Problem Set 1

Out: 23 Sept.; Due: 5 Oct.

Take time to write **clear** *and* **concise** *answers. None of the problem parts requires a long solution; if you find yourself writing a lot, you are either on the wrong track or giving an inappropriate level of detail. You are encouraged to form small groups (two or three people) to work through the problems, but you should write up your solutions on your own. Put your solutions in the box outside my office (677 Soda) by 5pm on Monday.*

1. Let G be a connected undirected graph. Show that random walk on G is periodic if and only if G is bipartite.

2. In this problem we'll prove the fundamental convergence theorem stated in Lecture 2: For a finite, irreducible, aperiodic Markov chain, $p_x^{(t)}$ converges to a unique stationary distribution π as $t \to \infty$. Note that we almost proved this in Lecture 3, where we showed by a coupling argument that $p_x^{(t)}$ converges at an exponential rate to π *assuming* that π exists. Thus essentially the only missing ingredient is to prove the existence of π. In what follows, assume that P has finite state space Ω and is irreducible and aperiodic.

(a) For states $x \neq y \in \Omega$, let $q_x(y)$ be the expected number of times the chain, started at x, visits y before its first return to x. Also, let $q_x(x) = 1$. Show that, for any x, the vector q_x is finite and strictly positive.
[HINT: Since P is aperiodic and irreducible, there exists a t such that $P^t(x, y) > 0$ for all x, y.]

(b) Show that, for any x, q_x satisfies $q_x P = q_x$.
[HINT: For a more believable proof, you might want to expand the expectation as $q_x(y) = \sum_{t\geq 1} r_x^{(t)}(y)$, where $r_x^{(t)}(y)$ is the probability that the chain is at y after t steps having not returned to x.]

(c) Parts (a) and (b) prove the existence of a stationary distribution π (which is just the normalized version of q_x). By Lecture 3, this implies that $\Delta(2t) \leq D(2t) \leq D(t)^2$. What additional simple fact is needed to deduce that $\Delta(t) \to 0$ (and hence that $p_x^{(t)} \to \pi$)? Prove this fact.
[HINT: What happens to $D(t)$ in a periodic Markov chain?]

3. The *East model* is a Markov chain with state space $\Omega = \{x = x_1 x_2 \ldots x_{n+1} \in \{0,1\}^{n+1} : x_{n+1} = 1\}$ with transitions defined as follows:

- Select a coordinate $i \in \{1, \ldots, n\}$ u.a.r.
- If $x_{i+1} = 1$ then flip the value at i (i.e., replace x_i by $1 - x_i$), else do nothing.

(a) Show that this Markov chain is irreducible and aperiodic, and that its stationary distribution is uniform.

(b) Show that the mixing time is *at least* $n^2 - cn^{3/2}$ for some constant c. [HINT: Consider the time it takes for the leftmost 1 to move one position to the left. You will need Chebyshev's inequality and the fact that a geometric random variable with parameter p has expectation $\frac{1}{p}$ and variance $\frac{1}{p^2} - \frac{1}{p}$.]

NOTE: It turns out that the mixing time is $O(n^2)$; this result uses methods we will see later in the course.

4. Let n, k be positive integers with $k \le \frac{n}{2}$, and let Ω denote the set of all subsets of $\{1, \dots, n\}$ of cardinality k. Thus $|\Omega| = \binom{n}{k}$. Consider the following Markov chain on Ω: from a subset S, pick an element $a \in S$ and an element $b \in \{1, \dots, n\} - S$, independently and u.a.r., and move to the set $S + b - a$.

(a) Show that this Markov chain (augmented if necessary by a self-loop probability) is ergodic with uniform stationary distribution.

(b) Devise a coupling argument that shows that the mixing time is asymptotically $O(n \log k)$.

5. Here is an unusual card trick. I take a shuffled deck and turn up the cards one by one. I ask you to select one of the first ten cards, without telling me; let $c_1 \in \{1, 2, \dots, 13\}$ be the numerical value of your card. You then count c_1 cards from the one you selected, and note that card; call its value c_2. You then count a further c_2 cards and note that card, and so on until the deck is exhausted. At that point, I am able to identify the last card you noted (at least most of the time).

Describe how I perform this amazing feat, and give a qualitative explanation for why it works. [HINT: think about coupling. You are not expected to perform any calculations to justify why the method works, though as a (tough) project topic you might think about it. As far as I am aware, there is no tight analysis of the success probability, though some partial results exist.]

6. This problem concerns the boundary case for coupling arguments, when the expected change in distance is zero rather than negative.

(a) By generalizing the end of the proof of Claim 5.3 in Lecture 5, prove the following. If d is an integer-valued distance on Ω, and if we can define a coupling $(X, Y) \mapsto (X', Y')$ such that, for some $\alpha > 0$,

$$\mathrm{E}[d(X', Y') \mid X, Y] \le (1 - \alpha) d(X, Y) \qquad \text{for all pairs } X, Y, \tag{1}$$

then the mixing time is $O(\alpha^{-1} \log D)$, where $D = \max_{x,y} d(x, y)$.

(b) Suppose now that we are only able to prove (1) with $\alpha = 0$ (i.e., the distance does not increase in expectation). (In our colorings example from Lecture 5, this corresponds to the case $q = 4\Delta$.) Show that the mixing time is bounded by $O(\beta^{-1} D^2)$, where $\beta = \min_{X,Y} \mathrm{E}[(d(X', Y') - d(X, Y))^2 \mid X, Y]$ bounds the *second moment* of the change in distance. [HINT: This follows by a standard martingale argument using the Optional Stopping Theorem. If you are unfamiliar with such arguments, do the proof for the special case where d changes by ± 1 in each step and use comparison with symmetric random walk on the integer interval $[0, D]$.] Hence deduce that the mixing time for the colorings Markov chain when $q = 4\Delta$ is $O(n^3)$. [NOTE: Of course, we already proved a stronger bound in Lecture 6.]

(c) We saw in our discussion of path coupling in Lecture 6 that, if we can prove (1) only for pairs (X, Y) that are *adjacent* in some premetric that extends to d, then (1) also holds for all pairs (X, Y). This implication is of course valid even when $\alpha = 0$ (corresponding to the case $q = 2\Delta$ for our colorings Markov chain). Now let $\beta = \min_{X,Y} \mathrm{E}[(d(X', Y') - d(X, Y))^2 \mid X, Y]$, where the min is taken only over *adjacent* pairs X, Y. Can we deduce, as in part (b) above, that the mixing time is $O(\beta^{-1} D^2)$? Explain your answer with a careful argument. [HINT: The following example may be instructive. Let G be a cycle with four vertices, and consider the colorings $X = (\mathrm{R}, \mathrm{G}, \mathrm{B}, \mathrm{Y})$ and $Y = (\mathrm{Y}, \mathrm{B}, \mathrm{R}, \mathrm{G})$ (where colors are listed in order around the cycle, starting with vertex 1). Note that $\Delta = 2$ and $q = 4 = 2\Delta$, and that the distance $d(X, Y) = 4$. What happens to the distance $d(X, Y)$ under *any* coupling?]

(d) **[Optional]** In the path coupling scenario of part (c), can you give a (clean) argument that proves a mixing time of $O(n^4)$ (or some other polynomial) for the colorings Markov chain in the boundary case $q = 2\Delta$?