Controlled Markov Chains, Graphs, and Hamiltonicity

Controlled Markov Chains, Graphs, and Hamiltonicity

Jerzy A. Filar

School of Mathematics and Statistics
University of South Australia
Mawson Lakes
SA 5095
Australia

j.filar@unisa.edu.au

the essence of knowledge

Boston – Delft

Foundations and Trends® in Stochastic Systems

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is J. A. Filar, Controlled Markov Chains, Graphs, and Hamiltonicity, Foundation and Trends® in Stochastic Systems, vol 1, no 2, pp 77–162, 2006

ISBN: 978-1-60198-088-5

Foundations and Trends® in Stochastic Systems
Volume 1 Issue 2, 2006

Editorial Board

Editorial Scope

Foundations and Trends® in Stochastic Systems will publish survey and tutorial articles in the following topics:

- Brownian Models
- Logistics and Supply Chain Applications
- Markov Decision Processes
- Mathematical Finance and Insurance
- Point Processes
- Queuing Theory
- Reliability Theory
- Risk Analysis
- Simulation
- Stochastic Control
- Stochastic Networks
- Stochastic Optimization
- Stochastic Programming
- Telecom and Internet Applications

Information for Librarians
Foundations and Trends® in Stochastic Systems, 2006, Volume 1, 4 issues. ISSN paper version 1551-3106. ISSN online version 1551-3092. Also available as a combined paper and online subscription.

Foundations and Trends® in
Stochastic Systems
Vol. 1, No. 2 (2006) 77–162

DOI: 10.1561/0900000003

now
the essence of knowledge

Controlled Markov Chains, Graphs, and Hamiltonicity

Jerzy A. Filar

School of Mathematics and Statistics, University of South Australia, Mawson Lakes, SA 5095, Australia, j.filar@unisa.edu.au

Abstract

This manuscript summarizes a line of research that maps certain classical problems of discrete mathematics — such as the Hamiltonian Cycle and the Traveling Salesman Problems — into convex domains where continuum analysis can be carried out. Arguably, the inherent difficulty of these, now classical, problems stems precisely from the discrete nature of domains in which these problems are posed. The convexification of domains underpinning the reported results is achieved by assigning probabilistic interpretation to key elements of the original deterministic problems.

In particular, approaches summarized here build on a technique that embeds Hamiltonian Cycle and Traveling Salesman Problems in a structured singularly perturbed Markov Decision Process. The unifying idea is to interpret subgraphs traced out by deterministic policies (including Hamiltonian Cycles, if any) as extreme points of a convex polyhedron in a space filled with randomized policies.

The topic has now evolved to the point where there are many, both theoretical and algorithmic, results that exploit the nexus between

graph theoretic structures and both probabilistic and algebraic entities of related Markov chains. The latter include moments of first return times, limiting frequencies of visits to nodes, or the spectra of certain matrices traditionally associated with the analysis of Markov chains. Numerous open questions and problems are described in the presentation.

Contents

1

Embedding of a Graph in a Markov Decision Process

1.1 Introduction

Arguably, the inherent difficulty of many problems of discrete mathematics and combinatorial optimization stems, precisely, from the discrete nature of the domains in which these problems are posed. This manuscript is devoted to a line of research that maps such problems into convex domains where continuum analysis can be easily carried out. This convexification of domains is achieved by assigning probabilistic interpretation to the key elements of the original problems even though these problems are deterministic.

While there are probably other instances of similar ideas being exploited elsewhere, our approach builds on the innovation introduced in Filar and Krass [35] where the Hamiltonian Cycle and the Traveling Salesman Problems were embedded in a structured singularly perturbed Markov Decision Process (MDP, for short). The unifying idea of [35] was to interpret subgraphs traced out by deterministic policies (including Hamiltonian Cycles, if any) as extreme points of a convex polyhedron in a space filled with randomized policies.

This approach was continued by Chen and Filar [20][1] and, independently, by Feinberg [32, 33]. Further results were obtained by Filar and Liu [37], Andramonov et al. [7], Filar and Lasserre [36], Ejov et al. [24, 25, 26, 27, 28, 29, 30, 31] and in Borkar et al. [16, 17]. Thus there is now an active group of mathematicians in a number of countries interested in this approach to discrete problems. Majority of these contributions focused on the classical Hamiltonian Cycle Problem but, in principle, many of the techniques used could be adapted to other problems of discrete mathematics (as, indeed, was done by Feinberg in [33]).

The essence of the Hamiltonian Cycle Problem (HCP, for short) is contained in the following — deceptively simple — single sentence statement: *given a graph on N nodes, find a simple cycle that contains all vertices of the graph* (*Hamiltonian Cycle* (*HC*)) *or prove that HC does not exist.* The HCP is known to be NP-hard and has become a challenge that attracts mathematical minds both in its own right and because of its close relationship to the famous *Traveling Salesman Problem* (TSP). An efficient solution of the latter would have an enormous impact in operations research, optimization, and computer science. However, from a mathematical perspective the underlying difficulty of the TSP is, perhaps, hidden in the Hamiltonian Cycle Problem. Hence we focus on the latter.

Just to indicate the flavor of the results reported in this survey, consider a key observation that led to the recent results presented in Borkar et al. [16, 17]. Namely, that: *the "correct" convex domain where Hamiltonian Cycles should be sought, is the set $\mathcal{DS}$ of doubly stochastic matrices*[2] *induced by a given graph.*

The above observation is nearly obvious, once we recall the famous (and nonobvious) Birkhoff-von Neumann Theorem which states that the set of all $N \times N$ doubly stochastic matrices is the convex hull of permutation matrices. Of course, in searching for a Hamiltonian Cycle of a given graph we need to restrict ourselves to the convex hull of only those permutation matrices that correspond to subgraphs of that

[1] Despite the fact that [20] appeared before [35], the latter preceded [20].
[2] A square nonnegative matrix is doubly stochastic if both its row-sums and column-sums are equal to 1.

graph. Results in [16, 17] imply that after a suitable perturbation and defining the random variable

$$\tau_1 := \textit{the first hitting time of the home node } 1 \textit{ (after time } 0),$$

the Hamiltonian Cycle Problem essentially reduces to "merely" minimizing the variance-like functional

$$E[(\tau_1 - N)^2]$$

over the space $\mathcal{DS}$. This probabilistic, almost statistical, interpretation should permit us to bring to bear a wide range of both analytical and algorithmic tools on the HCP.

Thus the theoretical aim of this line of research is to explain the essential difficulty of the Hamiltonian Cycle Problem — that is, its NP-hardness — in analytic terms such as a measure of variability, or the size of a gap between certain optimization problems, or by the nature of certain singularities.

The algorithmic aim of these studies is to construct a general purpose heuristic algorithm for the HCP and is based on the belief that some classical "static" optimization problems can be analyzed by embedding them in suitably constructed Markov Decision Processes.

In our setting, the theoretical and algorithmic aims are not separate. Indeed, results on one of these aims seem to influence progress on the other. For instance, the heuristic algorithm in Ejov et al. [24] follows directly from [35] and [20] but has identified difficulties that some of the theoretical developments in [16] are trying to resolve.

The general approach constitutes one of the few instances where probabilistic, continuous optimization and dynamic control methods are combined to deal with a hard problem of discrete mathematics. Arguably, simulated annealing could be seen as a precursor of this approach. However, it should be mentioned that relationships between Markov chains and graphs are also of recent interest to other researchers; notably Aldous and Fill [4] and Hunter [44].

Many of the successful classical approaches of discrete optimization to the HCP and TSP focus on solving a linear programming "relaxation" followed by heuristics that prevent the formation of sub-cycles

(e.g., see Lawler et al. [52]). In the present approach, we embedded a given graph in a singularly perturbed MDP in such a way that we can identify Hamiltonian Cycles with irreducible Markov chains and subcycles with non-exhaustive ergodic classes. This permitted a search for a Hamiltonian Cycle in the frequency space of an MDP that is a polytope with a nonempty interior, thereby converting the original discrete problem to a continuous one.

Next we shall, briefly, differentiate between our approach and some of the best known, well established, approaches to the HCP.

We first note that the present line of research is essentially different from that adopted in the study of *random graphs* where an underlying random mechanism is used to generate a graph (eg., see Karp's seminal paper [47]). In our approach, the graph that is to be studied is given and fixed but a *controller* can choose arcs according to a probability distribution and with a small probability (due to a perturbation) an arc may take you to a node other than its "head." Of course, random graphs played an important role in the study of Hamiltonicity, a striking result to quote is that of Robinson and Wormald [62] who showed that *with high probability* k-regular graphs[3] are Hamiltonian for $k \geq 3$.

Typical general purpose heuristic algorithms can, perhaps, be classified (we cite only few representative papers) as *rotational transformation* algorithms Posa [60], *cycle extension* Bollobas et al. [14], *long path* algorithms [50], *low degree vertices* algorithms Broder et al. [18], Brunacci [19], *multipath search* Kocay and Li [50], and *pruning* algorithms Christofides [21]. Of course, much research has been done on algorithms for finding a Hamiltonian Cycle on various restricted graph classes (e.g., see Parberry [58]). Clearly, algorithms designed for particular classes of graphs tend to outperform the best general purpose algorithms when applied to graphs from these classes.

Finally, the reported results open up many natural directions for further investigation. The recently implemented heuristic interior-point algorithm (see Ejov et al. [24] and Section 3.2) is based on the cited

[3] Namely, graphs where the in-degree and the out-degree at every node is equal to k.

stochastic embedding and is performing competitively with alternative — general purpose — algorithms on various test problems including the "Knight's tour" problem on chessboards of the size up to 32×32. See also the "Branch and Fix" heuristics of Sections 3.4 and 3.5.

1.2 A Graph and A Markov Decision Process

Consider a directed graph G with the node set S and the arc set $\mathcal{A}$. We can associate a Markov Decision Process Γ with the graph G as follows:

- The set of N nodes is the finite state space $S = \{1, 2, \ldots, N\}$ and the set of arcs in G is the total action space $\mathcal{A} = \{(i,j), i,j \in S\}$ where, for each state (node) i, the action space is the set of arcs (i,j) emanating from this node and will be denoted by $\mathcal{A}(i)$.
- $\{p(j|i,a) = \delta_{aj} | a = (i,j) \in \mathcal{A}(i),\ i,j \in S\}$, where δ_{aj} the Kronecker delta, is the set of (one-step) transition probabilities. Note that, we are adopting the convention that a equals to both arc (i,j) and its "head" j, whenever there is no possibility of confusion as to the "tail" i.

A *stationary policy* f in Γ is a set of N probability vectors $f(i) = (f(i,1), f(i,2), \ldots, f(i,N))$, where $f(i,k)$ denotes the probability of choosing an action k (arc emanating from i to k) whenever state (node) i is visited. Of course, $\sum_{k=1}^{N} f(i,k) = 1$ and if the arc $(i,k) \notin \mathcal{A}(i)$, then $f(i,k) = 0$. Equivalently, it will be sometimes convenient to represent a policy f as an $N \times N$ matrix whose (i,k)th entry is $f(i,k)$. The set of all stationary policies will be denoted by $\mathcal{F}$.

A *deterministic policy* f is simply a stationary policy that selects a single action with probability 1 in every state (hence, all other available actions are selected with probability 0). That is, $f(i,k) = 1$ for some $(i,k) \in \mathcal{A}(i)$. For convenience, we will write $f(i) = k$ in this case. The set of all deterministic policies will be denoted by $\mathcal{D}$.

It is easy to check that any stationary policy $f \in \mathcal{F}$ induces a probability transition matrix

$$P(f) = \big[p(j|i,f)\big], \quad i,j = 1,\ldots,N,$$

where for all $i,j \in S$

$$p(j|i,f) = \sum_{a=1}^{N} p(j|i,a)\ f(i,a).$$

In the above summation, we assume that $p(j|i,a) = 0$ if the arc $(i,a) \notin \mathcal{A}$.

A *doubly stochastic policy* $f \in \mathcal{F}$ is one which induces a probability transition matrix $P(f)$ that is doubly stochastic; namely all of its rows and columns sum to unity. The set of all doubly stochastic policies will be denoted by $\mathcal{D}S$. It should be clear from the construction that

$$\mathcal{D}S \subseteq \mathcal{F}.$$

Assume now that 1 is the initial state (home node). We shall say that a deterministic policy f in Γ is a Hamiltonian Cycle (HC) (or simply "is Hamiltonian") in G if the sub-graph G_f with the set of arcs $\big\{(1,f(1)),(2,f(2)),\ldots,(N,f(N))\big\}$ is a HC in G.

If an analogous sub-graph G_f induced by a deterministic policy f contains cycles of length less than N, say m, we say that f has an *m-sub-cycle*.

However, such a straightforward identification of G with Γ leads to an inevitable difficulty of confronting multiple ergodic classes induced by various deterministic policies.

Note that if $f \in \mathcal{DS} \cap \mathcal{D}$, then the Markov chain induced by f corresponds to either a Hamiltonian Cycle or to a policy tracing out a union of disjoint sub-cycles in the graph G.

Example 1.1. All of the above can be illustrated on a complete graph G_4 on 4 nodes (without self-loops) in Figure 1.1. A policy f_1 such that $f_1(1) = 2$, $f_1(2) = 3$, $f_1(3) = 4$, and $f_1(4) = 1$ induces a sub-graph $G_{f_1} = \big\{(1,2),(2,3),(3,4),(4,1)\big\}$ that is a Hamiltonian Cycle. Policy f_1

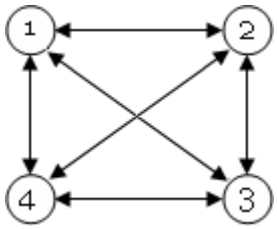

Fig. 1.1 Complete graph G_4.

also induces a Markov chain with the probability transition matrix

$$P(f_1) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

that has a single ergodic class containing all four states. A policy f_2 such that $f_2(1) = 2$, $f_2(2) = 1$, $f_2(3) = 4$, and $f_2(4) = 3$ induces a sub-graph $G_{f_2} = \{(1,2),(2,1),(3,4),(4,3)\}$ which contains two 2-sub-cycles (see Figure 1.2). Policy f_2 also induces a Markov chain with the probability transition matrix

$$P(f_2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

which has two ergodic classes corresponding to the sub-cycles of G_{f_2}.

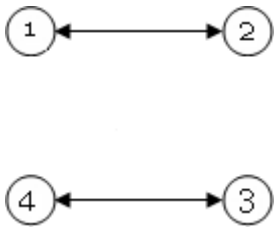

Fig. 1.2 Sub-graph G_{f_2}.

Of course, randomized stationary policies can be regarded as convex combinations of deterministic policies. For instance, if in this example we take a policy f_3 that is a set of 4 probability vectors

$$\begin{aligned} f_3(1) &= (f_3(1,2), f_3(1,3), f_3(1,4)) = (1,0,0), \\ f_3(2) &= (f_3(2,1), f_3(2,3), f_3(2,4)) = (0.8,0.2,0), \\ f_3(3) &= (f_3(3,1), f_3(3,2), f_3(3,4)) = (0,0,1), \text{ and} \\ f_3(4) &= (f_3(4,1), f_3(4,2), f_3(4,3)) = (0.2,0,0.8), \end{aligned}$$

then it is clear that $f_3 = 0.2f_1 + 0.8f_2$ which induces the Markov chain probability transition matrix

$$P(f_3) = 0.2P(f_1) + 0.8P(f_2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.8 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 1 \\ 0.2 & 0 & 0.8 & 0 \end{bmatrix}.$$

1.2.1 Classification of Deterministic Policies

We shall now describe a useful partition of $\mathcal{D}$ that is based on the graphs "traced out" in G by deterministic policies. As above, with each $f \in \mathcal{D}$ we associate a sub-graph G_f of G defined by

$$(i,j) \in G_f \iff f(i) = j.$$

We shall also denote a simple cycle of length m and beginning at 1 by a set of arcs

$$c_m^1 = \big\{(i_1 = 1, i_2), (i_2, i_3), \ldots, (i_m, i_{m+1} = 1)\big\}, \quad m = 2,3,\ldots,N.$$

Note that c_N^1 is a HC. If G_f contains a cycle c_m^1, we write $G_f \supset c_m^1$. Let

$$C_m^1 := \big\{f \in \mathcal{D} | G_f \supset c_m^1\big\},$$

namely, the set of deterministic policies that trace out a simple cycle of length m, beginning at node 1, for each $m = 2,3,\ldots,N$. Of course, C_N^1 is the (possibly empty) set of policies that correspond to HCs and any single C_m^1 can be empty depending on the structure of the original

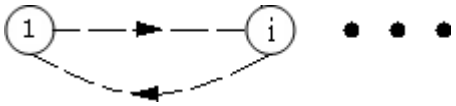

Fig. 1.3 A typical policy f in C_m^1.

graph G. Thus, a typical policy $f \in C_m^1$ traces out a graph G_f in G that might look as Figure 1.2 where the dots indicate the "immaterial" remainder of G_f in the sense that it corresponds to states/nodes that will never be observed if the process begins at node 1 and the policy f is adhered to. The broken arrows indicate a sequence of one or more arcs (Figure 1.3). We now introduce the partition of the deterministic policies of the form:

$$\mathcal{D} = \left[\bigcup_{m=2}^{N} C_m^1\right] \bigcup \mathcal{N}_c,$$

where $\mathcal{N}_c$ contains all deterministic policies that are not in any of the C_m^1s. A typical policy $f \in \mathcal{N}_c$ traces out a sub-graph G_f in G as in Figure 1.4, where the dots again denote the immaterial part of G_f. We shall call policies in $\mathcal{N}_c$ *noose cycles*. For many operations related to Markov chains induced by deterministic policies properties of interest to us will be invariant under permutations of states/nodes that leave the home node unchanged. Thus unless stated otherwise, and without loss of generality, it is sufficient to consider only $f_m \in C_m^1$ tracing out the graph in Figure 1.5 as the representative of the whole class C_m^1 and also, $f_m^k \in \mathcal{N}_c$ that traces out Figure 1.6 as the representative of the entire class $\mathcal{N}_c$.

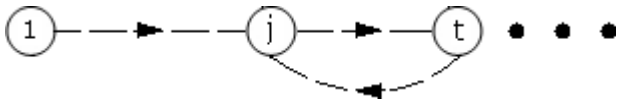

Fig. 1.4 A typical policy f in $\mathcal{N}_c$.

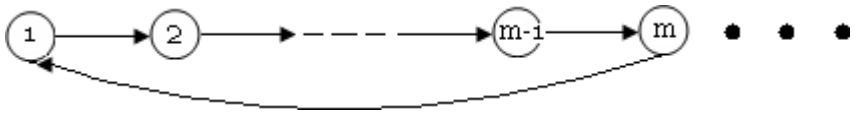

Fig. 1.5 A representative f_m of the whole class C_m^1.

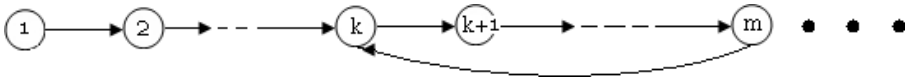

Fig. 1.6 A representative f_m^k of the whole class $\mathcal{N}_c$.

1.3 Perturbed Embedding Γ_ε

We have seen that the direct embedding of G in Γ, in general, induces a multi-chain ergodic structure. This and some other technical difficulties would vanish if we force the MDP to be *unichain*. The latter is a Markov Decision Process in which every stationary policy induces a Markov chain containing only a single ergodic class plus a (possibly empty) class of transient states. A *completely ergodic* MDP is a unichain MDP in which the class of transient states is empty no matter which stationary policy induces the Markov chain.

There are many possible ways of perturbing the MDP Γ discussed in the preceding section to obtain a parameterized family of perturbed unichain or completely ergodic MDP's Γ_ε, where $\varepsilon \in (0,1)$ will be called the *perturbation parameter*. However, all these perturbations share the characteristic of altering the ergodic structure of Markov chains induced by various stationary policies. Hence, they are so-called *singular perturbations*.

The question of what constitutes the "best" perturbation is potentially very interesting if the notion of optimality for such perturbations were formalized. In the results reported here three (generic) perturbations were considered.

1.3.1 The Symmetric Linear Perturbation

This is achieved by passing to a singularly perturbed MDP Γ_ε^s, that is obtained from Γ by introducing perturbed transition probabilities

$\{p_\varepsilon(j|i,a)|\ (i,j) \in \mathcal{A},\ i,j \in S\}$, where for any $\varepsilon \in \left(0, \frac{1}{N-1}\right)$

$$p_\varepsilon(j|i,a) := \begin{cases} 1-(N-1)\varepsilon & \text{if } a = j, \\ \varepsilon & \text{if } a \neq j. \end{cases}$$

Note that this perturbation ensures that every $f \in \mathcal{D}$ induces a Markov chain with a completely ergodic transition probability matrix $P_\varepsilon(f)$ whose dominant terms coincide with the 1-entries of the corresponding unperturbed probability transition matrix $P(f)$ that the same policy f induces in Γ.

For instance, if in the example of the preceding section we consider the policy f_2 that traces out the two sub-cycles depicted in Figure 1.2, it is clear that in Γ_ε^s the same policy induces the probability transition matrix

$$P_\varepsilon(f_2) = \begin{bmatrix} \varepsilon & 1-3\varepsilon & \varepsilon & \varepsilon \\ 1-3\varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & 1-3\varepsilon \\ \varepsilon & \varepsilon & 1-3\varepsilon & \varepsilon \end{bmatrix},$$

where, for instance, the second entry in the first row can be formally derived as

$$p_\varepsilon(2|1,f_2) = \sum_{a \in \mathcal{A}(1)} p_\varepsilon(2|1,a) f_2(1,a) = p_\varepsilon(2|1,2) f_2(1,2) = 1-3\varepsilon.$$

Of course, $P_\varepsilon(f_2)$ now has only a single ergodic class and no transient states. The latter is a desirable property but comes at a price of replacing a sparse probability transition matrix with one that is "dense," in the sense of not having any zero entries.

The above symmetric linear perturbation has been used in [16] and [17] where a perturbation was needed that also preserved double-stochasticity of a probability transition matrix.

1.3.2 The Asymmetric Linear Perturbation

In the sequence of papers that launched this topic (e.g., [20], [35], and [37]) an asymmetric linear perturbation was used. Its goal was not only to eliminate multiple ergodic classes but also to differentiate

the home node from all other nodes and to maintain the sparsity of probability transition matrices induced by deterministic policies. This was achieved by passing to a singularly perturbed MDP Γ_ε^a, that is obtained from Γ by introducing perturbed transition probabilities $\{p_\varepsilon(j|i,a)\big|\ (i,j) \in \mathcal{A},\ i,j \in S\}$, where for any $\varepsilon \in (0,1)$

$$p_\varepsilon(j|i,a) := \begin{cases} 1 & \text{if } i = 1 \text{ and } a = j, \\ 0 & \text{if } i = 1 \text{ and } a \neq j, \\ 1 & \text{if } i > 1 \text{ and } a = j = 1, \\ \varepsilon & \text{if } i > 1,\ a \neq j \text{ and } j = 1, \\ 1 - \varepsilon & \text{if } i > 1,\ a = j \text{ and } j > 1, \\ 0 & \text{if } i > 1,\ a \neq j \text{ and } j > 1. \end{cases}$$

Note that 1 denotes the home node. For each pair of nodes i, j (not equal to 1) corresponding to a (deterministic) arc (i,j), our perturbation replaces that arc by a pair of "stochastic arcs" $(i,1)$ and (i,j) (see Figure 1.7) with weights ε and $1-\varepsilon$, respectively. This stochastic perturbation has the interpretation that a decision to move along arc (i,j) results in movement along (i,j) only with probability of $(1-\varepsilon)$ and with probability ε it results in a return to the home node 1. We emphasize that the perturbation is chosen to ensure that the Markov chain defined by $P_\varepsilon(f)$ contains only a single ergodic class. On the other hand, the ε-perturbed process Γ_ε clearly "tends" to Γ as $\varepsilon \to 0$, in the sense that $P_\varepsilon(f) \to P_0(f)$ for every stationary policy f.

For instance, if in the example of the preceding section we consider the policy f_2 that traces out the two sub-cycles depicted in Figure 1.2,

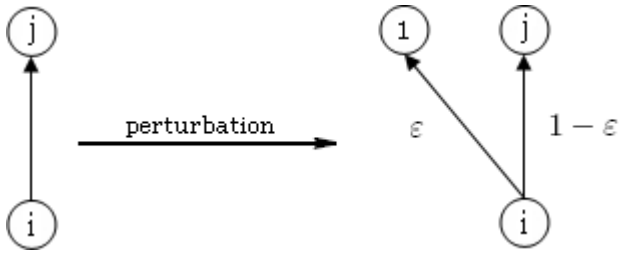

Fig. 1.7 Perturbation of a deterministic action (arc).

it is clear that in Γ_ε^a the same policy induces the probability transition matrix

$$P_\varepsilon(f_2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \varepsilon & 0 & 0 & 1-\varepsilon \\ \varepsilon & 0 & 1-\varepsilon & 0 \end{bmatrix}$$

that still has only a single ergodic class, but now has a nonempty class $\{3,4\}$ containing the two transient states created by this perturbation.

1.3.3 The Asymmetric Quadratic Perturbation

Of course, there are situations where the presence of transient states is undesirable. For that reason, in [29] and [24] the asymmetric perturbation was modified slightly by introducing a quadratic (in ε) term. This was achieved by passing to a singularly perturbed MDP Γ_ε^q, that is obtained from Γ by introducing perturbed transition probabilities $\{p_\varepsilon(j|i,a)\big|\ (i,j) \in \mathcal{A},\ i,j \in S\}$, where for any $\varepsilon \in \left(0, \frac{1}{\sqrt{N-2}}\right)$

$$p_\varepsilon(j|i,a) := \begin{cases} 1-(N-2)\varepsilon^2 & \text{if } i=1 \text{ and } a=j, \\ \varepsilon^2 & \text{if } i=1 \text{ and } a \neq j > 1, \\ 1 & \text{if } i>1 \text{ and } a=j=1, \\ \varepsilon & \text{if } i>1,\ a \neq j \text{ and } j=1, \\ 1-\varepsilon & \text{if } i>1,\ a=j \text{ and } j>1, \\ 0 & \text{if } i>1,\ a \neq j \text{ and } j>1. \end{cases}$$

In this instance, in the example of the preceding section we consider the policy f_2 that traces out the two sub-cycles depicted in Figure 1.2, it is clear that in Γ_ε^q the same policy induces the probability transition matrix

$$P_\varepsilon(f_2) = \begin{bmatrix} 0 & 1-2\varepsilon^2 & \varepsilon^2 & \varepsilon^2 \\ 1 & 0 & 0 & 0 \\ \varepsilon & 0 & 0 & 1-\varepsilon \\ \varepsilon & 0 & 1-\varepsilon & 0 \end{bmatrix},$$

which now has a single ergodic class and no transient states. The fact that this perturbation preserves much of the sparsity of $P(f)$ is easier to see in examples of higher dimensions.

Remark 1.1. It should be clear that the above perturbations link this topic with the well developed field of analytic perturbations of operators and Markov chains. The treatise by Kato [48] is the seminal reference for the latter topic. However, some of the techniques used here are, perhaps, somewhat more in the spirit of the recent developments such as those reported in Avrachenkov et al. [9, 10]. More particularly, in the context of perturbation and sensitivity analysis of MDP's the papers by Schweitzer [65, 66] and Veinott [68] are quite relevant to the results reported here.

1.4 Background from Markov Chains and MDPs

As before, let $f \in \mathcal{F}$ be a stationary policy and $P(f)$ be the corresponding probability transition matrix. By $P^*(f)$ we denote its *stationary distribution matrix*, that is defined as the *limit Cesaro-sum matrix*

$$P^*(f) := \lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} P^t(f), \quad P^0(f) = I,$$

where I is an $N \times N$ identity matrix. It is well known (e.g., see [13]) that the above limit exists and satisfies the identity

$$P(f)P^*(f) = P^*(f)P(f) = P^*(f)P^*(f) = P^*(f). \tag{1.1}$$

An important special case arises when the Markov chain corresponding to $P(f)$ contains only a single ergodic class. In this case, $P^*(f)$ consists of identical rows, each of which will be denoted by $\pi(f) = (\pi_1(f), \pi_2(f), \ldots, \pi_N(f))$, where $\sum_{i=1}^{N} \pi_i(f) = 1$. Hence, $\pi(f)$ constitutes a probability vector that is often called the *stationary or invariant distribution* of such a Markov chain. It follows from the preceding identity that $\pi(f)$ is a solution of the linear system of equations:

$$\pi(f)P(f) = \pi(f); \quad \pi(f)\mathbf{1} = 1, \tag{1.2}$$

where $\mathbf{1} := (1,1,\ldots,1)^T$. Indeed, $\pi(f)$ is the unique solution of (1.2).

An even more special but also extremely important case is that of an *irreducible* Markov chain, where $P(f)$ contains only a single ergodic class and no transient states. In this case, the invariant distribution vector $\pi(f)$ is still the unique solution of (1.2) and possesses the additional useful property that $\pi_i(f) > 0 \ \ \forall i = 1,2,\ldots,N$.

Another two very important matrices associated with the probability transition matrix $P(f)$ of a Markov chain induced by a policy $f \in \mathcal{F}$ are the *fundamental matrix* $G(f)$ that is defined by

$$G(f) := \big(I - P(f) + P^*(f)\big)^{-1} = \lim_{\beta \to 1^-} \sum_{t=0}^{\infty} \beta^t (P(f) - P^*(f))^t, \quad (1.3)$$

and the closely related *deviation matrix*

$$D(f) := G(f) - P^*(f). \quad (1.4)$$

The following identities are well known (e.g., see [13, 45, 49, 61])

$$D(f)P^*(f) = P^*(f)D(f) = 0, \quad (1.5)$$

where the 0 on the right-hand side above is an $N \times N$ matrix with $0s$ in all entries, and

$$\begin{aligned}(I - P(f))D(f) + P^*(f) - I &= D(f)(I - P(f)) + P^*(f) - I \\ &= P^*(f), \end{aligned} \quad (1.6)$$

$$D(f)\mathbf{1} = \mathbf{0} \quad \& \quad G(f)\mathbf{1} = \mathbf{1}, \quad (1.7)$$

where $\mathbf{1}$ and $\mathbf{0}$ are N-vectors consisting entirely of $1s$ and $0s$, respectively.

Another matrix, also induced by any policy $f \in \mathcal{F}$, that plays an important role in the theory of Markov Decision Processes is the resolvent-like matrix

$$[I - \beta P(f)]^{-1} = \sum_{t=0}^{\infty} \beta^t P^t(f), \quad (1.8)$$

where the parameter $\beta \in [0,1)$, is frequently called the *discount factor*. Note that this choice of the domain for β ensures that the spectral

radius of $\beta P(f)$ is strictly less than 1, thereby guaranteeing the existence of the above inverse and the power series expansion.

In a traditional, Markov Decision Process setting, there is also a *reward or a cost*, denoted by $r(i,a)$, associated with each state i and action a. The interpretation is that this is the reward/cost associated with action a if that action is selected in state i. However, if actions are being selected in accordance with a stationary policy $f \in \mathcal{F}$, then the "lottery" on the actions available in state i is prescribed by f and hence the expected reward in that state is given by

$$r(i,f) := \sum_{a=1}^{N} r(i,a) f(i,a), \quad i \in \mathcal{S}. \tag{1.9}$$

This immediately defines the *expected reward/cost vector*, $r(f)$, *induced by* f the transpose of which is defined by

$$r^T(f) := (r(1,f), r(2,f), \ldots, r(N,f)). \tag{1.10}$$

There are now two well-known MDP's that have been extensively studied in the literature (e.g., see [61] and [38]). They are differentiated by the manner of aggregating the infinite stream of expected rewards/costs induced by a policy[4] f. The first of these is the so-called *limiting* (*or long-run*) *average process* (*AMD, for short*) where the performance of the policy f is defined by the *value vector*

$$v(f) := P^*(f) r(f), \tag{1.11}$$

whose entries $v(i,f) := [P^*(f)r(f)]_i$ for each $i \in \mathcal{S}$ are simply the long-run average expected rewards induced by f when the process begins in state i.

Analogously, the second process is the so-called *discounted Markov Decision Process* (*DMD, for short*) where the performance of the policy

[4] Note that in the theory of MDP's policies can be more general than the stationary policies introduced here. For instance, they may depend on past histories of states and actions. However, for our purposes stationary policies suffice.

f is defined by the *value vector*

$$v^{\beta}(f) := [I - \beta P(f)]^{-1} r(f), \tag{1.12}$$

whose entries $v^{\beta}(i,f) := \{[I - \beta P(f)]^{-1} r(f)]\}_i$ for each $i \in \mathcal{S}$ are simply the discounted expected rewards induced by f when the process begins in state i.

The optimization problems normally associated with the AMD and DMD processes, respectively, are

$$\max_{f \in \mathcal{F}} v(f), \quad \text{and} \quad \max_{f \in \mathcal{F}} v^{\beta}(f), \tag{1.13}$$

where the maximization is taken componentwise in the above expressions.

The above optimization problems are well understood and, for most purposes, completely solved. In particular, it is a remarkable fact that in each case there exist deterministic policies $f^0, f^0_\beta \in \mathcal{D}$ that, respectively, attain the maxima in (1.13), componentwise. Furthermore, if β is sufficiently near 1, there exists a deterministic policy that is simultaneously optimal for both the AMD and DMD processes. While many outstanding researchers contributed to this topic Blackwell's 1962 paper (see [13]) is, perhaps, the authoritative reference. There are also many treatments of this problem in text books (e.g., see [23, 38, 61]).

Remark 1.2. In our embedding of the Hamiltonian Cycle Problem in Markov Decision Processes, in most instances, we use rewards/costs only to differentiate the home node 1 from the other nodes. For that purpose it is sufficient to assume that $r(i,a) \equiv 0$ for all actions/arcs emanating from nodes other than the home node, and that $r(1,a) \equiv 1$ for all actions/arcs emanating from the home node 1. Hence, unless explicitly stated otherwise, we shall assume that

$$r^T(f) = e_1^T = (1,0,\ldots,0), \quad \forall f \in \mathcal{F}.$$

Remark 1.3. The above notation was developed for the unperturbed MDP Γ, however, whenever we use one of the previously defined

ε-perturbed MDP's, we shall simply add a subscript ε to the relevant quantity. For instance, the probability transition, stationary distribution and fundamental matrices $P(f)$, $P^*(f)$, $G(f)$ are replaced by $P_\varepsilon(f)$, $P^*_\varepsilon(f)$, $G_\varepsilon(f)$, where $\varepsilon > 0$ is the perturbation parameter. Similar indexing by ε will also apply to other quantities. When the choice of the type of perturbation used is clear from context, the perturbed MDP will be denoted simply by Γ_ε.

References

[1] M. Abbad, T. Bielecki, and J. A. Filar, "Algorithms for singularly perturbed limiting average Markov control problems," *IEEE Transactions on Automatic Control*, vol. 37, pp. 1421–1425, 1992.

[2] M. Abbad and J. A. Filar, "Perturbations and stability theory for Markov control problems," *IEEE Transactions on Automatic Control*, vol. 37, pp. 1415–1420, 1992.

[3] W. Adams and P. Lostaunau, *An Introduction to Gröbner Bases*. Graduate Studies in Mathematics, American Mathematical Society, Vol. 3, 1994.

[4] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*. Book in progress, available at http://www.stat.berkeley.edu/users/aldous/RWG/book.html.

[5] A. Altman and J. Gondzio, "Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization," *Optimization Methods and Software*, vol. 11–12, pp. 275–302, 1999.

[6] E. D. Andersen, J. Gondzio, C. Mészáros, and X. Xu, "Implementation of interior point methods for large scale linear programming," in *Interior Point Methods in Mathematical Programming*, (T. Terlaky, ed.), pp. 189–252, Kluwer Academic Publishers, 1996.

[7] M. Andramonov, J. A. Filar, A. Rubinov, and P. Pardalos, "Hamiltonian cycle problem via Markov chains and min-type approaches," in *Approximation and Complexity in Numerical Optimization: Continuous and Discrete Problems*, (P. Pardalos, ed.), pp. 31–47, Dordrecht, The Netherlands: Kluwer Academic Publishers, 2000.

[8] M. Andramonov, A. Rubinov, and B. M. Glover, "Cutting angle methods in global optimization," *Applied Mathematics Letters*, vol. 12, pp. 95–100, 1999.

[9] K. E. Avrachenkov, J. A. Filar, and M. Haviv, "Singular perturbations of Markov chains and Decision Processes," in *Handbook of Markov Decision Processes*, (E. A. Feinberg and A. Shwartz, eds.), pp. 113–152, Boston: Kluwer, 2002.

[10] K. E. Avrachenkov, M. Haviv, and P. G. Howlett, "Inversion of analytic matrix functions that are singular at the origin," *SIAM Journal on Matrix Analysis and Applications*, vol. 22, pp. 1175–1189, 2001.

[11] A. M. Bagirov, A. M. Rubinov, and J. Zhang, "Local optimization method with global multidimensional search," *Journal of Global Optimization*, vol. 32, no. 2, DOI 10.1007/s10898-004-2700-0, June 2005.

[12] L. Batten and G. Beliakov, "Fast algorithm for the cutting angle method of global optimization," *Journal of Global Optimization*, vol. 24, pp. 149–161, 2003.

[13] D. Blackwell, "Discrete dynamic programming," *Annals of Mathematical Statistics*, vol. 33, pp. 719–726, 1962.

[14] B. Bollobas, T. Fenner, and A. Frieze, "An algorithm for finding Hamiltonian paths and cycles in random graphs," *Combinatorica*, vol. 7, no. 4, pp. 327–341, 1987.

[15] V. S. Borkar, *Probability Theory: An Advanced Course*. New York: Springer-Verlag, 1995.

[16] V. S. Borkar, V. Ejov, and J. A. Filar, "Directed graphs, Hamiltonicity and doubly stochastic matrices," *Random Structures and Algorithms*, vol. 25, pp. 376–395, 2004.

[17] V. S. Borkar, V. Ejov, and J. A. Filar, "On the Hamiltonicity gap and doubly stochastic matrices," submitted to *Random Structures and Algorithms*, 2006.

[18] A. Broder, A. Frieze, and E. Shamir, "Finding hidden Hamiltonian cycles," *Random Structures and Algorithms*, vol. 5, no. 3, pp. 395–410, 1994.

[19] F. Brunacci, "Two useful tools for constructing Hamiltonian circuits," *EJOR*, vol. 34, pp. 231–236, 1988.

[20] M. Chen and J. A. Filar, "Hamiltonian cycles, quadratic programming and ranking of extreme points," in *Global Optimization*, (C. Floudas and P. Pardalos, eds.), pp. 32–49, Princeton University Press, 1992.

[21] N. Christofides, *Graph Theory: An Algorithmic Approach*. New York: Academic Press, 1975.

[22] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. New York: Springer, Second ed., 1997.

[23] C. Derman, *Finite State Markovian Decision Processes*. New York: Academic Press, 1970.

[24] V. Ejov, J. A. Filar, and J. Gondzio, "An interior point heuristic for the Hamiltonian cycle problem via Markov decision processes," *Journal of Global Optimization*, vol. 29, pp. 315–334, 2003.

[25] V. Ejov, J. A. Filar, M. Haythorpe, and G. T. Nguyen, "Refined MDP-based branch-and-bound algorithm for the Hamiltonian cycles problem," submitted to *Mathematics of Operations Research*, 2007.

[26] V. Ejov, J. A. Filar, W. Murray, and G. T. Nguyen, "Determinants and longest cycles of graphs," submitted to *SIAM Journal on Discrete Mathematics*, 2007.

[27] V. Ejov, J. A. Filar, and M. Nguyen, "Hamiltonian cycles and singularly perturbed Markov chains," *Mathematics of Operations Research*, vol. 29, no. 1, pp. 114–131, 2004.

[28] V. Ejov, J. A. Filar, and F. Spieksma, "On regularly perturbed fundamental matrices," *Journal of Mathematical Analysis and Applications*, vol. 336, pp. 18–30, 2007.

[29] V. Ejov, J. A. Filar, and J. Thredgold, "Geometric interpretation of the Hamiltonian cycle problem via singularly perturbed Markov decision processes," *Optimization*, vol. 52, pp. 441–458, 2003.

[30] V. Ejov, J. A. Flar, S. K. Lucas, and J. Nelson, "Solving the Hamiltonian cycle problem using symbolic determinants," *Taiwanese Journal of Mathematics*, vol. 10, pp. 327–337, 2006.

[31] V. Ejov, J. A. Flar, S. K. Lucas, and P. G. Zograf, "Clustering of spectra and fractals of regular graphs," *Journal of Mathematical Analysis and Applications*, vol. 333, pp. 236–246, 2007.

[32] E. Feinberg, "Constrained discounted Markov decision process and Hamiltonian cycles," *Mathematics of Operations Research*, vol. 25, pp. 130–140, 2000.

[33] E. Feinberg and M. T. Curry, "Generalized pinwheel problem," *Mathematical Methods of Operations Research*, vol. 62, pp. 99–122, 2005.

[34] J. A. Filar, A. Gupta, and S. K. Lucas, "Connected co-spectral graphs are not necessarily both Hamiltonian," *The Australian Mathematical Society Gazette*, vol. 32, no. 3, p. 193, 2005.

[35] J. A. Filar and D. Krass, "Hamiltonian cycles and Markov chains," *Mathematics of Operations Research*, vol. 19, pp. 223–237, 1994.

[36] J. A. Filar and J.-B. Lasserre, "A non-standard branch and bound method for the Hamiltonian cycle problem," *ANZIAM Journal*, vol. 42, no. E, pp. C556–C577, 2000.

[37] J. A. Filar and K. Liu, "Hamiltonian cycle problem and singularly perturbed Markov decision process," in *Statistics, Probability and Game Theory: Papers in Honor of David Blackwell*, IMS Lecture Notes — Monograph Series, USA, 1996.

[38] J. A. Filar and K. Vrieze, *Competitive Markov Decision Processes*. New York: Springer-Verlag, 1996.

[39] J. Gondzio, "Multiple centrality corrections in a primal-dual method for linear programming," *Computational Optimization and Applications*, vol. 6, pp. 137–156, 1996.

[40] R. R. Greenlaw and R. Petreschi, "Cubic graphs," *ACM Computational Survey*, vol. 27, pp. 471–495, 1995.

[41] C. M. Grinstead and J. L. Snell, *Introduction to Probability*. American Mathematical Society, Second ed., 2003.

[42] A. Hordijk, R. Dekker, and L. C. M. Kallenberg, "Sensitivity-analysis in discounted Markovian decision problems," *OR Spektrum*, vol. 7, pp. 143–151, 1985.

[43] A. Hordijk and L. C. M. Kallenberg, "Transient policies in discrete dynamic programming: Linear programming including suboptimality tests and additional constraints," *Mathematical Programing*, vol. 30, pp. 46–70, 1984.

[44] J. J. Hunter, "Mixing times with applications to perturbed Markov chains," *Linear Algebra and its Applications*, vol. 417, no. 1, pp. 108–123, 2006.

[45] L. C. M. Kallenberg, *Linear programming and Finite Markovian Control Problems*. Mathematical Centre Tracts 148, Amsterdam, 1983.

[46] L. C. M. Kallenberg, "Survey of linear programming for standard and non-standard Markovian control problems, Part I: Theory," *ZOR — Methods and Models in Operations Research*, vol. 40, pp. 1–42, 1994.

[47] R. Karp, "Probabilistic analysis of partitioning algorithms for the travelling-salesman problem in the plane," *Mathematics of Operations Research*, vol. 2, no. 3, pp. 209–224, 1977.

[48] T. Kato, *Perturbation Theory for Linear Operators*. Berlin: Springer-Verlag, 1966.

[49] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. New York: Springer, 1983.

[50] W. Kocay and P.-C. Li, "An algorithm for finding a long path in a graph," *Utilitas Mathematica*, vol. 45, pp. 169–185, 1994.

[51] M. Kreuzer and L. Robbiano, *Computational Commutative Algebra 1*. New York: Springer, 2000.

[52] E. L. Lawler, J. K. Lenstra, A. R. Kan, and D. Shmoys, "The travelling salesman problem: A guided tour of combinatorial optimization," Wiley Series in *Discrete Mathematics and Optimization*. New York: John Wiley and Sons, 1985.

[53] N. Litvak and V. Ejov, "Optimality of the Hamiltonian cycle," preprint, University of Twente, Enschede, The Netherlands, 2007.

[54] M. Meringer, "Fast generation of regular graphs and construction of cages," *Journal of Graph Theory*, vol. 30, pp. 137–146, 1999.

[55] P. Mnëv, *Discrete Path Integral Approach to the Trace Formula for Regular Graphs*. Preprint, Steklov Math. Inst., St. Petersburg, 2006.

[56] W. Murray and K.-M. Ng, "Algorithms for global optimization and discrete problems based on methods for local optimization," in *Handbook of Global Optimization — Volume 2: Heuristic Approaches*, (P. Pardalos and H. E. Romeijn, eds.), Kluwer Academic Publishers, 2002.

[57] W. Murray and K.-M. Ng, *An Algorithm for Nonlinear Optimization Problems with Discrete Variables*. preprint, 2003.

[58] I. Parberry, "An efficient algorithm for the Knight's tour problem," *Discrete Applied Mathematics*, vol. 73, pp. 251–260, 1997.

[59] J. W. Pitman, "Occupation measures for Markov chains," *Advances in Applied Probability*, vol. 9, pp. 69–86, 1997.

[60] L. Posa, "Hamiltonian circuits in random graphs," *Discrete Mathematics*, vol. 14, pp. 359–364, 1976.

[61] M. L. Puterman, *Markov Decision Processes*. New York: John Wiley & Sons, 1994.

[62] R. Robinson and N. Wormald, "Almost all regular graphs are Hamiltonian," *Random Structures and Algorithms*, vol. 5, no. 2, pp. 363–374, 1994.

[63] R. Y. Rubinstein, "The cross-entropy method for combinatorial and continuous optimization," *Methodology and Computing in Applied Probability*, vol. 2, pp. 127–190, 1999.

[64] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entrropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning*. New York: Springer-Verlag, 2004.

[65] P. J. Schweitzer, "Perturbation theory and finite Markov chains," *Journal of Applied Probability*, vol. 5, pp. 401–413, 1968.

[66] P. J. Schweitzer, "Perturbation series expansion of nearly completely-decomposable Markov chains," in *Teletraffic Analysis and Computer Performance Evaluation*, (O. J. Boxma, J. W. Cohen, and H. C. Tijms, eds.), pp. 319–328, Elsevier Science Publishers B.V. (North Holland), 1986.

[67] R. S. Varga, *Matrix Iterative Analysis*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1962.

[68] A. F. Veinott, Jr., "Discrete dynamic programming with sensitive discount optimality criteria," *Annals of Mathematical Statistics*, vol. 40, pp. 1635–1660, 1969.

[69] S. J. Wright, *Primal-Dual Interior-Point Methods*. Philadelphia: SIAM, 1997.