

# **Threshold and Complexity Results for the Cover Pebbling Game**

Anant P. Godbole (East Tennessee State University),

Nathaniel G. Watson (Washington University, St. Louis),

and

Carl R. Yerger (Harvey Mudd College).

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 Feller, 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 Czygrinow et al. 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. Time does not permit us to share the underlying additive number theory connection

**SPECIAL CASES:** The pebbling number  $\pi(P_n)$  of the path is  $2^{n-1}$ . Chung proved that  $\pi(Q^d) = 2^d$  and  $\pi(P_n^m) = 2^{(n-1)m}$ . An easy pigeonhole principle argument yields  $\pi(K_n) = n$ . The pebbling number of trees has been determined (see Hurlbert).

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.**

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

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  $\pi(G) = |G|$ . Cubes are of Class 0, as are complete graphs, but what other families fall in this important class of graphs for which  $\pi$  is as low as it can possibly be? Here are two answers: For graphs of diameter 2, if  $G$  is 3-connected, then  $G$  is Class 0 (Clark, Hurlbert and Hochberg). In fact, they show that 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. Generalizations of this result to the case where  $p = p_n \rightarrow 0$  as  $n \rightarrow \infty$  are also available. Other authors, e.g. Chan and G, have obtained general pebbling bounds, while Bukh has proved almost-tight asymptotic bounds on the pebbling number of diameter three graphs.

In *Random Pebbling*, we seek the probability that  $G$  is pebbleable when  $t$  pebbles are placed randomly on it according to the B-E or M-B scheme. Numerous *threshold results* have been determined in Czygrinow et al. for B-E 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:

$$t = t_n \gg a_n \Rightarrow \mathbb{P}(G = G_n \text{ is pebbleable}) \rightarrow 1 \quad (n \rightarrow \infty)$$

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

Of course, we 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 (Czygrinow et al.) that  $a_n = b_n = \sqrt{n}$ . In many cases, however, the analysis is quite delicate; see Wierman et al. for some of the issues involved in finding the pebbling threshold for a family as basic as  $P_n$ , the path on  $n$  vertices.

A detailed survey of graph pebbling has been presented by Hurlbert, and it would probably not be an oversimplification to state that most results available to date fall in four broad categories: finding pebbling numbers for classes of graphs; 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 (Hurlbert and Kierstead). A survey of some not-open-anymore problems in graph pebbling may be found on Glenn Hurlbert's website; see <http://math.la.asu.edu/~hurlbert/HurlPebb.ppt>

. 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. For reasons that will become obvious, we focus only on analogs of the last two of the four general directions mentioned above.

The *cover pebbling number*  $\lambda(G)$  is defined as the minimum number of pebbles required such that it is possible, given any initial configuration of at least  $\lambda(G)$  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 (Crull et al.) that  $\lambda(K_n) = 2n - 1$ ;  $\lambda(P_n) = 2^n - 1$ ; and that for trees  $T_n$ ,

$$\lambda(T_n) = \max_{v \in V(T_n)} \sum_{u \in V(T_n)} 2^{\text{dist}(u,v)}. \quad (1)$$

Likewise, it was shown by Hurlbert and Munyon that  $\lambda(Q^d) = 3^d$  and by Watson and Yerger that  $\lambda(K_{r_1, \dots, r_m}) = 4r_1 + 2r_2 + \dots + 2r_m - 3$ , where  $r_1 \geq r_2 \geq \dots \geq r_m$ .

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  $T_n$  by Crull et al. led to their Open Question No. 10, 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 (paper submitted to EJC) were able to prove the

**STACKING THEOREM:** *For any connected graph  $G$ ,*

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

thereby proving that (1) holds for all graphs. In fact, the key result of Vuong and Wyckoff 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 fifth general direction, namely a study of which distributions are (cover) solvable and which are not. Finally, we mention that it is of great interest that many of the 60+ authors who have contributed to the theory of pebbling and cover pebbling are undergraduates. The 2004 Mathfest featured talks by Aparna Higgins and Zsuzsanna Szaniszlo that focused on contributions by undergraduates. Students at the Central Michigan University, East Tennessee State University, and University of Minnesota at Duluth REU sites, together with students at Valparaiso University and Arizona State University, have been among the key undergraduate contributors.

**Maxwell Boltzmann Cover Pebbling Threshold for  $K_n$**  Clearly  $n$  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  $\lambda(K_n) = 2n - 1$ . 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 M-B scheme.

Let  $X = X_{n,t}$  be the number of vertices on which an odd number of pebbles are placed (=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*

$$X + t \geq 2n. \quad (2)$$

Armed with Lemma 1, we now provide the heuristic reason why we believe there is a sharp cover pebbling threshold at  $t = (1.5238\dots)n$ . Our rationale is standard but likely to be somewhat surprising to the uninitiated. Given a random variable  $X$  with expected value  $\mathbb{E}(X)$ , we will say that  $X$  is *sharply concentrated* around  $\mathbb{E}(X)$  if  $X \sim \mathbb{E}(X)$  with high probability (w.h.p.)

Assuming therefore that  $X \sim \mathbb{E}(X)$  w.h.p., it makes sense to speculate that  $K_n$  is pebbleable 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

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

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

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

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

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

Since  $(1 - 2/n)^n \sim e^{-2}$  we see from (5) that a reasonable guess for a threshold value of  $t$  is  $A_0n$  where  $A_0$  is the solution of

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

or  $A_0 = 1.5238\dots\dots$

**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. Here, however, we establish our main result by using a baseline technique, Tchebychev's inequality. Azuma will be used later.

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \lambda) \leq \frac{\mathbb{V}(X)}{\lambda^2}.$$

We have, after much work,

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

We are now ready to state

**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  $A_0 = 1.5238\dots$ . Then*

$$t = A_0 n + \varphi(n)\sqrt{n} \Rightarrow$$

$$\mathbb{P}(K_n \text{ is cover pebbleable}) \rightarrow 1 \quad (n \rightarrow \infty)$$

and

$$t = A_0 n - \varphi(n)\sqrt{n} \Rightarrow$$

$$\mathbb{P}(K_n \text{ is cover pebbleable}) \rightarrow 0 \quad (n \rightarrow \infty),$$

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

**Poissonization, Exchangeability, and the Quest for a Central Limit Theorem at the Threshold** Theorem 2 raises a question. What is the probability of being able to successfully cover pebble  $K_n$  at the threshold? Specifically, if  $t = A_0 n + B\sqrt{n}$  for a constant  $B$  then what is  $\mathbb{P}(X \geq 2n - t)$ ?

One of the standard methods used in simplifying problems in discrete probability is that of *Poissonization*. Given a random structure that depends on an input of fixed size, say  $t$ , we study a related model in which the value of the input is random, specifically of size  $\text{Po}(t)$ , where  $\text{Po}(\lambda)$  represents the Poisson random variable with parameter  $\lambda$ . In our case, we will assign a  $\text{Po}(t)$  number of pebbles to the  $n$  vertices of  $K_n$ , hoping that our results will reveal something meaningful about the situation when the *actual* (as opposed to *expected*) number of pebbles is  $t$ . Returning to the fixed size (or Bernoulli) model from the Poissonized model is termed *dePoissonization*.

Specifically, we are interested in computing the quantity

$$q(\theta) = \sum_s p(s) \frac{e^{-\theta} \theta^s}{s!},$$

where  $p(s)$  is the “real” object of interest (in our case, the probability of cover pebbleability when  $s$  pebbles are used). We hope, moreover, in the spirit of Tauberian theorems, to recover the  $p(s)$ s from their averages. Poissonization has been studied in the context of “balls in boxes” models by Aldous and by Holst, and a heuristic is provided by Aldous that permits one to dePoissonize:

A major study of analytic dePoissonization has been conducted by Jacquet and Szpankowski, and a serious application of their work is provided by Janson. The key result we use requires that we traverse to the complex domain to search for rigorous sufficient conditions that permit dePoissonization:

**Theorem 3** (*Jacquet and Szpankowski*) Assume that the complex function  $q(z) = \sum_s p(s) \frac{e^{-z} z^s}{s!}$  defined as in (9) is an entire function, and that in some linear cone  $\mathcal{S}_\theta = \{z : -\theta \leq \arg(z) \leq \theta < \pi/2\}$  and for some  $A, B, R > 0; \alpha < 1; \beta > 1$  we have

- For  $z \in \mathcal{S}_\theta, |z| > R \Rightarrow |q(z)| \leq B|z|^\beta;$
- For  $z \notin \mathcal{S}_\theta, |z| > R \Rightarrow |q(z)e^z| \leq A \exp\{\alpha|z|\}.$

Then  $p(t) = q(t) + O(t^{\beta-1}), t \rightarrow \infty.$

**Theorem 4** Consider  $t$  distinct pebbles that are thrown onto  $K_n$  according to the M-B distribution. Set  $A_0 = 1.5238 \dots$ . Then  $\forall B \in \mathbf{R},$

$$t = A_0 n + B\sqrt{n} \Rightarrow$$

$$\mathbb{P}(K_n \text{ is cover pebbleable}) \rightarrow \int_C^\infty \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

$n \rightarrow \infty,$  where  $C$  is a well specified function of  $B.$

## Bose Einstein Cover Pebbling Threshold

In the Maxwell Boltzmann scheme, it is practically impossible to calculate  $\mathbb{P}(X = x)$  exactly. 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 5** *If  $t$  and  $x$  have the same parity, and if  $x \leq \min\{t, n\}$ , then*

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

## **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  $n$  colors, with the understanding that the number of balls of color  $j$  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 as a means of modeling contagious diseases and takes place as follows: Initially the urn contains one ball of each of  $n$  colors. After each draw, the selected ball is replaced together with another ball of the same color. 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 color  $j$  is drawn can be set to equal the number of pebbles on vertex  $j$ , but do these two procedures yield the same probability model? Yes!!

**Lemma 6** (*Azuma-Hoeffding*) For all  $\lambda > 0$ ,

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \lambda) \leq 2 \exp \left\{ -\frac{\lambda^2}{2 \sum \|d_i\|_\infty^2} \right\}. \quad (6)$$

where (in our case)  $\|d_i\|_\infty$  is bounded above by the worst case scenario change in  $X$  when the  $i$ th Polya draw is “redone.” We thus have  $\|d_i\|_\infty \leq 2$ . We thus get the following concentration for the number  $X$  of odd stacks in the Bose-Einstein scheme:

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

so that  $X$  is concentrated in an interval of length  $\sqrt{n}\varphi(n)$  around  $\mathbb{E}(X)$  whenever  $t \sim Kn$ . We are now ready to prove the main result of this section:

**Theorem 7** Consider  $t$  indistinguishable pebbles that are placed on the vertices of the complete graph  $K_n$  according to the Bose Einstein distribution. Then, with  $\gamma$  representing the golden ratio  $(1 + \sqrt{5})/2$ ,

$$t = \gamma n + \varphi(n)\sqrt{n} \Rightarrow$$

$$\mathbb{P}(K_n \text{ is cover pebbleable}) \rightarrow 1 \quad (n \rightarrow \infty)$$

and

$$t = \gamma n - \varphi(n)\sqrt{n} \Rightarrow$$

$$\mathbb{P}(K_n \text{ is cover pebbleable}) \rightarrow 0 \quad (n \rightarrow \infty),$$

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

## **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 10. Such results would provide a nice complement to those by Hurlbert and Kierstead. 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 pebbleability 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.

**Theorem 8** *Let  $G$  be a graph,  $C$  a configuration on  $G$ . Let  $|G| = m$  and label the vertices of  $G$  as  $v_1, v_2, \dots, v_m$ . Then  $C$  is cover solvable if and only if there exist integers  $n_{ij} \geq 0$  with  $1 \leq i, j \leq m$  and  $n_{ij} = 0$  and  $n_{ji} = 0$  whenever  $\{v_i, v_j\} \notin E(G)$  such that for all  $0 \leq k \leq m$ ,*

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

**Corollary 9** *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.*

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  $B$ , 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** *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  $n$  four-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*