Probabilistic Self-Stabilizing Algorithms, Markov Chains and Markov Decision Processes

Lecture Notes (MPRI: Course 2.18, Oct.-Nov. 2007)

Laurent Fribourg

Plan

1. Probability Review
2. Markov Chains
3. Time to Absorption: An Algebraic View
4. Markov Chains with Costs
5. Maximum Principle: Applications
6. Markov Decision Processes with Costs
7. Time to Absorption: A Martingale View
8. Time to Stabilization: A Coupling View
9. Stationary Distributions, Mixing and Coupling

Abstract

We see *probabilistic (self-)stabilizing algorithms* as *Markov chains* and the set of legitimate (or legal) states as a set of recurrent states of the Markov chain. The stabilization of the algorithm corresponds to the time taken by the Markov chain to enter into the recurrent space. Probabilistic classical techniques are used to infer upper bounds on the expected time to stabilization. In the case where the algorithm contains also non-deterministic features (choice of actions), the probabilistic algorithm is seen as a *Markov decision process*, and techniques of control theory are used to study the expected time to stabilization under the "worst " policy (or scheduler).

Bibliography

[KS] J.G. Kemeny and J.L. Snell *Finite Markov Chains*, Springer-Verlag, 1976
[BT] D.P. Bertsekas and J.N. Tsitsiklis *Introduction to Probability* Athena scientific, 2000.
[B] P. Brémaud *Markov Chains: Gibbs Fields, Monte Carlo Simulations, and Queues*, Springer, 2001
[GS] GR Grimmett and DR Stirzaker *Probability and Random Processes*, Oxford University Press, 1992
[MU] M Mitzenmacher and E Upfal *Probability and Computing* Cambridge University Press, 2005
[S] Alistair Sinclair *Convergence Rates for Monte Carlo Experiments* Numerical Methods for Polymeric Systems," SGWhittington ed, IMA Volumes in Mathematics & its Applications, 1997

1 Probability Review

1.1 Events (Brémaud,p.1)

Probability theory provides a mathematical framework for the study of random phenomena. It requires a precise description of the *outcome* of an observation when such a phenomenon is observed. The collection of all possibles outcome ω is called the *sample space* Ω.

Example 1. A Die. The experiment consists in tossing a die once. The possible outcomes are $\omega = 1, 2, ..., 6$ and the sample space is the set $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Example 2. Coin Tosses. The experiment is an infinite sequence of coin tosses. one can take for the sample space Ω the collection of all sequences $\omega = \{x_n\}_{n\geq 1}$, where $x_n = 1$ or 0 depending on whether the nth toss results in heads or tails.

Probability theory assigns to an event a number, the *probability* of the said event. For technical reasons, the collection $\mathcal{F}$ of events that are assigned a probability is not always identical to the collection of all subsets of Ω. The requirements on $\mathcal{F}$ are the following.

1. The impossible event $\emptyset$ and the certain event Ω are in $\mathcal{F}$
2. If A is in $\mathcal{F}$, so is $\overline{A}$
3. If $A_1, A_2, ...$ are in $\mathcal{F}$, then so is $\cup_{k=1}^{\infty} A_k$

One calls the collection of subsets $\mathcal{F}$ a *sigma field* on Ω, here the sigma-field of *events*. If the sample space Ω is finite, one usually considers any subset of Ω to be an event. The same is generally true for a countable sample space.

1.2 Random Variables (Brémaud,p.3)

Definition 1. *Random Variables (Brémaud,p.3).*
A random variable *is a function:* $\Omega \to \mathbb{R}$ *such that for all* $a \in \mathbb{R}$*, the event* $\{X \leq a\} = \{\omega; X(\omega) \leq a\}$ *can be assigned a probability, that is,* $\{X \leq a\} \in \mathcal{F}$.
A function $X : \Omega \to S$ *where* S *is a denumerable set is called a* discrete random variable *if for all* $i \in S$: $\{X = i\} \in \mathcal{F}$.

Sometimes a random variable is called a *random number*. This is an innocuous habit as long as one is aware that it is not a function X that is random, but the outcome ω, which in turn makes the number $X(\omega)$ random. If $X(\omega)$ is real for all ω, then X is called a *real random variable*.

Example 3. : A Die. Take for X the identity $X(\omega) = \omega$. In that sense X is a random number obtained by tossing a die.

Example 4. Coin Tosses. Here $\omega = \{x_n\}_{n\geq 1}$. Define the X_n to be the random number obtained at the nth toss: $X_n(\omega) = x_n$.

1.3 Probability (Brémaud,p.4)

The *probability* $P(A)$ of an event $A \in \mathcal{F}$ measures the likeness of its occurrence. as a function defined on $\mathcal{F}$, the probability P is required to satisfy a few properties, the *axioms of probability*.

Definition 2. *Axioms of Probability (Brémaud,p.4).*
A probability (measure) on $(\Omega, \mathcal{F})$ *is a mapping* $P : \mathcal{F} \to \mathbb{R}$ *such that*

1. $0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. $P(\sum_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$

The third property is called sigma additivity. *The triple* $(\Omega, \mathcal{F}, P)$ *is called a* probability space, *or* probability model.

The axioms of probability are motivated by the following heuristic interpretation of $P(A)$ as the empirical frequency of occurrence of event A. If n "independent" experiments are performed, among which n_A result in the realization of A, then the empirical frequency $F(A) = \frac{n_A}{n}$ should be close to $P(A)$ if is "sufficiently large". Clearly the function F satisfies the axioms.

The axiomatic presentation of probability theory is nevetheless logically independent of the frequency interpretation. As a matter of fact, its success is due to its apparent ignorance of the frequency interpretation, which blurs the picture because the empirical frequency F deends on too many things: on the number of experiments and on the experiments themselves. The axiomatic theory of probability connects to the frequency interpretation *a posteriori*: the latter appears as a *theorem*, the famous *strong law of large numbers (SLLN)* given in Section 1.10. To obtain it, all that is needed besides the axioms of probability and clever computations is a good definition of what is meant by independent experiments. This definition will be given in Section **??**.

Example 5. A Die. For $A \subset \Omega = \{1, 2, 3, 4, 5, 6\}$, the formula $P(A) = \frac{|A|}{6}$, where A is the *cardinal* of A, that is, the number of elements in A, defines a probability P. This choice suggest an unbiased die, where each outcome $1, 2, 3, 4, 5$ or 6 has the same probability.

Example 6. Coin Tosses. Choose a probability P such that for any event of the form $A = \{x_1 = a_1, ..., x_k = a_k\}$, where $a_1, ..., a_k$ are arbitrary in $\{0, 1\}$, $P(A) = \frac{1}{2^k}$. This choice for probability P implies an unbiased coin and independent tosses.

1.4 Expectation (GS,p.50)

Let $x_1, x_2, \cdots x_n$ be the numerical outcomes of N repetitions of some experiment. The average of these outcomes is $m = \frac{1}{N}\sum_i x_i$. In advance of performing these experiments we can represent their outcomes by a sequence $X_1, X_2, ..., X_N$ of random variables, and we shall suppose that these variables are discrete with a common probability density function f ($\equiv P(X = .)$). Then, roughly speaking, for each possible value x, about $NP(X = x)$ of the X_i will take that value x. So the average m is about $m \simeq \frac{1}{N}\sum_x xNP(X = x) = \sum_x xP(X = x)$ where the summation here is over all possible values of the X_i. This average is called the 'expectation' of the underlying distribution with probability density function $P(X = .)$.

Definition 3. *: Expectation (GS,p.50).*
The expectation *of X is defined to be $E[X] = \sum_x xP(X = x)$ whenever this sum is absolutely convergent.*

We require **absolute** convergence in order that $E[X]$ be unchanged by reordering the x_i. Note that, although the expression $\sum_x xP(X = x)$ may appear to be an uncountable sum, only countably many of its contributions are non zero. If the numbers $P(X = x)$ are regarded as masses at points x then $E[X]$ is just the position of the centre of gravity; we can speak of X as having an 'atom' or 'point mass' of size $P(X = x)$ at x. We sometimes omit the brackets and simply write EX.

Lemma 1. *(GS,p.51) If X has a probability density function $P(X = .)$ and $g : \mathcal{R} \to \mathbb{R}$, then $E[g(X)] = \sum_x g(x)P(X = x)$ whenever this sum is absolutely convergent.*

Example 7. Bernouilli variables (GS,p.52). Let X be a Bernouilli variable, taking the value 1 with probability p ($= 1 - q$). Then $E[X] = \sum_x xP(X = x) = 0.q + 1.p = p$.

Example 8. : (Counterexample,GS,p.54). Let X have probability density function $P(x) = \frac{A}{x^2}$ for $x = \pm 1, \pm 2, ...$ where A is chosen so that $\sum P(X = x) = 1$. The sum $\sum xP(X = x) = A\sum_{x \neq 0} \frac{1}{x}$ does not converge absolutely, because both the positive and the negative parts diverge.

The expectation has the linearity and monotonicity properties, and the triangle inequality is true:

- linearity: $E[aX + bY] = aE[X] + bE[Y]$, for all $a, b \in \mathbb{R}$.
  Proof. $E[aX + bY] = \sum_{x,y}(ax + by)P(X = x, Y = y) = a\sum_x x \sum_y P(X = x, Y = y) + b\sum_y y \sum_x P(X = x, Y = y) = a\sum_x xP(X = x) + b\sum_y yP(Y = y) = aEX + bEY$.
- linearitybis: $E[ag_1(X) + bg_2(X)] = aE[g_1(X)] + bE[g_2(X)]$,
  where $a, b \in \mathbb{R}$ and g_1 and g_2 satisfy the absolute convergence conditions.
- monotonicity: $(\forall x\ g_1(x) \leq g_2(x)) \Rightarrow E[g_1(X)] \leq E[g_2(X)]$.
- triangle inequality: $|E[g(X)]| \leq E[|g(X)|]$.

Definition 4. *Variance.*
The variance *$Var(X)$ of a random variable X is $E[(X - E[X])^2]$.*
The square root of the variance, written σ, is called standard deviation.

From the linearity of expectation, it follows:
$\mathrm{Var}(X) = E[(X - m)]^2 = E[X^2] - 2mE[X] + m^2 = E[X^2] - m^2$, where m denotes $E[X]$.

Lemma 2. *(Wald) (p.19).*
Let $\{X_n\}_{n\geq 1}$ be a sequence of integrable random variables such that $E[X_n] = E[X_1]$ for all $n \geq 1$. Let T be an integer-valued random variable such that for all $n \geq 1$, the event $\{T \geq n\}$ is independent of X_n. Then $E[\sum_{n=1}^{T} X_n] = E[X_1]E[T]$.

1.5 Conditional Probabilities (Brémaud,p.9)

In the frequency interpretation, the definition of independence ($P(A \cap B) = P(A)P(B)$) reads, in rough and imprecise terms $n_{A\cap B} \approx (n_A/n).(n_b/n)$, or $\frac{n_{A\cap B}}{n_B} \approx n_A n$ (here $\approx$ is a "fuzzy" version of the equality sign). Therefore statistics relative to A do not vary when performed on a neutral sample of population characterized by the property B. For example, the proportion of people with a family name beginning with H is the same among a large population with th esual mix of men and women as it would be among a large all-male population. This is very much the intuitive notion of independence. Dependence betewwen A and B occurs when $P(A \cap B) \neq P(A)P(B)$. In this case the relative frequency $n_{A\cap B}/n_B \approx P(A \cap B)/P(B)$ is different from the frequency n_A/n.

Definition 5. *: Conditional Probability (p.9).*
The conditional probability *of A given B is the number $P(A|B) = \frac{P(A\cap B)}{P(B)}$, defined when $P(B) > 0$.*

The quantity $P(A|B)$ represents our expectation of A being realized when the only available information is that B is realized. Indeed, this expectation would then be based upon the relative frequency $n_{A\cap B}/n_B$ alone. A symmetric form is:
$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$.

Lemma 3. *: (GS,p.10).*
For any events A and B, $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$.

Theorem 1. *Bayes's Rules (Brémaud,p.10).*

- Bayes's rule of retrodiction*: with $P(A) > 0$, we have $P(B|A) = \frac{P(A|B)P(B)}{P(A)}$.*
- Bayes's rules of exclusive and exhaustive causes*: For $B_1, B_2, ...$ such that $\sum_{i=1}^{\infty} B_i = \Omega$ and for all A, we have $P(A) = \sum_{i=1}^{\infty} P(A|B_i)P(B_i)$.*
- Bayes's sequential formula*: For any sequence of events $A_1, ..., A_n$, we have $P(\cap_{i=1}^{k} A_i) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_k| \cap_{i=1}^{k-1} A_i)$.*

(BT, 2002,p.28) We start with the following theorem, which is often useful for computing the probabilities of various events, using a "divide-and-conquer" approach.

Theorem 2. *(Total Probability Theorem) Let $A_1, ..., A_n$ be disjoint events that form a partition of the sample space (each possible outcome is included in exactly one of the events $A_1, ..., A_n$) and assume that $P(A_i) > 0$, for all i. Then, for any event B, we have*

$$P(B) = P(A_1 \cap B) + \cdots + P(A_n \cap B) = P(A_1)P(B|A_1) + \cdots + P(B|A_n).$$

Intuitively, we are partitioning the sample space into a number of scenarios (events) A_i. Then, the probability that B occurs is a weighted average of its conditional probability under each scenario, where each scenario is weighted according to its (unconditional) probability. One of the uses of the theorem is to compute the probability of various events B for which the conditional probabilities $P(B|A_i)$ are known easy to derive. The key is to choose appropriately the partition $A_1, ..., A_n$, and this choice is often suggested by the problem structure.

Theorem 3. *We have*

$$P(A|B) = P(C|B)P(A|B \cap C) + P(C^c|B)P(B \cap C^c),$$

assuming all the conditioning events have positive probability.

Definition 6. *The* conditional (probability) mass function *of X given particular event A is defined by*

$$p_{X|A}(x) = \mathbb{P}(X = x \mid A) = \frac{\mathbb{P}(\{X = x\} \cap A)}{\mathbb{P}(A)}.$$

Definition 7. *The* conditional (probability) mass function *of X given Y is defined by*

$$p_{X|Y}(x|y) = \mathbb{P}(X = x \mid Y = y) = \frac{\mathbb{P}(\{X = x\} \cap \{Y = y\})}{\mathbb{P}(Y = y)} = \frac{p_{X,Y}(x, y)}{p_Y(y)}.$$

Definition 8. *The* conditional (probability) mass function *of X given $Y = y$ is defined by*

$$p_{X|Y=y}(x) = \mathbb{P}(X = x \mid Y = y).$$

1.6 Conditional Expectation

(MU,p.26). Just as we have defined conditional probability, it is useful to define the *conditional expectation* of a random variable. The following definition is quite natural.

Definition 9.

$$E[Y \mid Z = z] = \sum_y \mathbb{P}(Y = y \mid Z = z),$$

where the summation is over all y in the range of Y.

The definition states that the conditional expectation of a random variable is, like the expectation, a weighted sum of the values it assumes. The difference is that now each value is weighted by the *conditional probability* that the variable assumes that value.

Proposition 1. *For any random variables X and Y,*

$$E[X] = \sum_y \mathbb{P}(Y = y)E[X \mid Y = y],$$

where the sum is over all values in the range of Y and all of the expectations exist.

Proof. $\sum_y \mathbb{P}(Y = y)E[X \mid Y = y] = \sum_y \mathbb{P}(Y = y) \sum_x x\mathbb{P}(X = x \mid Y = y)$
$= \sum_x \sum_y x\mathbb{P}(X = x \mid Y = y)\mathbb{P}(Y = y)$
$= \sum_x \sum_y x\mathbb{P}(X = x \cap Y = y)$
$= \sum_x x\mathbb{P}(X = x)$
$= E[X]$.

Perhaps somewhat confusingly, the conditional expectation is also used to refer to the following random variable.

Definition 10. *The expression $E[Y|Z]$ is a random variable $f(Z)$ that takes on the value $E[Y \mid Z = z]$ when $Z = z$.*

We emphasize that $E[Y|Z]$ is *not* a real value; it is actually a function of the random variable Z. Hence $E[Y|Z]$ is itself a function from the sample space to the real numbers and can therefore be thought of as a random variable.

Theorem 4.

$$E[Y] = E[E[Y \mid Z]].$$

Proof. $E[Y|Z]$ is a function $f(Z)$, where $f(Z)$ takes on the value $E[Y|Z = z]$ when $Z = z$. Hence

$$E[E[Y|Z]] = \sum_z E[Y|Z = z]\mathbb{P}(Z = z).$$

The right-hand side equals $E[Y]$ by Prop. 1.

This theorem provides us a useful method for calculating $E[Y]$, since it asserts that

$$E[Y] = \sum_x E[Y|X = x]\mathbb{P}(X = x).$$

Proposition 2. *(Successive conditioning)*

$$E[E(X \mid Y_1, Y_2) \mid Y_1] = E[X \mid Y_1].$$

Theorem 5. *(Total Expectation Theorem) Let $A_1, ..., A_n$ be disjoint events that form a partition of the sample space, and assume that $P(A_i) > 0$, for all i. Then,*

$$E[X] = \sum_{i=1}^{n} P(A_i)E[X|A_i].$$

Theorem 6. *Let $A_1, ..., A_n$ be disjoint events that form a partition of the an event B, and assume that $P(A_i \cap B) > 0$, for all i. Then,*

$$E[X|B] = \sum_{i=1}^{n} P(A_i|B)E[X|A_i \cap B].$$

The total expectation theorem basically says that "the unconditional average can be obtained by averaging the conditional averages".

1.7 Indicators (p.18)

The *indicator function* of an event A:

$$1_A(\omega) = \begin{cases} 1 \text{ if } \omega \in A \\ 0 \text{ if } \omega \notin A \end{cases}$$

The random variable $X = 1_A$ takes the value 1 with probability $P(X = 1) = P(A)$ and the value $p(X = 0) = P(\overline{A}) = 1 - P(A)$. Therefore, $E[X] = 0 \times P(X = 0) + 1 \times P(X = 1) = P(X = 1) = P(A)$, that is to say,

$$E[1_A] = P(A).$$

In particular, $E[1] = 1$.[1]

1.8 Telescope Formula

Theorem 7. *Telescope Formula (MU, p.31).*
For a random variable X taking its values in $\mathbb{N}$, $E[X] = \sum_{i=1}^{\infty} P(X \geq i)$.

Proof. $\sum_{i=1}^{\infty} \mathbb{P}(X \geq i) = \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} \mathbb{P}(X = j) = \sum_{j=1}^{\infty} \sum_{i=1}^{j} \mathbb{P}(X = j) = \sum_{j=1}^{\infty} j\mathbb{P}(X = j)$
$= E[X]$.
The interchange of (possibly) infinite summations is justified, since the terms being summed are all nonnegative.

[1] LF: Conditional expectation $E[1_A|Y] = P(A|Y)$, cf Brémaud, p.39.

1.9 Jensen's Inequality

Proposition 3. *(MU, p.24) If a function $f : \mathbb{R} \to \mathbb{R}$ is convex (i.e., $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$, for any x_1, x_2 and $0 \leq \lambda \leq 1$), then*

$$E[f(X)] \geq f(E[X]).$$

Proof. (Case f has a Taylor expansion). Let $\mu = E[X]$. By Taylor's theorem, there is a value c such that

$$f(x) = f(\mu) + f'(\mu)(x-\mu) + f''(c)\frac{(x-\mu)^2}{2} \geq f(\mu) + f'(\mu)(x-\mu),$$

since $f''(c) > 0$ by convexity. Taking expectations of both sides and applying linearity of expactations yields the result:

$$E[f(X)] \geq E[f(\mu) + f'(\mu)(X-\mu)] = E[f(\mu)] + f'(\mu)(E[X]-\mu) = f(\mu) = f(E[X]).^2$$

1.10 Almost-Sure Convergence (Brémaud,p.43)

A sequence $\{Z_n\}_{n\geq 1}$ of real random variables is said to converge P-almost surely (P-a.s.) to the real random variable Z if $\mathbb{P}(\lim_{n\uparrow\infty} Z_n = Z) = 1$. (Paraphrasing: For all ω outside a set N of null probability, $\lim_{n\uparrow\infty} Z_n(\omega) = Z(\omega)$).

Theorem 8. *: Criterion of Almost-Sure Convergence. (Brémaud,p.43).*
Let $\{Z_n\}_{n\geq 1}$ and Z be random variables. If $\sum_{n\geq 1} P(|Z_n - Z| \geq \varepsilon_n) < \infty$ for some sequence of positive numbers $\{\varepsilon_n\}_{n\geq 1}$ converging to 0, then the sequence $\{Z_n\}_{n\geq 1}$ converges P-a.s. to Z.

Theorem 9. Dominated Convergence. *(GS,p.160)*
If $\{X_n\}$ is a sequence of variables with $X_n \to X$ a.s. (i.e., $X_n(\omega) \to X(\omega)$ for all $\omega \in \Omega$ except possibly on some null event) and $\forall n, \omega\ |X_n(\omega)| \leq Y(\omega)$ a.s. and $E(Y) < \infty$, then, $E(X_n) \to E(X)$.

1.11 Markov's Inequality (p.45)

Theorem 10. Markov's inequality. *Let X be a random variable with values in $\mathbb{R}$, $f : \mathcal{R} \to \mathbb{R}_+$, and $a > 0$. We then have*

$$P(f(X) \geq a) \leq \frac{E[f(X)]}{a}.$$

Proof. From the inequality $f(X) \geq a 1_{\{f(X)\geq a\}}$, it follows by taking expectations that

$$E[f(x)] \geq aE[1_{\{f(X)\geq a\}}] = aP(f(X) \geq a).$$

---

[2] (Case X takes only finitely many values). By induction, one can prove that if f is convex then, for any $x_1, x_2, \ldots, x_n$ and $\lambda_1, \lambda_2, \ldots, \lambda_n$ with $\sum_{i=1}^n \lambda_i = 1$, $(\sum_{i=1}^n \lambda_i x_i) \leq \sum_{i=1}^n \lambda_i f(x_i)$. Then, it follows: $E[f(X)] \geq f(E[X])$.

2 Markov Chains (KS, p.25)

2.1 Stochastic Processes (Kemeny-Snell, p.14)

We wish to give a probability measure to describe an experiment which takes place in stages. The outcome at the nth stage is allowed to depend on the outcomes of the previous stages. It is assumed, however, that the possibility for each possible outcome at a particular stage is known when the outcomes of all the previous stages are known. From this knowledge we shall construct a possibility space and measure for the over-all experiment.

We require that there be a finite number of possible outcomes at each stage and that we know the probabilities for any particular outcome at the j-th stage, given the knowledge of the outcomes for the first $j-1$ stages. For each j we obtain a tree T_j. The set of paths of this tree serves as a possibility space for any statement relating to the first j experiments. On this tree we assign a measure to the set of all paths. We first assign branch probabilities. Then the weight assigned to a path is the product of all branch probabilities on the path. The tree measures are consistent in the following sense. A statement whose truth value depends only on the first j stages may be considered a statement relative to any tree T_i for $i \geq j$. Each of these trees has its own tree measure and the probability of the statement could be found from any one of these measures. However, in every case the same probability would be assigned.

Assume that we have a tree for an n stage experiment. Let X_j be a function with domain the set of paths T_n and value the outcome at the j-th stage. Then functions $X_1, X_2, \ldots, X_n$ are called the *outcome functions*. The set of functions $X_1, X_2, \ldots, X_n$ is called a *stochastic process*. (In Markov chain theory it is convenient to denote the first outcome by X_0 instead of X_1.)

There is a simple connection between the branch probabilities and the outcome functions. The branch probabilities at the first stages are:

$\mathbb{P}[X_1 = i]$

at the second stage

$\mathbb{P}[X_2 = j \mid X_1 = i]$

at the third stage

$\mathbb{P}[X_3 = k \mid X_2 = j, X_1 = i]$

etc.

A stochastic process for which the outcome functions all have ranges which are subsets of a given finite set is called a *finite stochastic process*.

Therefore, for a finite stochastic process we have a tree and a tree measure and a sequence of outcome functions $X_n, n = 0, 1, 2, \ldots$. The domain of X_n is the tree T_n and the range is the set S_n of possible outcomes for the n-th experiment. The value of X_n is j if the outcome of the n-th experiment is j.

A finite stochastic process is an *independent process* if

(I) For any statement p whose truth value depends only on the outcomes before the n-th:

$$\mathbb{P}[X_n = j \mid p] = \mathbb{P}[X_n = j].$$

For such a process the kowledge of the outcome of any preceding experiment does not affect the knowledge of the outcome of any preceding experiment does not affect our predictions for the next experiment. For a Markov process we weaken this to allow the knowledge of the immediate past to influence these predictions.

Definition 11. *A* finite Markov process *is a finite stochastic process such that*

(II) For any statement p whse truth value depends only on the outcomes before the n-th,

$$\mathbb{P}[X_n = j \mid X_{n-1} = i, p] = \mathbb{P}[X_n = j \mid X_{n-1} = i].$$

We shall refer to condition II as the *Markov property*. For a Markov process, knowing the outcome of the last experiment we can neglect any other information we have about the past in predicting the future. Itis important to realize that this is the case only if we know exactly the outcome of the last experiment. For example, if we know only that the outcome of the last experiment was either i or k then knowledge of the truth value of a statement p relating to earlier experiments may affect our future predictions.

Definition 12. *The* n-th step transition probabilities *for a Markov process, denoted by* $p_{ij}(n)$ *are*

$$p_{ij}(n) = \mathbb{P}[X_n = j \mid X_{n-1} = i].$$

Definition 13. *A* finite Markov chain *is a finite Markov process such that the transition probabilities* $p_{ij}(n)$ *do not depend on* n. *In this case they are denoted by* p_{ij}. *The elements of possible outcomes of the experiments (elements of* S*) are called states.*

Definition 14. *The* transition matrix *for a Markov chain is the matrix* P *with entries* p_{ij}. *The* initial probability vector *is the vector* $\pi_0 = \{p_j^{(0)}\} = \{\mathbb{P}[X_0 = j]\}$.

For a Markov chain we may visualize a process which moves from state to state. It sarts in j with probability $p_j^{(0)}$. If at any time it is in state j, then it moves on the next "step" to j with probability p_{ij}. The initial probabilities are thought of as giving the probabilities for the various possible starting states. The initial probability vector and the transition matrix completely determine the Markov chain process, since they are sufficient to build the entire tree measure.Thus, given any probability vector π_0 and any probability matrix P, there is a unique Markov chain (except possibly for renaming the states) which will have the π_0 as initial probability vector and P as transition matrix.

2.2 Recapitulation and Examples

Definition 15. *(*Markov Models*)*

- *A Markov chain model is specified by identifying:*
  (a) the set of states $S = \{1, ..., m\}$,
  (b) the set of possible transitions, namely, those pairs (i, j) *for which* $p_{ij} > 0$, *and,*
  (c) the numerical values of those p_{ij} *that are positive.*
- *The Markov chain specified by this model is a sequence of random variables* $X_0, X_1, X_2, ...,$ *that take values in* S, *and which satisfy*
  $\mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0) = \mathbb{P}(X_{n+1} = j | X_n = i) = p_{ij}$,
  for all times n, *all states* $i, j \in S$, *and all possible sequences* $i_0, ..., i_{n-1}$ *of earlier states.*

All the elements of a Markov chain model can be encoded in a *transition probability matrix*, which is simply a two-dimensional array whose element at the i-th row and j-th column is p_{ij}:

$$\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}.$$

It is also helpful to lay out the model in the so-called *transition probability graph*, whose nodes are the states and whose arcs are the possible transitions. By recording the numerical values of p_{ij} near the corresponding arcs, one can visualize the entire model in a way that can make some of its major properties readily apparent.

Example 9. Un exemple est donné en figure 1, correspondant à la matrice de transition suivante :

$$T = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}$$

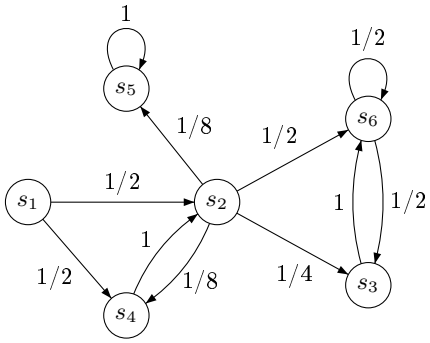

Fig. 1. Un exemple de graphe associé à une chaîne de Markov

Example 10. We consider Herman's mutual exclusion algorithm [?]. The topology is a cyclic graph (ring) of N vertices, and the scheduler synchronous. The set of states is $Q = \{0, 1\}$, and the number of machines N is odd. At each step, the state of every machine $x(i)$ $(1 \leq i \leq N)$ is changed into $x'(i)$ as follows:

- if $x(i) \neq x(i-1)$ then $x'(i) = \neg x(i)$,
- if $x(i) = x(i-1)$ then $x'(i) = \begin{cases} 0 & \text{with probability } 1/2, \\ 1 & \text{with probability } 1/2. \end{cases}$

(When $i = 1$, $(i-1)$ stands for N. As usual, $\neg 0$ stands for 1, and $\neg 1$ for 0.)

For two tokens, the Herman matrix is of the form:

$$\begin{pmatrix} 1 & 0 & 0 & . & . & . & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & & & & \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & & & \\ & . & . & . & . & & \\ & & & . & . & . & \\ & & & & . & . & \\ & & & & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ & & & & & \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

Example 11. We consider the problem of the Iterated Prisoner's Dilemma, as modeled in [?]. The topology is a cyclic graph (ring) of N vertices, and the scheduler randomized central. The set of states is $Q = \{-, +\}$. At each step, a vertex i $(1 \leq i \leq N)$ is chosen uniformly at random, and the values $x(i)$ and $x(i+1)$ are changed into $x'(i)$ and $x'(i+1)$ respectively as follows:

- if $x(i) = x(i+1)$, then $x'(i) = x'(i+1) = +$,
- if $x(i) \neq x(i+1)$, then $x'(i) = x'(i+1) = -$.

(When $i = N$, $(i+1)$ stands here for 1.)

2.3 The Probability of a Path (BT 2002,p.318)

Given a Markov chain model, we can compute the probability of any particular sequence of future states. This is analoguous to the use of the multiplication rule on sequential (tree) probability models. In particular, we have

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, ..., X_n = i_n) = \mathbb{P}(X_0 = i_0)p_{i_0 i_1}p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

If the initial state X_0 is given and is known to be equal to some i_0, a similar argument yields

$$\mathbb{P}(X_1 = i_1, ..., X_n = i_n | X_0 = i_0) = p_{i_0 i_1}p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

Graphically, a state sequence can be identified with a sequence of arcs in the transition probability graph, and the probability of such a path (given the initial state) is given by the product of the probabilities associated with the arcs traversed by the path.

2.4 n-Step Transition Probabilities (BT 2002,p.319)

Many Markov chain problems require the calculation of the probability law of the state at some future time, conditioned at the current state. This probability law is captured by the *n-step transition probabilities*, defined by

$$p_{ij}(n) = \mathbb{P}(X_n = j | X_0 = i).$$

In words, $p_{ij}(n)$ is the probability that the state after n time periods will be j, given that the current state is i.

(GS,p.196)Broadly speaking, we are interested in the evolution of X over two different time scales, the 'short term' and the 'long term' In the short term the random evolution of X is described by P, whilst long-term changes are described in the following way.

Definition 16. *(GS,p196) The* n-step transition matrix $P_n = (p_{ij}(n))$ *is the matrix of* n-step transition probabilities $p_{ij}(n) = \mathbb{P}(X_{m+n} = j | X_m = i)$,

Of course, $P_1 = P$

Theorem 11. *: Chapman-Kolmogorov Equations (GS,p196)*
$p_{ij}(m+n) = \sum_k p_{ik}(m)p_{kj}(n)$
Hence $P_{m+n} = P_m P_n$, *and so* $P_n = P^n$, *the nth power of* P.

This theorem relates long-term development to short-term development, and tells us how X_n depends on the initial variable X_0 Let $\mu_i^{(n)} = \mathbb{P}(X_n = i)$ be the density function of X_n, and write $\mu^{(n)}$ for the row vector with enties $(\mu_i^{(n)} : i \in S)$.

Lemma 4. *(GS,p196)* $\mu^{(m+n)} = \mu^{(m)} P_n$, *and hence* $\mu^{(n)} = \mu^{(0)} P^n$.

Thus we reach the important conclusion:

Theorem 12. *Distribution of an HMC*
The distribution of a discrete-time HMC is determined by its initial distribution $\mu^{(0)}$ *and its transition matrix* P.

Many questions about the chain can be expressed in terms of these quantities (P and $\mu^{(0)}$), and the study of the chain is thus largely reducible to the study of algebraic properties of matrices.

2.5 Classification of states (Kemeny-Snell, p.35)

(Brémaud, p.71) All the properties defined in the present section are *topological* in the sense that they concern only the *naked* transition graph (without labels).

Definition 17. *A set* C *of states is called:*

- closed *if for all* $i \in C$: $p_{ij} > 0 \Rightarrow j \in C$.
- irreducible *(or* strongly connected*) if* $i \leftrightarrow j$ *for all* $i, j \in C$.

Once the chain takes a value in a closed set C of states then it never leaves C subsequently. A closed set containing exactly one state is called *absorbing*. We have:

Proposition 4. *A state i is absorbing if and only if $p_{ii} = 1$.*

Definition 18. *(GS,p.204) We say i* communicate with *j, written $i \to j$, if the chain may ever visit state j with positive probability, starting from i (that is, $i \to j$ if $p_{ij}(n) > 0$ for some $n \geq 0$). We say i and j are* interconnected *if $i \to j$ and $j \to i$, in which case we write $i \leftrightarrow j$.*

It is easy to see that $\leftrightarrow$ is an equivalence relation (Two states are in the same equivalence class if they "communicate", i.e. if one can go from either state to the other one.) The state space S can be partitioned into equivalence classes of $\leftrightarrow$. The resulting partial ordering shows us the possible directions in which the process can proceed. The minimal elements of the partial ordering are of particular interest.

Definition 19. *The minimal elements of the partial ordering of equivalence classes are called* recurrent sets. *The remaining element are called* transient sets. *The elements of a transient set are called* transient states. *The elements of a recurrent set are said to be* recurrent *(or* non-transient*).*

Since every finite partial ordering must have at least one minimal element, there must be at least one recurrent set for every Markov chain. However, there need be no transient set. The latter will occur if the entire chain consists of a single recurrent set, or if there are several recurrent sets, which do not communicate with eithers.

If a process leaves a transient set it can never return to this set, while if it onces enters a recurrent set, it can never leave it.

Theorem 13. *: Decomposition of the State Space (GS,p.205).*
The state space S can be partitioned uniquely as $E = T \cup R_1 \cup R_2 \cup \cdots$ where T is the state of transient states, and the R_i are the recurrent sets.

Proposition 5. *Each recurrent set is closed and irreducible.*

The decomposition theorem gives us a canonical form for the transition matrix. We renumber the states as follows: The elements of a given equivalence class will receive consecutive numbers. The minimal sets will come first, the sets that are one level above the minimal sets, then sets that are one level above the minimal sets, then sets with two levels above the minimal sets, etc. This will assure us that we can go from a given state to another in the same class, or to a state in an earlier class, but not to a state in a later class. The communication structure of a transition matrix is therefore of the form $P = (P_1, P_2, P_3, ...)$.

$$\begin{array}{c} \\ P = \\ \\ \end{array} \begin{array}{c} \begin{array}{cccc} R_1 & R_2 & R_3 & T \end{array} \\ \begin{pmatrix} P_1 & 0 & 0 & 0 \\ 0 & P_2 & 0 & 0 \\ 0 & 0 & P_3 & 0 \\ x & x & x & x \end{pmatrix} \end{array} \begin{array}{l} R_1 \\ R_2 \\ R_3 \\ T \end{array} . \qquad (8.5)$$

Example 12. In the example of Fig. 1, there are two recurrent sets: $\{s_5\}$ and $\{s_3, s_6\}$. After reordering the set as $\{s_5, s_3, s_6, s_1, s_2, s_4\}$, with the recurrent sets at the beginning, the matrix of transition has the form:

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{2} & 0 & 0 & \frac{1}{8} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Every finite Markov chain has always at least one recurrent set (since finite partial ordering $\ll$ must have at least one minimal element). Moreover, two distinct recurrent sets are disjoint (since they are both strongly connected).

Example 13. Consider Herman's algorithm in the case where N is odd. In a configuration, a "token" at position i ($1 \leq i \leq N$) corresponds to the presence of two contiguous states of the same value (00 or 11) at position $i-1$ and i. Since N is odd, any configuration contains always at least one token. It is easy to see that such a set is recurrent:
The set $\mathcal{L}$ is the set of the configurations with a single token. There are $2N$ such configurations: they are of the form $x_i = 01\cdots010\mathbf{0}10\cdots101$ where token 00 is at position i, or $x'_i = 10\cdots101\mathbf{1}01\cdots010$ where token 11 is at position i, for all $1 \leq i \leq N$. (Letters in bold indicate that they are subjet to randomized transitions.) Let us show that $\mathcal{L}$ is recurrent, i.e. closed and strongly connected.

Applying a transition to an arbitrary element of $\mathcal{L}$, say x_i, leads to the 'dual' element $x'_i = 10\cdots101\mathbf{1}01\cdots010$ with probability 1/2, where token 11 is at the same position, or to $x_{i+1} = 10\cdots1010\mathbf{0}1\cdots010$ with probability 1/2, where token 00 is at position one more right. This shows that $\mathcal{L}$ is closed (since x'_i and x_{i+1} belong to $\mathcal{L}$). Moreover, this shows that one can go from x_i to x'_i and x_{i+1} in one step; two elements of $\mathcal{L}$ are thus connected together within at most N steps. Hence $\mathcal{L}$ is strongly connected.

2.6 Combining States (Kemeny-Snell, p. 123)

Assume that we are given an r-state Markov chain with transition matrix P and initial vector π. Let $A = \{A_1, A_2, \ldots, A_t\}$ be a partition of the set of states. We form a new process as follows. The outcome of the j-th experiment in the new process is the set A_k that contains the outcome of the j-th step in the original chain. We define branch probabilities as follows: At the zero level we assign

$$\mathbb{P}_\pi[X_0 \in A_i]. \tag{1}$$

At the first level we assign

$$\mathbb{P}_\pi[X_1 \in A_j \mid X_0 \in A_i].$$

In general, at the n-th level we assign branch probabilities,

$$\mathbb{P}_\pi[X_n \in A_t \mid X_{n-1} \in A_s, \ldots, X_1 \in A_j, X_0 \in A_i]. \tag{2}$$

The above procedure could be used to reduce a process with a very large number of states to a process with a smaller number of states. We call this process a *lumped process.* It is also often the case in applications that we are only interested in questions which relate to this coarser analysis of the possibilities. Thus it is important to able to determine whether the new process ca be treated by Markov chain methods.

Definition 20. *We shall say that a Markov chain is* lumpable *with respect to a partition $A = \{A_1, A_2, ..., A_r\}$ if for every starting vector π the lumped process defined by* (1) *and* (2) *is a Markov chain and the transition probabilities do not depend on the choice of π.*

Let $p_{iA_j} = \sum_{k \in A_j} p_{ik}$. Then p_{iA_j} represents the probability of moving from state i into set A_j in one step of the original Markov chain.

Theorem 14. *A necessary condition for a Markov chain to be lumpable with respect to a partition $A = \{A_1, A_2, \ldots, A_s\}$ is that for every pair of sets A_i and A_j, p_{kA_j} have the same value for every $k \in A_i$. These common values $\{\hat{p}_{ij}\}$ form the transition matrix for the lumped chain.*

Proof. For the chain to be lumpable it is clearly necesaary that

$$\mathbb{P}_\pi[X_1 \in A_j \mid X_0 \in A_i]$$

be the same for every $\hat{p}_{ij}$. In particular this must be the same for π having a 1 in its k-th component, for state $k \in A_i$. Hence $p_{kA_j} = \mathbb{P}_k[X_1 \in A_j] = \hat{p}_{ij}$ for every $k \in A_i$. Hence $p_{kA_j} = \mathbb{P}_k[X_1 \in A_j] = \hat{p}_{ij}$ for every $k \in A_i$. Thus the condition given is necessary. To prove it is sufficient, we must show that if the condition is satisfied the probability (2) depends only on A_s and A_t. The probability (2) may be written in the form

$$\mathbb{P}_{\pi'}[X_1 \in A_t]$$

where π' is a vector with non-zero components only on the states of A_s. It depends on π and on the first n outcomes. However, if $\mathbb{P}_k[X_1 \in A_t] = \hat{p}_{st}$ for all $k \in A_s$, then it is clear also that $\mathbb{P}_{\pi'}[X_1 \in A_t] = \hat{p}_{st}$. Thus the probability in (2) depends only on A_s and A_t.

NB: This notion of lumpable Markov chain is similar to the definition of probabilistic bisimulation of Larsen and Skou [?], in the sense that lumping consists in creating a new Markov chain bisimilar to the original one and having less states.

Example 14. Let us consider the following matrix

$$P = \begin{array}{c} \\ R \\ N \\ S \end{array} \begin{array}{c} \begin{array}{ccc} R & N & S \end{array} \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix} \end{array}.$$

We note that the probability of moving from either of states R and S to N is the same. Hence if we choose for our partition $A = (\{N\}, \{R, S\}) = (G, B)$, the condition for lumpability is satisfied. The new transition matrix is

$$P' = \begin{array}{c} \\ G \\ B \end{array} \begin{array}{c} \begin{array}{cc} G & B \end{array} \\ \begin{pmatrix} 0 & 1 \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix} \end{array}.$$

Note that the condition for lumpability is not satisfied for the partition $A = (\{R\}, \{N, S\})$ since $p_{NA_1} = p_{NR} = \frac{1}{2}$ and $p_{SA_1} = p_{SR} = \frac{1}{4}$.

3 Time to Absorption: An Algebraic View

(KS, p. 22+46)

Proposition 6. *In any finite Markov chain, no matter where the process starts, the probability after n steps that the process is in a recurrent state tends to 1 as n tends to ∞.*

Definition 21. *A set of states $\mathcal{O}$ is an* open set *if from every state in $\mathcal{O}$ it is possible to go to a state in the complementary set $\neg\mathcal{O}$.*

Proposition 7. *A set of states $\mathcal{O}$ is open if and only if no recurrent set is a subset of $\mathcal{O}$.*

Lemma 5. *If A^n tends to $\mathbf{0}$ (zero matrix) as n tends to infinity, then $(I-A)$ has an inverse, and*

$$(I-A)^{-1} = I + A + A^2 + \cdots = \sum_{k=0}^{\infty} A^k.$$

Proposition 8. *Let $\mathcal{O}$ be an open set of states. Let Q be the $s \times s$ submatrix of P corresponding to these states. Let the process start in i. Then:*

1. *$I-Q$ has an inverse, and*

$$(I-Q)^{-1} = I + Q + Q^2 + \cdots = \sum_{k=0}^{\infty} Q^k.$$

2. *The i-th component of $\tau = (I-Q)^{-1}\mathbf{1}$ is the mean number of steps needed to leave $\mathcal{O}$, where $\mathbf{1}$ denotes the column vector with all entries 1.*

(BT,p.150)

Proposition 9. *If M is a nonnegative matrix then the following are equivalent:*

1. *$\rho(M) < 1$*
2. *There exists some w such that $\|M\|_\infty^w < 1$*
   (where $\|M\|_\infty^w = \max_i \frac{1}{w_i} \sum_{j=1}^n |m_{ij}|w_j = \max_i \frac{1}{w_i}[Mw]_i = \|Mw\|_\infty^w$).
3. *There exists some $\lambda < 1$ and $w > 0$ such that $Mw \leq \lambda w$.*

Proposition 10. *Given any square matrix M, there exists some $w > 0$ such that $\|M\|_\infty^w < 1$ if and only if $\rho(|M|) < 1$.*

Proposition 11. *If M is row diagonally dominant (i.e. $\sum_{j \neq i} |a_{ij}| < |a_{ii}|$, for all i), then the Jacobi-type iteration $x := Mx + b$ (for solving $Mx = b$) converges, as well as the corresponding Gauss-Seidel iteration (i.e., the iteration where the components are updated one at a time).*

It is easy to show that the set of transient states of a Markov chain is open. Therefore, the above theorem gives us the expected time to absorption of a Markov chain under an algebraic form.

Proposition 12. *(BT89, p. 170) Consider the directed graph associated with P[3]. Let $\{X(t) \mid t = 0, 1, ...\}$ be a Markov chain whose one-step transition probabilities are given by P. Let T be a positive integer and let*

$$\delta_T = \min_{i=2,\ldots,n} \mathbb{P}(\text{there exists some } \tau \leq T \text{ such that } X(\tau) = 1 | X(0) = i). \qquad (10.1)^4$$

(a) *Assume that, for all state i, there exists a positive path from i to state 1. Then, there exists $K > 0$ such that $\delta_T > 0$ for all $T \geq K$.*

(b) *If $\delta_T > 0$, then*
$\rho(\tilde{P}) \leq (\|\tilde{P}^T\|_\infty)^{1/T} \leq (1-\delta_T)^{1/T} < 1.$

[3] i.e., the graph with the state space $S = \{1, \ldots, r\}$ as its set of vertices and an oriented edge from vertex i to vertex j if and only if $p_{ij} > 0$.

[4] Alternatively: $[P^T]_{i1} > 0$ for all $i > 1$.

Proof. (a) The positivity of δ_T is a straightforward consequence of the fact that for each $i \neq 1$, there exists a sequence of positive probability transitions leading from i to 1. We simply need to take T large enough so that for each $i \neq 1$, there exists at least one such path that uses no more than T arcs.

(b) Let

$$Q = \begin{pmatrix} 1 & 0 \\ b & \tilde{P} \end{pmatrix},$$

which is easily seen to be a stochastic matrix, and let $\{Y(t)|t = 0, 1, ...\}$ be an associated Markov chain. We notice that $Y(t)$ has the same transition probabilities with $X(t)$ except that state 1 is an *absorbing* state: once $Y(t)$ becomes 1, it never changes. It follows that

$$\delta_T \leq \mathbb{P}(\text{there exists some } \tau \leq T \text{ such that } Y(\tau) = 1 | Y(0) = i) = [Q^T]_{i1}, \quad i = 2, ..., n \tag{8.8}$$

We now notice that Q^T is of the form

$$Q^T = \begin{pmatrix} 1 & 0 \\ c & \tilde{P}^T \end{pmatrix}, \tag{8.9}$$

where c is an $(n-1)$-dimensional column vector with all its entries bounded below by δ_T. Since Q is a stochastic matrix, so is Q^T and each row sums to 1. It follows from Eq. (8.9) that the sum of the entries in any row of $\tilde{P}^T$ is bounded above by $1 - \delta_T$. Therefore, $\rho(\tilde{P}^T) \leq \|\tilde{P}^T\|_\infty \leq 1 - \delta_T$. The result follows because $\rho(\tilde{P}) = (\rho(\tilde{P}^T))^{1/T}$.

3.1 Application

Proposition 13. *The expected timed to absorption of a lumped chain is equal to the expected time to absorption of the original chain.*

Example 15. (Herman after lumping). Configurations with k tokens are represented by k-tuples of the distances between conscutive tokens.

Lumpability is ensured due to the invariance by rotation of the (minimal distance) between tokens.

Let us now explain the computation of the D-transition matrix $\mathcal{Q}_k$ for $k = 2$ in Herman's example [**?**]. $\Delta_2 = \{1, 2, ..., m\}$ with $m = N/2$. Q_2 is the $m \times m$ matrix of components $\xi(d, e)$ (with $d, e \in \Delta_2$), of the form:

$$\begin{pmatrix} 1/2 & 1/4 & & & & & \\ 1/4 & 1/2 & 1/4 & & & & \\ & . & . & . & & & \\ & & . & . & . & & \\ & & & . & . & & \\ & & & 1/4 & 1/2 & 1/4 & \\ & & & & 1/4 & 3/4 & \end{pmatrix}$$

Note that, component $\xi(1, \perp) = 1/4$ is excluded from the matrix Q_2, since the $\perp$-column has been truncated. We then compute $B_2 = (\mathcal{I} - \mathcal{Q}_2)^{-1}$, which gives:

$$\begin{pmatrix} 4 & 4 & . & . & . & . & 4 \\ 4 & 8 & 8 & . & . & . & 8 \\ . & 8 & 12 & . & . & . & 12 \\ . & . & . & . & & . & . \\ . & . & . & & . & & . \\ . & . & . & & & . & . \\ 4 & 8 & 12 & . & . & . & 4m \end{pmatrix}$$

By corollary **??**, we know that the result of applying B_2 to $\mathbf{1}$ gives a column vector of d-component $2d(N - d)$, for $d \in \{1, ..., \lfloor N/2 \rfloor\}$. Therefore $E_2(d \rightsquigarrow^* \perp) = 2d(N - d)$. The maximal expected time corresponds to $d = \lfloor N/2 \rfloor = m$, and is $2m(m+1) \simeq N^2/2$. This corresponds to $E[\min\{t : X_t \in \mathcal{L} \mid X_0 = i\}]$ for $i \in X_2$.

4 Markov Chains with costs (BT89, p.308-312)

4.1 Motivation: Expected Time to Absorption (Bertsekas-Tsitsiklis 2002, p.341)

We now turn our attention to the expected number of steps until a recurrent state is entered (an event that we refer to as "absorption"), starting from a particular transient state. For any state i, we denote

$v_i = E[\text{number of transitions until absorption, starting from } i]$

$= E[\min\{n \geq 0 \mid X_n \text{ is recurrent}\}|X_0 = i]$.

Note that if i is recurrent, then $v_i = 0$ according to this definition.

We can derive equations for the v_i by using the total expectation theorem.[5] We argue that the time to absorption starting from a transient state i is equal to 1 plus the expected time to absorption starting from the next state, which is j with probability p_{ij}. We then obtain a system of linear equations, stated below, which has a unique solution.

NB: *finite* solution.

Proposition 14. *The expected times to absorption, $v_1, ..., v_m$ are the unique solution to the equations*

$v_i = 0$, *for all recurrent states* i,

$v_i = 1 + \sum_{j=1}^{m} p_{ij} v_j$, *for all transient states* i.

Proof. The statement $v_i = 0$ for all recurrent states i, is obvious. Let $\mathcal{R}$ be the set of recurrent states. Consider the case where i is transient. We have:

$v_i = E[\min t : X_t \in \mathcal{R} \mid X_0 = i]$

$= \sum_j (E[\min t : X_t \in \mathcal{R} \mid X_0 = i, X_1 = j] \times \mathbb{P}(X_1 = j \mid X_0 = i])$

$= \sum_j (E[\min t : X_t \in \mathcal{R} \mid X_1 = j] \times p_{ij})$

$= \sum_j (1 + E[\min t : X_t \in \mathcal{R} \mid X_0 = j]) \times p_{ij}$

$= \sum_j (1 + v_j) p_{ij}$

$= 1 + \sum_j v_j p_{ij}$

Let $\tilde{P}$ be the transition matrix of probabilistic transitions restricted to the subset of transient configurations. The probability of *not* stabilizing (i.e., reaching a recurrent state) on the very next step is $\tilde{P}\tilde{1}$ (a column vector indexed by the initial state); and thus in general $\tilde{P}^k\tilde{1}$ gives the probabilities that stabilization will not occur within k steps. From elementary probability, the *expected time to stabilization* is a column vector $\tilde{v} = (\sum_{k=0}^{\infty} \tilde{P}^k)\tilde{1}$ where this summation exists, provided stabilization occurs with probability 1: each element of the vector gives the expected time from that initial state. Where the summation does exist, matrix algebra shows that in fact we have $\tilde{v} = \tilde{P}\tilde{v} + \tilde{1}$. Thus if from every initial configuration i in S the expected time $\tilde{v}$ to stabilisation is finite, we know that it satisfies

$$\tilde{x} = \tilde{P}\tilde{x} + \tilde{1}.$$

Conversely, if we have some $\tilde{x}$ that satisfies $\tilde{x} = \tilde{P}\tilde{x} + \tilde{1}$ *uniquely* then, provided we have established (by some other means) that the expected time to stabilisation is everywhere finite, we will know it is given by $\tilde{x}$.

Proposition 15. *Let $\tilde{P}$ be the transition matrix of probabilistic transitions restricted to the subset of transient configurations. We have:*

- *For all i, the expected time to stabilization v_i from i is finite.*
- *$\tilde{v}$ is the unique solution of $\tilde{x} = \tilde{P}\tilde{x} + \tilde{1}$.*

4.2 A More General Approach: Expected Cost Vector

Consider a stationary discrete-time Markov chain $\mu = (X_t)$ with state space $S = \{1, 2, ..., m\}$, and transition probability matrix $P(\mu)$ with elements p_{ij}. Suppose that if the state is $X(t) = j$ at time t, there is a cost c_j^μ incurred (sometimes denoted by $c^\mu(j)$ or $c^\mu(X_T)$), where c_j^μ is a given scalar (I.e.: c^μ is a function from S to $\mathbb{R}$).

c^μ is called the *one-step cost vector* associated with μ. Since we have

$\mathbb{P}(X_t = j|X_0 = i) = [P^t(\mu)]_{ij}$.

In the following, we will often abbreviate $P(\mu)$ as P, and c^μ as c.

[5] (BT 2002,p.105):$E[X] = \sum_{i=1}^{n} \mathbb{P}(A_i)E[X|A_i]$, where $A_1, ..., A_n$ are disjoint events that form a partition of the sample space, ans $\mathbb{P}(A_i) > 0$ for all i. Furthermore: $E[X|B] = \sum_{i=1}^{n} \mathbb{P}(A_i|B)E[X|A_i \cap B]$.

Definition 22. *The* expected cost *associated with the system starting at state i is*
$v_i = \sum_{t=0}^{\infty} E_i[c(X_t)] = \sum_{t=0}^{\infty} \mathbb{P}_i(X_t = j)c(j) = \sum_{t=0}^{\infty} P_{ij}^t c(j) = \sum_{t=0}^{\infty} [P^t c]_i$.
Equivalently, we have
$v = \sum_{t=0}^{\infty} P^t c$ *(2.1),*
where v and c are the vectors with coordinates v_i and c_i, respectively.
v is called the expected cost vector *associated with P.*

NB: The expected time to stabilization corresponds to the case $\tilde{c} = \tilde{1}$.

Given a closed set $\mathcal{L}$, we will say that the Markov chain $\mu = (X_t)$ (or more simply P) is *stabilizing* w.r.t. $\mathcal{L}$ if $\forall i \mathbb{P}_i(T_{\mathcal{L}} < \infty)$ where $T_{\mathcal{L}}$ is the first time where a state of $\mathcal{L}$ is reached, starting from i (i.e. for all i, $T_{\mathcal{L}} = \min\{t : X_t \in \mathcal{L} | X_0 = i\}$.[6]

We will consider the special case where state 1 is cost-free and absorbing, and that μ is stabilizing w.r.t. $\mathcal{L} = \{1\}$. (It is the case when all the recurrent states have been lumped in a single class.)

Lemma 6. *Suppose that state 1 be absorbing and cost-free and that μ is stabilizing w.r.t. 1. Let L be the operator defined by $L(v) = c + Pv$. Then the expected cost vector v is such that:*
(a0) $(I - P)^{-1}$ is invertible, and v is finite and is equal to $(I - P)^{-1}c$.
(a1) v is the unique fixed point of L within V.[7]
(a2) $\lim_{t\to\infty}(L)^t(u) = v$ for all $u \in V$.
(a3) If there exists u such that $u \geq Lu$ then $u \geq v$.

Proof. $(a0)$-$(a1)$-$(a2)$. Let us write P in the form

$$P = \begin{pmatrix} 1 & 0 \cdots 0 \\ p_{21} & \\ \cdot & \\ \cdot & \tilde{P} \\ \cdot & \\ p_{n1} & \end{pmatrix}. \qquad (2.6)$$

Since μ is stabilizing, there exists some $\bar{t} > 0$ such that state 1 is reached with positive probability after at most $\bar{t}$ transitions regardless of the initial state ($[P^{\bar{t}}]_{i1} > 0$ for all i). It was shown in Prop. 12 that for some $\delta > 0$, we have
$\rho(\tilde{P}^{\bar{t}}) \leq \|\tilde{P}^{\bar{t}}\|_\infty \leq 1 - \delta.$ (2.7)
Since $c_1 = 0$, it is seen from Eq. (2.6) that for all $t \geq 0$, we have

$$P^t c = \begin{pmatrix} 0 \\ \tilde{P}^t \tilde{c} \end{pmatrix},$$

where

$$\tilde{c} = \begin{pmatrix} c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_m \end{pmatrix}.$$

From Eq. (2.7) we can see that the spectral radius $\rho(\tilde{P})$ of $\tilde{P}$ satisfies $\rho(\tilde{P}) < 1$. Therefore, $\tilde{P}^n$ tends to $\mathbf{0}$, and, by lemma 5, the matrix $(I - \tilde{P})$ is invertible. It follows that the equation $\tilde{x} = \tilde{c} + \tilde{P}\tilde{x}$ has a unique solution. This solution, call it x^*, satisfies
$x^* = \tilde{c} + \tilde{P}(\tilde{c} + \tilde{P}x^*) = \cdots = \sum_{t=0}^{m-1} \tilde{P}^t \tilde{c} + \tilde{P}^m x^*, \quad \forall m.$ (2.9)
Since $\rho(\tilde{P}) < 1$, we have $\lim_{m\to\infty} \tilde{P}^m x^* = 0$. Therefore, by taking the limit in Eq. (2.9) as $m \to \infty$, we obtain that the series $\sum_{t=0}^{\infty} P^t c$ is convergent and is the (unique) solution of $x^* = Lx^*$. Now this series defines the cost vector $\tilde{v}$. Hence:

$$\tilde{x^*} = \begin{pmatrix} v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{pmatrix}.$$

[6] We know that P is stabilizing iff $[P^{\bar{t}}]_{i1} > 0$, which is equivalent itself to the connectivity of i to 1.
[7] Recall that V is the set of vectors of $\mathbb{R}^m$ with first component $v_1 = 0$.

So $\tilde{v} = \tilde{c} + \tilde{P}\tilde{v}$, $c_1 = 0$. Furthermore, $v_1 = 0$. It follows using Eq. (2.6) that v satisfies the equation $v = L(v)(= c + Pv)$. Furthermore, since $\rho(\tilde{P}) < 1$, the iteration

$\tilde{v} := L(\tilde{v})$,[8]

converges to $\tilde{v}$ starting from an arbitrary initial condition. Hence $\lim_{t\to\infty}(L)^t(v) = v$ for all $v \in V$.

$(a3)$. Suppose $v \geq Lv$, for some v. Hence: $\tilde{v} \geq L\tilde{v}$. Since $\tilde{v} = L\tilde{v}$, we have: $\tilde{v} - \tilde{v} \geq \tilde{P}(\tilde{v} - \tilde{v})$. Hence: $\tilde{v} - \tilde{v} \geq \tilde{P}^m(\tilde{v} - \tilde{v})$, for all $m \geq 1$. Now $\tilde{P}^m \to \mathbf{0}$ when m tends to ∞ (since $\rho(\tilde{P}) < 1$). Hence, by taking the limits of the sides of the above inequation: $\tilde{v} - \tilde{v} \geq 0$. It follows: $v - v \geq 0$ (assuming $v_1 = 0$).

Lemma 6 can be rephrased as a "maximum principle" as follows.

Proposition 16. *(Maximum Principle) Let c be a cost vector on $S = \{1, \ldots, m\}$, $\mathcal{L}$ a subset of S, and P the transition matrix of a Markov chain (X_t) stabilizing w.r.t. $\mathcal{L}$ (i.e., for all i, $\mathbb{P}_i(\min\{t : X_t \in \mathcal{L}\} < \infty) = 1$). Suppose that $\mathcal{L}$ is closed and cost-free. If for some column vector u of size $|S|$,*

$$u \geq (resp. =) \begin{cases} 0 & \text{on } \mathcal{L}, \\ Pu+c & \text{on } S \setminus \mathcal{L}. \end{cases}$$

Then $u \geq (resp. =)\ v$, where v is the expected cost of absorption to $\mathcal{L}$ (i.e., $v_i = E_i[\sum_{t=0}^{\infty} c(X_t)]$ for $i \in S$).

NB: The proposition can be generalized by replacing the hypothesis $u \geq 0$ on $\mathcal{L}$ by $u \geq \varphi$ on $\mathcal{L}$, where φ is a function called the *final costS*, and v is $v_i = E_i[\sum_{0\leq t\leq T} c(X_t) + \varphi(X_T)1_{\{T<\infty\}}]$ for $i \in S$).

4.3 Recapitulation

Summary:

- Series $\sum_{t=0}^{\infty} P^t c$ converges to, say, $\tilde{v}$.
- $\tilde{x} = \tilde{c} + \tilde{P}\tilde{x}$ has $\tilde{v}$ as a unique solution.
- Iteration of $\tilde{v} = \tilde{c} + \tilde{P}\tilde{v}$, starting from arbitrary initial vector, converges to $\tilde{v}$.

NB:

- Gives a linear programming scheme of resolution
- For $\tilde{c} = 1$, we compute the expected time of absorption (or exit from the transient space)
- using lumping, these results of exiting from the transient space generalize to the case where the absorbing states are not unique. It suffices to consider a set $\mathcal{L}$ closed and cost-free, such that from any $i \in S$ there exists a path from i to $\mathcal{L}$, and take for V, the set of vectors of components null on states of $\mathcal{L}$.

[8] It can be shown that this iteration converges even if executed in a totally asynchronous environment.

5 Maximum Principle: Applications

5.1 Application to Herman with three tokens (McIver-Morgan)

Example 16. Application to Herman with 3 tokens (McIver-Morgan) Consider the special case in which exactly three processors have tokens initially. we give an exact value for the expected time to stabilisation.

Lemma 7. *The expected time to stabilisation of a ring with initially three tokens is $4abc/N$, where a, b, c are the initial separations of the token. (Note that $a+b+c=N$).*

Proof. Let $\mathcal{L}$ be the set of one-token configurations, and $\tilde{S} = S \setminus \mathcal{L}$ be the set of three-token configurations of the ring. Define column vector $\tilde{u}$ over $\tilde{S}$ so that for r in $\tilde{S}$ the r-entry $\tilde{u}_r$ is $4a_r b_r c_r/N$, where a_r, b_r, c_r are the particular separations a, b, c in that configuration r; let $\tilde{P}$ be the reduced Markov matrix for $\tilde{S}$ only. Direct calculation shows that $\tilde{P}\tilde{u} = \tilde{u} - 1$, as explained hereafter. In a three-token system there are eight equiprobable outcomes for a single step, ranging from "all tokens kept" to "all tokens passed". Their effect transformed a, b, c into a', b', c' respectively, as tabulated below:

token movements	a'	b'	c'
no token passed:	a	b	c
one token passed:	$\begin{cases} a-1 \\ a+1 \\ a \end{cases}$	$\begin{cases} b \\ b-1 \\ b+1 \end{cases}$	$\begin{cases} c+1 \\ c \\ c-1 \end{cases}$
two tokens passed:	$\begin{cases} a \\ a+1 \\ a-1 \end{cases}$	$\begin{cases} b-1 \\ b \\ b+1 \end{cases}$	$\begin{cases} c+1 \\ c-1 \\ c \end{cases}$
all tokens passed:	a	b	c

Direct calculation of $\tilde{P}\tilde{u}$ for initial configuration a, b, c gives

$$1/8 \times 4/N \times \begin{cases} abc \\ +(a-1)b(c+1) + (a+1)(b-1)c + a(b+1)(c-1) \\ +a(b-1)(c+1) + (a+1)b(c-1) + (a-1)(b+1)c \\ +abc, \end{cases}$$

which via $a+b+c=N$ is readily simplified to $4abc/N - 1$ as required. It follows by Prop. 16 that the r-entry of $\tilde{u}$ gives the expected time to stabilisation for each r in $\tilde{S}$.

Example 17. Gambler's ruin

5.2 Application to Herman with *k* tokens

Example 18. Herman with k tokens via decreasing function (McIver-Morgan).

Proposition 17. *The expected time to stabilisation of Herman's Ring is $O(N^2)$.*

Proof. Choose for an upper bound the column vector $\tilde{u}$ of height $n-1$ whose r-entry is $2x_r(2N - x_r - 1)$ for each configuration r, where x_r is the *extent* of r, the minimum number of contiguous segments containing all tokens in r. Note that since $x_r \leq N-1$ for all configurations r, each entry of u is $O(N^2)$ as a function of N. Elementary (but detailed) calculation shows that $\tilde{u} \geq \tilde{P}\tilde{u} + \tilde{1}$.

Hence we have $\tilde{u} \geq \tilde{v}$ from Prop. 16, where $\tilde{v}$ is the vector of expected time to absorption, giving that each entry is $O(N^2)$.

5.3 Application to Israeli-Jalfon on a Ring

Avec Israeli-Jalfon, chaque politique (Markovienne) est caractérisée par une matrice de transition P particulière. Nous appliquons les critères ci-dessus pour montrer l'équivalence et l'optimalité en temps moyen des stratégies pour Israeli-Jalfon sur anneau, pour aller d'une classe de configurations initiales à une classe de configurations finales. On supposera prouvé que la probabilité d'aller d'une configuration initiale à une configuration finale est 1, et on "lumpera" toutes les configurations finales en un seul état (assimilés à un état "1" unique et cost-free).

L'idée pour montrer que, pour un problème donné, toutes les politiques sont *équivalentes* est de calculer le coût v^μ pour une politique μ particulière, et de montrer $v^\mu = Pv^\mu + 1$ sur $S \setminus \mathcal{L}$, où P représente la matrice générique associée à une politique quelconque.

Pour montrer qu'une politique particulière de coût v^μ est *maximale*, on montre $v^\mu \geq Qv^\mu + 1$, où Q représente la matrice générique Q associée à une politique quelconque.

NB: Pour trouver le coût v^μ d'une politique particulière, on peut soit y arriver par calcul, soit l'intuiter et vérifier qu'elle satisfait $v^\mu = Pv^\mu + 1$ où P désigne la matrice associée à la politique en question. (Le résultat provient de l'unicité de la solution de l'équation, en vertu de l'hypothèse de convergence presque sûre vers un état final.)

Nous obtenons des résultats qui se résument ainsi :

- Pour aller de 3 jetons distincts à un jeton, toutes les politiques sont équivalentes.
- Pour aller de 2 jetons distincts à 1 jeton, dans le cas où $p \neq 1/2$, la politique consistant à bouger le jeton le plus éloigné de l'autre pour le sens horaire, est optimale.

On considère Israeli-Jalfon sur un anneau avec 3 jetons. Soit 3 jetons sur un anneau de longueur N, et soit a, b, c les distances respectives des jetons 2 à 2 ($a + b + c = N$ et $a, b, c \geq 0$).

$a = 0$ (resp. $b = 0$, $c = 0$) signifie que deux jetons sont confondus; $a = b = 0$ (resp. $b = c = 0$, $c = a = 0$) signifie que les 3 jetons sont confondus.

$a = 1$ (resp. $b = 1$, $c = 1$) signifie que deux jetons sont contigus; $a = b = 1$ (resp. $b = c = 1$, $c = a = 1$) signifie que les 3 jetons sont contigus.

$a \geq 2 \wedge b \geq 2$ (resp. $b \geq 2 \wedge c \geq 2$, $c \geq 2 \wedge a \geq 2$) signifie qu'un jeton est isolé;

$a \geq 2 \wedge b \geq 2 \wedge c \geq 2$ signifie que les 3 jetons sont isolés.

Une configuration M sera caractérisée par le triplet (a, b, c) des distances respectives de ses 3 jetons.[9]

On considèrera une politiques particulière μ_0 consistant à bouger systématiquement un jeton donné, par exemple: un jeton isolé. Pour calculer v^{μ_0}, le temps moyen d'absorption associé à μ_0, on utilisera un résultat classique sur le temps moyen des marches aléatoires symétriques sur un segment (ou "gambler ruin"), pour dire : étant donnée une configuration de départ (a, b, c), si on bouge systématiquement le même jeton (par exemple, en faisant varier a et c, et laissant b invariant), le temps moyen pour que a ou c devienne 0 (resp. 1) est de ac (resp. $(a-1)(c-1)$).

Israeli-Jalfon 1 on a Ring Soit C_1 les configurations où les 3 jetons sont confondus ($a = b = 0$, $b = c = 0$ ou $c = a = 0$). Posons, pour tout point $M = (a, b, c)$:

$v^{\mu_0}(M) = ac + b(a + c) = ab + bc + ca$.

NB: Il est facile de voir que: $v^{\mu_0}(M)$ correspond à $E[\min\{t : Y_t \in C_1 \mid Y_0 = M\}]$, où Y_t correspond à la stratégie qui consiste à choisir, une fois pour toutes, un jeton de la configuration initiale, à déplacer systématiquement ce jeton jusqu'à ce qu'il se confonde avec un des deux autres jetons, puis à bouger systématiquement un des deux jetons distincts restants (s'il en existe deux) jusqu'à ce qu'il se confonde avec l'autre, produisant ainsi une configuration de C_1.

Considérons maintenant une stratégie Q qui consiste à bouger, à chaque étape, un jeton quelconque (pas nécessairement toujours le même comme dans (Y_t)). Montrons que, pour tout $M \notin C_1$ (i.e, tel que M contienne au moins deux jetons distincts), on a: $v^{\mu_0}(M) = Qv^{\mu_0}(M) + 1$.

- Si les 3 jetons dans $M : (a, b, c)$ sont distincts ($a, b, c \geq 1$), on peut supposer, pour fixer les idées, que Q bouge M en modifiant a et b, mais pas c. (Les autres cas sont analogues). On a:

[9] Il y a du "lumping" dans cette caractérisation, car on agrège des configurations qui ne diffèrent que par rotation.

$Qv^{\mu_0}(a,b,c)$
$= 1/2v^{\mu_0}(a-1,b+1,c) + 1/2v^{\mu_0}(a+1,b-1,c)$
$= 1/2((a+1)(b-1) + (b-1)c + c(a+1)) + 1/2((a-1)(b+1) + (b+1)c + c(a-1))$
$= ab + bc + ca - 1$
$= v^{\mu_0}(a,b,c) - 1.$

– Si deux jetons seulement sont distincts (par exemple $a = 0, b \geq 1, c \geq 1$), il est facile de voir que: $v^{\mu_0}(M) = Qv^{\mu_0}(M) + 1$, où Q est une matrice de transition qui fait bouger, à chaque étape, l'un des deux jetons de M, quel qu'il soit.

Par conséquent, en vertu de la proposition 16, $v^{\mu_0}(M)$ est égal au temps moyen d'absorption $v^Q(M)$ associé à la politique Q. On a ainsi montré l'équivalence de toutes les politiques du point de vue du temps moyen d'absorption.

NB1: Il resterait en fait à montrer, que toute politique Q atteint C_1 en temps fini avec probabilité 1.

Israeli-Jalfon on a Ring à 2 jetons with $q > p$ One supposes now that $q > p$. One considers configurations with two tokens only (configurations (d,e) with $d \geq 2, e \geq 2, d+e = N$). Recall that d is the distance from token 1 to token 2 in the clockwise sense. If $q > p$, it means that it is faster for a token to move anticlockwise than clockwise. Therefore a "maximal" strategy, say μ, should select token 2 (resp. 1) if its "anticlockwise" distance to the other one is the bigger: it selects token 1 if $d \leq e$ (anticlockwise distance from 1 to 2 equals e), and token 2 otherwise (anticlockwise distance from 2 to 1 equals d). Under such a policy, the system behaves as a biased random wal on segment $0, ..., N/2$ with a reflecting barrier at $N/2$. Let $\Phi_\mu(e,d)$ be the expected time to reach $\mathcal{L}_2$, i.e. a configuration with $e = 1$ or $d = 1$, under this policy. Let us sketch out how to show that μ is maximal indeed. By question 1, any arbitrary policy $\mathcal{P}$ is "terminating", therefore, in order to prove that $v^\mu(d,e)$ is slower than $\mathcal{P}$, it suffices, by the "maximum principle", to show, for all $(d,e) \notin \mathcal{L}_2$:

$$v^\mu(d,e) \geq 1 + pv^\mu(d_1,e_1) + qv^\mu(d_2,e_2) \qquad (*),$$

where (d_i,e_i) is the configuration resulting from (d,e) by moving the token selected by $\mathcal{P}$ clockwise for $i = 1$, anticlockwise for $i = 2$.
For (d,e) where the policy of $\mathcal{P}$ coincides with the policy of μ, checking the inequality $(*)$ is immediate. Consider now a configuration (d,e) where the policy of $\mathcal{P}$ differs from the policy of μ. In order to fix the idea, suppose $d \leq e$. (The other case is symmetrical.) Then $\mathcal{P}$ selects token 2, and $(d_1,e_1) = (d+1, e-1)$, $(d_2,e_2) = (d-1, e+1)$. The inequality $(*)$ becomes:
$v^\mu(d,e) \geq 1 + pv^\mu(d+1,e-1) + qv^\mu(d-1,e+1)$
On the other hand, by first step analysis, since μ moves the 1st token, we have:
$v^\mu(d,e) = 1 + pv^\mu(d-1,e+1) + qv^\mu(d+1,e-1)$
We have therefore to show, when $d \leq e$:
$(q-p)v^\mu(d+1,e-1) \geq (q-p)v^\mu(d-1,e+1)$, i.e.
$v^\mu(d+1,e-1) \geq v^\mu(d-1,e+1)$.
This is to say that μ takes more time to reach $\mathcal{L}_2$ from $(d+1,e-1)$ than from $(d-1,e+1)$. This is (intuitively) true because the time needed to go from a configuration (d,e) to $\mathcal{L}_2$ increases with d when d ranges from 0 to $N/2$.

NB: We have proved the optimality of μ without using an explicit formulation of the associated cost v, but using only the property: $v(d+1) \geq v(d-1)$ for $0 < d < N/2$ (resp.: $v(d-1) \geq v(d+1)$ for $N/2 < d < N-1$).

5.4 Application to Israeli-Jalfon on a Graph (Tetali-Winkler)

IJ on a Graph: Two Tokens The *hitting time* $H_G(x,y)$ from x to y is defined to be the expected number of steps for a random walk on G beginning at vertex x to reach vertex y for the first time. The *meeting time* $M_G(x,y)$ is defined to be the expected number of moves before tokens placed initially at vertices x and y of G meet, given optimal (delaying) play by the demon in deciding at each step which token will move.

Theorem 15. *Let G be any connected, undirected graph, and let t be a remote*[10] *vertex of G. Then for every pair x, y of vertices of G,*

$$M_G(x,y) \leq H_G(x,y) + H_G(y,t) - H_G(t,y).$$

Proof. Let μ a given policy deciding at each step which token will move. Let P_M^μ be the transition matrix associated to the Markov chain of playing x or y on G according to μ (until meeting occurs). Let $v_G^\mu(x,y)$ be the expected number of moves before tokens placed initially at vertices x and y meet ($M_G(x,y) = \max_\mu(v^\mu(x,y))$).

We define a vector u defined for each couple (x,y) of vertices of G (i.e., for each state) in accordance with the right-hand side of the above inequality, as follows:

$$u(x,y) = H_G(x,y) + H_G(y,t) - H_G(t,y) = H_G(y,x) + H_G(x,t) - H_G(t,x)$$

(u is symmetric on account of the fact: $H_G(x,y)+H_G(y,z)+H_G(z,x) = H_G(x,z)+H_G(z,y)+H_G(y,x)$, due to the reversibility of the Markov chain for random walks on an undirected graph.) Since u is symmetric, no matter which token μ decides to move, its expected value will decline by 1. Furthermore u is non-negative (on account of the remoteness of t). It follows by Prop. 16: $u(x,y) \geq v_G^\mu(x,y)$ for all μ, hence: $u(x,y) \geq M_G(x,y)$.

Corollary 1. $M_G(x,y) \leq \frac{8}{27}n^3$.

Proof. Follows from Theorem 15 and Brightwell-Winkler $(4/27)n^3$ upper bound for the worst-case value of $H_G(x,y)$.

IJ on a graph: k tokens Suppose we reinstate the demon but give him more than two tokens to work with; as in the description of Israeli and Jalfon self-stabilizing token management scheme above, the rule is that whenever two tokens meet one is eliminated. However, for our purposes it is more convenient to think of tokens not being eliminated but simply "glued together" when they meet; then when there are k tokens at the beginning, we can employ throughout the multivariate potential

$$u_k(x_1, ..., x_k) = \frac{1}{k-1} \sum_{i<j} M_G(x_i, x_j)$$

where the x_i's will cease to be all distinct once collisions begin. Suppose, for example, that tokens corresponding to indices $i \in I$ are currently on vertex v and are designated for movement by the demon; then the expected value of $M_G(x_i, x_j)$ drops by at least 1 when $|\{i,j\} \cup I| = 1$, which occurs for $|I|(k-|I|) \geq k-1$ of the terms, while the other terms remain constant. hence the expected value of u_k diminishes by at least 1 at every step, as desired.

Note that $M_G(x_i, x_j)$ remains at zero when i and j are both in I, since tokens i and j continue to travel to the same vertex. It is for this reason that we use M_G rather than our original v in the definition of u_k; the expected value of $M_G(v)$ can in some circumstances jump as v moves, e.g. when v is remote.

Now the application of the maximum principle shows that the expected time before reducing to a single token is at most the maximum value of u_k on G, which is in turn bounded by k times the maximum 2-token meeting time.

Theorem 16. *In any graph G on n vertices, the maximum meeting time is bounded by $(8/27)n^3$.*

Optimal Policy by Maximum Principle (Brémaud,p.184) We consider a stochastic process $\{X_n\}$ with values in S, that is controlled in the following way. Let A some set, the set of *actions*, and $\{p_{ij}(a); i,j \in S\}_{a \in A}$ a family of transition probabilities on S. A *control policy* μ is a (measurable) function $\mu : S \to A$ which prescribes to take action $\mu(i)$ when the process is in state i. Let Pol be the set of control policies.

Let $\{P(\mu)\}_{\mu \in Pol}$, be a family of transition matrices on S, with the interpretation that, if at time n the controlled process is in state i, and if the controller takes action $a = \mu(i)$, then at time $n+1$ the state will be j with probability $p_{ij}(a)$.

[10] A vertex t is *remote* if it is minimal for the vertex relation $\leq$, defined by $u \leq v$ iff $H_G(u,v) \leq H_G(v,u)$, which can be shown to be transitive (i.e., is a pre-order).

Therefore, under the strategy μ, the controlled process is an MC with transition matrix $P(\mu)$, where

$$(P(\mu))_{ij} = p_{ij}(\mu(i)).$$

There is a cost v_i^μ associated with each strategy μ and each initial state i, of the form

$$v_i^\mu = E_i^\mu[\sum_{0 \le k} c^\mu(X_k)],$$

where c^μ as in Maximum Principle Proposition 16, with $S \setminus \mathcal{L}$ fixed, and moreover, $c_i^\mu = c(i, \mu(i))$, for appropriate functions c. The problem of *optimal control* is that of finding, if it exists, an *optimal policy* μ^*, such that

$$v_i^{\mu^*} \ge v_i^\mu,$$

for all states i, all policies μ. We have the following result.

Suppose that there exists a function $v^* : S \to \mathcal{R}$ such that

$$v_i^* = sup_{a \in A}\{\sum_{j \in S} p_{ij}(a) v_j^* + c(i, a)\}$$

for all $i \in S \setminus \mathcal{L}$, and

$$v_i^* = 0$$

for all $i \in \mathcal{L}$, and that the suprema above are attained for $a = \mu^*(i)$, for some (measurable) function $\mu^* : S \to A$. Then μ^* is an optimal control and $v^* = v^{\mu^*}$.

Proof. Since for all controls μ, $v^* \ge P^\mu v^* + c^\mu$ on $S \setminus \mathcal{L}$, and $v^* \ge 0$ on $\mathcal{L}$, it follows from the (extended version of) maximum principle that $v^* \ge v^\mu$ for all controls μ. Also $v = v^{\mu^*}$ and therefore u^* is an optimal control.

6 Markov Decision Processes with Costs

6.1 Markov Decision Processes

A *Markov decision process* (MDP) is a generalization of a markov chain in which nondeterministic choice coexist with probabilistic one. Given a countable set C, let $\mathcal{D}(C)$ be the set of probability distributions over C, i.e., the set of functions $f : C \mapsto [0,1]$ such that $\sum_{x \in C} f(x) = 1$.

Definition 23. *An MDP $(S, Acts, A, p)$ consists of the following components:*

- *a finite set S of states;*
- *a finite set $Acts$ of actions;*
- *a function $A : S \mapsto 2^{Acts}$ that associates with each $i \in S$ a finite set $A(i) \subseteq Acts$ of actions available at s;*
- *a function $p : S \times Acts \mapsto \mathcal{D}(S)$ that associates with each $i, j \in S$ and $a \in A(i)$ the probability $p(i,a)(j)$, sometimes written $p_{ij}(a)$, of a transition from i to j when action a is selected.*

Il est souvent utile pour décrire un processus de décision markovien $\Pi = (S, Acts, A, p)$ d'avoir une représentation graphique adaptée. Nous allons décrire ici celle qui est le plus souvent utilisée. La représentation est celle d'un graphe dans lequel les noeuds sont les éléments $i \in S$. Pour tout $i \in S$ et $a \in A(i)$ nous dessinons un faisceau d'arcs pour chaque j tel que $p_{ij}(a) > 0$. Les arcs appartenant au même faisceau sont reliés ensemble par un petit arc de cercle, et chaque faisceau porte le label de l'action correspondante.

Example 19. **Le barman aveugle**

Cet exemple classique de théorie des automates va nous permettre d'illustrer la manière dont on représente graphiquement les processus de décision markoviens. Un barman a les yeux bandés et il y a quatre jetons sur son plateau disposés pour former un carré (voir figure 2 (une face des jetons est blanche l'autre face est noire). A chaque coup, il décide s'il retourne un jeton (action Un), deux jetons adjacents (action Adj) ou deux jetons opposés (action Opp), puis tire au sort aléatoirement le (ou les jetons qu'il retourne). Son but, est que les jetons soient tous retournés dans le même sens à la fin. Voici la représentation graphique du processus de décision markovien associé dans lequel chaque état est la classe des configurations à inversion de couleur près et à rotation près: Nous pouvons détailler ici l'ensemble des

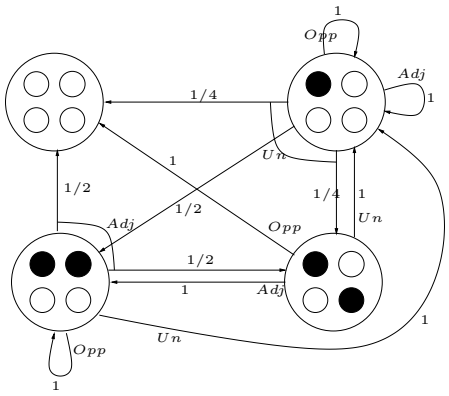

Fig. 2. représentation graphique du problème du barman aveugle

caractéristiques de ce processus de décision markovien:

- $S = \{1, 2, 3, 4\}$ où :
  - 1 est l'état qui représente les configurations dans lesquelles les quatre jetons sont de la même couleur.
  - 2 celui qui représente les configurations avec un jeton d'une couleur différente des trois autres.
  - 3 celui qui représente les configurations avec les deux jetons opposés de la même couleur.
  - 4 qui représente les configurations avec deux jetons adjacents de la même couleur.

– Pour tout $i \in \{2,3,4\} A(i) = \{Un, Opp, Adj\}$, $A(1) = \emptyset$.
– Et les différentes lois de probabilités:

$j =$	1	2	3	4
$p_{2,j}(Un)$	1/4	0	1/4	1/2
$p_{2,j}(Opp)$	0	1	0	0
$p_{2,j}(Adj)$	0	1	0	0
$p_{3,j}(Un)$	0	1	0	0
$p_{3,j}(Opp)$	1	0	0	0
$p_{3,j}(Adj)$	0	0	0	1
$p_{4,j}(Un)$	0	1	0	0
$p_{4,j}(Opp)$	0	0	0	1
$p_{4,j}(Adj)$	1/2	0	1/2	0

Definition 24. *Given an MDP $(S, Acts, A, p)$, a subset $\mathcal{L}$ of states is* closed *if:*
$\forall i \in \mathcal{L}, a \in A(i) : \ p_{ij}(a) > 0 \Rightarrow j \in \mathcal{L}$.

When a closed set $\mathcal{L}$ is reduced to one state, say 1, we say that 1 is "absorbing".

Definition 25. *Given an MDP $(S, Acts, A, p)$ and a closed set $\mathcal{L}$ a policy μ is* stabilizing w.r.t. $\mathcal{L}$ *if, for all $i \in S$:*
$\mathbb{P}_i^\mu(T_\mathcal{L} < \infty) = 1$, *where $T_\mathcal{L} = \min\{t : X_t \in \mathcal{L}\}$ is the first (random) time of entrance in $\mathcal{L}$ (starting from $X_0 = i$ under policy μ).*

Given a closed set $\mathcal{L}$, we denote by Pol_P the class of stabilizing policies (w.r.t. $\mathcal{L}$).

6.2 Markov Decision Processes with Costs (BT89, pp. 312+317)

Consider an MDP $(S, Acts, A, p)$. At each state i, we are given a set of actions $A(i)$. We now suppose that we are also given a *cost function* $c : S \times Acts \mapsto \mathbb{R}$, that associates with each state i and each action $a \in A(i)$ the scalar $c(i,a)$ or sometimes $c_i(a)$ If the state is i and action a is chosen at time t, the cost incurred is $c_i(a)$.

Consider the finite set of functions μ that map states i into actions $\mu(i) \in A(i)$, that is the set:

$$Pol_D = \{\mu \mid \mu(i) \in A(i), \ i = 1, \ldots, m\}.$$

We can regard each μ as a *memoryless* or *deterministic* policy that chooses always the same action when the system is at state i, whenever the time t is.

The cost associated to a path $i_0 a_0 i_1 \cdots$ corresponds to the sum $\sum_{k \geq 0} c_{i_k}(a_k)$.

Suppose that there is a cost-free state, say state 1, which is absorbing (i.e.: For all $a \in A(1), c_1(a) = 0$ and $p_{11}(a) = 1$). Our objective will be to find the "minimum expected cost" to reach such a state 1, and the associated policy. Formally, let: $T_{\{1\}} = \min\{k \mid X_k = 1\}$ be the random variable indicating the first time of entrance in state 1. We have:

Definition 26. *Given an MDP with cost, suppose that there is a cost-free state, say state 1, which is absorbing. The* (expected) cost v_i^μ *of a policy μ at $i \in S \backslash \{1\}$ is defined by*

$$v_i^\mu = E_i^\mu[\sum_{k=0}^{T_{\{1\}}-1} c(X_k, \mu(X_k))] = E_i^\mu[\sum_{k=0}^{\infty} c(X_k, \mu(X_k))].$$

The *(expected) cost vector* v^μ of policy μ is the vector whose i-th component is v_i^μ $(1 \leq i \leq m)$. Note that $v_1^\mu = 0$.

Definition 27. *Given an MDP with cost $(S, Acts, A, P, c)$, suppose that there is a cost-free state, say state 1, which is absorbing, the* minimum cost problem *is to determine, for all $i \in S \backslash \{1\}$*

$$v_i^* = \inf_{\mu \in Pol_P} v_i^\mu$$

The minimum cost vector *is the vector v^* whose i-th component is v_i^* (i.e., the minimum cost starting at state i).*

We say that the policy μ is *optimal* if μ is stabilizing and, for all i: $\quad v_i^\mu = v_i^*$.

The advantage of the minimum cost problem is to allow us to unify the problems of finding the *minimum* and the *maximum* expected time to absorption, which are seen as two instances of the minimum cost problem, with $c = \mathbf{1}$ for the minimum expected time and $c = -\mathbf{1}$ for the maximum one.

6.3 One-Step Cost Operators

Deterministic policies Let μ be a deterministic policy. Let $P(\mu)$ be the transition probability matrix corresponding to μ, that is, the matrix with elements

$$[P(\mu)]_{ij} = p_{ij}(\mu(i)), \text{ for } i, j = 1, ..., m.$$

Let also

$$c^\mu = \begin{pmatrix} c_1(\mu(1)) \\ . \\ . \\ . \\ c_m(\mu(m)) \end{pmatrix}$$

For a deterministic (memoryless) policy μ we have $\mathbb{P}(X_t = j \mid X_0 = i, \text{ and } \mu \text{ is used}) = [P^t(\mu)]_{ij}$. Therefore, if v_i^μ is the expected cost corresponding to initial state i and policy μ, and v^μ is the vector of coordinates $v_1^\mu, ..., v_n^\mu$, we have: $v^\mu = \lim_{k\to\infty} \sum_{t=0}^k (P^t(\mu))c^\mu$. [11]

Given a deterministic policy μ, it is convenient to introduce the mappings $L^\mu : \mathbb{R}^m \mapsto \mathbb{R}^m$ and $L^* : \mathbb{R}^m \mapsto \mathbb{R}^m$ defined by:
$L^\mu(v) = c^\mu + P(\mu)v$,
$L^*(v)$ is the vector whose ith component $[L^*(v)]_i$ is: $\min_{a\in A(i)}[c_i(a) + \sum_{j=1}^m p_{ij}(a)\ v_j]$.

Note that L^μ is the mapping involved in the iteration $v := c + Pv$ of Section 4 with c and P replaced by c^μ and $P(\mu)$, respectively. After k iterations, we have: $(L^\mu)^k(v) = P^k(\mu)v + \sum_{t=0}^k [P^t(\mu)]c^\mu$.
Therefore, for a deterministic policy μ, we can write the expected cost vector: $v^\mu = \lim_{k\to\infty}(L^\mu)^k(v^0)$, where v^0 is the zero vector $(0, 0, ..., 0)$.

The following proposition gives a basic property of L^*, and L^μ.

Proposition 18. *Let $\mu \in Pol_D$, $j \in S$ and $a \in A(j)$. Then L^* and L^μ (in the sense that $L^*(v) \le L^*(v')$ and $L^\mu(v) \le L^\mu(v')$ for all v, v' such that $v \le v'$).*

6.4 Computing the Minimal Cost Vector

We are going to show that, under reasonable conditions, the minimum cost vector, denoted by v^*, is a fixed point of the mapping L^* (i.e., $v^* = L^*(v^*)$), and can be obtained in the limit through the "dynamic programming" iteration $v := L^*(v)$.

The minimal cost problem can be seen as a generalisation of the shortest path problem in a (non-probabilistic) graph.[12]

[11] This is justified as follows: $v_i^\mu = E_i^\mu[\sum_t c(X_t, \mu(X_t))] = \sum_t E_i^\mu[c(X_t, \mu(X_t))] = \sum_t(\sum_j \mathbb{P}_i(X_t = j)c(j, \mu(j))) = \sum_t(\sum_j [P^t(\mu)]_{ij}c_j^\mu) = \sum_t (P^t(\mu)c^\mu)_i$. Hence $v^\mu = \sum_t (P^t(\mu))c^\mu$.

[12] The shortest path problem is an important example of dynamic programming problem where one makes assumptions similar to H1 and H2. It concerns a directed graph consisting of m nodes, numbered $1, ..., m$. Node 1 is a special node called the "destination", and has no outgoing arcs. Each arc (i, j) is given a scalar c_{ij} called "length". The problem is to find a path of minimum length (or shortest path) from each node to the destination. The Bellman-Ford algorithm solves the problem by iterating on v:
$v_i := \min_{j\in A(i)}(c_{ij} + v_j), \quad i = 2, ..., m$
$v_1 := 0$.
where $A(i)$ denotes here the set of all the nodes j for which there is an outgoing arc (i, j) from node i. This iteration converges to the solution provided that two assumptions on the graph are satisfied, that correspond to H1 and H2. This solution is then the (unique) solution of of "Bellman's equation":

We will operate under the following assumption:

Assumption H.
State 1 is absorbing and cost-free (i.e., $p_{11}(a) = 1$ and $c_1(a) = 0$ for all $a \in A(1)$), and furthermore:

- H1: there exists at least one stabilizing deterministic[13] policy.[14]
- H2: Each non-stabilizing deterministic policy yields infinite cost for at least one initial state, i.e.: for each non-stabilizing μ, there is a state $i \in S\backslash\{1\}$ such that $v_i^\mu = \lim_{k\to\infty}[\sum_{t=0}^{k} P^t(\mu)c^\mu]_i = \infty$.

We have the following lemmas.

Lemma 8. *Let Assumption* **H** *hold:*
(a1) Let μ a stabilizing deterministic policy. Then v^μ is the unique fixed point of L^μ within V.
(a2) Let μ a stabilizing deterministic policy. Then $\lim_{t\to\infty}(L^\mu)^t(v) = v^\mu$ for all $v \in V$.
(b) Let μ be a deterministic policy. If there exists $v \in V$ such that $v \geq L^\mu(v)$, then μ is stabilizing.

Proof. $(a1)$-$(a2)$: If μ is stabilizing, the conclusion follows from Lemma 6.
(b) : If $v \in V$ and $v \geq L^\mu(v)$, then by the monotonicity of L^μ,
$v \geq (L^\mu)^t(v) = P^t(\mu)v + \sum_{k=0}^{t-1} P^k(\mu)c^\mu, \quad \forall t \geq 1.$
If μ were non-stabilizing, then some subsequence of $\sum_{k=0}^{t-1} P^k(\mu)c^\mu$ would have a coordinate that tends to infinity (by H2), thereby contradicting the above inequality.

Lemma 9. *Let Assumption* **H** *hold. Then:*
L^* has at most one fixpoint within the subspace $V = \{v \in \mathbb{R}^m | v_1 = 0\}$.

Proof. Let us show that L has at most one fixed point within V. Indeed, if v and v' are two fixed points in V, then we select μ and μ' such that $v = L^*(v) = L^\mu(v)$ and $v' = L^*(v') = L^{\mu'}(v')$. By Lemma 8 (b), we have that μ and μ' are stabilizing, and furthermore $v = v^\mu$ and $v' = v^{\mu'}$. We have $v = (L^*)^t(v) \leq (L^{\mu'})^t(v)$ for all $t \geq 1$, and by Lemma 8 $(a2)$, we obtain $v \leq \lim_{t\to\infty}(L^{\mu'})^t(v) = v^{\mu'} = v'$. Similarly, $v' \leq v$, showing that $v = v'$ and that L^* has at most one fixed point within V.

Lemma 10. *Let Assumption* **H** *hold. Then:*
There exists a stabilizing deterministic policy μ^* such that $v = L^{\mu^*}(v)$ iff $v = L^*(v)$.
This shows thats L^* has a unique fixpoint, which is equal to v^{μ^*}.

Proof. Let us show that L^* has a fixed point within V, and exhibit a stabilizing deterministic policy μ^* such that $L^* = L^{\mu^*}$. Let μ be a stabilizing policy (whose existence is guaranteed by H1). Choose $\mu' \in Pol_D$ such that $L^{\mu'}(v^\mu) = L^*(v^\mu)$. Then we have $v^\mu = L^\mu(v^\mu) \geq L^{\mu'}(v^\mu)$. By the Maximum Principle, it follows:

$$v^\mu \geq v^{\mu'}. \tag{3.12}$$

If $v^\mu = v^{\mu'}$, then we obtain $v^\mu = v^{\mu'} = L^{\mu'}(v^{\mu'}) = L^{\mu'}(v^\mu) = L^*(v^\mu)$ and v^μ is a fixed point of L^*. If $v^\mu \neq v^{\mu'}$, then $v_i^\mu > v_i^{\mu'}$ for at least one state i. We then replace μ by μ' and continue the process. Since the set of stabilizing policies is finite, we must obtain eventually two successive stabilizing policies with equal cost vectors, thereby showing that L^* has a fixed point within V. Furthermore, the construction provides a stabilizing deterministic policy, say μ^*, such that $L^{\mu^*}(v^{\mu^*}) = v^{\mu^*} = L^*(v^{\mu^*})$.

$x_i^* = \min_{j\in A(i)}(a_{ij} + x_j^*), \quad i = 2, ..., m,$
$x_1 = 0.$
The first assumption (corresponding to H1) says: "There exists a path from every node $i = 2, ..., m$ to the destination node 1". The second assumption says (corresponding to H2): "Every cycle has positive length" (cf BT89, p318,293). It is easy to a construct "counterexample" to Bellman-Ford's algorithm when there exist non-positive cycles (Counterexample (BT89, p. 297), Fig.4.1.1: avec un état 1 absorbant, un arc allant de l'état 2 à 1 avec proba. 1 et coût 1, un cycle de cout nul (et proba. 1) entre 2 et 3). In this counterexample, the shortest distances are $v_1^* = 0, v_2^* = v_3^* = 1$ and satisfy Bellman's equation. The zero vector also satisfies Bellman's equation, and if the Bellman-Ford algorithm is started with that vector, it will make no progress towards the shortest distance vector. The problem comes from the multiplicity of solutions of the Bellman's equation. This counterexample can be seen (after straightforward adaptation) as a counterexample for the minimum cost problem when H2 is not satisfied (which entails multiple fixed points for L^*).

[13] In (de Alfaro 99), H1 is different: For all $s \in S$, $Prp(i) = \{\mu \in Pol | \mathbb{P}_i^\mu(T_{\{1\}} < \infty) = 1\} \neq \emptyset$, where Pol is the set of policies non necessarily deterministic.

[14] (BT 91) Note that under Assumption H1, the set $\{1\}$ consisting of just state 1 is a recurrent class under all $P(\mu)$, $\mu \in Pol_D$. Furthermore, μ is non-stabilizing if and only if under $P(\mu)$, there exists a recurrent class other than $\{1\}$.

Lemma 11. *Let Assumption* **H** *hold. Then:*
For all $v \in V$, $\lim_{k\to\infty} L^k(v) = v^{\mu^*}$

Proof. Let us show that $(L^*)^t(v) \to v^{\mu^*}$ for all $v \in V$. Let Δ be the vector with coordinates

$$\Delta_i = \begin{cases} 0 & \text{if } i = 1 \\ \delta & \text{if } i \neq 1, \end{cases}$$

where $\delta > 0$ is some scalar, and let v^Δ be the vector in V satisfying
$L^{\mu^*}(v^\Delta) = v^\Delta - \Delta$.
[There is a unique such vector because the equation $v^\Delta = c^{\mu^*} + \Delta + P(\mu^*)v^\Delta (= L^{\mu^*}(v^\Delta) + \Delta)$ has a unique solution within V by the analysis of Lemma 6.]

Since v^Δ is the cost vector corresponding to μ^* for c^{μ^*} replaced by $c^{\mu^*} + \Delta$, we have $v^\Delta \geq v^{\mu^*}$. Furthermore, for any $v \in V$, there exists $\Delta > 0$ such that $v \leq v^\Delta$. we have:

$$v^{\mu^*} = L^*(v^{\mu^*}) \leq L^*(v^\Delta) \leq L^{\mu^*}(v^\Delta) = v^\Delta - \Delta \leq v^\Delta.$$

Using the monotonicity of L and the previous relation, we obtain:
$v^{\mu^*} = (L^*)^t(v^{\mu^*}) \leq (L^*)^t(v^\Delta) \leq L^{t-1}(v^\Delta) \leq v^\Delta, \quad \forall t \geq 1)$
Hence, $(L^*)^t(v^\Delta)$ converges to some $\tilde{v} \in V$, and by continuity of L^*, we must have $\tilde{v} = v^{\mu^*}$. It is also seen using the fact that $v_1^{\mu^*} = 0$, that
$v^{\mu^*} - \Delta = L^*(v^{\mu^*}) - \Delta \leq L^*(v^{\mu^*} - \Delta) \leq L^*(v^{\mu^*}) = v^{\mu^*}$,
so $v^{\mu^*} - \Delta \leq \lim_{t\to\infty}(L^*)^t(v^{\mu^*} - \Delta) \leq v^{\mu^*}$. Similarly, as earlier, it follows that $\lim_{t\to\infty}(L^*)^t(v^{\mu^*} - \Delta) = v^{\mu^*}$. For any $v \in V$, we can find $\delta > 0$ such that

$$v^{\mu^*} - \Delta \leq v \leq v^\Delta.$$

By monotonicity of L^*, we then have
$(L^*)^t(v^{\mu^*} - \Delta) \leq (L^*)^t(v) \leq (L^*)^t(v^\Delta), \quad \forall t \geq 1,$
and since $\lim_{t\to\infty}(L^*)^t(v^{\mu^*} - \Delta) \leq (L^*)^t(v) \leq (L^*)^t(v^\Delta) = v^{\mu^*}$, it follows that $\lim_{t\to\infty}(L^*)^t(v) = v^{\mu^*}$.

Lemma 12. *Let Assumption* **H** *hold. Then:*
The unique fixpoint v^{μ^*} of L^* is equal to the minimum cost vector v^*.

As a recapitulation of the above lemmas, we have:

Proposition 19. *Let Assumption* **H** *hold. Then:*
(a1) L^* has at most one fixpoint within the subspace $V = \{v \in \mathbb{R}^m | v_1 = 0\}$.
(a2) There exists a stabilizing deterministic policy μ^* such that $v = L^{\mu^*}(v)$ iff $v = L^*(v)$.
This shows thats L^* has a unique fixpoint, which is equal to v^{μ^*}.
(a3) For all $v \in V$, $\lim_{k\to\infty} L^k(v) = v^{\mu^*}$
(a4) The unique fixpoint v^{μ^*} of L^* is equal to the minimum cost vector v^*.
(c) A deterministic policy μ is optimal if and only if $L^\mu(v^*) = L^*(v^*)$.
Furthermore, there exists an optimal (deterministic) stabilizing policy.

As a summary, we have, under Assumption **H**:
- The minimum cost vector v^* is the unique fixed point of L^* within the space V.
- For every $v \in V = \{w \in \mathbb{R}^m \mid w_1 = 0\}$, $\lim_{t\to\infty}(L^*)^t v = v^*$.
- There exists a deterministic policy μ which is optimal, and can be obtained by "policy iteration" (see below).

Policy Iteration The construction used in the proof of Lemma 10 to show that L^* has a fixed point constitutes an algorithm, known as *policy iteration*, for obtaining an optimal stabilizing policy starting with an arbitrary stabilizing policy. In the typical iteration of this algorithm, given a stabilizing deterministic policy μ and the corresponding cost vector v^μ, one obtains a new stabilizing deterministic policy μ' satisfying the equation $L^{\mu'}(v^\mu) = L^*(v^\mu)$, or, equivalently,
$\mu'(i) = \arg\min_{a\in A(i)}[c_i(a) + \sum_{j=2}^m p_{ij}(a)v_j^\mu], \quad i = 2, 3, ..., m.$
The new policy is strictly better if the current policy is nonoptimal; indeed, it was shown by Eq. (3.12) and the discussion following that equation $v^{\mu'} \leq v^\mu$, with strict inequality $v_i^{\mu'} < v_i^\mu$ for at least one state i, if the policy μ is nonoptimal. Because the number of deterministic policies is finite, it follows that this policy iteration algorithm terminates after a finite number of iterations with an optimal stabilizing (deterministic) policy. Note that each iteration involves a "policy evaluation" step, whereby, given $\mu \in Pol_D$, we obtain the corresponding cost vector v^μ by solving the system of equations $v^\mu = c^\mu + P(\mu)v^\mu$ subject to the constraint $v_1^\mu = 0$. This step can be very time-consuming when the number of states is large.

Example 20. Considérons le MDP de la Figure 3.

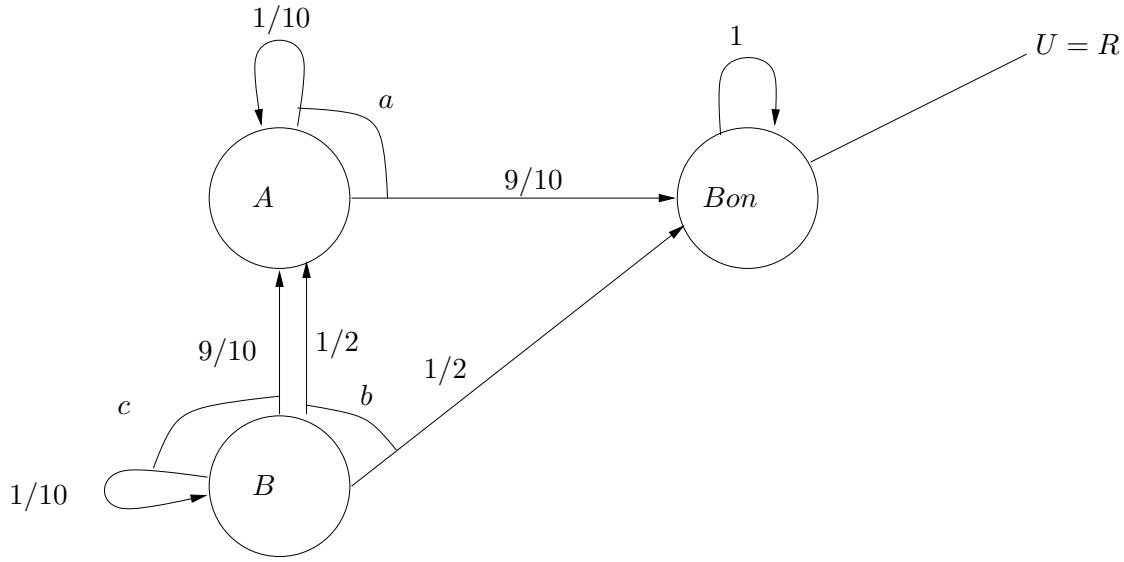

Fig. 3. Un exemple pour le problème du pire temps moyen

Calculons les pires temps moyen d'absorption v_A^*, v_B^* (avec $c_i(\alpha) = -1$ pour $i = A, B$ et $\alpha = a, b, c, d$). Tout d'abord,calculons le vecteur v_i^*, en trouvant le point-fixe de l'opérateur de Bellman. Partons de l'état initial $v_A = v_B = 0$. On obtient le système:

$v_A^{n+1} = p_{AA}(a)v_A^n - 1 = \frac{1}{10}v_A^n - 1$

$v_B^{n+1} = \min_{\alpha\in\{b,c\}}[p_{BB}(\alpha)v_B^n + p_{BA}(\alpha)v_A^n] - 1 = \min[\frac{1}{2}v_A^n, \frac{9}{10}v_A^n + \frac{1}{10}v_B^n] - 1.$

En partant du vecteur $v_A^0, v_B^0 = (0, 0)$, on trouve:

$v_A^1 = -1, v_B^1 = -1$

$v_A^2 = -\frac{11}{10}, v_B^2 = \min\{-\frac{1}{2}, -1\} - 1 = -2$

$v_A^3 = -\frac{111}{100}, v_B^3 = \min\{-\frac{11}{20}, -\frac{119}{100}\} - 1 = -\frac{219}{100}$

...

A chaque itération de v_A^n, il n' y a qu'un seul choix d'action possible, *viz.* a. A chaque itération de v_B^n, le minimum de l'opérateur de Bellman est obtenu pour l'action $\alpha = c$. (Cela pourrait se démontrer par récurrence.) Cela signifie que la politique (déterministe) optimale consiste à sélectionner l'action a en A et l'action c en B. La suite (v_A^n, v_B^n) converge vers $(-\frac{10}{9}, -\frac{20}{9})$ qui correspond à (v_A^*, v_B^*), et donne (l'opposé du) pire temps d'absorption en partant de A et B.

Example 21. Policy Iteration

(a5) Exercice 3.5: Show that, under **H**, if $x \in \mathbb{R}^n$ is such that $L^*(x) \geq x$, then $v^* \geq x$. Use this fact to show that v^* solves the linear program

maximize $\beta' x$

subject to $c(i,a) + \sum_{j=1}^{n} p_{ij}(a)\ x_j \geq x_i, \quad i = 1, ..., n,\ a \in A(i)$,

where β is a nonzero vector with nonnegative coordinates.

Here is a slightly different version of Prop. 19.

Proposition 20. *Let Assumption* **H** *hold. Then:*

- *The functional L^* admits exactly one fixpoint $v^\bullet$ such that $v^\bullet = L^* v^\bullet$.*
- *The fixpoint $v^\bullet$ of L is the single optimal solution of the following linear programming problem on the set $\{v_i\}_{i \neq 1}$ of variables: Maximize $\sum_{i \neq 1} v_i$ subject to*

$$v_i \leq \sum_{j \in S} p_{ij}(a)\ v_j + c(i,a) \qquad i \in S \backslash \{1\}.$$

- *Consider any Markovian policy η that selects at every $i \neq 1$ only actions a that minimize $\sum_{j \in S} pij(a)\ v_j + c(i,a)$. Then, policy η is stabilizing, and we have: $v_i^\eta = v_i^* = v_i^\bullet$ for all $i \neq 1$.*

NB: To prove H1: find $f : S \mapsto \mathbb{N}$ such that $\forall i \neq 1 \exists j :\ p_{ij} > 0 \wedge (j = 1 \vee f(j) < f(i))$

Recapitulation:

- Politique optimale et temps (coût) moyen associé. Méthode (BT):

(a) minimum cost vector is unique solution of Bellman's equation.

(b) successive approximation methods convenenient to the minimum cost vector for an arbitrary starting vector.

(c) the policy iteration algorithm yields an optimal *deterministic* (Markovian) policy.

NB: Under assumption 1, the set $\{1\}$ consisting of just state 1 is a recurrent class under all $P(\mu)$, $\mu \in Pol_D$. Furthermore, μ is non-stabilizing if and only if under $P(\mu)$, there exists a recurrent class other than $\{1\}$.

7 Time to Absorption: A Martingale View (Brémaud, p.178+185-191)

7.1 Stopping Time

Definition 28. *: Stopping Times (p.83).*
A stopping time *with respect to a stochastic process $\{X_n\}_{n\geq 0}$ is, by definition, a random variable τ taking its values in $\mathbb{N}\cup\{\infty\}$ and such that for all integers $m\geq 0$, the event $\{\tau=m\}$ can be expressed in terms of $X_0, X_1, ..., X_m$.*

The latter property is symbolized by the notation $\{\tau=m\}\in X_0^m$. When the state space is countable, this means that $1_{\{\tau=m\}}=\psi_m(X_0,...,X_m)$, for some function ψ_m with values in $\{0,1\}$. For a given stopping time τ, one can decide whether $\tau=m$ just by observing $X_0, X_1, ..., X_m$. This is why stopping times are said to be *nonanticipative.*

It is decidable if $\tau=m$ or not with knowledge of past and present only (no future).

A stopping time corresponds to a strategy for determining when to stop a sequence based only on the outcomes seen so far. For example, the first time the gambler wins five games in a raw is a stopping time, since this can be determined by looking at the outcomes of the games played. similarly, the first time the gambler has won at leat a hundred dollars is also a stopping time. Letting T be the *last* time the gambler wins five games in a row, however, would not be a stopping time, since determining whether $T=n$ cannot be done without knowing $X_{n+1}, X_{n+2}, ...$

Example 22. Successive Returns (p.84).
Let $\tau_1=T_i, \tau_2, ...$ be the successive return times to state i. If there are only r returns to state i, let $\tau_{r+1}=\tau_{r+2}=\cdots=\infty$. These random times are stopping times with respect to $\{X_n\}_{n\geq 0}$, since for any $m\geq 1$,
$\{\tau_k=m\}=\{\sum_{n=1}^{m}1_{\{X_n=i\}}=k, X_m=i\}$ is indeed expressible in terms of $X_0,...,X_m$.

Example 23. (*Counterexample*: p.84). The random time $\tau=\inf\{n\geq 0; X_{n+1}=i\}$ where $\tau=\infty$ if $X_{n+1}\neq i$ for all $n\geq 0$, is anticipative because $\tau=m\}=\{X_1\neq i,...,X_m\neq i, X_{m+1}\neq i\}$ for all $m\geq 0$. Knowledge of this random time provides information about the value of the process just after it. It is em not a stopping time.

7.2 Martingales

Definition 29. *: Martingales (Brémaud, p.179)*
A real-valued stochastic process $\{Y_n\}$ such that for each $n\geq 0$,
(i) Y_n is a function of $X_0,...,X_n$, and
(ii) $(E[|Y_n|]<\infty\vee Y_n\geq 0)$ is called a martingale *(resp* submartingale, supermartingale*) with respect to $\{X_n\}$ if, moreover, $E[Y_{n+1}|X_0,X_1,\ldots,X_n]=Y_n$ (resp $\geq Y_n$, $\leq Y_n$)* (25)

In the above definition, $\{X_n\}$ can be any stochastic process[15], not necessarily a Markov chain

A good way to think of martingales is as a model of "fair game" Here X_n represents what happens in the game at time n, and Y_n represents the player's wealth at time n (after the n-th play of the game) So, at time n, Y_n represents the current wealth, and $E[Y_{n+1}|X1,...,X_n]$ represents the expected wealth after one more play To say that these two are equal means that the next play of the game is "fair". One of the main results of martingale theory, which is the key to the recurrence (resp transience) criteria, is the probabilistic counterpart of the convergence of a bounded nondecreasing sequence of real numbers to a finite limit.

Theorem 17. *: Martingale Convergence (Brémaud, p.185)*
Let $\{Y_n\}$ be either a nonnegative supermartingale, or a bounded submartingale, with respect to $\{X_n\}$ Then almost surely, $lim_{n\uparrow\infty}\ Y_n$ exists[16] *and is finite.*

[15] A stochastic process with state space S is a sequence $\{X_n\}_{n\geq 0}$ of random variables with values in S
[16] i.e.: $\{\omega\in\Omega: Y_n(\omega)\to Y(\omega)\}$ as $n\to\infty\}$ is an event whose probability=1

Theorem 18. *Optional Sampling (Brémaud, p.185)*
Let $\{M_n\}$ be a martingale with respect to some process $\{X_n\}$, and let T be a stopping time of $\{X_n\}$ Suppose that at least one of the following conditions holds:
(α) P-as, $T \leq n_0$ for some $n_0 \geq 0$, or
(β) P-as, $T < \infty \wedge (n < T \Rightarrow |M_n| < K < \infty)$,
Then $E[M_T] = E[M_0]$.[17]

7.3 Time to Absorption with Martingales

Example 24. (GZbis, p.118+401) Let $\{S_n : n \geq 0\}$ be a simple symmetric random walk with $0 < S_0 < N$ and with absorbing barriers at 0 and N. Let us compute the mean time until absorption using the optional stopping theorem.

Let T be the time until absorption, and note that $\{S_n\}$ is bounded, and therefore uniformly integrable. Also $\mathbb{P}(T < \infty) = 1$ since T is no larger than the waiting time for N consecutive steps in the same direction.

Secondly, $\{S_n^2 - n : n \geq 0\}$ is a martingale, since
$E[S_{n+1}^2 - (n+1) \mid S_n] =$
$E[(S_n + X_{n+1})^2 - (n+1) \mid S_n] =$
$E[S_n^2 \mid S_n] + 2E[S_n \times X_{n+1} \mid S_n] + E[X_{n+1}^2] - (n+1) =$
$S_n^2 + 2S_n E[X_{n+1}] + E[X_{n+1}^2] - (n+1) =$
$S_n^2 + (\frac{1}{2} + \frac{1}{2}) - (n+1) =$
$S_n^2 - n,$
where X_n denotes the size of the nth jump (i.e., ± 1 with half probability).

The optional stopping theorem (if it may be applied) gives that

$$E[S_0^2] = E[S_T^2 - T] = N^2 \mathbb{P}(S_T = N) - E[T],$$

and hence $E[T] = NE[S_0] - E[S_0^2]$ as required (using the fact that $\mathbb{P}(S_T = N) = E[S_0]/N$, which can be proved by using the optional stopping theorem to $\{S_n : n \geq 0\}$, viewed as a martingale w.r.t. itself).[18]

It remains to check the conditions of the optional stopping theorem. Certainly $\mathbb{P}(T < \infty) = 1$, and in addition, if $n < T$, we have: $|S_n^2 - n| < K < \infty$ for $K = N^2 - N$.

Proposition 21. *Suppose that $D = (D_t)_{t=0}^{\infty}$ is a nonnegative stochastic process on $\{0, 1, \cdots, B\}$ such that*
$E[D_{t+1}|D_t] \leq \beta D_t$ (with $0 < \beta < 1$). Then if τ is the first time that $D_t = 0$, we have: $E[\tau] \leq B/(1-\beta)$.

Proof. The process $Z(t) = (B - D_t) - (1-\beta)\min(t, \tau)$ is a submartingale since $E[Z(t+1)] - Z(t) = D_t - E[D_{t+1}] - (1-\beta) \geq (1-\beta)(D_t - 1) \geq 0$. Moreover, τ is a stopping time for Z, and the differences $Z(t+1) - Z(t)$ are bounded. The Optional Stopping/Sampling theorem for submartingales (see e.g., [**?**]) then applies, which yields: $E[Z_\tau] \geq Z_0$, i.e: $B - (1-\beta)E[\tau] \geq 0$. Hence: $E[\tau] \leq \frac{B}{1-\beta}$.

Proposition 22. *Suppose that $D = (D_t)_{t=0}^{\infty}$ is a nonnegative stochastic process on $\{0, 1, \cdots, B\}$ such that*
$E[D_{t+1}|D_t] \leq D_t$. Furthermore suppose that $\mathbb{P}(D_{t+1} \neq D_t) \geq \alpha$ (with $\alpha > 0$) when $D_t > 0$. Then if τ is the first time that $D_t = 0$, we have: $E[\tau] \leq B^2/\alpha$.

Proof. (This proof follows that given in [**?**]; cf [**?**]):
The process $Z(t) = (B - D_t)^2 - \alpha t$ is a submartingale since $E[(D_{t+1} - D_t)^2] \geq \alpha$. (We have: $E[(D_{t+1} - D_t)^2] \geq \mathbb{P}((D_{t+1} - D_t)^2 \geq 1) = \mathbb{P}(D_{t+1} \neq D_t) \geq \alpha$.) Moreover, τ is a stopping time for Z, and the differences $Z(t+1) - Z(t)$ are bounded. The Optional Stopping theorem for submartingales then applies: $E[Z_\tau] = B^2 - \alpha E[\tau] \geq Z_0 = (B - D_0)^2$. Hence:
$E[\tau] \leq \frac{1}{\alpha}(B^2 - (B - D_0)^2) \leq \frac{B^2}{\alpha}$.

These theorems will allow us to analyse the behavior of Markov chains as random walks with a drift.

---

[17] M_T is obtained by (repeatedly) drawing an experience with (M_n) until $n = T$ is observed. This is a restriction of the martingale to the space $n = T$.

[18] We have indeed $E[S_{n+1} \mid S_n] = S_n$, hence, by the optional stopping theorem, $E[S_0] = E[S_T] = 0 \times \mathbb{P}(S_T = 0) + N \times \mathbb{P}(S_T = N) = N\mathbb{P}(S_T = N)$.

8 Time To Stabilization: A Coupling View

8.1 Coupling

(Sinclair) Coupling is an elementary probabilistic method for bounding the mixing time of a Markov chain $\mathcal{M}$ by relating it to the stopping time of an associated stochastic process.

(Brémaud,p.128) Coupling is an old idea of Doeblin (1938), revived in Markov-chain theory by the influential work of Griffeath (1975) and Pitman (1974), and brought to fame by Lindvall (1977) who gave a purely probabilistic proof of the renewal theorem. The coupling method has a wide range of applications (see the book (Lindvall,1992))

Definition 30. *A* coupling *of a Markov chain M_t with state space S is a Markov chain (X_t, Y_t) on the state pace $S \times S$ with the properties*

1. *Each of the processes (X_t) and (Y_t) is a faithful copy of M_t* [19]
2. *If $X_t = Y_t$, then $X_{t+1} = Y_{t+1}$.*

Although each of (X_t), (Y_t), viewed in isolation, behaves exactly like (M_t), they need not be independent; on the contrary, we will construct a joint distribution for the two processes in such a way that they tend to move closer together. By the second condition above, once they have met they must remain together at all future times.

We will sometimes denote such a coupling (X_t, Y_t) by $(X, Y) \mapsto (X', Y')$ (meaning that, given a pair $(X, Y) \in S \times S$, the coupling goes from (X_t, Y_t) to (X_{t+1}, Y_{t+1}) with $\mathbb{P}(X_{t+1} = X' | (X_t, Y_t) = (X, Y)) = p_{XX'}$ and $\mathbb{P}(Y_{t+1} = Y' | (X_t, Y_t) = (X, Y)) = p_{YY'}$).

Definition 31. *Given a coupling (X_t, Y_t), the* (expected) coupling time *is:*
$\mathbf{T} = \max_{i,j \in S} E[T_{i,j}]$,
where $T_{i,j} = min\{t : X_t = Y_t \mid X_0 = i, Y_0 = j\}$.

Note that $T_{i,j}$ is the (random) time until the processes meet.

Using Prop. 21 and 22, we can bound the coupling time via contractive distances as follows.

Theorem 19. *Given a Markov chain (M_t), suppose there exist a coupling $(X, Y) \mapsto (X', Y')$, a function δ on $S \times S$ which takes values in $\{0, 1, ..., B\}$, and a positive constant $\beta < 1$ such that, for all $(X, Y) \in S \times S$:*

- *$\delta(X, Y) = 0$ iff $X = Y$, and*
- *$E[\delta(X', Y') | X, Y] \leq \beta \delta(X, Y)$.* (1)

Then the coupling time satisfies: $\mathbf{T} \leq \frac{B}{1-\beta}$.

Proof. Consider two elements $i, j \in S$, and the coupling (X_t, Y_t) starting from $(X_0, Y_0) = (i, j)$. Let D_t be the process defined by $D_t = \delta(X_t, Y_t)$ for $t \geq 0$. Since $\delta(X_t, Y_t) = 0$ iff $X_t = Y_t$, the quantity $T_{i,j}$ is the time required for D_t to reach 0. Consider the coupling (X_t, Y_t) which starts from $(X_0, Y_0) = (i, j)$. Therefore by Prop. 21, we have, for all $i, j \in S$, $E[T_{i,j}] \leq B/(1 - \beta)$.

Theorem 20. *Given a Markov chain (M_t), suppose suppose there exist a coupling $(X, Y) \mapsto (X', Y')$, a function δ on $S \times S$ which takes values in $\{0, 1, ..., B\}$ and a constant $\alpha > 0$ such that, for all $(X, Y) \in S \times S$:*

- *$\delta(X, Y) = 0$ iff $X = Y$, and*
- *$E[\delta(X', Y') | X, Y] \leq \delta(X, Y) \quad \wedge \quad \mathbb{P}(\delta(X', Y') \neq \delta(X, Y)) \geq \alpha$.* (2)

Then the coupling time satisfies: $\mathbf{T} \leq \frac{B^2}{\alpha}$.

Proof. Consider two elements i, j of S and a coupling (X_t, Y_t) of initial element $(X_0, Y_0) = (i, j)$. Let $D_t = \delta(X_t, Y_t)$ for $t \geq 0$. Since $\delta(X_t, Y_t) = 0$ iff $X_t = Y_t$, the quantity $T_{i,j}$ is the time required for D_t to reach 0. Therefore by Prop. 22, we have, for all $i, j \in S$, $E[T_{i,j}] \leq B^2/\alpha$.

[19] This means (Mitzenmacher-Upfal,p.274): $\mathbb{P}(X_{t+1} = i' \mid (X_t, Y_t) = (i, j)) = \mathbb{P}(M_{t+1} = i' \mid M_t = i)$ and $\mathbb{P}(Y_{t+1} = j' \mid (X_t, Y_t) = (i, j)) = \mathbb{P}(M_{t+1} = j' \mid M_t = j)$.

Example 25. (*Coupling with contractive distance for Herman*).
Reprenons l'exemple de l'algorithme de Herman, et trouvons une distance sur les couples de configurations. Nous ne donnerons pas en détail l'ensemble des preuves ici, car le calcul du temps de convergence pour Herman est plus simple par la technique du path coupling que nous décrivons ci-après.

- *Coupling:* Le coupling que nous choisissons est extrêmement simple; nous ne forçons les choix probabilistes uniquement si la machine i a une transition probabiliste à faire à la fois dans la configuration X_t et dans la configuration Y_t (i.e., lorsque $X_t(i) = X_t(i-1)$ et $Y_t(i) = Y_t(i-1)$); dans ce cas les machines i de X_t et Y_t sont forcés de faire le même choix probabilistes, ainsi $X_{t+1}(i)$ et $Y_{t+1}(i)$ coïncident:
$$X_{t+1}(i) = Y_{t+1}(i) = \begin{cases} 0 & \text{avec probabilité } 1/2, \\ 1 & \text{avec probabilité } 1/2. \end{cases}$$
- *Fonction δ:* la fonction $\delta(X_t, Y_t)$ est assez simple elle correspond à la distance maximale entre deux machines dont les valeurs diffèrent entre X_t et Y_t. L'évolution de δ est donnée dans la figure 4. Ainsi

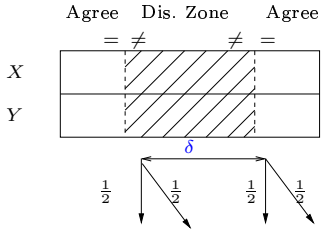

Fig. 4. L'évolution de δ sous le coupling

on obtient sans entrer dans les détails:

$$\delta \begin{cases} = \delta + 1 & \text{avec probabilité } 1/4 \\ \leq \delta - 1 & \text{avec probabilité } 1/4 \\ \leq \delta & \text{avec probabilité } 1/4 \\ = \delta & \text{avec probabilité } 1/4 \end{cases}$$
Donc, $E(\delta(X_{t+1}, Y_{t+1})|X_t, Y_t) \leq \delta(X_t, Y_t)$
Et, $\mathbb{P}(\delta(X_{t+1}, Y_{t+1}) \neq \delta(X_t, Y_t)) \geq \frac{1}{2}$
Il existe donc un coupling satisfait les propriétés désirées. Il s'ensuit, d'après le théorème 20 que le temps de coupling de Herman est donc majoré par $2N^2$.

8.2 Path coupling

As pointed out in [?], it is often cumbersome to measure the expected change in distance between two arbitrary configurations. The method of *path coupling*, introduced by Bubley and Dyer [?], simplifies the approach by showing that only pairs of configurations that are "close" need to be considered. Path coupling involves defining a coupling (X_t, Y_t) by considering a *path*, or sequence $X_t = Z_0, Z_1, ..., Z_r = Y_t$ between X_t and Y_t where the Z_i satisfy certain conditions. The following version of the path coupling method is convenient:

Lemma 13. (Dyer and Greenhill [?])
Let δ be a metric[20] *defined on $S \times S$ which takes value in $\{0, ..., B\}$.*
Let U be a subset of $S \times S$ such that, for all $(X, Y) \in S \times S$:
there exists a path $X = Z_0, Z_1, ..., Z_r = Y$ between X and Y
such that $(Z_i, Z_{i+1}) \in U$ for $0 \leq i < r$ and $\sum_{i=0}^{r-1} \delta(Z_i, Z_{i+1}) = \delta(X, Y)$. (3)
Suppose there exist a coupling $(X, Y) \mapsto (X', Y')$ for the Markov chain (M_t) on all pairs $(X, Y) \in U$, and a constant $\beta \leq 1$ such that, for all $(X, Y) \in U$:
$E[\delta(X', Y')|X, Y] \leq \beta\delta(X, Y)$. (4)
Then this coupling can be extended to a coupling for (M_t) on $S \times S$, which satisfies (4) for all $(X, Y) \in S \times S$.

[20] I.e., a function such that: $\delta(X, Y) = 0$ iff $X = Y$, and $\delta(X, Z) \leq \delta(X, Y) + \delta(Y, Z)$, for all $X, Y, Z \in S$.

Proof. Let us show that, for any $(X,Y) \in S \times S$, $E[\delta(X',Y')|X,Y] \leq \beta\delta(X,Y)$. Let $X = Z_0, Z_1, .., Z_r = Y$ be the path between X and Y, whose existence is stated by (3), i.e.: such that $(Z_i, Z_{i+1}) \in U$ for $0 \leq i < r$ and $\sum_{i=0}^{r-1} \delta(Z_i, Z_{i+1}) = \delta(X,Y)$. We have

$$\begin{aligned} E[\delta(X',Y')] &\leq \textstyle\sum_{i=0}^{r-1} E[\delta(Z'_i, Z'_{i+1})] && \text{(by triangular inequality and linearity of expectation)} \\ &\leq \textstyle\sum_{i=0}^{r-1} \beta\delta(Z_i, Z_{i+1}) && \text{(by (4), since each } (Z_i, Z_{i+1}) \text{ belongs to } U) \\ &= \beta\delta(X,Y). \end{aligned}$$

Two configurations X and Y are said to be *adjacent* if $(X,Y) \in U$.

NB: One cannot *a priori* go from one configuration to an adjacent configuration through a single-step transition of the Markov chain.

The advantage of this lemma is that it allows us to check the crucial property (4) on the set U of adjacent pairs only, rather than on the entire space $S \times S$. In particular, it allows us to find ma coupling and a contractive distance in a simpler way in order to bound the coupling time.

Example 26. (*Path coupling for Herman*). Let us come back to Herman's algorithm (see Example 10).

Lemma 14. *For Herman's algorithm and N odd, there exist a subset U of $S \times S$, a metric δ on $S \times S$ taking value in $\{0, ..., N\}$ and satisfying condition* (3)*, and a coupling such that:*

- $\forall (X_t, Y_t) \in U \quad E[\delta(X_{t+1}, Y_{t+1})|X_t, Y_t] \leq \delta(X_t, Y_t)$, *and*
- $\forall (X_t, Y_t) \in S \times S$ (*with* $X_t \neq Y_t$) : $\quad \mathbb{P}(\delta(X_{t+1}, Y_{t+1}) \neq \delta(X_t, Y_t)) \geq 1/2$.

Proof. – *Subset U and metric δ.* We define δ as the Hamming distance: $\delta(X_t, Y_t)$ is the number of positions at which X_t and Y_t differ. The pair (X_t, Y_t) belongs to U iff $\delta(X_t, Y_t) = 1$. It is immediate to check condition (3) of Lemma 13.

– *Coupling.* The coupling is defined in order to force X_t and Y_t to do the same probabilistic choice, when they both have to perform a random action. In other words, for all i $(1 \leq i \leq N)$:
If $X_t(i) = X_t(i-1)$ and $Y_t(i) = Y_t(i-1)$ then

$$X_{t+1}(i) = Y_{t+1}(i) = \begin{cases} 0 & \text{with probability } 1/2, \\ 1 & \text{with probability } 1/2. \end{cases}$$

– *Proof of $E[\delta(X_{t+1}, Y_{t+1})] = \delta(X_t, Y_t)$ on U.* Consider a pair $(X_t, Y_t) \in U$, and let ℓ be the position of disagreement between X_t and Y_t. In order to fix the ideas consider the following vector

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \nu_1 \; \nu_2 \; \cdots \; \nu_{\ell-2} \; 0 \; \mathbf{0} \; 0 \; \nu_{\ell+2} \; \cdots \; \nu_N \\ \nu_1 \; \nu_2 \; \cdots \; \nu_{\ell-2} \; 0 \; \mathbf{1} \; 0 \; \nu_{\ell+2} \; \cdots \; \nu_N \end{pmatrix}$$

where all the ν_i are in $\{0,1\}$, the figures in bold font correspond to positions ℓ. (The other cases are similar.) After one step, the state of all the machines at position $1, \cdots, N$ are updated. We have:

$$\begin{pmatrix} X_{t+1} \\ Y_{t+1} \end{pmatrix} = \begin{pmatrix} \nu'_1 \; \nu'_2 \; \cdots \; \nu'_{\ell-2} \; \nu'_{\ell-1} \; ? \; ? \; \nu'_{\ell+2} \; \cdots \; \nu'_N \\ \nu'_1 \; \nu'_2 \; \cdots \; \nu'_{\ell-2} \; \nu'_{\ell-1} \; \mathbf{0} \; 1 \; \nu'_{\ell+2} \; \cdots \; \nu'_N \end{pmatrix}$$

where '?' means "0 with prob. 1/2 and 1 with prob. 1/2". Note that, for $1 \leq i \leq \ell-1$ and $\ell+2 \leq i \leq N$, $X_{t+1}(i) = Y_{t+1}(i) = \nu'_i$ thanks to our coupling. So X_{t+1} and Y_{t+1} coincide everywhere except, perhaps, at positions ℓ or $\ell+1$. We have:

$$\delta(X_{t+1}, Y_{t+1}) = \begin{cases} 0 & \text{with probability } 1/4, \\ 1 & \text{with probability } 1/2, \\ 2 & \text{with probability } 1/4. \end{cases}$$

Hence $E[\delta(X_{t+1}, Y_{t+1})|X_t, Y_t] = \delta(X_t, Y_t)$, for all $(X_t, Y_t) \in U$.

– *Proof of $\mathbb{P}(\delta(X_{t+1}, Y_{t+1}) \neq \delta(X_t, Y_t)) \geq 1/2$.* Let us denote by q the number of disagreeing tokens (a disagreeing token is a position i such that $X_t(i-1) = X_t(i) \neq Y_t(i-1) = Y_t(i)$) and by p the number of zones of contiguous disagreeing positions. Let us identify the three sources of possible evolution of the set of disagreeing positions after a step (see Figure **??**):

1. Thanks to the coupling, each disagreeing token $X_t(i-1) = X_t(i) \neq Y_t(i-1) = Y_t(i)$ evolves in a new agreeing position $X_{t+1}(i) = Y_{t+1}(i)$ with probability 1.
2. Each first position in a disagreeing zone, say i, such that $X_t(i-1) = Y_t(i-1)$ and $X_t(i) \neq Y_t(i)$ can evolve in an agreeing position with probability 1/2. We denote by r the number of such i $(0 \leq r \leq p)$.

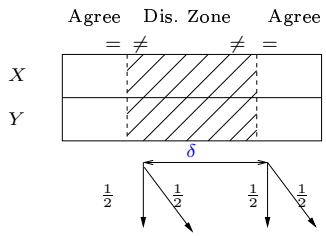

Fig. 5. Evolution of a disagreeing zone

3. Each first position in agreeing zone, say i, such that $X_t(i-1) \neq Y_t(i-1)$ and $X_t(i) = Y_t(i)$ can evolve in a disagreeing position with probability $1/2$. We denote by s the number of such i $(0 \leq r \leq p)$.

Cases 2 and 3 are depicted on Figure 5.

We have: $\delta(X_{t+1}, Y_{t+1}) = \delta(X_t, Y_t) - q - r + s$. Therefore the event $\delta(X_{t+1}, Y_{t+1}) = \delta(X_t, Y_t)$ corresponds to all the cases where $q + r = s$. If $q > p$, such an event can never occur (probability 0). Otherwise, its probability is $\frac{1}{4^p} \sum_{r=0}^{p-q} \binom{p}{r}\binom{p}{q+r} = \frac{1}{4^p} \sum_{r=0}^{p-q} \binom{p}{r}\binom{p}{p-q-r} \leq$
$\frac{1}{4^p}\binom{2p}{p-q} \leq \frac{1}{2^{2p}}\binom{2p}{p}$ (by Vandermonde's convolution [?])
$\leq \frac{1}{2}$ (by induction on p).
Hence $\mathbb{P}(\delta(X_{t+1}, Y_{t+1}) \neq \delta(X_t, Y_t)) \geq \frac{1}{2}$.

By Lemma 13, this coupling on U can be extended to a coupling on $S \times S$ with the same contractive properties. Hence, by Theorem 20, we have:

Theorem 21. *The coupling time of Herman is bounded above by* $2N^2$.

8.3 Stabilization Revisited

We show hereafter that the coupling time gives an upper bound on the stabilization time. More precisely, we assume that we are given a Markov chain $(M_t)_{t=0}^{\infty}$ and a recurrent set $\mathcal{L}$, and we consider the problem of proving the stabilization property of $(M_t)_{t=0}^{\infty}$ w.r.t. $\mathcal{L}$. We focus on a Markov chain with a unique recurrent set. This corresponds to the notion of "self-stabilizing" algorithm, as originally defined by Dijkstra in the deterministic framework [?], where all the legal configurations are strongly connected. More precisely, in the non-probabilistic context, we say that, given a set $\mathcal{L}$ of legal configurations, (M_t) is *stabilizing* w.r.t. $\mathcal{L}$ if, starting from any initial configuration, the system is guaranteed to reach a configuration of $\mathcal{L}$ within a finite number of transitions (see, e.g., [?]). For example, in mutual exclusion problems, a legal configuration is a configuration with a single token, which expresses the fact that only one machine can enjoy the resource. In the probabilistic context of Markov chains, the convergence property has to be guaranteed with probability 1. Formally:

Definition 32. *Given a closed set $\mathcal{L}$ of configurations, (M_t) is* (self-)stabilizing *w.r.t. $\mathcal{L}$, if (M_t) converges towards $\mathcal{L}$ (with probability 1), whatever the initial configuration is, i.e:*
$\forall i \in S \ \mathbb{P}_i(T_{\mathcal{L}} < \infty) = 1$ [21]

Given a Markov chain $(M_t)_{t=0}^{\infty}$ on a finite set S, and a closed set $\mathcal{L}$ of configurations, we are interested in methods for proving the stabilization property of (M_t) w.r.t. $\mathcal{L}$. We are also interested in evaluating the rate of convergence of (M_t) to $\mathcal{L}$. The expected stabilization time is the standard rate of convergence used in the self-stabilization community (see, e.g., [?], p. 118). It is the expected time for (M_t) to reach $\mathcal{L}$, starting from the "worst" configuration, i.e.:

Definition 33. *Given a Markov chain (M_t) and a set $\mathcal{L}$ of configurations, the* expected stabilization time *of $\mathcal{L}$ (or more simply the* stabilization time*) is:*

$$\mathbf{A}_{\mathcal{L}} = max_{i \in S} \ E_i[T_{\mathcal{L}}],$$

where $E[.]$ denotes expectation and $T_{\mathcal{L}} = min\{t : \ X_t \in \mathcal{L}\}$.

[21] Recall that by Markov's theorem (see, e.g., [?]) (M_t) is stabilizing iff
$\forall i \in S \ \mathbb{P}_i(T_{\mathcal{L}} > 0) = 1.$ $(\diamondsuit)$.

Let us recall that a finite Markov chain always converges towards the set of "recurrent" configurations, that is the union of the recurrent sets (see, e.g., (Kemeny-Snell)). [22] We have:

Proposition 23. *A finite Markov chain (M_t) is stabilizing w.r.t. the union of recurrent sets. Moreover: If (M_t) is stabilizing w.r.t. a subset $\mathcal{L}$ of S and $\mathcal{L}$ is recurrent, then $\mathcal{L}$ is the unique recurrent set.*

We assume that we are given a Markov chain (M_t) and a recurrent set $\mathcal{L}$, and we focus on the problem of proving the stabilization property of (M_t) w.r.t. $\mathcal{L}$. The following property will be useful.

Proposition 24. *Given a closed set $\mathcal{L}$, if $\mathbf{A}_{\mathcal{L}}$ is finite, then (M_t) is stabilizing w.r.t. $\mathcal{L}$.*

Proof. By contraposition. Suppose that (M_t) is not stabilizing. Then, from $(\Diamond)$, we know that there exists $i \in S$ such that

$$\mathbb{P}(M_t \in \mathcal{L} \mid M_0 = i) = 0 \text{ for all } t \geq 0.$$

So $A_{i\mathcal{L}}$ takes always an infinite value. Therefore $E[A_{i\mathcal{L}}]$ is infinite, and so is $\mathbf{A}_{\mathcal{L}}$.

In case $\mathcal{L}$ is not only closed, but recurrent (i.e., closed and irreducible), we have, using Prop. 23:

Proposition 25. *Given a recurrent set $\mathcal{L}$, if $\mathbf{A}_{\mathcal{L}}$ is finite, then (M_t) is stabilizing w.r.t. $\mathcal{L}$, and $\mathcal{L}$ is the unique recurrent set.*

In the following, we assume that we are given a Markov chain (M_t) and a recurrent set $\mathcal{L}$, and we focus on the problem of proving the stabilization property of (M_t) w.r.t. $\mathcal{L}$.

Theorem 22. *Given a Markov chain (X_t) and a recurrent set $\mathcal{L}$, if there exists a coupling of finite expected time $\mathbf{T}$, then:*

1. *The expected stabilization time $\mathbf{A}_{\mathcal{L}}$ satisfies: $\mathbf{A}_{\mathcal{L}} \leq \mathbf{T}$.*
2. *$\mathcal{L}$ is the unique recurrent set, and (X_t) is stabilizing w.r.t. $\mathcal{L}$.*

Proof. Let us suppose that there exists a coupling of finite expected $\mathbf{T}$, and let us show statements 1 and 2.

1. Recall that: $A_{i\mathcal{L}} = min\{t : \ X_t \in \mathcal{L} \mid X^0 = i\}$, and $T_{i,j} = min\{t : \ X_t = Y_t \mid X^0 = i, Y^0 = j\}$. Suppose now that $y \in \mathcal{L}$. Then $Y_t \in \mathcal{L}$ since $\mathcal{L}$ is closed. Hence: $A_{i\mathcal{L}} \leq T_{i,j}$ for all $i \in S, j \in \mathcal{L}$. And by taking the expectations, then the maxima of the two sides: $\mathbf{A}_{\mathcal{L}} \leq \mathbf{T}$.
2. Uniqueness of $\mathcal{L}$ and stabilization of (X_t) follow from the finiteness of $\mathbf{A}_{\mathcal{L}}$ (statement 1) by Prop. 25.

This theorem gives a method for showing stabilization via coupling.

8.4 Stabilization Time Using Coupling

By Theorem 22, finding an upper bound on the time of coupling T allows us at once to prove the stabilization property and to obtain an upper bound on the stabilization time. We give hereafter two sufficient conditions for bounding the coupling time. In each case, this provides us additionally with an upper bound for the stabilization time.

Theorem 23. *Given a Markov chain (M_t) and a recurrent set $\mathcal{L}$, suppose there exist a coupling $(X, Y) \mapsto (X', Y')$, a function δ on $S \times S$ which takes values in $\{0, 1, ..., B\}$ and a constant $\beta < 1$ such that, for all $(X, Y) \in S \times S$:*

- *$\delta(X, Y) = 0$ iff $X = Y$, and*
- *$E[\delta(X', Y')|X, Y] \leq \beta\delta(X, Y)$.* (1)

Then $\mathcal{L}$ is the unique recurrent set and (M_t) is stabilizing w.r.t. $\mathcal{L}$. Furthermore, the expected stabilization time satisfies: $\mathbf{A}_{\mathcal{L}} \leq \frac{B}{1-\beta}$.

Proof. Let us consider an integer-valued function δ satisfying the assumptions of Theorem 23, and let us show $\mathbf{A}_{\mathcal{L}} \leq \frac{B}{1-\beta}$. (The facts that $\mathcal{L}$ is the unique recurrent set, and (M_t) is stabilizing follow from statement 1, by Prop. 25.) By Theorem 19, we have: for all $i, j \in S$, $E[T_{i,j}] \leq B/(1-\beta)$. Now, from Theorem 22 (statement 1), we infer: $\mathbf{A}_{\mathcal{L}} \leq max_{i,j}\ E[T_{i,j}] \leq B/(1-\beta)$.

[22] Given a closed set $\mathcal{L}$, (M_t) is stabilizing w.r.t. $\mathcal{L}$ iff $Rec \subseteq \mathcal{L}$, where Rec is the set of recurrent states of (M_t).

A similar theorem exists even when $\beta = 1$, i.e.: $E[\delta(X_{t+1}, Y_{t+1})|X_t, Y_t] \leq \delta(X_t, Y_t)$, provided that the probability of $\delta(X_{t+1}, Y_{t+1}) \neq \delta(X_t, Y_t)$ can be bounded below.

Theorem 24. *Given a Markov chain (M_t) and a< recurrent set $\mathcal{L}$, suppose there exist a coupling $(X, Y) \mapsto (X', Y')$, a function δ on $S \times S$ which takes values in $\{0, 1, ..., B\}$ and a positive constant α such that, for all $(X, Y) \in S \times S$:*

- *$\delta(X, Y) = 0$ iff $X = Y$, and*
- *$E[\delta(X', Y')|X, Y] \leq \delta(X, Y) \quad \wedge \quad \mathbb{P}(\delta(X', Y') \neq \delta(X, Y)) \geq \alpha.$* (2)

Then $\mathcal{L}$ is the unique recurrent set and (M_t) is stabilizing w.r.t. $\mathcal{L}$. Furthermore, the expected stabilization time satisfies: $\mathbf{A}_{\mathcal{L}} \leq B^2/\alpha$.

The proof of Theorem 24 is analogous to that of Theorem 23, but relies on Prop. 22.

Proof. Let us consider an integer-valued function δ satisfying the assumptions of Theorem 24, and let us show $\mathbf{A}_{\mathcal{L}} \leq B^2/\alpha$. (The facts that $\mathcal{L}$ is the unique recurrent set, and (M_t) is stabilizing follow from statement 1, by Prop. 25.) From Theorem 20, we have: for all $i, j \in S$, $E[T_{i,j}] \leq B^2/\alpha$. Now, from Theorem 22, we infer: $\mathbf{A}_{\mathcal{L}} \leq max_{i,j}\ E[T_{i,j}] \leq B^2/\alpha$.

Therefore finding a coupling (X_t, Y_t) and a function δ such that (1) (resp. (2)) holds allows us to prove that (M_t) is stabilizing, and gives us an upper bound on the expected stabilization time. Thus, from example 26, Theorems 21 and 24, and Lemma 13, we have:

Theorem 25. *For N odd, Herman's algorithm is stabilizing w.r.t. the set $\mathcal{L}$ of configurations with a single token. Furthermore, the expected stabilization time satisfies:* $\mathbf{A}_{\mathcal{L}} \leq 2N^2$.

Note that the metric δ on $S \times S$ found here (Hamming distance) is much simpler than the decreasing function on S used by Herman, which involves the number of tokens of a configuration x together with the minimal distance between two tokens of x. The method here gives also directly an upper bound for the expected stabilization time with no need for a separate analysis as done in Herman's work [**?**].

9 Stationary Distributions (Steady-State Behavior), Mixing and Coupling

9.1 Computation of the Stationary Distribution (BT89, p.166-170)

Definition 34. *Given a Markov chain of transition matrix P, a* stationary distribution π *is a probability distribution π satisfying*

$$\pi = \pi P,$$

(where π is viewed as a row vector), i.e, for all $i \in S$: $\pi_i = \sum_{j \in S} \pi_j p_{ji}$.

The equations $\pi_i = \sum_{j \in S} \pi_j p_{ji}$ are called the *balance equation.*

Iteration of equation $\pi = \pi P$ gives $\pi = \pi P^n$ for all $n \geq 0$, and therefore (by the Chapman-Kolmogorov equation (???)), if the initial distribution $\nu = \pi$, then the distribution at the n-th step is still π, for all $n \geq 0$. Thus, if a chain is started with a stationary distribution, it keeps the same distribution forever. Thus, a stationary distribution represents a steady state or an equilibrium in the chain's behavior.

NB (Brémaud,p.76): The balance equation $\pi P = \pi$, together with the requirement that π be a probability vector, i.e., $\pi \mathbf{1} = 1$ (where $\mathbf{1}$ is a column vector with all its enries equal to 1), constitute $|S| + 1$ equations for S variables. One of the S equations in $\pi P = \pi$ is superfluous given the constraint $\pi \mathbf{1} = 1$. Indeed, summing up all equalities of $\pi P = \pi$ yields the equality $\pi P \mathbf{1} = \pi \mathbf{1}$, that is, $\pi \mathbf{1} = 1$.

Theorem 26. *A Markov chain with a single recurrence class (and possibly some transient states) has a unique stationary distribution π. Furthermore:*

- $\pi_i = 0$ *if i is a transient state.*
- $\pi_i > 0$ *if i is recurrent.*

Proof. A FAIRE ???

NB: (BT89,p.169) We now consider an algorithm "implementing" a variant of the iteration $\pi := \pi P$. The algorithm is described by

$\pi_1(t+1) = \pi_1(t)$, (8.4)

$\pi_i(t+1) = \sum_{j=1}^{n} \pi_j(t) p_{ji}, \qquad i = 2, ..., n,$ (8.5)

The initialization of the algorithm is arbitrary, provided that $\pi_1(0) \neq 0$. In order to represent the algorithm in matrix form, we partition the matrix P as shown:

$$P = \begin{pmatrix} p_1 & a \\ b & \tilde{P} \end{pmatrix}. \qquad (8.5)$$

Here, a (respectively, b) is a row (respectively, a column) vector of dimension $n-1$ and $\tilde{P}$ is the matrix of dimension $(n-1) \times (n-1)$ obtained by deleting the first row and the first column of P. Let $\tilde{\pi}(t)$ be the row vector $(\tilde{\pi_2}(t), ..., \tilde{\pi_n}(t))$. Then, Eq. (8.5) can be rewritten as

$\tilde{\pi}(t+1) = \tilde{\pi}(t)\tilde{P} + \pi_1(0)a.$ (8.6)

This iteration converges provided that $\rho(\tilde{P}) < 1$ (Prop. **??**). The following result provides conditions for this to be the case and characterizes the limit of $\pi(t)$.

9.2 Stationary Distribution As Limit of $p_{ij}(n)$ (BT 2002,p.326)

In Markov chain models, we are often interested in long-term state occupancy behavior, that is, in the n-step transition probabilities $p_{ij}(n)$ when n is very large. We have seen (???) that the $p_{ij}(n)$ may converge to steady-state values that are independent of the initial state. We wish to understand the extent to which this behavior is typical.

If there are two or more recurrent classes, it is clear that the limiting values of the $p_{ij}(n)$ must depend on the initial state (the possibility of visiting j far into the future depends on whether j is in the same class as the initial state i). We will, therefore, restrict attention to chains involving a single recurrent class, plus possibly some transient states. This is not as restrictive as it may seem, since we know that once the state enters a particular recurrent class, it will stay within that class. Thus, the asymptotic behavior of a multiclass chain can be understood in terms of the asymptotic behavior of a single-class chain.

Even for chains with a single recurrent class, the $p_{ij}(n)$ may fail to converge. To see this, consider a recurrent class with two states, 1 and 2, such that from state 1 we can only go to 2, and from 2 we can only go to 1 ($p_{12} = p_{21} = 1$). Then, starting at some state, we will be in that same state after any even number of transitions, and in the other state after any odd number of transitions. Formally,

$$p_{ii}(n) = \begin{cases} 1, & n \text{ even}, \\ 0, & n \text{ odd}. \end{cases}$$

What is happening here is that the recurrent class is periodic, and for such a class, it can be seen that the $p_{ij}(n)$ generically oscillate.

We now assert that for every state j, the probability $p_{ij}(n)$ of being a state j approaches a limiting value that is independent of the initial state i, provided we exclude the two situations discussed above (multiple recurrent classes and/or a periodic class). This limiting value, denoted by π_j, has the interpretation

$\pi_j \approx \mathbb{P}(X_n = j)$, when n is large,

and is called the *steady-state probability of* j. The following is an important theorem.

Theorem 27. *(*Steady-State Convergence*) Consider a Markov chain with a single recurrent class, which is aperiodic. Then, for each j, we have:*

$\lim_{n\to\infty} p_{ij}(n) = \pi_j$, *for all* i.

The steady-state probabilities π_j sum to 1 and form a probability distribution on the state space, called the *stationary distribution* of the chain. The reason for the qualification "stationary" is that if the initial state is chosen according to this distribution, i.e., if

$\mathbb{P}(X_0 = j) = \pi_j$, $j = 1, ..., m$,

then, using the total probability theorem, we have

$\mathbb{P}(X_1 = j) = \sum_{k=1}^m \mathbb{P}(X_0 = k)p_{kj} = \sum_{k=1}^m \pi_k p_{kj} = \pi_j$,

where the last equality follows from part (b) of the steady-state convergence theorem. Similarly, we obtain $\mathbb{P}(X_n = j) = \pi_j$, for all n and j. Thus, if the initial state is chosen according to the stationary distribution, the state at any future time will have the same distribution.

The equations

$\pi_j = \sum_{k=1}^m \pi_k p_{kj}$, $j = 1, ..., m$,

are called the *balance equations*. They are a simple consequence of part (a) of the theorem and the Chapman-Kolmogorov equation. Indeed, once the convergence of $p_{ij}(n)$ to some π_j is taken for granted, we can consider the equation,

$p_{ij}(n) = \sum_{k=1}^m p_{ik}(n-1)p_{kj}$,

take the limit of both sides as $n \to \infty$, and recover the balance equations.[23]

Together with the *normalization equation*

$\sum_{k=1}^m \pi_k = 1$,

the balance equations can be solved to obtain the π_j.

Example 27. Herman Prove that the uniform distribution on the configurations with a single token is stationary.

Example 28. A simple queue (Mitzenmacher-Upfal,p.173)

A *queue* is a line where customers wait for service We examine a model for a bounded queue where time is divided into steps of equal length at each time step, exactly one of the following occurs:

- If the queue has fewer than n customers, then with probability λ a new customer joins the queue
- If the queue is not empty, then with probability μ the head of the line is served and leaves the queue
- With the remaining probability, the queue is unchanged

If X_i is the number of customers in the queue at time t, then under the foregoing rules the X_t yield a finite-state Markov chain its transition matrix has the following nonzero entries:

$P_{i(i+1)} = \lambda$ if $i < n$;
$P_{i(i-1)} = \mu$ if $i > 0$;

$$P_{ii} = \begin{cases} 1-\lambda & \text{if } i = 0 \\ 1-\lambda-\mu & \text{if } 1 \le i \le n-1 \\ 1-\mu & \text{if } i = n \end{cases}$$

[23] According to the a famous and important theorem from linear algebra (called the Perron-Frobenius theorem), the balance equations always have a nonnegative solution, for any Markov chain. What is special about a chain that has a single recurrent class, which is aperiodic, is that given also the normalization equation, the solution is unique and is equal to the limit of the n-step transition probabilities $p_{ij}(n)$.

The Markov chain is irreducible finite and aperiodic, so it has a unique stationary distribution $\overline{\pi}$ We use $\overline{\pi} = \overline{\pi}P$ to write

$\pi_0 = (1-\lambda)\pi_0 + \mu\pi - 1,$

$\pi_i = \lambda\pi_{i-1} + (1-\lambda-\mu)\pi_i + \mu\pi_{i+1}, \quad 1 \leq i \leq n-1,$

$\pi_n = \lambda\pi_{n-1} + (1-\mu)\pi_n$

It is easy to verify that $\pi_i = \pi_0(\frac{\lambda}{\mu})^i$ is a solution to the preceding system of equations adding the requirement $\sum_{i=0}^{n}\pi_i = 1$, we have: $\sum_{i=0}^{n}\pi_i = \sum_{i=0}^{n}\pi_0(\frac{\lambda}{\mu})^i = 1$, or $\pi_0 = \frac{1}{\sum_{i=0}^{n}(\lambda/\mu)^i}$ For all $1 \leq i \leq n$, $\pi_i = \frac{(\lambda/\mu)^i}{\sum_{i=0}^{n}(\lambda/\mu)^i}$

Another way to compute the stationary probability in this case is to use cut-sets For any i, the transitions $i \rightarrow i+1$ and $i+1 \rightarrow i$ constitute a cut-set of the graph representing the Markov chain Thus, in the stationary distribution, the probability of moving from state i to $i+1$ must be equal to the probability of moving from state $i=1$ to i, or: $\lambda\pi_i = \mu\pi_{i+1}$ A simple induction now yields: $\pi_i = \pi_0(\frac{\lambda}{\mu})^i$

In the case where there is no upper limit n on the number of customers in a queue, the Markov chain has a countably infinite state space The Markov chain has a stationary distribution iff the following set of linear equations has a solution with all $\pi_i > 0$:

$\pi_0 = (1-\lambda)\pi_0 + \mu\pi - 1;$

$\pi_i = \lambda\pi_{i-1} + (1-\lambda-\mu)\pi_i + \mu\pi_{i+1}, \quad i \geq 1$

It is easy to verify that $\pi_i = \frac{(\lambda/\mu)^i}{\sum_{i=0}^{\infty}(\lambda/\mu)^i} = (\frac{\lambda}{\mu})^i(1-\frac{\lambda}{\mu})$ is a solution of this system This naturally generalizes the solution to the case where there is no upper bound n on the number of the customers in the finite system All of the π_i are greater than 0 iff $\lambda < \mu$, which corresponds to the situation when the rate at which customers arrive is lower than the rate at which they are served If $\lambda > \mu$, then the rate at which customers arrive is higher than the rate at which they depart hence there is no stationary distribution, and the queue length will become arbitrarily long In this case, each state in the Markov chain is transient The case of $\lambda = \mu$ is more subtle Again there is no stationary distribution and the queue length will become arbitrarily long, but now states are null recurrent.

9.3 Mixing

(Mitzenmacher-Upfal,p.271) In our study of discrete time Markov chains, we found that ergodic Markov chains converge to a stationary distribution. However, we did not determine how *quickly* they converge, which is important in a number of algorithmic applications, such as sampling using the Markov chain Monte Carlo technique. We will introduce the notion of coupling, a powerful method for bounding the rate of convergence of Markov chains.

(Brémaud,p.125) Consider an HMC that is irreducible and positive recurrent In particular, if its initial distribution is the stationary distribution, it keeps the same distribution at all times the chain is said to be in the *stationary regime*, or in *equilibrium*, or in *steady state* A question arises naturally: What is the long-run behavior of the chain when the initial distrribution μ is arbitrary? For instance, will it *converge to equilibrium*, and in which sense? When the HMC is reducible, another type of problem is of interest Suppose, for instance, that the set of transient states is not empty and that each remaining state is absorbing One may want to compute the probability of reaching a given absorbing state when the initial state is transient, or the probability of remaining forever in the transient state In this special case, where all the recurrent states are absorbing, the probability of leaving the transient set is exactly the property of converging We are dealing here with almost-sure convergence For an ergodic HMC, the type of convergence of interest is not almost-sure convergence but convergence in variation of the distribution at time n to the stationary distribution This type of convergence is relative to a metric structure that we proceed to define

Definition 35. *: Distance in Variation (Brémaud,p125)*
Let S be a countable space and let μ and ν be probability distributions on S The (total) distance in variation *$d_V(\mu,\nu)$ between μ and ν is defined by:*

$$d_V(\mu,\nu) = \frac{1}{2}|\mu-\nu| = \frac{1}{2}\sum_{i\in S}|\mu_i - \nu_i|.$$

The distance in variation between two random variables X and Y with values in S and respective distributions μ and ν is $d_V(\mu,\nu)$, and is denoted with a slight abuse of notation $d_V(X,Y)$.

That d_V is indeed a distance is clear.

This is just the L_1 norm, with the $1/2$ introduced so that the distance is always at most 1.

Lemma 15. *(MU,p272) The variation distance between two distributions μ and ν on S is given by*

$$d_V(\mu,\nu) = \max_{A\subseteq S} |\mu_A - \nu_A|.$$

Proof. Let $S^+ \subseteq S$ be the set of states i such that $\mu_i \geq \nu_i$, and let $S^- \subseteq S$ be the set of states i such that $\nu_i > \mu_i$.

Clearly,

$\max_{A\subseteq S} \mu_A - \nu_A = \mu_{S^+} - \nu_{S^+}$, and

$\max_{A\subseteq S} \nu_A - \mu_A = \nu_{S^-} - \mu_{S^-}$.

But since $\mu_S = \nu_S = 1$, we have

$\mu_{S^+} + \mu_{S^-} = \nu_{S^+} + \nu_{S^-} = 1$,

which implies that

$\mu_{S^+} - \nu_{S^+} = \nu_{S^-} - \mu_{S^-}$.

Hence

$\max_{A\subseteq S} |\mu_A - \nu_A| = |\mu_{S^+} - \nu_{S^+}| = |\mu_{S^-} - \nu_{S^-}|$.

Finally, since

$|\mu_{S^+} - \nu_{S^+}| + |\mu_{S^-} - \nu_{S^-}| = \sum_{i\in S} |\mu_i - \nu_i| = 2d_V(\mu,\nu)$,

we have

$\max_{A\subseteq S} |\mu_A - \nu_A| = d_V(\mu,\nu)$,

completing the proof.

NB: (Mitzenmacher-Upfal, p.272) Fig.11.1: The areas shaded by upward diagonal lines correspond to values i where $\mu_i < \nu_i$ (aka: $\mathbb{P}(X = i) < \mathbb{P}(Y = j)$); the areas shaded by downward diagonal lines correspond to values i where $\mu_i > \nu_i$ (aka: $\mathbb{P}(X = i) > \mathbb{P}(Y = j)$). The total area shaded by upward diagonal lines must equal the total area shaded by downward diagonal lines, and the variation distance equals one of these two areas.

Definition 36. *: Convergence in Variation (Brémaud,p128)*
Let $\{\mu^{(n)}\}$ and ν be probability distributions on a countable state space S If $\lim_{n\uparrow\infty}\ d_V(\mu^{(n)},\nu) = 0$, the sequence $\{\mu^{(n)}\}$ is said to converge in variation *to the probability distribution ν.*

Let $\{X_n\}$ be an S-valued stochastic process. If for some probability distribution ν on S, the distribution $\mu(X_n)$ of the random variable X_n converges in variation to ν, ie, if

$\lim_{n\uparrow\infty}\ \sum_{i\in S} |\mathbb{P}(X_n = i) - \nu_i| = 0$,

then $\{X_n\}$ is said to converge in variation *to ν.*

There is some abuse of terminology in the above definition (it is the state random variable, not the process, that converges in variation) However, such abuse turns out to be harmless and very convenient

If the process $\{X_n\}$ converges in variation to π, then
$\lim_{n\uparrow\infty}\ E[f(X_n)] = \pi(f)$ for all bounded function $f : S \to \mathbb{R}$, where $\pi(f) = \sum_{i\in S} \pi_i f(i)$
Indeed, if M is an upper bound of $|f|$, then:

$|E[f(X_n) - \pi(f)]| = |\sum_{i\in S} f(i)(\mathbb{P}(X_n = i) - \pi_i)| \leq M \sum_{i\in S} |\mathbb{P}(X_n = i) - \pi_i|$.

(Randall)
The time a Markov chains takes to converge to its stationary distribution, known as *mixing time* of the chain, is measured in terms of the variation distance between the distribution at time t and the stationary distribution. For a comparison of rates of convergence based on different measures of distances, in the case of reversible Markov chains, see (Aldous-Fill, to appear)((Lovasz-Winkler 1998).

Definition 37. *: Mixing Time.*
Let π be the stationary distribution of a Markov chain (X_t) with state space S. For $\varepsilon > 0$, the mixing time *$\tau(\varepsilon)$ is defined by:*

$\tau(\varepsilon) = \max_{i\in S} \tau_i(\varepsilon)$,

where $\tau_i(\varepsilon) = \min\{t : d_V(X_{t'},\pi) \leq \varepsilon,\ \forall t' \geq t \mid X_0 = i\}$.

We say a Markov chain is *rapidly mixing* if the mixing time is bounded above by a polynomial in n and $\log \varepsilon^{-1}$, where n is the size of each configuration in the state space.

9.4 Relating the Mixing Time to the Coupling Time

(Randall) It is well-known from probability theory that the eigenvalue gap of the transition matrix of the Markov chain provides a good bound on the mixing rate of a chain (see, e.g., (Sinclair, 1993)). However, for most algorithmic applications the size of the state space is exponentially large and we typically do not have a compact, mathematical representation of the adjacency matrix, so it is far too diffficult to determine the eigenvalues of the transition matrix. Another popular method for bounding mixing times is coupling, both because of its elegance and its simplicity. This was first introduced in computer science in the context of sampling spanning trees (Broder, 1989), and has since seen many more applications.

Observe that convergence in variation concerns only the marginal distributions of the process, not the process itself. Therefore, if there exists another process $\{X'_n\}_{n\geq 0}$ with $\mu(X_n) = \mu(X'_n)$ for all $n \geq 0$, and if there exists another process $\{X'_n\}_{n\geq 0}$ such that $\mu(X''_n) = \pi$ for all $n \geq 0$, then convergence in variation to π follows from

$\lim_{n\uparrow\infty} d_V(X'_n, X''_n) = 0.$ (1.9)

This trivial observation is useful because of the resulting freedom in the choice of $\{X'_n\}$ and $\{X''_n\}$. In particular, one can use dependent versions, and the most interesting case occurs when there exists a finite random time T such that $X'_n = X''_n$ for all $n \geq T$. It follows then, as will be proven later (see lemma below???) that:

$$d_v(X'_n, X''_n) \leq \mathbb{P}(T > n). \quad (1.10)$$

Finiteness of T is equivalent to $\lim_{n\uparrow\infty}\mathbb{P}(T > n) = 0$, and therefore (1.9) is a consequence of (1.10).

Lemma 16. *(MU,p.278-280) Given two distributions μ and ν on a state space S, Let $Z = (X, Y)$ be a random variable on $S \times S$. where X is distributed according to a distribution μ and Y is distributed according to a distribution ν. Then:*

$d_V(\mu, \nu) \leq \mathbb{P}(X \neq Y)$.

Moreover, there exists a joint distribution $Z = (X, Y)$, where X is distributed according to μ and Y is distributed according to ν, for which equality holds.

Proof. For each $s \in S$, we have

$\mathbb{P}(X = Y = i) \leq \min(\mathbb{P}(X = i), \mathbb{P}(Y = i))$.

Hence

$\mathbb{P}(X = Y) \leq \sum_{i\in S} \min(\mathbb{P}(X = i), \mathbb{P}(Y = i))$,

and therefore

$\mathbb{P}(X \neq Y) \geq 1 - \sum_{i\in S} \min(\mathbb{P}(X = i), \mathbb{P}(Y = i))$

$\sum_{i\in S}(\mathbb{P}(X = i) - \min(\mathbb{P}(X = i), \mathbb{P}(Y = i)))$.

Hence we are done if we can show

$d_V(\mu, \nu) = \sum_{i\in S}(\mathbb{P}(X = i) - \min(\mathbb{P}(X = i), \mathbb{P}(Y = i)))$. (11.2)

But $\mathbb{P}(X = i) - \min(\mathbb{P}(X = i), \mathbb{P}(Y = i)) = 0$ when $\mu_i < \nu_i$, and when $\mu_i \geq \nu_i$ it is

$\mathbb{P}(X = i) - \mathbb{P}(Y = i) = \mu_i - \nu_i$.

If we let S^+ be the set of all the states i for which $\mu_i \geq \nu_i$, then the right-hand side of Eqn (11.2) is equal to $\mu_{S^+} - \nu_{S^+}$, which is equal to $d_V(X, Y)$ from the argument in Lemma 15. This gives the first part of the lemma.

Equality holds in Eqn. (11.1) if we take a joint distribution where $X = Y$ as much as possible. Specifically, let $m_i = \min(\mathbb{P}(X = i), \mathbb{P}(Y = i))$. If $\sum_i m_i = 1$, then X and Y have the same distribution and we are done. Otherwise, let $Z = (X, Y)$ be defined by

$$\mathbb{P}(X = i, Y = j) = \begin{cases} m_i & \text{if } i = j; \\ \frac{(\mu_i - m_i)(\nu_j - m_j)}{1 - \sum_z m_z} & \text{otherwise.} \end{cases}$$

The idea behind this choice of Z is to first match X and Y as much as possible and then force X and Y to behave independently if they do not match.

For this choice of Z,

$\mathbb{P}(X = Y) = \sum_i m_i = 1 - d_V(\mu, \nu)$.

It remains to show that, for this choice of Z, $\mathbb{P}(X = i) = \mu_i$; the same argument will hold for $\mathbb{P}(Y = j)$. If $m_i = \mu_i$ then $\mathbb{P}(X = i, Y = i) = m_i$ and $\mathbb{P}(X = i, Y = j) = 0$ when $x \neq y$, so $\mathbb{P}(X = i) = \mu_i$. If $m_i = \nu_i$, then

$\mathbb{P}(X = i) = \sum_j \mathbb{P}(X = i, Y = j)$

$= m_i + \sum_{y\neq x} \frac{(\mu_i - m_i)(\nu_j - m_j)}{1 - \sum_z m_z}$

$= m_i + \frac{(\mu_i - m_i)\sum_{y \neq x}(\nu_j - m_j)}{1-\sum_z m_z}$
$= m_i + \frac{(\mu_i - m_i)(1-\nu_i-(\sum_z (m_z - m_i))}{1-\sum_z m_z}$
$= m_i + (\mu_i - m_i)$
$= \mu_i,$
completing the proof.

Again, examining a specific example (such as in Figure 11.1???) helps understand the above proof.

Example 29. Maximal Coupling for Herman

(Sinclair)

Theorem 28. *The mixing time $\tau(\varepsilon)$ of an ergodic*[24] *Markov chain satisfies: $\tau(\varepsilon) \leq \mathcal{T}\lceil \ln\ \varepsilon^{-1} \rceil$, where $\mathcal{T} = \min\{t :\ \mathbb{P}(T_{i,j} > t) \leq e^{-1}$ for all $i, j\}$.*

Recall that $T_{i,j} = \min\{t :\ X_t = Y_t \mid X_0 = i, Y_0 = j\}$. Thus, to obtain an upper bound on the mixing time, it suffices to find a coupling with a small "threshold" time $\mathcal{T}$.

Proof. Letting $Y_0 = j$ distributed according to the stationary distribution π, we have, using Lemma 16:
$d_V(X_t, \pi)) \leq \mathbb{P}(X_t \neq Y_t) \leq \max_{i,j} \mathbb{P}(T_{i,j} > t)$
Also, by definition of $\mathcal{T}$ we have, for any positive integer k and all pairs $x, y \in S$:
$\mathbb{P}(T_{i,j} > k\mathcal{T}) \leq e^{-k}$
To see this, consider a sequence of k epochs each of length $\mathcal{T}$, during each of which coupling fails to occur with probability at most e^{-1} The two inequalities yield the theorem.

Rather than with $\mathcal{T}$ itself, it is often convenient to work with the expected coupling time $\mathbf{T}$ (defined as $\mathbf{T} = \max_{i,j}\ E[T_{i,j}]$, see Def. 31). Using Markov's inequality,[25] we have:

Theorem 29. *The mixing time $\tau(\varepsilon)$ of an ergodic Markov chain satisfies: $\tau(\varepsilon) \leq e\mathbf{T}\lceil \ln\ \varepsilon^{-1} \rceil$.*

This is somewhat cruder than the previous theorem, but often easier to use in practice when an upper bound on the expectation of $T_{i,j}$ is readily available but its distribution is more complicated.

NB: Theorem 28 has a converse which (very loosely stated) says that there always exists a coupling that captures the time taken for the chain to converge (see (Griffeath,1978)).

Theorem 30. *(*Convergence to Steady State (Brémaud,p.130) *Let P be an irreducible and aperiodic transition matrix on S. For all probability distributions μ and ν on S,*
$\lim_{n\to\infty} d_V(\mu P^n, \nu P^n) = 0.$
In particular, if ν is the stationary distribution π, $\lim_{n\to\infty} d_V(\mu P^n, \pi) = 0$,
and with $\mu = \delta_j$, the probability distribution putting all its mass on j,
$\lim_{n\to\infty} \sum_{i\in S} |p_{ji}(n) - \pi_i| = 0.$

Proof. (sketch). By constructing two coupling chains with initial distributions μ and ν respectively, using the facts that the product chain is irreducible, aperiodic, with $p_{ij}(n)p_{k,\ell}(n)$ as probability of transition from (i, k) to (j, ℓ), and $\{\pi_i\pi_j\}_{(i,j)\in S^2}$ as a stationary distribution.

9.5 The Markov Chain Monte Carlo Method (MU, p.263)

The Monte Carlo method is absed on sampling. It is often difficult to generate a random sample with the required probability distribution. The Markov chain Monte Carlo (MCMC) method provides a very general approach to sampling from a desired probability distribution. The basic idea is to define an ergodic (irreducible aperiodic) Markov chain whose set of states is the sample space and whose stationary distribution is the required sampling distribution. Let $X_0, X_1, \ldots, X_n$ be a run of the chain. The Markov chan converges to the stationary distribution, so it can be used as a sample. Similarly, repeating this argument with X_r as the starting point, we can use X_{2r} as a sample, and so on. We can therefore use the sequence $X_r, X_{2r}, X_{3r}, \ldots$ as almost independent samples from the stationary distribution of the Markov

[24] LF: A-t-on besoin de l'ergodicité ?
[25] In order to ensure $\mathbb{P}(T_{i,j} > t) < e^{-1}$, it suffices to have $(E[T_{i,j}]/t) < e^{-1}$, i.e.: $t > eE[T_{i,j}]$.

chain. The efficiency of this approach depends on (a) how large r must be to ensure a suitably good sample and (b) how much computation is required for each step of the Markov chain.

The *Metropolis algorithm* refers to a general construction that transforms any ireducible Markov chain on a state space S to a time-reversible Marko chain with a required stationary distribution.

Let us assume that we have designed an irreducible state space for our Markov chain; now we want to construct a Markov chain on this state space with a stationary distribution $\pi_x = b(x)/B$, where for all $x \in S$ we have $b(x) > 0$ and such that $B = \sum_{x \in S} b(x)$ is finite. We will only need the ratios between the required probabilities; the sum B can be unknown.

Proposition 26. *For a finite state space S and neighborhood structure $\{N(X) \mid x \in S\}$, let $N = \max_{x \in S} |N(x)|$. Let M be any number such that $M \geq N$. For all $x \in S$, let $\pi_x > 0$ be the desired probability of state x in the stationary distribution. Consider a Markov chain where*

$$P_{x,y} = \begin{cases} (1/M)\min(1, \pi_y/\pi_x) & \text{if } x \neq y \text{ and } y \in N(x); \\ 0 & \text{if } x \neq y \text{ and } y \notin N(x); \\ 1 - \sum_{y \neq x} P_{x,y} & \text{if } x = y. \end{cases}$$

Then, if this chain is irreducible and aperiodic, the stationary distribution is given by the probabilities π_x.