DISJOINT TRANSVERSALS OF SUBSETS

P. J. HIGGINS

1. Introduction. Let $A_1, A_2, \ldots, A_n$ be a finite collection of subsets (not necessarily distinct) of a set A. By a *transversal*[1] of $A_1 A_2, \ldots, A_n$ we shall mean a set of n distinct elements $a_1, a_2, \ldots, a_n$ of A such that, for some permutation $i_1, i_2, \ldots, i_n$ of the integers $1, 2, \ldots, n$,

$$a_j \in A_{i_j} \qquad (j = 1, 2, \ldots, n).$$

More generally, we shall say that the set $\{a_1, a_2, \ldots, a_r\}$, $(r \leqslant n)$ is a *partial transversal* of $A_1, A_2, \ldots, A_n$ of *length* r if (i) $a_1, a_2, \ldots, a_r$ are distinct elements of A and (ii) there exists a set of distinct integers $i_1, i_2, \ldots, i_r$ such that

$$a_j \in A_{i_j} \qquad (j = 1, 2, \ldots, r).$$

A well-known theorem of P. Hall **(2)** states that the sets $A_1, A_2, \ldots, A_n$ have a transversal (of length n) if, and only if, every k of them contain collectively at least k distinct elements $(k = 1, 2, \ldots, n)$. A generalization of this theorem by Ore **(3)** states that the sets $A_1, A_2, \ldots, A_n$ have a partial transversal of length $r \leqslant n$ if, and only if, every k of them contain collectively at least $k + r - n$ distinct elements $(n - r + 1 \leqslant k \leqslant n)$.

In this paper we enquire under what conditions the sets $A_1, A_2, \ldots, A_n$ will have m mutually disjoint partial transversals of prescribed lengths $r_1, r_2, \ldots, r_m$. As in the two theorems quoted above, the obvious necessary conditions are again found to be sufficient. As a special case we deduce a theorem of Ryser **(4)** and Gale **(1)** concerning the existence of matrices of 0's and 1's with prescribed row sums and column sums.

2. Notation. Throughout our argument n will denote a fixed positive integer (the number of subsets A_j), and $r_1, r_2, \ldots, r_m$ will denote positive integers not exceeding n. We shall suppose that $r_1 \geqslant r_2 \geqslant \ldots \geqslant r_m > 0$ and think of these integers as a partition $[r_i]$ of $r_1 + r_1 + \ldots + r_m$. The conjugate partition $[r_j^*]$ is defined as usual:

$$r_j^* = \sum_{r_i \geqslant j} 1 \qquad (j = 1, 2, \ldots, r_1). \tag{1}$$

It is convenient also to define $r_j^* = 0$ if $r_1 < j \leqslant n$, which is in accord with (1) if we interpret empty sums as zero.

Received May 22, 1958.

[1]This term, due to P. Hall, is normally used when the sets $A_1, A_2, \ldots A_n$ are disjoint, but its use in this wider sense is convenient here.

We now write

$$\alpha_k = \sum_{j=n-k+1}^{n} r_j^* \qquad (k = 1, 2, \ldots, n). \tag{2}$$

An alternative expression for α_k can be obtained as follows. If s and t are integers, let $Ex(s, t)$ denote the excess, if any, of s over t, that is, $Ex(s, t) = s - t$ if $s \geqslant t$, and $Ex(s, t) = 0$ if $s \leqslant t$. Then

$$\alpha_k = \sum_{i=1}^{m} Ex(r_i, n - k) \qquad (k = 1, 2, \ldots, n). \tag{3}$$

The easiest way to see this is to draw a partition diagram for $[r_i]$, that is, an $m \times n$ matrix whose ith row has entries 1 in the first r_i places and 0 elsewhere. Then r_j^* is the number of 1's in the jth column, and α_k is the number of 1's in the last k columns. However, the ith row has exactly $Ex(r_i, n - k)$ 1's in the last k columns, and (3) follows.

3. Disjoint partial transversals. Suppose that $A_1, A_2, \ldots, A_n$ have disjoint partial transversals (D.P.T.'s) $R_1, R_2, \ldots, R_m$ of lengths $r_1, r_2, \ldots, r_m$ respectively. The elements of R_i represent r_i of the A's. Of these A's at least $Ex(r_i, n - k)$ must be included in any collection of k of the A's. It follows that every k of the A's must contain between them at least $Ex(r_i, n - k)$ distinct elements out of R_i and hence at least

$$\alpha_k = \sum_{i=1}^{m} Ex(r_i, n - k)$$

distinct elements altogether, since the R's are disjoint. Our theorem asserts that this necessary condition is also sufficient.

THEOREM. *A necessary and sufficient condition for $A_1, A_2, \ldots, A_n$ to have mutually disjoint partial transversals of lengths $r_1, r_2, \ldots, r_m$ is that, for $k = 1, 2, \ldots, n$, every k of the A's contain between them at least α_k distinct elements, where α_k is defined by* (1) *and* (2) *above.*

We observe here that the case $m = 1, r_1 = r$, is precisely Ore's theorem since we then have $\alpha_k = 0$ $(1 \leqslant k \leqslant n - r)$ and $\alpha_k = k + r - n$ $(n - r + 1 \leqslant k \leqslant n)$.

The proof of sufficiency proceeds by induction on n. It is trivial when $n = 1$, and from now on we shall assume the result for all collections of $n' < n$ sets and all sets of integers $r_i \leqslant n'$.

We distinguishtwo cases which are mutually exclusive and cover all possibilities:

Case 1. $m \geqslant 2$ and $r_m < n$, $r_{m-1} < n$;

Case 2. $r_1 = r_2 = \ldots = r_{m-1} = n$, $1 \leqslant r_m \leqslant n$.

We shall first reduce Case 1 to Case 2.

If $[r_i]$ is a partition falling under Case 1, we define a new partition $[\tilde{r}_i]$, the *reduction* of $[r_i]$, as follows. Let $r_1 = r_2 = \ldots = r_t = n$, $r_{t+1} < n$, where, by assumption, $0 \leqslant t \leqslant m-2$. Then $\tilde{r}_m = r_m - 1$, $\tilde{r}_{t+1} = r_{t+1} + 1$, and $\tilde{r}_i = r_i$ for all other values of i. Clearly $\tilde{r}_1 \geqslant \tilde{r}_2 \geqslant \ldots \geqslant \tilde{r}_m$, and by a finite number of such reductions any partition falling under Case 1 can be reduced to one falling under Case 2. Note that we may have $\tilde{r}_m = 0$, in which case the value of m is reduced by 1. It will, however, be convenient at times to retain such vanishing parts of a partition and interpret a partial transversal of length zero as the empty set. This will not affect the proof in any way.

To reduce Case 1 to Case 2 it is enough to prove

LEMMA. *If the theorem is true for the partition $[\tilde{r}_i]$ then it is also true for the partition $[r_i]$.*

Proof. Suppose that $[r_i]$ falls under Case 1, and every k of the A's contain between them at least α_k distinct elements $(k = 1, 2, \ldots, n)$.

Case 1 (a). First consider the possibility that for some k $(1 \leqslant k \leqslant n-1)$ there is a collection of k of the A's, say $A_1, A_2, \ldots, A_k$, which contain between them precisely α_k distinct elements. We construct two new partitions $[p_i]$ and $[q_i]$ where $p_i = Ex(r_i, n-k)$, $q_i = \min(r_i, n-k)$ $(i = 1, 2, \ldots, m)$. Then $p_i + q_i = r_i$ $(i = 1, 2, \ldots, m)$, $p_j^* = r_{j+n-k}^*$ $(j = 1, 2, \ldots, k)$, and $q_j^* = r_j^*$ $(j = 1, 2, \ldots, n-k)$. We apply our induction hypothesis to the sets $A, A_2, \ldots, A_k$ with the partition $[p_i]$ and to the sets $A_{k+1}, A_{k+2}, \ldots, A_n$ with the partition $[q_i]$. For this purpose let β_s and γ_s be the integers obtained from $[p_i]$ and $[q_i]$ in the same way that the α_s were obtained from $[r_i]$. Thus

$$\beta_s = \sum_{j=k-s+1}^{k} p_j^* = \sum_{j=n-s+1}^{n} r_j^* = \alpha_s \qquad (s = 1, 2, \ldots, k),$$

and

$$\gamma_s = \sum_{j=n-k-s+1}^{n-k} q_j^* = \sum_{j=n-k-s+1}^{n-k} r_j^* = \alpha_{k+s} - \alpha_k \qquad (s = 1, 2, \ldots, n-k).$$

By assumption, every s of the sets $A_1, A_2, \ldots, A_k$ contain between them at least $\alpha_s = \beta_s$ distinct elements. Also, any s of the sets $A_{k+1}, A_{k+2}, \ldots, A_n$ contain, together with all of $A_1, A_2, \ldots, A_k$, at least α_{k+s} distinct elements. Since $A_1 \cup A_2 \cup \ldots \cup A_k$ contains precisely α_k elements, any s of the sets $A_{k+1}, A_{k+2}, \ldots, A_n$ must contain between them at least $\alpha_{k+s} - \alpha_k = \gamma_s$ distinct elements not in $A_1 \cup A_2 \cup \ldots \cup A_k$. It follows that there exist D.P.T.'s $P_1, P_2, \ldots, P_m$ of $A_1, A_2, \ldots, A_k$ of lengths $p_1, p_2, \ldots, p_m$, and D.P.T.'s $Q_1, Q_2, \ldots, Q_m$ of $A_{k+1}, A_{k+2}, \ldots, A_n$ of lengths $q_1, q_2, \ldots, q_m$, none of the Q's having any elements in common with any of the P's. The sets $P_1 \cup Q_1, P_2 \cup Q_2, \ldots, P_m \cup Q_m$ are then D.P.T.'s of $A_1, A_2, \ldots, A_n$ of lengths $r_1, r_2, \ldots, r_m$.

Case 1 (b). If no such collection of A's exists then, for $k = 1, 2, \ldots, n-1$, every k of the A's must contain between them at least $\alpha_k + 1$ distinct elements, and we now appeal to the reduced partition. We observe that in passing from $[r_i]$ to $[\tilde{r}_i]$ one of the r_j^* is increased by 1, and one of them is decreased by 1, the others being unaltered. Hence, in the obvious notation,

$$\tilde{\alpha}_k = \sum_{j=n-k+1}^{n} \tilde{r}_j^* \leqslant 1 + \sum_{j=n-k+1}^{n} r_j^* = 1 + \alpha_k \qquad (k = 1, 2, \ldots, n),$$

while $\tilde{\alpha}_n = \alpha_n$. Thus every k of the A's contain between them at least $\tilde{\alpha}_k$ distinct elements ($k = 1, 2, \ldots, n$). Assuming the theorem for the partition $[\tilde{r}_i]$, we can find D.P.T.'s $\tilde{R}_1, \tilde{R}_2, \ldots, \tilde{R}_m$ of lengths $\tilde{r}_1, \tilde{r}_2, \ldots, \tilde{r}_m$. Now $\tilde{r}_{t+1} > r_{t+1} \geqslant r_m > \tilde{r}_m$ (t has the same meaning as before). Hence there must be in $\tilde{R}_{t+1}$ at least one element which represents a set A_j not represented by any element of $\tilde{R}_m$. If we transfer this element from $\tilde{R}_{t+1}$ to $\tilde{R}_m$ we obtain D.P.T.'s of lengths $r_1, r_2, \ldots, r_m$. This proves the lemma.

It remains to prove the theorem in Case 2, that is, under the assumptions $r_1 = r_2 = \ldots = r_{m-1} = n$, $1 \leqslant r_m \leqslant n$, $m \geqslant 1$. Then $r_j^* = m$ for $j = 1, 2, \ldots, r$, and $r_j^* = m - 1$ for $j = r + 1, r + 2, \ldots, n$, where for convenience we write $r_m = r$. We now make the further definition

$$\delta_k = \sum_{j=1}^{k} r_j^* \geqslant \alpha_k \qquad (k = 1, 2, \ldots, n).$$

Assume that $A_1, A_2, \ldots, A_n$ satisfy the conditions of the theorem.

Case 2 (a). First suppose that, for $k = 1, 2, \ldots, n - 1$, every k of the A's contain between them at least δ_k distinct elements. The same will be true for $k = n$ since $\delta_n = \alpha_n$. Consider a collection of sets $\{B_j\}$ consisting of m repetitions of each of the sets $A_1, A_2, \ldots, A_r$ and $m - 1$ repetitions of each of the sets $A_{r+1}, A_{r+2}, \ldots, A_n$ (if any). There are α_n sets altogether, and we shall show that, for $s = 1, 2, \ldots, \alpha_n$, any s of these sets contain between them at least s distinct elements. We must first count the number k of distinct* A's included amongst s of the B's. Clearly $k \geqslant s/m$; and if $s > mr$ then $k \geqslant r + (s - mr)/(m - 1)$. If $s \leqslant mr$, then $s/m \leqslant r$ and, if k' is the smallest integer such that $k' \geqslant s/m$, then $k' \leqslant r$. Hence $\delta_{k'} = k'm \geqslant s$, and any s of the B's must contain between them at least $\delta_k \geqslant \delta_{k'} \geqslant s$ distinct elements. On the other hand, if $s > mr$, then $k - r \geqslant (s - mr)/(m - 1)$, and $\delta_k = rm + (k - r)(m - 1) \geqslant rm + (s - mr) = s$. Thus again any s of the B's must contain between them at least s distinct elements. Applying Hall's theorem quoted in the introduction, we can find a complete transversal of the B's. The α_n distinct elements in this transversal comprise m distinct representatives of each of the sets $A_1, A_2, \ldots, A_r$ and $m - 1$ distinct representatives of each of the sets $A_{r+1}, A_{r+2}, \ldots, A_n$. It is easy to see that these elements can be arranged to form D.P.T.'s of $A_1, A_2, \ldots, A_n$, $m - 1$ of length n and one of length r.

*By "distinct" we mean here "having distinct suffixes." Thus distinct A's may have the same members.

Case 2 (b). The alternative to 2 (a) is that for some k $(1 \leqslant k \leqslant n-1)$ there is a collection of k of the A's whose union contains fewer than δ_k distinct elements (but at least α_k). From all such collections (for all possible values of k) we pick one collection consisting of, say, k A's whose union contains $\alpha_k + u$ distinct elements with *u as small as possible.* Thus every s of the A's $(s = 1, 2, \ldots, n)$ contain between them at least $\min(\delta_s, \alpha_s + u)$ distinct elements. (This statement for $s = n$ follows from the fact that $\delta_n = \alpha_n$.) Let the chosen sets be $A_1, A_2, \ldots, A_k$ (k is now fixed, $1 \leqslant k \leqslant n-1$). If $u = 0$ we may, of course, proceed as in Case 1 (a). This fails, however, if $u > 0$, and we must appeal again to the special form of the partition $[r_i]$.

Consider the sums of k successive r^*'s, that is, the integers $\epsilon_i = r_{i+1}^* + r_{i+2}^* + \ldots + r_{i+k}^*$ $(i = 0, 1, \ldots, n-k)$. Clearly $\delta_k = \epsilon_0 \geqslant \epsilon_1 \geqslant \ldots \geqslant \epsilon_{n-k} = \alpha_k$. Also $\epsilon_i - \epsilon_{i+1} \leqslant 1$ since $m = r_1^* \geqslant r_2^* \geqslant \ldots \geqslant r_n^* \geqslant m-1$. Now $\delta_k > \alpha_k + u > \alpha_k$; hence there is an integer t $(1 \leqslant t < n-k)$ such that $\epsilon_t = \alpha_k + u$. We may take $t \leqslant r$ since, if ϵ_r is defined, its value is $(m-1)k$ which must also be the value of α_k.

In the partition diagram of $[r_i]$ we now look at columns $t+1, t+2, \ldots, t+k$. They form the partition diagram of $[p_i]$ where $p_1 = p_2 = \ldots = p_{m-1} = k$, $p_m = r - t$, and

$$\sum_{i=1}^{m} p_i = \alpha_k + u.$$

The remaining columns form the partition diagram of $[q_i]$ where $q_1 = q_2 = \ldots = q_{m-1} = n - k$, $q_m = t$. The integers β_s and γ_s obtained from $[p_i]$ and $[q_i]$ in the same way that the α_s were obtained from $[r_i]$ are given by

$$\beta_s = \sum_{j=k-s+1}^{k} p_j^* = \sum_{j=t+k-s+1}^{t+k} r_j^* \qquad (s = 1, 2, \ldots, k),$$

$$\gamma_s = \sum_{j=n-k-s+1}^{n-k} q_j^* = \begin{cases} \alpha_s & \text{if } s \leqslant n-k-t \\ \alpha_{k+s} - (\alpha_k + u) & \text{if } n-k-t < s \leqslant n-k. \end{cases}$$

Consider a collection of $s \leqslant k$ of the sets $A_1, A_2, \ldots, A_k$. Between them they contain at least $\min(\delta_s, \alpha_s + u)$ distinct elements. Now

$$\delta_s = \sum_{j=1}^{s} r_j^* \geqslant \sum_{j=t+k-s+1}^{t+k} r_j^* = \beta_s.$$

Also

$$\begin{aligned} \alpha_s + u = (\alpha_k + u) - (\alpha_k - \alpha_s) &= \sum_{j=t+1}^{t+k} r_j^* - \sum_{j=n-k+1}^{n-s} r_j^* \\ &\geqslant \sum_{j=t+1}^{t+k} r_j^* - \sum_{j=t+1}^{t+k-s} r_j^* \qquad (\text{since } t \leqslant n-k) \\ &= \beta_s. \end{aligned}$$

Thus any s of $A_1, A_2, \ldots, A_k$ contain between them at least β_s distinct elements, and since $k < n$, we may apply our induction hypothesis to find D.P.T.'s $P_1, P_2, \ldots, P_m$ of $A_1, A_2, \ldots, A_k$ of lengths $p_1, p_2, \ldots, p_m$.

https://doi.org/10.4153/CJM-1959-030-9 Published online by Cambridge University Press

Now consider a collection of $s \leqslant n - k$ of the sets $A_{k+1}, A_{k+2}, \ldots, A_n$. Together with all of $A_1, A_2, \ldots, A_k$ they contain at least $\min(\delta_{k+s}, \alpha_{k+s} + u)$ distinct elements. Since $A_1 \cup A_2 \cup \ldots \cup A_k$ contains exactly $\alpha_k + u$ elements, the s sets from $A_{k+1}, A_{k+2}, \ldots, A_n$ must contain between them at least $\min(\delta_{k+s} - (\alpha_k + u), \alpha_{k+s} - \alpha_k)$ distinct elements not already used in $P_1, P_2, \ldots, P_m$. If we can show that, for $s = 1, 2, \ldots, n - k$,

$$\text{(i)} \quad \delta_{k+s} - (\alpha_k + u) \geqslant \gamma_s,$$

and

$$\text{(ii)} \quad \alpha_{k+s} - \alpha_k \geqslant \gamma_s,$$

we may apply our induction hypothesis to obtain D.P.T.'s $Q_1, Q_2, \ldots, Q_m$ of $A_{k+1}, A_{k+2}, \ldots, A_n$ of lengths $q_1, q_2, \ldots, q_m$ from elements not already used in $P_1, P_2, \ldots, P_m$. If $s > n - k - t$, these inequalities are obvious; for then $\gamma_s = \alpha_{k+s} - (\alpha_k + u)$, and clearly $\delta_{k+s} \geqslant \alpha_{k+s}$, $\alpha_k \leqslant \alpha_k + u$. On the other hand, if $s \leqslant n - k - t$, then $\gamma_s = \alpha_s$. In this case we observe that $\delta_{k+s} \geqslant \delta_k + \alpha_s > (\alpha_k + u) + \alpha_s$, from which (i) follows. Also $\alpha_{k+s} \geqslant \alpha_k + \alpha_s$, from which (ii) follows. This establishes the existence of $Q_1, Q_2, \ldots, Q_m$.

Finally, $P_1 \cup Q_1$, $P_2 \cup Q_2, \ldots, P_m \cup Q_m$ are D.P.T.'s of $A_1, A_2, \ldots, A_n$ of lengths $r_1, r_2, \ldots, r_m$, and the theorem is proved.

The application to matrices of 0's and 1's, mentioned in the introduction, is immediate. Let $n \geqslant r_1 \geqslant r_2 \geqslant \ldots \geqslant r_m \geqslant 0$ and $s_1 \geqslant s_2 \geqslant \ldots \geqslant s_n \geqslant 0$. The insertion of 1's in an $m \times n$ matrix so that there are at least r_i 1's in the ith row ($i = 1, 2, \ldots, m$) and not more than s_j in the jth column ($j = 1, 2, \ldots, n$) is equivalent to the construction of D.P.T.'s of lengths $r_1, r_2, \ldots, r_m$ of n *disjoint* sets containing respectively $s_1, s_2, \ldots, s_n$ elements. Our theorem gives as necessary and sufficient conditions for the existence of such D.P.T.'s the inequalities

$$\sum_{j=n-k+1}^{n} s_j \geqslant \alpha_k = \sum_{j=n-k+1}^{n} r_j^* \qquad (k = 1, 2, \ldots, n).$$

(The inclusion of zeros amongst the r's affects neither the hypotheses nor the conclusion.) If we require exactly r_i 1's in the ith row and exactly s_j in the jth column, we need only add the condition

$$\sum_{j=1}^{n} s_j = \sum_{i=1}^{m} r_i.$$

These are the conditions found by Ryser and Gale.

References

1. David Gale, *A theorem on flows in networks*, Pacific J. Math., 7 (1957), 1073–82.
2. Philip Hall, *On representatives of subsets*, J. Lond. Math. Soc., *10* (1935), 26–30.
3. Oystein Ore, *Graphs and matching theorems*, Duke Math. J., *22* (1955), 625–39.
4. H. J. Ryser, *Combinatorial properties of matrices of zeros and ones*, Can. J. Math., *9* (1957), 371–7.

Harvard University