Supplement to "Dynamics in stochastic evolutionary models"

(*Theoretical Economics*, Vol. 11, No. 1, January 2016, 89–131)

David K. Levine
Department of Economics, Washington University in St. Louis and European University Institute

Salvatore Modica
Department of Economics, Universita di Palermo Salvatore

Appendix S.A: Further bounds on the length of direct paths

We first prove two lemmas.

Lemma S.1. *For $0 < C < 1$ and $t_F \geq 0$, we have* $\sum_{t_s=0}^{\infty} t_s(1-C^{t_F})^{\lfloor t_s/t_F \rfloor} \leq 3t_F^2/C^{2t_F}$.

Proof. We have

$$\begin{aligned}
\sum_{t_s=0}^{\infty} t_s(1-C^{t_F})^{\lfloor t_s/t_F \rfloor} &= \sum_{k=0}^{\infty}\sum_{h=1}^{t_F}(kt_F+h)(1-C^{t_F})^k \\
&= \sum_{h=1}^{t_F}\left(t_F\sum_{k=0}^{\infty}k(1-C^{t_F})^k + h\sum_{k=0}^{\infty}(1-C^{t_F})^k\right) \\
&= \sum_{h=1}^{t_F(A)}(t_F(1-C^{t_F})/C^{2t_F}+h/C^{t_F}) \\
&= (t_F^2(1-C^{t_F})/C^{2t_F}+(t_F^2+t_F)/(2C^{t_F})) \\
&\leq 3t_F^2/C^{2t_F},
\end{aligned}$$

giving the desired result. □

Lemma S.2. *If $A \subseteq A_{xBW}$ is not empty and W is comprehensive, then for $t \geq 0$, we have*

$$P_\epsilon(t(a)=t+1, a\in A|x) \leq (1-C^{t_F(A)})^{\lfloor (t+1)/t_F(A)\rfloor},$$

and if B is a singleton, then

$$P_\epsilon(t(a)=t+1, a\in A|x) \leq \max_{(x,z_1,z_2,\ldots,z_{t-1},z_t)\in A} P_\epsilon(z_t(a)|z_{t-1}(a))(1-C^{t_F(A)})^{\lfloor t/t_F(A)\rfloor}.$$

David K. Levine: david@dklevine.com
Salvatore Modica: salvatore.modica@unipa.it

DOI: 10.3982/TE1978

Proof. The first inequality was proven in Lemma 7 in Appendix A. The second inequality makes use of the fact that in the course of proving that lemma, we used only the fact that all the loops ended at the same target and that all had the same transition probability at the end. If we replace the unique final transition probability with the maximum over all final transition probabilities, the same argument goes through. $\square$

In Appendix A to the paper, a better bound is given for least resistance paths that exploits the fact that they have a special structure. The idea is that long least resistance paths are not likely to be very long, because to be long they must contain long loops, and long loops are not very likely. For least resistance paths, these loops must have zero resistance; however, in a large state space, we could have zero resistance pieces of least resistance paths that are "unnecessarily" long but do not in fact loop. Our goal is to show that these too are unlikely. To do so, we introduce the idea of a *waypoint* of a path $a = (z_0, z_1, \ldots, z_t)$. Let $(z_{\tau-1}, z_\tau)$ be the first transition in the path that has positive resistance. The first waypoint is defined as z_τ. Similarly, the second waypoint is defined to be the end of the second transition in the path that has positive resistance and so forth. We say that two paths a, a' are equivalent, written $a \sim a'$, if they have the same waypoints. The idea is now to give conditions for least resistance paths under which the amount of time between waypoints is bounded independent of the size of the state space, and, consequently, to get a bound on the expected length of least resistance paths of order equal to the number of waypoints. Let $Y(A)$ be the set of sequences of waypoints derived from paths in A and, for any given sequence of waypoints $y \in Y(A)$, let $A_{\tau-1}(y)$ be the set of least resistance paths from $z_{\tau-1}$ to z_τ.

Theorem S.1. *If W is comprehensive and $A \subseteq A_{xBW}$ not empty is the set of all least resistance paths, then*

$$E_\epsilon(t(a)|x, A) \le \max_{y=(z_0,z_1,\ldots,z_{t-1})\in Y(A)} t\Big[\max_{0\le s\le t-1} 3Dt_F(A_s(y))^2/C^{2t_F(A_s(y))+t(A_S(y))}\Big].$$

Proof. Pick $y = (z_0, z_1, \ldots, z_{t-1}) \in Y(A)$, that is, a sequence of waypoints, and let A_y be the paths with those waypoints. Notice these sets form a partition of A. If a_τ is a sequence of states (indexed starting with 1), let $z_s(\tau)$ be the sth element of the sequence and let $s(\tau)$ be the length of the sequence. Since the paths in question are least resistance paths, they are exactly paths of the form $(a_0, a_1, \ldots, a_{t-1})$, where

- $z_1(0) = x$
- either $z_{s(t-1)}(t-1) \in B$ or $z_{t-1} \in B$, $a_{t-1} = \emptyset$
- any transitions in a_τ have zero resistance
- transitions $z_{s(\tau-1)}(\tau-1)$, $z_1(\tau)$ have positive resistance $r_{\tau-1}$ that depends only on τ
- $(a_{\tau-1}, z_1(\tau))$ is a least resistance path from $z_1(\tau-1)$ to $z_1(\tau)$ (with forbidden set W).

Put differently, setting $A_{\tau-1} = A_{\tau-1}(y)$ (the set of least resistance paths from $z_{\tau-1}$ to z_τ), then a path is a least resistance path if and only if $a_{\tau-1} \in A_{\tau-1}$ and $a_\tau \in A_\tau$ imply that any transitions in a_τ have zero resistance, and the transition $z_{s(\tau-1)}(\tau-1), z_1(\tau)$ has positive resistance equal to $r_{\tau-1}$ independent of which path in $A_{\tau-1}$ is chosen. Let $P_\tau(t) \equiv P_\epsilon(t((a, z_{\tau+1})) = t+1, a \in A_\tau | x)$. Then (using the same algebra as in the paper) we have

$$E(t(a)|x, A_y) = \frac{\sum_{t_0=0}^{\infty} \sum_{t_1=0}^{\infty} \cdots \sum_{t_{t-1}=0}^{\infty} (\sum_{s=0}^{t-1} t_s) \prod_{\tau=0}^{t-1} P_\tau(t_\tau)}{\prod_{\tau=0}^{t(a)-1} \sum_{t=0}^{\infty} P_\tau(t)}$$
$$= \sum_{s=0}^{t-1} \frac{\sum_{t=0}^{\infty} t_s P_s(t)}{\sum_{t=0}^{\infty} P_s(t)}.$$

As in Lemma 7 in Appendix A, by using Lemma S.2 and Lemma S.1, we find

$$\sum_{t=0}^{\infty} t_s P_s(t) \leq \sum_{t_s=0}^{\infty} t_s D \epsilon^{r_s} (1 - C^{t_F(A_s)})^{\lfloor t_s / t_F(A_s) \rfloor}$$
$$\leq D\epsilon^{r_s} 3 t_F(A_s)^2 / C^{2t_F(A_s)}.$$

Since, $\sum_{t=0}^{\infty} P_s(t) \geq C^{t(A_s)} \epsilon^{r_s}$, the desired bound holds. □

As we move away from a recurrent communicating class along a least resistance path, initially we are in the basin of the class and we encounter resistance. This gives a natural monotonicity to this part of the path: each time we encounter resistance, we cannot go back and do it again because to do so would add unnecessary resistance. The bounds in Theorem S.1 exploit this monotonicity and so are useful in bounding the time it takes to get out of the basin. However, once we leave the basin there will be zero resistance paths to other recurrent communicating classes, and so there will be no more waypoints and the bound is not useful. Indeed, as Appendix S.B shows, the length of time in this region may not scale. However, in applications such as the model of hegemony, once we get close enough to the recurrent communicating class that will be the end of the least resistance path, there may be a form of monotonicity: in the example there is a point at which the eventual hegemon can only gain land (along a least resistance path) and not lose it. If, in place of the natural monotonicity of Theorem S.1, we assume monotonicity, then we can get a bound for this final segment of the least resistance path.

To formalize this, we first give a bound on the probability of zero resistance paths in the basin. Suppose that for comprehensive W, the set $A \subseteq A_{xBW}$ of least resistance paths is not null. Define $r_{xBW} \equiv \min\{r(A_{x(W \setminus B)W}), r(A_{xBW} \setminus A)\}$ and $t_{xBW} \equiv \max\{t(A_{x(W \setminus B)W}), t(A_{xBW} \setminus A)\}$. Notice that $r_{xBW} > 0$ means that $r(A) = 0$, since there must be some zero resistance path from x to W, and that x is in the basin of B, since all 0 resistance direct routes from x to B are in A.

Theorem S.2. *If $r_{xBW} > 0$, then $P_\epsilon(A|x) \geq 1 - 2G(t_{xBW})\epsilon^{r_{xBW}}$.*

Proof. Since W is comprehensive, with probability 1 every path originating at x hits W with probability 1. Hence $P_\epsilon(A_{x(W\setminus B)W}|x) + P_\epsilon(A_{xBW}\setminus A|x) + P_\epsilon(A|x) = 1$. However, by the bound proven in Appendix A of the paper, we have $P_\epsilon(A_{x(W\setminus B)W}|x)$, $P_\epsilon(A_{xBW}\setminus A|x) \leq G(t_{xBW})\epsilon^{r_{xBW}}$, giving the desired result. □

Now consider a sequence of targets $B_1, B_2, \ldots, B_t$, where $B_t = B$. Also set $B_0 = \{x\}$. For any a starting at x, we may consider $t_1(a)$ the first time B_1 is hit before hitting W, possibly infinite, and if B_1 is hit before W, we may consider $t_2(a)$ the additional amount of time from first hitting B_1 until B_2 is hit before hitting W, again infinite if either target is not hit before reaching W, and so forth. We say that the sequence is a *Liapunoff* sequence for A if for every a we have $t_\tau(a) < \infty$. In this case, the sequence of states $(z_1, z_2, \ldots, z_t)$ that are hit are similar to waypoints. For $y \in B_\tau$, let $A_\tau(y) \equiv \mathcal{A}(y, B_{\tau+1}, W)$. Let $t_{FF}(A) \equiv \max_{0\leq\tau<t} t_F(A_\tau)$. Then we can state the following theorem.

Theorem S.3. *If $B_1, B_2, \ldots, B_t$ is a Liapunoff sequence for least resistance paths A, then*

$$E_\epsilon(t(a)|x, A) \leq t\frac{1}{P_\epsilon(A|x)}\frac{3t_{FF}(A)^2}{C^{2t_{FF}(A)}}.$$

Proof. Define $\underline{t}_\tau(a)$ to be $t_\tau(a)$ if it is finite and zero otherwise, and observe that for $a \in A$, we have $t_\tau(a) = \underline{t}_\tau(a)$. Hence we may write

$$\begin{aligned} E_\epsilon(t(a)|x, A) &= \sum_{\tau=0}^{t-1} E_\epsilon(\underline{t}_\tau(a)|x, A) \\ &= \sum_{\tau=0}^{t-1} \frac{E_\epsilon(\underline{t}_\tau(a)|x, A)P_\epsilon(A|x)}{P_\epsilon(A|x)} \\ &\leq \frac{1}{P_\epsilon(A|x)}\sum_{\tau=0}^{t-1} E_\epsilon(\underline{t}_\tau(a)|x). \end{aligned}$$

Moreover, $E_\epsilon(\underline{t}_\tau(a)|x) \leq \max_{y\in B_\tau} E_\epsilon(\underline{t}_\tau(a)|y)$, as either $\underline{t}_\tau(a)$ is zero or a hits some $y \in B_\tau$ before hitting $B_{\tau+1}$ by definition. The desired bound now follows from Lemma S.2 and the summation formula Lemma S.1. □

Appendix S.B: Expected passage time bounds

Let V_t a standard Weiner process with 0 drift and instantaneous variance 1 that starts at 0. Now let T be the first time that V_t leaves the region $[-A, +A]$. As usual, Φ is the standard normal. First we prove the next lemma.

Lemma S.3. *We have $ET \geq A^2/(2[\Phi^{-1}(\frac{1}{8})]^2)$.*

Proof. Let τ^+ be the first passage time for $A > 0$. We first establish a standard result: $\Pr(V_t > A) = \Pr(V_t > A \,\&\, \tau^+ < t) = \frac{1}{2}\Pr(\tau^+ < t)$. The first equality follows from the fact that if $V_t > A$, then certainly $\tau^+ < t$. The second follows from the reflection principle: starting at $V_{\tau^+} = A$, there is an equal probability of $\frac{1}{2}$ that $V_t > A$ and $V_t < A$; hence if $\tau^+ < t$, the probability that $V_t > A$ also is half the probability that $\tau^+ < t$.

Our goal is to establish a lower bound on the expectation of T. Let τ^- be the first passage time of $-A$. First we observe that

$$\Pr(\tau^+ < t) = \Pr(\tau^+ < t \,\&\, \tau^- > t) + \Pr(\tau^+ < t \,\&\, \tau^+ < \tau^- < t) + \Pr(\tau^+ < t \,\&\, \tau^- < \tau^+).$$

Using the reflection principle, we have

$$\Pr(\tau^+ < t \,\&\, \tau^- < \tau^+) = \Pr(\tau^- < t \,\&\, \tau^+ < \tau^-) = \Pr(\tau^+ < t \,\&\, \tau^+ < \tau^- < t)$$

so that

$$\begin{aligned}\Pr(\tau^+ < t) &= \Pr(\tau^+ < t \,\&\, \tau^- > t) + 2\Pr(\tau^+ < t \,\&\, \tau^+ < \tau^- < t)\\ &\geq \Pr(\tau^+ < t \,\&\, \tau^- > t) + \Pr(\tau^+ < t \,\&\, \tau^+ < \tau^- < t).\end{aligned}$$

Moreover,

$$\begin{aligned}\Pr(T < t) &= 2\Pr(\tau^+ < t \,\&\, \tau^- > t) + 2\Pr(\tau^+ < t \,\&\, \tau^+ < \tau^- < t)\\ &\leq 2\Pr(\tau^+ < t) = 4\Pr(V_t > A) = 4\Phi(-A/\sqrt{t}).\end{aligned}$$

Finally, $ET \geq t(1 - \Pr(T < t)) \geq t(1 - 4\Phi(-A/\sqrt{t}))$ for all t and, in particular, for $t = A^2/[\Phi^{-1}(\frac{1}{8})]^2$, which gives $ET \geq A^2/(2[\Phi^{-1}(\frac{1}{8})]^2)$. □

Now we consider a random walk with probability β of moving up or down by 1 and passage time K to $\pm\overline{\theta}L$.

THEOREM S.4. *The expected hitting time is bounded below by*

$$E\kappa \geq \frac{(\overline{\theta}/(2\beta))^2}{6[\Phi^{-1}(1/8)]^2} L^2.$$

PROOF. Let L_k be the random walk and consider the sums $S_L(t) = \sum_{k=1}^{t/L^2}(L_k - L_{k-1})/(2\beta L)$ as $L \to \infty$ converges weakly to a Weiner process with instantaneous variance 1. The random walk passes $\pm\overline{\theta}L$ when $S_L(t)$ passes $\pm\overline{\theta}/(2\beta)$. Considering the $\overline{T}$ truncated hitting time $\tilde{T}$, we have

$$E_S T \geq E_S\tilde{T} \geq E_W T - |E_W T - E_W\tilde{T}| - |E_W\tilde{T} - E_S\tilde{T}|,$$

where the final inequality is just the triangle inequality. However, $\lim_{L\to\infty} E_S\tilde{T} = E_W\tilde{T}$ and $\lim_{\overline{T}\to\infty} E_W\tilde{T} = E_W T$. So for all sufficiently large L, $\overline{T}$, we can make $|E_W T - E_W\tilde{T}|$, $|E_W\tilde{T} - E_S\tilde{T}|$ both less than or equal to $\frac{1}{3}$ the bound in Lemma S.3, giving the bound

$$E_S T \geq \left(1 - \frac{1}{3} - \frac{1}{3}\right)\frac{(\overline{\theta}/(2\beta))^2}{2[\Phi^{-1}(1/8)]}.$$

Finally, observe that the number of periods corresponding to T is $L^2 T$. □

Appendix S.C: Length of the fall, rise, and warring states

Here we prove the following proposition.

Proposition S.1. *For any K, there exists an $\overline{L}$ such that for all $L \geq \overline{L}$, there exists an $\overline{\epsilon}$ such that for all $\epsilon \leq \overline{\epsilon}$, the expected length of the warring states period exceeds that of either the fall or the rise by K periods.*

Proof. First the fall. From Appendix S.A, we see that the waypoints are where the hegemon loses a unit of land to opponents that consist entirely of a single society of zealots. Hence there are no more than $\overline{\theta}L$ waypoints. The time to failure is 1, since the hegemon can gain a unit of land with zero resistance and game over, and the least length of a least resistance path from the state after a waypoint to the next waypoint is 2: one transition to replace the society that initially gained the land with the zealots and one transition for the zealots to take a unit of land from the hegemon. Hence from Theorem S.1, we have the bound

$$E_\epsilon(t(a)|x, A) \leq \overline{\theta}LD3/C^6.$$

Turning to the rise, fix x such that a would be hegemon j has enough land $\theta_0 L$ to resist an opponent consisting entirely of zealots. Let r_z be that resistance. By Theorem S.2, we have the bound $P_\epsilon(A|x) \geq 1 - 2G(t_{xBW})\epsilon^{r_z}$. Moreover, the sets B_τ such that the hegemon has $\theta_0 L + \tau$ units of land form a Liapunoff sequence. Notice that for this sequence $t_{FF}(A) = 1$, since there is always zero resistance to the hegemon gaining a single unit of land, and along a least resistance path starting at x, he can never lose any land. Hence by Theorem S.3, we also have the bound

$$\begin{aligned} E_\epsilon(t(a)|x, A) &\leq (1-\theta_0)L\frac{1}{P_\epsilon(A|x)}\frac{3}{C^2} \\ &\leq (1-\theta_0)L\frac{1}{1-2G(t_{xBW})\epsilon^{r_z}}\frac{3}{C^2} \end{aligned}$$

during the rise.

Recall that at some point during the warring states period, there is a society with $L_{j\tau}$ units of land that follows a random walk with β chance of increasing by 1 or decreasing by 1 at least until either $L_{j\tau} \geq \overline{\theta}L$ or $L_{j\tau} \leq (1-\overline{\theta})L$. From Theorem S.4, we have the expected passage time bound

$$E_\epsilon \kappa \geq \frac{(\overline{\theta}/(2\beta))^2}{6[\Phi^{-1}(1/8)]^2}L^2.$$

Hence for L sufficiently large the expected amount of time in the warring states is $3K$ larger than an upper bound $\overline{\theta}LD3/C^6$ on the expected amount of time during a least resistance path during the fall and larger than $(1-\theta_0)L3/C^2$, which is not quite an upper bound on the expected amount of time during the rise. This is not quite the end of the story, since it is the expected amount of time of all paths during the rise or the fall that matters, and because we must account for dividing by the probability of the rise.

However, the expected length of all non-least-resistance paths is bounded above by the bound in Appendix A to the paper as is $G(t_{xBW})$, and while that bound increases quite rapidly with L, it is also weighted according to that theorem by a probability that goes to zero with ϵ. Hence once we fix L, we can choose a small enough ϵ that the expected length of all paths (during the rise or the fall) is at most K larger than that of the length of least resistance paths; that is, of total length at most $2K$. Hence the expected amount of time in the warring states period is at least K larger than during the rise or fall. □

Appendix S.D: Ergodic probabilities and circuits

We are given a finite set of nodes Ω^k and for $\psi, \phi \in \Omega^k$, a resistance function $r^k(\psi, \phi)$. For any $\psi \in \Omega^k$, we define the *least resistance* $r^k(\psi) = \min_{\phi \in \Omega^k \setminus \psi} r^k(\psi, \phi)$. We are interested in trees T on Ω^k. For any such tree and any ψ, let $T(\psi)$ denote the unique predecessor of ψ on the tree (which is null for the unique root). Note that we follow the standard game theory terminology that the predecessor is closer to the root—in contrast to Young, who follows the logic of the Markov process in imagining that the node closer to the root is the successor node. The *resistance of the tree* T is defined to be $r^k(T) = \sum_{\psi \in \Omega^k} r^k(\psi, T(\psi))$, where $r^k(\psi, \emptyset) \equiv 0$.

Our goal is to characterize least resistance trees by showing how they are constructed out of groups of nodes that we call circuits. As in the text, $\Omega_x^{k+1} \subseteq \Omega^k$ is a *circuit* if for each pair $\psi_1, \psi_y \in \Omega_x^{k+1}$, there is a path $\psi_1, \psi_2, \ldots, \psi_n \in \Omega_x^{k+1}$ with $\psi_n = \psi_y$ such that for $\tau = 2, 3, \ldots, n$ we have $r^k(\psi_{\tau-1}, \psi_\tau) = r^k(\psi_{\tau-1})$, that is, there is a path from ψ_1 to ψ_y within the circuit such that each connection has least resistance.

Definition S.1 (Consolidation). A circuit Ω_x^{k+1} is *consolidated within the tree* T if there is a $\phi \in \Omega_x^{k+1}$ that precedes all other $\psi \in \Omega_x^{k+1}$, and for these other $\psi \neq \phi$, we have $T(\psi) \in \Omega_x^{k+1}$ and $r^k(\psi, T(\psi)) = r^k(\psi)$.

In other words, in the consolidated tree, the circuit Ω_x^{k+1} forms a subtree with root ϕ, and each connection within the circuit has least resistance. We refer to ϕ as the *top of the circuit*.

Intuitively, if we think of the circuit as a circle of least resistance connections, then we will break that circle after ϕ to make a subtree and use ϕ to connect this subtree to the rest of the tree. Breaking the connection saves at least $r^k(\phi)$, while making the new connection costs $r^k(\phi, T(\phi))$; hence we define the *modified resistance from* ϕ *to* ψ as $R^k(\phi, \psi) = r^k(\phi, \psi) - r^k(\phi)$.

In the next lemma, we consolidate a circuit within a tree by breaking it after the node that minimizes modified resistance. By so doing, the resistance of the tree cannot increase.

Lemma S.4. *Suppose that T on Ω^k has root ψ and that Ω_x^{k+1} is a circuit on Ω^k. Then there is a tree T' with root ψ such that $r^k(T') \le r^k(T)$ and Ω_x^{k+1} is consolidated in T' with the additional properties that (i) if $\phi' \notin \Omega_x^{k+1}$, then $T'(\phi') = T(\phi')$, and (ii) if ϕ is the top of Ω_x^{k+1} in T', then $R^k(\phi, T'(\phi)) = \min\{R^k(\phi', T'(\phi)) \mid \phi' \in \Omega_x^{k+1}\}$.*

Proof. Let T have root ψ and let $\phi^* \in \Omega_x^{k+1}$ be such that the unique path from ϕ^* to the root ψ contains no element of Ω_x^{k+1}. If $\phi^* = \psi$, take $\phi = \phi^*$. Otherwise, choose as the top a $\phi \in \Omega_x^{k+1}$ such that $r^k(\phi, T(\phi^*)) - r^k(\phi) = \min\{r^k(\phi', T(\phi^*)) - r^k(\phi') \mid \phi' \in \Omega_x^{k+1}\}$. We now use tree surgery to create a sequence of new trees ending in the desired tree T'. As we proceed, we never cut a connection originating in any set other than Ω_x^{k+1}, so that property (i) will be satisfied.

At each step, Ω_x^{k+1} will be divided into two sets $\Phi_\phi, \Phi_{\sim\phi} = \Omega_x^{k+1} \setminus \Phi_\phi$. The first set Φ_ϕ will contain at least ϕ, and consists of those elements of Ω_x^{k+1} that are already consolidated with ϕ at the top and such that no element of $\Phi_{\sim\phi}$ appears between ϕ and the root. We will proceed constructing new trees by moving one element from $\Phi_{\sim\phi}$ to Φ_ϕ at a time, making sure that all properties are preserved.

We start the process. If $\phi = \psi$ or $\phi = \phi^*$, we do nothing. Otherwise, cut ϕ from the tree and paste it to $T(\phi^*)$. Observe that this increases the resistance of the tree by at most $r^k(\phi, T(\phi^*)) - r^k(\phi)$. Let Φ_ϕ be the maximal set consolidated with ϕ at the top: this set now contains at least ϕ.

We now continue the process until $\Phi_{\sim\phi}$ is empty. Pick an element $\phi' \in \Phi_{\sim\phi}$. Because Ω_x^{k+1} is a circuit, there is a least resistance path in Ω_x^{k+1} from ϕ' to ϕ. Let ϕ_τ be the last element in $\Phi_{\sim\phi}$ that is reached on this path. Then cut ϕ_τ from the tree and paste it to $\phi_{\tau+1}$. Notice that this cannot increase the resistance of the tree, since the connection from ϕ_τ to $\phi_{\tau+1}$ has least resistance. Moreover, if $\phi \neq \phi^*$, then at some step $\phi_\tau = \phi^*$ and at this step the resistance of the tree is decreased by exactly $r^k(\phi^*, T(\phi^*)) - r^k(\phi^*)$. Once again let Φ_ϕ be the maximal set consolidated with ϕ at the top: this set now contains at least one more element ϕ_τ.

When we are finished we end up with the new tree T'. Now observe that either $\phi = \phi^*$ or the resistance over the original tree was increased only in the first step, by at most $r^k(\phi, T(\phi^*)) - r^k(\phi)$, and it was decreased by $r^k(\phi*, T(\phi^*)) - r^k(\phi^*)$ when we pasted ϕ^*. By the choice of ϕ we have $r^k(\phi, T(\phi^*)) - r^k(\phi) \le r^k(\phi^*, T(\phi^*)) - r^k(\phi^*)$, and in all other cases, the resistance did not increase. Therefore, $r^k(T') \le r^k(T)$. Since, by construction, $T'(\phi) = T(\phi^*)$, we have $R^k(\phi, T'(\phi)) = \min\{R^k(\phi', T'(\phi)) \mid \phi' \in \Omega_x^{k+1}\}$. □

We now focus on least resistance trees. Let $\mathcal{T}(\psi)$ be the set of trees with root ψ, let $r_\psi^k = \min_{T\in\mathcal{T}(\psi)} r^k(T)$ be the least resistance of any tree with root ψ, and let $\mathcal{T}_\psi^k = \arg\min_{T\in\mathcal{T}(\psi)} r^k(T)$ be the set of least resistance trees with root ψ. First we prove a simple relation between least resistance of trees and of their roots.

Lemma S.5. *If ψ and ϕ are in the same circuit on Ω^k, then $r_\psi^k - r_\phi^k = r^k(\phi) - r^k(\psi)$.*

Proof. Suppose $\psi, \phi \in \Omega_x^{k+1}$, where Ω_x^{k+1} is a circuit. Then we can choose a path $\phi_1, \ldots, \phi_\nu, \ldots, \phi_n \in \Omega_x^{k+1}$ with $\phi_1 = \psi$, $\phi_\nu = \phi$, and $\phi_n = \psi$ such that for $\tau = 2, 3, \ldots, n$, we have $r^k(\phi_{\tau-1}, \phi_\tau) = r^k(\phi_{\tau-1})$. Choose $T_1 \in \mathcal{T}_{\phi_1}$ and, supposing that $T_{\tau-1}$ has root $\phi_{\tau-1}$, define T_τ as the tree in which we cut ϕ_τ from $T_{\tau-1}$, make it the root of T_τ, and paste the root of $T_{\tau-1}$ to ϕ_τ. This tree has root ϕ_τ and resistance $r^k(T_\tau) \le r^k(T_{\tau-1}) + r^k(\phi_{\tau-1}, \phi_\tau) - r^k(\phi_\tau) = r^k(T_{\tau-1}) + r^k(\phi_{\tau-1}) - r^k(\phi_\tau)$. Hence $r^k(T_\tau) \le r^k(T_1) + r^k(\phi_1) -$

$r^k(\phi_\tau)$. Since $\phi_n = \phi_1$, we conclude that $r^k(T_n) \leq r^k(T_1)$, and since T_1 had least resistance, it must be that $r^k(T_n) = r^k(T_1)$. Hence all the inequalities must hold with equality, that is, $r^k(T_\tau) = r^k(T_1) + r^k(\phi_1) - r^k(\phi_\tau)$. Choosing $\tau = \nu$, we then have $r^k(T_\tau) = r^k_\psi + r^k(\psi) - r^k(\phi)$, whence $r^k_\phi \leq r^k_\psi + r^k(\psi) - r^k(\phi)$; but by interchanging ϕ and ψ, and rearranging, we get $r^k_\phi \geq r^k_\psi + r^k(\psi) - r^k(\phi)$. This gives the conclusion. □

We now assume that for $\epsilon > 0$, P_ϵ is ergodic so that there is a unique ergodic probability distribution μ_ϵ on the state space Z. Let $\mathcal{T}_S(x)$ denote all trees over a set S with root x and set

$$\mathcal{M}_\epsilon(x) = \sum_{T \in \mathcal{T}_Z(x)} \prod_{z \in Z} P_\epsilon(T(z)|z).$$

Following Young (1993) and Friedlin and Wentzell (2012), we observe that

$$\mu_\epsilon(x) = \frac{\mathcal{M}_\epsilon(x)}{\sum_{z \in Z} \mathcal{M}_\epsilon(z)}.$$

Let the resistance $r(x, y)$ on Z be the ordinary resistance. Let r_x be the least resistance of trees on Z with root x. Observing from Cayley's formula that N^{N-2} is the number of trees with the same root over N nodes, Theorem S.5 follows.

Theorem S.5. *The ratio of ergodic probabilities satisfies the bounds*

$$\frac{C^N}{N^{N-2} D^N} \epsilon^{r_x - r_y} \leq \frac{\mu_\epsilon(x)}{\mu_\epsilon(y)} \leq \frac{N^{N-2} D^N}{C^N} \epsilon^{r_x - r_y}.$$

Proof. We may rearrange the Friedlin and Wentzell (2012) result to get

$$\mu_\epsilon(x) \sum_{z \in Z} \mathcal{M}_\epsilon(z) = \mathcal{M}_\epsilon(x)$$

so that

$$\frac{\mu_\epsilon(x)}{\mu_\epsilon(y)} = \frac{\mathcal{M}_\epsilon(x)}{\mathcal{M}_\epsilon(y)}.$$

Recall the bounds $C\epsilon^{r(x,z)} \leq P_\epsilon(z|x) \leq D\epsilon^{r(x,z)}$ on transition probabilities. Hence we have

$$C^N \epsilon^{r_x} \leq \sum_{T \in \mathcal{T}_Z(x)} C^N \prod_{x \in Z} \epsilon^{r(x,z)} \leq \mathcal{M}_\epsilon(x) \leq \sum_{T \in \mathcal{T}_Z(x)} D^N \prod_{x \in Z} \epsilon^{r(x,z)} \leq D^N \epsilon^{r_x} N^{N-2}.$$

Dividing by $\mathcal{M}_\epsilon(y)$ and using the corresponding bounds then gives the result. □

These bounds are in terms of resistances of least resistance trees. The next goal is to translate them in terms of appropriate resistances of least resistance paths.

Applying Lemma S.5 gives as an immediate corollary the following result, where recall that $r^0(\Omega_x)$ is defined in terms of direct routes.

Theorem S.6. *If the recurrent communicating classes Ω_x and Ω_y are in the same circuit on $\Omega^0 \equiv \Omega$ then*

$$\frac{C^N}{N^{N-2}D^N}\epsilon^{r^0(\Omega_y)-r^0(\Omega_x)} \le \frac{\mu_\epsilon(x)}{\mu_\epsilon(y)} \le \frac{N^{N-2}D^N}{C^N}\epsilon^{r^0(\Omega_y)-r^0(\Omega_x)}.$$

This goes one step in the desired direction but applies only to elements of a given circuit. In general, we can find the least resistance of trees in Z by finding the least resistance of trees in Ω. Recall that $r^0_{\Omega_x}$ is the least resistance of trees on Ω with root Ω_x, and r_x is the least resistance of trees on Z with root x. We next show that they are equal:

Lemma S.6. *If $x \in \Omega_x \in \Omega$ then $r_x = r^0_{\Omega_x}$.*

Proof. Young (1993) proves this lemma (Lemma 2 in his Appendix) for the case where the resistance, call it $r^*(\Omega_x, \Omega_y)$, is the least resistance of any path from Ω_x to Ω_y; that is, he allows the path to pass through recurrent communicating classes Ω_z, which are neither Ω_x nor Ω_z (Ellison 2000 does the same in his definition of the modified co-radius). Our resistance is, in general, larger than Young's, since we do not allow paths to pass through these other recurrent communicating classes. However, his proof requires only minor modification to yield the stronger result. Young first shows that the least resistance $r^*_{\Omega_x}$ of any tree on Ω with root Ω_x is greater than or equal to r_x. Since $r^0_{\Omega_x} \ge r^*_{\Omega_x}$, we have the immediate implication that $r^0_{\Omega_x} \ge r_x$.

The second part of Young's proof shows that $r^*_{\Omega_x} \le r_x$. Following Young, we show how to transform a least resistance tree $T \in \mathcal{T}_x$ on Z into a tree $T' \in \mathcal{T}(\Omega_x)$ over Ω such that $r^0(T') \le r^0(T)$. The easiest way to do this would be by simply taking one point from each irreducible class and using the resistance between those states to get a tree over Ω. However, this does not work because there can be double-counting if paths in T join between irreducible classes. Young shows how to avoid double-counting by reorganizing the tree. We can use his construction if we can avoid having or creating paths between irreducible classes that contain elements of a third irreducible class. This is the case if we start by choosing the "right" least resistance tree and the "right" point from each irreducible class before we apply Young's method.

Observe that each $\phi \in \Omega$ is a circuit, so by consolidating where needed as from Lemma S.4, we can assume that each $\phi \in \Omega$ is already consolidated in T. The first step of Young's proof is to choose one point $y' \in \phi$ for each $\phi \in \Omega$: these are what Young calls *special vertices*. We do this by choosing for each $\phi \in \Omega$, the top of ϕ in the tree. Observe that because the tree is consolidated, the path from any special vertex to the next special vertex y in the direction of the root cannot contain elements of any irreducible class other than Ω_y.

Now apply Young's construction to eliminate junctions (a *junction* in a tree T is any vertex y with at least two incoming T edges). Observe that when Young cuts a subtree T^* from a vertex y that is not in a recurrent communicating class, this preserves the consolidated structure, because those $\phi' \in \Omega$ that lie farther from the root than y are necessarily entirely contained in T^*. Consequently, we never need to cut junctions at y that are in recurrent communicating classes, for T is consolidated and, therefore, the path from y to the top of the circuit has zero resistance and no double-counting is involved.

Finally, when Young pastes cuts T^* from the junction y back into the tree T, he implicitly introduces new paths $a = (y, z_1, \ldots, z_{t-1}, z)$ from y to a special vertex z with $r(a) = 0$. However, these implicit paths cannot contain elements of any recurrent communicating class Ω_y other than Ω_z. If they did, the path could not have zero resistance, since there is no path from $\Omega_y \neq \Omega_z$ to Ω_z that has zero resistance. Hence at the end of Young's procedure, we find that the paths along which resistance is computed—those from one special vertex to the next special vertex in the direction of the root—do not contain a vertex from a third recurrent communicating class. By this procedure, we then obtain a tree in $\mathcal{T}(\Omega_x)$ with resistance not larger than T, whence $r^0_{\Omega_x} \leq r_x$. $\square$

Our next goal is to recursively compute r^k and by doing so find bounds on $\mu_\epsilon(x)/\mu_\epsilon(y)$ without the restriction that Ω_x and Ω_y be in the same circuit.

We take $\Omega^0 = \Omega$, so an element $\psi^1 \in \Omega^1$ will be a circuit of recurrent communicating classes and for $\psi, \phi \in \Omega^0$, the resistance $r^0(\psi, \phi)$ is just the least resistance along a direct route. We recursively define on Ω^{k-1} the modified resistance function $R^{k-1}(\psi^{k-1}, \phi^{k-1}) = r^{k-1}(\psi^{k-1}, \phi^{k-1}) - r^{k-1}(\psi^{k-1})$, and we define a resistance function on Ω^k by the least modified resistance:

$$r^k(\psi^k, \phi^k) = \min_{\psi^{k-1} \in \psi^k, \phi^{k-1} \in \phi^k} R^{k-1}(\psi^{k-1}, \phi^{k-1}).$$

Then the following formula holds, where we notice that the term $\sum_{\phi^{k-1} \in \Omega^{k-1}} r^{k-1}(\phi^{k-1})$ is a constant independent of the tree in question.

Lemma S.7. *If $\psi^{k-1} \in \psi^k$, then $r^{k-1}_{\psi^{k-1}} = r^k_{\psi^k} - r^{k-1}(\psi^{k-1}) + \sum_{\phi^{k-1} \in \Omega^{k-1}} r^{k-1}(\phi^{k-1})$.*

Proof. Suppose we have a tree T^{k-1} on Ω^{k-1} that is consolidated with respect to all the circuits in Ω^k, and let ψ^{k-1} be its root. The fact that T^{k-1} is consolidated means that the top of each circuit has a predecessor that belongs to a different circuit. For $\psi^k \in \Omega^k$, denote by $\Gamma(T^{k-1}, \psi^k) \in \Omega^{k-1}$ the top of circuit ψ^k in T^{k-1}. Then if $T^{k-1}(\Gamma(T^{k-1}, \psi^k)) = \phi^{k-1} \in \phi^k \neq \psi^k$ (where if ϕ^{k-1} is null, we set $\phi^k = \emptyset$ as well), we may define $T^k(\psi^k) = \phi^k$. In this way we define a tree on Ω^k. We have $r^{k-1}(T^{k-1}) = \sum_{\phi^{k-1} \in \Omega^{k-1}} r^{k-1}(\phi^{k-1}, T^{k-1}(\phi^{k-1}))$. However, since the tree is consolidated, for any ϕ^{k-1} not at the top of the corresponding circuit ϕ^k, we have $r^{k-1}(\phi^{k-1}, T^{k-1}(\phi^{k-1})) = r^{k-1}(\phi^{k-1})$; hence we may write

$$\begin{aligned} r^{k-1}(T^{k-1}) = \sum_{\phi^{k-1} \in \Psi^{k-1}} r^{k-1}(\phi^{k-1}) - r^{k-1}(\psi^{k-1}) \\ + \sum_{\phi^k \in \Psi^k} \rho^{k-1}\big(\Gamma(T^{k-1}, \phi^k), T^{k-1}(\Gamma(T^{k-1}, \phi^k))\big). \end{aligned}$$

Now start with a least resistance tree $T^{k-1} \in \mathcal{T}_{\psi^{k-1}}$. By Lemma S.4, we may consolidate this tree T^{k-1} with respect to all the circuits in Ω^k to get another least resistance tree $\tilde{T}^{k-1} \in \mathcal{T}_{\psi^{k-1}}$. By the previous computation and the definition of r^k, we see that

$$\begin{aligned}
r^{k-1}_{\psi^{k-1}} &= r^{k-1}(\tilde{T}^{k-1}) \\
&= \sum_{\phi^{k-1} \in \Psi^{k-1}} r^{k-1}(\phi^{k-1}) - r^{k-1}(\psi^{k-1}) \\
&\qquad\qquad + \sum_{\phi^k \in \Psi^k} \rho^{k-1}\big(\Gamma(T^{k-1}, \phi^k), T^{k-1}(\Gamma(T^{k-1}, \phi^k))\big) \\
&\geq \sum_{\phi^{k-1} \in \Psi^{k-1}} r^{k-1}(\phi^{k-1}) - r^{k-1}(\psi^{k-1}) + \sum_{\phi^k \in \Psi^k} r^k(\phi^k, T^k(\phi^k)) \\
&\geq \sum_{\phi^{k-1} \in \Psi^{k-1}} r^{k-1}(\phi^{k-1}) - r^{k-1}(\psi^{k-1}) + r^k_{\psi^k}.
\end{aligned}$$

Next start with a least resistance tree $T^k \in \mathcal{T}_{\Omega^k_x}$, where $\psi^{k-1} \in \psi^k$, and construct a tree on Ω^{k-1} as follows. For the root $\phi^k = \psi^k$, define $\phi^{k-1} = \psi^{k-1}$. For given non-root ϕ^k and $T^k(\phi^k)$ there are points $\phi^{k-1} \in \phi^k$ and $\tilde{\phi}^{k-1} \in T^k(\phi^k)$ such that $r^k(\phi^k, T^k(\phi^k)) = r(\phi^{k-1}, \tilde{\phi}^{k-1}) - r(\phi^{k-1})$. For each ϕ^k, consolidate the tree over ϕ^k with root ϕ^{k-1} to get a tree $T[\phi^k, \phi^{k-1}]$. Now define a tree on Ω^{k-1} by putting together these subtrees as follows: if $\hat{\phi}^{k-1}$ is in $T[\phi^k, \phi^{k-1}]$ but is not the root, set $T^{k-1}(\hat{\phi}^{k-1}) = T[\phi^k, \phi^{k-1}](\hat{\phi}^{k-1})$. For the root ϕ^{k-1}, set $T^{k-1}(\hat{\phi}^{k-1}) = \tilde{\phi}^{k-1}$. This is clearly a tree with root ψ^{k-1}, and we see that the resistance is

$$\begin{aligned}
r^{k-1}_{\psi^{k-1}} &\leq r^{k-1}(T^{k-1}) \\
&= \sum_{\phi^{k-1} \in \Psi^{k-1}} r^{k-1}(\phi^{k-1}) - r^{k-1}(\psi^{k-1}) + \sum_{\phi^k \in \Omega^k} r^k(\phi^k, T^k(\phi^k)) \\
&= \sum_{\phi^{k-1} \in \Omega^{k-1}} r^{k-1}(\phi^{k-1}) - r^{k-1}(\psi^{k-1}) + r^k_{\psi^k}.
\end{aligned}$$

Putting together the two inequalities gives the desired result. □

LEMMA S.8. *If Ω^k has at least two elements, it has at least one nontrivial circuit.*

PROOF. Starting at an arbitrary point $\psi^k \in \Omega^k$, choose a path of least resistance. Since Ω^k is finite, this must eventually have a loop and that loop is necessarily a circuit. □

We can now recursively define a class of reverse filtrations with resistances over the set $\Omega^0 = \Omega$ of recurrent communicating classes for P_0; assume Ω has N_Ω elements, with $N_\Omega \geq 2$. Starting with Ω^{k-1}, we observe that there is at least one nontrivial circuit and that every singleton element is trivially a circuit. Hence we can form a nontrivial partition of Ω^{k-1} into circuits and denote this partition Ω^k. All the resistances are defined as before. Note that since each partition is nontrivial, this construction has at most $k \leq N_\Omega$ layers before the partition has a single element and the construction stops.

The *modified radius* of $x \in \Omega_x$ of order k is defined by

$$\overline{R}^k(x) = \sum_{\kappa=0}^{k} r^\kappa(\Omega_x^\kappa),$$

where $\Omega_x^0 = \Omega_x$ and for each $\kappa > 0$, the element $\Omega_x^\kappa \ni \Omega_x^{\kappa-1}$. Then we can state another theorem.

Theorem S.7. *Let k be such that $\Omega_x^k = \Omega_y^k$. Then $r_x - r_y = \overline{R}^{k-1}(y) - \overline{R}^{k-1}(x)$ and, consequently,*

$$\frac{C^N}{N^{N-2}D^N}\epsilon^{\overline{R}^{k-1}(y)-\overline{R}^{k-1}(x)} \leq \frac{\mu_\epsilon(x)}{\mu_\epsilon(y)} \leq \frac{N^{N-2}D^N}{C^N}\epsilon^{\overline{R}^{k-1}(y)-\overline{R}^{k-1}(x)}.$$

Proof. From Lemma S.6, we know that $r_x - r_y = r^0_{\psi^0(x)} - r^0_{\psi^0(y)}$. Applying Lemma S.7 iteratively, we see that if $\psi^{k-1} \in \psi^k$, then

$$r^0_{\psi^0} = r^k_{\Omega_x^k} + \sum_{\kappa=0}^{k-1}\Bigg[\sum_{\phi^\kappa \in \Omega^\kappa} r^\kappa(\phi^\kappa)\Bigg] - \sum_{\kappa=0}^{k-1} r^\kappa(\psi^\kappa),$$

from which

$$r^0_{\psi^0(x)} - r^0_{\psi^0(y)} = -\sum_{\kappa=0}^{k-1} r^\kappa(\psi^\kappa(x)) + \sum_{\kappa=0}^{k-1} r^\kappa(\psi^\kappa(y)) = \overline{R}^{k-1}(y) - \overline{R}^{k-1}(x).$$

□

References

Ellison, Glenn (2000), "Basins of attraction, long-run stochastic stability, and the speed of step-by-step by evolution." *Review of Economic Studies*, 67, 17–45. [10]

Freidlin, Mark I. and Alexander D. Wentzell (2012), *Random Perturbations of Dynamical Systems*, third edition. Springer, Heidelberg. [9]

Young, H. Peyton (1993), "The evolution of conventions." *Econometrica*, 61, 57–84. [9, 10]

Submitted 2014-9-22. Final version accepted 2014-12-16. Available online 2014-12-17.