Journal of Statistical Physics, Vol. 59, Nos. 1/2, 1990

On The Stability of Crystal Growth

D. J. Gates[1] and M. Westcott[1]

Received March 3, 1989; revision received November 22, 1989

We investigate properties of solid-on-solid models for crystal growth, involving general microscopic rates of capture of atoms by the crystal surface and of escape of atoms. The rates in this Markov process influence the stability of the growing surface. We prove, for various different ranges of the rate parameters, stability (i.e., ergodicity) and instability (i.e., nullity) of the growth process. Symmetry properties of the process, such as reversibility, dynamic reversibility, and reflection invariance, are proved or disproved under various conditions. We give a measure of surface smoothness that distinguishes between stable and unstable growth.

KEY WORDS: Crystal growth; stability; Markov process; ergodicity; transience; null recurrence; dynamic reversibility; surface roughness.

1. INTRODUCTION

The growth of crystalline materials has been simulated by many workers using microscopic models of the growing crystal surface. One class of models, discussed here, involves Markov processes whose transitions are the captures of single atoms by the crystal surface and escapes of single atoms (e.g., refs. 8–10 and 18). Much physical insight has been gained from the simulation of such Markov models: it has provided quantitative information about the important influence of dislocations upon growth rate, and the influence of microscopic rates on surface topography and roughness.

There seems, however, to be no (rigorous) mathematical demonstration in three dimensions that steady crystal growth and a statistically stable surface structure occur in such models. Some computer simulations suggest

[1] CSIRO, Division of Mathematics and Statistics, GPO Box 1965, Canberra ACT 2601, Australia.

0022-4715/90/0400-0073$06.00/0

that this occurs (e.g., ref. 9, Fig. 4), but while a mathematical proof is lacking, there remains the alarming (or perhaps interesting) possibility that, in the longer term, growth might become unsteady and unstable in such models: the surface might become increasingly rough without limit. We shall find (Theorems 2–4, 8, and 9) that this latter effect indeed occurs for a range of microscopic rates.

Therefore, it becomes important to understand conditions under which stable growth does occur. Our Theorem 1 provides some conditions, but we believe them to be excessively stringent.

For crystal growth in two dimensions we studied a stable case in detail and gave an exact description of the growing crystal edge (its stochastic equilibrium probability function) and exact formulas for growth rates (refs. 5–7: these results are in accord with the extensive experimental studies of lamellar crystallization by polymers). Our method relied upon the *dynamic reversibility* of the growth process.

There appears to be no equivalent knowledge of three-dimensional crystal growth, where matters are more difficult. Our Theorem 5 highlights a basic difficulty, by showing that in three dimensions there are no dynamically reversible processes of our type. Certain standard techniques for studying stable growth are therefore not available. Other techniques, notably the general theorems of Foster,[4] based on inequalities, do, however, provide some useful information about stability.

One of our aims is to show that recurrence and transience, in the stochastic sense, provide natural mathematical characterizations of the physical notions of stability and instability.

2. THE MODEL

In the familiar solid-on-solid (SOS) model, atoms are regarded as unit cubes and are stacked on the unit squares of a portion of the integer lattice Z^2, a large chessboard, say. All such arrays of stacks are possible configurations. We are interested only in the surfaces of such arrays, so that surfaces differing only in absolute height are regarded as equivalent states. The centers of the stacks (lying on a shifted Z^2) are labeled (i, j), where $i = 1,..., M$ and $j = 1,..., N$. If $n_{i,j}$ is the height (number of cubes) at site (i, j), we put

$$g_{i,j} = n_{i+1,j} - n_{i,j} \tag{2.1a}$$

and

$$h_{i,j} = n_{i,j+1} - n_{i,j} \tag{2.1b}$$

representing height differences between stacks at (i, j) and stacks to the east and to the north, respectively. We assume periodic boundary conditions

$$g_{i+M,j+N} = g_{i,j} \qquad \text{for all} \quad i, j \tag{2.2}$$

and likewise for $h_{i,j}$. We can impose the conditions

$$\begin{aligned} \sum_i g_{i,j} &= 0 \qquad \text{for all } j \\ \sum_j h_{i,j} &= 0 \qquad \text{for all } i \end{aligned} \tag{2.3}$$

for all time, since the addition or removal of atoms from stacks preserves these sums. Conditions (2.2) and (2.3) imply that, effectively, growth occurs on the surface of a torus and that the base of the chessboard joins row by row and column by column at its opposite edges: so there is no dislocation in the lattice structure. For large M and N, edge effects are expected to be unimportant, so that placing the process on a torus should be an innocuous device. Some of our results can be generalized so as to include various idealized dislocations in the crystal lattice. In this paper, however, we keep matters as simple as possible.

With the periodic boundary conditions, the definitions (2.1) extend to all (i, j). Since the net height difference around any circuit is zero, we have

$$g_{i,j} + h_{i+1,j} - g_{i,j+1} - h_{i,j} = 0 \tag{2.4}$$

Then our state space comprises all vectors

$$\mathbf{h} = \{g_{i,j}, h_{i,j} : i = 1,..., M; j = 1,..., N\} \tag{2.5}$$

satisfying (2.2)–(2.4).

We consider a Markov process on this state space, whose transitions are the additions (captures) of single atoms to (by) stacks, i.e., transitions of the form

$$\mathbf{h} \to \mathbf{h}'$$

where, for some i, j,

$$\begin{aligned} g'_{i-1,j} &= g_{i-1,j} + 1 \\ h'_{i,j-1} &= h_{i,j-1} + 1 \\ g'_{i,j} &= g_{i,j} - 1 \\ h'_{i,j} &= h_{i,j} - 1 \end{aligned} \tag{2.6}$$

while other $g_{k,l}$ and $h_{k,l}$ are unchanged. The state $\mathbf{h}'$ of the form (2.6) is denoted $\mathbf{h}_{i,j}$, so that all transitions are of the form $\mathbf{h} \to \mathbf{h}_{i,j}$. Prior to Section 9, we do not allow atoms to escape from the surface, so we are currently confining our study to the case where crystal growth is rapid enough for escapes to be relatively improbable.

The probability rates $q(\mathbf{h}, \mathbf{h}')$ of these transitions are taken to depend only on the number of neighbor atomic bonds created, or equivalently on the number of faces that are covered when a new atom is captured, namely 1, 2, 3, 4, or 5. Since one horizontal face is always covered, we count only vertical faces, and specify rates

$$\beta_n\colon\ n = 0,\ 1,\ldots,\ 4$$

for transitions that cover n vertical faces, or, equivalently, increase the number of vertical faces on the surface by $4 - 2n$. We take $\beta_n > 0$ for all n, so that the process is irreducible (every state can be reached eventually from every other state), and $\beta_n < \infty$, so there are no instantaneous states. We choose the process to be Markovian so that the probability distribution $p_t(\mathbf{h})$ of state $\mathbf{h}$ at time t satisfies the Kolmogorov forward equation

$$\frac{\partial}{\partial t} p_t(\mathbf{h}) = \sum_{\mathbf{h}'} p_t(\mathbf{h}')\, q(\mathbf{h}', \mathbf{h}) - p_t(\mathbf{h}) \sum_{\mathbf{h}'} q(\mathbf{h}, \mathbf{h}') \tag{2.7}$$

The central question addressed in this paper is the possible existence of a stochastic equilibrium for this process, with probability distribution

$$p_{\mathrm{eq}}(\mathbf{h}) = \lim_{t \to \infty} p_t(\mathbf{h}) \tag{2.8}$$

Such an equilibrium describes a statistically stable surface structure and a steady growth rate[6]

$$\sum_{h} p_{\mathrm{eq}}(\mathbf{h})\, q(\mathbf{h})$$

where

$$q(\mathbf{h}) = \sum_{\mathbf{h}' \neq \mathbf{h}} q(\mathbf{h}, \mathbf{h}')$$

which is the total (probability) rate out of state $\mathbf{h}$.

Since our process has an infinite state space, such an equilibrium is not assured. If we constrain the state space so as to prevent large steps, e.g., take

$$|g_{ij}| \leqslant 1, \qquad |h_{ij}| \leqslant 1$$

and disallow transitions that violate these constraints, then the resulting finite state space guarantees the existence of an equilibrium (ref. 3, p. 392). Jackson,[12] for example, imposed such constraints, but Gilmer and Jackson[10] did not. Such constraints make certain mathematical problems more difficult. For example, in two dimensions we know the stationary distributions in the unconstrained cases,[5,6] but not in the constrained cases: the constraints destroy dynamic reversibility.

We emphasize that the stochastic equilibria studied here are quite distinct from the *thermodynamic equilibrium* between two phases (see Section 9). When crystal growth occurs, the fluid phase is not in thermodynamic equilibrium with the crystal phase.

3. THE MAIN RESULTS AND THEIR SIGNIFICANCE

Here we examine how conditions on the rates β_n influence the existence of (2.8). First we state the results and follow with some explanation and interpretation.

Theorem 1. If

$$(MN-1)^2 \beta_0 < \min(2K\beta_1, 4\beta_2, 2\beta_3, \beta_4) \tag{3.1}$$

where $K = \min(M, N)$, the process is ergodic. For $M = N = 2$, it is ergodic if $\beta_0 < \beta_2 < \beta_4$.

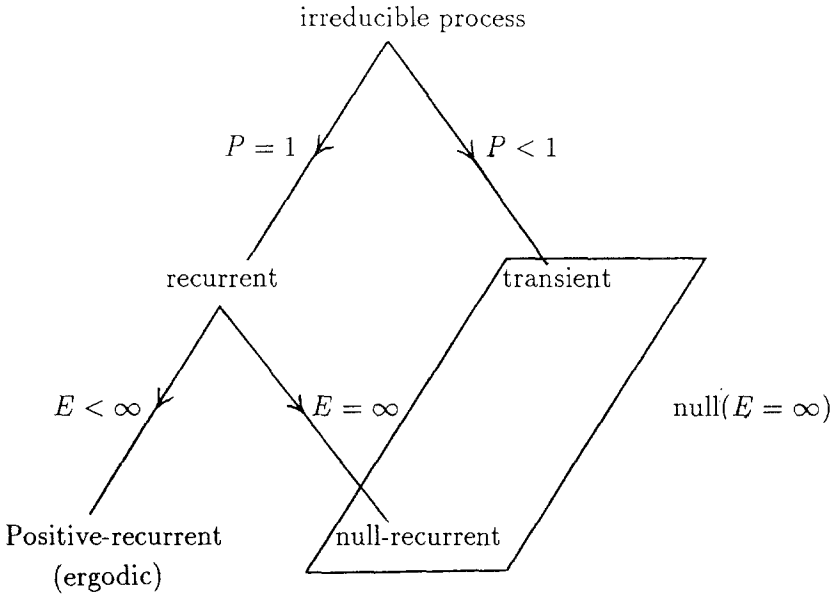
Theorem 2. If $\min(\beta_0, \beta_1) > \max(\beta_3, \beta_4)$, the process is transient.

Theorem 3. If $\min(\beta_0, \beta_1) = \max(\beta_3, \beta_4)$, the process is null.

Theorem 4. If $\beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4$, the process is null-recurrent for $MN = 2$ or 3 and transient for $MN \geqslant 4$.

Theorem 5. If $M = N = 2$, $\beta_0 < \beta_2 < \beta_4$, and $\beta_0 + \beta_4 = 2\beta_2$, the process is dynamically reversible with respect to conjugate states $-\mathbf{h}$. For $M \geqslant 2$ and $N \geqslant 3$ (or vice versa) it is not dynamically reversible for any choice of β_n.

Many standard texts deal with the concepts of ergodic, null, and transient processes.[2,3] We briefly summarize with reference to a simple tree.



Processes are termed *recurrent* if they are certain to return eventually to any starting state, i.e., with probability $P = 1$. If not, they are termed *transient*. Recurrent processes are termed *positive-recurrent* (also called *ergodic*) if their expected time E of first recurrence is finite: otherwise, they are termed *null-recurrent*. *Null* processes comprise both transient and null-recurrent processes: both have $E = \infty$, but for very different reasons. Only ergodic processes have equilibrium distributions. So transient processes are the most errant and unstable, while null-recurrent processes are intermediate in character between transient and positive-recurrent processes.

We emphasize the distinction between our present usage of the term *transient* and its other usage to describe initial or short-term behavior (cf. ref. 9).

The term *dynamically reversible* was coined by Whittle[19] in the stochastic context, and refers to a generalization of the concept of a reversible Markov chain. Details are given by Whittle[19,20] and Kelly[13] and its application to two-dimensional crystal growth by Gates and Westcott[6] and Gates.[5] For dynamically reversible processes, like reversible ones, there are explicit general formulas for equilibrium probability distributions. The significance of Theorem 5, then, is that (except for the trivial case $M = N = 2$) these formulas are inapplicable in three dimensions, so the equilibrium probabilities might be quite difficult to find.

The condition (3.1) of Theorem 1 can be tightened somewhat, as shown in Section 4. The condition can be interpreted as follows. We put

$\beta^* = \min(K\beta_1, 4\beta_2, 2\beta_3, \beta_4)$ and $L = MN$. We suppose that the surface is initially completely flat ($\mathbf{h} = \mathbf{0}$). Then the expected time to the first β_0 transition is $1/q(\mathbf{0}) = 1/(L\beta_0) \equiv T_1$, say. After this transition a new layer can be completed with $L-1$ non-β_0 transitions: this takes a time of order $(L-1)/\beta^* \equiv T_2$, say. Thus, condition (3.1) states that

$$T_2 < T_1$$

which means that layers tend to be completed after a single β_0 transition, before the next β_0 transition occurs. The growing surface therefore tends to be rather flat.

We note that, for two-dimensional crystal growth, a condition of the form $\beta_0 L^2 = \beta^*$ with $L \to \infty$ was shown[6] to lead to a continuum model. So we might expect analogous two-dimensional growth behavior in the near-equality case of condition (3.1). This implies that our proof of ergodicity applies only when the surface is no rougher than those occurring in the continuum regime.

The condition $\beta_0 < \beta_2 < \beta_4$ in the $M = N = 2$ case of Theorem 1 is close to optimal, as we shall show elsewhere. For general M and N we suspect that stable crystal growth occurs under the condition

$$\beta_0 < \beta_1 < \beta_2 < \beta_3 < \beta_4 \tag{3.2}$$

or even a somewhat weaker condition. The heuristic argument is that recurrence of the flat surface state $\mathbf{h} = \mathbf{0}$ is likely if those transitions that have the greatest flattening effect occur at the highest rate. Since β_n transitions have an increasing flattening effect as n increases (they remove more vertical faces), we are led to the conjecture (3.2). The condition of Theorem 1 therefore seems excessively strong.

The condition of Theorem 2 is tending against (3.2), since it includes, for example, the case

$$\beta_0 > \beta_1 > \beta_2 > \beta_3 > \beta_4 \tag{3.3}$$

This latter would seem to favor the creation of an increasingly rough surface, so the result seems intuitively natural. More surprising, perhaps, is that Theorem 2 includes the case

$$\beta_1 > \beta_0 > \beta_4 > \beta_3 \tag{3.4}$$

One might think that by sufficiently favoring β_1 transitions over β_0's, and β_4's over β_3's (with a free choice of β_2), one could induce stable growth.

Theorem 2 shows that this is not so. The condition of Theorem 2 might be closer to optimal.

Theorem 3 is something of an appendage to Theorem 2, but its method of proof is different. We cannot say whether the null process is null-recurrent or transient.

Theorem 4 deals with the case where stacks grow independently, and can be proved by elementary methods. The trivial case $MN = 1$ is static.

We shall not discuss in detail how physical conditions influence the relative sizes of the β_n (see, however, our comments at the end of Section 8 and at the beginning of Section 9). Interatomic attraction favors orderings like (3.2). Limitations on the mobility of atoms favor exposed sites, and hence an ordering more like (3.3). Such limitations are more severe when crystal growth is fast. So fast growth tends more toward a transient process and a surface of anomalous roughness (see Section 10). Our purpose here, however, is not to prejudice particular values of the β_n, but rather to examine the mathematical and physical consequences of various such values.

In proving these theorems, it is sufficient to consider the *embedded* (or *jump*) *chain*, defined as the discrete-time Markov chain whose transitions are the successive state changes in the original process whenever they occur (ref. 1, p. 259; ref. 13, p. 3). It has transition probabilities

$$p(\mathbf{h}, \mathbf{h}') = q(\mathbf{h}, \mathbf{h}')/q(\mathbf{h}) \tag{3.5}$$

Provided $0 < \beta_n < \infty$ for all n, the embedded chain is itself irreducible and has the same character (ergodic, null-recurrent, or transient) as the original process. This is because the process is now *regular*, i.e., uniquely determined by the q's (ref. 1, 1967, Corollary II.19.2), and hence:

(i) Chain and process are recurrent or not together [ref. 17, Lemma 4.2(iv)].

(ii) Chain and process are ergodic or not together (ref. 14, Theorem 3).

4. PROOF OF THEOREM 1

Proofs of Theorems 1 and 2 are based upon general theorems of Foster[4] and Theorem 3 upon a related theorem of Tweedie.[17] The general theorems all involve an unspecified test function whose existence and properties have to be proved either indirectly or, as in our approach,

by explicit choice of such a function. This difficulty limits the utility of these general theorems, and accounts for the incompleteness of our results.

According to a simple extension of Foster's Theorem 2 (cf. ref. 17, 1976, Theorem 6.1), our embedded chain is ergodic if we can find a positive function $y(\mathbf{h})$ such that, for all but a finite number of $\mathbf{h}$,

$$\Sigma(\mathbf{h}) \equiv \sum_{\mathbf{h}'} p(\mathbf{h}, \mathbf{h}')\, y(\mathbf{h}') \leqslant y(\mathbf{h}) - 1 \tag{4.1}$$

and such that

$$\Sigma(\mathbf{0}) < \infty \tag{4.2}$$

Intuitively, (4.1) is saying that the mean change in the function y between jumps of the process is strictly negative, so that, in an average sense, the process is drifting back toward $\mathbf{0}$. This makes the resultant ergodicity very plausible. Also, (4.1) is satisfied as an equality if $y(\mathbf{h})$ is the mean time of first arrival at $\mathbf{0}$ starting from $\mathbf{h}$: this (unknown) y is the optimal test function.

To apply this extended Foster's theorem, we choose

$$y(\mathbf{h}) = \sum_{i=1}^{M} \sum_{j=1}^{N} n^*_{i,j} \tag{4.3}$$

where

$$n^*_{i,j} = \max_{i,j} n_{i,j} - n_{i,j} \tag{4.4}$$

with the $n_{i,j}$ of Section 2. The $n^*_{i,j}$ represent the depths of sites (i, j) below the highest site(s), and $y(\mathbf{h})$ represents the total volume of space above the surface up to the level of the highest site(s). Evidently $y(\mathbf{h})$ is a very crude measure of the expected time of first arrival at $\mathbf{0}$ starting from $\mathbf{h}$.

Putting $L = MN$, we note that a transition at a site (i, j) where $n^*_{i,j} = 0$ increases $y(\mathbf{h})$ by $L - 1$. A transition at other sites decreases $y(\mathbf{h})$ by 1. We put

$$l(\mathbf{h}) = \#\{i, j \mid n^*_{i,j} = 0\}$$

i.e., the number of equal maximal sites on the surface. Then

$$\begin{aligned} \Sigma(\mathbf{h}) &= q(\mathbf{h})^{-1} \left[\beta_0 l(\mathbf{h})\{y(\mathbf{h}) + L - 1\} + \tilde{q}(\mathbf{h})\{y(\mathbf{h}) - 1\}\right] \\ &= y(\mathbf{h}) + q(\mathbf{h})^{-1} \{\beta_0 l(\mathbf{h})(L-1) - \tilde{q}(\mathbf{h})\} \end{aligned} \tag{4.5}$$

where

$$\tilde{q}(\mathbf{h}) = \sum_{i,j:\, n^*_{i,j} \neq 0} q(\mathbf{h}, \mathbf{h}_{i,j})$$

is the total rate out of the nonmaximal sites.

We now prove that, under (3.1), the second term on the right side of (4.5) is bounded below zero whenever $\mathbf{h} \neq \mathbf{0}$. Write $U(\mathbf{h}) = L - l(\mathbf{h})$, the number of nonmaximal sites, which is positive only when $\mathbf{h} \neq \mathbf{0}$.

If $U(\mathbf{h}) = 1$, this nonmaximal site is a β_4 site. So $\tilde{q}(\mathbf{h}) = \beta_4$, $l(\mathbf{h}) = (L-1)$, and the required result follows directly from (3.1).

If $U(\mathbf{h}) = 2$, the nonmaximal sites are either β_3 or β_4. Should there be any β_4 sites, the required result is already implied by the $U(\mathbf{h}) = 1$ case, since $\tilde{q}(\mathbf{h})$ is larger and $l(\mathbf{h})$ smaller here. So, in the new situation, $\tilde{q}(\mathbf{h}) = 2\beta_3$, $l(\mathbf{h}) = (L-2)$, and again (3.1) is sufficient for our needs.

If $U(\mathbf{h}) = 3$, there is either a β_4 site or two β_3 sites. By the argument in the preceding paragraph, this case is subsumed by earlier ones.

If $U(\mathbf{h}) = 4$, the only case not subsumed already is that of four β_2 sites. Here $\tilde{q}(\mathbf{h}) = 4\beta_2$, $l(\mathbf{h}) = (L-1)$, and (3.1) suffices again.

If $U(\mathbf{h}) > 4$, the boundary of the region of nonmaximal sites is either one or more closed curves entirely within the $M \times N$ lattice array, or it also includes regions which reach an edge and are then periodically continued onto the opposite edge. Any closed curve, which of course consists of linked horizontal and vertical pieces, must have at least four corners, hence have at least a β_4 or two β_3 sites, or a β_3 and three β_2 sites. Only the last case is new; but then $\tilde{q}(\mathbf{h}) > \beta_3 + 3\beta_2 > \frac{5}{4}(L-1)^2$ by (3.1), and yet again the required boundedness holds. So closed curves provide nothing new. A little thought shows that in this context, the most extreme case involving just β_1 sites is a two-site-wide "ditch" running straight across between the closer sides. Here $\tilde{q}(\mathbf{h}) = 2K\beta_1$, $l(\mathbf{h}) = (L-1)(L-2K)$, so once more (3.1) is adequate.

We conclude that, under condition (3.1), we can find $\delta > 0$ such that

$$\Sigma(\mathbf{h}) < y(\mathbf{h}) - \delta$$

The rescaled test function $y(\mathbf{h})/\delta$ then satisfies (4.1). Since $\Sigma(\mathbf{0}) = L - 1$, (4.2) also holds, so ergodicity is proved under condition (3.1).

In the case $M = N = 2$, the periodic boundary conditions imply that all transitions are of type β_0, β_2, or β_4. Now relabel sites and steps

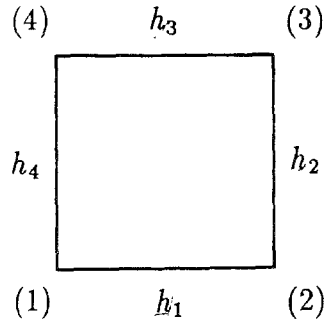
$$(1,1) = (1), \qquad g_{1,1} = h_1$$

$$(2,1) = (2), \qquad h_{2,1} = h_2$$

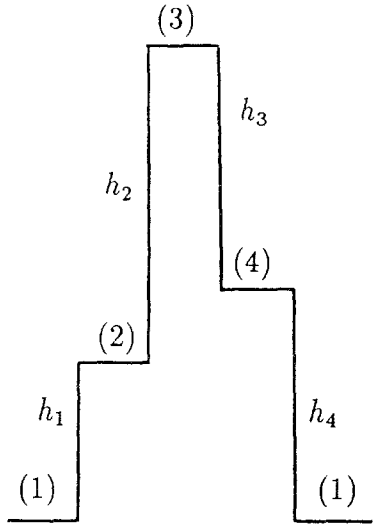
$$(2,2) = (3), \qquad -g_{1,2} = h_3$$

$$(1,2) = (4), \qquad -h_{1,1} = h_4$$

giving the picture



on every cell of the square lattice. Then (2.4) implies $h_1 + h_2 + h_3 + h_4 = 0$, while (2.3) makes no constraint. Hence the sites 1,..., 4 define, equivalently, the edge of a two-dimensional crystal comprising four columns of unit squares; thus



with height differences h_j between columns j and $j+1$ and periodic end conditions. [In this illustration, (3) is a β_0 site, (2) and (4) are β_2 sites, and (1) is a β_4 site.] This is a special case of the model studied extensively by Gates and Westcott[6,7] and Gates.[5]

Now we choose the test function

$$y(\mathbf{h}) = \sum_{i=1}^{4} h_i^2 \tag{4.6}$$

and note that, after a transition at site i,

$$y(\mathbf{h}) \to y(\mathbf{h}) + 2 + 2h_{i-1} - 2h_i \tag{4.7}$$

If $w_i(\mathbf{h})$ denotes the transition rate at i for state $\mathbf{h}$, we therefore have

$$\Sigma(\mathbf{h}) = y(\mathbf{h}) + 2 + z(\mathbf{h}) \tag{4.8}$$

where

$$\begin{aligned} z(\mathbf{h}) &= 2q(\mathbf{h})^{-1} \sum_i w_i(\mathbf{h})(h_{i-1} - h_i) \\ &= 2q(\mathbf{h})^{-1} \sum_i h_i\{w_{i+1}(\mathbf{h}) - w_i(\mathbf{h})\} \end{aligned} \tag{4.9}$$

Suppose $h_i > 0$. Then w_i can only be β_2 or β_4 and w_{i+1} can only be β_0 or β_2. In each of these four possible combinations, the condition $\beta_0 < \beta_2 < \beta_4$ of Theorem 1 implies that

$$w_{i+1} \leqslant w_i \tag{4.10}$$

If $h_i < 0$, the roles of sites i and $i+1$ are merely exchanged, so the inequality (4.10) is reversed. Thus

$$z(\mathbf{h}) \leqslant 0 \qquad \text{for all } \mathbf{h}$$

Now at least two of $h_1, \ldots, h_4$ must be arbitrarily large for all but a finite number of $\mathbf{h}$, since $h_1 + \cdots + h_4 = 0$. Further, at least one of these large h_i must be between β_0 and β_2 or β_0 and β_4 sites; they cannot both lie between a pair of β_2 sites. Therefore at least one summand in $z(\mathbf{h})$ is arbitrarily negative. So if $\beta_0 < \beta_2 < \beta_4$, (4.8) implies that, for some $\delta > 0$,

$$\Sigma(\mathbf{h}) < y(\mathbf{h}) - \delta \tag{4.11}$$

for all but a finite number of $\mathbf{h}$. This establishes (4.1), while (4.2) is true because $\Sigma(\mathbf{0}) = 2$. Thus, Theorem 1 is proved.

There are two observations from the proof. First, the condition (3.1) can clearly be slightly tightened, since, for example, $2\beta_2 > (L-1)(L-2)\,\beta_0$ will suffice in the proof. We chose to use (3.1) as stated because it has the right multiplying factors for the β_i yet is relatively uncluttered. Second, the test function $y(\mathbf{h})$ used for $M = N = 2$ cannot cope with general $M \times 2$ arrays precisely because it *is* then possible to have arbitrarily large h_i only between β_2 sites. So there will be an infinite set of $\mathbf{h}$ for which $z(\mathbf{h})$ cannot be made sufficiently negative to give (4.11).

5. PROOF OF THEOREM 3

We postpone the proof of Theorem 2 to Section 6 because the proof of Theorem 3 is much simpler and involves some concepts useful in the former.

Here we use Theorem 9.1(ii) of Tweedie.[17] A simplified version states that a process of our type is null if there exists a nonnegative function $y(\mathbf{h})$ such that

$$\Sigma(\mathbf{h}) \geqslant y(\mathbf{h}) \qquad \text{for} \quad \mathbf{h} \neq \mathbf{0} \tag{5.1}$$

and, for some $C > 0$,

$$\Delta(\mathbf{h}) \equiv \sum_{\mathbf{h}'} p(\mathbf{h}, \mathbf{h}') \, |y(\mathbf{h}) - y(\mathbf{h}')| < C \qquad \text{uniformly} \tag{5.2}$$

and

$$y(\mathbf{h}) > y(\mathbf{0}) \tag{5.3}$$

To apply the theorem, we choose

$$y(\mathbf{h}) = \sum_{i,j} (|g_{i,j}| + |h_{i,j}|) \tag{5.4}$$

which is simply the total area of all the vertical faces. Thus, if $\mathbf{h} \to \mathbf{h}'$ is a β_0 transition, then $y(\mathbf{h}') = y(\mathbf{h}) + 4$; in general, for a β_n transition $y(\mathbf{h}') = y(\mathbf{h}) + 4 - 2n$. Letting $N_n(\mathbf{h})$ denote the number of β_n sites on state $\mathbf{h}$, we have

$$\begin{aligned} \Sigma(\mathbf{h}) &= y(\mathbf{h}) + 2q(\mathbf{h})^{-1} (2\beta_0 N_0 + \beta_1 N_1 - \beta_3 N_3 - 2\beta_4 N_4) \\ &\geqslant y(\mathbf{h}) + 2q(\mathbf{h})^{-1} D(2N_0 + N_1 - N_3 - 2N_4) \end{aligned} \tag{5.5}$$

by the condition of Theorem 3, where

$$D = \min(\beta_0, \beta_1)$$

Lemma 1:

$$2N_0 + N_1 - N_3 - 2N_4 \geqslant 0 \tag{5.6}$$

Proof. Let $s(\mathbf{h})$ be the total number of nonzero $g_{i,j}$ and $h_{i,j}$, i.e., the total number of steps (of any height $\neq 0$) on the surface. Since every β_n site has exactly n steps facing it, we have

$$s = N_1 + 2N_2 + 3N_3 + 4N_4$$

But s does not exceed the total number of $g_{i,j}$ and $h_{i,j}$, so that

$$s \leqslant 2(N_0 + N_1 + N_2 + N_3 + N_4)$$

The lemma now follows.

The lemma and (5.5) give (5.1). We also have

$$\Delta(\mathbf{h}) = (4\beta_0 N_0 + 2\beta_1 N_1 + 2\beta_3 N_3 + 4\beta_4 N_4)/q(\mathbf{h}) \leqslant 4 \tag{5.7}$$

since

$$q(\mathbf{h}) = \sum_{n=0}^{4} \beta_n N_n \tag{5.8}$$

Noting that (5.3) holds, we complete the proof of Theorem 3. The proof obviously works under the much weaker condition

$$\min(\beta_0, \beta_1) \geqslant \max(\beta_3, \beta_4) \tag{5.9}$$

but the inequality case gives a weaker result than Theorem 2.

6. PROOF OF THEOREM 2

Here we use Theorem 6 of Foster,[4] which states that a Markov process of our type is transient if there is a bounded function $y(\mathbf{h})$ that satisfies, for all $\mathbf{h} \neq \mathbf{0}$,

$$\Sigma(\mathbf{h}) \leqslant y(\mathbf{h}) \tag{6.1}$$

and

$$y(\mathbf{h}) < y(\mathbf{0}) \tag{6.2}$$

for some $\mathbf{h}$. If (6.1) were an equality, $y(\mathbf{h})$ would be the probability of reaching $\mathbf{0}$ from $\mathbf{h}$, in which case (6.2) shows that this will be less than 1 [$= y(\mathbf{0})$ here] for some, hence all, $\mathbf{h} \neq \mathbf{0}$.

We choose

$$y(\mathbf{h}) = \rho^{\lambda/2} \tag{6.3}$$

where $\lambda(\mathbf{h})$ is the right side of (5.4) and $0 < \rho < 1$ is a number yet to be determined. Then $y(\mathbf{h}) < 1 = y(\mathbf{0})$ in accord with (6.2). The next step parallels the proof of Theorem 3. If $\mathbf{h} \to \mathbf{h}'$ is a β_0 transition, then $y(\mathbf{h}') = \rho^2 y(\mathbf{h})$; in general,

$$y(\mathbf{h}') = \rho^{2-n} y(\mathbf{h})$$

for a β_n transition. Thus, with the notation of Section 5,

$$\Sigma(\mathbf{h}) = q(\mathbf{h})^{-1}\, y(\mathbf{h}) \sum_{n=0}^{4} N_n \beta_n \rho^{2-n}$$

$$= y(\mathbf{h}) + q(\mathbf{h})^{-1}\, y(\mathbf{h}) \sum_{n=0}^{4} N_n \beta_n (\rho^{2-n} - 1)$$

using (5.8). Then, omitting $\mathbf{h}$ arguments,

$$\begin{aligned}\Sigma &= y + yq^{-1}(1-\rho)[-N_0\beta_0(1+\rho) - N_1\beta_1 + N_3\beta_3\rho^{-1} \\ &\quad + N_4\beta_4(1+\rho)\,\rho^{-2}] \\ &\leqslant y + yq^{-1}(1-\rho)[-(2N_0+N_1)\,b + (N_3+2N_4)\,B] \qquad (6.4)\end{aligned}$$

where

$$b = \min\{\tfrac{1}{2}\beta_0(1+\rho),\, \beta_1\} \qquad (6.5)$$

$$B = \max\{\beta_3\rho^{-1},\, \tfrac{1}{2}\beta_4(1+\rho)\,\rho^{-2}\} \qquad (6.6)$$

We show below that, under the condition of Theorem 2, we can find $0 < \rho < 1$ such that

$$b > B \qquad (6.7)$$

Then

$$\begin{aligned}\Sigma &< y + yq^{-1}B(1-\rho)(-2N_0 - N_1 + N_3 + 2N_4) \\ &\leqslant y\end{aligned}$$

by Lemma 1, and this is (6.1).

To prove (6.7), suppose first that

$$\min(\beta_0, \beta_1) > \beta_3 > \beta_4 \qquad (6.8)$$

We show that we can find ρ such that (6.7) holds with $B = \beta_3/\rho$. First we note that $B = \beta_3/\rho$ if we take

$$\rho > \beta_4/(2\beta_3 - \beta_4)$$

and this is possible in view of (6.8). Then we must satisfy the conditions

$$\beta_1 > \beta_3/\rho \qquad \text{and} \qquad \tfrac{1}{2}\beta_0(1+\rho) > \beta_3/\rho \qquad (6.9)$$

The former is possible with $\rho > \beta_3/\beta_1$. Since

$$\tfrac{1}{2}\rho(1+\rho) > \tfrac{1}{2}(3\rho - 1)$$

the latter condition in (6.9) is satisfied if

$$\rho > \tfrac{1}{3}(1 + 2\beta_3/\beta_0)$$

which is again admitted by (6.8). To summarize, (6.7) holds under condition (6.8) if we choose

$$\max\{\beta_4/(2\beta_3 - \beta_4), \beta_3/\beta_1, (\beta_0 + 2\beta_3)/(3\beta_0)\} < \rho < 1$$

It remains to consider the case

$$\min(\beta_0, \beta_1) > \beta_4 \geqslant \beta_3 \tag{6.10}$$

Here we show that we can find ρ such (6.7) holds with

$$B = \tfrac{1}{2}\beta_4(1+\rho)/\rho^2 \tag{6.11}$$

First we note that

$$\rho(2\beta_3 - \beta_4) \leqslant \rho\beta_4 < \beta_4$$

so that (6.11) holds for all $0 < \rho < 1$. Then we must satisfy the conditions

$$\tfrac{1}{2}\beta_0(1+\rho) > \tfrac{1}{2}\beta_4(1+\rho)/\rho^2 \qquad \text{and} \qquad \beta_1 > \tfrac{1}{2}\beta_4(1+\rho)/\rho^2$$

The former holds if $\rho > (\beta_4/\beta_0)^{1/2}$ and the latter if $\rho > (\beta_4/\beta_1)^{1/2}$. Thus (6.7) holds under condition (6.10) if we choose

$$\max\{(\beta_4/\beta_0)^{1/2}, (\beta_4/\beta_1)^{1/2}\} < \rho < 1$$

This completes the proof of Theorem 2.

7. PROOF OF THEOREM 4

For $MN = 1$ the process is entirely static with unique state $\mathbf{h} = \mathbf{0}$ (and hence is trivially positive-recurrent). For $MN = 2$, the embedded chain is a simple symmetric random walk if $\beta_0 = \beta_4$ and is null-recurrent (ref. 3, p. 397). Hence, for all $MN > 2$, the process is not positive-recurrent under the condition of the theorem, so it is certainly null. Theorem 4 makes a stronger statement.

If all the β_n are equal, the stochastic process of crystallization is equivalent to running $L = MN$ independent Poisson processes, at the common

rate, one for each possible site. So the embedded chain is equivalent to a multinomial scheme, where at each jump there is a probability $1/L$, independently between jumps, of the new atom going to a particular site.

Clearly, then, we return to that flat state at jump n (which must of course be a multiple of L) if and only if every "cell" of the multinomial distribution has the same number of atoms. Writing $n = kL$, we have

$$\begin{aligned} p^{(n)} &= \Pr\{\text{all cells have the same number of entries, which must be } k\} \\ &= \frac{(kL)!}{(k!)^L}\left(\frac{1}{L}\right)^{kL} \\ &\sim \frac{(kL)^{kL+1/2}\, e^{-kL}(2\pi)^{1/2}}{[k^{k+1/2}\, e^{-k}(2\pi)^{1/2}]^L}\left(\frac{1}{L}\right)^{kL} \qquad \text{by Stirling's formula} \\ &= Ck^{-(L-1)/2} \qquad \text{for some constant } C \end{aligned}$$

If $L = 2$ or 3, then

$$\sum_{n=0}^{\infty} p^{(n)} = \infty \qquad \text{but} \quad p^{(n)} \to 0$$

and so the embedded chain is null-recurrent (ref. 3, p. 389). If $L \geqslant 4$, then

$$\sum_{n=0}^{\infty} p^{(n)} < \infty$$

and so the embedded chain is transient. The process in continuous time then has the same behavior.

8. PROOF OF THEOREM 5

An ergodic process $\mathbf{h}(t)$ is defined as *dynamically reversible*, with conjugate states $-\mathbf{h}$, if the equilibrium process $\mathbf{h}(T-t)$ is probabilistically identical to the equilibrium process $-\mathbf{h}(t)$ for all T and t.[13,19,20] One can choose other conjugate processes, but $-\mathbf{h}(t)$ seems the natural choice for crystal growth models. Dynamic reversibility holds if any only if, for all $\mathbf{h}$,

$$q(\mathbf{h}) = q(-\mathbf{h}) \tag{8.1}$$

and, for any (closed) cycle of states $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_r, \mathbf{h}_1$,

$$\begin{aligned} &q(\mathbf{h}_1, \mathbf{h}_2)\, q(\mathbf{h}_2, \mathbf{h}_3) \cdots q(\mathbf{h}_r, \mathbf{h}_1) \\ &\quad = q(-\mathbf{h}_1, -\mathbf{h}_r)\, q(-\mathbf{h}_r, -\mathbf{h}_{r-1}) \cdots q(-\mathbf{h}_2, -\mathbf{h}_1) \end{aligned} \tag{8.2}$$

Thus, the conditional probability of the cycle equals the conditional probability of the reverse cycle through the conjugate states.

In the case $M=N=2$ we relate the model to the two-dimensional process of Section 4. Under the condition $\beta_0+\beta_4=2\beta_2$ of Theorem 5, dynamic reversibility was proved for such a process by Gates and Westcott.[6] The equilibrium probabilities were also given and these reduce here to

$$p(\mathbf{h})=Z^{-1}(\beta_0/\beta_4)^k \qquad \text{on} \quad g_{1,1}+h_{2,1}-g_{1,2}-h_{1,1}=0 \tag{8.3}$$

where

$$k=\tfrac{1}{2}(|g_{1,1}|+|h_{2,1}|+|g_{1,2}|+|h_{1,1}|)$$

and Z is a normalizing constant. This has a form similar to the familiar probability distribution of Gilmer and Jackson.[10] The major differences are that (a) microscopic rates are different, (b) the Gilmer and Jackson process is reversible (not dynamically reversible) and does not allow net crystal growth (it can describe only a thermodynamic equilibrium between crystal and fluid), and (c) (8.3) applies only if $M=N=2$.

For the remainder of Theorem 5 we need two lemmas.

Lemma 2. Suppose $M\geqslant 3$ and $N\geqslant 2$ (or $M\geqslant 2$ and $N\geqslant 3$). Let $k_n(C)$ $(n=0,\ldots,4)$ be the number of β_n transitions in a closed cycle C. Then:

(i) $2k_0+k_1-k_3-2k_4=0$ for all C.

(ii) There exists C such that $k_0\neq k_4$.

Proof. For (i) we note that $y(\mathbf{h})$ given by (5.4) increases by $4k_0+2k_1-2k_3-4k_4$ during C. Since C is closed, this must be zero. For (ii), consider the cycle that begins with state $\mathbf{h}=\mathbf{0}$ and then places an atom at sites (1, 1), then at (2, 2), i.e., two β_0 transitions. Then, unless $M=N=2$, one can add an atom to every other site without undergoing a β_4 or a β_0 transition until the very last site is occupied, at which point one has returned to state $\mathbf{h}=\mathbf{0}$. To achieve this, one can, for example, start at site (2, 1) and add atoms to neighboring atoms in some convenient pattern, ultimately "painting oneself into the corner" (1, 2) with a β_4 transition. Then, the left side of (8.2) has the form

$$\beta_0^2\beta_1^i\beta_2^j\beta_3^k\beta_4 \tag{8.4}$$

for some integers i, j, k. For example, if $M=3$ and $N=2$, one can add atoms to sites in the sequence (1, 1), (2, 2), (2, 1), (3, 1), (3, 2), (4, 2) = (1, 2) with rate product $\beta_0^2\beta_2\beta_3^2\beta_4$. This proves Lemma 2.

Lemma 3. Equation (8.1) holds for all $\mathbf{h}$ if and only if

$$\beta_n = \alpha + n\nu, \qquad n = 0,..., 4 \tag{8.5}$$

with $\alpha > 0$, $\alpha + 4\nu > 0$.

To prove Lemma 3, we note that (8.5) implies

$$q(\mathbf{h}) = MN\alpha + \nu s(\mathbf{h}) \tag{8.6}$$

with the $s(\mathbf{h})$ of Lemma 1. Then (8.1) follows immediately. For the reverse implication, we need look only at particular states $\mathbf{h}$. For a state comprising one atom on a large, flat layer we have

$$q(\mathbf{h}) = \beta_0 + 4\beta_1 + (MN - 5)\,\beta_0$$

and

$$q(-\mathbf{h}) = \beta_4 + (MN - 1)\,\beta_0$$

so that (8.1) implies

$$4\beta_1 = 3\beta_0 + \beta_4$$

The state with atoms at (1, 1) and (2, 1) on an otherwise flat layer gives, similarly,

$$3\beta_1 = 2\beta_0 + \beta_3$$

The state with atoms at (1, 1) and (2, 2) gives

$$2\beta_1 + \beta_2 = 2\beta_0 + \beta_4$$

and the state with atoms at (1, 1), (2, 2), and (3, 1) gives

$$5\beta_1 + 2\beta_2 + \beta_3 = 5\beta_0 + 3\beta_4$$

These four relations among the β_n reduce to

$$\beta_4 - \beta_3 = \beta_3 - \beta_2 = \beta_2 - \beta_1 = \beta_1 - \beta_0$$

This implies (8.5), which completes the proof of Lemma 3.

Now we note that if $\mathbf{h} \to \mathbf{h}'$ is a β_n transition, then $-\mathbf{h}' \to -\mathbf{h}$ is a β_{4-n} transition. Thus, (8.2) takes the form

$$\beta_0^{k_0}\beta_1^{k_1}\beta_2^{k_2}\beta_3^{k_3}\beta_4^{k_4} = \beta_4^{k_0}\beta_3^{k_1}\beta_2^{k_2}\beta_1^{k_3}\beta_0^{k_4}$$

Using Lemma 2(i), this reduces to

$$(\beta_0/\beta_4)^{k_0-k_4} = (\beta_1/\beta_3)^{2(k_0-k_4)}$$

and Lemma 2(ii) implies

$$\beta_0/\beta_4 = (\beta_1/\beta_3)^2 \tag{8.7}$$

This, together with (8.5), holds only if $v=0$. Thus, (8.1) holds and (8.2) holds for all closed cycles only if $v=0$. But Theorem 4 says this case is not ergodic and hence is not dynamically reversible. This completes the proof of Theorem 5.

One can give a physically plausible argument that (8.1), or equivalently (8.5), should apply to crystal and related growth processes, as follows. Suppose each face to which an atom can attach attracts that atom independently of other faces, horizontal faces with probability $\alpha\, dt$, vertical faces with probability $v\, dt$ in a small time dt. Then the probability that, in time dt, an atom attaches to a β_n site is the probability of attachment to the horizontal face or to any one of the n vertical faces at this site. The probability of this union of events is

$$\begin{aligned}\beta_n\, dt &= 1-(1-\alpha\, dt)(1-v\, dt)^n \\ &= (\alpha+nv)\, dt + O\{(dt)^2\}\end{aligned}$$

so that (8.5) follows.

9. ESCAPES AND TWO-PHASE (THERMODYNAMIC) EQUILIBRIUM

When crystal growth is slow, events in which atoms escape from the crystal surface have significant probability. These probabilities are of similar magnitude to the capture probabilities if the crystal–fluid system is in two-phase (thermodynamic) equilibrium, when there is no net crystal growth. Escape events dominate during the melting, dissolving, or vaporizing of the crystal surface.

Assuming that escape rates of atoms are determined (like capture rates) by the number of atomic contacts prior to escape leads to another five parameters analogous to the β_n. A convenient way to represent escapes is to first consider another process with capture event only. Let $q'(\mathbf{h}, \mathbf{h}')$ be its transition rate matrix taking values β'_n, $n=0,\ldots,4$, for captures which cover n vertical faces (as before).

If $q'(\mathbf{h}, \mathbf{h}')$ is a β'_n capture, then $q'(-\mathbf{h}, -\mathbf{h}')$ is an escape for which the surface gains $4-2n$ vertical faces. For example, a single atom sitting on top

of a flat layer escapes at rate β'_4, and an atom escapes from within a flat at rate β'_0. Then the new process with transition matrix

$$\tilde{q}(\mathbf{h}, \mathbf{h}') = q(\mathbf{h}, \mathbf{h}') + q'(-\mathbf{h}, -\mathbf{h}') \tag{9.1}$$

admits both captures at rates β_n and escapes at rates β'_n. There are no composite transitions, because the transitions $\mathbf{h} \to \mathbf{h}_{i,j}$ and $-\mathbf{h} \to -\mathbf{h}_{i,j}$ are distinct.

Our main results are minor extensions of Theorems 1–5. Theorems 6 and 7 offer something new.

Theorem 1E. Suppose that (3.1) holds separately for the β_n and the β'_n, with at least one of the two inequalities being strict. Then the process is ergodic.

Theorem 2E. If $\min(\beta_0, \beta_1) \geqslant \max(\beta_3, \beta_4)$ and $\min(\beta'_0, \beta'_1) \geqslant \max(\beta'_3, \beta'_4)$ and at least one of the inequalities is strict, then the process is transient.

Theorem 3E. If $\min(\beta_0, \beta_1) = \max(\beta_3, \beta_4)$ and $\min(\beta'_0, \beta'_1) = \max(\beta'_3, \beta'_4)$, then the process is null.

Theorem 4E. If $\beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4$ and $\beta'_0 = \beta'_1 = \beta'_3 = \beta'_4$, then the process is null-recurrent for $MN = 2$ or 3 and transient for $MN \geqslant 4$.

Theorem 5E. If $M = N = 2$, $\beta_0 - 2\beta_2 + \beta_4 = \beta'_0 - 2\beta'_2 + \beta'_4$, and $\beta_0/\beta_4 = \beta'_0/\beta'_4$, then the process is dynamically reversible with equilibrium probability function (8.3). If $M \geqslant 2$ and $N \geqslant 3$ (or vice versa), then the process is not dynamically reversible for any choice of β_n and β'_n.

Theorem 6. If $\beta'_n = \eta\beta_n$ for all n, and the β_n are given by (8.5), then the process has growth rate

$$(1 - \eta)\, G(t) \tag{9.2}$$

where $G(t)$ is the growth rate, at time t, in the $\eta = 0$ model.

Theorem 7. Suppose $\beta'_n = \beta_n$. Then for $M = N = 2$ and under the conditions of Theorem 5E, the process is reversible. For $M \geqslant 2$ and $N \geqslant 3$ (or vice versa) and the conditions of Theorem 6, the process is not reversible.

Theorem 1E is proved by the method of Theorem 1. Now (4.5) has added to it another similar sum involving escapes. Treating this in the same way gives the theorem.

The conditions of the theorem are physically plausible: exposed atoms with no side neighbors escape at the highest rate β'_4; atoms *locked* within a layer escape at the lowest rate β'_0, so there is a tendency toward a flatter

surface (and hence toward recurrence). We would expect, however, to find ergodicity under (at least) the conditions (3.2) and

$$\beta'_0 < \beta'_1 < \beta'_2 < \beta'_3 < \beta'_4 \tag{9.3}$$

Theorems 2E and 3E are proved by slight modifications of previous proofs (Sections 5 and 6). In both cases, one adds to $\Sigma(\mathbf{h})$ further terms involving the β'_n and the $N_n(-\mathbf{h})$. These terms can be dealt with in the same way as, and independently of, the previous terms. The results are physically plausible because they tend against (3.2) and (9.3) and hence toward increasingly rough surfaces.

Theorem 4E implies, like Theorem 4, independently growing columns and is proved by the same method.

For Theorem 5E, in the case $M=N=2$, one follows the proof of Theorem 5 and verifies (8.1) and (8.2) in the manner of Gates and Westcott.[6] For $M \geqslant 3$, $N \geqslant 2$ (or vice versa) we note first from (9.1) that

$$\tilde{q}(\mathbf{h}) - \tilde{q}(-\mathbf{h}) = q^*(\mathbf{h}) - q^*(-\mathbf{h})$$

where

$$q^*(\mathbf{h}) = q(\mathbf{h}) - q'(\mathbf{h})$$

So, following the proof of Lemma 3 (negative values of q^* do not matter) gives the following result.

Lemma 4. $\tilde{q}(\mathbf{h}) = \tilde{q}(-\mathbf{h})$ for all $\mathbf{h}$ if and only if

$$\beta_n - \beta'_n = \alpha + n\nu, \qquad n = 0,..., 4$$

Lemma 2 holds in this new context with the k_n replaced by $k_n + k'_n$, where $k'_n(C)$ is the number of β'_n transitions in a closed cycle C.

Now one can find cycles with only β_0, β_4, β'_0, and β'_4 transitions. For these, (8.2) gives

$$(\beta_0/\beta_4)^{k_0 - k_4} = (\beta'_0/\beta'_4)^{k'_4 - k'_0}$$

But $(k_0 + k'_0) = (k_4 + k'_4)$ and there are cycles with $k_0 \neq k_4$. Thus, (8.2) implies

$$\beta_0/\beta_4 = \beta'_0/\beta'_4$$

and similarly

$$\beta_1/\beta_3 = \beta'_1/\beta'_3$$

Using (8.5) and putting $b_n = \beta_n - \beta'_n$, we deduce that

$$b_0 b_3^2 = b_1^2 b_4$$

and, from Lemma 4, $b_n = \alpha + n\nu$. These conditions imply $\nu = 0$. The equality of the b_n implies $\beta_0 = \beta_4$, $\beta_1 = \beta_3$, $\beta'_0 = \beta'_4$, and $\beta'_1 = \beta'_3$. Then $\tilde{q}(\mathbf{h}, \mathbf{h}') = \tilde{q}(-\mathbf{h}', -\mathbf{h})$, so the equilibrium equations are satisfied by a $p(\cdot)$ that is independent of $\mathbf{h}$. Such a p is not normalizable, so the process is null (ref. 13, p. 3).

For Theorem 6, we note that the growth rate, in general, is the expectation of

$$q(\mathbf{h}) - q'(-\mathbf{h}) \tag{9.4}$$

over $p_t(\mathbf{h})$. If $\beta'_n = \eta\beta_n$, then

$$q'(\mathbf{h}, \mathbf{h}') = \eta q(\mathbf{h}, \mathbf{h}') \tag{9.5}$$

Also, the β'_n are then of the form (8.5) (since the β_n are), so that, by Lemma 2,

$$q'(-\mathbf{h}) = q'(\mathbf{h}) \tag{9.6}$$

Now (9.5) and (9.6) imply

$$q'(-\mathbf{h}) = \eta q(\mathbf{h})$$

from which (9.2) follows. The significance of Theorem 6 is that the growth rate is trivially modified by escape events under the conditions of the theorem, and that $\eta = 1$ gives zero growth. This idea was exploited in the two-dimensional case,[5] where growth rates can be found exactly.

For $M = N = 2$ the statement of Theorem 7 follows from Proposition 1 of ref. 5. For $M \geqslant 3$ and $N \geqslant 2$ (or vice versa) we note that (9.1) reduces here to

$$\tilde{q}(\mathbf{h}, \mathbf{h}') = q(\mathbf{h}, \mathbf{h}') + q(-\mathbf{h}, -\mathbf{h}') \tag{9.7}$$

whence

$$\begin{aligned} \tilde{q}(\mathbf{h}, \mathbf{h}') &= \tilde{q}(-\mathbf{h}, -\mathbf{h}') \\ &= \begin{cases} q(\mathbf{h}, \mathbf{h}') & \text{for captures} \\ q(-\mathbf{h}, -\mathbf{h}') & \text{for escapes} \end{cases} \end{aligned} \tag{9.8}$$

The reversibility of a process is equivalent to the Kolomogorov cycle condition,[20]

$$\tilde{q}(\mathbf{h}_1, \mathbf{h}_2)\, \tilde{q}(\mathbf{h}_2, \mathbf{h}_3) \cdots \tilde{q}(\mathbf{h}_r, \mathbf{h}_1) = \tilde{q}(\mathbf{h}_1, \mathbf{h}_r)\, \tilde{q}(\mathbf{h}_r, \mathbf{h}_{r-1}) \cdots \tilde{q}(\mathbf{h}_2, \mathbf{h}_1) \tag{9.9}$$

for all cycles $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_r, \mathbf{h}_1$. But (9.8) and our counterexamples to (8.2) show that (9.9) does not always hold.

The significance of Theorem 7 is that reversibility, and concomitant conditions like *detailed balance*, need not hold when there is zero growth rate, even in plausible models. Objectively this may seem unremarkable, but the notion of detailed balance is so entrenched in the physics literature that an exception seems to cut across convention. But we know of no compelling physical argument for reversibility.

Theorem 7 describes zero growth rate, but does it describe two-phase equilibrium? Probably, but we have no general proof. If $\nu > 1/4(MN-1)^2\alpha$, such an equilibrium exists by Theorem 1E. When an equilibrium $p(\mathbf{h})$ does exist, we do at least have one familiar reflection invariance property[10]

$$p(\mathbf{h}) = p(-\mathbf{h}) \tag{9.10}$$

To prove this, note that (9.8) implies

$$\tilde{q}(\mathbf{h}) = \tilde{q}(-\mathbf{h}) \tag{9.11}$$

The equilibrium equations [cf. (2.7)] are

$$p(\mathbf{h})\,\tilde{q}(\mathbf{h}) = \sum_{\mathbf{h}'} p(\mathbf{h}')\,\tilde{q}(\mathbf{h}', \mathbf{h}) \tag{9.12}$$

Changing $\mathbf{h}$ to $-\mathbf{h}$ and $\mathbf{h}'$ to $-\mathbf{h}'$ and using (9.8) and (9.11) gives

$$p(-\mathbf{h})\,\tilde{q}(\mathbf{h}) = \sum_{\mathbf{h}'} p(-\mathbf{h}')\,\tilde{q}(\mathbf{h}', \mathbf{h}) \tag{9.13}$$

The uniqueness of the solution $p(\mathbf{h})$ of (9.12) then implies (9.10).

By the same argument applied to (2.7), we find that (9.8) implies $p_t(\mathbf{h}) = p_t(-\mathbf{h})$ for all t provided $p_0(\mathbf{h}) = p_0(-\mathbf{h})$.

10. MEASURES OF SURFACE ROUGHNESS

We have already noted in Section 3 that ergodic processes eventually achieve statistical stability, with an equilibrium distribution over the possible states (2.5) of the surface. By contrast, no such distribution or stability occurs for null processes. Mathematically, this is because null processes spend too much time in states with arbitrarily large indices $\mathbf{h}$ (rough surface states); indeed, transient processes eventually never return to any finite set of states. In the ergodic case the chance of the surface being in a state far removed from $\mathbf{0}$ (the flat state) becomes small. That is, the

degree of irregularity, or *roughness*, of the surface is controlled in ergodic cases, but not in null cases.

To quantify this, we need to order the states $\mathbf{h}$ using a convenient index $j(\mathbf{h}) \in Z_+$. Two obvious condidates are the right sides of Eqs. (4.3) and (5.4), these being measures of roughness. For mathematical reasons, it is preferable to define a *measure of smoothness* $\sigma = \sigma(\mathbf{h})$, where $\sigma(\cdot)$ is a positive *nonincreasing* function of $j(\cdot)$ and $\sigma(\mathbf{h}) \downarrow 0$ as $j(\mathbf{h}) \to \infty$. So a high roughness measure corresponds to a small smoothness measure.

Suppose now that our process commences in some state $\mathbf{h}_0$ at time $t = 0$. Let $\mathbf{h}(t)$ represent the state of the process at time t, with the probability distribution $p_t(\mathbf{h}; \mathbf{h}_0)$ of (2.7) (we now acknowledge the initial state explicitly). Then the expected value of the measure of smoothness is

$$\begin{aligned} e(t) &= \langle \sigma\{\mathbf{h}(t)\} \mid \mathbf{h}(0) = \mathbf{h}_0 \rangle \\ &= \sum_{\mathbf{h}} \sigma(\mathbf{h})\, p_t(\mathbf{h}; \mathbf{h}_0) \end{aligned}$$

Theorem 8. (i) If the process is ergodic, $\lim_{t \to \infty} e(t)$ exists and is positive $\forall \mathbf{h}_0$. (i) If the process is null, $\lim_{t \to \infty} e(t) = 0$ $\forall \mathbf{h}_0$.

Proof. (i) In the ergodic case, $p_{\text{eq}}(\mathbf{h})$ exists and is strictly positive $\forall \mathbf{h}$ [ref. 1 (II.10)]. Since $\sum_{\mathbf{h}'} p_{\text{eq}}(\mathbf{h}')\, p_t(\mathbf{h}; \mathbf{h}') = p_{\text{eq}}(\mathbf{h})$, we have[11]

$$p_t(\mathbf{h}; \mathbf{h}_0) \leqslant p_{\text{eq}}(\mathbf{h}) / p_{\text{eq}}(\mathbf{h}_0) \tag{10.2}$$

Since $\sigma(\mathbf{h})$ is nonincreasing, it is bounded, so $\lim_{t \to \infty} e(t) = \sum_{\mathbf{h}} \sigma(\mathbf{h})\, p_{\text{eq}}(\mathbf{h})$ exists $\forall \mathbf{h}_0$ by dominated convergence and is strictly positive because p_{eq} is.

(ii) In the null case, $\lim_{t \to \infty} p_t(\mathbf{h}; \mathbf{h}_0) = 0$ [ref. 1, (II.10)]. Because $\sigma(\mathbf{h}) \downarrow 0$, given any $\varepsilon > 0$ there is a set H of states, whose complement H^c is finite, such that $\sigma(\mathbf{h}) < \varepsilon$ if $\mathbf{h} \in H$. Let $B = \sup_{\mathbf{h} \in H^c} \sigma(\mathbf{h}) < \infty$. Now

$$\begin{aligned} e(t) &= \left\{ \sum_{\mathbf{h} \in H^c} + \sum_{\mathbf{h} \in H} \right\} \sigma(\mathbf{h})\, p_t(\mathbf{h}; \mathbf{h}_0) \\ &\leqslant B \sum_{\mathbf{h} \in H^c} p_t(\mathbf{h}; \mathbf{h}_0) + \varepsilon \end{aligned}$$

so $\limsup_{t \to \infty} e(t) \leqslant \varepsilon$ $\forall \mathbf{h}_0$ and the result follows.

Thus, our average measure of smoothness remains positive for ergodic processes, but becomes arbitrarily small for null processes, as we would hope. Notice that the theorem applies to *any* measure σ with the appropriate properties.

It is possible to refine (ii) to distinguish between the null-recurrent and transient cases. Further analysis shows that, for a transient process, we have the stronger result

$$\lim_{t \to \infty} \sigma\{\mathbf{h}(t)\} = 0 \quad \text{for all } \mathbf{h}(t) \text{ outside a set of zero probability} \tag{10.3}$$

Clearly, this is not the case for any recurrent process, which will return to $\mathbf{0}$ infinitely often, so $\limsup_{t \to \infty} \sigma\{\mathbf{h}(t)\} = \sigma(\mathbf{0})$ with probability one.

One specific example of a smoothness measure that we have already met is the function defined at (6.3) and used in the proof of transience in Theorem 2. For this measure we can obtain a stronger (though less general) result for the regime of Theorem 2.

Theorem 9. If $\sigma(\mathbf{h})$ is given by (6.3) and $\min(\beta_0, \beta_1) > \max(\beta_3, \beta_4)$, then, for all $\mathbf{h}_0$,

$$e(t) < e^{-\mu t} \tag{10.4}$$

where $\mu = (1-\omega) \inf q(\mathbf{h}) > 0$ and ω $(0 < \omega < 1)$ is defined below.

This shows that surface roughness grows, in an average sense, at an exponential rate. So, physically, the surface development should be sharply distinct from the stable (ergodic) case.

Proof. In Section 6 we proved $b > B$ for large enough ρ, say $\rho_{\min} < \rho < 1$. Then from (6.4) and Lemma 1 we have

$$\Sigma \leqslant y - yq^{-1}(1-\rho)(b-B)(2N_0 + N_1)$$

Since $N_0 \geqslant 1$, we have $2N_0 + N_1 \geqslant 2$. Also,

$$\begin{aligned} q &= \sum_0^4 \beta_n N_n \\ &\leqslant \beta_0 N_0 + \beta_1 N_1 + \beta_2 N_2 + \min(\beta_0, \beta_1)(N_3 + N_4) \\ &\leqslant cMN \end{aligned}$$

where $c = \max(\beta_0, \beta_1, \beta_2)$. Thus,

$$\Sigma(h) \leqslant \omega y(\mathbf{h}) \tag{10.5}$$

where

$$\omega(\rho) = 1 - 2(1-\rho)(b-B)/(MNc) \tag{10.6}$$

so that $0 < \omega < 1$. If $p^{(n)}(\mathbf{h}, \mathbf{h}')$ are the n-step transition probabilities of the embedded chain, then iteration of (10.5) gives

$$e^{(n)} \equiv \sum_{\mathbf{h}'} p^{(n)}(\mathbf{h}, \mathbf{h}')\, y(\mathbf{h}') \leqslant \omega^n y(\mathbf{h}) \leqslant \omega^n \tag{10.7}$$

If $\rho \leqslant \rho_{\min}$, we define

$$\omega(\rho) = \min_{\rho_{\min} < \rho < 1} \omega(\rho) \tag{10.8}$$

and obtain for the n-step conditional expectation of y,

$$e^{(n)} \leqslant \omega^n \tag{10.9}$$

for all ρ. As in Theorem 8, we have the contrasting result that $e^{(n)}$ is bounded away from zero in the ergodic case.

To prove (10.4), use (10.5), (3.4), and the usual convention $q(\mathbf{h}, \mathbf{h}) = -q(\mathbf{h})$ to get

$$\begin{aligned} \sum_{\mathbf{h}'} q(\mathbf{h}, \mathbf{h}')\, \sigma(\mathbf{h}') &\leqslant -(1-\omega)\, q(\mathbf{h})\, \sigma(\mathbf{h}) \\ &\leqslant -\mu\sigma(\mathbf{h}), \qquad \forall \mathbf{h} \end{aligned} \tag{10.10}$$

So $\sigma(\cdot)$ is a μ-subinvariant function for $q(\cdot, \cdot)$, with $\mu \leqslant \inf q(\mathbf{h})$. Then from the dual to Proposition 1(i) of ref. 15, $\sigma(\cdot)$ is also μ-subinvarant for the process $\mathbf{h}(t)$, which is minimal since it is regular. That is,

$$\begin{aligned} \sum_{h} p_t(\mathbf{h}; \mathbf{h}_0)\, \sigma(\mathbf{h}) &\leqslant e^{-\mu t}\sigma(\mathbf{h}_0) \\ &\leqslant e^{-\mu t}, \qquad \forall t, \mathbf{h}_0 \end{aligned}$$

which is (10.4).

Note that our method also proves that $p_t(\mathbf{h}; \mathbf{h}_0) = O(e^{-\lambda t})$ for all $\mathbf{h}$, $\mathbf{h}_0$ and some $\lambda \geqslant \mu$ [cf. ref. 15, the proof of Proposition 1(ii)], which significantly supplements the general transience result $p_t(\mathbf{h}; \mathbf{h}_0) \to 0$. A consequence of this result is that, for *any* finite set of states H,

$$\Pr\{\mathbf{h}(t) \in H \mid \mathbf{h}(0) = \mathbf{h}_0\} = O(e^{-\lambda t}) \tag{10.11}$$

That is, the process has an exponentially small chance of being in any finite set of states as time increases. This is yet another indication of the rapid tendency to roughness in the transient regime of the theorem.

One consequence is that simulations with $\min(\beta_0, \beta_1) \gg \max(\beta_3, \beta_4)$ should contrast sharply with stable cases. On the other hand, null-recurrent cases might be quite hard to identify by simulation alone.

11. UNSOLVED PROBLEMS

In our approach, where the growth model is defined solely in terms of microscopic rates for single-atom transitions (discrete state space), under-

standing of the resulting surface growth behavior in mathematical terms has scarcely begun. If one wants to study stable (ergodic or steady-state) growth, all one has is a proof of its occurrence under Theorems 1 and 1E and in the constrained cases (Section 2). The conditions of these theorems are unsatisfactory, both physically (as discussed in Section 3) and also mathematically: Theorems 1–4 leave a large part of the $(\beta_0,..., \beta_4)$ parameter space unclassified.

Furthermore, there is no stable case, with net growth, where the stationary distribution (of **h**) is known. Worse still, Theorems 5 and 5E indicate why such a stationary distribution might be difficult to find. The equations for this distribution [i.e., $\partial p_t/\partial t = 0$ in (2.7)] have none of the nice balance properties that would render them tractable. Until one has such a distribution, the detailed study of surface properties, through partition functions and so on, cannot even begin.

How can one improve on Theorem 1? As we have already said (Section 4), Foster's theorems, though potentially powerful, offer no guide to the choice of test function $y(\mathbf{h})$. We chose our three different test functions by a combination of trial and error and a little physical intuition. There may be scope for much better choices: basically one wants a test function that *drifts* in the right direction under weak enough conditions on the rates β_n. Both physicists and mathematicians might be able to contribute to this problem.

The unstable (transient and null-recurrent) cases are more likely to be relevant to rapid growth. Only in the independent-sites case $(\beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4)$ can we say anything about the detailed surface structure. When neighboring sites are dependent, there is scope for a rich variety of surface structures as one moves through the parameter sets given in Theorems 2, 3, 2E, and 3E.

REFERENCES

1. K. L. Chung, *Markov Chains with Stationary Transition Probabilities*, 2nd ed. (Springer-Verlag, Berlin, 1967).
2. D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes* (Methuen, London, 1965).
3. W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. I, 3rd ed. (Wiley, New York, 1968).
4. F. G. Foster, On the stochastic matrices associated with certain queueing processes, *Ann. Math. Stat.* **24**:355–360 (1953).
5. D. J. Gates, Growth and decrescence of two-dimensional crystals: A Markov rate process, *J. Stat. Phys.* **52**:259–271 (1988).
6. D. J. Gates and M. Westcott, Kinetics of polymer crystallization. I. Discrete and continuum models, *Proc. R. Soc. Lond. A* **416**:443–461 (1988).

7. D. J. Gates and M. Westcott, Kinetics of polymer crystallization. II. Growth regimes, *Proc. R. Soc. Lond. A* **416**:463–476 (1988).
8. G. H. Gilmer, Growth on imperfect crystal surfaces, *J. Crystal Growth* **35**:15–28 (1976).
9. G. H. Gilmer, Transients in the rate of crystal growth, *J. Crystal Growth* **49**:465–474 (1980).
10. G. H. Gilmer and K. A. Jackson, Computer simulation of crystal growth, in *Crystal Growth and Materials*, E. Kaldis and H. J. Scheel, eds. (North-Holland, Amsterdam, 1976), pp. 80–114.
11. P. J. Holewijn and A. Hordijk, On the convergence of moments in stationary Markov chains, *Stochastic Proc. Appl.* **3**:55–64 (1975).
12. K. A. Jackson, On the theory of crystal growth: Growth of small crystals using periodic boundary conditions, *J. Crystal Growth* **34**:507–511 (1968).
13. F. P. Kelly, *Reversibility and Stochastic Networks* (Wiley, Chichester, 1979).
14. R. G. Miller, Jr., Stationarity equation in continuous time Markov chains, *Trans. Am. Math. Soc.* **109**:35–44 (1963).
15. R. L. Tweedie, Some ergodic properties of the Feller minimal process, *Q. J. Math. Oxford (2)* **25**:485–495 (1974).
16. R. L. Tweedie, Sufficient conditions for regularity, recurrence and ergodicity of Markov processes, *Math. Proc. Camb. Philos. Soc.* **78**:125–136 (1975).
17. R. L. Tweedie, Criteria for classifying general Markov chains, *Adv. Appl. Prob.* **8**:737–771 (1976).
18. J. D. Weeks and G. H. Gilmer, Dynamics of crystal growth, *Adv. Chem. Phys.* **40**:157–228 (1979).
19. P. Whittle, Reversibility and acyclicity, in *Perspectives in Probability and Statistics*, J. Gani, ed. (Academic Press, London, 1975).
20. P. Whittle, *Systems in Stochastic Equilibrium* (Wiley, New York, 1986).