# Classic Papers in Combinatorics

Edited by

Ira Gessel Gian-Carlo Rota

# Classic Papers in Combinatorics

Edited by
Ira Gessel   Gian-Carlo Rota

Birkhäuser
Boston · Basel · Stuttgart

Ira Gessel
Department of Mathematics
Brandeis University
Waltham, MA 02254
U.S.A.

Gian-Carlo Rota
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA 02139
U.S.A.



Library of Congress Cataloging in Publication Data
Classic papers in combinatorics
I. Combinatorial analysis. I. Gessel, Ira.
II. Rota, Gian-Carlo, 1932–
QA164.C56 1987 511.6 87-5138

CIP-Kurztitelaufnahme der Deutschen Bibliothek
Classic papers in combinatorics / ed. by Ira
Gessel and Gian-Carlo Rota.—Boston ; Basel ;
Stuttgart : Birkhäuser, 1987
ISBN 3-7643-3364-2 (Basel ...)
ISBN 0-8176-3364-2 (Boston)
NE: Gessel, Ira [Hrsg.]



ISBN 0-8176-3364-2
ISBN 3-7643-3364-2

Printed and bound by Quinn-Woodbine Inc., Woodbine, New Jersey.
Printed in the U.S.A.

9 8 7 6 5 4 3 2 1

We wish to thank the following persons who have made valuable suggestions of papers to be included in this volume: George Andrews, Richard Brualdi, Andrew Gleason, Phil Hanlon, David Jackson, Jeff Kahn, Adalbert Kerber, Joseph Kung, and Richard Stanley.

# Introduction

This volume surveys the development of combinatorics since 1930 by presenting in chronological order the fundamental results of the subject proved in the orginal papers.

We begin with the celebrated theorem of Ramsey [1930], originally developed to settle a special case of the decision problem for the predicate calculus with equality. It remains to this day the fundamental generalization of the classical pigeonhole principle. The paper by Erdös and Szekeres [1935a] was one of the first applications of Ramsey's theorem, and it is still one of the most elegant. Through the partition calculus of Erdös and Rado [1956a], Ramsey's theorem made inroads into set theory, where nowadays it holds the limelight. The next major advance along the lines initiated by Ramsey came with the work of Hales and Jewett [1963a], a result which has served as a foundation for much further work in the area. The categorical underpinning of Ramsey theory was worked out by Graham, Leeb, and Rothschild [1972a]. Here the original ideas of Ramsey are cleverly blended with the contribution of Hales and Jewett.

Whitney's paper [1932] marks the beginning of what is now the theory of matroids. Three years later the theory makes its appearance fully clad in another paper of Whitney [1935c] which remains the basic reference on the subject. The theory of matroids was also in the backround of Tutte's paper [1947]. Tutte's paper is couched in the language of graphs and was later generalized to arbitrary matroids by Brylawski. The motivation behind much of the work of the two outstanding graph theorists of the day, Whitney and Tutte, was the coloring problem for graphs. Two short and elegant results on coloring problems are Brooks's theorem [1941] relating the chromatic number of a graph to its maximal degree and Lovász's theorem [1972b] characterizing perfect graphs.

Philip Hall's paper [1935b] was the first in what is now called matching theory. A very short proof of Hall's marriage theorem was given by Halmos and Vaughan [1950b]. In the same year Dilworth [1950a] proved his famous decomposition theorem for partially ordered sets, which generalizes Hall's theorem. Several other minimax combinatorial theorems can be viewed as variants or generalizations of the marriage theorem. Such are Tutte's definitive work on factors in graphs [1952], Ford and Fulkerson's theory of flows in networks [1956b], and Edmonds's [1965] efficient algorithm for matching in graphs. Gale [1957a] used network flow theory to prove a result on matrices of 0's and 1's with given row and column sums also proved directly by Ryser [1957b].

De Bruijn and van Aardenne-Ehrenfest [1951], taking their lead from the early work of Kirchhoff, obtained a definitive result, now called the BEST theorem (de Bruijn-Ehrenfest-Stone-Tutte) concerning the enumeration of spanning trees and Eulerian circuits of a graph by determinants. Several years later, Kasteleyn [1961a] succeeded in solving a packing problem for dimers on a lattice by reducing the problem to the evaluation of Pfaffians.

Pólya's paper on picture-writing [1956c] foreshadows the notion of the incidence algebra, a term introduced years later by Rota [1964] in his theory of Möbius functions. Rota's work was substantially extended by Crapo [1968]. The mystery of the characteristic polynomial, defined in terms of the Möbius function of a partially ordered set, motivates Stanley's beautiful result [1973a] on acyclic orientations of graphs. Geissinger's three papers [1973b–d] are the definitive presentation of the theory of Möbius functions.

The subject that is now called extremal set theory is represented by Katona's paper [1966a]. The main result, independently proved by Kruskal, harks back to a theorem of Macaulay, and was generalized by Clements and Lindström [1969]. Lubell's blitz proof of Sperner's theorem [1966b] has been extensively generalized and applied to many problems. Kleitman's solution [1970b] of a long-standing problem of Erdős related to the Littlewood-Offord problem shows the power of a simple, but far from obvious, induction argument.

Brooks, Smith, Stone, and Tutte's paper [1940] on the decomposition of rectangles into squares was the first to use Kirchhoff's laws for the solution of a problem in combinatorics, a technique that has since become standard.

Kaplansky's solution [1943] of the *problème des ménages* using the inclusion-exclusion principle has developed into what is now the theory of permutations with restricted position. Lovász's contribution [1972c] to the Ulam reconstruction problem is another ingenious use of the inclusion-exclusion principle.

Erdös's paper on graph theory and probability [1959] is the first paper to show how probabilistic methods can lead to combinatorial existence theorems. A substantial number of theorems in combinatorics, for which no explicit construction is known, can be given existence proofs by this method.

Schensted's bijection [1961b] between permutations and pairs of standard Young diagrams has proved central in the seemingly unrelated topics of plane partitions and representations of the symmetric group.

Schützenberger's paper [1962] lays the foundation of what is now the theory of rational and algebraic power series in noncommutative variables.

Nash-Williams's striking proof [1963b] that finite trees form a well-quasi-ordered set has blossomed both in logic and in graph theory.

I. J. Good's short proof [1970a] of a conjecture of Dyson has since been widely generalized but his approach is still the good one.

Ira M. Gessel
Gian-Carlo Rota

# Contents

# Classic Papers in Combinatorics

# ON A PROBLEM OF FORMAL LOGIC

*By* F. P. RAMSEY.

[Received 28 November, 1928.—Read 13 December, 1928.]

This paper is primarily concerned with a special case of one of the leading problems of mathematical logic, the problem of finding a regular procedure to determine the truth or falsity of any given logical formula*. But in the course of this investigation it is necessary to use certain theorems on combinations which have an independent interest and are most conveniently set out by themselves beforehand.

## I.

The theorems which we actually require concern finite classes only, but we shall begin with a similar theorem about infinite classes which is easier to prove and gives a simple example of the method of argument.

THEOREM A. *Let $\Gamma$ be an infinite class, and $\mu$ and $r$ positive integers; and let all those sub-classes of $\Gamma$ which have exactly $r$ members, or, as we may say, let all $r$-combinations of the members of $\Gamma$ be divided in any manner into $\mu$ mutually exclusive classes $C_i$ $(i = 1, 2, \ldots, \mu)$, so that every $r$-combination is a member of one and only one $C_i$; then, assuming the axiom of selections, $\Gamma$ must contain an infinite sub-class $\Delta$ such that all the $r$-combinations of the members of $\Delta$ belong to the same $C_i$.*

Consider first the case $\mu = 2$. (If $\mu = 1$ there is nothing to prove.) The theorem is trivial when $r$ is 1, and we prove it for all values of $r$ by induction. Let us assume it, therefore, when $r = \rho - 1$ and deduce it for $r = \rho$, there being, since $\mu = 2$, only two classes $C_i$, namely $C_1$ and $C_2$.

* Called in German the *Entscheidungsproblem*; see Hilbert und Ackermann, *Grundzüge der Theoretischen Logik*, 72–81.

It may happen that $\Gamma$ contains a member $x_1$ and an infinite sub-class $\Gamma_1$, not including $x_1$, such that the $\rho$-combinations consisting of $x_1$ together with any $\rho-1$ members of $\Gamma_1$, all belong to $C_1$. If so, $\Gamma_1$ may similarly contain a member $x_2$ and an infinite sub-class $\Gamma_2$, not including $x_2$, such that all the $\rho$-combinations consisting of $x_2$ together with $\rho-1$ members of $\Gamma_2$, belong to $C_1$. And, again, $\Gamma_2$ may contain an $x_3$ and a $\Gamma_3$ with similar properties, and so on indefinitely. We thus have two possibilities: either we can select in this way two infinite sequences of members of $\Gamma$ $(x_1, x_2, \ldots, x_n, \ldots)$, and of infinite sub-classes of $\Gamma$ $(\Gamma_1, \Gamma_2, \ldots, \Gamma_n, \ldots)$, in which $x_n$ is always a member of $\Gamma_{n-1}$, and $\Gamma_n$ a sub-class of $\Gamma_{n-1}$ not including $x_n$, such that all the $\rho$-combinations consisting of $x_n$ together with $\rho-1$ members of $\Gamma_n$, belong to $C_1$; or else the process of selection will fail at a certain stage, say the $n$-th, because $\Gamma_{n-1}$ (or if $n=1$, $\Gamma$ itself) will contain no member $x_n$ and infinite sub-class $\Gamma_n$ not including $x_n$ such that all the $\rho$-combinations consisting of $x_n$ together with $\rho-1$ members of $\Gamma_n$ belong to $C_1$. Let us take these possibilities in turn.

If the process goes on for ever let $\Delta$ be the class $(x_1, x_2, \ldots, x_n, \ldots)$. Then all these $x$'s are distinct, since if $r > s$, $x_r$ is a member of $\Gamma_{r-1}$ and so of $\Gamma_{r-2}, \Gamma_{r-3}, \ldots$, and ultimately of $\Gamma_s$ which does not contain $x_s$. Hence $\Delta$ is infinite. Also all $\rho$-combinations of members of $\Delta$ belong to $C_1$; for if $x_s$ is the term of such a combination with least suffix $s$, the other $\rho-1$ terms of the combination belong to $\Gamma_s$, and so form with $x_s$ a $\rho$-combination belonging to $C_1$. $\Gamma$ therefore contains an infinite sub-class $\Delta$ of the required kind.

Suppose, on the other hand, that the process of selecting the $x$'s and $\Gamma$'s fails at the $n$-th stage, and let $y_1$ be any member of $\Gamma_{n-1}$. Then the $(\rho-1)$-combinations of members of $\Gamma_{n-1}-(y_1)$ can be divided into two mutually exclusive classes $C_1'$ and $C_2'$ according as the $\rho$-combinations formed by adding to them $y_1$ belong to $C_1$ or $C_2$, and by our theorem (A), which we are assuming true when $r=\rho-1$ (and $\mu=2$), $\Gamma_{n-1}-(y_1)$ must contain an infinite sub-class $\Delta_1$ such that all $(\rho-1)$-combinations of the members of $\Delta_1$ belong to the same $C_i'$; *i.e.* such that the $\rho$-combinations formed by joining $y_1$ to $\rho-1$ members of $\Delta_1$ all belong to the same $C_i$. Moreover, this $C_i$ cannot be $C_1$, or $y_1$ and $\Delta_1$ could be taken to be $x_n$ and $\Gamma_n$ and our previous process of selection would not have failed at the $n$-th stage. Consequently the $\rho$-combinations formed by joining $y_1$ to $\rho-1$ members of $\Delta_1$ all belong to $C_2$. Consider now $\Delta_1$ and let $y_2$ be any of its members. By repeating the preceding argument $\Delta_1-(y_2)$ must contain an infinite sub-class $\Delta_2$ such that all the $\rho$-combinations got by joining $y_2$ to $\rho-1$ members of $\Delta_2$ belong to the same $C_i$.

And, again, this $C_i$ cannot be $C_1$, or, since $y_2$ is a member and $\Delta_2$ a sub-class of $\Delta_1$ and so of $\Gamma_{n-1}$ which includes $\Delta_1$, $y_2$ and $\Delta_2$ could have been chosen as $x_n$ and $\Gamma_n$ and the process of selecting these would not have failed at the $n$-th stage. Now let $y_3$ be any member of $\Delta_2$; then $\Delta_2-(y_3)$ must contain an infinite sub-class $\Delta_3$ such that all $\rho$-combinations consisting of $y_3$ together with $\rho-1$ members of $\Delta_3$, belong to the same $C_i$, which, as before, cannot be $C_1$ and must be $C_2$. And by continuing in this way we shall evidently find two infinite sequences $y_1, y_2, \ldots, y_n, \ldots$ and $\Delta_1, \Delta_2, \ldots, \Delta_n, \ldots$ consisting respectively of members and sub-classes of $\Gamma$, and such that $y_n$ is always a member of $\Delta_{n-1}$, $\Delta_n$ a sub-class of $\Delta_{n-1}$ not including $y_n$, and all the $\rho$-combinations formed by joining $y_n$ to $\rho-1$ members of $\Delta_n$ belong to $C_2$; and if we denote by $\Delta$ the class $(y_1, y_2, \ldots, y_n, \ldots)$ we have, by a previous argument, that all $\rho$-combinations of members of $\Delta$ belong to $C_2$.

Hence, in either case, $\Gamma$ contains an infinite sub-class $\Delta$ of the required kind, and Theorem A is proved for all values of $r$, provided that $\mu=2$. For higher values of $\mu$ we prove it by induction; supposing it already established for $\mu=2$ and $\mu=\nu-1$, we deduce it for $\mu=\nu$.

The $r$-combinations of members of $\Gamma$ are then divided into $\nu$ classes $C_i$ $(i=1, 2, \ldots, \nu)$. We define new classes $C'_i$ for $i=1, 2, \ldots, \nu-1$ by

$$C'_i = C_i \quad (i=1, 2, \ldots, \nu-2),$$

$$C'_{\nu-1} = C_{\nu-1}+C_\nu.$$

Then by the theorem for $\mu=\nu-1$, $\Gamma$ must contain an infinite sub-class $\Delta$ such that all $r$-combinations of the members of $\Delta$ belong to the same $C'_i$. If, in this $C'_i$, $i \leqslant \nu-2$, they all belong to the same $C_i$, which is the result to be proved; otherwise they all belong to $C'_{\nu-1}$, *i.e.* either to $C_{\nu-1}$ or to $C_\nu$. In this case, by the theorem for $\mu=2$, $\Delta$ must contain an infinite sub-class $\Delta'$ such that the $r$-combinations of members of $\Delta'$ either all belong to $C_{\nu-1}$ or all belong to $C_\nu$; and our theorem is thus established.

Coming now to finite classes it will save trouble to make some conventions as to notation. Small letters other than $x$ and $y$, whether Italic or Greek (*e.g.* $n$, $r$, $\mu$, $m$) will always denote finite cardinals, positive unless otherwise stated. Large Greek letters (*e.g.* $\Gamma$, $\Delta$) will denote classes, and their suffixes will indicate the number of their members (*e.g.* $\Gamma_m$ is a class with $m$ members). The letters $x$ and $y$ will represent members of the classes $\Gamma$, $\Delta$, etc., and their suffixes will be used merely to distinguish them. Lastly, the letter $C$ will stand, as before, for classes of combinations, and its suffixes will not refer to the number of members, but serve merely to distinguish the different classes of combinations considered.

Corresponding to Theorem A we then have

THEOREM B. *Given any $r$, $n$, and $\mu$ we can find an $m_0$ such that, if $m \geqslant m_0$ and the $r$-combinations of any $\Gamma_m$ are divided in any manner into $\mu$ mutually exclusive classes $C_i$ $(i = 1, 2, \ldots, \mu)$, then $\Gamma_m$ must contain a sub-class $\Delta_n$ such that all the $r$-combinations of members of $\Delta_n$ belong to the same $C_i$.*

This is the theorem which we require in our logical investigations, and we should at the same time like to have information as to how large $m_0$ must be taken for any given $r$, $n$, and $\mu$. This problem I do not know how to solve, and I have little doubt that the values for $m_0$ obtained below are far larger than is necessary.

To prove the theorem we begin, as in Theorem A, by supposing that $\mu = 2$. We then take, not Theorem B itself, but the equivalent

THEOREM C. *Given any $r$, $n$, and $k$ such that $n+k \geqslant r$, there is an $m_0$ such that, if $m \geqslant m_0$ and the $r$-combinations of any $\Gamma_m$ are divided into two mutually exclusive classes $C_1$ and $C_2$, then $\Gamma_m$ must contain two mutually exclusive sub-classes $\Delta_n$ and $\Lambda_k$ such that all the combinations formed by $r$ members of $\Delta_n + \Lambda_k$ which include at least one member from $\Delta_n$ belong to the same $C_i$.*

That this is equivalent to Theorem B with $\mu = 2$ is evident from the fact that, for any given $r$, Theorem C, for $n$ and $k$, asserts more than Theorem B for $n$, but less than Theorem B for $n+k$.

The proof of Theorem C must be performed by mathematical induction, and can conveniently be set out as a demonstration that it is possible to define by recursion a function $f(r, n, k)$ which will serve as $m_0$ in the theorem.

If $r = 1$, the theorem is evidently true with $m_0$ equal to the greater of $2n-1$ and $n+k$, so that we may define

$$f(1, n, k) = \max(2n-1, n+k) \quad (n \geqslant 1,\ k \geqslant 0).$$

For other values of $r$ we define $f(r, n, k)$ by recursion formulae involving an auxiliary function $g(r, n, k)$. Suppose that $f(r-1, n, k)$ has been defined for a certain $r-1$, and all $n$, $k$ such that $n+k \geqslant r-1$, then we define it for $r$ by putting

$$\begin{aligned} f(r, 1, k) &= f(r-1, k-r+2, r-2)+1 && (k+1 \geqslant r), \\ g(r, 0, k) &= \max(r-1, k), \\ g(r, n, k) &= f\{r, 1, g(r, n-1, k)\} && (n \geqslant 1), \\ f(r, n, k) &= f\{r, n-1, g(r, n, k)\} && (n > 1). \end{aligned}$$

5

These formulae can be easily seen to define $f(r, n, k)$ for all positive values of $r$, $n$ and $k$ satisfying $n+k \geqslant r$, and $g(r, n, k)$ for all values of $r$ greater than 1, and all positive values of $n$ and $k$; and we shall prove that Theorem C is true when we take $m_0$ to be this $f(r, n, k)$. We know that this is so when $r = 1$, and we shall therefore assume it for all values up to $r-1$ and deduce it for $r$.

When $n = 1$, and $m \geqslant m_0 = f(r-1, k-r+2, r-2)+1$, we may take any member $x$ of $\Gamma_m$ to be sole member of $\Delta_1$ and there remain at least $f(r-1, k-r+2, r-2)$ members of $\Gamma_m-(x)$; the $(r-1)$-combinations of these members of $\Gamma_m-(x)$ can be divided into classes $C'_1$ and $C'_2$ according as they belong to $C_1$ or $C_2$ when $x$ is added to them, and, by our theorem for $r-1$, $\Gamma_m-(x)$ must contain two mutually exclusive classes $\Delta_{k-r+2}$, $\Lambda_{r-2}$ such that every combination of $r-1$ terms from $\Delta_{k-r+2}+\Lambda_{r-2}$ (since one of its terms must come from $\Delta_{k-r+2}$, $\Lambda_{r-2}$ having only $r-2$ members) belongs to the same $C'_i$. Taking $\Lambda_k$ to be this $\Delta_{k-r+2}+\Lambda_{r-2}$ all combinations consisting of $x$, together with $r-1$ members of $\Lambda_k$, belong to the same $C_i$. The theorem is therefore true for $r$ when $n = 1$.

For other values of $n$ we prove it by induction, assuming it for $n-1$ and deducing it for $n$. Taking

$$m \geqslant m_0 = f(r, n, k) = f\{r, n-1, g(r, n, k)\},$$

$\Gamma_m$ must, by the theorem for $n-1$, contain a $\Delta_{n-1}$ and a $\Lambda_{g(r, n, k)}$ such that every combination of $r$ members of $\Delta_{n-1}+\Lambda_{g(r, n, k)}$, at least one term of which comes from $\Delta_{n-1}$, belongs to the same $C_i$, say to $C_1$. If, now, $\Lambda_{g(r, n, k)}$ contains a member $x$ and a sub-class $\Lambda_k$ not including $x$, such that every combination of $x$ and $r-1$ members of $\Lambda_k$ belongs to $C_1$, then, taking $\Delta_n$ to be $\Delta_{n-1}+(x)$ and $\Lambda_k$ to be this $\Lambda_k$, our theorem is true. If not, there can be no member of $\Lambda_{g(r, n, k)}$ which has a sub-class of $k$ members of $\Lambda_{g(r, n, k)}$ connected with it in this way. But since

$$g(r, n, k) = f\{r, 1, g(r, n-1, k)\},$$

$\Lambda_{g(r, n, k)}$ must contain a member $x_1$ and a sub-class $\Lambda_{g(r, n-1, k)}$, not including $x_1$, such that $x_1$ combined with any $r-1$ members of $\Lambda_{g(r, n-1, k)}$ gives a combination belonging to the same $C_i$, which cannot be $C_1$, or $x_1$ and any $k$ members of $\Lambda_{g(r, n-1, k)}$ could have been taken as the $x$ and $\Lambda_k$ above. Hence the combinations formed by $x_1$ together with any $r-1$ members of $\Lambda_{g(r, n-1, k)}$ all belong to $C_2$. But now

$$g(r, n-1, k) = f\{r, 1, g(r, n-2, k)\},$$

and $\Lambda_{g(r, n-1, k)}$ must contain an $x_2$ and a $\Lambda_{g(r, n-2, k)}$, not including $x_2$, such that the combinations formed by $x_2$ and $r-1$ members of $\Lambda_{g(r, n-2, k)}$ all belong to the same $C_i$, which must, as before, be $C_2$, since $x_2$ and $\Lambda_{g(r,\,n-2,\,k)}$ are both contained in $\Lambda_{g(r,\,n,\,k)}$ and $g(r,\,n-2,\,k) \geqslant k$. Continuing in this way we can find $n$ distinct terms $x_1,\,x_2,\,\ldots,\,x_n$ and a $\Lambda_{g(r,\,0,\,k)}$ such that every combination of $r$ terms from $(x_1,\,x_2,\,\ldots,\,x_n)+\Lambda_{g(r,\,0,\,k)}$ belongs to $C_2$, provided that at least one term of the combination comes from $(x_1,\,x_2,\,\ldots,\,x_n)$. Since $g(r,\,0,\,k) \geqslant k$ this proves our theorem, taking $\Delta_n$ to be $(x_1,\,x_2,\,\ldots,\,x_n)$ and $\Lambda_k$ to be any $k$ terms of $\Lambda_{g(r,\,0,\,k)}$.

Theorem C is therefore established for all values of $r$, $n$, and $k$, with $m_0$ equal to $f(r,\,n,\,k)$. It follows that, if $\mu = 2$, Theorem B is true for all values of $r$ and $n$ with $m_0$ equal to $f(r,\,n-r+1,\,r-1)$, which we shall also call $h(r,\,n,\,2)$.

For other values of $\mu$ we prove Theorem B by induction, taking $m_0$ to be $h(r,\,n,\,\mu)$, where

$$h(r,\,n,\,2) = f(r,\,n-r+1,\,r-1)$$

$$h(r,\,n,\,\mu) = h\{r,\,h(r,\,n,\,\mu-1),\,2\} \quad (\mu > 2).$$

For, assuming the theorem for $\mu-1$, we prove it for $\mu$ by defining new classes of combinations

$$C_1' = C_1,$$

$$C_2' = \sum_{i=2}^{\mu} C_i.$$

If then $m \geqslant h(r,\,n,\,\mu) = h\{r,\,h(r,\,n,\,\mu-1),\,2\}$, by the theorem for $\mu = 2$, $\Gamma_m$ must contain a $\Gamma_{h(r,\,n,\,\mu-1)}$ the $r$-combinations of whose members belong either all to $C_1'$ or all to $C_2'$. In the first case there is no more to prove; in the second we have only to apply the theorem for $\mu-1$ to $\Gamma_{h(r,\,n,\,\mu-1)}$.

In the simplest case in which $r = \mu = 2$ the above reasoning gives $m_0$ equal to $h(2,\,n,\,2)$, which is easily shown to be $2^{n(n-1)/2}$. But for this case there is a simple argument which gives the much lower value $m_0 = n\,!$, and shows that our value $h(r,\,n,\,\mu)$ is altogether excessive.

For, taking Theorem C first, we can prove by induction with regard to $n$ that, for $r = 2$, we may take $m_0$ to be $k\,.\,(n+1)\,!$. ($k$ is here supposed greater than or equal to 1.) For this is true when $n = 1$, since, if $m \geqslant 2k$, of the $m-1$ pairs obtained by combining any given member of $\Gamma_m$ with the others, at least $k$ must belong to the same $C_i$. Assuming it, then, for $n-1$, let us prove it for $n$.

If $m \geqslant k\,.\,(n+1)\,! = k(n+1)\,.\,n\,!$, $\Gamma_m$ must, by the theorem for $n-1$, contain two mutually exclusive sub-classes $\Delta_{n-1}$ and $\Lambda_{k(n+1)}$ such that all pairs from $\Delta_{n-1}+\Lambda_{k(n+1)}$, at least one term of which comes from $\Delta_{n-1}$, belong to the same $C_i$, say $C_1$. Now consider the members of $\Lambda_{k(n+1)}$; in the first place, there may be one of these, $x$ say, which is such that there are $k$ other members of $\Lambda_{k(n+1)}$ which combined with $x$ give pairs belonging to $C_1$. If so, the theorem is true, taking $\Delta_n$ to be $\Delta_{n-1}+(x)$; if not, let $x_1$ be any member of $\Lambda_{k(n+1)}$. Then there are at most $k-1$ other members of $\Lambda_{k(n+1)}$ which combined with $x_1$ give pairs belonging to $C_1$, and $\Lambda_{k(n+1)}-(x_1)$ must contain a $\Lambda_{kn}$ any member of which gives when combined with $x_1$ a pair belonging to $C_2$. Let $x_2$ be any member of $\Lambda_{kn}$, then, since $x_2$ and $\Lambda_{kn}$ are both contained in $\Lambda_{k(n+1)}$, there are at most $k-1$ other members of $\Lambda_{kn}$ which when combined with $x_2$ give pairs belonging to $C_1$. Hence $\Lambda_{kn}-(x_2)$ contains a $\Lambda_{k(n-1)}$ any member of which combined with $x_2$ gives a pair belonging to $C_2$. Continuing in this way we obtain $x_1, x_2, \ldots, x_n$ and $\Lambda_k$, such that every pair $x_i, x_j$ and every pair consisting of an $x_i$ and a member of $\Lambda_k$ belongs to $C_2$. Theorem C is therefore proved.

Theorem B for $n$ then follows, with the $m_0$ of Theorem C for $n-1$ and 1, *i.e.* with $m_0$ equal to $n!$*; and it is an easy extension to show that, if in Theorem B $r=2$ but $\mu \neq 2$, we can take $m_0$ to be $n!!!$, ..., where the process of taking the factorial is performed $\mu-1$ times.

## II.

We shall be concerned with logical formulae containing variable propositional functions, *i.e.* predicates or relations, which we shall denote by Greek letters $\phi$, $\chi$, $\psi$, etc. These functions have as arguments individuals denoted by $x$, $y$, $z$, etc., and we shall deal with functions with any finite number of arguments, *i.e.* of any of the forms

$$\phi(x),\ \chi(x, y),\ \psi(x, y, z),\ \ldots.$$

In addition to these variable functions we shall have the one constant function of identity

$$x = y \quad \text{or} \quad =(x, y).$$

By operating on the values of $\phi$, $\chi$, $\psi$, ..., and $=$ with the logical operations

| | | |
|---|---|---|
| $\sim$ | meaning | *not*, |
| $\vee$ | ,, | *or*, |
| $\cdot$ | ,, | *and*, |
| $(x)$ | ,, | *for all x*. |
| $(Ex)$ | ,, | *there is an x for which*, |

* But this value is, I think, still much too high. It can easily be lowered slightly even when following the line of argument above, by using the fact that if $k$ is even it is impossible for every member of an odd class to have exactly $k-1$ others with which it forms a pair of $C_1$, for then twice the number of these pairs would be odd; we can thus start when $k$ is even with a $\Lambda_{k(n+1)-1}$ instead of a $\Lambda_{k(n+1)}$

we can construct expressions such as

$$[(x, y)\{\phi(x, y) \vee x = y\}] \vee \{(Ez)\chi(z)\}$$

in which all the individual variables are made "apparent" by prefixes $(x)$ or $(Ex)$, and the only real variables left are the functions $\phi$, $\chi$, .... Such an expression we shall call a *first order formula*.

If such a formula is true for all interpretations* of the functional variables $\phi$, $\chi$, $\psi$, etc., we shall call it *valid*, and if it is true for no interpretations of these variables we shall call it *inconsistent*. If it is true for some interpretations (whether or not for all) we shall call it *consistent*†.

The *Entscheidungsproblem* is to find a procedure for determining whether any given formula is valid, or, alternatively, whether any given formula is consistent; for these two problems are equivalent, since the necessary and sufficient condition for a formula to be consistent is that its contradictory should not be valid. We shall find it more convenient to take the problem in this second form as an investigation of *consistency*. The consistency of a formula may, of course, depend on the number of individuals in the universe considered, and we shall have to distinguish between formulae which are consistent in every universe and those which are only consistent in universes with some particular numbers of members. Whenever the universe is infinite we shall have to assume the axiom of selections.

The problem has been solved by Behmann‡ for formulae involving only functions of one variable, and by Bernays and Schönfinkel§ for formulae involving only two individual apparent variables. It is solved below for the further case in which, when the formula is written in "normal form", there are any number of prefixes of generality $(x)$ but none of existence $(Ex)$‖. By "normal form"¶ is here meant that all the prefixes stand at the beginning, with no negatives between or in front of them, and have scopes extending to the end of the formula.

* To avoid confusion we call a constant function substituted for a variable $\phi$, not a value but an *interpretation* of $\phi$; the *values* of $\phi(x, y, z)$ are got by substituting constant individuals for $x$, $y$, and $z$.

† German *erfüllbar*.

‡ H. Behmann, "Beiträge zur Algebra der Logik und zum Entscheidungsproblem", *Math. Annalen*, 86 (1922), 163–229.

§ P. Bernays und M. Schönfinkel, "Zum Entscheidungsproblem der mathematischen Logik", *Math. Annalen*, 99 (1928), 342–372. These authors do not, however, include identity in the formulae they consider.

‖ Later we extend our solution to the case in which there are also prefixes of existence provided that these all precede all the prefixes of generality.

¶ Hilbert und Ackermann, *op. cit.*, 63–4.