ORIGINAL SCIENTIFIC PAPER

CROATICA CHEMICA ACTA

Croat. Chem. Acta **2017**, *90*(*1*), 75–85
Published online: July 3, 2017
DOI: 10.5562/cca3099

HKD 1926

A Dual of the Cycle Theorem and its Application to Molecular Complexity†

Edward C. Kirby,[1] Roger B. Mallion,[2,*] Paul Pollak,[3] Paweł J. Skrzyński[4]

[1] Resource Use Institute, Fishersview Court, Pitlochry PH16 5AN, Scotland, United Kingdom
[2] School of Physical Sciences, University of Kent, Canterbury CT2 7NH, England, United Kingdom
[3] The King's School, Canterbury CT1 2ES, England, United Kingdom
[4] The Canterbury High School, Canterbury CT2 8QA, England, United Kingdom
* Corresponding author's e-mail address: R.B.Mallion@kent.ac.uk

RECEIVED: February 10, 2017 * REVISED: April 26, 2017 * ACCEPTED: May 8, 2017

Abstract: The duality alluded to in the title is that between the faces and vertices of a graph embedded on a surface. Its recognition in the context of the five Platonic solids is classic. Algebraically, it is present in the equation for Euler's Polyhedron Theorem and in the various extensions thereof. The Cycle Theorem (CT) establishes a formula for the number of spanning trees contained in a graph embedded on a surface. It is based on the mutual incidences of its *cycles* (circuits which also carry a sense of direction), *i.e.*, of sub-graphs of the C_n type, endorsed with a sense. These appear (though not exclusively) as the boundaries of faces, so that, so to speak, the Cycle Theorem establishes a result which is essentially about vertices *via* relations between faces. Among several possible duals of the Cycle Theorem there might thus be one that establishes a relation which is essentially about faces *via* relations between vertices. In order to formulate one, we define, for an embedded graph, a feature concerning faces that is dual to a spanning tree. We call it a *ladder*. A formula is presented for the number of ladders contained in a graph which, in some cases, introduces the concept of 'artificial vertices'. It is based on the mutual incidences of its vertices. Its form is clearly analogous, or 'dual', to the Cycle Theorem formula for spanning trees, previously proposed (in this journal — 2004) by three of the present authors, together with Klein and Sachs. A new index is proposed, which involves ladders. We call it the *Patency Index* of a graph; its numerical value may be related to molecular complexity. It is effectively the dual of, and is entirely analogous to, the Spanning-Tree Density Index which was earlier (2003) proposed, defined and applied to molecular graphs by one of the present authors and Trinajstić.

Keywords: Chemical graph theory, duality, molecular complexity, spanning trees, ladders, Cycle Theorem.

HISTORICAL BACKGROUND

A spanning tree of a one-component graph *G* is a connected acyclic sub-graph that contains all the vertices of *G*.[1] Over the years, considerable attention has been devoted to the relevance of the spanning trees of graphs[1–6] in various chemical contexts such as ring currents in conjugated systems (for example Refs. [7–10]) and the complexities of molecules (*e.g.*, Refs. [11–26]), as well as in physical situations — *e.g.*, electrical networks, such as in Ref. [27] and the papers that are cited therein. The work presented in this paper has evolved chronologically from the earlier formalisms of Gutman, Mallion and Essam[28] (hereafter 'GME') and of Kirby *et al.*[29,30] for spanning-tree enumeration. Both are alternatives to the traditional Matrix Tree Theorem ('MTT').[3–6,8,9,27] The GME approach is applicable only to graphs embedded in a planar manner — that is, without the use of generic circuits (defined in Ref. [29]) — and draws on the idea of an inner dual.[28,31–36] It has the drawback that it is applicable only to planar graphs[28,29,32,35] because GME relies on the fact that a planar graph and its dual have the same number of spanning trees.[3,37–39] In practice, this was no great privation to chemists because most structures of chemical interest can be represented by molecular graphs so embedded. Nevertheless, Kirby *et al.* later introduced[29] (and developed)[30] what they called the 'Cycle Theorem'[29] (hereafter 'CT') which, without adopting the device of an inner dual,[28,31–36,40] employs cycles-to-cycles incidence[40] in its formulation (such as in Refs. [41–43] and [27]) and has

This work is licensed under a Creative Commons Attribution 4.0 International License.

† We should like to dedicate this paper to the memory of the late Professor Horst Sachs (1927–2016) with whom we were privileged to work on the Cycle Theorem (2004) itself.

the advantage that it is valid for a graph of any genus embedded on a suitable surface. When they are applied to graphs representing molecules, both GME and CT usually result in smaller determinants than are encountered in the traditional Matrix Tree Theorem (MTT). We note in passing that the present authors have recently remarked,[27] in this journal, that, in the context of electrical networks,[41–45] (*i*) CT arises from Kirchhoff's[44] circuit law[27] whilst (*ii*) MTT arises from Kirchhoff's[44] law of current conservation at junctions.[27] Furthermore, the non-singular matrix **M** referred to in CT[29,30] is of dimension ($\mu \times \mu$), where μ is the circuit rank/cyclomatic number[1,44,45] of the graph in question, whilst the singular matrix **K** that features in MTT (*e.g.*, Refs. [3–6,8,9,27,36]) is of size ($v \times v$), where v is the number of vertices in that graph.

(For the avoidance of doubt we note in passing that when we make mention — above and throughout — of a 'planar graph' then, (a) this clearly denotes a graph that can be embedded, without edge-crossings, on a plane or sphere, but also, (b) if such a graph is embedded otherwise, the expression 'planar graph' implies that no generics are used in the course of the embedding (in the sense explained in Ref. [29]).)

It is the purpose of the present paper to propose a *dual* [3,27,28,31–40] of the Cycle Theorem (DCT) and, in the process, to introduce the concept of what we call a *ladder* — the dual of a spanning tree — and then to use it to define what we are here calling the 'Patency Index' for an embedded graph, a quantity which may have a part to play in assessing molecular complexity.[11–26] Effectively, this *Patency Index* expresses a probability and, furthermore, we emphasise that it is analogous to the previously defined[25] *Spanning-Tree Density*.

INTRODUCTORY REMARKS

1. The graphs considered are embedded on a single closed surface, the *host surface*. Acceptable surfaces are the sphere (which also does service for the plane), the torus, and more complicated surfaces of the torus type. These names are, of course, to be understood in their topological sense. We exclude the Klein Bottle[29] on account of its self-penetrating structure (in three-dimensional space), for simplicity rather than any deeper unsuitability.
2. The terms *vertex* and *edge* have their usual meanings (*e.g.*, Refs. [1–6]).
3. The term *face* refers to a single portion of the host surface that is completely *bounded* by one or more circuits.

   By 'circuits', we mean sub-graphs of the C_n type, including C_1, often called a self-loop.

   The host surface itself is a — or *the* — face if no other faces are present. At times the boundaries of a face may themselves be referred to as the face.
4. The term *internal edge* refers to an edge that is not part of the boundary of any face.
5. Host surfaces are represented on the page (or screen) in two-dimensional diagrams that preserve and exhibit all features essential to the argument. Simple ones for a torus (and also Klein Bottle and Möbius Band[29,46,47]) are well known. They depend on the presence of 'identified' (or 'counter-identified') lines,[29,46,47] arising from cuts in the original surface that enable it to be 'flattened-out'. In this paper we have extended this idea to provide further diagrams for the torus and also for more complicated surfaces. Each pair of such lines has its own colour, and arrows to show whether the identification is 'direct' or 'counter'.[29,46,47]
6. On host surfaces whose topological genus (g) is greater than 0, graphs may be embedded with circuits that, considered in the absence of the rest of the graph, do not divide the host surface into two separate parts. In our plane representations, they must therefore arise as paths between the two appearances of an identified point. Such circuits are called *generic circuits* or just *generics*. (Please see Ref. [29] for further elaboration of these terms.) Two generics are of different *kinds* (see Ref. [29]) if, when both are present simultaneously, the host surface is still not divided. The numbers g and k (the number of different kinds) satisfy $2g \geq k$.
7. We employ, at times, an equation that generalises the one expressing Euler's Polyhedron Theorem.[47] It is

$$f + v - e = p + s - k \,. \tag{1}$$

   The parameters f, v, e denote, respectively, the numbers of faces, vertices, and edges of a graph embedded on an unrestricted host surface. This means a surface which is not restricted to any particular kind of surface. In particular, 'unrestricted' means that Equation (1) applies to all surfaces, and not only to those specified in §1 of 'Introductory Remarks'. The parameter p denotes the number of *pieces* (or *components*) of the embedded graph and the parameter s the number of spaces enclosed by the host surface. The parameter k, which might be called the 'genus of the embedding', has already been defined in §6, above.
8. Throughout this paper, the value of s is 1: but, for algebraical clarity, the value of this parameter will usually be denoted by s.
9. When necessary, vertices are labelled using upper-case Roman letters, faces using lower-case Greek letters, and edges using numerals. To avoid clutter, over-labelling is shunned.

THE DEFINITION OF *LADDER*

An edge which belongs to the boundaries of two faces is said to *link* these faces or to be a *link* between them. A set of edges such that, by means of a succession of links that are members of the set, each face is linked to every other face, constitutes a *linking set*. A *ladder* (*cf.* p. 64 of Ref. [3]) is a minimal linking set, in the sense that the removal of even one edge from it would destroy its capacity to link fully. Evidently, this is dual to a possible definition of the set of edges belonging to a spanning tree.

If the embedded graph is such that the host surface is the only face, then there are no edges that might have formed a linking set. To avoid tedious exceptions, a unique *null-ladder* may be said to be present in these circumstances. (The empty set of edges of a spanning tree without edges, belonging to a graph consisting of a single vertex, might be its dual.)

SOME PROPERTIES OF LADDERS

The Number of Edges in a Ladder.

Denote by f_0, v_0, *etc.*, the values of these parameters in Equation (1) for an embedded graph G_0 and by f_1, v_1, *etc.*, those for the graph G_1 which remains when all the edges of a ladder, say l in number, have been deleted. We then have, in general — but remembering that our single host surface corresponds to $s = 1$ in Equation (1) —

$$f_1 = s_0,\ v_1 = v_0,\ e_1 = e_0 - l,\ p_1 = p_0$$

(since no isthmus edge can belong to a ladder), $s_1 = s_0$, $k_1 = k_0$ (since no edge essential for the existence of a *kind* of generic — see Ref. [29] for definitions — can belong to a ladder). Substituting these values in Equation (1) for the graph G_1, we obtain

$$s_0 + v_0 - (e_0 - l) = p_0 + s_0 - k_0\ ,$$

so that

$$l = e_0 - v_0 + p_0 - k_0\ .$$

By use of Equation (1) again, the right-hand side can be seen to equal

$$f_0 - s_0$$

so that

$$l = f - 1\ . \qquad (2)$$

This can be compared to $(v - 1)$ for the number of edges in a spanning tree. It also shows that all ladders in a given embedded graph contain the same number of edges. The case $l = 0$ is consistent with the concept of the null-ladder.

The Number of Ladders in an Embedded Graph.

With the embedded graph G may be associated its well-known complete ('geometric') dual,[3,27,28,31–40] say G^+. The faces of G are related 1:1 to the vertices of G^+ and the edges of G linking its faces are related 1:1 to the edges of G^+ linking its vertices. (Internal edges of G which would give rise to self-loops, *i.e.*, graphs of the type C_1, may be ignored for our purposes.) Denoting the number of ladders in G by $l(G)$ and the number of spanning trees in G^+ by $t(G^+)$, we obtain

$$l(G) = t(G^+)\ . \qquad (3)$$

If we define the *degree* of a face to be the number of edges in its boundary, then the degrees of faces and vertices in G and G^+ are related 1:1. Let $\mathbf{F}$ be the full $(f \times f)$ face–boundary overlap-matrix derived from the adjacencies of all the faces of G. Then $\det \mathbf{F} = 0$ and all its $((f-1)\times(f-1))$ sub-determinants have the same absolute value, $l(G)$. An example of this follows. (This corresponds to the well-known result for the number of spanning trees based on the likewise equi-cofactorial $(v \times v)$ Laplacian matrix[40,48–53] of a graph, also known as the Kirchhoff matrix — *e.g.* Refs [5,6,8,9,27,54,55], amongst many others — and as the Admittance matrix.[4])

Example 1. The complete graph K_5 embedded on a double ('figure-of-eight') torus.

In Figure 1 the complete graph K_5 is shown embedded on a double ('figure-of-eight') torus. Note that the circuit 1, 5, 4, (1) is not a generic:[29] in the absence of the other edges, it divides the surface into two separate parts.

Parameter census: $f = 5$, $v = 5$, $e = 10$, $p = 1$, $s = 1$ and $k = 2$.

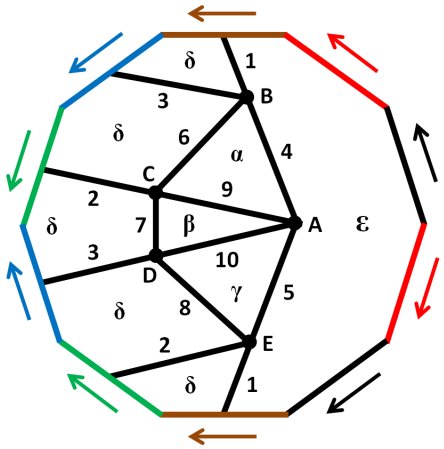

Figure 1. The complete graph K_5 embedded on a double ('figure-of-eight') torus.

For this embedding, by means of the following table,

	ε	α	β	γ	δ
ε	3	–1	0	–1	–1
α	–1	3	–1	0	–1
β	0	–1	3	–1	–1
γ	–1	0	–1	3	–1
δ	–1	–1	–1	–1	4

we construct the full face–face ($f \times f$) matrix, which is denoted $\mathbf{F_1}$:

$$\mathbf{F_1} = \begin{pmatrix} 3 & -1 & 0 & -1 & -1 \\ -1 & 3 & -1 & 0 & -1 \\ 0 & -1 & 3 & -1 & -1 \\ -1 & 0 & -1 & 3 & -1 \\ -1 & -1 & -1 & -1 & 4 \end{pmatrix},$$

and $\det \mathbf{F_1} = 0$, as all the rows (and columns) sum to zero. The matrix $\mathbf{F_1}$ is analogous to the ($v \times v$) Laplace[40,48–53] (Kirchhoff,[5,6,8,9,27,28,54,55] Admittance[4]) matrix ($\mathbf{K}$) that arises in the Matrix Tree Theorem.[3–6,8,9,36,54,55]

The sub-determinant obtained by striking out the row labelled δ and the column labelled δ in the table is

$$\det \begin{pmatrix} 3 & -1 & 0 & -1 \\ -1 & 3 & -1 & 0 \\ 0 & -1 & 3 & -1 \\ -1 & 0 & -1 & 3 \end{pmatrix} = 45.$$

THE CYCLE THEOREM (CT) AND ITS VERTEX-BASED DUAL, THE DUAL CYCLE THEOREM (DCT)

The Cycle Theorem establishes a formula for $t(G)$, the number of spanning trees in an embedded graph G, in terms of the adjacencies of its cycles. The formula is[29,30]

$$t(G) = \frac{\det \mathbf{M}}{\det \mathbf{M}^*}. \qquad (4)$$

$\mathbf{M}$ is the *cycle-overlap matrix*[27,29,30,40] for a full set of independent cycles, both boundary and generic; please see Ref. [29] for a definition of 'generic'. Disregarding the *direction* of the cycle — that is, considering only *circuits*, as previously defined (§3 of 'Introductory Remarks') — we see that such a set contains $e-(v-1)$, *i.e.*, $e-v+1$, circuits,[45] as it can be created by adding edges, one at a time, to a spanning tree of G. This number is usually denoted by μ, defined as follows:[1,44,45,56]

$$\mu = e - v + 1. \qquad (5)$$

$\mathbf{M}^*$ is a modification of $\mathbf{M}$ following the deletion from G of the edges of an arbitrarily chosen spanning tree. The conditions ensuring that $\det \mathbf{M}^* = 1$ have been investigated.[29,30]

μ can also be interpreted as giving the minimum number of edges that must be deleted from G in order to break every circuit, whether boundary[29,30] or generic.[29,30] This minimum number (usually called the *cyclomatic* number) is clearly equal to $f-s+k$ since $f-s$ is the number of edges in a ladder (so that their deletion just breaks, or opens, the boundaries of all faces) while k accounts for the opening of a set of independent generic circuits. This alternative expression for μ also follows from Equation (5) *via* Equation (1), automatically indicating that the sought circuits come from the ($f-s$) independent face-boundaries supplemented by k generics. This is because $\mu = e-v+p \Rightarrow \mu = (f-s)+k$, with $p = 1$ because we are here dealing with one-component graphs, and $s = 1$, as explained earlier.

With Equation (4) as model, we look for a formula

$$I(G) = \frac{\det \mathbf{M}}{\det \mathbf{M}^*}. \qquad (6)$$

Here, $\mathbf{M}$ would be the 'vertex-overlap' matrix for the appropriate number of independent vertices, namely, $e-f+s$ (guided by Equation (5)). In considering what we mean by 'independent vertices', it should be recalled that the Laplace/Kirchhoff matrix[9,27] for the adjacencies between the vertices of a graph has zero determinant, indicating, in some sense, 'dependence' of the vertices. In visual terms, if a vertex is deleted from a graph, the 'loose ends' of the affected edges could only have terminated in this vertex, which is, therefore, in that sense, 'dependent' on the others. In other words, only ($v-p$) vertices of a graph with p components are independent.[27] Note that ($f-s$) is the number of edges in any ladder of G, just as ($v-1$) is the number of edges in any spanning tree of G. Similarly, $\mathbf{M}^*$ would be a modification of $\mathbf{M}$ following the deletion from G of the edges of an arbitrarily chosen ladder. The 'appropriate' number just referred to, $e-f+s$, can also be expressed (from Equation (1)) as $v-p+\mathrm{k}$, indicating that the ($v-p$) independent vertices will be needed, and k more.

If G is a planar graph embedded without the use of generics,[29] we have $k = 0$ and the expressions needed for Equation (6) can be compiled.

Example 2. The Case $k = 0$.

Figure 2 (i) shows G as two C_3 graphs embedded on a torus and Figure 2 (ii) shows the same embedding after the deletion of the edges of a ladder.

Parameter Census: $f = 3$, $v = 6$, $e = 6$, $p = 2$, $s = 1$ and $k = 0$.

For this plane graph, $v-p = 6-2 = 4$: for the four independent vertices we must choose two pairs, which may be selected arbitrarily — say, A, B, and D, E.

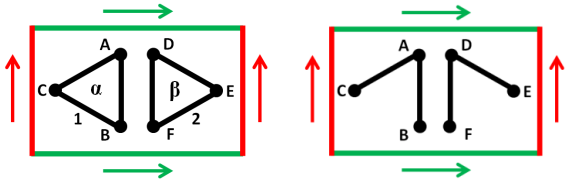

Figure 2. (i) Two C_3 graphs embedded on a torus; (ii) the same embedding after the deletion of the edges of a ladder.

For this embedding, by means of the following table,

	A	B	D	E
A	2	–1	0	0
B	–1	2	0	0
D	0	0	2	–1
E	0	0	–1	2

we construct the vertex-vertex overlap-matrix for the independent vertices, which is denoted $\mathbf{M}_2$.The matrix $\mathbf{M}_2$ is formed from the above table in an obvious manner, exactly analogous to the way in which the matrix $\mathbf{F}_1$ was obtained from the corresponding table preceding it in Example 1, so that $\det \mathbf{M}_2 = 9$. For $\mathbf{M}_2^*$, the corresponding table is

	A	B	D	E
A	2	–1	0	0
B	–1	1	0	0
D	0	0	2	–1
E	0	0	–1	1

and the matrix $\mathbf{M}_2^*$ formed from this table is such that $\det \mathbf{M}_2^* = 1$.

From Equation (6),

$$I(G) = \frac{9}{1} = 9 .$$

and that is clearly the correct number of ladders present.

A problem arises when $k > 0$. The number $v - p + k$ is greater than $v - p$, the number of independent vertices, and thus $\mathbf{M}$ cannot be compiled from their adjacencies. A somewhat counter-intuitive remedy is to supply k 'artificial vertices' (AV), suitably defined.

Each of the k AVs is associated, 1:1, with a generic cycle[29] that the host surface is capable of supporting. This generic[29] has the following features:

(i) It does not contain any vertex of G.

(ii) It crosses at least one — and preferably only one — edge that belongs to a generic[29] of G of a different kind from its own. Such crossing points do not make a vertex.

(iii) It is endowed, arbitrarily, with a direction of traversal, its 'sense'.

(iv) If X and X label the two appearances of the 'identified point' present in this cycle, the associated AV is labelled X. The cycle itself is denoted by XX, if required.

(v) The parameter census of the embedding is not altered by the introduction of AVs.

The artificial vertex X is considered to be adjacent to a vertex A of G if and only if the cycle XX crosses $n > 0$ edges that belong to A. In $\mathbf{M}$, this adjacency is marked by n if A lies to the right as XX is traversed in its sense, and by $-n$ if A lies to the left. This is, of course, an arbitrary convention and its opposite would serve just as well. The degree of X is equal to the number of edges crossed by XX.

(It might be thought that the artificial vertices could be identified as actual vertices, possibly in an additional dimension, 'floating' in the vicinity of the actual vertices. The necessary asymmetry of the rules for their adjacencies militates against this.)

Example 3. The Case $k > 0$.

Figure 3 (i) shows the complete graph K_5 embedded on a double torus (as in Figure 1) with the addition of two AVs, X and Y. Figure 3 (ii) is the same after the deletion of the edges of a ladder. The parameter census is the same as in Example 1 (Figure 1):

$f = 5$, $v = 5$, $e = 10$, $p = 1$, $s = 1$ and $k = 2$.

$\mathbf{M}_3$ and $\mathbf{M}_3^*$ can now be compiled as follows: the corresponding table for the matrix $\mathbf{M}_3$ is

	B	C	D	E	X	Y
B	4	–1	–1	–1	0	–1
C	–1	4	–1	–1	1	0
D	–1	–1	4	–1	0	1
E	–1	–1	–1	4	–1	0
X	0	1	0	–1	1	0
Y	–1	0	1	0	0	1

and the matrix $\mathbf{M_3}$ derived from it is such that $\det \mathbf{M_3} = 45$.

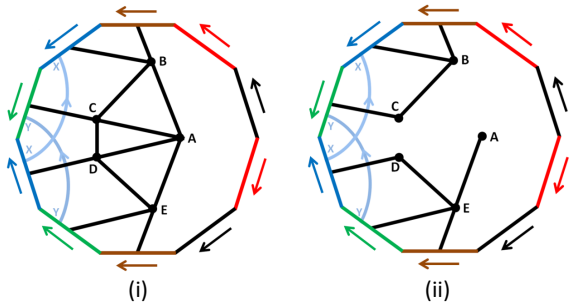

Figure 3. (i) The complete graph K_5 embedded on a double torus (as in Figure 1) with the addition of two artificial vertices, X and Y; (ii) the same as (i) after the deletion of the edges of a ladder.

The appropriate table for the matrix $\mathbf{M}_3^*$ is

	B	C	D	E	X	Y
B	3	–1	–1	–1	0	–1
C	–1	2	0	–1	1	0
D	–1	0	2	–1	0	1
E	–1	–1	–1	4	–1	0
X	0	1	0	–1	1	0
Y	–1	0	1	0	0	1

which leads to det $\mathbf{M}_3^* = 1$.

Hence

$$I(G) = \frac{\det \mathbf{M}_3}{\det \mathbf{M}_3^*} = \frac{45}{1} = 45\,,$$

in agreement with the value obtained earlier from a co-factor of $\mathbf{F}_1$.

Example 4. Another case of $k > 0$ with an embedding on a hexagonal representation of a torus.

Figure 4 (i) shows the graph $K_{3,3}$ (the 'utilities' graph) embedded on a torus and Figure 4 (ii) shows the same embedding after deletion of the edges of a ladder.

Parameter census: $f = 3$, $v = 6$, $e = 9$, $p = 1$, $s = 1$ and $k = 2$.

$\mathbf{M}_4$ can now be compiled as follows: the corresponding table for the matrix $\mathbf{M}_4$ is

	A	B	C	D	E	X	Y
A	3	–1	0	–1	0	0	1
B	–1	3	–1	0	–1	1	–1
C	0	–1	3	–1	0	–1	–1
D	–1	0	–1	3	–1	–1	1
E	0	–1	0	–1	3	1	0
X	0	1	–1	–1	1	2	0
Y	1	–1	–1	1	0	0	2

so that det $\mathbf{M}_4 = 27$.

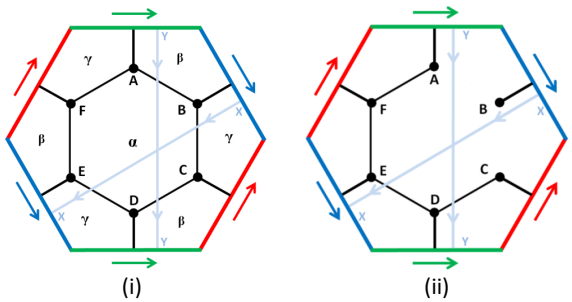

Figure 4. (i) The graph $K_{3,3}$ (the 'utilities' graph) embedded on a torus; (ii) the same embedding after deletion of the edges of a ladder.

The appropriate table for the matrix $\mathbf{M}_4^*$ is

	A	B	C	D	E	X	Y
A	2	0	0	–1	0	0	0
B	0	1	0	0	–1	0	0
C	0	0	2	–1	0	0	–1
D	–1	0	–1	3	–1	–1	1
E	0	–1	0	–1	3	1	0
X	0	0	0	–1	1	1	0
Y	0	0	–1	1	0	0	1

leading to det $\mathbf{M}_4^* = 1$.

Hence

$$I(G) = \frac{\det \mathbf{M}_4}{\det \mathbf{M}_4^*} = \frac{27}{1} = 27\,,$$

in agreement with the value obtained from the matrix of the overlap of the two independent faces chosen, namely α and β. See Figure 4 (i) — thence

$$\det \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix} = 27.$$

THE VALUE OF DET M*

That in the examples given the value of det $\mathbf{M}^*$ has consistently been 1 is no coincidence. It is known that in the Cycle Theorem (CT) calculations[29,30] the value of det $\mathbf{M}^*$ in Equation (4) can be assumed to be 1 whenever the cycles chosen for $\mathbf{M}$ originate from a so-called 'fundamental set'[1,44,45] of cycles.[27,29,30] In our vertex-based Dual Cycle Theorem (DCT) the criteria for a fundamental set in relation to the vertices available are met. So, as in the CT, the value of det $\mathbf{M}^*$ is inevitably 1. If, instead of single vertices, we had considered vertex-clusters, values of det $\mathbf{M}^*$ different from 1 would have occurred in some cases.

This point really goes back to the question raised by '|det $\mathbf{U}$|', defined in Ref. [29] (and '|det $\mathbf{M}^*$|' in Ref. [30]), in the original Cycle Theorem [29,30]. Usually — as was argued in Ref. [29] — the moduli of these determinants are guaranteed to have a value of 1. The moduli of these determinants in the Cycle Theorem are not equal to one only if the cycles are re-aggregated — as they are in the illustrative Example on Page 268 (right-hand column) of Ref. [29] — in what might be regarded as an artificially complicated, and somewhat *outré*, way. (In that latter example, for instance, det $\mathbf{U} = 2$.) In the Dual Cycle Theorem — the one which is being proposed in the present paper — it is individual vertices that are being dealt with, and not vertex clusters.

AN APPLICATION: THE PATENCY INDEX

In a one-component graph, G, with v vertices and e edges, any spanning tree must contain precisely $(v-1)$ edges. The number of ways of choosing any $(v-1)$ edges from the e available edges is ${}^{e}C_{(v-1)}$. (This implies that a graph containing only one vertex — *e.g.*, Figure 5 — has that vertex as a — or *the* — spanning tree.) In general, of course, not every set of $(v-1)$ edges so selected will constitute the edges of a spanning tree of G. With $t(G)$ representing the number of spanning trees within G, Trinajstić and one of the present authors (RBM) defined[25] the *Spanning-Tree Density* of G (denoted by the symbol $STD(G)$) as the ratio

$$STD(G) = \frac{t(G)}{{}^{e}C_{(v-1)}} . \qquad (7)$$

Mallion & Trinajstić[25] suggested that this may be thought of as representing the *probability* that, if *any* $(v-1)$ *edges* are randomly selected, and the other $(e-v+1)$ edges in G are deleted, the resulting entity is a spanning tree. This Index obeys the inequality:

$$0 < STD(G) \leq 1 .$$

The left-hand side of this inequality is correct (being '0 <' rather than '0 ≤') because we do not admit a 'null graph', *i.e.*, one with neither vertices nor edges. Such an entity would just be a vacuum, not a graph; hence, the numerator in the expression for $STD(G)$ (defined in expression (7), above) cannot be zero. The highest extremal value of 1 can occur for $STD(G)$ in various ways: for example, in the case of the circuit graphs, C_n, (including C_1) and also in the case of a graph without circuits, as such a graph is already a tree. $STD(G)$ is also equal to 1 in certain artificially constructed graphs such as a self-loop C_1 or a graph (illustrated in Figure 5 for $n = 6$) consisting of n C_1-sub-graphs, with one vertex in common. For the self-loop C_1, $t(C_1) = 1$, as the single vertex that results when the lone edge is deleted is admitted as a tree, being a connected graph with no circuits (and no edges, either)

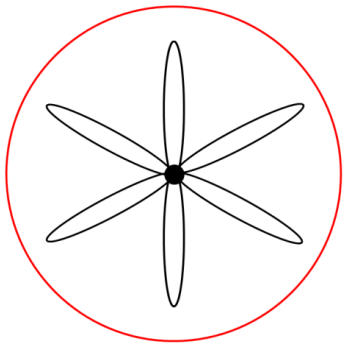

Figure 5. A graph (G) specifically designed to have $STD(G) = 1$ consists of n C_1-sub-graphs, with one vertex in common (illustrated here for $n = 6$), embedded on the surface of a sphere. The convention for illustrating the embedding is described below.

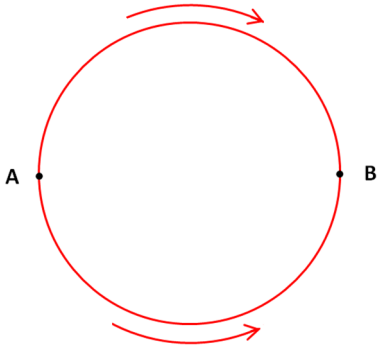

Figure 6. The representation of a spherical surface, as discussed in the text below.

Before we examine this and other examples, we explain our plane representation of a spherical surface. We shall represent a spherical surface as a circular disc bounded by two 'directly identified' semi-circumferences, say AB. By 'directly identified' we mean that each represents the same points of the spherical surface and in the same order. This is analogous to other well-known plane representations, for example, that of a torus.[29] The full diagram will be as in Figure 6.

The arrows, which indicate that the identification is direct, will in future be omitted (as they were in Figure 5) if clarity is not impaired, as will the points marked A and B (also as in Figure 5). This representation ensures that the points of the disc cannot sustain a *generic circuit*[29] (or cycle) — as is also the case with the original spherical surface, of course.

As argued above, $STD(G)$ is never precisely zero although, by a suitable choice of graph, it can be made as close to zero as desired. For example (Example 5), for what might be called the 'spotted' n-petal graph, illustrated — embedded on a sphere — in Figure 7 (for $n = 5$), it is easy to see that

$$STD(\text{'spotted' } n\text{-petal graph}) = \frac{2^n}{{}^{2n}C_n} .$$

A value of n may be found to make STD('spotted' n-petal graph) as small as is desired, and it is formally proved in the Appendix that

$$\frac{2^n}{{}^{2n}C_n} \to 0, \text{ as } n \to \infty ,$$

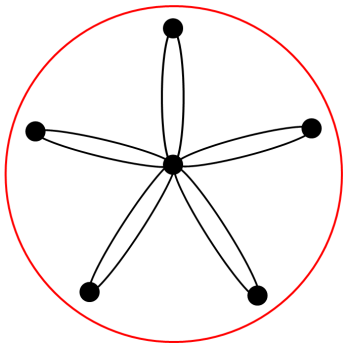

Figure 7. The 'spotted' *n*-petal graph (with $n = 5$), embedded on the surface of a sphere. The sphere is represented as a disc bounded by the red circle, as explained above in connection with Figures 5 and 6.

so that

$$STD(\text{'spotted'}\ n\text{-petal graph}) \to 0, \text{ as } n \to \infty .$$

In the context of chemical complexity[11–26] the reciprocal of *STD*(*G*) was suggested[25] as defining a quantitative measure of what Mallion & Trinajstić[25] called the *Intricacy* of *G* — the bigger the numerical value of this reciprocal the more *intricate* (in the sense of being 'complicated' or 'reticulated') the given graph, *G*, is.

By analogy with this idea, we here suggest defining what we call the *Patency Index*, *P*(*G*), for a one-component graph embedded on a closed host surface, as

$$P(G) = \frac{l(G)}{{}^{e}\mathrm{C}_{(f-1)}}, \qquad (8)$$

where *l*(*G*) is the number of ladders in *G* (defined in the expression labelled (6), above) and *f* (as throughout) is the number of faces in *G*. (If *G* has no bounded faces, then the host surface itself is a — or *the* — face, as already stated.) The denominator in (8) arises because the number of edges in a ladder is (*f*–1). (Note that, between definition (7) and definition (8), '*l*(*G*)' replaces '*t*(*G*)' and '*f*' replaces '*v*'; the expression labelled (8) is, therefore, properly considered as being, in that sense, the *dual* of the expression presented as (7), and hence within the area of interest of this paper's (partial) title, 'A Dual of the Cycle Theorem'.)

The value of the newly defined Patency Index gives the *probability* that an arbitrarily selected set of (*f*–1) edges of *G* constitutes a ladder. The index *P*(*G*) satisfies

$$0 < P(G) \leq 1 .$$

A high value of *P*(*G*) thus indicates an 'openness', or 'patency', in the structure of *G*.

This inequality is analogous to the one encountered earlier in the case of *STD*(*G*). Once again, as with the inequality on *STD*(*G*), the left-hand side of the inequality on *P*(*G*) is '0 <' rather than '0 ≤ ', but the reason for this is different. For *P*(*G*), the numerator cannot be zero because a null-ladder has been defined. A ladder is a *set* of edges, not a graph; an 'empty set' is acceptable and so, therefore, is a null-ladder. Graphs having *P*(*G*) as close as desired to zero do exist — the graph in Figure 7 is an example — but zero itself cannot be attained because *l*(*G*), the numerator of *P*(*G*), cannot equal zero, on account of the possibility of the null-ladder.

As for the right-hand part of the above equality, the extreme value *P*(*G*) = 1 is attained in the same cases as was observed with *STD*(*G*) — for example, if *G* is an acyclic graph (that is, a tree), a C_n-type graph, or a set of *n* C_1-sub-graphs with a single vertex in common (as illustrated in Figure 5, for the case of $n = 6$). For the first of these, we have

$$P(\text{a Tree}) = \frac{l(G)}{{}^{e}\mathrm{C}_{(f-1)}} = \frac{l(G)}{{}^{e}\mathrm{C}_{0}} = \frac{1}{1} = 1 ,$$

l(*G*) being equal to 1 on account of the null-ladder, defined earlier (in the Section 'The Definition of a *Ladder*'). [Note that, as observed above, the null-ladder is an (empty) set of edges, whereas the 'null-tree' would be a vacuum in space.]

The expression presented in definition (8) (above) does seem to fit in with the intuitive idea of 'openness' of graphs and molecular structures.[25] It may be observed that, for a *planar* embedding, the edges of a ladder are those edges that are not in a spanning tree; evidently, therefore, in these cases, the number of spanning trees, *t*(*G*), is the same as the number of ladders, *l*(*G*), and so the numerators in the definitions of *STD*(*G*) (the expression labelled (7)) and *P*(*G*), (the expression labelled (8)) will be the same. Furthermore, for a connected graph with such a planar embedding, because

(a) the Generalised Euler Polyhedron Theorem[47] (Equation (1)), in this case of a one-component (connected) graph, becomes

$$v + f = e + 2 \quad (\text{which implies that } v - 1 = e - f + 1)$$

and

(b) ${}^{n}\mathrm{C}_{r} = {}^{n}\mathrm{C}_{(n-r)}$,

it follows that the denominator in the definition of *STD*(*G*) (labelled (7)) — namely, $({}^{e}\mathrm{C}_{(v-1)})$ — is equal to the denominator $({}^{e}\mathrm{C}_{(f-1)})$ in the definition of *P*(*G*) (labelled (8)) since, from the above, we have that

$${}^{e}\mathrm{C}_{(v-1)} = {}^{e}\mathrm{C}_{(e-f+1)} = {}^{e}\mathrm{C}_{(e-\{e-f+1\})} = {}^{e}\mathrm{C}_{(f-1)}.$$

This means that the denominator of the definition of the Patency Index (labelled (8)) is equal to the corresponding denominator of the definition of Spanning Tree Density ((7)) and because, as already seen, the

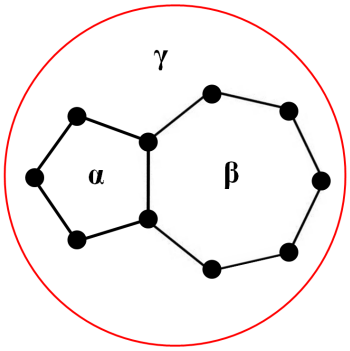

Figure 8. The molecular graph of the azulene molecule has three faces, α, β and γ, when embedded on the surface of a sphere. The sphere is represented as a disc bounded by the red circle, as explained above in connection with Figures 5 and 6.

numerators of the two expressions are likewise the same when the embedding is planar, it follows that, in such a case,

$$P(G) = STD(G).$$

Example 6. This situation is exemplified by the molecular graph of the conjugated hydrocarbon azulene, illustrated in Figure 8 as being embedded on a sphere.

The parameter census is:

$f = 3$, $v = 10$, $e = 11$, $p = 1$, $s = 1$ and $k = 0$.

Evidently — see, for example, the formulae in Ref. [32] — $t(\text{azulene}) = (5 \times 7) - 1 = 34$. Furthermore, ${}^{e}C_{(v-1)} = {}^{11}C_{(10-1)} = {}^{11}C_9 = {}^{11}C_2 = 55$. Hence, from Equation (7), $STD(\text{azulene}) = {}^{34}\!/_{55} = 0.618$ (to 3 significant figures). This may be thought of as representing the probability that if any set of $(v - 1) = (10 - 1) = 9$ edges in the azulene molecular graph is selected, and the other $(e - v + 1) = (11 - 10 + 1) = 2$ edges are deleted, the resulting entity is a spanning tree.

Now turning to the calculation of $P(\text{azulene})$, we note that the table for the face-to-face incidence matrix is

	α	β	γ
α	5	–1	–4
β	–1	7	–6
γ	–4	–6	10

The ensuing matrix, **A**, is given by

$$\begin{pmatrix} 5 & -1 & -4 \\ -1 & 7 & -6 \\ -4 & -6 & 10 \end{pmatrix}.$$

As is to be expected, $\det \mathbf{A} = 0$. The several (2×2) sub-determinants of **A** have the common absolute value of 34. Consequently, $I(\text{azulene}) = 34$. Furthermore, ${}^{e}C_{(f-1)} = {}^{11}C_{(3-1)} = {}^{11}C_2 = 55$. Hence, from definition (8), $P(\text{azulene}) = {}^{34}\!/_{55}$, which, as anticipated, is the same as $STD(\text{azulene})$.

The situation identified in Example 6 — namely, the equality of the indices $STD(G)$ and $P(G)$ — will occur whenever the graph, G, in question is (a) in a planar embedding, *and* (b) is a *connected* graph (that is, G consists of just one component: $p = 1$). This will in general be the case for the graph that represents any extant or potentially viable molecule.

EPILOGUE

The great Cayley's papers (for example, his classic one on trees[57] — see the comments on pp. 4 & 5 of Ref. [3]) have been described as follows (alluding to the propensity of contemporary botanists cumulatively to collect and deposit 'specimens'):[58]

'They [*i.e.*, Cayley's papers] are typically discursive, contain little formal proof, and many of them simply assemble the specimens.'

We have assembled many more 'specimens' than are included in this paper, all in agreement with the proffered result. We could not, therefore, be in better company.

It is, of course, a merit of the general concept of duality that a suitable proof of a proposition in one context ensures the existence of an analogous proof of its dual — even when that proof is not explicitly formulated.

Finally, the idea of the Patency Index, as defined in this paper, will be further applied and evaluated in a future publication so that its usefulness, or otherwise, may be critically assessed in the context of what has become known as 'molecular complexity'.[11–26] Already, a referee has suggested that possible candidates for application of the Patency Index might be the objects studied by Diudea and coworkers (in, for example, Ref. [59]) as well as various geometrical objects that are constructed for nano-technology needs from strands of DNA or polypeptides, such as those studied in the work of Jerala *et al.* (of which a representative example is Ref. [60].)

APPENDIX

Proof that

$$\frac{2^n}{{}^{2n}C_n} \to 0 \text{ as } n \to \infty.$$

$$\frac{2^n}{{}^{2n}C_n} = \frac{2^n n!n!}{(2n)!} = \frac{n!}{P_n},$$

where P_n is the product of all odd natural numbers $< 2n$. Now,

$$\frac{(n+1)!}{P_{n+1}} = \frac{(n+1)}{(2n+1)} \cdot \frac{n!}{P_n}.$$

For all $n \geq 1$,

$$\frac{n+1}{2n+1} \leq \frac{2}{3}.$$

So,

$$\frac{(n+1)!}{P_{n+1}} \leq \frac{2}{3} \frac{n!}{P_n}.$$

Now,

$$\frac{1!}{P_1} = 1,$$

and so:

$$\lim_{n \to 0} \left(\frac{2^n}{{}^{2n}C_n} \right) = 0. \qquad \square$$

REFERENCES

[1] R. J. Wilson, *Introduction to Graph Theory*, First Edition, Oliver & Boyd, Edinburgh, **1972**, pp. 46–48; Fifth Edition, Prentice Hall, Upper Saddle River, New Jersey, USA, **2010**, p. 63.

[2] N. Trinajstić, *Chemical Graph Theory*, First Edition, Vols. I & II, CRC Press, Boca Raton, Florida, USA **1983**; Revised Second Edition (all in one volume), CRC Press, Boca Raton, Florida, USA, **1992**.

[3] J. W. Moon, *Counting Labelled Trees*, Canadian Mathematical Monographs, Ottawa, **1970**, especially pp. 64–65 and pp. 4–5.

[4] D. M. Cvetković, M. Doob, H. Sachs, *Spectra of Graphs — Theory and Applications,* First Edition, Deutscher Verlag der Wissenschaften Berlin (East), **1979**; Second Edition, Academic Press, New York, **1980**; Third Edition, Johann Ambrosius Barth, Heidelberg & Leipzig, **1995**, especially pp. 38–41.

[5] H. Fleischner, *Eulerian Graphs and Related Topics* Part 1, Volume 2, North Holland, Amsterdam, **1991**, pp. IX 73–IX 78.

[6] W. T. Tutte, *Graph Theory As I Have Known It*, Oxford Lecture Series in Mathematics and its Applications, 11. (Series Eds.: G. Ball, D. Welsh), Clarendon Press, Oxford, UK, **1998**, especially pp. 7, 26, 27, 99 & 112.

[7] R. McWeeny, *Mol. Phys*, **1958**, *1*, 311.

[8] R. B. Mallion, *Proc. Royal Soc.* (*London*) *A* **1974/1975**, *341*, 429.

[9] T. K. Dickens, R. B. Mallion, *Croat. Chem. Acta* **2013**, *86*, 387.

[10] T. K. Dickens, R. B. Mallion, *MATCH Commun. Math. Comput. Chem.* **2016**, *76*, 297.

[11] A. Mowshowitz, *Bull. Math. Biophys*. **1968**, *30*, 175.

[12] D. Minoli, *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur*. (*Ser. 8*) **1975**, 59, 651.

[13] H. Sachs, 'On the number of spanning trees' in *Proceedings of the Fifth British Combinatorial Conference*, *Aberdeen*, *1975*, (Eds.: C. St. J. A. Nash-Williams, J. Sheehan), (*Congressus Numerantium XV*), Ultilitas Mathematica, Winnipeg, Manitoba, Canada, **1976**, pp. 529–535.

[14] S. H. Bertz, *J. Am. Chem. Soc*. **1981**, *103*, 3599.

[15] D. Bonchev, D. Kamenski, O. N. Temkin, *J. Math. Chem*. **1987**, *1*, 345.

[16] D. Bonchev, 'The problems of computing molecular complexity' in *Computational Chemical Graph Theory*, (Ed.: D. H. Rouvray), Nova Science Publishers, New York, **1990**, Chapter 2, pp. 34–63.

[17] M. M. Waldrop, *Complexity*, Touchstone/Simon & Schuster, New York, **1992**.

[18] B. Kaye, *Chaos and Complexity*, VCH, Weinheim, Germany, **1993**.

[19] S. Nikolić, N. Trinajstić, A. Jurić, Z. Mihalić, G. Krilov, *Croat. Chem. Acta* **1996**, *69*, 883.

[20] S. Nikolić, N. Trinajstić, I. M. Tolić, *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 920.

[21] R. Barone, M. Chanon, *J. Chem. Inf. Comput. Sci*. **2001**, *41*, 269.

[22] I. Gutman, C. Rücker, G. Rücker, *J. Chem. Inf. Comput. Sci*. **2001**, *41*, 739.

[23] M. Randić, *Croat. Chem. Acta* **2001**, *74*, 683.

[24] M. Randić, D. Plavšić, *Croat. Chem. Acta* **2002**, *75*, 107.

[25] R. B. Mallion, N. Trinajstić, *MATCH Commun. Math. Comput. Chem.* **2003**, *48*, 97.

[26] D. Bonchev, D. H. Rouvray, *Complexity in Chemistry, Biology and Ecology*, Mathematical and Computational Chemistry Series, (Ed.: P. G. Mezey), Springer, New York, **2005**.

[27] E. C. Kirby, R. B. Mallion, P. Pollak, P. J. Skrzyński, *Croat. Chem. Acta* **2016**, *89*, 403.

[28] I. Gutman, R. B. Mallion, J. W. Essam, *Mol. Phys.* **1983**, *50*, 859.

[29] E. C. Kirby, D. J. Klein, R. B. Mallion, P. Pollak, H. Sachs, *Croat. Chem. Acta* **2004**, *77*, 263.

[30] E. C. Kirby, R. B. Mallion, P. Pollak, 'Counting spanning trees in toroidal Fullerenes' in *The Mathematics and Topology of Fullerenes*, (Series Eds.: F. Cataldo, P. Milani), (Volume Eds.: F. Cataldo, A. Graovac, O. Ori), (Foreword: Sir Harold Kroto), Carbon Materials, Chemistry and Physics Volume 4, Springer, Dordrecht, Heidelberg, London & New York, **2010**, Chapter 10, pp. 187–204.

[31] D. M. Cvetković, I. Gutman, *Pub. Inst. Math*. (*Beograd*) **1981**, *29*, 49.

[32] B. O'Leary, R. B. Mallion, 'Counting the spanning trees of labelled, planar, molecular-graphs embedded on the surface of a sphere' in *Graph Theory and Topology in Chemistry*, (Eds.: R. B. King, D. H. Rouvray), Studies in Physical and Theoretical Chemistry, Vol. 51, Elsevier, Amsterdam, **1987**, pp. 544–551.
[33] T. J. N. Brown, R. B. Mallion, P. Pollak, B. R. M. de Castro, J. A. N. F. Gomes, *J. Comput. Chem.* **1991**, *12*, 1118.
[34] P. E. John, R. B. Mallion, *J. Math. Chem.* **1994**, *16*, 261. (See also *Erratum*, *J. Math. Chem.* **1994**, *16*, 389.)
[35] E. C. Kirby, R. B. Mallion, P. Pollak, *Mol. Phys*. **1994**.
[36] T. J. N. Brown, R. B. Mallion, P. Pollak, A. Roth, *Discrete Appl. Math.* **1996**, *67*, 51. (See also *Erratum, Discrete Appl. Math.* **1997**, *75*, 199.)
[37] J. J. Dambit, *Latvian Math. Yearbook*, Izdat. Zinatne, Riga, **1965**, pp. 337–345.
[38] J. Sedláček, *Časopis pro Pěstováni* **1966**, *91*, 221. [In Czech].
[39] N. L. Biggs, *Algebraic Graph Theory*, Cambridge University Press, Cambridge, UK, **1974**, p. 38.
[40] D. Janežič, A. Miličević, S. Nikolić, N. Trinajstić, *Graph Theoretical Matrices in Chemistry*, (Eds.: I. Gutman, B. Furtula), Mathematical Chemistry Monographs, Faculty of Science, University of Kragujevac, Serbia, **2007**, Chapter 2, pp. 5–50, especially pp. 46–49.
[41] P. R. Bryant, 'Graph theory applied to electrical networks' in *Graph Theory and Theoretical Physics*, (Ed.: F. Harary), Academic Press, London, **1967**, Chapter 3, pp. 111–137.
[42] P. R. Bryant, 'Graph theory and electrical networks' in *Applications of Graph Theory*, (Eds.: R. J. Wilson, L. W. Beineke), Academic Press, London, **1979**, pp. 81–119, especially pp. 84–85.
[43] J. A. Bondy, U. S. R. Murty, 'The cycle space and the bond space' in *Graph Theory with Applications*, Macmillan, London, **1976** *and* Elsevier, New York, **1976**, Chapter 12, pp. 212–226.
[44] G. Kirchhoff, *Annal. Physik u. Chemie* ('*Poggendorff's Annalen*') **1847**, *72*, 497.
[45] W. Ahrens, *Math. Annal.* **1897**, *49*, 311.
[46] E. C. Kirby, R. B. Mallion, P. Pollak, *J. Chem. Soc. Faraday Trans.***1993**, *89*, 1945.
[47] E. C. Kirby, R. B. Mallion, P. Pollak, P. J. Skrzyński, *Int. J. Chemical Modeling* **2014**, *6*, 255.
[48] B. Mohar, 'Laplacian matrices of graphs' in *MATH/CHEM/COMP 1988* (Ed.: A. Graovac), Elsevier, Amsterdam, **1989**, pp. 1–8.
[49] B. Mohar, 'The Laplacian spectrum of graphs' in *Graph Theory*, *Combinatorics and Applications*, (Eds.: Y. Alavi, G. Chartrand, O. T. Ollermann, A. J. Schwenk), John Wiley & Sons Inc., New York, **1991**, pp. 871–898.
[50] N. Trinajstić, Z. Mihalić, F. E. Harris, *Int. J. Quantum Chem*. **1994**, *52*, 525.
[51] Z. Mihalić, N. Trinajstić, *Fullerene Science & Technology* **1994**, *2*, 89.
[52] N. Trinajstić, D. Babić, S. Nikolić, D. Amić, Z. Mihalić, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 368.
[53] P. W. Fowler, *MATCH Comm. Math. Comput. Chem*. **2003**, *48*, 87.
[54] M. Kunz, *J. Math. Chem.* **1992,** *9*, 297.
[55] R. B. Mallion, *Chem. Phys. Letters* **1975**, *36*, 170.
[56] P. John, *Wissenschaftliche Zeitschrift der Technischen Hochschule Ilmenau* **1989**, *35*, 41.
[57] A. Cayley, *Quart. J. Pure Appl. Math.* **1889**, *23*, 376.
[58] T. Crilley, *Arthur Cayley*: *Mathematician Laureate of the Victorian Age*, Johns Hopkins University Press, Baltimore, Maryland, USA, **2006**, p. 195.
[59] M. V. Diudea, C. L. Nagy, *Periodic Nanostructures*. Series: Developments in Fullerene Science, Vol. *7*, Springer Netherlands, Springer Science & Business Media BV, Dordrecht, The Netherlands, **2007**.
[60] H. Gradišar, S. Božič, T. Doles, D. Vengust, I. Hafner-Bratković, A. Mertelj, B. Webb, A. Šali, S. Klavžar, R. Herala, *Nature Chem. Biol*. **2013**, *9*, 362.