Analysis of Complex Networks

From Biology to Linguistics

Edited by
Matthias Dehmer and Frank Emmert-Streib

WILEY-VCH Verlag GmbH & Co. KGaA

Analysis of Complex Networks

From Biology to Linguistics

Edited by
Matthias Dehmer and
Frank Emmert-Streib

Related Titles

B.H. Junker, F. Schreiber

Analysis of Biological Networks

2008

ISBN 978-0-470-04144-4

F. Emmert-Streib, M. Dehmer (Eds.)

Analysis of Microarray Data
A Network-Based Approach

2008

ISBN 978-3-527-31822-3

E. Keedwell, A. Narayanan

Intelligent Bioinformatics
The Application of Artificial Intelligence Techniques to Bioinformatics Problems

2005

ISBN 978-0-470-02175-0

F. Azuaje, J. Dopazo (Eds.)

Data Analysis and Visualization in Genomics and Proteomics

2005

ISBN 978-0-470-09439-6

Analysis of Complex Networks

From Biology to Linguistics

Edited by
Matthias Dehmer and Frank Emmert-Streib

WILEY-VCH

WILEY-VCH Verlag GmbH & Co. KGaA

The Editors

PD Dr. habil. Matthias Dehmer
Vienna University of Technology
Discrete Mathematics and Geometry
Wiedner Hauptstraße 8–10
1040 Vienna
Austria
and
University of Coimbra
Center for Mathematics
Apartado 3008
3001-454 Coimbra
Portugal

Prof. Dr. Frank Emmert-Streib
Computational Biology and Machine Learning
Center for Cancer Research and Cell Biology
School of Medecine, Dentistry and Biomedical
Sciences Queen's University Belfast
97 Lisburn Road
Belfast, BT9 7BL
UK

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data: A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de.

Printed in the Federal Republic of Germany
Printed on acid-free paper

Cover design Adam Design, Weinheim
Typesetting le-tex publishing services oHG, Leipzig
Printing Strauss GmbH, Mörlenbach
Bookbinding Litges & Dopf Buchbinderei GmbH, Heppenheim

ISBN 978-3-527-32345-6

Contents

Preface

This book, *Analysis of Complex Networks: From Biology to Linguistics*, presents theoretical and practical results on graph-theoretic methods that are used for modeling as well as structurally investigating complex networks. Instead of focusing exclusively on classical graph-theoretic approaches, its major goal is to demonstrate the importance and usefulness of network-based concepts for scientists in various disciplines. Further, the book advocates the idea that theoretical as well as applied results are needed to enhance our knowledge and understanding of networks in general and as representations for various problems. We emphasize methods for analyzing graphs structurally because it has been shown that especially data-driven areas such as web mining, computational and systems biology, chemical informatics, and cognitive sciences profit tremendously from this field.

The main topics treated in this book can be summarized as follows:

- Information-theoretic methods for analyzing graphs
- Problems in quantitative graph theory
- Structural graph measures
- Investigating novel network classes
- Metrical properties of graphs
- Aspects in algorithmic graph theory
- Analytic methods in graph theory
- Network-based applications.

Analysis of Complex Networks: From Biology to Linguistics is intended for an interdisciplinary audience ranging from applied discrete mathematics, artificial intelligence, and applied statistics to computer science, computational and systems biology, cognitive science, computational linguistics, machine learning, mathematical chemistry, and physics. Many colleagues, whether consciously or unconsciously, provided us with input, help, and support before and during the development of the present book. In particular we would like to thank Andreas Albrecht, Rute Andrade, Gökhan Bakır, Alexandru T. Balaban, Subhash Basak, Igor Bass, Natália Bebiano, Danail Bonchev, Stefan Borgert, Mieczyslaw Borowiecki, Michael Drmota, Abdol-Hossein Esfahanian, Bernhard Gittenberger, Earl Glinn, Elena Konstantinova, Dmitrii Lozovanu, Alexander Mehler, Abbe Mowshowitz, Max Mühlhäuser, Arcady Mushegian, Paolo Oliveira, João da Providência, Host Sachs, Heinz Georg Schuster, Helmut Schwegler, Chris Seidel, Fred Sobik, Doru Stefanescu, Thomas Stoll, John Storey, Kurt Varmuza, Bohdan Zelinka, and all the coauthors of this book and apologize to all those whose names have been mistakenly omitted. We would also like to thank our editors Andreas Sendtko and Gregor Cicchetti from Wiley-VCH; they were always available and extremely helpful. Last but not least, we would like to thank our families for their support and encouragement throughout the writing of the book.

Finally, we hope that this book helps the reader to understand that the presented field is multifaceted in depth and breadth and as such is inherently interdisciplinary. This is important to realize because it allows one to pursue a problem-oriented rather than field-oriented approach to efficiently tackling state-of-the-art problems in modern sciences.

Vienna and Belfast,
March 2009

Matthias Dehmer
Frank Emmert-Streib

List of Contributors

Michael J. Barber
Austrian Research Centers – ARC
Division Systems Research
Donau-City-Strasse 1
1220 Vienna
Austria

Danail Bonchev
Virginia Commonwealth University
Center for the Study of
Biological Complexity
P.O. Box 842030
Richmond, VA 23284-2030
USA

Horst Bunke
University of Bern
Institute of Computer Science
and Applied Mathematics
Neubrückstrasse 10
3012 Bern
Switzerland

Qiong Cheng
Georgia State University
Department of Computer Science
Atlanta, GA 30303
USA

Michel Deza
École Normale Supérieure
45 rue d'Ulm
75005 Paris
France
and
Japan Advanced Institute of
Science and Technology
1-1 Asahidai
Nomi
Ishikawa
Japan

Dimitris Dimitropoulos
European Bioinformatics Institute
Genome Campus, Hinxton
Cambridge CB10 1SD
UK

Michael Drmota
Vienna University of Technology
Institute of Discrete Mathematics
and Geometry
Wiedner Hauptstrasse 8/104
1040 Vienna
Austria

Ernesto Estrada
University of Strathclyde
Institute of Complex Systems
at Strathclyde
Department of Physics and
Department of Mathematics
Glasgow G1 1XH
UK

Maria Fonoberova
AIMdyn, Inc.
Santa Barbara, CA 93101
USA

Bernhard Gittenberger
Vienna University of Technology
Institute of Discrete Mathematics
and Geometry
Wiedner Hauptstrasse 8/104
1040 Vienna
Austria

Adel Golovin
European Bioinformatics Institute
Genome Campus, Hinxton
Cambridge CB10 1SD
UK

Ivan Gutman
University of Kragujevac
Faculty of Science
P.O. Box 60
34000 Kragujevac
Serbia

M. John
European Bioinformatics Institute
Genome Campus, Hinxton
Cambridge CB10 1SD
UK

Marcus Kaiser
Newcastle University
School of Computing Science
Newcastle-upon-Tyne NE1 7RU
UK
and
Newcastle University
Institute of Neuroscience
Newcastle-upon-Tyne NE2 4HH
UK

Alexander Kononov
Russian Academy of Sciences
Sobolev Institute of Mathematics
Novosibirsk
Russia

Matjaž Kovše
University of Maribor
Faculty of Natural Sciences
and Mathematics
Koroška cesta 160
2000 Maribor
Slovenia
and
Institute of Mathematics,
Physics and Mechanics
Jadranska 19
1000 Ljubljana
Slovenia

Eugene Krissinel
European Bioinformatics Institute
Genome Campus, Hinxton
Cambridge CB10 1SD
UK

Xueliang Li
Nankai University
Center for Combinatorics
LPMC-TJKLC
Tianjin 300071
P.R. China

Dmitrii Lozovanu
Moldovan Academy of Sciences
Institute of Mathematics
and Computer Science
Academiei str., 5
Chisinau, MD-2005
Moldova

Alexander Mehler
Goethe-Universität
Frankfurt am Main
Abteilung für geisteswis-
senschaftliche Fachinformatik/
Department for Computing
in the Humanities
Georg-Voigt-Strasse 4
60325 Frankfurt am Main
Germany

Valia Mitsou
The City University of New York
Doctoral Program in Computer
Science
365 Fifth Avenue
New York, NY 10016
USA

Abbe Mowshowitz
The City College of New York
Department of Computer Science
Convent Avenue at 138th Street
New York, NY 10031
USA
and
The City University of New York
Doctoral Program in Computer
Science
365 Fifth Avenue
New York, NY 10016
USA

Manfred Paier
Austrian Research Centers – ARC
Division Systems Research
Donau-City-Strasse 1
1220 Vienna
Austria

Christian M. Reidys
Nankai University
Center for Combinatorics
LPMC-TJKLC
Tianjin 300071
P.R. China

Kaspar Riesen
Universtity of Bern
Institute of Computer Science and
Applied Mathematics
Neubrückstrasse 10
3012 Bern
Switzerland

Thomas Scherngell
Austrian Research Centers – ARC
Division Systems Research
Donau-City-Strasse 1
1220 Vienna
Austria

Sergey Sevastyanov
Russian Academy of Sciences
Sobolev Institute of Mathematics
Novosibirsk
Russia

Mikhail Shtogrin
Steklov Mathematical Institute
Gubkina str. 8
117966 Moscow
Russia

Mathieu Dutour Sikirić
Institute Rudjer Bošković
Group for Satellite Oceanography
10000 Zagreb
Croatia

Jennifer Simonotto
Newcastle University
School of Computing Science
Newcastle-upon-Tyne NE1 7RU
UK
and
Newcastle University
Institute of Neuroscience
Newcastle-upon-Tyne NE2 4HH
UK

Stefan Thurner
Medical University of Vienna
Complex Systems Research Group
Währinger Gürtel 18–20
1090 Vienna
Austria
and
Santa Fe Institute
1399 Hyde Park Road
Santa Fe, NM 87501
USA

Alexander Zelikovsky
Georgia State University
Department of Computer Science
Atlanta, GA 30303
USA

Jianbin Zhang
Nankai University
Center for Combinatorics
LPMC-TJKLC
Tianjin 300071
P.R. China

1
Entropy, Orbits, and Spectra of Graphs

Abbe Mowshowitz and Valia Mitsou

1.1
Introduction

This chapter is concerned with the notion of entropy as applied to graphs for the purpose of measuring complexity.

Most studies of complexity focus on the execution time or space utilization of algorithms. The execution time of an algorithm is proportional to the number of operations required to produce the output as a function of the input size. Space utilization measures the amount of storage required for computation. Both time and space complexity measure the resources required to perform a computation for a specified input. Measuring the complexity of a mathematical object such as a graph is an exercise in structural complexity. This type of complexity does not deal directly with the costs of computation; rather, it offers insight into the internal organization of an object. The structural complexity of a computer program, for example, may indicate the difficulty of modifying or maintaining the program.

One approach to structural complexity involves the length of a code needed to specify an object uniquely (Kolmogorov complexity). The complexity of a string, for example, is the length of the string's shortest description in a given description language [27]. The approach taken in this chapter centers on finding indices of structure, based on Shannon's entropy measure. Unlike Kolmogorov complexity, such an index captures a particular feature of the structure of an object. The symmetry structure of a graph provides the basis for the index explored here.

The choice of symmetry is dictated by its utility in many scientific disciplines. D'Arcy Thompson's classic work [25] showed the relevance of symmetry in the natural world. Structure-preserving transformations based on symmetry play a role in physics, chemistry, and sociology as well as in biology. A symmetry transformation of a graph is typically an edge-preserving bijection of the vertices, i.e., an isomorphism of the graph onto itself. Such a transformation is called an *automorphism*. If the vertices of the graph are labeled, an automorphism can be viewed as a permutation of the vertices that preserves adjacencies. The set of all automorphisms forms a group whose orbits provide the foundation for applying Shannon's entropy measure.

The collection of orbits of the automorphism group constitutes a partition and thus defines an equivalence relation on the vertices of a graph. Two vertices in the same orbit are similar in some sense. In a social network, collections of similar vertices can be used to define communities with shared attributes. The identification of such communities is of interest in applications such as advertising, intelligence, and sensor networks.

Measures of structural complexity are useful for classifying graphs and networks represented by graphs. One is led to conjecture, for example, that the more symmetric a network is (or the lower its automorphism-based complexity is), the more vulnerable to attack it will be. These related issues are explored in [19] in relation to sensor networks modeled as dynamic distributed federated databases [2].

In what follows we define the measure of graph complexity, discuss algorithms and heuristics for computing it, and examine its relationship to another prominent entropy measure [11] defined on graphs.

1.2
Entropy or the Information Content of Graphs

Given a decomposition of the vertices or edges of a graph, one can construct a *finite probability scheme* [10] and compute its entropy. A finite probability scheme assigns a probability to each subset in the decomposition. Such a numerical measure can be seen to capture the information contained in some particular aspect of the graph structure.

The orbits of the automorphism group of a graph constitute a decomposition of the vertices of the graph. As noted above, this decomposition captures the symmetry structure of the graph, and the entropy of the finite probability scheme obtained from the automorphism group provides an index of the complexity of the graph relative to the symmetry structure.

Let $G = (V, E)$ be a graph with vertex set V (with $|V| = n$) and edge set E. The *automorphism group* of G, denoted by $Aut(G)$, is the set of all adjacency-preserving bijections of V. Let $\{V_i | 1 \le i \le k\}$ be the collection of orbits of $Aut(G)$, and suppose $|V_i| = n_i$ for $1 \le i \le k$. The *entropy* or *information content* of G is given by the following formula ([13]):

$$I_a(G) = -\sum_{i=1}^{k} \frac{n_i}{n} \log \frac{n_i}{n} \,.$$

For example, the orbits of the graph of Figure 1.1 are $\{1\}$, $\{2,5\}$, and $\{3,4\}$, so the information content of the graph is $I_a(G) = -\frac{1}{5}\log\frac{1}{5} - 2(\frac{2}{5}\log\frac{2}{5}) = \log 5 - \frac{4}{5}\log 2$.

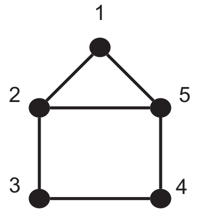

Figure 1.1 Information content of a graph.

Clearly, $I_a(G)$ satisfies $0 \le I_a(G) \le \log n$, where the minimum value occurs for graphs with the transitive automorphism group, such as the cycle C_n and the complete graph K_n on n vertices; the maximum is achieved for graphs with the identity group. The smallest nontrivial, undirected graph with an identity group is shown in Figure 1.2.

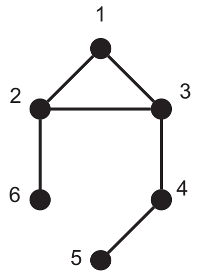

Figure 1.2 Smallest nontrivial graph with identity group.

The idea of measuring the information content of a graph was first presented in [21]; it was formalized in [26] and further developed in [13–16]. $I_a(G)$ is a function of the partition of the vertices of G determined by the orbits of $Aut(G)$. As such the measure captures the structure of vertex similarity. In the case of organic molecules, the lower the information content (or the greater the symmetry), the fewer the possibilities for different interactions with other molecules. If all the atoms are in the same equivalence class, then it makes no difference which one interacts with an atom of another molecule. The same can be said for social networks. Any member of an equivalence class of similar individuals can serve as a representative of the class.

The utility of the measure $I_a(G)$ can be seen from the following special case. The cartesian product $G \times H$ of graphs G and H is defined by $V(G \times H) = V(G) \times V(H)$ and for $(a, b), (c, d) \in V(G \times H)$, $[(a, b), (c, d)] \in E(G \times H)$ if $a = c$ and $[b, d] \in E(H)$ or if $b = d$ and $[a, c] \in E(G)$.

The hypercube Q_n with 2^n vertices is defined recursively by $Q_1 = K_2$ and for $n \ge 2$, $Q_n = K_2 \times Q_{n-1}$. Since Q_n has a transitive automorphism group, $I(Q_n) = 0$. The hypercube Q_n offers a desirable configuration for parallel computation because processors must exchange messages in executing an algorithm, and the distance between any two vertices (representing processors) in the hypercube is at most n.

By contrast, an $m \times m$ mesh configuration (formed by taking the cartesian product of two isomorphic line graphs, each with m vertices) consists of m^2 vertices and has a maximum distance of $2m$. A $2^{\frac{n}{2}} \times 2^{\frac{n}{2}}$ mesh for even n having the same number of vertices as Q_n has a maximum distance between vertices of $2(2^{\frac{n}{2}} - 1)$. At the same time the information content of such a mesh is $\frac{n}{2} - 1$ [13].

This example suggests that good graph configurations for parallel computation score low on information complexity or, alternatively, are highly symmetric. Information complexity is a coarse filter, but it is useful nonetheless.

Computing the group-based entropy or information content of a graph requires knowledge of the orbits of the automorphism group. An obvious approach to computing the orbits is to determine the automorphism group and then to observe the action of automorphisms on the vertices of the graph. This is not an efficient method in general, but the algebraic structure of a graph can be exploited to find the automorphism group efficiently in some cases. The general question of determining the automorphism group is taken up in Section 1.3; heuristics for finding the orbits of $Aut(G)$ are surveyed in Section 1.4.

1.3
Groups and Graph Spectra

Let $G = (V, E)$ be a graph with vertex set V of size n, edge set E of size m, and automorphism group $Aut(G)$. (See [3] for general coverage of algebraic aspects of graph theory and [12] for specific treatment of the automorphism group of a graph.) Since automorphisms are in effect relabelings of the vertices, they can be represented as permutation matrices. Let $A = A(G)$ be the adjacency matrix of G. Then a permutation matrix P is an automorphism of G if and only if $P^T AP = A$ or $PA = AP$.

Thus, one way to construct the automorphism group of a graph G is to solve the matrix equation $AX = XA$ for permutation matrices X. The Jordan canonical form of A as a matrix over the reals can be used to obtain the general solution X. Taking G to be undirected and thus A symmetric and letting $\tilde{A} = U^T AU$ be the Jordan form of A, we have $(U\tilde{A}U^T)X = X(U\tilde{A}U^T)$ or $\tilde{A}\tilde{X} = \tilde{X}\tilde{A}$, where $\tilde{X} = U^T XU$.

Thus the construction of $Aut(G)$ requires computing the orthogonal matrix U and finding all $\tilde{X}$ that commute with $\tilde{A}$. The matrix $\tilde{X}$ depends on the elementary divisors of A. With no additional information, this method of constructing the group is not too promising since it is necessary to find all those solutions that are permutation matrices.

In the special case where A has all distinct eigenvalues, $\tilde{X}$ has the form of a diagonal matrix with arbitrary parameters on the main diagonal. In this case, $X = U\tilde{X}U^T$. Clearly $U\tilde{X}U^T$ is symmetric, so if it is a permutation matrix, it must correspond to a product of disjoint transpositions. This means that every element of $Aut(G)$ has order 2 and the group is therefore abelian [12, 17]. The converse is not true since, for example, the graph G of Figure 1.3 has the characteristic polynomial $(x+1)^2(x^3-2x^2-5x+2)$.

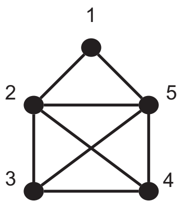

Figure 1.3 $Aut(G)$ is abelian, every element is of order 2, but the characteristic polynomial has repeated roots.

An analogous result holds for digraphs. Using the same analysis, Chao [5] showed that if the adjacency matrix of a digraph has all distinct eigenvalues, then its automorphism group is abelian. However, the automorphisms need not be of order 2. For example, the adjacency matrix of digraph D in Figure 1.4 has the characteristic polynomial $(x^3-1) = (x-1)(x^2+x+1)$ but the permutation (123) is an automorphism of D.

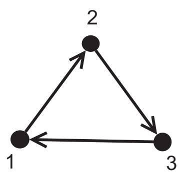

Figure 1.4 $Aut(D) = \langle(123)\rangle$, abelian but not every element has order 2.

Both of these results are special cases of the following:

Theorem 1.1 *Suppose the adjacency matrix $A = A(D)$ of a digraph D is non-derogatory with respect to a field F, i.e., its characteristic polynomial coincides with its minimal polynomial over F. Then $Aut(D)$ is abelian.*

Proof. The result is an immediate consequence of the fact that under the hypothesis of the theorem, every matrix over F commuting with A can be expressed as a polynomial in A.

In particular, if A has all distinct eigenvalues, it is non-derogatory over the complex number field. To see that every automorphism of an (undirected) graph has order 2 under this condition, it suffices to observe that any polynomial in a symmetric matrix is again symmetric.

If the adjacency matrix fails to be nonderogatory, then some leverage in constructing the automorphism group can be obtained by taking advantage of the fact that the matrix consists of zeroes and ones. In particular, the adjacency matrix can be interpreted as a matrix over $GF(2)$, thus reducing the solution space of the matrix equation $AX = XA$ to zero-one matrices at the outset.

Thus suppose $A = A(G)$ (for a graph G) is a matrix over $GF(2)$. To demonstrate a method for constructing automorphisms, we revisit the special case of A being nonderogatory over $GF(2)$.

In this case we know that:

1. $M \in Aut(G)$ implies $M^2 = I$ (the identity matrix) and
2. $M \in Aut(G)$ implies $M = \sum_{i=0}^{n-1} a_i A^i$.

So if $M \in Aut(G)$, then we can write

$$M = \sum_{i=0}^{n-1} a_i A^i$$

and

$$I = M^2 = \left(\sum_{i=0}^{n-1} a_i A^i \right)^2 = \sum_{i=0}^{n-1} a_i (A^i)^2.$$

Thus $\{M | M = \sum_{i=0}^{n-1} a_i A^i \text{ and } M^2 = I\} \supseteq Aut(G)$.

Constructing the group in this case reduces to finding all polynomials in A^2 that are equal to the identity matrix. These have the form

$$p(A)\mu_{A^2}(A^2) + I,$$

where $\mu_{A^2}(x)$ is the minimal polynomial of A^2.

Thus, if $M^2 = I$, then $M = p(A)\mu_{A^2}(A) + I$ for some polynomial $p(x)$, since $(p(A)\mu_{A^2}(A) + I)^2 = (p(A^2)\mu_{A^2}(A^2) + I) = 0 + I = I$.

The characteristic and minimal polynomials of graph G in Figure 1.5 coincide over the real numbers, i.e., $\phi(x) = \mu(x) = (x^3 - x^2 - 6x + 2)x(x + 1)$ and over $GF(2)$ with $\phi(x) = \mu(x) = x^3(x + 1)^2$. Hence, the adjacency matrix of G is nonderogatory over both fields. The minimal polynomial of A^2 is $\mu_{A^2}(x) = x^2(x + 1)$, which is of degree 3.

Therefore, $M \in Aut(G)$ implies $M = \mu_{A^2}(A)(b_0 I + b_1 A) + I$. There are four possible solutions for M corresponding to the four possible values for b_0 and b_1. All of these solutions, namely,

$$I, A^3 + A^2 + I, A^4 + A^3 + I, A^4 + A^2 + I,$$

turn out to be permutation matrices so that the automorphism group of G contains precisely these four elements.

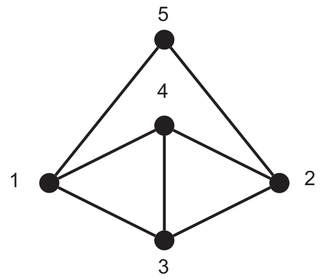

Figure 1.5 Computation of automorphisms over $GF(2)$.

Note that $\mu^2_{A^2}(x) = x\phi_A(x)$ if n is odd, or $\mu^2_{A^2}(x) = \phi_A(x)$ if n is even. Hence, if $m = \deg \mu_{A^2}(x)$ and M satisfies $AM = MA$ and $M^2 = I$, then $M = \mu_{A^2}(x) \sum_{i=0}^{n-m-1} b_i A^i + I$, where $b_i \in GF(2)$.

To determine $Aut(G)$, it suffices to examine $2^{n-m-1} \approx 2^{n/2}$ values of the parameters b_i, to pick out the permutation matrices (i.e., elements of $Aut(G)$).

However, some further simplification is possible. Let $Q = \mu_{A^2}(A)$ and $Z(b) = \sum_{i=0}^{n-m-1} b_i A^i$. Then $M = QZ(b) + I$. Multiplying by M on the right gives $MQ = Q^2 Z(b) + Q = Q$. Thus, if M is an automorphism of G, then $MQ = Q$, which means that similar vertices of G correspond to identical rows of Q. In addition, the identical rows must occur in *minimal pairs*, which gives a sufficient condition for $Aut(G)$ to be trivial.

If $\mu_{A^2}(A)$ has all distinct rows or no minimal pairs of identical rows, then $Aut(G)$ is trivial. The converse is not true. Both graphs in Figure 1.6 have trivial groups, but $\mu_{A^2}(A(G_1))$ has all distinct rows while $\mu_{A^2}(A(G_2))$ has three pairs of identical rows.

Theorem 1.2 *[18]; see also [6]. Let D be a digraph and $A = A(D)$ be its adjacency matrix. If $\phi_A(x)$ is irreducible over the integers, then $Aut(D)$ is trivial.*

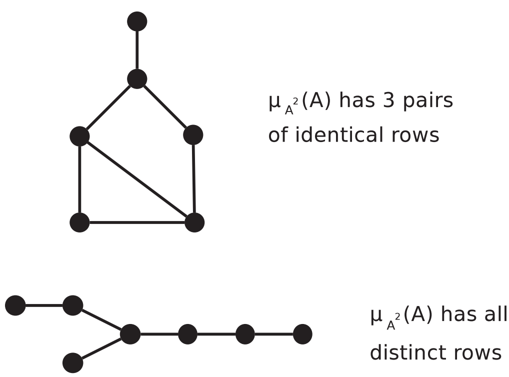

Figure 1.6 Identity graphs.

Proof. Suppose there is an $M(\neq I) \in Aut(D)$, and that the permutation corresponding to M consists of r disjoint cycles of lengths $k_1, \ldots, k_r$. Let z be a nonzero vector consisting of k_1 components equal to c_1, followed by k_2 components equal to c_2, followed by $\ldots k_r$ components equal to c_r. Consider $Az = xz$. This gives a system of r equations in the r unknowns $c_1, c_2, \ldots, c_r$. Thus $Az = xz$ reduces to $Bc = xc$, where $c = (c_1, c_2, \ldots, c_r)^T$. Now z and c are eigenvectors of A and B, respectively, and $\det(B - xc) | \det(A - xz)$, where $\deg(\det(B - xc)) < \deg(\det(A - xz))$. Hence, $\phi_A(x)$ has a nontrivial factorization, which completes the proof.

Figure 1.7 shows a digraph (D) and graph (G) (with the smallest number of vertices) satisfying the condition of the theorem. $\phi_{A(D)}(x) = x^3 - x - 1$ and $\phi_{A(G)}(x) = x^6 - 6x^4 - 2x^3 + 7x^2 + 2x - 1$.

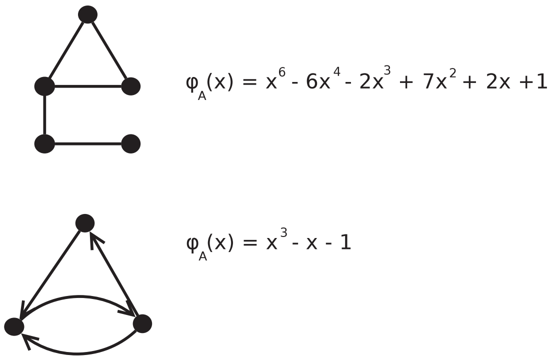

Figure 1.7 Smallest graph and digraph whose characteristic polynomials are irreducible over the integers.

Note that the theorem also holds if $\phi_A(x)$ is taken as a polynomial over a finite field. For example, over $GF(2)$, $x^3 - x - 1$ is irreducible, but $x^6 - 6x^4 - 2x^3 + 7x^2 + 2x - 1 = x^6 + x^2 + 1 = (x^3 + x + 1)^2$.

For graphs this criterion is not very useful since the characteristic polynomial of any graph is reducible over $GF(2)$. There are regular graphs and trees that have the trivial group, but the characteristic polynomial of any regular graph has a linear factor, as does the characteristic polynomial of a tree with an odd number of vertices.

The foregoing discussion suggests the utility of trying to relate the factorization of the characteristic polynomial to the structure of the automorphism group. For example, if G is a graph with an even number n of vertices and adjacency matrix $A = A(G)$, and if $\phi_A(x) = \alpha(x)\beta(x)$ with $\deg \alpha = \deg \beta$ and both α and β are irreducible, then either $Aut(G)$ is trivial or it is of order 2 and consists of the identity and (with a suitable labeling) the permutation $(1, 2)(3, 4) \cdots (n/2, n/2)$.

Table 1.1 contains a list of all 156 graphs on six vertices, showing factored characteristic polynomials and the sizes of their respective automorphism group orbits. Each graph is defined by its list of edges, shown as a sequence of pairs of numbers referring to a standard template with the vertices numbered from 1 to 6 in clockwise order. The last column shows the sizes of the orbits. Complements are not given explicitly, but their polynomials are listed. The orbits of G and $\bar{G}$ are the same.

Table 1.1 Characteristic polynomials and orbit sizes of all graphs on six vertices.

G: # edges: list	Polynomial of G	Polynomial of $\bar{G}$	Orbit sizes
0:	x^6	$(x+1)^5(x-5)$	6
3: 16 23 45	$(x+1)^3(x-1)^3$	$x^3(x+2)^2(x-4)$	6
6: 12 16 23 34 45 56	$(x-1)^2(x+1)^2(x+2)(x-2)$	$x^2(x-1)(x+2)^2(x-3)$	6
6: 15 16 23 24 34 56	$(x-2)^2(x+1)^4$	$x^4(x-3)(x+3)$	6
3: 15 16 56	$x^3(x+1)^2(x-2)$	$x^2(x+1)^2(x^2-2x-9)$	33
1: 12	$x^4(x-1)(x+1)$	$x(x+1)^3(x^2-3x-8)$	24
4: 12 15 24 45	$x^4(x-2)(x+2)$	$(x-1)(x+1)^3(x^2-2x-7)$	24
4: 12 16 34 45	$x^2(x^2-2)^2$	$(x+1)^2(x^4-2x^3-8x^2+6x-1)$	24
5: 12 15 16 23 24	$x^2(x-1)(x+1)(x-2)(x+2)$	$(x-1)(x+2)(x+1)^2(x^2-3x-2)$	24
5: 14 16 23 45 56	$x^2(x-1)(x+1)(x-2)(x+2)$	$x(x-1)(x+1)^2(x^2-x-8)$	24
6: 12 14 15 24 25 45	$x^2(x-3)(x+1)^3$	$x^3(x+1)(x^2-x-8)$	24
7: 12 16 23 25 34 45 56	$(x+1)(x-1)(x^2-2x-1)(x^2+2x-1)$	$x(x+2)(x^2-2)(x^2-2x-2)$	24
7: 14 15 16 23 45 46 56	$(x+1)^4(x-3)(x-1)$	$x^4(x^2-8)$	24
7: 15 16 23 24 34 45 56	$(x+1)^2(x^2-3)(x^2-2x-1)$	$x^2(x^2-2x-2)(x^2+2x-2)$	24
2: 12 56	$x^2(x-1)^2(x+1)^2$	$x^2(x+1)(x+2)(x^2-3x-6)$	222
3: 12 16 23	$x^2(x^2-x-1)(x^2+x-1)$	$(x+1)(x^2+x-1)(x^3-2x^2-8x-3)$	222
6: 14 15 16 23 45 56	$x(x-1)(x+1)^2(x^2-x-4)$	$x^2(x+1)(x^3-x^2-8x+4)$	222
7: 13 16 23 26 34 45 56	$x(x^2+x-1)(x^3-x^2-5x+4)$	$(x-1)(x+1)(x^2+x-1)(x^2-x-5)$	222
7: 12 15 23 24 25 45 56	$x^2(x^2+x-1)(x^2-x-5)$	$(x+1)(x^2+x-1)(x^3-2x^2-4x+1)$	222
5: 12 16 24 45 56	$x(x-2)(x^2+x-1)^2$	$(x^2+x-1)^2(x^2-2x-5)$	15
5: 12 23 24 25 26	$x^4(x^2-5)$	$x(x-4)(x+1)^4$	15
2: 12 16	$x^4(x^2-1)$	$(x+1)^3(x^3-3x^2-7x+3)$	123
3: 12 15 16	$x^4(x^2-3)$	$(x+1)^3(x^3-3x^2-6x+4)$	123
4: 15 16 23 56	$x(x+1)(x-2)(x+1)^3$	$x^3(x^3-11x-12)$	123
4: 12 15 16 34	$x^2(x+1)(x-1)(x^2-3)$	$x(x+1)^2(x^3-2x^2-8x+4)$	123

Table 1.1 (continued).

G: # edges: list	Polynomial of G	Polynomial of $\bar{G}$	Orbit sizes
5: 15 16 23 34 56	$x(x-2)(x+1)^2(x^2-2)$	$x^2(x+1)(x^3-x^2-9x+3)$	123
6: 12 15 23 24 35 45	$x^4(x^2-6)$	$(x+1)^3(x^3-3x^2-3x+7)$	123
6: 12 13 14 15 16 56	$x^2(x+1)(x^3-x^2-5x+3)$	$x^2(x+1)^2(x^2-2x-6)$	123
7: 12 14 15 16 25 45 56	$x^3(x-3)(x+1)(x+2)$	$x(x-1)^2(x^3-2x^2-5x+4)$	123
4: 12 14 15 16	$x^4(x-2)(x+2)$	$(x+1)^3(x^3-3x^2-5x+3)$	114
7: 15 16 26 34 35 45 56	$(x-1)(x+1)^2(x^3-x^2-5x+1)$	$x^3(x+2)(x^2-2x-4)$	114
3: 16 23 56	$x^2(x+1)(x-1)(x^2-2)$	$x(x+1)(x^4-x^3-11x^2-7x+4)$	1122
4: 12 15 16 56	$x^2(x+1)(x^3-x^2-3x+1)$	$x(x+1)(x+2)(x^3-3x^2-4x+2)$	1122
4: 12 16 23 56	$x^2(x-1)(x+1)(x^2-3)$	$x(x+1)(x+2)(x^3-3x^2-4x+4)$	1122
5: 12 15 16 45 56	$x^2(x^2-x-3)(x^2+x-1)$	$(x+1)(x^2+x-1)(x^3-2x^2-6x+1)$	1122
5: 12 14 15 16 56	$x^2(x+1)(x^3-x^2-4x+2)$	$x(x+1)(x^4-x^3-9x^2-5x+4)$	1122
5: 12 15 16 23 45	$(x-1)(x+1)(x^4-4x^2+1)$	$x(x+2)(x^4-2x^3-6x^2+2x+4)$	1122
5: 15 16 23 45 56	$(x-1)(x+1)^2(x^3-x^2-3x+1)$	$x^2(x^4-1x^2-8x+4)$	1122
6: 12 15 16 24 45 56	$x^2(x+2)(x^3-2x^2-2x+2)$	$(x-1)(x+1)(x^4-8x^2-8x+1)$	1122
6: 13 16 23 34 45 56	$(x-1)(x^2+x-1)(x^3-4x-1)$	$(x+2)(x^2+x-1)(x^3-3x^2-x+2)$	1122
6: 12 13 14 16 45 56	$x^2(x^4-6x^2+4)$	$(x-1)(x+1)^2(x^3-x^2-7x-3)$	1122
4: 12 16 23 45	$(x-1)(x+1)(x^2-x-1)(x^2+x-1)$	$x(x^2+x-1)(x^3-x^2-9x-4)$	222
5: 12 14 15 24 45	$x^3(x+1)(x^2-x-4)$	$x(x+1)^2(x^3-2x^2-7x+4)$	222
5: 12 16 34 45 56	$(x^3-x^2-2x+1)(x^3+x^2-2x-1)$	$(x^3-2x^2-5x+1)(x^3+2x^2-x-1)$	222
6: 12 14 16 34 45 56	$(x^3-2x^2-x+1)(x^3+2x^2-x-1)$	$(x^3-x^2-6x-3)(x^3+x^2-2x-1)$	222
6: 12 15 16 23 25 45	$(x^2-2x-1)(x^2+x-1)^2$	$(x^2-2x-4)(x^2+x-1)^2$	1122
6: 12 15 16 23 24 56	$x(x+1)(x^4-x^3-5x^2+3x+4)$	$x(x+1)(x^4-x^3-8x^2-2x+6)$	1122
7: 12 15 16 24 26 45 56	$x^2(x+1)(x^3-x^2-6x+2)$	$x(x+1)(x^4-x^3-7x^2+x+8)$	1122
7: 12 14 15 16 24 45 56	$x(x^2+x-1)(x^3-x^2-5x-2)$	$(x^2+x-1)(x^4-x^3-6x^2-x+1)$	1122
7: 12 16 23 24 34 45 56	$(x^2+x-1)(x^4-x^3-5x^2+2x+4)$	$(x^2+x-1)(x^4-x^3-6x^2+3x+1)$	1122
7: 14 16 23 24 34 45 56	$x(x+1)(x^4-x^3-6x^2+4x+4)$	$x(x-1)(x+1)(x^3-7x-4)$	1122
7: 12 13 15 24 34 45 56	$x^2(x^4-7x^2+4)$	$(x+1)^2(x^4-2x^3-5x^2+6x+4)$	1122
7: 12 14 16 23 24 45 56	$x^2(x-1)(x+2)(x^2-x-4)$	$(x+1)(x-1)(x+2)(x^3-2x^2-3x+2)$	1122