right)^k &< 8(2e^{c/9})^n \exp\left(2n \;-\; n(\frac{1}{2}n\log n + 2cn)/8n\right) \\ &= 8(2e^{c/9})^n \exp\left(2n \;-\; \frac{1}{16}n\log n \;-\; cn/4\right) . \end{aligned}$$

Now, if we choose n so large that $2n - \frac{1}{16}n\log n < 0$, this is less than

$$8(2e^{c/9})^n(e^{-c/4})^n < (16e^{-c/8})^n .$$

This is less than the factor $(Ke^{-c/8})^n$ which was required.

We continue to increase each of $a_1, a_2, \ldots, a_n$ by 1 until $a_1 = a_1^*$. Each time we pick up a factor smaller than $\left(16e^{-c/8}\right)^n$. Once we have $a_1 = a_1^*$, we then move each of $a_2, a_3, \ldots, a_n$ by 1 until $a_2 = a_2^*$. The argument above shows that for each of these moves, we again pick up a factor smaller than $\left(16e^{-c/8}\right)^n$. Indeed, each time $d_{\vec{a}}$ is multiplied by an amount bounded just as it was above. As for $r_{\vec{a}}^{(j)}$, the factor $\frac{1}{a_1+a_j}$ in $r_{\vec{a}}^{(j)}$ is not decreased by as much as it was above, but the factor $\frac{1}{a_j-a_1}$, which remained the same before, now decreases by enough to compensate.

Once we have $a_2 = a_2^*$, we then increase each of $a_3, a_4, \ldots, a_n$ by 1 until $a_3 = a_3^*$. We continue in this manner until finally we are increasing only each of $a_j, a_{j+1}, \ldots, a_n$ by 1 to make $a_j = a_j^*$. When this is done, we have that $\vec{a} = \vec{a}^*$. Furthermore, we have picked up a factor of $\left(252e^{-c/8}\right)$ or smaller for *each* time we have increased *any* a_m by 1. This completes the proof of the Proposition. ∎

To complete the proof of Theorem 2-2, we need only bound the sum of terms $d_{\vec{a}}^2 r_{\vec{a}}^{2k}$ with $a_n \geq 8n$. We do this as follows. For $1 \leq m \leq n$, and for any permissible index $\vec{a} = (a_1, \ldots, a_n)$, we set

$$d_{\vec{a}}^{[m]} = \frac{2^m}{1!3!\ldots(2m-1)!} \prod_{r=1}^{m} a_r \prod_{r=2}^{m} \prod_{s=1}^{r-1} \left(a_r^2 - a_s^2\right)$$

and

$$r_{\vec{a}}^{[m]} = \frac{(2m-1)!}{2^{2m-1}} \sum_{j=1}^{m} \frac{1}{\prod_{\substack{r=1 \\ r\neq j}}^{m} a_r \prod_{r=j+1}^{m} \left(a_r^2 - a_j^2\right) \prod_{s=1}^{j-1} \left(a_j^2 - a_s^2\right)} .$$

In other words, $d_{\vec{a}}^{[m]}$ and $r_{\vec{a}}^{[m]}$ are the values of $d_{\vec{a}}$ and $r_{\vec{a}}$ if we replace n by m and consider only the indices $a_1, \ldots, a_m$. It is easily verified that

$$d_{\vec{a}}^{[m]} \leq \frac{2}{(2m-1)!} (a_{m+1})^{2m-1} d_{\vec{a}}^{[m-1]}$$

and

$$r_{\vec{a}}^{[m]} \leq \frac{m}{a_m} r_{\vec{a}}^{[m-1]} .$$

We wish to prove that

$$\sum_{a_1,\ldots,a_m} d_{\vec{a}}^{[m]2} r_{\vec{a}}^{[m]2k} \ \le\ 1 + (2K^2+1)e^{-c/4}, \qquad m = 2,3,\ldots,n\ , \tag{6-1}$$

for sufficiently large n and c, with K as in Proposition 6-1. (If we prove this, then Theorem 2-2 follows by setting $m = n$. Note, however, that we are *not* proving Theorem 2-2 for all values of m, since the value of $k \ge \frac{1}{2} n \log n + 2cn$ does not get smaller with m.) We proceed by induction on m. For $m = 2$,

$$d_{\vec{a}}^{[2]} = \frac{2}{3} a_1 a_2 (a_2^2 - a_1^2) \le a_1 a_2^3$$

and

$$r_{\vec{a}}^{[2]} = \frac{3}{4} \left(\frac{1}{a_1(a_2^2 - a_1^2)} + \frac{1}{a_2(a_2^2 - a_1^2)} \right) \le \frac{3}{2} \frac{1}{a_1 a_2}\ .$$

The sum

$$\sum_{a_1,a_2} d_{\vec{a}}^{[2]2} r_{\vec{a}}^{[2]2k}$$

may now easily be bounded by a double integral (after summing the first few terms directly), and made to be much smaller than is required.

For the induction step, we first note that the argument given earlier in this Section shows that the sum

$$\sum_{\substack{a_1,\ldots,a_m \\ (a_m < 8n)}} d_{\vec{a}}^{[m]2} r_{\vec{a}}^{[m]2k} \ \le\ 1 + 2q, \qquad m = 2,3,\ldots,n\ ,$$

for c larger than some universal constant c_0, where $q = K^2 e^{-c/4}$. Indeed, Proposition 6-1 still holds if we replace $d_{\vec{a}}$ by $d_{\vec{a}}^{[m]}$ and $r_{\vec{a}}$ by $r_{\vec{a}}^{[m]}$, and set $b_r = 0$ for $m+1 \le r \le n$. To see this, note that this replacement clearly makes $d_{\vec{a}}$ smaller. As for $r_{\vec{a}}$, the only two estimates we made for changes in $r_{\vec{a}}^{(j)}$ during the proof of Proposition 6-1 were the estimates

$$\frac{1}{(1+\frac{1}{a_n})^2}$$

and

$$\left(\frac{1}{(1+\frac{1}{a_n})^2} \right)^n .$$

The first of these is actually made smaller by our replacement. As for the second, our replacement will leave it smaller than

$$\left(\frac{1}{(1+\frac{1}{a_n})^2} \right)^m ,$$

which is all that we require since in Corollary 6-4 we may now replace n by m.

Hence, we need only bound the sum

$$\sum_{\substack{a_1,\ldots,a_m \\ (a_m > 8n)}} d_{\vec{a}}^{[m]2} r_{\vec{a}}^{[m]2k} .$$

We have that

$$\sum_{\substack{a_1,\ldots,a_m \\ (a_m > 8n)}} d_{\vec{a}}^{[m]2} r_{\vec{a}}^{[m]2k} \le \left(\sum_{a_m = 8n+\frac{1}{2}}^{\infty} \left(\frac{m}{a_m} \right)^{2k} a_m^{2m} \right) \sum_{a_1,\ldots,a_{m-1}} d_{\vec{a}}^{[m-1]2} r_{\vec{a}}^{[m-1]2k}$$

where we now allow $a_1, \ldots, a_{m-1}$ to take on *all* admissible values, not just those which are less than some specified value of a_m. But by induction, the sum over the $a_1, \ldots, a_{m-1}$ is 1 + something small, so in any case we may take it to be less than 2. We then have that

$$\begin{aligned}
\sum_{\substack{a_1,\ldots,a_m \\ (a_m > 8n)}} d_{\vec{a}}^{[m]2} r_{\vec{a}}^{[m]2k} &\le 2 \left(\sum_{a_m = 8n+\frac{1}{2}}^{\infty} \left(\frac{m}{a_m} \right)^{2k} a_m^{2m} \right) \\
&\le 2n^{2k} \int_{8n-1}^{\infty} x^{2n-2k}\, dx \\
&= 2n^{2k}(8n-1)^{2n-2k+1}/(2k-2n-1) \\
&< 2n^{2k}(8n-1)^{2n-2k+1} \\
&= 2\exp\left(-2(\frac{1}{2} n \log n + cn) \log(8 - \frac{1}{n}) \ + \ (2n+1)\log(8n-1) \right) \\
&< e^{-2cn \log 8} ,
\end{aligned}$$

for sufficiently large n, where we have used the fact that $\log(8-\frac{1}{n}) > 2$ and $2k-2n-1 > 1$. This is smaller than, say, $e^{-c/4}$, establishing (6-1). This completes the induction step, and therefore the proof of Theorem 2-2.

Remarks:

1. *The cutoff phenomenon.* Theorems 2-1 and 2-2 provide an example of the "cutoff" or "threshold" phenomenon (for $\theta = \pi$), whereby the variation distance from Haar measure jumps from being near 1 to being near 0 over a relatively short interval of values of k (the number of rotations being done). See Diaconis [D], Aldous and Diaconis [AD], and Hildebrand [H] for a discussion of this phenomenon and for other examples, especially on finite groups. It is an open problem to explain why this phenomenon occurs in so many of the examples of random processes on finite and compact groups which have been studied.
2. *Other values of θ.* One suspects that the "cutoff phenomenon" described above also occurs for values of θ other than 180 degrees. Specifically, one suspects that if we do $\frac{1}{2(1-\cos\theta)} N \log N \ + \ cN$ rotations through an angle θ, then the variation distance to Haar measure will decrease to 0 exponentially as c increases, uniformly in N. The methods used in this section should suffice to show this, but the computations appear to be somewhat more difficult. Choosing θ from a more complicated distribution (say, uniform on $(0, \pi/2)$) is also possible.
3. *Connection with Random Reflections.* As mentioned in the Introduction, Diaconis and Shashahani [DS2] have carried out a similar analysis for a process of "random reflections" on $O(N)$. They determine that $\frac{1}{2} N \log N$ such reflections suffice to get close to Haar measure in total variation distance. This value is precisely *twice* the value obtained in Theorem 2-2 above. This is not surprising. Doing two consecutive reflections through axes making an angle α with each other is precisely the same as doing a rotation through an angle 2α in the plane spanned by the two axes. For large N, two random axes will make an angle of approximately 90 degrees with each other with high probability, and so 2α will be close to 180 degrees, so that two random reflections is roughly the same as one random rotation through an angle of 180 degrees.
4. *Possible Extensions.* The methods used in this paper would appear to be applicable to any conjugate-invariant random process on any compact Lie group. For example, the unitary and symplectic groups are promising candidates for further analysis.

Acknowledgements: I am very grateful to Persi Diaconis, my Ph.D. advisor at Harvard

University, for providing lots of valuable guidance, advice, and encouragement. I am also very grateful to Peter Magyar for extremely helpful discussions about the theory of Lie groups.

REFERENCES

[Ad] J. F. Adams (1969), *Lectures on Lie Groups,* University of Chicago Press, Chicago.

[Ar] E. Artin (1964), *The Gamma Function* (trans. M. Butler), Holt, Rinehart and Wilson, New York.

[AD] D. Aldous and P. Diaconis (1987), *Strong Stopping Times and Finite Random Walks,* Adv. in App. Math. **8**, 69-97.

[D] P. Diaconis (1988), *Group Representations in Probability and Statistics,* IMS Lecture Series volume **11**, Institute of Mathematical Statistics, Hayward, California.

[DS1] P. Diaconis and M. Shashahani (1981), *Generating a Random Permutation with Random Transpositions,* Z. Wahrscheinlichkeitstheorie Verw. Gebiete **57**, 159-179.

[DS2] P. Diaconis and M. Shashahani (1986), *Products of Random Matrices as they arise in the study of Random Walks on Groups,* Contemporary Mathematics **50**, 183-195.

[GS] A.E. Gelfand and A.F.M. Smith (1990), *Sampling-based approaches to calculating marginal densities,* J. Amer. Stat. Soc. **85**, 398-409.

[H] M.V. Hildebrand (1990), *Rates of Convergence of some Random Processes on Finite Groups,* Ph.D. thesis, Mathematics Department, Harvard University.

[HR] E. Hewitt and K.A. Ross (1963 and 1970), *Abstract Harmonic Analysis, Vol. I, II,* Springer-Verlag, Berlin.

[R] J.S. Rosenthal (1991), *Rates of convergence for Gibbs sampling for variance component models,* Technical Report, University of Minnesota.

School of Mathematics
University of Minnesota
Minneapolis, MN 55455, U.S.A.