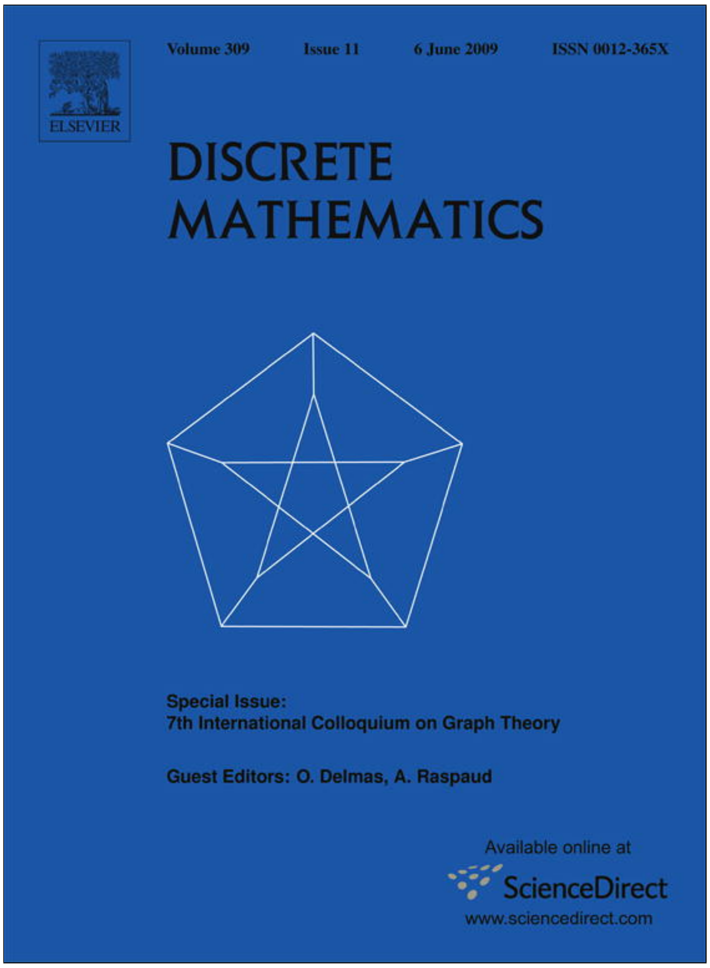
Volume 309 Issue 11 6 June 2009 ISSN 0012-365X
ELSEVIER
DISCRETE MATHEMATICS
Special Issue:
7th International Colloquium on Graph Theory
Guest Editors: O. Delmas, A. Raspaud
Available online at
ScienceDirect
www.sciencedirect.com

Available online at www.sciencedirect.com
ScienceDirect
Discrete Mathematics 309 (2009) 3609–3624
DISCRETE MATHEMATICS
www.elsevier.com/locate/disc

Threshold and complexity results for the cover pebbling game

Anant P. Godbole[a], Nathaniel G. Watson[b], Carl R. Yerger[c,*]

[a] *East Tennessee State University, United States*
[b] *University of California-Berkeley, United States*
[c] *Georgia Institute of Technology, United States*

Received 14 October 2005; accepted 20 December 2007
Available online 8 February 2008

Abstract

Given a configuration of pebbles on the vertices of a graph, a *pebbling move* is defined by removing two pebbles from some vertex and placing one pebble on an adjacent vertex. The cover pebbling number of a graph, $\gamma(G)$, is the smallest number of pebbles such that through a sequence of pebbling moves, a pebble can eventually be placed on every vertex simultaneously, no matter how the pebbles are initially distributed. We determine Bose–Einstein and Maxwell–Boltzmann cover pebbling thresholds for the complete graph. Also, we show that the cover pebbling decision problem is NP-complete.

Keywords: Cover pebbling; Solvable; Threshold; Complete graph

1. Games

There are several popular games that involve the movement of “objects” along a graph-like structure. These include Mancala, where beads are moved along a bent path, and Peg Solitaire, where pegs are moved across a triangular grid. In each case, some objects are removed from the game after a move is made. Both Mancala[1] and Peg Solitare[2] proceed according to a prescribed set of rules. Mancala is a well-defined game between two players, while the solo player in Peg Solitaire pits herself against “nature”. This paper concerns the *pebbling* and *cover pebbling* games, which take place between a highly intelligent Player 1 and a rather non-competitive opponent with a limited strategy. There is a version of pebbling, called “pegging”, which is far closer to Peg Solitaire on graphs than are our games; see [13]. See also [5] for a chessboard game related to pebbling.

The focus of our paper, and indeed of all previous research on the subject, is on deriving conditions under which Player 1 wins the game, or wins with probability that is asymptotic to one, or wins with probability that approaches zero as the size of the problem grows to infinity. Let us start with some baseline definitions and previously derived facts.

* Corresponding author.
E-mail addresses: godbolea@mail.etsu.edu (A.P. Godbole), watson@math.berkeley.edu (N.G. Watson), cyerger@math.gatech.edu (C.R. Yerger).

1 http://www.centralconnector.com/GAMES/mancala.html.
2 http://www.mazeworks.com/peggy/.

0012-365X/$ - see front matter
doi:10.1016/j.disc.2007.12.067

2. Preliminaries

Given a connected graph G, distribute t pebbles on its vertices in some configuration. Specifically, a configuration of weight t on a graph G is a function C from the vertex set $V(G)$ to $\mathbb{N} \cup \{0\}$ such that $\sum_{v \in V(G)} C(v) = t$. Clearly C represents an arrangement of pebbles on $V(G)$. If the pebbles are indistinguishable, there are $\binom{n+t-1}{t} = \binom{n+t-1}{n-1}$ configurations of t pebbles on n vertices. Using quantum mechanical terminology as in [11], we shall call this situation *Bose–Einstein* pebbling and posit that the underlying probability distribution is uniform, i.e. that each of the $\binom{n+t-1}{n-1}$ distributions are equally likely — should the pebbles be thrown randomly onto the vertices. This is the model studied in [8]. Now there is no reason to assume, *a priori*, that the pebbles are indistinguishable. Accordingly, if the pebbles are distinct, we shall refer to our process as *Maxwell–Boltzmann* pebbling, in which a random distribution of pebbles leads to n^t equiprobable configurations. Maxwell–Boltzmann pebbling does not appear to have been studied more than peripherally in the literature.

A *pebbling move* is defined as the removal of two pebbles from some vertex and the placement of one of these on an adjacent vertex. Given an initial configuration, a vertex v is called *reachable* if it is possible to place a pebble on it in finitely many pebbling moves. The graph G is said to be *pebbleable* (this is not standard nomenclature) if any of its vertices can be thus reached. Define the pebbling number $\pi(G)$ to be the minimum number of pebbles that are sufficient to pebble the graph regardless of the initial configuration. The pebbling game may thus be described as follows: Player 2 specifies a distribution C and a target vertex v. Player 1 wins the game iff she is able to reach vertex v using a sequence of pebbling moves. The pebbling number of G is the smallest number t_0 of pebbles so that Player 1 wins no matter what strategy Player 2 employs.

The origin of pebbling is rather interesting and somewhat unexpected. Lagarias and Saks [14] were considering a way to produce an alternative proof to a conjecture of Erdős and Lemke, which Lemke and Kleitman proved in 1989 [19]. It is known that for any set $N = \{n_1, n_2, \ldots, n_q\}$ of q natural numbers, there is a nonempty index set $I \subset \{1, \ldots, q\}$ such that $q \mid \sum_{i \in I} n_i$. The Erdős–Lemke conjecture states that the additional conclusion $\sum_{i \in I} n_i \leq \operatorname{lcm}(q, n_1, n_2, \ldots, n_q)$ could also be reached, see [14]. Unfortunately, Lemke and Kleitman's argument was detailed and contained a considerable amount of case analysis. This provoked Lagarias and Saks to invent graph pebbling as a way to produce a cleaner proof, since such a proof would follow easily if the pebbling number of the Cartesian product of paths was determined. This was accomplished in a landmark paper by Chung [4]. One generalization of pebbling, called p-pebbling, was utilized in Chung's proof, and is defined as the removal of p pebbles from some vertex and the placement of one pebble on an adjacent vertex. It turns out that a greedy-like condition, called the numerical pebbling operation for pebbling paths, can be used to prove Chung's theorem in [14]. In fact, one of the lemmas of Chung's proof actually uses the fact that for any set $N = \{n_1, n_2, \ldots, n_q\}$ of q natural numbers, there is a nonempty index set $I \subset \{1, \ldots, q\}$ such that $q \mid \sum_{i \in I} n_i$.

Special cases: The pebbling number $\pi(P_n)$ of the path is 2^{n-1}. Chung [4] proved that $\pi(Q^d) = 2^d$ and $\pi(P_n^m) = 2^{(n-1)m}$, where Q^d is the d-dimensional binary cube and P_n^m is the m-fold product of the n-path with itself. An easy pigeonhole principle argument yields $\pi(K_n) = n$. The pebbling number of trees has been determined, see [14].

One of the key conjectures in pebbling, now proved in several special cases, is due to Graham; its resolution would clearly generalize Chung's result for m-dimensional grids:

Graham's conjecture. *The pebbling number of the Cartesian product of two graphs is no more that the product of the pebbling numbers of the two graphs, i.e.*

$$\pi(G \square H) \leq \pi(G)\pi(H).$$

Structural characteristics of graphs have also been employed to determine the pebbling number of specific classes of graphs. For instance, a graph is said to be *Class 0* if $\pi(G)$ equals the number of vertices of G. Cubes are of Class 0, as are complete graphs, but what other families fall in this important class of graphs for which π is as low as it can possibly be? Here are two answers: For graphs of diameter 2, if G is 3-connected, that is, the removal of 2 or fewer vertices does not disconnect the graph, then G is Class 0 [6]. Further, the authors in [10] prove a more general theorem that a diameter d graph of high enough connectivity ($2^d/d$ is required, 2^{2d+3} is sufficient) is Class 0. In fact, if we consider $G(n, p)$, the class of random graphs on n vertices where the probability of each particular edge being present is a fixed constant $p \in (0, 1)$, then almost all such graphs are Class 0 [6]. Generalizations of this result to the case where $p = p_n \to 0$ as $n \to \infty$ are also available; see [10]. Other authors, e.g. [3], have obtained general pebbling bounds, while Bukh [2] has proved almost-tight asymptotic bounds on the pebbling number of diameter three graphs.

Another aspect of pebbling that has been explored is the random structure one obtains when placing pebbles randomly on graphs. Specifically, we seek the probability that a graph G is pebbleable when t pebbles are placed randomly on it according to the Bose Einstein or Maxwell–Boltzmann scheme. Numerous *threshold results* have been determined in [8] for Bose–Einstein pebbling of families of graphs such as K_n, the complete graph on n vertices; C_n, the cycle on n vertices; stars; wheels; etc. A threshold result is a theorem of the following kind, where a_n and b_n are non-negative sequences:

$$t = t_n \gg a_n \Rightarrow \mathbb{P}(G = G_n \text{ is pebbleable}) \to 1 \quad (n \to \infty)$$
$$t = t_n \ll b_n \Rightarrow \mathbb{P}(G = G_n \text{ is pebbleable}) \to 0 \quad (n \to \infty),$$

where we write, for non-negative sequences c_n and d_n, $c_n \gg d_n$ (or $d_n \ll c_n$) if $c_n/d_n \to \infty$ as $n \to \infty$. Of course, we will have reason to feel particularly gratified if we can show that $a_n = b_n$ in a result of this genre. For the families of complete graphs, wheels and stars, for example, we know [8] that $a_n = b_n = \Theta(\sqrt{n})$ for the Bose–Einstein scheme. In many cases, however, the analysis is quite delicate; see [27,9] for some of the issues involved in finding the pebbling threshold for a family as basic as P_n, the path on n vertices. The fundamental reference [1] contains general results on the existence of sharp pebbling thresholds for families of graphs.

A detailed survey of graph pebbling has been presented by Hurlbert [14], newer results are found in [15] and it would probably not be an oversimplification to state that most results available to date fall in five broad categories: finding pebbling numbers for classes of graphs; extending these results to work towards Graham's product conjecture; addressing the issue of when a family of graphs is of class 0; pinpointing graph pebbling thresholds; and seeking to understand the complexity issues in graph pebbling [16]. A survey of open problems in graph pebbling may be found on Glenn Hurlbert's pebbling website.[3] This site is the most complete reference on pebbling as of the writing of this article.

The above mini-survey on pebbling notwithstanding, we focus in this paper on a *variant* of pebbling called cover pebbling, first discussed by Crull et al. [7]. For reasons that will become obvious, we focus only on analogues of the last two of the five general directions mentioned above.

The *cover pebbling number* $\gamma(G)$ is defined as the minimum number of pebbles required such that it is possible, given any initial configuration of at least $\gamma(G)$ pebbles on G, to make a series of pebbling moves that *simultaneously* reaches *each* vertex of G. A configuration is said to be *cover solvable* if it is possible to place a pebble on every vertex of G starting with that configuration. Various results on cover pebbling have been determined. For instance, we now know [7] that $\gamma(K_n) = 2n - 1$; $\gamma(P_n) = 2^n - 1$; and that for trees T_n,

$$\gamma(T_n) = \max_{v \in V(T_n)} \sum_{u \in V(T_n)} 2^{\mathrm{dist}(u,v)}. \tag{1}$$

Likewise, it was shown in [17] that $\gamma(Q^d) = 3^d$ and in [25] that $\gamma(K_{r_1,\ldots,r_m}) = 4r_1 + 2(r_2 + \cdots + r_m) - 3$, where $r_1 \geq r_2 \geq \ldots \geq r_m$. The above examples reveal that for these special classes of graphs, at any rate, the cover pebbling number equals the "stacking number", or, put another way, the worst possible distribution of pebbles consists of placing all the pebbles on a single vertex. The intuition built by computing the value of the cover pebbling number for the families K_n, P_n, and T_n in [7] led to Open Question No. 10 in [7], which was christened the *Stacking Conjecture* by students at the Summer 2004 East Tennessee State University REU. In an exciting summer development, participants Annalies Vuong and Ian Wyckoff [23] were able to prove the claim

Stacking Theorem. *For any connected graph G,*

$$\gamma(G) = \max_{v \in V(G)} \sum_{u \in V(G)} 2^{\mathrm{dist}(u,v)},$$

[3] http://mingus.la.asu.edu/~hurlbert/pebbling/pebb.html.

thereby proving that (1) holds for all graphs. (The stacking theorem was independently proved soon after by Sjöstrand [21].) In fact, the key result in Vuong and Wyckoff's paper [23] is really a sufficient condition for a distribution to be cover solvable, so further investigations in the theory of (cover) pebbling might soon veer, we speculate, in a sixth general direction, namely a study of which distributions are (cover) solvable and which are not. Indeed, such a research thrust would be most consistent with our description of pebbling as a *game*, and is addressed in Section 5 of this paper.

3. Maxwell–Boltzmann cover pebbling threshold for K_n

It is evident that n is the smallest number of pebbles that might suffice to cover pebble K_n — in the unlikely event that they happen to be distributed one apiece on the vertices. On the other hand, we know that $2n-1$ pebbles always suffice, since $\gamma(K_n) = 2n-1$. We seek a sharp cover pebbling threshold that is somewhere in between these two extremes, when distinguishable pebbles are thrown onto the n vertices of K_n according to the Maxwell–Boltzmann scheme. To explain *why* there is a dramatic increase in the probability of the pebbleability of K_n at $t = (1.5238\cdots)n$, we first prove an important auxiliary result that gives necessary and sufficient conditions for a configuration to be cover solvable. Such results are not easy to come by, as we will further see in Section 5.

3.1. Necessary and Sufficient Conditions for Cover Solvability of K_n

Let $X = X_{n,t}$ be the number of vertices on which an odd number of pebbles are placed. We will often refer to X as the number of *odd stacks*.

Lemma 1. *A configuration of t pebbles on the n vertices of K_n is cover solvable if and only if*

$$X + t \geq 2n. \tag{2}$$

Proof. Given a configuration C, let Y_i, $0 \leq i \leq t$, be the number of vertices with i pebbles. Now a vertex can cover exactly i others if and only if it has either $2i+1$ or $2i+2$ pebbles on it. It follows that C is cover solvable iff

$$\sum_{i\geq 0} i(Y_{2i+1} + Y_{2i+2}) \geq Y_0,$$

or, iff

$$\sum_{i\geq 0}(2i+1)Y_{2i+1} + \sum_{i\geq 0}(2i+2)Y_{2i+2} \geq 2Y_0 + \sum_{i\geq 0} Y_{2i+1} + 2\sum_{i\geq 0} Y_{2i+2},$$

Now, since $t \geq 2E + X = 2n - X$, we thus have $X + t \geq 2n$, as required. □

Armed with Lemma 1, we now provide the heuristic reason why we believe there is a sharp cover pebbling threshold at $t = (1.5238\cdots)n$. Given a random variable X with expected value $\mathbb{E}(X)$, we will say that X is *sharply concentrated* around $\mathbb{E}(X)$ if $X \approx \mathbb{E}(X)$. Assuming, therefore, that $X \approx \mathbb{E}(X)$, it makes sense to speculate that K_n is cover solvable with high probability whenever $\mathbb{E}(X) \geq 2n-t$. But $X = \sum_{j=1}^{n} I_j$, where $I_j = 1$ (resp. 0) if there is an odd (resp. even) stack on vertex j, so that linearity of expectation yields

$$\begin{aligned}\mathbb{E}(X) &= n\mathbb{P}(I_1 = 1)\\ &= n\sum_{j \text{ odd}} \binom{t}{j}\left(\frac{1}{n}\right)^j\left(1-\frac{1}{n}\right)^{t-j}\\ &= \frac{n}{2}\left(1-\left(1-\frac{2}{n}\right)^t\right).\end{aligned} \tag{3}$$

Thus $\mathbb{E}(X) \geq 2n - t$ iff

$$t - \frac{n}{2}\left(1-\frac{2}{n}\right)^t \geq \frac{3n}{2}, \tag{4}$$

and, parametrizing by setting $t = An$, we see that (4) holds iff

$$A - \frac{1}{2}\left(1 - \frac{2}{n}\right)^{An} \geq \frac{3}{2}. \tag{5}$$

Since $(1 - 2/n)^n \sim \mathrm{e}^{-2}$, we see from (5) that a reasonable guess for a threshold value of t is $A_0 n$ where A_0 is the solution of

$$A - \frac{1}{2}\exp\{-2A\} = \frac{3}{2},$$

or $A_0 = 1.5238\cdots$ We next make this hunch more precise.

3.2. Main result

Various tools are used to establish concentration of measure results. Some of the more sophisticated techniques employed are the martingale method, a.k.a. Azuma's inequality (proved independently and a few years earlier by W. Hoeffding), and Talagrand's isoperimetric inequalities in product spaces; see [22] for an exposition of both. In this section, however, we establish our main result by using a baseline technique, *viz.* Tchebychev's inequality, also known in probabilistic combinatorics circles as the "second moment method." Theorem 2 below thus yields the "correct" result, but with a sub-optimal rate of convergence due to the crudeness of the method employed. Tchebychev's inequality states that for any random variable X and for any $\gamma > 0$,

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geq \gamma\right) \leq \frac{\mathbb{V}(X)}{\gamma^2},$$

where $\mathbb{V}(Z)$ denotes the variance of Z. We start by computing the variance of X; $\mathbb{E}(X)$ has already been evaluated in (3). We have

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{V}\left(\sum_{j=1}^{n} I_j\right)\\ &= \sum_{j=1}^{n}\mathbb{V}(I_j) + \sum_{i\neq j}\left(\mathbb{E}(I_i I_j) - \mathbb{E}(I_i)\mathbb{E}(I_j)\right)\\ &= n\mathbb{P}(I_1 = 1)(1 - \mathbb{P}(I_1 = 1)) + n(n-1)\left\{\mathbb{P}(I_1 I_2 = 1) - \mathbb{P}^2(I_1 = 1)\right\},\end{aligned} \tag{6}$$

So we focus on computing $\mathbb{P}(I_1 I_2 = 1)$, i.e., the probability that both vertex 1 and vertex 2 have odd stacks of pebbles. It is not too hard to verify that

$$\begin{aligned}\mathbb{P}(I_1 I_2 = 1) &= \sum_{r,s \text{ odd}} \binom{t}{r, s, t-r-s}\left(\frac{1}{n}\right)^r\left(\frac{1}{n}\right)^s\left(1 - \frac{2}{n}\right)^{t-r-s}\\ &= \frac{1}{4}\left(\Sigma_1 - \Sigma_2 - \Sigma_3 + \Sigma_4\right),\end{aligned}$$

where

$$\begin{aligned}\Sigma_1 &= \sum_{r,s} \binom{t}{r, s, t-r-s}\left(\frac{1}{n}\right)^r\left(\frac{1}{n}\right)^s\left(1 - \frac{2}{n}\right)^{t-r-s},\\ \Sigma_2 &= \sum_{r,s} \binom{t}{r, s, t-r-s}\left(\frac{1}{n}\right)^r\left(-\frac{1}{n}\right)^s\left(1 - \frac{2}{n}\right)^{t-r-s},\\ \Sigma_3 &= \sum_{r,s} \binom{t}{r, s, t-r-s}\left(-\frac{1}{n}\right)^r\left(\frac{1}{n}\right)^s\left(1 - \frac{2}{n}\right)^{t-r-s},\end{aligned}$$

and

$$\Sigma_4 = \sum_{r,s} \binom{t}{r, s, t-r-s} \left(-\frac{1}{n}\right)^r \left(-\frac{1}{n}\right)^s \left(1-\frac{2}{n}\right)^{t-r-s},$$

so that

$$\mathbb{P}(I_1 I_2 = 1) = \frac{1}{4}\left(1 + \left(1-\frac{4}{n}\right)^t - 2\left(1-\frac{2}{n}\right)^t\right),$$

and

$$\begin{aligned} \mathrm{Cov}(I_1, I_2) &= \mathbb{E}(I_1 I_2) - \mathbb{E}(I_1)\mathbb{E}(I_2) \\ &= \frac{1}{4}\left\{\left(1-\frac{4}{n}\right)^t - \left(1-\frac{2}{n}\right)^{2t}\right\}. \end{aligned}$$

It follows from (6) that

$$\mathbb{V}(X) = \frac{n}{4}\left(1 - \left(1-\frac{2}{n}\right)^{2t}\right) + \frac{n(n-1)}{4}\left\{\left(1-\frac{4}{n}\right)^t - \left(1-\frac{2}{n}\right)^{2t}\right\}. \tag{7}$$

We are now ready to state the following theorem:

Theorem 2. *Consider t distinct pebbles that are thrown onto the vertices of the complete graph K_n on n vertices according to the Maxwell–Boltzmann distribution. Set $A_0 = 1.5238\cdots$. Then*

$$t = A_0 n + \varphi(n)\sqrt{n} \Rightarrow \mathbb{P}(K_n \text{ is cover solvable}) \to 1 \quad (n \to \infty)$$

and

$$t = A_0 n - \varphi(n)\sqrt{n} \Rightarrow \mathbb{P}(K_n \text{ is cover solvable}) \to 0 \quad (n \to \infty),$$

where $\varphi(n) \to \infty$ is arbitrary.

Proof. Assume first that $t = A_0 n + \varphi(n)\sqrt{n}$. Then

$$\begin{aligned} \mathbb{P}(X \ge 2n - t) &= \mathbb{P}\left(X - \mathbb{E}(X) \ge 2n - t - \frac{n}{2}\left(1 - \left(1-\frac{2}{n}\right)^t\right)\right) \\ &= \mathbb{P}\left(X - \mathbb{E}(X) \ge \frac{3}{2}n - A_0 n + \frac{n}{2}\left(1-\frac{2}{n}\right)^{A_0 n + \varphi(n)\sqrt{n}} - \varphi(n)\sqrt{n}\right) \\ &= \mathbb{P}\left(X - \mathbb{E}(X) \ge -\frac{n}{2}\mathrm{e}^{-2A_0} + \frac{n}{2}\left(1-\frac{2}{n}\right)^{A_0 n + \varphi(n)\sqrt{n}} - \varphi(n)\sqrt{n}\right) \\ &\ge \mathbb{P}\left(X - \mathbb{E}(X) \ge -\frac{n}{2}\mathrm{e}^{-2A_0} + \frac{n}{2}\exp\left\{-2A_0 - \frac{2\varphi(n)}{\sqrt{n}}\right\} - \varphi(n)\sqrt{n}\right) \\ &= \mathbb{P}\left(X - \mathbb{E}(X) \ge \frac{n}{2}\mathrm{e}^{-2A_0}\left\{\exp\left\{\frac{-2\varphi(n)}{\sqrt{n}}\right\} - 1\right\} - \varphi(n)\sqrt{n}\right) \\ &\ge \mathbb{P}\left(X - \mathbb{E}(X) \ge \frac{n}{2}\mathrm{e}^{-2A_0} \cdot \frac{-2\varphi(n)}{\sqrt{n}}(1+o(1)) - \varphi(n)\sqrt{n}\right) \\ &= \mathbb{P}\left(X - \mathbb{E}(X) \ge -\varphi(n)\sqrt{n}\left(1 + \mathrm{e}^{-2A_0}\right)(1+o(1))\right) \\ &\ge \mathbb{P}\left(|X - \mathbb{E}(X)| \le \varphi(n)\sqrt{n}\left(1 + \mathrm{e}^{-2A_0}\right)(1+o(1))\right). \end{aligned} \tag{8}$$

In (8), the second and third inequalities follow from the facts that $1 - x \le \mathrm{e}^{-x}$ and $\mathrm{e}^{-x} - 1 \le -x/(1+x)$ respectively. We next analyze the variance as given by (7). The first component $\frac{n}{4}(1 - (1 - \frac{2}{n})^{2t})$ of $\mathbb{V}(X)$ can easily be verified to be of the form $\frac{n}{4}(1 + K(1 + o(1)))$ (for some constant K) when t is as specified; the second component

$$\frac{n(n-1)}{4}\left\{\left(1-\frac{4}{n}\right)^t - \left(1-\frac{2}{n}\right)^{2t}\right\} = \frac{n(n-1)}{4}\left\{\left(1-\frac{4}{n}\right)^t - \left(1-\frac{4}{n}+\frac{4}{n^2}\right)^t\right\},$$

on the other hand, may be bounded using the inequalities

$$t(b-a)a^{t-1} \le b^t - a^t \le t(b-a)b^{t-1}$$

for $b > a$ to yield, for t as above,

$$\begin{aligned}\frac{n(n-1)}{4}\left\{\left(1-\frac{4}{n}+\frac{4}{n^2}\right)^t - \left(1-\frac{4}{n}\right)^t\right\} &= \frac{n(n-1)}{4}\frac{4t}{n^2}\left(1-\frac{4}{n}+\Theta\left(\frac{1}{n^2}\right)\right)^{t-1}\\ &= \Theta(t)\\ &= \Theta(n).\end{aligned}$$

It follows that $\mathbb{V}(X) = \Theta(n)$ and thus we have, by (8) and Tchebychev's inequality, with K representing a generic constant, that

$$\mathbb{P}(X \ge 2n - t) \ge \mathbb{P}\left(|X - \mathbb{E}(X)| \le K\cdot\sqrt{n}\varphi(n)\right) \ge 1 - \frac{1}{K^2\varphi^2(n)} \to 1 \quad (n\to\infty),$$

as asserted.

The proof of the second half of the theorem is similar; we bound above instead of below to get, with $t = A_0 n - \varphi(n)\sqrt{n}$,

$$\begin{aligned}\mathbb{P}(X \ge 2n-t) &= \mathbb{P}\left(X - \mathbb{E}(X) \ge -\frac{n}{2}e^{-2A_0} + \frac{n}{2}\left(1-\frac{2}{n}\right)^{A_0 n - \varphi(n)\sqrt{n}} + \varphi(n)\sqrt{n}\right)\\ &\le \mathbb{P}\left(X - \mathbb{E}(X) \ge -\frac{n}{2}e^{-2A_0} + \frac{n}{2}\exp\left\{\left(-2A_0 + \frac{2\varphi(n)}{\sqrt{n}}\right)(1+o(1))\right\} + \varphi(n)\sqrt{n}\right)\\ &= \mathbb{P}\left(X - \mathbb{E}(X) \ge \frac{n}{2}e^{-2A_0}\left\{\exp\left\{\frac{2\varphi(n)}{\sqrt{n}}(1+o(1))\right\} - 1\right\} + \varphi(n)\sqrt{n}\right)\\ &\le \mathbb{P}\left(X - \mathbb{E}(X) \ge \frac{n}{2}e^{-2A_0}\cdot\frac{2\varphi(n)}{\sqrt{n}}(1+o(1)) + \varphi(n)\sqrt{n}\right)\\ &\le \mathbb{P}\left(|X - \mathbb{E}(X)| \ge \varphi(n)\sqrt{n}\left(1 + e^{-2A_0}(1+o(1))\right)\right)\\ &\le \frac{K}{\varphi^2(n)} \to 0.\end{aligned}$$

This completes the proof. □

Remarks: Note that $\mathbb{V}(X) = \Theta(n)$ implies that X is concentrated w.h.p. in an interval of length $\Omega(\sqrt{n})$ around $\mathbb{E}(X) = \Theta(n)$. This is because of the inequality $\mathbb{V}(X) \le \text{Range}^2(X)/4$.

4. Bose–Einstein cover pebbling threshold

4.1. Exact distributions

In the Maxwell–Boltzmann scheme, it is extremely difficult (though not impossible) to calculate $\mathbb{P}(X = x)$ exactly. Moreover, the formula for $\mathbb{P}(X = x)$ obtained thus is quite intractable. Surprisingly, this is not the case when we consider Bose–Einstein statistics; we can derive such a formula by counting the number of configurations of size t on K_n with x odd stacks. Call this value $\phi(x, t, n)$. Clearly if x and t have different parity, or $x > t$ or $x > n$, $\phi(x, t, n) = 0$. Suppose this is not the case. Then, we can find the configurations with x odd stacks of pebbles by placing one pebble on the x vertices that are to have odd stacks on them, and then distributing the remaining $t - x$ pebbles on the n vertices of G in $\frac{t-x}{2}$ indistinguishable pairs. Thus, since the vertices with odd stacks may be chosen in $\binom{n}{x}$ ways, we have proved

Proposition 3. *If t and x have the same parity, and if $x \le \min\{t, n\}$, then*

$$\mathbb{P}(X = x) = \frac{\binom{n}{x}\binom{\frac{t-x}{2}+n-1}{n-1}}{\binom{n+t-1}{n-1}}.$$

In fact, the above exact distribution was used fruitfully in [26] to prove a weaker version of the main result of the next section. We choose, however, to prove a cover pebbling threshold for Bose–Einstein pebbling by using more contemporary probabilistic tools.

4.2. Polya sampling and Azuma's inequality yield dividends

There is a natural and sequential probabilistic process associated with Maxwell–Boltzmann pebbling. We simply take t pebbles (balls) and throw them one by one onto (into) n vertices (urns) in the "natural" way that inspires many elementary problems in discrete probability texts. By contrast, the "global" Bose–Einsteinian positioning of t indistinct balls into n distinct urns – so that we obtain $\binom{n+t-1}{n-1}$ equiprobable configurations – does not *appear* to have a sequential process associated with it. *But it does.* We first rephrase the problem — not as one associated with throwing balls *into* boxes but, conversely, as a sampling problem, i.e., drawing balls *from* boxes. In this light, Maxwell–Boltzmann pebbling consists of drawing t balls "with replacement" from a box containing one ball of each of n colours, with the understanding that the number of balls of colour j drawn in the altered model equals the number of pebbles that are tossed onto vertex j *à la* the balls-in-boxes model. *Bose–Einstein pebbling can be recast in a similar fashion, but one needs to employ a process called Pólya sampling.* Pólya sampling (or the Pólya urn model) is described in [11] as a means of modelling contagious diseases and takes place as follows: Initially the urn contains one ball of each of n colours. After each draw, the selected ball is replaced together with another ball of the same colour. In this mode of sampling, we lose the independence inherent to the with-replacement procedure, and, as a matter of fact, the selection process is not even Markovian — but are able to "see" the sequentiality that will be critical in the sequel. As before, the number of times that colour j is drawn can be set to equal the number of pebbles on vertex j, but do these two procedures yield the same probability model? We claim so, and here is a proof of this rather well-known fact:

Lemma 4. *Let X_j be the number of times the color j is drawn among the t draws. Then for any $x_1, x_2, \ldots, x_n$ with $\sum x_j = t$,*

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \frac{1}{\binom{n+t-1}{t}}$$

Proof. First let us find the probability that the stated outcome appears in the following order: colour 1 is first drawn x_1 times, then colour 2 is drawn x_2 times, etc., until we finally draw colour n the last x_n times. Call this ordered event A. It is easily seen that

$$\begin{aligned}\mathbb{P}(A) &= \frac{(1 \cdot 2 \cdot \cdots \cdot x_1)(1 \cdot 2 \cdot \cdots \cdot x_2) \cdots (1 \cdot 2 \cdot \cdots \cdot x_n)}{n(n+1)\cdots(n+t-1)} \\ &= \frac{x_1! x_2! \cdots x_n!}{n(n+1)\cdots(n+t-1)}.\end{aligned}$$

Moreover, this probability is the same regardless of the order in which the balls are drawn. The total number of ways of ordering our configuration turns out to be $t!/x_1!x_2!\cdots x_n!$. Thus, we can write the probability of obtaining a given configuration as

$$\begin{aligned}\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) &= \frac{t!}{x_1!x_2!\cdots x_n!} \cdot \frac{x_1!x_2!\cdots x_n!}{n(n+1)\cdots(n+t-1)} \\ &= \frac{1}{\binom{n+t-1}{t}}.\end{aligned}$$

This concludes the proof. □

We next provide the reader with background concerning the Azuma martingale inequality. Consider the following generic set up: $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and (Y_n) a sequence of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ that may not, in general, be independent. In our case, the Y_js are the sequence of draws made in the Pólya urn scheme associated with the random cover pebbling problem. Let $X = X_t = X(Y_1, \dots, Y_t)$ be an objective function (in our case X is the number of odd stacks of pebbles), and consider the filtration (sequence of sigma algebras) $(\mathcal{F}_n)$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_i = \sigma(Y_1, \dots, Y_i)$. Let $\mathbb{E}_i X$ denote the conditional expectation of X with respect to $\mathcal{F}_i$ (with $\mathbb{E} = \mathbb{E}_0$) and set $d_i = \mathbb{E}_i X - \mathbb{E}_{i-1} X$. Then it is well known that $(d_i, \mathcal{F}_i)$ is a martingale difference sequence, and that we have $X - \mathbb{E}(X) = \sum_{i=1}^{t} d_i$. A key method used towards gaining an understanding of the concentration of X around $\mathbb{E}(X)$ is the method of bounded differences, also known as the Azuma (1967)–Hoeffding (1963) inequality [22]:

Lemma 5 (*Azuma–Hoeffding*). *For all $\gamma > 0$,*

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \gamma) \leq 2 \exp\left\{-\frac{\gamma^2}{2\sum \|d_i\|_\infty^2}\right\}. \tag{9}$$

Here $\|Z\|_\infty = \sup |Z(\omega)|$. Expression (9) may also be applied as follows: Letting (Y_i^*) be an independent copy of (Y_i), we have

$$\mathbb{E}_{i-1} X(Y_1, \dots, Y_t) = \mathbb{E}_i X(Y_1, \dots, Y_{i-1}, Y_i^*, Y_{i+1} \dots, Y_t),$$

so that d_i can be written as a single conditional expectation as follows:

$$d_i = \mathbb{E}_i \left(X(Y_1, \dots, Y_t) - X(Y_1, \dots, Y_{i-1}, Y_i^*, Y_{i+1} \dots, Y_t) \right),$$

and thus

$$\begin{aligned} \|d_i\|_\infty &= \|\mathbb{E}_i \{X(Y_1, \dots, Y_t) - X(Y_1, \dots, Y_{i-1}, Y_i^*, Y_{i+1} \dots, Y_t)\}\|_\infty \\ &\leq \|X(Y_1, \dots, Y_t) - X(Y_1, \dots, Y_{i-1}, Y_i^*, Y_{i+1} \dots, Y_t)\|_\infty. \end{aligned} \tag{10}$$

The philosophy behind the method of bounded differences is thus that the change in the value of X resulting from a change in a single input variable should be small. Moreover the bound in (10) shows us, after a moment's reflection, that we have, for our problem, $\|d_i\|_\infty \leq 2$. Furthermore, Lemma 5 and (10) yield the following concentration for the number X of odd stacks in the Bose–Einstein scheme:

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geq \gamma\right) \leq 2 \exp\left\{-\frac{\gamma^2}{8t}\right\}, \tag{Az}$$

so that X is concentrated in an interval of length $\sqrt{n}\varphi(n)$ around $\mathbb{E}(X)$ whenever $t \sim Kn$, where $\varphi(n) \to \infty$ is arbitrary. We are now ready to prove the main result of this section — one that features the golden ratio $g = (1 + \sqrt{5})/2$:

Theorem 6. *Consider t indistinguishable pebbles that are placed on the vertices of the complete graph K_n according to the Bose–Einstein distribution. Then, with g representing the golden ratio $(1 + \sqrt{5})/2$,*

$$t = gn + \varphi(n)\sqrt{n} \Rightarrow \mathbb{P}(K_n \text{ is cover solvable}) \to 1 \quad (n \to \infty)$$

and

$$t = gn - \varphi(n)\sqrt{n} \Rightarrow \mathbb{P}(K_n \text{ is cover solvable}) \to 0 \quad (n \to \infty),$$

where $\varphi(n) \to \infty$ is arbitrary.

Proof. We start by establishing tight bounds on $\mathbb{E}(X)$; unfortunately, the sum

$$\mathbb{E}(X) = n\mathbb{P}(I_1 = 1)$$

does not appear to be tractable. Consider first an upper bound on $\mathbb{P}(I_1 = 1)$: We have, with α_n and ℓ_n representing, respectively, generic functions of the form $O(1/n)$ and $O(\log n/n)$ whose exact form may vary from line to line,

$$\begin{aligned}
\mathbb{P}(I_1 = 1) &= \mathbb{P}(\text{vertex 1 has an odd number of pebbles}) \\
&= \sum_{j=1}^{\text{``}\infty\text{''}} \frac{\binom{n+t-2j-1}{n-2}}{\binom{n+t-1}{n-1}} \\
&= \sum_{j=1}^{\infty} \frac{n-1}{(n+t-1)} \frac{t}{(n+t-2)} \frac{t-1}{(n+t-3)} \frac{t-2}{(n+t-4)} \cdots \frac{t-2j+2}{(n+t-2j)} \\
&\le \frac{n}{(n+t)} \frac{t}{(n+t)} (1+\alpha_n) \sum_{j=1}^{\infty} \frac{t-1}{(n+t-3)} \frac{t-2}{(n+t-4)} \cdots \frac{t-2j+2}{(n+t-2j)} \\
&\le \frac{nt}{(n+t)^2} (1+\alpha_n) \sum_{j=1}^{\infty} \left(\frac{t-1}{n+t-3} \right)^{2j-2} \\
&= (1+\alpha_n) \frac{nt}{(n+t)^2} \frac{1}{1 - \left(\frac{t-1}{n+t-3} \right)^2} \\
&= (1+\alpha_n) \frac{nt}{(n+t)^2} \frac{1}{1 - \left(\frac{t}{n+t} \right)^2} \\
&= (1+\alpha_n) \frac{t}{n+2t}.
\end{aligned} \tag{11}$$

Next note that for R to be determined,

$$\begin{aligned}
\mathbb{P}(I_1 = 1) &\ge \sum_{j=1}^{R} \frac{n-1}{(n+t-1)} \frac{t}{(n+t-2)} \frac{t-1}{(n+t-3)} \frac{t-2}{(n+t-4)} \cdots \frac{t-2j+2}{(n+t-2j)} \\
&\ge \frac{nt}{(n+t)^2} (1+\alpha_n) \sum_{j=1}^{R} \left(\frac{t-2j+2}{n+t-2j} \right)^{2j-2} \\
&\ge \frac{nt}{(n+t)^2} (1+\alpha_n) \sum_{j=1}^{R} \left(\frac{t-2R+2}{n+t-2R} \right)^{2j-2} \\
&= \frac{nt}{(n+t)^2} (1+\alpha_n) \frac{1 - \left(\frac{t-2R+2}{n+t-2R} \right)^{2R}}{1 - \left(\frac{t-2R+2}{n+t-2R} \right)^2}.
\end{aligned} \tag{12}$$

We now pick R in (12) so that

$$1 - \left(\frac{t-2R+2}{n+t-2R} \right)^{2R} = 1 - \alpha_n;$$

This may be done, e.g., with $R = K \log n$, since in the range of t's that we are dealing with (namely $t = gn \pm \varphi(n)\sqrt{n}$), we have

$$1 - \left(\frac{t-2R+2}{n+t-2R} \right)^{2R} = 1 - \left(\frac{g\left(1 + O\left(\frac{\varphi(n)}{\sqrt{n}}\right)\right)}{(1+g)\left(1 + O\left(\frac{\varphi(n)}{\sqrt{n}}\right)\right)} \right)^{2R} = 1 + \alpha_n$$

if R is an appropriate multiple of $\log n$. Eq. (12) thus yields

$$\begin{aligned}\mathbb{P}(I_1 = 1) &\geq \frac{nt}{(n+t)^2}(1+\alpha_n)\frac{(n+t-2R)^2}{(n-2)(n+2t-4R+2)} \\ &= (1+\alpha_n)(1+\ell_n)\frac{t}{n+2t} \\ &= (1+\ell_n)\frac{t}{n+2t}. \end{aligned} \tag{13}$$

Eqs. (11) and (13) thus give

$$(1+\ell_n)\frac{nt}{n+2t} \leq \mathbb{E}(X) \leq (1+\alpha_n)\frac{nt}{n+2t}, \tag{14}$$

and we are ready to derive our cover pebbling threshold. We first set $t = gn + \varphi(n)\sqrt{n}$ to get, for some (and this is important) *positive* A,

$$\begin{aligned}\mathbb{P}(X \geq 2n-t) &\geq \mathbb{P}\left(X - \mathbb{E}(X) \geq 2n - t - \frac{nt}{n+2t}(1+\ell_n)\right) \\ &= \mathbb{P}\left(X - \mathbb{E}(X) \geq (2-g)n - \varphi(n)\sqrt{n} - \frac{n(gn+\varphi(n)\sqrt{n})}{n+2(gn+\varphi(n)\sqrt{n})}(1+\ell_n)\right) \\ &= \mathbb{P}\left(X - \mathbb{E}(X) \geq (2-g)n - \varphi(n)\sqrt{n} - \frac{gn^2\left(1+\frac{\varphi(n)}{g\sqrt{n}}\right)}{n(1+2g)\left(1+\frac{2\varphi(n)}{(1+2g)\sqrt{n}}\right)}(1+\ell_n)\right) \\ &\geq \mathbb{P}\left(X - \mathbb{E}(X) \geq (2-g)n - \varphi(n)\sqrt{n} - \frac{gn}{(1+2g)}\left(1 + A\left(\frac{\varphi(n)}{\sqrt{n}}\right)\right)(1+\ell_n)\right). \end{aligned} \tag{15}$$

Now $(2-g)n - gn/(1+2g) = 0$, so (15) yields

$$\begin{aligned}\mathbb{P}(X \geq 2n-t) &\geq \mathbb{P}\left(X - \mathbb{E}(X) \geq -\varphi(n)\sqrt{n} - \frac{Ag}{1+2g}\varphi(n)\sqrt{n} + O(\log n)\right) \\ &\geq \mathbb{P}(X - \mathbb{E}(X) \geq -B\varphi(n)\sqrt{n} + O(\log n)) \\ &\geq \mathbb{P}(|X - \mathbb{E}(X)| \leq B\varphi(n)\sqrt{n} + O(\log n)) \\ &\geq 1 - 2\exp\left\{-\frac{(B\varphi(n)\sqrt{n}) + O(\log n)^2}{8t}\right\} \\ &\to 0, \end{aligned} \tag{16}$$

by Azuma's inequality as given by Eq. (Az).

Conversely, for $t = gn - \varphi(n)\sqrt{n}$, we have for some $C > 0$,

$$\begin{aligned}\mathbb{P}(X \geq 2n-t) &\leq \mathbb{P}\left(X - \mathbb{E}(X) \geq (2-g)n + \varphi(n)\sqrt{n} - \frac{n(gn-\varphi(n)\sqrt{n})}{n+2(gn-\varphi(n)\sqrt{n})}(1+\alpha_n)\right) \\ &= \mathbb{P}\left(X - \mathbb{E}(X) \geq (2-g)n + \varphi(n)\sqrt{n} - \frac{gn^2\left(1-\frac{\varphi(n)}{g\sqrt{n}}\right)(1+\alpha_n)}{(1+2g)n\left(1-\frac{2\varphi(n)}{(1+2g)\sqrt{n}}\right)}\right) \\ &\leq \mathbb{P}\left(X - \mathbb{E}(X) \geq (2-g)n + \varphi(n)\sqrt{n} - \frac{ng}{1+2g}\left(1 - \frac{C\varphi(n)}{\sqrt{n}}\right)(1+\alpha_n)\right) \\ &= \mathbb{P}(X - \mathbb{E}(X) \geq D\varphi(n)\sqrt{n} + O(1)) \\ &\leq \mathbb{P}(|X - \mathbb{E}(X)| \geq D\varphi(n)\sqrt{n} + O(1)) \\ &\to 0, \end{aligned} \tag{17}$$

again by (Az). Eqs. (16) and (17) complete the proof. □

5. NP-completeness of the cover pebbling problem

One of the obvious open problems that can be formulated as a result of our work is the following: What are cover pebbling thresholds for families of graphs other than K_n? It would certainly advance the theory of cover pebbling if one could uncover a host of results similar, e.g. to Theorems 2 and 6. Such results would provide a nice complement to those in [8]. Our results in this section show, however, that this task might not be as easy as one might imagine. Necessary and sufficient conditions for the cover solvability of a graph are likely to be complicated, and the best hope might thus be to establish necessary conditions and sufficient conditions that are not too far apart.

The normal way to formalize the concept of the difficulty of a problem is to use the concept of computational complexity. Formally, we imagine a *decision problem* to be a set of infinite strings of characters (like data represented by bits in a computer.) A decision problem is said to *accept* a string if this set contains the string. Usually, we look for the best possible asymptotic upper bound (in terms of the length of the string) for the number of steps the fastest possible algorithm takes to determine whether a given string is in the set. Informally, we think of decision problems being yes-no questions about a property of some class of finite mathematical structures (graphs, integer matrices, etc.), and we ask how fast it is possible to correctly determine the yes or no answer in terms of the size of the input.

For instance, some problems can be solved by an algorithm which takes only a number of steps which is bounded by a polynomial in the size of the input, while others take at least an exponential amount of time to solve. The former class of decision problems is called P for "polynomial." The class NP, for "nondeterministic polynomial" is a bit more complicated; roughly speaking it is the set of decision problems for which a "yes" answer can be "checked" in polynomial time, given an appropriate piece of information. That is, if we call the class of inputs to the decision problem X, and the class of inputs which the decision problem accepts X', there exists a class Y of objects (called the *certificates*) and a function $A : X \times Y \to \{0, 1\}$ which is computable in polynomial time, such that for any instance $x \in X$ of the decision problem, there exists a $y \in Y$ such that $A(x, y) = 1$ if and only if $x \in X'$. For instance, the decision problem which asks whether a given number is composite is easily seen to be in NP, because the composite numbers are exactly those with nontrivial divisor, and, given two numbers, it is easy to determine by division whether one is a divisor of the other. Also, any problem in P is also in NP, because any polynomial-time method of solving a problem is trivially also a polynomial-time method of verifying a yes answer. However, it is a celebrated open problem if the converse also holds and $P = NP$.

Within NP, there is a class of problems, called the NP-*complete* problems, which is thought of being a set of problems which are at least as hard as any other problem in NP. This is because any instance x of a problem D in NP can be translated by a polynomial-time algorithm to an instance x' of any NP-complete problem D' such that x' is accepted by D' if and only if x is accepted by D. Therefore, if we could solve any NP-complete problem in polynomial time, we could solve any problem in NP in polynomial time by translating it to an instance of this problem. Thus, the question of whether $P = NP$ reduces to the question of whether any particular NP-complete problem can be solved in polynomial time.

We now show that the problem which asks if a configuration of pebbles on a graph is cover solvable is NP-complete. It is worth noting that most complexity theorists speculate that $P \neq NP$, and therefore, when a problem is classified as NP-complete, it is usually thought of as evidence of its difficulty. See [12] for a comprehensive theory of NP-completeness, and Watson [24] and Milans et al. [20] for a more general exposition on the complexity of cover pebbling.

Theorem 7. *Let G be a graph, and C a configuration on G. Label the vertices of G $v_1, v_2, \ldots v_n$. Then C is cover solvable if and only if there exist integers $a_{ij} \geq 0$ with $1 \leq i, j \leq m$ and $a_{ij} = 0$ and $a_{ji} = 0$ whenever $\{v_i, v_j\} \notin E(G)$ such that for all $1 \leq k \leq m$,*

$$C(v_k) + \sum_{l=1}^{m} a_{lk} - 2\sum_{l=1}^{m} a_{kl} \geq 1.$$

Proof. First, suppose C is cover solvable. Then find some sequence of pebbling moves which cover solves C. Let a_{ij} be the total number of pebbling moves from v_i to v_j in this sequence. Then after all the moves, there are exactly

$$C(v_k)+\sum_{l=1}^{m}a_{lk}-2\sum_{l=1}^{m}a_{kl}$$

pebbles left on v_k, which is always at least 1 because of the fact that this sequence of moves cover solves C.

Conversely, suppose such numbers a_{ij} exist. Then if we can execute a sequence of pebbling moves starting at C for which the total number of moves from v_i to v_j is exactly a_{ij} for all i and j, the sequence will cover solve C because it will leave vertex $C(v_k)$ with $C(v_k)+\sum_{l=1}^{m}a_{lk}-2\sum_{l=1}^{m}a_{kl}$ pebbles.

So we imagine the a_{ij} give us a list of pebbling moves, and to show that C is solvable, we must show it is possible to execute all the moves on the list in some order without ever leaving any vertex with negative pebbles at an intermediate stage.

It turns out that no strategic foresight is needed to successfully execute the moves; we may proceed by haphazardly making moves from the list, with only the restriction that we only make moves that are currently legal, that is, which start from a vertex with at least two pebbles. For suppose we proceed in this manner but are forced to come to a standstill before exhausting the list of moves because every remaining move on our list (if any) involves moving pebbles from a vertex that now has less than 2 pebbles. Then let C' be this configuration of pebbles, let b_{ij} denote the number of moves we have thus far made from v_i to v_j, and let $c_{ij}=a_{ij}-b_{ij}$

Let S denote the set of vertices from which we must still move pebbles. That is, S is the set of vertices v_k for which $c_{kj}\neq 0$ for some j (abusing notation, S will also denote the set of indices of these vertices). Since we have assumed it is not possible to execute any more of the moves on our list, $C'(v_k)\leq 1$ for each vertex in S. For any vertex v_k, we can easily compute that

$$C'(v_k)=C(v_k)+\sum_{l=1}^{m}b_{lk}-2\sum_{l=1}^{m}b_{kl}.$$

Also, since $b_{ij}+c_{ij}=a_{ij}$ we can use this formula to see

$$1\leq C(v_k)+\sum_{l=1}^{m}a_{lk}-2\sum_{l=1}^{m}a_{kl}=C'(v_k)+\sum_{l=1}^{m}c_{lk}-2\sum_{l=1}^{m}c_{kl}.$$

We simplify both sides of the inequality over the vertices of S to get

$$|S|\leq\sum_{k\in S}C'(v_k)+\sum_{k\in S}\sum_{l=1}^{m}c_{lk}-2\sum_{k\in S}\sum_{l=1}^{m}c_{kl}.$$

Using $C'(v_k)\leq 1$ for $k\in S$ and that $c_{lk}=0$ when $;\notin S$, we get

$$|S|\leq|S|+\sum_{k\in S}\sum_{l\in S}c_{lk}-2\sum_{k\in S}\sum_{l=1}^{m}c_{kl}.$$

Combining terms which occur in both summations we get

$$0\leq-\sum_{k\in S}\sum_{l\in S}c_{kl}-2\sum_{k\in S}\sum_{l\notin S}c_{kl}$$

which implies that $c_{kl}=0$ for all $k\in S$ and therefore for all $1\leq k\leq n$ and so $b_{ij}=a_{ij}$ for all i and j. Therefore $C'(v_k)\geq 1$ for all k, so we have solved C. □

Corollary 8. *The cover solvability decision problem which accepts pairs* (G,C) *if and only if* G *is a graph and* C *is a configuration which is cover solvable on* G *is in* NP.

Proof. The above theorem gives the appropriate certificate of cover solvability: any list of integers a_{ij} which satisfy the equation in Theorem 7. Indeed, Theorem 7 shows that cover pebbling is equivalent to a special case of the NP-complete problem of integer programming, which asks, given an $n\times m$ integer matrix A and an n-dimensional integer vector b if there exists an m-dimensional integer vector x such that $Ax\geq b$ holds componentwise. Having reduced cover solvability to a special case of this NP problem, we know that cover solvability is also in NP.

We now pause to point out another corollary which will be needed later and which is interesting in its own right:

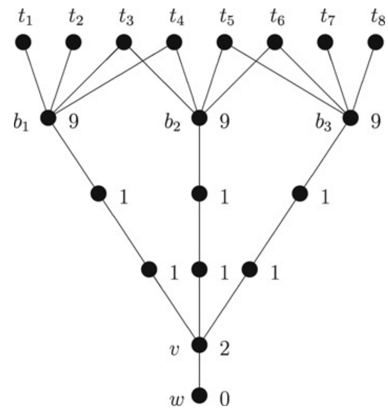

Fig. 1. A graph that corresponds to the exact cover by four 4-sets problem, $a_1 = \{s_1, s_2, s_3, s_4\}$, $a_2 = \{s_3, s_4, s_5, s_6\}$, $a_3 = \{s_5, s_6, s_7, s_8\}$.

Corollary 9. *Let G be a graph, C a configuration on G. If the sequence of pebbling moves $Q = (q_1, q_2 \dots q_k)$ cover solves C, and it is possible to make the sequence $Q' = (q_{i_1}, q_{i_2}, \dots q_{i_l})$ of moves (with $1 \le i_j \le k$ for all j but with no particular requirement on the order of the i_j,) then the configuration C' obtained from C after the moves $(q_{i_1}, q_{i_2}, \dots q_{i_l})$ is cover solvable.*

Proof. The point of this corollary is that the order of our pebbling moves can't matter. To show this, we simply note that if it were possible to somehow execute the remaining moves from Q which are not in Q', they would solve C'. By Theorem 7, it is thus possible to solve C'. □

Now we turn our attention to showing that the cover solvability decision problem is NP-hard, that is, that any instance of any problem in NP can be translated to an instance of cover solvability in polynomial time. The usual method of showing that a problem A is NP-hard is to find an NP-complete problem B for which any instance of B can be translated into an instance of A in polynomial time. Then for any instance of any problem in NP, we can translate it in polynomial time to an instance of B, and then translate this instance into an instance of A. For cover solvability, we will use a known NP-complete problem known as "exact cover by 4-sets." Indeed, the corresponding and seemingly simpler problem of perfect cover by 3-sets is also NP-complete, but for our purposes, the 4-set problem is more useful.

Theorem 10 (*Karp [18]*). *Let the exact cover by 4-sets problem be the decision problem which takes as input a set S with $4n$ elements and a class A of at least n 4-element subsets of S, accepting such a pair if there exists an $A' \subseteq A$ such that A' is a class of disjoint subsets which make a partition of S, that is they are n subsets containing every element of S. This problem is NP-complete.*

Theorem 11. *The cover solvability decision problem is NP-complete*

Proof. Having shown that this decision problem is in NP, it remains to be shown that it is NP-hard. Given a set $S = \{s_1, s_2, \dots s_{4n}\}$ and a class $A = \{a_1, a_2, \dots a_m\}$ of four element subsets of S, that is, an instance of the exact cover by 4-sets problem, we construct a graph G' and a configuration C' on G' in the following manner: We create a set of vertices $T = \{t_1, t_2, \dots t_{4n}\}$ which will be thought of as corresponding to the elements of S, and a set of vertices $B = \{b_1, b_2, \dots b_m\}$ which will be thought of as corresponding to the members of A. Let $C'(t) = 0$ for all $t \in T$ and let $C'(b) = 9$ for all $b \in B$. We make edges between B and T in the intuitive way, including $\{b_i, t_j\}$ if $t_j \in b_i$. Additionally, create a vertex v and a path of length $m - n$ which has one terminal vertex v and the other called w. Note that $w = v$ if $m = n$. Let $C'(v) = 2^{m-n} - (m-n) + 1$, $C'(w) = 0$, and $C'(u) = 1$ for all u between v and w on the path. Finally, create vertex classes $B' = \{b'_1, b'_2 \dots b'_m\}$ and $B'' = \{b''_1, b''_2 \dots b''_m\}$, creating edges $\{b_i, b'_i\}$, $\{b'_i, b''_i\}$ and $\{b''_i, v\}$ for all i. Let $C'(u) = 1$ for all $u \in B' \cup B''$ (Fig. 1).

Clearly, this construction can be made in polynomial time in the size of the pair $\{S, A\}$. Indeed, we have made only $5|A| + 1$ vertices and $8|A| - |S|$ edges. In order to finish the proof, we claim that C' is solvable if and only if A contains a perfect cover of S.

First suppose that A contains a perfect cover $A' = \{a_{i_1}, a_{i_2}, \dots a_{i_n}\}$ of S. Then for each vertex in B which is a b_{i_j} for some $1 \le j \le n$, we use 8 of the pebbles on this b_{i_j}, two each to cover the four vertices of T to which it is adjacent. Because of the fact that A' is a perfect cover and the way we constructed G', we now have exactly one pebble on every vertex of T. Furthermore, we have $m-n$ vertices in B which still have 9 pebbles each on them. Because v is at distance 3 from each of these vertices, we can use 8 pebbles from each of these vertices to move one pebble each onto v from these $m-n$ vertices. This leaves $2^{m-n}+1$ pebbles on v, which is enough to move one pebble onto w while leaving one pebble on v. After this is done, we have exactly one pebble on every vertex of G', and we thereby know that C' is solvable.

To show the converse, suppose that A does not contain a perfect cover of S. Suppose as well that C' is solvable on G'. Clearly, the sequence of pebbling moves which solves C' must contain (at least) one move to t for every $t \in T$. Clearly, each of these moves must originate from B, and no more than 4 can originate from any one vertex of B. Since A does not contain a perfect cover of S, it cannot be the case that these moves originate from exactly n vertices in B.

We make these $4n$ moves immediately from C', using Corollary 9 to see that the resulting configuration must be solvable (we use the fact that no more than 4 of these moves can originate from any one vertex in B to see that it is indeed possible to make these moves). In addition to the one pebble left on every vertex of B to ensure they remain covered, there are now $8(m-n)$ pebbles on B, but they are not in $m-n$ groups of 8 pebbles because the moves we made originated from more than n vertices of B. In order to reach w, we clearly need to move $m-n$ pebbles onto v while leaving the rest of the graph covered. Clearly, this is only possible if all $8(m-n)$ extra pebbles are moved by a path of length 3 onto v. However, only one such path is available for any group of these pebbles. But any group of less than 8 pebbles cannot increase the number of pebbles on v by moving along this path, while leaving the vertices of the path covered. Since there are not indeed $m-n$ groups of 8 pebbles on B, we see that it is impossible to gather the pebbles necessary to reach w, and thus our configuration is not solvable, which is a contradiction.

Acknowledgements

This work was done under the supervision of Anant Godbole at the East Tennessee State University Research Experience for Undergraduates (REU) site during the summer of 2004, while Watson and Yerger were undergraduate students at Washingon University in St. Louis, and Harvey Mudd College respectively. Support from NSF Grant DMS-0139291 is gratefully acknowledged by all three authors, as is the support offered by fellow REU participants — "pebblers" and "non-pebblers" alike.

References

[1] A. Bekmetjev, G. Brightwell, A. Czygrinow, G. Hurlbert, Thresholds for families of multisets, with an application to graph pebbling, Discrete Math. 269 (2003) 21–34.
[2] B. Bukh, Maximum pebbling number of graphs of diameter three, J. Graph Theory 52 (2006) 353–357.
[3] M. Chan, A. Godbole, Improved pebbling bounds, Discrete Math. (2005). Preprint (in press).
[4] F. Chung, Pebbling in hypercubes, SIAM J. Discrete Math. 2 (1989) 467–472.
[5] F. Chung, R. Graham, J. Morrison, A. Odlyzko, Pebbling a chessboard, Amer. Math. Monthly 102 (1995) 113–123.
[6] T. Clarke, R. Hochberg, G. Hurlbert, Pebbling in diameter 2 graphs and products of paths, J. Graph Theory 25 (1997) 119–128.
[7] B. Crull, T. Cundiff, P. Feltman, G. Hurlbert, L. Pudwell, Z. Szaniszlo, Z. Tuza, The cover pebbling number of graphs, Discrete Math. 296 (2005) 15–23.
[8] A. Czygrinow, N. Eaton, G. Hurlbert, P.M. Kayll, On pebbling thresholds for graph sequences, Discrete Math. 247 (2002) 93–105.
[9] A. Czygrinow, G. Hurlbert, On the pebbling threshold of paths and the pebbling threshold spectrum, Discrete Math. (2007) (in press).
[10] A. Czygrinow, G. Hurlbert, H. Kierstead, W.T. Trotter, A note on graph pebbling, Graphs Combin. 18 (2002) 219–225.
[11] W. Feller, An Introduction to Probability Theory and its Applications, 3rd edition, vol. 1, John Wiley, New York, 1968.
[12] M.R. Garey, D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, W. H. Freeman, New York, 1983.
[13] G. Helleloid, M. Khalid, P. Matchett, D. Moulton, Graph pegging numbers, 2005. Preprint.
[14] G. Hurlbert, A survey of graph pebbling, Congr. Numer. 139 (1999) 41–64.
[15] G. Hurlbert, Recent progress in graph pebbling, Graph Theory Notes of New York 49 (2005) 25–37.
[16] G. Hurlbert, H. Kierstead, On the complexity of graph pebbling, 2005. Preprint.
[17] G. Hurlbert, B. Munyan, Cover pebbling hypercubes, Bull. Inst. Combin. Appl. 47 (2006) 71–76.
[18] R.M. Karp, Reducibility among combinatorial problems, in: Complexity of Computer Computations, in: Proc. Sympos. IBM Thomas J. Watson Res. Center, Yorktown Heights, NY, New Plenum, New York, 1972, pp. 85–103.
[19] D. Kleitman, P. Lemke, An addition theorem on the integers modulo n, J. Number Theory 31 (1989) 335–345.
[20] K. Milans, B. Clark, The complexity of graph pebbling, SIAM J. Discrete Math. 20 (2006) 769–798.
[21] J. Sjöstrand, The cover pebbling theorem, Electron. J. Combin. 12 (2005) #N22.
[22] J.M. Steele, Probability Theory and Combinatorial Optimization, in: NSF-CBMS Regional Research Conference Lecture Notes Series, vol. 69, Society for Industrial and Applied Mathematics, Philadelphia, 1997.

[23] A. Vuong, I. Wyckoff, Conditions for weighted cover pebbling of graphs, 2005. http://arxiv.org/abs/math/0410410 Manuscript.
[24] N. Watson, The complexity of pebbling and cover pebbling, 2005. http://www.arxiv.org/abs/math.CO/0503511 Preprint.
[25] N. Watson, C. Yerger, Cover pebbling numbers and bounds for certain families of graphs, Bull. Inst. Combin. Appl. 48 (2006) 53–62.
[26] N. Watson, C. Yerger, Cover pebbling thresholds for the complete graph, 2004. Manuscript.
[27] A. Wierman, J. Salzman, M. Jablonski, A. Godbole, An improved upper bound for the pebbling threshold of the n-path, Discrete Math. 275 (2004) 367–373.