THE DISTRIBUTION OF RAMSEY NUMBERS

LANE CLARK AND FRANK GAITAN

Abstract. We prove that the number of integers in the interval \([0, x]\) that are non-trivial Ramsey numbers \(r(k, n)\) \((3 \leq k \leq n)\) has order of magnitude \(\sqrt{x \ln x}\).

1. Introduction

Suppose function \(f: A \to \mathbb{N}\) where \(A \subseteq \mathbb{N}^k\). Then \(f\) has distribution function \(D_f: [0, \infty) \to \mathbb{N}\) defined by \(D_f(x) = \# f^{-1}(\mathbb{N}) \cap [0, x]\). Here \(D_f(x)\) counts each value of \(f\) at most \(x\) precisely once ignoring its multiplicity. Related is the distribution function \(M_f: [0, \infty) \to \mathbb{N}\) defined by \(M_f(x) = \# \{ a \in A : f(a) \leq x \} = \sum_{n \leq x} \# f^{-1}(n)\). Here \(M_f(x)\) counts each value of \(f\) at most \(x\) according to its multiplicity. Then \(D_f(x) \leq M_f(x)\) with equality when \(f\) is injective. Estimating \(D_f\) (or \(M_f\)) is fundamental for many counting functions \(f\). Here \(\mathbb{N}\) (respectively, \(\mathbb{P}\)) denotes the non-negative (respectively, positive) integers. A set \(S\) has cardinality \(\# S\).

The distribution of values of number-theoretic functions is a primary research area in number theory. The seminal example is the distribution of the prime numbers: Let \(p: \mathbb{P} \to \mathbb{N}\) where \(p(n)\) is the \(n^{th}\) prime. The function \(\pi(x) = D_p(x) = M_p(x)\) is the number of primes at most \(x\). Chebyshev [8] determined the order of magnitude of \(\pi(x) = \Theta(x/\ln x)\). (See Mathematical Reviews Mathematics Subject Classification 11N for many further examples.)

Combinatoric-theoretic numbers may be viewed as functions \(f: A \to \mathbb{N}\), where \(A \subseteq \mathbb{N}^k\), and, hence, their distribution function \(D_f\) (or \(M_f\)) investigated. Very few have however: The function \(b: \mathbb{N}^2 \to \mathbb{N}\) where \(b(n, k) = \binom{n}{k}\) has distribution function \(D_b(x) = \sqrt{2x} + o(\sqrt{x})\) (cf. [11; pp. 76-77]). Clark and Dabab [10] determined the distribution function \(D_{\mathfrak{R}}\) of the function \(\mathfrak{R}: \mathbb{P}^2 \to \mathbb{N}\) where the \(\mathfrak{R}(n, k)\) are the Narayana numbers. Erdős and Niven [14] proved that the number of distinct multinomial coefficients at most \(x > 0\) is \((1+\sqrt{2})\sqrt{x} + o(\sqrt{x})\). Andrews, Knopfmacher, and Zimmermann [2] also investigated this topic. Clark [9] examined the multiplicity problem and proved best possible bounds for \(\# f^{-1}(n)\) for a large class of functions \(f\) including \(b, \mathfrak{R}\).

The Ramsey numbers, \(r(k, l)\) where \(k, l \in \mathbb{P}\) (cf. West [22]), are among the most important of combinatoric-theoretic numbers. Despite substantial efforts spanning decades our knowledge of the Ramsey numbers is quite limited. Only nine non-trivial values of \(r(k, l)\), with \(3 \leq k, l \leq 5\), and non-trivial bounds for certain other \(r(k, l)\), with \(3 \leq k, l \leq 19\), are known (cf. Radziszowski [19]). General bounds for all \(r(k, l)\) are known. At present the best lower bounds are due to Bohman [4] and Bohman and Keevash [5] and the best upper bounds are due to Conlon [12]. They
have very different orders of magnitude. The order of magnitude of only one infinite family of Ramsey numbers is known: Kim [18] proved that $r(3, n) = \Theta(n^2 / \log n)$ by improving the lower bound $r(3, n) \geq c_1 n^2 / \log^2 n$ of Erdős [13] to match the upper bound $r(3, n) \leq c_2 n^2 / \log n$ of Ajtai, Komlós, and Szemerédi [1]. His proof used many modern tools from probabilistic combinatorics. See also [16].

The computational complexity of determining $r(k, l)$ is not known although clearly hard (cf. Haanpää [17] and Schweitzer [21]). More is known about certain generalizations. Burr [6] proved that determining whether the graph Ramsey number $r(G, H) \leq m$ is NP-hard. He [7] proved that determining whether the arrow relation $F \rightarrow (G, H)$ holds is coNP-hard. Schaefer [20] proved that determining whether $F \rightarrow (G, H)$ holds is $\Sigma_2$–complete. A quantum algorithm in complexity class QMA for computing $r(k, l)$ was given by the authors [15] who proved that its solution can be found using adiabatic evolution. See also Bian et al. [3].

Trivially $r(1, n) = r(n, 1) = 1$ ($n \geq 1$) and $r(2, n) = r(n, 2) = n$ ($n \geq 2$). Hence every positive integer is a trivial Ramsey number. We consider the non-trivial Ramsey numbers $r(k, n)$ where $3 \leq k \leq n$ since all $r(k, n) = r(n, k)$. We note that $k_1 \leq k_2$ and $n_1 \leq n_2$ imply $r(k_1, n_1) \leq r(k_2, n_2)$. The function $r : A \rightarrow \mathbb{N}$, where $A = \{(k, n) \in \mathbb{Z}^2 : 3 \leq k \leq n\}$, defined by $r(k, n) = r(k, n)$ gives the non-trivial Ramsey numbers. In this note we prove that its distribution function $D_k(x) = \Theta(\sqrt{x \log x})$. It is rather remarkable that a property of Ramsey numbers as fundamental as their distribution has not been studied prior to this work. Our result for Ramsey numbers is the direct analog of Chebyshev’s result for prime numbers namely $\pi(x) = \Theta(x / \log x)$. We deduce that the density of non-trivial Ramsey numbers is roughly the square root of the density of prime numbers.

2. Distribution of the Ramsey Numbers

Bohman [4] proved there exists a constant $c > 0$ and an integer $N_1 \geq 4$, such that $r(4, n) \geq cn^{3/2} / \ln^2 n$ ($n \geq N_1$). It follows that there exists a constant $1 \geq d > 0$ such that $r(4, n) \geq dn^{3/4}$ ($n \geq 4$). Further $r(k, n) \geq r(4, n)$ for all $n \geq k \geq 4$. Hence

$$r(k, n) \geq dn^{3/4} \quad (n \geq k \geq 4)$$

Define $R_k := \{r(k, n) : k \leq n\}$ ($k \geq 3$) and $R := \{r(k, n) : 3 \leq k \leq n\} = \bigcup_{k=3}^{\infty} R_k$. Hence

$$R_3 \subseteq R = R_3 \cup \bigcup_{k=4}^{\infty} R_k.$$  

For $x \in [0, \infty)$, define $R_k(x) = R_k \cap [0, x], r_k(x) = \#R_k(x)$ and $R(x) = R \cap [0, x], r(x) = \#R(x)$. Then $R(x)$ is the set of non-trivial Ramsey numbers and $r(x)$ is the number of non-trivial Ramsey numbers, ignoring their multiplicity, in the interval $[0, x]$. Notice that $r(x) = D_k(x)$ from the introduction. Then (2.2) gives

$$R_3(x) \subseteq R(x) = R_3(x) \cup \bigcup_{k=4}^{\infty} R_k(x),$$

hence,

$$r_3(x) \leq r(x) \leq r_3(x) + \sum_{k=4}^{\infty} r_k(x).$$
Lemma 2.1. Suppose $k \geq 4$. If $x \geq d^{81/16} k^{9/4}$, then $r_k(x) \leq d^{-9/4} x^{4/9}$.

Proof. If $x \geq d^{81/16} k^{9/4}$, then $d^{-9/4} x^{4/9} \geq k$. Suppose integer $l > d^{-9/4} x^{4/9}(\geq k)$. Inequality (2.1) gives $r(k, l) \geq d^{3/4} x$. Hence $r(k, l) \notin R_k(x)$. This implies the result. □

Lemma 2.2. Suppose $x \geq 8$. If $k \geq \lceil 2 \log_2 x \rceil (\geq 3)$, then $r_k(x) = 0$.

Proof. Suppose $k \geq \lceil 2 \log_2 x \rceil$. Erdős’ classic result gives $r(k, n) \geq r(k, k) > 2^{k/2} \geq 2^{\log_2 x} = x$. This implies the result. □

The result of Kim [18] implies there exist positive constants $c_1 < c_2$ such that $c_1 n^2 / \ln n \leq r(3, n) \leq c_2 n^2 / \ln n$ ($n \geq 3$). It follows that there exists $x_1 \geq 6$ such that

$$c_3 \sqrt{x \ln x} \leq r_3(x) \leq c_4 \sqrt{x \ln x} \quad (x \geq x_1)$$

where, say, $c_3 = (4e_2)^{-1/2} > 0$ and $c_4 = c_1^{-1/2} > 0$ which is adequate for our needs.

Theorem 2.3. We have $r(x) = D_t(x) = \Theta(\sqrt{x \ln x})$.

Proof. There exists $x_2 \geq 8$ such that $x \geq d^{9/4} d^{81/16} \log_2^{9/4} x$ for all $x \geq x_2$. Fix $x \geq \max\{x_1, x_2\}$ from $x_1$ from (2.5). Then (2.4) and (2.5) give

$$c_3 \sqrt{x \ln x} \leq r_3(x) \leq r(x) \leq r_3(x) + \sum_{k=4}^{\infty} r_k(x) \leq c_4 \sqrt{x \ln x} + \sum_{k=4}^{\infty} r_k(x).$$

Each $r_k(x) = 0$ for $k \geq \lceil 2 \log_2 x \rceil$ by Lemma 2.2. Hence

$$\sum_{k=4}^{\infty} r_k(x) = \sum_{k=4}^{\lceil 2 \log_2 x \rceil} r_k(x).$$

If $x \geq d^{9/4} d^{81/16} \log_2^{9/4} x$, then $x \geq d^{81/16} k^{9/4}$ for $k = 4, \ldots, \lceil 2 \log_2 x \rceil$. Lemma 2.1 gives

$$\sum_{k=4}^{\lceil 2 \log_2 x \rceil} r_k(x) \leq \sum_{k=4}^{\lceil 2 \log_2 x \rceil} d^{-9/4} x^{4/9} \leq 2 d^{-9/4} (\log_2 x) x^{4/9}. $$

For $x \geq \max\{x_1, x_2\}$, (2.6)–(2.8) give

$$c_3 \sqrt{x \ln x} \leq r(x) \leq c_4 \sqrt{x \ln x} + 2d^{-9/4} (\log_2 x) x^{4/9}$$

which implies our result. □

3. Conclusion

We have proved that the number of non-trivial Ramsey numbers at most $x$ is $\Theta(\sqrt{x \ln x})$. It is interesting that this fundamental fact can be determined at present. Our result for the Ramsey numbers is the direct analog of Chebyshev’s result that the number of prime numbers at most $x$ is $\Theta(x / \ln x)$. Roughly then the density of non-trivial Ramsey numbers is the square root of the density of prime numbers.

As noted in the introduction, very little work has been done to determine the distribution of other significant families of combinatoric-theoretic numbers. We think this is an interesting, and important, direction for future work.
References

[1] M. Ajtai, J. Komlós and E. Szemerédi, A Note on Ramsey Numbers, *Journal of Combinatorial Theory, Series A* **29** (1980), no. 3, 354–360.

[2] G.E. Andrews, A. Knopfmacher and B. Zimmermann, On the Number of Distinct Multinomial Coefficients, *Journal of Number Theory* **118** (2006), no.1, 15–30.

[3] Z. Bian, F. Chudak, W. Macready, L. Clark and F. Gaitan, Experimental Determination of Ramsey Numbers with Quantum Annealing, arXiv:1201.1842v2 (10 January 2012), 6 pages.

[4] T. Bohman, The Triangle–Free Process, *Advances in Mathematics* **221** (2009), no. 5, 1653–1677.

[5] T. Bohman and P. Keevash, The Early Evolution of the $H$–Free Process, *Inventiones Mathematicae* **181** (2010), no. 2, 291–336.

[6] S.A. Burr, Determining Generalized Ramsey Numbers is NP–hard, *Ars Combinatoria* **17** (1984), 21–25.

[7] S.A. Burr, On the Computational Complexity of Ramsey–Type Problems, in *Mathematics of Ramsey Theory* (J. Nešetřil and V. Rödl, Eds.), Springer, Berlin (1990), pp. 46–52.

[8] P. Chebyshev, Memoire sur les Nombres Premiers, *Journal de Mathematiques Pures et Applications* **17** (1852), 366–390.

[9] L. Clark, Multiplicities of Integer Arrays, *Integers* **10** (2010), A 14, 187–199.

[10] L. Clark and Y. Dabab, The Distribution of Certain Combinatorial Arrays, *Congressus Numerantium* **210** (2011), 33–39.

[11] L. Comtet, *Advanced Combinatorics*, D. Reidel, Boston, MA (1974).

[12] D. Coulon, A New Upper Bound for Diagonal Ramsey Numbers, *Annals of Mathematics* **170** (2009), no. 2, 941–960.

[13] P. Erdős, Graph Theory and Probability II, *Canadian Journal of Mathematics* **13** (1961), 346–352.

[14] P. Erdős and I. Niven, The Number of Multinomial Coefficients, *American Mathematical Monthly* **61** (1954), no.1, 37–39.

[15] F. Gaitan and L. Clark, Ramsey Numbers and Adiabatic Quantum Computing, *Physical Review Letters* **108** (2012), no. 1, 010501-1 to 010501-4.

[16] R. Graham, B. Rothschild and J. Spencer, *Ramsey Theory. Second Edition*, John Wiley & Sons, New York, NY (1990).

[17] H. Haanpää, Computational Methods for Ramsey Numbers, Helsinki University of Technology, Laboratory for Theoretical Computer Science, Research Report 65 (November 2000), 50 pages.

[18] J.H. Kim, The Ramsey Number $R(3, t)$ has Order of Magnitude $t^2 / \log t$, *Random Structures and Algorithms* **7** (1995), 173–207.

[19] S. Radziszowski, Small Ramsey Numbers, *Electronic Journal of Combinatorics* **1** (1994), Dynamic Survey DS1.

[20] M. Schaefer, Graph Ramsey Theory and the Polynomial Hierarchy, *Journal of Computer and System Sciences* **62** (2001), no. 2, 290–322.

[21] P. Schweitzer, Problems of Unknown Complexity–Graph Isomorphism and Ramsey Theoretic Numbers, Ph.D. Dissertation, Universität des Saarlandes, Saarbrücken, Germany (24 July 2009), 144 pages.

[22] D. West, *Introduction to Graph Theory*, Prentice-Hall, Upper Saddle River, NJ (1996).

LANCE CLARK: DEPARTMENT OF MATHEMATICS, SOUTHERN ILLINOIS UNIVERSITY, CARBONDALE, IL 62901-4408
E-mail address: lclark@math.siu.edu

FRANK GAITAN: LABORATORY FOR PHYSICAL SCIENCES, 8050 GREENMEAD DRIVE, COLLEGE PARK, MD 20740
E-mail address: fgaitan@lps.umd.edu