Characterizing Stable States of Perturbed Markov Chains *

Jeffrey C. Ely †

November 12, 1998

*Department of Economics, Northwestern University. ely@nwu.edu

†Financial assistance from the Alfred P. Sloan foundation is gratefully acknowledged. Any mistakes in the paper remain despite, and not because of, the helpful comments from Bob Anderson, Paolo Ghirardato, Matthew Rabin, and Chris Shannon.

1 Introduction

Many models of evolution are based on deterministic dynamic processes (natural selection) augmented by stochastic mutation. The original papers in this literature were Kandori, Mailath, and Rob (1993) and Young (1993), many subsequent papers have built on the basic techniques developed there.

These models give rise to Markov chains on the set of population strategy profiles. These Markov chains are parameterized by a noise parameter, ε, the probability of a mutation. Young (1993) called these parameterized families of Markov chains *regular perturbations*. In order to characterize the asymptotic relative frequencies when mutations are rare, the limit properties of the Markov chain is analyzed as ε is taken to zero. Freidlin and Wentzell (1984) presented a way of calculating these limit relative frequencies. In this paper, the Friedlin and Wentzell technique is reviewed and a new algorithm for this calculation, called the *order decomposition* is presented. This algorithm was used in Ely (1996).

Throughout this chapter we will be considering finite-state discrete-time Markov chains. The set of states will be V and we will identify a Markov chain with its transition matrix P, denoting by $P(v_1, v_2)$ the Markovian probability of transition from state v_1 to state v_2.

2 Regular Perturbations

Definition 1 *Let P be a Markov Chain. A* **perturbation** *of P with noise parameter ε is a class of Markov Chains $\{P_\varepsilon\}_{\varepsilon \in (0,a)}$ with $0 < a < \infty$ such that*

- $\lim_{\varepsilon \to 0} P_\varepsilon = P := P_0$
- *If $P_\varepsilon(v_1, v_2) > 0$ for some $\varepsilon \in (0, a)$ then there exist real numbers $C(v_1, v_2) \in [0, \infty), k(v_1, v_2) \in (0, \infty)$ s.t.*

$$k(v_1, v_2) = \lim_{\varepsilon \to 0} \varepsilon^{-C(v_1, v_2)} P_\varepsilon(v_1, v_2)$$

If in addition P_ε is regular for every $\varepsilon \in [0, a)$, then $\{P_\varepsilon\}_{\varepsilon \in [0,a)}$ is called a **regular perturbation**.

In the case of probabilities $P_\varepsilon(v_1, v_2)$ which are zero on $[0, a)$ set $C(v_1, v_2) = \infty$. Now the following simple fact motivates a classification of states.

Proposition 1 *Let $v_1, v_2 \in V$. Then $C(v_1, v_2) = 0$ iff $P_0(v_1, v_2) > 0$.*

Proof: The if part follows immediately from the definitions, to show the only if part, we have $P_0(v_1, v_2) = \lim_{\varepsilon\to 0} P_\varepsilon(v_1, v_2) = \lim_{\varepsilon\to 0} \varepsilon^{-C(v_1,v_2)} P_\varepsilon(v_1, v_2) = k(v_1, v_2) > 0$ where the last inequality follows by definition since $C(v_1, v_2) < \infty$. ∎

Accordingly we have the following definitions

Definition 2 *If* $C(v_1, v_2) \in (0, \infty)$ *then* $P_\varepsilon(v_1, v_2)$ *is a* **vanishing probability**. *If* $C(v_1, v_2) = 0$, *then* $P_\varepsilon(v_1, v_2)$ *is called a* **non-vanishing probability**. *The values* $C(v_1, v_2)$ *and* $k(v_1, v_2)$ *are the* **order** *and* **rate** *of the probability* $P_\varepsilon(v_1, v_2)$.

A perturbation of P is a class of "neighboring" Markov chains. In the case of a regular perturbation of a non-regular underlying chain P, all chains in the class are regular and hence, perhaps unlike P, admit characterization by a unique invariant distribution μ_ε. But while the map $\varepsilon \to P_\varepsilon$ is continuous at zero, the correspondence $P_\varepsilon \to \mu_\varepsilon$ need not be. That is μ_ε is unique for every $\varepsilon \in (0, a)$, there is a continuum of invariant distributions of $P \equiv P_0$ if it is not regular. This property being non-generic in the set of Markov chains, we can interpret a regular perturbation as an approximation of P by a neighboring class of generic chains. As the following proposition formalizes, the purpose of this approximation is to select a particular element of μ_0. The suggested interpretation is that none of the remaining elements of μ_0 are robust to the perturbation.

Proposition 2 *Let* P_ε *be a regular perturbation of a Markov chain* P. *For every* $\varepsilon \in (0, a)$ *let* μ_ε *be the invariant distribution of* P_ε. *Then*

- $\mu^* := \lim_{\varepsilon\to 0} \mu_\varepsilon$ *exists and*
- $\mu^* \in \mu_0$

Naturally, then, we have the following definition.

Definition 3 *Let* P_ε *be a regular perturbation. The* **limit distribution** *is given by* $\mu^* = \lim_{\varepsilon\to 0} \mu_\varepsilon$. *A state* v *is* **strongly recurrent** *if* $\mu^*(v) > 0$.

The set of strongly recurrent states will be denoted $\mathcal{R}$. The technique of characterizing $\mathcal{R}$ involves a combinatorial argument due to Friedlin and Wentzell (Freidlin and Wentzell 1984). In the next section we introduce some background definitions which will be used in the statement of this and subsequent results.

3 Preliminaries

3.1 Graphs

A graph G consists of a finite set V of *vertices*, and a collection E of ordered pairs of elements of V, called edges.[1] Edges $e = (v_1, v_2) \in E$ of are interpreted as arrows drawn from the point v_1 to the point v_2. v_1 and v_2 are the *origin,* and *endpoint* of the edge e, respectively.

Given a graph $G = (V, E)$, a *subgraph* of G is a (V, E') where $E' \subset E$. Since the underlying set V is assumed fixed throughout, it is sufficient to identify a subgraph simply by its edges.

Definition 4 *A* **path** *in G from v_1 to v_n is a subgraph $\hat{E} = \{(v_{i-i}, v_i)\}_{i=2}^{n}$ where all the v_i are distinct. We say v_1 and v_n are the* **origin** *and* **endpoint** *of $\hat{E}$, respectively.*

We will write $x \succeq_G y$, read "x succeeds y in G," to indicate that there exists a path in G from y to x. Denote by $\mathcal{P}_G(x, y)$ the set of paths in G from x to y.

3.2 Markov Graphs

A Markov chain P generates a graph on V when E is taken to be the set of all transitions which have positive probability. In the case of a perturbation of P, include as edges all transitions (v_1, v_2) such that $C(v_1, v_2) < \infty$.

Definition 5 *The graph of a perturbed Markov chain P_ε is a graph on the state space V with edge set given by $E = \{(v_1, v_2) : C(v_1, v_2) < \infty\}$. When G is derived from a perturbed Markov chain, G is called a* **perturbed Markov graph**.

When V is fixed, a graph is simply a collection of edges. The order and rate can be viewed as functions whose domain is the set of edges of a perturbed Markov graph. The following demonstrates how these functions can be extended to the domain of subgraphs, i.e. subsets of edges.

Definition 6 *Let t be a subgraph of a perturbed Markov graph. Define the* **order** *and* **rate** *of t respectively as follows:*

- $C(t) = \sum_{e \in t} C(e)$

[1] Technically, when the edges in E are ordered pairs, the graph is known as a *directed graph.*

- $k(t) = \prod_{e \in t} k(e)$.

The following simple result demonstrates that these define consistent extensions of $C()$ and $k()$ to the domain of subgraphs.

Lemma 1 *Let t be a subgraph of a perturbed Markov graph and define $P_\varepsilon(t) = \prod_{e \in t} P_\varepsilon(e)$. Then if $P_{\bar{\varepsilon}}(t) > 0$ for some $\bar{\varepsilon} \in (0, a)$ then for every $0 \le x < \infty$*

$$\lim_{\varepsilon \to 0} \varepsilon^{-x} P_\varepsilon(t) = \begin{cases} 0 & \text{if } x < C(t) \\ k(t) & \text{if } x = C(t) \\ \infty & \text{if } x > C(t) \end{cases} \tag{1}$$

Remark: The "probability" $P_\varepsilon(t)$ is easiest to interpret when t is a path. In that case, we can view the path as a single extended transition (e.g. a transition in the $|t|$-step transition matrix.), and $P_\varepsilon(t)$ as its probability. Then the case $x = C(t)$ demonstrates that $C(t)$ and $k(t)$ as defined are the appropriate values.

Proof:

Write $t = \{e_1, \ldots, e_n\}$ and define

$$r_i = \begin{cases} C(e_i) & i = 1, \ldots, n-1 \\ x - \sum_{j=1}^{n-1} C(e_j) & i = n \end{cases}$$

Since $P_{\bar{\varepsilon}(t)} > 0$, we have $P_{\bar{\varepsilon}(e_i)} > 0$ implying $C(e_i) < \infty$ for every i. Hence $r_i < \infty \; \forall i$.

$$\begin{aligned} \lim_{\varepsilon \to 0} \varepsilon^{-x} P_\varepsilon(t) &= \lim_{\varepsilon \to 0} \prod_1^n \varepsilon^{-r_i} P_\varepsilon(e) & (2) \\ &= \prod_1^n \lim_{\varepsilon \to 0} \varepsilon^{-r_i} P_\varepsilon(e) & (3) \\ &= \lim_{\varepsilon \to 0} \varepsilon^{-r_n} P_\varepsilon(e_n) \prod_1^{n-1} k(e) & (4) \end{aligned}$$

The conclusion follows from the observation that $x(\text{resp. } <, =, >)C(t) \iff r_n(\text{resp. } <, =, >)C(e_n) \iff \lim_{\varepsilon \to 0} \varepsilon^{-r_n} P_\varepsilon(e_n) = (\text{resp.} \infty, k(e_n), 0)$. ■

Having extended C to graphs, we now demonstrate some of its useful properties. First we define some simple operations on graphs. If h and g are subgraphs of G, their composition, $h \vee g$ is the subgraph $h \cup g$.

For any graph h, denote by h°, the set of vertices which are origins of edges in h. The set of vertices which are either origins or endpoints of edges in h is denoted $\overline{h}$. If h,h' are graphs, then $h \wedge h'$ is the graph $h \cup h'|_{V \setminus \overline{h}}$. The set of roots of a forest h are $r(h) = \overline{h} \setminus h^\circ$.

The notation $G|_Z$ represents the subgraph $h = \{(s,t) \in G : s \in Z\}$.

Lemma 2 *Let G be a graph*

1. *Fix a graph h. For any disjoint subsets $S, T \subset V$, $C(h|_{S\cup T}) = C(h|_S) + C(h|_T)$.*

2. *For any subgraphs h, g, $C(h \vee g) = C(h) + C(g)$.*

Proof:

1. $C(h|_{S\cup T}) = \sum_{e\in h|_{s\cup T}} C(e) = \sum_{e\in h|_S} C(e) + \sum_{e\in h|_T} C(e) = C(h|_S) + C(h|_T)$.

2. $C(h \vee g) = \sum_{h\cup g} C(e) = \sum_h C(e) + \sum_g C(e) = C(h) + C(g)$.

■

Remark: Thus a graph g determines a measure (i.e. $C(g|.)$) on the set of all subsets of V (set $C(g|_\emptyset) = 0$), and $\vee$ is an addition operator for measures.

4 Spanning Trees and Regular Markov Chains

Definition 7 *Given a subset of vertices F, a* **tree** *is a subgraph h, together with a* root *$r \in V$ satisfying:*

1. *Every point $v \in F \setminus r$ is the origin of exactly one edge.*

2. *$r \succeq_h v$ for each $v \in F$*

When $F = V$, h is called a **spanning tree**.

Remark: Note that in the definition of a tree, r is not required to be an element of F.

It is necessary and sufficient for a graph G to contain a spanning tree that there exist an r such that $r \succeq_G v$ for all $v \neq r$. (Sufficiency can be seen by ordering the vertices, $v_1, \ldots v_n$ and considering the graph $\wedge_n t_j$ where t_j is the path from v_j to r.)

There is an important connection between regularity of a Markov chain and properties of its corresponding graph. It follows from a standard result in the theory of Markov chains. Let $P^{(m)}$ denote the m-step transition matrix of a Markov chain P.

Theorem 1 (Romanovsky (1970)) *A Markov chain P is regular if and only if for some state r and for some positive integer m,*

$$P^{(m)}(v,r) > 0 \quad \forall\, v \in V$$

Corollary 1 *The graph of a regular Markov chain contains a spanning tree. A regular perturbed Markov graph contains a finite cost spanning tree.*

Proof: It follows immediately from Theorem 1 that if G is the graph of a regular Markov chain, then there is a path to r from every state v. IF P_ε is a regular perturbation, there thus exists a spanning tree of the graph obtained by including edges for all transitions whose probability is positive for all $\varepsilon \in (0, a)$. But these are the transitions which have finite order, hence the edges of the graph of P_ε. ■

Corollary 2 *If P is a regular Markov chain with graph G, then any Markov chain with graph G is regular.*

Proof: By Theorem 1, a chain is regular if and only if there exists a state r and a positive integer m such that $P^{(m)}(v,r) > 0$ for every $v \neq r$. This is equivalent to the existence of a sequence of vertices $\{v_j\}_1^m$ such that $v_1 = v$, $v_m = r$ and $P(v_j, v_{j+1}) > 0 \iff v_{j+1} \succeq_G v_j$. Thus any chain whose graph satisfies this property is regular. ■

Let H_v denote the set of spanning trees roots at v.

Theorem 2 (Friedlin and Wentzell, 1984) *Let P be a regular Markov Chain, and G its associated graph. Then μ is the invariant distribution of P if and only if for every $v \in V$,*

$$\mu(v) = \frac{q(v)}{\sum_{v' \in V} q(v')} \tag{5}$$

where $q(v) = \sum_{h \in H_v} P(h)$

Remark: Because P is regular, by Corollary 1 its graph contains a spanning tree, i.e. H_v is nonempty for at least one vertex v. This guarantees that (5) defines a probability distribution.

The quantity $q(v)$ is a measure of the likelihood of all possible paths to v. Theorem 2 implies that the more likely are the paths to a state v, the greater weight v receives in the equilibrium distribution. This result, and Lemma 1 greatly simplify the calculation of the limit distribution μ^* of a regular perturbation.

Theorem 3 *Let P_ε be a regular perturbation and G its associated graph. For each $v \in V$ define the set $Q_v = \text{argmin}\{h \in H_v\}C(h)$. Then*

$$\mu^*(v) = \frac{\sum_{h\in Q_v} k(h)}{\sum_{v'\in V}\sum_{h\in Q_{v'}} k(h)} \tag{6}$$

Remark: Because P_ε is a regular perturbation, by corollary 1 $\cup_{v\in V} H_v \neq \emptyset$, and this ensures that (6) defines a probability distribution.
Proof: Let $\kappa = \min_{\substack{v\in V \\ h\in H_v}} C(h)$. By the definition of a regular perturbation, P_ε is a regular Markov chain for every positive ε. Thus Theorem 2 implies for every $\varepsilon > 0$

$$\begin{aligned}
\mu_\varepsilon(v) &= \frac{q_\varepsilon(v))}{\sum_{v'\in V} q_\varepsilon(v')} && (7)\\
&= \frac{\sum_{h\in Q_v} P_\varepsilon(h) + \sum_{h\in H_v\setminus Q_v} P_\varepsilon(h)}{\sum_{v'\in V}\sum_{h\in Q_{v'}} P_\varepsilon(h) + \sum_{v'\in V}\sum_{h\in H_{v'}\setminus Q_{v'}} P_\varepsilon(h)} && (8)\\
&= \frac{\varepsilon^{-\kappa}\left[\sum_{h\in Q_v} P_\varepsilon(h) + \sum_{h\in H_v\setminus Q_v} P_\varepsilon(h)\right]}{\varepsilon^{-\kappa}\left[\sum_{v'\in V}\sum_{h\in Q_{v'}} P_\varepsilon(h) + \sum_{v'\in V}\sum_{h\in H_{v'}\setminus Q_{v'}} P_\varepsilon(h)\right]} && (9)
\end{aligned}$$

Taking limits with ε approaching zero and applying Lemma 1 yields

$$\mu^*(v) = \frac{\sum_{h\in Q_v} k(h)}{\sum_{v'\in V}\sum_{h\in Q_{v'}} k(h)} \tag{10}$$

since $k(h) > \kappa$ for every $h \notin Q_v$. ■

This result allows us to characterize the limit distribution as the solution to an optimization problem on a graph. In most applications all that is sought is the support of the limit distribution $\mathcal{R}$. In such cases we make use of the following corollary

Corollary 3 *A state v is strongly recurrent if and only if v is the root of a C-minimizing spanning tree.*

Proof: Equation (10) implies that $\mu^*(v) > 0$ iff $Q_v \neq \emptyset$ and by definition, Q_v is the set of spanning trees rooted at v which achieve the C-minimum among all spanning trees. ∎

In what follows, notation is simplified by considering only perturbations whose order functions are integer-valued. The following result demonstrates that this entails no loss of generality.

Proposition 3 *For every regular perturbation P_ε of P, there exists a regular perturbation $\tilde{P}_\varepsilon$ of P with the same limit distribution, and with order function $\tilde{C}$ such that $\tilde{C}(v_1, v_2) < \infty$ implies $\tilde{C}(v_1, v_2)$ is an integer.*

Proof: The proof is by construction. Given a regular perturbation P_ε of P, denote by C, k, and μ^* its order and rate functions, and limit distribution respectively. Let M be any positive real number, and set

$$\tilde{P}_\varepsilon(v_1, v_2) = \begin{cases} \Phi_{v_1} k(v_1, v_2)\varepsilon^{MC(v_1,v_2)} & \text{if } C(v_1, v_2) < \infty \\ 0 & \text{otherwise} \end{cases}$$

where Φ_{v_1} is the normalization constant for transitions originating at v_1. Notice first that

$$\lim_{\varepsilon \to 0} \Phi_{v_1} = \sum_{v_2} \lim_{\varepsilon \to 0} [k(v_1, v_2)\varepsilon^{MC(v_1,v_2)}]^{-1} \tag{11}$$

$$= \sum_{\{v_2 : C(v_1,v_2)=0\}} k(v_1, v_2)^{-1} \tag{12}$$

$$= \sum_{\{v_2 : P_0(v_1,v_2)>0\}} P_0(v_1, v_2)^{-1} \tag{13}$$

$$= 1 \tag{14}$$

We now demonstrate that $\tilde{P}_\varepsilon$ is a regular perturbation of P. The first condition is that $\lim_{\varepsilon \to 0} \tilde{P}_\varepsilon = P$. If $P_\varepsilon(v_1, v_2)$ is vanishing, then $\lim_{\varepsilon \to 0} \tilde{P}_\varepsilon(v_1, v_2) = 0 = P(v_1, v_2)$. Now suppose $P_\varepsilon(v_1, v_2)$ is non-vanishing. Then $\lim_{\varepsilon \to 0} \tilde{P}_\varepsilon(v_1, v_2) = k(v_1, v_2) = \lim_{\varepsilon \to 0} \varepsilon^{-C(v_1,v_2)} P_\varepsilon(v_1, v_2) = \lim_{\varepsilon \to 0} P_\varepsilon(v_1, v_2) = P(v_1, v_2)$. The case of $P_\varepsilon(v_1, v_2) = \tilde{P}_\varepsilon(v_1, v_2) \equiv 0$ is obvious.

Secondly, $\tilde{P}_\varepsilon$ must be a regular Markov chain for every $\varepsilon \in [0, a)$. To show regularity, let E be the graph of perturbation P_ε. Then $(v_1, v_2) \in$

$E \iff C(v_1, v_2) < \infty \iff \tilde{P}_\varepsilon(v_1, v_2) > 0 \; \forall \; \varepsilon > 0$. Thus, for every $\varepsilon > 0$, the graph $\tilde{E}$ of the Markov chain $\tilde{P}_\varepsilon$ coincides with E. Since E contains a spanning tree, so does $\tilde{E}$.

Lastly, we must show existence of order and rate functions for $\tilde{P}_\varepsilon$. By construction, if $\tilde{P}_\varepsilon(v_1, v_2)$ is greater than zero on $[0, a)$, then

$$\lim_{\varepsilon \to 0} \varepsilon^{-MC(v_1,v_2)} \tilde{P}_\varepsilon = \lim_{\varepsilon \to 0} k(v_1, v_2) = k(v_1, v_2)$$

Hence, $\tilde{C}(v_1, v_2) = MC(v_1, v_2)$ and $\tilde{k}(v_1, v_2) = k(v_1, v_2)$ for $C(v_1, v_2) < \infty$. Otherwise $\tilde{C}(v_1, v_2) = \infty$.

To calculate $\tilde{\mu}^*$, we use Theorem 3,

$$\tilde{\mu}^*(v) = \frac{\sum_{h \in \tilde{Q}_v} \tilde{k}(h)}{\sum_{v' \in V} \sum_{h \in \tilde{Q}_{v'}} \tilde{k}(h)} \tag{15}$$

By the above argument, $\tilde{C}(h) = MC(h)$, hence the composition of $\tilde{Q}_v$ is unchanged. Moreover, $\tilde{k}(h) = k(h)$ for every h, and hence we conclude $\tilde{\mu}^* = \mu^*$.

We have shown that C can be rescaled by any factor $M > 0$ in a way that leaves the limit distribution unchanged. Since the domain of C is finite, there exists a rescaling which results in an integer-valued C. ■

Remark: The construction of $\tilde{P}_\varepsilon$ in the proof underscores the point that, for the purposes of the limit distribution, all that is relevant are the orders and rates of the probabilities. In particular, there would be no loss of generality in restricting attention to probabilities which are polynomials in ε. In fact this is a universal assumption in the evolutionary literature.

5 The Order Decomposition

Some additional definitions about graphs are useful at this stage.

Definition 8 *Given a graph G, an* **absorbing set** *is a nonempty subset of vertices A such that if $v \in A$ then $v' \in A \iff v' \succeq_G v$. The* **domain** *of A is the set $D := \{v : A \succeq_G v\}$.*

In the above definition, it is understood that for sets A, the notation $A \succeq_G v$ indicates $v' \succeq v$ for some $v'inv$. Denote by $A(G)$ the collection of absorbing sets of the graph G. It is straightforward to verify that $A(G) \neq \emptyset$.

Definition 9 *Let V be a set of vertices. A* **graph hierarchy** *is a sequence of graphs $G_i = (V_i, L_i)$ such that* i . $V_0 = V$

ii . $V_i = A(G_{i-1})$

We will use the shorthand notation $A_i = A(G_i)$ when $\{G_i\}$ is a graph hierarchy. Similarly $D_i(Z)$ denotes the domain of Z in graph G_i, $\succeq_i$ replaces $\succeq_{G_i}$ etc. The first observation to make about this definition is that $V_1 = A(G_0)$ is a collection of subsets of V. Thus V_i is a family of collections of ... of subsets of V. We require some terminology to refer to these types of objects. For any set X, set $P(X)$ be the set of subsets of of X. Write $P^2(X) = P(P(X))$ and $P^n(X) = P(P^{n-1}(X))$.

Definition 10 *An* i**-grouping** *of* V *is an element of* $P^i(P(V))$. *An* **ancestor** *of an* i*-grouping* Z *is an element* $v \in V$ *such that*

$$v \in \cup_{A_i \in Z} \cup_{A_{i-1} \in A_i} \dots \cup_{A_0 \in A_1} A_0.$$

Let $\bar{Z}$ denote the set of ancestors of an i-grouping Z.

A graph hierarchy $\{G_i\}$ is said to converge if there exists an l such that A_l is a singleton.

Proposition 4 *A graph hierarchy* $\{G_i\}$ *converges if for every* j *such that* V_j *is not a singleton, there exists* $i \geq j$ *such that* $L_i \neq \emptyset$.

Proof: For any graph $G = (V, E)$, $|A(G)| \leq V$ with strict inequality whenever $E \neq \emptyset$. Since by definition of a graph hierarchy, $V_{j+1} = A_j$, we have $|V_{j+1}| = |A_j| \leq |V_j|$ with strict inequality whenever $L_j \neq \emptyset$. The proposition now follows. ■

Definition 11 *Let* G_i *be a graph hierarchy on* V. *For* $A \in A_i$, *define* $D_i^*(A)$ *the* **i-th order domain** *of of* A *recursively as follows*

i . $D_{-1}^*(v) = \{v\}$

ii . $D_i^*(A) = \cup_{Z \in D_i(A)} D_{i-1}^*(Z)$

Definition 12 *Let* G_i *be a graph hierarchy and* $A \in A_i$. *For any family* $\mathcal{F} \subset A_i \setminus A$, *the* **quasi-domain** *generated by* $\mathcal{F}$ *is the subset* $\mathcal{D} \subset D_i(A)$ *defined by* $\mathcal{D} = D_i(A) \setminus \cup_{F \in \mathcal{F}} D_i(F)$. *The associated* **i-th order quasi-domain** *is the subset of vertices* $S = D_i^*(A) \setminus \cup_{F \in \mathcal{F}} D_i^*(F)$.

A special type of i-order quasi domain is the ith order ***basin of attraction*** obtained by setting $\mathcal{F} = A_i \setminus A$:

$$B_i^*(A) = D_i^*(A) \setminus \cup_{Z \in A_i \setminus A} D_i^*(A)$$

Lemma 3 *Let $A \in A_i$, and $\mathcal{D}$ and S a quasi-domain, and associated i-th order quasi-domain of A.*

1. *For all $Z \in A$, there exists a tree τ of G_i on $\mathcal{D}$ rooted at Z.*
2. *For each $Z \in \mathcal{D}$, $D_{i-1}^*(Z) \cap S$ is an $i-1$ order quasi-domain.*
3. *$S = \cup_{Z \in \mathcal{D}} [D_{i-1}^*(Z) \cap S]$*

Proof: 1. It is sufficient that for each $Y \in \mathcal{D}$, there is a path t_Y from Y to Z in G_i which does not exit $\mathcal{D}$. That $Z \succeq_i Y$ follows immediately from the definition of $\mathcal{D}$. Suppose there were some $F \notin \mathcal{D}$ along the path. Then $F \notin \mathcal{D}$ implies $F \in D_i(F')$ for some $F' \in \mathcal{F}$where $\mathcal{F}$ is the family of A_i generating the quasi-domain $\mathcal{D}$. But $F \succeq_{t_Y} Y \Rightarrow F \succeq_{G_i} Y \Rightarrow Y \in D_i(F') \Rightarrow Y \notin \mathcal{D}$, a contradiction.

2. Using the definitions

$$\begin{aligned} D_{i-1}^*(Z) \cap S &= D_{i-1}^*(Z) \setminus \cup_{F \in \mathcal{F}} D_i^*(F) & (16)\\ &= D_{i-1}^*(Z) \setminus \cup_{F \in \mathcal{F}} \cup_{Y \in D_i^*(F)} D_{i-1}^*(Y) & (17) \end{aligned}$$

Thus, $D_{i-1}^*(Z) \cap S$ is the $i-1$ order quasi-domain generated by the family $\mathcal{F}' = \cup_{F \in \mathcal{F}} D_i^*(F) \subset A_{i-1}$.

3. Taking unions in (17) over $\mathcal{D}$, we have

$$\begin{aligned} \cup_{Z \in \mathcal{D}} [D_{i-1}^*(Z) \cap S] &= \cup_{Z \in \mathcal{D}} D_{i-1}^*(Z) \setminus \cup_{F \in \mathcal{F}} \cup_{Y \in D_i(F)} D_{i-1}^*(F) & (18)\\ &= \cup_{Z \in D_i(A)} D_{i-1}^*(Z) \setminus \cup_{F \in \mathcal{F}} D_i^*(F) & (19)\\ &= D_i^*(A) \setminus \cup_{F \in \mathcal{F}} D_i^*(F) & (20) \end{aligned}$$

where the second equation follows from the observation that $Z \in D_i(A) \setminus \mathcal{D} \Rightarrow Z \in D_i(F)$ for some $F \in \mathcal{F}$ and hence $D_{i-1}^*(Z) \setminus \cup_{\mathcal{F}} \cup_{D_i(F)} D_{i-1}^*(Y) = \emptyset$. ■

We now turn to a specific type of graph hierarchy which will be used to characterize $\mathcal{R}$.

Definition 13 *A* **forest** *is any subgraph of a tree. Given a subset of vertices F, an* **exit forest** *of F is a forest h such that $F = h^\circ$. If an exit forest is a tree, it is called an* **exit tree**

For any $A \in A_i$ and for each $v \in \bar{A}$, write $q_A(v) = \text{argmin}_{h|_{B_i^*}(A):h \in H_v} C(h)$

Definition 14 *Let G be a regular perturbed Markov graph. The* **order decomposition** *of G is the graph hierarchy $\{G_i\}$ satisfying $(Y, Z) \in L_i$ iff there exists $v \in D_{i-1}^*(Z)$ and an exit tree $h \subset G$ of $B_i^*(Y)$ rooted at v satisfying:*

$$C(h) = \min_{\substack{y \in \bar{Y} \\ \hat{h} \in q_A(y)}} C(h) + i$$

Definition 15 *Let $\{G_i\}$ be the order decomposition of the graph of a regular perturbation. For each $i \geq 0$ and for each $A \in A_i$, define the i-th level* **conditional distribution** *on $\bar{A}$ as follows:*

$$\mu_i(v|\bar{A}) = \frac{\sum_{h \in q_A(v)} k(h)}{\sum_{v' \in \bar{A}} \sum_{h \in q_A(v')} k(h)}$$

Theorem 4 *Let $\{G_i\}$ be the order decomposition of the graph of a regular perturbation. Then $\{G_i\}$ converges and for every $i \geq 0$ and $A \in A_i$,*

1. $\forall\, v \in \bar{A}$, $\mu^*(v) > 0 \iff \mu^*(\bar{A}) > 0$
2. $\mu^*(\cup_{A \in A_i} \bar{A}) = 1$
3. $\forall\, Z \in A$, $\mu^*(\bar{A}) > 0 \Rightarrow \mu^*(Z) = \mu^*(\bar{A}) \mu_i(\bar{Z}|\bar{A})$

The first result is that absorbing sets are equivalence classes of a sort. If some element of $\bar{A}$ has positive weight in the limit distribution, then so does every element of $\bar{A}$. The second result is that a state cannot have positive weight in the limit distribution if it is not contained in some $\bar{A}$. Thus if some $A \in A_i$ is eliminated in G_{i+1}, then $\bar{A}$ has no weight in the limit distribution. Since the order decomposition converges, the set ancestors of the resulting absorbing set is the support of the limit distribution. The last result provides a technique for calculating the actual limit probabilities. If we know the conditional distribution on a set $\bar{A}$ at step i of the order decomposition, then provided $\bar{A}$ has positive weight in the limit distribution, then the conditional distribution on $\bar{A}$ is unchanged.

The remainder of the Chapter is devoted to proving this theorem. We begin by stating without proof the following useful facts about forests and the operations $\vee$ and *wedge*.

Lemma 4 *Let g,h be forests.*

- *If $\overline{g} \cap h^\circ = \emptyset$, then the graph $g \vee h$ is a forest whose roots are $[r(h) \setminus g^\circ] \cup r(g)$.*
- *The graph $g \wedge h$ is a forest with $r(g \wedge h) = r(g) \cup [r(h) \setminus g^\circ]$.*

Remark: In particular, in either case if g is a tree and $r(h) \subset g^\circ$, then the resulting graph is a tree.

Lemma 5 *Let $A \in A_i$ and suppose h is an exit forest of some i-th order quasi-domain S of A. For any vertex r such that $r \succeq_h v$ for some $v \in \bar{A}$, there exists a tree h' on S rooted at r such that $C(h') \leq C(h)$. The inequality can be made strict if h contains an edge (s, s') satisfying the following two conditions*

1. $r \succeq_h s \Rightarrow v \succeq_h s$
2. *There exists $E \in A_i$ such that $s' \in D_i^*(E)$ but $s \notin D_i^*(E)$.*

Finally, If $r \notin S$, then q is the unique immediate predecessor of r in t where q is the immediate predecessor in h of r along the path from v to r.

The proof of Lemma 5 is by induction on i. First suppose $v \in A \in A_0$, and let h be an exit forest of S. Notice that $G_0 \subset G$ and therefore a 0-order quasi-domain is a quasi-domain of G_0. Thus by Lemma 3, there exists a tree $\tau \subset G_0$ on S whose root is v.

Let t be the path in h from v to r. Consider the following graph

$$h := t \wedge \tau$$

By Lemma 4 this is a tree rooted at r and since $\tau \subset G_0$, $C(\tau) = 0$ implying

$$C(h) = C(t) \leq C(h) \tag{21}$$

Clearly $t \subset h$, hence the immediate successor of r along this path, q, immediately precedes r in h. Moreover, since τ is a tree on S, if $r \notin S$ there is no edge (v, r) in τ, and hence q is the unique immediate predecessor.

Finally, suppose conditions 1 and 2 are satisfied. If $C(s, s') = 0$ then $(s, s') \in G_0$. But then $s' \in D_i^*(E)$ implies $s \in D_i^*(E)$ a contradiction. Hence, $C(s, s') > 0$. It follows from 2 that $(s, s') \notin t$, hence $C(h) \geq C(t) + C(s, s')$ and the inequality in (21) is strict.

Having proven Lemma 5 for the case of $i = 0$, we now derive a corollary which will be useful in the inductive step. The following will be true for $i = \bar{i}$ whenever Lemma 5 is true for $i = \bar{i}$. We then use these results to prove the sequence for $i = \bar{i} + 1$.

Lemma 6 *Let $A \in A_i$ and S the i-th order quasi-domain of A generated by some family $\mathcal{F} \subset A_i$. There exists h^* on S rooted at some $v \in \bar{A}$ such that $C(h^*) \leq C(h) - (i+1)$ for any exit forest h of S, with strict inequality if h satisfies one of*

1. $\exists r \in r(h|_{B_i^*(A)}) \cap D_i^*(F)$ *for some* $F \not\succeq_{i+1} A$.
2. $\exists (s, s') \in h|_{S \setminus B_i^*(A)}$ *such that for some* $E \in A_i \setminus \{A\}$, $s' \in D_i^*(E)$ *but* $s \notin D_i^*(E)$.

Secondly, let $U \in \mathcal{F}$. Iff $U \succeq_{i+1} A$, there exists an exit tree e^* of S whose root is in $D_i^*(U)$ and such that $C(e^*) \leq C(h)$ for any exit forest h of S, with strict inequality if h satisfies one of the above conditions.

Proof: Let h be any exit forest of S, and y a root of $h|_{B_i^*(A)}$ such that $y \succeq_h a$ for some $a \in \bar{A}$. The graph $h|_{B_i^*(A)}$ is an exit forest. By Lemma 5, we can find a tree b on $B_i^*(A)$ rooted at y such that

$$C(b) \leq C(h|_{B_i^*(A)}) \tag{22}$$

Suppose condition 1 is satisfied and assume $y \neq r$. Let q be an immediate predecessor of r in $h|_{B_i^*(A)}$. Then $q \in B_i^*(A)$ implying $q \notin D_i^*(F)$.

Since r is a root of $h|_{B_i^*(A)}$, $y \not\succeq r$ and thus the edge (q, r) and the set F satisfy the hypotheses for strict inequality in Lemma 5, hence we then have strict inequality in (22).

Since $B_i(A)$ is a quasi-domain, for any $Z_n \in A$, there exists a tree $\tau \subset G_i$ of $B_i(A)$ whose root is Z_n (by Lemma 3.) Let $\{Z_1, \ldots, Z_n\}$ be an ordering of $B_i(A)$ satisfying $Z_l \not\succeq_\tau Z_{l+1}$, and define the following collection:

$$S_j = [D_{i-1}^*(Z_j) \setminus \cup_{k>j} D_{i-1}^*(Z_k)] \cap B_i^*(A)$$

This is clearly a disjoint collection and since $\cup S_j = \cup_{B_i(A)} D_{i-1}^*(Z_j) \cap B_i^*(A) = B_i^*(A)$ (by Lemma 3), it is a partition of $B_i^*(A)$. Moreover,

$$S_j = [D_{i-1}^*(Z_j) \cap B_i^*(A)] \setminus \cup_{k>j} D_{i-1}^*(Z_k)$$

and by Lemma 3, the term in brackets is an $i-1$ order quasi-domain and hence there exists $\mathcal{F}' \subset A_{i-1} \setminus Z_j$ such that

$$S_j = [D_{i-1}^*(Z_j) \setminus \cup_{F \in \mathcal{F}'} D_{i-1}^*(F)] \setminus \cup_{k>j} D_{i-1}^*(Z_k)$$

and therefore each S_j is an $i-1$ order quasi-domain.

Let x be an immediate predecessor of y in b and let S_{j^*} be the element containing x. Note that $S_{j^*} \subset B_i^*(A)$ and $y \notin B_i^*(A)$ imply that y is a root of $b|_{S_{j^*}}$.

Since $y \notin B_i^*(A)$, there is some $U \in A_i \setminus A$ and $E \in A_{i-1}$ such that $U \succeq_i E$ and $y \in D_{i-1}^*(E)$.

I claim that $b|_{S_{j^*}}$ satisfies either 1 or 2 of the induction hypothesis. First suppose $x \in S_{j^*} \setminus B_{i-1}^*(Z_{j^*})$. Then since $x \notin D_{i-1}^*(E)$ (else $x \notin B_i^*(A)$), condition 2 is satisfied by the edge (x, y).

Suppose instead $x \in B_{i-1}^*(Z_{j^*})$. Then y is a root of $b|_{B_{i-1}^*(Z_{j^*})}$. It cannot be that $E \succeq_i Z_{j^*}$ else $Z_{j^*} \in D_i(U)$ implying $Z_{j^*} \notin B_i(A)$. In this case, therefore, condition 1 is satisfied.

We will now apply the induction hypothesis to each graph $b|_{S_j}$. The preceding arguments show that we have the conditions for strict inequality in the case of $b|_{S_{j^*}}$.

Since τ is a tree of G_i, $Z_{j+1} \succeq_i Z_j$ for every $j < n$ and therefore, the second part of the induction hypothesis implies that for each $j < n$, there exists an exit tree e_j of S_j rooted in $D_{i-1}^*(Z_j)$ such that

$$C(e_j) \leq C(b|_{S_j})$$

with strict inequality for j^* (if $j^* < n$).

The root r_j of e_j must be contained in $B_i^*(A)$. If not, then $r_j \in D_i^*(F)$ for some $F \neq A$. By the only if part of the second part of the induction hypothesis, $F \succeq_i Z_j$ contradicting $Z_j \in B_i(A)$.

(Notice that it is not necessarily the case that $r_j \in S_{j+1}$, but it must be that $r_j \in S_k$ for some $j < k \leq n$.)

The first part of the induction hypothesis implies that there exists a tree e_n on S_n rooted at some $v \in \bar{A}$ such that

$$C(e_n) \leq C(b|_{S_n}) - i$$

with strict inequality if $n = j^*$.

Now let $b' = \vee_1^n e_n$. We have $C(b') < C(b) - i$ implying $C(b') \leq C(b) - (i + 1)$ since we have assumed $C(\cdot)$ is integer-valued. Let b^* be a C-minimizing tree on $B_i^*(A)$ rooted within $\bar{A}$. Then

$$C(b^*) \leq C(b') \leq C(b) - (i+1) \tag{23}$$

Suppose equality. Then since y is the root of the exit tree b, by the definition of the order decomposition, $U \succeq_{i+1} A$ for every U such that $y \in D_i^*(U)$. But this would contradict $r \in r(h|_{B_i^*(A)}) \cap D_i^*(F)$ (i.e. condition 1) and $r = y$, hence in this case we have strict inequality in (23).

To summarize the argument to this point, we have demonstrated inequalities (22) and (23), and have shown that if condition 1 is satisfied then one of them holds with strict inequality (in 22 when $y \neq r$, otherwise in 23).

Now since $i < l$, there exists $U \in A_i \setminus \{A\}$. Let $w \in \bar{U}$. The graph z_1 obtained by adding the edge (v, w) to $b^* \wedge h$ is an exit forest of S, with root w succeeding $v \in \bar{A}$. Lemma 5 therefore applies yielding an exit tree z_2 whose root, w, has unique immediate predecessor v, with $C(z_1) \leq C(z_2)$. Suppose condition 2 is satisfied. Then, $(s, s') \in z_1$ since $s \notin \overline{b^*} = B_i^*(A)$, and by the construction of z_1, if $w \succeq_{z_1} s$ then $v \succeq_{z_1} s$ and thus the conditions for strict inequality in Lemma 5 are satisfied. Letting h^* be the graph obtained by removing (v, w) from z_2, we have

$$\begin{aligned} C(h^*) &\leq C(b^* \wedge h) && (24)\\ &= C(b^*) + C(h|_{S \setminus B_i^*(A)}) && (25)\\ &\leq C(h) - (i+1) && (26) \end{aligned}$$

Where the first inequality is strict if 2 is satisfied and the second follows from combining equations (23) and (22), and therefore is strict if condition 1 holds. This establishes the first part of the lemma.

Finally, by the definition of the order decomposition, if and only if $F \succeq_{i+1} A$ there exists an exit tree e of $B_i^*(A)$ rooted within $D_i^*(F)$ such that $C(e) = C(h^*|_{B_i^*(A)})$. Hence, taking $e^* = e \wedge h^*$ gives the second part. ■

Proof of Lemma 5 We now turn to the inductive step. Since S is an i-th order quasi-domain, by Lemma 3 there exists a quasi-domain $\mathcal{D} \subset A_{i-1}$ of A such that $S = \cup_{Z \in \mathcal{D}} D_{i-1}^*(Z) \cap S$. Also by Lemma 3, there exists a tree τ of G_i on $\mathcal{D}$ rooted at the element Z_n of A whose ancestor is v. Let t be the path in h from v to r.

Now, fix an ordering of $\mathcal{D}$, $\{Z_1, \ldots, Z_n\}$ such that $Z_l \not\succeq_\tau Z_{l+1}$. Consider the family defined by

$$S_j := [D_{i-1}^*(Z_j) \setminus \cup_{k>j} D_{i-1}^*(Z_k)] \cap S \quad j \in \{1, \ldots, n\}$$

As was argued in the proof of Lemma 6, this is a partition of S, with each S_j an $i-1$ order quasi-domain. Let S_{j^*} be the element containing s, if it exists.

For each $j = 1, \ldots, n$, let $t_j = h|_{S_j} \cap t$. If $t_j \neq \emptyset$ then t_j is a path, so let o_j and r_j be the origin and endpoint, respectively. Notice that r_j is a root of $h|_{S_j}$. Define the sub-family

$$\Omega := \{S_j : \exists z_j \in \bar{Z}_j \text{ s.t. } z_j \succeq_h o_j\}$$

By assumption, $o_n = v \in \bar{Z}_n$ and therefore $S_n \in \Omega$. By the induction hypothesis, for each $S_j \in \Omega$ we can find a tree h_j on S_j rooted at r_j, satisfying

$$C(h_j) \leq C(h|_{S_j}) \tag{27}$$

If the S_j which contains q is in Ω, then, per the induction hypothesis, we choose h_j such that q is the unique immediate predecessor of r. Notice that $s' \in D^*_{i-1}(K)$ for some $K \in D_i(E)$ and since $s \notin D^*_i(E)$ we also have $s \notin D^*_{i-1}(K)$. Furthermore, the condition $r \succeq_h s \Rightarrow v \succeq_h s$ implies the analogous $r_j \succeq_{h_j} s \Rightarrow v_j \succeq_{h_j} s$ Thus we can make the inequality in (27) strict in the case of S_{j^*} (if it is contained in Ω).

Now consider $S_j \notin \Omega$. If $t_j \neq \emptyset$, let q_j be the immediate predecessor of r_j along the path t_j and define

$$T_j = \{v \in S_j : q_j \succeq_{h|_{S_j}} v\}$$

If $t_j = \emptyset$, then set $T_j = \emptyset$. Notice that the graph $h|_{T_j}$ is a tree rooted at r_j and it contains the path t_j. The graph $h|_{S_j \setminus T_j}$ is an exit forest of its domain of definition, none of whose roots are in T_j. (If $h|_{S_j \setminus T_j}$ had a root in T_j, then there would be at least one $v \in S_j \setminus T_j$ such that $r_j \in S_j \setminus T_j \succeq_h v$ implying $v \in T_j$.)

By Lemma 6 there exists an exit tree h^*_j of S_j with root $y_j \in D^*_{i-1}(Z_k)$ for some $k > j$ (specifically the successor of Z_j in τ) which minimizes the order of any exit forest of S_j. In particular, since no roots of the exit forest $h|_{S_j \setminus T_j}$ are in T_j, it follows from Lemma 4 that the graph $h^*_j|_{T_j} \wedge h|_{S_j}$ is an exit forest of S_j and therefore

$$\begin{aligned} C(h^*) &\leq C(h^*|_{T_j} \vee h|_{S_j \setminus T_j}) && (28)\\ C(h^*_j|_{T_j} \vee h^*_j|_{S_j \setminus T_j}) &\leq C(h^*|_{T_j} \vee h|_{S_j \setminus T_j}) \Rightarrow && (29)\\ C(h^*_j|_{S_j \setminus T_j}) &\leq C(h|_{S_j \setminus T_j}) && (30) \end{aligned}$$

Consider the case of $S_{j^*} \notin \Omega$. Suppose $s \in T_{j^*}$. Then $r \succeq_h s$ implying by hypothesis that $v \succeq_h s$. In particular, because there is a path within S_j from s to r_j, v must lie along this path. But this implies that $j^* = n$, a contradiction since we have already shown that $S_n \in \Omega$. Thus, if $S_{j^*} \notin \Omega$ then $s, s' \notin T_{j^*}$.

It follows that $(s, s') \in h^*|_{T_{j^*}} \vee h|_{S_{j^*} \setminus T_{j^*}}$. Suppose $s \in B^*_{i-1}(Z_{j^*})$. Then since $s \notin D^*_i(E)$, it cannot be that $K \succeq_i Z_{j^*}$, hence since $s' \in r(h|_{B^*_{i-1}(Z_{j^*})}) \cap D^*_{i-1}(K)$, we have condition 1 for strict inequality in 29 according to Lemma 6.

If on the other hand, $s \in S_{j^*} \setminus B^*_{i-1}(Z_{j^*})$, then we have condition 2 of Lemma 6 by condition 2 of the induction hypothesis.

Thus, the inequality in 29 and hence 30 is strict in the case of S_{j^*} if not in Ω

with a strict inequality in (29) in the case of j^* since $s, s' \notin T_{j^*}$ implies $(s, s') \in h^*|_{T_j} \wedge h|_{S_j}$.

For $S_j \notin \Omega$, define $h_j = h|_{T_j} \wedge h^*|_{S_j}$. By construction, r_j is the root of $h_j|_{T_j} = h|_{T_j}$ and therefore Lemma 4 implies that h_j is a forest of S_j whose roots are a subset of $\{y_j, r_j\}$.

Define $h' := \vee_j h_j$. Notice that for every j, such that $t_j \neq \emptyset$, $r_j \succeq_{h_j} o_j$, hence $r \succeq_{h'} v$. Moreover, for every $S_j \in \Omega$, $r_j \succeq_{h_j} S_j$ hence $r \succeq_{h'} S_j$.

For each vertex $u \in S_j \notin \Omega$, either $r_j \succeq_{h_j} u$ in which case $r \succeq_{h'} u$ or $y_j \succeq_{h_j} u$ in which case $S' \succeq_{h'} u$ where S' is the successor of S_j in the tree τ. Since τ is a tree rooted at $Z_n \in \Omega$, it follows that $r \succeq_{h'} u$.

Corollary 4 *Suppose h is a forest on $B^*_i(A)$ with a single root r within $B^*_i(A)$. Then there is a tree h' on $B^*_i(A)$ rooted at r with no greater order and strictly less whenever h is not a tree.*

Proof: Let q be a vertex outside $B^*_i(A)$ and consider the exit forest g_1 created by adding the edge (r, q) to h. Lemma 5 applies and we can construct a tree g_2 on $B^*_i(A)$ whose root is q with $C(g_2) \leq C(g_1)$ and such that the unique immediate predecessor of q is r. If h is not a tree, then there is some root z of h such that $z \notin B^*_i(A)$. But then the edge (y, z) satisfies the hypotheses for strict inequality, where y is any predecessor of z in h. Let h_1 be the graph that results from deleting the edge (r, q) from g_2. We have

$$\begin{aligned} C(h_1) &= C(g_2) - C(r, q) \\ &\leq C(g_1) - C(r, q) \\ &= C(h) \end{aligned}$$

with strict inequality as needed. Moreover, since r is the unique immediate predecessor of the root q of g_2, we have $q \succeq_{g_2} r \succeq_{g_2} s$ for every $s \in B^*_i(A)$ and thus h_1 is a tree rooted at r.

■

Lemma 7 *Let $A \in A_i$ and $Z \in D_i(A)$. For any $z \in \bar{Z}$, $h \in H_z$, and for any $a \in \bar{A}$, there exists $h_a \in H_a$ such that $C(h_a) \leq C(h_z)$ with strict inequality when $Z \notin A$.*

Proof: The proof is by induction on i. First let $A \in A_0$ and $z \in D_0(A)$ and fix a spanning tree $h \in H_z$. $z \in D_0(A)$ implies that for any $a \in A$, there is a path t from z to a such that $C(t) = 0$. The tree $t \wedge h$ is rooted at a and has no greater order than h.

Suppose now $z \notin A$. By the definition of A_0, $A = \{v \in V : v \succeq_0 a\}$. This implies that the path in h from a to z must go contain some edge (a', v) where $a' \in A$ and $C(a', v) > 0$. Since $a' \in A$, there is a path t' from a' a such that $C(t') = 0$. Thus $t' \wedge t \wedge h$ is a tree rooted at a with strictly smaller order than h.

For the inductive step, we begin with the following claim. If $Z \notin A$, then there exists some $F \in A$ such that $h|_{D^*_{i-1}(F)}$ satisfies either condition 1 or 2 of Lemma 6. To prove this claim, suppose $Z \notin A$. Then $B^*_{i-1}(Z) \cap \cup_{U \in A} D^*_{i-1}(U) = \emptyset$ and thus since $z \in \bar{Z} \subset B^*_{i-1}(Z)$, the graph $g := h|_{\cup_A D^*_{i-1}(U)}$ is an exit forest.

For every $q \in \cup_A B^*_{i-1}(U)$, there is a root r of g such that $r \succeq_g q$. Choose q and r such that q is a root of $g|_{\cup_A B^*_{i-1}(U)}$ and there is no root q' of $g|_{\cup_A B^*_{i-1}(U)}$ along the path from q to r in g. Let F be the element of A such that $q \in g|_{B^*_{i-1}(F)}$.

Since r is a root of g, $r \notin \cup_A D^*_{i-1}(U)$ hence $r \in D^*_{i-1}(E)$ for some $E \notin A$. All vertices between q and r are contained in $\cup_A D^*_{i-1}(U) \setminus \cup_{F \in A} B^*_{i-1}(F) = \cup_A [D^*_{i-1}(F) \setminus B^*_{i-1}(F)]$ and therefore if condition 2 is not satisfied for any $F \in A$, then each of these vertices must belong to $D^*_{i-1}(E)$. But in particular, this implies $q \in D^*_{i-1}(E)$. Since q is a root of $h|_{B^*_{i-1}(F)}$ and since $E \notin A$ implies $E \not\succeq_i F$, Condition 1 is satisfied for F

Thus, if $Z \notin A$, there exists some element, call it Z_n of A satisfying either condition 1 or 2 from Lemma 6. (If $Z \in A$, then let Z_n be any element of A.) Since $Z \in D_i(A)$, there is a path in G_i from Z to Z_n. Let $Z_1, \ldots, Z_{n-1}$ be the vertices between Z and Z_n along this path. Define the family of quasi-domains

$$S_j := D^*_{i-1}(Z_j) \setminus \cup_{k>j} D^*_{i-1}(Z_k) \quad j = 1, \ldots, n.$$

For each $j = 1, \ldots, n-1$, the graph $h|_{S_j}$ is an exit forest and therefore by the second part of Lemma 6, there exists a tree h_j on S_j rooted in $D^*_{i-1}(Z_{j+1})$ such that

$$C(h_j) \leq C(h|_{S_j}) \tag{31}$$

Consider the graph $h|_{B^*_i(Z)}$. Since h is a tree rooted at $z \in \bar{Z} \subset B^*_i(Z_1)$, $h|_{B^*_i(Z)}$ is a forest whose only root within S_1 is z. Applying Corollary 4, we obtain a tree t on $B^*_i(A)$ whose root is z with $C(t) \leq C(h|_{B^*_i(Z)})$.

Thus by the definition of the order decomposition, there exists an exit tree h_0 of $B_i^*(Z)$ rooted within $D_{i-1}^*(Z_1)$ such that

$$C(h_0) \leq C(h|_{B_{i-1}^*(Z)}) + i \tag{32}$$

Finally, $h|_{S_n}$ is an exit forest, and therefore by the first part of Corollary 6, there exists a tree h_n on S_n rooted at some $a \in \bar{Z}_n$ such that

$$C(h_n) \leq C(h|_{S_n}) - i \tag{33}$$

with strict inequality when $Z_1 \notin A$.

Define the graph $h' = (\vee_0^n h_j) \wedge h$. It follows by iterative application of Lemma 4 that h' is a tree whose root is a, and by combining equations (31) , (32) , and (33), we have

$$\begin{aligned} C(h') - C(h) &= \sum_{j=1}^{n}(C(h_j) - C(h|_{S_j})) \\ &\leq 0 \end{aligned}$$

with a strict inequality when $Z_1 \notin A$.

In the case $Z \in A$, Z_n was an arbitrary element of A. Thus, we have shown that for each $F \in A$, there is some element $a \in \bar{F}$ and a tree $h_a \in H_a$ which has no greater order than h. But by the induction hypothesis, for every $a' \in \bar{F}$, there is a tree $h_{a'} \in H_{a'}$ whose root is a' with no greater order than h_a. Thus, since $\bar{A} = \cup_{F \in A}\bar{F}$, the lemma is proven. ■

Proof of Theorem 4 We first show that the order decomposition converges. Suppose not, then by Proposition 4, there exists j such that $|V_j| > 1$ and $L_i = \emptyset$ for every $i \geq j$. Thus

$$\{B_i^*(A) : A \in A_i\} = \{B_{i+1}^*(A) : A \in A_{i+1}\} := \beta$$

for every $i \geq j$. But then by the definition of the order decomposition, $L_i = \emptyset \ \forall i \geq j$ implies that for every $B \in \beta$ and for every exit tree h of B, $C(h) = \infty$.

We will show that this leads to a contradiction. By the definition of a regular perturbation, there exists a tree h on V with finite order. Let B be an element of β which does not contain the root of h. (There must be some such B since $|\beta| > 1$ and by the definition of basins of attraction, the elements of β are disjoint.) The graph $h|_\beta$ is an exit forest. By lemma 5 there exists an exit tree t of B with no greater order. Hence there exists a finite order exit tree of B.

Define the map ϕ on V as follows.

$$\phi(v) = \min_{h \in H_v} C(h)$$

By Corollary 3, $\mu^*(v) > 0 \iff v \in \text{argmin}\phi(\cdot)$.

1. Suppose $\mu^*(v) > 0$ and $v, v' \in A$. Then by Lemma 7, $\phi(v') = \phi(v)$ and $v \in \text{argmin}\phi(\cdot) \Rightarrow v' \in \text{argmin}\phi(\cdot)$.

2. Let $v \notin \cup_{A \in A_i} \bar{A}$. Let j^* be the largest j such that $v \in \bar{A}$ for some $A \in V_j$. Since by definition $V_0 = V$, we have $0 \leq j^* < i$. For some $Z \in A_{j+1}$, $A \in D_{j+1}(Z)$. By Lemma 7 then, $\phi(\bar{Z}) < \phi(v)$. Thus $v \notin \text{argmin}\phi(\cdot)$.

3.Let $v \in \bar{A}$ and define the following sets

$$\begin{aligned} q(v) &= \text{argmin}_{h \in H_v} C(h) \\ d(v) &= \text{argmin}_{h \in H_v} C(h|_{V \setminus B_i^*(A)}) \end{aligned}$$

I claim $q(v) = \{b \wedge d : b \in q_A(v), d \in d(v)\}$. By Lemma 4, all graphs in this set are trees, and clearly they have equal order. Suppose there is some $h \in H_v$ which minimizes C. By Corollary 4, it is wlog to assume $h|_{B_i^*(A)}$ is a tree. But then it must be that $h|_{B_i^*(A)} \in q_A(v)$. Finally $h|_{V \setminus B_i^*(A)} \in d(v)$, else the graph $h|_{B_i^*(A)} \wedge d$ is a tree with strictly smaller order for any $d \in d(v)$.

We now calculate

$$\begin{aligned} \mu_i(v|\bar{A}) &= \frac{\sum_{b \in q_A(v)} k(b)}{\sum_{v' \in \bar{A}} \sum_{b \in q_A(v')} k(b)} \frac{\sum_{d \in d(v)} k(d)}{\sum_{d \in d(v)} k(d)} \\ &= \frac{\sum_{b \in q_A(v)} \sum_{d \in d(v)} k(b)k(d)}{\sum_{v' \in \bar{A}} \sum_{b \in q_A(v')} \sum_{d \in d(v)} k(b)k(d)} \\ &= \frac{\sum_{h \in q(v)} k(h)}{\sum_{v' \in \bar{A}} \sum_{h \in q(v')} k(h)} \\ &= \frac{\mu^*(v)}{\sum_{v' \in \bar{A}} \mu^*(v')} \\ &= \frac{\mu^*(v)}{\mu^*(\bar{A})} \end{aligned}$$

■

References

ELY, J. C. (1996): "Local Conventions," mimeo.

FREIDLIN, M., AND A. WENTZELL (1984): *Random Perturbations of Dynamical Systems.* Springer-Verlag, New York.

KANDORI, M., G. MAILATH, AND R. ROB (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 65, 383–414.

ROMANOVSKY, A. (1970): *Discrete Markov Chains.* Wolters-Noordhoof, Groningen.

YOUNG, H. P. (1993): "The Evolution of Conventions," *Econometrica*, 61, 57–84.