LONG CYCLES IN ABC–PERMUTATIONS

IGOR PAK* AND AMANDA REDLICH*

ABSTRACT. An *abc-permutation* is a permutation $\sigma_{abc} \in S_n$ obtained by exchanging an initial block of length a and a final block of length c of $\{1, \ldots, n\}$, where $n = a+b+c$. In this note we compute the limit of the probability that a random *abc*-permutation is a long cycle. This resolves Arnold's open problem [A, p. 144].

1. INTRODUCTION

For every $n = a + b + c$, an *abc-permutation* $\sigma_{abc} \in S_n$ is defined as

$$\sigma_{abc} \,=\, (a+b+1, a+b+2, \ldots, n, a+1, a+2, \ldots, a+b, 1, 2, \ldots, a).$$

Denote by $\Omega(n)$ the set of such permutations. A *long cycle* is a permutation $\sigma \in S_n$ with one cycle of length n. Denote by $\Lambda(n) \subseteq \Omega(n)$ the set of *abc*-permutations that are long cycles, and by

$$\mathbf{p}(n) \,:=\, \mathrm{P}\big(\sigma_{abc} \in \Lambda(n)\big) \,=\, \frac{|\Lambda(n)|}{|\Omega(n)|}$$

the probability that a random permutation is a long cycle. The main result of this paper is the following theorem:

Main Theorem. *The probability* $\mathbf{p}(n) \to 6/\pi^2$ *as* $n \to \infty$.

The proof of the main theorem is based on the following criterion, an explicit formula for $p(n)$ which we present in Section 3, and on a asymptotic analysis given in Section 4.

Lemma 1. *An abc-permutation* $\sigma_{abc} \in \Lambda(n)$ *if and only if* $(a+b, b+c) = 1$.

To better understand the implications of the lemma, recall the problem of finding the probability P that two "random" integers are relatively prime. Setting aside a formal definition of P, this probability is equal to the probability that no prime p divides both integers, so we obtain:

$$P \,=\, \prod_p \left(1 - \frac{1}{p^2}\right) \,=\, \zeta(2)^{-1} \,=\, \frac{6}{\pi^2}\,,$$

where here and throughout the paper p will denote a prime. Now, heuristically, the main theorem is saying that integers $a+b$ and $b+c$ are "nearly random", so the probability that they are relatively prime is the same as that of two "random" integers. Making this argument precise occupies most of the paper.

Let us say a few words on the history of the problem. The *abc*-permutations can be viewed as discrete analogues of *interval exchange transformations*, which play an important role in ergodic theory. These transformations go back to one of the first Arnold's problems [A, p. 2, 182–183] and were first studied in the 1960s in [Os, KS] (see also [Ke]). An interesting and important connection to combinatorics of Sturmian words was discovered by Rauzy in [R]. In recent years, significant advancements in the study of asymptotics of generic interval exchange transformations

Date: May 8, 2007.

*Department of Mathematics, MIT, Cambridge, MA 02139; {pak,aredlich}@math.mit.edu.

have been made in [Ko, Z], and for the three-interval transformations a detailed analysis has appeared in [FHZ1, FHZ2].

In the language of exchange transformations, this paper studies the asymptotic behavior of the number of transitive orbits of rational three-interval transformations. The problem, phrased in terms of *abc*-permutations, was formulated by Arnold in 2002 [A, pp. 144, 626]. Interestingly, numerical investigations reported by Arnold do not seem to suggest that $\mathbf{p}(n)$ is converging. This makes the (possibly, very slow) convergence in the main theorem even more surprising.

2. Proof of Lemma 1

Clearly, σ_{abc} is a long cycle if and only if $\sigma_{cba} = \sigma_{abc}^{-1}$ is a long cycle. Throughout the paper, we assume that $c \geq a$. We use the standard notation $(k, l) = \gcd(k, l)$. When a, b, c are clear, we use σ in place of σ_{abc}.

To prove that $(a + b, b + c) = 1$ is necessary, consider the orbit of a single element x under $\sigma_{abc} = \sigma$. As stated above, we will assume (without loss of generality) that $a \leq c$. Let $c - a = d$. Suppose $x \leq a$. Then

$$\sigma(x) = c + b + x\,.$$

Applying σ again brings us back to

$$\sigma^2(x) = \sigma(x + c + b) = c + b + x - a - b = x + d\,.$$

In general, while $x + (i - 1)d \leq a$,

$$\sigma^{2i+1}(x) = x + id + c + b \quad \text{and} \quad \sigma^{2i}(x) = x + id\,.$$

Similarly,

$$\sigma(x) = x + c - a = x + d \quad \text{if} \quad a < x \leq a + b,$$

and

$$\sigma^i(x) = x + id \quad \text{while} \quad x + (i - 1)d \leq a + b.$$

Also

$$\sigma(x) = x - a - b \quad \text{for} \quad a + b < x,$$

and

$$\sigma^i(x) = x - i(a + b) \quad \text{while} \quad (i - 1)(a + b) \leq x - a - b.$$

Therefore, for every element $x \in \{1, \ldots, n\}$, the intersection of its orbit with $\{1, \ldots, a + b\}$ contains only elements of the form $x + id - j(a + b)$. If $m := (d, a + b) = (a + b, b + c) \neq 1$, then x's orbit contains at most one equivalence class modulo m within $\{1, \ldots, a + b\}$. Therefore, σ is a long cycle only if $(a + b, b + c) = 1$.

To prove that $(a+b, b+c) = 1$ is sufficient, consider the orbit of an element $y < a$ under iterations of σ. As above, $\sigma^k(y) < a$ implies that

$$\sigma^k(y) = y + id - j(a + b)\,,$$

for some i and j. Suppose $(d, a + b) = 1$ and y is in a cycle of length k in σ. Then

$$\sigma^k(y) = y + id - j(a + b) = y\,,$$

which gives

$$id - j(a + b) = 0\,.$$

If $(c - a, a + b) = (d, a + b) = 1$, then $a + b \mid i$ and $d \mid j$. Therefore, $i \geq a + b$ and $j \geq d$. Since $i + j \leq k$, we obtain

$$d + a + b = b + c \leq k\,.$$

By assumption, $c \geq a$, and $c \neq a$ from above. Thus, $c > a$ and $k > n/2$. A similar argument for $y > a$ shows its orbit has length greater than $n/2$ as well. Therefore, every cycle in σ has length greater than $n/2$ and σ is a long cycle. This completes the proof. $\square$

3. Exact expression for $\mathbf{p}(n)$

The main result of this section is the following explicit formula for the desired probability $\mathbf{p}(n)$.

Lemma 2.

$$\mathbf{p}(n) = \prod_{p<n,\ p\nmid n} \left(1 - \frac{\lfloor \frac{n}{p} \rfloor \big(\lfloor \frac{n}{p} \rfloor + 1\big)}{n(n+1)}\right) \prod_{p<n,\ p\mid n} \left(1 - \frac{\big(\frac{n}{p}+1\big)\big(\frac{n}{p}+2\big)}{n(n+1)}\right)$$

Proof. By Lemma 1, we need to compute the probability $(a+b, b+c) = 1$. Find this by looking at the probability that $a+b$ and $b+c$ have some common divisor k. Suppose $k \mid (b+c)$ and $k \mid (a+b)$. Since $a+b+c = n$, that implies

$$c = a = n \mod k.$$

Let $c' = n - c$ and $a' = n - a$. Then the original conditions become

$$2n \geq a' + c' \geq n \quad \text{and} \quad c' = a' = 0 \mod k.$$

To count how many a', c' pairs are valid, fix $a' = ik$ and count valid c'. Valid c' are those such that $n - ik \leq c' \leq n$ and $c' = 0 \mod k$. Therefore, for each $a' = ik$, there are i possibilities for c' when $k \nmid n$ and $i+1$ possibilities when $k \mid n$. Thus, there are

$$\sum_{i=0}^{\lfloor n/k \rfloor} i = \frac{\lfloor \frac{n}{k} \rfloor \big(\lfloor \frac{n}{k} \rfloor + 1\big)}{2} \quad \text{pairs } (a', c') \text{ for each } k \nmid n, \text{ and}$$

$$\sum_{i=0}^{n/k} (i+1) = \frac{\big(\frac{n}{k}+1\big)\big(\frac{n}{k}+2\big)}{2} \quad \text{pairs when } k \mid n.$$

Let $f(n,k)$ be the probability that $(b+c)$ and $(a+b)$ are both divisible by an integer k. From above,

$$f(n,k) = \frac{\lfloor \frac{n}{k} \rfloor \big(\lfloor \frac{n}{k} \rfloor + 1\big)}{n(n+1)} \quad \text{for} \quad k \nmid n,$$

and

$$f(n,k) = \frac{\big(\frac{n}{k}+1\big)\big(\frac{n}{k}+2\big)}{n(n+1)} \quad \text{for} \quad k \mid n.$$

Finally, use the inclusion-exclusion principle to obtain

$$\mathbf{p}(n) = \frac{\Lambda(n)}{\Omega(n)} = 1 - \sum_{k=2}^{n} \mu(k)\, f(n,k) = \prod_{p \leq n} \big(1 - f(n,p)\big),$$

where μ is the Möbius function. This completes the proof. $\square$

4. Proof of the Main Theorem

In notation of the proof of Lemma 2, write $\mathbf{p}(n) = A(n) \cdot B(n)$, where

$$A(n) \;=\; \prod_{p \le \log n} \big(1 - f(n,p)\big)\,, \qquad B(n) \;=\; \prod_{\log n < p \le n} \big(1 - f(n,p)\big),$$

and $\log n$ denotes the natural logarithm. We estimate each term separately.

First, for $p = O(\log n)$, we have

$$1 - f(n,p) \;=\; \left(1 - \frac{1}{p^2}\right)\left(1 + O\left(\frac{p}{n}\right)\right) \quad \text{as } \; n \to \infty.$$

Since

$$\prod_{p < \log n} \left(1 + O\left(\frac{p}{n}\right)\right) \;=\; 1 + O\left(\frac{\log^2 n}{n}\right),$$

we obtain

$$A(n) \;=\; \prod_{p \le \log n} \big(1 - f(n,p)\big) \;=\; \prod_{p \le \log n} \left(1 - \frac{1}{p^2}\right) \prod_{p \le \log n} \left(1 + O\left(\frac{p}{n}\right)\right) \;=\; \frac{6}{\pi^2}\big(1 + o(1)\big).$$

Similarly, observe that $B(n) \le 1$, and that for all $p \le n$, we have

$$f(n,p) \;\le\; \frac{n + 5p}{p^2(n+1)}\,.$$

Therefore, $B(n) \ge F(n)$, where

$$F(n) \;=\; \prod_{\log n < p \le n} \left(1 - \frac{n+5p}{p^2(n+1)}\right) \;\le\; \prod_{\log n < p \le n} \big(1 - f(n,p)\big).$$

We will show that $F(n) \to 1$ as $n \to \infty$. Let $m = \lceil \log(n/\log n) \rceil$ and

$$F_i(n) \;=\; \prod_{e^{i-1}\log n \le p < e^i \log n} \left(1 - \frac{n+5p}{p^2(n+1)}\right), \quad \text{for } \; 1 \le i \le m.$$

Then

$$F(n) \;\ge\; \prod_{i=1}^{m} F_i(n) \ge \big(F_{j_n}(n)\big)^m,$$

where j_n is the index such that $F_i(n)$ is minimized.

For each F_i, the number of terms in its product is the number of primes between $e^{i-1}\log n$ and $e^i \log n$, i.e. $\pi(e^i \log n) - \pi(e^{i-1}\log n)$. Recall that $\pi(n) \sim n/\log n$, and note that $e^{j_n}\log n \to \infty$ as $n \to \infty$. Therefore, as $n \to \infty$, the number of terms in F_{j_n} is equal to

$$e^{j_n - 1}\log n \,\frac{\log\log n + j_n - 2}{(\log\log n + j_n)(\log\log n + j_n - 1)}\,\big(1 + o(1)\big).$$

Finally, note that the smallest term in each F_i is the first term

$$\left(1 - \frac{n + 5e^{i-1}\log n}{(e^{i-q}\log n)^2(n+1)}\right).$$

This implies that

$$F(n) \ge \left(1 - \frac{n + 5e^{i-1}\log n}{(e^{i-1}\log n)^2(n+1)}\right)^{\log(n/\log n)\, e^{j_n - 1}\log n \frac{\log\log n + j_n - 2}{(\log\log n + j_n)(\log\log n + j_n - 1)}\,(1 + o(1))}.$$

A direct calculation shows that the r.h.s. $\to 1$ as $n \to \infty$. Together with $F(n) \le B(n) \le 1$, this implies that $B(n) \to 1$. Therefore, $\mathbf{p}(n) = A(n) \cdot B(n) \to 6/\pi^2$ as $n \to \infty$, which completes the proof. $\square$

5. Final remarks and open problems

5.1. There is a striking similarity between our Lemma 1 and the Proposition 4 in [AMP]. In fact, one can think of the former as of a discrete analogue of the latter. It would be interesting to make this observation precise.

5.2. One can consider random permutations corresponding to a given pattern, defined as fixed permutation of k blocks. These would be natural discrete analogues of k interval exchange transformations. We conjecture that for a given pattern there is always a limit as in the main theorem. It would be nice to see if these limits can be computed exactly.

5.3. By analogy with the continuous case, for every *abc*-permutation $\sigma = \sigma_{abc} \in S_n$ one can define a word $w_1 \dots w_n$ in the alphabet $\{A, B, C\}$ corresponding to the orbit of element 1. Formally, let $w_k = A$ if $\sigma^k(1) \le a$, let $w_k = B$ if $a < \sigma^k(1) \le a + b$, and let $w_k = C$ otherwise, for all $1 \le k \le n$. The study of asymptotic behavior of various statistics on these words would be a natural approach to understand the behavior of random *abc*-permutations (cf. [L]).

Acknowledgments. We are grateful to Luca Zamboni for useful remarks and help with the references. The first named author was partially supported by the NSF.

References

[AMP] P. Ambrož, A. Masáková and E. Pelantová, Matrices of 3iet preserving morphisms, `arXiv: math.CO/0702336`.

[A] V. I. Arnold, *Arnold's problems*, Springer, Berlin, 2004.

[FHZ1] S. Ferenczi, C. Holton and L. Q. Zamboni, Structure of three interval exchange transformations. I. An arithmetic study, *Ann. Inst. Fourier (Grenoble)* **51** (2001), no. 4, 861–901.

[FHZ2] S. Ferenczi, C. Holton and L. Q. Zamboni, Structure of three-interval exchange transformations. II. A combinatorial description of the trajectories, *J. Anal. Math.* **89** (2003), 239–276.

[KS] A. B. Katok and A. M. Stepin, Approximations in ergodic theory, *Uspehi Mat. Nauk* **22** (1967), no. 5, 81–106.

[Ke] M. Keane, Interval exchange transformations, *Math. Z.* **141** (1975), 25–31.

[Ko] M. Kontsevich, Lyapunov exponents and Hodge theory, in *The mathematical beauty of physics (Saclay, 1996)*, 318–332, World Sci., River Edge, NJ, 1997.

[L] M. Lothaire, *Combinatorics on words*, Cambridge University Press, Cambridge, 1997.

[Os] V. I. Oseledec, The spectrum of ergodic automorphisms, *Dokl. Akad. Nauk SSSR* **168** (1966), 1009–1011.

[R] G. Rauzy, Échanges d'intervalles et transformations induites (in French), *Acta Arith.* **34** (1979), no. 4, 315–328.

[Z] A. Zorich, How do the leaves of a closed 1-form wind around a surface?, in *Pseudoperiodic topology*, 135–178, *AMS Transl. Ser. 2*, vol. **197**, AMS, Providence, RI, 1999.