Strong emergence of wave patterns on Kadanoff sandpiles*

Kévin Perrot
Aix Marseille Université
LIF UMR CNRS 7279
13288 Marseille, France
kevin.perrot@lif.univ-mrs.fr

Éric Rémila
Université de Lyon
GATE LSE UMR CNRS 5824
42023 St-Etienne, France
eric.remila@univ-st-etienne.fr

Submitted: Oct 7, 2015; Accepted: Mar 28, 2017; Published: Apr 13, 2017
Mathematics Subject Classifications: 68R05, 74H05, 76T25

Abstract

Emergence is easy to exhibit, but very hard to formally explain. This paper deals with square sand grains moving around on nicely stacked columns in one dimension (the physical sandpile is two dimensional, but the support of sand columns is one dimensional). The Kadanoff sandpile model is a discrete dynamical system describing the evolution of finitely many sand grains falling from an hourglass (or equivalently from a finite stack of sand grains) to a stable configuration. The repeated application of a simple local rule let grains move until reaching a fixed point. The difficulty of understanding its behavior, despite the simplicity of its rule, is the main interest of the model. In this paper we prove the emergence of exact wave patterns periodically repeated on fixed points. Remarkably, those regular patterns do not cover the entire fixed point, but eventually *emerge* from a seemingly disordered segment: grains are added on the left, triggering avalanches that become regular as they fall down the sandpile. The proof technique we set up associated arguments of linear algebra and combinatorics, which interestingly allow to formally demonstrate the emergence of regular patterns without requiring a precise understanding of the non-regular initial segment's dynamic.

1 Introduction

Understanding and explaining regularity properties on discrete dynamical systems (DDS) rapidly becomes a puzzling problem, and formally proving the global behavior of a DDS

*An extended abstract [32] of this work has been presented at the conference MFCS'15. This work is partially supported by IXXI (Complex System Institute, Lyon), ANR projects Subtile, Dynamite and QuasiCool (ANR-12-JS02-011-01), Modmad Federation of U. St-Etienne, Fondecyt Postdoctoral Project 3140527, and Núcleo Milenio Información y Coordinación en Redes (ACGO).

defined with local rules is at the heart of our comprehension of natural phenomena [21, 35]. A lot of simply stated conjectures, often issued from simulations, remain open (among the famous examples is the Langton's Ant [14, 15]). Sandpile models are a class of DDS describing how grains move according to local rules in discrete space and time. We start from a finite number of grains stacked on a single column (the analogy with an hourglass is formally presented in Subsection 2.4), and try to predict the asymptotic shape of the stable configurations reached at the end of the dynamical evolution.

Bak, Tang and Wiesenfeld introduced sandpile models as simple examples of systems presenting *self-organized criticality* (SOC), a property of dynamical systems having critical points as attractors [1]. Informally, they considered a discretized flat surface on which grains are added one by one. Each addition possibly triggers an avalanche, consisting in grains falling from column to column according to one elementary local rule, and after a while a heap of sand has formed. SOC is related to the fact that a single grain addition on a stabilized sandpile has a hardly predictable consequence on the system, on which surprising fractal structures appear [5]. This model can naturally be extended to any number of dimensions.

A one-dimensional sandpile configuration can be represented as a sequence $(h_i)_{i \in \mathbb{N}}$ of non-negative integers, h_i being the number of sand grains stacked on column i. The initial configuration is composed of a finite number of stacked grains on column 0: the configuration h where $h_0 = N$ and $h_i = 0$ for $i > 0$, and in the classical sandpile model one grain can fall from column i to column $i+1$ if and only if the height difference $h_i - h_{i+1} > 1$. Different update policies may be applied: sequential, parallel, block-sequential, *etc.* One-dimensional sandpile models were well studied in recent years [8, 10, 12, 17, 18, 29, 34].

In this paper, we study the Kadanoff sandpile model (KSM) which generalizes classical sandpile models. A fixed parameter, denoted p, indicates the number of grains falling at each rule application. The results we develop in this paper present an interesting feature: we asymptotically completely describe the form of stable configurations: they are made out of exact wave patterns, though there is a part of asymptotically null relative size (and asymptotically infinite absolute size) on the left of fixed points, apparently complicated and non-regular, which remains unexplained. Furthermore, proven regularities are directly issued from this unordered initial segment, which precise understanding is bypassed by the proof technic. Indeed, the fixed point can be reached by adding grains one by one, each grain addition triggering an avalanche letting grains fall down the sandpile until they eventually create and maintain very regular wave shapes. This point is discuted in the conclusion.

We formally define the model in Subsection 1.1, and Subsection 1.2 introduces the results we expound in the rest of the paper. We present useful folklore in Section 2: in particular an interesting and useful way of computing fixed points in Subsection 2.4. The proof of the main result is presented in Section 3, and at the light of those developments, Section 4 discusses that sandpile models exhibit a behavior at the edge between discrete and continuous phenomena.

1.1 Definition of the Kadanoff sandpile model

In 1989, Kadanoff *et al* proposed a generalization of classical models in which a fixed parameter p denotes the number of grains falling at each step [23]. Starting from the initial configuration composed of N stacked grains on column 0, we iterate the following rule: if the difference of height (the slope) between columns i and $i+1$ is greater than p, then p grains can fall from column i, and one grain reaches each of the p columns $i+1, i+2, \ldots, i+p$. Note that when $p = 1$ we get the classical sandpile model. The rule is applied once (non-deterministically) during each time step. The words *column* and *index* are synonyms, and for the sake of imagery we always consider indices (column numbers) to be increasing on the right as it is presented on Figure 1.

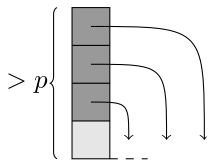


Figure 1: KSM(p) rule. When p grains leave column i, the slope b_{i-1} is increased by p, b_i is decreased by $p+1$ and b_{i+p} is increased by 1. The slope of other columns are not affected.

Formally, the rule is defined on the space of ultimately null decreasing integer sequences, where each integer represents a column of stacked sand grains. A *configuration* is denoted $h = (h_i)_{i \in \mathbb{N}}$, and h_i is the number of grains on column i. It is very convenient to consider only the relative height between columns, so we will mainly represent configurations as sequences of *slopes* $b = (b_i)_{i \in \mathbb{N}}$, where for all $i \geqslant 0$, $b_i = h_i - h_{i+1}$, within the space of ultimately null non-negative integer sequences (also called *partitions* in the litterature). We use 0^ω to denote the infinite sequence of zeros at the end of a configuration.

Let us now give two definitions of the model, obviously isomorphic. Definition 1 is more natural and uses *heights*, but Definition 2 is more convenient and uses *slopes*, this latter is the main one we will use throughout the paper.

Definition 1. The Kadanoff sandpile model with parameter $p > 0$, KSM(p), is defined by two sets:

- *height configurations.* Ultimately null non-negative and decreasing integer sequences.
- *transition rules.* We have a transition from a configuration h to a configuration h' on column i, and we note $h \xrightarrow{i} h'$ when
  - ◦ $h'_i = h_i - p$
  - ◦ $h'_{i+k} = h_{i+k} + 1$ for $0 < k \leqslant p$
  - ◦ $h'_j = h_j$
    for $j \notin \{i, i+1, \ldots, i+p\}$.

In this case we say that i is *fired*. Remark that according to the definition of the transition rules, i may be fired if and only if $h_i - h_{i+1} > p$, otherwise h'_i is negative or the sequence h' is not decreasing.

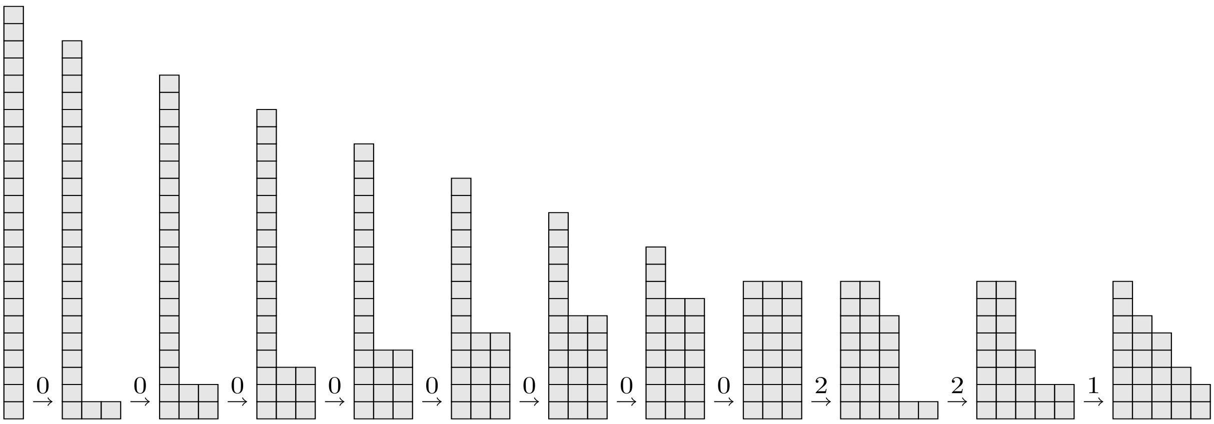


Figure 2: A possible evolution in KSM(2) from the initial configuration for $N = 24$ to $\pi(24)$. $\pi(24) = (2, 1, 2, 1, 2, 0^\omega)$ and its shot vector (defined in Subsection 3.1) is $(8, 1, 2, 0^\omega)$.

Definition 2. The Kadanoff sandpile model with parameter $p > 0$, KSM(p), is defined by two sets:

- *slope configurations.* Ultimately null non-negative integer sequences.
- *transition rules.* We have a transition from a configuration b to a configuration b' on column i, and we note $b \xrightarrow{i} b'$ when
  - $b'_{i-1} = b_{i-1} + p$ (for $i \neq 0$)
  - $b'_i = b_i - (p + 1)$
  - $b'_{i+p} = b_{i+p} + 1$
  - $b'_j = b_j$ for $j \notin \{i - 1, i, i + p\}$.

Again, remark that according to the definition of the transition rules, i may be fired if and only if $b_i > p$, otherwise b'_i is negative.

We denote $b \to b'$ when there exists an integer i such that $b \xrightarrow{i} b'$. The reflexive transitive closure of $\to$ is denoted by $\xrightarrow{*}$, and we say that b' is *reachable* from b when $b \xrightarrow{*} b'$. KSM model has the *diamond property*: if there exists i and j such that $b \xrightarrow{i} b'$ and $b \xrightarrow{j} b''$, then there exists a configuration b''' such that $b' \xrightarrow{j} b'''$ and $b'' \xrightarrow{i} b'''$.

$$\begin{tikzcd} & b \arrow[dl, "i"'] \arrow[dr, "j"] & \\ b' \arrow[dr, dashed, "j"'] & & b'' \arrow[dl, dashed, "i"] \\ & b''' & \end{tikzcd}$$

A configuration b is *stable*, or a *fixed point*, if no transition is possible from b. As a consequence of the diamond property, for each configuration b there exists a unique stable configuration, denoted by $\pi(b)$, such that $b \xrightarrow{*} \pi(b)$. Moreover, for any configuration b' such that $b \xrightarrow{*} b'$, we have $\pi(b') = \pi(b)$ (see [20] for details). Figure 2 pictures an example of evolution. For convenience, we abusively denote by N the initial configuration $(N, 0^\omega)$, such that $\pi(N)$ is the sequence of slopes of the fixed point associated to the initial

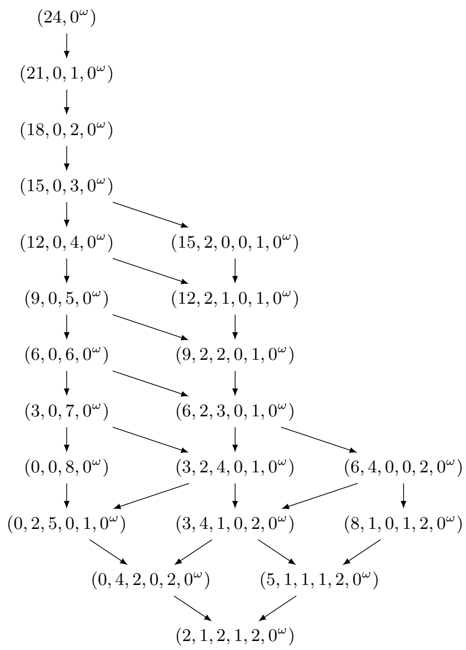


Figure 3: The set of reachable configurations for $p = 2$ and $N = 24$ represented as sequences of slopes. The initial configuration is on the top, and the unique fixed on the bottom.

configuration composed of N stacked grains on column 0 (see Figure 3 for an illustration). This paper is devoted to the study of $\pi(N)$ according to N and the fixed parameter p.

1.2 Objective of the paper

Let us briefly recall classical notations and state the main result of this paper. Regarding regular expressions, let ϵ denote the empty word, and $\cdot$ (respectively $+$) denote the *concatenation* (respectively *or*) operator. $\boldsymbol{+}$ is to the *or* what Σ is to the *sum*[1], and * is the Kleene star denoting finite repetitions of a regular expression (see for example [22] for details). Finally, for a configuration b we denote $b_{[n\,,\infty[}$ the infinite subsequence of b starting from index n to ∞.

In this paper we prove the following precise asymptotic form of fixed points (expressed in terms of slope configuration). Importantly, note that the support of $\pi(N)$ is in $\Theta(\sqrt{N})$, as it is a non-degenerated rectangular triangle whose area is N (details in Subsection 2.2).

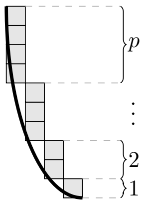
Theorem 3. *For a fixed $p > 1$, the smallest n such that*

$$\pi(N)_{[n\,,\infty[} \in (p \cdot \ldots \cdot 2 \cdot 1)^* \, 0 \, (p \cdot \ldots \cdot 2 \cdot 1)^* \, 0^\omega$$

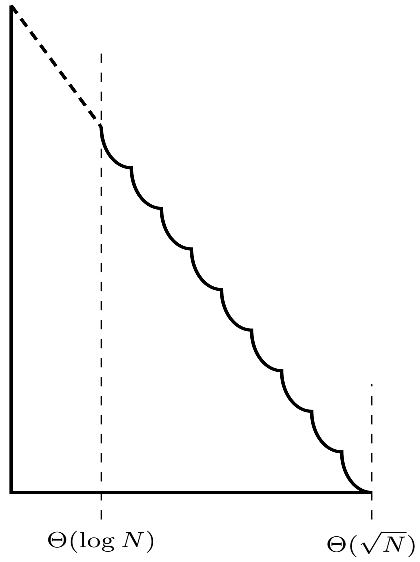
is in $\Theta(\log N)$.

---

[1] $\displaystyle\mathop{\Large +}_{i=0}^{k} 0^i = \epsilon + 0 + 00 + \cdots + \underbrace{0\ldots0}_{k}.$



(a) We call *wave* the pattern $p \cdot \ldots \cdot 2 \cdot 1$ in a sandpile configuration.



(b) From an index in $\Theta(\log N)$ the fixed point $\pi(N)$ consists of waves all consecutive to each other, except at at most one place where two waves may be separated by two columns of same height.

Figure 4: A *wave* (4a), and a graphical representation of Theorem 26 (4b).

A graphical representation of our main result providing an exact bound for wave appearance (Theorem 26) is given on Figure 4. The asymptotic form of fixed points shows a very regular structure made out of wave, *emerging* from a seemingly complex initial segment. Again, since we can see this fixed point as the configuration reached after all the grains have been added one by one on the leftmost column, any single grain part of a wave pattern (on the right part of the fixed point) has travelled through the left part until it adopts, together with the sandpile upper layer it walks on, a regular behavior leading to perfect wave patterns.

The result above and the proof technics exposed thereafter present a promising feature: we asymptotically completely describe the form of stable configurations, though there is a part of asymptotically null relative size (but asymptotically infinite absolute size) which remains unknown. This part, on the left of fixed points, does not seem to have a highly ordered structure. It may look like a drawback that we can't explain any single column of the fixed point, but this remark can also be seen in a reversed way: the proof technic we develop does not require to understand precisely this complex initial segment to formally explain the emergence of very regular patterns periodically repeated on fixed points. As it has been argued above, we think it is natural to qualify the wave patterns as *emergent* regularities; and this paper will explain how to prove that those regularities emerge from a transitional less ordered phase, without requiring a fully accurate understanding of the dynamic.

A series of previous works ([30, 31, 33]) lead us to a similar result, but only for the smallest "new" parameter $p = 2$ (the case $p = 1$ is the well known sandpile model), using exclusively arguments of combinatorics. However, for the general case, we introduce a completely different approach. The proof of the tight $\Theta(\log N)$ bound is decomposed into two parts:

- in a first part (Subsections 3.1, 3.2, 3.3 and 3.4) we prove the $\mathcal{O}(\log N)$ bound, stated in Theorem 22 (upper bound theorem);
- in a second part (Subsection 3.5) we prove the $\Omega(\log N)$ bound, stated in Corollary 25 and leading to the exact bound theorem (Theorem 26).

The main ideas are the following: we establish a relation between different representations of a sandpile configuration (Subsection 3.1), leading to a recurrence relation of the form

$$a_{i+1} = \frac{1}{p}\, a_{i-p} + \frac{p+1}{p}\, a_i + \frac{1}{p}\, \pi(N)_i$$

where a_i is the number of times the column i is fired to reach $\pi(N)$ from $(N, 0^\omega)$.

- During the first step, the term $\pi(N)_i$ is in some sens neglected. The recurrence is rewritten in term of the discrete derivative $y_i = a_i - a_{i-1}$. It appears that the recurrence describes an averaging process on p successive y_i such that in $\mathcal{O}(\log(N))$ iterations the difference between the maximal and minimal value among $(y_{i-p+1}, \ldots, y_i)$ becomes lower than a constant α (Lemma 10). This upper bounds result requires linear algebra and a study of eigenvalues of a related matrix.
- During the second step, the arithmetic properties of the $\pi(N)_i$ are considered to show that in a linear time with respect to the constant α, the recurrence leads to p consecutive equal values $y_{i-p+1} = y_{i-p+2} = \cdots = y_i$ (Lemmas 12 and 13). We say that the sequence $(y_i)_{i\in\mathbb{N}}$ is p-uniform at i.
- The third step requires to identify some regular patterns modulo p in the exact analysis in all possible cases of the arithmetic properties of the $\pi(N)_i$ starting from the step n where $(y_i)_{i\in\mathbb{N}}$ is p-uniform. It appears that either one reads $\pi(N)_i = 0$ and $(y_i)_{i\in\mathbb{N}}$ is also p-uniform at $n+1$ or $\pi(N)_i = p, \pi(N)_{i+1} = p-1, \ldots, \pi(N)_{i+p-1} = 1$ and then $(y_i)_{i\in\mathbb{N}}$ is p-uniform at $n+p$ (Lemma 16). This observation is at the origin of the wave pattern.
- The fourth step relies explicitly on fine properties of the Kadanoff sandpile model to show that $\pi(N)_i = 0$ is chosen at most once before the last occurrence of the pattern $p \cdot p{-}1 \cdot \ldots \cdot 1$ in $(\pi(N)_i)_{i\in\mathbb{N}}$ (Lemma 20). At this stage, the upper bounds part of Theorem 26 is proved (Theorem 22).
- The fifth step discusses the lower bounds. The origin of these bounds is the averaging process in step 1 which can not be too quick (Lemma 23) and the additional fact that if the final wave patterns appear at i it conversely implies that $(y_i)_{i\in\mathbb{N}}$ is p-uniform at i (Lemma 24).

1.3 The context

Finding and formally proving global regularity properties suggested by numerical simulations, for models defined by elementary local rules, is a present challenge for physicists, mathematicians, and computer scientists. A lot of conjecture have been proposed on such discrete dynamical systems (sandpile models [6] or chip firing games, but also rotor router [26], the famous Langton's ant [14, 15], *etc*) but very few results have actually been proved.

Regarding KSM(1), the *prediction problem* (the computational complexity of computing the fixed point $\pi(k)$), has been proven in [28] to be in $\mathbf{NC}^2 \subseteq \mathbf{AC}^2$ for the one dimensional case ($\mathbf{NC}$ is Nick's class, of problems efficiently computable in parallel), the model of our purpose (improved to $\mathbf{LOGCFL} \subseteq \mathbf{AC}^1$ in [27]), and $\mathbf{P}$-complete when the dimension is $\geqslant 3$.

The two-dimensional case has been recently studied in [19], where the authors have shown that the *avalanche problem* (given a configuration and two columns i and j, does adding one grain on column i have an influence on column j?) is $\mathbf{P}$-complete for KSM(p) with $p > 1$, which points out an inherently sequential behavior. It is also shown in [11] that the avalanche problem of KSM(p) is in $\mathbf{NC}^1$ (computable in logarithmic time on a polynomial number of processors, *i.e.* efficiently computable in parallel) for all $p \geqslant 1$ in one dimension. The two dimensional case for $p = 1$ is still open, though we know from [13] that simple wires cannot cross.

2 Useful folklore

This section is devoted to the presentation of useful pieces of folklore, simple extensions of known results. We begin with two straightforward results concerning the KSM model (Subsections 2.1 and 2.2), included for the sake of self-containedness. We then present how KSM can be seen as a Chip Firing Game in Subsection 2.3, and Subsection 2.4 gives an inductive way of computing fixed points, via avalanches, which will be used to complete the proof of the upper bound theorem (Theorem 22).

2.1 There is no plateau of length larger than $p + 1$

A plateau is a set of at least two non-empty and consecutive columns of equal height in terms of height configuration, or a set of at least two consecutive slope values equal to 0 in terms of slope configuration. The length of a plateau is the number of columns composing it.

Proposition 4. *For any N and any configuration σ such that $(N, 0^\omega) \xrightarrow{*} \sigma$, in σ there is no plateau of length strictly greater than $p + 1$.*

Proof. This proof proceeds by contradiction, using the fact that configurations are sequences of non-negative integers ($\mathcal{H}_1$). Suppose there exists a slope configuration σ reachable from $(N, 0^\omega)$ for some N, such that there is a plateau of length at least $p + 2$ in σ. Since there is no plateau in the initial configuration $(N, 0^\omega)$, and there is a finite number of steps to reach σ, there exists two slope configurations ρ and τ such that $\rho \to \tau$ and such that there is a plateau of length at least $p+2$ in τ, and none in ρ ($\mathcal{H}_2$). Let k be the leftmost column of the plateau of length at least $p+2$ in τ, *i.e.* for all column j between k and $k+p$, $\tau_j = 0$ ($\mathcal{H}_3$). We will now see that there is no i such that $\rho \xrightarrow{i} \tau$, which completes the proof.

- if $i < k-p$ or $i > k+p+1$ then a firing at i has no influence on columns between k and $k+p+1$ and there is a plateau of length at least $p+2$ in ρ, contradicting $\mathcal{H}_2$.
- if $k-p \leqslant i \leqslant k$ then according to the rule definition we have $\tau_{i+p} = \rho_{i+p-1} - 1$, and from $\mathcal{H}_3$ $\rho_{i+p} = 0$ therefore $\tau_{i+p} < 0$ which is not possible from $\mathcal{H}_1$.
- if $k < i \leqslant k+p+1$ then according to the rule definition we have $\tau_{i-1} = \rho_{i-1} - p$ and from $\mathcal{H}_3$ $\rho_{i-1} = 0$ therefore $\tau_{i-1} < 0$ which again is not possible from $\mathcal{H}_1$. $\square$

2.2 The support of $\pi(N)$ is in $\Theta(\sqrt{N})$

We give bounds for the maximal index of a non-empty column in the fixed point $\pi(N)$ according to the number N of grains, denoted $w(N)$. $w(N)$ can be interpreted as the support or width or size of $\pi(N)$. We consider a general model KSM(p) with p a constant integer greater or equal to 1. A formal definition of $w(N)$ is for example $w(N) = w(\pi(N)) = \min\{i \mid \forall j \geqslant i, \pi(N)_j = 0\}$. See Figure 5.

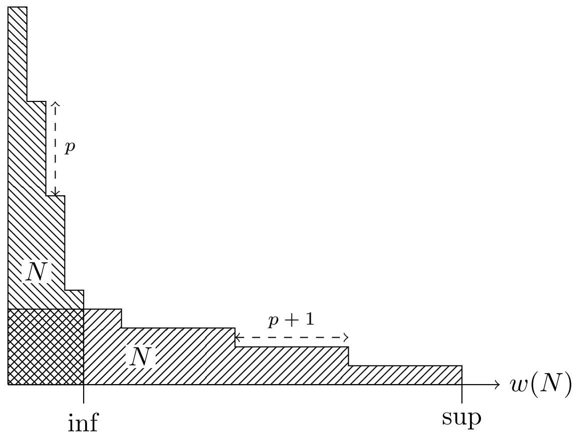


Figure 5: The support of $\pi(N)$ is in $\Theta(\sqrt{N})$. It is lower bounded by the fact that each slope on a stable configuration is at most p, and upper bounded by the fact that there is no plateau of length larger than $p+1$.

Proposition 5. *The support of* $\pi(N)$ *is in* $\Theta(\sqrt{N})$.

Proof. The support of $\pi(N)$ is denoted $w(N)$.

Lower bound: $\pi(N)$ is a fixed point, therefore by definition for all index i we have $\pi(N)_i \leqslant p$. Then,

$$N \leqslant \sum_{i=0}^{w(N)} p \cdot i = p\frac{w(N) \cdot (w(N)+1)}{2} < p^2(w(N)+1)^2$$

hence $\frac{1}{p}\sqrt{N} - 1 < w(N)$.

Upper bound: From Proposition 4, there is no plateau of length greater than $p+1$. Therefore, for $w(N) \geqslant p$ we have

$$N \geqslant \sum_{i=0}^{\lfloor \frac{w(N)}{p+1} \rfloor} (p+1) \cdot i \geqslant (p+1) \left(\frac{\left(\frac{w(N)}{p+1} - 1\right) \frac{w(N)}{p+1}}{2} \right) > \left(\frac{w(N)}{p+1} - 1 \right)^2$$

hence $(p+1)\sqrt{N} + p + 1 > w(N)$. □

2.3 KSM is a Chip Firing Game

Chip Firing Games (CFG) are sandpile models on arbitrary graphs ([2, 3]) (or just sandpile models as Dhar defined them in [7]). A CFG is played on a directed graph in which each vertex v has a load $l(v)$ and a threshold $t(v) = deg^+(v)$ where $deg^+(v)$ denotes the out-degree of v, and the iteration rule is: if $l(v) \geqslant t(v)$ then v gives one unit to each of its neighbors (we say v is fired). As a consequence, we inherit all properties of CFGs.

Kadanoff sandpile is referred to as a *linear chip firing game* in [20]. The authors show that the set of reachable configurations endowed with the order induced by the successor relation has a lattice structure, in particular it has a unique *fixed point*. Since the model is non-deterministic, they also prove *strong convergence i.e.* the number of iterations to reach the fixed point is the same whatever the evolution strategy is. The morphism from KSM(2) to CFG is depicted on Figure 6.

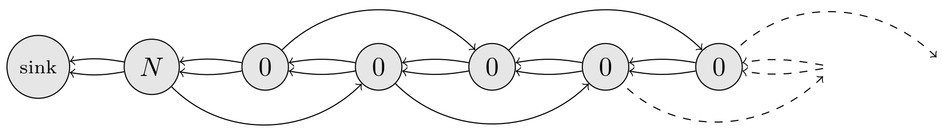


Figure 6: The initial configuration of KSM(2) is presented as a CFG where each vertex corresponds to a column (except the sink, vertices from left to right corresponds to columns $0, 1, 2, 3, \dots$) with a load equal to the slope at i (the height difference $b_i = h_i - h_{i+1}$). For example the vertex with load N is the difference of height between column 0 (N grains) and column 1 (0 grain). The content of the sink is ignored.

When reasoning and writing formal developments about KSM, it is convenient to think about its CFG representation where local rules let units of slope move between columns.

2.4 Hourglass, inductive computation and avalanches

In order to compute $\pi(N)$, the basic procedure is to start from the initial configuration $(N, 0^\omega)$ and perform all the possible transitions. However, it also possible to start from the configuration (0^ω), add one grain on column 0 and perform all the possible transitions, leading to $\pi(1)$, then add another grain on column 0 and perform all the possible transitions, leading to $\pi(2)$, etc... And repeat this process until reaching $\pi(N)$. This inductive definition of fixed points will play a role in Subsection 3.4.

Formally, let b be a slope configuration, $b^{\downarrow 0}$ denotes the configuration obtained by adding one grain on column 0. In other words, if $b = (b_0, b_1, \dots)$ then $b^{\downarrow 0} = (b_0{+}1, b_1, \dots)$. The correctness of the process described above relies on the fact that

$$(k, 0^\omega) \xrightarrow{*} \pi(k-1)^{\downarrow 0}$$

Indeed, there exists a sequence of firings, named a *strategy*, $(s_i)_{i=1}^{i=l}$ such that $(k-1, 0^\omega) \xrightarrow{s_1} \dots \xrightarrow{s_l} \pi(k-1)$. It is obvious that using the same strategy we have $(k, 0^\omega) = (k-1, 0^\omega)^{\downarrow 0} \xrightarrow{s_1} \dots \xrightarrow{s_l} \pi(k-1)^{\downarrow 0}$ since we only drag one more grain on column 0 along the evolution, which does not prevent any firing to occur (see Figure 7). Thus, with the uniqueness of the fixed point reachable from $(k, 0^\omega)$, we have the recurrence formula

$$\pi(\pi(k-1)^{\downarrow 0}) = \pi(k)$$

with the initial condition $\pi(0) = 0^\omega$, enabling an inductive computation of $\pi(k)$.

$$\begin{tikzcd}
(k-1, 0^\omega) \arrow[r, "*"] \arrow[d, "\downarrow 0"'] & \pi(k-1) \arrow[d, "\downarrow 0"'] & \\
(k, 0^\omega) \arrow[r, "*"] & \pi(k-1)^{\downarrow 0} \arrow[r, line width=1.2pt, "*"] & \pi(k)
\end{tikzcd}$$

Figure 7: Inductive computation of $\pi(k)$ from $\pi(k-1)$. The bold arrow represents an avalanche.

The strategy from $\pi(k-1)^{\downarrow 0}$ to $\pi(k)$ is called an *avalanche*. Note that from the non-determinacy of the model, this strategy is not unique. To overcome this issue, it is natural to distinguish a particular one which we think is the simplest: the k^{th} *avalanche* s^k is the leftmost strategy from $\pi(k-1)^{\downarrow 0}$ to $\pi(k)$, where *leftmost* is the minimal strategy according to the lexicographic order. This means that at each step, the leftmost possible firing is performed (example on Figure 8). We reproduce below a preliminary result of [30], which allows to write without ambiguity for an index i: $i \in s^k$ or $i \notin s^k$.

Lemma 6 ([30]). *During an avalanche each column is fired at most once.*

Proof. We prove the result by induction on the number of iterations of the avalanche. At each time step, in order to be fired twice a column must have received at least $p+2$ units of slope, since:

- it has initially a slope at most p on the fixed point,

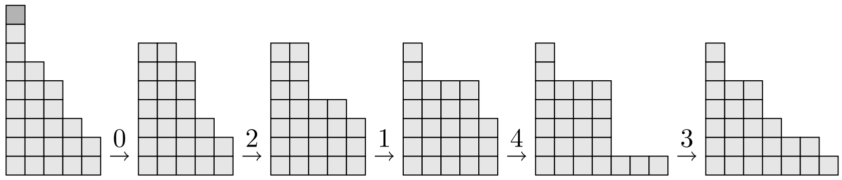


Figure 8: An example of avalanche: starting from $\pi(24)$, we add one grain on column 0 (darkened on the leftmost configuration) and apply the iteration rule until reaching $\pi(25)$. Arrows are labelled by the index of the fired column (the leftmost unstable column is fired at each step). The 25^{th} avalanche $s^{25} = (0, 2, 1, 4, 3)$.

- it looses $p + 1$ units the first time it is fired,
- and it needs to have a slope at least equal to $p + 1$ to fire again.

However, if every other column has been fired only once (by induction hypothesis), any column i received at most p units of slope: 1 unit from the firing at $i - p$, and p units from the firing at $i + 1$. The case $i = 0$ is slightly different: it receives the additional sand grain, but cannot receive a unit of slope from $i - p$. □

3 Analysis

Upper bounding the column from which waves appear on fixed points is the major difficulty this paper overcomes. This section mainly concentrates on this issue, expressed in the upper bound theorem (Theorem 22) (providing an index n in $\mathcal{O}(\log N)$). The last Subsection 3.5 is devoted to lower bounding it (showing that n is in $\Omega(\log N)$). Let us remark that the known results about the case $p = 1$ matches the upper bound result we develop, but not the lower bound result [17].

We consider the parameter $p > 1$ to be fixed, and static constraints on fixed points: the possible sand content on a column has important relations (restrictions) according to its local neighborhood. We therefore study the "*internal dynamic*" of fixed points, via the construction of a DDS in $\mathbb{Z}^{p+1}$, such that the orbit of a well chosen point according to the number of grains N describes the fixed point $\pi(N)$ (Subsection 3.1). The aim is then to prove the convergence of this orbit in $\mathcal{O}(\log N)$ steps, such that the values it takes involve regular wave patterns, as described in the exact bound theorem (Theorem 26).

The reader can refer to Subsection 3.6 and Figure 10 for an illustration of the representations that will be used along the Analysis section.

3.1 Internal dynamic of fixed points

A useful representation of a configuration reachable from $(N, 0^\omega)$ is its *shot vector* $(a_i)_{i\in\mathbb{N}}$, where a_i is the number of times that the rule has been applied on column i from the initial configuration [9] (see Figure 2 for an example). In the following, $(a_i)_{i\in\mathbb{N}}$ and

$(\pi(N)_i)_{i\in\mathbb{N}}$ respectively denote the shot vector and sequence of slopes of the unique fixed point reached from $(N, 0^\omega)$. Those two representations are obviously linked in various ways. In particular for any i we can compute the slope at index i provided the number of firings at $i-p$, i and $i+1$, because the slope at i is initially equal to 0 (the case $i=0$ is discussed below), and

- a firing at $i-p$ increases the slope at i by 1;
- a firing at i decreases the slope at i by $p+1$;
- a firing at $i+1$ increases the slope at i by p;
- any other firing has no consequence on the slope at i.

Therefore, $\pi(N)_i = a_{i-p} - (p+1)\,a_i + p\,a_{i+1}$, with $0 \leqslant \pi(N)_i \leqslant p$ since $\pi(N)$ is a fixed point, and thus

$$a_{i+1} = -\frac{1}{p}\,a_{i-p} + \frac{p+1}{p}\,a_i + \frac{1}{p}\,\pi(N)_i \tag{1}$$

This equation expresses the value of the shot vector at position $i+1$ according to its values at positions $i-p$ and i, and a bounded perturbation $0 \leqslant \frac{\pi(N)_i}{p} \leqslant 1$. As an initial condition, we consider a virtual column of index $-p$ that has been fired N times: $a_{-p} = N$ and $a_i = 0$ for $-p < i < 0$, representing the fact that column 0 is the only one receiving N times 1 unit of slope.

The following lemma states that the value of $\pi(N)_i$ is *nearly determined*.

Lemma 7. *Given a_{i-p} and a_i, there is only one possible value of $\pi(N)_i$, except when $-a_{i-p} + (p+1)\,a_i \equiv 0 \mod p$ in which case $\pi(N)_i$ equals 0 or p.*

Proof. Note that $a_{i+1} \in \mathbb{N}$, thus $-a_{i-p} + (p+1)\,a_i + \pi(N)_i \equiv 0 \mod p$. The Lemma holds since on fixed points we have $0 \leqslant \pi(N)_i \leqslant p$. □

For example, consider $\pi(2000)$ for $p=4$ (see Figure 10). We have $a_8 = 120$ and $a_4 = 189$, so $-a_4 + 5a_8 = 411 \equiv 3 \mod 4$. From this knowledge, $\pi(N)_8$ is determined to be equal to 1, so that $a_9 = -\frac{1}{4}a_4 + \frac{5}{4}a_8 + \frac{1}{4}\pi(N)_8 = 103$ is an integer.

Let us first manipulate Equation 1 so that the resulting recurrence relation is suitable for a study of the convergence. Adding zero sum terms to Equation 1 gives

$$a_{i+1} - a_i = \frac{1}{p}\,(a_i - a_{i-1}) + \frac{1}{p}\,(a_{i-1} - a_{i-2}) + \cdots + \frac{1}{p}\,(a_{i-p+1} - a_{i-p}) + \frac{\pi(N)_i}{p} \tag{2}$$

We rewrite this relation as a linear system we can manipulate easily. Let $y_i = a_i - a_{i-1}$, and $Y_i = {}^t(y_{i-p+1}, \ldots, y_i)$, with ${}^tv$ the transpose of v, be vectors in $\mathbb{Z}^p$. Note that we consider only finite configurations, so there always exists an integer i_0 such that $a_i = 0$

for $i_0 \leqslant i$, and therefore $Y_i = \vec{0}$ for $i_0 + p \leqslant i$, with $\vec{0} = {}^t(0, \ldots, 0)$. Given Y_i and $\pi(N)_i$ we can compute Y_{i+1} with the relation

$$Y_{i+1} = M\,Y_i + \frac{\pi(N)_i}{p} K \qquad \text{with} \qquad M = \begin{pmatrix} 0 & 1 & & 0 \\ & & \ddots & \\ 0 & 0 & & 1 \\ \frac{1}{p} & \frac{1}{p} & \ldots & \frac{1}{p} \end{pmatrix} \quad K = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{3}$$

M is a square matrix[2] of size p, and we call this partially defined recurrence relation (the sequence $(\pi(N)_i)_{i \in \mathbb{N}}$ is not known) the *averaging system*. This system expresses the difference of shot vector around position $i+1$ (via Y_{i+1}) in terms of the shot vector around position i (via Y_i) and the slope at i (via $\pi(N)_i$). Thus there exists an *orbit* of the point $Y_0 = (-N, 0, \ldots, 0, a_0)$ of $\mathbb{Z}^p$, *i.e.* a sequence of non-deterministic iterations (since $\pi(N)_i$ may take two possible values as stated in Lemma 9, of which only one corresponds to the real value of $\pi(N)_i$) starting from Y_0, which describes the sequence of differences of shot vector of the fixed point composed of N grains.

The system we get is a linear map plus a perturbation induced by the discreteness of values of the slope. Note that it may look odd to study the sequence $(\pi(N)_i)_{i \in \mathbb{N}}$ using a DDS whose iterations presuppose the knowledge of $(\pi(N)_i)_{i \in \mathbb{N}}$. It is actually helpful because of the underlined fact that values $\pi(N)_i$ are *nearly determined* (Lemma 7): in a first phase we will make no assumption on the sequence $(\pi(N)_i)_{i \in \mathbb{N}}$ (except that $0 \leqslant \pi(N)_i \leqslant p$ for all i) and prove that the system converges exponentially quickly in N; and in a second phase we will see that from an n in $\mathcal{O}(\log N)$ such that the system has converged, the sequence $(\pi(N)_i)_{i \geqslant n}$ is *determined* to have a regular wavy shape.

Now that we have the averaging system, the proof of the upper bound theorem (Theorem 22) is done in two steps:

1. first prove that it converges exponentially quickly (in $\mathcal{O}(\log N)$ steps) to a uniform vector (Subsection 3.2);

2. then prove that as soon as the vector is uniform, then the wavy shape of the upper bound theorem (Theorem 22) takes place (Subsections 3.3 and 3.4).

3.2 Convergence of the averaging system

The averaging system is understandable in simple terms. From Y_i in $\mathbb{Z}^p$, we obtain Y_{i+1} by:

1. shifting all the values one row upward;

2. for the bottom component, computing the mean of values of Y_i, and adding a small perturbation (a multiple of $\frac{1}{p}$ between 0 and 1) to it.

[2]As a convention, blank spaces are 0s and dotted spaces are filled by the sequence induced by its endpoints.

Notation 8. Let m_i (respectively $\overline{m}_i$, $\underline{m}_i$) denote the mean (respectively maximal, minimal) of values of Y_i.

Lemma 9. *If m_i is not an integer then $\pi(N)_i$ equals $p\left(\lceil m_i \rceil - m_i\right)$, otherwise $\pi(N)_i$ equals 0 or p.*

Proof. $(Y_i)_{i\in\mathbb{N}}$ are integer vectors, hence the perturbation added to the last component is again nearly determined (see Lemma 7), we have $(m_i + \frac{\pi(N)_i}{p}) \in \mathbb{Z}$ and $0 \leqslant \frac{\pi(N)_i}{p} \leqslant 1$. □

For example, consider $\pi(2000)$ for $p = 4$ (see Figure 10a, be careful that it pictures $a_i - a_{i+1}$ at position i). We have $Y_{13} = {}^t(-3, -5, -7, -7)$, then $m_{13} = -\frac{11}{2}$ and $\pi(N)_{13}$ is forced to be equal to 2 so that $y_{14} = -5$ and $Y_{14} = {}^t(-5, -7, -7, -5)$ is an integer vector.

We can foresee what happens as we iterate this dynamical system and new values are computed: on a *large scale* (when values are large compared to p) the system evolves roughly toward the mean of values of the initial vector Y_0, and on a *small scale* (when values are small compared to p) the perturbation lets the vector wander a little around.

The study of the convergence of the averaging system works in three steps:

(i). state a linear convergence of the whole system; then express Y_n in terms of Y_0 and $(\pi(N)_i)_{0\leqslant i\leqslant n}$;

(ii). isolate the perturbations induced by $(\pi(N)_i)_{0\leqslant i\leqslant n}$ and bound them;

(iii). prove that the other part (corresponding to the linear map M) converges exponentially quickly.

From (ii) and (iii), a point converges exponentially quickly into a ball of constant radius, then from (i) this point needs a constant number of extra iterations in order to reach the center of the ball, that is, a uniform vector.

Recall notations 8, we will prove that $\overline{m}_i - \underline{m}_i$ converges exponentially quickly to 0 (N is fixed and i runs). For $p > 1$ we start with $Y_0 = {}^t(-N, 0, \ldots, 0, a_0)$, thus $\overline{m}_0 - \underline{m}_0 = N + a_0$. Furthermore
$$\frac{N}{p+1} \leqslant a_0 \leqslant \frac{N}{p}.$$
Indeed, recall that a_0 is the number of times column 0 has been fired: we can begin the evolution from $(N, 0^\omega)$ by firing it $\frac{N}{p+1}$ times, and after $\frac{N}{p}$ firings there are no more grains on column 0. Consequently $\overline{m}_0 - \underline{m}_0$ is in $\Theta(N)$.

Firstly, we prove that the system converges exponentially quickly on a large scale. Intuitively, when $\overline{m}_i - \underline{m}_i$ is large compared to p, the perturbation induced by $(\pi(N)_i)_{i\in\mathbb{N}}$ is negligible.

Lemma 10. *There exists a constant α and a n_0 in $\mathcal{O}(\log N)$ s.t. $\overline{m}_{n_0} - \underline{m}_{n_0} < \alpha$.*

Proof. The relation linking Y_i to Y_{i+1} is (M and K are defined in Equation 3)
$$Y_{i+1} = M\,Y_i + \frac{\pi(N)_i}{p}\,K$$

Let $M_i = {}^t(m_i, \ldots, m_i)$ in $\mathbb{Z}^p$. Since Y_i converges roughly towards the mean of its values, we consider the evolution of $Z_i = Y_i - M_i$ which tends to $\vec{0}$. We have

$$Z_{i+1} = O\,Z_i + \frac{\pi(N)_i}{p}\,L \text{ where } \begin{cases} O = D\,M \\ L = D\,K \end{cases} \text{with } D = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} - \frac{1}{p}\begin{pmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{pmatrix}$$

because $D\,M\,Y_i = D\,M\,D\,Y_i$.

The aim is thus to prove that there exists an n_0 in $\mathcal{O}(\log N)$ such that the norm of Z_{n_0} is bounded by a constant.

We express Z_n in terms of Z_0 and $(\pi(N)_i)_{0 \leqslant i \leqslant n}$:

$$Z_n = O^n Z_0 + \frac{1}{p}\sum_{i=0}^{n-1} \pi(N)_i O^{n-1-i} L$$

The main argument for the proof of the lemma is in the following fact.

Fact 11. *The matrix O is diagonalizable, and the spectral radius r (the maximal modulus of an eigenvalue of O) is such that $r \leqslant \frac{p-1}{p} < 1$.*

Let us postpone the demonstration of Fact 11. The rest of the proof is straightforward, using the infinity norm induced by a basis of proper vectors. With such a norm, we have:

$$\|Z_n\| \leqslant r^n\|Z_0\| + \frac{1}{p}\sum_{i=0}^{n-1}\pi(N)_i r^{n-1-i}\|L\| \leqslant r^n\|Z_0\| + \sum_{i=0}^{n-1} r^{n-1-i}\|L\| \leqslant r^n\|Z_0\| + \frac{\|L\|}{1-r}.$$

A classical theorem of linear algebra (see for example [4]) claims that, in finite dimension, all norms are equivalent. In particular, there exists a positive constant c such that, for each vector Z, we have: $\|Z\|_\infty \leqslant c\,\|Z\|$. Using it, we get:

$$\overline{m}_n - \underline{m}_n \leqslant 2\,\|Z_n\|_\infty \leqslant 2\,c\,\|Z_n\| \leqslant 2\,c\,\left(r^n\,\|Z_0\| + \frac{\|L\|}{1-r}\right).$$

Now, if we take $\alpha = 2\,c\,\left(1 + \frac{\|L\|}{1-r}\right)$, and n_0 the lowest integer such that $r^{n_0}\,\|Z_0\| < 1$, *i.e.* $n_0 > -\frac{\ln(\|Z_0\|)}{\ln(r)}$, then we have the result (note that $\|Z_0\|_\infty$ is in $\mathcal{O}(N)$ because $Y_0 = {}^t(-N, 0, 0, \ldots, 0, a_0)$ with $a_0 < \frac{N}{p}$, so this also holds for $\|Z_0\|$ since all norms are equivalent).

To complete the proof, we now have to prove Fact 11 above. M is a *companion matrix* and its characteristic polynomial is

$$x^p - \sum_{k=0}^{p-1}\frac{1}{p}x^k = (x-1)R(x)$$

with $R(x) = \frac{1}{p}\Big(p\, x^{p-1} + (p-1)\, x^{p-2} + \cdots + 2\, x + 1\Big)$.

Consider $S(x) = px^{p-1} R(\frac{1}{x}) = x^{p-1} + 2x^{p-2} + \cdots + (p-1)x + p$. We have the Bezout equality: $a\, S + b\, S' = 1$, with $a(x) = \frac{-p+1}{p(p+1)} x + \frac{1}{p}$ and $b(x) = \frac{1}{p(p+1)} x^2 - \frac{1}{p(p+1)} x$. Thus, S and S' are co-prime, which ensures that S has $p-1$ distinct roots. Therefore, R also has $p-1$ distinct roots.

Moreover, a classical result due to Eneström and Kakeya (see for example [16]) states that, if λ is a root of a polynomial P such that $P(x) = \sum_{i=0}^{n} \alpha_i\, x^i$ with each $\alpha_i > 0$, then we have $\min\{\frac{\alpha_i}{\alpha_{i+1}} \,|\, 0 \leqslant i < n\} \leqslant |\lambda| \leqslant \max\{\frac{\alpha_i}{\alpha_{i+1}} \,|\, 0 \leqslant i < n\}$. Applying this result to $R(x)$, we obtain that all roots of $R(x)$ have a modulus between $\frac{1}{2}$ and $\frac{p-1}{p}$.

As a consequence M is diagonalizable, and the set of eigenvalues of M is $M_\lambda = \{1, \lambda_1, \ldots, \lambda_{p-1}\}$, with all λ_i pairwise distinct. Now, let $v_0, \ldots, v_{p-1}$ be non null eigenvectors respectively associated to the eigenvalues $1, \lambda_1, \ldots, \lambda_{p-1}$ of M.

- We have $DM\, v_0 = D\, v_0$ by definition of v_0. Moreover, if $M\, v_0 = v_0$, then, by an easy computation, we obtain that all components of v_0 are equal. This fact enforces $D\, v_0 = \vec{0}$. In other words, 0 is an eigenvalue of DM.

- For the other eigenvectors, that is, for $1 \leqslant i \leqslant p-1$, let c_i be the uniform vector with all its components equal to $\frac{1}{p}\sum_{k=0}^{p-1} v_{i_k}$, with v_{i_k} the k^{th} component of the vector v_i. Notice that $M\, c_i = c_i$, which ensures that $v_i - c_i \neq \vec{0}$. We have

$$\begin{aligned} DM\,(v_i - c_i) &= D\,(M\, v_i - M\, c_i) \\ &= D\,(\lambda_i\, v_i - c_i) \\ &= \lambda_i\, D\, v_i - D\, c_i \\ &= \lambda_i (v_i - c_i) - \vec{0} \end{aligned}$$

  where the last equality is obtained from the fact that by definition of D, we have $D\, v_i = v_i - c_i$. As a consequence, λ_i is an eigenvalue of DM.

Finally, the set DM_λ of eigenvalues of DM is $DM_\lambda = \{0, \lambda_1, \ldots, \lambda_{p-1}\}$, and the spectral radius of DM is at most $\frac{p-1}{p}$. □

Secondly, on a small scale, the system converges linearly. Note that the convergence is not "stable", in the sense that once a uniform vector is reached ($\underline{m}_i = \overline{m}_i$), the next y_{i+1} may be incremented so that after one more iteration the vector may not be uniform anymore (as can be seen on the example of Figure 10a).

Lemma 12. *When distinct from 0, the value of $\overline{m}_i - \underline{m}_i$ decreases linearly: if $\underline{m}_i \neq \overline{m}_i$, then there is an integer c, with $1 \leqslant c \leqslant p$ such that $\overline{m}_{i+c} - \underline{m}_{i+c} < \overline{m}_i - \underline{m}_i$.*

Proof. If $\underline{m}_i \neq \overline{m}_i$, that is, if the vector Y_i is not uniform, the mean value is strictly between the greatest and smallest values: $\underline{m}_i < m_i < \overline{m}_i$. Consequently $\underline{m}_i < y_{i+p} = m_i + \frac{\pi(N)_i}{p} \leqslant \overline{m}_i$ (since the perturbation added is at most one and the resulting number is an integer, we cannot reach a greater integer). Therefore, we get $\underline{m}_i \leqslant \underline{m}_{i+1} \leqslant \overline{m}_{i+1} \leqslant \overline{m}_i$.

This reasoning applies while $\underline{m}_{i+j} \neq \overline{m}_{i+j}$, from which we get $\underline{m}_{i+j} < y_{i+p+j} \leqslant \overline{m}_{i+j}$ and $\underline{m}_{i+j} \leqslant \underline{m}_{i+j+1} \leqslant \overline{m}_{i+j+1} \leqslant \overline{m}_{i+j}$.

If there exists $c \leqslant p$ such that $\underline{m}_{i+c} = \overline{m}_{i+c}$, then, we are done. Otherwise, for $0 \leqslant j < p$, we have, $\underline{m}_i \leqslant \underline{m}_{i+j} < y_{i+j+p} \leqslant \overline{m}_{i+j} \leqslant \overline{m}_i$), thus $\underline{m}_i < \underline{m}_{i+p} \leqslant \overline{m}_{i+p} \leqslant \overline{m}_i$. □

Lemmas 10 and 12 allow to conclude the exponential convergence of Z_n to a uniform vector, stated in the following convergence lemma.

Lemma 13. *There exists an n in $\mathcal{O}(\log N)$ such that Y_n is a uniform vector.*

Proof. We start with $\overline{m}_0 - \underline{m}_0$ in $\mathcal{O}(N)$, we have a constant α and a n_0 in $\mathcal{O}(\log N)$ such that $\overline{m}_{n_0} - \underline{m}_{n_0} < \alpha$ thanks to the exponential decrease on a large scale (Lemma 10). Then after p iterations the value of $\overline{m}_{n_0+p} - \underline{m}_{n_0+p}$ is decreased by at least 1 (Lemma 12), hence there exists β with $\beta \leqslant p\,\alpha$ such that after β extra iterations we have $\overline{m}_{n_0+\beta} - \underline{m}_{n_0+\beta} = 0$. Thus $Y_{n_0+\beta}$ is a uniform vector, and $n_0 + \beta$ is in $\mathcal{O}(\log N)$. □

In this proof, neither the continuous (Lemma 10) nor the discrete (Lemma 12) studies is conclusive by itself. On one hand, the discrete study (combinatorial arguments) gives a linear convergence to a precise pattern, but not an exponential convergence. On the other hand, the continuous study (arguments of linear algebra) gives an exponential convergence towards an almost uniform vector, but in itself the continuous part never reaches a uniform vector (because of the uncertainties in the values of $(\pi(N)_i)_{i\in\mathbb{N}}$). Furthermore the uniform vector towards which the averaging system tends are not stable: once a uniform vector is reached, one more iteration may lead to a non-uniform vector. The combination of results on those two modalities (discrete and continuous) appears as an important step for proving the convergence lemma (Lemma 13).

Remark 14. Note that for $p = 1$, the averaging system has a trivial dynamics. For $p = 2$, the behavior is a bit more complex, but major simplifications are found: the computed value is equal to the mean of two values, hence in this case the difference $\overline{m}_i - \underline{m}_i$ decreases by a factor of two *at each time step.*

Remark 15. Y_i is uniform implies that Y_j is almost uniform for all $j \geqslant i$. Indeed, the mean m_i is an integer, consequently $y_{i+1} = m_i + \frac{\pi(N)_i}{p}$ is either equal to y_i, or equal to $y_i + 1$. In the latter case $y_{i+1} = y_{i+2} = \cdots = y_{i+p}$, as it is detailed in Subsection 3.3.

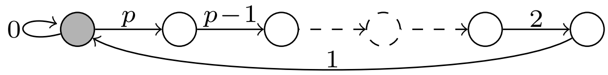
3.3 Emergence of a loosely wavy shape

We call *wave* the pattern $p \cdot \ldots \cdot 2 \cdot 1$ in the sequence of slopes. The convergence lemma (Lemma 13) shows that there exists an n in $\mathcal{O}(\log N)$ such that Y_n is a uniform vector. In this subsection, we prove that if Y_n is a uniform vector, then from index n the shape of the sandpile configuration is exclusively composed of waves and 0s.

Lemma 16. *Y_n is a uniform vector of $\mathbb{Z}^p$ implies*

$$\pi(N)_{[n\,,\infty[} \in \big(0 + (p \cdot p{-}1 \cdot \ldots \cdot 1)\big)^* 0^\omega$$

Proof. The idea of this proof follows Lemma 9. If Y_i is a uniform vector, we notice that the value of $\pi(N)_i$ is 0 or p. If it is 0, then Y_{i+1} is still a uniform vector; if it is p, then the sequence $(\pi(N)_j)_{i \leqslant j < i+p}$ is determined to be equal to $p \cdot p-1 \cdot \ldots \cdot 1$, and Y_{i+p} is again a uniform vector (recall that $y_{i+1} = m_i + \frac{\pi(N)_i}{p} \in \mathbb{Z}$). The following diagram illustrates this observation: the grey node corresponds to a uniform vector Y_i, and at each iteration we follow an arc whose label gives the value for $\pi(N)_i$. If we start from the grey node, any path's labels verify the statement of the lemma.



We concentrate on the sequence of values of Y_i. The fact that its components are integers, and especially the last one, will play a crucial role in the determination of the value of $\pi(N)_i$ because $0 \leqslant \pi(N)_i \leqslant p$.

We start from the hypothesis of uniformity: that $Y_i = {}^t(\gamma, \ldots, \gamma)$ for some γ, thus from the averaging system's Equation 3 we have $Y_{i+1} = {}^t(\gamma, \ldots, \gamma, \gamma + \frac{\pi(N)_i}{p})$. Y_{i+1} is an integer vector and γ is an integer, hence $\pi(N)_i$ equals 0 or p.

- If $\pi(N)_i = 0$ then $Y_{i+1} = {}^t(\gamma, \ldots, \gamma)$ and we are back to the same situation, the dilemma goes on: the value of $\pi(N)_{i+1}$ is not determined, it can be 0 or p.

- If $\pi(N)_i = p$ then $Y_{i+1} = {}^t(\gamma, \ldots, \gamma, \gamma + 1)$ from the relation above. A regular pattern then emerges:

  - if $Y_{i+1} = {}^t(\gamma, \ldots, \gamma, \gamma + 1)$, then $Y_{i+2} = {}^t(\gamma, \ldots, \gamma, \gamma + 1, \frac{p\gamma+1+\pi(N)_{i+1}}{p})$ and it determines $\pi(N)_{i+1} = p - 1$ so that $Y_{i+2} = (\gamma, \ldots, \gamma, \gamma + 1, \gamma + 1)$ is an integer vector (Lemma 9);
  - if $Y_{i+2} = {}^t(\gamma, \ldots, \gamma, \gamma{+}1, \gamma{+}1)$, then $Y_{i+3} = {}^t(\gamma, \ldots, \gamma, \gamma{+}1, \gamma{+}1, \frac{p\gamma+2+\pi(N)_{i+2}}{p})$ and it determines $\pi(N)_{i+2} = p - 2$ so that $Y_{i+3} = {}^t(\gamma, \ldots, \gamma, \gamma + 1, \gamma + 1, \gamma + 1)$ is an integer vector (Lemma 9);
  - *et cetera* we have $\pi(N)_{i+j} = p - j$ for $0 \leqslant j < p$, and eventually $Y_{i+p} = {}^t(\gamma + 1, \ldots, \gamma + 1)$ is a uniform vector (note that Y_0 has a negative mean, hence γ is negative, which is consistent with the $\gamma + 1$ we obtain).

  Therefore we are back to the initial grey node, after one complete wave pattern has formed.

This process goes on until $Y_i = \vec{0}$. □

Composing the convergence lemma (Lemma 13) and Lemma 16, we can show the exponentially quick emergence of loosely wavy patterns:

Corollary 17. *Let p be fixed. There exists a column n in $\mathcal{O}(\log N)$ such that*

$$\pi(N)_{[n\,,\infty[} \in \big(0 + (p \cdot p-1 \cdot \ldots \cdot 1)\big)^* 0^\omega$$

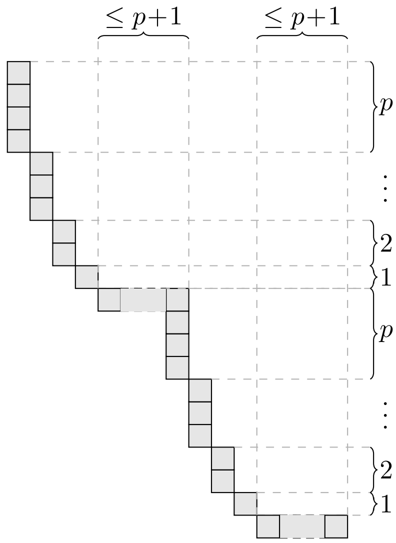


Figure 9: Emergence of loosely wavy shapes (Corollary 18).

Note that Proposition 4 of Subsection 2.1, stating the impossibility of having more than $p+1$ consecutive columns with the same height in a configuration reachable from $(N, 0^\omega)$, allows to refine the corollary above to get:

Corollary 18. *Let* p *be fixed. There exists a column* n *in* $\mathcal{O}(\log N)$ *such that*

$$\pi(N)_{[n\,,\infty[} \in \left(p \cdot \ldots \cdot 2 \cdot 1 \cdot \left(\underset{i=0}{\overset{p}{+}} 0^i\right)\right)^* 0^\omega$$

This corollary is illustrated on Figure 9.

3.4 Avalanches to complete the proof

In order to prove the upper bound theorem (Theorem 22), we refine Corollary 18 to show that there is at most one set of two non-empty and consecutive columns of equal height, called a *plateau* of size two, and corresponding to a slope equal to 0. It seems necessary to overcome the *static* study -(or a given fixed point) presented above, and consider the *dynamic* of sand grains from $\pi(0)$ to $\pi(N)$. These developments use the notion of avalanche introduced in Subsection 2.4: from the relation

$$\text{for all } k > 0,\ \pi(\pi(k-1)^{\downarrow 0}) = \pi(k)$$

the k^{th} avalanche is the minimal strategy from $\pi(k-1)^{\downarrow 0}$ to $\pi(k)$ according to the lexicographic order. We will use the fact that the structure of an avalanche on a wave pattern is very constrained.

We say that there is a *hole* at position i in an avalanche s^k if and only if $i \notin s^k$ and $(i+1) \in s^k$. An interesting property of an avalanche is the absence of hole from an index

l, which tells that

$$\text{there exists an } m \text{ such that } \begin{cases} \text{for all } i \text{ with } l \leqslant i \leqslant m, \text{ we have } i \in s^k \\ \text{for all } i \text{ with } m < i, \text{ we have } i \notin s^k \end{cases}$$

namely, from column l, a set of consecutive columns is fired, and nothing else. We say that an avalanche s^k is *dense starting from* an index l when s^k contains no hole i with $i \geqslant l$. We studied the structure of avalanches in [30, 31, 33], and saw that this property leads to important regularities in successive fixed points. It induces a kind of "pseudo linearity" on avalanches, it somehow "breaks" the criticality of avalanche's behavior and let them flow smoothly along the sandpile. Let us introduce a formal notation:

Definition 19. $\mathcal{L}'(p,k)$ is the minimal column such that the k^{th} avalanche is dense starting at $\mathcal{L}'(p,k)$:

$$\mathcal{L}'(p,k) = \min\{l \in \mathbb{N} \mid \exists m \in \mathbb{N} \text{ such that } \forall\, l \leqslant i \leqslant m, i \in s^k \text{ and } \forall\, i > m, i \notin s^k\}$$

Then, the *global density column* $\mathcal{L}(p,N)$ is defined as:

$$\mathcal{L}(p,N) = \max\{\mathcal{L}'(p,k) \mid k \leqslant N\}$$

The global density column $\mathcal{L}(p,N)$ is the smallest column number starting from which the N first avalanches are dense (*i.e.* contain no hole). A conjecture of [30], proven only for $p = 2$, is solved by Corollary 21 of the following lemma. This latter shows how constrained the avalanche is on wave patterns: it is forced to stop on the first slope of value 0 it encounters. The column on which the avalanche stops is the maximal index fired within the $N{+}1^{th}$ avalanche, and is denoted $\max s^{N+1}$.

Lemma 20. *Let n and N be such that*

$$\pi(N)_{[\![n\,,\infty[\![} \in (p \cdot p{-}1 \cdot \ldots \cdot 1)\Big(0 + (p \cdot p{-}1 \cdot \ldots \cdot 1)\Big)^{*} 0^{\omega}.$$

s^{N+1} is the $(N+1)^{th}$ avalanche, and we have

- $\mathcal{L}'(p, N+1) \leqslant n$ *;*
- *and if* $\max s^{N+1} \geqslant n$ *then* $\max s^{N+1} = \min\{i > n : \pi(N)_i = 0\} - p$.

Proof. If $\max s^{N+1} < n$, the two parts of the lemma are straightforwardly true. Let us consider the case where $\max s^{N+1} \geqslant n$.

We prove the two parts of the lemma by induction on the number of consecutive (without column of slope 0 in-between) wave patterns. Considering an avalanche, it is easy to observe that (one can refer to [30] for details):

1. if there is a set of p consecutive columns with slopes all strictly smaller than p then the avalanche stops before it;

2. the greatest fired column, $\max s^{N+1}$ has slope value p in $\pi(N)$.

From the two remarks above it is now straightforward to conclude: n has value p and is fired (since $\max s^{N+1} \geqslant n$), and then by induction if the next wave is consecutive (without symbol 0) then the avalanche goes on one wave further, otherwise is stops, and in both cases the two parts of the lemma are verified. More precisely, let $\pi(N)_q = p$ with $q \geqslant n$ be a fired column, we will prove one induction step (initialization is done with n).

- If $q+p \neq 0$, that is if there is a wave on columns $q+p$ to $q+2\,p-1$, then $q+p$ is fired since $\pi(N)_{q+p} = p$ and receives one unit of slope when column q is fired. We can now notice that a chain reaction fires columns $q+p-1$ to $q+1$:

$$\ldots (\underset{\overset{\uparrow}{q}}{p} \cdot \ldots \cdot 2 \cdot 1)(\underset{\overset{\uparrow}{q+p}}{p} \cdot \ldots \cdot 2 \cdot 1) \ldots$$

  - the tumbling of $q+p$ gives p units of slope to $q+p-1$, but $\pi(N)_{q+p-1} = 1$ therefore it becomes unstable and will be fired ;
  - the tumbling of $q+p-1$ gives p units of slope to $q+p-2$, but $\pi(N)_{q+p-2} = 2$ therefore it becomes unstable and will be fired ;
  - ...
  - the tumbling of $q+2$ gives p units of slope to $q+1$, but $\pi(N)_{q+1} = p-1$ therefore it becomes unstable and will be fired.

  Columns $q+1$ to $q+p-1$ are all fired, and $q+p$ is the next column considered in the induction (recall Lemma 6, each column is fired at most once).

- If $q+p = 0$,

$$\ldots (\underset{\overset{\uparrow}{q}}{p} \cdot \ldots \cdot 2 \cdot 1) \underset{\overset{\uparrow}{q+p}}{0} \ldots$$

  from observation 1 the avalanche stops at $q+p$ and the lemma holds. □

From Corollary 17, this lemma applies for a column n in $\mathcal{O}(\log N)$. We therefore have $\mathcal{L}'(p,N)$ in $\mathcal{O}(\log N)$, *i.e.*, the N^{th} avalanche is dense starting from a logarithmic index in N, and we deduce the following corollary on the emergence of regularities in the structure of avalanches.

Corollary 21. *Let p be fixed, $\mathcal{L}(p,N)$ is in $\mathcal{O}(\log N)$.*

We are now able to conclude the proof of the first part of our main theorem.

Theorem 22. *For a fixed p, there exists an n in $\mathcal{O}(\log N)$ such that*

$$\pi(N)_{[n\,,\,\infty[} \in (p \cdot \ldots \cdot 2 \cdot 1)^* \, 0 \, (p \cdot \ldots \cdot 2 \cdot 1)^* \, 0^\omega.$$

Proof. We prove the result by induction on the number of grains N. From Corollary 18, for all $N' \leqslant N$ there exists a column n' in $\mathcal{O}(\log N')$ such that $\pi(N')_{[n',\infty[}$ has the form

$$\left(p \cdot p{-}1 \cdot \ldots \cdot 1 \cdot \left(\mathop{+}_{i=0}^{p} 0^i\right)\right)^* 0^\omega.$$

Let $\mathcal{N}$ be the set of such indices n' (for N' from 0 to N). We will consider the column $\ell = \max \mathcal{N} + p$, which is in $\mathcal{O}(\log N)$, and prove that $\pi(N)_{[\ell,\infty[}$ has at most one symbol 0.

We prove the result by induction, supposing that for a N' with $N' < N$ we have at most one value 0 within $\pi(N')_{[\ell,\infty[}$ (apart from the final 0^ω), and showing that this fact still holds for $N'+1$. Let $q = \min\{i \geqslant \ell : \pi(N')_i = 0\}$ be the index of slope 0 among the wave patterns, or the first in the infinite sequence of 0. We depict $\pi(N')$ as follows

$$\ldots \underset{\substack{\uparrow \\ q-p}}{(}p \cdot p-1 \cdot \ldots \cdot 1)\, \underset{\substack{\uparrow \\ q}}{0} (p \cdot p-1 \cdot \ldots \cdot 1)^* \, 0^\omega$$

with $\ell \leqslant q$.

- If $\max s^{N'+1} + p < \ell$ then this avalanche does not modify the slopes on the right of ℓ (included), thus $\pi(N'+1)_{[\ell,\infty[}$ still has at most one value 0.
- If $\max s^{N'+1} + p \geqslant \ell$ then by definition of ℓ we can apply Lemma 20 for a column smaller of equal to $\ell - p$, which gives
  - $\mathcal{L}'(p, N'+1) \leqslant \ell - p$ (hypothesis $\mathcal{H}_{dense}$)
  - $\max s^{N'+1} = \min\{i > \ell - p : \pi(N')_i = 0\} - p$

  Since $\max s^{N'+1} + p \geqslant \ell$, we have

  $$\min\{i > \ell - p : \pi(N')_i = 0\} \geqslant \ell$$

  which means according to the induction hypothesis that $\max s^{N'+1} = q - p$.

  Thanks to the density of avalanches before column $\ell - p$ (hypothesis $\mathcal{H}_{dense}$), we know all the fired columns ($\max s^{N'+1} = q - p$), and we can simply compute the values of $\pi(N'+1)$ on the right of column ℓ by summing the gain and loss of units of slope:

  - $\pi(N'+1)_{q-p} = 0$;
  - $\pi(N'+1)_{q-p+i} = \pi(N')_{q-p+i} + 1$ for $i \in [1\,,\,p]$;
  - $\pi(N'+1)_i = \pi(N')_i$ for $i \geqslant \ell$ and $i \notin [q-p\,,\,q]$.

  Consequently, the value of slope 0 has "climbed" one wave on the left, which concludes the proof. □

3.5 Lower bounding the wave appearance

Few more developments allow us to prove a tight asymptotic bound of $\Theta(\log N)$ for the wave appearance. This subsection is devoted to it. We first prove an $\Omega(\log N)$ bound on the n such that Y_n is uniform, and then that it is not possible to get wave patterns as described in the upper bound theorem (Theorem 22) if Y_n is not uniform.

Lemma 23. *For a fixed $p > 1$, the smallest n such that Y_n is a uniform vector is in $\Omega(\log N)$.*

Proof. This lemma follows from rough bounds on the difference between the greatest and smallest value of Y_{i+1} compared to that of Y_i, respectively denoted $\overline{m}_{i+1}$, $\underline{m}_{i+1}$, $\overline{m}_i$ and $\underline{m}_i$. From $Y_i = {}^t(y_{i-p+1}, \dots, y_i)$ to $Y_{i+1} = {}^t(y_{i-p+2}, \dots, y_{i+1})$ in the averaging system, the value y_{i-p+1} is "forgotten" and the value $y_{i+1} = m_i + \frac{\pi(N)_i}{p}$ is "appended", where m_i is the mean of the values of Y_i. Let us do a case disjunction.

- If neither $\overline{m}_i$ nor $\underline{m}_i$ equals y_{i-p+1} (the value that is forgotten) then nothing moves: $\overline{m}_{i+1} = \overline{m}_i$ and $\underline{m}_{i+1} = \underline{m}_i$ since the new value y_{i+1} verifies $\underline{m}_i < y_{i+1} \leqslant \overline{m}_i$ (when $\overline{m}_i \neq \underline{m}_i$, with the same argument as in the proof of Lemma 12).

- If $y_{i+p-1} = \overline{m}_i$, that is, when the maximal value of Y_i is forgotten, then $\underline{m}_{i+1} = \underline{m}_i$ and
$$\overline{m}_{i+1} \geqslant y_{i+1} = m_i + \frac{\pi(N)_i}{p} \geqslant \frac{p-1}{p}\,\underline{m}_i + \frac{1}{p}\,\overline{m}_i$$
where the left inequality comes from the fact that the greatest value of Y_{i+1} is at least the newly appended y_{i+1}, and the right inequality is the worst case if all the other values $y_{i-p+2}, \dots, y_i$ equal the smallest $\underline{m}_i$. We can conclude that
$$\overline{m}_{i+1} - \underline{m}_{i+1} \geqslant \frac{p-1}{p}\,\underline{m}_i + \frac{1}{p}\,\overline{m}_i - \underline{m}_i = \frac{1}{p}\,(\overline{m}_i - \underline{m}_i).$$

- If $y_{i-p+1} = \underline{m}_i$, that is, when the minimal value of Y_i is forgotten, then $\overline{m}_{i+1} = \overline{m}_i$ and
$$\underline{m}_{i+1} \leqslant y_{i+1} = m_i + \frac{\pi(N)_i}{p} \leqslant \frac{1}{p}\,\underline{m}_i + \frac{p-1}{p}\,\overline{m}_i + 1$$
where the left inequality comes from the fact that the smallest value of Y_{i+1} is at least the newly appended y_{i+1}, and the right inequality is the worst case if all the other values $y_{i-p+2}, \dots, y_i$ equal the greatest $\overline{m}_i$. We can conclude that
$$\overline{m}_{i+1} - \underline{m}_{i+1} \geqslant \overline{m}_i - \frac{1}{p}\,\underline{m}_i + \frac{p-1}{p}\,\overline{m}_i - 1 = \frac{1}{p}\,(\overline{m}_i - \underline{m}_i) - 1.$$

In any case, we have $\overline{m}_{i+1} - \underline{m}_{i+1} \geqslant \frac{1}{p}\,(\overline{m}_i - \underline{m}_i) - 1$. This can be rewritten as
$$\overline{m}_{i+1} - \underline{m}_{i+1} + \frac{p}{p-1} \geqslant \frac{1}{p}\left(\overline{m}_i - \underline{m}_i + \frac{p}{p-1}\right)$$

A straightforward induction on the above inequality gives that, for each non-negative integer i,

$$\overline{m}_i - \underline{m}_i + \frac{p}{p-1} \geqslant \frac{1}{p^i} \left(\overline{m}_0 - \underline{m}_0 + \frac{p}{p-1} \right).$$

Consequently, for i such that $\frac{1}{p^i} \left(\overline{m}_0 - \underline{m}_0 + \frac{p}{p-1} \right) - \frac{p}{p-1} > 0$, *i.e.*, for

$$i < \frac{1}{\ln p} \left(\ln \left(\overline{m}_0 - \underline{m}_0 + \frac{p}{p-1} \right) - \ln \left(\frac{p-1}{p} \right) \right),$$

we have $\overline{m}_i - \underline{m}_i > 0$. The fact that $\overline{m}_0 - \underline{m}_0$ is in $\Theta(N)$ gives the result. $\square$

The proof of the necessity for uniform vectors in order to get wave patterns is very similar to the proof of its reciprocal, Lemma 16, which told that uniform vectors imply wave pattern. We will use the contrapositive of the following statement.

Lemma 24. $\pi(N)_{[n\,,\infty[} \in (p \cdot \ldots \cdot 2 \cdot 1)^*\, 0\, (p \cdot \ldots \cdot 2 \cdot 1)^*\, 0^\omega$ *implies that* Y_n *is a uniform vector.*

Proof. On the left end of the configuration, for a column i inside the 0^ω part, we have $Y_{i+p} = {}^t(0, \ldots, 0)$. From right to left, let us see what happens when we encounter wave patterns. We have $Y_i = M\, Y_{i-1} + \frac{\pi(N)_{i-1}}{p} K$ (Equation 3), hence

$$Y_{i-1} = M^{-1}\, Y_i - \frac{\pi(N)_{i-1}}{p} L \text{ with } M^{-1} = \begin{pmatrix} -1 & \ldots & -1 & p \\ 1 & & 0 & 0 \\ & \ddots & & \\ 0 & & 1 & 0 \end{pmatrix} \text{ and } L = M^{-1}\, K = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ p \end{pmatrix}.$$

Let us immediately present the general case: we start from a uniform $Y_i = {}^t(\gamma, \ldots, \gamma)$ for some γ.

- If $\pi(N)_{i-1} = 0$ then $Y_{i-1} = {}^t(\gamma, \ldots, \gamma)$ and we are back to the same situation, this is the case when we encounter a 0 between two wave patterns.
- If $(\pi(N)_{i-1}, \pi(N)_{i-2}, \ldots, \pi(N)_{i-p}) = (1 \cdot \ldots \cdot p)$, *i.e.* we encounter a wave, then from the relation above:
  - $Y_{i-1} = {}^t(\gamma - 1, \gamma, \ldots, \gamma)$;
  - $Y_{i-2} = {}^t(\gamma - 1, \gamma - 1, \gamma, \ldots, \gamma)$;
  - $Y_{i-3} = {}^t(\gamma - 1, \gamma - 1, \gamma - 1, \gamma, \ldots, \gamma)$;
  - *et cetera*, and eventually $Y_{i-p} = {}^t(\gamma - 1, \ldots, \gamma - 1)$ is a uniform vector.

This process goes on while we are on wave patterns: we always get back to uniform vectors. $\square$

Combining Lemmas 23 and 24 we get the corollary below.

Corollary 25. *For a fixed $p > 1$, the smallest n such that*

$$\pi(N)_{[n\,,\infty[} \in (p \cdot \ldots \cdot 2 \cdot 1)^*\, 0\, (p \cdot \ldots \cdot 2 \cdot 1)^*\, 0^\omega$$

is in $\Omega(\log N)$.

The upper bound theorem (Theorem 22) and Corollary 25 give the final result.

Theorem 26. *For a fixed p, the smallest n such that*

$$\pi(N)_{[n\,,\infty[} \in (p \cdot \ldots \cdot 2 \cdot 1)^*\, 0\, (p \cdot \ldots \cdot 2 \cdot 1)^*\, 0^\omega$$

is in $\Theta(\log N)$.

3.6 An illustrative example

Figure 10 presents some representations of $\pi(2000)$ for $p = 4$ used in the developments of the paper: differences of shot vector (Figure 10a), shot vector (Figure 10b) and height (Figure 10c).

$$\begin{aligned}\pi(2000) = \;& (4,0,4,1,3,2,4,1,1,3,4,3,4,2,0,1,4,2,2,1,\\ & 4,3,2,1,0,4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,0^\omega)\end{aligned}$$

We can notice on Figure 10a that the shot vector differences contract towards some "steps" of length p, which corresponds to the statement of the convergence lemma (Lemma 13) that the vector Y_n becomes uniform exponentially quickly (note that this graphic plots the opposite of the values of the components of Y_n). Figure 10c pictures the sandpile configuration on which the wavy shape appears starting from column 20 over the 40 non-empty columns.

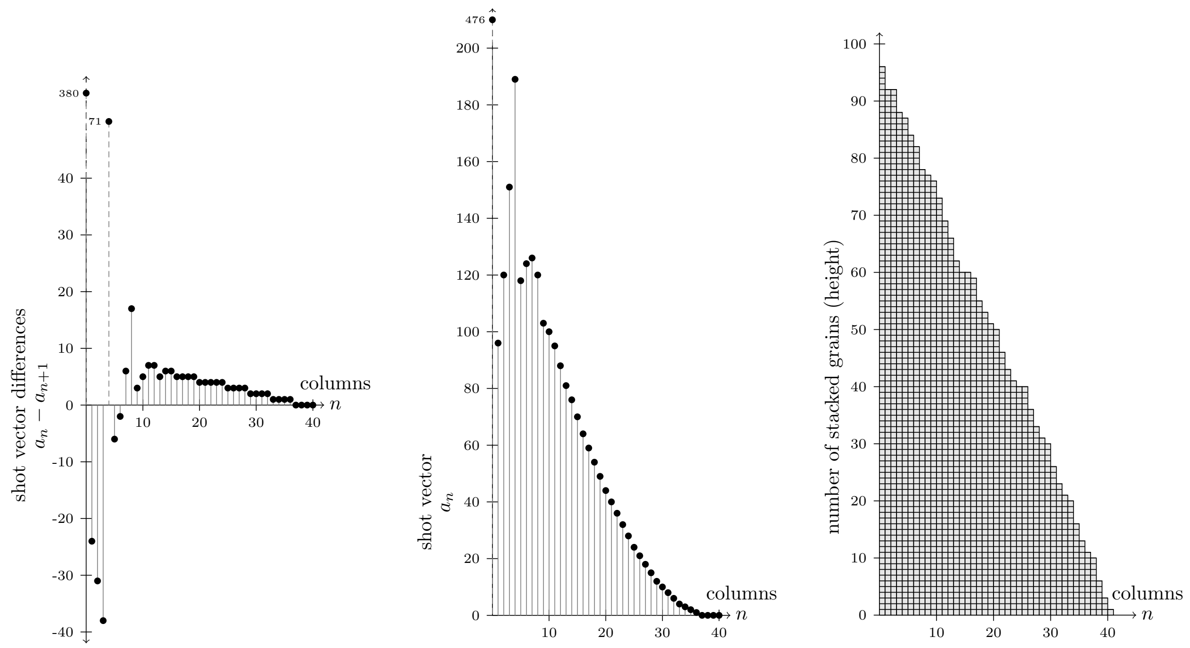
To see that wave patterns cover asymptotically completely the fixed point, let us briefly present the fixed point for $p = 4$ and $N = 20000$:

$$\begin{aligned}\pi(20000) = \;& (3,0,0,4,2,4,3,3,4,2,1,4,4,3,4,3,2,2,4,0,2,4,4,3,1,0,\\ & 4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,\\ & 4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,0,\\ & 4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,\\ & 4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,4,3,2,1,0^\omega)\end{aligned}$$

where the wavy shape appears starting from column 26 over the 126 non-empty columns.

4 Concluding Discussion

The proof technic we set up in this paper allowed to prove the emergence of regular patterns periodically repeated on fixed points, *stemming* from an initial segment presenting a disordered structure. The absolute size of this segment tends to infinity, and its



(a) $\pi(2000)$ for $p = 4$ represented as shot vector differences.

(b) $\pi(2000)$ for $p = 4$ represented by its shot vector.

(c) $\pi(2000)$ for $p = 4$ represented as stacked sand grains.

Figure 10: Representations of $\pi(2000)$ for $p = 4$.

relative size (compared to the width of the fixed point) tends to zero. Consequently, the result asymptotically completely describes the fixed point. Arguments of linear algebra allowed to prove a rough convergence of the system (in an imprecise but coarsely bounded representation), completed with the use of combinatorial arguments exploiting the discrete nature of the model, leading to the description and explanation of precise and regular wave patterns. Interestingly, this main result is proved without requiring a precise understanding of the initial segment's dynamic.

Let us argue again that the word *emergence*, though not formally defined, feels right to describe the phenomenon captured in the exact bound theorem (Theorem 26). As presented in Subsection 2.4, the fixed point configuration can be reached by adding sand grains one by one on the leftmost column: each grain addition triggers an avalanche (that may involve more grains in its way) traveling along the upper layer of the sandpile. During their movements, as they go rightward and downward, they *self-organize* into regular structures that do not appear immediately, but only after a transient of asymptotically unbounded size (though, relatively, very small). Since all the grains creating and maintaining the regular wave patterns on the right cross the left and weakly ordered part, it feels natural to say that the regularities *emerge* from the dynamic of sand grains: naively, some kind of order is destroyed (the initial column may be seen as extremely regular) and recreated in this process.

Trying to extend the proof technics of Theorem 26 (especially the upper bound) to other models is a promising track of research. It is easy to construct such models from the CFG representation of KSM. An example where similar regularities experimentally emerge is to apply the KSM rule for all $k \leqslant p$ when the height difference between two columns is strictly greater than $\sum_{k=1}^{p} k$. It would shade some light on the concept of *emergence* to find families of models of increasing expressiveness on which the technics allow to understand the asymptotic shape of fixed point, or the asymptotic dynamic.

It seems that an alternative proof can be derived from the developments of [24, 25], using others tools like harmonic functions and the least action principle. We think that the technic presented here is nonetheless of interest, as it clearly separates two ingredients: the continuous one linked to linear algebra arguments, and the discrete one linked to combinatorial arguments.

This result stresses the idea that sandpile models are on the edge between discrete and continuous systems. As we pour grains one by one, they start to self-organize and create regular wave structures. Hypothetically, the result suggests a separation of the discrete and continuous parts of the system. On one hand, there is a seemingly unordered initial segment, interpreted as reflecting the discrete behavior. On the other hand, the asymptotic and ordered part, interpreted as reflecting the continuous behavior, lets a regular and smooth pattern emerge. Can we interpret the *continuous* part as a liquid limit of the *discrete* part?

There is a slight bias appearing on the continuous part: it is not fully homogeneous (that is, with exactly the same slope at each index) which would have been expected for a continuous system, but a (very small) pattern is repeated. It looks like this bias does not come from the unicity of the initial column, because we still observe the appearance of wave patterns starting from variations of the initial configuration (for example starting from p consecutive columns of height N, thus $p\,N$ grains, or starting with a whole sandpile of wave patterns and adding grains one by one on the leftmost column). Rather, we think that this bias comes from the gap between the unicity of the *border* column on the left side (at index -1) compared to the rule which affects p columns.

Finally, the emergence of regularities in this system hints at a clear qualitative distinction between *some* sand grains and a *heap* of sand, providing a solution to the *sorite paradox* in the Kadanoff sandpile model: we have a *heap* as soon as a wave pattern appears.

Acknowledgments

The authors are thankful to the referees for their careful read of the paper and the improvements they proposed.

References

[1] P. Bak and K. Tang, C. Wiesenfeld. Self-organized criticality. *Phys. Rev. A*, 38(1):364–374, Jul 1988.

[2] A. Björner, L. Lovász, and P. W. Shor. Chip-firing games on graphs. *Eur. J. Comb.*, 12(4):283–291, 1991.

[3] A. Bjöer and L. Lovász. Chip-firing games on directed graphs. *Journal of Algebraic Combinatorics*, 1:305–328, 1992.

[4] N. Bourbaki. *General Topology.* Elements of mathematics. Springer, 1998.

[5] M. Creutz. Cellular automata and self organized criticality. In *in Some New Directions in Science on Computers*, 1996.

[6] A. Dartois and C. Magnien. Results and conjectures on the sandpile identity on a lattice. In *Discrete Model for C.S.*, pages 89–102. Discrete Math. and T.C.S., 2003.

[7] D. Dhar. Self-organized critical state of sandpile automaton models. *Phys. Rev. Lett.*, 64:1613–1616, 1990.

[8] J. O. Durand-Lose. Parallel transient time of one-dimensional sand pile. *Theor. Comput. Sci.*, 205(1-2):183–193, 1998.

[9] K. Eriksson. *Strongly convergent games and coexeter groups.* PhD thesis, KTH Stockholm, 1993.

[10] E. Formenti, B. Masson, and T. Pisokas. Advances in symmetric sandpiles. *Fundam. Inform.*, 76(1-2):91–112, 2007.

[11] E. Formenti, K. Perrot, and E. Rémila. Computational complexity of the avalanche problem on one dimensional kadanoff sandpiles. *LNCS*, 8996:21–30, 2014. AUTOMATA'2014.

[12] E. Formenti, T. Van Pham, T. H. D. Phan, and T. T. H. Tran. Fixed point forms of the parallel symmetric sandpile model. *Theor. Comput. Sci.*, 533:1–14, 2014.

[13] A. Gajardo and E. Goles. Crossing information in two-dimensional sandpiles. *Theor. Comput. Sci.*, 369(1-3):463–469, 2006.

[14] A. Gajardo, A. Moreira, and E. Goles. Complexity of Langton's ant. *Discrete Applied Mathematics*, 117(1-3):41–50, 2002.

[15] D. Gale, J. Propp, S. Sutherland, and S. Troubetzkoy. Further travels with my ant. *Mathematical Entertainments column, Mathematical Intelligencer*, 17:48–56, 1995.

[16] R. B. Gardner and N. K. Govil. Some generalizations of the Eneström-Kakeya theorem. *Acta Math. Hungar.*, 74(1-2):125–134, 1997.

[17] E. Goles and M.A. Kiwi. Games on line graphs and sand piles. *Theor. Comput. Sci.*, 115:321–349, 1993.

[18] E. Goles, M. Latapy, C. Magnien, M. Morvan, and T. H. D. Phan. Sandpile models and lattices: a comprehensive survey. *Theor. Comput. Sci.*, 322(2):383–407, 2004.

[19] E. Goles and B. Martin. Computational Complexity of Avalanches in the Kadanoff Two-dimensional Sandpile Model. In *Proc. of JAC 2010*, pages 121–132, 12 2010.

[20] E. Goles, M. Morvan, and T. H. D. Phan. The structure of a linear chip firing game and related models. *Theor. Comput. Sci.*, 270(1-2):827–841, 2002.

[21] S. Grauwin, É. Bertin, R. Lemoy, and P. Jensen. Competition between collective and individual dynamics. *Nat. Ac. of Sciences USA*, 106(49):20622–20626, 2009.

[22] J. E. Hopcroft, R. Motwani, and J. D. Ullman. *Introduction to automata theory, languages, and computation - international edition (2. ed).* Addison-Wesley, 2003.

[23] L. P. Kadanoff, S. R. Nagel, L. Wu, and S. Zhou. Scaling and universality in avalanches. *Phys. Rev. A*, 39(12):6524–6537, Jun 1989.

[24] L. Levine, W. Pegden, and C. K. Smart. Apollonian structure in the Abelian sandpile. arXiv:1208.4839, 2012.

[25] L. Levine, W. Pegden, and C. K. Smart. The Apollonian structure of integer superharmonic matrices. arXiv:1309.3267, 2013.

[26] L. Levine and Y. Peres. Spherical asymptotics for the rotor-router model in z d. *Indiana Univ. Math. J*, pages 431–450, 2008.

[27] P. B. Miltersen. The computational complexity of one-dimensional sandpiles. *Theory Comput. Syst.*, 41(1):119–125, 2007.

[28] C. Moore and M. Nilsson. The computational complexity of sandpiles. *Journal of Statistical Physics*, 96:205–224, 1999.

[29] K. Perrot, T. H. D. Phan, and T. Van Pham. On the set of Fixed Points of the Parallel Symmetric Sand Pile Model. *DMTCS*, pages 17–28, 2011. AUTOMATA'2011.

[30] K. Perrot and É. Rémila. Avalanche Structure in the Kadanoff Sand Pile Model. *LNCS*, 6638:427–439, 2011. LATA'2011 arXiv:1101.5940.

[31] K. Perrot and É. Rémila. Transduction on Kadanoff Sand Pile Model Avalanches, Application to Wave Pattern Emergence. *LNCS*, 6907:508–520, 2011. MFCS'2011.

[32] K. Perrot and É. Rémila. Emergence on decreasing sandpile models. *LNCS*, 9234:419-431, 2015. MFCS'2015.

[33] K. Perrot and É. Rémila. Kadanoff sand pile model. Avalanche structure and wave shape. *Theor. Comput. Sci.*, 504:52–72, 2013.

[34] T. H. D. Phan. Two sided sand piles model and unimodal sequences. *Theor. Inf. Appl.*, 42(3):631–646, 2008.

[35] W. Weaver. Science and Complexity. *American Scientist*, 36(536), 1948.