An Upper Bound for the Number of Complete Mappings

Colin Cooper [*] Igor N. Kovalenko [†]

October 16, 2005

Abstract

A *complete mapping* on a set G with binary operation $\circ$ is a bijection $\theta : G \to G$ such that the mapping $\eta : G \to G$ defined by $\eta(x) = x \circ \theta(x)$ is again a bijection.

We give an asymptotic upper bound of $\exp\{-0.08854n\}$ for the proportion of permutations in S_n which form complete mappings under addition modulo n.

1 Introduction

Following the notation of Dénes and Keedwell [2], we define a *complete mapping* on a set G with binary operation $\circ$ as a bijection $\theta : G \to G$ such that the mapping $\eta : G \to G$ defined by $\eta(x) = x \circ \theta(x)$ is again a bijection.

We restrict our attention here to the case where $(G, \circ)$ is $(\mathbb{Z}_n, +)$ the group of integers under addition modulo n. For example, in $(\mathbb{Z}_5, +)$, the mapping θ, given below, is complete with associated mapping $\eta(x) = (x + \theta(x))(\text{mod } 5)$, given by respectively

$$\theta = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix} \text{ and } \eta = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 1 & 3 \end{pmatrix}.$$

The set of bijections from $\mathbb{Z}_n$ to $\mathbb{Z}_n$ is S_n, the set of permutations on $\{0, 1, \ldots, n-1\}$. We ask the following natural question:

If $G_n = \{\sigma \in S_n : \sigma \text{ is a complete mapping on } (\mathbb{Z}_n, +)\}$, what can we say about $|G_n|/|S_n|$?

Any group G of odd order has a complete mapping (see Theorem 1.4.3 of [2]) this being a consequence of the fact that every element of G has a unique square root. In the case of

[*]School of Mathematical Studies, University of North London, London N7 8DB. Research performed while on secondment to the STORM Research Group, UNL.

[†]Academician, Academy of Sciences of Ukraine. Research performed whilst visiting the STORM Research Group, UNL.

$(\mathbb{Z}_n, +)$, n odd, it is trivial to construct n complete maps based on the identity permutation, by cyclically rotating the bottom row.

When n is even, there are no complete mappings on $(\mathbb{Z}_n, +)$ as may be seen from the following theorem of I.J. Paige [4] (quoted from Theorem 1.4.5 of [2])

Theorem 1 *If $(G, \circ)$ is a finite group of order n which has a complete mapping, there exists an ordering of its elements, say, $a_1, a_2, \ldots, a_n$ such that $a_1 a_2 \cdots a_n = e$, where e is the identity of G.*

Considering $(\mathbb{Z}_n, +)$ we can only have $\frac{1}{2}n(n-1) \equiv 0 (\mathrm{mod}\ n)$ if n divides $\frac{1}{2}n(n-1)$, which is true only when n is odd.

Our main result is the following upper bound.

Theorem 2 *There exists a constant $c \geq 0.08854$ such that for sufficiently large n,*

$$\frac{|G_n|}{|S_n|} \leq \exp\{-cn\}.$$

We offer the reader two proofs of this theorem. The first uses a martingale inequality to establish a slightly weaker value of 0.06766 for c. A subsequent analysis from first principles improves the value of c to 0.08854. We are at present unable to further improve this bound, although intuitive considerations of Poisson occupancy would suggest that c is at least 1.

For the purposes of discussion, we refer to the set G_n as the set of *good* permutations.

2 Proof that c is at least 0.06766

The proof is an application of a martingale inequality which may be found (for example) in the survey paper by McDiarmid [3]. For completeness we give the necessary definitions and statement of the inequality.

Let (V, d) be a finite metric space. A *partition sequence* $((\mathcal{P}_k, c_k) : k = 0, .., n)$ consists of a sequence $\mathcal{P}_0, \mathcal{P}_1, ..., \mathcal{P}_n$ of *increasingly refined partitions* of V, starting with the trivial partition $\mathcal{P}_0$ with single subset V, and ending with the discrete partition $\mathcal{P}_n$ of V into singleton sets, and with the general property that if $A \in \mathcal{P}_k$ then $A \subseteq B \in \mathcal{P}_{k-1}$.
The sequence $c_0, c_1, ..., c_k, ..., c_n$ of positive reals has the following property. For each $k, (k = 1, ..., n)$, whenever $A, B \in \mathcal{P}_k$ and $A, B \subseteq C \in \mathcal{P}_{k-1}$ for some C, then there is a bijection $\phi : A \to B$ with $d(x, \phi(x)) \leq c_k$ for all $x \in A$.

For the case we consider here, $V = S_n$, and if $\alpha, \beta \in S_n$, where $\alpha = (\alpha(i) : i = 1, ..., n)$ we use the metric

$$d(\alpha, \beta) = \sum_{j:\alpha(j) \neq \beta(j)} 1.$$

Two permutations α, β are placed in the same subset A in the partition $\mathcal{P}_k$ of S_n if they are identical on the first k entries. In other words $\alpha(1) = \beta(1), ..., \alpha(k) = \beta(k)$. Clearly such a partition is increasingly refined, for if $A, B \in \mathcal{P}_k$ where $A \neq B$ but $A, B \subset C \in \mathcal{P}_{k-1}$, then the elements of α and β differ for the first time at the $k-$th entry. Of course every $\alpha \in A$ has identical value $\alpha(k)$, and similarly for every $\beta \in B$.

Any $\alpha \in A$ is such that $\alpha(i) = \beta(k)$ for some $i > k$, and we define $\phi(\alpha)$ to be that $\beta \in B$ for which the elements $\alpha(k), \alpha(i)$ are transposed and all other elements remain unaltered. But $d(\alpha, \phi(\alpha)) \leq 2$, as α, and $\phi(\alpha)$ are identical apart from the stated transposition, and thus $c_k = 2$.

Theorem 3 *Suppose that the finite metric space (V, d) has a partition sequence $((\mathcal{P}_k, c_k) : k = 0, ..., n)$. Let the function f on V satisfy $|f(x) - f(y)| \leq d(x, y)$ for all $x, y \in V$. Let X be uniformly distributed over V. Then for any $t > 0$,*

$$\Pr(f(X) - E(f(X)) \geq t) \leq \exp\{-2t^2 / \sum c_k^2\}.$$

In our case let $J_\alpha = \{j \in \{0, 1, ..., n-1\} : i + \alpha(i) = j\}$, and set $f(\alpha) = |J_\alpha|$. Note that S_n has the uniform (counting) measure. We must show that our choice of f satisfies

$$|f(\alpha) - f(\beta)| \leq d(\alpha, \beta), \tag{1}$$

for any choice of α, β. Denote $J_\alpha \cap J_\beta$ by $J_{\alpha\beta}$ and $J_\alpha \cap \overline{J}_\beta$ by $J_{\alpha\overline{\beta}}$. Let $j \in J_{\alpha\overline{\beta}}$. Then for some i we have $i + \alpha(i) = j$, and obviously $i + \beta(i) \neq j$. Thus $\alpha(i) \neq \beta(i)$. The indices i are distinct for $j \in J_{\alpha\overline{\beta}}$, and therefore

$$|J_{\alpha\overline{\beta}}| \leq d(\alpha, \beta).$$

On the other hand,

$$f(\alpha) - f(\beta) = |J_{\alpha\overline{\beta}}| - |J_{\overline{\alpha}\beta}| \leq |J_{\alpha\overline{\beta}}|.$$

We conclude that

$$f(\alpha) - f(\beta) \leq d(\alpha, \beta),$$

and symmetry implies the truth of (1).

The result that $E(|J|)$ is asymptotic to $n(1 - e^{-1})$ follows from the result that the number of $i + \alpha(i)$ equal to $0 (\text{mod } n)$ is asymptotically Poisson with parameter 1. This may be seen by considering k-th factorial moments or appealing directly to Theorem 4A of [1] which bounds the error term in this approximation by $2/n$.

Thus

$$\begin{aligned} \Pr(\alpha \in G_n) &= \Pr(|J_\alpha| = n) \\ &\leq \Pr\left([|J_\alpha| - E(|J|)] \geq ne^{-1}(1 + o(1))\right) \\ &\leq \exp\{-(\frac{e^{-2}}{2})(1 + o(1))n\}, \end{aligned}$$

as required.

3 Proof that c is at least 0.08854

We first describe a method for generating permutations by sampling without replacement. Choose (x_1, y_1) at random so that

$$\Pr(x_1 = i, y_1 = j) = \frac{1}{n^2} \qquad 0 \le i, j \le n-1,$$

and set $X_1 = \{x_1\}, Y_1 = \{y_1\}$. In general, choose x_k from $\{0, ..., n-1\} \setminus X_{k-1}$, where $X_{k-1} = \{x_1, ..., x_{k-1}\}$, and similarly for y_k. Thus

$$\Pr\left((x_k, y_k) = (i, j) \mid X_{k-1}, Y_{k-1}\right) = \frac{1}{(n-(k-1))^2},$$

for $i \notin X_{k-1}, j \notin Y_{k-1}$. As long as no confusion arises we will identify the set X_k with the vector $(x_1, ..., x_k)$.

Set $z_k = x_k + y_k (\mathrm{mod}\, n)$, and say that a failure occurs at time k if $z_k \in Z_{k-1}$, where $Z_{k-1} = \{z_1, ..., z_{k-1}\}$.

If no failure occurs at times $1, ..., n$ then the permutation is good. Evidently,

$$|G_n| = n!\, E\left[p_2(X_1, Y_1) p_3(X_2, Y_2) ... p_n(X_{n-1} Y_{n-1})\right],$$

where $p_k(X_{k-1}, Y_{k-1})$ is the conditional probability of success at time k given (X_{k-1}, Y_{k-1}).

The entire set $I = \{(i, j) : 0 \le i, j \le n-1\}$ consists of *admissable* and *inadmissable* positions for the new point (x_k, y_k). It is easy to see there are $(n-(k-1))^2$ admissable positions (i, j). We now estimate how many of these admissable positions satisfy $i + j (\mathrm{mod}\, n) \notin Z_{k-1}$.

For a point $(i, j) \in I$ each of the following three events may occur,

$$\begin{aligned} A &= \{i \in X_{k-1}\} \\ B &= \{j \in Y_{k-1}\} \\ C &= \{i + j (\mathrm{mod}\, n) \in Z_{k-1}\}. \end{aligned}$$

Denote by $N(A)$ then number of points in the set A, and similarly for $N(B), N(C)$ etc. We observe that

$$p_k(X_{k-1}, Y_{k-1}) = \frac{n^2 - N(A \cup B \cup C)}{(n-k+1)^2}.$$

By the Inclusion-Exclusion principle, we have

$$N(A \cup B \cup C) = N(A) + N(B) + N(C) - N(AB) - N(BC) - N(AC) + N(ABC).$$

All the terms except the last one are easy to obtain precisely, namely

$$N(A) = N(B) = N(C) = n(k-1),$$

$$N(AB) = N(AC) = N(BC) = (k-1)^2.$$

For the last term , one can at least assert the trivial inequality,

$$(k-1) \leq N(ABC) \leq (k-1)^2,$$

and thus

$$1 - \frac{k-1}{n-k+1} \leq p_k(X_{k-1}, Y_{k-1}) \leq 1 - \frac{(k-1)(n-2k+3)}{(n-k+1)^2}.$$

Considering the upper bound, one obtains,

$$\begin{aligned} E(p_2(X_1,Y_1)...p_n(X_{n-1},Y_{n-1})) &\leq \prod_{1\leq k\leq (n+3)/2} \left(1 - \frac{(k-1)(n-2k+3)}{(n-k+1)^2}\right) \\ &= \exp\left\{ \sum_{1\leq k\leq (n+3)/2} \ln\left(1 - \frac{(k-1)(n-2k+3)}{(n-k+1)^2}\right)\right\} \\ &\leq \exp\left\{(1+o(1))n\int_0^{1/2} \ln\left(1 - \frac{x(1-2x)}{(1-x)^2}\right) dx\right\}. \end{aligned}$$

Simple inequalities allow the integral to be bounded by $(3\ln 2 - 2)$ and numerical integration gives a more precise value for the integral of -0.0885474.

4 Counting good permutations in a restricted class

Let $\alpha = \{(i, \alpha(i)), i = 1, ..., n\}$ be a permutation, and denote by $N_k(a)$ the number of entries $(i, \alpha(i))$ of α occuring in the rectangle $\{1, ..., k\} \times \{a, ..., a+m-1\}$

We define the deviation from the expected value

$$\Delta_k(a) = N_k(a) - \frac{km}{n},$$

and maximal deviation

$$\Delta_m^+ = \max\left\{\Delta_k(a) \ : \ a \in \mathbb{Z}_n, k = m, 2m, ..., \lfloor\frac{n-m}{m}\rfloor m\right\}.$$

Let $G_n(z)$ be the subsetof good permutations G_n for which $\Delta_m^+ \leq z$.

Theorem 4 *Let $n \to \infty$ and let m, z vary in such a way that $m = o(n)$, $z = o(m)$. Then there exists $\epsilon > 0$, such that for sufficiently large n,*

$$\frac{|G_n(z)|}{|S_n|} \leq e^{-(1-\epsilon)n}.$$

Proof Let $\alpha \in G_n(z)$. Assume that $\alpha(1), ..., \alpha(rm)$ are already chosen. How many possibilities exist for the choice of $\alpha(rm+1), .., \alpha(rm+m)$? This number equals

$$mn - 2m^2 r + L, \tag{2}$$

where L is the number of points (i, j) in the rectangle

$$\{rm+1, ..., rm+m\} \times \{1, ..., n\}$$

for which

$$j \in \{\alpha(1), ..., \alpha(rm\},$$

and

$$i + j \in \{\nu + \alpha(\nu) (\mathrm{mod}\ n),\ 1 \le \nu \le rm\}.$$

Thus L equals the number of intersections of rm 'horizontal' and rm 'inclined' segments.

There are $N_{rm}(a)$ intersections along an inclined segment incident with the point $(rm+m, a)$. Thus if

$$a_\nu = [\nu + \alpha(\nu) - (rm+m)] \bmod n,$$

then

$$L = N_{rm}(a_1) + \cdots + N_{rm}(a_{rm}),$$

and

$$L \le \frac{r^2 m^3}{n} + rm\Delta_m^+.$$

Let x_i denote the number of choices for $\alpha(rm+i)$, then

$$\max\{x_1 x_2 \cdots x_m \ :\ x_1 + \cdots + x_m = c,\ x_1 \ge 0, ..., x_m \ge 0\}$$

is achieved at the point $x_i = c/m,\ 1 \le i \le m$, where c is given by (2). Hence the number of possibilities for the choice of $\alpha(rm+1), ..., \alpha(rm+m)$ does not exceed

$$\begin{aligned}\left(n - 2mr + \frac{r^2 m^2}{n} + r\Delta_m^+\right)^m &= n^m \left((1 - \frac{rm}{n})^2 + \frac{r}{n}\Delta_m^+\right)^m \\ &\le n^m \left((1 - \frac{rm}{n})^2 + \frac{rz}{n}\right)^m.\end{aligned}$$

Hence for $sm \le n$,

$$|G_n(z)| \le n^{sm} \prod_{r=0}^{s-1} \left((1 - \frac{rm}{n})^2 + \frac{rz}{n}\right)^m (n - sm)!, \tag{3}$$

where $(n-sm)!$ bounds the number of choices for $\alpha(n-sm+1), ..., \alpha(n)$.

Choose $\delta > 0,\ 0 < \delta < 1$, and let s be such that

$$sm \le (1-\delta)n \le (s+1)m.$$

Then

$$m\sum_{r=0}^{s-1}\ln\left((1-\frac{rm}{n})^2+\frac{rz}{n}\right)=n(I_0+I_1),$$

where

$$\begin{aligned} I_0 &= \frac{2m}{n}\sum_{r=0}^{s-1}\ln(1-\frac{rm}{n}),\\ I_1 &= \frac{m}{n}\sum_{r=0}^{s-1}\ln\left(1+\frac{rz}{n}(1-\frac{rm}{n})^{-2}\right). \end{aligned}$$

Since $m=o(n)$ we have

$$I_0\sim 2\int_0^{1-\delta}\ln(1-t)dt,$$

where

$$\int_0^1\ln(1-t)dt=-1.$$

Hence δ can be chosen in such a way that

$$I_0<-2+\frac{\epsilon}{2}. \tag{4}$$

For sufficiently large n

$$\begin{aligned} I_1 &\le \frac{m}{n}\sum_{r=0}^{s-1}\frac{rz}{n\delta^2}\\ &\sim \frac{zm}{2n^2\delta^2}s^2\\ &\sim \frac{zm}{2n^2\delta^2}\left(\frac{n(1-\delta)}{m}\right)^2\\ &< \frac{z}{m\delta^2}. \end{aligned}$$

Since $z=o(m)$ we have

$$I_1<\frac{\epsilon}{2}. \tag{5}$$

Setting (4) and (5) into (3) and noting that $(n-sm)!\le n^{n-sm}$ we obtain

$$|G_n(z)|<n^ne^{-(2-\epsilon)n},$$

for n sufficiently large. Hence for n sufficiently large,

$$\frac{|G_n(z)|}{|S_n|}<e^{-(1-\epsilon)n},$$

and the theorem follows. □

We now consider the following question.

Let α be a random permutation. How can the probability of the event $\{\Delta_m^+ \leq z\}$ be estimated?

For the complement of the event we have

$$\mathbf{Pr}(\Delta_m^+ > z) \leq n^s \max_{0 \leq k \leq n} \left\{ \mathbf{Pr}(N_k(0) > \frac{km}{n} + z) \right\}.$$

Let $p(i) = \mathbf{Pr}(N_k(0) = i)$. From a combinatorial argument we have

$$p(i) = \binom{m}{i} \frac{(k)_i (n-k)_{m-i}}{(n)_m}.$$

Setting $c = n/(n-m),\ k = nt,\ \tau = 1-t$, we obtain

$$p(i) \leq \binom{m}{i} t^i (\tau c)^{m-i},$$

and

$$\sum_{i=0}^{m} e^{\lambda i} p(i) \leq (te^{\lambda} + \tau c)^m,$$

for arbitrary real λ. Hence if $\lambda > 0$,

$$\mathbf{Pr}(N_k(0) \geq u) \leq e^{-\lambda u} (te^{\lambda} + \tau c)^m.$$

Setting $u = mt + z, e^{\lambda} = c(mt\tau + z\tau)/(mt\tau - zt)$, and assuming that $z = o(m)$, we obtain

$$\ln\ \mathbf{Pr}(N_k(0) \geq tm + z) \leq m \left(\frac{\tau m}{n-m} - \frac{z^2}{2m^2 + \tau} \right) (1 + o(1)),$$

and hence

$$\ln\ \mathbf{Pr}(\Delta_m^+ \geq z) \leq \left(\frac{n}{m} \ln n + \frac{m^2}{n-m} - \frac{2z^2}{m} \right) (1 + o(1)).$$

For example, assume that $m \sim n\sigma,\ z \sim m\sqrt{\sigma}$ when $\sigma = \sigma_m$ is a slowly decreasing function, then

$$\mathbf{Pr}(\Delta_m^+ \geq z) \leq \exp\left\{ -\sigma^2 n (1 + o(1)) \right\}.$$

Hence if z is chosen appropriately, a random permutation which satisfies $\Delta_m^+ \leq z$ belongs to $G_n(z)$ with probability tending to 1.

5 References

[1] Barbour, A.D., Holst, L. and Janson, S. (1992) *Poisson Approximation.* (Oxford).

[2] Dénes, J. and Keedwell, A.D. (1974) *Latin Squares and Their Applications.* (English University Press).

[3] McDiarmid, C. (1989) *On the method of bounded differences.* Surveys in Combinatorics 1989 (Editor J. Siemons). LMS Lecture Note Series 141. (Cambridge University Press).

[4] Paige, L.J. (1951) *Complete mappings of finite groups.* Pacific J. Math 1, 111-116.