Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Combinatorics 26 (2005) 607–616
European Journal of Combinatorics
www.elsevier.com/locate/ejc

Integer sets with prescribed pairwise differences being distinct

Béla Bollobás[a,b], Oleg Pikhurko[c]

[a] *University of Memphis, Memphis, TN 38152-3240, USA*
[b] *Trinity College, Cambridge CB2 1TQ, UK*
[c] *Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA*

Received 20 January 2004; received in revised form 25 April 2004; accepted 26 April 2004
Available online 10 June 2004

Abstract

We label the vertices of a given graph G with positive integers so that the pairwise differences over its edges are all distinct. Let $\mathcal{D}(G)$ be the smallest value that the largest label can have.

For example, for the complete graph K_n, the labels must form a Sidon set. Hence, $\mathcal{D}(K_n) = (1 + o(1))n^2$. Rather surprisingly, we demonstrate that there are graphs with only $n^{\frac{3}{2}+o(1)}$ edges achieving this bound.

More generally, we study the maximum value of $\mathcal{D}(G)$ that a graph G of the given order n and size m can have. We obtain bounds which are sharp up to a logarithmic multiplicative factor. The analogous problem for pairwise sums is considered as well. Our results, in particular, disprove a conjecture of Wood.

1. Introduction

Let G be a graph. A *difference-magic labelling* of G is an injective mapping $l : V(G) \to \mathbb{N}$ (into positive integers) such that the $e(G)$ numbers

$$|l(x) - l(y)|, \qquad \{x, y\} \in E(G),$$

are pairwise distinct.

It is trivial to see that every graph admits a difference-magic labelling, so a natural question to ask is how economical it can be. More precisely, we should like to determine

E-mail address: bollobas@msci.memphis.edu (B. Bollobás).
URL: http://www.math.cmu.edu/~pikhurko/ (O. Pikhurko).

0195-6698/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejc.2004.04.008

the *difference-magic number* $\mathcal{D}(G)$ which is the smallest k such that a difference-magic labelling of G into $[k] := \{1, \dots, k\}$ exists.

For example, it is easy to see that if G is the complete graph of order n, then $\mathcal{D}(G)$ is precisely s_n, the smallest s such that $[s]$ contains a *Sidon* subset of size n. (A set $A \subset \mathbb{Z}$ is *Sidon* if all sums $a + b$ with $a, b \in A$ and $a \le b$ are distinct.) The latter problem is well studied; the results of Singer [13] and Erdős and Turán [8] (see e.g. Halberstam and Roth [10, Chapter II]) imply that $s_n = (1 + o(1))n^2$. Erdős [5] offered \$500 for proving or disproving that $s_n = n^2 + O(n)$.

Here we deal with

$$\mathcal{D}(n, m) := \max\{\mathcal{D}(G) : v(G) = n,\ e(G) = m\},$$

the maximum value of $\mathcal{D}(G)$ for a graph G of order n and size m.

It turns out that

$$\mathcal{D}(n, m) = (1 + o(1))n^2 \qquad \text{if } m \Big/ \sqrt{n^3 \ln n} \to \infty\ . \tag{1}$$

In fact, a random graph of order n with the appropriate edge probability demonstrates (1). We find it surprising that graphs so sparse (with only $n^{\frac{3}{2}+o(1)}$ edges) have the $\mathcal{D}$-function asymptotically the same as that of the complete graph.

What happens for smaller m? The obvious choice is to consider random graphs of suitable density. This, indeed, leads to interesting results. Let $G \in \mathcal{G}(n, p)$, that is, G is a random graph on n vertices where each edge is included in G independently of others and with probability p. If $p = O((\ln n/n)^{1/2})$ and $p > n^{-1+\varepsilon}$, then

$$\mathcal{D}(G) = \Theta(n^3 p^2 / \ln n). \tag{2}$$

A lower bound on $\mathcal{D}(n, m)$ can be obtained by adding isolated vertices to a random graph and figuring out the best parameters to choose. On the other hand, the simple labelling procedure described in Section 4 gives an upper bound that is within an $O((\ln n)^{2/3})$-factor of the lower bound. Roughly, we obtain

$$\mathcal{D}(n, m) = m^{4/3+O\left(\frac{\ln\ln m}{\ln m}\right)}, \qquad \text{if } m = O\left(\sqrt{n^3 \ln n}\right), \tag{3}$$

unless $m = o(n^{3/4})$ when $\mathcal{D}(n, m) = (1 + o(1))n$. All details (with more precise expressions for the error terms) can be found in the corresponding sections.

Let us define a *sum-magic labelling* of a graph G as an injection $l : V(G) \to \mathbb{N}$ such that all $e(G)$ sums $l(x) + l(y)$, $\{x, y\} \in E(G)$, are pairwise distinct. We ask for the *sum-magic number* $\mathcal{S}(G)$, the smallest value that the largest label can have, and for

$$\mathcal{S}(n, m) := \max\{\mathcal{S}(G) : v(G) = n, e(G) = m\}.$$

It is not surprising that most of the methods on the $\mathcal{D}$-function transfer to $\mathcal{S}$, giving similar bounds. (In particular, (3) holds for $\mathcal{S}(n, m)$ as well.) However, there is one peculiar distinction. While Corollary 2 states that $\mathcal{S}(K_n) = (1 + o(1))n^2$, Theorem 3 shows that there is a constant $c > 0$ such that $\mathcal{S}(n, m) < (1 - c)n^2$ whenever $m \le cn^2$. Random graphs are far worse in hitting $(1 + o(1))n^2$: this happens only when the random graph is almost complete.

Wood [15] defines an *edge-magic injection* with the *magic sum* s as an injection $l : V(G) \cup E(G) \to \mathbb{N}$ such that for any edge $\{a, b\} \in E(G)$ we have $s = l(a) + l(b) + l(\{a, b\})$. Let $\mathcal{E}(G)$ be the smallest possible value of s. Wood [15, Section 7] conjectured that there is an absolute constant C such that for any graph G we have $\mathcal{E}(G) \le C(v(G) + e(G))$. Clearly, the vertex labels of any edge-magic injection form a sum-magic labelling, so $\mathcal{E}(G) \ge \mathcal{S}(G)$ and random graphs disprove Wood's conjecture.

One can also ask what is the value of, for example,

$$\mathcal{S}_{\min}(n, m) := \min\{\mathcal{S}(G) : v(G) = n, e(G) = m\}.$$

This is the inverse problem to maximising the number of distinct pairwise sums that a set $A \subset [s]$ of given size n can have. This question is investigated by Pikhurko [12].

2. Some preliminary results

Let $A \in \binom{[m]}{n}$, meaning that A is an n-subset of $[m]$.

Recall that A is called a *Sidon set* if the sums $a + b$, $a, b \in A$ with $a \ge b$, are pairwise distinct, which is equivalent to all differences $a - b$, $a, b \in A$ with $a > b$, being pairwise distinct. Erdős and Turán [8] proved that this property implies that $m \ge (1 + o(1))n^2$. The following results show that, in a sense, it is the condition on differences (rather than that on sums) which pushes $\max A$ upwards.

For $i \in [m - 1]$ let g_i be the number of representations $i = a - b$ with $a, b \in A$. Thus, if $n^2 \ge (1+\varepsilon)m$ (and n is large), then there must be i with $g_i \ge 2$. Although the following theorem strengthens this claim considerably, its proof goes via an easy modification of the original argument of Erdős and Turán [8]. A similar result (in a more precise form) was independently obtained by Ferrara, Kohayakawa and Rödl [9, Lemma 12].

Let $f_+ = f$ if $f > 0$ and $f_+ = 0$ otherwise.

Theorem 1. *Let $\varepsilon > 0$ be fixed and $n \to \infty$. Let $A \in \binom{[m]}{n}$. If $n^2 \ge (1 + \varepsilon)m$, then $g = \Omega(n^2)$, where $g := \sum_{i=1}^{m-1}(g_i - 1)_+$.*

Proof. Let $t := cn^2$, where $c = c(\varepsilon) > 0$ is a small constant. Assume $t \in \mathbb{N}$. Define

$$A_i := A \cap [i, i + t - 1] \qquad \text{and} \qquad a_i := |A_i|, \qquad i \in [2 - t, m],$$

where $[i, j] := \{i, i + 1, \dots, j\}$.

Let $\mathcal{X}$ consist of all quadruples (a, b, i, x) such that $x = a-b > 0$ and $a, b \in A_i$. Using the identity $\sum_{i=2-t}^{m} a_i = nt$ and the quadratic-arithmetic mean inequality, we obtain

$$|\mathcal{X}| = \sum_{i=2-t}^{m} \binom{a_i}{2} = \frac{1}{2}\sum_{i=2-t}^{m} a_i^2 - \frac{nt}{2} \ge \frac{(nt)^2}{2(m + t - 1)} - \frac{nt}{2}. \tag{4}$$

Each $x \in [t - 1]$ is included in $g_x \cdot (t - x) \le (t - x) + t(g_x - 1)_+$ quadruples. Hence,

$$|\mathcal{X}| \le \sum_{x=1}^{t-1}(t - x + t(g_x - 1)_+) = \frac{t(t - 1)}{2} + gt. \tag{5}$$

By choosing c sufficiently small, we can ensure that the right-hand side of (4) is, for example, at least $(1+\frac{\varepsilon}{2})\frac{t^2}{2}$, which together with (5) implies the theorem. $\square$

We will need Theorem 1 in Section 3. Here we demonstrate another application.

Corollary 2. $\mathcal{S}(K_n) = (1+o(1))n^2$.

Proof. Let A be the label set of a sum-magic labelling. Note that A need not be Sidon as it may well happen that $a-c=c-b$ for $a,b,c \in A$. However, if $a-b=c-d$ with $a \notin \{b,c\}$, then either $a=d$ or $b=c$. It follows that $g_x \le 2$ for any $x>0$ and, if $g_x = 2$, then there are $a,b,c \in A$ with $a-b=c-a=x$. If $a-b'=c'-a \ne 0$, then we have $b'+c' = 2a = b+c$ and thus $\{b',c'\} = \{b,c\}$. Hence, no a can appear for more than one x in the above manner. We conclude that $g \le |A|$, implying the claim by Theorem 1. $\square$

The natural analogue of Theorem 1 in terms of the number of solutions to $x = a+b$, $a,b \in A$, is not true, as the following construction of Erdős and Freud [6] demonstrates. Let $\mathcal{S} \in \binom{[t]}{s}$ be a Sidon set with $t = (1+o(1))s^2$. (Such sets were constructed by Singer [13].) Let $X = \mathcal{S} \cup \mathcal{S}'$, where

$$S' := 3t+1-S := \{3t+1-a : a \in S\} \subset [2t+1, 3t].$$

Clearly, $S+S \subset [2,2t]$, $S+S' \subset [2t+2,4t]$ and $S'+S' \subset [4t+2,6t]$ are disjoint. Hence, all sums $a+b$, $a,b \in X$ with $a \le b$, are pairwise distinct except those s sums which are equal to $3t+1$. If the complement of an order-n graph G has a matching covering all but $r = o(n)$ vertices, then considering the first n elements of the set X constructed above for $s := \frac{n+r}{2}$, we conclude

$$\mathcal{S}(G) \le (3/4 + o(1))n^2. \tag{6}$$

By modifying the above construction, we can show one of the results claimed in the Introduction.

Theorem 3. *There is a constant $c > 0$ such that if $m \le cn^2$, then*

$$\mathcal{S}(n,m) \le (1-c)n^2. \tag{7}$$

Proof. Let $\alpha = 0.9$, for example. In the above construction of $X = S \cup S'$ let $Y \subset X$ consist of the first $n := \lfloor (1+\alpha)s \rfloor$ elements of X. As it was shown by Erdős and Freud [6, Lemma 1], any asymptotically maximum Sidon subset of $[t]$ is almost uniformly distributed. This implies that $\max Y = (2+\alpha+o(1))t$.

Now, all sums in $Y+Y$ are distinct except those sums which equal $3t+1$. The number of these exceptional sums is $\lfloor \alpha s \rfloor = (\frac{\alpha}{1+\alpha} + o(1))n$. So, if the complement of an order-n graph G has a matching of size bigger than $0.48n > (\frac{\alpha}{1+\alpha} + o(1))n$, then

$$\mathcal{S}(G) \le (2+\alpha+o(1))t = \frac{2+\alpha+o(1)}{(1+\alpha)^2} n^2 < 0.9n^2.$$

It follows from the Tutte 1-factor theorem [14] that a matching of size $0.48n$ in the complement $\overline{G}$ is guaranteed if $e(G) \le \delta n^2$ for some constant $\delta > 0$. Now, the theorem follows. $\square$

Remark. Random graphs do not provide good examples if we want to achieve $\mathcal{S}(G) = (1+o(1))n^2$: this happens only when $1-p = O(\frac{\ln n}{n})$. Indeed, Erdős and Rényi [7] (cf. Bollobás and Thomason [3]) showed that if $p \le 1-(1+\varepsilon)\frac{\ln n}{2n}$ then with high probability the complement of $G \in \mathcal{G}(n, p)$ has an almost perfect matching; so then (6) holds.

3. Random graphs

Theorem 4. *Fix any $\delta > 0$. Let $G \in \mathcal{G}(n, p)$, where $n \to \infty$ and $p \in (0, 1)$ is a function of n such that $np/\ln n \to \infty$. Let $\lambda := p\sqrt{n/\ln n}$. Then almost surely $\mathcal{D}(G) \ge d$ and $\mathcal{S}(G) \ge s$, where*

$$d := \begin{cases} (1-\delta)n^2 \\ \left(\dfrac{\lambda^2}{16+2\lambda^2} - \delta\right) n^2, \\ \left(\dfrac{1}{16} - \delta\right)\dfrac{n^3p^2}{\ln n}, \end{cases} \qquad s := \begin{cases} \left(\dfrac{1}{4} + \dfrac{1}{(\pi+2)^2} - \delta\right) n^2, & \textit{if } \lambda \to \infty, \\ \left(\dfrac{\lambda^2}{32+4\lambda^2} - \delta\right) n^2, & \textit{if } \lambda = \Theta(1), \\ \left(\dfrac{1}{32} - \delta\right)\dfrac{n^3p^2}{\ln n}, & \textit{if } \lambda = o(1). \end{cases}$$

Proof. We prove the lower bound on $\mathcal{D}(G)$. Let $[n]$ be the vertex set. Let $\varepsilon > 0$ be a small constant depending on δ. Assume $d \in \mathbb{N}$.

Fix an injective mapping $l : [n] \to [d]$. Now, let us choose $G \in \mathcal{G}(n, p)$. We want to bound the probability p' that all differences $l(i) - l(j)$, with $\{i, j\} \in \binom{[n]}{2}$ being an edge of $\mathcal{G}(n, p)$ and $l(i) > l(j)$, are pairwise distinct. If u is an upper bound on p' for any l, then the probability that $G \in \mathcal{G}(n, p)$ satisfies $\mathcal{D}(G) \le d$ is at most $n!\binom{d}{n}u < d^n u$. Hence, if we can show that $p' = o(d^{-n})$, then almost surely $\mathcal{D}(G) > d$.

For $k \in [d]$, let g_k be the number of representations $k = l(i) - l(j)$ with $i, j \in [n]$. Let $t := \binom{n}{2} = \sum_{k=1}^{d} g_k$. Clearly,

$$p' = \prod_{k=1}^{d} p_k, \tag{8}$$

where $p_k = (1-p)^{g_k} + g_k p(1-p)^{g_k-1}$ is the probability of selecting at most one edge with difference k. (Note that the formula is also valid for $g_k = 0$ and $g_k = 1$, when $p_k = 1$.) It is routine to see that

$$p' = \prod_{k=1}^{d} p_k \le ((1-p)^{t/d} + (pt/d)(1-p)^{(t/d)-1})^d. \tag{9}$$

Case 1. $p = o(\sqrt{\ln n/n})$, that is, $\lambda = o(1)$.

We have $t/d \to \infty$ and $pt/d = o(1)$. Using the inequality $e^{-x}(1+x) \le 1 - (\frac{1}{2} - \varepsilon)x^2$ valid if $x > 0$ is small, we deduce from (9) the required bound on p':

$$\begin{aligned} p' &\le ((1-p)^{t/d}(1 + pt/d + 2p^2t/d))^d \le (e^{-pt/d}(1 + pt/d + 2p^2t/d))^d \\ &\le (1 - (1/2-\varepsilon)(pt/d)^2 + 2p^2t/d)^d \le e^{-(1/2-2\varepsilon)(pt)^2/d} = o(e^{-n\ln d}). \end{aligned}$$

Case 2. $p = \Theta(\sqrt{\ln n/n})$, that is, $\lambda = \Theta(1)$.

We have $t/d = O(1)$ so we can simply take the Taylor expansion of (9) to obtain the required bound:

$$p' \le \left(1 + \left(\frac{t}{2d} - \frac{t^2}{2d^2}\right) p^2 + O(p^3)\right)^d \le \mathrm{e}^{\frac{1}{2}(t - \frac{t^2}{d} + \varepsilon) p^2} = o(\mathrm{e}^{-n \ln d}).$$

Case 3. $p\sqrt{n/\ln n} \to \infty$, that is, $\lambda \to \infty$.

By Theorem 1 we know that $g := \sum_{k=1}^{d} (g_k - 1)_+ = \Omega(n^2)$. It is routine to see that if $g_i \ge g_j + 2$, then the right-hand side of (8) increases if we replace g_i and g_j by $g_i - 1$ and $g_j + 1$ respectively. Hence,

$$p' \le ((1-p)^2 + 2p(1-p))^g = (1-p^2)^g = o(d^{-n}),$$

as required.

Let us turn to the sum-magic number. Fix an injection $l : [n] \to [s]$. For $k \in [2s]$ define g_k as the number of representations $k = l(i) + l(j)$ with $1 \le i < j \le n$. Let $t := \binom{n}{2} = \sum_{k=1}^{2s} g_k$. The remainder of the proof goes via the obvious modification of the argument for $\mathcal{D}(G)$ except that for $\lambda \to \infty$ we use the result by Pikhurko [12, Theorem 2] which implies that $\sum_{k=1}^{2s} (g_k - 1)_+ = \Omega(n^2)$. (If we are content with $s = (\frac{1}{4} - \delta)n^2$, then $\sum_{k=1}^{2s} (g_k - 1)_+ = \Omega(n^2)$ follows by trivial counting.) The reader should have little difficulty in filling in all missing details. □

Remark. There is a jump in the lower bounds when we change from the case $\lambda = \Theta(1)$ to $\lambda \to \infty$. It should be possible to 'smoothen' this by improving our bounds for large but bounded λ. However, the calculations seem to be rather unpleasant, so we do not go into the details.

Remark. As it was mentioned in the introduction, Theorem 4 disproves the conjecture of Wood in view of the inequality $\mathcal{E}(G) \ge \mathcal{S}(G)$. Indeed, if we take $G \in \mathcal{G}(n, n^{-1/2})$ for example, then almost surely $e(G) = (\frac{1}{2} + o(1))n^{3/2}$ while $\mathcal{D}(G) = \Omega(n^2/\ln n)$. With a bit of extra work it is possible to show that under the assumptions of Theorem 4 we have almost surely $\mathcal{E}(G) \ge 2s$. To do this, prove that, almost surely, any sum-magic labelling of G has $\Omega(n)$ labels which are greater than s and there is an edge connecting two such labels. We leave the details to the interested reader.

Now let us turn to upper bounds.

Theorem 5. *Let $\delta > 0$ be fixed. Let $G \in \mathcal{G}(n, p)$, where $n \to \infty$ and $p \in (0, 1)$ is a function of n such that $\frac{\ln(np)}{\ln \ln n} \to \infty$ and $p = O((\ln n/n)^{1/2})$. Then almost surely $\mathcal{D}(G) \le 2m$ and $\mathcal{S}(G) \le m$, where*

$$m := (1 + \delta)\frac{n^3 p^2}{\ln(np)}. \tag{10}$$

Proof. Let us estimate $\mathcal{S}(G)$. (The case of $\mathcal{D}(G)$ is dealt with almost identically.)

We can assume that δ is sufficiently small and $m \in \mathbb{N}$. Let n be large and $\varepsilon > 0$ be a small constant depending on δ. Let $V(G) = [n]$ be the vertex set. Chernoff's bound [4] implies that almost surely we have

$$\Big|\big|\Gamma(i+1)\cap[i]\big| - ip\Big| \le \varepsilon np, \qquad \text{for all } i\in[0,n-1], \tag{11}$$

where $\Gamma(i+1)$ is the set of neighbours of $i+1\in V(G)$.

Consider the conditional distribution of G given (11). We have gained the very useful control over the edges while some important properties of $G\in\mathcal{G}(n,p)$ are preserved. (That is, almost sure events stay so; the random set $\Gamma(i+1)\cap[i]$ is independent from $G[[i]]$, etc.)

We choose vertex labels one by one, doing the label arithmetic in $M=\mathbb{Z}/m\mathbb{Z}$ (that is, modulo m). Our labelling $l: V(G)\to[m]$ will have the property that the sums $l(x)+l(y)$, $\{x,y\}\in E(G)$, will be pairwise distinct modulo m.

Suppose that we have already chosen labels for the vertices in $I:=[i]$.

Let

$$K := \{l(x)+l(y) : \{x,y\}\in E(G), x,y\in I\}\subset M,$$

and $k:=|K|$. By (11),

$$k\le\left(\frac{1}{2}+\varepsilon\right) inp. \tag{12}$$

Clearly, we can find a suitable label for $i+1$ if

$$M\backslash l(I) \not\subset \bigcup_{x\in I\cap\Gamma(i+1)} (K-l(x)), \tag{13}$$

that is, if the translates $K-l(x)$, $x\in I\cap\Gamma(i+1)$, do not cover $M\backslash l(I)$.

This is obviously the case if

$$|M\backslash(\cup_{x\in I}(K-l(x))|\ge n,$$

so let us assume otherwise. Then we have $m-ik\le n$, which implies by (12) that

$$i\ge n\sqrt{2p/\ln(np)}. \tag{14}$$

Now, we have to overcome the difficulty that i is large enough to potentially refute (13). In outline, we fix the labelling l of I and then choose the random set $I\cap\Gamma(i+1)$. The labels $l(x)$, $x\in I\cap\Gamma(i+1)$, are random variables. If the translates $K-l(x)$ cover the whole of $M\backslash l(I)$, then for every $z\in M\backslash l(I)$ at least one element $l(x)\in K-z$ is chosen. We prove that this is unlikely.

Let S consist of those elements from $M\backslash l(I)$ which are covered by at most

$$t:=\lfloor(1+\varepsilon)kn/m\rfloor$$

of the translates $K-l(x)$, $x\in I$. Clearly,

$$|M\backslash(S\cup l(I))|\times(1+\varepsilon)kn/m\le kn,$$

so

$$s:=|S|\ge\frac{\varepsilon m}{1+\varepsilon}-n\ge\varepsilon m/2.$$

Let $\gamma' = \lfloor ip + \varepsilon np \rfloor$ and $\gamma = \lceil ip + 2\varepsilon np \rceil$. Let us choose $y_1, \ldots, y_\gamma \in I$, one by one, independently and uniformly distributed. Of course, some of these might coincide. Let $Y := \{y_1, \ldots, y_\gamma\}$, ignoring multiple occurrences of the same vertex. The probability that $|Y| \le \gamma'$ is at most

$$\binom{i}{\gamma'}\left(\frac{\gamma'}{i}\right)^{\gamma} \le \left(\frac{\mathrm{e}i}{ip+\varepsilon np}\right)^{ip+\varepsilon np}\left(\frac{ip+\varepsilon np}{i}\right)^{ip+2\varepsilon np}$$
$$= \mathrm{e}^{ip+\varepsilon np}\left(p+\frac{\varepsilon np}{i}\right)^{\varepsilon np}.$$

The last expression, as a function of a real-valued argument $i \ge 0$, is first decreasing and then increasing in i so it is maximised if either $i = n$ or i achieves the lower bound (14). In either case, the result can be bounded by $o(n^{-1})$. Thus, the set Y has at least γ' elements with probability $1 - o(n^{-1})$.

Let the random variable U count the number of $x \in S$ which belong to none of $K - l(y)$, $y \in Y$.

We consider the martingale $(U_0, \ldots, U_\gamma)$, where U_j is the expected value of U after having exposed the first j vertices $y_1, \ldots, y_j$. Clearly, each new vertex changes U by at most k.

It is easy to estimate U_0, the expectation of U:

$$E(U) \ge |S|\frac{\binom{i-t}{\gamma}}{\binom{i}{\gamma}} \ge s\left(\frac{i-t-\gamma+1}{i-\gamma+1}\right)^{\gamma} \ge s\mathrm{e}^{-(1+\varepsilon)\gamma t/i} \ge s(np)^{-\frac{1}{2}+\frac{\delta}{4}}.$$

(Note that $t = o(i)$ by the definition of t and $\gamma = o(i)$ by (14).)

By applying the Hoeffding–Azuma inequality [2, 11] (see e.g. Alon and Spencer [1, Theorem 7.2.1]) we obtain

$$Pr\{U = 0\} \le Pr\{|U - E(U)| \ge E(U)\} \le \exp\left(-\frac{(E(U))^2}{2k^2\gamma}\right)$$
$$= \exp(-\Omega((np)^{\frac{\delta}{2}}/\ln^2(np))) = o(n^{-1}).$$

Hence, the event that $|Y| < \gamma'$ for some i or $U = 0$ has probability $o(1)$. Of course, when we select a random a-subset of Y, we obtain a uniformly distributed a-subset of I. Note that $|\Gamma(i+1) \cap [i]| \le \gamma'$ by (11). We can find a distribution for $a \in [0, i]$ such that when we first choose a, then Y as above, then a random a-subset of Y, we obtain precisely the distribution of $\Gamma(i+1) \cap [i]$, conditioned on (11).

Hence, almost surely for any i, condition (13) holds; that is, we can always choose an appropriate label. $\square$

4. General graphs

Let us prove upper bounds that apply to arbitrary graphs. The obvious greedy algorithm gives the following (cf. Wood [15, Theorem 4]).

Lemma 6. *For any graph G we have*

$$\mathcal{D}(G) \le 2\Delta(G)e(G) + v(G),$$
$$\mathcal{S}(G) \le \Delta(G)e(G) + v(G).$$

Proof. Let us bound $\mathcal{S}(G)$, for example. We choose vertex labels one by one. When we consider a vertex $i \in V(G)$, we are forbidden to choose a previously used label as well as any number of the form $l(u) + l(v) - l(w)$ where $\{u, v\}, \{w, i\} \in E(G)$ and the labels of u, v and w have already been chosen. This forbids at most $v(G) - 1 + d(i)e(G)$ elements so we can always proceed. □

Remarkably, the trivial Lemma 6 is not far from the truth: if applied to $G \in \mathcal{G}(n, p)$, with $\frac{np}{\ln n} \to \infty$, it gives a bound within the multiplicative factor of $O(\ln n)$ from the actual value. It is an interesting open problem to determine the maximum value of $\mathcal{D}(G)$ and $\mathcal{S}(G)$ over all graphs of order n and maximum degree at most d.

For the functions $\mathcal{D}(n, m)$ and $\mathcal{S}(n, m)$ we obtain the following upper bounds.

Theorem 7. *Let $n \to \infty$ and $m \le \binom{n}{2}$. Then*

$$\mathcal{D}(n, m) \le n + (2^{4/3} + o(1))m^{4/3},$$
$$\mathcal{S}(n, m) \le n + (2^{2/3} + o(1))m^{4/3}.$$

Proof. Let us deal with $\mathcal{D}(n, m)$ here. Let n be large and G be an arbitrary graph of order n and size m. It is easy to see that $\mathcal{D}(n, m) = n$ if $m = O(1)$, so assume that $m \to \infty$.

Order the vertices of G by their degrees: $d(x_1) \ge \cdots \ge d(x_n)$. Let $k = \lceil (2m)^{2/3} \rceil$. Label vertices $x_1, \ldots, x_k$ by a Sidon k-subset of $[s]$, $s = \lfloor (1 + \varepsilon)k^2 \rfloor$. We try to label the remaining vertices one by one using labels from $[n + s]$. When choosing a label for x_i, the forbidden values are the already assigned labels, $i - 1$ of them, as well as the numbers $l(x_u) \pm (l(x_v) - l(x_w))$, where $u, v, w \in [i - 1]$ with $\{x_i, x_u\}, \{x_v, x_w\} \in E(G)$, at most $2md(x_i)$ numbers. But $d(x_j) \ge d(x_i)$ for any $j < i$, hence $d(x_i) \le \frac{2m}{i} \le \frac{2m}{k}$ and the total number of forbidden labels is at most

$$n - 1 + \frac{4m^2}{k} < n + s;$$

that is, we can always find a suitable label. □

Needless to say, we have a trivial upper bound, namely $(1 + o(1))n^2$.

Good lower bounds on $\mathcal{D}(n, m)$ and $\mathcal{S}(n, m)$ are provided by random graphs plus isolated vertices. Our aim is to choose $v \le n$ such that, if we define p by $p\binom{v}{2} = (1-\varepsilon)m$, the bound of Theorem 4 for $G \in \mathcal{G}(v, p)$ is as large as possible. In order not to clutter this paper with details we compute only the order of magnitude, not bothering about multiplicative constants.

If $m = \Omega(n^{3/2}\sqrt{\ln n})$, then we take $v = n$. Almost surely $\mathcal{D}(G), \mathcal{S}(G) = \Omega(n^2)$. Otherwise, take $v = \Theta(m^{2/3}(\ln m)^{-1/3}) < n$. Now, the lower bound is

$$\mathcal{D}(n, m), \mathcal{S}(n, m) = \Omega(m^{4/3}(\ln m)^{-2/3}), \qquad \text{for } m = o(n^{3/2}\sqrt{\ln n}).$$

Also, note the trivial lower bound $\mathcal{D}(n, m), \mathcal{S}(n, m) \ge n$.

A little more careful analysis shows that there is an absolute constant C such that our lower and upper bounds on $\mathcal{D}(n, m)$ and $\mathcal{S}(n, m)$ are within factor $C(\ln n)^{2/3}$ for any m, n. This poses an intriguing problem of closing this gap.

References

[1] N. Alon, J. Spencer, The Probabilistic Method, 2nd edition, Wiley Interscience, 2000.
[2] K. Azuma, Weighted sums of certain dependent variables, Tohoku Math. J. 3 (1967) 357–367.
[3] B. Bollobás, A.G. Thomason, Random graphs of small order, Ann. Discrete Math. (1985) 47–97.
[4] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, Ann. Math. Statist. 23 (1952) 493–507.
[5] P Erdős, A survey of problems in combinatorial number theory, Ann. Discrete Math. 6 (1980) 89–115.
[6] P Erdős, R. Freud, On sums of a Sidon-sequence, J. Number Theory 38 (1991) 196–205.
[7] P. Erdős, A. Rényi, On the existence of a factor of degree one of a connected random graph, Acta Math. Acad. Sci. Hungar. 17 (1966) 359–368.
[8] P Erdős, P. Turán, On a problem of Sidon in additive number theory, and on some related problems, J. London Math. Soc. 16 (1941) 212–215.
[9] M. Ferrara, Y. Kohayakawa, V. Rödl, Distance graphs on the integers, Combin. Probab. Comput. (2004) (in press).
[10] H. Halberstam, K.F. Roth, Sequences, Springer Verlag, 1983.
[11] W. Hoeffding, Probability inequalities for sums of bounded random variables, J. Amer. Statist. Assoc. 58 (1963) 13–30.
[12] O. Pikhurko, Dense edge-magic graphs and thin additive bases, 2003 (submitted for publication).
[13] J. Singer, A theorem in finite projective geometry and some applications to number theory, Trans. Amer. Math. Soc. 43 (1938) 377–385.
[14] W.T. Tutte, The factorization of linear graphs, J. London. Math. Soc. 22 (1947) 107–111.
[15] D.R. Wood, On vertex-magic and edge-magic total injections of graphs, Australas. J. Combin. 26 (2002) 49–63.