A COMBINATORIAL THEOREM

P. ERDÖS AND R. RADO†.

[*Extracted from the Journal of the London Mathematical Society, Vol.* 25, 1950.]

1. F. P. Ramsey (**1**) proved the following theorem. *Let n be a positive integer, and let Δ be an arbitrary distribution of all sets of n positive integers into a finite number of classes. Then there exists an infinite set M of positive integers which has the property that all sets of n numbers of M belong to the same class of Δ.* Apart from its intrinsic interest the theorem possesses applications in widely different branches of mathematics. Thus in (**1**) the theorem is used to deal with a special case of the "Entscheidungsproblem" in formal logic. In (**2**) the theorem serves to establish the existence of convex polygons having any number of vertices when these vertices are to be selected from an arbitrary system of sufficiently many points in a plane. In (**3**) it is a principal tool in finding all extensions of the distributive law

$$(a+b)(c+d) = ac+ad+bc+bd$$

to the case where the factors on the left-hand side are replaced by convergent infinite series. Finally, Ramsey's theorem at once leads to Schur's result (**4**), which asserts the existence of a number n_k such that, whenever the numbers $1, 2, \ldots, n_k$ are arbitrarily distributed over k classes, at least one class contains three numbers x, y, z satisfying $x+y = z$. The estimate of n_k obtained in this way is, however, inferior to Schur's estimate $n_k < e\,k!$

The object of the present note is to prove a generalisation of Ramsey's theorem in which the number of classes of Δ need not be finite. We consider the term "distribution of the set Ω" as synonymous with "binary, reflexive, symmetrical, transitive relation in Ω". Let $N = \{1, 2, 3, \ldots\}$, and denote, for $n \varepsilon N$, by Ω_n the set of all subsets $\{a_1, a_2, \ldots, a_n\}$ of N, where $a_1 < a_2 < \ldots < a_n$. Let $k, \nu_1, \nu_2, \ldots, \nu_k$ be integers, $0 \leqslant k \leqslant n$, $0 < \nu_1 < \nu_2 < \ldots < \nu_k \leqslant n$. Consider the following special distribution of Ω_n, called the *canonical distribution* $\Delta^{(k)}_{\nu_1 \nu_2 \ldots \nu_k}$ of Ω_n. Two elements $\{a_1, \ldots, a_n\}$, $\{b_1, \ldots, b_n\}$ of Ω_n are in the same class of $\Delta^{(k)}_{\nu_1 \ldots \nu_k}$ if, and only if,

$$a_1 < a_2 < \ldots < a_n; \quad b_1 < b_2 < \ldots < b_n,$$

$$a_{\nu_1} = b_{\nu_1}; \quad a_{\nu_2} = b_{\nu_2}; \ldots; \quad a_{\nu_k} = b_{\nu_k}.$$

There are exactly 2^n canonical distributions of Ω_n. We mention the following two extreme cases of such distributions: (i) $\Delta^{(0)}$, the distribution

† Received 26 August, 1949; read 17 November, 1949.

in which all elements of Ω_n form one single class, (ii) $\Delta^{(n)}_{1,2,\ldots,n}$, the distribution in which every element of Ω_n forms a class by itself. We shall prove

THEOREM I. *Let n be a positive integer. Let Δ be an arbitrary distribution of all sets of n positive integers into classes. Then there is an infinite set N^* of positive integers and a canonical distribution $\Delta^{(k)}_{\nu_1\ldots\nu_k}$ such that, as far as subsets of N^* are concerned, the given distribution Δ coincides with the canonical distribution $\Delta^{(k)}_{\nu_1\ldots\nu_k}$.*

If, in particular, Δ has only a finite number of classes then the canonical distribution of Theorem I, having itself only a finite number of classes, must be $\Delta^{(0)}$, so that Ramsey's theorem follows from Theorem I.

2. It may be worth while to state explicitly the special case $n = 2$ of Theorem I.

THEOREM II. *Suppose that all pairs of positive integers (a, b), where $a < b$, are arbitrarily distributed into classes. Then there is an increasing sequence of integers $x_1, x_2, x_3, \ldots$ such that one of the following four sets of conditions holds, where it is assumed that $\alpha < \beta$; $\gamma < \delta$:*

(i) *All (x_α, x_β) belong to the same class.*

(ii) *(x_α, x_β) and (x_γ, x_δ) belong to the same class if, and only if, $\alpha = \gamma$.*

(iii) *(x_α, x_β) and (x_γ, x_δ) belong to the same class if, and only if, $\beta = \delta$.*

(iv) *(x_α, x_β) and (x_γ, x_δ) belong to the same class if, and only if, $\alpha = \gamma$; $\beta = \delta$.*

We shall deduce Theorem I from Ramsey's theorem. Our argument does not make any use of Zermelo's axiom. Ramsey stated explicitly† that his proof assumes Zermelo's axiom. It is, however, very easy to modify his proof in such a way that this axiom is not required. In order to establish Theorem I without the use of Zermelo's axiom, we give a brief account, in §5, of such a modified proof of Ramsey's theorem.

3. We introduce some notations and definitions. The letter Δ denotes distributions of objects into classes. The relation $X \equiv Y\ (.\Delta)$ expresses the fact that X and Y are objects distributed by Δ, and that X and Y belong to the same class of Δ. Letters A, B, C, D denote typical finite subsets of N. The number of elements of A is $|A|$. A relation‡

$$A_1 : A_2 : \ldots : A_m = B_1 : B_2 : \ldots : B_m \tag{1}$$

† (1), Theorem A.

‡ Set-theoretical operations are denoted by the common algebraic symbols, and brackets { } are only used in order to define sets by means of a list of their elements.

means that there exists a function $f(x)$, defined for $x \varepsilon A_1 + \ldots + A_m$ and having functional values in N, which has the properties:

$$\textit{if } x < y, \textit{ then } f(x) < f(y),$$

$$\sum_{x \varepsilon A_\mu} \{f(x)\} = B_\mu \quad (1 \leqslant \mu \leqslant m).$$

Thus (1) simply means that, as far as the order relation in N is concerned, the relative position of the sets A_μ to each other is the same as that of the B_μ. A relation

$$A_1 : A_2 = B_1 : B_2 = C_1 : C_2$$

is, by definition, equivalent to the simultaneous validity of the two relations

$$A_1 : A_2 = B_1 : B_2, \quad B_1 : B_2 = C_1 : C_2.$$

4. Using the notation and definitions of §3, we can state Theorem I, for a fixed $n \geqslant 0$, as follows†.

PROPOSITION P_n. *Let Δ be a distribution of Ω_n. Let C_0 be fixed, $|C_0| = n$. Then there is an infinite subset N^* of N and a subset C_0^* of C_0 such that the following condition holds: if*

$$A + B \subset N^*; \ |A| = |B| = n; \ A^* : A = B^* : B = C_0^* : C_0,$$

then $A \equiv B$ $(.\Delta)$ if, and only if, $A^* = B^*$.

A corollary of the proposition P_n is the following test for a distribution to be canonical.

THEOREM III. *A distribution Δ of Ω_n is canonical if, and only if, whenever*

$$A \equiv B\,(.\Delta); \ A : B = C : D,$$

then $C \equiv D\,(.\Delta)$.

5. Our "choice-free" version of Ramsey's proof of his theorem runs as follows. Let Δ be a distribution of Ω_n into a finite number of classes. We want to define an infinite subset $M(\Delta)$ of N such that, for some class κ of Δ, we have $A \varepsilon \kappa$ whenever $A \subset M(\Delta)$, $|A| = n$.

If $n = 0$, then we may put $M(\Delta) = N$. Let $n > 0$, and use induction with respect to n. If $a \varepsilon N - M$, where

$$M = \{x_1, x_2, \ldots\} \subset N; \quad x_1 < x_2 < \ldots,$$

† The case $n = 0$ is included merely in order to have an easy start of the induction proof which is to follow.

we define the distribution $\Delta(M, a)$ by putting

$$A \equiv B\big(.\Delta(M, a)\big)$$

if, and only if,

$$A+B \subset M, \quad |A|=|B|=n-1,$$
$$\{a\}+A \equiv \{a\}+B \,(.\Delta).$$

By induction hypothesis, applied† to the set M, in place of N, and the distribution $\Delta(M, a)$, there is a well-defined infinite subset $\sigma(M, a)$ of M and a class $\kappa(M, a)$ of Δ satisfying

$$\{a\}+A \in \kappa(M, a) \text{ whenever } A \subset \sigma(M, a),\ |A|=n-1.$$

Now we define, inductively, numbers a_k and sets M_k not containing a_k. Put $a_1 = 1$; $M_1 = N-\{1\}$. Let a_{l+1} be the least number of $\sigma(M_l, a_l)$, and put $M_{l+1} = \sigma(M_l, a_l)-\{a_{l+1}\}$ $(l = 0, 1, \ldots)$. Let k_0 be the least number such that $\kappa(M_{k_0}, a_{k_0}) = \kappa(M_k, a_k)$ for infinitely many k, and let $k_0, k_1, \ldots$ be all numbers k satisfying this last equation, $k_0 < k_1 < \ldots$. Then we may put $M(\Delta) = \{a_{k_0}, a_{k_1}, \ldots\}$. This proves Ramsey's theorem.

6. We now prove P_n. Clearly, P_0 is true. For we may put $N^* = N$; $C_0^* = C_0$. Let $n > 0$, and use induction with respect to n. Let Δ be a distribution of Ω_n. Choose some fixed D_0 satisfying $|D_0| = 2n$. Define, for any A such that $|A| = 2n$, the set $\phi(A)$ of pairs of subsets of D_0 by putting

$$\phi(A) = \sum_{\substack{A'+A'' \subset A \\ A' \equiv A''\,(.\Delta) \\ A' : A'' : A \,=\, D' : D'' : D_0}} \{(D', D'')\}.$$

The set $\phi(A)$ characterizes the effect of Δ on the subsets of A. We define Δ^* by putting

$$A \equiv B(.\Delta^*) \tag{2}$$

if, and only if, $|A| = |B| = 2n$; $\phi(A) = \phi(B)$.

Since Δ^* has only a finite number of classes it follows from Ramsey's theorem that there is an infinite subset M of N such that (2) holds whenever $A+B \subset M$; $|A|=|B|=2n$.

Without loss of generality we may assume that $M = N$. For all our arguments are only based on order relations in N.

---

† The relation $\nu \leftarrow\rightarrow x_\nu$ sets up a one-one mapping of N on M. By means of this mapping there corresponds to every well-defined subset of N a well-defined subset of M, and vice versa.

Consider any sets A', B', C', D' satisfying

(3) $$A' \equiv B' \,(.\Delta),$$

(4) $$A' : B' = C' : D'.$$

We want to deduce that

(5) $$C' \equiv D' \,(.\Delta).$$

According to (3), (4), one can choose A and B satisfying

$$A' + B' \subset A; \quad C' + D' \subset B; \quad |A| = |B| = 2n,$$

(6) $$A' : B' : A = C' : D' : B.$$

Then (2) holds and therefore, in view of (6), (3), and the definition of Δ^*, also (5). The fact that (3) and (4) imply (5) will briefly be described by saying that Δ is *invariant*.

Case 1. Suppose that $A \equiv B \,(.\Delta)$ only holds if $A = B$. Then the conclusion of P_n is true if we put $N^* = N$; $C_0^* = C_0$.

Case 2. Suppose that there are sets A_0, B_0 satisfying $A_0 \equiv B_0 \,(.\Delta)$; $A_0 \neq B_0$. Put

$$A_1 = \sum_{x \varepsilon A_0} \{2x\}; \quad B_1 = \sum_{x \varepsilon B_0} \{2x\}.$$

Then

$$A_0 : B_0 = A_1 : B_1,$$

and therefore, since Δ is invariant,

(7) $$A_1 \equiv B_1 \,(.\Delta).$$

As $A_0 \neq B_0$, we can choose $x_0 \,\varepsilon\, B_0 - A_0 B_0$. Put

$$B_2 = (B_1 - \{2x_0\}) + \{2x_0 + 1\}.$$

Then

$$A_0 : B_0 = A_1 : B_2$$

and hence, again on account of the invariance of Δ

(8) $$A_1 \equiv B_2 \,(.\Delta).$$

From (7) and (8),

(9) $$B_1 \equiv B_2 \,(.\Delta).$$

There is C_0' satisfying

$$(B_0 - \{x_0\}) : B_0 = C_0' : C_0.$$

Now consider any sets A_3, A_4 such that

$$|A_3| = |A_4| = n; \quad A_3 \neq A_4; \quad A' : A_3 = A' : A_4 = C_0' : C_0,$$

where A' is some suitable set. We shall show that

$$A_3 \equiv A_4\,(.\Delta). \tag{10}$$

We may assume that

$$A' = A_3 - \{x_3\} = A_4 - \{x_4\};\ x_3 < x_4.$$

Then $B_1 : B_2 = A_3 : A_4$, and therefore, since (9) holds and Δ is invariant, (10) follows. In other words, if sets A and A' satisfy $A' : A = C_0' : C_0$, then the class of Δ which contains A only depends on A' and not on $A - A'$. Hence every set A'' satisfying

$$|A''| = n-1;\quad A'' \subset \{2, 4, 6, \ldots\} = N'',$$

say, determines a unique class $K(A'')$ of Δ, namely that class which contains all A satisfying $A'' : A = C_0' : C_0$. Such sets A always exist.

Define Δ'' by putting

$$A'' \equiv B''\,(.\Delta'')$$

if, and only if,

$$|A''| = |B''| = n-1;\quad A'' + B'' \subset N'';\quad K(A'') = K(B'').$$

By induction hypothesis, the proposition P_{n-1} is true for Δ''. Thus there is an infinite subset N''' of N'' and a subset C_0''' of C_0' such that the following conditions are satisfied. Let

$$A'' + B'' \subset N''';\quad A''' : A'' = B''' : B'' = C_0''' : C_0'.$$

Then $A'' \equiv B''\,(.\Delta'')$ if, and only if, $A''' = B'''$. In view of the definition of Δ'', this means that the conclusion of P_n holds for the given distribution Δ if we put

$$N^* = N''';\ C_0{}^* = C_0'''.$$

This proves Proposition P_n and hence Theorems I and II.

7. We now prove Theorem III. First of all, suppose that Δ is a canonical distribution of Ω_n, say $\Delta = \Delta^{(k)}_{\nu_1 \ldots \nu_k}$. Then the relation $A \equiv B\,(.\Delta)$ means that

$$A = \{a_1, \ldots, a_n\};\ B = \{b_1, \ldots, b_n\},$$
$$a_1 < \ldots < a_n;\ b_1 < \ldots < b_n,$$
$$a_{\nu_\kappa} = b_{\nu_\kappa} (1 \leqslant \kappa \leqslant k).$$

If now

$$C = \{c_1, \ldots, c_n\};\ D = \{d_1, \ldots, d_n\},$$
$$c_1 < \ldots < c_n;\ d_1 < \ldots < d_n,$$

then the validity of $A:B=C:D$ implies that $c_{\nu_\kappa}=d_{\nu_\kappa}$ $(1\leqslant\kappa\leqslant k)$, *i.e.* that $C\equiv D\,(.\Delta)$. Hence Δ is invariant.

Vice versa, suppose that Δ is invariant. By Theorem I, we can find an infinite subset N^* of N, so that Δ is canonical in N^*, say $\Delta=\Delta^{(k)}_{\nu_1\ldots\nu_k}$, as far as subsets of N^* are concerned. But, since Δ is invariant in the whole set N, this obviously implies that $\Delta=\Delta^{(k)}_{\nu_1\ldots\nu_k}$ as far as all subsets of N are concerned. This proves Theorem III.

References.

1. F. P. Ramsey, " On a problem in formal logic ", *Proc. London Math. Soc.* (2), 30 (1930), 264–286.
2. P. Erdös and G. Szekeres, " A combinatorial problem in geometry ", *Compositio Mathematica*, 2 (1935), 463–470.
3. R. Rado, " The distributive law for products of infinite series ", *Quart. J. of Math.*, 11 (1940), 229–242.
4. I. Schur, " Ueber die Kongruenz $x^m+y^m\equiv z^m$ (mod p) ", *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 25 (1916), 114–117.

University of Illinois,
Urbana, Illinois, U.S.A.
King's College, London.

Printed by C. F. Hodgson & Son, Ltd., Pakenham St., London, W.C.1.