I.3
Combinatorial methods

We shall now consider non-ergodic operators; in fact we construct operators that have more than one invariant measure. Our proofs (except for those of Chapter 13) follow the same general lines that have been developed for Ising-like models of statistical physics and random fields. The key part of such a proof is a combinatorial estimation of the number of graphs of a particular sort. In Chapter 8 this is done for the Stavskaya system which was described in Chapter 1 as Example 1.2 and for similar systems. Chapter 9 is about deterministic operators. It investigates their property of being 'eroders' which is relevant to ergodicity of those stochastic operators which are compositions of these deterministic operators and the standard independent homogeneous random noise, and this relevance is the theme of Chapter 10. In Chapter 11 we prove non-ergodicity of random walk operators (introduced in Chapter 6) with random noise added. Chapter 12 presents two counter-examples which show impossibility of some generalisations. Chapter 13 is a special one. It is an introduction to the very interesting but difficult problem of whether there are one-dimensional independent local homogeneous non-degenerate non-ergodic operators.

In most cases $X = \{0;1\}^V$. Often we concentrate on the case when the measure δ_1 concentrated in the state 'all ones' is invariant for an operator P. In this case we consider the family of operators $P_\varepsilon = PS_\varepsilon$ with one parameter $\varepsilon \in [0,1]$ where S_ε is the random noise defined in Chapter 2 which turns any 0 into 1 with probability ε independently from others. Of course, δ_1 is invariant for P_ε with all ε. Thus, P_ε is ergodic if it contracts X to δ_1 in the limit $t \to \infty$.

Of course, P_ε does it with $\varepsilon = 1$, and most quickly. If P is homogeneous and independent, P_ε is ergodic with all $\varepsilon > 1 - \frac{1}{R}$ where $R = |U(h)|$. This just follows from Proposition 2.17 and will be proved here in another way.

We shall construct such P for which P_ε are non-ergodic with some $\varepsilon > 0$. For every such P there exists some (at least one) *critical value* of ε, which separates those values of ε which make P_ε ergodic from those values of ε which make P_ε non-ergodic. Example 11.1 presents a non-monotone P having no less than two critical values of ε.

For a monotone P it is easy to prove uniqueness of the critical value. By definition, this is such value ε^* of ε that operator P_ε is ergodic with all ε, $\varepsilon^* < \varepsilon \leqslant 1$ and non-ergodic with all ε, $0 \leqslant \varepsilon < \varepsilon^*$. If P is ergodic with all $\varepsilon > 0$, we define $\varepsilon^* = 0$.

Obviously, $\varepsilon^* = 0$ if V is finite. In fact, in this case P_ε is an operator of a finite Markov chain which is ergodic with $\varepsilon > 0$ and has the absorbing state 'all ones'.

Chapter 8
Percolation operators

Let us have a graph $\Gamma(V, \mathcal{U})$. Then D_0 stands for the deterministic operator on $X = \{0;1\}^V$ defined by the condition

$$(xD_0)_h = \begin{cases} 1 & \text{if} \quad x_{U(h)} = 1, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, the hth component of xD_0 is 1 if and only if all the components of x in the neighbourhood $U(h)$ are ones. Of course, on any $\Gamma(V, \mathcal{U})$ there is just one D_0. Note that $D_0 \prec P$ for any $P \in \mathcal{P}_1$ on the same $\Gamma(V, \mathcal{U})$.

The operator D_0 is always monotone, hence it always has just one critical value which we denote $\varepsilon_\Gamma^* \in [0;1]$. By definition of ε_Γ^*, the composition $D_0 S_\varepsilon$ is ergodic with all $\varepsilon > \varepsilon_\Gamma^*$ and non-ergodic with all $\varepsilon < \varepsilon_\Gamma^*$.

This chapter is about compositions $D_0 S_\varepsilon$ which we term 'percolation operators'. Remember that the evolution measure $\tilde{\mu}$ on the evolution space (defined in Chapter 2) for the special case of $D_0 S_\varepsilon$ is representable as induced by the Bernoulli measure on the auxiliary space Ω with the mapping defined by the following formulae:

$$x_h^{t+1} = \max\{\omega_h^t, \min_{k \in U(h)} \{x_k^t\}\}, \; h \in V, \; t \in \mathbb{Z}_+.$$

This mapping defines any x_H^T, $H \in V$, $T \in \mathbb{Z}_+$ as a function of a finite set of arguments ω_h^t, x_h^0. The following proposition is about this function.

Proposition 8.1 For percolation operators $x_H^T = 0$ as a function of ω_h^t, x_h^0, $h \in V$, $t \in \mathbb{Z}_+$ if and only if there is such a sequence $h_0, h_1, \ldots, h_T$ of elements of V for which the following three conditions hold:

(a) $h_T = H$ and $h_{t-1} \in U(h_t)$ for all t from 1 to T.
(b) $\omega^t_{h_t} = 0$ for all t from 1 to T.
(c) $x^0_{h_0} = 0$.

Proof can be easily performed by induction. Instead of doing this, let us describe a 'physical' interpretation of this proposition which warrants our use of the term 'percolation'. Imagine that all bonds of the evolution graph (defined in Chapter 2) are semi-conducting pipes which can transmit some fluid to any point $(h,t+1)$ from the points (k,t), $k \in U(h)$. Any point (h,t) is free for the fluid to pass it if $\omega^t_h = 0$ and is corked with a stopper if $\omega^t_h = 1$. Initially the fluid comes to those points $(h,0)$ where $x^0_h = 0$. Then the value of x^T_H (as a function of arguments x^0_h and ω^t_h) equals zero if and only if the fluid can percolate to the point (H,T) and through it.

Percolation problems have been treated in many works ([17,41,54] in particular) and have led to the development of methods such as those to be used in our proof of Theorem 8.4. But let us first prove two simple Propositions 8.2 and 8.3, the former of which is a special case of Proposition 2.17.

Proposition 8.2 For any homogeneous graph $\Gamma(V, \mathscr{U})$ its critical value $\varepsilon^*_\Gamma \leqslant 1 - \frac{1}{R}$ where $R = |U(h)|$. In other words, any percolation operator $D_0 S_\varepsilon$ is ergodic with $\varepsilon > 1 - \frac{1}{R}$.

Proof. Let $\varepsilon > 1 - \frac{1}{R}$. It is sufficient to prove that

$$\lim_{T \to \infty} \mathbf{0} P^T_\varepsilon (x_0 = 0) = 0.$$

The sub-limit expression is the probability that the fluid percolates to the vertex $(0,T)$ of the evolution graph where 0 stands for an arbitrary element of V.

There are R^T oriented paths from $t = 0$ to the point $(0,T)$ for the fluid to percolate. The probability that a given path is free for flow is $(1 - \varepsilon)^T$. Then the probability that at least one path is free does not exceed $R^T(1 - \varepsilon)^T$. If $\varepsilon > 1 - \frac{1}{R}$ then this value tends to 0 with $T \to \infty$. q.e.d.

Using monotonicity considerations, one can infer from it the ergodicity of PS_ε, with $\varepsilon > 1 - \frac{1}{R}$ for any $P \in \mathscr{P}_1$.

Note that this upper estimation for ε_Γ^* cannot in general be improved because $\varepsilon_\Gamma^* = 1 - \dfrac{1}{R}$ for the tree-like graphs (see Chapter 2).

Proposition 8.3 Let two graphs Γ_1 and Γ_2 have the common set of vertices V and their neighbourhood systems be $\mathscr{U}_1$ and $\mathscr{U}_2$. Let

$$\forall h \in V\colon U_1(h) \subseteq U_2(h).$$

Then

$$\varepsilon_{\Gamma_1}^* \leqslant \varepsilon_{\Gamma_2}^*.$$

Proof. Let $D_0^1 S_\varepsilon$ and $D_0^2 S_\varepsilon$ stand for percolation operators on graphs Γ_1 and Γ_2. From percolation considerations

$$\mathbf{0}(D_0^1 S_\varepsilon)^t > \mathbf{0}(D_0^2 S_\varepsilon)^t,$$

which proves the proposition.

The following is the main result of the present chapter.

Theorem 8.4 [71,85]. Let $\Gamma(V, \mathscr{U})$ be a homogeneous graph where $V = G = \mathbb{Z}^d$, the neighbourhood system $\mathscr{U}$ being translation invariant, and $|U(h)| > 1$. Then $\varepsilon_\Gamma^* > 0$. In other words, percolation operators are non-ergodic on such a graph with ε small enough.

Proof. Proposition 8.3 provides

$$\varepsilon_\Gamma^* \geqslant \varepsilon_{\Gamma_1}^*$$

where Γ_1 is the simplest one-dimensional graph with $V = \mathbb{Z}$ and $U(h) = \{h-1, h\}$. In fact, eliminating from Γ some bonds in such a way as to conserve homogeneity we can ensure that $|U(h)| = 2$ and the value ε_Γ^* will not increase with this. Now, in the case $|U(h)| = 2$, all the connected components of the evolution graph are isomorphic to the evolution graph of Γ_1 which makes $\varepsilon_\Gamma^* = \varepsilon_{\Gamma_1}^*$. Thus it is sufficient to prove our theorem for the special case of Γ_1.

Two quite different proofs have been proposed for this case. One of them, described in [71], is cumbersome and we shall not describe it. The other [85] is based on a combinatorial estimation. It is simple enough and gives a better estimation for $\varepsilon_{\Gamma_1}^*$ (although it is very far from precision). Let us describe the latter proof.

Due to Proposition 8.1 we have to estimate the probability that the fluid cannot percolate to the point $(0, T)$. We are going to show that to prevent the fluid from reaching $(0, T)$, the stoppers must form at least one 'fence' in a certain sense (as we have mentioned in describing Example 1.2). It is most convenient to fix the ideas in terms of dual graphs.

First let us describe the necessary notions in general. A planar graph is a graph placed in a certain way in a plane without intersections. Let Δ stand for a planar graph with two marked vertices A and B. We say that a state of the graph Δ is given, if it has been said about every KL bond of it whether it passes (the fluid) from K to L and whether it passes from L to K. Let $\bar{\Delta}$ stand for the dual planar graph of Δ. (This means that vertices of $\bar{\Delta}$ are the domains into which Δ cuts the plane and vice versa; bonds of Δ and $\bar{\Delta}$ cross each other in the one-one correspondence.) For any state of the graph Δ we define the corresponding state of $\bar{\Delta}$ by the following rule. Let the bond MN of $\bar{\Delta}$ cross the bond KL of Δ and M be on the left side (N being on the right side) of KL when moving along KL from K to L. Then MN passes from M to N if and only if KL does not pass from K to L.

The following is true for all planar graphs: the graph Δ does not pass (allow the fluid to percolate) from A to B if and only if in the graph $\bar{\Delta}$ there is a circular oriented open (for the flow of the fluid) path separating domain A from domain B and oriented clockwise around A. This can be proved by induction over the number of bonds.

Now let us apply this general statement to our case. Luckily, the evolution graph of Γ_1 is planar – it is a square lattice. We need only a finite triangular part of it, namely the part which the fluid can use to reach $(0,T)$. (Fig. 8.1. shows this part for $T = 4$; to make the picture symmetric we inclined the lines h = const.) Now we substitute every vertex (i,t) by a vertical bond and keep the same notation (i,t) for it. Thus we have turned our triangular piece of square lattice into another planar graph which we denote Δ_T. It is a homogeneous lattice with hexagonal cells shown in Figure 8.1 by unbroken lines. Let us merge all the lower ends of bonds $(i,0)$ into one vertex A which is one pole of our graph. The upper end of the bond $(0,T)$ will be the opposite pole B. As before, the inclined bonds of this hexagonal lattice always pass the fluid upward and never downward. Any vertical bond (i,t) passes upward if and only if $\omega_i^t = 0$. As a result of all this, our graph Δ_T passes the fluid from A to B just in those cases in which $x_0^T = 0$.

For technical reasons we assume that vertical bonds always pass (allow to flow) downward; in fact, this changes nothing because inclined bonds never pass downward and prevent the fluid from moving downward.

Thus we have the planar graph Δ_T and rules for its bonds to pass. This defines uniquely the dual graph $\bar{\Delta}_T$ and its bonds' passing behaviour. After minor amendments, $\bar{\Delta}_T$ becomes a piece of homogeneous lattice with triangular cells. It is shown by dotted lines in Figure 8.1. According to the general rule, the former graph Δ_T does not pass from A to B if and only if the amended $\bar{\Delta}_T$ passes from EE to FF. (Essentially the way from EE to FF is circular because EE and FF are in one domain and ought to be identified.)

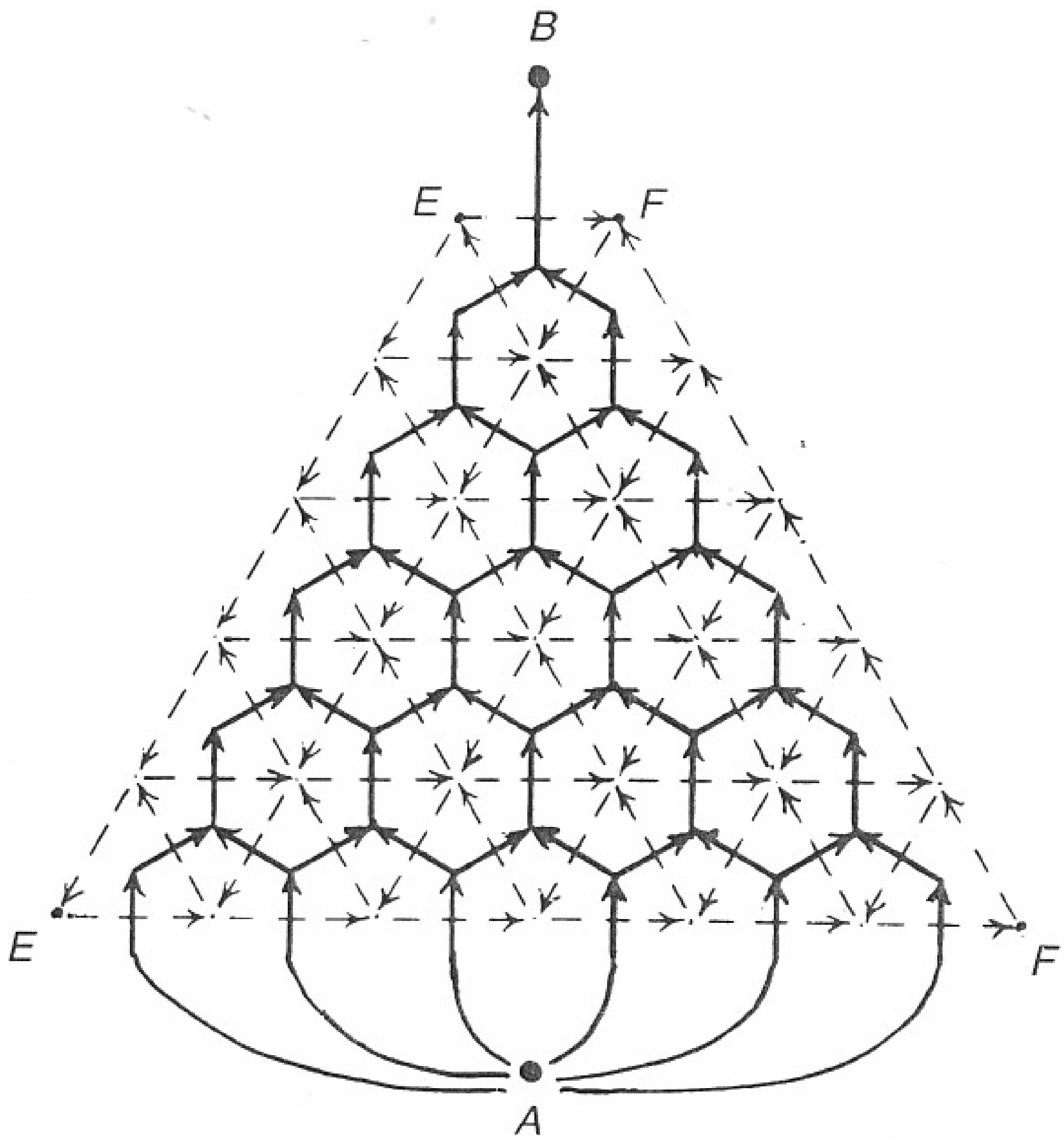

Fig. 8.1 Percolation graph and its dual.

Thus our aim, expressed in the new terms, is to estimate the probability that $\bar{\Delta}_T$ does pass from *EE* to *FF*. The main point is that this probability does not exceed the sum of the probabilities of all the events 'the way passes from *EE* to *FF*'. Let us estimate this sum.

The inclined bonds of $\bar{\Delta}_T$ are dual to the inclined bonds of Δ_T. Hence they always pass leftward and never pass rightward. The horizontal bonds of $\bar{\Delta}_T$ are dual to the vertical bonds of Δ_T. Hence they pass rightward with a probability ε independently from each other. As we have ensured that the vertical bonds of Δ_T always pass downward, the horizontal bonds of $\bar{\Delta}_T$ never pass leftward.

Thus any path in $\bar{\Delta}_T$ from *EE* to *FF* that can ever pass, consists of just three types of steps, shown by arrows in Figure 8.1. These steps are shifts of three vectors, the sum of which is 0. We may consider only those paths which never pass the same point twice. For these paths the probability of being pass-free is ε^k where k is the number of horizontal bonds in the path. Thus

$$\tilde{\mu}(x_0^T = 1) \leqslant \sum_{k=1}^{\infty} N_k \varepsilon^k, \tag{8.1}$$

where N_k is the number of different paths from EE to FF of the kind we consider having k horizontal bonds. It remains to estimate N_k. It is convenient to continue every path beyond its ends along EE and FF as far as the point where EE meets FF. Then every path will have equal numbers of steps of all the three directions. Thus it has $3k$ steps and moving along it we have no more than three choices at every point. Hence $N_k \leqslant 3^{3k}$. Thus $\varepsilon < \frac{1}{54}$ makes the sum (8.1) less than 1. q.e.d. Theorem 8.4 is proved.

Thus $\varepsilon^*_{\Gamma_1} \geqslant \frac{1}{54}$. Let us discuss how to get a better estimation.

Proposition 8.5 [85]. The value of $\varepsilon^*_{\Gamma_1}$ is no less than the radius of convergence of the series (8.1).

Proof. Proof is based on the following generalisation of (8.1):

$$\tilde{\mu}(x_1^T = x_2^T = \ldots = x_m^T = 1) \leqslant \sum_{k=1}^{m} N_k \varepsilon^k. \tag{8.2}$$

First let us prove (8.2). That part of the supergraph (evolution graph) which the fluid uses to reach our points

$$(1,T), \ldots, (m,T)$$

is an equilateral trapezoid. Hence we turn all vertices (i,t) into vertical bonds (i,t) too and introduce pole A as before. But now we identify all the upper ends of the vertical bonds $(1,T), \ldots, (m,T)$ into the other pole B. The resulting graph does not pass the fluid from A to B if and only if $x_1^T = \ldots = x_m^T = 1$. So the dual graph does pass in the same case. This dual graph is also a piece of triangular lattice, but truncated. The shortest relevant path in it contains m horizontal bonds. Thus (8.2) holds.

Now let us prove our Proposition 8.5. Let ε be less than the radius of convergence of the series (8.1). Then (8.1) converges. Then we can choose such a large m that the sum on the right-hand side of (8.2) will be less than 1. Since this estimation does not depend on T, we can apply Corollary 2.8, where C_1 is the totality of those $\mu \in \mathcal{M}(X)$ in which the probability of $x_1^T = \ldots = x_m^T = 1$ does not exceed the right-hand side of (8.2), to prove existence of invariant $\mu \in C_1$ which differs from $\boldsymbol{\delta_1}$.

Proposition 8.5 is proved. This yields $\varepsilon^*_{\Gamma_1} \geqslant \frac{1}{27}$ at once. But we can do better. Let us code paths from EE to FF by sequences of signs 1, 2, 3 where 1 denotes a left-down-directed bond, 2 denotes a right-directed bond, 3 denotes a left-up-directed bond. Now we take into consideration the fact that codes with 13 or 31 entries need not be considered (because in those cases the path can be shortened without including any new horizontal bond). This leads to a better estimation:

$$\varepsilon^*_{\Gamma_1} > 0.09.$$

Let us explain the estimation in detail to show our method which will be used in a more complicated form to prove Theorem 11.1 below.

We introduce a coordinate system (i,t) in Figure 8.1. Let the left-upper corner of the trapezoid be the origin 0. The i-axis goes to the right, the t-axis goes up. We choose the scale to make vectors corresponding to the three kinds of steps in our paths have coordinates $(-1,-1)$ for type 1, $(2,0)$ for type 2, $(-1,1)$ for type 3. Let a 'correct path' be a path starting at the origin, passing several bonds in the direction of the arrows, loopless and avoiding 13 and 31 entries. (Note that we have not yet specified the end location.)

Every correct path has a 'weight' ε^k where k is the number of type-2 bonds in this path. We have $\sigma^q_n(i,t)$ standing for the sum of weights of all the correct paths with n bonds which end in the point (i,t) and have the last bond of the type $q \in \{1,2,3\}$. Of course

$$\tilde{\mu}(x^T_1 = \ldots = x^T_m = 1) \leqslant \sum_{n=1}^{\infty} \sum_{q=1}^{3} \sigma^q_n(2m,0). \qquad (8.3)$$

The sums $\sigma^q_n(i,t)$ satisfy the following recurrent inequalities:

$$\begin{aligned} \sigma^1_{n+1}(i,t) &\leqslant \sigma^1_n(i+1,t+1) + \sigma^2_n(i+1,t+1), \\ \sigma^2_{n+1}(i,t) &\leqslant \varepsilon[\sigma^1_n(i-2,t) + \sigma^2_n(i-2,t) + \sigma^3_n(i-2,t)], \qquad (8.4) \\ \sigma^3_{n+1}(i,t) &\leqslant \sigma^2_n(i+1,t-1) + \sigma^3_n(i+1,t-1). \end{aligned}$$

Now we introduce some new sums:

$$\Sigma^q_n(\alpha,\beta) = \sum_{i=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} \alpha^i \beta^t \sigma^q_n(i,t),$$

where α and β are positive parameters. Equation (8.3) implies

$$\tilde{\mu}(x^T_1 = \ldots = x^T_m = 1) \leqslant \alpha^{-2m} \sum_{n=1}^{\infty} \sum_{q=1}^{3} \Sigma^q_n(\alpha,\beta), \qquad (8.5)$$

for all $\alpha > 0$, $\beta > 0$.

The relations (8.4) provide relations for the new values:

$$\begin{aligned} \Sigma^1_{n+1}(\alpha,\beta) &\leqslant \alpha^{-1}\beta^{-1}[\Sigma^1_n(\alpha,\beta) + \Sigma^2_n(\alpha,\beta)], \\ \Sigma^2_{n+1}(\alpha,\beta) &\leqslant \varepsilon\alpha^2[\Sigma^1_n(\alpha,\beta) + \Sigma^2_n(\alpha,\beta) + \Sigma^3_n(\alpha,\beta)], \\ \Sigma^3_{n+1}(\alpha,\beta) &\leqslant \alpha^{-1}\beta[\Sigma^2_n(\alpha,\beta) + \Sigma^3_n(\sigma,\beta)]. \end{aligned}$$

Let M stand for the matrix of coefficients of the right-hand sides:

$$M = \begin{pmatrix} \alpha^{-1}\beta^{-1} & \alpha^{-1}\beta^{-1} & 0 \\ \varepsilon\alpha^2 & \varepsilon\alpha^2 & \varepsilon\alpha^2 \\ 0 & \alpha^{-1}\beta & \alpha^{-1}\beta \end{pmatrix}.$$

We want to make all the eigenvalues of M less than 1 in modulo because this makes the series on the right-hand side of (8.5) convergent. If $\alpha > 1$ too, there exists such m that the right-hand side of (8.5) is less than 1, which allows us to refer to Corollary 2.8 to prove non-ergodicity of P.

According to a well-known criterion (see [26]) all the eigenvalues of M are less than 1 in modulo if all the main corner minors of the matrix $E - M$ are positive. This condition leads to the following system (in transformations it is convenient to transpose the second line and column with the third ones):

$$\begin{cases} \alpha > 1, \\ \alpha^{-1} < \beta < \alpha, \\ \varepsilon < \dfrac{(1 - \alpha^{-1}\beta^{-1})(1 - \alpha^{-1}\beta)}{\alpha^2 - 1}. \end{cases}$$

To obtain from this the best possible estimation for ε^*, we should choose α and β so as to maximise the right-hand side of the last inequality, namely $\alpha = \frac{1}{2}(1 + \sqrt{2}) > 1$ and $\beta = 1$. They make the right-hand side in question equal to

$$\tfrac{1}{2}(5\sqrt{5} - 11) \approx 0.09017.$$

Thus

$$\varepsilon^*_{\Gamma_1} > 0.09$$

Analogously one can take into consideration the impossibility of some longer entries in the codes of the paths. For example, we may ignore all codes where 123, 321 and so on occur. But such computations, even in the limit, certainly will not lead us to the exact computation of $\varepsilon^*_{\Gamma_1}$. The point is that various events of the type 'this path from EE to FF is free to pass' are compatible. One can show that the probability of the dual graph passing from EE to FF equals the sum of probabilities of events of the type 'this path passes from EE to FF, and there is no free path which is at least partially above it'. We cannot see a way of estimating the sum of probabilities of these events which would provide lower estimations for $\varepsilon^*_{\Gamma_1}$ tending to $\varepsilon^*_{\Gamma_1}$ in the limit.

To end this chapter let us survey some other results about percolation operators.

Let $\tilde{\mu}$ be the evolution measure on $\mathbb{Z} \times \mathbb{Z}_+$ produced by the percolation operator $P_\varepsilon = D_0 S_\varepsilon$ on the standard graph Γ_1 and such initial measure μ that $\mu(\mathbf{1}) = 0$ (where $\mathbf{1}$ is the single state 'all ones'). The percolation interpretation of $\tilde{\mu}$ (our Proposition 8.1) allows an upper estimation of probability $\tilde{\mu}$ of a two-dimensional 'island' of ones surrounded by zeros: it does not exceed $Ae^{-B\varepsilon L}$ where A and B are positive constants and L is the length of the boundary. This estimation is used in [100] to work out a system of correlation equations for the probabilities of islands which is similar to systems of correlation equations in models of lattice gas [49,73] and to prove uniqueness of its solution for small values of ε (say, $\varepsilon < 0.001$). Hence P_ε has just one invariant measure with $\mu(\mathbf{1}) = 0$ for small ε and this unique measure μ_ε depends analytically on ε. (Uniquenes of ε_ε among homogeneous measures is proved for all $\varepsilon < \varepsilon^*_{\Gamma_1}$ by another method; see [47] and our Proposition 4.9.)

Symmetry of percolation allows us to prove for small values of ε that any homogeneous initial measure with $\mu(\mathbf{1}) = 0$ produces μ_ε in the limit:

$$\lim_{t \to \infty} \mu P^t_\varepsilon = \lim_{t \to \infty} \mathbf{0} P^t_\varepsilon = \mu_\varepsilon$$

and in general this symmetry allows us to describe $\lim xP^t_\varepsilon$ for all initial states x [98].

Let Γ_d stand for the d-dimensional graph with $V = \mathbb{Z}^d$ and

$$U(h) = \{h, h-e_1, \ldots, h-e_d\},$$

where $e_1, \ldots, e_d$ are unitary coordinate vectors. Due to Proposition 8.3, the sequence $\varepsilon^*_{\Gamma_d}$ is non-decreasing. In [53] it is proved that this sequence tends to 1 in the following way:

$$\varepsilon^*_{\Gamma_d} = 1 - \frac{\alpha}{d} + o\left(\frac{1}{d}\right),$$

where $\alpha \geqslant 1$ is a constant; but we do not know the value of α.

It seems in general that there are many graphs for which $\varepsilon^*_\Gamma > 0$, that is, on which the percolation operators are non-ergodic for small values of ε. Perhaps this is true for all infinite connected (in the sense defined in Chapter 2) homogeneous graphs with $|U(h)| > 1$. Now $\varepsilon^*_\Gamma > 0$ is proved in the following cases:

(a) If $h \in U(h)$ for all $h \in V$. See proof in [55, pp. 28–9]. This is based on $\varepsilon^*_{\Gamma_1} > 0$.

(b) If Γ has a commutative transitive group of automorphisms. This follows easily from (a).

(c) If for any $h \in V$ there are such $\{h_1, h_1\} \subset U^\infty(h)$ that $U^\infty(h_1) \cap U^\infty(h_2) = \emptyset$. For example, this is true for a tree-like graph. See proof in [55, pp. 26–8].

As has been said in Chapter 1, the non-ergodicity of percolation operators on Γ_1 with small values of ε was first discovered in [78] by computer simulation. In fact, a finite system was simulated with various boundary conditions. In a series of computer experiments this finite system was a circular chain of automata: the right-hand end automaton was made the left-hand neighbour of the left-hand end one. Applied to these finite systems, our combinatorial method shows that they behave for small ε in a special way too. To give exact formulations let us denote $V = \{1,\ldots,m\}$ and

$$U(h) = \begin{cases} \{h-1;h\} & \text{for} \quad 2 \leqslant h \leqslant m, \\ \{m;1\} & \text{for} \quad h = 1. \end{cases}$$

For finite V, all percolation operators on this graph are ergodic for all m, $\varepsilon > 0$. Thus, to discern differences in their behaviour we have to introduce a parameter relevant to the speed of their inevitable convergence to **1**. Let $\tau_\varepsilon(m)$ stand for the mean time t at which the system reaches the state

$$x_1^t = \ldots = x_m^t = 1$$

if it has started from the state

$$x_1^0 = \ldots = x_m^0 = 0.$$

Then:

(a) If $\varepsilon > \frac{1}{2}$ then

$$\tau_\varepsilon(m) \leqslant \log_{\frac{1}{2(1-\varepsilon)}} m + \frac{2\varepsilon}{2\varepsilon - 1}.$$

(b) If $\varepsilon < \frac{1}{27}$ then

$$\tau_\varepsilon(m) \geqslant \frac{1 - 27\varepsilon}{(3\varepsilon)^m}.$$

The estimation (a) was first proved in [78]. It is deducible from our proof of Proposition 2.17.

To prove the estimation (b) (see [85]) note that the evolution graph in this case is planar too. You may imagine it drawn on a cylinder. Thus the job boils down to estimation of the number of paths around this cylinder.

The comparision of estimations (a) and (b) shows that in the limit $m \to \infty$ the finite systems behave in qualitatively different ways for different ε. So let $\varepsilon_{\exp}$ stand for the supremum of those ε for which $\tau_\varepsilon(m)$ grows exponentially with $m \to \infty$. Let $\varepsilon_{\log}$ stand for the infimum of those ε for which $\tau_\varepsilon(m)$ grows logarithmically. Clearly $\varepsilon_{\exp} \leqslant \varepsilon_{\log}$. From estimations (a) and (b),

$$0 < \varepsilon_{\exp} \leqslant \varepsilon_{\log} < 1.$$

In the computer simulation we assumed similarity between the finite and infinite cases, and this assumption is very common. But to prove is better than to assume. The following proposition is a step in this direction.

Proposition 8.6 (L. G. Mityushin).

(a) $\varepsilon^*_{\Gamma_1} \leqslant \varepsilon_{\text{lin}}$ where ε_{lin} is the infimum of those ε with which $\tau_\varepsilon(m) = o(m)$.

(b) $\varepsilon_{\log} \leqslant \varepsilon^*_{\text{walk}}$ where $\varepsilon^*_{\text{walk}}$ has been defined in Chapter 6.

Proof. Let us prove (a). Let $\tilde{\mu}_m$ and $\tilde{\mu}_\infty$ stand for the evolution measure in the finite and infinite cases. Note that for $t \leqslant m$

$$\tilde{\mu}_m(x^t_h = 0) = \tilde{\mu}_\infty(x^t_h = 0).$$

Hence

$$\begin{aligned}\tau_\varepsilon(m) &= \sum_{t=0}^{\infty} \tilde{\mu}_m\{x: x^t \neq \mathbf{1}\} \geqslant \\ &\geqslant \sum_{t=0}^{\infty} \tilde{\mu}_m\{x: x^t_h = 0\} \geqslant \\ &\geqslant m\tilde{\mu}_m\{x: x^m_h = 0\} = m\tilde{\mu}_\infty\{x: x^m_h = 0\}.\end{aligned}$$

Multiplying this by $1/m$ and proceeding to the limit $m \to \infty$, we obtain the required estimation.

Let us prove (b). Let $\varepsilon > \varepsilon^*_{\text{walk}}$. Then (see Note 6.7),

$$\tilde{\mu}_\infty\{x: x^t_h = 0\} = o(\mathrm{e}^{-\text{const}\cdot t}).$$

On the other side

$$\begin{aligned}\tau_\varepsilon(m) &= \sum_{t=0}^{\infty} \tilde{\mu}_m\{x: x^t \neq \mathbf{1}\} = \\ &= \sum_{t=0}^{\infty} \tilde{\mu}_m\left\{\bigcup_{h\in V} \{x^t_h = 0\}\right\} \leqslant \\ &\leqslant \sum_{t=0}^{\infty} \min\left(1, m\tilde{\mu}_m\{x: x^t_h = 0\}\right).\end{aligned}$$

From the results of the paper [54],

$$\tilde{\mu}_m\{x: x^t_h = 0\} \leqslant \tilde{\mu}_\infty\{x: x^t_h = 0\}$$

for all m, t. Thus

$$\tau_\varepsilon(m) \leqslant \sum_{t=0}^{\infty} \min\left(1, 0(m\mathrm{e}^{-\text{const}\cdot t})\right) = o(\ln t)$$

q.e.d.

Finally, let us resume all the inequalities between the various introduced critical values of ε, all of which refer to percolation operators on Γ_1 and its finite analogues (see end of Chapter 6):

$$
\begin{array}{l}
0 < \varepsilon^* \left\{ \begin{array}{l} \leqslant \varepsilon^*_{\mathrm{syst}} < 1 \\ \leqslant \end{array} \right. \\
\qquad\qquad\qquad\qquad \left. \begin{array}{r} \\ \end{array} \right\} \varepsilon_{\mathrm{lin}} \leqslant \varepsilon_{\mathrm{log}} \leqslant \varepsilon^*_{\mathrm{walk}} < 1. \\
0 < \varepsilon_{\mathrm{exp}} \leqslant
\end{array}
$$

However, it seems most plausible to us that all these critical value equal. It remains to prove it.

Chapter 9
Eroders

In Chapters 3–5 we proved that operators having approximately equal transition probabilities θ_z^y are ergodic. As we look for non-ergodic operators here, we should do it among those having sharply different θ_z^y. Best of all would be for some of these parameters to be near 0 the others being near 1. In this case, it pays to treat a stochastic operator as a perturbed deterministic one, as we shall do. So we have something to say about deterministic operators first. The letter D will now stand for a deterministic operator on X.

Definition 9.1 Let $x = \{0;1\}^V$. In this case $I(x)$ stands for the set of those $h \in V$ where $x_h = 1$:

$$I(x) = \{h \in V: x_h = 1\}.$$

A state $x \in X$ is termed an 'island' if $I(x)$ is finite. A deterministic operator $D: X \to X$ is said to 'erode' a state $x \in X$ if after a finite number (t) of iterations the state x is turned into the state 'all zeros', that is, if

$$\exists t: I(xD^t) = \varnothing.$$

A deterministic operator $D: X \to X$ is an eroder if it erodes all islands in X.

For example, the percolation operators D_0 introduced in Chapter 8 certainly are eroders in the case most interesting for us: when Γ is infinite, connected, homogeneous, and $|U(h)| > 1$. In particular, the deterministic part of the Example 1.2 operator is an eroder, the deterministic part of the Example 1.1b operator is an eroder, but that of the Example 1.1c operator is not.

We are interested in these chapters not just in D's being eroder or non-eroder as such but rather in the behaviour of random perturbations of D. Often we shall consider the behaviour of DS_ε where S_ε is the random noise introduced in Chapter 2. (Remember that S_ε changes all zeros into ones with probability ε independently from each other, and leaves ones unchanged.)

Proposition 9.2 [88]. Let $D: X \to X$ be homogeneous, monotone (in the sense of Chapter 2) and a non-eroder. Then DS_ε are ergodic and have $\delta_\mathbf{1}$ as the invariant measure with all $\varepsilon > 0$.

Proof. Let $\tilde{\mu}$ be the evolution measure produced by DS_ε and the initial measure $\delta_\mathbf{0}$ concentrated in the state $\mathbf{0}$ 'all zeros'. The value of x_0^T (where 0 stands for an arbitrary element of V) is a deterministic function of auxiliary variables ω_h^t, $h \in V$, $0 \leqslant t \leqslant T$ (which were introduced in Chapter 2). Since D is a non-eroder, there is an island y, which D does not erode. This implies that for any natural t there is such $k_t \in V$ that

$$(yD^t)_{k_t} = 1.$$

Let $g_t \in G$ be such an automorphism of V that $g_t(k_t) = 0$.

The value of x_0^T is certainly 1 if there is such a t, $0 < t \leqslant T$, that $\omega_h^t = 1$ for all

$$h \in g_{T-t}(I(y)). \tag{9.1}$$

The probability of (9.1) being true for a given t is $\varepsilon^{|I(y)|}$. As the events (9.1) with different values of t are independent, the probability of (9.1) being true for at least one t, $0 < t \leqslant T$ is

$$1 - (1 - \varepsilon^{|I(y)|})^T.$$

This value tends to 1 with $T \to \infty$. The value of $\tilde{\mu}(x_0^T = 1)$ is no less than this value; hence it tends to 1 too. Thus

$$\lim_{T\to\infty} \mathbf{0}(DS_\varepsilon)^T = \delta_\mathbf{1},$$

q.e.d.

The question of ergodicity of DS_ε with D eroders is more complicated. Some of them have DS_ε ergodic with all $\varepsilon > 0$, some have not. Our main theorem proves that some eroders have DS_ε non-ergodic with $\varepsilon > 0$ small enough. Before formulating it, we have to classify homogeneous monotone eroders on $\{0;1\}^{\mathbb{Z}^d}$, and before that we need Theorem 9.4 and Proposition 9.5. Theorem 9.4 allows us to know whether any given homogeneous monotone D on $\{0;1\}^{\mathbb{Z}^d}$ is an eroder or not.

Remember that any D is defined by functions $f_h: X_{U(h)} \to X_h$. In the homogeneous case all these functions are one and the same, provided we have enumerated their arguments in the coherent way. Let us use the following notation:

$$\forall h \in \mathbb{Z}^d\colon\ U(h) = \{h+u_1,\ldots,h+u_R\} = h + U(0)$$

where $u_1, \ldots, u_R$ are fixed in $\mathbb{Z}^d$ and 0 is the origin in $\mathbb{Z}^d$.

$$(xD)_h = f(x_{h+u_1},\ldots,x_{h+u_R})$$

where the function f is one and the same for all $h \in \mathbb{Z}^d$. We have that D is monotone if and only if f is monotone, that is,

$$a_1 \leqslant b_1, \ldots\ a_R \leqslant b_R \Rightarrow f(a_1,\ldots,a_R) \leqslant f(b_1,\ldots,b_R).$$

Definition 9.3 Let us have a monotone homogeneous $D\colon X \to X$ where $X = \{0;1\}^V$ and $V = \mathbb{Z}^d$.

(a) We-term a subset $K \subset U(0)$ a zero-set if the condition '$x_h = 0$ for all $h \in K$' guarantees

$$(xD)_0 = 0.$$

We assume that function f does not constantly equal 1, whence at least one zero-set exists.

(b) We think of $\mathbb{Z}^d$ as a subset of the real space $\mathbb{R}^d$. Let $\sigma_D \subset \mathbb{R}^d$ stand for the intersection of convex hulls (in $\mathbb{R}^d$) of all zero-sets of D.

Theorem 9.4 [87]. A monotone homogeneous $D\colon X \to X$ where $X = \{0;1\}^{\mathbb{Z}^d}$ is an eroder if and only if σ_D is empty.

Note. Since $U(0)$ is finite, it has a finite number of subsets. Thus, for any given D with not too many elements in $U(0)$ it is quite possible to actually construct σ_D and check if it is empty or not. You can do it for the deterministic parts of Examples 1.1, 1.2 and 1.5 and see the theorem's correctness in those cases. To prove the theorem in general we need the following.

Proposition 9.5 [87]. The set σ_D is empty if and only if there are $m + 1$ such non-homogeneous linear functionals $L_0, \ldots, L_m$ on $\mathbb{Z}^d$ where $1 \leqslant m \leqslant d$, that

$$\sum_{k=0}^{m} L_k \equiv -\ell \tag{9.3}$$

where ℓ is some positive number, and

$$\max_{h \in I(xD)} L_k(h) \leqslant \max_{h \in I(x)} L_k(h) + L_k(0) \tag{9.4}$$

for all $k = 0, \ldots, m$ and all islands x, provided $I(xD) \neq 0$.

This proposition enables us to formulate the following:

Definition 9.6 Every monotone eroder D on $x = \{0;1\}^{\mathbb{Z}^d}$ has a *range* which is defined as the minimal value of m for which there are $m + 1$ functionals $L_0, \ldots, L_m$ with the conditions (9.3) and (9.4).

The two following examples illustrate the geometrical interpretation of Proposition 9.5 and Definition 9.6. In both examples $V = \mathbb{Z}^2$, and elements of V are denoted $h = (h^1,h^2)$ where h^1, h^2 are integers.

Example 9.7 (This is the deterministic part of Example 1.5.) Let $R = |U(h)| = 4$,

$$U(0) = \{(0,1),(1,0),(0,1),(1,1)\},$$
$$f(x_1,x_2,x_3,x_4) = x_1x_2 \vee x_3x_4$$

where $\vee$ is logical disjunction (which coincides with maximum). Note that the order in which the elements of $U(0)$ are enumerated corresponds to the order of x_1, x_2, x_3, x_4.

You can check that σ_D is empty. The range of D is 1. This means that (9.3) and (9.4) hold with $m = 1$. In fact,

$$L_0(h) = -h^1, \; L_1(h) = h^1 - 1$$

will do. The condition (9.4) becomes

$$\min_{h \in I(xD)} h^1 \geqslant \min_{h \in I(x)} h^1$$

$$\max_{h \in I(xD)} h^1 \geqslant \max_{h \in I(x)} h^1 - 1$$

Geometrically this means that evolution of any island x (that is, behaviour of the sequence x, xD, xD^2, ...) is subject to the following condition: the set $I(xD^t)$ is contained between two vertical lines of which the left one remains immovable and the right one moves left by one unit of the lattice in one unit of time. When the right line reaches the left one, the island certainly gets eroded.

In general, an eroder D of range 1 erodes islands in a special way: it squeezes them between two parallel hyperplanes which move uniformly so that the distance between them diminishes by a constant every unit of time.

Example 9.8 (Deterministic part of Example 1.1b). Let $R = |u(h)| = 3$,

$$U(0) = \{(0;0),(0;1),(1;0)\},$$
$$f(x_1,x_2,x_3) = x_1x_2 \vee x_1x_3 \vee x_2x_3.$$

You can check that σ_D is empty. The range of D is 2. The three functionals may be taken in the form:

$$L_0(h) = -h^1, \; L_1(h) = -h^2, \; L_2(h) = h^1 + h^2 - 1.$$

This operator erodes in another way: for any island x the set $I(xD')$ is enclosed between three lines of which two remain immovable and the third one at every step of time moves a constant distance towards the intersection of the two immovable lines.

Now let us look at the main points of the proofs of Theorem 9.4 and Proposition 9.5. In fact, instead of 9.4 and 9.5 we prove the following three assertions:

(1) If σ_D is empty, then there exist functionals subject to conditions (9.3) and (9.4).
(2) If there are functionals subject to (9.3) and (9.4), then D is an eroder.
(3) If σ_D is non-empty, then D is a non-eroder.

Let us prove (1). Let a halfspace mean a closed part of $\mathbb{R}^d$ bounded by a hyperplane. Let a zero-halfspace mean such a halfspace that contains a zero-set. The set σ_D is an intersection of a finite number of zero-halfspaces. In fact, let K be a zero-set. Its convex hull can be represented as an intersection of several halfspaces which clearly are zero-halfspaces. Now σ_D is the intersection of all these zero-halfspaces over all zero-sets.

Let $\pi_1, \ldots, \pi_q$ stand for those zero-halfspaces whose common part is just σ_D. To each π_k of them a normed linear functional L_k corresponds which is non-negative just at π_k. We have supposed that σ_D is empty. Then by the theorem 21.3 in [66] (which is a version of Helly's theorem), among $L_1, \ldots, L_q$ there are $m + 1$ such, hence denoted $L_0, \ldots, L_m$, such that

$$\sum_{k=0}^{m} \lambda_k L_k + \ell \equiv 0 \tag{9.5}$$

where $\lambda_0, \ldots, \lambda_m, \ell$ are positive numbers and

$$2 \leqslant m + 1 \leqslant d + 1.$$

The functionals $\lambda_k L_k$, $0 \leqslant k \leqslant m$, are taken as those in Proposition 9.5. The truth of (9.3) for them follows from (9.5). The formula (9.4) can be inferred from the fact that all the halfspaces defined by inequalities $L_k \geqslant 0$ are zero-halfspaces.

Proof of (2) is obvious.

Now to prove (3). Let us say that an island x fills a set $M \subset \mathbb{R}^d$ if

$$I(x) \supseteq (M \cap \mathbb{Z}^d),$$

that is, if x has ones in all the integer points of M. We can construct a set $M \subset \mathbb{R}^d$ having the following property: if x fills M then xD fills $M + v$, which is a shift of M. Clearly this proves (3), provided this set M is so large that any its shift meets $\mathbb{Z}^d$. It remains to find M.

Let us term a set $A \subset \mathbb{R}^d$ 'obtuse' for a set $B \subset \mathbb{R}^d$ if

$$\forall c \in \mathbb{R}^d((A + c) \cap B = \varnothing \Rightarrow (A + c) \cap \operatorname{conv}(B) = \varnothing),$$

where $\operatorname{conv}(B)$ is the convex hull of B and $A + c = \{a + c | a \in A\}$.

Let $Z_1, \ldots, Z_Q$ be all the zero-sets. Suppose that for every Z_q, $1 \leqslant q \leqslant Q$, we have found a bounded set $\overline{Z}_q$ which is obtuse for Z_q. Then

$$M = \overline{Z}_1 + \ldots + \overline{Z}_Q + C$$

will do, where the unit cube C is added to guarantee that any shift of M meets $\mathbb{Z}^d$.

Indeed, for any set $Z \subset \mathbb{R}^d$ the set $-d \operatorname{conv}(Z)$ is obtuse for Z. So we may take $\overline{Z}_q = -d \operatorname{conv}(Z_q)$. (See details in [90].) Further, if A is obtuse for B, any $A + A'$ is obtuse for B too. Thus our M is obtuse for all zero-sets. Hence M can be proved to have the claimed property: if x fills M then xD fills $M + v$ where v is any point of σ_D.

Finally, we still need to explain the importance of σ_D for the behaviour of non-eroders. Informally, for any large enough initial island the set of ones after time t is something like $-t\sigma_D$ which is the image of homothety of σ_D with coefficient $-t$. To be precise, the following holds.

Proposition 9.9 [87]. Let σ_D be non-empty. Define an island x by

$$I(x) = \operatorname{Sph}(r) \cap \mathbb{Z}^d$$

where $\operatorname{Sph}(r)$ is the ball with centre 0 and radius $r > r_0$. Then

$$(-t\sigma_D) \cap \mathbb{Z}^d \subseteq I(xD^t) \subseteq (-t\sigma_D) + \operatorname{Sph}(kr)$$

for all natural t where r_0 and k are constants depending only on D.

This proposition answers many natural questions about behaviour of operators: islands growing infinitely exist if σ_D consists of more than one point; islands expanding to fill the whole lattice exist if 0 is the inner point of σ_D, and so on.

But some other no less natural questions cannot be solved algorithmically even in the minimal dimension $d = 1$. There is such a monotone homogeneous D on $\{0;1\}^{\mathbb{Z}}$ for which there is no algorithm to say for any island whether D erodes it or not [92]. There is an island x in $\{0;1\}^{\mathbb{Z}}$ for which there is no algorithm to say for any monotone homogeneous operator whether it erodes x or not [92]. Both results are proved by making operators simulate Turing machines. And there is no algorithm to say for any homogeneous operator whether it is an eroder or not [57].

Analogous problems about eroding may be examined for $X_0 = \{0;1,\ldots,m\}$ too. For the one-dimensional case this is done in [23–25].

Chapter 10

Non-ergodic non-degenerate operators

The main aim of this chapter is to present homogeneous non-degenerate non-ergodic operators. We can do it for $d \geqslant 2$. The following is a key theorem. A more general theorem is published in [90] as theorem 1.

Theorem 10.1 This theorem concerns eroders with noise. For any monotone eroder D: $X \to X$ where $X = \{0;1\}^{\mathbb{Z}^d}$ the operator DS_ε has an invariant measure which tends to $\mathbf{0}$ with $\varepsilon \to 0$.

The proof is not simple – it takes more than half of this chapter. So let us first explain its idea informally. In this proof $\tilde{\mu}$ will stand for the evolution measure produced by DS_ε with the initial measure $\delta_{\mathbf{0}}$, 'all zeros'. We shall estimate the probability $\tilde{\mu}(x_0^t = 1)$:

$$\tilde{\mu}(x_0^t = 1) \leqslant \psi(\varepsilon)$$

where $\psi(\varepsilon)$ does not depend on t and tends to 0 with ε. Due to Corollary 2.8, this guarantees that DS_ε has an invariant measure μ_ε such that

$$\mu_\varepsilon(x_0 = 1) \leqslant \psi(\varepsilon) \to 0,$$

whence $\mu_\varepsilon \to \delta_{\mathbf{0}}$ with $\varepsilon \to 0$.

To achieve this, we present $\tilde{\mu}$ as induced by the Bernoulli measure β on Ω (see Chapter 2) with the mapping

$$x_h^0 \equiv 0,\ x_h^{t+1} = \max(\omega_h^{t+1}, f(x_{U(h)}^t)),\ t \geqslant 0.$$

Every ω_h^t equals 1 with probability ε and equals 0 with probability $(1-\varepsilon)$ independently from the others, according to the measure β. So we have to estimate the value of the measure β of the set

$$\{\omega \in \Omega: x_0^t = 1\} \subset \Omega.$$

We cover it by a system of cylinder sets, the number of which we estimate. The most cumbersome part of the proof is the construction of special combinatorial objects for all the cylinder sets in question which enable us to estimate their number. The range of D plays an essential role in this construction. If the range equals 1, these combinatorial objects turn into contours like those we used to prove non-ergodicity of Example 1.2.

Now we start to introduce these constructions. Let D have the range m. Let $L_0, \ldots, L_m$ be the linear functionals on $\mathbb{Z}^d$ satisfying (9.3) and (9.4) whose existence is provided by Proposition 9.5. Denote $M = \{0,1,\ldots,m\}$.

Let us reword the condition (9.4) as follows: if $f(x_{u(h)}) = 1$, then for any $k \in M$ there is such $a \in U(h)$ that $x_a = 1$ and $L_k(a) \geqslant L_k(h) - L_k(0)$. Now we denote

$$r = \frac{\ell}{m + 1} \tag{10.1}$$

and introduce functionals on the evolution space $\mathbb{Z}^d \times \mathbb{Z}_+$:

$$\mathscr{L}_k(h,t) = L_k(h) + (L_k(0) - r)t + r$$

for $k \in M$ Clearly,

$$\sum_{k=0}^{m} \mathscr{L}_k(h,t) \equiv 0.$$

Also, if $f(x_{u(h)}) = 1$ then for any $k \in M$ there is such a point $a \in U(h)$ that $x_a = 1$ and $\mathscr{L}_k(a,t) \geqslant \mathscr{L}_k(h,t+1) + r$. Denote also

$$\rho = \max_{a \in U(0),\, K \in M} |\mathscr{L}_k(a,t) \times \mathscr{L}_k(0,t+1)|. \tag{10.2}$$

The further proof consists of four parts using the notations introduced here.

Part I

Remember that $M = \{0;1;\ldots,m\}$. Let a *polar* π on a set A stand for any mapping π: $M \to A$, and the kth *pole* of π stand for $\pi(k)$ for any $k \in M$. Let $\pi(M)$ stand for the range of π and more generally $\pi(M') = \{\pi(k), k \in M'\}$ for any $M' \subseteq M$. A polar π on the set of vertices of a graph g is termed a bonder on g if $\pi(M)$ consists of two ends of a bond. In any graph throughout this proof two ends of a bond are always different from each other, and any two vertices a and b are connected with not more than one bond denoted as (a,b). In that special case where two vertices may be connected with more than one bond we speak of a multigraph instead of a graph.

First we need polars on our evolution space $\mathbb{Z}^d \times \mathbb{Z}_+$ of the two following special types.

A *k-arrow* (or just *arrow* if we are not interested in the value of k) stands for a polar π on $\mathbb{Z}^d \times \mathbb{Z}$ where

$$\pi(k) = (a,t),\ \pi(M\backslash k) = (b,t+1),\ a \in U(b),$$
$$\mathscr{L}_k(a,t) - \mathscr{L}_k(b,t+1) \geqslant r.$$

A *fork* stands for a polar π on $\mathbb{Z}^d \times \mathbb{Z}_+$, for which

$$\pi(M) = \{(a,t),(b,t)\}$$

where $a \neq b$ and

$$\exists c\colon \{a;b\} \subset U(c);$$

clearly, an arrow cannot be a fork. It is convenient to introduce a new graph γ with $\mathbb{Z}^d \times \mathbb{Z}_+$ as the set of vertices; its two vertices (a,t) and (a',t') are connected with a bond if

(1) either $a \in U(a')$ and $t = t' - 1$,
(2) or $a' \in U(a)$ and $t' = t - 1$,
(3) or $t = t'$ and there is such b that $\{a;a'\} \subset U(b)$.

Clearly, γ is defined so as to make all arrows and forks bonders on γ.

We assume two finite sequences of polars to be equivalent if some permutation turns one of them into the other. Thus all finite sequences of polars separate into equivalence classes which will be termed *trusses*. A truss π will be written down as a sequence:

$$\Pi = (\pi_1,\ldots,\pi_k).$$

There is an empty truss which contains no polar. For any truss $\Pi = (\pi_1,\ldots,\pi_k)$ we denote

$$\Pi(M) = \bigcup_{\ell=1}^{k} \pi_\ell(M).$$

$\Pi_1 * \Pi_2$ stands for the concatenation of Π_1 and Π_2 (just writing one sequence after the other).

A truss $\Pi = (\pi_1,\ldots,\pi_k)$ on some set A will be termed *p-even* (or just *even* if we are not interested in the value of p) on some $B \subseteq A$ if for every $\ell \in M$ just p members of the sequence $\pi_1(\ell), \ldots, \pi_k(\ell)$ belong to B. A truss Π on the set A is termed *overall even* on $B \subseteq A$ if it is even on every single element $b \in B$ (which implies Π being even on every $B' \subseteq B$).

For any truss $\Pi = (\pi_1,\ldots,\pi_k)$ on $\mathbb{Z}^d \times \mathbb{Z}_+$ consisting only of bonders on γ we introduce the multigraph $\gamma(\Pi)$ as follows. $\Pi(M)$ is its set of vertices. For every π_ℓ, $1 \leqslant \ell \leqslant k$ there is the corresponding bond of $\gamma(\Pi)$

which connects those two vertices of γ that are poles of the bonder π_ℓ. We term Π connected if $\gamma(\Pi)$ is connected.

Lemma 10.2 Let an overall even truss on $\mathbb{Z}^d \times \mathbb{Z}_+$ consist of k arrows and ℓ forks. Then

$$rk \leqslant 2m\rho\ell$$

where r and ρ are defined in (10.1) and (10.2).

Proof. For any polar π on $\mathbb{Z}^d \times \mathbb{Z}_+$ the value of

$$\sum_{k=0}^{m} \mathscr{L}_k(\pi(k))$$

will be termed the *span* of π. One can check that if a truss on $\mathbb{Z}^d \times \mathbb{Z}_+$ is overall even, then the sum of spans of its polars equals zero.

Thus the sum of spans of our k arrows and ℓ forks equals zero.

It is easy to estimate that the span of any arrow is no less than r. The span of any fork is no more than $2m\rho$ in modulo. Hence the sum of spans of our truss's polars is no less than $rk - 2m\rho\ell$. The resulting inequality $0 \geqslant rk - 2m\rho\ell$ proves the lemma.

Lemma 10.3 Consider all connected trusses Π on $\mathbb{Z}^d \times \mathbb{Z}_+$ consisting of n bonders on graph γ whose $\Pi(M)$ contain a given point. Their number is no more than C^n where C depends only on m and R. One may take

$$C = (2^{m+1}(R^2 + 2R))^2.$$

Proof. Remember that any truss Π in question has its own connected multigraph. It is known that in any connected multigraph there is a closed path which passes every bond twice. In our case this path has $2n$ steps. For every Π we choose such a path that begins and ends in the given point and use it to code our truss by a sequence of $2n$ symbols, one for each step. Each symbol codes one polar, and their order coincides with the order of steps in the path. Now to estimate how many values each symbol must have. Suppose we have gone through our path up to a point (h,t). To define the next polar we need to know:

(1) Along which bond of γ starting from (h,t) we make the next step. This makes no more than $R^2 + 2R$ variants. In fact, just from definition of γ one can see that every vertex of γ meets no more than $R^2 + 2R$ bonds.
(2) At which end of this bond is every pole of the next polar. This makes less than 2^{m+1} variants.

Thus, every symbol having $2^{m+1}(R^2 + 2R)$ variants, we can code our truss by a sequence of $2n$ symbols, which proves the lemma.

Lemma 10.4 For any polar π_0 on the set of vertices of a connected graph g there is such a truss $(\pi_1,\ldots,\pi_\ell)$ of bonders on g that the truss

$$(\pi_0,\pi_1,\ldots,\pi_\ell)$$

is either 0-even or 1-even on every vertex of g.

Substitute g by g', some minimal connected subgraph of g whose set of vertices still contains $\pi_0(M)$. Clearly g' is a tree. For every bond (a,b) of the tree g' we define a bonder $\pi_{(a,b)}$ on g as follows: its pole $\pi_{(a,b)}(k)$ is that element of $\{a;b\}$ which would no longer be connected with $\pi_0(k)$ if bond (a,b) were rejected from tree g'; and this is done for all $k \in M$. So $\pi_1, \ldots, \pi_\ell$ are thus defined bonders for all bonds of the tree g' and the assertion of the lemma can be proved.

Part II

Fix a natural number T. The value of x_0^T is a certain function of arguments ω_h^t, where

$$(h,t) \in \widetilde{U}^\infty(0,T).$$

These are the only points we need and only these will be called 'points' till the end of the proof. Often we shall denote a point by one letter, say $a = (h,t)$, in which case we shall write $t(a) = t$, $x_a = x_h^t$, $\omega_a = \omega_h^t$.

Now we fix some $\omega \in \widetilde{\Omega}$ such that it makes $x_0^t = 1$. All the constructions of Parts II and III of our proof are made for the fixed T, ω. That is, in Parts II and III we shall build some truss corresponding to the fixed ω. In Part IV T will remain fixed and we shall estimate the number of different trusses built.

For every point (h,t) where $x_h^t = 1$ but $\omega_h^t = 0$ we do the following. In this case

$$f(x_{\widetilde{U}(h,t)}) = 1,$$

that is, the value 1 of x_h^t is not spontaneous but inherited. This implies, as we noted just after introducing the functionals $\mathcal{L}_k$, that for every $k \in M$ there is a point

$$\bar{u}_k(h,t) \in \widetilde{U}(h,t)$$

in which

$$x_{\bar{u}_k(h,t)} = 1$$

and

$$\mathcal{L}_k(\bar{u}_k(h,t)) - \mathcal{L}_k(h,t) \geqslant r.$$

We fix these $\overline{u}_k(h,t)$ and denote

$$\overline{U}(h,t) = \{u_k(h,t), k \in M\}.$$

If $x_h^t = 0$ or $\omega_h^t = 0$, the set $\overline{U}(h,t)$ is empty by definition. Now we denote

$$\overline{U}(A) = \bigcup_{a \in A} \overline{U}(a),$$

$$\overline{U}^{k+1}(A) = U(U^k(A)),$$

where $\overline{U}^0(A) = A$, and,

$$\overline{U}^\infty(A) = \bigcup_{k=0}^{\infty} U^k(A).$$

Denote also

$$\hat{U} = \{(h,t) \in \overline{U}^\infty(0,T) \colon \overline{U}(h,t) = \varnothing\}.$$

For all $(h,t) \in \hat{U}$, $t \geqslant 1$ and $\omega_h^t = 1$. Informally speaking, elements of $\hat{U}$ are some of those points where ones first appeared spontaneously (that is, due to $\omega_h^t = 1$, not to $f(x_{u(h,t)}) = 1$) which after the iterative process resulted in $x_0^T = 1$.

This iterative process goes along the set $\overline{U}^\infty(0,t)$. Hence we concentrate on this set and a point will mean a point of this set.

Let us define an equivalence relation $\sim$ between some points (of $\overline{U}^\infty(0,T)$, of course) by the following three rules:

(1) If $t(a) = t(b)$ and

$$\overline{U}^\infty(a) \cap \overline{U}^\infty(b) \neq \varnothing$$

then $a \sim b$.

(2) If $a \sim b$ and $b \sim c$ then $a \sim c$.

(3) Only those points are equivalent which are made so by the two former conditions.

The resulting equivalence classes will be called just *classes*. Now we define an oriented graph τ whose vertices are all these classes. In this graph an oriented bond goes from class A to class B if and only if there are such points $a \in A$ and $b \in B$ that $a \in \overline{U}(b)$. Let $U_\tau(A)$ stand for the set of those classes whence oriented bonds go to this class A. For any set S of classes we define

$$U_\tau(S) = \bigcup_{A \in S} U_\tau(A),$$

then

$$U_\tau^{k+1}(S) = U_\tau(U_\tau^k(S)),$$

where $U_\tau^0(S) = S$ and

$$U_\tau^\infty(S) = \bigcup_{k=0}^{\infty} U_\tau^k(S).$$

Of course, our point $(0,T)$ is not equivalent to any other point because it has the greatest value of t. So $(0,T)$ forms its own class $\{(0,T)\}$. Also any point $a \in \hat{U}$ is not equivalent to any other because its $\bar{U}^\infty(a)$ is empty.

Let us prove that τ is a tree. In fact, let A, B and C be such classes that

$$A \in U_\tau(C),\ B \in U_\tau(C) \quad \text{and} \quad A \neq B.$$

Then

$$U_\tau^\infty(A) \cap U_\tau^\infty(B)$$

is empty because otherwise A and B would be included into one class. So, starting from $\{(0,T)\}$ and moving against the direction of bonds, we shall never come to one class by two different ways, and moving in the direction of bonds of τ we never have more than one way. So τ is a tree, all bonds of which are so directed that moving in this direction we come to $\{(0;T)\}$.

For every class A which has non-empty $U_\tau(A)$ we define a non-oriented graph $g(A)$ as follows. Its set of vertices is $U_\tau(A)$. Two classes B, $D \in U_\tau(A)$ are connected with a bond in $g(A)$ if and only if $B \neq D$ and there are such points a, b, d that

$$b \in (B \cap \bar{U}(a)) \quad \text{and} \quad d \in (D \cap \bar{U}(a)).$$

Lemma 10.5 Every graph $g(A)$ is connected.

Proof. Let B and B' be two different vertices of $g(A)$. Let us construct a path from B to B' in $g(a)$. Since B, $B' \in U_\tau(A)$, there are such $b \in B$, $b' \in B'$, a, $a' \in A$ that

$$b \in \bar{U}(a),\ b' \in \bar{U}(a').$$

If $a = a'$, then B and B' are connected with some bond which is what we want. So let $a \neq a'$. But $a \sim a'$ since they have got into one class A. This means by definition of equivalence $\sim$ that there is such a sequence

$$a_0 = a,\ a_1,\ a_2,\ \ldots,\ a_{n-1},\ a_n = a'$$

of elements of A (all of which we may and shall assume different) such that every intersection

$$\bar{U}^\infty(a_{\ell-1}) \cap \bar{U}^\infty(a_\ell),\ 1 \leqslant \ell \leqslant n$$

is non-empty. Hence there are such

$$d_{\ell-1} \in \bar{U}(a_{\ell-1}) \quad \text{and} \quad d'_\ell \in \bar{U}(a_\ell)$$

that

$$\bar{U}^{\infty}(d_{\ell-1}) \cap \bar{U}^{\infty}(d_{\ell})$$

is non-empty. Hence $d_{\ell-1} \sim d'_{\ell}$ and belong to one class which we denote D_{ℓ}, $1 \leqslant \ell \leqslant n$. Now let us consider the sequence of classes B, D_1, D_2, ..., D_n, B'). In this sequence every two next classes either coincide or are connected with a bond of $g(A)$. q.e.d.

Part III

Here q is an integer parameter which grows from 0 to Q; the value of Q will be chosen later. For every q in this range we shall construct two trusses Π_q^1 and Π_q^2 on $\bar{U}^{\infty}(0,T)$, a set S_q of classes and a non-oriented graph G_q which depends on Π_q^1, Π_q^2, S_q in the following way. Denote

$$\Pi_q = \Pi^0 * \Pi_q^1 * \Pi_q^2$$

where Π^0 is the truss consisting of one polar, all poles of which are in the point $(0;T)$. The set of vertices of G_q is $\Pi_q(M)$. Two different vertices of G_q are connected with a bond if at least one of the two following conditions holds:

(1) Either both of them belong to some $\pi(M)$ where π is a polar in Π_q;
(2) or both of them belong to one class which enters S_q.

We shall build Π_q^1, Π_q^2, S_q (and consequently G_q too) in the inductive way with q growing from 0 to Q. Having just constructed Π_q^1, Π_q^2, S_q and G_q, we shall prove the following seven properties of them:

1 The truss Π_q^1 consists only of arrows; the truss Π_q^2 consists only of forks.
2 If two different classes A and B belong to S_q, then A does not belong to $U_T^{\infty}(B)$.
3 For any point $a \in \Pi_q(M)$ there is such a class $A \in S_q$ that $\bar{U}^{\infty}(a) \cap A \neq \varnothing$.
4 The truss Π_q is overall even on the set

$$\bar{U}^{\infty}(0,T) \backslash \bar{U}^{\infty}\left(\bigcup_{A \in Sq} A\right).$$

5 The truss Π_q is 1-even on every class belonging to S_q.
6 The number of elements of S_q equals the number of forks in Π_q^2 plus one.
7 The graph G_q is connected.

These properties, as soon as they get proved, warrant the next step of construction of Π_{q+1}^1, Π_{q+1}^2, S_{q+1} and G_{q+1}. Thus the induction steps of construction and proof alternate. However, the proofs are obvious and we omit them. Let us describe the construction induction step in the following three items.

1 In the initial case $q = 0$ the trusses Π_0^1 and Π_0^2 are empty. The set S_0 consists of one class $\{(0;T)\}$ Clearly, all the seven properties hold in this case.
2 When the induction stops. Suppose that at some construction step we have got such S_q that

$$\forall A \in S_q \colon U_\tau(A) = \varnothing.$$

Then this step is the last one and the present value of q is Q.
3 The induction step. Suppose that Π_q^1, Π_q^2, S_q and G_q have been built and possess the seven properties. Suppose also that there is such a class $A \in S_q$, for which $U_\tau(A)$ is non-empty.

Then we build Π_{q+1}^1, Π_{q+1}^2 and S_{q+1} in the following way. We make

$$\Pi_{q+1}^1 = \Pi_q^1 * \Pi^1 \quad \text{and} \quad \Pi_{q+1}^2 = \Pi_q^2 * \Pi^2$$

where it remains to define Π^1 and Π^2. Remember that A is fixed. Property 5 claims that for any $k \in M$ just one polar of Π_q has its kth pole in A. Let $\alpha_k \in A$ stand for this kth pole. We form an arrow π_k by the following rule: its kth pole is $\bar{u}_k(\alpha_k)$ and all its other poles are in the point α_k. The truss Π_1 consists of these $m + 1$ arrows:

$$\Pi_1 = (\pi_0, \ldots, \pi_m).$$

Note that the truss $\Pi q * \Pi^1$ is overall even on A and 1-even on the union of all elements of

$$U_\tau^\infty(A) \backslash A.$$

The last assertion allows us to define a polar π_0 on $U_\tau(A)$ by the following rule: $\pi_0(k) = B$ where B is that element of $U_\tau(A)$ for which the union of all elements of $\bar{U}^\infty(B)$ contains the kth pole of some polar in $\Pi_q * \Pi_1$. Having thus formed polar π_0, we use Lemmas 10.5 and 10.4 to obtain such a truss $(\pi_1, \ldots, \pi_k)$ of bonders on $g(A)$ that the truss $(\pi_0, \pi_1, \ldots, \pi_k)$ is either 0-even or 1-even at every vertex of $g(A)$. Having got bonders $\pi_1, \ldots, \pi_k$ we form the corresponding forks $\bar{\pi}_1, \ldots, \bar{\pi}_k$ by the following rule. Let $\pi_\ell(M) = \{B, B'\} \subset U_\tau(A)$. Since B and B' are connected with a bond in $g(A)$, there are such points a, b, b' that $b \in B \cap \bar{U}(a)$ and $b' \in B' \cap \bar{U}(a)$. These define fork $\bar{\pi}_\ell$ in the following way: if $\pi_\ell(k) = B$ then $\bar{\pi}_\ell(k) = b$ and if $\pi_\ell(k) = B'$ then $\bar{\pi}_\ell(k) = b'$ for all $k \in M$. The sequence of forks $\bar{\pi}_1, \ldots, \bar{\pi}_k$ thus defined presents the truss Π^2.

Now define the set S_{q+1}. It results from S_q by excluding A and including those classes $B \in U_\tau(A)$ on which the truss Π_{q+1} is 1-even (instead of being 0-even).

The construction induction step is described. It is easy to prove that if Π^1_q, Π^2_q, S_q and G_q possess the seven properties, then Π^1_{q+1}, Π^2_{q+1}, S_{q+1} and G_{q+1} do too.

Note that every construction induction step diminishes the number of classes in $U^\infty_\tau(S_q)$. This guarantees that the induction will terminate. After that we shall have the trusses Π^1_Q and Π^2_Q and the set S_Q. Clearly, every class entering S_Q consists of one point. This allows us to define

$$S = \{a: \{a\} \in S_Q\}.$$

Clearly, $S \subset \hat{U}$. Denote also $\Pi = \Pi^1_Q * \Pi^2_Q$. Applying our seven properties to the case $q = Q$ we see that Π is connected, overall even, and consists of arrows and forks, the number of forks in Π being $|S| - 1$.

Informally, elements of S are some of those points where ones first emerged spontaneously (that is, due to ω-influence, not to inheritance). Of course, the probability of spontaneous emergence of ones in all elements of S is $\varepsilon^{|S|}$. Summing these probabilities over all S gives us an upper estimate for $\tilde{\mu}(x^T_0 = 1)$. This summing is based on a combinatorial estimation which is made using the corresponding Π.

Part IV

The point $(0,T)$ remains fixed. But ω is no longer fixed; instead we consider all ω which make $x^T_0 = 1$. For every such ω we have constructed the truss Π and the set S which we now denote $\Pi(\omega)$ and $S(\omega)$. It is important that $\Pi(\omega)$ determines $S(\omega)$, that is,

$$(\Pi(\omega) = \Pi(\omega')) \Rightarrow (S(\omega) = S(\omega')). \tag{10.3}$$

In fact, applying our properties 3 and 5 to the case $q = Q$ proves that S may be defined as the set of those points $a \in \Pi(M)$ for which there is no such arrow π in Π that

$$\pi(M) = \{a;b\} \quad \text{where} \quad b \in U(a).$$

Now let W_k stand for the set of different sets $S(\omega)$ consisting of $k + 1$ points. We are going to estimate W_k.

Since Π^2_Q contains k forks, $\Pi(\omega)$ contains no less than k and no more than $k(1 + 2m\rho/r)$ polars due to Lemma 10.2. This and Lemma 10.3 allow us to estimate the number of different $\Pi(\omega)$ in this case. Due to (10.3) the same estimation holds for W_k:

$$W_k \leqslant \sum_{\ell=k}^{k(1+2m\rho/r)} (2^m(R^2 + 2R))^{2\ell}.$$

Now we are ready to estimate $\tilde{\mu}(x^T_0 = 1)$. For any ω which makes $x^T_0 = 1$, we introduce the cylinder set C_ω:

$$C_\omega = \{\omega': \omega'_a = 1 \quad \text{for all} \quad a \in S(\omega)\}.$$

Every ω enters its C_ω due to Lemma 10.5. Hence

$$\{\omega: x_0^T = 1\} \subset UC_\omega$$

where the union on the right-hand side is over such ω that $x_0^T = 1$. Of course, different ω's have equal C_ω in a lot of instances. So

$$\tilde{\mu}(x_0^T = 1) \leqslant \Sigma\tilde{\mu}(C_\omega)$$

where the sum on the right-hand side is taken over all different C_ω. According to the definition of W_k this turns into

$$\tilde{\mu}(x_0^T = 1) \leqslant \sum_{k=0}^{\infty} W_k \varepsilon^{k+1}.$$

Substituting here our estimation for W_k we see that the series on the right-hand side converges for small enough $\varepsilon > 0$ and its sum tends to zero with ε. This proves our main theorem about eroders with noise.

The main job of the present chapter was to prove this theorem. Since it is proved, it is easy to construct non-ergodic non-degenerate operators on $\{0;1\}^{\mathbf{Z}^d}$.

Remember, first, our Examples 1.1a and 1.5. Their non-ergodicity with small $\varepsilon > 0$ just follows from our theorem. In fact, both can be presented as DS_ε^+ where corresponding D are described in Examples 9.7 and 9.8 and S_ε^+ is defined in Chapter 2. In both cases D is an eroder and DS_ε has an invariant measure which tends to $\mathbf{0}$ with $\varepsilon \to 0$ due to our theorem. But $S_\varepsilon^+ \prec S_\varepsilon$ whence $DS_\varepsilon^+ \prec DS_\varepsilon$. So DS_ε^+ has an invariant measure which tends to $\mathbf{0}$ too.

On the other hand we can rename zeros as ones and ones as zeros in $X = \{0;1\}$ and follow the same argument. Our D's are chosen in such a way that they are eroders in this case too. So DS_ε^+ have invariant measures which tend to $\mathbf{1}$ with $\varepsilon \to 0$ too. Of course, measures which tend to $\mathbf{0}$ and measures which tend to $\mathbf{1}$ are different with small enough $\varepsilon > 0$, and our operators are non-ergodic.

A larger variety of examples can be constructed with any $d \geqslant 2$. The following example shows one way to do this in the case $d = 2$.

Example 10.6 [90]. Let $x = \{0;1\}^V$, where $V = \mathbb{Z}^2$. Choose any natural n. We are going to construct a non-degenerate P_ε which has invariant measures tending to different states $y^1, \ldots, y^n$ with $\varepsilon \to 0$. Of course, with $\varepsilon > 0$ small enough these measures have to be different and P will have no less than n different invariant measures.

We term $x \in X$ a translate of $y \in X$ if there is such $a \in \mathbb{Z}^2$ that $x_h = y_{h+a}$ for all $h \in \mathbb{Z}^2$. Define a periodic state as having only a finite number of different translates. So we choose our $y^1, \ldots, y^n$ to be periodic. Without loss of generality we assume that all translates of $y^1, \ldots, y^n$ are among these $y^1, \ldots, y^n$ too. Now we choose ρ large enough to ensure the following: for any circle C of radius ρ and any k, ℓ there exists $h \in C \cap \mathbb{Z}^2$ such that $y_h^k \neq y_h^\ell$. Then we choose three homogeneous linear functionals $L_1, L_2, L_3: \mathbb{Z}^2 \to \mathbb{R}$ such that any two of them are independent and

$$L_1 + L_2 + L_3 \equiv 0.$$

Finally, we choose three circles C_1, C_2, C_3 of radius ρ in the plane so that

$$\begin{aligned} h \in C_1 &\Rightarrow (L_2(h) \geqslant 1 \quad \text{and} \quad L_3(h) \geqslant 1), \\ h \in C_2 &\Rightarrow (L_3(h) \geqslant 1 \quad \text{and} \quad L_1(h) \geqslant 1), \\ h \in C_3 &\Rightarrow (L_1(h) \geqslant 1 \quad \text{and} \quad L_2(h) \geqslant 1). \end{aligned}$$

Now we define our neighbourhood system:

$$U(0) = (C_1 \cup C_2 \cup C_3) \cap \mathbb{Z}^2, \; U(h) = U(0) + h.$$

So we have a graph $\Gamma = (V, \mathcal{U})$. To define a deterministic operator D it remains only to define a function

$$f\colon X_{U(0)} \to X_0.$$

This we do as follows:

$f(x_{U(0)}) = y_0^k$ if at least two of the following three conditions hold for some $k \in \{1,\ldots,n\}$:

$$\begin{aligned} &(1)\ \forall h \in (C_1 \cap \mathbb{Z}^2)\colon x_h = y_h^k, \\ &(2)\ \forall h \in (C_2 \cap \mathbb{Z}^2)\colon x_h = y_h^k, \\ &(3)\ \forall h \in (C_3 \cap \mathbb{Z}^2)\colon x_h = y_h^k, \end{aligned}$$

and $f(x_{U(0)})$ defined arbitrarily if the previous condition holds for no $k \in \{1,\ldots,n\}$.

You can check that this definition is consistent, whence D is defined; y_1, $\ldots$, y_n are its invariant states. D is defined in such a way that any small independent stochastic perturbation of it has measures which tend to y_1, $\ldots$, y_n when this perturbation vanishes. To express this in a rigorous way we can formulate the following theorem. But first we introduce the random noise operator S_ν where ν is the noise matrix:

$$\nu = \|\nu_{ij}\|$$

where

$$1 \leqslant i \leqslant n, \; 1 \leqslant j \leqslant n, \; \nu_{ij} \geqslant 0, \; \sum_j \nu_{ij} = 1.$$

Action of S_ν at any state $x \in X = \{1,\ldots,n\}^{\mathbb{Z}^d}$ consists in the following: any x_h changes its state from i to j with probability ν_{ij} independently. Denote

$$|S_\nu| = \max_{i \neq j} \nu_{ij}.$$

Theorem 10.7 (A. L. Toom). We have a homogeneous deterministic operator D: $X \to X$ where $X = \{1,\ldots,n\}^V$, $V = \mathbb{Z}^d$, defined in the standard way:

$$(Dx)_h = f(x_{U(h)})$$

where f: $\{1,\ldots.n\}^{|U(h)|} \to \{1,\ldots,n\}$, $U(h) = U(0) + h$. We have a state $y \in X$ which is invariant for D. There are m linear functionals L_1, ..., L_m: $\mathbb{Z}^d \to \mathbb{R}$ with

$$L_1 + \ldots + L_m \equiv 0$$

and a number $r > 0$. For any $v \in \mathbb{Z}^d$, any $k \in \{1,\ldots,m\}$ and any $x_{U(v)}$ such that $f(x_{U(v)}) \neq y_v$ there is such a point $a \in U(v)$ that $x_a \neq y_a$ and

$$L_k(a) - L_k(v) \geqslant r.$$

Then the composition DS_ν has an invariant measure which tends to the state y with $|S_\nu| \to 0$. To deduce this theorem from our main Theorem 10.1 is mere technicality, and the theorem proves our assertion about Example 10.6.

Chapter 11
Random walk operators: non-ergodicity

The task of this chapter is twofold. First, we shall prove a theorem about non-ergodicity of random walk operators composed with the standard random noise S_ε. Second, we shall apply this theorem to prove non-ergodicity of some more independent operators on $\{0;1\}^{\mathbb{Z}}$.

Random walk operators have been defined in Chapter 6. Throughout the present chapter, W stands for any random walk operator in question. We shall use the following notation for the random variables that control the action of W:

$$\xi_{i+1/2},\ \eta_{i+1/2} \quad \text{where} \quad i \in \mathbb{Z}.$$

The value of $\xi_{i+1/2}$ determines the distance at which will move (right if $\xi_{i+1/2} > 0$, left if $\xi_{i+1/2} < 0$) the left-hand end of the massif of ones if its leftmost 1 was in the point $i + 1$. The value of $\eta_{i+1/2}$ determines the distance at which will move (in the same sense) the right-hand end of the massif of ones if its rightmost 1 was in the point i. All $\xi_{i+1/2}$ have equal distributions and $\mathbb{E}\xi$ stands for its mean, as does $\mathbb{E}\eta$ for all $\eta_{i+1/2}$. If $\mathbb{E}\xi < \mathbb{E}\eta$ the massifs grow in the mean; if $\mathbb{E}\xi > \mathbb{E}\eta$ the massifs diminish in the mean.

Since any W is monotone, it has a critical value ε_W^* defined as follows: WS_ε is ergodic with all $\varepsilon > \varepsilon_W^*$ and WS_ε is non-ergodic with all $\varepsilon < \varepsilon_W^*$. Lemma 6.5 proves that $\mathbb{E}\xi < \mathbb{E}\eta$ makes $\varepsilon_W^* = 0$. This is not unexpected, because in this case interactions of W alone (without S_ε) contract 'almost all' $M(x)$ into **1**.

Theorem 11.1 [83]. $\mathbb{E}\xi > \mathbb{E}\eta$ makes $\varepsilon_W^* > 0$.

The formula (11.5) gives estimation for ε_W^* in this case. The theorem just follows from the formula

$$\lim_{\varepsilon\to 0}\lim_{T\to\infty}\mathbf{0}(WS_\varepsilon)^t = \delta_{\mathbf{0}}.$$

To prove it, we shall estimate

$$\tilde{\mu}(x_0^T = x_1^T = \ldots = x_n^T = 1) \tag{11.1}$$

where $\tilde{\mu}$ is the evolution measure generated by WS_ε and $\delta_{\mathbf{0}}$. Let $\xi_{h+1/2}^{t-1}$, $\eta_{h+1/2}^{t-1}$ and ω_h^t stand for the auxiliary variables which control the tth application of our operators W and S_ε where $t = 1, 2, \ldots$. Then any x_H^T is a function (in the usual deterministic sense) of a finite set of these arguments. To estimate (11.1), we shall cover the set of those (ξ,η,ω) which make

$$x_0^T = x_1^T = \ldots = x_n^T = 1$$

with a countable family of cylinder sets up to a set of measure 0. Then we shall estimate (11.1) by the sum of measures of these cylinder sets. Assuming $\varepsilon < 1$, as we shall do, guarantees that infinite massifs of ones at any time have the measure 0. Hence (11.1) equals the probability of having such $k \leqslant 0$ and $\ell \geqslant n$ that

$$x_{k-1}^T = 0,\ x_k^T = x_{k+1}^T = \ldots = x_{\ell-1}^T = x_\ell^T = 1,\ x_{\ell+1}^T = 0. \tag{11.2}$$

So let us estimate the sum of probabilities of (11.2) over all $k \leqslant 0$, $\ell \geqslant n$. For that we shall build a polygonal line which is a contour of a sort surrounding the two-dimensional cluster of ones which contains the segment $\{(h,T)\colon k \leqslant h \leqslant \ell\}$. To every step of such a polygonal line there corresponds a certain auxiliary variable $\xi_{h+1/2}^t$, $\eta_{h+1/2}^t$, or ω_h^t which must have a certain value. Let us describe in detail which steps may occur. All steps begin and end in points $(h+\frac{1}{2},t)$ where h and t are integers. Any step will be written down as

$$(h_1+\tfrac{1}{2},t_1) \to (h_2+\tfrac{1}{2},t_2)$$

where $(h_1+\frac{1}{2},t_1)$ is its starting point and $(h_2+\frac{1}{2},t_2)$ is its final point. There are four types of steps:

Type 1 step is

$$(h-\tfrac{1}{2},t) \to (h+\tfrac{1}{2},t).$$

The corresponding restriction is $\omega_h^t = 1$.

Type 2 step is

$$(h+\tfrac{1}{2}+j,t+1) \to (h+\tfrac{1}{2},t) \quad \text{where} \quad -S \leqslant j \leqslant S.$$

The corresponding restriction is $\xi_{h+1/2}^t = j$.

Type 3 step is

$$(h+\tfrac{1}{2},t) \to (h+\tfrac{1}{2}+j,t+1) \quad \text{where} \quad -S \leqslant j \leqslant S.$$

The corresponding restriction is $\eta^t_{h+1/2} = j$.
Type 4 step is

$$(h+\tfrac{1}{2},t) \to (h+\tfrac{1}{2}+j,t) \quad \text{where} \quad 1 \leqslant j \leqslant r.$$

There is no corresponding restriction.

It is essential that type 2 and type 3 steps should not be next to each other, and any type 4 step demands that there be type 1 steps just before and just after it. The polygonal line begins in the point $(k-\frac{1}{2},T)$ and ends in the point $(\ell+\frac{1}{2},T)$.

You can guess the reasons for such definitions. In fact, type 1 restrictions reflect the spontaneous emergence of ones due to S_ε. Type 4 steps reflect the action of W_1 which blots out short sequences of zeros (turning all their zeros into ones). Type 2 and type 3 steps reflect the action of W_2. They imitate the movement of left-hand (type 2) and right-hand (type 3) ends of massifs of ones.

Now we are going to estimate the sum of measures of all resulting cylinder sets over all polygonal lines allowed by us and all $k \leqslant 0$ and $\ell \geqslant n$. To this end, let us for a while allow the polygonal lines to end in all points $(h+\frac{1}{2},t)$, h, $t \in \mathbb{Z}$. For every polygonal line with corresponding restrictions we define a weight which is the measure of the cylinder set resulting from all the restrictions corresponding to its steps. Let $\sigma^N_\Delta(h+\frac{1}{2},t)$ where h, $t \in \mathbb{Z}$, $N \in \mathbb{Z}$, $\Delta \in \{1,2,3,4\}$ stand for the sum of weights of all the polygonal lines which have N steps, begin in any point $(i+\frac{1}{2},T)$ where $i < 0$, end in the point $(h+\frac{1}{2},t)$, and have the type Δ last step. Of course,

$$\tilde{\mu}(x^T_0 = \ldots = x^T_n = 1) \leqslant \sum_{N=1}^{\infty} \sum_{\Delta=1}^{4} \sum_{h=n}^{\infty} \sigma^N_\Delta(h+\tfrac{1}{2},T). \qquad (11.3)$$

It remains only to estimate the right-hand side. Introducing special variables $\alpha > 1$, $\beta > 0$, we may claim that the right-hand side of (11.3) does not exceed the following value:

$$\alpha^{-n} \sum_{N=1}^{\infty} \sum_{\Delta=1}^{4} \sum_{h=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} \sigma^N_\Delta(h+\tfrac{1}{2},T)\alpha^h\beta^{t-T} =$$
$$= \alpha^{-n} \sum_{N=1}^{\infty} \sum_{\Delta=1}^{4} \sum\nolimits^N_\Delta \qquad (11.4)$$

where $\sum^N_\Delta$ stand for the sums over t and h. The four-dimensional vector

$$\sum{}^N = \left(\sum\nolimits^N_1, \sum\nolimits^N_2, \sum\nolimits^N_3, \sum\nolimits^N_4\right)$$

results from the vector $\sum^{N-1}$ by multiplication by the matrix

$$M = \begin{pmatrix} \varepsilon\alpha & \varepsilon\alpha & \varepsilon\alpha & \varepsilon\alpha \\ \beta^{-1}A & \beta^{-1}A & 0 & \beta^{-1}A \\ \beta B & 0 & \beta B & 0 \\ C & 0 & C & 0 \end{pmatrix}$$

where A, B and C stand for

$$A = \sum_{i=-s}^{s} q_i^L \alpha^i,\ B = \sum_{i=-s}^{s} q_i^R \alpha^i,\ C = \sum_{i=1}^{r} \alpha^i.$$

Our proof will be completed if we find such $\alpha > 1$ and $\beta > 0$ that the series in the right-hand side of (11.4) converge. Indeed, in this case we can take n so large that (11.4) will be less than 1, whence our operator will have an invariant measure which differs from δ_1 due to Corollary 2.8.

The series in the right-hand side of (11.4) will certainly converge if all the eigenvalues of M are less than 1 in modulo. According to a corollary of the Frobenius theorem [26] the necessary and sufficient condition of it is positiveness of all the main corner minors of the matrix $E - M$. It is convenient to start from the lower right angle of the matrix when calculating them; this results in the following three conditions:

$$\begin{cases} \beta^{-1}A < 1, \\ \beta B < 1, \\ |E - M| > 0. \end{cases}$$

We have to choose $\alpha > 1$ and $\beta > 0$ to fulfil these conditions, ε being positive and as large as possible. The best is to put $\beta = \sqrt{A/B}$. Then the conditions turn into

$$\begin{cases} AB < 1, \\ \varepsilon < \dfrac{(1 - \sqrt{AB})^2}{\alpha(1 + C - AB)}. \end{cases}$$

The denominator in the last expression is certainly positive. Note that $\alpha = 1$ makes

$$AB = 1 \quad \text{and} \quad (AB)'_\alpha = \mathbb{E}\eta - \mathbb{E}\xi < 0.$$

Hence there is such $\alpha > 0$ which makes $AB < 1$. This proves Theorem 11.1. Along with that we have obtained the estimation

$$\varepsilon_W^* > \sup_{\substack{\alpha > 1 \\ AB < 1}} \frac{(1 - \sqrt{AB})^2}{\alpha(1 + C - AB)}. \tag{11.5}$$

The remaining part of this chapter applies Theorem 11.1 to prove non-ergodicity of some one-dimensional operators $P \in \mathscr{P}_1$. We shall describe a method (which is not claimed to be the best one) to majorise any given $P \in \mathscr{P}_1$ with a composition WS_ε:

$$P \prec WS_\varepsilon.$$

If this WS_ε happens to be non-ergodic, the non-ergodicity of the P in question gets proved. Without loss of generality we assume that P is defined on the graph $\Gamma(V, \mathscr{U})$ where $V = \mathbb{Z}$ and

$$U(h) = \{h-s, \ldots, h+s\}.$$

First of all we majorise P with a composition $\widetilde{P}S_\varepsilon$ where $\widetilde{P} \in \mathscr{P}_1$ too. In addition the measure δ_0 must be invariant for $\widetilde{P}$ because later $\widetilde{P}$ will be majorised by some W. First we put

$$\varepsilon = \theta^1_{0,\ldots,0} \tag{11.6}$$

and define the transitional probabilities of $\widetilde{P}$ as follows:

$$\tilde{\theta}^1_{x_{-s},\ldots,x_s} = \max\{0, (1-\varepsilon)^{-1}(\theta^1_{x_{-s},\ldots,x_s} - \varepsilon)\}. \tag{11.7}$$

It is easy to check that this $\widetilde{P}$ makes $P \prec \widetilde{P}S_\varepsilon$. But for technical reasons we have to introduce yet another operator $\widetilde{\widetilde{P}}$ which has rather different transitional probabilities:

$$\tilde{\tilde{\theta}}^1_{x_{-s},\ldots,x_s} = \begin{cases} (\tilde{\theta}_{x_{-s},\ldots,x_s})^{1/2} & \text{if} \quad x_{-s} = x_s = 0, \\ \tilde{\theta}^1_{x_{-s},\ldots,x_s} & \text{otherwise.} \end{cases} \tag{11.8}$$

Of course, we define

$$\tilde{\tilde{\theta}}^0_{x_{-s},\ldots,x_s} = 1 - \tilde{\tilde{\theta}}^1_{x_{-s},\ldots,x_s}. \tag{11.9}$$

Now we introduce

$$\left.\begin{aligned} Q^L_{k+1} &= \min_{y_1,\ldots,y_k} \prod_{i=0}^{k} \tilde{\tilde{\theta}}^0_{0,\ldots,0,1,y_1,\ldots,y_i}, \\ Q^R_{k+1} &= \min_{y_1,\ldots,y_k} \sum_{i=0}^{k} \tilde{\tilde{\theta}}^0_{y_i,\ldots,y,1,0,\ldots,0}. \end{aligned}\right\} \tag{11.10}$$

for all $k = 0, \ldots, 2s-1$, and

$$q^j_k = \begin{cases} 1 - Q^j_1 & \text{for} \quad k = s, \\ Q^j_{s-k} - Q^j_{s-k+1} & \text{for} \quad k \in [-s+1, s-1], \\ Q^j_{2s} & \text{for} \quad k = -s. \end{cases} \tag{11.11}$$

Here j stands for L (left) or R (right). The values (11.11) define the operator W_2:

$$q^{\mathrm{L}}_k = \mathbb{P}(\xi^t_{i+1/2} = k),$$
$$q^{\mathrm{R}}_k = \mathbb{P}(\eta^t_{i+1/2} = k).$$

Putting $r = 2s$ we define W_1. Thus we have defined the random walk operator

$$W = W_1 W_2.$$

Proposition 11.2 [83]. Let ε and q_k^j be defined by the formulae (11.6)–(11.11). Then $P \prec WS_\varepsilon$.

The proof is cumbersome and we omit it (see [83]). We note only that

$$y\tilde{\tilde{P}} \prec yW$$

holds for any state y in which all the massifs of ones have length no less than $2S$. However, this does not hold for all states. But, due to the difference between $\tilde{P}$ and $\tilde{\tilde{P}}$

$$y\tilde{P} \prec yW$$

holds for all $y \in X$. Thus $\tilde{P} \prec W$ whence

$$P \prec \tilde{P}S_\varepsilon \prec WS_\varepsilon$$

q.e.d.

The following simple example shows that the results of this chapter really have extended the range of provable non-ergodicity as compared with results which had been available before.

Example 11.3 A homogeneous independent operator P on $\{0;1\}^{\mathbb{Z}}$ with

$$U(h) = \{h-1, h, h+1\}$$

has transitional probabilities

$$\theta^1_{x_{-1},x_0,x_1} = \begin{cases} \frac{1}{3} & \text{if} \quad x_{-1} = x_0 = 0,\ x_1 = 1, \\ 1 & \text{if} \quad x_{-1} = x_0 = 1, \\ 0 & \text{in the other cases.} \end{cases}$$

This operator P is majorable with the random walk operator W with $s = 1$ and

$$q_1^{\mathrm{L}} = \tfrac{1}{3},\ q_0^{\mathrm{L}} = 0,\ q_{-1}^{\mathrm{L}} = \tfrac{2}{3},\ q_1^{\mathrm{R}} = 0,\ q_0^{\mathrm{R}} = 1,\ q_{-1}^{\mathrm{R}} = 0$$

which has

$$\mathbb{E}\xi = -\tfrac{1}{3};\ \mathbb{E}\eta = 0.$$

Thus any composition PS_ε with sufficiently small $\varepsilon > 0$ (say $\varepsilon < 10^{-4}$) is non-ergodic.

Chapter 12
Some counter-examples

After reading the preceding chapters, a usual question naturally arises: what generalisations of the results presented are possible through weakening of their conditions. This chapter gives three counter-examples which show that some generalisations are impossible. Proposition 9.2 and Theorem 11.1 suggest certain connections between an operator P's being an eroder and non-ergodicity of PS_ε with some $\varepsilon > 0$. The following examples breach this connection in various ways. The first example presents a non-eroder P that has nevertheless an $\varepsilon > 0$ that makes PS_ε non-ergodic. The same example shows the possibility of non-uniqueness of critical value.

The second example is an analogue of the third one but every automaton has three states which allows us to use the simplest case $V = \mathbb{Z}$.

The third example presents an eroder P on $\{0;1\}^V$ such that PS_ε are ergodic with all $\varepsilon > 0$. This does not contradict Theorem 11.1 because the graph of the third example is not $\mathbb{Z}^d$.

Example 12.1 (I. I. Piatetski-Shapiro, L. G. Mityushin and A. L. Toom). There is a one-dimensional non-monotone deterministic operator D on $x = \{0;1\}^{\mathbb{Z}}$ which has the following properties. First, D is a non-eroder but has some $\varepsilon > 0$ (for example, $\varepsilon = \frac{1}{2}$) which makes DS_ε non-ergodic. This shows that the condition of monotonicity is essential in Proposition 9.2. Second, D has no less than two critical values of ε. This is because DS_ε is non-ergodic with $|\varepsilon - \frac{1}{2}|$ small enough and is ergodic with ε near 0 or 1. So the uniqueness of critical value which has been proved for monotone operators does not hold in general.

The main idea is the following. The function f equals 1 if the sequence of values of its arguments is sufficiently ordered in some sense. That is why D^2 turns any state into the state 'all ones'. Almost the same occurs in the 'approximately deterministic' cases. In the contrary case $\varepsilon \approx \frac{1}{2}$ 'ordered' sequences are rare and always remain rare.

Now for exact definitions. Take a finite sequence $x_1, \ldots, x_n$ of zeros and ones; we term it a 'chaos' if there are such a, b, c that $1 \leqslant a < b < c \leqslant n$ and $x_a = 1, x_b = 0, x_c = 1$. In other words, a sequence is a chaos if it has at least two massifs of ones.

Now put

$$U(h) = \{h-2m+1,\ldots,h+m-1\}$$

where m is some natural number. The exact value of m will be chosen later. Define $f: x_{U(h)} \to x_h$ as follows:

$$f(x_{-2m+1},\ldots,x_{m-1}) = \begin{cases} 0 & \text{if there is such } i \in [-2m+1,0] \\ & \text{that the sequence } x_i, \ldots, x_{i+m-1} \text{ is a chaos,} \\ 1 & \text{otherwise.} \end{cases}$$

Thus D is defined.

Note that $f(1,\ldots,1) = 1$, whence $\mathbf{1}D = \mathbf{1}$ and the measure $\delta_{\mathbf{1}}$ is invariant for DS_ε with any ε. Proposition 2.17 proves ergodicity of DS_ε with

$$\varepsilon > 1 - \frac{1}{|U(h)|} = 1 - \frac{1}{3m-1}.$$

You can check that D is a non-eroder.

Let us prove that , $|\varepsilon - \frac{1}{2}|$ being small enough, DS_ε has another invariant measure besides $\delta_{\mathbf{1}}$. This needs m large enough. In this proof we take $m = 10$. We separate $\mathbb{Z}$ into segments:

$$[10k+1;10k+10],\ k \in \mathbb{Z}$$

of length 10, which we shall term 'blocks'. Let Q stand for the percolation operator on Γ_1. We have proved its non-ergodicity with small enough $\varepsilon > 0$ in Theorem 8.4. Now we are going to use it to prove that the proportion of blocks being in the state of a 'chaos' does not tend to 0 as $t \to \infty$. For that we define a mapping $F\colon X \to X$ as follows:

$$(xF)_i = \begin{cases} 0 & \text{if } x_{10i+1}, \ldots, 10i + 10 \text{ is a chaos,} \\ 1 & \text{otherwise.} \end{cases}$$

First take $\varepsilon = \frac{1}{2}$ for simplicity. Suppose that at least one of the blocks $y_{10(i-1)+1}, \ldots, y_{10(i-1)+10}$ or $y_{10i+1}, \ldots, y_{10i+10}$ is a chaos in some state $y \in X$. Then yD certainly has zeros at places $10i+1, \ldots, 10i+10$. Then the measure $yDS_{1/2}$ makes all the 2^{10} states at these places equiprobable. Direct counting shows that only 65 of these 2^{10} states are not chaoses. Note that

$$65/2^{10} < 0.07 < \varepsilon^*_{\Gamma_1}.$$

In fact we have proved

$$DS_{1/2}F \prec FQ_{0.07}.$$

Hence it follows, using Proposition 2.14, that any initial state y, in which all blocks are chaoses, generates measures

$$y(DS_{1/2})^t$$

in which the proportion of chaoses is not less than

$$\mathbf{0}(Q_{0.07})^t(x_i = 0) \tag{12.1}$$

which is known to exceed a positive constant c at all times $t \in \mathbb{Z}_+$. Hence, according to Corollary 2.8, $DS_{1/2}$ has an invariant measure μ in which

$$\mu(x_{10i+1}, \ldots, x_{10i+10} \text{ is a chaos}) \geqslant c.$$

Thus $\mu \neq \delta_{\mathbf{1}}$. q.e.d. With $|\varepsilon - \frac{1}{2}|$ small enough the same argument works.

Now let us prove that DS_ε is ergodic with ε small enough. Note that all sequences of zeros in any xD are not shorter than $2m$. Hence any sequence

$$(XD)_h, \ \ldots, \ (xD)_{h+m-1}$$

is a non-chaos, whence $xD^2 = \mathbf{1}$. Now we consider the evolution measure generated by DS_ε and any initial state x^0. As before, we introduce auxiliary variables ω^t_h which control the action of S_ε at time t, and $\tilde{\mu}$ is induced by their independent distribution.

Lemma 12.2 Let us have an $\omega \in \Omega$ which makes $x^t_0 = 0$. Then there is some h in the range

$$-2m + 1 \leqslant h \leqslant m - 1,$$

where

$$x^{t-1}_h = 0 \quad \text{and} \quad \omega^{t-1}_h = 1.$$

Proof. Assume the contrary:

$$\omega^{t-1}_h \leqslant x^{t-1}_h \quad \text{for all} \quad h \in [-2m+1, m-1].$$

Then

$$x^{t-1}_h \vee \omega^{t-1}_h = x^{t-1}_h$$

in the same range of h. Then

$$x^t_0 = f(x^{t-1}_{-2m+1}, \ldots, x^{t-1}_{m-1}) \vee \omega^t_0$$

which is no less than

$$(x^{t-2}D^2)_0$$

which equals 1 as we know.

Lemma 12.3 Let us have an $\omega \in \Omega$ which makes $x_0^T = 0$. Then there is a sequence of integers $h_1, \ldots, h_T$, where $h_T = 0$, such that

$$h_{t+1} - 2m + 1 \leqslant h_t \leqslant h_{t+1} + m - 1$$

and

$$\omega_{h_t}^t = 1$$

for all $t = 1, \ldots, T - 1$.

Proof is made by induction with Lemma 12.2 serving as the induction step.

Now we are ready to prove ergodicity of DS_ε with small ε. In fact, we shall prove

$$\lim_{T\to\infty} \tilde{\mu}(x_0^T = 0) = 0.$$

Lemma 12.3 helps us to estimate $\tilde{\mu}(x_0^T = 0)$. In fact, the sequence $h_1, \ldots, h_T$ being fixed, the probability of

$$\omega_{h_1}^1 = \ldots = \omega_{h_{T-1}}^{T-1} = 1$$

is ε^{T-1}. There are $(3m - 1)^{T-1}$ such sequences. Thus

$$\tilde{\mu}(x_0^T = 0) \leqslant [(3m - 1)\varepsilon]^{T-1}.$$

With $\varepsilon < \dfrac{1}{3m - 1}$ the right-hand side tends to 0. q.e.d.

Example 12.4 [88]. Every automaton has three states 0, 1, 2. They form a line $\mathbb{Z}$ and

$$X = \{0;1;2\}^{\mathbb{Z}}.$$

To define a deterministic operator D: $X \to X$ we put

$$U(h) = \{h-1;h;h+1\}$$

and f: $\{0,1,2\}^3 \to \{0,1,2,\}$, which is

$$f(x,y,z) = \begin{cases} 0 & \text{if} \quad x = 0,\ y = z = 1, \\ 1 & \text{if} \quad x = y = 2,\ z = 1, \\ 2 & \text{if} \quad x + y = z = 2, \\ \text{the nearest integer to } (x + y + z)/3 \text{ in the other cases.} \end{cases} \qquad (12.2)$$

Let us mention some properties of this function. First, f is monotone in the sense

$$x \leqslant x',\ y \leqslant y',\ z \leqslant z' \Rightarrow f(x,y,z) \leqslant f(x',y',z').$$

Second, D is an eroder, in the sense that every state having only a finite set of non-zero components is turned into 'all zeros' by some degree of D. Let us prove it. Due to monotonicity we need only to apply degrees of D to

$$\ldots 0001222 \ldots 2221000 \ldots$$

The second line in the right-hand side of (12.2) turns 'twos' into 'ones' one after the other from right to left, which results in

$$\ldots 000111 \ldots 111000 \ldots$$

Now the first line in the right-hand side of (12.2) turns 'ones' into 'zeros' one after the other from left to right, which results in $\mathbf{0}$.

The noise operator $S_{\gamma,\delta,\varepsilon}$ which is appropriate here transforms states of single automata independently with the probabilities shown in Table 12.1. We assume that $\gamma + \delta > 0$ and $\varepsilon > 0$. Then the composition $DS_{\gamma,\delta,\varepsilon}$ is ergodic and its only invariant measure is concentrated in the state 'all twos'. Thus the direct generalisation of our Theorem 10.1 to the three-states automata is not true.

Let us try to explain informally why $DS_{\gamma,\delta,\varepsilon}$ is ergodic. Suppose we are iterating this operator with the initial state having a finite but large segment filled with 'twos', the other components being zeros. As we have said, D turns 'twos' into 'ones' one after the other from right to left. But $S_{\gamma,\delta,\varepsilon}$ turns a few of them back into 'twos'. Due to the third line in the right-hand side of (12.2) these 'twos' move one place left at every unit of time. Now $S_{\gamma,\delta,\varepsilon}$ continues to work and the 'twos' get thicker. Near the left end the probability of 1 is about $(1 - \varepsilon)^n$ where n is the length of the non-zero blot. Moreover, n grows linearly with t because the left end of the blot drifts to the left due to $S_{\gamma,\delta,\varepsilon}$. Estimation shows that no 1 will reach the left end of the blot with some positive probability. Moreover, this probability tends to 1 with the growing of the initial blot. Thus such a blot is a stochastic analogue of a non-erodable island.

Table 12.1

from \ to	0	1	2
0	$1 - \gamma - \delta$	γ	δ
1	0	$1 - \varepsilon$	ε
2	0	0	1

Example 12.5 [88]. Unlike typical constructions in our survey, V is not an integer lattice here. Instead

$$V = \{h\},\ h = (h^1,h^2),\ h^1 \in \mathbb{Z},\ h^2 \in \{-1,1\}.$$

You may imagine V as two parallel integer lines. But our graph Γ and operator D will be homogeneous under the transitive group G of automorphisms which consists of shifts

$$g\colon (h^1,h^2) \to (h^1+\text{const},h^2)$$

and 'inverted shifts'

$$g\colon (h^1,h^2) \to (-h^1+\text{const},-h^2).$$

As we presented in Chapter 9, this G is not commutative; this is essential because the case of commutative G boils down to the $\mathbb{Z}^d$ case which is exhausted (assuming $\{0,1\}$ and monotony) by our Theorem 10.1.

Now we define D. For any $h = (h^1,h^2)$

$$U(h) = \{U_1(h),U_2(h),U_3(h)\}$$

where

$$U_1(h) = h,\ U_2(h) = (h^1,-h^2),\ U_3(h) = (h^1-h^2,h^2).$$

Finally, $X_h = \{0;1\}$ for any h and

$$f(x,y,z) = (x \vee y) \wedge z,$$

which you may take as $f(x,y,z) = \min(\max(x,y),z)$.

Evidently, f is monotone and

$$f(1,1,1) = 1.$$

So D is monotone and $\mathbf{1}D = \mathbf{1}$.

To see that D is an eroder it is sufficient to examine what D^t makes out of the state

$$\begin{matrix} \ldots\ 000111\ \ldots\ 111000\ \ldots \\ \ldots\ 000111\ \ldots\ 111000\ \ldots \end{matrix}$$

or

$$x_h = \begin{cases} 1 & \text{if } \ a \leqslant h^1 \leqslant b, \\ 0 & \text{otherwise.} \end{cases}$$

This island is eroded in $b - a + 1$ units of time. But DS_ε is ergodic with any positive ε which is proved along the same lines as in Example 12.4.

Chapter 13
Quasi-non-ergodicity in the one-dimensional case

In Chapter 10 we have constructed simple enough non-ergodic non-degenerate homogeneous independent operators with $V = \mathbb{Z}^d$, where d was any natural number other than 1. It has been and remains a very intriguing question whether these are possible in the one-dimensional case.

In [14] it was conjectured that all non-degenerate homogeneous independent operators with $X = X_0^V$ where X_0 is finite and $V = \mathbb{Z}$, are ergodic. The computer simulation of one-dimensional voting operators (our Example 1.1) was an attempt to refute this conjecture, but it failed.

Later, Tsirelson constructed a hierarchical system of unreliable elements which did not forget its initial state [94,95]. However, this system was non-homogeneous both in space and time, and neither did it refute the conjecture.

Later, Kurdyumov published a paper [43] in which he claimed to have constructed a suitable non-ergodic operator thus having refuted the conjecture. But his description and proof were informal and never achieved a formal status.

Later Gács published a very complicated construction [21] which he claimed to be a non-ergodic operator refuting the conjecture. We have not yet examined Gács's construction sufficiently to subscribe to its validness, which is due partly to its excessive complication and partly to some unclear points in his description. Thus we shall not attempt to present it here.

Our task is more modest. All the constructions we present here are simple enough. They probably do not refute the conjecture, after all. But we think they are provocative.

First we are going to describe three one-dimensional deterministic operators having the common property of being eroders and remaining eroders after renaming 0 into 1 and 1 into 0 in $X_h = \{0;1\}$. So these operators blot out any finite perturbation of the states **0** and **1**. These operators with small independent noise S_ε^+ added were computer simulated. It is interesting that within the computer time available they behaved as if they were non-ergodic. This does not imply that they actually are non-ergodic. It is quite plausible (but not proved) that they are ergodic with any $\varepsilon \in (0;1)$.

In fact, multi-processing has shown their convergence. But unusually slow convergence is interesting as such too. It may have applications, including biological ones.

Now for exact definitions. As before, our state space is $X = X_0^{\mathbb{Z}}$ where X_0 is the finite set of every single automaton's states. A deterministic operator is a mapping D: $X \to X$. A state x is termed *stationary* for D if $xD = x$. A stationary state is termed *attractive* if for any y which differs from x only in a finite set of components there is such t that

$$yD^t = x.$$

In particular, D is an eroder if **0** is attractive for D. Following [22], a *conservator* is an operator which has at least two different periodic attractive states. It is convenient here to use more mnemonic symbols to denote elements of X_0. In all cases we call a finite perturbation of a stationary state 'an island'. The subset of $\mathbb{Z}$ where components differ from the stationary state components will be called an island too, the rest of $\mathbb{Z}$ being called 'the sea'. We shall also speak about the two boundaries of an island (its leftmost and rightmost perturbated points), and about the length of an island (which is the distance between the boundaries).

Example 13.1 We consider an operator, D_{VI}. [22]. Let

$$X_0 = \{+,-,\rightarrow,\leftarrow,\triangleright,\triangleleft\},$$
$$U(h) = \{h-1,h,h+1\}.$$

The function f: $X_{U(h)} \to X_h$ is defined by Table 13.1 where y and z are arbitrary elements of X_0. Numbering of the lines is essential in this table because some lines contradict each other; in these cases the lower numbered line prevails.

Under this operator the states 'all pluses' and 'all minuses' are attractive.

Note the way in which D_{VI} erodes a finite sequence of n minuses surrounded by pluses. The first moments at the islands' ends appear as $\triangleright$ and $\triangleleft$ directed from plus to minus. They turn into $\rightarrow$ and $\leftarrow$ every of which proliferates to both sides with the speed $\frac{1}{2}$. After n units of time, $\rightarrow$ and $\leftarrow$ massifs meet in the middle and produce a + there which prolifer-

Table 13.1

Number of the line	x_{h-1}^t	x_h^t	x_{h+1}^t	x_h^{t+1}
1	y	z	y	y
2	$+$	z	$-$	$\triangleright$
3	$-$	z	$+$	$\triangleleft$
4	y	$\triangleright$	z	$\rightarrow$
5	y	$\triangleleft$	z	$\leftarrow$
6	$\leftarrow$	z	$\rightarrow$	$-$
7	$\rightarrow$	z	$\leftarrow$	$+$
8	$\rightarrow$	$-$	z	$\triangleright$
9	z	$-$	$\leftarrow$	$\triangleleft$
10	$\leftarrow$	$+$	z	$\triangleleft$
11	z	$+$	$\rightarrow$	$\triangleright$
12	$-$	$\rightarrow$	z	$-$
13	z	$\leftarrow$	$-$	$-$
14	z	$\rightarrow$	$+$	$+$
15	$+$	$\leftarrow$	z	$+$

ates to both sides with the speed 1, thus catching up with the arrows. The total time of erosion of this island is $2.5n$ units of time.

Example 13.2 Operator D_{IV} [22]. It has

$$X_0 = \{+,-,\rightarrow,\triangleright\},\ U(h) = \{h-1,h,h+1\}.$$

Its function is given by Table 13.1 without the lines which contain two omitted symbols $\leftarrow$ and $\triangleleft$.

Proposition 13.3 The states 'all pluses' and 'all minuses' are attractive for D_{IV} and D_{VI} of Examples 13.2 and 13.1.

The proofs are analogous for the two operators. Let us prove that 'all pluses' is attractive for D_{IV}. Take any finite perturbation x^0 of it with length n. For any island x (finite perturbation of 'all pluses') $L(x)$ and $R(x)$ will stand for the leftmost and rightmost points of $\mathbb{Z}$ whose states are not pluses. Thus

$$n = R(x^0) - L(x^0).$$

Let us show that already $x^0 D_{\text{IV}}^{6n}$ is 'all pluses'. Let $M(x^0)$ stand for the leftmost point of $\mathbb{Z}$ occupied with a minus (if such a point exists). One can check the following:

(1) $R(xD_{\mathrm{IV}}) \leqslant R(x)$.
This means that the island never expands to the right.

(2) $L(xD^2_{\mathrm{IV}}) \geqslant L(x) - 1$.
This means that the island expands to the left with a speed not higher than 1/2.

(3) $M(xD^2_{\mathrm{IV}}) \geqslant M(x) + 1$, if these are defined.
This means that the minuses retreat on their left-hand end with a speed not less than 1/2. So after $2n$ units of time there will be no more minuses.

(4) If there are no minuses,

$$R(xD_{\mathrm{IV}}) < F(x).$$

This means that the right-hand end of the island moves left with a speed not less than 1.

So the erosion process goes as follows: first minuses die out in $2n$ units of time. The island is not longer than $2n$ by this time. Then the island's right-hand end catches up with the left end in no more than $4n$ units of time. The total time of erosion is no more than $6n$. q.e.d.

Example 13.4 Operator D_{II} [22]. Here

$$X_0 = \{\rightarrow,\leftarrow\},$$
$$U(h) = \{h-3, h-1, h, h+1, h+3\}.$$

The function f is defined as follows. If $x^t_h = \rightarrow$, the value of x^{t+1}_h coincides with the value of the majority of

$$x^t_{h-3},\ x^t_{h-1},\ x^t_h.$$

If $x^t_h = \leftarrow$ then the value of x^t_h coincides with the value of the majority of

$$x^t_h,\ x^t_{h+1},\ x^t_{h+3}.$$

Here 'all $\rightarrow$' and 'all $\leftarrow$' are attractive too, which is proved along the same lines.

Now we introduce the noise. Let N_α stand for the independent operator which acts upon every automaton independently. Every automaton either keeps its state with probability $1 - \alpha$, or goes into all other states with equal probabilities. Behaviour of $D_{\mathrm{II}}N_\alpha$, $D_{\mathrm{IV}}N_\alpha$, $D_{\mathrm{VI}}N_\alpha$ was computer simulated in [22].

In fact, a finite number (2000–5000) of automata, which formed a ring, functioned during 1000–2000 units of model time (that is, applications of the operator in question). The usual pseudo-random numbers were used. The proportions of states were outputed from the computer.

Figure 13.1 shows plots of Q_+ and Q_- which are proportions of pluses

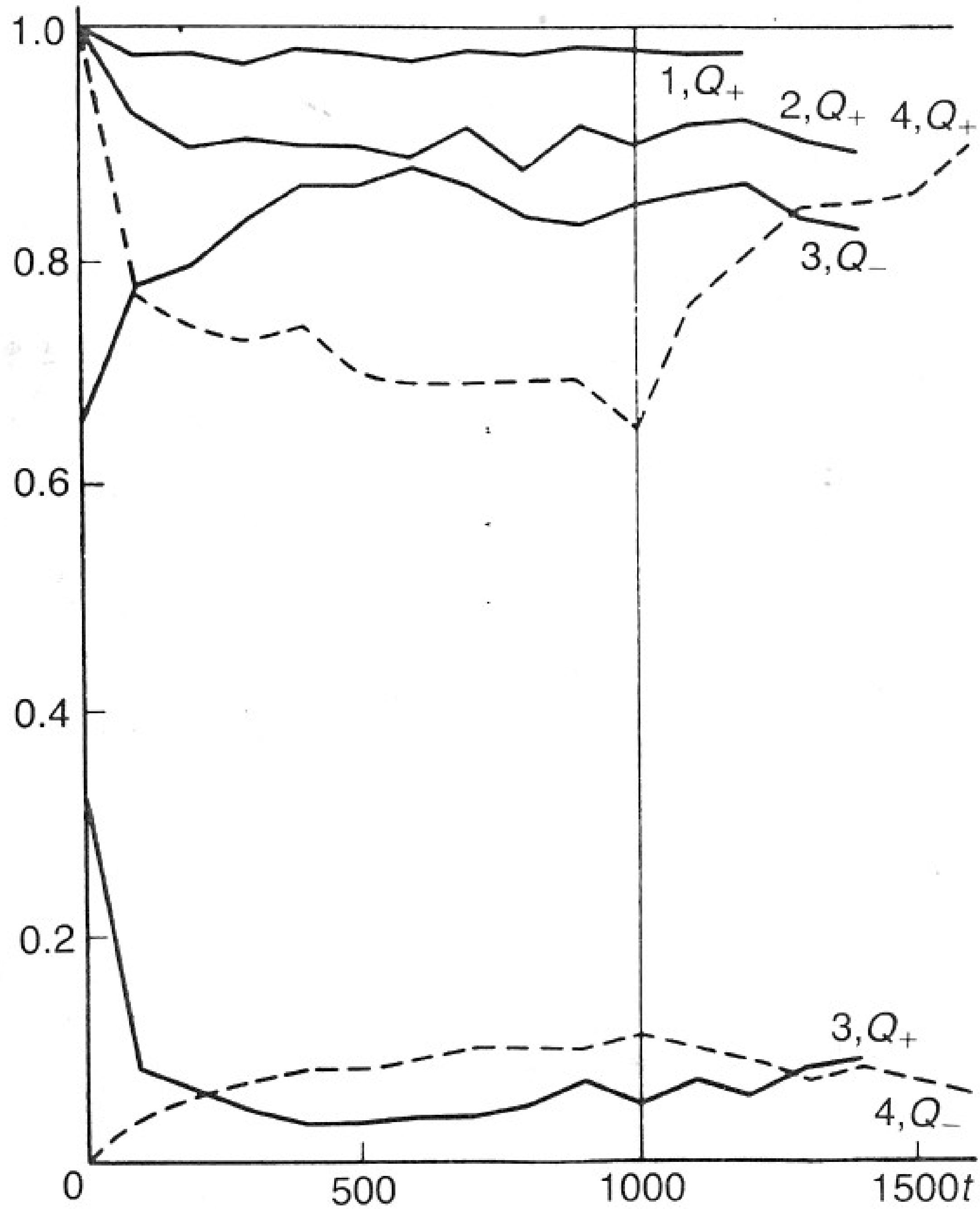

Fig. 13.1 Proportions of automata Q_+ and Q_- being in the state + or − depending on time t.

and minuses in the action of $D_{\mathrm{VI}}N_\alpha$ with various values of parameters given in Table 13.2. Due to the symmetry of D_{VI} and N_α, the ergodicity needs the equality of limits of Q_+ and Q_- as $t \to \infty$. So, during the simulation, they are expected to approach each other. In fact they behaved quite differently (see Table 13.2).

In experiments 1 and 2, the initial state was 'all pluses'. All values of Q_+ output (which was done after every 100 units of time) exceeded 0.85. The difference in the number of automata seemed not to influence the data.

Table 13.2 Parameters of $D_{\mathrm{VI}}N_\alpha$ simulation.

Number of the experiment	Number of automata	Initial state	α
1	5000	$Q_+ = 1$	0.01
2	2000	$Q_+ = 1$	0.03
3	2000	$Q_+ = 1/3$ $Q_- = 2/3$	0.03
4	5000	$Q_+ = 1$	When $t \leqslant 1000$ $\alpha = 0.05$ when $t > 1000$ $\alpha = 0.01$

In experiment 3, the initial state was a random realisation of the Bernoulli measure with proportion of pluses $\frac{1}{3}$ and proportion of minuses $\frac{2}{3}$. Here, Q_- grew in the first 300 units of time and remained more than 0.8 all the subsequent time. But in this case Q_- was less than Q_+ was in experiment 2. This seems to be due to some large island which happened to occur in the initial random sequence and was hard to erode.

In experiment 4 the initial state 'all pluses' was first deliberately spoilt by action of $D_{\mathrm{VI}}N_\alpha$ with $\alpha = 0.05$ during 1000 units of time. In this period of time Q_+ clearly decreased, which may be interpreted as an argument for the ergodicity of $D_{\mathrm{VI}}N_{0.05}$. But when the value of α was made 0.01 the Q_+ sharply increased. We see in this a hint that $D_{\mathrm{VI}}N_\alpha$ possesses some property of 'quasi-non-ergodicity'.

The $D_{\mathrm{VI}}N_\alpha$ with $\alpha = 0.15$ was also simulated with the initial state 'all pluses'. In this case Q_+ and Q_- were almost equal after only 100 units of time and in the subsequent simulation their plots crossed each other many times.

Thus the results of simulation suggest that $D_{\mathrm{VI}}N_\alpha$ behaves differently with different values of α. With $\alpha \geqslant 0.5$ it converges quickly but with $\alpha \leqslant 0.03$ it converges (if it does) so slowly that our computer simulation did not show it.

It is natural to think that large islands, once having randomly emerged, play an important part in the process of convergence. This has led us to simulate $D_{\mathrm{VI}}N_{0.03}$ with the following initial state:

$$x_h = \begin{cases} - & \text{for} \quad h \in [10;1000] \\ + & \text{for all other } h. \end{cases}$$

In spite of many local distortions, this island remained of about the same length for the whole 1000 time units of its simulation.

The operators $D_{II}N_\alpha$ and $D_{IV}N_\alpha$ were simulated too and showed the same properties still more. The $D_{IV}N_{0.05}$ was simulated with the initial state 'all →' during 2000 units of time. All this time $Q_\rightarrow$ was no less than 0.78 and had already stopped decreasing at $t = 100$. In the simulation of $D_{II}N_{0.03}$ the initial island of length 100 was eroded in the first 450 units of time, after which the values of $Q_\rightarrow$ remained more than 0.95 all the time.

All the simulations mentioned above used symmetric noise. We restrict our conjecture of the quasi-non-ergodicity to this case. Plausible reasoning suggests that in the presence of a non-symmetric noise the convergence of our operators is the usual quick one.

Note that the deterministic part of our Example 1.1a is not a conservator; so it is not surprising that this example converged in the usual quick way.

But let us consider the following example which is a generalisation of the voting operator of Example 1.1a.

Example 13.5 Operator $\Pi_{\alpha,\gamma}$ where α and γ are non-negative parameters, $\alpha + \gamma \leqslant 1$. It has

$$X_0 = \{0,1\},\ V = \mathbb{Z},\ U(h) = \{h-1,h,h+1\}.$$

Instead of directly writing down the transitional probabilities we shall use independent auxiliary variables ω_h, each of which has the following distribution:

$$\omega_h = \begin{cases} 1 & \text{with probability} & \alpha/2, \\ 2 & \text{with probability} & \alpha/2, \\ 3 & \text{with probability} & \gamma, \\ 4 & \text{with probability} & (1-\alpha-\gamma)/3, \\ 5 & \text{with probability} & (1-\alpha-\gamma)/3, \\ 6 & \text{with probability} & (1-\alpha-\gamma)/3. \end{cases}$$

These control the action of $\Pi_{\alpha,\gamma}$ in the following way:

$$(y\Pi_{\alpha,\gamma})_h = \begin{cases} 1 \text{ if } \omega_h = 1, \\ 0 \text{ if } \omega_h = 2, \\ \text{the nearest integer to } (y_{i-1} + y_i + y_{i+1})/3 \quad \text{if} \quad \omega_h = 3, \\ y_{i-1} \quad \text{if} \quad \omega_h = 4, \\ y_i \quad \text{if} \quad \omega_h = 5, \\ y_{i+1} \quad \text{if} \quad \omega_h = 6. \end{cases}$$

In other words, the projection of $y\Pi_{\alpha,\gamma}$ to the space X_h is a mixture of six measures, three of which are arguments, two others are constants, and the sixth one is the corresponding measure of the voting operator. In the case $\alpha + \gamma = 1$ our $\Pi_{\alpha,\gamma}$ turns into this voting operator (Example 1.1a) with $\varepsilon = \alpha/2$.

This operator was computer simulated in the following three cases:

(1) $\alpha = 0.1$, $\gamma = 0.6$,
(2) $\alpha = 0.05$, $\gamma = 0.35$,
(3) $\alpha = 0.01$, $\gamma = 0.06$,

with a ring of 2000 automata up to $t = 2100$. In all these cases convergence was quick and evident.