The Annals of Probability
1983, Vol. 11, No. 4, 909–930

LIMIT THEOREMS FOR CERTAIN BRANCHING RANDOM WALKS ON COMPACT GROUPS AND HOMOGENEOUS SPACES

BY SVANTE JANSON

Uppsala University

A certain branching random walk, $\{X_i\}$, on a compact group or a compact homogeneous space is studied. It is proved that the sums $\sum_0^n f(X_i)$ are asymptotically normally distributed for all nice functions f if and only if the Fourier coefficients of the transition probability distribution have real parts not exceeding ½.

1. Introduction. We will study a certain branching random walk consisting of a sequence $\{X_n\}_0^\infty$ of random points in a suitable state space. The state space will be a compact group or, more generally, a compact homogeneous space, e.g. a sphere. In the somewhat simpler case of a group, G, the exact definition is as follows.

Let $X_0, Y_1, Y_2, Y_3, \cdots, I_1, I_2, \cdots$ be independent random variables such that X_0 has the distribution ν on G, each Y_n has the distribution μ on G and I_n has a uniform distribution on the set $\{0, 1, \cdots, n-1\}$, $n \geq 1$. Define $\{X_n\}$ inductively by

$$X_n = X_{I_n} Y_n, \quad n \geq 1. \tag{1.1}$$

Here ν and μ are two arbitrary probability measures on G. In order to avoid trivial complications, we will assume that μ is not supported on any proper closed subgroup of G. (We do not impose this condition on ν; ν may e.g. be the point mass at the identity element.)

This is equivalent to the following description: First X_0 is chosen according to an initial distribution ν. Then, at the nth step, one of $X_0 \cdots X_{n-1}$ is selected at random to be the parent X_{I_n} of X_n. The daughter X_n then appears with a displacement Y_n from its parent. All displacements are identically distributed and all choices are independent. As an example, we may assume that every point, from its birth on, generates new points (with independent random displacements as above) according to a Poisson process with fixed intensity. (Cf. [1].)

The definition for a homogeneous space is essentially the same, the only difference being that (1.1) does not make sense, and thus another definition of the displacements is required. We will discuss this in detail in Section 5. In the particular case of a sphere, we may simply say that the displacement from a parent point to a daughter has a length according to some fixed distribution and a uniformly distributed direction.

We may compare the branching random walk with a simple random walk which has the same definition except that $I_n = n - 1$. It is well-known that, except in degenerate cases, the distributions of the points, X_n, in a simple random walk converge to the uniform distribution, i.e. the Haar measure, henceforth denoted by m, [3], [5]. It will be shown below that the same is true for the branching random walk. Thus $X_0 \cdots X_n$ will be rather uniformly spread out over the state space (for n large). The purpose of this paper is to study the fluctuations from the uniform distribution; more precisely we will study the asymptotic distribution of the sums $S_n(f)$, defined by

$$S_n(f) = \textstyle\sum_0^n f(X_i) \qquad (f \in L^2(m)). \tag{1.2}$$

Note that if A is a Borel subset of the state space and χ_A its indicator function, $S_n(\chi_A) = \#\{k \leq n : X_k \in A\}$.

In the trivial case when $\mu = m$, $\{X_n\}_1^\infty$ are uniformly distributed and independent. Thus

Received September 1982.

AMS 1980 *subject classifications*. 60J80, 60B15, 60J15, 60F05.

Key words and phrases. Branching random walks, compact groups, limit theorems.

$S_n(f)$ is a sum of i.i.d. random variables with finite second moments and, by the Central Limit Theorem, $S_n(f)$ is asymptotically normally distributed.

The main result is that this is typical for a large class of displacement measures μ, but not for all of them. In fact (assuming a technical smoothness condition on μ, ν or f), $n^{-1/2}(S_n(f) - n \int f\,dm)$ converges in distribution to a normal distribution—provided that all Fourier coefficients (omitting the trivial one which always equals one) have real parts less than ½. (This is slightly simplified. In general, the Fourier coefficients are matrices, see Sections 4 and 5, and the correct condition is that the eigenvalues of these matrices should have real parts less than ½.) Moreover, if the real part of (an eigenvalue of) some Fourier coefficient exceeds ½, this utterly fails: $S_n(f)$ will in general have a variance growing faster than n, and with the correct normalization it may converge to a non-normal distribution, which furthermore depends on the distribution ν of X_0. (The borderline case when the real part of some Fourier coefficient equals ½, but none is larger, is intermediate.) This expresses the intuitive feeling that if the displacements are small, the points will tend to cluster. The sharp division at Re $\lambda = ½$ between weak and strong clustering for various distributions of the displacements is vaguely reminiscent of phase transitions in statistical mechanics. (This phenomenon does not occur for simple random walks. A theorem similar to Theorem 4.3 holds (for the simple random walk on a compact group) for all measures μ that are not supported on any coset of a proper closed subgroup.)

In Sections 2 and 3 we will only study the technically simpler case of a commutative compact group, although most of the results will be generalized in Section 4. Section 2 treats $S_n(f)$ for characters while Section 3 treats general functions f. Section 4 treats a general compact group. In Section 5 we apply and adapt the preceding results to homogeneous spaces. Section 6 contains some examples.

2. Commutative groups: characters. Throughout Sections 2 and 3, G is a commutative compact group.

We recall the following facts from abstract harmonic analysis. A character γ is a continuous complex-valued function on G such that $|\gamma(g)| = 1$ and $\gamma(gg') = \gamma(g)\gamma(g')$, g, $g' \in G$, i.e. a continuous homomorphism into the circle group. The product of two characters is a character, and so is the complex conjugate of a character. The set Γ of all characters form a group, the dual group of G. Γ is an orthonormal basis in $L^2(m)$, thus $f = \Sigma \hat{f}(\gamma)\gamma$ with $\hat{f}(\gamma) = \int f\bar{\gamma}\,dm$ for any $f \in L^2$. Similarly, we define $\hat{\mu}(\gamma) = \int \gamma\,d\mu$ for measures μ. (For notational convenience we use γ and not $\bar{\gamma}$ here.) A special rôle is played by the trivial character ι, defined by $\iota(g) = 1$. Note that $\int \gamma\,dm = \int \gamma\bar{\iota}\,dm = 0$ for $\gamma \not\in \iota$. Since μ is a probability measure $|\mu(\gamma)| \le 1$, $\gamma \in \Gamma$ and our assumption that μ is not supported on any proper closed subgroup is equivalent to $\mu(\gamma) \ne 1$ for $\gamma \ne \iota$.

We will distinguish between real and complex characters and define $\Gamma_r = \{\gamma \in \Gamma : \gamma = \bar{\gamma} \text{ and } \gamma \ne \iota\}$. Let Γ_c consists of one character from each pair $(\gamma, \bar{\gamma})$, $\gamma \ne \bar{\gamma}$. Then $\Gamma = \{\iota\} \cup \Gamma_r \cup \Gamma_c \cup \bar{\Gamma}_c$. If $\gamma \ne \bar{\gamma}$ then $\int \gamma^2\,dm = 0$ while $\int \gamma\bar{\gamma}\,dm = 1$; hence $\int (\text{Re}\,\gamma)^2 = \int (\text{Im}\,\gamma)^2 = ½$ and $\int \text{Re}\,\gamma\,\text{Im}\,\gamma = 0$. Consequently, $\{\iota\} \cup \Gamma_r \cup \{\sqrt{2}\,\text{Re}\,\gamma, \sqrt{2}\,\text{Im}\,\gamma : \gamma \in \Gamma_c\}$ is a real orthonormal basis in $L^2(m)$.

We will study the asymptotic distribution of the complex random variables $S_n(\gamma) = \sum_0^n \gamma(X_n)$ $(\gamma \in \Gamma)$ by estimating moments. Therefore we define, for $m \ge 1$ and $\gamma_1 \cdots \gamma_m \in \Gamma$,

$$F_m(\gamma_1, \cdots, \gamma_m; z) = \textstyle\sum_0^\infty z^n E(S_n(\gamma_1)S_n(\gamma_2) \cdot \cdots \cdot S_n(\gamma_m)). \tag{2.1}$$

This generating function is analytic for $|z| < 1$. We have

$$F_m(\gamma_1, \cdots, \gamma_m; 0) = E(\gamma_1(X_0) \cdot \cdots \cdot \gamma_m(X_0)) = \hat{\nu}(\gamma_1 \cdot \cdots \cdot \gamma_m). \tag{2.2}$$

By the definition (1.1) of the branching random walk,

$$E(\gamma(X_n) \mid X_0 \cdots X_{n-1}) = E(\gamma(X_{I_n})\gamma(Y_n) \mid X_0 \cdots X_{n-1})$$

$$(2.3)\qquad = E(\gamma(X_{I_n}) \mid X_0 \cdots X_{n-1})E(\gamma(Y_n)) = \frac{1}{n} S_{n-1}(\gamma) \int \gamma(g)\, d\mu(g) = \frac{1}{n} S_{n-1}(\gamma)\hat{\mu}(\gamma).$$

Thus,

$$(2.4)\qquad E\gamma(X_n) = \frac{1}{n} ES_{n-1}(\gamma)\hat{\mu}(\gamma)$$

$$(2.5)\qquad ES_{n+1}(\gamma) = ES_n(\gamma) + E\gamma(X_{n+1}) = ES_n(\gamma)\left(1 + \frac{1}{n+1}\hat{\mu}(\gamma)\right)$$

$$(2.6)\qquad \begin{aligned}(1-z)\frac{d}{dz}F_1(\gamma; z) &= \textstyle\sum_0^\infty n(z^{n-1} - z^n)ES_n(\gamma) = \sum_0^\infty z^n((n+1)ES_{n+1}(\gamma) - nES_n(\gamma))\\ &= \textstyle\sum_0^\infty z^n(1 + \hat{\mu}(\gamma))ES_n(\gamma) = (1 + \hat{\mu}(\gamma))F_1(\gamma; z).\end{aligned}$$

This differential equation yields

$$(2.7)\qquad F_1(\gamma; z) = F_1(\gamma; 0)(1-z)^{-1-\hat{\mu}(\gamma)} = \hat{\nu}(\gamma)(1-z)^{-1-\hat{\mu}(\gamma)}.$$

We use the binomial expansion

$$(2.8)\qquad (1-z)^{-\alpha} = \textstyle\sum_0^\infty (-1)^n\binom{-\alpha}{n}z^n = \sum_0^\infty \binom{\alpha+n-1}{n}z^n.$$

Hence

$$(2.9)\qquad ES_n(\gamma) = \hat{\nu}(\gamma)\binom{n + \hat{\mu}(\gamma)}{n}$$

which asymptotically yields

$$(2.10)\qquad ES_n(\gamma) \sim \frac{\hat{\nu}(\gamma)}{\Gamma(\hat{\mu}(\gamma) + 1)} n^{\hat{\mu}(\gamma)}.$$

THEOREM 2.1. *X_n converges to the uniform distribution, i.e. for every continuous function f on G,*

$$(2.11)\qquad \lim_{n\to\infty} Ef(X_n) = \lim_{n\to\infty} \frac{1}{n+1} ES_n(f) = \int f\, dm.$$

PROOF. We note that $|\hat{\mu}(\gamma)| \le 1$ for every character γ and that $\hat{\mu}(\gamma) = 1$ if and only if μ is supported on the subgroup $\{g \in G: \gamma(g) = 1\}$, but for $\gamma \ne \iota$ this is eliminated by our assumption on μ. Thus, if $\gamma \ne \iota$, Re $\hat{\mu}(\gamma) < 1$ and by (2.10) (for some $C < \infty$)

$$\left|\frac{1}{n+1} ES_n(\gamma)\right| \le Cn^{\operatorname{Re}\hat{\mu}(\gamma)}/n \to 0 = \int \gamma\, dm \quad \text{as} \quad n \to \infty.$$

By (2.4), also $E\gamma(X_n) \to 0$, $\gamma \ne \iota$. Since (2.11) is trivial for $f = \iota$, it holds for all characters. Thus, it holds for every continuous f by [3, Theorem 3.2.2].

The second moments of $\{S_n(\gamma)\}$ are computed by the same method.

$$S_{n+1}(\gamma_1)S_{n+1}(\gamma_2) = S_n(\gamma_1)S_n(\gamma_2) + S_n(\gamma_1)\gamma_2(X_{n+1}) + \gamma_1(X_{n+1})S_n(\gamma_2) + \gamma_1\gamma_2(X_{n+1})$$

$$(2.12)\qquad \begin{aligned}E(S_{n+1}(\gamma_1)S_{n+1}(\gamma_2) \mid X_1 \cdots X_n) &= S_n(\gamma_1)S_n(\gamma_2) + \frac{1}{n+1} S_n(\gamma_1)S_n(\gamma_2)\hat{\mu}(\gamma_2)\\ &\quad + \frac{1}{n+1} S_n(\gamma_1)S_n(\gamma_2)\hat{\mu}(\gamma_1) + \frac{1}{n+1} S_n(\gamma_1\gamma_2)\hat{\mu}(\gamma_1\gamma_2)\end{aligned}$$

$$(1-z)\frac{d}{dz}F_2(\gamma_1,\gamma_2;z)$$

$$\begin{aligned}&=\textstyle\sum_0^\infty z^n((n+1)E(S_{n+1}(\gamma_1)S_{n+1}(\gamma_2))-nE(S_n(\gamma_1)S_n(\gamma_2)))\\ &=\textstyle\sum_0^\infty z_n((1+\hat\mu(\gamma_1)+\hat\mu(\gamma_2))ES_n(\gamma_1)S_n(\gamma_2)+\hat\mu(\gamma_1\gamma_2)ES_n(\gamma_1\gamma_2))\\ &=(1+\hat\mu(\gamma_1)+\hat\mu(\gamma_2))F_2(\gamma_1,\gamma_2;z)+\hat\mu(\gamma_1\gamma_2)F_1(\gamma_1\gamma_2;z)\\ &=(1+\hat\mu(\gamma_1)+\hat\mu(\gamma_2))F_2(\gamma_1,\gamma_2;z)+\hat\mu(\gamma_1\gamma_2)\hat\nu(\gamma_1\gamma_2)(1-z)^{-1-\hat\mu(\gamma_1\gamma_2)}\end{aligned}\tag{2.13}$$

using (2.7) for $\gamma_1\gamma_2$. Hence

$$\begin{aligned}F_2(\gamma_1,\gamma_2;z)=\hat\nu(\gamma_1\gamma_2)&\frac{\hat\mu(\gamma_1\gamma_2)}{\hat\mu(\gamma_1\gamma_2)-\hat\mu(\gamma_1)-\hat\mu(\gamma_2)}(1-z)^{-1-\hat\mu(\gamma_1\gamma_2)}\\ -\hat\nu(\gamma_1\gamma_2)&\frac{\hat\mu(\gamma_1)+\hat\mu(\gamma_2)}{\hat\mu(\gamma_1\gamma_2)-\hat\mu(\gamma_1)-\hat\mu(\gamma_2)}(1-z)^{-1-\hat\mu(\gamma_1)-\hat\mu(\gamma_2)}\end{aligned}\tag{2.14a}$$

or, when $\hat\mu(\gamma_1\gamma_2)=\hat\mu(\gamma_1)+\hat\mu(\gamma_2)$,

$$\begin{aligned}F_2(\gamma_1,\gamma_2;z)=&-\hat\nu(\gamma_1\gamma_2)\hat\mu(\gamma_1\gamma_2)\log(1-z)(1-z)^{-1-\hat\mu(\gamma_1\gamma_2)}\\ &+\hat\nu(\gamma_1\gamma_2)(1-z)^{-1-\hat\mu(\gamma_1\gamma_2)}.\end{aligned}\tag{2.14b}$$

Define

$$p_n(x)=x\binom{n+x}{n}=x(1+x)\left(1+\frac{x}{2}\right)\cdot\cdots\cdot\left(1+\frac{x}{n}\right)\sim\frac{1}{\Gamma(x)}n^x.\tag{2.15}$$

Then (2.14) and (2.8) yield

$$ES_n(\gamma_1)S_n(\gamma_2)=\hat\nu(\gamma_1\gamma_2)\frac{p_n(\hat\mu(\gamma_1\gamma_2))-p_n(\hat\mu(\gamma_1)+\hat\mu(\gamma_2))}{\hat\mu(\gamma_1\gamma_2)-\hat\mu(\gamma_1)-\hat\mu(\gamma_2)}\tag{2.16a}$$

or, when $\hat\mu(\gamma_1\gamma_2)=\hat\mu(\gamma_1)+\hat\mu(\gamma_2)$,

$$ES_n(\gamma_1)S_n(\gamma_2)=\hat\nu(\gamma_1\gamma_2)p_n'(\hat\mu(\gamma_1\gamma_2)).\tag{2.16b}$$

Now, we specialize to the case $\gamma_2=\bar\gamma_1$. Then $\gamma_1\gamma_2=\iota$, $\hat\mu(\gamma_1\gamma_2)=\hat\nu(\gamma_1\gamma_2)=1$ and $\hat\mu(\gamma_1)+\hat\mu(\gamma_2)=2\operatorname{Re}\hat\mu(\gamma_1)$.

We recognize three different cases:

If $\operatorname{Re}\hat\mu(\gamma)<\frac12$,

$$\begin{aligned}E|S_n(\gamma)|^2&=\frac{p_n(1)-p_n(2\operatorname{Re}\hat\mu(\gamma))}{1-2\operatorname{Re}\hat\mu(\gamma)}\\ &=\frac{n+1}{1-2\operatorname{Re}\hat\mu(\gamma)}+O(n^{2\operatorname{Re}\hat\mu(\gamma)})\sim\frac{n}{1-2\operatorname{Re}\hat\mu(\gamma)}.\end{aligned}\tag{2.17a}$$

If $\operatorname{Re}\hat\mu(\gamma)=\frac12$,

$$E|S_n(\gamma)|^2=p_n'(1)=p_n(1)\left(1+\frac12+\cdots+\frac{1}{n+1}\right)\sim n\log n.\tag{2.17b}$$

If $\operatorname{Re}\hat\mu(\gamma)>\frac12$,

$$E|S_n(\gamma)|^2=\frac{p_n(2\operatorname{Re}\hat\mu(\gamma))-p_n(1)}{2\operatorname{Re}\hat\mu(\gamma)-1}\sim\frac{1}{(2\operatorname{Re}\hat\mu(\gamma)-1)\Gamma(2\operatorname{Re}\hat\mu(\gamma))}n^{2\operatorname{Re}\hat\mu(\gamma)}.\tag{2.17c}$$

Thus, we see the phenomenon claimed in the introduction; $E|S_n(\gamma)|^2$ is roughly a constant times n if $\operatorname{Re}\hat\mu(\gamma)<\frac12$, but grows faster otherwise.

We proceed to prove that if $\operatorname{Re}\hat\mu(\gamma)\le\frac12$, the distribution of $S_n(\gamma)$ is asymptotically normal.

To begin with, there is a differential equation for $F_m(\gamma_1,\cdots,\gamma_m;z)$, $m\ge3$, corresponding to, and proved as, (2.6) and (2.13). For $m = 3$,

$$(1-z)\frac{d}{dz}F_3(\gamma_1, \gamma_2, \gamma_3; z)$$
$$= (1 + \hat{\mu}(\gamma_1) + \hat{\mu}(\gamma_2) + \hat{\mu}(\gamma_3))F_3(\gamma_1, \gamma_2, \gamma_3; z) + \hat{\mu}(\gamma_1\gamma_2)F_2(\gamma_1\gamma_2, \gamma_3; z) \tag{2.18}$$
$$+ \hat{\mu}(\gamma_1\gamma_3)F_2(\gamma_1\gamma_3, \gamma_2; z) + \hat{\mu}(\gamma_2\gamma_3)F_2(\gamma_1, \gamma_2\gamma_3; z) + \hat{\mu}(\gamma_1\gamma_2\gamma_3)F_1(\gamma_1\gamma_2\gamma_3; z).$$

In general, the right-hand side contains the homogeneous term $(1 + \hat{\mu}(\gamma_1) + \cdots + \hat{\mu}(\gamma_m))F_m(\gamma_1, \cdots, \gamma_m; z)$ and a linear combination of F_k's $(k < m)$ where a subset of $\gamma_1 \cdots \gamma_m$ has been multiplied together.

An easy induction proves the following lemma, cf. (2.14).

LEMMA 2.1. *$F_m(\gamma_1, \cdots, \gamma_m; z)$ is a linear combination of terms*

$$(1-z)^{-1-\alpha}(-\log(1-z))^{\beta}$$

where $\beta = 0, 1, \cdots$ and α equals $\sum_j \hat{\mu}(\prod_{i\in E_j} \gamma_i)$ for some partition $E_1, E_2, \cdots$ of $\{1 \cdots m\}$.

The next lemma, generalizing (2.8), is proved in [11, Theorem V.2.31].

LEMMA 2.2. *The Taylor coefficients of $(1-z)^{-1-\alpha}(-\log(1-z))^{\beta}$ are*

$$\sim (1/\Gamma(\alpha+1))n^{\alpha}(\log n)^{\beta}$$

as $n \to \infty$.

Hence, we can for large n ignore all terms in F_m but those for which Re α is maximal and β is maximal for these α's.

LEMMA 2.3. *If* Re $\hat{\mu}(\gamma_j) < 1/2$, $j = 1 \cdots m$, *then*

$$F_m(\gamma_1, \cdots, \gamma_m; z) = A(1-z)^{-1-m/2} + \textit{lower order terms} \tag{2.19}$$

where A equals $(m/2)!$ times the sum of $(1 - 2\,\mathrm{Re}\,\hat{\mu}(\gamma_{j_1}))^{-1} \cdot \cdots \cdot (1 - 2\,\mathrm{Re}\,\hat{\mu}(\gamma_{j_{m/2}}))^{-1}$ for all arrangements of $\gamma_1 \cdots \gamma_m$ into $m/2$ (unordered) pairs $(\gamma_{j_1}, \bar{\gamma}_{j_1}), (\gamma_{j_2}, \bar{\gamma}_{j_2}) \cdots$. (Thus, $A = 0$ when there is no such arrangement, e.g. when m is odd.) (The "lower order terms" are $(1-z)^{-1-\alpha}(-\log(1-z))^{\beta}$ with Re $\alpha < m/2$.)

PROOF. The homogeneous part of the differential equation like (2.18) is $(1-z)(dF/dz) = (1 + \sum_1^m \hat{\mu}(\gamma_j))F$ with the solution $C(1-z)^{-1-\sum\hat{\mu}(\gamma_j)}$, one of the lower order terms. The non-homogeneous terms also are of lower order, except possibly the leading terms of $\hat{\mu}(\gamma_i\gamma_j)F_{m-1}(\gamma_i\gamma_j, \cdots; z)$, $\gamma_i\gamma_j = \iota$. Now,

$$F_{m+1}(\iota, \gamma_1, \cdots, \gamma_m; z)$$
$$= \sum z^n E(S_n(\iota) \cdot \cdots \cdot S_n(\gamma_m)) = \sum z^n(n+1)E(S_n(\gamma_1) \cdot \cdots \cdot S_n(\gamma_m))$$
$$= \frac{d}{dz}(zF_m(\gamma_1, \cdots, \gamma_m; z))$$

and

$$\frac{d}{dz}z(1-z)^{-1-\alpha}(-\log(1-z))^{\beta} = (1+\alpha)(1-z)^{-2-\alpha}(-\log(1-z))^{\beta} + \text{lower order terms.}$$

Thus, by induction, each arrangement $(\gamma_{j_1}, \bar{\gamma}_{j_1}), (\gamma_{j_2}, \bar{\gamma}_{j_2}), \cdots$ contributes to $m/2$ terms on the right-hand side; its total contribution is $(1-z)^{-1-m/2}$ times

$$\sum_{k=1}^{m/2} \frac{m}{2}\left(\frac{m-2}{2}\right)! \prod_{i\neq k} (1-2\operatorname{Re}\hat{\mu}(\gamma_{j_i}))^{-1}$$

$$= \left(\frac{m}{2}\right)! \sum_1^{m/2} (1-2\operatorname{Re}\hat{\mu}(\gamma_{j_i})) \prod_1^{m/2} (1-2\operatorname{Re}\hat{\mu}(\gamma_{j_i}))^{-1}$$

$$= \left(\frac{m}{2} - \sum_1^m \hat{\mu}(\gamma_j)\right)\left(\frac{m}{2}\right)! \prod (1-2\operatorname{Re}\hat{\mu}(\gamma_{j_i}))^{-1}.$$

Hence, the leading terms on the right hand side sum to $(m/2 - \sum \hat{\mu}(\gamma_j))A(1-z)^{-1-m/2}$, which confirms the inductive hypotheses.

LEMMA 2.4. *If* $\operatorname{Re}\hat{\mu}(\gamma_j) = ½$, $j = 1 \cdots m$, *then*

(2.20) $$F_m(\gamma_1, \cdots, \gamma_m; z) = A(-\log(1-z))^{m/2}(1-z)^{-1-m/2} + \textit{lower order terms}$$

where A *equals* $(m/2)!$ *times the number of arrangements of* $\gamma_1 \cdots \gamma_m$ *into pairs* $(\gamma_{j_1}, \bar{\gamma}_{j_1}) \cdots$.

PROOF. Similar to Lemma 2.3.

THEOREM 2.2. *Suppose that* $\operatorname{Re}\hat{\mu}(\gamma) < ½$.

(i) *If* γ *is real,*

(2.21) $$n^{-1/2}S_n(\gamma) \to_d N(0, (1-2\hat{\mu}(\gamma))^{-1})$$

(ii) *If* γ *is complex,*

(2.22) $$n^{-1/2}(\operatorname{Re} S_n(\gamma), \operatorname{Im} S_n(\gamma)) \to_d N\left(0, \frac{1}{2}(1-2\operatorname{Re}\hat{\mu}(\gamma))^{-1}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right).$$

The joint distribution of several $n^{-1/2}S_n(\gamma_i)$ ($\operatorname{Re}\hat{\mu}(\gamma_i) < ½$) *converges to a normal distribution with independent components.*

PROOF. (i) By Lemmas 2.3 and 2.2,

(2.23) $$F_{2k}(\gamma, \cdots, \gamma; z) = k!\frac{(2k)!}{2^k k!}(1-2\hat{\mu}(\gamma))^{-k}(1-z)^{-1-k} + \cdots,$$

(2.24) $$E(S_n(\gamma))^{2k} \sim \frac{(2k)!}{2^k k!}(1-2\hat{\mu}(\gamma))^{-k}n^k$$

and

(2.25) $$E(S_n(\gamma))^{2k+1} = o(n^{k+1/2}).$$

Thus, the moments of $n^{-1/2}S_n(\gamma)$ converge to the moments of $N(0, (1-2\hat{\mu}(\gamma))^{-1})$, whence the distribution converges.

(ii) Similarly

(2.26) $$F_{2k}(\gamma, \cdots, \gamma, \bar{\gamma}, \cdots, \bar{\gamma}) = k!k!(1-2\operatorname{Re}\hat{\mu}(\gamma))^{-k}(1-z)^{-1-k} + \cdots \quad (k\ \gamma:s \text{ and } \bar{\gamma}:s).$$

Thus, $n^{-k}E((S_n(\gamma))^k(S_n(\bar{\gamma}))^k) \to k!(1-2\operatorname{Re}\hat{\mu}(\gamma))^{-k}$, while all other mixed moments of $n^{-1/2}(S_n(\gamma), S_n(\bar{\gamma}))$ tend to zero. Thus, the moments of $n^{-1/2}(\operatorname{Re} S_n(\gamma), \operatorname{Im} S_n(\gamma))$ converge to the moments of the given normal distribution.

Lemma 2.3 also shows that mixed moments of $S_n(\gamma_1)$, $S_n(\gamma_2)$, $\cdots$ converge to the corresponding products of moments, which proves the statement for joint distributions.

THEOREM 2.3. *Suppose that* $\operatorname{Re}\hat{\mu}(\gamma) = ½$.

(i) *If* γ *is real,* $(n \log n)^{-1/2} S_n(\gamma) \to_d N(0, 1)$.

(ii) *If* γ *is complex,* $(n \log n)^{-1/2}(\operatorname{Re} S_n(\gamma), \operatorname{Im} S_n(\gamma)) \to_d N\left(0, \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right)$.

The joint distribution of several $(n \log n)^{-1/2}S_n(\gamma_i)$ $(\operatorname{Re} \hat{\mu}(\gamma_i) = \frac{1}{2})$ *converges to a normal distribution with independent components.*

PROOF. As for Theorem 2.2, using Lemma 2.4.

Let us finally briefly consider the case $\operatorname{Re} \hat{\mu}(\gamma) > \frac{1}{2}$. The leading term of $F_m(\gamma, \cdots, \gamma; z)$ now is $C_m(1 - z)^{-1-m\hat{\mu}(\gamma)}$. Hence, all moments of $n^{-\hat{\mu}(\gamma)}S_n(\gamma)$ converge, but there are several complications: Unless $\hat{\mu}(\gamma)$ is real or $\hat{\nu}(\gamma) = 0$, $En^{-\operatorname{Re}\hat{\mu}(\gamma)}S_n(\gamma)$ will oscillate, and $n^{-\operatorname{Re}\hat{\mu}(\gamma)}(S_n(\operatorname{Re} \gamma), S_n(\operatorname{Im} \gamma))$ does not converge. The limits of the moments depend on $\hat{\mu}(\gamma), \hat{\mu}(\gamma^2), \cdots$ and on $\hat{\nu}(\gamma), \hat{\nu}(\gamma^2), \cdots$. Different $S_n(\gamma)$ do not have to be asymptotically uncorrelated.

If γ is real ($\gamma^2 = \iota$), crude estimates show that the sequence of limits of the moments determines a unique distribution. Hence, $n^{-\hat{\mu}(\gamma)}S_n(\gamma)$ converges in distribution. A straightforward computation of $F_1 \cdots F_4$ shows that the first four moments of the limiting distribution are (with $\hat{\mu} = \hat{\mu}(\gamma)$)

$$\frac{\hat{\nu}(\gamma)}{\Gamma(\hat{\mu}+1)}, \quad \frac{2\hat{\mu}}{(2\hat{\mu}-1)\Gamma(2\hat{\mu}+1)}, \quad \frac{3\hat{\nu}(\gamma)(1+\hat{\mu})}{(2\hat{\mu}-1)\Gamma(3\hat{\mu}+1)} \quad \text{and} \tag{2.27}$$
$$\frac{24\hat{\mu}(2\hat{\mu}^2+2\hat{\mu}-1)}{(2\hat{\mu}-1)^2(4\hat{\mu}-1)\Gamma(4\hat{\mu}+1)}.$$

No simple pattern is seen and we have not been able to identify the asymptotic distribution.

3. Commutative groups: general functions. Throughout this section we assume that $\operatorname{Re} \hat{\mu}(\gamma) \le \frac{1}{2}$ for every $\gamma \ne \iota$.

We will study $S_n(f)$ for a real function $f \in L^2(dm)$. If the Fourier series $f = \sum_\Gamma \hat{f}(\gamma)\gamma$ happens to have only a finite number of non-zero terms, $S_n(f)$ is asymptotically normally distributed by Theorems 2.2 and 2.3. When G is finite, every f is such a finite sum, but when G is infinite, some further argument is needed. It is clear that a convergence theorem requires some extra smoothness condition on f, μ or ν. One reason is that the distribution of X_n may be singular with respect to m while f may be defined only a.e. Also, it is possible to have $\operatorname{Re}\hat{\mu}(\gamma) < \frac{1}{2}$ for every $\gamma \ne \iota$, but $\sup \operatorname{Re} \hat{\mu}(\gamma) = \frac{1}{2}$. Then σ^2 in Theorem 3.1 below may be infinite even for a continuous f.

LEMMA 3.1. (i) *Suppose that* $\operatorname{Re} \hat{\mu}(\gamma_1)$, $\operatorname{Re} \hat{\mu}(\gamma_2) < \frac{1}{2}$. *Then*

$$n^{-1}|ES_n(\gamma_1)S_n(\gamma_2)| \le C|\nu(\gamma_1\gamma_2)|(1 - \operatorname{Re} \hat{\mu}(\gamma_1) - \operatorname{Re} \hat{\mu}(\gamma_2))^{-1}. \tag{3.1}$$

(ii) *Suppose that* $\operatorname{Re} \hat{\mu}(\gamma_1)$, $\operatorname{Re} \hat{\mu}(\gamma_2) \le \frac{1}{2}$. *Then*

$$(n \log n)^{-1}|ES_n(\gamma_1)S_n(\gamma_2)| \le C|\hat{\nu}(\gamma_1\gamma_2)| \quad (n \ge 2). \tag{3.2}$$

PROOF. By (2.16),

$$|ES_n(\gamma_1)S_n(\gamma_2)| \le |\hat{\nu}(\gamma_1\gamma_2)|\sup_{0\le t\le 1}|p_n'((1-t)\hat{\mu}(\gamma_1\gamma_2) + t(\hat{\mu}(\gamma_1) + \hat{\mu}(\gamma_2)))|. \tag{3.3}$$

If $|x| \le 2$ and $\operatorname{Re} x \ge 0$ we obtain (for $n \ge 2$)

$$\left|\frac{p_n(x)}{x}\right|^2 = \prod_1^n \left|1 + \frac{x}{k}\right|^2 = \prod_1^n \left(1 + \frac{2\operatorname{Re} x}{k} + \frac{|x|^2}{k^2}\right) \le \prod_1^n \left(1 + \frac{4}{k^2}\right)\left(1 + \frac{2\operatorname{Re} x}{k}\right)$$
$$\le C\prod_1^n \left(1 + \frac{1}{k}\right)^{2\operatorname{Re} x} = C(n+1)^{2\operatorname{Re} x}$$

(3.4) $$|p_n(x)| \le C|x|n^{\operatorname{Re} x}$$

(3.5) $$|p_n'(x)| = \left| p_n(x) \sum_0^n \frac{1}{x+k} \right| \le |p_n(x)| \sum_0^n \frac{1}{|x+k|} \le |p_n(x)| \left(\frac{1}{|x|} + \sum_1^n \frac{1}{k} \right) \le C \log n n^{\operatorname{Re} x}.$$

These estimates hold also for $\operatorname{Re} x < 0$, $|x| \le 2$. This follows e.g. by the identity $p_n(x) = (x/n)p_{n-1}(x+1)$.

(3.2) follows from (3.3) and (3.5). To prove (3.1), let $\alpha = 1 - \operatorname{Re} \hat{\mu}(\gamma_1) - \operatorname{Re} \hat{\mu}(\gamma_2)$. If $\operatorname{Re} \hat{\mu}(\gamma_1\gamma_2) \le 1 - \alpha/2$, (3.3) and (3.5) yield

$$n^{-1}E|S_n(\gamma_1)S_n(\gamma_2)| \le C|\hat{\nu}(\gamma_1\gamma_2)| \log n n^{-\alpha/2} \le C|\hat{\nu}(\gamma_1\gamma_2)|\alpha^{-1},$$

and if $\operatorname{Re} \hat{\mu}(\gamma_1\gamma_2) > 1 - \alpha/2$, (2.16a) and (3.4) yield

$$E|S_n(\gamma_1)S_n(\gamma_2)| \le C|\hat{\nu}(\gamma_1\gamma_2)|(n^{\operatorname{Re}\hat{\mu}(\gamma_1\gamma_2)} + n^{\operatorname{Re}\hat{\mu}(\gamma_1)+\operatorname{Re}\hat{\mu}(\gamma_2)})/(\alpha/2) \le C|\hat{\nu}(\gamma_1\gamma_2)|n/\alpha.$$

THEOREM 3.1. *Suppose that* $\operatorname{Re} \hat{\mu}(\gamma) < \frac{1}{2}$ *for* $\gamma \neq \iota$ *and that* $f \in L^2(m)$ *is real. Let* $\sigma^2 = \sum_{\gamma\neq\iota} |\hat{f}(\gamma)|^2(1 - 2\operatorname{Re} \hat{\mu}(\gamma))^{-1}$. *Suppose further that either*

(i) $\sum|\hat{f}(\gamma)|(1 - 2\operatorname{Re}\hat{\mu}(\gamma))^{-1/2} < \infty$, *or*

(ii) $\sigma^2 < \infty$ *and* ν *is absolutely continuous with* $d\nu/dm$ *bounded, or*

(iii) μ *is absolutely continuous with* $d\mu/dm$ *bounded.*

Then

(3.6) $$n^{-1/2}\left(S_n(f) - n\int f\,dm \right) \to_d N(0, \sigma^2).$$

PROOF. By subtracting a constant, we may assume that $\hat{f}(\iota) = \int f\,dm = 0$. Denote the characters such that $\hat{f}(\gamma) \neq 0$ by $\gamma_1, \gamma_2, \cdots$ and let f_N be the partial sum $\sum_1^N \hat{f}(\gamma_i)\gamma_i$. We assume that every complex γ_i is immediately followed (or preceded) by its conjugate. Then, restricting attention to a subsequence, we may assume that f_N is real. By Theorem 2.2,

$$n^{-1/2}S_n(f_N) \to_d N(0, \sigma_N^2) \text{ as } n \to \infty,$$

where $\sigma_N^2 = \sum_1^N |\hat{f}(\gamma_i)|^2(1 - 2\operatorname{Re}\hat{\mu}(\gamma_i))^{-1}$.

Since $\sigma_N^2 \to \sigma^2$ and thus $N(0, \sigma_N^2) \to N(0, \sigma^2)$ as $N \to \infty$, the theorem will follow from [2, Theorem 4.2] once we have showed that $n^{-1/2}S_n(f_N)$ converges in probability uniformly in n to $n^{-1/2}S_n(f)$ as $N \to \infty$.

We study the three cases separately.

(i) $f_N \to f$ uniformly and thus $S_n(f_N)$ converges to $S_n(f)$ in square mean. By Lemma 3.1 (i), $E|S_n(\gamma)|^2 \le Cn(1 - 2\operatorname{Re}\hat{\mu}(\gamma))^{-1}$, whence

(3.7) $$\|S_n(f)\|_2 = \|\sum \hat{f}(\gamma_i)S_n(\gamma_i)\|_2 \le \sum|\hat{f}(\gamma_i)|\,\|S_n(\gamma_i)\|_2 \le Cn^{1/2}\sum|\hat{f}(\gamma_i)|(1 - 2\operatorname{Re}\hat{\mu}(\gamma_i))^{-1/2}.$$

Applying this to $f - f_N$ we obtain

$$\sup_n \|n^{-1/2}S_n(f) - n^{-1/2}S_n(f_N)\|_2 = \sup_n n^{-1/2}\|S_n(f - f_N)\|_2 \le C\sum_{N+1}^\infty |\hat{f}(\gamma_i)|(1 - 2\operatorname{Re}\hat{\mu}(\gamma_i))^{-1/2}.$$

Since the last sum tends to zero as $N \to \infty$, we have proved the sought uniform convergence in square mean, and hence in probability.

(ii) First we assume that $\nu = m$. Then, every X_n is uniformly distributed and $S_n(f_N)$

converges to $S_n(f)$ in square mean as $N \to \infty$. By Lemma 3.1 (i),

$$E|S_n(f)|^2 = \sum_\Gamma \sum_\Gamma \hat{f}(\gamma)\hat{f}(\gamma')ES_n(\gamma)S_n(\gamma') = \sum_\Gamma |\hat{f}(\gamma)|^2 ES_n(\gamma)S_n(\bar{\gamma}) \le Cn \sum_\Gamma |\hat{f}(\gamma)|^2(1 - 2\,\mathrm{Re}\,\hat{\mu}(\gamma))^{-1} = Cn\sigma^2. \tag{3.8}$$

For an arbitrary ν satisfying the condition, this yields

$$E|S_n(f)|^2 = \int E(|S_n(f)|^2|X_0)\,d\nu(X_0) \le \sup\frac{d\nu}{dm}\int E(|S_n(f)|^2|X_0)\,dm(X_0) \le Cn\sup\frac{d\nu}{dm}\sigma^2. \tag{3.9}$$

Again, we apply this to $f - f_N$.

$$\sup_n E(n^{-1/2}S_n(f) - n^{-1/2}S_n(f_N))^2 \le C\sup\frac{d\nu}{dm}(\sigma^2 - \sigma_N^2) \to 0 \text{ as } N \to \infty.$$

This completes the proof as in case (i).

(iii) Since $\mu \in L^2$, $\hat{\mu} \in l^2$ and $\sup_{\gamma\neq 1}\mathrm{Re}\,\hat{\mu}(\gamma) < ½$, which implies that $\sigma^2 < \infty$.

We may, by conditioning on X_0, assume that X_0 is constant. We separate $X_1 \cdots X_n$ into different branches, each one comprising the progeny of a single daughter of X_0. Let B_n be the number of branches and $N_1, N_2, \cdots, N_{B_n}$ the numbers of elements in the branches. Thus $\sum_1^{B_n} N_j = n$. Also $EB_n = E \# \{k \le n : I_k = 0\} = \sum_1^n (1/k) \le \log n + 1$ and

$$EB_n^2 = (EB_n)^2 + \sum_1^n \frac{1}{k}\left(1 - \frac{1}{k}\right) \le 2(\log n + 1)^2. \tag{3.10}$$

The branches develop independently of each other and each branch is a process of the type studied in this paper with initial distribution $\nu' = X_0\mu$. Note that ν' satisfies (ii).

Let $S_m^{(j)}(f)$ be the sum of $f(X)$ for the first m points of the jth branch, and $a_m = ES_m^{(j)}(f)$. Thus $S_n(f) = f(X_0) + \sum_1^{B_n} S_{N_j}^{(j)}(f)$. By (3.9).

$$E(S_m^{(j)}(f) - a_m)^2 \le ES_m^{(j)}(f)^2 \le Cm\sup\frac{d\nu'}{dm}\sigma^2. \tag{3.11}$$

Since the branches are independent,

$$E((\sum_1^{B_n}(S_{N_j}^{(j)} - a_{N_j}))^2 | B_n, N_1 \cdots N_{B_n}) = \sum_1^{B_n} E((S_{N_j}^{(j)} - a_{N_j})^2 | N_j) \le C\sum_1^{B_n} N_j \sup\frac{d\nu'}{dm}\sigma^2 = Cn\sup\frac{d\mu}{dm}\sigma^2. \tag{3.12}$$

Put $\kappa = \sup\{0, \mathrm{Re}\,\hat{\mu}(\gamma) : \gamma \neq 1\}$. By (2.9) and Plancherel's formula

$$|a_m| = |\sum \hat{f}(\gamma)ES_m^{(j)}(\gamma)| = \left|\sum f(\gamma)\hat{\nu}'(\gamma)\binom{m - 1 + \hat{\mu}(\gamma)}{m-1}\right| \le C\sum|\hat{f}(\gamma)\hat{\nu}'(\gamma)|m^{\mathrm{Re}\hat{\mu}(\gamma)} \le Cm^\kappa\|f\|\,\|d\hat{\nu}'/dm\|$$

($\|\ \|$ denotes the norm in L^2.) Thus,

$$|\sum_1^{B_n} a_{N_j}| \le C\|f\|\ \|d\mu/dm\|\sum_1^{B_n} N_j^\kappa. \tag{3.13}$$

By Hölder's inequality,

$$\sum_1^{B_n} N_j^\kappa \le (\sum_1^{B_n} N_j)^\kappa(\sum_1^{B_n} 1)^{1-\kappa} = n^\kappa B_n^{1-\kappa} \tag{3.14}$$

and, by (3.10),

$$E(\sum_1^{B_n} N_j^\kappa)^2 \le n^{2\kappa}EB_n^{2(1-\kappa)} \le n^{2\kappa}(EB_n^2)^{1-\kappa} \le 2n^{2\kappa}(1 + \log n)^{2-2\kappa}. \tag{3.15}$$

By (3.13) and (3.15), $n^{-1}E\,|\sum_1^{B_n} a_{N_j}|^2 \le C\|f\|^2$, where C depends on μ only. We combine this with (3.12) and obtain

$$n^{-1}E(S_n(f) - f(X_0))^2 = n^{-1}E(\textstyle\sum_1^{B_n} S_{N_j}^{(j)})^2 \le C\sigma^2.$$

The same argument as for the other cases shows that $n^{-1/2}(S_n(f) - f(X_0)) \to_d N(0, \sigma^2)$ as $n \to \infty$. Since $n^{-1/2}f(X_0) \to_p 0$, this completes the proof.

REMARKS.
1. It follows from the proof that $ES_n(f) = n\int f + o(n^{-1/2})$ and $\operatorname{Var} S_n(f) = n\sigma^2 + o(n)$. (In case (iii) this requires the extra assumption $E(f(X_0))^2 < \infty$.) Similar remarks apply to later theorems.
2. If $\sup_{\gamma\neq\iota}\operatorname{Re}\hat\mu(\gamma) < \frac{1}{2}$, then $\sigma^2 < \infty$ and (i) reduces to $\sum|\hat f(\gamma)| < \infty$. On the unit circle, this holds e.g. if f has a bounded derivative.

If some $\operatorname{Re}\hat\mu(\gamma) = \frac{1}{2}$ we have the following substitute.

THEOREM 3.2. *Suppose that* $\operatorname{Re}\hat\mu(\gamma) \le \frac{1}{2}$ *for* $\gamma \neq 1$ *and that* $f \in L^2(m)$ *is real. Suppose further that either*
(i) $\sum|\hat f(\gamma)| < \infty$, *or*
(ii) ν *is absolutely continuous with* $d\nu/dm$ *bounded, or*
(iii) μ *is absolutely continuous with* $d\mu/dm$ *bounded.*
Then

$$(n\log n)^{-1/2}\left(S_n(f) - \int f\,dm\right) \to_d N(0, \textstyle\sum_{\operatorname{Re}\hat\mu(\gamma)=1/2}|\hat f(\gamma)|^2).$$

PROOF. As for Theorem 3.1, using Lemma 3.1 (ii).

Note that only $\hat f(\gamma)$ for γ such that $\operatorname{Re}\hat\mu(\gamma) = \frac{1}{2}$ matter. Thus, even if $f_1 \cdots f_m$ are linearly independent, $(n\log n)^{-1}(S_n(f_1), \cdots, S_n(f_m))$ may converge to a degenerate distribution. In fact, the rank is at most the number of $\gamma \in \Gamma$ such that $\operatorname{Re}\hat\mu(\gamma) = \frac{1}{2}$. (This number is often, but not always, finite.) This does not happen when $\operatorname{Re}\hat\mu(\gamma) < \frac{1}{2}$ for every $\gamma \neq \iota$.

4. General compact groups. In this section we adapt the preceding arguments to non-commutative groups. Instead of characters, we will work with the more general concept of group representations. The following facts on representations of compact groups may be found e.g. in [8]. See also [4].

A representation of G is a continuous homomorphism of G into the group of unitary operators on some finite-dimensional complex vector space V. We will not distinguish between an operator and the corresponding matrix for some fixed orthonormal basis. Thus, we may equivalently say that a representation is a continuous matrix valued function $R(g) = (r_{ij}(g))_{i,j=1}^d$ such that $R(g)$ is unitary and $R(g_1g_2) = R(g_1)R(g_2)$, $g_1, g_2 \in G$. Thus $R(g^{-1}) = R(g)^{-1} = R(g)^*$.

We define the conjugate representation as $\bar R(g) = (\overline{r_{ij}}(g))_{i,j=1}^d$.

Two representations are equivalent if they differ by a change of coordinates only, i.e. $R_2(g) = UR_1(g)U^{-1}$ for some fixed unitary matrix U. Equivalent representations can be considered identical for all purposes. A simple but important example is the trivial representation in a one-dimensional space defined by $R(g) = I$, the identity operator.

Given two representations R_1 and R_2 in V_1 and V_2 of dimensions d_1 and d_2, respectively, we may form new representations $R_1 \oplus R_2$ in the $d_1 + d_2$-dimensional space $V_1 \oplus V_2$ and $R_1 \otimes R_2$ in the d_1d_2-dimensional space $V_1 \otimes V_2$. In coordinate form

$$(R_1 \otimes R_2(g))_{ikjl} = r_{ij}^{(1)}(g)r_{kl}^{(2)}(g). \tag{4.1}$$

A representation is irreducible if the space V does not contain any proper subspace, W, which is invariant for every $R(g)$, e.g. $R(g)(W) \subset W$. Every representation is equivalent to a direct sum of irreducible representations; equivalently V is a direct sum of some invariant subspaces V_i such that $R|_{V_i}$ is irreducible. (If G is commutative, the irreducible representations are one-dimensional and given by the characters, cf. Section 2.)

For a representation R in V, we define $\hat{f}(R) = \int f(g)R(g)^*\, dm(g)$ (for $f \in L^1(m)$) and $\hat{\mu}(R) = \int R(g)\, d\mu(g)$ (for a measure μ); these are operators on V.

We let $\{R_\alpha\}_{\alpha\in\hat{G}} = \{(r_{ij}^\alpha)\}$ be a complete set of irreducible representation (i.e. every irreducible representation is equivalent to exactly one R_α), and let d_α be the dimension of R_α. Then $\cup_\alpha\{d_\alpha^{1/2} r_{ij}^\alpha\}_{i,j=1}^{d_\alpha}$ is an orthonormal basis in $L^2(m)$. Thus, we have the Fourier series expansion

$$\begin{aligned} f(g) &= \textstyle\sum_\alpha \sum_{ij} d_\alpha \langle f, r_{ij}^\alpha\rangle r_{ij}^\alpha(g) = \sum_\alpha d_\alpha \sum_{ij} \hat{f}(R^\alpha)_{ji} r_{ij}^\alpha(g) \\ &= \textstyle\sum_\alpha d_\alpha \mathrm{Tr}(\hat{f}(R_\alpha)R_\alpha(g)) \end{aligned} \tag{4.2}$$

and the Plancherel formula

$$\|f\|^2 = \textstyle\sum_\alpha d_\alpha \|\hat{f}(R_\alpha)\|_{\mathrm{HS}}^2, \tag{4.3}$$

where $\|f\|$ denotes the normal in $L^2(m)$ and $\|\ \|_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm, $\|(a_{ij})\|_{\mathrm{HS}}^2 = \sum_{ij}|a_{ij}|^2$.

Returning to the branching random walk, we define for each representation R, $S_n(R) = \sum_0^n R(X_k)$. This is a matrix-valued stochastic variable with components $(S_n(R))_{ij} = S_n(r_{ij})$.

For representations $R_1 \cdots R_m$ in $V_1 \cdots V_m$, respectively, we define

$$F_m(R_1, \cdots, R_m; z) = \textstyle\sum_0^\infty z^n E(S_n(R_1) \otimes S_n(R_2) \otimes \cdots \otimes S_n(R_m)). \tag{4.4}$$

This generating function is an analytic function on $\{z : |z| < 1\}$ taking values in the space of linear operators on $V_1 \otimes \cdots \otimes V_m$. Its components are generating functions for mixed moments $E(S_n(r_{i_1j_1}^{(1)})S_n(r_{i_2j_2}^{(2)}) \cdots)$ (cf. (4.1)). By the definition,

$$F_m(R_1, \cdots, R_m; 0) = E(R_1(X_0) \otimes \cdots \otimes R_m(X_0)) = \hat{\nu}(R_1 \otimes \cdots \otimes R_m). \tag{4.5}$$

We obtain differential equations for F as in the commutative case,

$$\begin{aligned} E(R(X_n)|X_0 \cdots X_{n-1}) &= E(R(X_{I_n})R(Y_n)|X_0 \cdots X_{n-1}) \\ &= \frac{1}{n} S_{n-1}(R)ER(Y_n) = \frac{1}{n} S_{n-1}(R)\hat{\mu}(R) \end{aligned} \tag{4.6}$$

$$E(S_{n+1}(R)) = ES_n(R)\left(I + \frac{1}{n+1}\hat{\mu}(R)\right) \tag{4.7}$$

$$\begin{aligned} (1 - z)\frac{d}{dz}F_1(R; z) &= \textstyle\sum z^n((n + 1)ES_{n+1}(R) - nES_n(R)) \\ &= F_1(R; z)(I + \hat{\mu}(R)). \end{aligned} \tag{4.8}$$

We will soon solve this equation. Formally, $F_1(R; z) = \hat{\nu}(R)(1 - z)^{-I-\hat{\mu}(R)}$. Similarly, we obtain for F_2,

$$\begin{aligned} S_{n+1}(R_1) \otimes S_{n+1}(R_2) &= S_n(R_1) \otimes S_n(R_2) + S_n(R_1) \otimes R_2(X_{n+1}) \\ &\quad + R_1(X_{n+1}) \otimes S_n(R_2) + R_1 \otimes R_2(X_{n+1}) \end{aligned}$$

$$\begin{aligned} E(S_{n+1}(R_1) \otimes S_{n+1}(R_2)) &= E(S_n(R_1) \otimes S_n(R_2))\left(I + \frac{1}{n+1} I \otimes \hat{\mu}(R_2) + \frac{1}{n+1}\hat{\mu}(R_1) \otimes I\right) \\ &\quad + \frac{1}{n+1} ES_n(R_1 \otimes R_2)\hat{\mu}(R_1 \otimes R_2) \end{aligned} \tag{4.9}$$

$$(4.10)\quad (1-z)\frac{d}{dz}F_2(R_1, R_2; z) = F_2(R_1, R_2; z)(I + \hat{\mu}(R_1)\otimes I + I\otimes \hat{\mu}(R_2)) + F_1(R_1\otimes R_2; z)\hat{\mu}(R_1\otimes R_2).$$

For $m \geq 2$ we obtain similar equations. We give the differential equation for F_3 in coordinate form, cf. (2.18).

$$(4.11)\quad \begin{aligned}(1-z)&\frac{d}{dz}F_3(R_1, R_2, R_3; z)_{i_1i_2i_3j_1j_2j_3}\\ &= (F_3(R_1, R_2, R_3; z)(I + \hat{\mu}(R_1)\otimes I\otimes I + I\otimes\hat{\mu}(R_2)\otimes I + I\otimes I\otimes\hat{\mu}(R_3)))_{i_1i_2i_3j_1j_2j_3}\\ &+ (F_2(R_1\otimes R_2, R_3; z)(\hat{\mu}(R_1\otimes R_2)\otimes I))_{i_1i_2i_3j_1j_2j_3}\\ &+ (F_2(R_1\otimes R_3, R_2; z)(\hat{\mu}(R_1\otimes R_3)\otimes I))_{i_1i_3i_2j_1j_3j_2}\\ &+ (F_2(R_2\otimes R_3, R_1; z)(\hat{\mu}(R_2\otimes R_3)\otimes I))_{i_2i_3i_1j_2j_3j_1}\\ &+ (F_1(R_1\otimes R_2\otimes R_3; z)\hat{\mu}(R_1\otimes R_2\otimes R_3))_{i_1i_2i_3j_1j_2j_3}.\end{aligned}$$

To solve these equations we will use the following facts on Jordan's normal form of a matrix, see e.g. [7].

Let A be a linear operator. Then there is a basis (not necessarily orthogonal) such that A decomposes as a direct sum of Jordan boxes, i.e. matrices $(\lambda\delta_{ij} + \delta_{ij-1})$. The λ's are eigenvalues of A and each eigenvalue occurs in at least one Jordan box.

LEMMA 4.1. *Assume that A is a matrix with eigenvalues $\lambda_1 \cdots \lambda_l$ and let $d_1 \cdots d_l$ be the dimensions of the largest corresponding Jordan boxes.*

(i) *Any solution of the homogeneous equation*

$$(4.12)\quad (1-z)\frac{d}{dz}F = FA$$

is of the form

$$(4.13)\quad F = \textstyle\sum_{k=1}^{l}\sum_{j=1}^{d_k-1} A_{kj}(1-z)^{-\lambda_k}(-\log(1-z))^j.$$

(ii) *The non-homogeneous equation*

$$(4.14)\quad (1-z)\frac{d}{dz}F = FA + B(1-z)^{-\alpha}(-\log(1-z))^m \qquad (m = 0, 1, \cdots)$$

has a solution

$$(4.15)\quad F = \textstyle\sum_{j=0}^{M} BA_j(1-z)^{-\alpha}(-\log(1-z))^j,$$

where

(a) *if α is not an eigenvalue of A, then $M = m$ and $A_m = (\alpha I - A)^{-1}$,*

(b) *if $\alpha = \lambda_k$, then $M = m + d_k$ and $A_{m+n} = (m!/(m+n)!)(A - \alpha I)^{n-1}P_k$, $n = 1, \cdots d_k$, where P_k is the projection (defined by the Jordan decomposition) onto* $\ker(A - \alpha I)^{d_k}$.

PROOF. (i) Choose a basis such that A is of Jordan form. The equation decomposes correspondingly. Hence it suffices to prove (4.13) when A is a Jordan box. Furthermore, since (4.12) is equivalent to $(1-z)(d/dz)(F(1-z)^a) = F(1-z)^a(A - aI)$, we may assume that the eigenvalue is zero. Thus $A = (\delta_{ij-1})_{i,j=1}^{d}$ and (4.12) becomes $(1-z)\, dF_{ij}/dz = F_{ij-1}$, $i, j = 1 \cdots d$, or with $x = -\log(1-z)$, $dF_{ij}/dx = F_{ij-1}$, which implies that F is of the stated form.

(ii) It suffices to prove this for $B = I$. (The resulting equation may be left-multiplied by B.) As in part (i), we may assume that $A = (\delta_{ij-1})_{i,j=1}^{d}$. Then $(d/dx)F_{ij} = F_{ij-1} + e^{\alpha x}x^m\delta_{ij}$, $i, j = 1 \cdots d$.

If $\alpha \neq 0$, the equation may be written $(d/dx)(F + (A - \alpha)^{-1}e^{\alpha x}x^m) = (F + (A - \alpha)^{-1}e^{\alpha x}x^m)A + m(A - \alpha)^{-1}e^{\alpha x}x^{m-1}$. (a) follows by induction.

If a = 0 then one solution is $F_{ij} = (m!/(m + j - i + 1)!)x^{m+j-i+1}, j \geq i$. This proves (b).

We will in the sequel denote the eigenvalues of $\hat{\mu}(R)$ by $\lambda_k(R)$, $k = 1, 2, \cdots$, and the dimensions of the largest corresponding Jordan boxes by $d_k(R)$. We assume that they are ordered so that $\operatorname{Re} \lambda_1(R) \geq \operatorname{Re} \lambda_2(R) \geq \cdots$.

Lemma (4.1) (i) and (4.8) yield

$$F_1(R; z) = \sum_k \sum_{j=0}^{d_k(R)-1} A_{kj}(1 - z)^{-1-\lambda_k(R)}(-\log(1 - z))^{-j} \tag{4.16}$$

for some A_{kj}. This gives, using Lemma 2.2, the asymptotic behaviour of $ES_n(R)$. If we for simplicity assume that $\operatorname{Re} \lambda_1(R) > \operatorname{Re} \lambda_2(R)$ we obtain

$$ES_n(R) \sim n^{\lambda_1(R)}(\log n)^{d_1(R)-1}A,$$

for some matrix A. If $\operatorname{Re} \lambda_1(R) > ½$, it follows from the lemma and (4.10) that the leading term of $F_2(R, \bar{R}; z)$ is

$$B(1 - z)^{-1-2\operatorname{Re}\lambda_1(R)}(-\log(1 - z))^{2d_1(R)-2}$$

and thus

$$ES_n(R) \otimes S_n(\bar{R}) \sim n^{2\operatorname{Re}\lambda_1(R)}(\log n)^{2d_1(R)-2} \cdot B_1,$$

for some matrices B and B_1, which furthermore may be shown to be nonzero. We obtain the same misbehaviour as in the commutative case.

Another consequence of (4.16) is that Theorem 2.1 holds verbatim for non-commutative groups, the proof being the same except for simple modifications. (The assumption on μ is equivalent to $\operatorname{Re} \lambda_1(R) < 1$ for every non-trivial irreducible representation.)

Now, we are prepared to study the well-behaved case in detail. Assume that R_1 and R_2 are two representations such that $\operatorname{Re} \lambda_1(R_1)$, $\operatorname{Re} \lambda_1(R_2) < ½$, $k = 1, \cdots$. Let P be the orthogonal projection of $V_1 \otimes V_2$ onto $\{x \in V_1 \otimes V_2 : R_1 \otimes R_2 x = x\}$. Since $R_1 \otimes R_2$ restricted to ker P decomposes into a direct sum of non-trivial irreducible representations,

$$\hat{m}(R_1 \otimes R_2) = P \tag{4.17}$$

and

$$P\hat{\mu}(R_1 \otimes R_2) = P\hat{\nu}(R_1 \otimes R_2) = P. \tag{4.18}$$

Using (4.8), Lemma 4.1 (i) and the decomposition $V_1 \otimes V_2 = \operatorname{im} P \oplus \ker P$ we obtain

$$F_1(R_1 \otimes R_2; z) = P(1 - z)^{-2} + \text{lower order terms.} \tag{4.19}$$

Since the eigenvalues of $I + \hat{\mu}(R_1) \otimes I + I \otimes \hat{\mu}(R_2)$ are $\{1 + \lambda_k(R_1) + \lambda_l(R_2)\}$ which by assumption have real parts smaller than 2, (4.10), (4.18), (4.19) and Lemma 4.1 yield

$$F_2(R_1, R_2; z) = P(I - \hat{\mu}(R_1) \otimes I - I \otimes \hat{\mu}(R_2))^{-1}(1 - z)^{-2} + \text{lower order terms.} \tag{4.20}$$

We specialize to the case $R_1 = \bar{R}_2 = R$ and define

$$\Sigma(R) = (\sigma_{ikjl}(R)) = P(I - \hat{\mu}(R) \otimes I - I \otimes \hat{\mu}(\bar{R}))^{-1}. \tag{4.21}$$

Then (4.20) implies

$$ES_n(R) \otimes S_n(\bar{R}) = \Sigma(R)n + o(n). \tag{4.22}$$

This can be written

$$ES_n(r_{ij})S_n(\bar{r}_{kl})/n \to \sigma_{ikjl}(R), \quad n \to \infty. \tag{4.23}$$

If furthermore R is an irreducible representation in the space V with the orthonormal basis $e_1 \cdots e_d$, then Im P is the one-dimensional subspace of $V \otimes V$ spanned by $e_1 \otimes e_1$

$+ \cdots + e_d \otimes e_d$, and thus

$$P = \left(\frac{1}{d}\,\delta_{ik}\delta_{jl}\right)_{i,k,j,l=1}^{d}. \tag{4.24}$$

(This follows by the theory of representations, or from the special case $\mu = m$ in (4.23).) Hence

$$\begin{aligned}\sigma_{ikjl}(R) &= \langle \Sigma\,(R) e_j \otimes e_l,\, e_i \otimes e_k \rangle \\ &= \frac{1}{d}\,\delta_{ik} \textstyle\sum_{m=1}^{d} \langle (I - \hat\mu(R) \otimes I - I \otimes \hat\mu(\bar R))^{-1} e_j \otimes e_l,\, e_m \otimes e_m \rangle.\end{aligned} \tag{4.25}$$

If $\hat\mu(R)$ is diagonal, with diagonal entries $\lambda_1 \cdots \lambda_m$, this simplifies to

$$\begin{aligned}\sigma_{ikjl}(R) &= \frac{1}{d}\,\delta_{ik} \textstyle\sum_{m=1}^{d} \langle (1 - \lambda_j - \bar\lambda_l)^{-1} e_j \otimes e_l,\, e_m \otimes e_m \rangle \\ &= \frac{1}{d}\,(1 - 2\,\mathrm{Re}\,\lambda_j)^{-1} \delta_{ik}\delta_{jl}.\end{aligned} \tag{4.26}$$

Similarly, we obtain for higher moment, using the same argument as in Lemma 2.3, the following result.

LEMMA 4.2. *If* $\mathrm{Re}\,\lambda_1(R) < 1/2$ *and* $R = \bar R$ *(i.e. every* r_{ij} *is real), then*

$$\begin{aligned}F_m\,(R, R, \cdots, R; z)_{i_1 \cdots i_m j_1 \cdots j_m} &= \left(\frac{m}{2}\right)!\, \textstyle\sum \prod_{k=1}^{m/2} \sigma_{i'_k i''_k j'_k j''_k} (1 - z)^{-1-m/2} \\ &\quad + \textit{lower order terms},\end{aligned} \tag{4.27}$$

where the sum is taken over all partitions of $(1 \cdots m)$ *into* $m/2$ *pairs* (l'_k, l''_k) *and* $i'_k = i_{l'_k}$, $j'_k = j_{l'_k}$, $i''_k = i_{l''_k}$, $j''_k = j_{l''_k}$.

This lemma implies that

$$E n^{-m/2} S_n(r_{i_1 j_1}) \cdots S_n(r_{i_m j_m}) \to \textstyle\sum \prod \sigma_{i'_k j''_k j'_k j''_k}.$$

Thus $\{n^{-1/2} S_n(r_{ij})\}$ asymptotically satisfies the relations $E\xi_{\alpha_1} \cdots \xi_{\alpha_m} = \Sigma\, \Pi_1^{m/2}\, E(\xi_{\alpha'_k} \xi_{\alpha''_k})$ (where the sum is taken over all partitions of $(\alpha_1 \cdots \alpha_m)$ into $m/2$ pairs (α'_k, α''_k)), characterizing central joint normal distributions [9, Proposition I.2]. Consequently, if every r_{ij} is real,

$$\{n^{-1/2} S_n(r_{ij})\}_{i,j=1}^{d} \to_d N(0, \Sigma). \tag{4.28}$$

If some r_{ij} is complex-valued, we may apply this to $R \oplus \bar R$ (which is equivalent to a representation with real coefficients) and conclude that $\{n^{-1/2}\mathrm{Re}\, S_n(r_{ij}),\, n^{-1/2}\mathrm{Im}\, S_n(r_{ij})\}$ is asymptotically normally distributed.

Finally, for a finite number of representations, we apply this to their direct sum. This shows that $n^{-1/2} S_n(R_\alpha)$ converge jointly for several representations R_α.

If R_1 and R_2 do not contain any common irreducible representation, $P = 0$ in (4.20). Thus $E S_n(R_1) \otimes S_n(R_2) = o(n)$.

We have proved the following theorem.

THEOREM 4.1. *Suppose that* $\mathrm{Re}\,\lambda_1(R) < 1/2$ *for all irreducible representations but the trivial one. Then* $\{n^{-1/2} S_n(r_{ij}^{\alpha})\}$ *converge jointly to normal distributions (i.e. convergence in* $C^\infty = (R^2)^\infty$*).*

The covariance structure is given by (4.23). In particular, terms corresponding to non equivalent, or conjugate-equivalent, representations are asymptotically independent.

If $f \in L^2$ is real and $\hat f(R)$ vanishes for all but a finite number of representations, then $n^{-1/2}(S_n(f) - \int f\, dm)$ converges to a normal distribution.

In order to prove limit theorems for more general functions, we will need uniform estimates of $E\,|\,S_n(f)|^2$ and not only asymptotical results. We will do that with a somewhat stronger condition on μ. We introduce two further matrix norms; the standard operator norm $\|A\| = \sup\{\|Ax\| : \|x\| = 1\}$ and the trace class norm $\|A\|_{\mathrm{TC}} = \mathrm{Tr}((A^*A)^{1/2})$. If A is a normal matrix with eigenvalues $\lambda_1 \cdots \lambda_d$, $\|A\| = \sup|\lambda_k|$, $\|A\|_{\mathrm{TC}} = \Sigma\,|\,\lambda_k|$ and $\|A\|_{\mathrm{HS}} = (\Sigma\,|\,\lambda_k|^2)^{1/2}$. In particular $\|R(g)\| = 1$, whence $\|\hat{\mu}(R)\|$, $\|\hat{\nu}(R)\| \le 1$ for every representation R and the crude estimates

$$\|ES_n(R)\| \le \textstyle\sum_0^n \|ER(X_k)\| \le \sum_0^n E\|R(X_k)\| = n+1 \tag{4.29}$$

$$\|E(S_n(R_1) \otimes S_n(R_2))\| \le (n+1)^2. \tag{4.30}$$

LEMMA 4.3. *Assume that there exist $t \ge 0$ and $\kappa < \frac{1}{2}$ such that*

$$\|tI + \hat{\mu}(R)\| \le t + \kappa \tag{4.31}$$

for every irreducible representation of G except the trivial one. Then, for all $f \in L^2(m)$ with $\int f\,dm = 0$,

$$\text{(i)} \qquad \|S_n(f)\|_2 \le Cn^{1/2} \textstyle\sum_{\alpha \in \hat{G}} d_\alpha \|\hat{f}(R_\alpha)\|_{\mathrm{TC}}. \tag{4.32}$$

(ii) *If furthermore ν equals the Haar measure m,*

$$\|S_n(f)\|_2 \le Cn^{1/2}\|f\|. \tag{4.33}$$

(iii) *If ν is absolutely continuous and $d\nu/dm \in L^2$,*

$$|ES_n(f)| \le Cn^{\kappa}\|f\|\,\|d\nu/dm\|. \tag{4.34}$$

The constants C depend on t and κ only.

PROOF. (i) First we note that if (4.31) holds for some t, it holds for all larger numbers as well. Let R be a non-trivial irreducible representation acting on the d-dimensional space V. Then, if $n+1 \ge 2t$,

$$\begin{aligned} &\left\| I + \frac{1}{n+1}\hat{\mu}(R) \otimes I + \frac{1}{n+1} I \otimes \hat{\mu}(\bar{R}) \right\| \\ &\quad = \left\| \left(\frac{1}{2}I + \frac{1}{n+1}\hat{\mu}(R)\right) \otimes I + I \otimes \left(\frac{1}{2}I + \frac{1}{n+1}\hat{\mu}(\bar{R})\right) \right\| \\ &\quad \le 2\left\| \frac{1}{2}I + \frac{1}{n+1}\hat{\mu}(R) \right\| = \frac{2}{n+1}\left\| \frac{n+1}{2}I + \hat{\mu}(R) \right\| \le 1 + \frac{2\kappa}{n+1}. \end{aligned} \tag{4.35}$$

By (4.9), (4.35) and (4.29), for $n+1 \ge 2t$,

$$\begin{aligned} &\|E(S_{n+1}(R) \otimes S_{n+1}(\bar{R}))\| \\ &\quad \le \|E(S_n(R) \otimes S_n(\bar{R}))\| \left\| I + \frac{1}{n+1}\hat{\mu}(R) \otimes I + \frac{1}{n+1} I \otimes \mu(\bar{R}) \right\| \\ &\qquad + \frac{1}{n+1}\|ES_n(R \otimes \bar{R})\|\,\|\hat{\mu}(R \otimes \bar{R})\| \\ &\quad \le \|E(S_n(R) \otimes S_n(\bar{R}))\| \left(1 + \frac{2\kappa}{n+1}\right) + 1. \end{aligned} \tag{4.36}$$

Let p_n be as in (2.15) and define

$$q_n(x) = \frac{p_n(1) - p_n(x)}{1-x}. \tag{4.37}$$

Then

$$q_0(x) = 1, \quad q_{n+1}(x) = q_n(x)\left(1 + \frac{x}{n+1}\right) + 1. \tag{4.38}$$

If we assume, as we may, that $\kappa \geq 0$, $q_n(2\kappa) \geq n + 1$. By (4.30) we have for $n \leq 2t$, with $C = 2t + 1$,

$$\| E(S_n(R) \otimes S_n(\bar{R})) \| \leq C q_n(2\kappa). \tag{4.39}$$

But by (4.36), (4.38) and induction, (4.39) then holds for all n.

Let $a_1 \cdots a_d$ and $b_1 \cdots b_d$ be complex numbers. Then, with $a = \Sigma a_i e_i$ and $b = \Sigma b_i e_i \in V$,

$$\begin{aligned} E\,|\,S_n(\Sigma a_i b_j r_{ij})\,|^2 &= \textstyle\sum_{ijkl} a_i b_j \bar{a}_k \bar{b}_l E(S_n(r_{ij}) S_n(\bar{r}_{kl})) \\ &= \textstyle\sum a_i \bar{a}_k (ES_n(R) \otimes S_n(\bar{R}))_{ikjl} b_j \bar{b}_l \\ &= \langle E(S_n(R) \otimes S_n(\bar{R})) b \otimes \bar{b}, a \otimes \bar{a} \rangle \\ &\leq \| a \otimes \bar{a} \| \, \| E(S_n(R) \otimes S_n(\bar{R})) \| \, \| b \otimes \bar{b} \| \\ &= \| E(S_n(R) \otimes S_n(\bar{R})) \| \textstyle\sum_1^d |a_i|^2 \sum_1^d |b_i|^2. \end{aligned} \tag{4.40}$$

By a well-known property of the trace class norm, there exists for any matrix $A = (a_{ij})$ numbers λ_k and a_i^k, b_i^k, $i, k = 1 \cdots d$, such that

$$\Sigma\,|\,\lambda_k\,| = \| A \|_{\mathrm{TC}}, \quad \Sigma_i |\,a_i^k\,|^2 = \Sigma_i |\,b_i^k\,|^2 = 1 \quad \text{and} \quad a_{ij} = \Sigma_k \lambda_k a_j^k b_i^k.$$

Then, by Minkowski's inequality, (4.40) and (4.39),

$$\begin{aligned} \| S_n(\mathrm{Tr}(AR)) \|_2 &= \| S_n(\Sigma_k \lambda_k \Sigma_{ij} a_i^k b_j^k r_{ij}) \|_2 \leq \Sigma_k |\,\lambda_k\,| \, \| E(S_n(R) \otimes S_n(\bar{R})) \|^{1/2} \\ &\leq C \| A \|_{\mathrm{TC}} q_n(2\kappa)^{1/2}. \end{aligned} \tag{4.41}$$

It follows from (4.37) that (assuming $\kappa \geq 0$)

$$q_n(2\kappa) \leq p_n(1)/(1 - 2\kappa) = (n + 1)/(1 - 2\kappa). \tag{4.42}$$

The Fourier expansion (4.2), (4.41) and (4.42) yield (for some C depending on κ and t)

$$\| S_n(f) \|_2 \leq \Sigma d_\alpha \| S_n \,\mathrm{Tr}(\hat{f}(R_\alpha) R_\alpha) \|_2 \leq C n^{1/2} \Sigma d_\alpha \| \hat{f}(R_\alpha) \|_{\mathrm{TC}}.$$

(ii) We continue to use the same notation. By (4.17) $\hat{\nu}(R \otimes \bar{R}) = P$ and (using (4.18)) (4.8) and (4.5) are solved by

$$F_1(R \otimes \bar{R}; z) = P(1 - z)^{-2}. \tag{4.43}$$

Consequently $ES_n(R \otimes \bar{R}) = (n + 1)P$ and (4.9), (4.18) yield, with W denoting $\hat{\mu}(R) \otimes I + I \otimes \hat{\mu}(\bar{R})$,

$$E(S_{n+1}(R) \otimes S_{n+1}(\bar{R})) = E(S_n(R) \otimes S_n(\bar{R}))\left(I + \frac{1}{n+1} W\right) + P. \tag{4.44}$$

By (4.38) and the fact that $E(S_0(R) \otimes S_0(\bar{R})) = \hat{\nu}(R \otimes \bar{R}) = P$, (4.44) is solved by

$$E(S_n(R) \otimes S_n(\bar{R})) = P q_n(W). \tag{4.45}$$

Let $\Xi = (\xi_{ij})$ be a $d \times d$ matrix and set $f = \sum_{i,j=1}^d \xi_{ij} r_{ij}$. Then, by (4.1), (4.45), (4.24),

$$\begin{aligned} E\,|\,S_n(f)\,|^2 &= \textstyle\sum_{iji'j'} \xi_{ij} \bar{\xi}_{i'j'} ES_n(r_{ij}) S_n(\bar{r}_{i'j'}) \\ &= \textstyle\sum \xi_{ij} \bar{\xi}_{i'j'} E(S_n(R) \otimes S_n(\bar{R}))_{ii'jj'} = \sum \xi_{ij} \bar{\xi}_{i'j'} (P q_n(W))_{ii'jj'} \\ &= \textstyle\sum_{ii'jj'kk'} \xi_{ij} \bar{\xi}_{i'j'} \dfrac{1}{d} \delta_{ii'} \delta_{kk'} q_n(W)_{kk'jj'} = \dfrac{1}{d} \sum_{ijj'k} \xi_{ij} \bar{\xi}_{ij'} q_n(W)_{kkjj'}. \end{aligned}$$

We define, for any linear operator T on $V \otimes V$,

$$\varphi_f(T) = \frac{1}{d} \sum_{ijlk} \xi_{ij} \bar{\xi}_{il} T_{kkjl}. \tag{4.46}$$

Thus

$$E \mid S_n(f) \mid^2 = \varphi_f(q_n(W)). \tag{4.47}$$

If $A = (a_{ij})$ and $B = (b_{ij})$ are operators on V, then

$$\varphi_f(A \otimes B) = \frac{1}{d} \sum_{ijkl} \xi_{ij} \bar{\xi}_{il} a_{kj} b_{kl} = \frac{1}{d} \operatorname{Tr}(\Xi A^t B \Xi^*). \tag{4.48}$$

Hence, using standard matrix norm inequalities and (4.3),

$$\begin{aligned} \mid \varphi_f(A \otimes B) \mid &\le \frac{1}{d} \| \Xi \|_{\mathrm{HS}} \| A^t \| \| B \| \| \Xi^* \|_{\mathrm{HS}} \\ &= \frac{1}{d} \| \Xi \|_{\mathrm{HS}}^2 \| A \| \| B \| = \| f \|^2 \| A \| \| B \|. \end{aligned} \tag{4.49}$$

By the definition of W,

$$\begin{aligned} \mid \varphi_f(W^m) \mid &= \textstyle\sum_0^m \binom{m}{k} \varphi_f(\hat{\mu}(R)^k \otimes \hat{\mu}(\bar{R})^{m-k}) \mid \le \sum_0^m \binom{m}{k} \| \hat{\mu}(R)^k \| \| \hat{\mu}(\bar{R}) \|^{m-k} \| f \|^2 \\ &\le \textstyle\sum_0^m \binom{m}{k} \| \hat{\mu}(\bar{R}) \|^k \| \hat{\mu}(\bar{R}) \|^{m-k} \| f \|^2 = 2^m \| \hat{\mu}(R) \|^m \| f \|^2. \end{aligned} \tag{4.50}$$

By (4.38), q_n is a polynomial with positive coefficients. If $q_n(x) = \Sigma c_m x^m$, (4.47) and (4.50) yield

$$\begin{aligned} E \mid S_n(f) \mid^2 &= \varphi_f(q_n(W)) = \varphi_f(\textstyle\sum c_m W^m) = \sum c_m \varphi_f(W^m) \\ &\le \textstyle\sum c_m (2 \| \hat{\mu}(R) \|)^m \| f \|^2 = q_n(2 \| \hat{\mu}(R) \|) \| f \|^2. \end{aligned} \tag{4.51}$$

This yields good estimates if $\| \hat{\mu}(R) \| < \kappa$, i.e. if $t = 0$ in (4.31). In general, we modify the calculation as follows. Since $2tI + W = (tI + \hat{\mu}(R)) \otimes I + I \otimes (tI + \hat{\mu}(\bar{R}))$, we obtain exactly as (4.50), using (4.31)

$$\mid \varphi_f((2tI + W)^m) \mid \le 2^m \| tI + \hat{\mu}(R) \|^m \| f \|^2 \le 2^m (t + \kappa)^m \| f \|^2.$$

Set $\tilde{q}_n(x) = q_n(x - 2t)$. Thus $\tilde{q}_0(x) = 1$ and

$$\tilde{q}_{n+1}(x) = \tilde{q}_n(x) \left(1 - \frac{2t}{n+1} + \frac{x}{n+1} \right) + 1. \tag{4.52}$$

Define $\tilde{\tilde{q}}_n$ by

$$\tilde{\tilde{q}}_0(x) = 1 \quad \text{and} \quad \tilde{\tilde{q}}_{n+1}(x) = \tilde{\tilde{q}}_n(x) \left(\left| 1 - \frac{2t}{n+1} \right| + \frac{x}{n+1} \right) + 1. \tag{4.53}$$

Let $\tilde{q}_n(x) = \Sigma \tilde{c}_m x^m$ and $\tilde{\tilde{q}}_n(x) = \Sigma \tilde{\tilde{c}}_m x^m$. It is clear by (4.52) and (4.53) that $\mid \tilde{c}_m \mid \le \tilde{\tilde{c}}_m$. Thus

$$\begin{aligned} E \mid S_n(f) \mid^2 &= \varphi_f(q_n(W)) = \varphi_f(\tilde{q}_n(2tI + W)) = \textstyle\sum \tilde{c}_m \varphi_f((2tI + W)^m) \\ &\le \textstyle\sum \mid \tilde{c}_m \mid 2^m (t + \kappa)^m \| f \|^2 \le \sum \tilde{\tilde{c}}_m (2t + 2\kappa)^m \| f \|^2 = \tilde{\tilde{q}}_n(2t + 2\kappa) \| f \|^2. \end{aligned}$$

We may assume that $\kappa \ge 0$. Then $\tilde{q}_n(2t + 2\kappa) = q_n(2\kappa) \ge 1$. If $C_1 = \sup_{n \le 2t} \tilde{\tilde{q}}_n(2t + 2\kappa) / \tilde{q}_n(2t + 2\kappa)$, it follows from (4.52) and (4.53) by induction that $\tilde{\tilde{q}}_n(2t + 2\kappa) \le C_1 \tilde{q}_n(2t + 2\kappa)$ for all $n \ge 0$. Hence

$$E \mid S_n(f) \mid^2 \le C_1 \tilde{q}_n(2t + 2\kappa) \| f \|^2 = C_1 q_n(2\kappa) \| f \|^2.$$

By (4.42), $E|S_n(f)|^2 \le Cn\|f\|^2$ for functions of this type. In general, decompose f by (4.2) as $\Sigma_{\hat{G}} f_\alpha$ where $f_\alpha = d_\alpha \sum_{i,j=1}^{d_\alpha} \hat{f}(R_\alpha)_{ji} r_{ij}^\alpha$. If $\alpha \neq \beta$, then $R_\alpha \otimes \bar{R}_\beta$ does not contain the trivial representation and $\hat{\nu}(R_\alpha \otimes \bar{R}_\beta) = 0$. It follows that $ES_n(R_\alpha) \otimes S_n(\bar{R}_\beta) = 0$, $\alpha \neq \beta$. Thus, if $\alpha \neq \beta$,

$$ES_n(f_\alpha)\overline{S_n(f_\beta)} = \Sigma_{ijkl}\, d_\alpha \hat{f}(R^\alpha)_{ji}\, d_\beta \overline{\hat{f}(R^\beta)}_{lk} E(S_n(r_{ij}^\alpha)S_n(\bar{r}_{kl}^\beta)) = 0.$$

Consequently,

$$E|S_n(f)|^2 = \Sigma_\alpha E|S_n(f_\alpha)|^2 \le Cn\Sigma_\alpha \|f_\alpha\|^2 = Cn\|f\|^2$$

and (4.33) is proved.

(iii) By (4.7), for $n > t$,

$$\|ES_{n+1}(R)\|_{\mathrm{HS}} \le \|ES_n(R)\|_{\mathrm{HS}} \left\| I + \frac{1}{n+1}\hat{\mu}(R) \right\| \le \left(1 + \frac{\kappa}{n+1}\right) \|ES_n(R)\|_{\mathrm{HS}}.$$

We obtain by (2.15) and induction as above

$$\|ES_n(R)\|_{\mathrm{HS}} \le Cp_n(\kappa)\|ES_0(R)\|_{\mathrm{HS}} \le Cn^\kappa \|\hat{\nu}(R)\|_{\mathrm{HS}}$$

(a slight modification is needed if $\kappa \le 0$).

(4.2) and (4.3) now yield

$$\begin{aligned} |ES_n(f)| &= |\Sigma\, d_\alpha \mathrm{Tr}(\hat{f}(R_\alpha)ES_n(R_\alpha))| \le \Sigma\, d_\alpha \|\hat{f}(R_\alpha)\|_{\mathrm{HS}} \|ES_n(R_\alpha)\|_{\mathrm{HS}} \\ &\le Cn^\kappa \Sigma\, d_\alpha \|\hat{f}(R_\alpha)\|_{\mathrm{HS}} \|\hat{\nu}(R_\alpha)\|_{\mathrm{HS}} \\ &\le Cn^\kappa (\Sigma\, d_\alpha \|\hat{f}(R_\alpha)\|^2)^{1/2} (\Sigma\, d_\alpha \|\hat{\nu}(R_\alpha)\|^2)^{1/2} = Cn^\kappa \|f\| \, \|d\nu/dm\|. \end{aligned}$$

The proof is complete.

The substitution $\varepsilon = 1/t > 0$ shows that (4.31) is equivalent to $\|I + \varepsilon\hat{\mu}(R)\| \le 1 + \varepsilon\kappa$. Now, suppose that the Hermitian matrix $\frac{1}{2}(\hat{\mu}(R) + \hat{\mu}(R)^*)$ has no eigenvalues exceeding $\kappa < \frac{1}{2}$. Then,

$$\begin{aligned} \|I + \varepsilon\hat{\mu}(R)\|^2 &= \|(I + \varepsilon\hat{\mu}(R))^*(I + \varepsilon\hat{\mu}(R))\| = \|I + \varepsilon\hat{\mu}(R) + \varepsilon\hat{\mu}(R)^* + \varepsilon^2\hat{\mu}(R)^*\hat{\mu}(R)\| \\ &\le \|I + \varepsilon(\hat{\mu}(R) + \hat{\mu}(R)^*)\| + \varepsilon^2 \le 1 + 2\varepsilon\kappa + \varepsilon^2. \end{aligned}$$

Taking ε small enough, we see that (4.31) holds with κ replaced by some $\kappa' < \frac{1}{2}$. Also, note that if (4.31) holds,

$$|t + \lambda_1(R)| \le \|t + \hat{\mu}(R)\| \le t + \kappa$$

and thus $\mathrm{Re}\,\lambda_1(R) \le \kappa < \frac{1}{2}$.

REMARK. In fact, (4.31) implies $\mathrm{Re}\langle\hat{\mu}(R)x, x\rangle \le \kappa\langle x, x\rangle$, which is equivalent to the above condition on $\frac{1}{2}(\hat{\mu}(R) + \hat{\mu}(R)^*)$. Another formulation is that the numerical range of $\hat{\mu}(R)$ is contained in $\{z : \mathrm{Re}\, z \le \kappa\}$. It is well-known that this is strictly stronger than just assuming the eigenvalues not to exceed κ (cf. Theorem 4.1).

After these preparations, the following theorem is proved as Theorem 3.1.

THEOREM 4.2. *Suppose that there exists a $\kappa < \frac{1}{2}$ such that the eigenvalues of $\frac{1}{2}(\hat{\mu}(R) + \hat{\mu}(R^*))$ are less than κ for every non-trivial irreducible representation. Let f be a real function in $L^2(m)$. Suppose further that either*

(i) $\Sigma_{\hat{G}} d_\alpha \|\hat{f}(R_\alpha)\|_{\mathrm{TC}} < \infty$, *or*
(ii) *ν is absolutely continuous with $d\nu/dm$ bounded, or*
(iii) *μ is absolutely continuous with $d\mu/dm$ bounded.*

Then

$$n^{-1/2}\left(S_n(f) - n\int f\,dm\right) \to_d N(0, \sigma^2),$$

with

$$\sigma^2 = \sum_\alpha d_\alpha^2 \sum_{ijkl} \langle f, r_{ij}^\alpha\rangle\langle r_{kl}^\alpha, f\rangle\sigma_{ikjl}(R_\alpha) < \infty.$$

REMARKS.

1. We have assumed that (4.31) holds uniformly in R. It is possible to obtain similar results also when t and κ are allowed to depend on R, cf, Theorem 3.1.

2. If every $\hat{\mu}(R_\alpha)$ is diagonal, (4.26) yields

$$\sigma^2 = \sum_\alpha d_\alpha \sum_{ij} \langle f, r_{ij}^\alpha\rangle^2(1 - 2\,\mathrm{Re}(\hat{\mu}(R_\alpha))_{jj})^{-1}$$
$$= \sum d_\alpha \sum |\hat{f}(R^\alpha)_{ij}|^2(1 - 2\,\mathrm{Re}\,\hat{\mu}(R_\alpha)_{ii})^{-1}.$$

There are similar results when the real part of some Fourier coefficient equals ½. We state two theorems, omitting the proofs which are similar to the preceding ones (although the combinatorics gets more complicted when d_1 below exceeds one).

THEOREM 4.3. *Assume that R is a real representation with $\lambda_1(R) = ½$. Let d be the dimension of R and d_1 the dimension of the largest Jordan box of $\hat{\mu}(R)$ for any eigenvalue with real part ½. Then, for some $\Sigma \neq 0$,*

$$(n(\log n)^{2d_1-1})^{-1/2}S_n(R) \to_d N(0, \Sigma).$$

If the eigenvalues of $\hat{\mu}(R)$ with real parts ½ are $\lambda_1 \cdots \lambda_l$ with corresponding projections P_k so that $P_k\hat{\mu}(R) = \hat{\mu}(R)P_k = \lambda_k P_k + N_k$, N_k nilpotent, Σ is given by

$$\sigma_{ii'jj'} = \frac{1}{d(2d_1 - 1)((d_1 - 1)!)^2}\,\delta_{ii'} \sum_{k=1}^{l} \langle P_k N_k^{d_1-1}e_j, P_k N_k^{d_1-1}e_{j'}\rangle.$$

THEOREM 4.4. *Suppose that there exist $t \geq 0$ such that $\| t + \hat{\mu}(R) \| \leq t + ½$ for every non-trivial irreducible representation. Let P_α be the projection commuting with $\hat{\mu}(R_\alpha)$ onto the eigenspace $\ker(\hat{\mu}(R_\alpha) - ½I)$. ($P_\alpha = 0$ is possible.) Let $f \in L^2(m)$. Suppose further that*

(i) $\Sigma\, d_\alpha \| \hat{f}(R_\alpha) \|_{\mathrm{TC}} < \infty$ *or,*

(ii) *ν is absolutely continuous with $d\nu/dm$ bounded, or*

(iii) *μ is absolutely continuous with $d\mu/dm$ bounded.*

Then

$$(n \log n)^{-1/2}\left(S_n(f) - n\int f\,dm\right) \to_d N(0, \Sigma_\Gamma\, d_\alpha \| P_\alpha \hat{f}(R_\alpha) \|_{\mathrm{HS}}^2).$$

REMARK. The assumption in the last theorem implies that the only eigenvalue on the critical line is ½, that $d_1 = 1$ and that P_α is an orthogonal projection.

5. Homogeneous spaces. In this section we assume that the state space K is a compact homogeneous space, i.e. there exists a compact group acting transitively on K. We fix an arbitrary element $x_0 \in K$ and let $H = \{g \in G : gx_0 = x_0\}$ be the subgroup fixing x_0. The mapping $\pi : g \to \pi(g) = gx_0$ maps G onto K, and K may be identified with G/H.

The unit sphere S^{n-1} in R^n is an important example; we take $G = O(n)$, the orthogonal group consisting of all rotations of R^n, and $H = O(n - 1)$. Another example is $G = O(n)$ and $H = O(k) \times O(n - k)$, $1 \leq k \leq n - 1$, which yields the Grassmann manifolds.

Let $Y_1, Y_2, \cdots$ be i.i.d. random elements of G with the distribution μ and assume that μ is invariant under left and right multiplication by elements of H, i.e. hY_nh' and Y_n are identically distributed when $h, h' \in H$.

If ν is a probability measure on K, let ν' be a measure on G inducing ν on K and let X_0', X_1', $\cdots$ be the branching random walk on G defined in Section 1 using ν' and μ. The branching random walk on K then is defined by

$$X_n = \pi(X_n').$$

Because of the invariance property of μ, we may alternatively use the definition of Section 1, replacing (1.1) by:

$$\text{If} \quad X_{I_n} = gx_0, \quad \text{then} \quad X_n \quad \text{is distributed as} \quad gY_nx_0.$$

A third possible definition could be formulated employing transition probabilities $p_x(y)$ that are invariant, i.e. $p_{gx}(gy) = p_x(y)$. Given X_{I_n}, X_n is given the distribution $p_{X_{I_n}}(\cdot)$.

Denote the Haar measures on G, H and K by m_G, m_H and m_K, respectively. Then π maps m_G to m_K and the mapping $f \to f \circ \pi$ yields an isometry between $L^2(K, m_K)$ and $\{\varphi \in L^2(G, m_G) : \varphi(gh) = \varphi(g) \text{ when } h \in H\} = \{\varphi \in L^2(G) : \varphi = \varphi * m_H\}$. We define the Fourier transform on $L^2(K)$ by $\hat{f}(R_\alpha) = \widehat{f \circ \pi}(R_\alpha)$, R_α a representation of G. Then $\hat{f}(R_\alpha) = \widehat{m_H}(R_\alpha)\hat{f}(R_\alpha)$; in particular, $\hat{f}(R_\alpha) = 0$ if R_α (restricted to H) does not contain the trivial representation of H.

If $f \in L^2(K)$, $S_n(f) = S_n'(f \circ \pi)$, where $S_n'(\varphi) = \sum_0^n \varphi(X_i')$. Thus, we may apply all results of the preceding section to the branching random walk $\{X_i'\}$ on G and obtain results for $S_n(f)$. Note that, by assumption,

$$\mu = m_H * \mu * m_H, \quad \text{whence} \quad \hat{\mu}(R_\alpha) = \hat{m}_H(R_\alpha)\hat{\mu}(R_\alpha)\hat{m}_H(R_\alpha). \tag{5.1}$$

Thus only representations with $\hat{m}_H(R_\alpha) \neq 0$ are important.

We develop the theory in some detail for the sphere S^{N-1}, cf. [4], [10]. $G = O(N)$ acts on the space of spherical harmonics of degree k, $k = 0, 1, \cdots$, by $Of(x) = f(O^{-1}x)$, $x \in S^{N-1}$. This yields an irreducible representation R_k of $O(N)$ of dimension $d_k = ((N + 2k - 2)/k)\binom{N+k-3}{k-1}$. The restriction to $H = O(N-1)$ of this representation contains the trivial representation of H exactly once; the corresponding eigenfunction ψ_k of $\hat{m}_H$ is known as the zonal harmonic of degree k. Since $\hat{m}_H$ has rank one, (5.1) implies that $\hat{\mu}(R_k) = \hat{\mu}(k)\hat{m}_H(R_k)$ for some real number $\hat{\mu}(k)$. $\hat{\mu}(k)$ is given by $\hat{\mu}(k)\psi_k = \hat{\mu}(R_k)\psi_k = E(\psi_k(X_n) \mid X_{I_n} = x)$. Thus

$$\hat{\mu}(k) = E(\psi_k(X_1) \mid X_0 = x_0)/\psi_k(x_0). \tag{5.2}$$

There are other irreducible representations of $O(N)$, but these can be ignored since they do not contain the trivial representation of H.

Any f in L^2 can be decomposed as $\sum_0^\infty f_k$ (convergence in L^2), where each f_k is a spherical harmonics of degree k. It follows by some computations that the operator $\hat{f}(R_k)$ on the space of spherical harmonics of degree k is given by (if ψ_k is normalized by $\| \psi_k \| = 1$)

$$\hat{f}(R_k)\varphi = d_k^{-1/2} \int f\varphi \, dm \, \psi_k = d_k^{-1/2}\langle \varphi, \bar{f}_k \rangle \psi_k.$$

Hence

$$\| \hat{f}(R_k) \|_{\text{TC}} = \| \hat{f}(R_k) \|_{\text{HS}} = \| \hat{f}(R_k) \| = d_k^{-1/2} \| f_k \|.$$

Since $d_k \sim Ck^{N-2}$ as $k \to \infty$, we obtain the following theorem from the results of Section 4.

THEOREM 5.1. *Supppose that* $\sup_{k \ge 1} \hat{\mu}(k) \le ½$ *and that* $f \in L^2(S^{N-1})$. *Let* $f = \sum_0^\infty f_k$ *be the expansion of f into spherical harmonics. Suppose further that either*

(i) $\sum_1^\infty k^{(N-2)/2} \| f_k \| < \infty$, *or*

(ii) ν *is absolutely continuous and* $d\nu/dm_{S^{N-1}}$ *is bounded, or*

(iii) μ *is absolutely continuous and* $d\mu/dm_{O(N)}$ *is bounded.*

Then, if $\sup \hat{\mu}(k) < \frac{1}{2}$,

$$n^{-1/2}\left(S_n(f) - n\int f\right) \to_d N(0, \textstyle\sum_1^\infty (1 - 2\hat{\mu}(k))^{-1} \| f_k \|^2)$$

while, if $\sup \hat{\mu}(k) = \frac{1}{2}$,

$$(n \log n)^{-1/2}\left(S_n(f) - n\int f\right) \to_d N(0, \textstyle\sum_{\hat{\mu}(k)=1/2} \| f_k \|^2).$$

6. Examples.

1. Let $G = \{1, -1\}$ be a group with two elements and let $\mu = p\delta_1 + (1 - p)\delta_{-1}$. The branching random walk is equivalent to the following urn model. (A special generalized Pólya-Eggenberger model [6].) Begin with one ball, black or white according to the distribution ν, in the urn. Draw a ball, replace it and add a new ball, where the new ball has the same colour as the drawn one with probability p. If $f(X) = X$, $S_n = S_n(f)$ is the difference between the numbers of black and white balls in the urn when the total number is $n + 1$. The only non-trivial Fourier coefficient is $p - (1 - p) = 2p - 1$. By the results of Section 2, if $p < \frac{3}{4}$, then $n^{-1/2}S_n \to_d N(0, (3 - 4p)^{-1})$, if $p = \frac{3}{4}$ then $(n \log n)^{-1/2}S_n \to_d N(0, 1)$ and if $p > \frac{3}{4}$, $n^{-(2p-1)}S_n$ converges to a distribution whose first four moments are given by (2.27) ($\hat{\mu} = 2p - 1$).

2. Let G be any compact group or homogeneous space and let $\mu = p\delta_e + (1 - p)m$, $0 \le p \le 1$ (e is the unit element). Thus, with probability p, X_n equals some previous point, thus enlarging the "colony" at that point, and with probability $1 - p$, X_n begins a new colony at a random point. $\hat{\mu}(R) = p \cdot I$ for every non-trivial irreducible representation. Consequently, for $f \in L^2(G)$ and $\int f = 0$, if $p < \frac{1}{2}$, $n^{-1/2}S_n(f) \to N(0, \| f \|^2(1 - 2p)^{-1})$, if $p = 1/2$, $(n \log n)^{-1/2}S_n(f) \to N(0, \| f \|^2)$ and if $p > \frac{1}{2}$ a stronger clustering occurs.

3. Assume that the points represent N different types $1 \cdots N$ and that if a parent is of type k, all her daughters are of type $k + 1$ (cyclically). Thus G is the cyclic group Z_n and $\mu = \delta_1$, a point mass. There are N characters and $\hat{\mu}(\gamma_j) = e^{2\pi ij/N}$, $j = 0 \cdots N - 1$. Hence $\sup_{j\neq 0}\mathrm{Re}\ \hat{\mu}(\gamma_j) = \cos(2\pi/N)$. Consequently, with $Z_n^{(k)}$ denoting the number of the first n points that are of type k, if $N \le 5$ then $n^{-1/2}(Z_n^{(1)} - n/N, \ldots, Z_n^{(N)} - n/N)$ converges to a non-degenerate normal distribution, but if $N = 6$ then we have to normalize by $(n \log n)^{-1/2}$ and the limiting distribution is only two-dimensional, and if $N > 6$, the distribution is not asymptotically normal.

4. We take the sphere S^2 in R^3 and let each daughter have the distance $\pi/2$ from her parent. (If the parent is at the north pole, the daughter will be uniformly distributed along the equator.) An explicit computation with the spherical harmonics shows that $\hat{\mu}(2k - 1) = 0$, $\hat{\mu}(2k) = \binom{-1/2}{k}$. Thus, the largest (non-trivial) Fourier coefficient is $\hat{\mu}(4) = \frac{3}{8} < \frac{1}{2}$ and $n^{-1/2}S_n(f)$ is asymptotically normally distributed for every sufficiently nice function f.

5. We keep the sphere but let each daughter be uniformly distributed over the half-sphere centered at the parent. Then $\hat{\mu}(1) = \frac{1}{2}$ and $\hat{\mu}(k) < \frac{1}{2}$, $k > 1$. Hence, $(n \log n)^{-1/2}S_n(f)$ is asymptotically normally distributed for nice functions f. If f is orthogonal to the linear functions, the limiting normal distribution is degenerate and, in fact, $n^{-1/2}S_n(f)$ converges. If the daughter is uniformly distributed over a smaller cap, $\hat{\mu}(1) > \frac{1}{2}$ and strong clustering occurs. If the cap is larger, $\sup \hat{\mu}(k) < \frac{1}{2}$ and we are in the well-behaved case.

REFERENCES

[1] ATHREA, K. B. and NEY, P. E. (1972). *Branching Processes.* Springer, Berlin.
[2] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.
[3] GRENANDER, U. (1963). *Probabilities on Algebraic Structures.* Almqvist and Wiksell, Uppsala.
[4] HANNAN, E. J. (1965). Group representations and applied probability. *J. Appl. Probab.* **2** 1–68.
[5] HEYER, H. (1977). *Probability Measures on Locally Compact Groups.* Springer, Berlin.
[6] JOHNSON, N. L., and KOTZ, S. (1977). *Urn Models and Their Application.* Wiley, New York.
[7] RAO, R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York.
[8] SERRE, J.-P. (1977). *Représentations Linéaires des Groupes Finis.* Hermann, Paris.

[9] SIMON, B. (1974). *The $P(\phi)_2$ Euclidean (Quantum) Field Theory*. Princeton University Press, Princeton.

[10] STEIN, E. M., and WEISS, G. (1971). *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, Princeton.

[11] ZYGMUND, A. (1959). *Trigonometric Series*, 2nd ed. Cambridge University Press, Cambridge.

DEPARTMENT OF MATHEMATICS
UPPSALA UNIVERSITY
THUNBERGSVÄGEN 3
752 38 UPPSALA, SWEDEN