

FINITE FULL TRANSFORMATION SEMIGROUPS AS COLLECTIONS OF RANDOM FUNCTIONS

by B. BROWN and P. M. HIGGINS

(Received 29 November, 1986)

1. Preliminaries. The collection of all self-maps on a non-empty set X under composition is known in algebraic semigroup theory as the full transformation semigroup on X and is written \mathcal{T}_X . Its importance lies in the fact that any semigroup S can be embedded in the full transformation semigroup \mathcal{T}_{S^1} (where S^1 is the semigroup S with identity 1 adjoined, if S does not already possess one). The proof is similar to Cayley's Theorem that a group G can be embedded in S_G , the group of all bijections of G to itself. In this paper X will be a finite set of order n , which we take to be $\bar{n} = \{1, 2, \dots, n\}$, and so we shall write T_n for \mathcal{T}_X .

We investigate certain random variables which arise from the random selection of $\alpha \in T_n$ (meaning that each such α is equally likely to be chosen). Various aspects of T_n , regarded as a set of random functions, have been investigated before: see for instance [9], [3], [6]. Harris for example studies the order of the monogenic semigroup generated by a random $\alpha \in T_n$.

As has often been observed, the members of T_n can be regarded as labelled directed graphs in which each vertex has outdegree one (the digraph of $\alpha \in T_n$ has n labelled vertices with ij an arc if $i\alpha = j$). The components of the digraph of α (see [5]) are *functional*, meaning that each consists of a unique cycle together with a number of labelled trees rooted around the vertices of the cycle. Two vertices labelled i and j respectively ($i, j \in \bar{n}$) are in the same component iff i and j are in the same orbit of α : in other words $i\alpha^r = j\alpha^s$ for some positive integers r and s . The direction of arcs within a tree of the digraph of α is towards the root. Therefore if we adopt the convention that the cycles of the digraph are directed anti-clockwise, we may delete the arrows from the picture of the digraph, with the proviso that one-cycles (corresponding to fixed points) should be distinguished (by shading say) in order to avoid ambiguity. The graphical representation of members of T_n has been used effectively and often by Howie and others in studying the full transformation semigroup (see for example [7] and [8]). It is a natural visual aid, and concepts arising from members of T_n often have clear counterparts in the digraph. To illustrate this last point, consider the *stable range* of $\alpha \in T_n$, denoted by $\text{stran } \alpha$, which is defined by Howie as

$$\bigcap_{i=0}^{\infty} \nabla \alpha^i \quad (\nabla \alpha \text{ denotes the range of } \alpha),$$

in other words, $\text{stran } \alpha$ consists of all those points contained in the range of all the iterates of α . (We conventionally allow α^0 to represent the identity mapping on \bar{n}). A little reflection reveals that $i \in \text{stran } \alpha$ is equivalent to the statement that the vertex i of the

digraph is a point of one of the cycles. Indeed it follows that the descending chain $\nabla\alpha \supseteq \nabla\alpha^2 \supseteq \dots$, stabilises at $\text{stran } \alpha$, and that $\text{stran } \alpha = \nabla\alpha^i$, where i is least such that $\nabla\alpha^i = \nabla\alpha^{i+1}$. Moreover $\alpha|_{\text{stran } \alpha}$ is a permutation, called the *main permutation of α* in [3].

An obvious random variable associated with α is R_n , whose value is $|\nabla\alpha|$, the order of the range of α [9]. Two other numerical quantities associated with α are the order of the stable range of α , and the number of orbits of α . Denote the corresponding random variables by X_n and C_n respectively.

The value of ER_n is easily found. The selection of $\alpha \in T_n$ corresponds in an obvious way to the random tossing of n labelled balls into n labelled boxes. The probability that $i \notin \nabla\alpha$ equals the probability that box i remains empty, and this is plainly given by $(1 - 1/n)^n$. Since $ER_n = nP(i \in \nabla\alpha)$, where i is any member of \bar{n} , it follows that $ER_n = n(1 - (1 - 1/n)^n)$. This gives the first of the following trio of results.

$$\lim_{n \rightarrow \infty} \frac{ER_n}{n} = 1 - e^{-1}; \tag{A}$$

$$\lim_{n \rightarrow \infty} \frac{EX_n}{\sqrt{n}} = \sqrt{\pi/2}; \text{ and} \tag{B}$$

$$\lim_{n \rightarrow \infty} \frac{EC_n}{\log n} = \frac{1}{2}. \tag{C}$$

Results (B) and (C) will be derived in sections 2 and 3 respectively.

The study of X_n will be facilitated by the introduction of another random variable, Y_n . Let $i \in \bar{n}$. Define the *iterative range of i under α* as $\{i\alpha^k : k \geq 0\}$. Let the value of Y_n be the order of the iterative range of a randomly chosen $i \in \bar{n}$ under α . For temporary purposes, define another random variable Y'_n whose value is the order of the iterative range of a fixed member of \bar{n} , denoted by 1, under α . Clearly $Y_n \stackrel{\mathcal{D}}{=} Y'_n$, and we shall henceforth not distinguish them. Although Y_n is the more natural object of investigation, it is sometimes convenient to deal with Y'_n . Furthermore, by the *iterative range of α* , denoted by $i\alpha$, we shall mean the iterative range of 1 under α .

2. Asymptotic results for the stable and iterative ranges. Let X_n and Y_n denote the random variables as defined in the first section. The key result is that X_n and Y_n have the same distribution.

THEOREM 2.1. $X_n \stackrel{\mathcal{D}}{=} Y_n$ for all $n \geq 1$.

In the course of the proof we use the following identity.

LEMMA 2.2

$$\sum_{i=1}^k \binom{m}{i} \frac{(n - m + i)i!}{n^i} = m - \frac{m(m - 1) \dots (m - k)}{n^k} \quad (1 \leq k \leq m \leq n). \tag{1}$$

Proof. By induction on k . For $k = 1$ we get

$$\binom{m}{1} 1! \frac{(n-m+1)}{n} = m - \frac{m(m-1)}{n}, \text{ as required.}$$

In general, the left hand side of (1) is

$$\sum_{i=1}^{k-1} \binom{m}{i} \frac{(n-m+i)i!}{n^i} + \binom{m}{k} k! \frac{n-m+k}{n^k}, \text{ which inductively equals}$$

$$m - \frac{m(m-1)\dots(m-k+1)}{n^{k-1}} + \frac{m(m-1)\dots(m-k+1)(n-m+k)}{n^k},$$

which simplifies to the right hand side of (1).

LEMMA 2.3. Let $\alpha \in T_n$ with $|\text{stran } \alpha| = k$. Let t_k be the number of extensions β of $\alpha \mid \text{stran } \alpha$ to a member of T_n such that $\text{stran } \beta = \text{stran } \alpha$. Then

$$t_k = kn^{n-k-1} = \frac{k}{n} \cdot n^{n-k} \quad (2)$$

REMARK. This lemma is of independent interest. It says that the proportion of all possible extensions β of $\alpha \mid \text{stran } \alpha$ to a member of T_n such that $\text{stran } \beta = \text{stran } \alpha$ is exactly $\frac{k}{n}$, that is, is directly proportional to the order of the given stable range. Moreover, as will be shown in the proof, the use of the notation t_k is justified, as t_k depends on k , but not on the permutation $\alpha \mid \text{stran } \alpha$. This plausible fact can be verified independently with a little effort.

Proof of 2.3. We prove (2) by induction on $m = n - k$. For $m = 0$, statement (2) says $t_n = 1$, which is obvious. Now $|\{\beta \in T_n : \beta \mid \text{stran } \alpha = \alpha \mid \text{stran } \alpha\}| = n^{n-k}$ can be expressed as,

$$n^{n-k} = t_k + \binom{n-k}{1} 1! t_{k+1} + \dots + \binom{n-k}{i} i! t_{k+i} + \dots + \binom{n-k}{n-k} (n-k)! t_n. \quad (3)$$

In detail, the term $\binom{n-k}{i} i! t_{k+i}$ equals the number of mappings β such that $\text{stran } \beta \supset \text{stran } \alpha$, $\beta \mid \text{stran } \alpha = \alpha \mid \text{stran } \alpha$ and $|\text{stran } \beta| = k + i$. Rearranging statement (3) gives

$$\frac{t_k}{n^{n-k}} = 1 - \sum_{i=1}^{n-k} \binom{n-k}{i} i! \frac{t_{k+i}}{n^{n-k}} \quad (4)$$

By the inductive hypothesis we have

$$\frac{t_{k+i}}{n^{n-k-i}} = \frac{k+i}{n} (1 \leq i \leq n-k). \quad (5)$$

Using (5) we rewrite (4) as

$$\frac{t_k}{n^{n-k}} = 1 - \sum_{i=1}^m \binom{m}{i} i! \frac{(n-m+i)}{n^{i+1}}. \tag{6}$$

The lemma requires us to show that

$$\frac{t_k}{n^{n-k}} = \frac{k}{n}, \tag{7}$$

and from (6) we see that (7) is equivalent to

$$\sum_{i=1}^m \binom{m}{i} i! \frac{(n-m+i)}{n^i} = m. \tag{8}$$

But (8) follows from Lemma 2.2 upon putting $k = m$.

Proof of Theorem 2.1. We need to show that

$$|\{\alpha \in T_n : |\text{stran } \alpha| = k\}| = |\{\alpha \in T_n : |\text{it } \alpha| = k\}| \tag{9}$$

for all $1 \leq k \leq n$. The left hand side of (9) is given by

$$\binom{n}{k} k! t_k \tag{10}$$

while the right hand side is

$$\binom{n-1}{k-1} (k-1)! k n^{n-k}. \tag{11}$$

The equality of (10) and (11) now follows easily from Lemma 2.3. This completes the proof.

The probability distribution for Y_n , and thus for X_n , can be written down by inspection:

$$P(Y_n = k) = \frac{k}{n} \prod_{i=0}^{k-1} \left(1 - \frac{i}{n}\right) = \frac{k(n-1)!}{(n-k)! n^k}, \quad (1 \leq k \leq n) \tag{12}$$

As incidental corollaries of this we obtain two graphical enumerations.

COROLLARY 2.4 (Cayley). *The number of labelled trees on n vertices is n^{n-2} .*

Proof. Interpret the statement $P(X_n = 1) = P(Y_n = 1)$. This is equivalent to

$$|\{\text{labelled rooted trees on } n \text{ vertices}\}| n^{-n} = n^{-1},$$

whereupon Cayley's theorem follows, as there are n rooted and labelled trees for each labelled tree on n vertices.

Replacing 1 by k in the above argument gives a more general result which is probably well known.

COROLLARY 2.5. *The number of forests on n vertices with k rooted labelled trees is $\binom{n-1}{k-1}n^{n-k}$.*

Proof. A forest on n vertices with k rooted, labelled trees corresponds to a member of T_n whose stable range consists of k fixed points. The number of such maps is

$$\binom{n}{k}t_k = \binom{n-1}{k-1} \cdot \frac{n}{k} \cdot \frac{k}{n} \cdot n^{-k} = \binom{n-1}{k-1}n^{n-k}.$$

Unlike the expectation of $|\nabla\alpha|$ which increases linearly with n , the expected value of X_n increases as \sqrt{n} . Indeed it is an elementary matter to find the limiting distribution of $V_n = X_n/\sqrt{n} = Y_n/\sqrt{n}$. For $v \geq 0$,

$$P(Y_n \geq v\sqrt{n}) = \prod_{j=0}^i \left(1 - \frac{j}{n}\right) \tag{13}$$

where i is least such that $i \geq v\sqrt{n} - 1$. Now in general

$$\left(1 - \frac{j}{n}\right)\left(1 - \frac{i-j}{n}\right) = 1 - \frac{i}{n} + \frac{j(i-j)}{n^2} \tag{14}$$

Also $j(i-j)/n^2 \leq (i/2n)^2$. Hence by pairing the terms of (13) in the fashion suggested by (14), and deleting the middle term if i is odd, we obtain the inequality

$$P(V_n \geq v) \leq \left(1 - \frac{i}{n} + \frac{i^2}{4n^2}\right)^{i/2} = \left(1 - \frac{i}{2n}\right)^i.$$

Taking logarithms, and using the fact that $i < v\sqrt{n} \leq i + 1$ we obtain

$$\log P(V_n \geq v) \leq v\sqrt{n} \log\left(1 - \frac{v\sqrt{n}-1}{2n}\right) \tag{15}$$

Since $\log(1-x) \sim -x$ as $x \rightarrow 0$, we rewrite the right hand side of (15) as

$$\begin{aligned} & -v\sqrt{n} \frac{(v\sqrt{n}-1) \log(1 - (v\sqrt{n}-1)/2n)}{2n} \\ & = \left(-\frac{v^2}{2} + \frac{v}{2\sqrt{n}}\right) \frac{\log\left(1 - \frac{v\sqrt{n}-1}{2n}\right)}{-(v\sqrt{n}-1)/2n} \rightarrow -\frac{v^2}{2} \text{ as } n \rightarrow \infty. \end{aligned}$$

Taking exponentials, we obtain that provided the limit exists,

$$\lim_{n \rightarrow \infty} P(V_n \geq v) \leq e^{-v^2/2}. \tag{16}$$

On the other hand, from (14), it is certainly true that $\left(1 - \frac{j}{n}\right)\left(1 - \frac{i-j}{n}\right) \geq 1 - \frac{i}{n}$,

which gives

$$P(V_n \geq v) \geq \begin{cases} \left(1 - \frac{i}{n}\right)^{i/2} & \text{if } i \text{ is even} \\ \left(1 - \frac{i}{n}\right)^{i/2} \cdot \left(1 - \frac{(i+1)/2}{n}\right) & \text{if } i \text{ is odd.} \end{cases}$$

In any case we have $P(V_n \geq v) \geq (1 - (i/n))^{(i+2)/2}$. Again we take logarithms, use the facts that $i < v\sqrt{n} \leq i + 1$ and $\log(1 - x) \sim -x$ as $x \rightarrow 0$ to obtain

$$\lim_{n \rightarrow \infty} P(V_n \geq v) \leq e^{-v^2/2} \tag{17}$$

provided the limit exists. Combining (16) and (17) gives us the limiting distribution of V_n . Indeed the following is true.

THEOREM 2.6. *The sequence of random variables V_n defined above approaches in distribution the random variable V , with distribution function $F(v) = 1 - e^{-v^2/2}$, $v \geq 0$. Moreover the moments of the V_n approach those of V , that is*

$$\lim_{n \rightarrow \infty} E(V_n^k) = E(V^k) = \int_0^\infty v^{k+1} e^{-v^2/2} dv, \text{ and in particular } \lim_{n \rightarrow \infty} E(V_n) = \sqrt{\pi/2}.$$

It remains to prove the statement about moments which is not immediate, as in general the moments of a sequence of random variables do not approach the moments of the limiting distribution. In order to complete the proof we need to establish the uniform integrability of the sequence $\{V_n^k\}$, for all $k \geq 1$. A sufficient condition for this is that $E V_n^k \leq C_k$, where C_k is independent of n . (For details see [1] p. 32 or [2] pp. 89–91.). Now

$$\begin{aligned} E X_n^k &= \sum_{r=1}^\infty r^k P(X_n = r) = \sum_{r=1}^\infty r^k [P(X_n > r - 1) - P(X_n > r)] \\ &= 1^r \cdot P(X_n > 0) + \sum_{r=1}^\infty [(r + 1)^k - r^k] P(X_n > r). \end{aligned}$$

Next we use the inequality $a^k - b^k \leq (a - b)ka^{k-1}$ for $a > b$, which can be easily verified by writing $a = b + h$ and expanding both sides. We thus have

$$E X_n^k \leq 1 + \sum_{r=1}^\infty k(r + 1)^{k-1} P(X_n > r).$$

Invoking the inequality $1 - x \leq e^{-x}$ for all $x \geq 0$ gives

$$P(X_n > r) = \prod_{j=1}^r \left(1 - \frac{j}{n}\right) \leq \exp\left(-\sum_{j=1}^r \frac{j}{n}\right) = \exp\left(-\frac{r(r+1)}{2n}\right).$$

Thus

$$EX_n^k - 1 \leq k \sum_{r=1}^{\infty} (r+1)^{k-1} \exp\left(-\frac{r(r+1)}{2n}\right).$$

Since $V_n = X_n/\sqrt{n}$, our task is to bound

$$kn^{-k/2} \sum_{r=1}^{\infty} (r+1)^{k-1} \exp\left(-\frac{r(r+1)}{2n}\right)$$

by a constant which is independent of n . Clearly it suffices to accomplish this for

$$\begin{aligned} S_k &= n^{-k/2} \sum_{r=1}^{\infty} r^{k-1} \exp\left(-\frac{r(r+1)}{2n}\right) \\ &= n^{-k/2} e^{1/8n} \sum_{r=1}^{\infty} r^{k-1} \exp\left(-\frac{(r+1/2)^2}{2n}\right). \end{aligned}$$

Now

$$\begin{aligned} \int_{(r-1/2)/\sqrt{n}}^{(r+1/2)/\sqrt{n}} (x+1)^{k-1} e^{-x^2/2} dx &\geq \int_{(r-1/2)/\sqrt{n}}^{(r+1/2)/\sqrt{n}} \left(x + \frac{1}{\sqrt{n}}\right)^{k-1} e^{-x^2/2} dx \\ &> \frac{1}{\sqrt{n}} \left(\frac{r+1/2}{\sqrt{n}}\right)^{k-1} \exp\left(-\frac{(r+1/2)^2}{2n}\right) > n^{-k/2} r^{k-1} \exp\left(-\frac{(r+1/2)^2}{2n}\right). \end{aligned}$$

Hence

$$\begin{aligned} S_k &< e^{1/8n} \sum_{r=1}^{\infty} \int_{(r-1/2)/\sqrt{n}}^{(r+1/2)/\sqrt{n}} (x+1)^{k-1} e^{-x^2/2} dx \\ &< e^{1/8} \int_0^{\infty} (x+1)^{k-1} e^{-x^2/2} dx = C_k, \quad \text{a constant independent of } n. \end{aligned}$$

The final assertion that EV_n approaches $\sqrt{\pi/2}$ follows upon evaluating $\int_0^{\infty} v^2 e^{-v^2/2} dv$ by parts. Indeed all the limiting values of the moments of the V_n can be explicitly calculated. This completes the proof of the theorem.

3. The distribution of the component number. Let S_n denote the symmetric group on n symbols, that is the group of all permutations on $\{1, 2, \dots, n\}$. Let τ_n denote the random variable whose value is the cycle number of a random $\pi \in S_n$. It is shown in [4] that $\tau_n \stackrel{\text{d}}{=} X_1 + X_2 + \dots + X_n$ where the X_i are independent indicator random variables (i.e. their only value is 0 or 1) with $P(X_i = 0) = \frac{n-i}{n-i+1}$, ($1 \leq i \leq n$). From this it follows that

$$E\tau_n = \sum_{i=1}^n \frac{1}{n} \sim \log n$$

and

$$\text{Var } \tau_n = \sum_{i=1}^n \frac{n-i}{(n-i+1)^2} \sim \log n.$$

Consider the random variable C_n , whose value is the component number of a random $\alpha \in T_n$, or equivalently the cycle number of the main permutation of α . Now

$$\begin{aligned} P(C_n = k | X_n = i) &= \frac{|\{\alpha \in T_n : |\text{stran } \alpha| = i, |\text{cycles of stran } \alpha| = k\}|}{|\{\beta \in T_n : |\text{stran } \beta| = i\}|} \\ &= \frac{\binom{n}{i} t_i |\{\pi \in S_i : |\text{cycles of } \pi| = k\}|}{\binom{n}{i} i! t_i} \\ &= P(\tau_i = k). \end{aligned}$$

Another way of starting this is

PROPOSITION 3.1. $C_n \stackrel{\mathcal{D}}{=} \tau_{X_n}$.

Since $E\tau_n \sim \log n$ and $EX_n = O(\sqrt{n})$ one might conjecture that $EC_n \sim \log \sqrt{n} = \frac{1}{2} \log n$, and indeed it is possible to show by an elementary argument that

$$\lim_{n \rightarrow \infty} \frac{EC_n}{\log n} = \frac{1}{2}.$$

However more can be said about C_n from known facts about the limiting distribution of τ_n . It is stated in [4, p. 258] that

$$Z_n = \frac{\tau_n - \log n}{\sqrt{\log n}} \stackrel{\mathcal{D}}{\rightarrow} N(0, 1). \quad (17)$$

It can be verified that Z_n can be written as the sum of n independent random variables each with zero mean: $Z_n = \sum_{i=1}^n Y_i$ with $EY_i = 0$, $\text{var } Y_i = \sigma_i^2$. Writing s_n^2 for $\sum_{i=1}^n \sigma_i^2 = \text{var } Z_n$,

it can then be verified that $\lim_{n \rightarrow \infty} s_n^{-2-\delta} \sum_{i=1}^n E(|Y_i|^{2+\delta}) = 0$, for some fixed $\delta > 0$ ($\delta = 2$ suffices), which is sufficient in order to draw the conclusion of (17), see [2] p. 191.

Let us return to our investigation of the component number. Write $X_n = n^{1/2}V_n$, where the distribution of V_n approaches that of V given in Theorem 2.6. Write $\tau_i = (\log i)^{1/2}U_i + \log i$ with $U_i \stackrel{\mathcal{D}}{\rightarrow} N(0, 1)$. Hence, by 3.1,

$$C_n \stackrel{\mathcal{D}}{=} \tau_{X_n} \stackrel{\mathcal{D}}{=} U_{X_n} (\log X_n)^{1/2} + \log X_n = \frac{1}{2} \log n + \log V_n + U_{X_n} (\frac{1}{2} \log n + \log V_n)^{1/2}.$$

Therefore

$$(\frac{1}{2} \log n)^{-1/2} (C_n - \frac{1}{2} \log n) = (\frac{1}{2} \log n)^{-1/2} \log V_n + U_{X_n} (1 + (\frac{1}{2} \log n)^{-1/2} \log V_n)^{1/2};$$

and since $(\frac{1}{2} \log n)^{-1/2} \log V_n \xrightarrow{P} 0$ and $X_n \rightarrow \infty$ as $n \xrightarrow{P} \infty$, this yields

THEOREM 3.2. As $n \rightarrow \infty$,

$$(\frac{1}{2} \log n)^{-1/2} (C_n - \frac{1}{2} \log n) \stackrel{\mathcal{D}}{\rightarrow} N(0, 1).$$

Also, since all moments of U_j and V_n converge to those of their limit distribution, it can be shown that all moments of the quotient in Theorem 3.2 converge to those of $N(0, 1)$ using the concavity of the logarithm function.

REFERENCES

1. P. Billingsley, *Convergences of Probability Measures*, (Wiley, 1968).
2. K. L. Chung, *A Course in Probability Theory*, (Harcourt, Brace and World, 1968).
3. J. Denes, 'Some Combinatorial Properties of Transformations and their Connections with the Theory of Graphs.' *J. Comb. Thry.* **9** (1969) 108–116.
4. W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd ed., (Wiley and Sons, 1968).
5. F. Harary, *Graph Theory*, (Addison-Wesley, 1969).
6. B. Harris, The Asymptotic Distribution of the Order of Elements in Symmetric Semigroups, *J. Comb. Thry (A)* **15** (1973) 66–74.
7. P. M. Higgins, A method for constructing square roots in finite full transformation semigroups, *Canad. Math. Bull.* **29** (1986).
8. J. M. Howie, Idempotent generators in finite full transformation semigroups, *Proc. Roy. Soc. Edin.* **81A** (1978) 317–323.
9. K. H. Kim and F. W. Roush, The average rank of a product of transformations, *Semigroup Forum* **19** (1980), 79–85.

BRUCE BROWN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TASMANIA
SANDY BAY, TASMANIA 7001
AUSTRALIA

PETER M. HIGGINS
MATHEMATICS SECTION
DEAKIN UNIVERSITY
VICTORIA, 3217
AUSTRALIA.

The second named author acknowledges the support of a Deakin University Post Doctoral Fellowship.