Simple One-Dimensional Interaction Systems with Superexponential Relaxation Times

Andrei Toom

Received April 25, 1994; final February 9, 1995

Finite one-dimensional random processes with local interaction are presented which keep some information of a topological nature about their initial conditions during time, the logarithm of whose expectation grows asymptotically at least as $M^3$, where $M$ is the “size” of the set $R_M$ of states of one component. Actually $R_M$ is a circle of length $M$. At every moment of the discrete time every component turns into some kind of average of its neighbors, after which it makes a random step along this circle. All these steps are mutually independent and identically distributed. In the present version the absolute values of the steps never exceed a constant. The processes are uniform in space, time, and the set of states. This estimation contributes to our awareness of what kind of stable behavior one can expect from one-dimensional random processes with local interaction.

KEY WORDS: Random processes; one-dimensional local interaction; relaxation time; smoothing; Cramér–Edgeworth expansion; harnesses.

1. INTRODUCTION

It is well known that many one-dimensional systems lack those qualitative properties that systems whose dimension is greater than one may have and to which students of statistical physics pay most attention. For example, Lieb and Mattis wrote in the introduction to their still valuable collection, (11) “In one dimension bosons do not condense, electrons do not superconduct, ferromagnets do not magnetize, and liquids do not freeze” (p. vi). Another example: §152 of Landau and Lifshitz’s famous monograph (9) was called, “The impossibility of the existence of phases in one-dimensional systems” and an argument of a physical nature (which

---

1 Partially supported by NSF grant #DMS-932 1216.
2 Incarnate Word College, San Antonio, Texas 78209. E-mail: toom@the-college.iwctx.edu.
Lieb and Mattis justly call "heuristic") was presented in support of this impossibility.

However, for a long time most models which moved physicists to single out the one-dimensional case described equilibrium states. Non-equilibrium phenomena may be described using uniform random processes with local interaction, a class of which we describe here. One important characteristic of random processes is how long they can remember something about their initial conditions. Let us call this time their "relaxation time" (below we define if for our processes). Some infinite processes, even in the presence of random noise, can keep some knowledge about their initial condition forever; these are often called nonergodic, as opposed to ergodic ones which forget everything about their initial condition as \( t \to \infty \). Nondegenerate finite processes cannot remember anything forever, but their relaxation times varies enormously depending on particulars of the interaction.

Examples of ergodic processes or processes with small relaxation times are easy to present: it is sufficient to make the interaction "strong enough" (see, for example, Chapters 3 and 4 of ref. 3, where other references can be found). On the other hand, for all \( d > 1 \), nondegenerate \( d \)-dimensional processes have been proposed which are nonergodic in the infinite case and have large relaxation times in the finite case (see, for example, ref. 13, where other references can be found). If is no wonder, however, that nonergodicity in the one-dimensional case has always presented special difficulties.

Much work has been done on random cellular automata, whose components have finite sets of states. The positive rates conjecture, which was proposed by several authors, claims that all nondegenerate one-dimensional random cellular automata are ergodic, that is, have a unique limit behavior (see, for example, Chapter 4, Section 3 of ref. 10, p. 115 of ref. 3, and ref. 6). However, the systems which mathematicians consider are much more general than those which arise from physical considerations, and may well contradict physical intuition. Now the positive rates conjecture seems to have been refuted: after some preliminary work,\(^{(1,8)}\) Péter Gács proposed a nonergodic, nondegenerate one-dimensional system.\(^{(4)}\) However, Gács's construction is very elaborate, which makes it difficult to apply in physics.

Several one-dimensional cellular automata which display some kind of stability were proposed in ref. 5. Now all of us seem to be (intuitively) sure that all the systems proposed in ref. 5 are ergodic (unless the noise is degenerate). However, their time of relaxation seems to grow unusually fast when \( e \), the probability of errors, tends to 0. For example, de Sa and Maes\(^{(2)}\) simulated one of these models (known as the "soldiers model") and
claimed that the relaxation time grows as an exponent of $1/\varepsilon$. (Note, however, that it is difficult to make a reliable conclusion about the way this time grows from computer experiments alone, because of the very nature of the question.) The relaxation time may be expected to grow fast also for the "two-line voting" system described in ref. 14.

Considering all that has been said, it seems worthwhile to present (as we do) simple one-dimensional systems which display properties which may seem to need more than one dimension. We prove that our systems keep some information about their initial conditions for a very long time, which grows at least as an exponent of $M^3$, when $M$, the "size" of the set of states of a single component, tends to infinity. A physicist might expect $M^2$; the unusual exponent of $M^3$ in our estimations is briefly commented upon in Note 4 below and will be discussed in more detail elsewhere.$^{(15)}$

For technical reasons our components' sets of states are continuous, but computer experiments suggest that the same is true for systems with discrete sets of states of components (see Note 3).

Constructions presented here are not models of any particular physical phenomenon. However, they seem to be closer to physics than to computer science, since we avoid any sophisticated mechanism which the human mind might intentionally create to preserve information about the initial condition. The functioning of our systems involves only smoothing (which is quite imaginable in natural systems) and a symmetric random noise which adds a random increment to every component at every step of the discrete time. This is what we mean by calling our systems "simple" and in this sense the approach of this paper is different from refs. 1, 8, 5, and 4 because we ask what kind of stability may be expected of systems which lack any special mechanism of preservation except "smoothing."

This simplicity has a price, however. It is typical of the nonergodic or stable systems presented until now that their components are mostly in the "correct" states in spite of the random noise. In our systems, however, there is no such thing as a "correct" or "incorrect" state of a single component. What is remembered instead is a topological property of the whole system's state. This makes it more difficult to generalize our results to systems with an infinite set of components. Another property of our systems which is natural from the physical point of view but useless in computer science is uniformity in the set of states in addition to uniformity in space and time.

Our main objects are systems whose elements have $\mathbb{R}_M$, a circle of length $M$, as the set of states. As a tool we consider systems with a linear interaction, the states of whose elements are real numbers. There is some analogy between the latter systems and those described in Chapter IX of ref. 10, since both are linear. Indeed, if we exclude noise, our linear systems become (rather trivial) discrete analog of the "smoothing processes"
described in Chapter IX of ref. 10. However, our main questions are quan-
titative rather than qualitative. On one hand the existence of a process,
which is a major concern throughout ref. 10, is trivial in our case due to the
discreteness of time. On the other hand, our main estimations (3) and (5)
seems to have no analogs for systems examined in Chapter IX of ref. 10.

2. DEFINITIONS AND THEOREMS

Components are indexed by 0, ..., L - 1, elements of the additive group
of residues modulo L \geq 2. A state of the process is an L-tuple
a = (a_0, ..., a_{L-1}), where all a_s belong to the set \( R_M \), which is defined as
follows. Choose a positive number M and define \( R_M \) as the additive group
of classes of equivalence if we declare two real numbers equivalent whenever
diff erence equals M, multiplied by an integer number. Our
processes may be thought os as linear operators \( P \) which act on the set of
probability measures on the set \( R_M^L \) of states, more exactly, on the
\( \sigma \)-algebra generated by its cylinder subsets.

We use mutually independent “hidden” real random variables \( \eta'_s \),
\( s = 0, ..., L - 1 \) and \( t = 1, 2, 3, ... \). Every \( \eta'_s \) is distributed as a real random
variable \( v \), called noise. We assume that \( v \) is not a constant and that there
is a constant \( v_{\text{max}} \) such that

\[
\text{Prob}(|v| > v_{\text{max}}) = 0
\]  

Choose a natural number \( N \) and different integer numbers \( v_1, ..., v_N \in \mathbb{Z} \)
such that the differences \( v_i - v_1 \) generate \( \mathbb{Z} \). Components \( s + v_1, ..., s + v_N \) are
those that influence the component \( s \) at every time step. Also choose
positive numbers \( w_1, ..., w_N \) whose sum equals one (intensities of this
influence). A system \( (v_1, w_1, v) \) is given by \( v_1, ..., v_N, w_1, ..., w_N, \) and the
distribution of \( v \). All the values which depend only on these parameters
will be called constants. To every system there correspond processes,
parametrized by \( L \) and \( M \).

Let us define the transition function \( F: R_M^N \mapsto R_M \). (Note that the
following definition is consistent.) Given a real number \( x, \mu(x) \in R_M \)
denotes its class. Given a class \( y \in R_M \), its intersection with the segment
\((-M/2, M/2]\) consists of one number, denoted \( \mu_0^{-1}(y) \). For any
\( x_1, ..., x_N \in R_M \):

- If there is \( q \in R_M \) such that

\[
\forall j = 1, ..., N: -M/4 < \mu_0^{-1}(x_j - q) < M/4
\]
the value of the transition function equals

\[ F(x_1, \ldots, x_N) = \mu \left( \sum_{i=1}^{N} w_i \cdot \mu_0^{-1}(x_i - q) \right) + q \]

- Otherwise the value of \( F(x_1, \ldots, x_N) \) is undefined.

A process \((a_i^t)\) is the distribution of random variables \(a_i^t \in \mathbb{R}_M\) induced by the distribution of the hidden variables \(\eta_i^t\) with the map defined in the following inductive way:

\[
a_i^t = \begin{cases} 
F(a_{i+1}^{t-1}, \ldots, a_{i+v}^{t-1}) + \mu(\eta_i^t) & \text{if } F(a_{i+1}^{t-1}, \ldots, a_{i+v}^{t-1}) \text{ is defined} \\
\text{arbitrary} & \text{otherwise}
\end{cases}
\]

where \(a_0^0\) are components of the initial condition.

We also define an integer-valued function \(\text{rot}(\cdot)\) on \(\mathbb{R}_M^L\) as follows, where \(a = (a_0, \ldots, a_{L-1})\):

\[
\text{rot}(a) = \begin{cases} 
M^{-1} \cdot \sum_{s=0}^{L-1} \rho(a_s, a_{s+1}) & \text{if all the addends in this sum are defined} \\
\text{undefined} & \text{otherwise}
\end{cases}
\]

Here the real function \(\rho(\cdot)\) is defined by \(\mathbb{R}_M^2\) as follows:

\[
\rho(x, y) = \begin{cases} 
\mu_0^{-1}(y - x) & \text{if } \mu_0^{-1}(y - x) \neq M/2 \\
\text{undefined} & \text{otherwise}
\end{cases}
\]

Informally speaking, \(\text{rot}(a)\) shows how many times we go around the continuous circle \(\mathbb{R}_M\) following the components of our states \(a = (a_0, \ldots, a_{L-1})\) as stepping stones as their index goes around the "discrete circle" \(\{0, \ldots, L-1\}\).

Consider a process \((a_i^t)\) with an initial condition \(a^0\) for which \(\text{rot}(a^0)\) is defined. Let \(a^t\) denote the state of the process at time \(t\). The first time \(t = t^*\) when \(\text{rot}(a^t)\) is different from \(\text{rot}(a^0)\) or undefined is called the relaxation time of this process. The following is our main theorem.

**Theorem 1.** For any system \((v_i, w_i, v)\) with \(N > 1\) there are positive constants \(M_0, C, K\) such that for any initial condition \(a^0\) of the form

\[
a_s^0 = \mu(R \cdot s/L)
\]

where \(R\) is an integer number, the expectation \(\mathbb{E}(t^*)\) of the relaxation time in all corresponding processes with \(M \geq M_0\) is bounded from below by

\[
L^{-1} \cdot \exp\left[ C(M - K |R|)^3 \right] \leq \mathbb{E}(t^*)
\]
Note that for states \( a^0 \) of the form (2) with \( |R| < ML \) the function \( \text{rot}(a^0) \) is defined and equals \( R \). The following theorem gives the opposite estimate for a special case.

**Theorem 2.** Consider the system with \( N = 2, \ v_1 = 0, \ v_2 = 1, \ w_1 = w_2 = 1/2, \) and

\[
v = \begin{cases} 
1 & \text{with probability } 1/2 \\
-1 & \text{with probability } 1/2 
\end{cases}
\]  

(4)

There are positive constants \( K \) and \( C \) such that for any initial condition \( a^0 \) the expectation of the relaxation time in all corresponding processes with \( L > KM \) is bounded from above by

\[
E(t^*) \leq \exp(CM^3)
\]  

(5)

### 3. PROOF OF THEOREM 1

The processes described above will be called *finite* processes now. They are finite in two respects: (a) the number of components is finite and equals \( L \); (b) the set of states is a circle. Our main tool are *infinite* processes whose set of components is \( \mathbb{Z} \) and whose set of states is \( \mathbb{R} \). However, the number \( L \) still serves as a parameter in their definition. Given a system \((v_i, w_i, v)\) and a number \( L \), an infinite process \((b'_s)\) is a distribution of real random variables \( b'_s \), where \( s \in \mathbb{Z} \), which is induced by the distribution of the same hidden variables \( \eta'_s \) which we used in finite processes, where \( s \in \{0, \ldots, L-1\} \), with the map defined in the following inductive way:

\[
b'_s = \sum_{i=1}^{N} w_i \cdot b'_{s+v_i} \pm \eta'_r(s)
\]

for all \( s \in \mathbb{Z}, \ t = 1, 2, 3, \ldots \), where \( r(s) \) is the residue when \( s \) is divided by \( L \) and \( b'_s \) are components of the initial condition. Given an infinite process \((b'_s)\), we denote \( \Delta b'_s = b'_{s+1} - b'_s \).

The following proposition shows which properties of infinite processes underlie Theorem 1.

**Proposition 1.** For any system \((v_i, w_i, v)\) with \( N > 1 \) there is a positive constant \( C \) such that for all \( L, \) all \( D > 0, \) and all corresponding infinite processes \((b'_s)\) with the initial condition \( b'_0 \equiv 0 \) the expectation \( E(t^0) \) of the first time \( t^0 \) when \( \sup_s \Delta b'_s > D \) is bounded from below by \( L^{-1} \cdot \exp(CD^3) \).

**Proof.** Note that the noise in the infinite processes, as we define them, is space-periodic. If the initial condition is also periodic, that is,
b^0_{s+L} = b^0_s (which is certainly true if b^0_s \equiv 0), then the process is also periodic, that is, b^0_{s+L} = b^0_s for all s, t. Note also that in the infinite process every variable b^0_s and every difference Δb^0_s = b^0_{s+1} - b^0_s is a linear combination of some hidden variables. Let us write these formulas for b^0_s and Δb^0_s:

\[
b^0_s = \sum_{n=0}^{t-1} \sum_{s} p^n_s \cdot η^{t-n}_s
\]

\[
Δb^0_s = \sum_{n=0}^{t-1} \sum_{s} Δp^n_s \cdot η^{t-n}_s, \quad \text{where} \quad Δp^n_s = p^n_{s+1} - p^n_s
\]

Let us prove that

\[
\max_s |Δp^n_s| = O(n^{-1}) \quad \text{and} \quad \sum_s (Δp^n_s)^2 = O(n^{-3/2})
\]

Proof of the First Statement in (7). Given v_1, ..., v_N and nonnegative w_1, ..., w_N whose sum equals 1, we can define a real random variable V:

\[
V = \begin{cases} 
v_1 & \text{with probability } w_1 \\
& \cdots \\
v_N & \text{with probability } w_N 
\end{cases}
\]

We can prove by induction that for all s, n the value of p^n_s equals the probability that V_n ≡ s modulo L, where V_n is a sum of n independent random variables, every one of which is distributed as V. (This observation is analogous to the duality between smoothing processes and potlatch processes, which is described, e.g., in Chapter IX of ref. 10.) Thus

\[
p^n_s = \sum_{m=-\infty}^{\infty} q^n_{s+mL}
\]

where \( q^n_s = \text{Prob}(V_n = s) \). The Cramér–Edgeworth expansion for convolutions of identical lattice distributions (for example, refer to Theorem 13 in ref. 12, Chapter VII, p. 205) gives us for any s, n, k \( \geq 1 \)

\[
q^n_s = Q_k(x(s, m)) + o(n^{-k/2})
\]

where

\[
x(s, m) = \frac{s + m \cdot L - n \cdot μ}{σ \sqrt{n}}, \quad Q_k(x) = \frac{\exp\left(-x^2/2\right)}{σ (2πn)^{1/2}} \left(1 + \sum_{j=1}^{k} P_j(x) n^{-j/2}\right)
\]
\( \mu \) and \( \sigma \) are the expectation and standard deviation of \( V \), and every \( P_j(x) \) is a polynomial whose degree is \( j \) and whose coefficients are determined by the first \( j \) moments of the distribution of \( V \).

Since our interaction has a finite range, the sum in (9) actually is always finite and contains \( O(n) \) terms. [In fact it is sufficient to add its terms only from \( m = -rn \) to \( m = rn \), where \( r = \max(|v_1|, ..., |v_N|) \).]

Therefore

\[
\Delta P_s = \sum_{m = -O(n)}^{O(n)} [Q_k(x(s + 1, m)) - Q_k(x(s, m))] + \sum_{m = -O(n)}^{O(n)} o(n^{-k/2}) \tag{12}
\]

Let \( k = 4 \). Then the summing of \( o(n^{-k/2}) \) results in \( O(n^{-1}) \). It remains to prove that the first sum in the right side of (12) is also \( O(n^{-1}) \). For any real function \( f \) denote

\[
\Sigma(f) = \sup_{d, p \neq 0} \left| \sum_{m = -\infty}^{\infty} f(d + m \cdot p) \right|
\]

Note that for any \( f \) and any positive constant \( C \)

\[
\Sigma(C \cdot f) = C \cdot \Sigma(f) \quad \text{and} \quad \Sigma(f(C \cdot x)) = \Sigma(f(x))
\]

**Lemma.** If a real function \( f \) is differentiable, \( |f'(x)| \leq 1 \), and \( |f(x)| \leq 1/(1 + x^2) \) for all \( x \), and \( \int_{-\infty}^{\infty} f(x) \, dx = 0 \), then \( \Sigma(f) \leq 8 \).

**Proof.** Without loss of generality we assume that \( p > 0 \). If \( p \geq 1 \),

\[
\left| \sum_{m = -\infty}^{\infty} f(d + mp) \right| \leq 2 \sum_{n = 0}^{\infty} \frac{1}{1 + n^2} < 8
\]

Now let \( 0 < p \leq 1 \). Then

\[
\left| \sum_{m = -\infty}^{\infty} f(d + mp) \right| \leq \sum_{|d + mp| \geq 1/p} |f(d + mp)| + \sum_{|d + mp| < 1/p} f(d + mp) \tag{13}
\]

The first sum in the right side of (13) does not exceed

\[
2 \sum_{n = 0}^{\infty} \frac{1}{1 + (np + 1/p)^2} \leq 2 \int_{1/p}^{\infty} \frac{dx}{1 + x^2} \leq 2 \int_{1/p}^{\infty} \frac{dx}{x^2} = 2
\]

Let us estimate the second sum in the right side of (13). Let \( m_1 \) and \( m_2 \) denote the smallest and the largest integer values of \( m \) for which
\[ |d + mp| < 1/p. \] Then the second term (without the absolute value sign) can be interpreted as \( p \) times the rectangle rule approximation of the integral

\[
\int_{d + m_1 p}^{d + (m_2 + 1) p} f(x) \, dx
\]

The difference between this finite integral and the infinite integral (which equals zero) does not exceed

\[
\int_{-\infty}^{\infty} \frac{1}{1 + x^2} \, dx = \pi
\]

Since \( |f'| \leq 1 \), the error of the rectangle rule approximation, multiplied by \( p \), does not exceed \( \frac{1}{2}(m_2 - m_1 + 1) p^2 \leq 2. \) Thus \( \Sigma(f) \leq 2 + \pi + 2 < 8 \) in this case also. The lemma is proved.

Let us use the properties of \( \Sigma(\cdot) \) to estimate the first sum in the right side in (12). Let us expand its summand into a Taylor polynomial of degree 3:

\[
Q_4(x(s + 1, m)) - Q_4(x(s, m)) = Q'_4(x(s, m)) + \frac{Q''_4(x(s, m))}{\sigma \sqrt{n}} + \frac{Q'''_4(x(s, m))}{6 \sigma^3 n \sqrt{n}} + O(n^{-2}) \tag{14}
\]

Note that for every \( i \) the function \( n^{i/2} Q_4(i/n^{i/2} x) \), multiplied by a suitable positive constant, satisfies all the conditions of the lemma. Hence the sum of the three first terms in (14) is \( O(n^{-1}) \). Summing the last term also gives \( O(n^{-1}) \), as before. Thus the first statement in (7) is proved.

**The Second Statement in (7).** This follows from the first one and

\[
\sum_s |\Delta p_s^n| = O(n^{-1/2}) \tag{15}
\]

Let us prove (15). Again we may neglect the remainder term in (10). Without it the right side of (10) certainly has a finite variation, because its derivative changes its sign only a finite number of times. This assures (15). Thus both statements in (7) are proved.

For any real random variable \( \xi \) let us define its moment generation function (MGF) as

\[
\phi(z \mid \xi) = \int_{-\infty}^{\infty} \exp(zx) \, dF(x) \tag{16}
\]
where $F(x) = \text{Prob}(\xi \leq x)$. The MGF $\phi(z \mid \xi)$ is a real function of a real argument $z$, the distribution of $\xi$ serving as a parameter. We use the MGF only when the absolute value of $\xi$ never exceeds a constant; in these case the integral (16) certainly converges. Let us use the MGF for $\xi = \Delta b_i^\prime$ to estimate for any $D > 0$ and $z > 0$

$$\text{Prob}(\Delta b_i^\prime > D) = \int_D^\infty dF(x) \leq \int_{-\infty}^\infty \text{exp}[z(x - D)] \, dF(x)$$

$$= \text{exp}(-zD) \phi(z \mid \Delta b_i^\prime)$$

Note that $\phi(z \mid c \cdot \xi) = \phi(c \cdot z \mid \xi)$ for any random variable $\xi$ and any numbers $z$ and $c$ and that the MGF of a sum of several independent random variables equals the product of their MGFs. Also remember that all $\eta_i^\prime$s are distributed as $v$. All this allows us to rewrite the last expression as

$$\text{exp}(-zD) \prod_{n=0}^{t} \prod_{s} \phi(\Delta p^n_s \cdot \eta_i^\prime \cdot n) = \text{exp}(-zD) \prod_{n=0}^{t} \prod_{s} \phi(\Delta p^n_s \cdot z \mid v)$$

Choosing some integer $v$ between 1 and $t$, we can rewrite this expression as

$$\text{exp}(-zD) \prod_{n=0}^{v-1} \prod_{s} \phi(\Delta p^n_s \cdot z \mid v) \prod_{n=v}^{t} \prod_{s} \phi(\Delta p^n_s \cdot z \mid v)$$

(17)

Let $A$ be a positive constant, whose value will be chosen later. Since we are interested in the asymptotics as $D \to \infty$, we may assume that $AD^2 \geq 1$. Take

$$z = AD^2 \quad \text{and} \quad v = \begin{cases} \lceil AD^2 \rceil & \text{if } 1 \leq \lceil AD^2 \rceil \leq t \\ t & \text{if } t \leq \lceil AD^2 \rceil \end{cases}$$

Let us estimate each product in (17).

**First Product.** From (1) there is a constant $E$ such that $\phi(x \mid v) \leq \text{exp}(E \mid x\rvert)$. Hence

$$\prod_{n=0}^{v-1} \prod_{s} \phi(AD^2 \cdot \Delta p^n_s \mid v) \leq \prod_{n=0}^{v-1} \prod_{s} \exp(AD^2E \cdot |\Delta p^n_s|)$$

$$= \exp \left( AD^2E \cdot \sum_{n=0}^{v-1} \sum_{s} |\Delta p^n_s| \right)$$

$$\leq \exp \left( AD^2E \cdot \sum_{n=0}^{v-1} O(n^{-1/2}) \right)$$

$$\leq \exp(AD^2E \cdot v^{1/2})$$

$$\leq \exp\{AD^2E \cdot O(A^{1/2}D)\}$$

$$= \exp\{O(A^{3/2})\}$$

(18)
Second Product. Due to (7) for $n \geq v$

$$|AD^2 \Delta p^n_s| \leq AD^2 \max_s |\Delta p^n_s| \leq AD^2 \cdot O(n^{-1})$$

$$\leq AD^2 \cdot O(v^{-1}) \leq AD^2 \cdot O\left(\frac{1}{AD^2}\right) = O(1)$$

If we add a constant to the noise of a process, the distribution of its relaxation time does not change. Based on this, we may assume without loss of generality that $E(v) = 0$. (This turns our processes into harnesses, introduced in ref. 7.) Hence from (1) there is a positive constant $E$ such that

$$\phi(x \mid v) \leq \exp(E \cdot x^2).$$

This assures

$$\prod_{n=v}^t \prod_s \phi(AD^2 \cdot \Delta p^n_s \mid v) \leq \prod_{n=v}^t \prod_s \exp\{A^2D^4E \cdot (\Delta p^n_s)^2\}$$

$$= \exp\left(A^2D^4E \sum_{n=v}^t \sum_s (\Delta p^n_s)^2\right)$$

$$\leq \exp\left(A^2D^4E \sum_{n=v}^t O(n^{-3/2})\right)$$

$$\leq \exp\{A^2D^4 \cdot O(v^{-1/2})\}$$

$$= \exp\{O(A^{3/2}D^3)\}$$

(19)

Together (17)–(19) give

$$\operatorname{Prob}(\Delta b'_s > D) \leq \exp\{-AD^3 + O(A^{3/2}D^3)\}$$

Now we can choose $A$ so small that the last expression will not exceed $\exp(-CD^3)$, where $C$ is a positive constant. We have proved that $\operatorname{Prob}(\Delta t'_s > D) \leq \exp(-CD^3)$ uniformly in $s$ and $t$. Summing this over $s$ and $t$, we get for any $T$

$$\operatorname{Prob}(t^0 \leq T) = \operatorname{Prob}(\exists s, t \leq T: \max_s \Delta b'_s > D) \leq TL \exp(-CD^3)$$

whence Proposition 1 follows. \qed

For any finite process $(a'_s)$ with an initial condition $a^0$ for which $\operatorname{rot}(a^0)$ is defined, let us choose an infinite process $(b'_s)$ which we shall call coupled with $a'_s$. As before, $(b'_s)$ is a distribution of real random variables $b'_t$, where $s \in \mathbb{Z}$, induced by the same hidden variables $\eta'_s$, where $s \in \{0, \ldots, L-1\}$, with the map

$$b'_s = \sum_{i=1}^N w_i \cdot b'_{s+v_i} + \eta'_s, \quad \text{where } s \in \mathbb{Z}, \quad t \in \{1, 2, 3, \ldots\}$$
Here components of the initial condition $b^0_s$ for $s \in \mathbb{Z}$ equal

$$b^0_s = a^0_{r(s)} + q(s) \cdot M \cdot \text{rot}(a^0) \quad (20)$$

where $q(s)$ and $r(s)$ are the quotient and the residue when $s \in \mathbb{Z}$ is divided by $L$.

**Proposition 2.** Take any finite process $(a'_t)$ and let $(b'_s)$ be the coupled infinite process. Denote $\Delta v = \max(v_1, ..., v_N) - \min(v_1, ..., v_N)$. While

$$\max_s |\Delta b'_s| < \frac{M - 2v_{\text{max}}}{2(\Delta v + 1)} \quad (21)$$

rot$(a')$ remains defined and equals its initial value rot$(a^0)$.

**Proof.** Let $t$ be less than or equal to the last time when (21) holds. Then rot$(a')$ is defined, because for all $s$

$$|\Delta b'_s| < \frac{M - 2v_{\text{max}}}{2(\Delta v + 1)} < \frac{M}{2}$$

It remains to prove that rot$(a') = \text{rot}(a'-1)$. For any $a, b, c \in R_M$

\[
\begin{cases}
\text{if } \rho(a, b) \text{ and } \rho(b, c) \text{ are defined and } |\rho(a, b)| + |\rho(b, c)| < M/2, \\
\text{then } \rho(a, c) \text{ is also defined and equals } \rho(a, b) + \rho(b, c)
\end{cases}
\quad (22)
\]

Therefore for any $a, b, c, d \in R_M$, if

$$|\rho(a, b)| + |\rho(b, c)| < M/2 \quad \text{and} \quad |\rho(a, d)| + |\rho(d, c)| < M/2$$

then

$$\rho(a, b) + \rho(b, c) = \rho(a, c) = \rho(a, d) + \rho(d, c) \quad (23)$$

Note also that for all $s$

$$\max(b'_{s+v_1}, ..., b'_{s+v_N}) - \min(b'_{s+v_1}, ..., b'_{s+v_N}) < \frac{(M - 2v_{\text{max}}) \Delta v}{2(\Delta v + 1)}$$

whence for all $s$

$$|\rho(a'_{s+v_1}, a'_s)| < \frac{(M - 2v_{\text{max}}) \Delta v}{2(\Delta v + 1)} + v_{\text{max}} \quad (24)$$

Thus for any $s$ both

$$|\rho(a'_{s+v_1}, a'_s)| + |\rho(a'_s, a'_{s+1})| \quad \text{and} \quad |\rho(a'_{s+v_1}, a'_{s+1+v_1})| + |\rho(a'_{s+1+v_1}, a'_{s+1})|$$
are less than the sum of the right sides of (21) and (4), which equals $M/2$. This allows us to apply (23) to obtain

$$\rho(a_{s+1}^{-1}, a_s^i) + \rho(a_s^i, a_{s+1}^i) = \rho(a_{s+1}^{-1}, a_{s+1}^i + v_l^i) + \rho(a_{s+1}^{-1} + v_l^i, a_{s+1}^i)$$

Summing this over $s = 0, \ldots, L - 1$ gives $\text{rot}(a_{s+1}^{-1}) = \text{rot}(a_s^i)$.

Now to prove Theorem 1. Proposition 2 reduces our task to estimating the time while (21) holds in the infinite process coupled with the finite process in question. The initial condition $b_0^i$ of this infinite process is $b_0^i = Rs/L$. [The result of application of (20) to (2).] Proposition 1 can be easily extended to this initial condition, because it adds only a constant $R/L$ to $\Delta b_j^i$. Thus we obtain the estimation

$$L^{-1} \exp \left\{ C_1 \left( D - \left| \frac{R}{L} \right| \right)^3 \right\} \leq E(T)$$

Substitution here of the right side of (21) as $D$ proves Theorem 1.

4. PROOF OF THEOREM 2

Proposition 3. There are positive constants $C_0$, $C_1$, and $K$ such that in every finite process $(a_s^i)$ which corresponds to the system in Theorem 2 with $L > KM$ and with any initial condition $a_0^i$ for which $\text{rot}(a_0^i)$ is defined, the value of $\text{rot}(a^i)$ changes or becomes undefined in the time span $[0, C_1 M^2]$ with a probability which is not less than $\exp(-C_0 M^3)$.

Proof. As before, consider the coupled infinite process $(b_s^i)$. Now

$$\Delta b_0^i = I_i + \sum_{n=0}^{t-1} \sum_{s \geq 0} \Delta p_s^n \cdot \eta_s^{i-n}$$

where $I_i$ is the contribution from the initial condition. Note that for the system in Theorem 2 the random variable $V$ [defined in (8)] has expectation $\mu = 1/2$, and that is why the following definition is useful. Given some $C_1, C_2 > 0$ (whose values will be chosen later), let us denote $T = [C_1 M^2]$ and classify all pairs $(s, n)$, where $0 \leq n < T$, into two classes, "relevant" and "irrelevant," according to the following rule:

$$(s, n) \begin{cases} \text{relevant} & \text{if } |s + n/2| \leq C_2 M \\ \text{irrelevant} & \text{otherwise} \end{cases}$$

Given two functions $f$ and $g$, let $f \asymp g$ mean that $f = O(g)$ and $g = O(f)$. Using the expansion (9)–(11), we can prove that $\sum_s |\Delta p_s^n| \asymp n^{-1/2}$ [which
is a stronger version of (15)]. Based on this, we can choose $K$ and $C_2$ so large that the sum of $|\Delta p_s^n|$ over all irrelevant $(s, n)$ will be less than half of the sum of $|\Delta p_{s}^n|$ over all relevant $(s, n)$. Now let us classify all relevant pairs $(s, n)$ into two classes: “left,” for which $\Delta p_s^n$ is positive, and “right,” for which $\Delta p_s^n$ is negative. Now denote by $E$ the following event:

$$E = \begin{cases} 
\text{if } I_r \geq 0, \text{ then every left variable equals } -1 \\
\text{and every right variable equals } 1 \\
\text{otherwise every left variable equals } 1 \\
\text{and every right variable equals } -1
\end{cases}$$

We can choose $C_1$ so large that, event $E$ assumed, the contribution of hidden variables $\eta_T^{-n}$, where pairs $(s, n)$ are relevant, to $\Delta b_0^r$ will exceed $M$ and therefore $|\Delta b_0^r|$ will certainly exceed $M/2$. Thus we can choose our constants $C_1$, $C_2$, and $K$ in such a way that $E$ will assure $|\Delta b_0^r| > M/2$.

Now let us classify all real numbers $x$ into three groups:

$$x \text{ is } \begin{cases} 
\text{subcritical} & \text{if } |x| < M/2 - 2 \\
\text{critical} & \text{if } M/2 - 2 \leq |x| < M/2 \\
\text{supercritical} & \text{if } M/2 \leq |x|
\end{cases}$$

Denote $y_s' = (b_s' + b_{s+1}')/2$ and $\Delta y_s' = y_{s+1}' - y_s'$. If $t + 1$ is the first time when at least one among $\Delta b_s'^{t+1}$ is supercritical, then

$$\begin{cases} 
\text{none among } \Delta b_s'^t \text{ and } \Delta y_s'^t \text{ is supercritical} \\
\text{but at least one among } \Delta y_s'^t \text{ is critical}
\end{cases}$$

(25)

Of course, $\text{Prob}(E) \geq \exp(-C_0 M^3)$ for any choice of our constants with a suitable $C_0 = \text{const} > 0$. Now let $E_t$ denote the event (25) for any given $t$. We have proved that the union of $E_t$ over $t < T = C_1 M^2$ has probability which is not less than $\exp(-C_0 M^3)$.

To prove Proposition 3, it is sufficient to prove that, $E_t$ assumed, $\text{rot}(a'^{r+1})$ is different from $\text{rot}(a'^r)$ or undefined with a probability which is not less than a positive constant. Let us assume $E$, and denote $x'_s = \mu(y'_s)$. Since none of the $\Delta b_s'$ are supercritical,

$$y_s' - b_s' = b_{s+1}' - y_s' < M/4$$

Since $|\rho(\mu(x), \mu(y))| \leq |y - x|$ for any real $x, y$,

$$|\rho(a'_s, x'_s)| < M/4 \quad \text{and} \quad |\rho(x'_s, a'_{s+1})| < M/4$$
Hence and from (22)

\[ \rho(a_s, a_{s+1}) = \rho(a_s', x_s') + \rho(x_s', a_{s+1}) \]

\[ \rho(x_{s-1}', x_s') = \rho(x_{s-1}', a_s') + \rho(a_s', x_s') \]

Summing this over \( s \) gives \( \text{rot}(x') = \text{rot}(a') \). Now let us consider two cases.

**Case 1.** At least one among \( \Delta y_s' \) is subcritical. Then there is \( s \) such that \( \Delta y_s' \) is critical, but \( \Delta y_{s+1}' \) is subcritical (or vice versa, which is analogous). In this case let us consider two new events:

\[ E^+ : \ \eta_{s-1}' = -1 \quad \text{and} \quad \eta_s' = 1 \]

\[ E^- : \ \eta_{s-1}' = -1 \quad \text{and} \quad \eta_s' = -1 \]

Both events are independent of \( E_t \) and have probability 1/4. Values of \( \text{rot}(a_{t+1}) \) (if defined) are different from each other in these two cases for any particular prehistory and values of all \( \eta_u' \) for \( u \neq s - 1, s \). Therefore in this case the conditional probability (given \( E_t \)) that \( \text{rot}(a_{t+1}) \) is different from \( \text{rot}(a') \) or undefined is not less than 1/4.

**Case 2.** All \( \Delta y_s' \) are critical. Suppose that not all of them have one and the same sign and come to a contradiction. We can find \( s \) such that \( \Delta y_s' \) is positive and \( \Delta y_{s+1}' \) is negative (or vice versa, which is analogous). Then \( y_s' \geq M/2 - 2 \) and \( y_{s+1}' \leq -(M/2 - 2) \), which means that \( (b_s' + b_{s+1}')/2 \geq M/2 - 2 \) and \( (b_{s+1}' + b_{s+2}')/2 \leq -(M/2 - 2) \), whence \( b_s' - b_{s+2}' \geq 2(M - 4) \).

But, according to our choice of \( t \), none of \( \Delta b_s' \), and \( \Delta b_{s+1}' \) is supercritical. This provides a contradiction if \( M \geq 8 \). Thus all \( \Delta y_s' \) have one and the same sign. In this case \( \text{rot}(a_{t+1}) \) is defined and equals \( \text{rot}(x') = \text{rot}(a') \) only if all \( \eta_s' \) have one and the same value for all \( s \). The probability of this events is \( 2^{\frac{1}{1 - \lambda}} \), which is less than or equal to 1/2 for \( \lambda \geq 2 \).

Proposition 3 is proved.
5. FINAL NOTES

Note 1. Theorem 1, which is true whenever $N \geq 2$, contrasts with the degenerate case $N = 1$ (in which components essentially do not interact). For simplicity let us assume that $a_s^0 = 0$ and that the noise $\nu$ is distributed as in Theorem 2. In this case there is a constant $M_0$ such that for all $M > M_0$,

$$\mathbb{E}(t^*) \geq \frac{M^2}{\ln L}$$

(26)

where $t^*$ is the relaxation time, as before.

Let us prove (26). We may assume that $L > L_0 = \text{const}$. Since $N = 1$, we may assume that $\nu = 0$. As before, let $(b_s^t)$ be the coupled infinite process. For every $s \in \{0, \ldots, L - 1\}$, as $t$ grows, $b_s^t$ performs a random walk, every step of which is distributed as (4) independently from all the other walks. Therefore every $\Delta b_s^t$ is a sum of $t$ independent, identically distributed random variables. Every one of these variables has zero expectation, because it is distributed as a difference of two independent random variables, every one of which is distributed as $\nu$. For any $p \in (0, 1)$ let $t(p)$ be the first time $t$ when $\text{Prob}(|\Delta b_0^t| \geq M/2) \geq p$. From the central limit theorem $t(p) \approx M^2/(-\ln p)$. Now we estimate $\mathbb{E}(t^*)$.

Estimation from Below. While

$$\sup_s |\Delta b_s^t| < M/2$$

(27)

rot($a^t$) certainly remains defined and equal to rot($a^0$). So it is sufficient to estimate the expectation of the time while (27) holds. For simplicity we may assume that $L$ is even. Then for any $t$ all the variables $\Delta b_s^t$ may be separated into two groups, those with odd $s$ and those with even $s$, all the variables of each group being mutually independent. Note that, if $p = L^{-2}$, then

$$t(p) = t(L^{-2}) \approx \frac{M^2}{\ln L}$$

(28)

For every $s$ the probability that the absolute value of $\Delta b_s^t$ reaches $M/2$ in time (28) does not exceed $L^{-2}$, therefore the probability that the absolute value of at least one of these variables for even $s$ reaches $M/2$ does not exceed $1 - (1 - L^{-2})^{L/2}$, which tends to zero when $L \to \infty$. The same is true for odd values of $s$, whence the probability that the absolute value of some $\Delta b_s^t$ for $t$ given by (28) reaches $M/2$ also tends to zero. Thus for large enough $L$ this probability is less than $1/2$, whence $\mathbb{E}(t^*)$ is not less than half of (28), which gives the desired estimation of $\mathbb{E}(t^*)$ from below.
**Estimation from Above.** Arguing as in Theorem 2, one can show that, starting from any initial condition, after every $\text{const} \cdot M^2 / \ln L$ steps, the probability that $\text{rot}(a')$ changes its value or becomes undefined exceeds a positive constant, whence the upper estimation for $E(t^*)$ can be deduced.

**Note 2.** It seems natural to consider a nonperiodic analog of our infinite processes, which is a distribution of real random variables $b'_s$, where $s \in \mathbb{Z}$ and $t = 0, 1, 2,...$, induced by independent "hidden" variables $\eta'_s$, where $s \in \mathbb{Z}$ and $t = 1, 2,...$, all of which are distributed as a nonconstant noise $v$, which satisfies (1), with the map

$$b'_s = \sum_{i=1}^{N} w_i \cdot b'_{s+v_i} + \eta'_s$$

(29)

for all $t > 0$, where $b'_0$ are components of the initial condition. Normally more basic objects should be examined first. So it looks like an omission that systems like (29) have not been given enough special consideration (ref. 7 stands out as a valuable study of such systems). Being linear and therefore certainly simpler than our processes, the processes (29) display a property which anticipates our results: On one hand, for any initial condition and any constants $C_1$ and $C_2$ the probability $\text{Prob}(C_1 \leq b'_t \leq C_2)$ tends to zero as $t \to \infty$. But, on the other hand (and this is our point) the distribution of differences $\Delta b'_t$ tends to one and the same probability distribution as $t \to \infty$ when we start from any initial condition $b'_0$ for which $|b'_0| \leq \text{const.}$ (Only ref. 7 anticipated this.) To prove this, let us write an analog of (6) for the present case:

$$\Delta b'_0 = \sum_{n=0}^{r-1} \sum_s \Delta p^n_s \cdot \eta'^{-n}_s + \sum_s \Delta p'_s \cdot b'_s$$

If $|b'_0| \leq \text{const.}$ then due to (15) (which is true in the present case also), the contribution of the initial condition tends to zero as $t \to \infty$. We discuss these systems elsewhere. (15)

**Note 3.** A superexponential relaxation time seems to take place also for some systems with a discrete set of states. This is suggested by our computer simulations, where the set of states of a single component was the set $\mathbb{Z}_M$ of residues modulo $M$, where $M$ was a positive integer number and the noise $v$ was the same as (4). When the transition function $F(a'_{s+v_1}, a'_{s+v_2}, a'_{s+v_3})$ was not defined, $a'_s$ was made equal to all the elements of $\mathbb{Z}_M$ with equal probabilities independent of all the prehistory. We took $N = 3$, $v_1 = -1$, $v_2 = 0$, $v_3 = 1$, and the transition function (acting on the set of residues...
modulo $M$) was the finite-valued version of one the following infinite-valued transition functions:

- $(a_{s-1}' + a_s' + a_{s+1}')/3$, rounded to the nearest integer.
- The median of the same three arguments $a_{s-1}', a_s', a_{s+1}'$. (A median of three numbers is the middle one if they are sorted in increasing order.)

**Note 4.** Let us discuss the unusual non-Gaussian exponent $\exp(const \cdot D^3)$ in Proposition 1 in an informal manner. A physicist would define an energy of a state $b$ of our infinite system as $\sum_s (\Delta b_s)^2$ and, assuming Gibbs distribution, would expect that the probability that $\Delta b_s > D$ would be of the order $\exp(D^2)$, like a Gaussian one. However, in our case, first, energy, as defined, is not conserved in our systems, and, second, due to condition (1), a large value of $\Delta b_s$ cannot emerge in one time step; it has to be accumulated in about $D^2$ steps, and meanwhile energy both decreases as "dissipates" at a distance about $D$, so that the total energy spent is about $D^3$, in accordance with our estimation. We discuss further elsewhere.\(^{(15)}\)

**ACKNOWLEDGMENTS**

The author thanks Deepak Dhar and Eugene Speer for valuable discussions.

**REFERENCES**

One-Dimensional Interaction Systems