Structure Indicators for Transportation Graph Analysis I: Planar Connected Simple Graphs
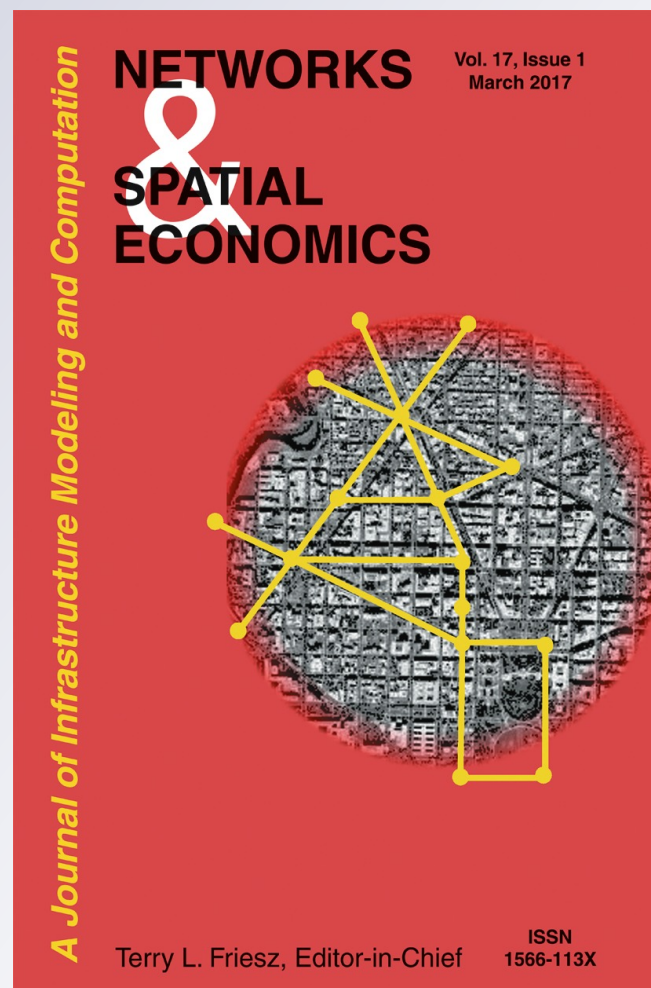
Lorenzo Mussone & Roberto Notari

Springer

Springer

CrossMark

Structure Indicators for Transportation Graph Analysis I: Planar Connected Simple Graphs

Lorenzo Mussone[1] · Roberto Notari[2]

Abstract The paper deals with the representation of a transportation infrastructure by a planar connected simple graph and aims at studying its features through the analysis of graph properties. All planar and connected graphs with 4 up to 7 edges are analysed and compared to extract the most suitable parameters to investigate some network features. Then, a set of 41 graphs representing some actual underground networks are also analysed. Besides, as a third scenario, the underground network of Milan, along its development in years, is proposed in order to apply the proposed methodology. Many parameters are taken into consideration. Some of them are already discussed in literature, such as the eigenvalues and gaps of adjacency matrix or such as the "classical" parameters α, β, γ. Others, such as the first two Betti numbers, are new for these applications. In order to overcome the problem of comparing features of graphs with different size, the normalisation of these parameters is considered. Some relationships between Betti numbers, eigenvalues, and classical parameters are also investigated. Results show that the eigenvalues and gaps of the adjacency matrix well represent some features of the graphs while combining them with the Betti numbers, a more significant interpretation can be achieved. Particularly, their normalised values are able to describe the increasing complexity of a graph.

Keywords Transportation graph · Topological interpretation · Graph theory · Adjacency matrix · Eigenvalues · Betti numbers

✉ Lorenzo Mussone
    mussone@polimi.it

    Roberto Notari
    roberto.notari@polimi.it

[1]   Department of Architecture, Built Environment, and Construction Engineering, Politecnico di Milano, Milan, Italy

[2]   Department of Mathematics, Politecnico di Milano, Milan, Italy

1 Introduction

Every transportation network (transit, road, maritime, aerial) can be represented with a graph, where nodes become vertices and links become (undirected or directed, accordingly) edges. Obviously, what can be subjective in the conversion is the amount of details taken into consideration. Only in transit networks and especially in underground ones, this problem is much less crucial because no at level crossings are allowed and possible paths are perfectly defined.

In general, a transportation network can have one-way links or can require the two ways of a link to be separated. In those cases, the network must be translated into a directed graph in order to reflect that property. In many other cases (like for underground networks) this is not generally needed and an undirected graph is adequate.

Other network properties concern the performance or property of a link or a node (e.g., length, travel time and speed curves, fares, and so on) and this can be inserted into a graph through weights.

Since 1950s, many attempts to analyse networks through graphs are made, without really discerning between the graph and the network; many papers dealing with network properties (e.g., Kansky 1963) are often cited together with papers based on graph properties.

The studies on graph properties aim at defining some numerical indicators to pick and synthesize the considered properties. This approach has a twofold justification for what concerns transportation studies:

– to a short term, to understand a possible different behaviour of networks, for example when they are subjected to dynamic changes in demand,
– to a medium-long term, how, why and where a network should be modified to better satisfy transport needs.

However, it must be stressed that network loading depends much on demand and its composition; usually, interaction between supply (what is represented by a graph) and demand is faced by assignment approaches and therefore we do not take into consideration that feature when we analyse networks through graphs.

A comprehensive analysis of the quite large literature on the study of transportation networks through graphs would require a dedicated review paper; therefore, we recall only a few papers among the many relevant ones, to outline briefly past and present researches with a special regard to the aims of the paper.

Seminal papers are those by Garrison and Marble (1962, 1964) who really pioneered the field by introducing three parameters, computed from the graph, directly linked to network design: circuits α, degree of connectivity γ, and complexity β, that we recall and comment in Section 2.2.

Actually, the number of circuits is the cyclomatic number proposed by Berge (1962) to count the number of independent circuits; it also shows the maximum number of edges that can be removed without disconnecting the (planar) graph. Variants of these indices can be found; e.g., Prihar (1956), working in telecommunication networks, proposes the use of the degree of connectivity calculated as $1/\alpha$. Garrison and Marble (1962) prove also that the indices α and γ are quite redundant and, for a constant number of vertices, the correlations between α, β and γ are equal to one. Black (2003)

states that, since the strong redundancy between these parameters, only one of them is sufficient to contain graph information. He prefers to use γ since it is the easiest to interpret and it ranges in [0, 1].

The paper by Gattuso and Miriello (2005) faces the problem of a topological and geographical evaluation of metro networks through the definition of a methodological approach based on a set of indicators, at graph and geographical levels.

At graph level they use, other than the ones by Garrison and Marble (1962),

- the local degree, i.e., the number of concurrent links on the generic node;
- the number of destinations directly connected by i nodes;
- the number of nodes; generally it does not correspond to the number of network stations (two or three different line stations, placed at the same point, are represented by a unique network node);
- the number of links (a link covered by several lines is counted only once);
- the number of cycles.

A list of parameters at geographical level is also proposed. Actually, this study is very notable for the development of two indicators at geographical level.

Derrible and Kennedy (2010, 2011, 2012) aim at characterizing the network features of metro systems by adapting various concepts of graph theory to describe their characteristics clustered into three classes: State, Form and Structure. In their study, 33 metro mode networks (meaning urban rapid rail transit whether it is underground, at grade or elevated), real working over six major world areas (North and Latin America, Eastern and Western Europe, Africa and Asia) and with different sizes, are taken into consideration. The authors transform the networks into graphs by coupling a node to each station only if the station is a transfer station (where transfer station means a station connected to at least three other stations, both transfer and termini). Obviously, this choice can affect the analysis related to land and service accessibility but not the functional properties of connection between nodes. Variables taken into account are the number of vertices, the number of edges, and the network diameter, i.e., the total number of edges linking the two furthest vertices, using the shortest path. The state indicators are two: the first one is the complexity β and the second one is the degree of connectivity γ both by Garrison and Marble (1962). Other indicators are calculated by using physical characteristics of the network (such as the lengths of edges).

The long-standing interest in measuring the spatial structure of road networks has produced many other papers with different aims and tools. Newman (2003), Gastner and Newman (2006) and Nadakuditi and Newman (2012) find some clear patterns that imply a connection between network structure and geography by analysing empirical data and overcoming past limitations on computer resources. Blumenfeld-Lieberthal (2009) analyse the topology of transportation networks within different systems of cities using their characteristics as indicator of economic activity between them.

Xie and Levinson (2007, 2008) propose to calculate three parameters, heterogeneity, connection pattern and continuity, to measure structural characteristics of road networks. Heterogeneity is derived from the probability, p_k, that a vertex is connected to other k vertices and it is proportional to $1/k^a$, where k is the degree

of a vertex and a is a coefficient to be calibrated. The distribution of p_k is called power law degree and networks with this property are called scale-free. On this topic a particular interest assumes the paper by Barabasi and Albert (1999) who study the ways networks can grow but without the constraint of planarity. Connection pattern uses Garrison and Marble's (1962) definition but it mainly focuses on comparing the total length of roads belonging to a ring or circuit to the total length of roads. However, it appears of great interest the effort of defining a methodology capable of singling out circuit blocks, subsets of a graph containing no bridge or articulation point and remaining connected after deleting a vertex or an edge. Continuity is based on a definition of edge hierarchy which again implies a certain quantification of a network property.

A thorough discussion about structure and dynamics in transportation networks is proposed by Ducruet and Lugo (2013); static dimension of those networks is analysed with regard to topology, geometry, morphology, and spatial structure from a global and local point of view, according to the required application (e.g., looking for interconnectivity or vulnerability, respectively). Kurant and Thiran (2006) propose to extract topology of mass transportation networks from their timetables. Many other papers focus on dynamic features (mainly the growth) of transportation networks [see, for example, Xie and Levinson (2009a, b), Levinson (2009)], so many and so specialized that this topic deserves a dissertation apart.

Given a graph, with its size we mean the number of vertices and of edges, while with complexity we mean not only the size, but the topology of the graph, too, that is to say, the ways edges connect vertices. With this premise in mind, our research focuses on the following two questions:

1. given a transportation graph, are there indicators that measure its complexity?
2. once the indicators are found, how can we compare the complexity of transportation graphs, different both for size and complexity?

It is well known that graphs have shown very useful in many application areas, other than transportation, but it is remarkable that the study of graph theory is still an active research area in pure mathematics, despite the fact that graph were first considered by L. Euler in 18th century. In the last 15 years, the development of both scientific calculus software and computer algebra systems has allowed researchers to study graphs from the point of view of linear and commutative algebra. Even if the properties under consideration in pure mathematics may not be of immediate interest for transport scholars, the selected indicators, e.g., eigenvalues, codimension, Betti numbers, among the others, could be used for measuring the complexity of graphs. E.g., Sáenz-de-Cabezón and Wynn (2009) and Maruri-Aguilar et al. (2012) use Betti numbers to study the reliability of networks and systems.

So, the above first question can be rewritten as: can the indicators coming from algebra measure the complexity of graphs? By applying the proposed indices to three different datasets, we give good evidences that the question has a positive answer. So, we tackle the second problem by introducing suitable normalization factors, main novelty for what at our knowledge. Certainly, future studies on different datasets could give new insights on the interpretation of the proposed indicators.

In this paper, we consider only planar, connected and undirected graphs. Extensions to non-planar, weighted and/or directed graphs will be the topic of future research. Our hope is that the linear algebraic and algebraic indicators (defined and discussed in Section 2) could support planners to integrate traditional planning factors such as demand, demography, geography, demand assignment, and costs. They could also be used to compare existing systems especially for what concerns their complexity.

The paper has four other sections and three appendices: Section 2 presents the methodology and the parameters proposed for the analysis; Section 3 describes the datasets used to answering our two research questions; Section 4 shows and discusses the results giving evidence that the proposed indicators really represent the complexity of transportation graphs. Finally, in Section 5 we summarize the conclusions. Appendix 1 presents a short review on polynomial ideals on which the proposed methodology is based, Appendix 2 contains a mathematical description of the algebraic parameters, and finally, Appendix 3 is a sample of computer session in which we compute the algebraic parameters for a given graph. We stress that the session can be easily adapted to other graphs, and that Singular, the used software, is a computer algebra system for polynomial computations and is free and open-source under the GNU General Public License.

2 Methodology

In this section we introduce the parameters we are interested in, and we explain their meaning in terms of the geometrical properties of graphs. As we have to compare graphs with different numbers of vertices and edges, we need to pinpoint some suitably computed functions in order to normalise these parameters. These functions are empirical, in the sense that we modelled and checked them starting from the exhaustive analysis of planar connected graphs with 4, 5, 6, and 7 vertices.

To help the reader, we subdivide the parameters into three classes: the classical, the linear algebraic, and the algebraic parameters (see the list in Subsection 2.1). The classical ones were introduced in Garrison and Marble (1962) and widely used since then (see Subsection 2.2). The linear algebraic parameters owe their name to the fact that they are computed from the adjacency matrix of graphs by linear algebra techniques (see Subsection 2.3). We call algebraic the last set of parameters because we compute them by using the algebra of polynomial rings (see Subsection 2.4). In order not to exceed here with mathematical details but to make clear the basic concepts of the theory, we propose a brief introduction on the algebra of polynomial rings in Appendix 1, definitions and proved statements on the algebraic parameters in Appendix 2, and a computer session in which the algebraic parameters of a given graph are computed in Appendix 3.

2.1 The Considered Parameters

Herein, we report the list of the parameters, both classical and new, that we take into consideration.

Let G be a simple planar connected graph with v vertices and e edges.

Classical parameters:

- α
- β
- γ

Linear algebraic parameters:

- Average degree of vertices
- λ and normalised λ
- Gap g and normalised g

Algebraic parameters:

- Codimension $codim(G)$ and normalised $codim(G)$
- Degree
- deg (G)
- Betti numbers $b_{21}(G)$, $b_{22}(G)$, and normalised $b_{21}(G)$

2.2 Classic Parameters

As said before, with classical parameters we mean the ones introduced in Garrison and Marble (1962), and, since then, widely used in the description and analysis of transportation graphs.

Let G be a simple planar connected graph with v vertices and e edges.

- α

   It is calculated according to equation

$$\alpha = \frac{e - v + 1}{2v - 5}.$$ (1)

α is the normalized number of independent cycles of a graph with v vertices and e edges. Of course, $0 \leq \alpha \leq 1$. We obtain the same range even if we observe that $v - 1 \leq e \leq 3v - 6$, because the graph is planar and connected. The main weakness of α is that it gives information neither on the number of vertices in a cycle, nor on the topology of the graph, as the following example shows. In Fig. 1 some examples are proposed

- β

   It is calculated according to equation
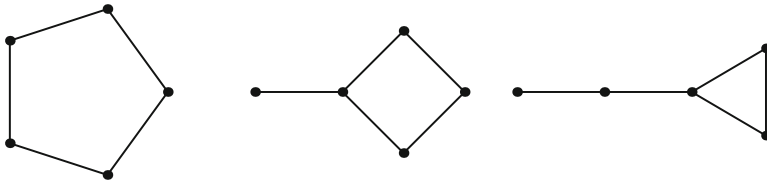
$$\beta = \frac{e}{v}.$$ (2)

Fig. 1 Three non-isomorphic graphs with $v = e = 5$ and $\alpha = 0.2$ despite the evident topological differences

Its graph-theoretical meaning in terms of graph-theory is not so evident. At a first look, it measures a density of edges with respect to vertices. As $v - 1 \leq e \leq 3v - 6$, we have that

$$1 - \frac{1}{v} \leq \beta \leq 3 - \frac{6}{v}.$$

Secondly, it is half the average degree of vertices of the graph, as explained in the next subsection. As for α, also β gives no information on the topology of the graph.

- Γ

It is calculated according to equation

$$\gamma = \frac{e}{3v - 6}. \tag{3}$$

It gives the number of edges normalized on the maximum number of edges for a connected planar graph with v vertices. We have that $\frac{1}{3} < \gamma \leq 1$. Once again, γ gives no information on the topology of the graph.

2.3 Linear Algebraic Parameters

In this paragraph, we introduce three parameters computed from the adjacency matrix of a graph. For the sake of completeness, we recall the construction of the adjacency matrix. Furthermore, we introduce two classes of graphs, one for the second parameter and one for the third one, that maximise the parameter in the set of the planar connected graphs with the same numbers of vertices and edges. Hence, we get a normalised parameter, more suitable in comparing graphs with different characteristics, because it is independent on the numbers v of vertices and e of edges.

Let $G = (V, E)$ be a simple graph on v vertices. It is possible to associate some matrices to G. These matrices encode all the information one needs to reconstruct the graph G. The aim of algebraic graph theory (Chung 1997; Lovasz 2007) is the study of the properties of G by using those matrices. In the following, we focus on a matrix, the adjacency matrix, and on its eigenvalues. Among the other matrices that can be considered, we quote the Laplacian matrix. However, the eigenvalues of the Laplacian matrix and those of the adjacency matrix are related each other (see Zumstein 2005 for the mathematical relationship) and for this reason, in the present paper, we consider the adjacency matrix only.

Definition 1 Assume $V = \{1, \ldots, v\}$. The adjacency matrix $A(G)$ is the $v \times v$ matrix whose $ij - th$ entry is equal to 1 if $\{i, j\} \in E$, to 0 otherwise, that is

$$A(G)_{ij} = \begin{cases} 1 & \text{if } \{i, j\} \in E \\ 0 & \text{if } \{i, j\} \notin E \end{cases}.$$

As said before, the adjacency matrix of a graph contains the same information as the graph itself. For example, the sum of the elements on the i-th row of $A(G)$ is the valence or degree of the i-th vertex of G. Hence, the sum of the degrees of all the vertices is twice the number of edges, as $A(G)$ is symmetric. So, the average degree of the vertices of G is equal to $\frac{2e}{v} = 2\beta$.

A graph is said to be d-regular if every vertex has degree d. Other than the degree of a single vertex, one can consider the largest degree of a vertex as well as the average degree of the vertices of the graph.

Two graphs on the same set of vertices are isomorphic if one can be obtained from the other one by relabeling the vertices. The relabeling can be encoded in a permutation matrix P, an orthogonal $v \times v$ matrix, and the adjacency matrix of the new graph is $P^T A(G) P$. Then, the characteristic polynomial $p(t) = \det(A(G) - tI)$ is invariant under isomorphism, and therefore it is associated to G, more than to $A(G)$. For the same reason, the eigenvalues of $A(G)$ are called the eigenvalues of G, and the set of the eigenvalues is called the spectrum of the graph G.

Since G is undirected, the matrix $A(G)$ is symmetric and real-valued. So, its eigenvalues are real and they are v if counted with multiplicity. This is the content of the celebrated Spectral Theorem from linear algebra. We can assume that the eigenvalues of a graph G are ordered increasingly as $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_v$. In the next Theorem 2, we recall some known results for the spectrum of a graph G.

Theorem 2 Let G be a graph, and let Δ (resp. d) be the maximum (resp. average) degree of a vertex of G. Then, the largest eigenvalue λ_v of G satisfies $d \le \lambda_v \le \Delta$. If G is a $d-$ regular graph, i.e., each vertex has degree d, then $d = \lambda_v = \Delta$.

We assume that G has no loop. Hence, $A(G)_{ii} = 0$ for every i, and so $\lambda_1 + \ldots + \lambda_v = 0$. In particular, the sum of the positive eigenvalues is the opposite of the sum of the negative ones. Moreover, it can be proved that the range $[\lambda_1, \lambda_v]$ of the eigenvalues is contained in $[-\Delta, \Delta]$.

The spectrum of a graph G does not characterize the graph uniquely, i.e., there are non-isomorphic graphs with the same spectrum (Fig. 2).
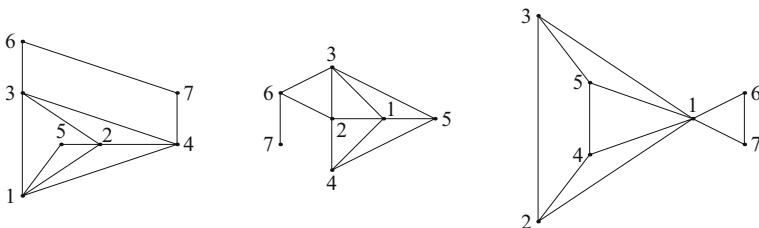

Fig. 2 Non-isomorphic graphs with the same spectrum but with different planar layout

This has obviously a negative impact on the graph analysis since there is no chance to reconstruct the graph from its spectrum. However, we have to consider also that the representation of a geographical region as a planar graph is not unique, since it is strongly affected by the construction rules one chooses.

One of the main open problems in algebraic graph theory is the meaning of the eigenvalues of a graph in terms of the geometrical properties of the graph itself. In particular, the largest eigenvalue λ_v and the difference between the two largest eigenvalues $g = \lambda_v - \lambda_{v-1}$, the so-called *gap*, have no precise interpretation, also if their importance is related to the study of expanders (Goldreich 2011; Dwork 2005). As for λ_v, also the gap of a regular graph can be calculated exactly. The theoretical exact computation of the gap is not possible, however, for every other graph. So, in general, one looks for bounds or estimates for both λ_v and g.

To compare eigenvalues and gaps of graphs with different numbers of vertices or edges, we proceed as follows: we look for graphs L (v,e) with special planar layouts with respect to the largest eigenvalue, we compute their largest eigenvalue Λ (v,e), and we use it to compute the relative version of the eigenvalue, namely $\frac{\lambda_v}{\Lambda}$ (v,e). We perform the same job for the gap too, and we call M (v,e) those graphs.

Now, we define the graph L (v,e) by giving its adjacency matrix.

Definition 3 Let v, e be integers with $v - 1 \leq e \leq 3v - 6$. Let M be the $v \times v$ matrix whose elements are

$$M_{ij} = \begin{cases} 1 & \text{if } (i,j) \in B_{ve} \\ 0 & \text{if } (i,j) \notin B_{ve} \end{cases}$$

where

- if $e = v - 1$, then

$$B_{ve} = \left\{ (1,j) \,\middle|\, j = 2, \ldots, v \right\};$$

- if $e = v + 2q$, $q \geq 0$, then

$$B_{ve} = \left\{ (1,i), \; (2,j), \; (2+h, 3+h) \,\middle|\, i = 2, \ldots, v, \; j = 3, \ldots, 3+q, \; h = 0, \ldots, q \right\};$$

- if $e = v - 1 + 2q$, $q \geq 1$, then

$$B_{ve} = \left\{ (1,i), \; (2,j), \; (1+h, 2+h) \,\middle|\, i = 2, \ldots, v, \; j = 3, \ldots, 3+q, \; h = 1, \ldots, q \right\}.$$

The adjacency matrix of L (v,e) is defined as A $(L$ $(v,e)) = M + M^T$.
In the following Fig. 3, we draw the graphs L $(6,5)$, L $(6,9)$ and L $(6,10)$.
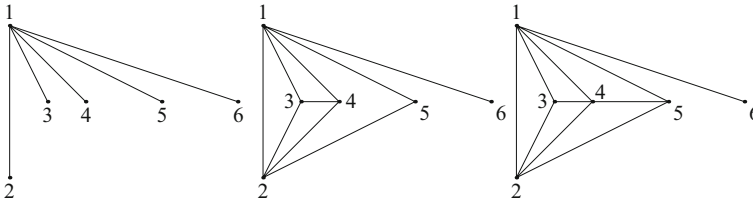
Fig. 3 Graphs that have the largest maximal eigenvalue for the planar graphs with the same number of vertices and edges

We checked that the largest eigenvalue $\Lambda\,(v,e)$ of $L\,(v,e)$ is the maximum of the largest eigenvalues of planar graphs with at most 7 vertices. Hence, we use it as a normalization factor.

Now, we define the graph $M\,(v,e)$ for $v \geq 6$.

Definition 4 Let v, e be integers with $v \geq 6$ and $v - 1 \leq e \leq 3v - 6$. Let M be a $v \times v$ matrix whose elements are

$$M_{ij} = \begin{cases} 1 & \text{if } (i,j) \in B_{ve} \\ 0 & \text{otherwise} \end{cases}$$

where

- if $e = v - 1$,   then $B_{ve} = \{(1,2), \dots, (1,v)\}$,
- if $v \leq e \leq 2v - 5$,   then $B_{ve} = \{(1,2), \dots, (1,v), (2,3), \dots, (2, e - v + 3)\}$,
- if $e = 2v - 4$,   then $B_{ve} = \{(i,3), \dots, (i,v) \mid i = 1,2\}$;
- if $e = 2v - 3$,   then $B_{ve} = \{(1,2), \dots, (1,v), (2,3), \dots, (2,v)\}$;
- if $2v - 2 \leq e \leq 3v - 8$,   then
  $B_{ve} = \{(1,2), \dots, (1,v), (2,3), \dots, (2, v - 1), (3,4), (4,5), \dots, (e - 2v + 6, e - 2v + 7)\}$;
- if $e = 3v - 7$,   then
  $B_{ve} = \{(1,2), \dots, (1,v), (2,3), \dots, (2,v), (3,4), (4,5), \dots, (v - 2, v - 1)\}$;
- if $e = 3v - 6$,   then
  $B_{ve} = \{(1,3), \dots, (1,v), (2,3), \dots, (2,v), (3,4), (4,5), \dots, (v - 1, v), (3,v)\}$.

The adjacency matrix of $M\,(v,e)$ is defined as $A\,(M\,(v,e)) = M + M^T$.

In the following Fig. 4, we draw the graphs $M\,(6,5)$, $M\,(6,7)$ and $M\,(6,10)$.

We checked that the gap $\Gamma\,(v,e)$ of the graph $M\,(v,e)$ is as large as possible for every (v,e) with $v = 6, 7$, and so we use it as a normalization factor.
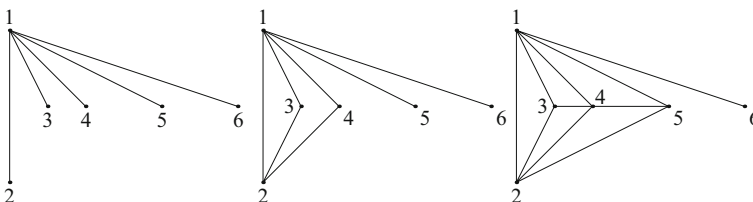
To summarise, the linear algebraic parameters are

Fig. 4 Graphs with the largest gap among planar graphs with the same number of vertices and edges

- d

  It is the average degree of the vertices of the graph G.
- λ

  It is the largest eigenvalue λ_v of the adjacency matrix of the graph G.
- Normalised λ

  It is the ratio between the largest eigenvalue λ_v and $\Lambda\ (v, e)$.
- g

  It is the difference g between the two largest eigenvalues of the adjacency matrix of the graph G.
- Normalised g

  It is the ratio between the gap g and $\Gamma\ (v, e)$.

2.4 Algebraic Parameters

In this paragraph, we introduce other parameters, that we call algebraic, more related to the geometry of the graph G under consideration. Herein, we define them from a computational point of view. In the Appendix 2, we define them from a mathematical point of view. For the interested reader, the parameters are defined in terms of the edge ideal associated to the graph, and the computations are performed in the framework of commutative algebra. As previously told, in Appendix 3 we give a computer session in which we compute the algebraic parameters for a given graph. We stress that the computation of codimension and degree for a graph is quite hard when the number of vertices increases (ninety vertices seems to be close to an upper bound).

2.4.1 Codimension

Let us consider a graph G. An edge is a couple of vertices, usually called the endpoints of the edge. A vertex meets an edge if the vertex is one of its endpoints. A vertex of a graph G meets all the edges of G if and only if G has a star shape. Hence, for graphs that are not star shaped, we need more than a single vertex to meet all the edges. Of course, if we consider all the vertices, they do meet all the edges, but not all of them are necessary. Hence, it is meaningful to look for the minimum number of vertices that meet all the edges of a given graph. We call codimension, codim (G), of the graph G this minimum number. Given two graphs with the same numbers of vertices and edges, but with different codimension, we would like to say that the graph having smaller codimension will have more vertices with high degree than the vertices of the graph with larger codimension. Unfortunately it not always the case, as the example in Fig. 5 shows.

As depicted in the previous paragraphs, the general strategy we try to follow is to introduce a numerical invariant associated to a graph G, and then to propose an upper bound for it. So we can compute the normalized invariant, better suited to compare graphs with different numbers of vertices and edges. As in the application we are concerned with planar connected graphs, we restrict ourselves to bounding the numerical invariants in this class of graphs. In such a class, we empirically computed two functions, according to the parity of the number of vertices, that give the maximal
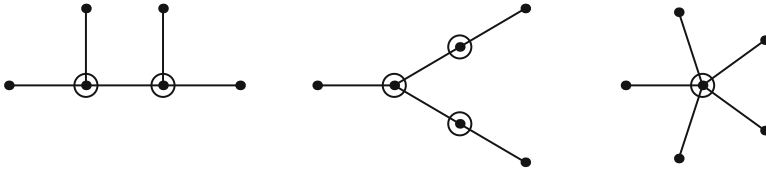
Fig. 5 Three graphs with $v = 6$, $e = 5$ and codimensions 2, 3, 1 from the left to the right. The degrees of the vertices are $(1, 1, 1, 1, 3, 3)$, $(1, 1, 1, 2, 2, 3)$, $(1, 1, 1, 1, 1, 5)$, respectively. The circled vertices are the ones that meet all the edges

codimension of a graph with a given number of vertices and edges. These functions are pinpointed on the exhaustive list of planar connected graphs with 5, 6, and 7 vertices.

Definition 5. Assume the number of vertices v to be even, i.e., $v = 2m$ for some $m \geq 3$. For $v - 1 \leq e \leq 3v - 6$ we define the function

$$codim\_max\,(v = 2m, e) = \begin{cases} m & \text{if } e = v - 1, v \\ m + k + 1 & \text{if } \begin{matrix} v + 1 \leq e \leq 3v - 11 \\ e - v - 1 = 4k + r \text{ with } 0 \leq r \leq 3 \end{matrix} \\ v - 2 & \text{if } e > 3v - 11 \end{cases}$$

Assume the number of vertices v to be odd, i.e., $v = 2m + 1$ for some $m \geq 2$. For $v - 1 \leq e \leq 3v - 6$ we define the function

$$codim\_max\,(v = 2m + 1, e) = \begin{cases} m & \text{if } e = v - 1 \\ m + 1 & \text{if } e = v, v + 1, v + 2 \\ m + k + 2 & \text{if } \begin{matrix} v + 3 \leq e \leq 3v - 11 \\ e - v - 3 = 4k + r \text{ with } 0 \leq r \leq 3 \end{matrix} \\ v - 2 & \text{if } e > 3v - 11 \end{cases}$$

Some examples of graphs with maximal codimension are shown in Fig. 6.

So, given a planar connected graph G with v vertices and e edges, we define its normalized codimension as

$$ncodim(G) = \frac{codim(G)}{codim\_max(v, e)}. \tag{4}$$

In graph theory, there are other indices concerning the minimum of some property. E.g., the distance between two vertices in a graph is shortest path connecting them. Even if it is not possible to compare the codimension with all those indices, we spend a few words about the spanning trees of a graph. We recall
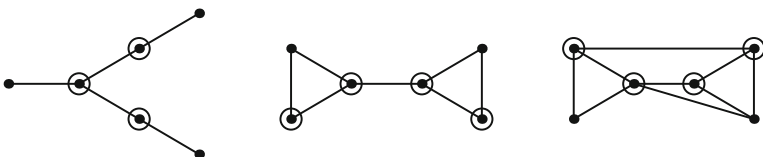

Fig. 6 Three graphs with 6 vertices, $e = 5, 7, 9$ edges, and maximal codimension

that a tree is a connected graph with no cycle, while a spanning subgraph G' of a graph G is a graph with $V(G') = V(G)$ and $E(G') \subseteq E(G)$. Combining the two notions, we get that a spanning trees G' of a graph G is a subgraph of G with $V(G') = V(G)$ and no cycle (see Godsil and Royle (2001)). The comparison between codimension and spanning trees is now evident: every spanning tree has the same vertices as the whole graph, and is a subgraph, while the codimension selects a subset of vertices and no edge.

2.4.2 Degree

Another parameter related to the codimension of a graph G is the degree $deg(G)$ of G. Before defining it, we stress that it is not related to the degrees of the vertices of G. It is defined as follows. At first, one considers all the subsets of $V(G)$ containing $codim(G)$ vertices exactly. Secondly, one discards the subsets containing vertices that do not meet all the edges. The number of non-discarded subsets is the degree $deg(G)$ of the graph. As example, we compute the degree of the graph G in Fig. 7.

The codimension of G is 2. Between the $\binom{6}{2} = 15$ subsets containing 2 vertices, only $\{2, 4\}$ meets all the vertices. In fact, if the subset contains neither 1 nor 2, then its vertices do not meet the edge 1–2 (we discard 6 subsets). Assume the subset is $\{1, j\}$. If $j = 2$ or 3, then they do not meet the edge 4–5. If $j = 4$, 5 or 6, then they do not meet the edge 2–3 (we discard 5 more subsets). Finally, the vertices 2 and 3 do not meet the edge 4–5, while the vertices 2,5 and 2,6 do not meet the edge 3–4 (we discard 3 more subsets). So, it remains the only subset $\{2, 4\}$ whose vertices meet all the edges, i.e., $deg(G) = 1$

We were not able to compute an upper bound for the degree of a graph in terms of the numbers of vertices and edges, and so the degree has not a normalized value.

2.4.3 Betti Numbers

When dealing with ideals generated by homogeneous polynomials in polynomial rings, it is natural to consider their graded Betti numbers. With abuse of notation, we call them Betti numbers of the graph, more than of its edge ideal. Even if for any edge ideal there are many Betti numbers, we consider only the first three. More precisely, we consider the number of edges, that is exactly the number of generators of the edge ideal, and we
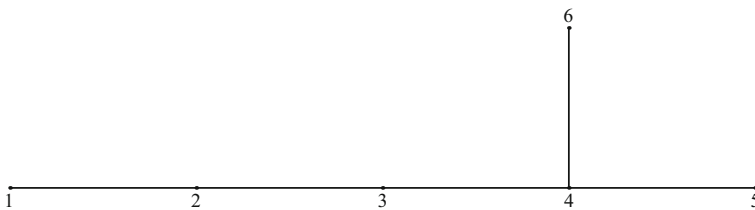

Fig. 7 The drawn graph has codimension 2 and degree 1: the only subset of 2 vertices that meet all the edges is $\{2, 4\}$

call it $b_{11}(G)$, and two others, that we call $b_{21}(G)$, $b_{22}(G)$, that depend on the geometry of the graph, and that we are going to define.

Let d_i be the degree of the i-th vertex, for each $i = 1, \ldots, v$, and let $t(G)$ be the number of independent triangles in G. Then

$$b_{21}(G) = \sum_{i=1}^{v} \frac{d_i(d_i-1)}{2} - t(G), \tag{5}$$

where a triangle is a complete sub-graph with 3 vertices.

Let e_1, e_2 be two edges. We say that they are close together if they have a common vertex, or if there exists an edge e_3 such that both e_1, e_3 and e_2, e_3 have a common vertex. We say that e_1, e_2 are far away if they are not close together. Finally, we define

$$b_{22}(G) = \text{number of couples of edges that are far away.} \tag{6}$$

The Betti numbers do not represent uniquely graphs, in the sense that it is possible to construct non-isomorphic graphs with the same Betti numbers. See, for example, the two graphs in Fig. 8.

They both have 7 vertices and 11 edges, $b_{21}(G) = 26$, $b_{22}(G) = 0$, but they are not isomorphic because in the one on the left there are two vertices of degree 1, while in the one on the right, there is only a vertex of degree 1.

Among the many graphs we studied, we did not find two non-isomorphic graphs with the same spectra and the same Betti numbers.

To compare Betti numbers of graphs with different numbers of vertices and of edges, we need a normalized version of the Betti numbers of a graph. Of course, $b_{11}(G) \leq 3v - 6$ as G is a planar connected graph. So the normalized value of b_{11} (G) is equal to γ. To normalise $b_{21}(G)$, we construct particular reference graphs for each number v of vertices and e of edges, we compute their Betti number $b_{21}(v, e)$, and we get the relative version as $\frac{b_{21}(G)}{b_{21}(v,e)}$. This approach is the same we have adopted in the previous paragraph (2.3) to get a normalised value of the maximal eigenvalue and of the gap of graphs. We did not succeed in finding the reference graphs for the Betti number $b_{22}(G)$, that remains an open task. The normalised Betti numbers can be considered as a measure of the difference of shape between the given graph and the one chosen as a reference.
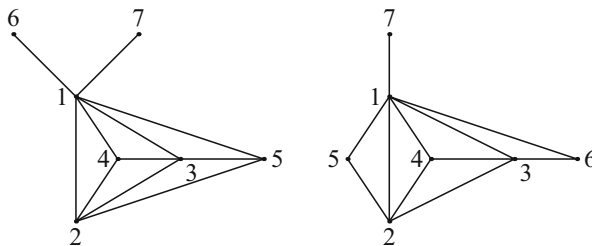
Fig. 8 Examples of non-isomorphic graphs with the same number of vertices, edges and Betti numbers b_{21} (G), $b_{22}(G)$

Definition 6 Let v, e be integers with $v - 1 \leq e \leq 3v - 6$. Let M be the $v \times v$ matrix whose elements are

$$M_{ij} = \begin{cases} 1 & \text{if } (i,j) \in B_{ve} \\ 0 & \text{if } (i,j) \notin B_{ve} \end{cases}$$

where

- if $e = v - 1$, then

$$B_{ve} = \left\{ (1,j) \,\middle|\, j = 2, \ldots, v \right\};$$

- if $v \leq e \leq 2v - 3$, then

$$B_{ve} = \left\{ (1,i), \ (2,j) \,\middle|\, i = 2, \ldots, v, \ j = 3, \ldots, e{-}v{+}3 \right\};$$

- if $2v - 2 \leq e \leq 3v - 6$, then

$$B_{ve} = \left\{ (1,i), \ (2,j), \ (2+h, 3+h) \,\middle|\, i = 2, \ldots, v, \ j = 3, \ldots, v, \ h = 1, \ldots, e{-}2v{+}3 \right\}.$$

The adjacency matrix of $H\,(v,e)$ is defined as $A\,(H\,(v,e)) = M + M^T$.

In the following Fig. 9, we draw the graphs $H\,(6,5)$, $H\,(6,8)$, $H\,(6,11)$, from the left to the right.

By using its definition, $b_{21}(v,e)$ of $H\,(v,e)$ is equal to

$$b_{21}(v,e) = \begin{cases} \dbinom{v-1}{2} + \dbinom{e-v+2}{2} & \text{if } v{-}1 \leq e \leq 2v{-}3 \\ v^2 - 9v + 10 + 3e & \text{if } 2v{-}2 \leq e \leq 3v{-}6 \end{cases}.$$

The graph $H\,(v,e)$ has the largest $b_{21}(G)$ for $v \leq 7$ and every e among the planar graphs with the same numbers of vertices and edges. Hence, we use $b_{21}(v,e)$ as a normalization factor.

Summarizing the content of the paragraph, we defined the following algebraic parameters:

Fig. 9 H(6,5) on the left, H(6,8) in the center, H(6,11) on the right, to illustrate the planar layout of the reference graphs for the Betti number $b_{21}(G)$

- $codim(G)$ and its normalization $ncodim(G)$

    $codim\ (G)$ (codimension) is the minimum number of vertices that meet all the edges of G and its normalization is the ratio between its value and the reference function $codim - \max(v, e)$.
- $\deg(G)$

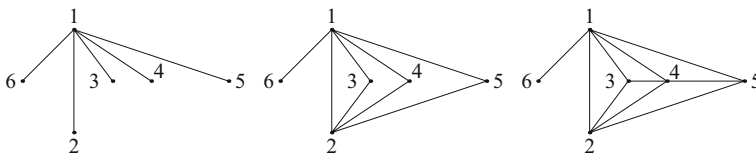 $\deg(G)$ (degree) is the number of subsets of vertices having cardinality $codim$ (G), which meet all the edges of G.
- $b_{11}(G)$ and its normalization γ

    $b_{11}(G)$ is the number of edges of the graph G.
- $b_{21}(G)$ and its normalization $b_{21}(G)/b_{21}(v, e)$

    $b_{21}(G)$ is related to the degrees of the vertices and to the number of independent triangles of the graph. Its normalized value is obtained by means of the function $b_{21}(v, e)$.
- $b_{22}(G)$

    $b_{22}(G)$ is the number of couples of far away edges in the graph.

3 The Application Scenarios

In this section we present the set of graphs used for the analysis and the application of the previous described indicators.

Three types of scenario are taken into consideration: the first one concerns all planar connected undirected graphs with four up to seven vertices, the second one concerns 41 graphs representing a functional description of operating underground networks (33 graphs are taken from Derrible and Kennedy (2010) and the remaining eight ones are their variations[1]), the third and last one concerns the underground network of Milan, analysed during its development along the last five decades.

In Table 1, we report the average values of the algebraic parameters *codim*, *degree*, b_{21}, b_{22} and of the linear algebraic parameters, maximum eigenvalue λ_v and the gap, and the average values of their normalizations, when computed, for each set of graphs with fixed number of vertices and edges. These 771 graphs represent the complete set of planar connected graphs with v vertices, $4 \leq v \leq 7$, and were computed by complete enumeration from the set of all graphs with that number of vertices. Since this approach is particularly time and memory consuming we limited the scenario to graphs with up to 7 vertices. We remark that the parameters α, β, γ are constant on each set of graphs with given number of vertices and edges (see Eqs. 1, 2, and 3).

Tables 2 and 3 reports the list of networks of the second scenario with the number of vertices and edges. 33 of these graphs were built up by Derrible and Kennedy (2010) to functionally represent the underground networks of some famous cities all over the world. The way they are carried out is to consider only stations with at least three entering/exiting connections; other stations are discarded since they do not modify the connectivity structure of the network. Variants of 8 graphs from the above ones are also

---

[1] The changes mainly consist in the insertion of one vertex for a different interpretation of the underground network structure. Other changes concern other eight graphs that are transformed in order to be planar.

Table 1 Graphs per number of vertices and edges for the first scenario (all graphs from 4 up to 7 vertices) with their algebraic parameters (*norm.* stays for *normalized*)

No. of vertices	No. of edges	No. of graphs	Average degree	Average b22	Average codim	Average norm. codim	Variancenorm. codim	Average b_{21}	Average norm. b_{21}	Variance norm. b_{21}	Total no. of graphs
4	3	2	2.000	0.000	1.50	0.750	0.125	2.500	0.8333	0.056	6
	4	2	2.000	0.000	2.00	1.000	0.000	4.000	1.0000	0.000	
	5	1	1.000	0.000	2.50	1.000	0.000	6.000	1.0000	0.000	
	6	1	4.000	0.000	3.00	1.000	0.000	8.000	1.0000	0.000	
5	4	3	1.333	0.333	1.60	0.833	0.083	4.333	0.7222	0.065	20
	5	5	2.800	0.200	2.00	0.800	0.033	5.800	0.8286	0.014	
	6	5	2.800	0.200	2.40	0.866	0.033	8.400	0.9333	0.004	
	7	4	2.500	0.000	2.80	0.916	0.028	11.250	0.9375	0.002	
	8	2	2.000	0.000	3.20	1.000	0.000	14.000	1.0000	0.000	
	9	1	1.000	0.000	3.60	1.000	0.000	17.000	1.0000	0.000	
6	5	6	2.333	1.000	1.66	0.722	0.063	6.167	0.6167	0.046	99
	6	13	2.538	1.231	2.00	0.897	0.026	7.769	0.7063	0.015	
	7	19	2.368	1.105	2.33	0.723	0.013	10.474	0.8057	0.009	
	8	22	2.682	0.682	2.66	0.772	0.017	13.682	0.8551	0.005	
	9	19	3.053	0.561	3.00	0.828	0.021	17.000	0.8527	0.004	
	10	13	3.538	0.154	3.33	0.903	0.016	20.462	0.9301	0.001	
	11	5	3.400	0.200	3.66	0.950	0.013	24.167	0.9680	0.000	
	12	2	3.000	0.000	4.00	1.000	0.000	28.000	1.0000	0.000	
7	6	11	1.818	2.545	1.71	0.818	0.053	8.000	0.5333	0.036	646
	7	33	2.970	3.091	2.00	0.757	0.021	9.697	0.6061	0.014	
	8	67	3.090	3.134	2.28	0.817	0.022	12.493	0.6940	0.009	
	9	107	3.178	2.636	2.57	0.873	0.019	15.832	0.7539	0.006	
	10	130	2.977	2.231	2.85	0.735	0.011	19.435	0.7775	0.004	

Table 1 (continued)

No. of vertices	No.of edges	No. of graphs	Average gap	Average norm.gap	Variance norm.gap	Average λ_v	Average norm. λ_v	Variance norm. λ_v	Total no. of graphs
	11	130	2.777	1.738	3.14	0.767	0.009	23.256	0.7756
	12	96	2.680	1.402	3.42	0.795	0.007	27.225	0.8544
	13	51	2.980	1.059	3.71	0.835	0.008	31.448	0.9003
	14	16	3.375	0.938	4.00	0.875	0.010	35.724	0.9441
	15	5	4.750	0.250	4.28	0.950	0.010	40.286	0.9817
4	3	2	1.366	0.788	0.089	1.675	0.967	0.002	6
	4	2	1.930	0.964	0.002	2.085	0.960	0.003	
	5	1	2.562	1.000	0.000	2.562	1.000	0.000	
	6	1	4.000	1.000	0.000	3.000	1.000	0.000	
5	4	3	1.272	0.635	0.107	1.860	0.930	0.005	20
	5	5	1.525	0.814	0.019	2.199	0.938	0.003	
	6	5	2.014	0.822	0.024	2.564	0.954	0.001	
	7	4	2.627	0.875	0.009	2.969	0.962	0.001	
	8	2	3.101	0.958	0.003	3.280	0.986	0.000	
	9	1	3.646	1.000	0.000	3.646	1.000	0.000	
6	5	6	1.115	0.498	0.071	1.991	0.890	0.005	99
	6	13	1.315	0.676	0.029	2.281	0.907	0.003	
	7	19	1.621	0.709	0.025	2.602	0.924	0.002	
	8	22	2.076	0.733	0.019	2.926	0.921	0.002	
	9	19	2.442	0.729	0.016	3.229	0.933	0.002	
	10	13	2.798	0.855	0.007	3.528	0.950	0.001	
	11	5	3.155	0.890	0.005	3.816	0.979	0.000	
	12	2	3.751	0.937	0.008	4.060	0.985	0.000	
7	6	11	0.996	0.406	0.053	2.087	0.851	0.005	646

Table 1 (continued)

7	33	1.146	0.561	0.026	2.349	0.875	0.003
8	67	1.408	0.617	0.024	2.629	0.892	0.002
9	107	1.732	0.652	0.020	2.914	0.890	0.002
10	130	2.026	0.640	0.015	3.181	0.912	0.002
11	130	2.329	0.629	0.011	3.447	0.913	0.001
12	96	2.614	0.760	0.011	3.699	0.936	0.001
13	51	2.899	0.807	0.008	3.948	0.914	0.001
14	16	3.191	0.802	0.004	4.181	0.970	0.000
15	5	3.529	0.911	0.006	4.426	0.979	0.000

Table 2 Number of vertices, edges and some results for the graphs of the second scenario – underground of some relevant cities (*norm.* stays for *normalized*)

No.	Network	Vertices	Edges	Alpha	Beta	Gamma	Degree	b_{22}	Codim	Norm. codim	b_{21}	Norm. b_{21}	Gap	Norm. gap	λ_v	Norm. λ_v
1	Athens	9	9	0.077	1.000	0.429	1	8	3	0.600	14	0.483	1.070	2.268	2.58	0.860
2	Athens_modified	11	12	0.118	1.091	0.444	2	15	4	0.667	23	0.479	1.070	2.504	2.92	0.843
3	Barcelona	29	42	0.264	1.448	0.519	2	507	14	0.778	109	0.226	1.030	4.641	4.16	0.661
4	Berlin	32	43	0.203	1.344	0.478	6	584	14	0.737	104	0.192	0.470	4.371	3.56	0.563
5	Boston	21	22	0.054	1.048	0.386	5	129	8	0.727	41	0.212	0.800	3.420	2.97	0.644
6	Boston_modified	22	24	0.077	1.091	0.400	5	165	9	0.750	44	0.204	0.490	3.409	2.98	0.620
7	Brussels	9	9	0.077	1.000	0.429	4	6	4	0.800	14	0.483	1.120	2.268	2.59	0.863
8	Bucharest	11	12	0.118	1.091	0.444	5	16	5	0.833	21	0.438	1.090	2.504	2.77	0.800
9	Buenos Aires	12	13	0.105	1.083	0.433	3	8	4	0.571	31	0.534	1.430	2.588	3.16	0.881
10	Cairo	6	5	0.000	0.833	0.417	1	0	2	0.667	6	0.600	1.000	2.236	2.00	0.894
11	Cairo_modified	7	7	0.111	1.000	0.467	1	0	2	0.500	12	0.750	1.560	2.039	2.56	0.955
12	Chicago_modified	25	30	0.133	1.200	0.435	2	255	10	0.714	67	0.226	0.700	3.621	3.39	0.634
13	Chicago	24	28	0.116	1.167	0.424	8	215	10	0.769	61	0.228	0.840	3.526	3.31	0.639
14	Delhi	8	7	0.000	0.875	0.389	1	0	2	0.500	12	0.571	1.000	2.646	2.30	0.869
15	Hong Kong	17	18	0.069	1.059	0.400	21	81	8	0.889	30	0.244	0.490	3.046	2.75	0.658
16	Lisbon	11	11	0.059	1.000	0.407	2	6	4	0.667	21	0.457	1.250	2.510	2.67	0.810
17	London	87	121	0.207	1.398	0.475	96	6360	41	0.788	286	0.067	0.342	6.854	4.04	0.400
18	Lyon	10	10	0.067	1.000	0.417	3	4	4	0.800	18	0.486	1.340	2.389	2.60	0.825
19	Madrid	73	124	0.369	1.699	0.582	744	6293	38	0.760	369	0.094	0.274	8.711	4.52	0.444
20	Marseille	6	5	0.000	0.833	0.417	1	0	2	0.667	6	0.600	1.000	2.236	2.00	0.894
21	Marseille_modified	7	7	0.111	1.000	0.467	1	0	2	0.500	12	0.750	1.560	2.039	2.56	0.955

Table 3 Number of vertices, edges and some results for the graphs of the second scenario – underground of some relevant cities (norm. means normalized)

No.	Network	Vertices	Edges	Alpha	Beta	Gamma	Degree	b22	Codim	Norm.codim	b21	Norm. b21	Gap	Norm. gap	λ_v	Norm. λ_v
22	Mexico City	35	52	0.277	1.486	0.525	5	901	16	0.727	134	0.183	0.620	5.168	3.78	0.547
23	Milan	14	15	0.087	1.071	0.417	6	36	6	0.750	28	0.346	1.210	2.766	3.06	0.799
24	Montreal	10	10	0.067	1.000	0.417	3	4	4	0.800	18	0.486	1.340	2.389	2.60	0.825
25	Montreal_modified	11	12	0.118	1.091	0.444	2	7	4	0.667	25	0.521	1.410	2.504	2.92	0.843
26	Moscow	43	66	0.296	1.535	0.537	278	1526	22	0.786	181	0.156	0.593	5.885	4.03	0.525
27	New York	86	136	0.296	1.563	0.533	20	8002	43	0.768	352	0.072	0.233	8.214	4.05	0.383
28	Osaka	36	51	0.239	1.417	0.500	29	876	17	0.773	127	0.174	0.610	4.880	3.78	0.553
29	Paris	104	165	0.305	1.587	0.539	1544	11892	50	0.735	471	0.065	0.383	8.964	4.40	0.380
30	Prague	9	9	0.077	1.000	0.429	1	0	3	0.600	17	0.586	1.730	2.268	2.73	0.910
31	Rome	5	4	0.000	0.800	0.444	1	0	1	0.500	6	1.000	2.000	2.000	2.00	1.000
32	Seoul	78	124	0.311	1.590	0.544	2748	6565	40	0.784	322	0.079	0.393	7.951	4.04	0.398
33	Shanghai	22	28	0.179	1.273	0.467	2	164	9	0.692	69	0.290	1.450	3.590	3.73	0.715
34	Singapore	12	13	0.105	1.083	0.433	3	23	5	0.714	23	0.397	1.110	2.588	2.85	0.794
35	Stockolm	20	19	0.000	0.950	0.352	7	110	8	0.800	28	0.164	0.280	4.359	2.43	0.557
36	Stockolm_modified	20	20	0.029	1.000	0.370	1	110	7	0.700	34	0.198	0.730	3.535	2.83	0.639
37	St Petersburg	14	16	0.130	1.143	0.444	1	32	5	0.625	34	0.405	1.310	2.806	3.16	0.790
38	Tokyo	131	233	0.401	1.779	0.602	**	24103	67	0.736	758	0.055	0.415	12.302	5.10	0.377
39	Toronto	10	9	0.000	0.900	0.375	3	15	4	0.800	11	0.306	0.500	3.000	2.15	0.717
40	Toronto_modified	10	10	0.067	1.000	0.417	3	11	4	0.800	16	0.432	1.120	2.389	2.63	0.835
41	Washington D.C.	17	18	0.069	1.059	0.400	1	70	6	0.667	33	0.268	1.150	3.046	2.98	0.713

**the used software was not able to calculate this parameter due to the dimension of the graph

taken into consideration (shown with the label 'modified' in Tables 2 and 3) and add to the list. Non planar graphs are transformed into planar ones. It must be underlined that in this scenario the set of graphs has not the same homogeneity and comprehensiveness of the first one since it presents cases rather different each other. However, this scenario allows us to appreciate how the normalization of some parameters works when comparing so different graphs. It must be underlined that also in this scenario the computation of algebraic parameters (e.g., *codim*, degree, b_{21} and b_{22}) is quite challenging and time consuming for graphs with a large number of vertices.

In Table 4, we report the underground network of Milan. We have considered the different construction steps in the last decades (from 1960s up to nowadays), so to mimic an increasing complexity of a transportation network. In detail, in the rows, from the first to the fifth one, we reported the parameters of the only red line, of the red and yellow lines, of the red, yellow and green ones, of the red, yellow, green and lilac lines, and finally, adding the railway bypass, of the whole network. This sequence with the lines for each step is drawn in Fig. 10.

4 Results

The first scenario includes all possible simple, connected, undirected planar graphs by number of vertices (from 4 to 7) and edges, and allows us to investigate whatever relationship between variables. In Table 1, for the reason of summary, only the average values per classes (obtained by considering the couples of number of vertices and edges) together with the variance of normalised algebraic parameters are reported.

From Table 1 the following results can be extracted:

– the average of algebraic parameters and the average of normalised ones increase with the number of vertices and edges;
– the average of the normalised values for each class is progressively far from 1 when the number of vertices and edges decrease;
– the variance of the normalised algebraic parameters is generally very low (about the 3 % of the average values) and becomes lower when the number of vertices and edges increases.

The correlations between the algebraic parameters and the number of vertices are always positive and quite large (greater than 0.5); on the opposite the same correlations for the normalized parameters are always negative, though they are lower and around −0.2. The same behaviour occurs when the correlation with degree and b_{22} is analysed. *Codim* is always negatively correlated with the other algebraic parameters. The number of edges is always positively correlated with all variables except with b_{22}. In general correlations between algebraic parameters and classic parameters are very large, in the range 0.7–0.9, but they are much lower when considering the normalised algebraic parameters, in the range 0.1–0.4. The maximum value of the algebraic parameters is highly correlated with all other variables except with b_{22}.

The third column of Table 1 gives the number of different graphs with the same number v of vertices and e of edges. While the classical parameters are constant in each class, as they depend on v and e only, both the linear algebraic, and the algebraic ones

Table 4 Number of vertices, edges and some results for the graphs of the third (Milan) scenario (norm. means normalized)

No.	Network	Vertices	Edges	Alpha	Beta	Gamma	Degree	b_{22}
1	Red Line	4	3	0.000	0.750	0.500	1	0
2	+ Yellow Line	7	6	0.000	0.857	0.400	1	0
3	+ Green Line	16	17	0.074	1.063	0.405	2	58
4	+ Lilac Line	21	24	0.108	1.143	0.421	1	142
5	+ railway bypass (whole network)	39	53	0.205	1.359	0.477	9	942

No.	Network	Codim	Norm. codim	b_{21}	Norm. b_{21}	Gap	Norm.gap	λ_v	Norm.- λ_v
1	Red Line	1	0.333	3	1.000	1.732	1.000	1.73	1.000
2	+ Yellow Line	2	0.667	9	0.600	1.049	0.428	2.18	0.888
3	+ Green Line	6	0.667	31	0.295	1.143	0.387	3.07	0.756
4	+ Lilac Line	8	0.667	49	0.245	1.020	0.308	3.25	0.676
5	+ railway bypass (whole network)	17	0.739	135	0.164	0.763	0.160	3.96	0.570
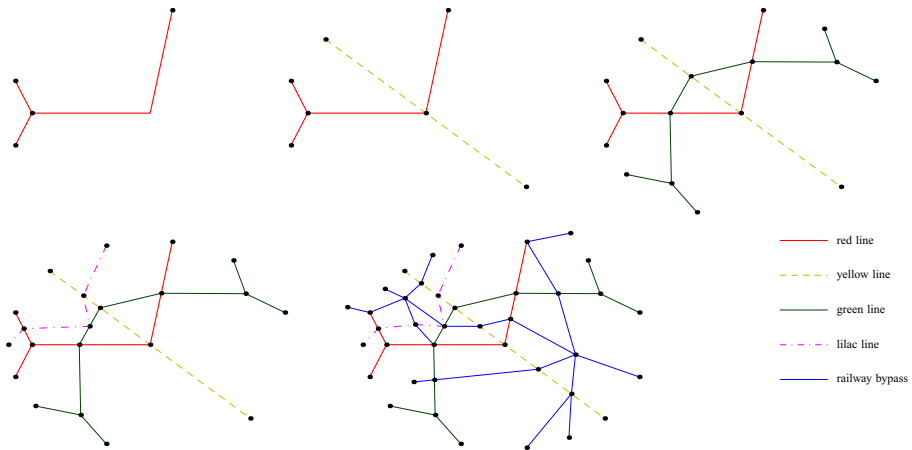
Fig. 10 The five graphs describing the increasing complexity of Milan underground network along the last decades

vary in each class. Hence, they are topological sensitive. Moreover, naively, between graphs with the same v, but with different e, the ones with more edges have a greater complexity. The averages of the proposed indicators, both absolute and normalized, show a good level of monotonicity with respect to e. So, we can conclude that the proposed indicators can measure the complexity of planar connected undirected graphs.

For what concerns the second scenario we present the relationships between the normalised algebraic parameters and the number of vertices in Figs. 11, 12, 13, and 14. In Fig. 11 this relationship for *codim* is drawn. This is the only case where the normalised parameter increases with the number of vertices.

In the other cases, that is for b_{21}, gap, and λ_v, the values decrease to zero with the number of vertices except for the max eigenvalue which has a lower bound at about 0.4.

The correlation of algebraic parameters is generally very different and opposite in sign from the correspondent normalised ones. In fact, algebraic parameters are highly
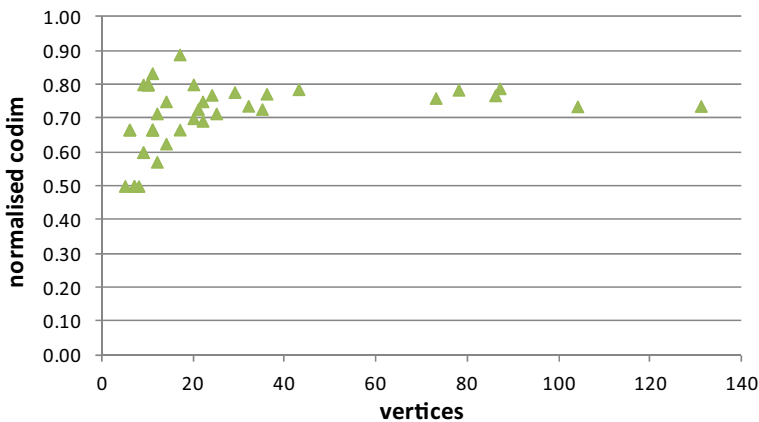

Fig. 11 Relationship between the number of vertices and normalized *codim* for planar connected graphs of undergrounds of some relevant cities (second scenario)
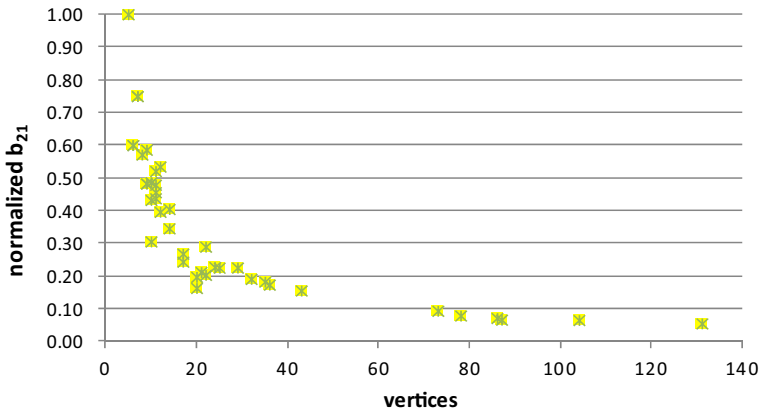
Fig. 12 Relationship between the number of vertices and normalized b_{21} for planar connected graphs of undergrounds of some relevant cities (second scenario)

positively correlated with classical parameters; by contrast, normalised algebraic parameters are highly negatively correlated with all the other variables except between themselves (in these cases the correlation is greater than 0.9).

This second scenario proves that the indicators we consider are suitable for measuring the complexity not only of planar connected undirected graphs, but also of real world transportation graphs. While they have an evident asymptotic behaviour as the number v of vertices, and so the complexity, increases, they give also different values for underground networks with (almost) the same size. Hence, they actually measure the changes in the topology of the given transportation graphs. E.g., compare Athens and Brussels, or Athens-modified, Bucharest, and Montreal-modified.

In Fig. 15 the relationships between the normalised algebraic parameters and the number of vertices of the third scenario are shown. Though the number of points is considerably lesser than in the previous scenario, trends for normalised algebraic parameters is very similar to that observed for the second scenario. Normalised *codim* increases with the number of vertices while the other normalised parameters decrease.
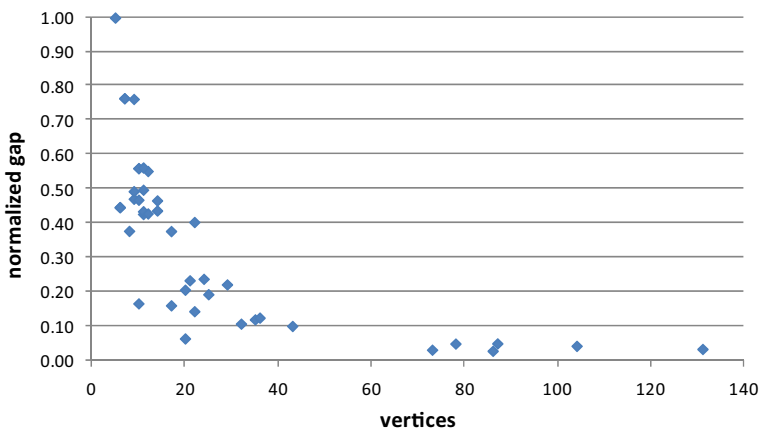
Fig. 13 Relationship between the number of vertices and normalized *gap* for planar connected graphs of undergrounds of some relevant cities (second scenario)
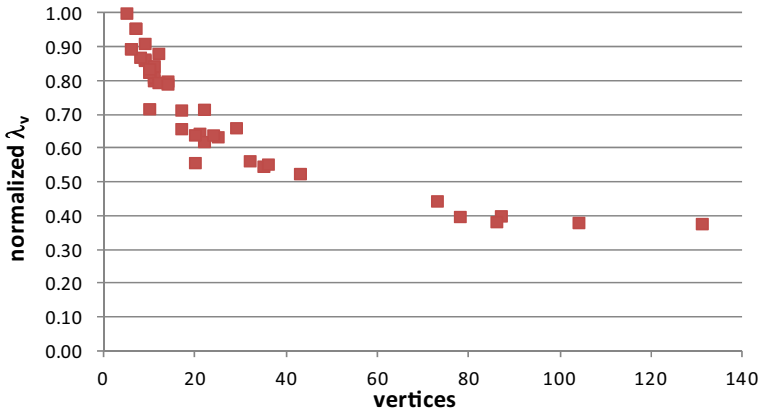
Fig. 14 Relationship between the number of vertices and normalized λ_v for planar connected graphs of undergrounds of some relevant cities (second scenario)

Correlation between normalised parameters and the other variables is almost always negative except for *codim* in which case it is positive. Results are indeed very similar to those of the second scenario and these examples confirm the opportunity of using the new proposed parameters in order to analyse transportation graphs.

5 Conclusions

The need of a more refined knowledge on transportation networks and their underlying graphs, on properties lying under their structure is a still challenging task, especially because we can observe their behaviour only when interacting with demand which is generally highly dynamical and changes day by day and within day. Since each transportation network can be assumed to have different performance, the aim is to understand and to separate how much of that
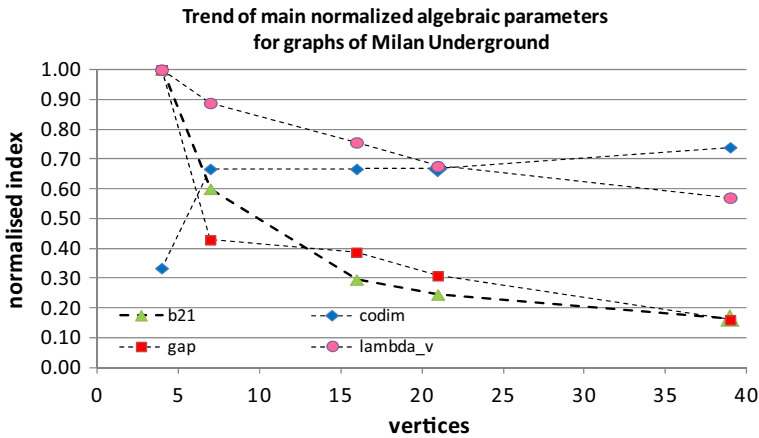

Fig. 15 Relationship between the number of vertices and normalized algebraic parameters for graphs of Milan underground network (third scenario)

performance depends on transportation demand and how much can be attributed to the particular structure of the network. In this sense the studies of graph properties can be considered a first step in this direction.

In this paper, we present some new parameters for investigating the graph properties together with some well-known others. The introduction of these new parameters aims at explaining better the complexity of graphs which is not simply related to the number of vertices and/or of edges, but how the edges topologically link vertices. In order to take into account this aspect and to overcome the problem of comparing graphs with different numbers of vertices or edges, we propose also to use their normalised values. To this aim we introduce some new definitions in order to calculate the maximum value of parameters for a graph given the number of vertices and edges.

In more detail, we consider the algebraic parameters (codimension, degree and Betti numbers) that are computed from the edge ideal associated to the graph, and the linear algebraic ones (average degree of vertices, maximum eigenvalue and gap) that are computed from the adjacency matrix of the graph. We remark that the edge ideal can be computed from the adjacency matrix and conversely.

We use three scenarios to analyse the behaviour of the proposed parameters together with the classical ones. The first scenario (all the 771 planar graphs with number of vertices from 4 up to 7) gives some comprehensive answers about the relationships between variables. The problem is that real transport graphs generally have a larger number of vertices and therefore results are not so simply transferable and generalizable. To overcome this problem a second scenario is considered. This second scenario is made up by 41 graphs derived from underground networks and offers more realistic cases. The third scenario represents the evolution in time of the underground network of Milan (in Northern Italy) and mimics the increasing complexity of a network.

Results show that the normalised parameters are able to describe the increasing complexity of the graph underlying an underground transportation network. Correlation of these normalised parameters with other variables (classical and non-normalised parameters) is low, meaning that their information content is rather different. It must be stressed that at this stage of research the territorial component of the transportation network cannot be taken into consideration and some further results by considering metric graphs are forthcoming in a near work. However, the comparison between the results of the second and third scenario (related to real world) and the first one (related to the complete list of graphs with 4 up to 7 vertices) allows us to draw some conclusions about increasing complexity of real networks. We observe that while network size increases, normalized algebraic parameters converge to a constant value. A possible explanation is that the way the network growths tends to repeat, by addition, similar basic structures and not to realize more complex structures, or at least not to realize the most complex structure achievable by a given number of vertices.

Future research will concern the extension of the proposed approach to directed, weighted and metric graphs which can represent a transportation network with a higher detail. Besides, we think to extend the methodology and to enhance computation tools to graphs of great dimensions (like the whole road network of Milan) to be able to face also the analysis of other types of transportation networks.

Appendix 1: A Short Review on Polynomial Ideals

Let K be a numerical field. Examples of such objects are the fields of rational numbers Q, of real numbers R, and of complex ones C. Let $x_1, \ldots, x_n$ be distinct unknowns. A term is a product of powers of the unknowns, i.e., $\underline{x}^a = x_1{}^{a_1} \ldots x_n{}^{a_n}$ with a_i non-negative integer. By definition, if $a_i = 0$ for every i=1, …,n, we write $1 = \underline{x}^0$. A monomial is the product of an element of K and a term, e.g., $c_a \underline{x}^a$ with $c_a \in K$. We set T the set of terms, with the product as operation. Finally, a polynomial is a finite sum of monomials with different terms. Sum and product of two polynomials is defined as usual. The set of polynomials with coefficients in K, with the operations of sum and product, is referred to as the polynomial ring $A = K[x_1, \ldots, x_n]$ (see Lang 2002).

Definition A1. A non-empty subset I of A is an ideal if the following conditions hold:

1. $P - Q \in I$ for every $P, Q \in I$;
2. $P Q \in I$ for every $P \in I$ and every $Q \in A$.

    Remark. If I is an ideal, the null polynomial 0 is in I. In fact, $0 = P - P \in I$. Moreover, if $P \in I$, then $-P \in I$ too, because $-P = 0 - P$ and then both belong to I.
    Lemma A2. Let $P_1, \ldots, P_s$ be polynomials. Then $<P_1, \ldots, P_s> = \{Q_1 P_1 + \ldots + Q_s P_s | Q_i \in A\}$ is the smallest ideal containing $P_1, \ldots, P_s$.
    Proof. It is a straightforward computation, and we leave it to the reader.

Definition A3 The ideal $<P_1, \ldots, P_s>$ is called the ideal generated by $P_1, \ldots, P_s$, and if we set $I = <P_1, \ldots, P_s>$, the polynomials $P_1, \ldots, P_s$ are called generators of I. Furthermore, the ideal I is finitely generated.
    A very deep result concerning the ideals in the polynomial rings is the following one.

Theorem A4 Every ideal in A is finitely generated.
    For the proof, see Erath et al. (2007).
    Theorem A4, first proved by D. Hilbert, is also stated as: A is a Noetherian ring, where a ring is said to be Noetherian exactly if every ideal in such a ring is finitely generated.
    To better understand ideals in A and to define numerical invariants for each ideal, we define two special kinds of ideals.

Definition A5 The ideal $I \subset A$ is prime if $PQ \in I, P \notin I \Rightarrow Q \in I$. Furthermore, the ideal J is primary if $P Q \in J, P \notin J \Rightarrow Q^n \in J$ for a suitable integer n.

Definition A6 Let $I \subset A$ be an ideal. The radical of I is the ideal $\sqrt{I}$ containing the polynomials P such that $P^n \in I$ for some integer n.

Proposition A7 The radical of a primary ideal is a prime ideal.

Proof. Let I be the primary ideal and let J be its radical.

At first, we observe that if $Q \notin J$ then $Q^m \notin J$ for every integer m, by Definition A5.

Assume now that P, Q are polynomials such that $PQ \in J$ and $Q \notin J$. From Definition A6, it follows that $P^n Q^n \in I$ for some integer n, but $Q^n \notin I$. Hence, again from Definition A5, there exists an integer h such that $P^{nh} \in I$ and so $P \in J$, as claimed. Q.E.D.

Primary and prime ideals in A play a very special role, as the following result shows.

Theorem A8 For every ideal $I \subset A$ there exist $J_1, \ldots, J_t$ primary ideals, with different radicals, such that $I = J_1 \cap J_2 \cap \ldots \cap J_t$.

The previous Theorem, known as primary decomposition of an ideal, shows that the elementary ideals with which one can construct all the ideals in A are the primary ones.

The proof of Theorem 8 can be found in Erath et al. (2007).

Now, we can define the first numerical invariant we are interested in.

Definition A9 A chain of length r of prime ideals is a sequence of prime ideals $I_0 \subset I_1 \subset \ldots \subset I_r$ with proper inclusions.

Definition A10 Let I be a prime ideal. The codimension *codim* (I) of I is the maximal length of a chain of prime ideals with I as last item of the chain. If I is not a prime ideal then its codimension *codim* (I) is the minimum of the codimensions of the radicals of the ideals that appear in a primary decomposition of I.

A linear ideal of codimension r is an ideal generated by r linearly independent degree 1 polynomials.

We introduce some tools for making computations with ideals.

Definition A11 A term-ordering $>$ is a well-ordering on T compatible with the product of terms, that is to say, if $\underline{x}^a > \underline{x}^b$ then $\underline{x}^a \underline{x}^c > \underline{x}^b \underline{x}^c$ for every $\underline{x}^c \in T$. The leading or initial term *in* (P) of a polynomial P is the largest term with respect to $>$ with non-zero coefficient in P. The initial ideal *in* (I) of the ideal I is the ideal generated by the initial terms of the elements of I.

A first relation among the defined concepts is the following one.

Theorem A12 An ideal I has codimension n if, and only if, $T \setminus \{\underline{x}^a \in in\ (I)\}$ is a finite set.

Definition A13 The degree deg(I) of an ideal of codimension n is the cardinality of $T \setminus \{\underline{x}^a \in in\ (I)\}$.

Proposition A14 Let I be an ideal of codimension *codim* $(I) < n$. Then, the degree of $I + L$ for a general linear ideal L of codimension *codim* $(L) = n - codim\ (I)$ does not depend on L.

So, we can define the degree of an ideal of whatever codimension.

Definition A15 The degree deg (I) of an ideal I is the degree of $I + L$, where L is a general linear ideal of codimension $n - codim\ (I)$.

To define the next numerical data associated to an ideal, we need the notion of free module.

Definition A16 Let t be an integer. The free A-module of rank t is A^t with the following operations

1. $(f_1, \ldots, f_t) + (g_1, \ldots, g_t) = (f_1 + g_1, \ldots, f_t + g_t)$;

2. $f(f_1, \ldots, f_t) = (ff_1, \ldots, ff_t)$ with $f \in A$.

Definition A17 A map $F : A^s \to A^t$ is A-linear if both $F(u+v) = F(u) + F(v)$ for every $u, v \in A^s$ and $F(fu) = f\,F(u)$ for every $f \in A$ and $u \in A^s$.

The free A-modules and A-linear maps between free A-modules are a generalization of vector spaces and linear maps between them. In that context, the basic object is a field, in ours, it is a ring.

As for linear maps between vector spaces, associated to F there is its kernel and its image, where $\ker(F) = \{u \in A^s|\ F(u) = 0\}$ and $Im(F) = \{F(u)|\ u \in A^s\}$.

Of course, both the kernel and the image are sub-modules of the right free A-module. A sub-module is a non-empty subset closed for both the sum of its elements and the product of its elements and polynomials of A. Furthermore, every A-linear map of free A-modules can be written as the product of a suitable matrix and the element of A^s as a column.

The free A-module of rank 1 is the polynomial ring A, and the sub-modules of A are exactly its ideals. As for ideals in a polynomial ring, the sub-modules of a free A-module are finitely generated, too.

Proposition A18 Let $I \subset A$ be the ideal generated by $P_1, \ldots, P_s$. Then, there exists a map $F : A^s \to A$ such that $I = Im(F)$.

Proof. Let $u = (f_1, \ldots, f_s) \in A^s$ and define $F(u) = f_1 P_1 + \ldots + f_s P_s$. It is easy to check that F is a A-linear map, and that $Im(F) = I$. Q.E.D.

Definition A19 Every element in $\ker(F)$ is called syzygy of I, and $\ker(F)$ is called the 1st syzygy module of I. The minimal number of generators of I is called 1st Betti number of I, and the minimal number of generators of the 1st syzygy module of I is called 2nd Betti number of I.

By definition, a syzygy is a collection of polynomials $f_1, \ldots, f_s$ such that $f_1 P_1 + \ldots + f_s P_s = 0$ where the equality means that on the left side there is the null polynomial, too.

Now, we specialize the previous general concepts to the case of monomial ideals, the only class of ideals we use.

Definition A20 A monomial ideal is an ideal that can be generated by monomials.

From Definition A1, it follows that the generators of a monomial ideal can be chosen with numerical coefficient equal to 1. In fact, if $c\underline{x}^c$, $c \neq 0$, is a generator of I, then $c^{-1} c\underline{x}^a \in I$ by property (2), and so we get the claim.

Proposition A21 Sum, intersection and radical of monomial ideals are monomial ideals.

Sketch of the proof. We give a set of monomial generators for each ideal.

The sum of ideals is generated by the union of the generators of both the ideals, and so the sum of monomial ideals is generated by monomials.

To prove the claim about the intersection, let I and J be monomial ideals generated by $\underline{x}^{a_1}, \ldots, \underline{x}^{a_s}$ and $\underline{x}^{b_1}, \ldots, \underline{x}^{b_t}$, respectively. Then, $I \cap J$ is generated by $\underline{x}^{c_{ij}}$ with $i = 1, \ldots, s, j = 1, \ldots, t$, where $\underline{x}^{c_{ij}} = lcm\left(\underline{x}^{a_i}, \underline{x}^{b_i}\right)$ i.e., the least common multiple of $\underline{x}^{a_i}$ and $\underline{x}^{b_i}$, and so it is a monomial ideal, too.

Finally, if I is generated by $\underline{x}^{a_1}, \ldots, \underline{x}^{a_s}$ then its radical $\sqrt{I}$ is generated by $\underline{x}^{d_1}, \ldots, \underline{x}^{d_s}$ where $\underline{x}^{d_1}$ is obtained from $\underline{x}^{a_1}$ by setting the non-zero exponents to 1. Q.E.D.

From its construction, it is clear the reason why radical monomial ideals are called square-free.

Proposition A22 A monomial ideal I is prime if, and only if, it is generated by degree 1 monomials, that is to say, it is generated by variables.

Proof. A standard argument allows us to reduce the proof to products of terms.

If $x_1^{a_1} \ldots x_n^{a_n} \in I$ with $a_1 \backslash not = 0$ and I is prime, then either $x_1 \in I$ or $x_1^{a_n - 1} \ldots x_n^{a_n} \in I$ by Definition A5. If the first item fails, by iterating the argument, we have that a different variable must be in I. Hence, if I is monomial and prime, then it is generated by variables.

Conversely, it is evident that if I is generated by variables, then it is prime. Q.E.D.

Theorem A23 The codimension of a monomial prime ideal I is equal to the minimal number of variables that generate I.

Proof. Assume that I is generated by $x_1, \ldots, x_t$. Then, a chain of maximal length of prime ideals is $0 \subset <x_1> \subset <x_1 x_2> \subset \cdots \subset <x_1, \ldots, x\{t-1\}> \subset I$ and so the claim follows. Q.E.D.

Theorem A24 The degree of a monomial prime ideal I is equal to 1.

Proof. Assume that I is generated by $x_1, \ldots, x_t$. Then, the general linear ideal L to be added to I to get an ideal of codimension n is generated by $x_{t+1} - c_{t+1}, \ldots, x_n - c_n$ with $c_{t+1}, \ldots, c_n \in K$. Then, $I + L$ is generated by $x_1, \ldots, x_t, x_{t+1} - c_{t+1}, \ldots, x_n - c_n$ and so $in(I+L)$ is generated by all the variables, i.e., 1 is the only term not in $in(I+L)$ and the claim follows. Q.E.D.

A consequence of Theorems A23 and A24 is that we can characterize the codimension and the degree, also for non-prime monomial ideals.

Proposition A25 The codimension of a radical monomial ideal I is the minimum of the codimensions of the prime ideals in a primary decomposition of I, while the degree of a radical monomial ideal I is equal to the number of prime ideals of minimal codimension associated to I.

Proof. The first claim is true by Definition A10. Assume that the primary decomposition of I is given by $I = J_1 \cap J_2 \cap \ldots \cap J_t$ with $J_1, \ldots, J_s$ of codimension equal to $codim(I)$, and $J_{s+1}, \ldots, J_t$ of codimension strictly larger. Let L be a general linear ideal of codimension equal to $n - codim(I)$. Then, $L + J_i, i = s+1, \ldots, t$, is equal to the whole ring A, as it follows from the theory of linear systems. In fact, L is generated by $n - codim(I)$ independent linear equations, and J_i is generated by more than $codim(I)$ variables. Hence, the resulting linear system has no solutions, that is to say, one of its

equations is $1 = 0$, or, in a different language, $1 \in L + J_i$ for each $i = s + 1, \ldots, t$. By using the same argument with the ideals $J_1, \ldots, J_s$, it follows that each ideal $L + J_i$, $i = 1, \ldots, s$, has degree 1, and the claim follows. Q.E.D.

Finally, we describe how to compute the first and second Betti number for a monomial ideal.

Proposition A26 The non-minimal generators of a monomial ideal are the ones that can be divided by a different generator of the same ideal.

Hence, to compute the first Betti number of a monomial ideal, one has to check if a given generator can be divided by one of the others. If it cannot be divided, then it is part of a minimal set of generators, otherwise it is unnecessary, and can be eliminated.

To compute the syzygies of a monomial ideal generated by $\underline{x}^{a_i}, \ldots, \underline{x}^{a_t}$, the first remark is that the generators of the syzygy module have the form $\left(0, \ldots, 0, \underline{x}^b, 0, \ldots, 0, -\underline{x}^c, 0, \ldots, 0\right)$, where the non—zero items appear at the i—th and j—th position and $\underline{x}^b = lcm\ (\underline{x}^{a_i}, \underline{x}^{a_j})\ /\underline{x}^{a_i}$ and $\underline{x}^c = lcm\ (\underline{x}^{a_i}, \underline{x}^{a_j})\ /\underline{x}^{a_i}$. The second Betti number is the number of minimal generators for the first syzygy module.

Appendix 2: Algebraic Parameters

Let $G = (V, E)$ be an undirected graph, with V set of its vertices, and E set of its edges. Usually, if G has v vertices, then $V = \{1, \ldots, v\}$ but for our purpose, we choose $V = \{x_1, \ldots, x_v\}$ and we think of $x_1, \ldots, x_v$ as independent variables. Each edge is a set $\{i, j\}$ with $i, j \in V$ but we represent it with the monomial $x_i x_j$. Hence, E is represented as a monomial ideal $I \subset A = Q[x_1, \ldots, x_v]$ with Q field of the rational numbers. If we want to stress the graph G, then we write I_G for the ideal I.

Definition B1 The ideal $I = < x_i x_j \mid (i, j) \in E >$ is the edge ideal of the undirected graph $G = (V, E)$.

Conversely, every monomial ideal generated by degree 2 monomials is the edge ideal of a graph G. We assume the graph G to have no loop and so I is a square—free monomial ideal generated by a set of degree 2 monomials. In the remaining part of the section, we give a graph—theoretical interpretation of the notions described in the Appendix 1: A short review on polynomial ideals.

We recall that two graphs are isomorphic if they have the same vertices and the same edges, up to relabeling the vertices of one of the two graphs. Of course, two isomorphic graphs have the same edge ideal, up to reorder the variables.

At first, we discuss the codimension of an edge ideal. We say that a vertex meets an edge if the vertex is one of the two vertices of the edge. Furthermore, a subset of vertices meets a set of edges if each edge of the set meets at least a vertex in the set.

Proposition B2 Let I be an edge ideal, and let $V' \subset V$ be a subset of the vertex set. If V' meets E, then $codim\ (I) \leq |V'|$. Conversely, there exists a subset V' of V containing $codim\ (I)$ vertices that meets E.

Proof. Algebraically, the vertex x_i meets the edge $x_h x_k$ if x_i divides $x_h x_k$, i.e., $x_i = x_h$ or $x_i = x_k$. Let $V' = \{x_{i_1}, \ldots, x_{i_t}\}$. So, $\dot{V}$ meets E if and only if $I \subset J = \langle x_{i_1}, \ldots, x_{i_t} \rangle$. Hence, $codim(I) \leq codim(J) = t = |V'|$.

Conversely, if $codim(I) = t$, there exists a prime ideal J containing I generated by t variables, as it follows from Definition A10 or Proposition A25 in Appendix 1. Hence, the claim follows from the previous argument. Q.E.D.

The general strategy we try to follow is to introduce a numerical invariant associated to a graph G and then to propose an upper bound for it, so we can compute the normalized invariant, better suited to compare graphs with different number of vertices and edges. As in the application we are concerned with planar graphs, we restrict ourselves to bounding the numerical invariants in the class of planar graphs. In such a class, for example, the number e of edges verifies $v - 1 \leq e \leq 3v - 6$ where v is the number of vertices.

Definition B3 Assume the number of vertices v to be even, i.e., $v = 2m$ for some $m \geq 3$. For $v - 1 \leq e \leq 3v - 6$ we define the function

$$
codim\_max(v, e) = \begin{cases} m & if \quad e = v - 1, \quad v \\ m + k + 1 & if \quad \begin{array}{l} v + 1 \leq e \leq 3v - 11 \\ e - v - 1 = 4k + r \text{ with } 0 \leq r \leq 3 \end{array} \\ v - 2 & if \quad e > 3v - 11 \end{cases}
$$

Assume the number of vertices v to be odd, i.e., $v = 2m + 1$ for some $m \geq 2$. For $v - 1 \leq e \leq 3v - 6$ we define the function

$$
codim\_max(v, e) = \begin{cases} m & if \quad e = v - 1 \\ m + 1 & if \quad e = v, v + 1, v + 2 \\ m + k + 2 & if \quad \begin{array}{l} v + 3 \leq e \leq 3v - 11 \\ e - v - 3 = 4k + r \text{ with } 0 \leq r \leq 3 \end{array} \\ v - 2 & if \quad e > 3v - 11 \end{cases}
$$

From the study of planar graphs with 5,6,7 vertices, the functions *codim_max* are an upper bound for the codimension of planar graphs. Then, given a planar graph G with v vertices, e edges and codimension c, we define its normalized codimension as

$$
ncodim(G) = \frac{codim(G)}{codim\_max(v, e)}.
$$

The graph-theoretical meaning of the degree (see Definitions A13 and A15 in Appendix 1) is explained in next Proposition.

Proposition B4 Let I be the edge ideal of a graph $G = (V, E)$, and assume that *codim* $(I) = t$. Then, deg (I) is equal to the number of subsets V' of V with t elements that meet E.

Proof. It is enough to apply Proposition A25 and Proposition B2. Q.E.D.

We were not able to compute an upper bound for the degree of an edge ideal in terms of the number of vertices and of edges, and so the degree hasn't a normalized value.

Also the first Betti number has a clear meaning in terms of graphs.

Proposition B5 The first Betti number of I is equal to the number of edges of G.

Proof. All the generators of I have degree 2. One of them is non-minimal if and only if it can be divided by a different generator. The degree argument proves that a generator is non-minimal if, and only if, it appears twice among the generators of I, i.e., all the edges are minimal generators. Q.E.D.

As explained when we discussed the meaning of γ, it is the normalized value of $b_{11}(I_G)$. With a change of notation, we call $b_{11}(G)$ the Betti number $b_{11}(I_G)$. The same notation will be adopted for the other Betti numbers.

Now, we discuss the second Betti numbers.

Let $x_i x_j$ and $x_h x_k$ be the first two generators of the edge ideal I. If the two edges have no common vertex, then $(x_h x_k, -x_i x_j, 0, \ldots, 0)$ is a syzygy of I, of degree 2, while, if they have a common vertex, say $x_i = x_h$, then $(x_k, -x_i, 0, \ldots, 0)$ is a syzygy of I, of degree 1. Analogously, each couple of edges gives a syzygy. It is possible to prove that every syzygy of I is a combination of the ones we just described, and so, if $b_2(G)$ is the second Betti number of I, we can write $b_2(G) = b_{21}(G) + b_{22}(G)$ where $b_{21}(G)$ is the number of minimal generators of the second syzygy module having degree 1, while $b_{22}(G)$ is the number of minimal generators of the second syzygy module having degree 2. Then, we have

Proposition B6 Let d_i be the degree of the vertex x_i for each $i = 1, \ldots, n$, and let $t(G)$ be the number of triangles in G. Then $b_{21}(G) = \sum_{\{i=1\}}^{n} \frac{d_i(d_i - 1)}{2} - t(G)$.

A triangle is a complete sub-graph with 3 vertices, while the degree of a vertex is the number of edges that meet the vertex.

Proof. Assume that the degree of x_1 is d_1. Then, there exist d_1 vertices, say $x_2, \ldots, x_{d_1+1}$, such that $x_1 x_2, \ldots, x_1 x_{d_1+1}$ are edges of G. Hence, for each couple of edges among them we get one degree 1 syzygy. So, each vertex contributes to $b_{21}(I)$ at most with $d \frac{(d-1)}{2}$ syzygies, where d is the degree of the vertex.

Assume that x_1, x_2, x_3 are the 3 vertices of a triangle. Then, $x_1 x_2, x_1 x_3, x_2 x_3$ are edges of I and so $s_1 = (x_3, -x_2, 0, 0, \ldots, 0)$, $s_2 = (x_3, 0, -x_1, 0, \ldots, 0)$, $s_3 = (0, x_2, -x_1, 0, \ldots, 0)$ are degree 1 syzygies of I. It is evident that the last syzygy is the difference of the second and the first, that is to say, $s_1 - s_2 + s_3 = 0$ and so for a triangle only two syzygies are part of a minimal set of generators.

Now, we prove that only the triangles decrease the number of minimal degree 1 generators.

Let $s_1, \ldots, s_r$ be degree 1 syzygies such that $a_1 s_1 + \ldots + a_r s_r = 0$ for some suitable non-zero $a_1, \ldots, a_r$ rational numbers. Up to multiplying for the common denominator, we can assume that they are integers. Assume that $s_1 = (x_i, -x_j, 0, \ldots, 0)$ This is always possible, up to a permutation of the generators of I. Then, the first two generators of I have a common vertex, say x_k, and the first two generators of I are $x_j x_k, x_i x_k$. From the equality $a_1 s_1 + \ldots + a_r s_r = 0$, we get that x_i appears as the first entry of at least 2 syzygies, and so, up to renumbering them, and eventually permuting the minimal generators of I, we can assume that $s_2 = (x_i, 0, -x_h, 0, \ldots, 0)$. Hence, the first and the third generators of I have a common vertex. If it were x_k again, then the third generator of I is $x_i x_k$ and this is not possible, because there are no repetitions. So, the common vertex is x_j and $h = k$. Hence, the third generator of I is $x_i x_j$ and the three generators form a triangle. We can subtract $a_1(s_1 - s_2 + s')$ from the previous equality, with

$s' = (0, x_j, -x_k, 0, \ldots, 0)$. By iterating the argument, we get a new equality in which x_i does not appear. Going on, we get that every such equality is a combination of similar equalities associated to triangles. Q.E.D.

Definition B7 Let $S \subset E \times E$ be the set containing the couples (e, e') such that $e \neq e'$, e, e' have no common vertex, and there is no edge e'' such that both e, e'' and e', e'' have a common vertex. In Subsection 2.4, the edges e, e' of a couple $(e, e') \in E$ are called far away.

Proposition B8 $b_{22}(G)$ is equal to the cardinality of S.

Proof. A degree 2 syzygy is minimal if it cannot be obtained as combination of degree 1 syzygies. In fact, if the edges $x_i x_j$ and $x_h x_k$ have no common vertex, the syzygy has $x_h x_k$ in correspondence of the edge $x_i x_j$ and conversely. Hence, each edge appears only once in a specific position of the syzygies. Assume that the syzygy is $s = (x_h x_k, -x_i x_j, 0, \ldots, 0)$ and that it is combination of degree 1 syzygies. There exist two syzygies $s_1 = (x_h, \ldots)$, $s_2 = (0, x_i, \ldots)$ so that $s - x_k s_1 + x_j s_2$ is a degree 2 syzygy with the first two entries equal to 0. The first edge is $x_i x_j$ and so either $x_j x_h$ or $x_i x_h$ is an edge of E, as it follows from the syzygy s_1. Analogously, being s_2 a syzygy, either $x_i x_k$ or $x_i x_h$ is an edge of E. If $x_i x_h$ is an edge, then the 3 edges $x_i x_j$, $x_i x_h$, $x_h x_k$ have the property of the first two ones and the last two ones have a common vertex. If $x_i x_h$ is not an edge, then $x_i x_j$, $x_j x_h$, $x_h x_k$, $x_i x_k$ are edges, and it is easy to check that s is a combination of degree 1 syzygies obtained from $x_i x_j$, $x_j x_h$, $x_h x_k$ with the property that the first two ones and the last two ones have a common vertex. Q.E.D.

As said in Subsection 2.4, Betti numbers don't represent uniquely graphs, in the sense that it is possible to construct non-isomorphic graphs with the same Betti numbers.

To compare the Betti numbers of graphs with different numbers of vertices and of edges, we need a relative or normalized version of the Betti numbers of an edge ideal. To this end, we construct particular reference graphs for each number v of vertices and e of edges, we compute their Betti number $b_{21}(v, e)$, and we get the relative version as $\frac{b_{21}(G)}{b_{21}(v,e)}$. We didn't succeed in finding the reference graphs for the Betti number b_{22}, that remains an open task. The relative Betti numbers can be considered as a measure of the difference of shape between the given graph and the one chosen as a reference.

Definition B9 The reference graph with v vertices and e edges for the Betti number b_{21} is $H(v, e)$ and it is associated to the ideal

$$I_{H(v,e)} = \begin{cases} < \quad x_1 x_2, \quad \ldots \quad , \quad x_1 x_v > & \text{if } e = v - 1 ; \\ < \quad x_1 x_2, \quad \ldots \quad , \quad x_1 x_v, \quad x_2 x_3, \quad \ldots \quad , \quad x_2 x_t > & \text{if } \begin{matrix} v \le e \le 2v-3 \\ t = e-v+3 \end{matrix} ; \\ < x_1 x_2, \ldots, x_1 x_v, x_2 x_3, \ldots, x_2 x_v, x_3 x_4, x_4 x_5, \ldots, x_{t-1} x_t > & \text{if } \begin{matrix} 2v-2 \le e \le 3v-6 \\ t = e-2v+6 \end{matrix} . \end{cases}$$

In Fig. 2, we drew the graphs $H(6, 5)$, $H(6, 8)$, $H(6, 11)$. In Definition 6, we defined the graph $H(v, e)$ in terms of its adjacency matrix.

Proposition B10 The Betti number $b_{21}(v, e)$ of $H\ (v, e)$ is equal to

$$b_{21}(v, e) = \begin{cases} \binom{v-1}{2} + \binom{e-v+2}{2} & \text{if } v-1 \le e \le 2v-3 \\ v^2 - 9v + 10 + 3e & \text{if } 2v-2 \le e \le 3v-6 \end{cases}.$$

Proof. It is enough to apply Proposition B5 to the graph $H\ (v, e)$. Q.E.D.

The graph $H\ (v, e)$ has the largest b_{21} for $v \le 7$ and every e among the planar graphs with the same number of vertices and edges. Hence, $b_{21}(v,\ e)$ can be considered as a normalization factor.

Appendix 3: Computing the Algebraic Parameters

Algebraic computations can be performed by using dedicated software. Among the available many ones, we mention CoCoA, developed at Genova University, and Singular, developed at Kaiserslautern University (Decker et al. 2015), a free software available on the web.

As example, we report the Singular session in which, starting from the adjacency matrix of a graph, we compute the various algebraic parameters associated to it. After the session, we comment it.

To fix ideas, we consider the graph $M\ (6, 7)$ in the middle of Fig. 4.

Command Line Number	Command line
line 1	> ring R=0,(x(1.0.6)),dp;
line 2	> int nvar=6;
line 3	> matrix ag[nvar][nvar]=0,1,1,1,1,1, 1,0,1,1,0,0, 1,1,0,0,0,0, 1,1,0,0,0,0, 1,0,0,0,0,0, 1,0,0,0,0,0;
line 4	> int i; int j;
line 5	> ideal ig;
line 6	> for(i=1;i<nvar;i=i+1){for(j=i+1;j<1+nvar;j=j+1) {if(ag[i,j]==1){ig=ig, x(i)*x(j);}}}; ig=std(ig);
line 7	> int codim=nvar-dim(ig);
line 8	> codim;
	2
line 9	> degree(ig);
	// dimension (proj.) = 3
	// degree (proj.) = 1
line 10	> matrix mg=mres(ig,2);
line 11	> matrix bg=betti(mres(ig,2));
line 12	> poly b21=bg[2,3]; b21;
	13
line 13	> poly b22=ncols(mg)-bg[2,3]; b22;
	0
line 14	> exit;

The first line command defines the polynomial ring in which the computation is performed. In every case, the variables are as many as the vertices, the ground field is the field of rational numbers (coded as the 0 at the first place), and the term-ordering is the degree reverse lexicographic (coded as dp).

In the second line, we input the number of variables, or of vertices.

In the third line, we input the adjacency matrix of M $(6, 7)$. We call ag this matrix, whose type is nvar × nvar.

The remaining code lines have to be typed verbatim independently from the graph, and they produce the following outputs: line 8 prints the codimension computed in the previous line, the next line computes and prints the degree (second output line), while the Betti numbers b_{21} and b_{22} are computed and printed at 12th and 13th lines.

So, M $(6, 7)$ has codimension 2, degree 1, and second Betti numbers 13 and 0, respectively.

References

Barabasi A-L, Albert R (1999) Emergence of scaling in random networks. Science 286(5439):509–512. doi: 10.1126/science.286.5439.509

Berge C (1962) The theory of graphs and its applications. Methuen, London

Black WR (2003) Transportation: a geographical analysis. The Guilford Pres, New York

Blumenfeld-Lieberthal E (2009) The topology of transportation networks: a comparison between different economies. Netw Spat Econ 9:427–458. doi:10.1007/s11067-008-9067-6

Chung FRK (1997) Spectral graph theory. CBMS regional conference series in mathematics, No. 92, American Mathematical Society

Decker W, Greuel G-M, Pfister G, Schönemann H (2015) Singular 4-0-2 A computer algebra system for polynomial computations. http://www.singular.uni-kl.de

Derrible S, Kennedy C (2010) Characterizing metro networks: state, form, and structure. Transportation 37:275–297. doi:10.1007/s11116-009-9227-7

Derrible S, Kennedy CA (2011) Applications of graph theory and network science to transit network design. Transp Rev 31:495–519. doi:10.1080/01441647.2010.543709

Derrible S, Kennedy CA (2012) A network analysis of subway systems in the world using updated graph theory. Transp Res Rec 2112:17–25

Ducruet C, Lugo I (2013) Structure and dynamics of transportation networks: Models, concepts, and applications. In: Rodrigue JP, Notteboom TE, Shaw J (eds) The SAGE Handbook of transport studies. SAGE Publications, pp. 347–364

Dwork C (2005) Lecture2: eigenvalues and expanders. CS369E: expanders, scribe: Geir Helleloid. http://www.tcs.tifr.res.in/~prahladh/teaching/05spring/. Accessed 28 Jul 2014

Erath A, Löchl M, Axhausen KW (2007) Graph-theoretical analysis of the Swiss road and railway networks over time. Conference paper STRC2007

Garrison WL, Marble DF (1964) Factor-analytic study of the connectivity of a transportation network. Reg Sci 12(1):231–238

Garrison WL, Marble DF (1962) The structure of transportation networks. Transportation Center Northwestern University, Evanston

Gastner MT, Newman MEJ (2006) The spatial structure of networks. Eur Phys J B 49:247–252. doi:10.1140/epjb/e2006-00046-8

Gattuso D, Miriello E (2005) Compared analysis of metro networks supported by graph theory. Netw Spat Econ 5(4):395–414

Godsil C, Royle GF (2001) Algebraic graph theory. GTM 207. Springer, New York

Goldreich O (2011) Basic facts about expander graphs. Studies in complexity and cryptography, LNCS6650. Springer, Berlin, pp 451–464

Kansky KJ (1963) Structure of transportation networks: relationships between network geometry and regional characteristics. University of Chicago Press, Chicago

Kurant M, Thiran P (2006) Extraction and analysis of traffic and topologies of transportation networks. Phys Rev E 74(036114):1–10

Levinson D (2009) Introduction to the special issue on the evolution of transportation network infrastructure. Netw Spat Econ 9:289–290. doi:10.1007/s11067-009-9105-z

Lang S (2002) Algebra. Springer Verlag, New York

Lovasz L (2007) Eigenvalues of graphs. November 2007, http://www.cs.elte.hu/~lovasz/eigenvals-x.pdf. Accessed 28 Jul 2014

Maruri-Aguilar H, Saenz-de-Cabezón E, Wynn HP (2012) Betti numbers of polynomial hierarchical models for experimental designs. Ann Math Artif Intell 64(4):411–426

Nadakuditi RR, Newman MEJ (2012) Graph spectra and the detectability of community structure in networks. Phys Rev Lett PRL 108:188701,1–5

Newman MEJ (2003) The structure and function of complex networks. Siam Rev 45(2):167–256

Prihar Z (1956) Topological properties of telecommunications. Proc IRE 44(7):927–933. doi:10.1109/JRPROC.1956.275149

Sáenz-de-Cabezón E, Wynn HP (2009) Betti numbers and minimal free resolutions for multi-state system reliability bounds. J Symb Comput 44(9):1311–1325

Xie F, Levinson D (2007) Measuring the structure of road networks. Geogr Anal 39:336–356. doi:10.1111/j.1538-4632.2007.00707

Xie F, Levinson D (2008) The use of road infrastructure data for urban transportation planning: issues and opportunities. In: Amekudzi A, McNeil S (eds) Infrastructure reporting and asset management. American Society of Civil Engineers, Reston, pp 93–98

Xie F, Levinson D (2009a) Modeling the growth of transportation networks: a comprehensive review. Netw Spat Econ 9:291–307. doi:10.1007/s11067-007-9037-4

Xie F, Levinson D (2009b) Jurisdictional control and network growth. Netw Spat Econ 9:459–483. doi:10.1007/s11067-007-9036-5

Zumstein P (2005) Comparison of spectral methods through the adjacency matrix and the laplacian of a graph. Diploma Thesis. ETH Zurich