EXTRA APPENDIX FOR REFEREES: Proofs of two lemmas by Pemantle (1990) for "A Herding Perspective..."

James S. Costain
Bank of Spain and Univ. of California, Davis

First version: March 2006
This version: July 2006

The proof of Proposition 4 in "A Herding Perspective..." relies on two results from Pemantle (1990).[1] Pemantle's paper shows that convergence to an unstable crossing occurs with probability zero. For "A Herding Perspective...", we need stronger results: we need to know the speed at which the stochastic process escapes the unstable steady state. Actually, a mild strengthening of Pemantle's arguments suffices to characterize the speed of escape. But since he did not need this information for the purposes of his paper, he stated his results in a simpler way that did not make the speed of escape explicit. In these notes, I prove versions of the two Pemantle lemmas that spell out the speed of escape explicitly, following Pemantle's calculations wherever possible. I use the notation and definitions from "A Herding Perspective..." without restating all the assumptions of that paper.

The first lemma states that for sufficiently large i, there is a strictly positive probability of escaping from the vicinity of any unstable steady state S^u. The second shows that conditional on this escape, there is a strictly positive probability of never returning to a smaller neighborhood of S^u.

[1] Robin Pemantle (1990), "Nonconvergence to Unstable Points in Urn Models and Stochastic Approximatinons." *Annals of Probability* 18 (2), pp. 698-712.

Pemantle (1990), Lemma 1. For any $B > 0$, $i' > i$, and arbitrary initial conditions S_i,

$$prob\left(|S_{i'} - S^u| > \frac{B}{\sqrt{i'}}\middle| S_i\right) \geq \frac{\left(\frac{i'}{i}\right)^{2\lambda}(S_i - S^u)^2 + \frac{c^*}{2\lambda}\left(\left(\frac{i'}{i}\right)^{2\lambda} - 1\right) - \frac{B^2}{i'}}{\left(\log\left(\frac{i'}{i}\right) + |S_i - S^u|\right)^2 - \frac{B^2}{i'}} \tag{1}$$

Proof. Write the stochastic process as $S_{i+1} = S_i + \frac{1}{i+1}(T(S_i) - S_i) + \frac{1}{i+1}\epsilon_{i+1}$, where $\epsilon_{i+1} \equiv x_{i+1} - T(S_i)$ is a mean zero shock independent of S_i. Note that under our assumptions,

$$0 \leq \min\{1 - F(\sigma^*|0.5 - \bar{Z}), F(-\sigma^*|0.5 + \bar{Z})\} \leq |\epsilon| \leq 1 \tag{2}$$

and so $0 \leq c^* \equiv [\min\{1 - F(\sigma^*|0.5 - \bar{Z}), F(-\sigma^*|0.5 + \bar{Z})\}]^2 \leq E(\epsilon^2) \leq 1$.

Consider a neighborhood of S^u in which $\frac{\partial T}{\partial S} \geq 1 + \lambda \geq 1$. Since by definition $T(S^u) = S^u$, we have $\frac{T(S)-S}{S-S^u} = \frac{T(S)-S^u}{S-S^u} + \frac{S^u-S}{S-S^u} \geq \lambda$ at all points in the neighborhood. To simplify exposition, we assume for the rest of this proof that $\frac{\partial T}{\partial S} \geq 1 + \lambda$ at all $S \in [0, 1]$ (which implies that T crosses the 45^o degree line only once, in an unstable fashion). This assumption is incorrect, but harmless. Non-convergence to S^u is a local property, so it cannot depend on the properties of T at points far from S^u (see Pemantle 1990).

We next investigate how the expected distance of S_i from S^u varies with i.

$$E\{(S_k - S^u)^2 \,|\, S_{k-1}\} = E\left\{\left(S_{k-1} - S_u + \frac{1}{k}(T(S_{k-1}) - S_{k-1}) + \frac{\epsilon_k}{k}\right)^2 \middle| S_{k-1}\right\}$$

$$= (S_{k-1} - S_u)^2 + \left(\frac{T(S_{k-1}) - S_{k-1}}{k}\right)^2 + E\left\{\left(\frac{\epsilon_k}{k}\right)^2 \middle| S_{k-1}\right\} + 2(S_{k-1} - S^u)\left(\frac{T(S_{k-1}) - S_{k-1}}{k}\right)$$

$$\geq \frac{c^*}{k^2} + \left(1 + \frac{\lambda}{k}\right)^2 (S_{k-1} - S^u)^2 \tag{3}$$

Now iterate out to some $i' > i$:

$$E\left\{(S_{i'} - S^u)^2 \middle| S_i\right\} = E\left\{E\left[(S_{i'} - S^u)^2 | S_{i'-1}\right] \middle| S_i\right\} \geq \left\{\frac{c^*}{(i')^2} + \left(1 + \frac{\lambda}{i'}\right)^2 E\left[(S_{i'-1} - S^u)^2 \middle| S_i\right]\right\}$$

$$\geq \frac{c^*}{(i')^2} + \left(1 + \frac{\lambda}{i'}\right)^2 \frac{c^*}{(i'-1)^2} + \left(1 + \frac{\lambda}{i'}\right)^2 \left(1 + \frac{\lambda}{i'-1}\right)^2 \frac{c^*}{(i'-2)^2} + \ldots$$

$$+ \left(1 + \frac{\lambda}{i'}\right)^2 \ldots \left(1 + \frac{\lambda}{i+1}\right)^2 \frac{c^*}{(i+1)^2} + \left(1 + \frac{\lambda}{i'}\right)^2 \left(1 + \frac{\lambda}{i'-1}\right)^2 \ldots \left(1 + \frac{\lambda}{i+1}\right)^2 (S_i - S^u)^2 \tag{4}$$

Repeatedly using the approximation

$$\left(1+\frac{\lambda}{i'}\right)\left(1+\frac{\lambda}{i'-1}\right)...\left(1+\frac{\lambda}{k+1}\right) \;\approx\; \exp\left[\lambda\sum_{j=k+1}^{i'}\frac{1}{j}\right] \;\approx\; \exp\left[\lambda(\log i' - \log k)\right] \;=\; \left(\frac{i'}{k}\right)^{\lambda} \tag{5}$$

(which is arbitrarily accurate for sufficiently large k), we obtain

$$E\left\{(S_{i'}-S^u)^2\Big| S_i\right\} \;\geq\; \frac{c^*}{2\lambda}\left(\left(\frac{i'}{i}\right)^{2\lambda}-1\right) \;+\; \left(\frac{i'}{i}\right)^{2\lambda}(S_i-S^u)^2 \tag{6}$$

We also need an upper bound to accompany this lower bound. Since each step is of size $1/i$, we have

$$|S_{i'}-S^u| \;\leq\; \sum_{j=i+1}^{i'}\frac{1}{j}+|S_i-S^u| \;\approx\; \log\left(\frac{i'}{i}\right)+|S_i-S^u|$$

Putting together the lower and upper bounds, we have

$$\frac{c^*}{2\lambda}\left(\left(\frac{i'}{i}\right)^{2\lambda}-1\right) \;+\; \left(\frac{i'}{i}\right)^{2\lambda}(S_i-S^u)^2 \;\leq\; E\left\{(S_{i'}-S^u)^2\Big| S_i\right\}$$

$$\equiv\; prob\left((S_{i'}-S^u)^2>\frac{B^2}{i'}\Bigg| S_i\right) E\left\{(S_{i'}-S^u)^2\Big| S_i,(S_{i'}-S^u)^2>\frac{B^2}{i'}\right\}$$

$$+\, prob\left((S_{i'}-S^u)^2<\frac{B^2}{i'}\Bigg| S_i\right) E\left\{(S_{i'}-S^u)^2\Big| S_i,(S_{i'}-S^u)^2<\frac{B^2}{i'}\right\}$$

$$\leq\; prob\left((S_{i'}-S^u)^2>\frac{B^2}{i'}\Bigg| S_i\right)\left[\log\left(\frac{i'}{i}\right)+|S_i-S^u|\right]^2 + \left[1-prob\left((S_{i'}-S^u)^2>\frac{B^2}{i'}\Bigg| S_i\right)\right]\frac{B^2}{i'} \tag{7}$$

Rearranging, we obtain the desired formula.
QED Pemantle Lemma 1.

Pemantle (1990), Lemma 2. For any $\epsilon \in (0,1)$, if we set $K-k = \sqrt{1-\epsilon}/\epsilon$, then for sufficiently large i:

$$prob\left(\inf_{j\geq i} |S_j - S^u| > \frac{k}{\sqrt{i}} \;\middle|\; S_i\,, |S_i - S^u| > \frac{K}{\sqrt{i}}\right) \;\geq\; 1-\epsilon \quad (8)$$

Proof. Suppose the distance of S_i from S^u satisfies $|S_i - S^u| \geq K/\sqrt{i}$ for some $K > 0$. Define $t \equiv \inf\{j \geq i : |S_j - S^u| \leq \frac{k}{\sqrt{i}}\}$. That is, t is the time of first return to the smaller interval $S^u \pm k/\sqrt{i}$. We wish to find a lower bound on $prob(t = \infty)$.

Note that the expected change in the distance of S_k from S^u is positive:

$$\begin{aligned} E\{|S_j - S^u|\,|S_{j-1}\} &= E\left\{\left|S_{j-1} - S^u + \frac{1}{j}[T(S_{j-1}) - S_{j-1} + \epsilon_j]\right|\;\middle|\; S_{j-1}\right\} \\ &\geq \left|E\left\{(S_{j-1} - S^u)\left[1 + \frac{1}{j}\left(\frac{T(S_{j-1}) - S_{j-1} + \epsilon_j}{S_{j-1} - S^u}\right)\right]\;\middle|\; S_{j-1}\right\}\right| \\ = \left|(S_{j-1} - S^u)\left[1 + \frac{1}{j}\left(\frac{T(S_{j-1}) - S_{j-1}}{S_{j-1} - S^u}\right)\right]\right| &\geq (1+\lambda/j)|(S_{j-1} - S^u)| \quad (9) \end{aligned}$$

Here we have used the fact that the absolute value is a convex function, and the fact that $\frac{\partial T}{\partial S} - 1 \geq \lambda \geq 0$. Now let Y_j be the martingale difference sequence representing unexpected changes in the distance of S_j from S^u. That is,

$$Y_j \;\equiv\; |S_j - S^u| - |S_{j-1} - S^u| - E\{|S_j - S^u| - |S_{j-1} - S^u|\,|S_{j-1}\}$$

By the triangle inequality,

$$|S_j - S^u| \equiv |S_{j-1} - S^u + \frac{1}{j}(x_j - S_{j-1})| \leq |S_{j-1} - S^u| + \frac{1}{j}|(x_j - S_{j-1})|$$

Since $|x_j - S_{j-1}| \leq 1$, we conclude that $|S_j - S^u| - |S_{j-1} - S^u| \leq \frac{1}{j}$, and therefore also that $|Y_j| \leq \frac{1}{j}$.

If we define Z_j for $j > i$ as $Z_j \equiv |S_i - S^u| + \sum_{l=i+1}^{j} Y_l$, then Z_j is a square-integrable martingale, and (9) implies that it satisfies $Z_j < |S_j - S^u|$ strictly for all $j > i$. Therefore, we cannot have $|S_j - S^u| \leq k/\sqrt{i}$ (which is the return condition) unless $Z_j \leq k/\sqrt{i}$ as well. Since $|S_i - S^u| > K/\sqrt{i}$ by assumption, we conclude that return therefore also requires $\sum_{l=i+1}^{j} Y_l \leq -(K-k)/\sqrt{i}$. Therefore we next investigate the distribution of $\sum_{l=i+1}^{j} Y_l$.

Consider $var(\sum_{j=i+1}^{t} Y_j)$. Given that $|Y_j| < \frac{1}{j}$, we have

$$var\left(\sum_{j=i+1}^{t} Y_j \middle| S_i\right) \;\leq\; var\left(\sum_{j=i+1}^{\infty} Y_j \middle| S_i\right) \;\leq\; \sum_{j=i+1}^{\infty} \frac{1}{j^2} \;\leq\; \int_i^{\infty} \frac{1}{j^2} dj \;=\; \frac{1}{i}$$

To obtain lower bound on the variance, note that

$$var\left(\sum_{j=i+1}^{t} Y_j \middle| S_i\right) \;=\; E\left[\left(\sum_{j=i+1}^{t} Y_j - E\left(\sum_{j=i+1}^{t} Y_j \middle| S_i\right)\right)^2 \middle| S_i\right] \;=\; E\left[\left(\sum_{j=i+1}^{t} Y_j\right)^2 \middle| S_i\right]$$

$$= \; E\left[\left(\sum_{j=i+1}^{t} Y_j\right)^2 \middle| S_i\,, t < \infty\right] prob(t < \infty|S_i) + E\left[\left(\sum_{j=i+1}^{\infty} Y_j\right)^2 \middle| S_i\,, t = \infty\right] prob(t = \infty|S_i)$$

$$\geq \; E\left[\left(\sum_{j=i+1}^{\infty} Y_j\right)^2 \middle| S_i\,, t = \infty\right] prob(t = \infty|S_i) \;\;\geq\;\; \left(E\left[\sum_{j=i+1}^{\infty} Y_j \middle| S_i\,, t = \infty\right]\right)^2 prob(t = \infty|S_i)$$

$$= \;\; \left(E\left[\sum_{j=i+1}^{\infty} Y_j \middle| S_i\,, t < \infty\right] \frac{prob(t < \infty|S_i)}{prob(t = \infty|S_i)}\right)^2 prob(t = \infty|S_i)$$

The second-to-last equality follows from the fact that the square is a convex function. The last one follows from the fact that the unconditional mean of $\sum_{j=i+1}^{t} Y_j$ is zero. Next, note that

$$E\left[\sum_{j=i+1}^{\infty} Y_j \middle| S_i\,, t < \infty\right] \;=\; E\left(E\left[\sum_{j=i+1}^{\infty} Y_j \middle| S_t\right] \middle| S_i\,, t < \infty\right)$$

$$= \; E\left(\sum_{j=i+1}^{t} Y_j \middle| S_i\,, t < \infty\right) \;\leq\; -\frac{(K-k)}{\sqrt{i}} \;<\; 0$$

by the definition of t. Therefore,

$$var\left(\sum_{j=i+1}^{t} Y_j \middle| S_i\right) \;\geq\; \frac{(K-k)^2}{i} \left(\frac{prob(t < \infty|S_i)^2}{prob(t = \infty|S_i)}\right)$$

Our lower and upper bounds on $var\left(\sum_{j=i+1}^{t} Y_j \middle| S_i\right)$ now imply that

$$\frac{[1 - prob(t = \infty|S_i)]^2}{prob(t = \infty|S_i)} \;\leq\; \frac{1}{(K-k)^2}$$

Note that the function $(1-x)^2/x$ is infinite at $x=0$, and is decreasing on the interval $[0,1]$, and equals 0 at $x=1$. Therefore the equation $(1-x)^2/x = 1/(K-k)^2$ has a unique solution x^* strictly between zero and one. In fact, since we imposed $(K-k) = \sqrt{1-\epsilon}/\epsilon$, we simply obtain $x^* = 1-\epsilon$. Therefore,

$$prob(t=\infty|S_i) \quad \geq \quad x^* \quad = \quad 1-\epsilon \quad > \quad 0$$

Q.E.D. Pemantle Lemma 2.