2. The averaging process

David Aldous

July 10, 2012

The two models in lecture 1 has only rather trivial interaction. In this lecture I discuss a simple FMIE process with less trivial interaction. It is intended as a prototype of the type of results one might seek to prove for other FMIE processes.

A write-up has been published as “A Lecture on the Averaging Process” (with Dan Lanoue) in recent *Probability Surveys*. The style of lecture and write-up are intended as illustration of how one might teach this subject in a graduate course.

Background meeting model with rates (ν_{ij}).

Model: averaging process.
Each agent initially has some amount of money;
whenever agents meet they share their money equally.
$X_i(t)$ is the amount of money agent i has at time t.

Formally, the states are the real numbers $\mathbb{R}$; initially $X_i(0) = x_i$, and the update rule, when agents i and j meet at t, is

$$(X_i(t+), X_j(t+)) = (\tfrac{1}{2}(X_i(t-) + X_j(t-)), \tfrac{1}{2}(X_i(t-) + X_j(t-))).$$

Your immediate reaction to this model should be (cf. General Principle 1) "obviously the individual values $X_i(t)$ converge to the average of initial values, so what is there to say?".

Exercise: write a one-sentence outline proof that a post-first-year-grad student could easily turn into a complete proof.

Curiously, while this process has been used as an ingredient in more elaborate models, the only place it appears by itself is in some "gossip algorithms" literature which derives a version of the "global bound" later – see paper for citations.

We will show

- If the initial configuration is a probability distribution (i.e. unit money split unevenly between individuals) then the vector of expectations in the averaging process evolves precisely as the probability distribution of the associated (continuous-time) Markov chain with that initial distribution (Lemma 1).
- There is an explicit bound on the closeness of the time-t configuration to the limit constant configuration (Proposition 1).
- Complementary to this global bound there is a "universal" (i.e. not depending on the meeting rates) bound for an appropriately defined *local* roughness of the time-t configuration (Propostion 2).
- There is a duality relationship with coupled Markov chains (Lemma 3).
- An entropy bound (Proposition 3).

The analysis in several ways parallels analysis of the well-known **Voter model** – will compare and contrast in the next lecture.

Basic properties of the averaging process

Write $I = \{i, j \ldots\}$ for the set of agents and $n \geq 2$ for the number of agents. Recall that the array of non-negative meeting rates $\nu_{\{i,j\}}$ for unordered pairs $\{i, j\}$ is assumed to be irreducible. We can rewrite the array as the symmetric matrix $\mathcal{N} = (\nu_{ij})$ in which

$$\nu_{ij} = \nu_{\{i,j\}}, \ j \neq i; \quad \nu_{ii} = -\sum_{j \neq i} \nu_{ij}. \tag{1}$$

Then $\mathcal{N}$ is the generator of the Markov chain with transition rates ν_{ij}; call this the *associated Markov chain*. The chain is reversible with uniform stationary distribution.

Throughout, we write $\mathbf{X}(t) = (X_i(t), i \in I)$ for the averaging process run from some non-random initial configuration $\mathbf{x}(0)$. Of course the sum is conserved: $\sum_i X_i(t) = \sum_i x_i(0)$.

Relation with the associated Markov chain

Write $\mathbf{1}_i$ for the initial configuration $(1_{(j=i)}, j \in I)$, that is agent i has unit money and other agents have none, and write $p_{ij}(t)$ for the transition probabilities of the associated Markov chain.

Lemma

For the averaging process with initial configuration $\mathbf{1}_i$ we have $\mathbb{E}X_j(t) = p_{ij}(t/2)$. More generally, from any deterministic initial configuration $\mathbf{x}(0)$, the expectations $\mathbf{x}(t) := \mathbb{E}\mathbf{X}(t)$ evolve exactly as the dynamical system

$$\frac{d}{dt}\mathbf{x}(t) = \tfrac{1}{2}\mathbf{x}(t)\mathcal{N}.$$

The time-t distribution $\mathbf{p}(t)$ of the associated Markov chain evolves as $\frac{d}{dt}\mathbf{p}(t) = \mathbf{p}(t)\mathcal{N}$. So if $\mathbf{x}(0)$ is a probability distribution over agents, then the *expectation* of the averaging process evolves as the distribution of the associated Markov chain started with distribution $\mathbf{x}(0)$ and slowed down by factor $1/2$. But keep in mind that the averaging process has more structure than this associated chain.

Proof. The key point is that we can rephrase the dynamics of the averaging process as

when two agents meet, each gives half their money to the other.

In informal language, this implies that the motion of a random penny - which at a meeting of its owner agent is given to the other agent with probability $1/2$ – is as the associated Markov chain at half speed, that is with transition rates $\nu_{ij}/2$.

To say this in symbols, we augment a random partition $\mathbf{X} = (X_i)$ of unit money over agents i by also recording the position U of the "random penny", required to satisfy

$$\mathbb{P}(U = i \mid \mathbf{X}) = X_i.$$

Given a configuration $\mathbf{x}$ and an edge e, write $\mathbf{x}^e$ for the configuration of the averaging process after a meeting of the agents comprising edge e. So we can define the augmented averaging process to have transitions

$$\begin{array}{lll} (\mathbf{x}, u) \to (\mathbf{x}^e, u) & \text{rate } \nu_e, & \text{if } u \notin e \\ (\mathbf{x}, u) \to (\mathbf{x}^e, u) & \text{rate } \nu_e/2, & \text{if } u \in e \\ (\mathbf{x}, u) \to (\mathbf{x}^e, u') & \text{rate } \nu_e/2, & \text{if } u \in e = (u, u'). \end{array}$$

This defines a process $(\mathbf{X}(t), U(t))$ consistent with the averaging process and (intuitively at least – see below) satisfying

$$\mathbb{P}(U(t) = i \mid \mathbf{X}(t)) = X_i(t). \tag{2}$$

The latter implies $\mathbb{E}X_i(t) = \mathbb{P}(U(t) = i)$, and clearly $U(t)$ evolves as the associated Markov chain slowed down by factor $1/2$. This establishes the first assertion of the lemma. The case of a general initial configuration follows via the following *linearity property* of the averaging process. Writing $\mathbf{X}(\mathbf{y}, t)$ for the averaging process with initial configuration $\mathbf{y}$, one can couple these processes as $\mathbf{y}$ varies by using the same realization of the underlying meeting process. Then clearly

$$\mathbf{y} \to \mathbf{X}(\mathbf{y}, t) \text{ is linear.}$$

How one writes down a careful proof of (2) depends on one's taste for details. We can explicitly construct $U(t)$ in terms of "keep or give" events at each meeting, and pass to the embedded jump chain of the meeting process, in which time m is the time of the m'th meeting and $\mathcal{F}_m$ its natural filtration. Then on the event that the m'th meeting involves i and j,

$$\mathbb{P}(U(m) = i \mid \mathcal{F}_m) = \tfrac{1}{2}\mathbb{P}(U(m-1) = i \mid \mathcal{F}_{m-1}) + \tfrac{1}{2}\mathbb{P}(U(m-1) = j \mid \mathcal{F}_{m-1})$$

$$X_i(m) = \tfrac{1}{2}X_i(m-1) + \tfrac{1}{2}X_j(m-1)$$

and so inductively we have

$$\mathbb{P}(U(m) = i \mid \mathcal{F}_m) = X_i(m)$$

as required.

For a configuration $\mathbf{x}$, write $\bar{\mathbf{x}}$ for the "equalized" configuration in which each agent has the average $n^{-1} \sum_i x_i$. Lemma 1, and convergence in distribution of the associated Markov chain to its (uniform) stationary distribution, immediately imply $\mathbb{E}\mathbf{X}(t) \to \overline{\mathbf{x}(0)}$ as $t \to \infty$.

Amongst several ways one might proceed to argue that $\mathbf{X}(t)$ itself converges to $\overline{\mathbf{x}(0)}$, the next leads to a natural explicit quantitative bound.

A function $f : I \to \mathbb{R}$ has (with respect to the uniform distribution) average $\overline{f}$, variance $\operatorname{var} f$ and L^2 norm $\|f\|_2$ defined by

$$\begin{aligned} \overline{f} &:= n^{-1} \sum_i f_i \\ \|f\|_2^2 &:= n^{-1} \sum_i f_i^2 \\ \operatorname{var} f &:= \|f\|_2^2 - (\overline{f})^2. \end{aligned}$$

The L^2 norm will be used in several different ways. For a possible time-t configuration $\mathbf{x}(t)$ of the averaging process, the quantity $\|\mathbf{x}(t)\|_2$ is a number, and so the quantity $||\mathbf{X}(t)||_2$ appearing in the proposition below is a random variable.

Proposition (Global convergence theorem)

From an initial configuration $\mathbf{x}(0) = (x_i)$ with average zero, the time-t configuration $\mathbf{X}(t)$ of the averaging process satisfies

$$\mathbb{E}||\mathbf{X}(t)||_2 \le ||\mathbf{x}(0)||_2 \exp(-\lambda t/4), \quad 0 \le t < \infty \tag{3}$$

where λ is the spectral gap of the associated MC.

Before starting the proof let us recall some background facts about reversible chains, here specialized to the case of uniform stationary distribution (that is, $\nu_{ij} = \nu_{ji}$) and in the continuous-time setting. See Chapter 3 of Aldous-Fill for the theory surrounding (4) and Lemma 2. The associated Markov chain, with generator $\mathcal{N}$ at (1), has *Dirichlet form*

$$\mathcal{E}(f, f) := \tfrac{1}{2} n^{-1} \sum_i \sum_{j \neq i} (f_i - f_j)^2 \nu_{ij} = n^{-1} \sum_{\{i,j\}} (f_i - f_j)^2 \nu_{ij}$$

where $\sum_{\{i,j\}}$ indicates summation over *unordered* pairs. The spectral gap of the chain, defined as the gap between eigenvalue 0 and the second eigenvalue of $\mathcal{N}$, is characterized as

$$\lambda = \inf_f \left\{ \frac{\mathcal{E}(f, f)}{\mathrm{var}(f)} : \ \mathrm{var}(f) \neq 0 \right\}. \tag{4}$$

Digression: An intuitive way to think about this. Recall the 3rd definition of variance:

$$\mathrm{var}\, Z = \tfrac{1}{2} \mathbb{E}(Z_1 - Z_2)^2, \quad Z_i \text{ i.i.d. } \stackrel{d}{=} Z.$$

For a stationary discrete-time chain (Z_t), the Dirichlet form is defined as

$$\mathcal{E}(f,f) = \tfrac{1}{2}\mathbb{E}(f(Z_1) - f(Z_0))^2$$

and we can think of variance as

$$\operatorname{var} f(Z_0) = \tfrac{1}{2}\mathbb{E}(f(Z_\infty) - f(Z_0))^2.$$

So the ratio is comparing "local" and "global" fluctuations of $f(Z_t)$.

Here are 3 of the many things the spectral gap λ tells you about a (general) reversible chain with stationary distribution π.

1. For the stationary chain,

$$\max_{f,g} \operatorname{cor}(f(Z_0), g(Z_t)) = \exp(-\lambda t).$$

2. $\mathbb{P}_i(Z_t = j) - \pi_j \sim c_{ij} \exp(-\lambda t)$ as $t \to \infty$.

3. Define a "distance from stationary" $d_2(\rho, \pi)$ for a probability dist. ρ to be the L^2 norm of the function $i \to \frac{\rho_i - \pi_i}{\pi_i}$. Then

Lemma (L^2 contraction lemma)

The time-t distributions $\rho(t)$ of the associated Markov chain satisfy

$$d_2(\rho(t), \pi) \leq e^{-\lambda t} d_2(\rho(0), \pi)$$

where λ is the spectral gap of the associated MC.

This is optimal, in the sense that the rate of convergence really is $\Theta(e^{-\lambda t})$ as $t \to \infty$.

We don't actually use this lemma, but our global convergence theorem for the averaging process is clearly analogous.

Notation for FMIE process dynamics. We will write

$$\mathbb{E}(dZ(t) \mid \mathcal{F}(t)) = [\leq] \;\; Y(t)dt$$

to mean

$$Z(t) - Z(0) - \int_0^t Y(s)ds \text{ is a martingale [supermartingale]},$$

– the former "differential" notation seems much more intuitive than the integral notation. In the context of a FMIE process we typically want to choose a functional Φ and study the process $\Phi(\mathbf{X}(t))$, and write

$$\mathbb{E}(d\Phi(\mathbf{X}(t)) \mid \mathbf{X}(t) = \mathbf{x}) = \phi(\mathbf{x})dt \tag{5}$$

so that $\mathbb{E}(d\Phi(\mathbf{X}(t)) \mid \mathcal{F}(t)) = \phi(\mathbf{X}(t))dt$. We can immediately write down the expression for ϕ in terms of Φ and the dynamics of the particular process; for the averaging process,

$$\phi(\mathbf{x}) = \sum_{\{i,j\}} \nu_{ij}(\Phi(\mathbf{x}^{ij}) - \Phi(\mathbf{x})) \tag{6}$$

where $\mathbf{x}^{ij}$ is the configuration obtained from $\mathbf{x}$ after agents i and j meet and average. This is just saying that agents i, j meet during $[t, t + dt]$ with chance $\nu_{ij}dt$ and such a meeting changes $\Phi(\mathbf{X}(t))$ by the amount $\Phi(\mathbf{x}^{ij}) - \Phi(\mathbf{x})$.

Proof of Proposition 1. A configuration $\mathbf{x}$ changes when some pair $\{x_i, x_j\}$ is replaced by the pair $\{\frac{x_i+x_j}{2}, \frac{x_i+x_j}{2}\}$, which preserves the average and reduces $||\mathbf{x}||_2^2$ by exactly $\frac{(x_j-x_i)^2}{2n}$. So, writing $Q(t) := ||\mathbf{X}(t)||_2^2$,

$$\begin{aligned} \mathbb{E}(dQ(t) \mid \mathbf{X}(t) = \mathbf{x}) &= -\sum_{\{i,j\}} \nu_{ij} \cdot n^{-1}(x_j - x_i)^2/2\ dt \\ &= -\mathcal{E}(\mathbf{x}, \mathbf{x})/2\ dt \\ &\leq -\lambda ||\mathbf{x}||_2^2/2\ dt. \end{aligned} \tag{7}$$

The first equality is by the dynamics of the averaging process, the middle equality is just the definition of $\mathcal{E}$ for the associated MC, and the final inequality is the extremal characterization

$$\lambda = \inf\{\mathcal{E}(g,g)/||g||_2^2\ :\ \overline{g} = 0,\ \mathrm{var}(g) \neq 0\}.$$

So we have shown

$$\mathbb{E}(dQ(t) \mid \mathcal{F}(t)) \leq -\lambda Q(t)\ dt/2.$$

The rest is routine. Take expectation:

$$\frac{d}{dt}\mathbb{E}Q(t) \leq -\lambda \mathbb{E}Q(t)/2$$

and then solve to get

$$\mathbb{E}Q(t) \leq \mathbb{E}Q(0)\exp(-\lambda t/2)$$

in other words

$$\mathbb{E}||\mathbf{X}(t)||_2^2 \leq ||\mathbf{x}(0)||_2^2 \exp(-\lambda t/2), \quad 0 \leq t < \infty.$$

Finally take the square root.

A local smoothness property

Thinking heuristically of the agents who agent i most frequently meets as the "local" agents for i, it is natural to guess that the configuration of the averaging process might become "locally smooth" faster than the "global smoothness" rate implied by Proposition 1. In this context we may regard the Dirichlet form

$$\mathcal{E}(f,f) := \tfrac{1}{2} n^{-1} \sum_i \sum_{j \neq i} (f_i - f_j)^2 \nu_{ij} = n^{-1} \sum_{\{i,j\}} (f_i - f_j)^2 \nu_{ij}$$

as measuring the "local smoothness", more accurately the local roughness, of a function f, relative to the local structure of the particular meeting process. The next result implicitly bounds $\mathbb{E}\mathcal{E}(\mathbf{X}(t), \mathbf{X}(t))$ at finite times by giving an explicit bound for the integral over $0 \leq t < \infty$. Note that, from the fact that the spectral gap is strictly positive, we can see directly that $\mathbb{E}\mathcal{E}(\mathbf{X}(t), \mathbf{X}(t)) \to 0$ exponentially fast as $t \to \infty$; Proposition 2 is a complementary non-asymptotic result.

Proposition

For the averaging process with arbitrary initial configuration $\mathbf{x}(0)$*,*

$$\mathbb{E} \int_0^\infty \mathcal{E}(\mathbf{X}(t), \mathbf{X}(t)) \ dt = 2 \operatorname{var} \mathbf{x}(0).$$

This looks slightly magical because the bound does not depend on the particular rate matrix $\mathcal{N}$, but of course the definition of $\mathcal{E}$ involves $\mathcal{N}$.

Proof. By linearity we may assume $\overline{\mathbf{x}(0)} = 0$. As in the proof of Proposition 1 consider $Q(t) := ||\mathbf{X}(t)||_2^2$. Using (7)

$$\frac{d}{dt} \mathbb{E} Q(t) = -\mathbb{E}\mathcal{E}(\mathbf{X}(t), \mathbf{X}(t))/2$$

and hence

$$\mathbb{E} \int_0^\infty \mathcal{E}(\mathbf{X}(t), \mathbf{X}(t)) \ dt = 2(Q(0) - Q(\infty)) = 2||\mathbf{x}(0)||_2^2 \qquad (8)$$

because $Q(\infty) = 0$ by Proposition 1.

General Principle 3: Duality

Notions of *duality* are one of the interesting and useful tools in classical IPS, and equally so in the social dynamics models we are studying. The duality between the voter model and coalescing chains (recalled later) is the simplest and most striking example. The relationship we develop here for the averaging model is less simple but perhaps more representative of the general style of duality relationships.

The technique we use is to extend the "random penny" (augmented process) argument used in Lemma 1. Now there are two pennies, and at any meeting there are independent decisions to hold or pass each penny. The positions $(Z_1(t), Z_2(t))$ of the two pennies behave as the following MC on product space, which is a particular coupling of two copies of the (half-speed) associated MC. Here i, j, k denote distinct agents.

$$
\begin{aligned}
(i,j) \to (i,k) \quad &: \quad \text{rate } \tfrac{1}{2}\nu_{jk} \\
(i,j) \to (k,j) \quad &: \quad \text{rate } \tfrac{1}{2}\nu_{ik} \\
(i,j) \to (i,i) \quad &: \quad \text{rate } \tfrac{1}{4}\nu_{ij} \\
(i,j) \to (j,j) \quad &: \quad \text{rate } \tfrac{1}{4}\nu_{ij} \\
(i,j) \to (j,i) \quad &: \quad \text{rate } \tfrac{1}{4}\nu_{ij} \\
(i,i) \to (i,j) \quad &: \quad \text{rate } \tfrac{1}{4}\nu_{ij} \\
(i,i) \to (j,i) \quad &: \quad \text{rate } \tfrac{1}{4}\nu_{ij} \\
(i,i) \to (j,j) \quad &: \quad \text{rate } \tfrac{1}{4}\nu_{ij}.
\end{aligned}
$$

For comparison, for two independent chains the transitions $(i,j) \to (j,i)$ and $(i,i) \to (j,j)$ are impossible (because of the continuous time setting) and in the other transitions above, all the $1/4$ terms become $1/2$. Intuitively, in the coupling the pennies move independently except for moves involving an edge between them, in which case the asynchronous dynamics are partly replaced by synchronous ones.

Repeating the argument around (2) – an exercise for the dedicated student – gives the following result. Write $\mathbf{X}^a(t) = (\mathbf{X}_i^a(t))$ for the averaging process started from configuration $\mathbf{1}_a$.

Lemma (The duality relation)

For each choice of a, b, i, j, not requiring distinctness, and for each t,

$$\mathbb{E}(X_i^a(t) X_j^b(t)) = \mathbb{P}(Z_1^{a,b}(t) = i, Z_2^{a,b}(t) = j)$$

where $(Z_1^{a,b}(t), Z_2^{a,b}(t))$ denotes the coupled process started from (a, b).

By linearity the duality relation implies the following – apply $\sum_a \sum_b x_a(0) x_b(0)$ to both sides.

Corollary (Cross-products in the averaging model)

For the averaging model $\mathbf{X}(t)$ started from a configuration $\mathbf{x}(0)$ which is a probability distribution over agents, and for each t,

$$\mathbb{E}(X_i(t) X_j(t)) = \mathbb{P}(Z_1(t) = i, Z_2(t) = j)$$

where $(Z_1(t), Z_2(t))$ denotes the coupled process started from random agents $(Z_1(0), Z_2(0))$ chosen independently from $\mathbf{x}(0)$.

Open Problem. One can define the averaging process on the integers – that is, $\nu_{i,i+1} = 1, -\infty < i < \infty$ – started from the configuration with unit total mass, all at the origin. By Lemma 1 we have

$$\mathbb{E}X_j(t) = p_j(t)$$

where the right side is the time-t distribution of a continuous-time simple symmetric random walk, which of course we understand very well.

But what can you say about the second-order behavior of this averaging process? That is, how does $\mathrm{var}(X_j(t))$ behave and what is the distributional limit of $(X_j(t) - p_j(t))/\sqrt{\mathrm{var}(X_j(t))}$? Note that duality gives an expression for the variance in terms of the coupled random walks, but the issue is to find an exact formula, or to somehow analyze asymptotics without an exact formula.

Quantifying convergence via entropy

Parallel to Lemma 2 are quantifications of reversible Markov chain convergence in terms of the log-Sobolev constant of the chain, defined (cf. (4)) as

$$\alpha = \inf_f \left\{ \frac{\mathcal{E}(f,f)}{L(f)} : L(f) \neq 0 \right\}. \tag{9}$$

where

$$L(f) = n^{-1} \sum_i f_i^2 \log(f_i^2/\|f\|_2^2).$$

See Montenegro and Tetali (2006) for an overview, and Diaconis and Saloff-Coste (1996) for more details of the theory, which we do not need here. One problem posed in the Spring 2011 course was to seek a parallel of Proposition 1 in which one quantifies closeness of $\mathbf{X}(t)$ to uniformity via entropy, anticipating a bound in terms of the log-Sobolev constant of the associated Markov chain in place of the spectral gap. Here is one solution to that problem.

For a configuration $\mathbf{x}$ which is a probability distribution write

$$\mathrm{Ent}(\mathbf{x}) := -\sum_i x_i \log x_i$$

for the entropy of the configuration. Consider the averaging process where the initial configuration is a probability distribution. By concavity of the function $-x \log x$ it is clear that in the averaging process $\mathrm{Ent}(\mathbf{X}(t))$ can only increase, and hence $\mathrm{Ent}(\mathbf{X}(t)) \uparrow \log n$ a.s. (recall $\log n$ is the entropy of the uniform distribution). So we want to bound $\mathbb{E}(\log n - \mathrm{Ent}(\mathbf{X}(t)))$. For this purpose note that, for a configuration $\mathbf{x}$ which is a probability distribution,

$$nL(\sqrt{\mathbf{x}}) = \log n - \mathrm{Ent}(\mathbf{x}). \tag{10}$$

Proposition

For the averaging process whose initial configuration is a probability distribution $\mathbf{x}(0)$*,*

$$\mathbb{E}(\log n - \mathrm{Ent}(\mathbf{X}(t))) \leq (\log n - \mathrm{Ent}(\mathbf{x}(0))) \ \exp(-\alpha t/2)$$

where α *is the log-Sobolev constant of the associated Markov chain.*

The format closely parallels that of Proposition 1, though the proof is a little more intricate. See the paper for proof.

Open Problem. A standard test bench for Markov chain related problems is the Hamming cube graph with vertex-set $\{0,1\}^d$ and rates $\nu_{ij} = 1/d$ for adjacent vertices. In particular its log-Sobolev constant is known. Can you get stronger results for the averaging process on this cube than are implied by our general results?

I have shown all that's explicitly known about the averaging process itself, though more elaborate variant models have been studied. Here is one variant.

Averaging process with noise.

This variant model can be described as

$$dX_i(t) = \sigma dW_i(t) + \text{(dynamics of averaging model)}$$

where the "noise" processes $W_i(t)$ are defined as follows. First take n independent standard Normals conditioned on their sum equalling zero – call them $(W_i(1), 1 \leq i \leq n)$. Now take $\mathbf{W}(t)$ to be the n-dimensional Brownian motion associated with the time-1 distribution $\mathbf{W}(1) = (W_i(1), 1 \leq i \leq n)$.

By copying the proof of Proposition 1, easy to show that the limit distribution $\mathbf{X}(\infty)$ of this process satisfies

$$\mathbb{E}||\mathbf{X}(\infty)||_2^2 \leq \frac{2\sigma^2(n-1)}{\lambda n}.$$

Here's an "opposite" process.

The Compulsive Gambler process.

Initially each agent has some (non-negative real-valued) amount of money. Whenever two agents meet, they instantly play a fair game: one agent acquires the combined money. In other words, if one has a and the other has b then the first acquires all $a + b$ with chance $a/(a + b)$.

Note that on the complete graph geometry, this process is just an augmentation of the Kingman coalescent process. On a general geometry, a configuration in which the set of agents with non-zero money forms an "independent set" (no two are adjacent in the weighted graph) is obviously an absorbing configuration, and conversely.

This process has apparently not been studied. I mention it because it provides a simple example of the "disordered" limit behavior mentioned in lecture 1. Here are two questions you might like to investigate.

Take as geometry some r-regular n-vertex graph. Give each agent 1 unit initially.

1. First consider the absorption time T. Clearly

$$T \leq \max_e (\text{time of first meeting across } e).$$

This leads quickly to the upper bound

$$\mathbb{E}T = O(r \log n).$$

For the n-cycle we have $\mathbb{E}T = \Omega(\log n)$. But what is the optimal bound for $r = r_n$?

2. In the absorbed configuration there is some mean proportion ρ of agents with non-zero money. In the $n \to \infty$ limit for fixed r, what are the maximum and minimum possible proportions $\rho^*(r), \rho_*(r)$?

3. Conditional on being non-zero, an agent's final fortune has expectation $= 1/\rho$, but what can you say about its distribution? For large r, an agent's fortune seems rather like a "double-or-quits" martingale.