Spectral Approaches to the Hamiltonian Cycle Problem

Scott Weber

November 3, 2012

Contents

Chapter 1

Introduction

1.1 Hamiltonian Cycle Problem

Consider a standard eight by eight chessboard. Starting with a knight in an arbitrary square, is it possible to move the knight (in the standard L-shaped manner) so that it visits every square on the board exactly once?

Consider a dodecahedron. Pick an arbitrary face, and say we can transition to other faces sharing an edge. Is it possible, moving in this way, to visit each face exactly once?

Given a group of people, some of whom are friends, some of whom are not, is it possible to seat them about a table in such a manner that each person will be sitting next to two friends?

These problems are all instances of a famous computational problem, the

Hamiltonian Cycle Problem. The general setting for this problem is a formal *graph*, which we elaborate in the next chapter.

We note briefly that this problem, besides its intrinsic interest to graph theorists, is imbued with significance by its status as a member of the special class of problems known as *NP-complete* problems. This is class of problems is fundamental to the area of Complexity Theory. Given an algorithm for a member of this class, or in particular an algorithm for efficiently solving HCP problems, we could solve *any* NP-complete problem by converting it (in a known, efficient way) to a HCP problem. Many practical problems are members of the NP-class, and hence an algorithm for solving HCP would solve many other problems as well, including but not limited to automatically giving proofs of mathematical theorems (provided a proof of reasonable length can be given), and breaking most known cryptography schemes.

Graphs are fundamentally discrete structures, and it is a general principle that if mathematicians are presented with a discrete problem, they will try to embed it in some continous context. In our particular case, Filar et al [3] took the approach of embedding the Hamiltonian Cycle Problem within the probabilistic framework of Markov Chains, and approach that has lead to several interesting results and conjectures.

The main technical contribution of this thesis is a proof of one of these conjectures of Filar et al. Specificly, they conjectured Hamiltonian Cycles

could be characterised in terms of the spectra of matrices with certain properties. This is surprising, because for general matrices it is difficult to make statements about their entries based on their spectra, but by working in a space well matched to the HCP, we are able to prove strong results, demonstrating the power of this formulation.

In the chapter 2, we will give some background, explaining the underlying concepts and notation of graphs, Markov chains and matrices. In chapter **??** we will discuss our main result, first introducing and summarising previous work, then moving onto the proof. We will then briefly remark on some consequences and future directions. In the final chapter

Chapter 2

Background

2.1 Graphs

We begin with the standard definition of a *graph*:

Definition 1. *A* graph *is a an ordered pair* (V, E) *with* $E \subset V^2$. *The elements of* V *are called the* vertices *(singular* vertex*), and the elements of* E *are called the* edges*.*

Intuitively, we think of a graph as being a set of entities (the elements of V), and some sort of 'connection' between them. We often illustrate graphs using dots, circles, or some other small geometric figure for vertices; and using curves, originating and terminating on vertices, for edges. If we wish to distinguish vertices on some basis other than their connections, we label

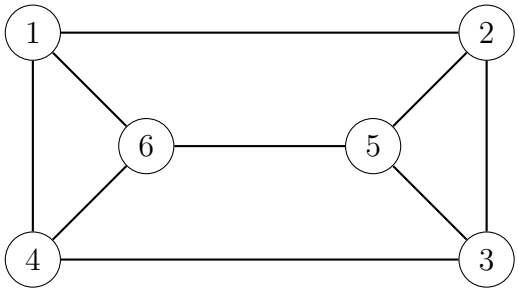

Figure 2.1: Graph with 6 vertices and 9 edges.

them, generally with the integers.

Generally we will denote a graph by $G = (V, E)$. Also, sometimes we will write $(u, v) \in G$ when we really mean $(u, v) \in E$.

The least complex variety of graphs are *simple* graphs. In these graphs, there are no edges of the form (v, v), referred to as *loops*; and there is no distinction made between the edges (v, u) and (u, v), we say the graph is *undirected*. These restrictions imply an obvious extension to *directed* graphs, graphs in which we distinguish between (v, u) and (u, v), interpreting the edge as having a direction going from the vertex v to vertex u. Diagrammatically, we depict this by adding arrows to the curves we use to depict edges.

A further extension is to associate with each edge (v, u) a real number w, referred to as the *weight*. As a motivating example, consider a electricity grid, modelled as a graph by letting substations be modelled as vertices and powerlines modelled by edges. In this case, the weight of an edge might represent the carrying capacity of the cable it represents.

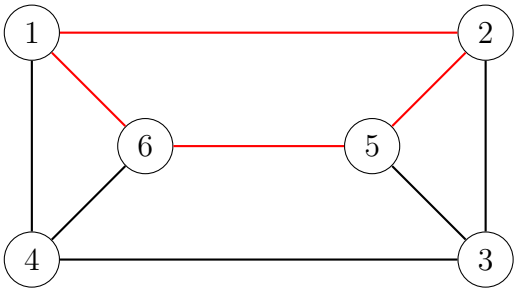

Figure 2.2: In red: the cycle $1, 2, 5, 6$

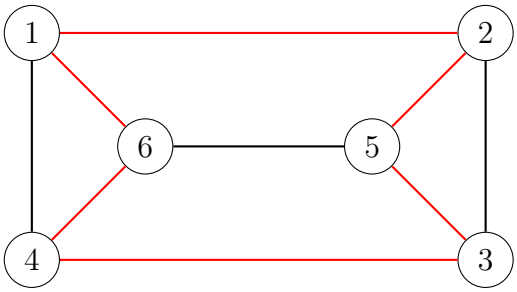

Figure 2.3: In red: the cycle $1, 2, 5, 3, 4, 6$ is Hamiltonian

Let us define a few more terms. A *walk* is a sequence of not necessarily distinct vertices $L_0, L_1, L_2, \ldots, L_k$ such each $(L_i, L_{i+1}) \in E$, that is, there is an edge between the successive vertices in the sequence. A *path* is walk where all the vertices are distinct. A *cycle* is a sequence of distinct vertices $C_0, C_1, \ldots, C_n$ that is a path, and in addition (C_n, C_0) is an edge.

A *Hamiltonian* cycle is a cycle that visits every vertex in a graph. If there is a Hamiltonian cycle in a graph, then we say the graph itself is Hamiltonian as well; and similiarly if no Hamiltonian cycle can be given, we say the graph itself is non-Hamiltonian.

The *adjacency matrix* of a graph with N vertices is a $N \times N$ matrix

defined by:

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

As an example, the adjacency matrix of the undirected graph in Figure 2.1 is:

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

2.2 Markov Processes

A *Markov Chain* is a formalisation of a process occurring over time which in some sense only depends on its present state. An example would be a game of snakes and ladders: it doesn't matter if we got to square 12 by rolling 6 twice or 4 three times, the probability of going to square 16 is the same. A bit more formally, we say that the probability of a future state is *conditionally independent* given the present state. In notation, we have some finite state space S, we and we have a sequence of random variables $X_0, X_1, \ldots, X_n$, interpreting the random variable $X_t \in S$ as "the state at time t". The *Markov property* is:

$$\Pr(X_{n+1} = x_{n+1} | X_n = x_n, \ldots, X_0 = x_0) = \Pr(X_{n+1} = x_{n+1} | X_n = x_n). \tag{2.2}$$

We also generally want that the specific time doesn't matter, we say that the process is *stationary*:

$$\Pr(X_{n+1} = x_{n+1}|X_n = x_n) = \Pr(X_{m+1} = x_{m+1}|X_m = x_m) \quad \forall m, n. \quad (2.3)$$

A finite markov process can naturally be represented by a matrix, called the *transition* matrix. Given that we have a finite number of states, we might as well represent them by the integers $1, ..., N$. Then we can define a matrix P by $\mathbf{p}_{ij} = \Pr(X_{t+1} = j|X_t = i)$, where we are using the Markov Chain theory convention that capitals represent matrices and the subscripted row vectors and entries are written in lower case boldface. This is a very nice representation, because if we define a vector x_t by $(x_t)_i = \Pr(X_t = i)$, then we have $x_{t+1} = Px_t$. In fact this matrix represents essentially all the information about the Markov Chain, and we will focus mainly on these matrices.

The probablistic origins of the transition matrix lead to the simple but useful property:

$$\sum_i \mathbf{p}_{ij} = 1 \quad \forall i.$$

Or, more compactly:

$$P\mathbf{1} = \mathbf{1}.$$

where $\mathbf{1}$ is the all ones vector. This property is referred to as the *stochastic property*, and we also say that P is a *stochastic matrix*. Given a stochastic matrix, we can interpret it as *defining* a Markov chain, and we will generally proceed in this direction, from a stochastic matrix to a Markov chain.

We will now describe the link between Markov chains and graphs. Given

a graph, consider its adjacency matrix A. We can use this matrix to define a natural class of stochastic matrices, which satisfy:

$$P \leq A$$

$$P\mathbf{1} = \mathbf{1}.$$

or more explicitly:

$$\sum_{i=1}^{N} \mathbf{p}_{ij} = 1 \quad \forall j$$

$$\mathbf{p}_{ij} \in [0, 1] \quad \forall i, j$$

$$\mathbf{p}_{ij} = 0 \text{ if } (i, j) \notin G.$$

We say that the space of such stochastic matrices is the space of matrices *induced* by the graph. We will sometimes denote this space by $\mathcal{P}(G)$.

We can naturally visualise a Markov chain where the states are the vertices, and the actions correspond to directed, weighted edges. Let us consider an example, again on the graph in Figure 2.1. A possible matrix induced by this graph might be:

$$\begin{bmatrix} 0 & 0.5 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0.5 & 0 & 0 \end{bmatrix}.$$

Which has the natural visualisation as Figure 2.4.

An important special class of stochastic matrices are the *doubly-stochastic*

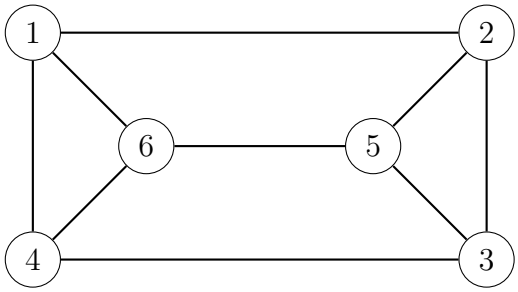

Figure 2.4: Visualisation of a stochastic matrix induced by a graph

matrices. In addition to satisfying $P\mathbf{1} = \mathbf{1}$, they satisfy the additional constraint:

$$\mathbf{1}^T P = \mathbf{1}^T.$$

Or, alternatively:

$$\sum_{j=1}^{N} \mathbf{p}_{ij} = 1 \quad \forall i.$$

Note that the previously given matrix also satisfies this constraint. Again, we may consider those doubly stochastic P that satisfy $P \leq A$ where A is the adjacency matrix of a graph G, and we will refer to these as the *doubly-stochastic matrices induced by a graph*, sometimes denoting them by $\mathcal{DS}(G)$.

These doubly-stochastic matrices are of interest because they have an elegant decomposition into *permutation* matrices, which are doubly-stochastic matrices with only zero-one entries, which intuitively do not "randomize". The theorem guaranteeing this decomposition exists is known as the Birkhoff-von Neumann theorem, or just the Birkhoff theorem, an exposition and proof of which can be found in most linear algebra texts covering non-negative matrices [1].

Theorem 1. *Given a doubly stochastic matrix P, there exists* $\alpha_i \in [0, 1]$ *and*

permutation matrices B_i such that:

$$\sum_i \alpha_i = 1$$

$$\sum_i \alpha_i B_i = P.$$

Another way of stating this theorem is to say that "the doubly stochastic matrices are the convex hull of the permutation matrices".

Important to our case is that if the matrix P is in the set of doubly-stochastic matrices induced by a graph, that is, $P \in \mathcal{DS}(G)$, then each of the B_i are also in $\mathcal{DS}(G)$.

There is also a natural graphical interpretation of this result, which will make clear the link to Hamiltonian cycles. The permutation matrices can be interpreted as directed *cycle covers*. If we visualise these matrices as previously described, we can see that the graph is decomposed into a set of directed cycles. It is now natural to perceive the graph of a doubly-stochastic matrix as a weighted sum of cycle covers.

Again, an example will prove illuminating. Consider again our graph from Figure 2.1, and the doubly-stochastic matrix we previously gave. This matrix can be decomposed as:

$$
\begin{bmatrix}
0 & 0.5 & 0 & 0 & 0 & 0.5 \\
0 & 0 & 0.5 & 0 & 0.5 & 0 \\
0 & 0 & 0 & 0.5 & 0.5 & 0 \\
0.5 & 0 & 0 & 0 & 0 & 0.5 \\
0 & 0.5 & 0.5 & 0 & 0 & 0 \\
0.5 & 0 & 0 & 0.5 & 0 & 0
\end{bmatrix}
$$

$$
= 0.5 \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0
\end{bmatrix} + 0.5 \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.
$$

which in turn gives the visualisation of Figure 2.5.

The attentive reader might notice that the graph of the second permutation is a Hamiltonian cycle. We call a permutation matrix that corresponds to a Hamiltonian cycle in this way a Hamiltonian matrix. That Hamiltonian matrices are included in the continuous domain of doubly-stochastic matrices is what allows the our embedding. If we could find a simple condition that characterised Hamiltonian matrices, then we could hopefully use it as a guide to some continuous optimisation procedure, moving through the space of doubly-stochastic matrices to a Hamiltonian cycle.

The principle contribution of this thesis will be a proof of a characterisation of Hamiltonian matrices in terms of their spectra. This will lead to a continuous, differentiable objective function that we can optimise to try and find Hamiltonian cycles in a graph. Before we get our hopes up too much, it turns out that this objective function is non-linear and non-convex, making such a scheme impractical in general; but nonetheless this characterisation of Hamiltonian cycles in continuous terms is interesting, and we hope it will

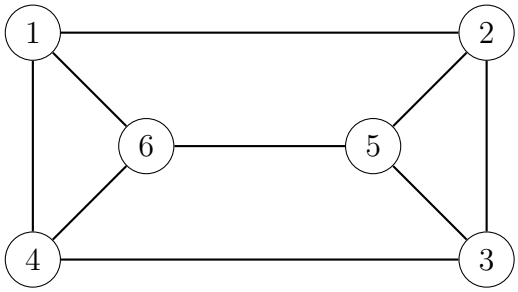

Figure 2.5: Visualisation of a decomposition of a stochastic matrix

lead to further insight into the general problem.

2.3 Vector and Matrix Norms

2.3.1 p-norms

Given a N dimensional vector $\mathbf{x}$, we can define its p-norm, $||\mathbf{x}||_p$, is defined by:

$$||\mathbf{x}||_p = (\sum_{i=1}^{N} |\mathbf{x}|^p)^{\frac{1}{p}}. \tag{2.4}$$

The most common norms are the 2-norm, the 1-norm and the ∞-norm. The 2-norm is just the ordinary euclidean norm. The 1-norm is sometimes referred to as the 'manhattan' norm [1]. The ∞-norm is also known as the 'maximum' norm, because it equals the magnitude of the largest element of the vector, indeed this is usually taken as its definition, though one can see easily enough that taking the limit as p goes to infinity yields the same result. We will mainly use the first two norms, the 2-norm and the 1-norm.

We will use a few useful inequalities relating different p-norms.

[1] The explanation for this terminology being that the 1-norm is the distance one would have to travel between two points if restricted to a rectangular grid, resembling the streets of Manhattan

Lemma 1. *For any N-dimensional vector* $\mathbf{x}$

$$||\mathbf{x}||_2 \leq ||\mathbf{x}_1||.$$

Proof. As we are only have non-negative terms, it suffices that the set of inequalities follow inclusions in the set of terms for each norm.

$$\begin{aligned}
||\mathbf{x}||_1^2 &= (\sum_{i=1}^{N} |\mathbf{x}_i|)^2 \\
&= \sum_{i=1}^{N} \sum_{i=1}^{N} |\mathbf{x}_i||\mathbf{x}_j| \\
&= \sum_{i=1}^{N} |\mathbf{x}_i|^2 + \sum_{i \neq j} |\mathbf{x}_i||\mathbf{x}_j| \\
&\geq \sum_{i=1}^{N} |\mathbf{x}_i|^2 \\
&= ||\mathbf{x}||_2^2.
\end{aligned}$$

□

Note that if we desire equality in above, we will need that $\sum_{i \neq j} |\mathbf{x}_i||\mathbf{x}_j| = 0$. This will obtain only if there is exactly one non-zero entry in $\mathbf{x}$.

Lemma 2. *Let* $N \in \mathbb{N}_+$ *and* $\lambda_k = e^{\frac{2\pi i k}{N}}, k = 1, \ldots, N$. *Then*

$$\prod_{k=1}^{N} \lambda_k = \pm 1.$$

Proof. Intuitively, this is the case because all the complex factors come in conjugate pairs, because they are the roots of a equation with real coefficients,

$1 - \lambda^N = 0$. Formally we evaluate the product as a sum in the exponent:

$$\begin{aligned}
&\prod_{k=1}^{N} e^{\frac{2\pi i k}{N}} \\
=&e^{\frac{2\pi i}{N}\sum_{k=1}^{N} k} \\
=&e^{\frac{2\pi i}{N}\frac{1}{2}N(N+1)} \\
=&e^{\pi i(N+1)} \\
=&\pm 1.
\end{aligned}$$

□

2.3.2 Frobenius norm

Just as we can define norms for vectors, we can also define them for matrices. Because matrices have additional structure beyond that of vectors, we can define more kinds of norms of matrices, that reflect this structure to different extents. One of the most widely used, and the one we will use here, is the *Frobenius norm*, defined in terms of the entries of a matrix. For an $M \times N$ matrix A, we write:

$$||A||_F = \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} A_{ij}^2}. \tag{2.5}$$

While this might not seem to capture much of interest about a matrix, we can rewrite it in the following way:

$$\begin{aligned} ||A||_F &= \sqrt{\mathbf{tr}(AA^T)} \\ &= \sum_{i=1}^{\min\{M,N\}} \sigma_i^2. \end{aligned}$$

Where the σ_i are the much studied *singular values* of A, defined as the square roots of the eigenvalues of AA^T.

2.4 Non-negative Matrix Theory

The theory of general non-negative matrices in based on the celebrated Perron-Frobenius theorem, which in its simplest form is:

Theorem 2. *Suppose A is a non-negative, irreducible matrix. Then there is a positive, real number r that is an eigenvalue of A, such that for all eigenvalues λ, $|\lambda| \leq r$, and the associated eigenvector of r is non-negative.*

Of course, if we are working in the domain of stochastic matrices, this does not tell us anything we aren't already assuming: the r is just $\mathbf{1}$, by the definition of stochasticicity. Nonetheless the general theory that has been worked out on basis of this theorem will turn out to be useful to us.

We now define a few terms. A stochastic (or more generally non-negative) matrix is said to be *irreducible* if for all i, j, there exists some k such that $(P^k)_{ij} > 0$, that is to say, there is some non-zero probability of transitioning between every pair of states eventually. The *period* of a stochastic matrix is defined to be $h = \gcd\{k \,|\, (P^k)_{ii} > 0 \,\forall i\}$.

With these definitions we can state the following theorem of non-negative matrices, specialised to our stochastic case:

Theorem 3. *Let A be a irreducible stochastic matrix with period h. Then A has exactly h eigenvalues located on the boundary of the unit disc, and these all take the form $e^{\frac{2\pi in}{h}}$.*

Note that this is tantalisingly close to the result we are seeking. The eigenvalues mentioned are roots of unity, just as in the condition of our conjecture. The implication in the theorem goes both directions, so the theorem allows us to say:

Corrolary 1. *Suppose A is a $N \times N$ irreducible stochastic matrix with eigenvalues $e^{\frac{2\pi ik}{N}} \quad k = 1, \ldots, N$. Then A has period N.*

It would remain then to connect the period to hamiltonicity. This seems reasonable: the period if defined in terms of cyclic walks. One way of stating the preceding corrollary might be "if a stochastic matrix has eigenvalues the roots of $1 - \lambda^N$, then all the cyclic walks have lengths multiples of N". It seems fairly intuitive that this is equivalent to hamiltonicity, but proving it turns out to be subtle. Seneta[**?**] uses a complex combinatorial argument to establish this result, and though this approach can be simplified somewhat (see appendix), it stills seems somehow unnecessarily complex. In the next section, we will give a proof that avoids the need for combinatorial arguments, and instead uses linear algebraic notions.

Nonetheless, this theorem imposes a restriction on the possible forms the eigenvalues on the boundary of the unit disc take: namely that they must consist of roots of unity of the same order. For example, there is no way to

have the only the eigenvalues $e^{\frac{2\pi i}{7}}, e^{\frac{2\pi i}{3}}$ on the boundary. This mean we will be able to concluded a nice corrolary of our main result later.

2.5 History of the Hamiltonian Cycle Problem

Possibly the first well-known problem of the Hamiltonian-cycle type is the *Knight's Tour* problem. This problem asks: given a chessboard and a knight on it, is it possible, moving the knight in the standard L-shaped way, to have the knight visit each square on the board exactly once. It is easy to convert this to an explicitly graph theoretic problem: one has a vertex for each square, and edges between those vertices whose corresponding squares could be reached in a single knight move.

This problem was studied by eminent Leonard Euler, who gave a number of solutions to the puzzle, described some techniques for constructing new tours from old, gave a somewhat inefficient scheme for attempting to extend a path to a tour, and another procedure for attempting to convert a Hamiltonian path to a Hamiltonian cycle. See [5] for more details.

One of the first effective heuristics for solving the Knight's Tour was Warnsdorff's Rule, named for H. C. von Warnsdorff. The heurstic simply proposes that given a path, or simply a starting point, one should select the next vertex in the cycle to minimise the number of usable neighbours that vertex has. This rule turns out to be useful in general, and on a surprisingly large set of graphs this rule can find a Hamiltonian cycle in time linear in

the number of vertices in the graph **??**.

The next famous historical Hamiltonian cycle type problem, and the one that gave the problem its name, is the Icosian Game devised by William Rowan Hamilton.

Chapter 3

Spectral Characterisation of Hamiltonian Matrices

?? In this section we will present our main result, the proof of a conjecture of Ejov and Filar. We begin by stating the result:

Theorem 4. *Suppose P is a $N \times N$ doubly stochastic matrix induced by a graph, and $e^{\frac{2\pi i}{N}}$ is an eigenvalue of P. Then P corresponds to a Hamiltonian cycle in the graph.*

This strong characterisation given only a single eigenvalue of the matrix is surprising. Our proof will build on the structural theory of non-negative matrices. In process we will establish a extension of Lemma 2.3 from [**?**, p. 52], a that could formerly only be obtained by combinatorial arguments. We will establish this lemma in a novel, using only tools from linear algebra. Indeed, we arrived at our proof of this lemma independently of Seneta, only discovering the proof in the referenced work after our success with our approach.

3.1 Previous Work

In this section we will give a brief summary of the relevant material from [**?**]. In particular, the authors prove the following fact.

Lemma 3. *Suppose P is a $N \times N$ permutation. Then P is a Hamiltonian if and only if its characteristic equation is:*

$$1 - \lambda^N = 0.$$

This fact was previously known to experts in Markov Chains, but a proof is difficult to find written down, so we include the reference for the sake of completeness.

Note that this condition is equivalent to restricting the eigenvalues to:

$$\lambda_k = e^{\frac{2\pi ik}{N}}, k = 0, \ldots, N-1.$$

Ejov and Filar made the observation that there did not seem to be any non-Hamiltonian, doubly stochastic matrices with these eigenvalues, even if the restriction to zero-one entries was removed. Consequently they conjectured:

Conjecture 1. *Let P be a $N \times N$ doubly-stochastic matrix. Then P is Hamiltonian if and only if it has the characteristic equation*

$$1 - \lambda^N = 0.$$

Furthermore, they observed that even if the hypothesis was weakened to only requiring a single eigenvalue on the boundary of the unit disc, the conclusion would still hold.

Conjecture 2. *Let P be a $N \times N$ doubly-stochastic matrix. Then P is Hamiltonian if and only if it has $e^{\frac{2\pi i}{N}}$ as an eigenvalue.*

This conjecture is noteworthy, in that it makes a strong statement about the entries of a matrix in terms of its eigenvalues, whereas most theorems make statements about the eigenvalues in terms of the entries. In general matrices very little can be said about the entries in terms of the eigenvalues, but in the doubly-stochastic context we are able to make this strong statement.

First we will discuss some preliminaries, then move on to the main argument.

3.2 Spectral Characterisation of Permutation Matrices

In this section we will prove the following lemma, which will be the main ingredient in establishing [1].

Lemma 4. *Suppose P is an $N \times N$ doubly stochastic matrix, and all the eigenvalues λ_i of P satisfy $|\lambda_i| = 1$. Then $\mathbf{p}_{ij} \in \{0, 1\}$ for all elements $\mathbf{p}_{ij}$ of P.*

Proof. We will prove this theorem using vector and matrix norms. First, we will give a condition that will ensure $\mathbf{p}_{ij} \in \{0, 1\} \quad \forall i, j$. Recall that the

entries of doubly stochastic matrices P satisfy the equations

$$\forall i, j : 0 \leq \mathbf{p}_{ij} \leq 1$$
$$\forall i : \sum_{j=1}^{N} \mathbf{p}_{ij} = 1.$$

The above imply that the 1-norm of each row vector $\mathbf{p}_i$ satisfies

$$||\mathbf{p}_i||_1 = 1.$$

Now, we have the norm inequality for each row vector:

$$||\mathbf{p}_i||_2 \leq ||\mathbf{p}_i||_1 = 1, \tag{3.1}$$

and equality holds only if the entries of $\mathbf{p}_i$ are all zero-one.

We now consider the Frobenius norm of the matrix. The Frobenius norm is the entry-wise 2-norm of the matrix, and so we have

$$||P||_F^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} |\mathbf{p}_{ij}|^2 = \sum_{i=1}^{N} ||\mathbf{p}_i||_2^2 \leq N.$$

where the inequality comes from [**??**], and equality holds if and only if we all the entries are zero-one.

The Frobenius norm can also be expressed as:

$$||P||_F^2 = \mathbf{tr}(PP^T) = \sum_{i=1}^{N} \sigma_i^2.$$

Where the σ_i are the much studied *singular values* of the underlying matrix

P, the square roots of the eigenvalues of PP^T. Here we will not need any property of these except that they are real and non-negative.

Because P is a doubly stochastic matrix, so is its tranpose; and as the product of two doubly stochastic matrices is doubly stochastic, so PP^T is doubly stochastic. Hence by the Gershgorin Circle Theorem, we have a bound on the magnitude of the eigenvalues of PP^T, namely:

$$\sigma_i^2 \leq 1 \quad \forall i.$$

We wish to ensure $\mathbf{p}_{ij} \in \{0, 1\}$, so by [3.2] we will need that $||P||_F^2 = N = \sum_i \sigma_i^2$, which can only be achieved if we equality in the previous, that is:

$$\sigma_i^2 = 1 \quad \forall i.$$

Now that we have connected the determinicty of P to the eigenvalues of PP^T, we will need a relationship between the singular values and the original eigenvalues. By standard linear algebra, we can deduce that

$$\prod_{i=1}^{N} \sigma_i^2 = \det(PP^T) = \det(P)\det(P^T) = [\det(P)]^2 = \prod_{i=1}^{N} \lambda_i^2.$$

Now we wish to exploit the assumption of our lemma, that all the eigenvalues will lie on the unit disc. Because the σ_i^2 are real and non-negative, we may write

$$\prod_i \sigma_i^2 = \prod_i |\sigma_i^2| = |\prod_i \sigma_i^2| = |\prod_i \lambda_i^2| = \prod_i |\lambda_i|^2 = 1.$$

□

To recap, the location of the eigenvalues of P on the boundary of the unit disc cause all the singular values σ_i^2 to equal one, which in turn forces the Frobenius norm to be N, equalling the 1-norm and hence making P a zero-one matrix.

We also note that this lemma establishes slightly more than we require, in that it makes strong statements about not only matrices with the

3.3 Single Eigenvalue Characterisation

We now turn to the conjecture that a matrix is Hamiltonian if and only if it has the particular eigenvalue $e^{\frac{2\pi i}{N}}$. That is, the presence of this single eigenvalue is enough to ennsure that all the other Nth roots of unity are in the spectrum, and that the matrix is Hamiltonian. For this, we need to exploit the Perron-Frobenius theorem introduced in the previous chapter.

Recall that this theorem gives a strong restriction on what values the eigenvalues might take on the boundary of the unit disc. In particular, it says that eigenvalues on the boundary of the unit disc must all be roots of unity of the same order, and that all the roots of unity of this order must also be present. This implies that if the root $e^{\frac{2\pi i}{N}}$ is present, then all $e^{\frac{2\pi ik}{N}}, k = 1, \ldots, N$ must be present. This reduces the single eigenvalue characterisation to the result of the previous.

However, we first need to address a technical concern: the Perron-Frobenius theorem is stated for irreducible matrices, while we would like our result to apply in general. We can resolve this difficulty by a partitioning argument. Because our matrix is required to be doubly-stochastic, we can find a per-

mutation C that transforms P to a block upper triangular form:

$$CPC^{-1} = \begin{bmatrix} A_1 & \dots & \dots & \dots \\ 0 & A_2 & \dots & \dots \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & A_m \end{bmatrix},$$

where each of the A_1 is an irreducible matrix. We may then apply the Perron-Frobenius theorem to each of these blocks. If the matrix is reducible and there is more than one distinct block, then the block must of be size $M \times M$, where M strictly less than N. Suppose for contradictions sake that the P had $e^{\frac{2\pi i}{N}}$ as an eigenvalue, and P is reducible. Then one of these blocks A_n which dimension strictly less than P must be the source of this eigenvalue. As this A_n is irreducible, by the Perron-Frobenius theorem it must have all the roots of $1 - \lambda^N$ as eigenvalues. But there are N of these, and an $M \times M$ matrix must have M strictly less than N eigenvalues. Hence we have a contradiction, so $\det(P - e^{\frac{2\pi i}{N}} I) = 0$ implies that the P is irreducible, and we are done.

The reason we might be interested in this single eigenvalue characterisation as it allows us to formulate the problem of seeking a Hamiltonian cycle in a graph as an optimization problem, with the doubly-stochastic graph conditions forming the constraints, and an eigenvalue condition taking the

role of the objective. We can write:

$$\begin{aligned} &\text{minimise } [\det(P - e^{\frac{2\pi i}{N}} I)]^2 \\ \text{subject to } &P\mathbf{1} = 1 \\ &\mathbf{1}^T P = \mathbf{1}^T \\ &0 \leq P \leq A. \end{aligned}$$

Where A is the adjacency matrix of the underlying graph.

The constraints are linear, which makes them ideal for an optimization approach. However, the objective function is highly non-linear and difficult to evaluate, in general it requires the evaluation of the adjugate of P, which is computationally expensive. Nonetheless, there is hope that some of this particular difficulty can be overcome by use of sophisticated linear algebraic algorithms, and this represents a possible direction for further research. However, we should note that the issue of non-global optima is likely unavoidable, owing to the underlying NP-completeness of the problem.

3.4 Eigenvectors of Permutation Matrices

In this section we will investigate the eigenvectors associated with the eigenvalues of permutation matrices. As discussed previously, the eigenvalues of a Hamiltonian matrix are the roots of $1 - \lambda^N$. It turns out the corresponding eigenvectors also take a particular form.

Bibliography

[1] Richard S. Varga *Matrix Iterative Analysis*. Prentice-Hall, New Jersey, 1994

[2] V. Ejov, J.A. Filar, S.K. Lucas, J.L. Nelson *Solving the Hamiltonian Cycle Problem with Symbolic Determinants* Taiwanese Journal of Mathematics Vol 10 No. 2 pp. 327-338 2006

[3] V.S. Borkar, V. Ejov, J. A. Filar, G. T. Nguyen Hamiltonian Cycle Problem and Markov Chains Springer 2012

[4] Seneta *Non-negative Matrices* Beans 1995

[5] Ed Sandifer *How Euler Did It: Knight's Tour* MAA Online March 2006 Available online: http://www.maa.org/editorial/euler/How%20Euler%20Did%20It%2030%20Knights%20tour

[6] Ira Pohl *A method for finding Hamilton paths and Knight's tours* Communications of the ACM Volume 10, Issue 7, July 1967 Pages 446-449 New York, USA