Deranged Socks

SALLY COCKBURN
Hamilton College
Clinton, NY 13323
scockbur@hamilton.edu

JOSHUA LESPERANCE
Oberlin College
Oberlin, OH 44074
joshua.p.lesperance@gmail.com

Gloves and derangements

In the spring semester of 2005, the first author presented the following problem (and solution) as an example for her graph theory and combinatorics class, as found in Alan Tucker's *Applied Combinatorics* [**5**]:

GLOVE PROBLEM. Given five pairs of gloves, how many ways are there for five people each to choose two gloves with no one getting a matching pair?

In this problem, we assume that we can distinguish between left and right gloves. If we require that each person choose one left glove and one right glove, then the answer is 5,280; if we allow for the possibility of a person choosing two gloves for the same hand, then it shoots up to 65,280. Obtaining these answers is a nice illustration of combinatorial techniques.

Mismatched gloves may remind us of derangements. A *derangement* of n objects is a permutation in which every object gets moved. To count the number of derangements of n distinct objects, we turn to the principle of inclusion-exclusion, used for counting the number of elements in some universal set $\mathcal{U}$, which satisfy none of n different properties. We let A_i denote the elements of $\mathcal{U}$ satisfying property i, and S_k the sum of the cardinalities of all k-fold intersections of the A_i; that is,

$$S_k = \sum_{i_1<\cdots<i_k} |A_{i_1} \cap \cdots \cap A_{i_k}|.$$

For example,

$$S_1 = |A_1| + |A_2| + \cdots + |A_n|$$

and

$$S_2 = |A_1 \cap A_2| + |A_1 \cap A_3| + \cdots + |A_{n-1} \cap A_n|$$

(including all $\binom{n}{2}$ pairs). Then the principle states that

$$\left|\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_n}\right| = |\mathcal{U}| - S_1 + S_2 - S_3 + \cdots + (-1)^n S_n. \tag{1}$$

The left side of (1) is the number of elements of $\mathcal{U}$ that are not contained in any of the sets A_i.

Math. Mag. **86** (2013) 97–109. doi:10.4169/math.mag.86.2.97. © Mathematical Association of America

To count derangements, we let $\mathcal{U}_n$ denote the set of all permutations of $\{1, \ldots, n\}$, and A_i denote the subset of these permutations in which object i is in its original position. Then $|\mathcal{U}_n| = n!$ and $|A_i| = (n-1)!$ for all $1 \le i \le n$. More generally, $|A_{i_1} \cap \cdots \cap A_{i_k}| = (n-k)!$ for all subsets $\{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$, and since there are $\binom{n}{k} = n!/[k!\,(n-k)!]$ such subsets, we get $S_k = \binom{n}{k}(n-k)! = n!/k!$. Hence the nth derangement number is

$$D_n = n!\left[\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \cdots + \frac{(-1)^n}{n!}\right].$$

The first few values are $D_1 = 0$, $D_2 = 1$, $D_3 = 2$, $D_4 = 9$, $D_5 = 44$. An immediate consequence is that the fraction of all permutations that are derangements quickly converges to $1/e$ as $n \to \infty$.

Derangements can be used to solve the version of the Glove Problem in which n people each choose one left glove and one right glove. We begin by labeling the gloves $1_L, 1_R, \ldots, n_L, n_R$. To create n mismatched pairs of gloves, we first line up all of the left-hand gloves in their natural order: $1_L, 2_L, \ldots, n_L$. We then pair them up with a derangement of the right-hand gloves, and there are D_n ways to do this. Finally, there are $n!$ ways for us to distribute all n pairs of gloves, since no two glove-pairs will be the same, giving us the solution

$$n!D_n = [n!]^2\left[\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \cdots + \frac{(-1)^n}{n!}\right].$$

When $n = 5$, this formula gives $(5!)D_5 = 5{,}280$, one answer to the original Glove Problem.

If we allow people to choose gloves regardless of handedness, then the answer is more subtle, but we can again invoke the principle of inclusion-exclusion. Let $\mathcal{U}_n$ be the set of all possible ways of distributing 2 gloves to each of n people, from a set of $2n$ distinct gloves. To compute $|\mathcal{U}_n|$, we line up the gloves in order $1_L, 1_R, \ldots, n_L, n_R$. We then assign to this glove lineup a permutation of the multiset consisting of two copies of each person's name, which gives

$$|\mathcal{U}_n| = (2n)!/2^n.$$

If we let A_i denote the subset of glove distributions in which somebody gets matching pair i, then to compute S_k, we first choose which k gloves are matched, distribute these matching pairs to k lucky people, then distribute the remaining $n-k$ pairs in any fashion. If we use $\binom{n}{k}$ to denote the number of k-combinations of n objects and $P(n,k) = n!/(n-k)!$ to denote the number of k-permutations of n objects, then we get as our answer

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}P(n,k)\frac{[2(n-k)]!}{2^{n-k}} = \sum_{k=0}^{n}(-1)^k\binom{n}{k}P(n,k)|\mathcal{U}_{n-k}|.$$

When $n = 5$, this gives a value of 65,280, which is the other answer to the original Glove Problem.

The sock problem

Suppose that we remove the crucial assumption that right- and left-hand gloves are distinguishable. As suggested by one of our students, this happens naturally if we phrase the problem in terms of socks, rather than gloves. The resulting problem is surprisingly more difficult.

DERANGED SOCK PROBLEM. Given n distinct pairs of socks, how many ways are there for n people each to choose two socks with no one getting a matching pair? For the remainder of this paper, the solution to this problem shall be denoted d_n.

Note that there is only one version of this problem: If left and right socks are indistinguishable, then it doesn't make sense to consider the version where each person ends up with one left sock and one right sock. If we attack this problem with the principle of inclusion-exclusion, then by analogy to the glove problem, we can start with what might seem to be an easier problem.

SOCK DISTRIBUTION PROBLEM. Given n distinct pairs of socks, how many ways are there for n people each to choose two socks? The solution to this problem shall be denoted u_n.

Inclusion-exclusion allows us to compute d_n as a function of u_j for $1 \le j \le n$, but we can also compute u_n as a function of d_j by dividing into mutually exclusive, exhaustive cases according to how many people get a matching pair. This leads to the pair of formulas:

$$d_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} P(n,k) u_{n-k}; \tag{2}$$

$$u_n = \sum_{k=0}^{n} \binom{n}{k} P(n,k) d_{n-k}. \tag{3}$$

Thus, if we can compute d_n for all n or u_n for all n, then we can compute the other sequence. It seems a good bet to try for u_n first. However, sock distribution is not as easy as glove distribution. To illustrate the difficulty, consider the case $n = 4$ and suppose that the people are named A, B, C, and D. We begin by trying to mimic the approach we used for gloves above: line up the socks in canonical order, minus the L and R labels, and then assign to that list an arrangement of two copies of each name. We quickly see that this fails, for

$$\frac{\begin{matrix}1 & 1 & 2 & 2 & 3 & 3 & 4 & 4\end{matrix}}{\begin{matrix}A & B & A & B & C & C & D & D\end{matrix}} \quad \text{and} \quad \frac{\begin{matrix}1 & 1 & 2 & 2 & 3 & 3 & 4 & 4\end{matrix}}{\begin{matrix}B & A & A & B & C & C & D & D\end{matrix}}$$

result in the same people having the same socks, as will the name permutations $ABBACCDD$ and $BABACCDD$. Similarly, there are 8 name permutations that result in the same sock distribution as $ABACBCDD$, and 16 for $ABCDABCD$. Surely there is a simpler way.

As a stab at computing d_n directly, we break it down by first forming sock pairs and then distributing them, as we did in the first version of the glove problem. (Here and in the rest of this paper, a *sock pair* refers to any two socks to be distributed to a single person, whether matching or not.) Line up one sock of each type in order, then pair them up with a derangement of the remaining socks. For example,

$$\frac{\begin{matrix}1 & 2 & 3 & 4\end{matrix}}{\begin{matrix}4 & 1 & 2 & 3\end{matrix}} \quad \text{or} \quad \frac{\begin{matrix}1 & 2 & 3 & 4\end{matrix}}{\begin{matrix}2 & 3 & 4 & 1\end{matrix}} \quad \text{or} \quad \frac{\begin{matrix}1 & 2 & 3 & 4\end{matrix}}{\begin{matrix}4 & 3 & 2 & 1\end{matrix}}.$$

Even in this small case, several problems become apparent. The first two derangements, which we hoped would be distinct, both result in the same four sock pairs:

{1, 4}, {1, 2}, {2, 3}, and {3, 4}. Furthermore, the third derangement results in repetition of sock pairs: {1, 4}, {1, 4}, {2, 3}, and {2, 3}. Thus, while there are 4! ways to distribute the sock pairs of the first two derangements, there are only $4!/2^2$ ways to distribute the sock pairs from the third derangement. As the number of socks and people increase, these problems only get worse.

These cursory attempts demonstrate that switching from gloves to socks significantly complicates matters. In fact, this problem has been addressed before, as one case of a larger problem. In the most general form of the problem, we define $H(n, r)$ to be the number of ways of distributing r copies of n distinct objects among n people, each person getting exactly r objects in total. The first correct solution to the $r = 2$ version of the problem was given by Kenji Mano in 1961 (see [**2**]), in the form of a pair of rather complicated recursive formulas. In this paper, we show how the problem can be attacked using a variety of basic weapons in the discrete mathematics arsenal. We make use of partitions, cyclic permutations, recurrence relations, and generating functions. We obtain solutions in the form of both recursive and (two different) nonrecursive formulas, closed forms of an exotic variation on the generating functions, and finally, with the help of a couple of big guns from complex analysis, asymptotic formulas.

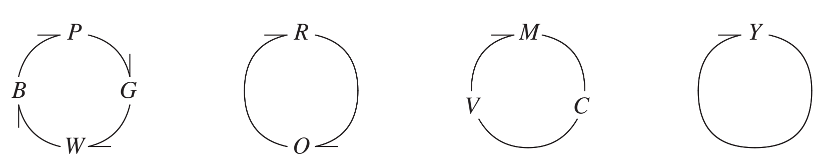
Partitions and cyclic permutations

In this section we return to our earlier approach for computing both d_n and u_n. We will first form n sock pairs and then distribute them to n people. We will find that both parts offer challenges that we did not encounter when sorting gloves. We begin with an example.

Consider the following pairing of socks, using the colors Red, Blue, Green, Yellow, Orange, Pink, White, Violet, Magenta, and Carbon:

$$\{\{R,O\},\{M,C\},\{P,G\},\{P,B\},\{R,O\},\{W,B\},\{C,V\},\{M,V\},\{Y,Y\},\{G,W\}\}.$$

We see that every color appears exactly twice. Additionally, we have a matched pair, which is allowable for u_n but not for d_n. Chains, or *cycles*, of colors can naturally be formed from this set as follows: We pick any color to begin with, say Pink, and find a color that is paired with Pink. There is a pair $\{P,G\}$, so we might choose Green. Then we find another color that is paired with Green, and so on, using each pair only once. In this case, Pink is paired with Green, Green is paired with White, White is paired with Blue, and Blue is paired with Pink, ending the first cycle. After a cycle ends, we pick a color that we haven't used yet and repeat, until we have used all of the colors. Red is paired with orange, which is paired with Red again. Magenta is paired with Carbon, Carbon is paired with Violet, and Violet is paired with Magenta. Finally, Yellow is paired only with Yellow.



These are cycles, not just sets.

Since every color appears exactly twice in this set of sock pairs and the number of colors is finite, this process will always produce cycles (beginning and ending with the same color). Furthermore, a cycle accounts for both copies of each color, each cycle can be thought of as a permutation of the colors involved, and no two cycles have any colors in common. Therefore, the entire set of sock pairs naturally corresponds to a permutation of the color set, written as a product of disjoint cycles: $(PGWB)(RO)(MCV)(Y)$. In fact, this set of sock pairs corresponds to more than one permutation, since each cycle can be reversed. (Order in the set of sock pairs, as well as in each individual sock pair, does not matter.)

Conversely, any permutation of the color set, which we know can be written as a product of disjoint cycles, naturally corresponds to a set of sock pairs, by placing adjacent colors into a pair. For example, the permutation

$$(RYV)(GMO)(CP)(B)(W)$$

gives us the set of sock pairs

$$\{\{R,Y\},\{Y,V\},\{V,R\},\{G,M\},\{M,O\},\{O,G\},\{C,P\},\{P,C\},\{B,B\},\{W,W\}\}$$

This correspondence between the set of sock pairs and the set of permutations of the color set holds for all such examples, and while it is not one-to-one, we can still use it to reach our goal. But first, we simplify our notation, so that we can more easily address the general problem.

We label the pairs of socks from 1 to n (that is, we have two socks labeled 1, two socks labeled 2, etc), referring to these as the *colors* of the socks. Consider any possible pairing of the $2n$ socks, $M = \{\{i_1, i_2\}, \{i_3, i_4\}, \ldots, \{i_{2n-1}, i_{2n}\}\}$. Order in each pair does not matter. For the moment we will allow matched pairs. Thus, for any given color i we have three possibilities: i is paired with itself; i is paired with j twice, for some $j \neq i$; or i is paired once with j and once with k, where i, j, and k are distinct. We notice that every color appears exactly twice in M.

We will say that i and j are *directly linked by M* if $\{i, j\} \in M$, and we will say that i and j are *linked by M* if we can get from i to j in a finite number of direct links. The relation of being linked is clearly symmetric and transitive, and since there are only finitely many colors, it must also be reflexive. Thus we have an equivalence relation on the set of colors $\{1, 2, \ldots, n\}$. Now let $P(M)$ be the corresponding set of equivalence classes; that is, $P(M)$ is the collection of subsets of $\{1, 2, \ldots, n\}$ where i and j are in the same subset if and only if they are linked by M. Thus, $P(M)$ is a partition of the set $\{1, 2, \ldots, n\}$.

As we saw in the example above, M induces a cyclic structure on the subsets in $P(M)$, by placing the colors that are directly linked by M adjacent to one another. That is, every subset of $P(M)$ corresponds to an element of S_n, the permutation group on n elements, written as a single cycle. For example, if $\{i, i\} \in M$, then $\{i\} \in P(M)$, corresponding to the 1-cycle (i). Likewise, if $\{i, j\}$ appears in M twice, then $\{i, j\} \in P(M)$, corresponding to the 2-cycle (ij). Finally, if $\{i_1, i_2\}, \{i_2, i_3\}, \ldots, \{i_k, i_1\} \in M$, $k \geq 3$, then $\{i_1, i_2, \ldots, i_k\} \in P(M)$, corresponding to *two distinct* k-cycles: $(i_1 i_2 \cdots i_k)$ and its inverse $(i_1 i_k i_{k-1} \cdots i_2)$. Thus, every set of sock pairs M corresponds to an element of S_n, written as a product of disjoint cycles. Likewise, since every permutation of S_n can be written as a product of disjoint cycles, every permutation in S_n corresponds to a set of sock pairs.

Let $C(a_1, a_2, \ldots, a_n)$ denote the number of permutations in S_n, written as a product of disjoint cycles, with a_i i-cycles for all $1 \leq i \leq n$. Notice that all of the a_i are nonnegative, and it has to be the case that $1a_1 + 2a_2 + \cdots + na_n = n$. To calculate

$C(a_1, a_2, \ldots, a_n)$ we first partition the set $\{1, 2, \ldots, n\}$ into a_1 subsets of size 1, a_2 subsets of size 2, and so on. The number of ways in which we can do this is

$$\binom{n}{1} \cdots \binom{n-a_1+1}{1} \frac{1}{a_1!} \binom{n-a_1}{2} \cdots \binom{n-a_1-2a_2+2}{2} \frac{1}{a_2!} \cdots$$
$$= \frac{n!}{a_1! \cdots a_n! \, (1!)^{a_1} \cdots (n!)^{a_n}}.$$

The factors $1/(a_i!)$ arise because, when we write a permutation as a product of disjoint cycles, the order in which the cycles appear does not matter. To complete the calculation of $C(a_1, a_2, \ldots, a_n)$ we need to make a cycle out of every subset in the partition. Recalling that there are $(j-1)!$ ways to make a cycle of length j from j distinct elements, we get

$$C(a_1, a_2, \ldots, a_n) = \frac{n!}{a_1! \cdots a_n! \, (1!)^{a_1} \cdots (n!)^{a_n}} (0!)^{a_1} \cdots ((n-1)!)^{a_n}$$
$$= \frac{n!}{a_1! \cdots a_n! \, 1^{a_1} \cdots n^{a_n}}.$$

Now let $C^*(a_1, a_2, \ldots, a_n)$ denote the number of ways to form a set of n sock pairs, corresponding to a permutation with a_i i-cycles. We saw above that for cycles of length 3 or greater, there are exactly two cycles that correspond to the same set of sock pairs. Therefore,

$$C^*(a_1, a_2, \ldots, a_n) = C(a_1, a_2, \ldots, a_n) \cdot \frac{1}{2^{a_3+\cdots+a_n}}.$$

After we have formed our sock pairs, distributing them is relatively easy. For the most part, one sock pair will be distinct from another. However, every time we have a cycle of length 2, we get two identical sock pairs. Thus, the number of ways we can distribute a set M of n sock pairs corresponding to a permutation with a_i i-cycles, $1 \le i \le n$, is $n!/2^{a_2}$.

To calculate u_n, we multiply $n!/2^{a_2}$ by $C^*(a_1, a_2, \ldots, a_n)$ and then sum over all possible nonnegative values $a_1, a_2, \ldots, a_n$ such that $1a_1 + 2a_2 + \cdots na_n = n$. To calculate d_n we do the same, but additionally we require that $a_1 = 0$, so that no sock is paired with one of the same color. Thus,

$$u_n = \sum \frac{(n!)^2}{a_1! \cdots a_n! \, 1^{a_1} \cdots n^{a_n} 2^{a_2+\cdots+a_n}} \tag{4}$$

and

$$d_n = \sum \frac{(n!)^2}{a_2! \cdots a_n! \, 2^{a_2} \cdots n^{a_n} 2^{a_2+\cdots+a_n}}. \tag{5}$$

In these sums, there is one term for each sequence $(a_1, \ldots, a_n)$ satisfying $a_1 + 2a_2 + 3a_3 + \cdots + na_n = n$.

The difficulty with this approach is in finding all possible nonnegative values $a_1, a_2, \ldots, a_n$ such that $1a_1 + 2a_2 + \cdots + na_n = n$. Below we work out the case when $n = 6$. For u_6 we sum all of the entries in the fourth column of the table, and for d_6 we exclude the rows where $a_1 \neq 0$.

This yields $u_6 = 202{,}410$ and $d_6 = 67{,}950$.

To conclude this section, we outline a recursive algorithm for finding all the n-tuples needed for the calculations of d_n and u_n. Actually, we show how to solve a

TABLE 1: Calculating u_6 and d_6

$(a_1, a_2, a_3, a_4, a_5, a_6)$	$C^*(a_1, a_2, a_3, a_4, a_5, a_6)$	$\frac{n!}{2^{a_2}}$	$C^*(a_1, a_2, a_3, a_4, a_5, a_6)\frac{n!}{2^{a_2}}$
(0, 0, 0, 0, 0, 1)	60	720	43200
(1, 0, 0, 0, 1, 0)	72	720	51840
(0, 1, 0, 1, 0, 0)	45	360	16200
(2, 0, 0, 1, 0, 0)	45	720	32400
(0, 0, 2, 0, 0, 0)	10	720	7200
(1, 1, 1, 0, 0, 0)	60	360	21600
(3, 0, 1, 0, 0, 0)	20	720	14400
(0, 3, 0, 0, 0, 0)	15	90	1350
(2, 2, 0, 0, 0, 0)	45	180	8100
(4, 1, 0, 0, 0, 0)	15	360	5400
(6, 0, 0, 0, 0, 0)	1	720	720

slightly more general problem. Let M and n be nonnegative integers. We wish to find all solutions $(a_1, a_2, \dots, a_n)$ of nonnegative integers to the equation

$$1a_1 + 2a_2 + \cdots + na_n = M. \tag{6}$$

First, note that $a_i = 0$ whenever $i > M$. To organize the process, we break down the solutions in terms of the position of last nonzero entry.

Let $1 \le i \le n$ such that $a_i \neq 0$, and $a_j = 0$ for all $j > i$. (We will start with $i = n$ and work our way down to $i = 1$.) It follows that $1 \le a_i \le \lfloor \frac{M}{i} \rfloor$. More importantly, a_i can have any value in this range with a single exception: if $i = 1$, then $a_1 = M$ is the only solution. For $i > 1$, we fix a value for a_i, and let $m = M - ia_i$. (Note that $0 \le m < M$.) If we have $m = 0$, then all the remaining variables in the solution $(a_1, a_2, \dots, a_n)$ must be zero. If $m > 0$, then we complete the solution by solving the equation

$$1a_1 + 2a_2 + \cdots + (i-1)a_{i-1} = m. \tag{7}$$

This, of course, is the recursive part. Since the number of variables in equation (7) is strictly smaller than the number of variables in equation (6), we know the process will end in a finite number of steps. After we have found all the solutions for each allowable value of a_i, we decrease i by one and start all over again.

Recurrence relations

Although we now have a solution, the difficulty of the calculations involved motivates us to try another approach. A common combinatorial strategy is to model problems with recurrence relations, and this tactic was used successfully by the Indian mathematicians Anand, Dumir, and Gupta in 1966 (see [**1**]). In this section, we present a streamlined version of their solution. Beginning with u_n, we divide into cases according to what happens to the last pair of socks.

Case 1. One of n people gets $\{n, n\}$; this can happen in nu_{n-1} ways.

Case 2. If pair n is split, there are two subcases to consider.

Subcase (i). We have two sock pairs of the form $\{i, n\}$ for some $1 \le i < n$; we choose 2 people to distribute them to, then distribute the remaining $n - 2$ pairs of socks. There are $\binom{n-1}{1}\binom{n}{2}u_{n-2}$ such distributions.

Subcase (ii). Pair n gets split into $\{i, n\}$ and $\{j, n\}$ for some $1 \leq i < j < n$ and these get distributed to 2 people. Temporarily consider the remaining sock i and sock j as a matching pair. If we distribute these $n-2$ 'pairs' in u_{n-2} ways, we are under-counting by a factor of 2 those distributions in which i and j go to different people. By the reasoning in Case 1, i and j go to the same person in exactly $(n-2)u_{n-3}$ ways. Hence, there are $\binom{n-1}{2}P(n, 2)[2u_{n-2} - (n-2)u_{n-3}]$ such distributions.

After some algebraic simplification, we get the recurrence relation

$$u_n = nu_{n-1} + \frac{n(n-1)^2}{2}\left[(2n-3)u_{n-2} - (n-2)^2 u_{n-3}\right], \quad \text{for } n \geq 3.$$

We require three initial conditions; on the basis that "there is always one way to do nothing" (as one of our students puts it), we have $u_0 = 1$, $u_1 = 1$, $u_2 = 3$.

A similar analysis gives a recurrence relation for d_n. We can ignore Case 1, and in subcase (i) of Case 2, we simply replace u_{n-2} with d_{n-2}. The latter part of subcase (ii) requires some modification because d_{n-2} will not count distributions in which somebody gets the 'pair' $\{i, j\}$. We can count these separately by giving $\{i, j\}$ to one of $n-2$ people and then distributing the remaining pairs in d_{n-3} ways. The distributions in which $\{i, j\}$ is split number exactly $2d_{n-2}$. A bit of algebra yields

$$d_n = \frac{n(n-1)^2}{2}\left[(2n-3)d_{n-2} + (n-2)^2 d_{n-3}\right], \quad \text{for } n \geq 3.$$

The corresponding initial conditions are $d_0 = 1$, $d_1 = 0$, $d_2 = 1$.

Armed with these relations and *Maple*, we can easily generate values for u_n and d_n; we display the first ten in the table below.

TABLE 2: Some values of u_n and d_n

n	u_n	d_n
3	21	6
4	282	90
5	6210	2040
6	202,410	67,950
7	9,135,630	3,110,940
8	545,007,960	187,530,840
9	41,514,583,320	14,398,171,200
10	3,930,730,108,200	1,371,785,398,200

There are several things to notice about these numbers. First, the values for $n = 6$ agree with those in the previous section, happily. Second, like many combinatorial quantities, they grow amazingly quickly. Third, u_n is roughly three times d_n. More precisely, as n increases, further computations show that d_n/u_n approaches $1/e$, the same ratio that $D_n/n!$ approaches, although the former at a slower rate than the latter. This is the first of several intriguing connections to the derangement numbers.

The Online Encyclopedia of Integer Sequences contains both of these sequences, described somewhat differently: (u_n) is sequence $A000681$, which counts the number of $n \times n$ nonnegative integer matrices such that every row sums to 2 and every column sums to 2, and (d_n) is sequence $A001499$, which counts the number of $n \times n$, $0-1$ matrices with exactly two 1's in each row and each column. The connection becomes obvious if we formulate the sock problem graph theoretically. Define a bipartite graph $G = (X, Y, E)$ with $|X| = |Y| = n$; each vertex in X represents a matching sock pair, and each vertex in Y represents a person. If each vertex has degree two, then everybody gets exactly two mismatched socks, and the incidence matrix is precisely an $n \times n$, $0-1$ matrix with exactly two 1's in each row and each column. The number of such incidence matrices is the solution to the Deranged Sock Problem. For the more general Sock Distribution Problem, we keep the requirement that G be 2-regular, but allow multiple edges.

Generating functions

Once we have recurrence relations for u_n and d_n, a natural impulse is to attempt to solve them to obtain non-recursive formulas, with the hope that they are simpler than the ones we found in our first approach. That the recurrence relations are nonlinear suggests that we try a generating function approach: That both u_n and d_n involve arrangements suggests that we use exponential generating functions. To this end, we define the formal power series

$$F(x) = \sum_{n=0}^{\infty} u_n \frac{x^n}{n!} \qquad \text{and} \qquad G(x) = \sum_{n=0}^{\infty} d_n \frac{x^n}{n!},$$

or equivalently,

$$F(x) = \sum_{n=0}^{\infty} \frac{u_n}{n!} x^n \qquad \text{and} \qquad G(x) = \sum_{n=0}^{\infty} \frac{d_n}{n!} x^n.$$

The point of this slight rewriting is to take advantage of a slight rewriting of equation (3) from the introduction. Since $P(n,k) = n!/(n-k)!$, we get

$$\frac{u_n}{n!} = \sum_{k=0}^{n} \binom{n}{k} \frac{d_{n-k}}{(n-k)!}.$$

If we let $\upsilon_n = u_n/n!$ and $\delta_n = d_n/n!$, then the equation above can be expressed as

$$\upsilon_n = \sum_{k=0}^{n} \binom{n}{k} \delta_{n-k}. \tag{8}$$

At this point, a notational trick from combinatorial analysis pays great dividends. When working with generating functions of the sequence $a_0, a_1, a_2, \ldots$, we can stipulate that $a^n \equiv a_n$. Then the ordinary and exponential generating functions of the sequence become, respectively,

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a^n x^n = \frac{1}{1-ax} \qquad \text{and} \qquad \sum_{n=0}^{\infty} a_n \frac{x^n}{n!} = \sum_{n=0}^{\infty} \frac{a^n x^n}{n!} = e^{ax}.$$

Remarkably, all formal operations with power series carry through with the constant indices treated as powers. If we set $\upsilon^n \equiv \upsilon_n$ and $\delta^n \equiv \delta_n$, then equation (8) becomes

$$\upsilon^n = \sum_{nk=0}^{n} \binom{n}{k} \delta^n = (1+\delta)^n. \tag{9}$$

This yields $\sum \upsilon^n x^n = \sum (1+\delta)^n x^n$, which can be rearranged to produce

$$\frac{1}{1-\upsilon x} = \frac{1}{1-(1+\delta)x} = \frac{1}{(1-\delta x)-x} = \frac{1}{1-\delta x}\left[1 + \frac{x}{1-\upsilon x}\right].$$

By definition, the ordinary generating function of the υ_n is $F(x)$, the exponential generating function of the u_n; similarly, the ordinary generating function of the δ_n is $G(x)$. Thus we have shown

$$F(x) = G(x)\,[1 + xF(x)]\,.$$

As pretty as this relationship is, it is not clear how it can help us find non-recursive formulas for u_n and d_n. However, another way we can use (9) is to obtain

$$\sum_{n=0}^{\infty} \frac{\upsilon^n x^n}{n!} = \sum_{n=0}^{\infty} \frac{(1+\delta)^n x^n}{n!},$$

which simplifies to

$$e^{\upsilon x} = e^{(1+\delta)x} = e^x e^{\delta x}. \tag{10}$$

This gives a relationship between what we term the *double-exponential* generating functions,

$$f(x) = e^{\upsilon x} = \sum_{n=0}^{\infty} \frac{u_n}{n!}\frac{x^n}{n!} \qquad \text{and} \qquad g(x) = e^{\delta x} = \sum_{n=0}^{\infty} \frac{d_n}{n!}\frac{x^n}{n!}.$$

We can now observe another tantalizing connection to the derangement numbers. The exponential generating function of the D_n is known to be $D(x) = e^{-x}/(1-x)$. Since $1/(1-x)$ can be interpreted as the exponential generating function $U(x)$ of the permutation numbers, $n!$, we have $U(x) = e^x D(x)$, the same relationship as (10). This suggests that we investigate these double-exponential generating functions further.

Double-exponential generating functions

We now exploit our earlier analysis in terms of partitions and cyclic permutations. Recall that the number of permutations of n objects consisting of a_i i-cycles is given by

$$C(a_1, a_2, \ldots, a_n) = \frac{n!}{a_1! \cdots a_n!\, 1^{a_1} \cdots n^{a_n}}.$$

In [**4**], Riordan shows that a multivariable ordinary generating function for these numbers is

$$C_n(t_1, t_2, \ldots, t_n) = \sum \frac{n!}{a_1! \cdots a_n!} \left(\frac{t_1}{1}\right)^{a_1} \cdots \left(\frac{t_n}{n}\right)^{a_n},$$

where the sum is over all nonnegative integers $a_1, \ldots, a_n$ satisfying $a_1 + 2a_2 + \cdots + na_n = n$. Riordan further demonstrates that

$$\sum_{n=0}^{\infty} C_n(t_1, \ldots, t_n)\frac{x^n}{n!} = \exp\left(t_1\frac{x}{1} + t_2\frac{x^2}{2} + t_3\frac{x^3}{3} + t_4\frac{x^4}{4} + \cdots\right).$$

Expressing the results (4) from our earlier approach in this new notation, we obtain

$$\frac{u_n}{n!} = C_n\left(1, \frac{1}{2}, \ldots, \frac{1}{2}\right) \qquad \text{and} \qquad \frac{d_n}{n!} = C_n\left(0, \frac{1}{2}, \ldots, \frac{1}{2}\right),$$

which in turn leads to

$$\begin{aligned}
f(x) &= \exp\left(x + \frac{1}{2}\left[\frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \cdots\right]\right) \\
&= \exp\left(\frac{1}{2}\left[\log\frac{1}{1-x} + x\right]\right) = \frac{e^{x/2}}{\sqrt{1-x}}; \\
g(x) &= \exp\left(\frac{1}{2}\left[\frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \cdots\right]\right) \\
&= \exp\left(\frac{1}{2}\left[\log\frac{1}{1-x} - x\right]\right) = \frac{e^{-x/2}}{\sqrt{1-x}}.
\end{aligned}$$

(An alternate derivation of these formulas appears in section 8 of [**1**].)

We can now use the closed expressions of the double-exponential generating functions to derive another set of non-recursive formulas for u_n and d_n. Recall the binomial series expansion

$$\frac{1}{\sqrt{1-x}} = (1-x)^{-1/2} = \sum_{n=0}^{\infty}\binom{-1/2}{n}(-x)^n,$$

where $\binom{-1/2}{0} = 1$ and for $n \geq 1$,

$$\begin{aligned}
\binom{-1/2}{n} &= \frac{(-1/2)(-3/2)(-5/2)\cdots(-1/2-n+1)}{n!} \\
&= \frac{(-1)^n 1\cdot 3\cdot 5\cdots(2n-1)}{2^n n!}.
\end{aligned}$$

Observe that this series can be viewed as the exponential generating function of the sequence $a_n = [1\cdot 3\cdot 5\cdots(2n-1)]/2^n$. Using the formal procedure for multiplying exponential series, and remembering that $f(x)$ and $g(x)$ are *double*-exponential generating functions, we obtain

$$\begin{aligned}
u_n &= \frac{n!}{2^n}\left(1 + \sum_{k=1}^{n}\binom{n}{k}1\cdot 3\cdot 5\cdots(2k-1)\right), \\
d_n &= \frac{n!}{2^n}\left((-1)^n + \sum_{k=1}^{n}\binom{n}{k}1\cdot 3\cdot 5\cdots(2k-1)(-1)^{n-k}\right).
\end{aligned}$$

Asymptotic behavior

We now have both recursive and non-recursive formulas for u_n and d_n, but neither shed much light on the observation that, just like $D_n/n!$, the ratio d_n/u_n approaches $1/e$ as n increases. To gain insight into this behavior, we must delve more deeply into the theory of generating functions.

By merely replacing the variable x with z, we can mentally transform a formal power series into a function of complex variables, which must be analytic on some disk in the complex plane centered at 0. In particular, $f(z) = e^{z/2}/\sqrt{1-z}$ and $g(z) = e^{-z/2}/\sqrt{1-z}$ are each analytic in the unit disk, with a single algebraic singularity at $z_0 = 1$. Fortunately, some heavy machinery from complex analysis exists for exactly this sort of function. The following can be found in [**6**].

THEOREM 1. *(Darboux's Lemma) Let $v(z)$ be analytic in some disk $|z| < 1 + \eta$ where $\eta > 1$. Let $\beta \in \mathbb{R}\backslash\{0, 1, 2, \ldots\}$. Suppose that in a neighborhood of $z = 1$, $v(z) = \sum v_n(1 - z)^n$. Then for every integer $m \geq 0$, the coefficient of z^n in the expansion of $v(z)(1 - z)^\beta$ is*

$$\left[\sum_{k=0}^{m} \binom{\beta + k}{n} (-1)^n v_k\right] + O(n^{-m-\beta-2}).$$

To apply this to $f(z)$, we use $v(z) = e^{z/2}$. This is an entire function; its Taylor expansion centered at $z = 1$ is

$$e^{z/2} = \sum_{n=0}^{\infty} \frac{v^{(n)}(1)}{n!}(z - 1)^n = \sum_{n=0}^{\infty} \frac{(-1)^n e^{1/2}}{2^n n!}(1 - z)^n.$$

Then using Darboux's Lemma with $m = 2$, we obtain

$$\frac{u_n}{(n!)^2} = e^{1/2}\left[\binom{-1/2}{n} + \frac{1}{2}\binom{-3/2}{n} + \frac{1}{8}\binom{-5/2}{n}\right] + O(n^{-7/2}). \tag{11}$$

For large values of n, we can simplify the fractional binomial coefficients in this expression with this result, also in [**6**].

THEOREM 2. *Let $\alpha \in \mathbb{R}\backslash\{0, 1, 2, \ldots\}$. Then as $n \to \infty$,*

$$\binom{\alpha}{n} \sim \frac{(-1)^n n^{-\alpha-1}}{\Gamma(-\alpha)}.$$

The appearance of the gamma function need not overly alarm us, as we can compute all the values we need from the simple pair of properties $\Gamma(1/2) = \sqrt{\pi}$ and $\Gamma(z)\Gamma(1 - z) = \pi/\sin \pi z$ (see [**3**]). Substituting into (11), we get for large values of n,

$$u_n \sim \frac{(n!)^2 e^{1/2}}{\sqrt{n\pi}}\left[1 + \frac{1}{4n} + \frac{3}{32n^2}\right].$$

Replacing $v(z) = e^{z/2}$ with $w(z) = e^{-z/2}$, we get an asymptotic formula for d_n:

$$d_n \sim \frac{(n!)^2 e^{-1/2}}{\sqrt{n\pi}}\left[1 - \frac{1}{4n} + \frac{3}{32n^2}\right].$$

The table below illustrates how accurate these approximations are, even for small values of n.

TABLE 3: Asymptotic values of u_n and d_n.

n	u_n	$\frac{(n!)^2 e^{1/2}}{\sqrt{n\pi}}\left[1+\frac{1}{4n}+\frac{3}{32n^2}\right]$	d_n	$\frac{(n!)^2 e^{-1/2}}{\sqrt{n\pi}}\left[1-\frac{1}{4n}+\frac{3}{32n^2}\right]$
5	6210	6312.3	2040	2101.8
10	3.931×10^{12}	3.974×10^{12}	1.371×10^{12}	1.391×10^{12}
20	1.239×10^{36}	1.247×10^{36}	4.444×10^{35}	4.474×10^{35}
30	1.200×10^{64}	1.205×10^{64}	4.341×10^{63}	4.360×10^{63}
40	9.823×10^{94}	9.853×10^{94}	3.568×10^{94}	3.580×10^{94}
50	1.220×10^{128}	1.223×10^{128}	4.443×10^{127}	4.454×10^{127}

The asymptotic formulas finally clear up the mystery:

$$\lim_{n\to\infty}\frac{d_n}{u_n}=\lim_{n\to\infty}\frac{\frac{(n!)^2e^{-1/2}}{\sqrt{n\pi}}\left[1-\frac{1}{4n}+\frac{3}{32n^2}\right]}{\frac{(n!)^2e^{1/2}}{\sqrt{n\pi}}\left[1+\frac{1}{4n}+\frac{3}{32n^2}\right]}=\frac{1}{e}.$$

Acknowledgment We would like to thank the students of Hamilton's 2005 combinatorics class for their useful discussions on this problem.

REFERENCES

1. Harsh Anand, Vishwa Chander Dumir, and Hansraj Gupta, A combinatorial distribution problem, *Duke Mathematical Journal* **33** (1966) 757–769. `http://dx.doi.org/10.1215/S0012-7094-66-03391-6`
2. Kenji Mano, On the formula of ${}_nH_r$, *Scientific Reports of the Faculty of Literature and Science: Hirosaki University* **8** (1961) 58–60.
3. J. E. Marsden and M. J. Hoffman, *Basic Complex Analysis*, W. H. Freeman, New York, 1999.
4. John Riordan, *Introduction to Combinatorial Analysis*, Dover, Mineola, NY, 2002.
5. Alan Tucker, *Applied Combinatorics*, John Wiley, New York, 2007.
6. Herbert S. Wilf, *Generatingfunctionology*, A. K. Peters, Natick, MA, 2006.

AuQ: Please confirm title format as ${}_nH_r$

Summary It is an elementary combinatorial problem to determine the number of ways n people can each choose two gloves from a pile of n distinct pairs of gloves, with nobody getting a matching pair. Change the gloves to socks (with right socks being indistinguishable from left socks), however, and the problem becomes surprisingly more difficult. We show how this problem can be solved using a wide range of discrete mathematics tools: the principle of inclusion-exclusion; partitions; cyclic permutations; recurrence relations; as well as both ordinary and exponential generating functions. We even draw on a result from complex analysis to show that the fraction of all sock distributions that are deranged in this sense converges to $1/e$.

SALLY COCKBURN is an associate professor in the mathematics department at Hamilton College, with interests in combinatorics, graph theory, linear optimization and set theory. For relief from mathematical dilettantism, she plays squash and tries to keep up with her menagerie of dogs.

JOSHUA LESPERANCE is currently a visiting assistant professor at Oberlin College, and his mathematical interests are varied and ever-evolving. He received a B.S. from Rochester Institute of Technology and both an M.S. and Ph.D. from the University of Notre Dame. In his free time, he enjoys playing the guitar and hiking with his wife and two Siberian Huskies.