Elementary Graphtheoretical Proof of Frobenius’ Theorem on Non-negative Matrices

Kurt Rosenbaum, Horst Sachs *

Department of Mathematics, Ilmenau Technical University, P.O. Box 100565, 98684 Ilmenau, Germany

Abstract

Continuing O. Perron’s work of 1907 on positive square matrices, in 1912 G. Frobenius published his celebrated theorem on the eigenvalues of irreducible non-negative square matrices. His proof is purely algebraic in character.

In 1950 H. Wielandt gave another proof of Frobenius’ theorem, elegant and much shorter than the original one. Since many generalizations and other approaches have been published.

Using graphtheoretical concepts and methods the authors could further simplify certain parts of the proof. For their elementary approach, the adequate language is that of graph theory avoiding in particular the clumsiness of traditional matrix notation.

The paper is, more or less, self-contained.

Key words: Characteristic polynomial; Cycle; Eigenvalue; Index of imprimitivity (orbit); Irreducible matrix; Non-negative matrix; Strongly connected graph; Weighted directed graph

1 Introduction

Let \( U, V, W \) be non-empty sets. A matrix \( M \) is a mapping \( w \) from \( U \times V \) onto \( W \) where \( U, V \) are the sets of indices (subscripts) and \( W \) is the set of entries. An edge-weighted directed graph \( G \) – briefly: a graph – is a mapping \( w \) from \( V \times V \) onto \( W \) where \( V \) is the set of vertices and \( W \) is the set of weights. In those areas of matrix theory where \( U \) and \( V \) coincide – e.g., in

* Corresponding author.

Email address: horst.sachs@tu-ilmenau.de (Horst Sachs).
spectral theory – there is no need to distinguish between matrix and graph: these are different coinages of the same concept. Thus the matrix $M$, or graph $G$, is defined as

$$M = G = \{w(u, v) \mid u, v \in V\}.$$ 

For our purposes it is convenient to use graph theoretical terminology and notation. $V$, the set of vertices, is assumed to be non-empty and finite, and $W$ is a subset of $\mathbb{C}$, the set of complex numbers. For $u, v \in V$ and $w \in W$, $a = (u, v; w)$ is a weighted arc (directed edge) – briefly: an arc – from $u$ to $v$ with weight $w = w(u, v) = w(a)$. The “zero arc” $(u, v; 0)$ is considered a non-arc. $A(G) := \{(u, v; w) \mid w = w(u, v) \neq 0; u, v \in V\}$ is the arc set of $G$. The (non-weighted) underlying structure graph, in the narrower sense of the word, with $0, 1$ adjacency matrix, has $\{(u, v) \in V \times V \mid w(u, v) \neq 0\}$ as its (non-weighted) arc set.

A matrix (vector) or graph all of whose weights (components) are real, or non-negative, or positive, is itself called real, or non-negative, or positive, respectively.

Graph $G$ is strongly connected iff, for any pair of distinct vertices $u, v$, there is in $G$ a path from $u$ to $v$; it is easy to see that this is the case if and only if the corresponding matrix is irreducible \(^1\) (indecomposable, unzerlegbar) [3, Th. 2.2.7], [7, Th. 3.2.1], [16, Rem. 8], [29, Th. 6.2.24], [33, p. 130], [40, Th. 3.2] [50, p. 159], [60, Th. 1.17]. In a strongly connected graph with at least one arc every vertex and every arc is contained in some cycle.

The front and rear neighbourhood of a vertex $u \in V$ in $G$ are the sets

$$F(u) = \{v \in V \mid w(u, v) \neq 0\} \quad \text{and} \quad R(u) = \{v \in V \mid w(v, u) \neq 0\}.$$ 

The front (or out-)degree and the rear (or in-)degree of $u$ are

$$d_F(u) = \sum_{v \in V} w(u, v) = \sum_{v \in F(u)} w(u, v), \quad \text{and} \quad d_R(u) = \sum_{v \in V} w(v, u) = \sum_{v \in R(u)} w(v, u);$$

the extremal degrees are $d_F^{\min}, d_F^{\max}, d_R^{\min}, d_R^{\max}$.

If $G$ has at least one arc, is non-negative and strongly connected then, clearly, $d_F^{\min} > 0$ and $d_R^{\min} > 0$.

Graph $G$ is cyclically $q$-partite ($q \in \{1, 2, 3, \ldots\}$) iff its vertex set $V$ can be partitioned into non-empty subsets $V_1, V_2, \ldots, V_q = V_0; V_{q+1} = V_1$ such that every arc originating in a vertex in $V_i$ terminates in a vertex in $V_{i+1}$ ($i = 0, 1, 2, \ldots, q$) [16, p. 170], [20, § 8], [23, Chap. 13], [31, Chap. 4], [40, Def. 3.4], [61, p. 642]. The maximum number $q$ such that graph $G$ is cyclically $q$-partite is called the orbit of $G$ and denoted by $\omega$; as, trivially, every graph is cyclically 1-partite, $\omega$ is a well-defined graph invariant. It is not difficult to

\(^1\) For a comprehensive discussion of the reducibility concept for matrices and its historical roots see [53].
show that if $G$ is strongly connected then the cyclic partitioning of $V$ into parts $V_1, V_2, \ldots, V_\omega$ (up to a cyclic permutation) is unique [16, p. 172]. For non-negative strongly connected graphs the orbit coincides with the index of imprimitivity ([20, § 8], [40, Def. 3.4]) or cyclicity ([3, Def. 2.26], [60, Def. 2.23]) of the corresponding matrix.

* 

Between 1907 and 1909 O. Perron [44], [45] and G. Frobenius [18], [19] proved some theorems on positive matrices which the latter in 1912 extended to non-negative matrices [20]. We reformulate Frobenius' renowned theorem in graph-theoretical terms.

**Theorem 1** (G. Frobenius [20] (1912)) Let $G$ denote a strongly connected non-negative graph with orbit $\omega$ that has at least one arc.

**Part 1**

$G$ has a positive eigenvalue $r$ with the following properties.

(I) $r$ maximizes the absolute value of the eigenvalues of $G$;

(II) $r$ has geometric (algebraic) multiplicity 1;

(III) to $r$ there belongs a positive eigenvector (which is, by (II), up to a positive factor, unique);

(IV) $r$ is the only eigenvalue of $G$ with a non-negative eigenvector;

(V) $d^{\min}_R \leq r \leq d^{\max}_R$, \hspace{1cm} (1)

$\quad d^{\min}_F \leq r \leq d^{\max}_F$, \hspace{1cm} (2)

if in (1) or (2) one of the equality signs holds then $d_r(v) = r$ or $d_F(v) = r$, respectively, for all $v \in V$.

**Part 2**

(VI) $G$ has precisely $\omega$ eigenvalues of absolute value $r$, namely, $r \cdot e^{\mu \frac{2\pi i}{\omega}}, \mu = 0, 1, \ldots, \omega - 1$.

(VII) The collection $S$ of eigenvalues of $G$, including geometric (algebraic) multiplicities, is invariant under a rotation of the complex plane, with its center at 0, by the angle $\frac{2\pi}{\omega}$ (but, because of (VI), by no smaller angle); briefly: $S = e^{\frac{2\pi i}{\omega}} S$, $S \neq e^{\frac{2\pi i}{\omega}} S$ for $\alpha > \omega$.

\[\begin{align*}
\text{2} \text{ Frobenius formulated items (IV) and (V) in his papers [19] and [18], respectively.} \\
\text{3} \text{ We make this distinction because we shall first prove the (weaker) geometric version of Frobenius' theorem which does not need the concept of the determinant (Cf. also [56, remark on p. 11])}.
\end{align*}\]
An ingenius, now also famous, proof of Frobenius' theorem was given in 1950 by H. Wielandt [61]. A crucial role in this proof plays a function \( \delta \) which had already been used in 1942 by L. Collatz [10].

Since then Wielandt's proof – enriched by an argument due to F.R. Gantmacher 1954 (see footnote 7) – has been reproduced in many monographs and textbooks (e.g., [3], [8], [29], [31], [40], [51], [56], [60]) and generalized in various directions (see [2], [26], [32], [34], [35], [36], [41], [48], [59]). For other approaches to Perron's or Frobenius' theorem, often in connection with some generalization, see, e.g., [1, p. 480], [2], [4, pp. 264, 265], [5], [14], [15], [17], [22], [26], [30], [38], [48], [52], [57]. More information and references can be found in H. Schneider's commentary [54] on Wielandt's 1950 paper.

Already in Frobenius' work some passages (e.g. [20, p. 474]) may be interpreted as a first hint to a graphtheoretical argument. Later, proving Frobenius' theorem several authors pointed out, and more or less used as a vehicle of explanation, the correspondence between matrix and graph, however, after a brief excursion, always returned to matrix language.  

*  
The aims of this paper are

- to provide a new, entirely elementary proof of Frobenius' theorem. No matrix arithmetic (addition, multiplication, inversion or transposition) \(^5\) is used, the variational argument is applied in its simplest form. The concept of the determinant is introduced only in the last part of the proof because the algebraic multiplicity of an eigenvalue cannot be defined without it. A theorem and its significance are not really comprehended unless the proof has been reduced to its conceptually and logically simplest form.

- to rid the theory of its clumsiness caused by the index notation of classical matrix theory. This is accomplished through the use of graphtheoretical invariant (i.e., label-free) concepts and notation.

- to demonstrate that graph theory, with its inherent quasi intuitiveness, provides efficient tools also for other disciplines, wherever appropriate \(^6\).

In Section 2 we prove the geometric version of Frobenius' theorem, in Section 3 we redefine the determinant concept in graphtheoretical terms and prove the algebraic versions of items (II) and (VII). Section 4 contains the proofs of

---

\(^4\) The situation is somewhat different in the theory of primitive matrices, cf. the concluding remark of this paper.

\(^5\) Expressions as \(Gx, w(u)x, \lambda I - G\) are used as convenient abbreviations only.

\(^6\) For the classical controversy between G. Frobenius, the algebrist, and D. König, the graph theorist, about this topic ([21, last paragraph], [37, Chap. XIV, footnote on p. 240]) see H. Schneider [53].
two related theorems of Wielandt and Collatz, and in concluding a remark concerning the theory of primitive graphs (or matrices) is made.

We have endeavored to keep the presentation self-contained.

2 Proof of Frobenius’ theorem, geometric version

2.1 Proof of Part 1

Next we shall prove some propositions which, together, provide a proof of the geometric version of Part 1 of Frobenius’ theorem. A vector \( x \) on \( G \) is a mapping from \( V \) into \( \mathbb{C} \): \( x = \{ x(v) \mid v \in V \} \). If \( x \) and \( y \) are real then, as usual, \( x < y \) means that \( x(v) < y(v) \) for all \( v \in V \), etc. An eigenvalue \( \lambda \) of \( G \) and an eigenvector \( x \) belonging to \( \lambda \) are defined as a pair \( (\lambda, x) \), \( x \neq 0 \), satisfying

\[
\lambda x(u) = \sum_{v \in V} w(u, v)x(v) = \sum_{v \in E(u)} w(u, v)x(v), \quad u \in V.
\]

Wherever convenient, the last sum will be abbreviated as \( w(u)x \) and the last equation as \( \lambda x = Gx \).

Let \( \text{spec} G \) denote the spectrum of \( G \), that is the collection of its eigenvalues regarding their algebraic multiplicities.

The Collatz-Wielandt function ([10], [61])

Let \( X \) denote the set of non-negative real vectors on \( G \) distinct from the zero vector. To \( x \in X \) there corresponds uniquely a normalized vector

\[
x^* = cx \in X \quad (c > 0) \quad \text{satisfying} \quad \sum_{v \in V} x^*(v) = 1; \quad \text{let} \quad X^* = \{ x^* \mid x \in X \}.
\]

Observation 2. The set \( X^* \) is closed and bounded (immediate).

The Collatz-Wielandt function \( \delta(x) \) (\( x \in X \)) is defined by any of the three following (evidently equivalent) properties.

(i) \( \delta(x) \cdot x \leq Gx; \quad d \cdot x \leq Gx \) implies \( d \leq \delta(x) \).

(ii) \( \delta(x) \cdot x(u) \leq w(u)x \) for all \( u \in V \), and for some \( u_0 \in V \), \( x(u_0) > 0 \) and \( \delta(x) \cdot x(u_0) = w(u_0)x \).
(iii) \( \delta(x) = \min \left\{ \frac{1}{x(u)} \sum_{v \in F(u)} w(u, v)x(v) \mid u \in V, \ x(u) \neq 0 \right\} \).

**Observation 3** \( \delta(x) = \delta(x^*) \) (immediate).

**Proposition 4** ([56, p. 4], see also [6, Lemma 4.2])

\( \delta(x) \) is upper-semicontinuous on \( X^* \). \(^7\)

**Proof.** For \( x^0 \in X^* \), let \( V^+(x^0) = \{ v \in V \mid x^0(v) > 0 \} \). Let

\[ \epsilon = \frac{1}{2} \min \{ x^0(v) \mid v \in V^+(x^0) \}, \ X^*_\epsilon(x^0) = \{ x \in X^* \mid |x(v) - x^0(v)| \leq \epsilon, v \in V \} \]

and define \( \delta^+(x) = \min \left\{ \frac{1}{x(u)} w(u)x \mid u \in V^+(x^0) \right\}, \ x \in X^*_\epsilon(x^0) \).

Then

(a) \( \delta(x^0) = \delta^+(x^0) \),

(b) \( \delta^+(x) \) is continuous on \( X^*_\epsilon(x^0) \),

(c) \( \delta(x) \leq \delta^+(x) \) on \( X^*_\epsilon(x^0) \).

Let \( \{ x_k \in X^*_\epsilon(x^0) \mid k = 1, 2, 3, \ldots \} \) be any sequence such that \( \lim x_k = x^0 \).

Using (c), (b), (b) again, and (a), we obtain

\[ \limsup \delta(x_k) \leq \limsup \delta^+(x_k) = \lim \delta^+(x_k) = \delta^+(x^0) = \delta(x^0). \quad \Box \]

From Observations 2, 3 and Proposition 4 we draw the following conclusion.

**Proposition 5** \( \delta(x) \) attains a maximum value on \( X \).

Set \[ \max_{x \in X} \delta(x) = r. \quad (3) \]

---

\(^7\) A function \( f \) defined on a set \( P \) is upper-semicontinuous in \( x^0 \in P \) iff \[ \limsup f(x_i) \leq f(x^0) \] for every sequence \( \{ x_i \in P \mid \lim_{i \to \infty} x_i = x^0 \} \). Recall that if \( P \) is closed and bounded and if \( f \) is upper-semicontinuous on \( P \) then \( \max \{ f(x) \mid x \in P \} \) exists (see [56, App. C]).

In his 1950 paper Wielandt does not mention that \( \delta(x) \) may be discontinuous at the boundary of \( X^* \) ([40, p. 10], [23, Sec. 13.2]), however it still is upper-semicontinuous guaranteeing the existence of \( \max \{ \delta(x) \mid x \in X^* \} \) which is needed for Wielandt's proof.

Using a contraction argument F.R. Gantmacher [23, Sec. 13.2] (1954) provided a simple proof for the existence of \( \max \delta(x) \) on \( X \).
With \( j(v) = 1 \) for all \( v \in V \) we obtain from property (ii) of \( \delta(x) \):

\[
\delta(j) = \sum_{v \in F(u_0)} w(u_0, v) = d_F(u_0) > 0 \quad \text{implying} \quad r > 0.
\]

**The Main Lemma**

**Proposition 6** Let \( z \in X \) maximize \( \delta(x) \), i.e., \( \max_{x \in X} \delta(x) = \delta(z) = r > 0 \).

Then \( rz(u) = w(u)z, u \in V \), i.e., \( rz = Gz \);

\[
(4)
\]

\[
(5)
\]

**Proof.** By property (ii) of \( \delta(x) \), \( V \) decomposes into

\[
V'_z = \{ u \in V \mid rz(u) = w(u)z \} \neq \emptyset,
\]

\[
(6)
\]

\[
V''_z = \{ u \in V \mid rz(u) < w(u)z \}.
\]

Assume that (4) does not hold. Then \( V''_z \neq \emptyset \). By the strong connectedness of \( G \), there is an arc from some vertex \( u' \in V'_z \) to some vertex \( u'' \in V''_z \), i.e., \( u'' \in F(u') \). Choose \( \varepsilon > 0 \) such that \( 0 < r \varepsilon < w(u'')z - rz(u'') \) and define \( z_\varepsilon \in X \) by

\[
z_\varepsilon(u) = z(v) \text{ for } v \in V - \{u''\}, \quad z_\varepsilon(u'') = z(u'') + \varepsilon.
\]

Then

\[
w(u)z_\varepsilon \geq w(u)z \geq rz(u) = rz_\varepsilon(u) \text{ for all } u \in V - \{u''\},
\]

\[
w(u')z_\varepsilon = w(u')z + \varepsilon w(u', u'') = rz(u') + \varepsilon w(u', u'') > rz(u') = rz_\varepsilon(u'),
\]

\[
w(u'')z_\varepsilon \geq w(u'')z > rz(u'') + r\varepsilon = rz_\varepsilon(u'').
\]

These inequalities imply \( |V'_z| < |V'_z| \) and \( rz_\varepsilon \leq Gz_\varepsilon \).

By property (i) of \( \delta(x) \) and (3), \( r \leq \delta(z_\varepsilon) \leq r \), thus \( \delta(z_\varepsilon) = r \).

Repeating this operation we eventually obtain, for some \( \tilde{z} \in X, \delta(\tilde{z}) = r \) and \( V'_z = \emptyset \), contradicting (6). This proves (4).

Assume that (5) does not hold. Then \( rz(u_0) = \sum_{v \in F(u_0)} w(u_0, v)z(v) \) for some \( u_0 \in V \) with \( z(u_0) = 0 \). Because of \( z \geq o \), then, necessarily, \( z(v) = 0 \) for all \( v \in F(u_0) \). \( G \) being strongly connected, this implies \( z(v) = 0 \) for all \( v \in V \), i.e., \( z = o \), contradicting \( z \in X \). \( \square \)
Lemma 7 (Main Lemma) Assume that for some \( x_0 \in X \), \( r x_0 \leq G x_0 \). Then \( r x_0 = G x_0 \) and \( x_0 > 0 \).

**Proof.** By property (i) of \( \delta(x) \), \( r \leq \delta(x_0) \leq \max_{x \in X} \delta(x) = r \) implying \( \delta(x_0) = \max_{x \in X} \delta(x) \). The assertion now follows from Proposition 6. \( \square \)

**Proof of (I)–(V)**

For any complex vector \( x \), define \( x^a \) by \( x^a(v) = |x(v)| \), \( v \in V \).

**Proposition 8** \( r = \max_{\lambda \in \text{spec } G} |\lambda| \).

**Proof.** \( \lambda x = G x \) implies \( |\lambda|^a x^a \leq G x^a \), thus, by property (i) of \( \delta(x) \),
\[ |\lambda| \leq \delta(x^a) \leq r. \] \( \square \)

**Proposition 9** Let \( y \) be a real non-negative eigenvector of \( G \) belonging to the eigenvalue \( s : G y = s y \), \( y \in X \). Then \( s = r = \max_{\lambda \in \text{spec } G} |\lambda| \), \( y > 0 \).

**Proof.** For some \( u_0 \in V \), \( y(u_0) > 0 \), thus \( s = \frac{1}{y(u_0)} w(u_0) y \geq 0 \). By Proposition 8, \( s \leq r \). By Proposition 6, there is a \( z > 0 \) such that \( G z = r z \). Let \( x = z - \varepsilon y \) with \( \varepsilon > 0 \) such that \( x > 0 \). Then \( G x = G z - \varepsilon G y = r z - \varepsilon s y = r x + \varepsilon (r - s) y \geq r x \). By the Main Lemma 7 this implies \( G x = r x \), thus \( \varepsilon (r - s) = 0 \), \( s = r \), \( G y = r y \) and, again by Lemma 7, \( y > 0 \). \( \square \)

**Proposition 10** To the eigenvalue \( r \) there belongs only one linearly independent eigenvector: \( r \) has geometric multiplicity 1.

**Proof.** Let vector \( x \neq 0 \) satisfy \( r x = G x \). Then \( r x^a \leq G x^a \). By Lemma 7, \( x^a \) is a positive eigenvector belonging to \( r \). Thus \( x(v) \neq 0 \) for all \( v \in V \).

Assume there are two linearly independent eigenvectors \( x_1, x_2 \) both belonging to \( r \). Then, with an arbitrary \( v_0 \in V \), \( x = x_2(v_0)x_1 - x_1(v_0)x_2 \) is an eigenvector satisfying \( \hat{x}(v_0) = 0 \), a contradiction. \( \square \)

From now on, the symbols \( r \) and \( z \) will be reserved for the main root, that is the positive eigenvalue maximizing \( |\lambda| \) (\( \lambda \in \text{spec } G \)), and the main vector, that is the unique normalized positive eigenvector belonging to \( r \).

---

8 Cf. [61, p. 644].

9 For positive matrices see [19, p. 514].
Proposition 11 ([3, Th. 2.2.35]; for positive matrices: [18, p. 476])

\[ d_{R}^{\min} \leq r \leq d_{R}^{\max}, \quad (8) \]
\[ d_{F}^{\min} \leq r \leq d_{F}^{\max}, \quad (9) \]

\[ r = d_{R}^{\min} \text{ or } r = d_{R}^{\max} (r = d_{F}^{\min} \text{ or } r = d_{F}^{\max}) \text{ implies } d_{R}(v) = r \text{ (} d_{F}(v) = r \text{) for all } v \in V. \]

Proof. 1) From \( r z(u) = \sum_{v \in V} w(u, v)z(v) \quad (u \in V) \) and \( \sum_{v \in V} z(v) = 1 \) we obtain

\[ r = \sum_{u \in V} r z(u) = \sum_{u \in V} \sum_{v \in V} w(u, v)z(v) = \sum_{v \in V} \left( \sum_{u \in V} w(u, v) \right) z(v) \]
\[ = \sum_{v \in V} d_{R}(v)z(v) = \bar{d}_{R} \]

where \( \bar{d}_{R} \) is a mean value between \( d_{R}^{\min} \) and \( d_{R}^{\max} \). Thus (8) holds.

Assume \( r = d_{R}^{\min} \). Then

\[ r = \sum_{v \in V} d_{R}(v)z(v) \geq \sum_{v \in V} d_{R}^{\min}z(v) = d_{R}^{\min} = r. \]

Since \( z(v) > 0 \) for all \( v \in V \), this implies \( d_{R}(v) = d_{R}^{\min} = r \) for all \( v \in V \).

Analogously for \( r = d_{R}^{\max} \).

2)\(^{10}\) Let \( u_{0} \) minimize and \( u^{0} \) maximize \( z(u) \). We have

\[ d_{F}^{\min} \leq d_{F}(u_{0}) = \sum_{v \in V} w(u_{0}, v) \leq \sum_{v \in V} w(u_{0}, v) \frac{z(v)}{z(u_{0})} = r \]
\[ = \sum_{v \in V} w(u^{0}, v) \frac{z(v)}{z(u^{0})} \leq \sum_{v \in V} w(u^{0}, v) = d_{F}(u^{0}) \leq d_{F}^{\max}. \]

Thus (9) holds. Assume \( r = d_{F}^{\min} \). Then

\[ r = d_{F}^{\min} \leq d_{F}(u_{0}) = \sum_{v \in F(u_{0})} w(u_{0}, v) \leq \sum_{v \in F(u_{0})} w(u_{0}, v) \frac{z(v)}{z(u_{0})} = r \]

\(^{10}\) Of course, (9) is an immediate consequence of (8) if the characteristic polynomial is available and transposition is applied, but we can easily do without these concepts.
implying \( z(v) = z(u_0) \) for all \( v \in F(u_0) \) and \( d_F(u_0) = d_{\text{min}} = r \). By the strong connectedness of \( G \), \( d_F(v) = d_{\text{min}} = r \) for all \( v \in V \).
Analogously for \( r = d_{\text{max}} \). □

Part 1 of Frobenius’ theorem (geometric version) is now proved.

2.2 Proof of Part 2

Proposition 12 ¹¹ Let \( G \) be a strongly connected graph with vertex set \( V \) and orbit \( \omega \); let \( m \) be a positive integer. The following statements are equivalent.

(i) The length of any closed walk in \( G \) is a multiple of \( m \).

(i') The length of any cycle in \( G \) is a multiple of \( m \).

(ii) For any pair of vertices \( u, v \in V \), the lengths of all walks from \( u \) to \( v \) belong to the same residue class modulo \( m \).

(iii) \( G \) is cyclically \( m \)-partite.

(iv) \( m \) is a divisor of \( \omega \).

Proof. \((i)\Rightarrow(i')\Rightarrow(i)\): immediate.

\((i)\Rightarrow(ii)\): Let \( W_1,W_2 \) be walks from \( u \) to \( v \) and \( W^* \) a walk from \( v \) to \( u \) with lengths \( l_1, l_2, l^* \), respectively. Then \( W_1W^* \) and \( W_2W^* \) are closed walks. By (i): \( l_1 + l^* \equiv l_2 + l^* \equiv 0 \mod m \) implying \( l_1 \equiv l_2 \mod m \).

\((ii)\Rightarrow(iii)\): Fix some vertex \( v_0 \in V \). By (ii), every vertex \( v^* \in V \) determines a residue class \( l^* \mod m \) where \( l^* \) is the length of some walk from \( v_0 \) to \( v^* \). This yields a partitioning \( \{V'_1, V'_2, \ldots, V'_m = V'_0\} \) of \( V \). Clearly, any arc that originates in a vertex of \( V'_\mu \) necessarily terminates in a vertex of \( V'_{\mu+1} (\mu = 1,2,\ldots,m) \), i.e., \( G \) is cyclically \( m \)-partite with parts \( V'_1, V'_2, \ldots, V'_m = V'_0 \).

\((iii)\Rightarrow(iv)\Rightarrow(i)\): immediate. □

Proposition 13 ¹² Let \( G \) be a strongly connected cyclically \( m \)-partite graph with parts \( V_1, V_2, \ldots, V_m = V_0 \) and let \( \varepsilon = e^{2\pi i/m} \). If \( \lambda \) is any eigenvalue of \( G \) then all the numbers \( \varepsilon^{\mu} \lambda (\mu = 0,1,\ldots,m-1) \) are eigenvalues of \( G \), each having the same geometric multiplicity as \( \lambda \).

¹¹ Cf. [16, Remarks 1,3,7], [7, Lemma 3.4.1].
¹² Cf. [46], [47].
Corollary 14 If in particular $G$ is non-negative then, together with the main root $\tau$, the $\omega$ numbers $r \cdot e^{\frac{\beta n \pi i}{\omega}} \ (\mu = 0, 1, \ldots, \omega - 1)$ are eigenvalues of $G$, each with geometric multiplicity 1.

Proof of Proposition 13 Let $x$ be an eigenvector belonging to $\lambda$. Then

$$\lambda x(u) = \sum_{v \in V_{\mu}} w(u, v)x(v), \quad u \in V_{\mu-1} \ (\mu = 1, 2, \ldots, m).$$

Let $\tau$ be the Transformation from $x$ to $x'$ defined by $x'(v) = e^{\mu}x(v), \ v \in V_{\mu} \ (\mu = 1, 2, \ldots, m)$. For $u \in V_{\mu-1} \ (\mu = 1, 2, \ldots, m)$ we obtain

$$\varepsilon \lambda \cdot x'(u) = e^{\mu} \cdot \lambda x(u) = e^{\mu} \sum_{v \in V_{\mu}} w(u, v)x(v) = \sum_{v \in V_{\mu}} w(u, v)x'(v), \ u \in V_{\mu-1},$$

i.e., $\varepsilon \lambda$ is an eigenvalue of $G$. $\tau$ and $\tau^{-1}$ preserving linear independence, $\lambda$ and $\varepsilon \lambda$ have the same geometrical multiplicity.

The assertion is proved by iterating $\tau$. $\square$

Proposition 15 Let $G$ be a strongly connected non-negative graph with orbit $\omega$ and assume that $\lambda = r \eta$ with $|\eta| = 1$ is an eigenvalue of $G$. Then $\eta$ is an $\omega$th root of unity.

Proof. Let $\eta = e^{i\varphi} \ (0 \leq \varphi < 2\pi)$ and let $x$ be an eigenvector belonging to $\lambda$. From $\lambda x = r e^{i\varphi}x = Gx$, we obtain $|\lambda|^a = r x^a \leq G x^a$, which by the Main Lemma 7 implies $r x^a = G x^a$.

By Proposition 10, $x^a = c z$ with some constant $c > 0$. We may w.l.o.g. assume $c = 1$ and obtain $x^a(v) = |x(v)| = z(v), \ v \in V$, therefore $x(v) = \zeta(v)z(v)$ where $|\zeta(v)| = 1, \ v \in V$ and, for $u \in V$,

$$r \eta \cdot \zeta(u)z(u) = \lambda x(u) = \sum_{v \in F(u)} w(u, v)x(v) = \sum_{v \in F(u)} w(u, v)\zeta(v)z(v) \quad (10)$$

implying $r \eta \cdot \sum_{v \in F(u)} w(u, v)z(v)\zeta(v) = rz(u) = \sum_{v \in F(u)} w(u, v)z(v), \ u \in V$

where all terms $w(u, v)z(v)$ are positive. This is possible only if $\zeta(v)$ is the same for all $v \in F(u)$. (10) now yields $\eta \cdot \zeta(u) = \zeta(v)$ whenever $v \in F(u)$, i.e., if there is an arc from $u$ to $v$ then $\zeta(u) = \eta \cdot \zeta(v)$. Consider any closed walk of length $l$ passing through some vertex $u_0$. Then $\zeta(u_0) = \eta^l \zeta(u_0)$, thus

$$\eta^l = e^{i\varphi l} = 1, \ l \varphi = k \cdot 2\pi, \ k \in \{0, 1, \ldots, l - 1\}, \varphi = \frac{k}{l} \cdot 2\pi, \ 0 \leq \frac{k}{l} < 1.$$

11
Let \( \frac{p}{q} = \frac{p}{q}, 0 \leq p < q, \) where \( p \) and \( q \) are relatively prime. Then \( \eta = e^{\frac{2\pi i}{q}} \) is a primitive \( q^{th} \) root of unity and \( l \) is a multiple of \( q \).
We conclude that the length of any closed walk in \( G \) is a multiple of \( q \), thus, by Proposition 12 ((i) \( \Rightarrow \) (iv)), \( q \) is a divisor of \( \omega \). \( \square \)

Part 2 (geometric version) now follows from Propositions 13, 15 and Corollary 14.

3 Proof of the algebraic version of (II) and (VII)

3.1 Determinant and characteristic polynomial of a graph

A linear graph \( L \) is a graph all of whose vertices have front degree and rear degree equal to 1. Let \( n = n(G), L_i(G) \) and \( c(L) \), \( c(L) \) denote the number of vertices of \( G \), the set of linear subgraphs \( L \) on \( i \) vertices of \( G \) and the set and the number of components (cycles) of \( L \), respectively. From linear algebra we take the concept of the determinant with its Leibniz definition as a sum extended over a set of permutations. To each of these permutations with its cycle decomposition there corresponds in \( G \) a linear subgraph \( L \in L_n(G) \) with the analog cycle decomposition, and the Leibniz definition takes the following form.\(^{13}\)

**Definition 16** \(^{14}\) ([7, Sec. 9.1], [9], [11], [12, Chap. 4], [24], [25])

\[
\det G = \sum_{L \in L_n(G)} \prod_{C \in c(L)} \left( \prod_{a \in A(C)} (-w(a)) \right) = (-1)^n \sum_{L \in L_n(G)} (-1)^{c(L)} \prod_{a \in A(L)} w(a).
\]

A principal minor of graph \( G \) (corresponding to a principal minor of matrix \( G \)) is an induced subgraph of \( G \). Let \( M_i(G) \) denote the set of all principal minors of \( G \) on \( i \) vertices. Note that \( M_{n-1}(G) = \{G - \{v \} | v \in V \} \).

The characteristic polynomial of \( G \) is defined as \( f_G(\lambda) = \det(\lambda I - G) \).

Clearly, \( \lambda \) is an eigenvalue of \( G \) if and only if it satisfies the characteristic equation \( f_G(\lambda) = 0 \). Using Definition 16 we obtain the coefficients theorem:

**Proposition 17** \(^{14}\)

Let \( f_G(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n \) be the characteristic polynomial of \( G \). Then

\(^{13}\)Here we use the fact that a permutation represented by a cycle is even (odd) if the length of the cycle is odd (even).

\(^{14}\)For a more detailed discussion of the determinant formula and the coefficients theorem in conventional terms including history and references see [13, Sec. 1.4].
\[
\begin{align*}
a_i &= \sum_{L \in \mathcal{L}(G)} \prod_{C \in \mathcal{C}(L)} \left( -\prod_{a \in \mathcal{A}(C)} w(a) \right) \\
&= \sum_{L \in \mathcal{L}(G)} (-1)^{c(L)} \prod_{a \in \mathcal{A}(L)} w(a) = (-1)^i \sum_{M \in \mathcal{M}_i(G)} \det M.
\end{align*}
\]

Note that \( f_G(\lambda) \) does not depend on the weight of any arc that does not lie on a cycle of \( G \) (cf. [58]).

### 3.2 The cycle key and the spectral key of a graph

Next we formulate a theorem concerning the cyclic structure of a graph which is partly present (but scattered) in the existing literature (see below) and can easily be derived from the preceding results.

Let \( G \) be a graph with cycle set \( \mathcal{C} \) and characteristic polynomial
\[
f_G(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_n.
\]

The cycle key \( k_C \) of \( G \) is the greatest common divisor of the lengths of all cycles in \( G \):
\[
k_C = \gcd \{ l(C) \mid C \in \mathcal{C} \} \quad (k_C = 0 \text{ iff } G \text{ has no cycles}).
\]

The spectral key \( k_S \) of \( G \) is the greatest common divisor of all subscripts \( i \) such that \( a_i \neq 0 \):
\[
k_S = \gcd \{ i \mid a_i \neq 0 \} \quad (k_S = 0 \text{ iff } a_1 = a_2 = \cdots = 0).
\]

**Proposition 18** For any graph \( G \),

(i) \( k_C \) is a multiple of \( \omega \),

(ii) \( k_S \) is a multiple of \( k_C \);

(iii) if \( G \) is non-negative then \( k_C = k_S \);

(iv) if \( G \) is strongly connected and has at least one arc then \( \omega = k_C \).

For (iii) and (iv) cf. [3, Theorems 2.2.27, 2.2.30], [16, Remarks 2, 16], [20, § 8], [49, pp. 161–166], [55, Th. 7].

For an example see the figure. Note that \( \omega \) and \( k_C \) do not depend on the weights.

Reversing the direction of arc \( a \) in this figure we obtain a strongly connected graph with \( \omega = k_C = 4, k_S = 0 \).

For non-negative graphs we define the key \( k \) to be the common value of \( k_C \) and \( k_S \).

**Corollary 19** ([16, Rem. 16, Th. 1], see also [49, pp. 161–166]) Let \( G \) be a non-negative graph on \( n \) vertices with key \( k \). Then \( f_G(\lambda) = h(\lambda^k) \cdot \lambda^q \) with some polynomial \( h(x) = x^d + \cdots (h(0) \neq 0) \) where \( dk + q = n \). In particular,
if $G$ is strongly connected and has at least one arc then

$$f_G(\lambda) = h(\lambda^\omega) \cdot \lambda^q \quad (h(0) \neq 0).$$

(11)

Proof of Proposition 18

(i) Immediate.

(ii) If $i$ is not a multiple of $k_C$ then $L_i(G)$ is empty, thus, according to Proposition 17, $a_i = 0$, i.e.: if $a_i \neq 0$ then $i$ is a multiple of $k_C$, implying (ii).

(iii) We shall show that if $G$ is non-negative then $k_S$ divides $k_C$. Assume that $G$ contains a cycle whose length is not a multiple of $k_S$ and let $C$ be a cycle with this property having the minimum length $l$: then $l \neq 0$, mod $k_S$. Clearly, $C \in L_l(G)$ and every $L \in L_l(G)$ consists of a single cycle because otherwise $l$ were not minimum. Therefore, the contributions of the members of $L_l(G)$ to the coefficient $a_i$ (Proposition 17) all have the same sign (the minus sign) implying $a_i \neq 0$ and therefore, by the definition of $k_S$, $l \equiv 0$, mod $k_S$ which is a contradiction.

(iv) follows immediately from Proposition 12. □

3.3 Proof of the algebraic version of (VII)

The algebraic version of (VII) is an immediate consequence of equation (11). Let $x_\sigma = \lambda_\sigma^\omega$ with some suitable $\lambda_\sigma$ ($\sigma = 1, 2, \ldots, s$) be the distinct roots of the equation $h(x) = 0$ where $x_\sigma$ has algebraic multiplicity $\mu_\sigma$. Then the numbers $\lambda_{\sigma \kappa} = \lambda_\sigma e^{\kappa \omega}$ ($\sigma = 1, 2, \ldots, s$; $\kappa = 0, 1, \ldots, \omega - 1$) are the distinct non-zero roots of the equation $f_G(\lambda) = 0$ where $\lambda_{\sigma \kappa}$ has algebraic multiplicity $\mu_\sigma$. □

3.4 The algebraic multiplicity of $r$: proof of the algebraic version of (II)

The relation "there are a path from $u$ to $v$ and a path from $v$ to $u$, or $u = v$" is an equivalence relation on $V$ partitioning $V$ into classes which are the
vertex sets of the strongly connected components, namely, the maximal strongly connected principal minors, of $G$. Let $S(G)$ denote the set of these principal minors.

**Proposition 20** The characteristic polynomial of a graph $G$ is the product of the characteristic polynomials of its strongly connected components:

$$f_G(\lambda) = \prod_{S \in S(G)} f_S(\lambda) \quad (12)$$

**Corollary 21** $\text{spec } G = \bigcup_{S \in S(G)} \text{spec } S$

**Corollary 22** If $G$ is non-negative then it has a non-negative real eigenvalue $\rho$ such that $|\lambda| \leq \rho \quad (\lambda \in \text{spec } G)$.

The proof of Proposition 20, based on Proposition 17, is straightforward by comparison of the coefficients on both sides of (12); we skip the details.

**Proposition 23** The algebraic multiplicity of the main root $r$ of a strongly connected non-negative graph $G$ is one.

**Proof.** Set $\mu = \lambda - r$. Then, because of $f_G(r) = 0$, polynomial $f_G(\lambda) = f_G(\mu + r)$ takes the form $\mu^n + b_1 \mu^{n-1} + \cdots + b_{n-1} \mu$. The assertion is that $b_{n-1} \neq 0$. We have $f_G(\mu + r) = \det((\mu + r)I - G) = \det(\mu I - (G - rI))$. According to Proposition 17,

$$b_{n-1} = (-1)^{n-1} \sum_{M' \in M_{n-1}(G-rI)} \det M' = \sum_{M \in M_{n-1}(rI-G)} \det M.$$

We shall show that $\det M = r^{n-1} + \cdots$ is positive for every $M \in M_{n-1}(rI-G)$. Deleting an arbitrary vertex $u^0 \in V$ we obtain from $G$ the principal minor $G^0$ and from $rI - G$ the principal minor $M^0 = rI^0 - G^0$ with vertex set $V^0 = V - \{u^0\}$. Let $\varphi(x) = \det(xI^0 - G^0) = f_{G^0}(x) = x^{n-1} + \cdots$. Clearly, $\varphi(x) > 0$ for large $x$. We shall show that $\varphi(x) \neq 0$ for $x \geq r$ implying $\det M^0 = \varphi(r) > 0$.

Let $S^*$ with vertex set $V^*$ be one of the strongly connected components of $G^0$ that has maximum main root $\rho = r^*$ with main vector $z^*$; then $\varphi(x) = f_{G^0}(x) \neq 0$ for $x > r^*$. We will show that $r^* < r$ implying $\varphi(x) \neq 0$ for $x \geq r$. Define vector $x$ on $G$ by

$$x(v) = z^*(v) \quad \text{for} \quad v \in V^*, \quad x(v) = 0 \quad \text{for} \quad v \in V - V^*;$$

note that $x$ is non-zero and non-negative but not positive since $x(u^0) = 0$.

Assume $r^* \geq r$. Then

It follows readily from the theory of determinants with dominant main diagonal in connection with the strong connectedness of $G$ that $\det M^0 > 0$, but we will give an elementary ad hoc proof of this fact.
\[ 0 < r x(u) \leq r^* x(u) = w(u)x \quad \text{for} \quad u \in V^*, \]
\[ 0 = r x(u) = r^* x(u) \leq w(u)x \quad \text{for} \quad u \in V - V^*, \]
i.e., \( r x \leq G x \). By the Main Lemma 7, this implies \( x > 0 \) which is a contradiction. Therefore, \( r^* < r \). \( \square \)

Frobenius’ theorem is now proved.

4 Appendix: two more theorems

For the sake of completeness we shall prove two classical results closely related to Frobenius’ theorem.

4.1 A theorem of Wielandt’s on arbitrary square matrices

**Proposition 24** \(^{16}\) Let \( G \) be a non-negative strongly connected graph with at least one arc, with vertex set \( V \), arc weights \( w(u, v) \) and main root \( r \). Let \( \tilde{G} \) be an arbitrary graph on \( V \) with complex arc weights \( \tilde{w}(u, v) \) satisfying
\[ |\tilde{w}(u, v)| \leq w(u, v), \quad u, v \in V, \]
and let \( \rho = \max \{ |\tilde{\lambda}| \mid \tilde{\lambda} \in \text{spec } \tilde{G} \} \) denote the spectral radius of \( \tilde{G} \). Then

(A) \( \rho \leq r; \)

(B) the following statements are equivalent:

(i) \( \rho = r; \)

(ii) \( \tilde{w}(u, v) = e^{i\varphi} \zeta(u) w(u, v) \zeta^{-1}(v) \)
for some angle \( \varphi \) and some vector \( \zeta \) satisfying \( |\zeta(v)| = 1, \quad v \in V; \)

(iii) \( \text{spec } \tilde{G} \) is obtained from \( \text{spec } G \) through a rotation of the complex plane around the origin by the angle \( \varphi \): \( \text{spec } \tilde{G} = e^{i\varphi} \text{spec } G. \)

**Proof.** (A) Let \( x \) be an eigenvector of \( \tilde{G} \) belonging to \( \tilde{\lambda} \in \text{spec } \tilde{G} \). Then
\[ \tilde{\lambda} x = \tilde{G} x, \quad |\tilde{\lambda}| x^a \leq \tilde{G}^a x^a \leq G x^a \quad (13) \]

\(^{16}\) Satz III in Wielandt’s 1950 paper [61, p. 645]. As Wielandt remarks, part (A) had already been proved by Frobenius [19, p. 516]. Cherubino [8, p. 140] attributes the whole of Proposition 24 to Frobenius which is incorrect.
Cf. also [30, Th. 14], [39, pp. 18–20], [42, Th. 3 and p. 361].
which by property (i) of the Collatz–Wielandt function $\delta$ (Section 2) implies $|\bar{\lambda}| \leq \delta(x^a) \leq r$.

(B) (i)⇒(ii). Let in (13) $\bar{\lambda} = \bar{\lambda}^0$ be such that $|\bar{\lambda}^0| = \max |\bar{\lambda}| = \rho = r$, i.e., $\bar{\lambda}^0 = e^{i\varphi} \cdot r$ with some $\varphi$, $0 \leq \varphi < 2\pi$. Then $|\bar{\lambda}^0|x^a = rx^a \leq Gx^a \leq Gx^a$. By the Main Lemma 7, $Gx^a = r x^a$ and $x^a = c z$ with some $c > 0$; we may w.l.o.g. assume $c = 1$. From the above we obtain $\bar{G}^a = G$, i.e., $|\bar{w}(u, v)| = w(u, v)$ $(u, v \in V)$. Thus if $w(u, v) \neq 0$ (i.e., if $v \in F(u)$) then $\bar{w}(u, v) = \gamma(u, v)w(u, v)$ with $|\gamma(u, v)| = 1$ $(u \in V, v \in F(u))$. Letting $x(v) = \zeta(v)z(v)$ where $|\zeta(v)| = 1$ $(v \in V)$ we have the following two relations.

$\lambda^0x(u) = e^{i\varphi}\zeta(u)z(u) = \sum_{v \in F(u)} \bar{w}(u, v)x(v) = \sum_{v \in F(u)} \gamma(u, v)w(u, v)\zeta(v)z(v)$

or, equivalently,

$rz(u) = \sum_{v \in F(u)} \left(e^{-i\varphi}\zeta^{-1}(u)\gamma(u, v)\zeta(v)\right)w(u, v)z(v)$, and

$rz(u) = \sum_{v \in F(u)} w(u, v)z(v)$ $(u \in V)$.

The products $w(u, v)z(v)$ being positive in all terms in the above two sums, we conclude that $e^{-i\varphi}\zeta^{-1}(u)\gamma(u, v)\zeta(v) = 1$, thus $\gamma(u, v) = e^{i\varphi}\zeta(u)\zeta^{-1}(v)$ and $w(u, v) = \gamma(u, v)w(u, v) = e^{i\varphi}\zeta(u)\zeta^{-1}(v)$ for all $u \in V, v \in F(u)$. The last equation holds also if $v \not\in F(u)$ since in this case $w(u, v) = \bar{w}(u, v) = 0$.

This was to be proved.

(ii)⇒(iii). Let $f_{\bar{G}}(x) = x^n + \bar{a}_1x^{n-1} + \cdots + \bar{a}_n$ be the characteristic polynomial of $G$. Clearly, for every cycle $C$ of length $l$, $\prod_{a \in A(C)} \bar{w}(a) = e^{il\varphi} \prod_{a \in A(C)} w(a)$, thus by the coefficients theorem (Proposition 17), $\bar{a}_\nu = e^{i\nu\varphi}a_\nu, \nu = 1, 2, \ldots, n$ implying $e^{-i\nu\varphi}f_{\bar{G}}(x) = f_{\bar{G}}(e^{-i\varphi}x)$. The assertion follows.

(iii)⇒(i): immediate. □

4.2 Collatz’ inclusion theorem

Proposition 25 (Collatz [10] (1942)). Let $G$ be a non-negative strongly connected graph with vertex set $V$, arc weights $w(u, v)$ and main root $r$, and let $x$ be a positive vector. Define graph $G^x$ by its arc weights $w^x(u, v) = x^{-1}(u)w(u, v)x(v)$. Then

(i) $G^x$ has the same structure graph as $G$, is non-negative and strongly connected,

(ii) $\text{spec} G^x = \text{spec} G$.

17In fact, Collatz proved his theorem for arbitrary non-negative square matrices, a version easily obtainable from Proposition 25.
(iii) \( d^x_{\min} = \min_{u \in V} \sum_{v \in V} x^{-1}(u)w(u, v)x(v) \)
\[ \leq r \leq \max_{u \in V} \sum_{v \in V} x^{-1}(u)w(u, v)x(v) = d^x_{\max}, \]

(iv) \( d^x_{\min} = \min_{u \in V} \sum_{v \in V} x^{-1}(v)w(v, u)x(u) \)
\[ \leq r \leq \max_{u \in V} \sum_{v \in V} x^{-1}(v)w(v, u)x(u) = d^x_{\max}. \]

Proof. (i): immediate.

(ii) For each cycle \( C \), clearly, \( \prod_{a \in A(C)} w(a) = \prod_{a \in A(C^x)} w^x(a) \). Therefore, by the coefficients theorem (Proposition 17), \( G \) and \( G^x \) have the same characteristic polynomial.

(iii) and (iv) are immediate consequences of (i), (ii) and Proposition 11. \( \Box \)

**Remark 26** With \( x = z \) we have

\[ d^z_{\min}(u) = \sum_{v \in V} z^x(u, v) = \sum_{v \in V} z^{-1}(u)w(u, v)z(v) = r \quad \text{for all} \quad u \in V. \]

**Concluding remark**

So far we have avoided any matrix calculations (powers etc.). However, it should be pointed out that matrix arithmetic, too, can appropriately be expressed in graphtheoretic terms: thus also the theory of primitive (and imprimitive) matrices where powers of matrices are considered has its analogue within graph theory (viz., the theory of strongly-connected non-negative graphs and their single or multiple step walks). As much of this is already present in the current literature ([7], [16], [27], [28], [33], [43]) we did not pursue questions of primitivity or imprimitivity in this paper.

**References**


19


[27] B.R. Heap, and M.S. Lynn, The index of primitivity of a non-negative matrix. Numerische Mathematik 6 (1964) 120–141


20


