Chapter 2
Intriguing Properties

2.1 Preliminaries and Notation

In this book, all graphs are connected and undirected, unless otherwise stated. We follow the graph terminology and conventions from Harary [59], where the reader can find an excellent introduction to graph theory. Consider a graph $G = (V(G), E(G)) = (V, E)$, where V is the set of vertices of G, $|V| = N$, and E is the set of edges of G. A graph with N vertices is said to be a graph of *order* N. As the graph G is undirected, for every edge $(i, j) \in E$ there exists an opposite edge $(j, i) \in E$, where $i \neq j$. A *loop* is an edge (i, i) joining a vertex to itself. We do not consider *multi-edges*, which are distinct edges that connect the same pair of vertices, and we use the terms *edge* and *arc* interchangeably.

Subgraphs and Regularity Consider a graph G', and let $V(G')$ be the set of its vertices and $E(G')$ be the set of its edges. Then G' is a *subgraph* of G if $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$. A subgraph G' is a *spanning subgraph* if $V(G') = V(G)$. From now on, the term *subgraph* refers to a spanning subgraph, unless otherwise stated. A vertex j is a *neighbour* of i, or vertex j is *adjacent* to i, if there exists an edge between them, that is, $(i, j) \in E(G)$. A vertex v has a *degree* d if it has d neighbours, and we write $\deg(v) = d$. A graph is *k-regular* if every vertex $i \in V$ has the same degree k, and a *cubic* graph is a 3-regular graph.

Walks, Paths and Cycles A *walk* is a sequence of vertices $(v_0, v_1, \ldots, v_n)$ where each edge $(v_i, v_{i+1}) \in E$ for $i = 0, \ldots, n-1$. A walk is said to be *closed* if $v_0 = v_n$, and *open* otherwise. A walk is a *path* if all vertices in the sequence are distinct, that is, $v_i \neq v_j$ for all $i \neq j$. A path is a *cycle* if it is closed. The *length* of a walk, a path or a cycle is the number of edges on the walk, the path or the cycle, respectively. The *girth* of a graph is the length of a shortest cycle on the graph, excluding cycles of length two. On the other hand, the

V. S. Borkar et al., *Hamiltonian Cycle Problem and Markov Chains,*
International Series in Operations Research & Management Science 171,
DOI 10.1007/978-1-4614-3232-6_2, © Springer Science+Business Media, LLC 2012

circumference of a graph is the length of a longest cycle on the graph. The circumference of a Hamiltonian graph of order N is N, as any Hamiltonian cycle is a longest cycle of the graph.

Example 2.1 *We give examples of an open walk (Figure 2.1), an open path (Figure 2.2) and a cycle (Figure 2.3).*

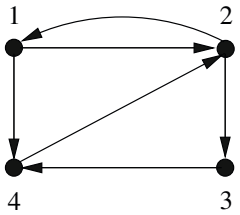

Fig. 2.1: An open walk $(1, 2, 3, 4, 2, 1, 4)$

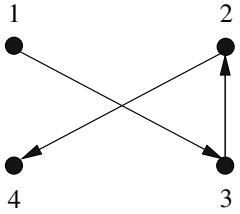

Fig. 2.2: An open path $(1, 3, 2, 4)$

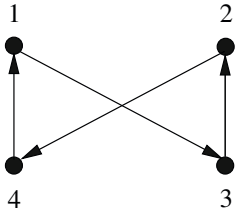

Fig. 2.3: A cycle $(1, 3, 2, 4, 1)$

Adjacency Matrices The *adjacency matrix* $\mathbf{A} = [a_{ij}]$ of a graph G has elements

$$a_{ij} = \begin{cases} 1 & \text{for } (i,j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The adjacency matrix of an undirected graph is always symmetric. Because it is a one-to-one representation of a (labelled) graph, the adjacency matrix is one of the most investigated matrices in graph theory. We can obtain much interesting information about a given graph by studying its adjacency matrix. For instance, the number of walks of length k from vertex i to vertex j is the (i, j)th entry of the matrix $\mathbf{A}^k$ [28].

Laplacian Matrices Another well-studied matrix in graph theory is the *Laplacian matrix* $\mathbf{L} = [\ell_{ij}]$ where

$$\ell_{ij} = \begin{cases} \deg(i) & \text{if } i = j, \\ -1 & \text{if } i \neq j \text{ and } (i, j) \in E, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

The degree matrix $\mathbf{D}$ of G has diagonal elements $d_{ii} = \deg(i)$ for $i = 1, \ldots, N$. It is clear that

$$\mathbf{L} = \mathbf{D} - \mathbf{A}. \tag{2.2}$$

A *normalised Laplacian matrix* is defined as $\widetilde{\mathbf{L}} = [\widetilde{\ell}_{ij}]$, where

$$\widetilde{\ell}_{ij} = \begin{cases} 1 & \text{if } i = j, \\ \dfrac{-1}{\sqrt{\deg(i)\deg(j)}} & \text{if } i \neq j \text{ and } (i, j) \in E, \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

Spectrum of a Graph The *spectrum of a graph* can refer to either the set of eigenvalues of the adjacency matrix, or of the Laplacian matrix, or of the normalised Laplacian matrix of the graph. For an excellent discussion of normalised Laplacian matrices, we refer the interested reader to Chung [24].

2.2 Fractal-like Structure of Graphs

In this section, we present an intriguing self-similar structure that groups cubic graphs according to the numbers of closed walks of various sizes in each graph. The procedure of obtaining this self-similar structure is: for each cubic graph G and its adjacency matrix $\mathbf{A}$, we evaluate the sample mean $\mu(\mathbf{A})$ and variance $\sigma^2(\mathbf{A})$ of the exponentials of all eigenvalues of $\frac{1}{3}\mathbf{A}$. As the eigenvalues of the adjacency matrix of a cubic graph are real and belong to the interval $[-3, 3]$, the eigenvalues of $\frac{1}{3}\mathbf{A}$ belong to the interval $[-1, 1]$. Thus, each cubic graph G corresponds to a point $(\mu(\mathbf{A}), \sigma^2(\mathbf{A}))$ in these

mean-variance coordinates. We illustrate this procedure with the following example.

Example 2.2 *Consider the labelled Petersen graph.*

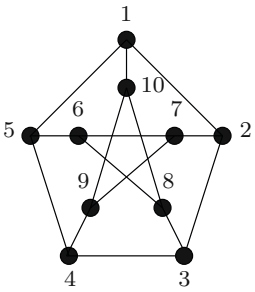

Fig. 2.4: Labelled Petersen graph

Its adjacency matrix $\boldsymbol{A}$ *is given by*

$$\mathbf{A} = \begin{bmatrix} \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & 1 & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & \cdot \end{bmatrix}, \tag{2.4}$$

where '·' denotes 0*. The eigenvalues of* $1/3\mathbf{A}$ *are* $1, -2/3, -2/3, -2/3, -2/3, 1/3, 1/3,\ 1/3, 1/3, 1/3$*. Consequently,*

$$\mu(\mathbf{A}) = \frac{1}{10}(e^1 + 4e^{-2/3} + 5e^{1/3}) = 1.1750, \tag{2.5}$$

and

$$\sigma^2(\mathbf{A}) = 0.4376, \tag{2.6}$$

and we obtain the point $(0.1750, 0.4376)$ *for the Petersen graph.*

Continuing the procedure of obtaining the self-similar structure, we then plot the mean-variance coordinates for all cubic graphs of the same order. We identify a very interesting behavior for N taking values from 10 to 18. For example, the structures for $N = 14$ and $N = 16$ appear as in, respec-

tively, Figures 2.5 and 2.6.

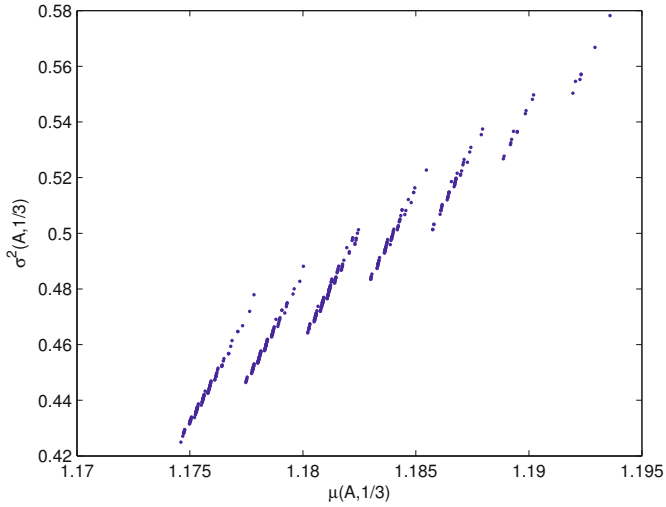

Fig. 2.5: Mean-variance plot for cubic graphs of order 14

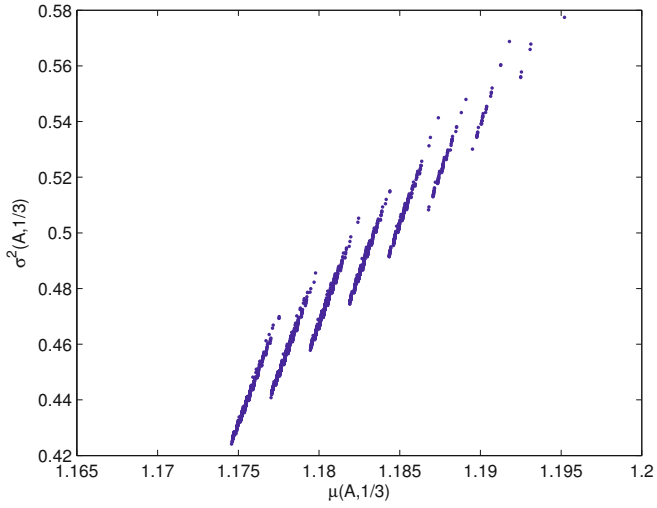

Fig. 2.6: Mean-variance plot for cubic graphs of order 16

In these two figures, it is evident that the resulting scattergram in the mean-variance coordinates consists of thread-like clusters with similar slopes of and distances between consecutive clusters. Moreover, the value of variance at the bottom of each segment is strictly increasing from left to right in the plot. Ejov *et al.* [33] use the term *multifilar* to refer to this thread-like structure, with each approximately linear cluster being called a *filar*. The authors make an important observation that the overall structure is self-similar. In particular, zooming in on each of these filars shows us similar but smaller sub-filars that are also made up of approximately straight and parallel segments, shifted gradually from left to right. We illustrate this self-similarity in Figures 2.7 and 2.8, by showing plots of two successive enlargements of the first filar of Figure 2.5.

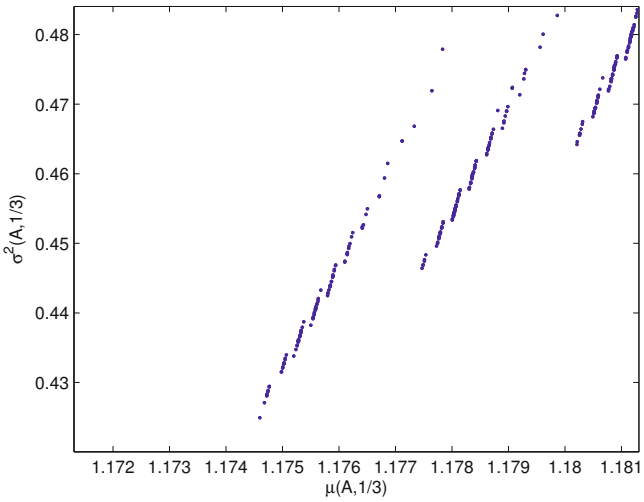

Fig. 2.7: Zooming in on the first, from the left, filar in Figure 2.5—first level

Using a form of Ihara-Selberg trace formula derived in Mnev [80], Ejov *et al.* [33] explain the filar memberships for each graph. In the overall clustering, all graphs belonging to each filar have the same number of triangles (cycles of length three) and these numbers strictly increase from the left most filar to the right most, starting from zero. In the first level of zooming-in, all graphs in a particular sub-filar have the same number of quadrangles (cycles of length four) while the number of triangles over all these sub-filars is fixed. This pattern repeats itself, with each higher level of zooming-in corresponding to a larger cycle size.

Consider another frequently used matrix function in the spectral theory of

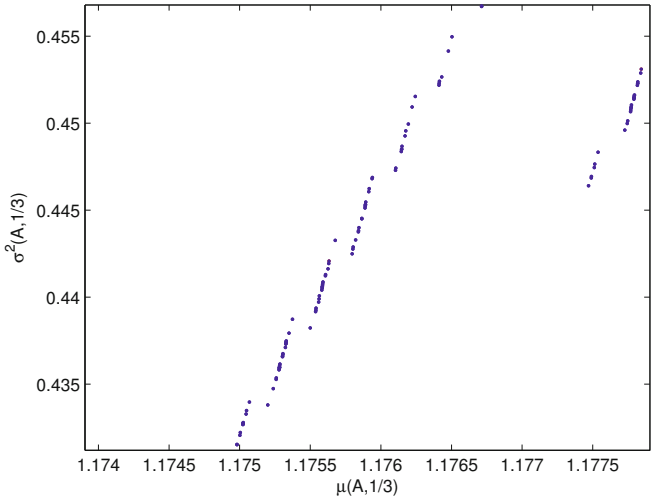

Fig. 2.8: Zooming in on the first, from the left, filar in Figure 2.5—second level

linear operators: the resolvent of $t\mathbf{A}$ for $t \in (0, 1/3)$. It appears that we can reproduce the self-similar phenomenon described above in the mean-variance coordinates with different slopes of and distances between segments [36]. In Chapter 9, we use a modification of the Ihara-Selberg trace formula [80] to justify the multifilar structure of the observed plots and to estimate the slopes of and distances between filars, consistently with numerical evidence.

It is worth noting that in the aforementioned self-similar phenomenon, non-Hamiltonian graphs seem to be separated in two groups. The first group contains *easy non-Hamiltonian* graphs that seem to be located at the tops of (the most zoomed in) sub-filars. We call a non-Hamiltonian graph *easy* if it contains one or more *bridges*. A *bridge* is an edge the removal of which disconnects the graph. A *bridge graph* is an easy non-Hamiltonian graph because these bridges, and consequently the bridge graphs, can be identified in polynomial time [72] (in fact, in linear time in N). We call a non-Hamiltonian graph that is not a bridge graph a *hard* non-Hamiltonian graph. In the self-replicating structure, *hard non-Hamiltonian* graphs (the second group) seem to be found at the bottom ends of (the most zoomed in) sub-filars. In these sub-filars, the Hamiltonian graphs are strictly in between these two groups of non-Hamiltonian graphs. To illustrate this observation, we plot the mean-variance coordinates for the trace of the matrix resolvent, over the set of all cubic graphs of order 14, in Figure 2.9, where dots represent Hamiltonian graphs and crosses represent non-Hamiltonian graphs.

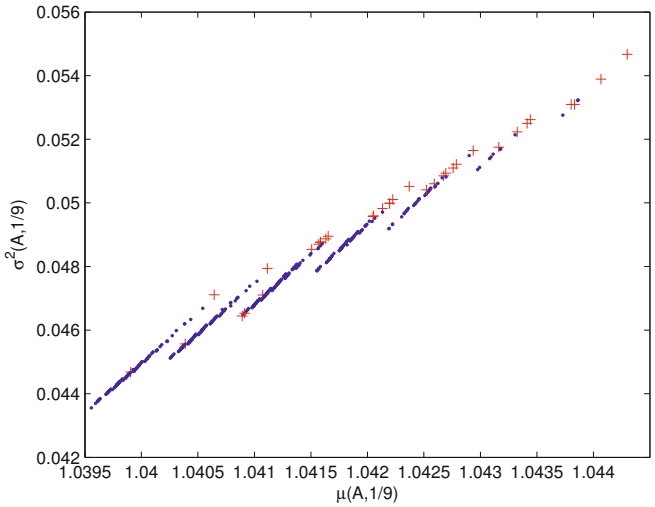

Fig. 2.9: Mean-Variance plot for cubic graphs of order 14. Here, the dots represent Hamiltonian graphs and crosses represent non-Hamiltonian graphs

We also include a zooming-in plot (see Figure 2.10) for more clarity. All crosses that can be seen clearly in this plot are either at the top or the bottom of their sub-filars.

In Chapter 9, we revisit this observation in more detail, and explore further interesting properties that arise from it.

2.3 Invariants of Graphs

A *graph invariant* is a function that maps the set of graphs to some other set, such as the set of natural numbers in such a way that certain "similar" graphs are all mapped onto the same number. Informally, we can think of a graph invariant as a numerical property associated with a graph that does not depend on the graph labelling or drawing. Examples of graph invariants include the number of vertices, the number of edges, the number of connected components, the graph spectrum, and the *chromatic number*, which is the minimum number of colours it requires to colour the vertices in a given graph in such a way that no two connected vertices share the same colour.

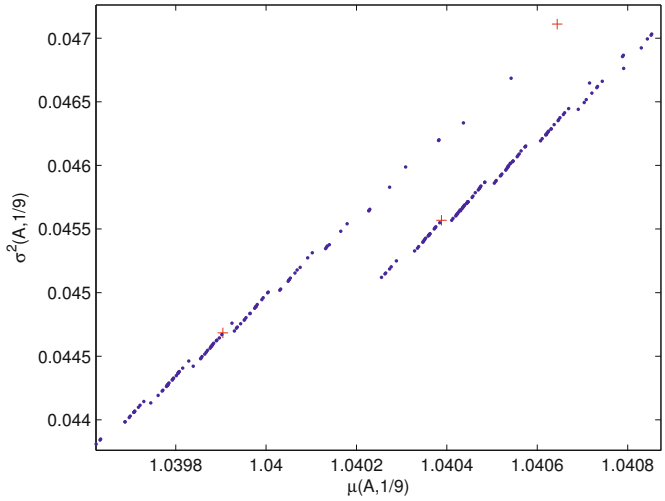

Fig. 2.10: Mean-Variance plot—zooming in. Here, the dots represent Hamiltonian graphs and crosses represent non-Hamiltonian graphs

For each graph G of order N, we introduce a set $\mathcal{F}(G)$ of matrices on G, where

$$\mathcal{F}(G) = \left\{ \mathbf{P} \in \mathbb{R}^{N\times N} \,\middle|\, p_{ij} = 0 \text{ if } (i,j) \notin E, \sum_{j=1}^{N} p_{ij} = 1, \right.$$
$$\left. p_{ij} \geq 0 \text{ for all } (i,j) \in E \right\}. \tag{2.7}$$

In Chapter 3, we show that these matrices have probabilistic interpretations. For now, we refer to $\mathcal{F}(G)$ as the set of *feasible matrices* on G, and drop the dependency on G when no confusion can arise. If p_{ij} is either 1 or 0 for all $i, j \in V$, the matrix $\mathbf{P}$ is said to be *deterministic*. Every deterministic matrix $\mathbf{P} \in \mathcal{F}(G)$ corresponds to a spanning subgraph of G that has exactly one edge coming out of each vertex.

For each $\mathbf{P} \in \mathcal{F}$, we define the matrix

$$\mathbf{W}(\mathbf{P}) = \mathbf{I} - \mathbf{P} + 1/N\mathbf{J}, \tag{2.8}$$

where $\mathbf{I}$ is the $N \times N$ identity matrix, and $\mathbf{J}$ is an $N \times N$ matrix of which every entry is unity. Next, we note that the maximum value of the determinant of $\mathbf{W}(\mathbf{P})$, over the set $\mathcal{F}$, is a graph invariant. It is in fact equal to the circumference of the graph, that is, the longest cycle of the graph. Conse-

quently, solving the Hamiltonian cycle problem for a graph G is equivalent to maximising the determinant of $\mathbf{W}(\mathbf{P})$ over the set $\mathcal{F}(G)$. In Chapter 5, we prove the above property in a more general case, where $1/N$ is replaced by any constant $\alpha \in \mathbb{R}^+$. In the same chapter, we also show that when we apply a linear, singular perturbation to the determinant functional, Hamiltonian cycles remain the maximisers over $\mathcal{F}$.

Example 2.3 *Consider a cubic graph G of order 6, which we call the* envelope graph *(Figure 2.11). One of the only two 6-vertex cubic graphs, both of which are Hamiltonian, the envelope graph has six Hamiltonian cycles.*

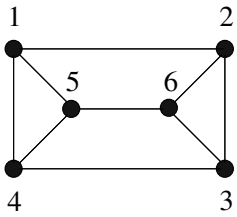

Fig. 2.11: The envelope graph

Let H_1 and H_2 be the two Hamiltonian cycles on the envelope graph, depicted in Figures 5.3 and 5.4, respectively.

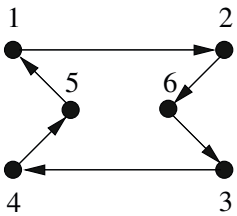

Fig. 2.12: The Hamiltonian cycle H_1

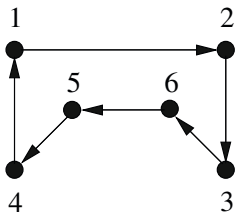

Fig. 2.13: The Hamiltonian cycle H_2

The transition matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ associated with the Hamiltonian cycles H_1 and H_2, respectively, are

$$\mathbf{P}_1 = \begin{bmatrix} \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot \end{bmatrix} \quad \textit{and} \quad \mathbf{P}_2 = \begin{bmatrix} \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot \end{bmatrix}.$$

Simple calculations give us

$$\det(\mathbf{I} - \mathbf{P}_1 + 1/6\mathbf{J}) = \det(\mathbf{I} - \mathbf{P}_2 + 1/6\mathbf{J}) = 6.$$

Let T_1 and T_2 be two subgraphs of the envelope graph, depicted respectively in Figures 2.14 and 2.15. Subgraph T_1 has one cycle, which is of length 3. Subgraph T_2 has one cycle, which is of length 4.

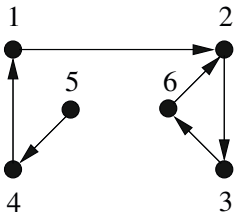

Fig. 2.14: Subgraph T_1 with one cycle

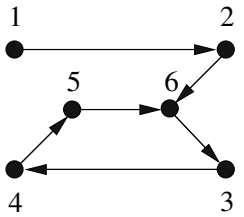

Fig. 2.15: Subgraph T_2 with one cycle

Because of their appearance, later on, we refer to subgraphs such as T_1 and T_2 as noose cycles. *The transition matrices $\mathbf{P}_3$ and $\mathbf{P}_4$ corresponding to subgraphs T_1 and T_2, respectively, are*

$$\mathbf{P}_3 = \begin{bmatrix} \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \quad \textit{and} \quad \mathbf{P}_4 = \begin{bmatrix} \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot \end{bmatrix}.$$

It is easy to check that

$$\det(\mathbf{I} - \mathbf{P}_3 + 1/6\mathbf{J}) = 3,$$

and

$$\det(\mathbf{I} - \mathbf{P}_4 + 1/6\mathbf{J}) = 4.$$

Let S_1 and S_2 be two subgraphs of the envelope, depicted in Figures 5.7 and 5.8. Each of subgraphs S_1 and S_2 has two disjoint cycles, both disjoint cycles in S_1 are of length 3, while one cycle in S_2 is of length 4 and the other is of length 2.

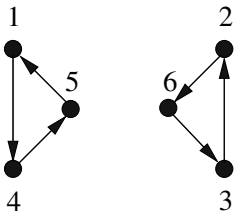

Fig. 2.16: Subgraph S_1

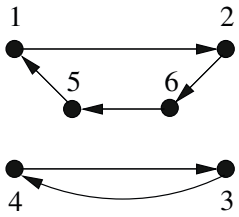

Fig. 2.17: Subgraph S_2

The transition matrices $\mathbf{P}_5$ *and* $\mathbf{P}_6$ *corresponding to subgraphs* T_1 *and* T_2*, respectively, are*

$$\mathbf{P}_5 = \begin{bmatrix} \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot \end{bmatrix} \quad \textit{and} \quad \mathbf{P}_6 = \begin{bmatrix} \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot \end{bmatrix}. \tag{2.9}$$

Again, it is easy to check that

$$\det(\mathbf{I} - \mathbf{P}_{S_1} + 1/6\mathbf{J}) = \det(\mathbf{I} - \mathbf{P}_{S_2} + 1/6\mathbf{J}) = 0. \tag{2.10}$$

This example suggests that Hamiltonian cycles, indeed, yield the maximal values of the determinant functional $\det \mathbf{W}(\mathbf{P})$*.*

In Example 2.3, it is not coincidental that the determinant functional has a value of 3 (or 4) at a subgraph containing a single cycle of length 3 (or 4), or that the determinant functional is zero at subgraphs containing two disjoint cycles. In Chapter 5, we prove that the determinant of $\mathbf{W}(\mathbf{P})$ is always k for a deterministic matrix $\mathbf{P}$ that corresponds to a subgraph containing a single cycle of length k. On the other hand, this determinant functional is always zero for a deterministic matrix $\mathbf{P}$ that corresponds to a subgraph containing two or more disjoint cycles.

In Chapter 3, we discuss the probabilistic interpretations of matrices $\mathbf{P}$, $\mathbf{W}(\mathbf{P})$, the inverse of the latter and other relevant matrices. These probabilistic interpretations exhibit the connection between the theory of Markov chains and the Hamiltonian cycle problem. Chapter 4 extends probabilistic approaches to the Hamiltonian cycle problem to include Markov decision processes. In Chapter 6, we use probabilistic arguments to prove that Hamiltonian cycles are minimisers of the trace functional of the inverse of a matrix similar to $\mathbf{W}(\mathbf{P})$.

http://www.springer.com/978-1-4614-3231-9

Hamiltonian Cycle Problem and Markov Chains
Borkar, V.S.; Ejov, V.; Filar, J.A.; Nguyen, G.T.
2012, XIV, 202 p., Hardcover
ISBN: 978-1-4614-3231-9