JOURNAL OF COMBINATORIAL THEORY (A) 16, 398–400 (1974)

Note

The Asymptotic Number of Geometries*

DONALD E. KNUTH

Stanford University and University of Oslo

Communicated by Gian-Carlo Rota

Received October 25, 1972

A simple proof is given that $\lim_{n\to\infty}(\log_2 \log_2 g_n)/n = 1$, where g_n denotes the number of distinct combinatorial geometries on n points.

Let g_n be the number of combinatorial geometries on n points (i.e., matroids with all 1-point sets independent). The following values of g_n for small n have been determined by Blackburn, Crapo, and Higgs [1]:

$$\begin{array}{llllllllll} n & = & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ g_n & = & 1 & 1 & 2 & 4 & 9 & 26 & 101 & 950. \end{array}$$

Crapo and Rota [2, p. 3.3] noted that the law

$$g_{n+1} \approx g_n^{3/2} \tag{1}$$

"seems approximately correct, on the basis of this data alone." If this "approximation" were valid for all n, we would have

$$\log_2 g_{n+1} \approx (3/2) \log_2 g_n ,$$

so that $\log_2 \log_2 g_n \approx n \log_2(3/2)$. On the other hand, every geometry is defined by specifying a certain set of subsets of the n points (e.g., the closed sets, the bonds, the bases, or the circuits), hence obviously

$$g_n \leqslant 2^{2^n}. \tag{2}$$

In other words, $\log_2 \log_2 g_n \leqslant n$.

* This research was supported in part by the National Science Foundation grant GJ-992, the Office of Naval Research contract ONR 00014-67-A-0112-0057 NR 044-402, and Norges Almenvitenskapelige Forskningsråd.

AMS categories: Primary 05B35, Secondary 05B40.

M. J. Piff and D. J. A. Welsh [4] have shown that g_n eventually grows more rapidly than the first few values would indicate. In fact, they have proved that

$$\log_2 \log_2 g_n \geqslant n - (5/2) \log_2 n + O(\log \log n), \tag{3}$$

so g_{n+1} will be roughly $g_n{}^2$ when n is large. On the other hand, Piff [3] has recently improved the upper bound (2) to

$$\log_2 \log_2 g_n \leqslant n - \log_2 n + O(\log \log n). \tag{4}$$

The purpose of this note is to narrow the gap a little further, by showing that

$$\log_2 \log_2 g_n \geqslant n - (3/2) \log_2 n + O(\log \log n). \tag{5}$$

The precise result to be proved is that

$$g_n \geqslant 2^{\binom{n}{\lfloor n/2 \rfloor}/2n}/n!, \tag{6}$$

from which (5) follows by Stirling's approximation. Indeed, the lower bound in (6) is less than or equal to the number of geometries of a very special kind, namely the so-called "partitions of type $\lfloor n/2 \rfloor - 1$." This confirms a remark of Crapo and Rota [2, p. 3.17], who conjectured that partition geometries would probably predominate in any asymptotic enumeration.

The factor $n!$ in (6) accounts for any isomorphisms between the geometries we shall construct, so we shall ignore isomorphisms in what follows.

Let M be a family of subsets of $\{1, 2,..., n\}$, where each subset contains exactly $\lfloor n/2 \rfloor$ elements, and where no two different subsets have more than $\lfloor n/2 \rfloor - 2$ elements in common. This set M, together with the set of all $\lfloor n/2 \rfloor - 1$ element subsets which are not contained in any member of M, constitutes a set of blocks such that every subset of size $\lfloor n/2 \rfloor - 1$ is contained in a unique block; therefore it defines a partition geometry.

If M contains m members, each of the 2^m subfamilies of M will define a partition geometry in the same way. Therefore (6) will follow if we can find such a family M of subsets, containing at least $m \geqslant \binom{n}{\lfloor n/2 \rfloor}/2n$ members. This is essentially the approach used in [4], although the authors of [4] did not construct such a large family M.

The problem is solved by realizing that it is the same as finding m binary code words of length n, each containing exactly $\lfloor n/2 \rfloor$ 1 bits, and single-error correcting. This characterization suggests the following "Hamming code" construction: Let $k = \lfloor \log_2 n \rfloor + 1$, and construct the $n \times k$ matrix H of 0's and 1's whose rows are the numbers from 1 to n

expressed in binary notation. For $0 \leqslant j < 2^k$, consider the set M_j of all row vectors x of 0's and 1's such that x contains exactly $\lfloor n/2 \rfloor$ 1 bits and the vector xH mod 2 is the binary representation of j. Note that if x and y are distinct elements of M_j, they cannot differ in just two places; otherwise we would have $(x + y) H \bmod 2 = (0 \cdots 0)$, contradicting the fact that no two rows of H are equal. Therefore M_j defines a family of subsets of $\{1, 2,..., n\}$ having the desired property. Furthermore, at least one of these 2^k families M_j will contain $\binom{n}{\lfloor n/2 \rfloor}/2^k$ or more elements, since they are disjoint and they exhaust all of the $\binom{n}{\lfloor n/2 \rfloor}$ possible $\lfloor n/2 \rfloor$-element subsets. This completes the proof, since $2^k \leqslant 2n$.

References

1. J. E. Blackburn, H. H. Crapo, and D. A. Higgs, A catalogue of combinatorial geometries, *Math. Comp.* **27** (1973), 155–166.
2. H. H. Crapo and G.-C. Rota, "On the Foundations of Combinatorial Theory: Combinatorial Geometries" Preliminary edition, MIT Press, 1970.
3. M. J. Piff, An upper bound for the number of matroids, *J. Combinatorial Theory Series B* **14** (1973), 241–245.
4. M. J. Piff and D. J. A. Welsh, The number of combinatorial geometries, *Bull. Lond Math. Soc.* **3** (1971), 55–56.