Claremont Colleges
Scholarship @ Claremont

WM Keck Science Faculty Papers | W.M. Keck Science Department

6-11-2007

Nonlinear Dynamics in Combinatorial Games: Renormalizing Chomp

Eric J. Friedman
Cornell University

Adam S. Landsberg
Claremont McKenna College; Pitzer College; Scripps College

Recommended Citation
Friedman, E.J., and A.S. Landsberg. "Nonlinear Dynamics in Combinatorial Games: Renormalizing Chomp." Chaos 17.2 (2007): 023117. DOI: 10.1063/1.2725717

This Article is brought to you for free and open access by the W.M. Keck Science Department at Scholarship @ Claremont. It has been accepted for inclusion in WM Keck Science Faculty Papers by an authorized administrator of Scholarship @ Claremont. For more information, please contact scholarship@cuc.claremont.edu.

Nonlinear dynamics in combinatorial games: Renormalizing Chomp

Eric J. Friedman
School of ORIE, Cornell University, Ithaca, New York 14853

Adam Scott Landsberg
Joint Science Department, Claremont McKenna, Pitzer, and Scripps Colleges, Claremont, California 91711

(Received 8 January 2007; accepted 20 March 2007; published online 11 June 2007)

We develop a new approach to combinatorial games that reveals connections between such games and some of the central ideas of nonlinear dynamics: scaling behaviors, complex dynamics and chaos, universality, and aggregation processes. We take as our model system the combinatorial game Chomp, which is one of the simplest in a class of "unsolved" combinatorial games that includes Chess, Checkers, and Go. We discover that the game possesses an underlying geometric structure that "grows" (reminiscent of crystal growth), and show how this growth can be analyzed using a renormalization procedure adapted from physics. In effect, this methodology allows one to transform a combinatorial game like Chomp into a type of dynamical system. Not only does this provide powerful insights into the game of Chomp (yielding a complete probabilistic description of optimal play in Chomp and an answer to a longstanding question about the nature of the winning opening move), but more generally, it offers a mathematical framework for exploring this unexpected relationship between combinatorial games and modern dynamical systems theory. © *2007 American Institute of Physics*. [DOI: 10.1063/1.2725717]

Combinatorial games, which include Chess, Go, Checkers, Chomp, Dots-and-Boxes, and Nim, have both captivated and challenged mathematicians, computer scientists, and players alike. These are two-player games with no randomness (e.g., no rolling of dice or dealing of cards) and no hidden information (unlike poker, for instance). Apart from their obvious entertainment value, combinatorial games have long been the subject of much serious study owing to the deep mathematical questions they raise in areas such as computational complexity, graph theory, and surreal numbers.[1–3] Using the game of Chomp as a prototype, we report here on a new geometrical approach which unveils unexpected parallels between combinatorial games and some of the central ideas of dynamical systems theory, most notably notions of scaling, renormalization, universality, and chaotic attractors. We show in particular that the game of Chomp can behave in ways analogous to a chaotic dynamical system. This insight and subsequent analysis allows us not only to answer a number of open questions about the game of Chomp, but also provides a new perspective on complex combinatorial games more generally and offers a mathematical framework for understanding this connection between combinatorial games and dynamical systems. Our central finding is that underlying the game of Chomp is a geometric structure that encodes essential information about the game, and that this structure exhibits a type of spatial scale invariance: Loosely speaking, the geometry of "small" and "large" winning positions in the game look the same, after rescaling. We demonstrate that the geometries on different spatial scales can be related to one another via a set of recursion operators, which in turn allows us to recast the game as a type of dynamical system. We then analyze this "dynamical system" using tools and concepts from nonlinear dynamics theory—most notably, a suitably adapted renormalization technique to analyze the scale invariance.

I. INTRODUCTION

Introduced over 30 years ago,[4,5] the combinatorial game Chomp is an ideal candidate for study, since it represents one of the simplest of the unsolved (and presumably "hard") combinatorial games. Its history is marked by some significant theoretical advances,[5–9] but it has yet to succumb to a complete analysis, and numerous intriguing features and open questions surround the game.[10] For instance, while it is possible to show (via a clever strategy-stealing argument by Gale[5]) that there always exists a winning opening move for the first player in Chomp, the determination of what that winning move is (or even whether it is unique) has remained elusive. Standard mathematical techniques in combinatorial game theory[2,3,11,12] have proven largely ineffective in analyzing Chomp due to its intrinsic indecomposability. The renormalization analysis of Chomp presented here will provide an explanation for virtually all numerical observations about the game previously reported in the literature, including some recent conjectures by Brouwer,[10] and, as we will show, leads directly to the discovery of new properties of optimal play in Chomp. For example, our technique will yield a precise probabilistic description and the asymptotic location of the long-sought-after winning opening move of the game. Moreover, the renormalization approach reveals hidden connections between a game's underlying geometric structure and some key ideas of nonlinear dynamics, allowing us to discuss for the first time the notions of universality and sensitivity to initial conditions in the context of combinatorial games.

 © 2007 American Institute of Physics

Our work follows a series of recent papers by Zeilberger[6,8] and Sun,[7] which examined three-row Chomp, a special case of Chomp. Three-row Chomp, as Zeilberger points out, is one of the simplest of combinatorial games that has resisted substantial analysis (two-row Chomp being trivial). In those works, three-row Chomp is shown to be remarkably complex, but also displays intriguing patterns and regularities. Brouwer[10] has provided an extensive listing of many of these properties, and Zeilberger discusses the computational complexity of solving Chomp[6] and whether the game is compressible in the sense of Chaitin.[13] Zeilberger[8] even suggests that three-row Chomp appears to "have 'chaotic' behavior, but in a vague, yet-to-be-made-precise, sense." As we will see, our approach provides a natural framework for discussing such questions. For instance, our results suggest that three-row Chomp displays some of the features of a chaotic attractor and that it appears to be partially compressible, i.e., being neither completely solvable nor completely incompressible.

Our key technical tool, renormalization, is adapted from a broad class of techniques originally used to study phase transitions in statistical mechanics,[14,15] and which now permeates most branches of modern physics, nonlinear dynamics, and chaos theory (see, e.g., Refs. 16 and 17). These techniques are widely used to analyze systems that exhibit "scaling behavior," i.e., similarity on different spatial scales. Renormalization has met with great success throughout physics and provides a standard technique for computing many system properties. However, in spite of their success, renormalization techniques are generally not rigorous from the standpoint of formal mathematical proof, and the renormalization approach we develop here is no exception in this regard. Nonetheless, as we hope to show here, the simplicity of the explanations it offers about the game of Chomp, its agreement with numerical simulations, its predictive capabilities, and the connections it draws with fundamental concepts from nonlinear dynamics theory, speak to its utility as a new tool for analyzing games (despite its nonrigorous status at present).

Lastly, we remark that one of the surprising (but essential) features of this renormalization methodology is that it incorporates probabilistic elements into the analysis even though the game of Chomp is itself completely deterministic. Specifically, from elementary game theory it is well known that every position in a game like Chomp can be uniquely classified as being either a "winner" or "loser" (also called "*N*" and "*P*" positions, respectively). Nonetheless, in our formulation, to every such position we assign a *probability* that it is a winner or loser. As we will see, this probabilistic characterization of the game's winners and losers allows one to capture the overall structure of the geometric patterns underlying the game while bypassing consideration of the precise (computationally hard to determine) locations of these winning and losing positions.

II. THE GAME OF CHOMP

The rules of Chomp are easily explained. In the general case, play begins with an $M \times N$ array of counters [Fig. 1(a)], with the counter in the southwest corner being "poison." On each turn a player selects a counter and removes it along with all counters lying northeast of it [Fig. 1(b)]. Play alternates between the two players until one player takes the last (i.e., poison) counter, thereby losing the game. As noted previously, a fundamental result of Gale's is that the first player to move in Chomp can always win[5] (under optimal play), yet what this first player's opening move should be is not known in general. The argument behind this is quite elegant: Consider the "nibble" move, which removes the counter in the northeast corner. Either this is a winning opening move (i.e., it leads to a win under optimal play) or it is not. If it is a winning move, then we are done. If it is not a winning move, then the second player must have a winning response. However, in this case, the first player could have chosen this response move as her opening move instead of the nibble move, leading to a win. This form of argument, known as strategy stealing, demonstrates the existence of a winning opening move for the first player, but, remarkably, provides no information as to what this winning move actually is.

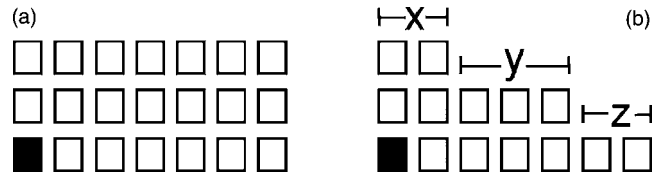

FIG. 1. The game of Chomp. (a) Play begins with an $M \times N$ rectangular array of counters (three-row Chomp is illustrated). The dark counter in the lower-left corner is "poison." (b) A sample game configuration after play has begun. The state of the game at any stage of play is specified by coordinates $[x,y,z]$.

In this paper, we will focus exclusively on the case of three-row (i.e., $3 \times N$) Chomp (generalizations to four-row Chomp and higher are analogous). In three-row Chomp, the configuration of the counters at any stage of play can be conveniently described in Zeilberger's coordinates[6] by the triplet of integers $[x,y,z]$, where the x,y,z specify the number of columns of height three, two, and one, respectively [Fig. 1(b)]. We will refer to each such triplet $[x,y,z]$ as a "position" of the game. Under the game rules, from an arbitrary position $p=[x,y,z]$ one may move to any one of the following positions:

$$[x,y-t,z+t] \quad 0 < t \leq y \quad \text{(M1)},$$

$$[x,y-t,0] \quad 0 < t \leq y \quad \text{(M2)},$$

$$[x,y,z-t] \quad 0 < t \leq z \quad \text{(M3)},$$

$$[x-t,y+t,z] \quad 0 < t \leq x \quad \text{(M4)},$$

$$[x-t,0,z+y+t] \quad 0 < t \leq x \quad \text{(M5)},$$

$$[x-t,0,0] \quad 0 < t \leq x \quad \text{(M6)}.$$

We refer to these positions as the "children" of position p. By inverting these relations, the "parents" of position p, i.e., the set of positions from which it is possible to reach p in a single move, are also readily determined.

Every position in the game may be uniquely classified as either an N position (**n**ext player to move can win), if a player starting from that position can always force a win under optimal play, or as a P position (**p**revious player wins) otherwise. This classification is well defined by Zermelo's theorem. (In this paper we will at times refer to N positions as winners, and the P positions as losers, in keeping with terminology introduced previously by Zeilberger in his study of Chomp.) Normally, one does not know (without exhaustive computation) whether a given position is N or P. (Gale's simple strategy-stealing argument showing that the starting position $[x,0,0]$ is an N position is exceptional in this regard.)

Before proceeding, we mention three elementary but useful facts from combinatorial game theory:

(F1) The children of any N position must include at least one P position;
(F2) the parents of any P position must all be N positions;
(F3) if all children of a position are N positions, then the position must be a P position.

A player who starts at an N position can, in theory, always win a game via a simple strategy: on each turn, the player should move the opponent into a P position. Repeated application of this strategy will eventually force the opponent to take the poison counter. (For this reason, a legal move from an N position to a P position is called a "winning move.") The main difficulty with implementing such a simple strategy is, of course, that for computationally challenging games like Chomp, players typically do not generally know which positions are P and which are N.

III. RENORMALIZATION

A. The renormalization framework: Overview

In the following sections we show how the analysis of the combinatorial game can be recast and transformed into a type of renormalization problem commonly seen in physics. The basic idea is as follows: Consider the abstract, three-dimensional space of all positions $[x,y,z]$ in the game of Chomp, and imagine marking all P positions in this space. We will refer to the set of all P positions as the "P set" of the game. It represents a critical entity: Complete knowledge of a game's P set would imply that the game has been effectively "solved," in the sense that a winning strategy for a player at an N position would be to simply move the opponent to any accessible (known) P position, as noted earlier.

The renormalization methodology is based on the discovery that the P set forms a type of "geometric object" (in position space). For complex games like Chomp, we find that this object is neither a diffuse set of seemingly randomly scattered points, nor is it entirely well structured with an easily characterized order, but rather a combination of the two. Informally, the P set possesses a well-defined global geometric structure but locally exhibits disorder. The renormalization analysis described here allows one to precisely characterize this global geometry, and handles the disorder through a probabilistic treatment. What makes this all possible is the observation that the P set's geometry exhibits a type of (linear) scaling behavior: if one slices the P set with a series of appropriately defined two-dimensional planes ("sheets") that foliate position space, the patterns on these different sheets are geometrically similar to one another up to an overall scale factor. Analytical recursion operators that relate these (scale-invariant) sheets to one another can be constructed, and, via the renormalization process, the asymptotic geometric structure of the sheets can ultimately be determined.

B. Construction of sheets

Two distinct types of sheets will play a crucial role in the analysis. The first are the "loser sheets" $\{L_0,L_1,L_2,\ldots\}$, which mark the location of the P positions of the game, grouped according to their x values; e.g., loser sheet L_2 describes all losing positions having an x value of 2. (Here, we are reverting to Zeilberger's terminology of calling P positions losers and N positions winners.) Formally, the loser sheet L_x is defined as an infinite, two-dimensional matrix whose (y,z)-th component is a 1 if position $[x,y,z]$ is a P position, and a 0 otherwise. Taken together, the loser sheets completely characterize the game's P set. (Informally, we think of $L_0,L_1,L_2,\ldots$ as forming a set of stacked sheets, ordered according to their "x level.")

A direct analysis of these loser sheets is not feasible, for reasons that will become apparent shortly. Rather, it is necessary to first introduce a special subclass of the game's winning (i.e., N) positions, so-called "instant winners" (this terminology is borrowed from Zeilberger[6]). As we will show, these instant winners essentially encode all critical information about the game—from them we can construct the loser sheets and hence the game's P set. An N position $[x,y,z]$ is defined to be an instant winner if, from that position, a player can legally move to a P position with a smaller x value; i.e., $[x,y,z]$ is an instant winner if there is some accessible loser $[x',y',z']$ with $x'<x$. Just as was done for the losers, it is useful to take two-dimensional slices through the (three-dimensional) space in which the instant winners live; we thus define an instant-winner sheet W_x to be the infinite, two-dimensional matrix marking the location of all instant winners with the specified x value, i.e., the (y,z)-th component of matrix W_x is a 1 if position $[x,y,z]$ is an instant winner, and a 0 otherwise. (The instant-winner sheets considered here are closely related to, and indeed inspired by, a similar set of objects first studied by Zeilberger;[6] however, the present formulation is required for the renormalization analysis which follows.)

C. Operators and recursion relations

In this section we demonstrate that the game's (loser/instant-winner) sheets are related to one another via a set of operators. Specifically, we derive here the two key relations that will form the basis of our renormalization analysis: (a) $W_{x+1}=\mathbf{R}W_x$ and (b) $L_x=\mathbf{M}W_x$, with $\mathbf{R}$ (a recursion operator) and $\mathbf{M}$ (a "supermex" operator) to be defined shortly. The significance of these relations is (a) that the entire set of instant-winner sheets $\{W_1,W_2,\ldots\}$ can be recursively generated, starting from the lowest-level (initial) sheet W_0, by re-

peated applications of the operator **R**, and (b) that once an instant-winner sheet at a given x level is determined, then the corresponding loser sheet at the same x level can be determined using the **M** operator. (This is the basis of our earlier claim that the instant-winner sheets encode all critical information about the loser sheets, and hence the game.) We mention that these two operator relations (whose derivations are given below) are both analytically exact relations. It is only much later, after we have begun to analyze these relations, that certain nonrigorous aspects of our renormalization scheme will be introduced.

The derivation begins by first defining a set of basic operators which can act on sheets (here, sheets can refer to the loser sheets, the instant-winner sheets, or, more generally, to any semi-infinite, two-dimensional matrix consisting of 0's and 1's); these operators will be used to construct the desired recursion and supermex operators. In what follows we denote a general two-dimensional sheet by A, and its (y,z)-th component by $A(y,z)$, where y specifies the column of matrix A, and z its row. The labeling begins with zero (i.e., $y,\ z \in \{0,1,2,\ldots\}$); $y=0$ refers to the leftmost column and $z=0$ to the bottom-most row:

- Define **I** to be the *identity* operator: $\mathbf{I}A=A$ for any sheet A.
- Define **L** to be the *left-shift* operator, which shifts all elements of a sheet to the left: $(\mathbf{L}A)(y,z)=A(y+1,z)$ (i.e., the effect of **L** on A is to eliminate the leftmost column of A).
- Define an *addition* operator + on sheets by the logical OR: Given two sheets A and B, define $(A+B)(y,z)=1$ if either $A(y,z)=1$ or $B(y,z)=1$.
- Define the *diagonal element-adding* operator **D**, which acts on loser sheets, as follows: Note first that for each loser sheet L_x there is a unique losing position with $y=0$, i.e., there is a unique z value [denoted $z^*(x)$] such that $L_x(0,z^*(x))=1$. [Such a loser must exist since the set of instant winners is bounded in z; uniqueness holds by (F2) in conjunction with (M3).] Now define the action of operator **d** on loser sheet L_x as follows: $\mathbf{d}L_x$ is a matrix, which is zero everywhere, except at the following points: $(\mathbf{d}L_x)(t,z^*(x)-t)=1$ for all $0\le t\le z^*(x)$. [Geometrically, the nonzero components of matrix $\mathbf{d}L_x$ form a diagonal line of 1's extending downwards and to the right at a 45° angle from the losing entry at $(0,z^*(x))$.] Finally, the desired diagonal element-adding operator **D** is defined by $\mathbf{D}=\mathbf{I}+\mathbf{d}$. Geometrically, $\mathbf{D}L_x$ is thus identical to L_x except for the addition of the diagonal row of 1's. Henceforth, we will call the diagonal elements associated with **D** the "deadly diagonals" (since, as will become clear, they tend to complicate the analysis considerably).

Using the above operators, we next derive a relationship between a given instant winner sheet W_x and the set of loser sheets $\{L_{x-1},L_{x-2},\ldots,L_0\}$. Recall that, by definition, position $p=[x,y,z]$ is an instant winner in W_x [i.e., $W_x(y,z)=1$] if there exists a P position $[x',y',z']$ at a lower x level (i.e., with $x'<x$) that is reachable from p in a single move. So, to construct W_x we need only determine which positions $[x,y,z]$ are capable of reaching a P position in $\{L_{x-1},L_{x-2},\ldots,L_0\}$ (i.e., we seek the parents, at level x, of the P positions in sheets $\{L_{x-1},L_{x-2},\ldots,L_0\}$). For this purpose, only game moves (M4)–(M6) (see Sec. II) are relevant [since moves (M1)–(M3) maintain the same x value]. From (M4) (or its inverse), it is easily verified that the terms $\mathbf{L}L_{x-1}$, $\mathbf{L}^2L_{x-2},\ldots,\mathbf{L}^xL_0$ will all contribute to W_x. (Here, $\mathbf{L}^n$ denotes repeated application of the left-shift operator n times.) From (M5) (and its inverse), terms $\mathbf{L}\mathbf{d}L_{x-1},\mathbf{L}^2\mathbf{d}L_{x-2},\ldots,\mathbf{L}^x\mathbf{d}L_0$ will contribute. There are no contributions from rule (M6) because, by Gale's argument, positions of the form $[x,0,0]$ are never P positions (for any $x>0$). Thus, combining all contributions we have

$$W_x=\sum_{t=1}^{t=x}\mathbf{L}^t(\mathbf{I}+\mathbf{d})L_{x-t}=\sum_{t=1}^{t=x}\mathbf{L}^t\mathbf{D}L_{x-t}. \tag{1}$$

We next demonstrate that instant-winner sheet W_x contains all the necessary information for constructing the corresponding loser sheet L_x. Specifically, we will construct an operator **M** (dubbed "supermex") and show that

$$L_x=\mathbf{M}W_x. \tag{2}$$

As will be described more fully below, the supermex operator **M** is a generalization of the standard mex ("minimal excluded value") operator defined on sets of non-negative integers [e.g., mex $(\{0,1,2,5,7\})=3$; mex$(\{1,4\})=0$].

We begin our derivation of relation (2) with a preliminary observation: If we rank all positions in the game by size using the standard dictionary ordering [i.e., define $[x,y,z]>[x',y',z']$ if either $(x>x')$ or $(x=x'$ and $y>y')$ or $(x=x',\ y=y'$, and $z>z')$], then as play progresses the size of successive positions will strictly decrease, as may be verified by considering the game rules (M1)–(M6) (i.e., children are always smaller than their parents). Now the basic intuition behind (2) is as follows.

Start with the instant-winner sheet W_x, and locate the smallest position which is not an instant winner (i.e., the smallest position which is not marked by a 1 in the W_x matrix). Call this position q. q must be a P position, as shown by the following argument: Since q is not an instant winner, then by definition none of its children can be P positions with a lower x value. Moreover, none of its children can be P positions with the same x value either, since q is (by construction) the *smallest* noninstant winner in W_x, and so its children in W_x (which under the dictionary ordering must be smaller in size) can only be instant winners. Hence, all the children of q are N positions, and by (F3) of Sec. II it follows that q is a P position. Having thus identified the first P position in L_x, we can then find the next P position as follows: Using rules (M1)–(M3), find all parents of q which are in the same sheet as q (i.e., same x level), and imagine marking these positions in the W_x matrix (by setting the appropriate matrix elements equal to 1). In practice, we do not actually want to alter W_x itself [since it is well defined via (1)], so instead we create a duplicate sheet T_x given by $T_x=W_x$, and make the alterations to T_x. Having marked the parents of q (in T_x), we again search for the smallest remaining unmarked position in T_x (excluding q itself). By the same argument used above, this smallest position must be a P position. Through this iterative process (of marking the parents of P positions and finding the smallest unmarked position which

remains), we can generate all P positions in L_x. This iterative procedure defines the supermex operator $\mathbf{M}$.

More formally, we algorithmically define the action of $\mathbf{M}$ on an instant-winner sheet W_x to generate L_x [Eq. (2)] as follows:

Supermex algorithm (steps):

(1) Set $L_x=0$ [i.e., $L_x(y,z)=0$ for all $y,z \in \{0,1,2,\ldots\}$];
(2) set $T_x=W_x$;
(3) set $y=0$ (i.e., we will start with the first column of T_x);
(4) let $z_{\text{small}}(y)=\text{mex}(\{z|T_x(y,z)=1\})$ (i.e., find the z value of the smallest unmarked position in T_x);
(5) set $L_x(y,\ z_{\text{small}}(y))=1$ (i.e., mark this point as a P position in the L_x matrix);
(6) set $T_x(y,t)=1$ for all $t \geq z_{\text{small}}(y)$ and set $T_x(y+t,\ z_{\text{small}}(y)-t)=1$ for all $t \leq z_{\text{small}}(y)$;
(7) set $y \rightarrow y+1$;
(8) if $z_{\text{small}}(y)=0$, stop; else go to step 4.

Several remarks are in order. (i) In step 6, one finds the parents of the current smallest P position via rules (M3) and (M1) and marks them. (Borrowing Zeilberger's terminology, we refer to these parents as "implied winners," in contrast to the instant winners.) Geometrically, the implied winners associated with a given P position fill up the entire column in the matrix directly above the P position [by rule (M3)], while the implied winners from rule (M1) mark a diagonal line in the matrix running down and to the right from the P position at a 45° angle. This diagonal line will prove important for later considerations, and in support of Zeilberger's terminology, we will refer to this diagonal as an "implied diagonal." (Implied diagonals are distinct from the "deadly diagonals" associated with the operator $\mathbf{D}$ discussed previously.) (ii) If $z_{\text{small}}(y)=0$ (step 8), then by rule (M2) no other P positions in L_x exist and the search terminates. (iii) If $z_{\text{small}}(y) \neq 0$ for any y, then the computation will not terminate. (As we show later, the probability of it terminating is $\gamma=\sqrt{2}-1$.) However, Byrnes[9] has shown that even when it does not terminate it does eventually becomes periodic and hence predictable.

Now, combining our expressions for W_x and L_x [Eqs. (1) and (2)], we have

$$W_x = \sum_{t=1}^{t=x} \mathbf{L}^t \mathbf{D} \mathbf{M} W_{x-t}.$$

Finally, we make the algebraic observation that if we create W_{x+1} by substituting $x \rightarrow x+1$ into the above expression and then compare it to the original expression for W_x, it becomes possible to re-express W_{x+1} in terms of W_x as follows:

$$W_{x+1} = \mathbf{L}(\mathbf{I}+\mathbf{DM})W_x \equiv \mathbf{R}W_x, \tag{3}$$

where we have defined the recursion operator

$$\mathbf{R} \equiv \mathbf{L}(\mathbf{I}+\mathbf{DM}). \tag{4}$$

This recursion relation (3) along with our earlier result (2) are the key operator relations that will prove fundamental to our analysis.

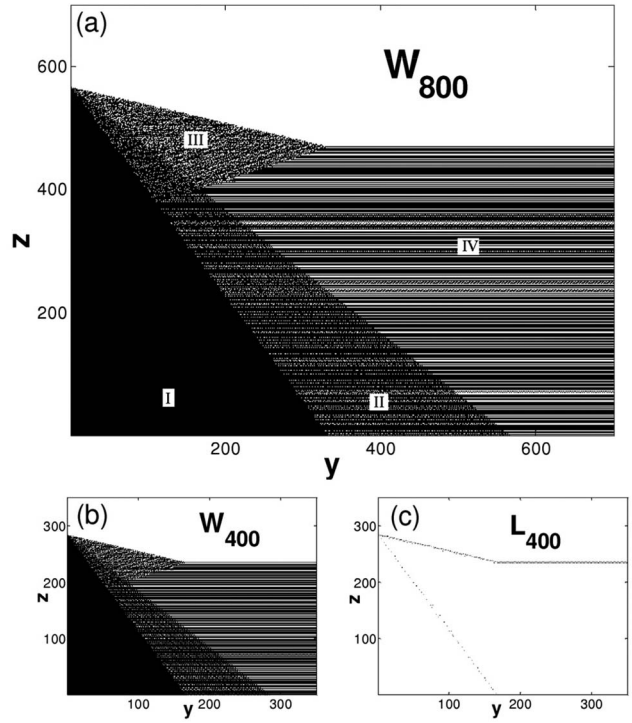

FIG. 2. The geometry of Chomp. (a) The instant-winner-sheet geometry for three-row Chomp, shown for $x=800$. The instant-winner locations in the y-z plane are marked in black. Distinct regions are labeled I-IV. (b) The instant-winner sheet for $x=400$. Comparison of W_{400} to W_{800} highlights the central scaling property of the instant-winner sheets $\{W_x\}$: as they "grow" with increasing x, they remain geometrically identical up to a scale factor; i.e., shapes, densities, and boundary-line slopes remain fixed (though the point-by-point location of the instant winners varies from one sheet to the next). (c) The loser-sheet geometry L_x, shown for $x=400$, marking the location of the game's P positions at the specified x level. The loser sheets, characterized by several distinct but slightly diffuse lines, also exhibit geometric scale invariance—i.e., they grow linearly with increasing x, but their overall structure is preserved. As described in the text, this invariant geometric structure can be characterized by six parameters: m_L, λ_L, m_U, λ_U, α, and γ. (The renormalization analysis to follow will yield precise values for these parameters.) Note that the three lines making up the loser sheets lie along boundaries of the associated instant-winner sheet (compare L_{400} with W_{400}). Note also the (narrow) scatter of points around these lines of losers.

D. Numerical analysis

We can now use the (analytically exact) operator relations from the preceding section to construct the game's sheets. Our first insight comes from numerically iterating relation (3) to generate a visual representation of the instant-winner sheets $\{W_x\}$ for various x values. [This is carried out by starting with sheet W_0, and repeatedly applying the recursion operator $\mathbf{R}$ until the desired x level is reached. Algorithmically, the recursion operator in (4) is constructed from the supermex algorithm $\mathbf{M}$, the diagonal element-adding operator $\mathbf{D}$, and the left-shift operator $\mathbf{L}$ defined in the preceding section.] Figures 2(a) and 2(b) show the structure of W_{800} and W_{400}, respectively, and are representative of what is observed at other x values. Each sheet exhibits a nontrivial internal structure characterized by several distinct regions: a solid (filled) triangular region at the lower left (region I), a series of horizontal bands extending to the right (region IV), and two other triangular regions of different densities (regions II and III). Most striking and significant, however, is

the observation that the entire set of instant-winner sheets $\{W_1, W_2, W_3, \ldots\}$ possesses a remarkable scaling property: the overall geometric structure from one sheet to the next is identical, up to a scaling factor [e.g., compare Figs. 2(a) and 2(b)]. In particular, as x increases, the boundary-line slopes, densities of points, and shapes of the various regions are preserved from one sheet to the next (although the actual point-by-point locations of the instant winners within a sheet will differ from one sheet to the next, a point we will discuss later). Hence, upon rescaling, the overall geometric structure of these sheets is identical (in a probabilistic sense).

We mention here that if one scans the instant-winner sheets W_x's (in the direction of increasing x), the "growth" of the observed geometric pattern shares some similarities to certain crystal-growth and aggregation processes found in physics: in both cases, one observes structures that grow through the accumulation of new points along existing boundaries, with the overall geometric integrity of the pattern being preserved during the growth process. We will return briefly to this point later in the paper.

The loser sheets $\{L_x\}$ can be numerically constructed in a similar manner via relation (2); their characteristic geometry is revealed in Fig. 2(c). It is found to consist of three (diffuse) lines: a lower line of slope m_L and density of points (per unit y) λ_L, an upper line of slope m_U and density λ_U, and a flat line extending to infinity. The upper and lower loser lines originate from a point whose height (i.e., z value) is αx. The flat loser line (with density one) is only present in some loser sheets, depending on their x values; the probability that it will be present in sheet L_x for randomly chosen x is denoted γ. The flat loser line, when present, begins at the same horizontal location as where the lower loser line ends. Like the instant winner sheets, the loser sheets exhibit a fundamental geometric scaling property: as x increases, the geometric structure of L_x grows linearly, but its overall form remains unchanged [the only caveat being that, as noted above, the flat line seen in Fig. 2(c) is absent in some of the loser sheets]. We emphasize that while these are only numerical findings, they will provide critical intuition for the results that follow. In particular, we will describe shortly how the precise values of the six key parameters m_L, λ_L, m_U, λ_U, α, γ, that characterize the geometry of the loser sheets can be analytically determined via renormalization.

E. Renormalization computation: Overview and summary

Stepping back for a moment, what we have here is a renormalization problem akin to those so often encountered in physics and the nonlinear sciences, such as the period-doubling cascade made famous by May[16] in a biological mapping, and analyzed by Feigenbaum[17] using renormalization techniques. In particular, we have objects (instant-winner matrices) that exhibit the same overall structure at different size scales [cf. Figs. 2(a) and 2(b)], and a recursion operator (3) relating them. Our task therefore is to determine the invariant geometry of the system, i.e., we seek an invariant geometrical structure W such that if we apply the recursion operator $\mathbf{R}$ to it (thereby growing W) and then act with an appropriately defined linear rescaling operator $\mathbf{S}$ (shrinking it back down), we get W back again: $W = \mathbf{SR}\, W$. In the language of renormalization, we seek a fixed point of the "renormalization-group operator" $\mathbf{S}\,\mathbf{R}$. This can be done, but before proceeding we point out one critical feature of the analysis: Even though the recursion operator $\mathbf{R}$ is exact and the game itself has absolutely no stochastic aspects to it, it will prove necessary to adopt a probabilistic framework in order to solve this fixed-point equation. Namely, our renormalization procedure will demonstrate that the slopes of all boundary lines and densities of all regions in the W_x's (and L_x's) are preserved—not that there exists a point-by-point equivalence. In essence, we will bypass consideration of the random-looking "scatter" of points surrounding the various lines and regions of W_x and L_x (see Fig. 2) by effectively averaging over these fluctuations.

The key to implementing the renormalization analysis is to observe that the P positions in L_x [Fig. 2(c)] are constrained to lie along certain boundary lines of the corresponding W_x plot [Fig. 2(b)], and are conspicuously absent from the various interior regions of W_x (for all x). In effect, the interior regions of each W_x remain "forbidden" to the P positions. This in turn implies that the allowable geometry of the W_x's must be very tightly constrained, in order that the forbidden regions be preserved as the system evolves under the recursion operator $W_x \xrightarrow{R} W_{x+1}$ (for otherwise the W_x geometry would not remain scale invariant). Each forbidden region in W_x imposes a constraint on the permissible structural form that the W_x's can take, and each such constraint can be formulated as an algebraic equation relating the hitherto unknown parameters m_L, λ_L, m_U, λ_U, γ, α that characterize the loser-sheet geometry. Detailed renormalization calculations, described in the following section, show that there are six independent conditions in all associated with the forbidden regions.

Renormalization constraints

Constraint 1: $$\lambda_U + \lambda_L = 1,$$

Constraint 2: $$\frac{\lambda_U}{1+m_U} = 1,$$

Constraint 3: $$\frac{1}{\alpha+1} - \frac{\lambda_L}{m_L+1} = 1,$$

Constraint 4: $$(\gamma-1)\frac{m_L}{\alpha-m_L} + \frac{1}{\alpha+1} = 1,$$

Constraint 5: $$\frac{\alpha\lambda_L}{\alpha-m_L}\left(\frac{m_U-m_L}{m_U\alpha-m_L\alpha+m_L\gamma}\right) + \frac{1}{\alpha+1} = 1,$$

Constraint 6: $$\frac{\lambda_L}{\alpha-m_L} - \frac{\alpha}{\alpha+1}\left(1-\frac{\lambda_U}{\alpha-m_U}\right) = 0.$$

These are the necessary conditions for the instant-winner sheets to be fixed points of the renormalization operator $\mathbf{S}\,\mathbf{R}$. Solving the renormalization constraints yields

$$\alpha = \frac{1}{\sqrt{2}}, \quad \lambda_L = 1 - \frac{1}{\sqrt{2}}, \quad \lambda_U = \frac{1}{\sqrt{2}},$$
$$m_L = -1 - \frac{1}{\sqrt{2}}, \quad m_U = -1 + \frac{1}{\sqrt{2}}, \quad \gamma = \sqrt{2} - 1. \tag{5}$$

Thus, the methodology yields analytical values for the six key geometric parameters that characterize the overall structure of the loser sheets $\{L_x\}$, thereby providing us with a complete (probabilistic) description of the global geometry of the P set of the game.

F. Renormalization computation: Detailed calculations

We now derive the six analytical constraint relations above, which define the game's geometry. The calculation proceeds in several steps.

1. Computation of line densities

Recall that the geometry of the loser sheets L_x [Fig. 2(c)] consists of three diffuse lines: a lower line of slope m_L and density of points (per unit y) λ_L, an upper line of slope m_U and density λ_U (per unit y), and a flat line extending to infinity. The upper and lower lines originate from a point whose height (i.e., z value) is αx. The flat line (with density one) is altogether absent in some loser sheets; we define γ to be the probability that a flat loser line is present in a randomly selected loser sheet. In this section we clarify how this loser-sheet geometry is related to the instant-winner sheet geometry, and explicitly calculate various line densities for the instant-winner sheets.

To begin, we express the recursion relation (3) as

$$W_{x+1} = \mathbf{L}(W_x + \mathbf{DM}W_x) \tag{6}$$

and observe that the new instant-winner sheet W_{x+1} is generated from the old W_x by the following sequence of steps: First, the supermex operator M acts to create the loser sheet L_x (since $\mathbf{M}W_x = L_x$). This loser sheet is then modified by the operator $\mathbf{D}$, which adds a diagonal line of 1's into the matrix (the deadly diagonals). We then add this modified loser sheet ($\mathbf{D}L_x$) to the original W_x sheet, and then left shift the whole sheet, yielding W_{x+1}.

The most critical stage of this growth of W_x into W_{x+1} is when the original instant-winner sheet W_x is altered by the addition of the modified loser sheet $\mathbf{D}L_x$. The details of this process are best understood by decomposing $\mathbf{D}L_x$ into its four basic components (Fig. 3): the lower loser line in L_x, the upper loser line in L_x, the flat loser line in L_x (when it exists), and the deadly diagonal line created by $\mathbf{D}$. Thus, we can think of the recursion operator $\mathbf{R}$ as building up the new instant-winner sheet via the addition of four new lines to the old sheet (followed by a left shift). Now, since we can recursively generate an arbitrary sheet W_x (starting from W_0) through repeated applications of the recursion operator (i.e., $W_0 \xrightarrow{R} W_1 \xrightarrow{R} \cdots \xrightarrow{R} W_x$), we can think of each W_x as being comprised of a series of nested lines coming from repeated addition of the four (modified loser-sheet) lines at each step in the recursion. Turning this around, any instant-winner sheet W_x can be decomposed into its constituent parts, wherein we separate out the contributions that came from lower loser lines, upper loser lines, flat lines, and deadly diagonals. Figure 4 shows a typical W_x (for $x=100$); Fig. 5 shows its decomposition into its four elemental sets (as defined by which of the four components of the modified loser sheets contributed). We label these sets LL (for lower loser lines), U (for upper loser lines), F (for flat loser lines), and DD (for deadly diagonals). Note that DD overlaps with U and with F, but that no other sets overlap. In particular, LL, U, and F cannot overlap with one another for the simple reason that as the three loser lines (in a modified loser sheet $\mathbf{D}L_x$) are being laid down, they cannot intersect the current instant-winner sheet to which they are being added (since P and N positions constitute disjoint sets). DD and LL do not overlap for a simple geometrical reason: Each time a new DD line is laid down, it is located above the existing LL region (since the LL lines have slope $m_L < -1$, whereas the DD lines have slope -1). We remark here that while the left-shift operator $\mathbf{L}$ appearing in Eq. (6) has been ignored in the above discussion, it is clear that its inclusion would not alter any of the preceding conclusions.

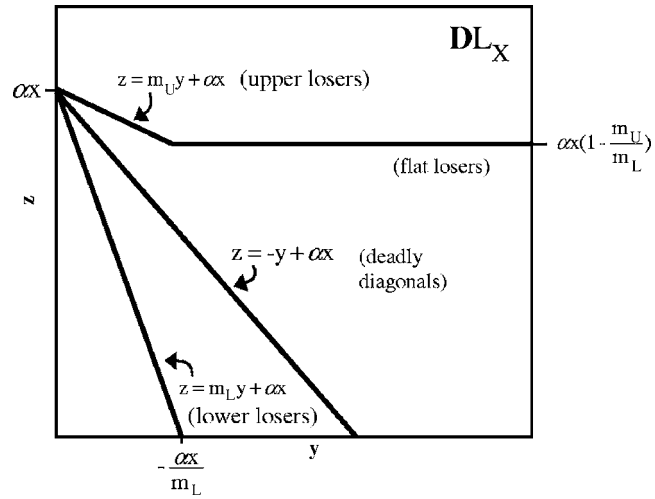

FIG. 3. Composition of $\mathbf{D}L_x$. An instant-winner sheet at level $x+1$ is constructed from the instant-winner sheet at level x by the addition of the modified loser sheet $\mathbf{D}L_x$ [via the relation $W_{x+1} = \mathbf{L}(W_x + \mathbf{D}L_x)$]. The basic structure of a typical $\mathbf{D}L_x$ sheet is shown. It is comprised of four lines: a lower loser line, an upper loser line, a flat loser line, and the deadly diagonal line.

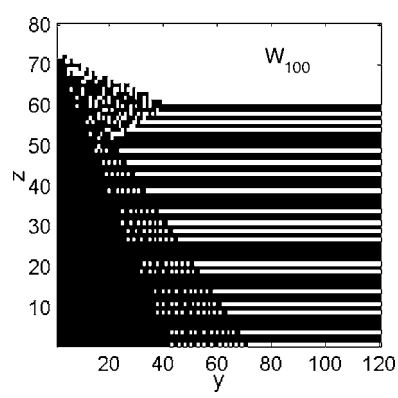

FIG. 4. Instant-winner sheet W_{100}.

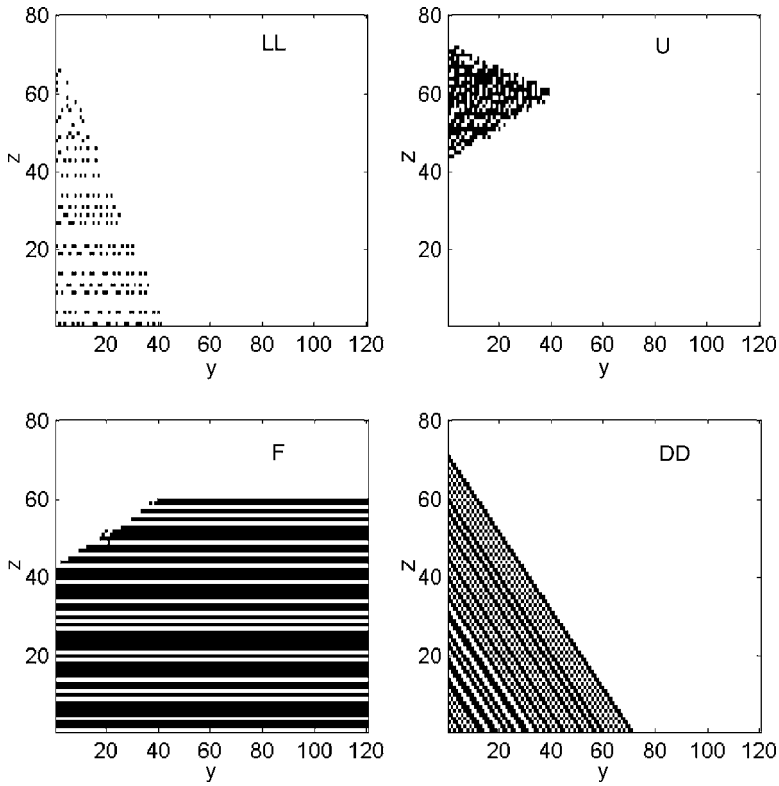

FIG. 5. Decomposition of W_{100}. The figure shows contributions to W_{100} from the lower loser lines (LL), upper loser lines (U), flat loser lines (F), and deadly diagonals (DD). Sets LL, U, and F are mutually disjoint; set DD overlaps with U and with F.

Each of the regions LL, U, F, and DD in Fig. 5 is comprised of a nested series of lines. [The sets of parallel deadly diagonal lines and the flat lines making up DD and F, respectively, are clearly visible in the figure; the parallel line structures for LL and U are somewhat less apparent because the lower and upper loser lines (Fig. 3) that are laid down during the recursion process are not solid (i.e., their density of points is less than one).] We now calculate the density of lines (per unit horizontal) for LL, DD, U, and F. We will start with LL, DD, and U, since they all follow from the same general argument (the flat loser lines comprising F must be handled separately). Consider a given W_x. The lower loser, upper loser, and deadly diagonal lines in $\mathbf{D}L_x$ (that are to be added to W_x to create W_{x+1}) all originate from the same point $(y,z)=(0,\alpha x)$. Thus, their initial "height" in the y–z plane is αx, and so all three lines have the form $z=my+\alpha x$ (where $m=m_L$, m_U, -1 for the lower loser line, upper loser line, and deadly diagonal line, respectively). Now, in the construction of W_{x+1} from W_x [Eq. (6)], these lines are added to W_x and then the resulting sheet is left shifted. These added lines will become the leading edges of the new instant-winner sheet W_{x+1}. The height of these lines (i.e., their z intercept) in the newly formed sheet W_{x+1} will be $\alpha x+m$ (since they come in at height αx when they are first added, but then subsequently drop in height by $|m|$ when they are left shifted one unit by the operator $\mathbf{L}$). Now consider the construction of sheet W_{x+2} from W_{x+1}. This is done by the addition of the three new lines from $\mathbf{D}L_{x+1}$. These new lines start at height $\alpha(x+1)$. When these are added to W_{x+1} and then left shifted, they become the leading edge of W_{x+2} at height $\alpha(x+1)+m$. Meanwhile, the old lines which had been the leading edge of W_{x+1} will drop in height again by $|m|$ owing to the second left shift (to height $\alpha x+2m$), and will now form the next-to-leading edge of W_{x+2}.

Comparing the heights of the lines comprising the leading edge and next-to-leading edge of W_{x+2} yields a height difference of $[\alpha(x+1)+m]-[\alpha x+2m]=\alpha-m$. This represents the vertical spacing between successive parallel lines comprising the LL, U, and DD regions of an instant-winner sheet. Hence, the density of lines (per unit z) in these regions is

$$\frac{1}{\alpha-m} \quad \text{(vertical line density)}, \tag{7}$$

where $m=m_L$, m_U, -1 for lower losers, upper losers, and deadly diagonals, respectively. It follows from elementary geometry that the density of lines per unit y is

$$\frac{-m}{\alpha-m} \quad \text{(horizontal line density)}, \tag{8}$$

where, as before, $m=m_L$, m_U, -1 for lower losers, upper losers, and deadly diagonals, respectively. We emphasize that results (7) and (8), when applied to LL and U, give the average density of the lower and upper loser *lines* in an instant winner sheet—they do not give the density of *points* making up the individual lines [recall that the lower and upper loser lines are not in fact solid, but are comprised of points with densities (per unit y) λ_L, λ_U, respectively].

Lastly, we calculate the density (per unit z) of the flat loser lines comprising F. Recall that the probability that a flat loser line exists in a randomly selected L_x is γ. Hence, the

total number of flat lines that have been generated during the recursive construction of instant-winner sheet W_x is simply γx. From Fig. 3, we see that, at a given x level, the total vertical span of the flat lines is $\alpha x(1-m_U/m_L)$. Hence, the vertical density of these lines is simply

$$\frac{\gamma}{\alpha\left(1-\frac{m_U}{m_L}\right)} \quad \text{(vertical density of flat loser lines)}. \tag{9}$$

2. Analysis of forbidden regions

With the preceding relations about line densities in hand, we now derive the six fundamental algebraic conditions [renormalization constraints (1)–(6) of Sec. III E] that characterize the structure of the instant-winner and loser sheets, and which in turn yield the analytical values of the six key geometric parameters m_L, λ_L, m_U, λ_U, γ, α [Eq. (5)]. Each of the six renormalization constraints (with the exception of the first one) arises from an analysis of a particular forbidden region in W_x (i.e., regions in which losers are excluded). Our labeling of forbidden regions (I-IV) in W_x is shown in Fig. 2(a).

Constraint 1: Existence and uniqueness of losers. Consider the portion of a loser sheet L_x, where both the upper and lower loser lines exist (i.e., $y<-\alpha x/m_L$). In this region there will be exactly one loser in each column of L_x. This critical observation follows directly from the supermex algorithm described in Sec. III C, wherein the z coordinate of the unique losing position in the y-th column of L_x was denoted $z_{\text{small}}(y)$. Accordingly, since λ_U, λ_L represent the density of losing positions (per unit y) along the upper and lower loser lines, respectively, we have

$$\lambda_U+\lambda_L=1.$$

Constraint 2: Region III is forbidden. Notice that the upper triangular region of the W_x's [labeled region III of Fig. 2(a)] is devoid of losers for all x. In particular, when the losers L_x are constructed from the instant-winner sheet W_x via the supermex operator ($L_x=\mathbf{M}W_x$), they are forbidden from appearing in any of the existing holes in region III [i.e., locations (y,z), where $W_x(y,z)=0$]. The mechanism preventing their appearance there is the implied diagonals generated during the supermex operation. Specifically, each time a new loser along the upper loser line is created, it casts down an implied diagonal (see step 6 of the supermex algorithm) and thereby fills in some of remaining holes in region III. These implied diagonals cannot overlap with one another, and as a set must effectively fill up all holes in region III (for otherwise the geometric invariance of the W_x's would not be preserved as they grow). The condition that the implied diagonals fill the gaps is actually stronger than it might appear at first glance: since the locations of the gaps are not well correlated with the locations of the upper losers, in fact, the implied diagonals which are cast down not only just fill the gaps, but in fact densely fill region III. Now, since the slope of the upper loser line is m_U, and the losers are scattered with density λ_U (per unit y) along this line, we can calculate the density (per unit vertical z) of the implied diagonals they cast

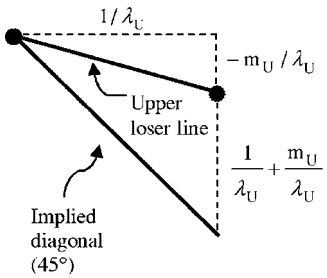

FIG. 6. Implied diagonal density. The figure shows geometric considerations leading to a determination of the density of the implied diagonal lines that are cast down by the upper losing positions during the supermex algorithm.

off via a simple geometric argument (see Fig. 6): Since the average horizontal separation between successive losing positions on the upper loser line is $1/\lambda_U$, and since the slope of this line is m_U, it follows that the average vertical separation of successive losers along this line is simply $-m_U/\lambda_U$. From elementary geometry, the average vertical spacing between the implied diagonal lines that are cast down by these losers is $1/\lambda_U+m_U/\lambda_U$, and so their density per unit vertical is simply the reciprocal of this, namely $\lambda_U/(1+m_U)$. Finally, demanding that these implied diagonals fill region III entirely, we set their density equal to unity, yielding constraint 2,

$$\frac{\lambda_U}{1+m_U}=1.$$

(An entirely analogous argument applied to region F of Fig. 5 shows that the density of losers along the flat loser line equals unity.)

Constraint 3: Region II is forbidden. Region II [Fig. 2(a)] is made up of contributions from DD (along with F in the lower part of the region and U in the upper part), with DD playing the key role. In particular, losers are prevented from appearing inside region II, because the implied diagonals which are cast out by P positions on the lower loser line during the supermex process mesh perfectly with the existing DD, thereby completely filling the region and making it forbidden to losers. The underlying reason behind this perfect meshing of the implied diagonals and DD is that every loser created on the lower loser line by the supermex algorithm must necessarily be filling in an existing hole in the current instant-winner sheet, and so there could not have been a DD line already there. Thus, the implied diagonal emanating from any such loser will not overlap any DD. Moreover, as we sweep across the columns of W_x and lay down lower losers during the supermex operation, the lowest remaining hole in the current column must necessarily get filled by a loser. Hence, the implied diagonals and the DD's will together entirely fill every column of region II. Hence the densities (per unit vertical) of the DD's and the implied diagonals must sum to unity. Now, the density of the DD's is given by Eq. (7) above with $m=-1$, while the density of the implied diagonals coming from the lower losers can be found via the same geometric argument used to find the density of implied diagonals for the upper losers in constraint 2 (see the

construction in Fig. 6), yielding $-\lambda_L/(1+m_L)$. Thus, the constraint for region II that these densities sum to unity is

$$\frac{1}{\alpha+1} - \frac{\lambda_L}{1+m_L} = 1.$$

Constraint 4: Bottom row of region I is forbidden. Focusing on the bottom-most row ($z=0$) of region I of an arbitrary instant-winner sheet, we see that it is always completely filled by instant winners, and hence forbidden to losers. (The same is true for all other rows of region I as well, but those will be handled separately.) We now examine how this condition arises, and derive the associated constraint. To begin, recall (Fig. 5) that only LL and DD contribute to the bottom row of region I, and that they do not overlap. (F would seemingly contribute too were it not for the easily verifiable fact that the $z=0$ row of F is always empty.) Since the bottom row of region I is completely filled, it follows that the densities of the DD's and LL positions (in the bottom row) must sum to unity. The density of the DD's has been calculated previously [Eq. (8) with $m=-1$]. The density of the lower lose lines is given by Eq. (8) with $m=m_L$; however, what we require here is not the density of the LL *lines*, but rather a measure of the rate at which these lines contribute actual points to the bottom row (since the reader will recall that each LL line is only sparsely filled with points, so not all lines will make an actual contribution to the bottom row). The desired density of points along the bottom row coming from the LL lines follows from a crucial observation: Whenever a flat loser line is absent in a loser sheet (which occurs with probability $1-\gamma$), it is because a losing position with height $z=0$ has been generated in the lower loser line (see step 8 of the supermex algorithm). Hence, the average density (per unit horizontal) of the LL positions along the bottom row of a given W_x must equal the average density (per unit horizontal) of the LL lines [Eq. (8)] times the probability that a height-zero loser is generated $(1-\gamma)$. This yields a density of points $(\gamma-1)m_L/(\alpha-m_L)$. Setting the sum of the densities of the DD's and LL positions in the bottom row equal to unity yields the desired constraint,

$$\frac{(\gamma-1)m_L}{\alpha-m_L} + \frac{1}{\alpha+1} = 1.$$

Constraint 5: Lower region I is forbidden. Observe that *all* rows of region I (not just the bottom-most row, as discussed above) are completely filled by instant winners, and hence forbidden to losers. Consider now a row in the lower part of region I, where LL, DD, and F all contribute (but not U). The sum of these three contributions must densely fill the row. For clarity of argument, we select a row in which the horizontal band of F is entirely empty. Hence LL and DD alone must fill this row, and since they do not overlap, it follows that the sum of the densities (per unit y) of DD and LL must equal unity for that row. Now, the density of the DD's is already known [Eq. (8) with $m=-1$], so we need only compute the density of points contributed by LL. This proves to be slightly subtle. At first glance, it might seem that we could simply multiply the horizontal density of the LL lines [Eq. (8) with $m=m_L$] by the average number of points contributed to a given row by each LL line $(-\lambda_L/m_L)$. However, this naïve argument would miss important correlations that exist. In particular, the actual location of the losing points on a given LL line is strongly correlated with the empty/filled rows of F. Indeed, points in LL can only exist in empty bands of F. Since we have been restricting consideration here to a row in region I in which the horizontal band coming from F happens to be empty, we must explicitly deal with these correlations. The proper calculation goes as follows: Consider a segment of a LL line. Let its vertical extent be H, so its horizontal extent is $-H/m_L$ (since it has slope m_L). The expected total number of points distributed along this LL line segment is $(-H/m_L)\lambda_L$. Now, the number of empty horizontal bands of F that will "cross" this line segment is given by the product of the vertical density of empty bands and the vertical height of the segment. (Note: the vertical density of empty bands is just one minus the density of filled bands [Eq. (9)].) Hence the number of points on the line segment per empty band is $-\lambda_L\alpha(m_L-m_U)/(m_L(\alpha m_L-\alpha m_U-\gamma m_L))$. Multiplying this by the density of LL lines [Eq. (8)] yields the desired horizontal density of points contributed by LL along the row. Finally, setting the sum of this density and the density of the DD's to unity, we find

$$\frac{\lambda_L\alpha(m_U-m_L)}{(\alpha-m_L)(\alpha m_U-\alpha m_L+\gamma m_L)} + \frac{1}{\alpha+1} = 1.$$

Constraint 6: Upper region I is forbidden. Consider a row in the upper portion of region I, where U, DD, and LL all contribute (but not F). Note that U and DD overlap with one another, though not with LL. Since the row is completely filled, we have density(LL)+density(U+DD)=1. Now, the density of points contributed to a row by the LL lines is easily found: The density of the LL lines themselves is given by Eq. (8), while the average number of points contributed to a given row by each LL line is just $-\lambda_L/m_L$. So the (horizontal) density of points in a row from LL is simply $\lambda_L/(\alpha-m_L)$. [Note: The intermediate expression $(-\lambda_L/m_L)$ used in the above calculation is easily obtained: Consider a segment of a LL line. Let H denote its vertical extent H and $-H/m_L$ its horizontal extent. So the total number of points along this segment is $(-H/m_L)\lambda_L$. Dividing by H yields the expected number of points per row contributed by the line.] Using a similar argument, the (horizontal) density of points contributed in a row contributed from U is $\lambda_U/(\alpha-m_U)$. Lastly, the density from DD is $1/(\alpha+1)$. Assuming no correlations between the contributions from U and DD, the expected density of their combined contributions is $1-(1-\lambda_U/(\alpha-m_U))\times(1-1/(\alpha+1))$. Adding this to the density from LL, the requirement that the row is completely filled becomes

$$\frac{\lambda_L}{\alpha-m_L} - \frac{\alpha}{\alpha+1}\left(1-\frac{\lambda_U}{\alpha-m_U}\right) = 0.$$

Solving these six algebraic constraint relations yields exact values for the parameters characterizing the geometric structure of the game. These values are given in (5) above.

G. Assumptions

As noted earlier, renormalization techniques in science, despite their great success, typically lack the strict level of mathematical rigor needed to constitute formal proof. In the present context, there are several components of our analysis which make strict proof difficult. One of the main assumptions in the preceding renormalization calculations (particularly apparent in the derivation of constraint 6 above) is the assumed lack of correlations between different lines of losers. Our main (nonrigorous) justification for this is that when x is large, these lines arise from distant sheets. If, as we argue later is true for Chomp, the renormalization operator has "sensitive dependence on initial conditions" (i.e., small changes in an instant-winner sheet W lead to essentially random changes in the sheets after many iterations, i.e., $\mathbf{R}^k W$ for large k) then this assumption will be asymptotically valid. A formal verification of this might be possible, but would require computing the spectrum of the linearization of $\mathbf{R}$ at the fixed point, a difficult though potentially tractable computation. In a related vein, while our renormalization scheme has allowed us to calculate the overall geometric structure of the sheets [culminating in Eq. (5)], a more formal analysis in which the renormalization fixed point is rigorously proven to be stable would be needed to properly address the observed scatter/fluctuations within this geometry. (Unlike standard renormalization schemes for studies of critical phenomena wherein one encounters unstable fixed points, here the fixed points can be stable; in this sense our renormalization scheme is perhaps more closely akin to those used to study self-organized critical systems, where attractive fixed points are also seen.[18])

At present then, it is perhaps more appropriate to regard the six renormalization constraints [Eq. (5)] not as formally proven mathematical propositions, but rather as a set of heuristically argued self-consistency conditions that the geometry of the instant-winner sheets must satisfy if it is to remain invariant under the (exact) recursion operator $\mathbf{R}$ for the game. Once this geometry is established in the (lower-level) sheets, it becomes self-perpetuating under the recursion dynamics, yielding a consistent theory.

IV. IMPLICATIONS FOR CHOMP

The new geometric picture that emerges from the renormalization analysis [Eq. (5); Fig. 2] provides powerful insights into the game, as we now describe.

A. Scalings and patterns

Our findings allow us to easily explain virtually all prior numerical observations (that we are aware of) about the game of Chomp that have been reported in the literature to date, including key numerical conjectures on the game's P-position properties by Brouwer.[10] As one illustration, we note that Brouwer numerically determined winning and losing positions in Chomp for x up to about 80 000, and observed that for every x there seems to be a losing position of the form $[x,0,z]$ where $z=\alpha x \pm 1.75$ with $\alpha \sim 0.7$. From the geometric picture emerging from the renormalization analysis, it becomes trivial to verify that the exact asymptotic value of α is in fact $1/\sqrt{2}$. A host of related, numerically reported observations in the literature can be similarly explained with ease, and we will not elaborate further here. We turn instead to the more novel aspects and deeper implications of the renormalization findings.

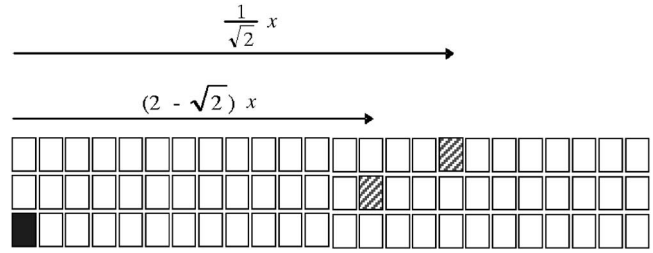

FIG. 7. Winning opening move. As revealed by the renormalization analysis, the two marked locations show the location of the two possible winning opening moves of the game (from starting configuration $[x,0,0]$). For a given x, only one of the two marked locations is the actual winning move: the probability that it is the upper one is $\sqrt{2}-1$; the probability of it being the lower one is $2-\sqrt{2}$. We note that although these results are asymptotic [in the sense that the narrow scatter of points around the loser lines of Fig. 2(c) has been neglected], numerical simulations indicate that the exact winning opening move locations always lie extremely close to (i.e., within 3 units of) these predicted values.

B. Winning opening move

Based on the renormalization results, it is now possible to answer a long-standing, open question about the winning opening move in the game of Chomp.[5] Specifically, we can for the first time show that:

(i) From starting position $[x,0,0]$, the sought-after winning opening move must (asymptotically) be to one of the following two positions (Fig. 7):

$$[x/\sqrt{2},x(2-\sqrt{2})/2,0] \quad \text{or} \quad [x(2-\sqrt{2}),0,x(\sqrt{2}-1)].$$

(ii) The winning opening move is *unique* (i.e., for a given x, only one of the two positions above corresponds to the actual winning move).

(iii) For a randomly chosen x, the probability that the winning opening move is to position $[x/\sqrt{2},x(2-\sqrt{2})/2,0]$ is $\sqrt{2}-1$; the probability that it is to position $[x(2-\sqrt{2}),0,x(\sqrt{2}-1)]$ is $2-\sqrt{2}$.

To show result (i), recall that the legal opening moves from starting position $[x,0,0]$ are to positions of the form $[x-r,r,0]$, $[x-s,0,s]$, and $[x-t,0,0]$, and that the desired (i.e., winning) opening move must be to a P position. It is easy to see that the last of these $[x-t,0,0]$, can never be a P position, by Gale's strategy-stealing argument. Next consider $[x-r,r,0]$; call this an r-type position. From the known geometric structure of the loser sheets, a simple calculation [Eqs. (5); Fig. 2] shows that the only possible P position of this form is for $r(x) \approx \alpha x(\alpha - m_L)$. Likewise, the only possible P position of the form $[x-s,0,s]$ (i.e., an s-type position) is for $s(x) \approx \alpha x/(\alpha+1)$. [Note that we use the notation $\approx$ here since these are asymptotic values; for any finite x there will be small deviations owing to the slight scatter of P positions in Fig. 2(c).] Hence, the knowledge of the P-set geometry

coming from the renormalization analysis has yielded a description of the only two possible winning opening moves in Chomp.

Moreover, we can show that the winning opening move is in fact *unique*—i.e., from a given starting position $[x,0,0]$ (which is always an N position), either the r-type position is a P position or the s-type position is a P position, but never both. Uniqueness of the winning opening move in three-row Chomp was previously only known numerically.[10] Result (ii) follows simply by observing that since $r(x)<s(x)$, the s position is reachable from the r position; however, the child of a P position can never be another P position by (F2), so only one of these two positions can in fact be a P position.

Further still, we can compute the probabilities (for a randomly selected x) that this unique winning opening move will be to the r-type position or to the s-type position, as follows: For each opening position $[x,0,0]$ for $x=1\cdots x_{\max}$ there is an associated r position $[x-r(x),r(x),0]$, which may or may not be a P position. However, we know that the total number of actual r-type P positions having an x value less than or equal to $x_{\max}-r(x_{\max})$ is just $\gamma(x_{\max}-r(x_{\max}))$, since this is identical to the total number of P positions with $z=0$ in the loser sheets $\{L_0,L_1,\ldots,L_{x_{\max}-r(x_{\max})}\}$. So the fraction of r positions which are actually P positions is $\gamma(x_{\max}-r(x_{\max}))/x_{\max}=\sqrt{2}-1$. Thus, the winning opening move is to an r position with probability $\sqrt{2}-1$, and to an s position with probability $2-\sqrt{2}$.

Lastly, we point out that result (i) above specifying the two possible forms of the winning opening move is really an asymptotic result. For any finite x, the actual location of the winning move will lie in the vicinity of the specified asymptotic positions, where vicinity is determined by the width of the scatter of the P positions surrounding the analytical loser lines [Fig. 2(c)]. Numerical simulations indicate that this width is quite narrow and appears to have a global maximum bound (i.e., for all x) of less than about 2.5. Though this bound has yet to be proven analytically, it is in agreement with the general heuristic argument that the supermex operator **M**, by its very construction, tends to place new P positions as close as possible to existing boundaries in the instant-winner sheets (i.e., the P positions "hug" the analytical loser lines)—hence the extremely accurate agreement between numerics and the asymptotic form.

C. Computation

Using the renormalization results, one can readily compute the possible winning moves from any N position in the game (not just from the game's starting configuration $[x,0,0]$). Here, provided the N position is not near a boundary in W_x (or, equivalently, near a loser line in L_x), one need only determine which loser lines [see Eqs. (5)] are accessible from that position under the game rules (M1)–(M6). Explicit exposition of these possible moves is straightforward (but cumbersome), and we do not elaborate here. We do note, however, that unlike for the starting configuration $[x,0,0]$, the winning moves from an arbitrary N position need not be unique.

In conjunction with this, knowledge of the overall geometrical structure of the loser sheets suggests a natural pathway to more efficient algorithms for Chomp by simply designing the search algorithm to aim directly for the analytically determined P-position lines in L_x [given by Eqs. (5); Fig. 2(c)] and ignore positions in the forbidden regions. More generally, these results hint at a new probabilistic approach to solving other hard combinatorial problems, a topic we are currently exploring (see Sec. V D for a related discussion).

V. GENERAL IMPLICATIONS

Thus far, we have seen how the probabilistic renormalization framework has allowed for a fairly systematic analysis of some of the key features of the combinatorial game Chomp. In this section, we look at this methodology from a somewhat broader perspective, and show how it leads to some entirely new classes of questions and paradigms about Chomp and combinatorial games more generally.

A. Universality

Borrowing a fundamental idea from modern dynamical systems theory, we begin with an examination of the robustness of a game like Chomp to perturbations. The idea here is to create variants of the original game by making certain rule changes to it, and then asking how the properties of these variant games relate to those of the original game. The type of perturbation we consider here is obtained by adding one (or more) new points to a game's instant-winner sheet. (The new points being added are always taken to be losing positions of the original game.) Through this perturbation, we are in effect considering what happens when various losing positions of the original game are simply declared (by fiat) to be automatic winners. Specifically, we wish to examine how the underlying geometry (i.e., instant-winner sheets and P sets) of these variant games compares to that of the original game. (Note that the underlying recursion operator of a original game is unaffected by the perturbation, and hence applicable to the game variants.)

Simulations show that, for a sizeable class of variants of Chomp, the original geometric structure of Fig. 2 reemerges—i.e., the renormalization analysis reveals that, although the instant-winner sheets for the various variants will differ from one another on a local, point-by-point basis, they nevertheless all share global geometric features [cf. Eq. (5)] identical to that of unperturbed Chomp. Hence, the geometry of Chomp appears to be *structurally stable* (in a probabilistic sense). In the language of renormalization, we would say that such game variants fall into the same universality class as the original game. Thus, the geometrically based renormalization framework gives us a general way of grouping combinatorial games into universality classes on the basis of their structural stability and shared geometry.

We mention that while Chomp and its variants lie in the same universality class, this is *not* a general property of all games. For instance, we have applied the renormalization methodology developed in this paper to the well-known game of Nim (a much studied, easily solvable combinatorial

game). We find that the underlying geometric structure of (three-heap) Nim is unstable to perturbations. Interestingly, in that case one finds that nearly all the variants of Nim appear to belong to a single universality class (i.e., they share the same underlying geometry, albeit different from that of unperturbed game). Hence, in a certain sense regular (unperturbed) Nim seems to be highly nongeneric: its geometric structure differs from the common structure shared by all its variants. This leads us to conjecture that solvable (i.e., computationally simple) games (like Nim) will have unstable geometries, while more complex, generic games such as Chomp, variants of Nim, and others, will be structurally stable to perturbations. These results will be described in more detail elsewhere.[19]

B. Sensitivity to initial conditions

One of the hallmarks of chaos in dynamical systems theory is the notion of sensitivity to initial conditions. The renormalization scheme provides an avenue for addressing these issues in the context of combinatorial games. We start, as before, by making a small perturbation to one of the game's instant-winner sheets W_x through the addition of one or more new points to the sheet, and then construct the higher-level sheets through repeated iterations of the game's recursion operator $\mathbf{R}$ (i.e., $W_{x+n}=\mathbf{R}^n W_x$). (Here, $\mathbf{R}$ can refer to the recursion operator of a general game, not necessarily Chomp.) Now, provided the game in question is structurally robust, the perturbed and unperturbed games will lie in the same universality class and hence their instant-winner and loser sheets will exhibit the same overall geometric features (in the probabilistic sense). However, sheets of the perturbed and unperturbed game will nonetheless differ on a point-by-point basis. Thus, we can examine the rate at which the initial variation (due to the perturbation to W_x) grows and spreads to higher-level sheets; i.e., in dynamical systems language, what is the sensitivity of the "mapping" $W_{x+1}=\mathbf{R}W_x$ to changes in initial conditions?

Simulations reveal that small initial perturbations are capable of significantly altering the P-position locations quite dramatically. For example, for the game of Chomp, adding just a single new point to W_{50} can, after only seven iterations, alter the locations of over half of all losing positions in all subsequent sheets $\{L_x | x \geq 57\}$, as revealed by the blue data of Fig. 8. This analysis helps to clarify Zeilberger's observation that Chomp seems to exhibit "'chaotic' behavior, but in a vague, yet-to-be-made-precise, sense."[8] The high degree of sensitivity in Chomp is reminiscent of what occurs in chaotic attractors (despite the linear appearance of the growth seen in the plot).

Other measures for quantifying a game's sensitivity would also be desirable here. Though it might seem natural to define *Lyapunov-like exponents* for games such as Chomp (in analogy with what is typically done for mappings in dynamical system theory)—there remain unresolved technical obstacles stemming from the fact that, unlike in ordinary mappings, the phase-space structure arising from the game's recursion operator effectively grows with x, thus leaving this an open problem to date.

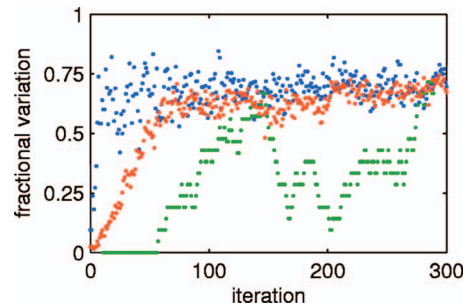

FIG. 8. (Color) Dependence on initial conditions. The figure illustrates how perturbing an instant-winner matrix by a single point subsequently spreads and "infects" the loser sheets at higher x values (i.e., altering the precise locations of the losing positions compared to the unperturbed case). The blue data points show the fraction of P positions along the upper and lower loser lines that are affected when one adds a single point to W_{50} and then iterates. The red data points show the corresponding effect when the initial perturbation is to W_{300}. The green data shows a rolling average of the spread of the infection to P positions lying along the flat loser line (for an initial perturbation to W_{300}). Note that the effects can be pronounced in spite of the linear appearance of the initial growth for small iteration numbers (see blue, red data). For example, the blue data shows that changing just a single point in the instant-winner sheet W_{50} will, after only seven iterations of the recursion operator, shift the location of over half of all losing positions.

Nonetheless, renormalization has provided us with a general means of discussing sensitivity to initial conditions in combinatorial games: Given a structurally stable game, we can think of its instant-winner-sheet or loser-sheet geometry as being a type of complex attractor, and can study the sensitivity of this attractor to a change in initial conditions via the action of the game's recursion operator. In this manner, renormalization provides a natural framework extending the notion of "sensitivity to initial conditions" in dynamical systems theory to combinatorial games.

C. Accretion and crystal growth

Here, we briefly remark upon an interesting issue which is suggested by the renormalization analysis but which has yet to be fully explored. In particular, we note that the growth (with increasing x) of the geometric structures W_x (Fig. 2) is suggestive of certain crystal-growth and aggregation processes in physics[20] (e.g., diffusion-limited aggregation models, solidification problems, etc.). This semblance arises because the recursion operators governing the game evolution ($W_x \xrightarrow{R} W_{x+1}$) typically act by attaching new points to the boundaries of the existing (instant-winner-sheet) structures. This type of attachment-to-boundaries process is a common feature of many physical growth models. Viewed in this way, then, the renormalization procedure offers a means of transforming the study of a combinatorial game into that of a shape-preserving growth process—and with it, the hope that some of the tools which physicists have developed for analyzing such growth models might potentially be brought to bear on combinatorial games.

D. Compressibility and computation

In his paper[8] Zielberger discusses the computational compressibility of Chomp, and how according to Chaitin's theory[13] most results are either trivial or incompressible;

however, our analysis suggests that Chomp is neither. There is no simple formula for P positions, yet the information in the P positions can be significantly (albeit probabilistically) compressed, using the knowledge of the renormalization analysis. Indeed, in a certain sense what the renormalization methodology does is to capture the informationally compressible component of the game (by revealing the overall geometry of the game's P set), and handles the difficult, noncompressible component (representing the precise, point-by-point locations of the P positions along the loser lines) through a probabilistic description. In other words, the coexistence of order (i.e., analytically well-defined loser lines) and disorder (i.e., the scatter of points around these lines) signifies that combinatorial games such as Chomp may be unsolvable yet still informationally compressible, in the language of Chaitin.[13]

Indeed, in the spirit of Frankael,[1] one might argue that this situation, wherein a game lies in the nebulous hinterland between solvability and incompressibility, is in fact the typical one for most "interesting games"—i.e., games that are solvable are uninteresting and those that are incompressible are unplayable.

E. Application to other games

The application of the renormalization methodology discussed in this paper to other games is a significant issue. Although renormalization provided new insights and answers to some long-standing questions about the game of Chomp, were this approach limited solely to Chomp then its value as a tool for analysis would be greatly diminished. However, preliminary work has shown that this is not the case. Indeed, we have successfully applied these techniques to several other low-dimensional combinatorial games, including: Nim and its variants (as already noted), three-dimensional Wythoff's game, a restricted version of Frogs and Toads, and the computation of Sprague-Grundy values for two-dimensional Wythoff's game. (Discussion of these examples will appear elsewhere.[19]) The extension to higher-dimensional games, such as Chomp with an arbitrary number of rows, is possible in principle but problematic in practice without further advances in renormalization techniques; perhaps even automated renormalization techniques (such as those used by the authors in a related paper[21]). For example, in four-row Chomp, it is straightforward to derive an exact recursion operator **R** relating (the now three-dimensional) instant-winner sheets, in complete analogy with what we did here for three-row Chomp. However, the additional dimension complicates the visualization process, thus rendering the analysis more difficult. In addition, it would be very interesting to extend these techniques to partizan games,[3] like chess and checkers (we have made preliminary progress in generalizing this methodology to very simple partizan games), or even to consider games with intrinsic randomness, such as the recent analysis of hex with coin flips,[22] or even backgammon.

VI. CONCLUSIONS

The physics-inspired renormalization technique developed here has provided some insights into the complexities of the combinatorial game Chomp. More importantly, it shows that there are deep connections between combinatorial games and dynamical systems. We have demonstrated some of the applications of dynamical systems theory to combinatorial games and presented many areas for further research; however, we also believe that one might be able to proceed in the other direction—applying the ideas from combinatorial games to increase our understanding of chaotic systems. This remains a key challenge.

[1] A. S. Fraenkel, "Complexity, appeal and challenges of combinatorial games," Theor. Comput. Sci. **313** 393 (2004).
[2] J. H. Conway, *On Numbers and Games* (AK Peters, Natick, MA, 2000).
[3] E. R. Berlekamp, J. H. Conway, and R. K. Guy, *Winning Ways for Your Mathematical Plays* (Academic, London, 1982).
[4] F. Schuh, "Spel van delers," Nieuw Tijdschrift voor Wiskunde **39**, 299 (1952).
[5] D. Gale, "A curious Nim-type game," Am. Math. Monthly **81**, 876 (1974).
[6] D. Zeilberger, "Three-rowed Chomp," Adv. Appl. Math. **26**, 168 (2001).
[7] X. Sun, "Improvements on Chomp," Integers **2**, G01 (2002).
[8] D. Zeilberger, "Chomp, recurrences, and chaos," J. Differ. Equations **10**, 1281 (2004).
[9] S. Byrnes, "Poset games periodicity." Integers **3**, G03 (2003).
[10] A. E. Brouwer, "The game of Chomp," http://www.win.tue.nl/~aeb/games/chomp.html (2004).
[11] R. Sprague, "Uber mathematische Kampfspiele," Tohoku Math. J., **41**, 438 (1936).
[12] P. M. Grundy, "Mathematics and games," Eureka **2**, 6 (1939).
[13] G. Chaitin, *Information, Randomness and Incompleteness*, 2nd ed. (World Scientific, Singapore, 1990).
[14] L. P. Kadanoff, "Scaling laws for Ising models near T_c," Physics (Long Island City, N.Y.) **2**, 263 (1966).
[15] K. G. Wilson, "The renormalization group: critical phenomena and the Kondo problem," Rev. Mod. Phys. **47**, 773 (1975).
[16] R. M. May, "Biological populations obeying difference equations: Stable points, stable cycles, and chaos," J. Theor. Biol. **51**, 511 (1975).
[17] M. Feigenbaum, "Universal behavior in nonlinear systems," Los Alamos Sci. **1**, 4 (1980).
[18] L. Pietronero, A. Vespignani, and S. Zapperi, "Renormalization scheme for self-organized criticality in sandpile models," Phys. Rev. Lett. **72**, 1690 (1994).
[19] E. J. Friedman and A. S. Landsberg, *Games of No Chance 3*, MSRI Series, edited by R. J. Nowakowski (Cambridge University Press, Cambridge, in press).
[20] Y. Bar-Yam, *Dynamics of Complex Systems* (Addison-Wesley, Reading, MA, 1997).
[21] E. J. Friedman and A. S. Landsberg, "Large-scale synchrony in weakly interacting automata," Phys. Rev. E **63**, 051303 (2001).
[22] Y. Peres, O. Schramm, S. Sheffield, and D. Wilson, "Random-turn hex and other selection games," Am. Math. Monthly (in press).

Chaos is copyrighted by the American Institute of Physics (AIP). Redistribution of journal material is subject to the AIP online journal license and/or AIP copyright. For more information, see http://ojps.aip.org/chaos/chocr.jsp