ON THE NUMBER OF SPARSE PAVING MATROIDS

DILLON MAYHEW AND DOMINIC WELSH

Dedicated to Geoff Whittle, in appreciation of his friendship and wisdom.

Abstract. Let $sp(n)$ be the number of sparse paving matroids on the ground set $\{1,\ldots,n\}$. We prove that $\log\log sp(n) = n - (3/2)\log n + O(\log\log n)$, and we conjecture that the same equality applies to the number of all matroids on the set $\{1,\ldots,n\}$.

1. Introduction

In 1973 Piff [4] proved the following upper bound on $m(n)$, the number of matroids on the ground set $\{1,\ldots,n\}$:

$$m(n) \le n^{k2^n n^{-1}}, \tag{1}$$

when $n \ge 2$, and where k is a fixed constant.

A year later, Knuth [2] showed that

$$2^{\binom{n}{\lfloor n/2\rfloor}(2n)^{-1}} \le m(n).$$

By adapting his argument, we can establish the following very slight improvement.

$$2^{\binom{n}{\lfloor n/2\rfloor}n^{-1}} \le m(n). \tag{2}$$

To see that Equation (2) holds, note that Theorem 1 of Graham and Sloane [1] implies that for any positive integer n, there is a code of at least

$$\binom{n}{\lfloor n/2\rfloor}n^{-1}$$

words with length n, constant weight $\lfloor n/2\rfloor$, and minimum distance at least 4. Therefore, there exists a family $\mathcal{C}$ of at least $\binom{n}{\lfloor n/2\rfloor}n^{-1}$ subsets of $\{1,\ldots,n\}$, such that $|C| = \lfloor n/2\rfloor$ for every $C \in \mathcal{C}$, and $|C \cup C'| \ge \lfloor n/2\rfloor + 2$ for every pair, $\{C, C'\}$, of distinct members of $\mathcal{C}$. Thus $\mathcal{C}$ is the family of non-spanning circuits of a paving matroid with rank $\lfloor n/2\rfloor$. The same statement is true of any subfamily of $\mathcal{C}$, so there are at least $2^{|\mathcal{C}|}$ distinct paving matroids on the set $\{1,\ldots,n\}$. Equation (2) follows. (Recall that a rank-r matroid is *paving* if every set with cardinality $r-1$ is independent.)

2010 *Mathematics Subject Classification.* 05B35.
Key words and phrases. Matroids, sparse paving, asymptotic, enumeration.

It is relatively straightforward to prove that $2^{n-1}n^{-1/2} \leq \binom{n}{\lfloor n/2 \rfloor}$ for all positive integers n. By combining this fact with Equations (1) and (2), we see that

$$n - (3/2)\log n - 1 \leq \log\log m(n) \leq n - \log n + \log\log n + O(1).$$

This represents the current state of knowledge on the matroid enumeration question. (Note that throughout this paper, logarithms will be taken to the base 2.)

Recall that a matroid is *sparse paving* if both it and its dual are paving. Let $sp(n)$ be the number of sparse paving matroids on the ground set $\{1, \ldots, n\}$. In a recent paper [3], the authors conjecture that asymptotically almost every matroid is paving, and point out that this implies that asymptotically almost every matroid is sparse paving. That is, they make the following conjecture:

Conjecture 1.1. *The limit* $\lim_{n\to\infty} sp(n)/m(n)$ *exists, and is equal to one.*

The purpose of this note is to show that when we apply Piff's techniques [4] to sparse paving matroids, we arrive at the following result.

Theorem 1.2. $\log\log sp(n) \leq n - (3/2)\log n + \log\log n + O(1)$.

It is easy to see that the matroids we constructed when establishing Equation (2) are all sparse paving. Combining this observation with Theorem 1.2 gives the following corollary.

Corollary 1.3. $\log\log sp(n) = n - (3/2)\log n + O(\log\log n)$.

This result, and our belief that sparse paving matroids predominate, lead us to make the following conjecture [1].

Conjecture 1.4. $\log\log m(n) = n - (3/2)\log n + O(\log\log n)$.

Although Corollary 1.3 determines $\log\log sp(n)$ with quite a high level of precision, it doesn't come close to providing us with an asymptotic formula for $sp(n)$. Even determining $\log\log sp(n)$ to within an additive constant would fail to achieve this goal. Therefore Conjecture 1.4 may be significantly weaker than Conjecture 1.1 (and perhaps easier to prove). Although $\lim_{n\to\infty} sp(n)/m(n) = 1$ would certainly imply Conjecture 1.4 (by virtue of Corollary 1.3), it is *a priori* possible that $sp(n)$ and $m(n)$ are not asymptotically equal, even though

$$\log\log sp(n) = n - (3/2)\log n + O(\log\log n) = \log\log m(n).$$

[1]Since the time of writing, Conjecture 1.4 has been proved by Bansal, Pendavingh, and Van der Pol, who have shown that $\log\log m(n) \leq n - (3/2)\log n + (1/2)\log(2/\pi) + 1 + o(1)$.

2. Proof of the main theorem

The proof depends on the following intermediate lemmas.

Lemma 2.1. *Let n be a positive integer. Then*

$$\binom{n}{\lfloor n/2 \rfloor} \leq \left(\sqrt{\frac{2}{\pi}}\right) 2^n n^{-1/2}.$$

We believe that Lemma 2.1 is likely to be known, but we sketch the argument for the sake of completeness, as we have been unable to locate a proof in the literature.

Sketch proof of Lemma 2.1. For any positive integer n, define $f(n)$ to be

$$\frac{\binom{n}{\lfloor n/2 \rfloor}}{2^n n^{-1/2}}.$$

It is routine to check that both

$$f(1), f(3), f(5), \ldots \quad \text{and} \quad f(2), f(4), f(6), \ldots$$

are increasing sequences. Moreover, Stirling's formula implies that $f(1), f(2), f(3), \ldots$ converges to $\sqrt{2/\pi}$. Therefore $f(n) \leq \sqrt{2/\pi}$ for every n, as desired. □

Note that Lemma 2.1 implies that

$$\binom{n}{\lfloor n/2 \rfloor} \leq 2^n n^{-1/2}. \tag{3}$$

The following fact is Lemma 1 of [4].

Lemma 2.2. *Let n and r be integers satisfying $1 \leq r \leq n$. Then*

$$\binom{n}{r} \leq \left(\frac{en}{r}\right)^r.$$

For integers $0 \leq r \leq n$, let $sp_r(n)$ denote the number of sparse paving matroids on the set $\{1, \ldots, n\}$ with rank r.

Lemma 2.3. *Let n and r be integers satisfying $0 \leq r \leq n$. Let $M(n, r)$ be*

$$\left\lfloor \frac{1}{n-r+1} \binom{n}{r} \right\rfloor.$$

Then

$$sp_r(n) \leq \sum_{i=0}^{M(n,r)} \binom{\binom{n}{r}}{i}.$$

Proof. Consider a sparse paving matroid on the set $\{1, \ldots, n\}$ with rank r. Let h be the number of non-spanning circuits. Sparse paving matroids are characterized by the fact that each non-spanning circuit is a hyperplane. Therefore each non-spanning circuit contains r sets of size $r - 1$, and any set of size $r-1$ is contained in at most one non-spanning circuit. It follows that

$$rh \leq \binom{n}{r-1},$$

and therefore

$$h \leq \left\lfloor \frac{1}{r}\binom{n}{r-1} \right\rfloor = M(n,r).$$

Since a sparse paving matroid is completely determined by its non-spanning circuits, the number of sparse paving matroids on the set $\{1,\ldots,n\}$ with rank r and i non-spanning circuits is clearly no greater than

$$\binom{\binom{n}{r}}{i}.$$

Summing this formula as i ranges from 0 to $M(n,r)$ gives the result. □

Lemma 2.4. *Let $n \geq 0$ be an integer. Then*

$$sp(n) \leq (n+1) \sum_{i=0}^{M(n,\lfloor n/2 \rfloor)} \binom{\binom{n}{\lfloor n/2 \rfloor}}{i}.$$

Proof. Note $sp(n) = sp_0(n) + \cdots + sp_n(n)$. Therefore it suffices to show that

$$sp_r(n) \leq \sum_{i=0}^{M(n,\lfloor n/2 \rfloor)} \binom{\binom{n}{\lfloor n/2 \rfloor}}{i}$$

for every $r \in \{0,\ldots,n\}$. By duality, $sp_r(n) = sp_{n-r}(n)$, so we assume that $r \leq n/2$.

Since

$$\binom{n}{r} \leq \binom{n}{\lfloor n/2 \rfloor}, \tag{4}$$

the result will follow from Lemma 2.3, if we can show that

$$M(n,r) \leq M(n,\lfloor n/2 \rfloor).$$

This is true by Equation (4), and because $0 \leq r \leq n/2$ implies

$$\frac{1}{n-r+1} \leq \frac{1}{n-\lfloor n/2 \rfloor + 1}. \quad □$$

Proof of Theorem 1.2. Since

$$\frac{1}{n-\lfloor n/2 \rfloor + 1}$$

is equal to either

$$\frac{2}{n+2} \quad \text{or} \quad \frac{2}{n+3}$$

depending on whether n is even or odd, it follows that

$$\frac{1}{n-\lfloor n/2 \rfloor + 1} \leq \frac{2}{n+2}. \tag{5}$$

We can assume that $n \geq 2$, so this implies

$$M(n, \lfloor n/2 \rfloor) \leq \frac{1}{2}\binom{n}{\lfloor n/2 \rfloor}.$$

Therefore

$$\binom{\binom{n}{\lfloor n/2 \rfloor}}{i} \leq \binom{\binom{n}{\lfloor n/2 \rfloor}}{M(n, \lfloor n/2 \rfloor)}$$

when $0 \leq i \leq M(n, \lfloor n/2 \rfloor)$.

Lemma 2.4 implies that

$$sp(n) \leq (n+1)(M(n, \lfloor n/2 \rfloor) + 1)\binom{\binom{n}{\lfloor n/2 \rfloor}}{M(n, \lfloor n/2 \rfloor)}.$$

It follows from Equation (3) that

$$sp(n) \leq (n+1)(M(n, \lfloor n/2 \rfloor) + 1)\binom{\lfloor 2^n n^{-1/2} \rfloor}{M(n, \lfloor n/2 \rfloor)}. \tag{6}$$

Claim 1.

$$\binom{\lfloor 2^n n^{-1/2} \rfloor}{M(n, \lfloor n/2 \rfloor)} \leq \binom{\lfloor 2^n n^{-1/2} \rfloor}{\lceil e 2^{n+1} n^{-3/2} \rceil}.$$

Proof. By Equations (3) and (5), we see that

$$M(n, \lfloor n/2 \rfloor) \leq \frac{2}{n+2}\binom{n}{\lfloor n/2 \rfloor} \leq \frac{2}{n}\left(2^n n^{-1/2}\right)$$
$$\leq e 2^{n+1} n^{-3/2} \leq \lceil e 2^{n+1} n^{-3/2} \rceil.$$

Therefore the claim will be proved as long as we can certify that

$$\lceil e 2^{n+1} n^{-3/2} \rceil \leq (1/2)\lfloor 2^n n^{-1/2} \rfloor.$$

It is not difficult to show that this is true for sufficiently large n. □

Applying Claim 1 to Equation (6) produces the following:

$$sp(n) \leq (n+1)(M(n, \lfloor n/2 \rfloor) + 1)\binom{\lfloor 2^n n^{-1/2} \rfloor}{\lceil e 2^{n+1} n^{-3/2} \rceil}.$$

Now we apply Lemma 2.2, and deduce that

$$sp(n) \leq (n+1)(M(n, \lfloor n/2 \rfloor) + 1)\left(\frac{e\lfloor 2^n n^{-1/2} \rfloor}{\lceil e 2^{n+1} n^{-3/2} \rceil}\right)^{\lceil e 2^{n+1} n^{-3/2} \rceil}$$
$$\leq (n+1)(M(n, \lfloor n/2 \rfloor) + 1)\left(\frac{e 2^n n^{-1/2}}{e 2^{n+1} n^{-3/2}}\right)^{\lceil e 2^{n+1} n^{-3/2} \rceil}$$
$$\leq (n+1)(M(n, \lfloor n/2 \rfloor) + 1)\left(\frac{n}{2}\right)^{e 2^{n+1} n^{-3/2} + 1}.$$

By Equations (3) and (5), we see that

$$\begin{aligned} sp(n) &\le (n+1)\left(\frac{2}{n+2}\binom{n}{\lfloor n/2\rfloor}+1\right)\left(\frac{n}{2}\right)^{e2^{n+1}n^{-3/2}+1} \\ &\le (n+1)\left(\frac{2}{n+1}2^n n^{-1/2}+1\right)\left(\frac{n}{2}\right)^{e2^{n+1}n^{-3/2}+1} \\ &\le (n+1)\left(\frac{2^{n+1}}{n+1}+1\right)\left(\frac{n}{2}\right)^{e2^{n+1}n^{-3/2}+1} \\ &\le (n+1)\left(\frac{2^{n+1}}{n+1}+\frac{2^{n+1}}{n+1}\right)\left(\frac{n}{2}\right)^{e2^{n+1}n^{-3/2}+1} \\ &= 2^{(n+2)-e2^{n+1}n^{-3/2}-1}n^{e2^{n+1}n^{-3/2}+1} \end{aligned}$$

But $(n+2)-e2^{n+1}n^{-3/2}-1$ is negative for sufficiently large n, so

$$2^{(n+2)-e2^{n+1}n^{-3/2}-1} \le 1$$

and therefore

$$sp(n) \le n^{e2^{n+1}n^{-3/2}+1}.$$

Hence

$$\begin{aligned} \log sp(n) &\le (e2^{n+1}n^{-3/2}+1)\log n \\ &\le (e2^{n+1}n^{-3/2}+e2^{n+1}n^{-3/2})\log n \\ &= e2^{n+2}n^{-3/2}\log n \end{aligned}$$

and

$$\log\log sp(n) \le n-(3/2)\log n+\log\log n+\log e+2.$$

This completes the proof of Theorem 1.2. □

3. Acknowledgements

We thank David Stirzaker for his helpful comments.

References

[1] R. L. Graham, N. J. A. Sloane, Lower bounds for constant weight codes, IEEE Trans. Inform. Theory 26 (1980) 37–43.

[2] D. E. Knuth, The asymptotic number of geometries, J. Combin. Theory Ser. A 16 (1974) 398–400.

[3] D. Mayhew, M. Newman, G. Whittle, D. Welsh, On the asymptotic proportion of connected matroids, European J. Combin. 32 (2011) 882–890.

[4] M. J. Piff, An upper bound for the number of matroids, J. Combin. Theory Ser. B 14 (1973) 241–245.