Kekulé Cells for Molecular Computation

W.H. Hesselink [1], J.C. Hummelen [2], H.T. Jonkman [2],
H.G. Reker [1], G.R. Renardel de Lavalette [1], M.H. van der Veen [2]
{w.h.hesselink,j.c.hummelen,harry.jonkman}@rug.nl,
rekerh@cs.man.ac.uk, g.r.renardel.de.lavalette@rug.nl,
marleen.vanderveen@yale.edu
February 17, 2013
[1] Computer Science, University of Groningen, The Netherlands
[2] Molecular Electronics, Zernike Institute of Advanced Materials,
University of Groningen, The Netherlands

Abstract. The configurations of single and double bonds in polycyclic hydrocarbons are abstracted as Kekulé states of graphs. Sending a so-called soliton over an open channel between ports (external nodes) of the graph changes the Kekulé state and therewith the set of open channels in the graph. This switching behaviour is proposed as a basis for molecular computation. The proposal is highly speculative but may have tremendous impact.

Kekulé states with the same boundary behaviour (port assignment) can be regarded as equivalent. This gives rise to the abstraction of Kekulé cells. The basic theory of Kekulé states and Kekulé cells is developed here, up to the classification of Kekulé cells with ≤ 4 ports. To put the theory in context, we generalize Kekulé states to semi-Kekulé states, which form the solutions of a linear system of equations over the field of the bits 0 and 1. We briefly study so-called omniconjugated graphs, in which every port assignment of the right signature has a Kekulé state. Omniconjugated graphs may be useful as connectors between computational elements. We finally investigate some examples with potentially useful switching behaviour.

1 Introduction

In the quest for smaller and smaller computational elements, we may hope to arrive one day at the level of molecules. Controllable electrical conductance within molecules is the realm of chemistry. One proposal in this direction is to use so-called π-conjugation in polycyclic hydrocarbons, as studied in Marleen van der Veen's PhD thesis "π-Logic" [17].

The development of this field of π-conjugation and π-logic needs several abstractions that belong to branches of mathematics like graph theory and linear algebra. In mathematics, the symbol π is primarily associated with the circumference and the area of the circle. The term "conjugation" has also several connotations in mathematics. In this paper, conjugation means the constructive interaction between a pair of neighbouring (carbon-carbon) π-orbitals, leading to delocalisation of the electrons.

The basic physical idea is that the configuration of single and double bonds in certain polycyclic 'aromatic' [note: these are polyunsaturated hydrocarbons, usually referred to as polycyclic aromatic hydrocarbons (PAHs), but they need not be aromatic in the strict chemical sense] hydrocarbons influences the electrical conductivity between points of the molecule, and can be influenced by electrical signals over channels in the molecule. In other words, the molecule can serve as a switch. Since it was Kekulé who proposed, in 1865, the alternating single and double bonds in the benzene ring (one of the simplest cyclic hydrocarbons), we prefer to associate the basic ideas to be exposed here with the name Kekulé.

The polycyclic hydrocarbons we are considering have boundary atoms that can serve as *ports* to probe and modify the electronic properties of the molecules. The electrical resistance between two ports is low when there is a path of alternating single and double bonds between them [20]. By sending a so-called *soliton* over the alternating path, the single and double bonds along the path are toggled [6]. Such toggling of an alternating path may open or close other alternating paths in the molecule. This is the switching behaviour alluded to. The toggling can also be done by chemical means (for example a redox reaction [1,19]).

In our abstraction of the molecule, the graph of the atoms and bonds is kept fixed, while it is allowed to change the multiplicities (single or double) of the bonds. A configuration of bonds such that every internal node has precisely one double bond is called a *Kekulé state*. Nodes with precisely one edge to the remainder of the graph are called *ports*. The *port assignment* of a Kekulé state describes the multiplicities of the bonds at the ports. A pair of ports is called a *channel*. A channel is called *open* (low resistance) in a Kekulé state if there is an alternating path between its ports.

Another reason for naming the configurations Kekulé states is that they are to represent closed shell molecules ('Kekulé structures'), i.e., molecules in which all electrons are paired, as opposed to non-Kekulé states.

The term 'Kekulé state' has its chemical equivalent in 'resonance structure' (or 'resonance contributor'), with the restriction of being a system that is closed-shell and without charges. The number of different Kekulé states of the graph is a measure of the stability of the molecule. In this sense, Kekulé states represent all structures that are regarded in the Valence Bond Theory [8,7,2,15] of molecules (as opposed to Molecular Orbital Theory, that also includes charged and open shell configurations as parts for the total quantum mechanical description of the electronic structure of a molecule).

The Kekulé state is a debatable abstraction. The actual quantum-mechanical state is a weighted superposition of many states, in which the Kekulé states have high weights; e.g., the two Kekulé states of the benzene hexagon are just two components of a single quantum-mechanical state. This does not matter for the switching behaviour described, however, because it turns out that all Kekulé states with the same port assignment have the same open (closed) channels between ports, see Theorem 2 in section 2.2. It follows that the effect of sending

a soliton over an open channel only effects the port assignment, and consists of toggling the port assignment only at the ports of the channel.

We introduce Kekulé cells to capture this behaviour. More precisely, we introduce a mathematical concept *cell* that captures the behaviour, and Kekulé cells are those that can be obtained from graphs with single and double bonds in them.

A serious physical objection is that the Kekulé states form a qualitative characteristic of the state, whereas the precise energy levels of the various eigenstates are quantitative. This objection must be dealt with when the qualitative investigations are leading to actual technical proposals.

Van der Veen et al. [18] have first proposed certain π-conjugated systems that can act as 'soldering points' for molecular wires in the sense that linear π-conjugated pathways between all ports ('terminals') exist; i.e., 'omniconjugated' systems. Subsequently, π-logic [17] was proposed as a way to perform any Boolean operation within certain π-conjugated hydrocarbon frameworks, bearing a number of ports. We now formalize this approach in a rigorous mathematical and topological way.

The application of graph theory to chemistry is not new. One of the research areas is the search for quantitative structure-property relationships based on molecular connectivity invariants of the hydrogen-suppressed chemical graphs, see the overview article [13]. The invariants are numerical functions of graphs that can be related to properties of classes of compounds like, e.g., solubility, crystal density, melting point, etc. In [16], it is proposed that one of these invariants (a generalized Randic index) for a certain class of hydrocarbons can be interpreted as an energy functional that depends on the π-electron density. The paper [4] goes beyond numerical invariants and studies energy levels of π-conjugated systems by means of the symmetry groups of the graphs. In our application of graph theory, numerical invariants and the symmetry of the graphs play no roles yet.

1.1 Overview

The formalization steps we need are first sketched in section 1.2. In section 1.3, we present the mathematical notations that we need. The graph theory needed is presented in section 1.4.

In section 2, we present the theory of Kekulé states and Kekulé cells, up to the classification of Kekulé cells with four ports. Section 3 contains the theory of semi-Kekulé states. Here it is proved that the number of Kekulé states for a given port assignment is $\leq 2^{e+1-v}$ where e is the number of edges and v is the number of nodes of the graph. This section serves as an aside to provide context for the Kekulé states, but is not really needed for the application. In section 4, we discuss graphs for which the Kekulé cell is as large as possible; such graphs are said to be omniconjugated. In section 5, we investigate the functionality of cells for switching behaviour. Conclusions are drawn in section 6.

1.2 The formalization steps

We now briefly sketch the central concepts of the theory. Precise definitions are postponed.

It is natural and conventional to represent molecules by graphs with the nodes for atoms and the edges for bonds. Polycyclic polyunsaturated hydrocarbons are usually almost flat, so that their graphs are planar. We use a subgraph to indicate the set of double bonds of a given molecule.

Nodes of the graph that are linked via only one edge to the remainder of the graph are called *ports*. A *port assignment* is a set of ports used to indicate the ports that are to be attached by a double bond. The *Kekulé cell* associated to a graph with port set P is the set of port assignments in P that allow Kekulé states.

A *cell* in P is defined to be an arbitrary set of port assignments in P. A cell is a *Kekulé cell* if there is some graph for which the cell is the set of Kekulé port assignments. A *functional cell* is a cell together with an initial port assignment and a system of channels to probe and modify the current port assignment.

1.3 Mathematical concepts and notations

In this paper, a number of mathematical theorems (and lemmas and corollaries) are proved. The end of such a proof is indicated by the symbol □. The Theorems 1, 2, and 3 and Corollary 1 have been verified with the mechanical theorem prover PVS [12] as a student project (the report and the proof scripts are available on request). The symbol □ is also used at the end of the examples and remarks.

If S is a set, we write $x \in S$ to denote that x is an element of S. We write $\#S$ to denote the number of elements of S.

If S and T are sets, we write $S \cup T$ for the *union* of S and T, which is the set of the elements x with $x \in S$ or $x \in T$ (or both). The *intersection* $S \cap T$ consists of the elements x with $x \in S$ and $x \in T$. We write $S \setminus T$ for the *difference* of S and T, i.e., the set of elements of S that are not in T.

A set S is called a *subset* of a set U (notation $S \subseteq U$) if every element of S is an element of U. The *power set* $Pow(U)$ of U is defined as the set of subsets of U.

We write $S \oplus T$ for the *symmetric difference*, which is $S \oplus T = (S \setminus T) \cup (T \setminus S)$. It is easy to verify that $\oplus$ is commutative: $S \oplus T = T \oplus S$ and associative: $(S \oplus T) \oplus R = S \oplus (T \oplus R)$, and that the empty set $\emptyset$ satisfies $\emptyset \oplus S = S$ and $S \oplus S = \emptyset$. Therefore, $Pow(U)$ with operation $\oplus$ is a commutative group with neutral element $\emptyset$.

1.4 Undirected graphs

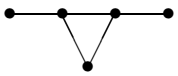
There is no standard terminology in graph theory, cf. [5]. In this note, all graphs are finite, undirected graphs without isolated nodes, multiple edges, or self-loops.

We formalize our graphs in the following way. The basic constituents of graphs are nodes (vertices). An *edge* is defined to be a set that consists of two

distinct nodes. A *graph* is defined to be a finite set of edges. The *nodes* of a graph are the elements of its edges. If G is a graph, we write nG for its set of nodes. The *ports* of graph G are the nodes that occur in precisely one edge of G (ports are called sites or terminals in [17]). We write pG to denote the set of ports of graph G. A node of G is called *internal* if it is not a port. We write iG to denote the set of internal nodes of G. So we have $iG = nG \setminus pG$.

The *degree* of a node v of graph G is the number of edges of G that contain v. Therefore, a node of G is a port if and only if its degree is 1. In view of the application to conjugation in carbon chemistry, we could restrict attention to graphs where all nodes have degrees ≤ 4. Since we have no use for this restriction, we do this only in some examples.

We often represent graphs by drawing the nodes as bullets and the edges as line segments between the nodes. The graph drawn here has two ports and three internal nodes, one of degree 2 and two of degree 3.



Two distinct nodes p and q are defined to be *connected* in graph G if there is a sequence of nodes $(p_0, \ldots, p_n)$ with $p_0 = p$ and $p_n = q$ and $(p_i, p_{i+1}) \in G$ for all $0 \leq i < n$. Graph G is called *connected* if it is nonempty and every pair of distinct nodes of G is connected in G. Graph G is called a *simple path* between nodes p and q if it is connected, p and q are its only ports, and all its other nodes have degree 2. Graph G is called a *cycle* if it is connected and all its nodes have degree 2.

Since a graph is just a set of edges, every subset of a graph G is itself a graph, and can therefore be called a *subgraph*. A subgraph C of G is called a *curve* in G if every node of C which is internal in G has degree 2 in C.

The empty subgraph is a curve. More interesting examples of curves are cycles, and simple paths between ports. Conversely, every connected component of a curve is a cycle or a simple path between ports.

2 Kekulé States

The concept of Kekulé states is motivated by their correspondence to π-conjugated systems in chemistry. Polycyclic polyunsaturated hydrocarbons are molecules in which carbon atoms are organized in rings, mainly hexagons, but possibly also pentagons and heptagons. The primary example would be naphthalene, consisting of two fused benzene rings.

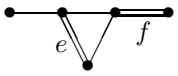
In 1865, August Kekulé proposed that a benzene molecule should consist of a hexagon of carbon atoms with attached hydrogen in an alternating cycle of double bonds and single bonds. In our graph theoretical abstraction, we neglect all hydrogen atoms. All carbon atoms are represented by nodes of the graph, all bonds are represented by edges. So, they are the hydrogen suppressed chemical graphs of [13].

We represent the set of double bonds by a subgraph of the graph of the molecule. In common polycyclic polyunsaturated hydrocarbons, the rule is that every carbon atom is sp^2 hybridized and has precisely one double bond. The ports represent other atoms that are linked to the graph by single or double bonds. This leads to the following definition.

We define a *Kekulé state* W in graph G to be a subgraph of G such that every internal node of G is a port of W.

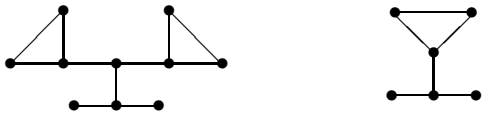
In graph theory, a subgraph W of G is called a *perfect matching* if every node of G is a port of W [5, p. 203]. It follows that W is a perfect matching if and only if W is a Kekulé state and W contains all edges to ports.

Example 1. The graph shown in section 1.4 has no perfect matchings, and has two Kekulé states, one of which ($\{e, f\}$) is shown below by doubling its edges e and f. If we give the bottom node an additional edge to a new port, the resulting graph has one perfect matching and four Kekulé states.



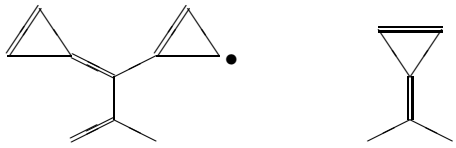
Efficient algorithms exist to determine whether a given graph has a perfect matching [10]. These algorithms can easily be adapted to yield Kekulé states. In the remainder of this paper, we have no need to consider perfect matchings.

Remark 1. There are graphs, even with several ports, that have no Kekulé states. The lefthand graph below is an example, as is easily verified. In this graph, the lower part with the two ports can be replaced by any other subgraph with any number of ports.



The righthand graph in the figure is a case with precisely one Kekulé state, even though the graph has two ports. In either case, the triangles can be replaced by pentagons to better meet the possibilities of carbon chemistry.

The chemical equivalents of the two structures are depicted below. The structure at the left is an open shell (non-Kekulé) radical structure with the bullet at a Kekulé violation. The structure at the right is a closed shell (Kekulé) structure.



2.1 Alternating curves

The electronic properties of Kekulé states are observed at and controlled by alternating paths, or more generally alternating curves.

An *alternating curve* is a pair (C, W) of subgraphs of G such that C is a curve in G and that the intersection $W \cap C$ is a Kekulé state of C. Since all internal nodes of C have degree 2, this means that every internal node of C belongs to one edge of W and one edge of $C \setminus W$, whence the name "alternating".

Roughly speaking, the next result says that the difference between any two Kekulé states is an alternating curve, and that, if at least one Kekulé state exists, every alternating curve is obtained in this way.

Theorem 1. *(a) Let W and W' be Kekulé states of a graph G. Then the subgraph $C = W \oplus W'$ is a curve in G and (C, W) is an alternating curve.*
(b) Let W be a Kekulé state of graph G and let (C, W) be an alternating curve in G. Then $W' = W \oplus C$ is a Kekulé state of G.

Proof. (a) Let v be a node of C which is an internal node of G. Since W and W' are Kekulé states of G, each contains precisely one edge, say $e \in W$ and $e' \in W'$, that contains v. Since v is a node of $C = W \oplus W'$, these edges are different. Therefore, both belong to C. It also follows that C has no other edges that contain v. Therefore v has degree 2 in C. This proves that C is a curve in G. Moreover, since we have $e \in W \setminus W'$ and $e' \in W' \setminus W$, the pair (C, W) is alternating.

(b) In order to show that W' is a Kekulé state, we need to show that every internal node v of G is a port of W'. Since W is a Kekulé state of G, there is a unique edge $e \in W$ with $v \in e$. If v is not a node of C, then e is also the unique edge of $W \oplus C$ that contains v. Assume therefore that v is a node of C. Since C is a curve, v is internal in C and has degree 2 in C. Since (C, W) is an alternating curve, it follows that there is a unique edge $e \in W \cap C$ and a unique edge $e' \in C \setminus W$ with $v \in e$ and $v \in e'$. It follows that e' is the unique edge of $W \oplus C = W'$ that contains v. Therefore v is a port of W'. □

2.2 Port assignments

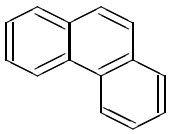
We now want to prescribe the port behaviour of Kekulé states. Let G be a graph with a set of ports P. We define a *port assignment* ("archetype" in [17]) to be a subset of P. For any subgraph W of G, we define $W|P$ as the set of nodes of W that also belong to P, i.e., $(W|P) = P \cap nW$. A port assignment g is called *Kekulé* if there is a Kekulé state W of G with $g = (W|P)$.

In chemistry, the different Kekulé states with the same port assignment are called resonance structures. The actual quantum mechanical state is a superposition of these resonance structures. In the Valence Bond Theory picture, the actual quantum mechanical state is taken as a weighted superposition of these (neutral) resonance structures. The number of neutral closed shell resonance structures is determined by the number of alternating curves in the following way:

Corollary 1. *Let G be a graph with $pG = P$. Let g be a port assignment in P. Let n be the number of Kekulé states W of G with $(W|P) = g$. For every Kekulé state W with $(W|P) = g$, there are precisely n curves C without ports such that (C, W) is alternating (including the empty curve).*

Proof. This follows from Theorem 1 by the observation that W and W' give the same port assignments if and only if the curve $W \oplus W'$ has no ports. □

Example 2. Phenantrene consists of three hexagons without any ports. One of its Kekulé states is depicted here.



This Kekulé state has five alternating curves. Four of them are easy to find: the empty curve, two cycles of length 6 around the extreme hexagons, and one cycle of length 10 around the two horizontal neighbouring hexagons. The remaining alternating curve is disconnected: it is the union of the two cycles of length 6. Phenantrene has therefore five Kekulé states. □

In chemistry, i.e., within Valence Bond Theory, the number of Kekulé states of a port assignment is an indication of its stability. Here the Kekulé structures having the highest number of aromatic rings (rings with $4n+2$ electrons; usually 6, as in benzene) are considered as the most important resonance contributors.

We now come to the central result of this paper that says that the existence of an alternating path between two given ports is completely determined by the set of Kekulé port assignments. It follows that the switching behavour of the graph is also completely determined by the set of Kekulé port assignments and hence independent of all other aspects of the graph.

In order to formulate this result, we note that, for distinct ports p and q, the doubleton set $\{p, q\}$ is a port assignment.

Theorem 2. *Let port assignment k satisfy $k = (W|P)$ for some Kekulé state W. Let p and q be distinct ports. There is a simple alternating path (C, W) from p to q in G if and only if the port assignment $k \oplus \{p, q\}$ is Kekulé.*

Proof. If there is a simple alternating path (C, W) from p to q in G, Theorem 1(b) implies that $W' = W \oplus C$ is a Kekulé state of G. It satisfies $(W'|P) = k \oplus \{p, q\}$.

Conversely, if the port assignment $k \oplus \{p, q\}$ is Kekulé, let W' be a Kekulé state with $(W'|P) = k \oplus \{p, q\}$ and let $C' = W \oplus W'$. Then (C', W) is an alternating curve by Theorem 1(a). The Kekulé states W and W' agree on all ports except for p and q. Therefore, the graph C' only contains the ports p and q. Let C be the connected component of p in C'. Then it also contains q and (C, W) is a simple alternating path from p to q in G. □

Theorem 2 implies that all Kekulé states corresponding to a given Kekulé port assignment agree on the question which ports are connected by alternating

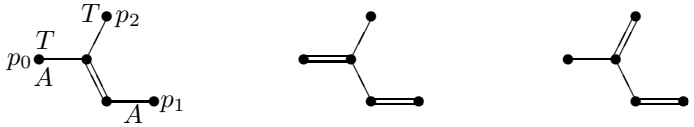
paths, and that the result is determined completely by the set of Kekulé port assignments.

The physical relevance of Theorem 2 is that an alternating path between a pair of ports is, electrically, an open channel (low resistance), and that the channel is closed in the absence of alternating paths [11,9,3]. This leads to the following formal definitions.

In view of Theorem 2, we define a *channel* to be a port assignment of the form $\{p, q\}$ for distinct ports p and q. A channel c is defined to be *open* with respect to a set K of port assignments and an element $k \in K$ if and only if $k \oplus c \in K$.

The idea is that K stands for the set of Kekulé port assignments of some graph. If a channel c is open for $k \in K$, one can send a so-called soliton through the channel with the effect that state k is replaced by $k \oplus c$ [17]. Indeed, the element $k \in K$ is regarded as a memory state. It is up to the physics to determine how long such memory states can be preserved.

Example 3. One of the simplest cases is an ethene molecule with three attached ports p_0, p_1, p_2. The initial port assignment is $\emptyset$. We use the channels $A = \{p_0, p_1\}$ and $T = \{p_0, p_2\}$. There are three Kekulé states, as depicted below. The corresponding port assignments are $\emptyset$, A and $A \oplus T = \{p_1, p_2\}$. Port assignment T is non-Kekulé.



Initially, channel A is open and T is closed. When A is signalled, we arrive at the middle state: this opens T, while A remains open. If A is signalled again, we return to the initial state where T is closed. We may use the system as a simple switch to open and close T by means of A. Note, however, that if A opens T, and T is then read, i.e., is signalled, we arrive at the righthand state where channel A is closed. Therefore, in order to preserve the system as a switch, reading this memory cell must be done by signalling the channel twice. □

2.3 Abstraction from the graphs

At this point we take the abstraction step to concentrate on the ports and the port assignments and treat the specific graph itself merely as auxiliary. From here on, P is an arbitrary finite set and the elements of P are called *ports*. The subsets of P are called *port assignments*. Just as before, a *channel* is a port assignment of the form $\{p, q\}$ for distinct ports p and q.

Recall from section 1.3 that the set $Pow(P)$ of the port assignments is a commutative group with operation $\oplus$. We define $Even(P)$ and $Odd(P)$ to be the subsets of $Pow(P)$ that consist of the port asignments with an even (odd) number of ports. It is easy to verify that $Even(P)$ is a subgroup of $Pow(P)$ and that $Odd(P)$ is not.

For $\#P = 4$, the sets $Even(P)$ and $Odd(P)$ are depicted in Figure 5.5 of [17]. Table 5.10 of [17] gives the table of the operation $\oplus$ in $Even(P)$. The group $Even(P)$ *acts* on $Odd(P)$ in the sense that $g \oplus h \in Odd(P)$ for all $g \in Even(P)$ and $h \in Odd(P)$. All channels are elements of $Even(P)$.

For any two port assignments k, $k' \in Pow(P)$, we define the *Hamming distance* $dist(k, k')$ as the number of ports in $k \oplus k'$. It is easy to verify that $dist(k, k')$ is always ≥ 0, and that it is zero if and only if $k = k'$. Recall that, in any group, *translation* is the operation that adds a fixed element to its argument. The Hamming distance is translation invariant: $dist(g \oplus k, g \oplus k') = dist(k, k')$ for all g, k, $k' \in Pow(P)$.

2.4 Cells and Kekulé cells

A *cell* over P is defined to be a subset of $Pow(P)$. A channel c is defined to be *open* with respect to a cell K and a state $k \in K$, if and only if $k \oplus c \in K$. In accordance with Theorem 2, the interpretation is that a signal can be sent over channel c if and only if the channel is open and that, in that case, the state is transformed into $k \oplus c$.

For any graph G, we let $KP(G)$ denote the set of the Kekulé port assignments of G. Clearly $KP(G)$ is a cell over P, i.e., a subset of $Pow(P)$. An arbitrary cell K over P is called a *Kekulé cell* if and only if there is a graph G with $P = pG$ and $K = KP(G)$.

At this point the question is: how arbitrary are Kekulé cells? Our first result in this direction is that Kekulé cells are preserved by translation. If K is a cell and g is a port assignment, the *translated cell* $g \oplus K$ is defined by $g \oplus K = \{g \oplus k \mid k \in K\}$.

Theorem 3. *Let K be a Kekulé cell over P. Let $g \in Pow(P)$. Then $g \oplus K$ is also a Kekulé cell over P.*

Proof. Choose a graph G with $P = pG$ and $K = KP(G)$. It suffices to construct a graph G' with $g \oplus K = KP(G')$. We construct G' from G. For every port $p \in g$, we replace the (unique) edge of G that contains p by two edges linked by a new internal node. More precisely, let each edge $e = \{p, v\}$ with $p \in g$ be replaced by the edges $\{p, v'\}$ and $\{v', v\}$. If W is a Kekulé state of G, let W' be the subgraph of G' that consists of $W \cap G'$ together with all edges $\{v', v\}$ with $p \in g$ and $\{p, v\} \in W$, and all edges $\{p, v'\}$ with $p \in g$ and $\{p, v\} \in G \setminus W$. Then W' is a Kekulé state of G' with $W'|P = g \oplus (W|P)$. Moreover, every Kekulé state W' of G' is obtained in this way. This shows that $g \oplus K = KP(G')$. □

We now prove a result that shows that every Kekulé cell is "connected by channels" in a certain, rather strong, sense. For any set D of port assignments we use the notation $\bigoplus D$ to denote the $\oplus$ sum of the elements of D.

Theorem 4. *Let K be a Kekulé cell over P. For every pair of elements g, $g' \in K$, the Hamming distance $dist(g, g')$ is even and there is a set D of disjoint channels such that $g' = g \oplus (\bigoplus D)$ and $dist(g, g') = 2 \cdot \#D$ and, in addition, $g \oplus (\bigoplus D') \in K$ for every subset D' of D.*

Proof. Since K is a Kekulé cell, we can choose a graph G with $P = pG$ and $K = KP(G)$. Let g, $g' \in K$. We can take Kekulé states W and W' of G so that $g = (W|P)$ and $g' = (W'|P)$. Let (C, W) be the alternating curve of Theorem 1(a) with $C = W \oplus W'$. Curve C is the disjoint union of simple paths between ports, and cycles. Assume that $C_0, \ldots, C_{d-1}$ are the components of C that are the simple paths. Then (C_i, W) is an alternating path between ports for every $i < d$.

For each index i, let c_i be the channel that consists of the endpoints of path C_i. We take $D = \{c_i \mid i < d\}$. Since the paths C_i are components of C, they are disjoint. Therefore, the channels in D are disjoint. It is easy to see that $g' = g \oplus (\bigoplus D)$ and $dist(g, g') = 2 \cdot \#D$. Repeated application of Theorem 1(b) yields that $g \oplus (\bigoplus D') \in KP(G) = K$ for every subset D' of D. □

This result implies that, in a Kekulé cell, every two elements differ by a set of disjoint channels in such a way that all intermediate combinations (of one element with some of the channels) also belong to the cell.

A port $p \in P$ is called *flexible for a cell* K if there are g, $g' \in K$ with $p \in g \setminus g'$. Nonflexible ports do not participate in any switching behaviour. We are therefore only interested in the flexible ports. Now, Theorem 4 also implies the following result.

Corollary 2. *Let K be a Kekulé cell in $Pow(P)$.*
(a) Let p be a flexible port for K. For any $g \in K$, there is a channel c with $p \in c$ and $g \oplus c \in K$.
(b) When K has a flexible port, it has to have at least two flexible ports.

Since nonflexible ports are useless for the application we have in mind, we define a cell to be *flexible* if all ports are flexible. We eliminate the nonflexible ports in the following way. Let Q be the set of flexible ports of K. Then the *flexible cell* $flex(K) \subseteq Pow(Q)$ is defined by $flex(K) = \{g \cap Q \mid g \in K\}$.

It is easy to see that, indeed, cell $flex(K)$ is flexible, i.e., all ports $p \in Q$ are flexible for $flex(K)$. The restriction function $g \mapsto g \cap Q$ is a bijective correspondence between K and $flex(K)$.

Theorem 5. *The cell $flex(K)$ is Kekulé if and only if the cell K is Kekulé.*

Proof. First assume that cell K is Kekulé. We can therefore choose a graph G with $P = pG$ and $K = KP(G)$. We define an edge of graph G to be flexible if there are Kekulé states W and W' of G with $e \in W \setminus W'$. Let G' be the subgraph of the flexible edges of G.

We will show that $flex(K)$ is a Kekulé cell by demonstrating that $flex(K) = KP(G')$. We first need to demonstrate that the ports of G' are the flexible ports of K. If node v of G is a port of G', it is incident with precisely one edge $e \in G'$; then there are Kekulé states W and W' of G with $e \in W \setminus W'$; if node v is internal in G, it follows that there is also an edge e' that contains v such that $e' \in W' \setminus W$, and hence $e' \in G'$, so that v is not a port of G', a contradiction. This shows that v is a flexible port of $KP(G) = K$. Conversely, every flexible port of $KP(G)$ clearly is a port of G'.

We now claim that, for every Kekulé state W of G, the intersection $W \cap G'$ is a Kekulé state of G'. Let v be an internal node of G'. Then it is an internal node of G and hence contained in a unique edge $e \in W$. If $e \notin G'$, then $e \in W'$ for all Kekulé states W' of G. Since v is contained in a flexible edge of G, this results in a contradiction. This proves that $e \in G'$ and, therefore, that $W \cap G'$ is a Kekulé state of G'. This implies that $\mathit{flex}(K) \subseteq KP(G')$.

Conversely, every Kekulé state W_0 of G' is the restriction of a Kekulé state W of G. In fact, choose an arbitrary Kekulé state W_1 of G. Define W on G by $W = W_0 \cup (W_1 \setminus G')$. Since, as shown in the first paragraph, $e \notin W_0$ for every edge $e \notin G'$ that is incident with a node v of G', the subgraph W is a Kekulé state of G. This implies that $\mathit{flex}(K) \supseteq KP(G')$ and, hence $\mathit{flex}(K) = KP(G')$. Therefore, $\mathit{flex}(K)$ is a Kekulé cell.

Conversely, assume that $\mathit{flex}(K)$ is a Kekulé cell, say $\mathit{flex}(K) = KP(G')$ for some graph G'. We construct a graph G with $K = KP(G)$ by extending G' with handles for the nonflexible ports. The nonflexible ports fall in two classes. Let P_1 consist of the ports $p \in P$ with $p \in g$ for all $g \in K$ and let P_0 be the remainder $P \setminus (Q \cup P_1)$.

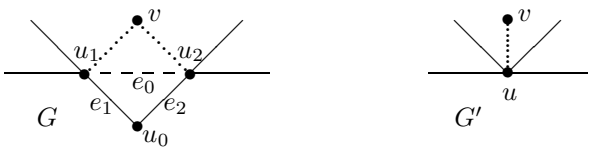
In the diagram, we depict the handles to be attached to the nonflexible ports $p_1 \in P_1$ and $p_0 \in P_0$. The verification that $K = KP(G)$ is straightforward. □

2.5 Transforming the graph while preserving its Kekulé cell

The next result allows us to simplify graphs while retaining the Kekulé cell.

Theorem 6. *Let graph G have a node u_0 with precisely two neighbour nodes u_1 and u_2. Assume that u_1 and u_2 are both internal. Let graph G' be obtained from G by removing u_0 and its two incident edges and merging the nodes u_1 and u_2 into a new node u. If there is an edge between u_1 and u_2, it is removed since it would become a self-loop. For every node $v \neq u_0$ that is a common neighbour of u_1 and u_2, the two edges linking v to u_1 and u_2 are identified, since multiple edges are not allowed.*

Consequently, the degree of u is $d_1 + d_2 - 2 - 2a - b$, where d_1 and d_2 are the degrees of u_1 and u_2, respectively, and $a \in \{0, 1\}$ is the number of edges between u_1 and u_2, and $b \in \mathbb{N}$ is the number of common neighbours of u_1 and u_2 different from u_0. In all cases, $KP(G') = KP(G)$.



Proof. Let e_1 and e_2 be the edges linking u_0 with u_1 and u_2, respectively. If it exists, let e_0 be the edge between u_1 and u_2. Let E_1 be the set of the other edges incident with u_1 and let E_2 be the set of the other edges incident with u_2. Every Kekulé state W of G has $e_1 \in W$ if and only if $e_2 \notin W$. It follows that W contains precisely one of the edges in $E_1 \cup E_2$ and that $e_0 \notin W$ if e_0 exists. We can therefore transform W into a Kekulé state of G' with the same values on all remaining edges. The details about neighbour v are left to the reader. It is easy to see that all Kekulé states of G' are obtained in this way. Because the transformation does not change the ports, the Kekulé cells of G and G' are equal. □

This result can also be used in the other direction, i.e., starting with graph G'. In this way, it can serve to split a node with a high degree into two nodes with a lower degree. By first merging and then splitting, it can also be used to shift edges between nodes u_1 and u_2 in a similar way as with operation **Op.vii** of [17, section 3.3.1].

2.6 Classification of small Kekulé cells

We define the *Hamming diameter* of a cell K as the maximum Hamming distance between its elements: $\mathit{diam}(K) = \max\{\mathit{dist}(k, k') \mid k, k' \in K\}$. We clearly have $\mathit{diam}(K) \leq \#P$. Cell K has flexible ports if and only if $\mathit{diam}(K) > 0$. By Theorem 4, the diameter of a Kekulé cell is always even. We use these observations here to determine all flexible Kekulé cells with at most 4 ports.

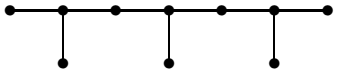
For a fixed set P of ports, we define the cell $K_1 = \{\{p\} \mid p \in P\}$. If P has more than one element, K_1 has diameter 2.

Lemma 1. *Let K be a flexible cell in $\mathrm{Pow}(P)$ with diameter 2. Assume that $\#P \neq 3$. Then K is a Kekulé cell if and only if there is $g \in \mathrm{Pow}(P)$ such that $K = g \oplus K_1$.*

Proof. We first show that K_1 is a Kekulé cell. For this purpose, we take the graph G with one internal node connected to all ports (a so-called star). Every Kekulé state W of G contains precisely one edge. The corresponding port assignment contains precisely one port. Since Kekulé cells can be translated, this proves the *if* part of the assertion.

Conversely, let K be a Kekulé cell with diameter 2. By translation, we may assume that $\emptyset \in K$. By Corollary 2(a) with $g = \emptyset$, for every $p \in P$, there is some $q \neq p$ with $\{p, q\} \in K$. Using $\#P \neq 3$ and $\mathit{diam}(K) = 2$, one can show that there is one port p_0 such that K consists of the channels $\{p_0, q\}$ for all ports $q \neq p_0$. This implies that $K = \{p_0\} \oplus K_1$. □

Remark 2. A star graph with > 4 ports is not commonly found in carbon chemistry, but by (repeated) application of Theorem 6, one can easily find alternatives. An example where one can use a tree graph for the five port situation is given below. □



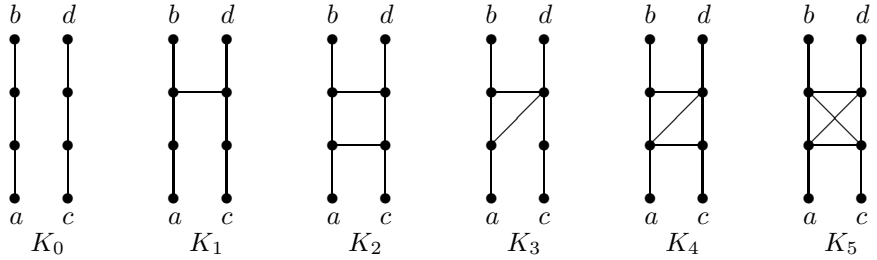
Remark 3. In the case of $\#P = 3$, all flexible Kekulé cells have diameter 2. In this case, there are two classes of flexible Kekulé cells: the translates of K_1 of Lemma 1 and the translates of $Even(P)$. □

Lemma 2. *Let P have 4 ports. Let K be a cell in $Pow(P)$. Then K is a Kekulé cell with diameter 4 if and only if there is $g \in Pow(P)$ and an enumeration a, b, c, d of the ports such that*

$$\{\emptyset, \{a,b\}, \{c,d\}, \{a,b,c,d\}\} \subseteq g \oplus K \subseteq Even(P)\ .$$

Proof. First, assume that K is a Kekulé cell with diameter 4. By translation, say over g, we may assume that $\emptyset \in K$ and $dist(\mathbf{0}, k) = 4$ for some $k \in K$. By Theorem 4, there are two disjoint channels $\{a,b\}$ and $\{c,d\}$ such that K contains the port assignments $\emptyset$, $\{a,b\}$, $\{c,d\}$, and $\{a,b,c,d\}$. Also, by Theorem 4, all elements of K have even distance to $\emptyset$. Therefore, K is contained in $Even(P)$.

For the converse implication (only if), we may assume that K satisfies $K_0 \subseteq K \subseteq Even(P)$ where $K_0 = \{\emptyset, \{a,b\}, \{c,d\}, \{a,b,c,d\}\}$. The only elements for which it is not known whether they are contained in K, are $\{a,c\}$, $\{a,d\}$, $\{b,c\}$, $\{b,d\}$. This leaves $2^4 = 16$ possibilities for K. Using symmetry, we can reduce it to 6 cases, viz. $K_1 = K_0 \cup \{\{a,c\}\}$, $K_2 = K_1 \cup \{\{b,d\}\}$, $K_3 = K_1 \cup \{\{b,c\}\}$, $K_4 = K_2 \cup K_3$, and $K_5 = Even(P)$. It remains to show that the cells K_0 up to K_5 indeed occur as the Kekulé cells of certain graphs.



It is easy to verify that the six graphs depicted above represent the Kekulé cells K_0 up to K_5, as required. □

Remark 4. This proof shows that the Kekulé cells with 4 ports and diameter 4 can all be realized by subgraphs of the graph which is called Δ_4 in section 4. It seems likely that this result can be generalized in some way, but at this point we do not see how.

3 Semi-Kekulé States

This section serves as an aside to provide a context in which Kekulé states are special cases. We weaken the condition for Kekulé states to a linear condition

over the field with two elements. This simplifies the mathematics, since the whole theory of linear equations becomes available.

For an integer x, we write x **mod** 2 to denote the remainder of x upon integer division by 2. It follows that x **mod** $2 = 0$ (or 1) if and only if x is even (or odd, respectively). For integers x and y, we define $x \oplus y = (x + y)$ **mod** 2.

A node in a graph is called *even/odd* if its degree in the graph is even/odd. A subgraph W of G is called a *semi-Kekulé* state of G if every internal node of G is an odd node of W. Clearly, every Kekulé state is a semi-Kekulé state, but not *vice versa*.

For any graph G, we define $odd(G)$ to be the set of odd nodes of G. Therefore, W is a semi-Kekulé state of G if and only if $iG \subseteq odd(W)$. Since every edge contributes twice to the degree of a node, the sum of the degrees of the nodes is twice the number of edges. It is therefore even. Since the sum of the degrees at the even nodes is also even, it follows that the sum of the degrees of the odd nodes is even. Since every odd node contributes an odd number, it follows that the number of odd nodes is even: $\#odd(G)$ **mod** $2 = 0$ for every graph G.

3.1 There are enough semi-Kekulé states

Let the *signature* of a graph G be defined as the number $sG = (\#iG)$ **mod** 2. In other words, the signature is 0 or 1, depending on whether the number of internal nodes is even or odd. To formulate the next result conveniently, we define $Pow(P)_0 = Even(P)$ and $Pow(P)_1 = Odd(P)$, see section 2.3.

Theorem 7. *Let G be a graph with port set $P = pG$ and signature $\varepsilon = sG$.*
(a) Let W be a semi-Kekulé state of G. Then $(W|P) \in Pow(P)_\varepsilon$.
(b) Conversely, let $g \in Pow(P)_\varepsilon$. Assume that graph G is connected. Then G has a semi-Kekulé state W with $(W|P) = g$.

Proof. For any subgraph W of G, we define the set of violating nodes by $vio(W) = iG \setminus odd(W)$. Clearly, W is semi-Kekulé if and only if $vio(W)$ is empty.

On the other hand, we have $(W|P) = P \cap nW = odd(W) \setminus iG$, since ports are nodes that belong to precisely one edge. It follows that

$$\begin{aligned} &\#vio(W) + \#(W|P) = \#(iG \oplus odd(W)) \\ &= \#iG + \#odd(W) - 2 \cdot \#(iG \cap odd(W)) \ . \end{aligned}$$

Since $\#odd(W)$ is even and $\#iG$ **mod** $2 = \varepsilon$, this implies that

$$(*) \qquad (\#vio(W) \textbf{ mod } 2) \oplus (\#(W|P) \textbf{ mod } 2) = \varepsilon \ .$$

If W is semi-Kekulé, $vio(W)$ is empty, so that $\#(W|P)$ **mod** $2 = \varepsilon$, i.e., $(W|P) \in Pow(P)_\varepsilon$. This proves part (a).

Conversely, assume that $\#g$ **mod** $2 = \varepsilon$. We can define some subgraph W of G with $(W|P) = g$. Then the righthand summand of (*) equals ε. It follows that the number of elements of $vio(W)$ is even. If it is positive, then we can choose two different nodes u and $v \in vio(W)$. Since V is connected, we can choose a simple path C from u to v in G, and replace W by $W' = W \oplus C$, which satisfies $vio(W') = vio(W) \setminus \{u, v\}$ and $W'|P = g$. We can continue in this way until $vio(W)$ is empty. □

For a port assignment g in a connected graph, Theorem 7 implies that the condition $g \in Pow(P)_{\mathcal{E}}$ is necessary and sufficient for the existence of a semi-Kekulé state W with $(W|P) = g$. This also implies that this is a necessary condition for the existence of a Kekulé state W with $(W|P) = g$. In section 4, we will describe graphs in which this condition is not only necessary but also sufficient.

3.2 The homogeneous semi-Kekulé kernel

Let $\mathbb{B}$ (pronounced *Bit*) be the set of the two elements 0 and 1, with the addition $\oplus$ described above and an ordinary multiplication. This is a field, i.e., an algebraic structure with the usual laws for addition, subtraction, multiplication, and division by nonzero elements. It follows that the whole theory of vector spaces and linear equations is applicable over $\mathbb{B}$.

In particular, for a set P, the additive group $Pow(P)$ is a vector space over $\mathbb{B}$ of dimension $\#P$. The easiest way to see this is to identify a subset $g \subseteq P$ with its characteristic function χ_g given by $\chi_g(p) = 1$ if $p \in g$ and $\chi_g(p) = 0$ if $p \notin g$. In this way $Pow(P)$ is identified with the set of functions $P \to \mathbb{B}$. The symmetric difference operator $\oplus$ corresponds to the operator on functions given by $(f \oplus g)(p) = f(p) \oplus g(p)$.

The vector space $Pow(G)$ is introduced in the same way. We define the *homogeneous semi-Kekulé kernel*, $HSK(G)$, of G to be the set of subgraphs W for which all nodes are even:

$$HSK(G) = \{W \in Pow(G) \mid odd(W) = \emptyset\} \ .$$

It is easy to see that HSK is a linear subspace of $Pow(G)$.

If an inhomogeneous system of linear equations has a solution W_0, the set of all its solutions is of the form $W_0 \oplus H$, where H is the set of solutions of the homogeneous linear system. It follows that, given one semi-Kekulé state W_0 of G, the set of semi-Kekulé states with the same port assignment as W_0 is equal to $W_0 \oplus HSK(G) = \{W_0 \oplus W \mid W \in HSK(G)\}$.

A graph is called a *tree* if it is connected and has no connected proper subgraph with the same set of nodes. Every connected graph has a spanning tree, i.e., a subgraph with the same nodes which is a tree (just remove edges one by one until you would cause disconnectedness).

Theorem 8. *Let graph G be connected. The vector space $HSK(G)$ over $\mathbb{B}$ has dimension $\#G + 1 - \#nG$. It has a basis over $\mathbb{B}$ that consists of cycles of G.*

Proof. First, assume that G is a tree. Then the number of nodes is one plus the number of edges, as is well-known in graph theory. So we have $\#nG = \#G + 1$. On the other hand, since G is a tree, one can prove by induction on the size of the tree that $HSK(G) = \{\emptyset\}$. This implies the assertion for the case that G is a tree.

We now proceed by induction on the number of edges. Let graph G be given. We may assume that G is connected and not a tree. Therefore G contains an edge

e, such that removal of e yields a connected graph G'. By induction, the space $HSK(G')$ of the subgraphs W of G' with $odd(W)$ empty has dimension $\#G - \#nG$ and a basis B that consists of cycles of G'. Since graph G' is connected, edge e is contained in a cycle C of graph G. Now C is an element of $HSK(G)\backslash HSK(G')$. Therefore $HSK(G')$ is a proper subset of $HSK(G)$. Since $HSK(G')$ is the subset of $HSK(G)$ given by the single equation that corresponds to the condition $e \notin W$, it follows that $dim(HSK(G)) = 1 + dim(HSK(G')) = \#G + 1 - \#nG$ and that $B \cup \{C\}$ is a basis of $HSK(G)$. □

This result was inspired by the observation that, for each of the graphs considered in Figures 5.7 and 5.8 of [17], the set $HSK(G)$ has 8 elements.

Corollary 3. *Let G be connected. Let $C_0, \ldots, C_{r-1}$ be cycles in G such that $r = \#G + 1 - \#nG$ and that $C_0, \ldots, C_{r-1}$ are linearly independent over $\mathbb{B}$.*
(a) Then $C_0, \ldots, C_{r-1}$ form a basis of $HSK(G)$.
(b) Let $P = pG$ and $\varepsilon = sG$. For any $g \in Pow(P)_\varepsilon$, there are precisely 2^r semi-Kekulé states W with $(W|P) = g$.

From part (b) it follows that the number of Kekulé states with a given port assignment is $\leq 2^r$.

Usually, it is very easy to find enough linearly independent cycles to obtain a basis for HSK with Corollary 3, and to find simple paths between different ports. Since the constructive proof of Theorem 7 is not very efficient, we prefer to apply Gauss elimination to the inhomogeneous system to find semi-Kekulé states.

Complexity. The problem to determine some Kekulé state for a given graph with given port assignment, has a complicated but efficient solution with complexity $\mathcal{O}(\#E \cdot \sqrt{\#V})$ [10]. All semi-Kekulé states can easily be found in time $\mathcal{O}(\#E \cdot \#V)$. The above approach leads to an algorithm to determine all Kekulé states in time $\mathcal{O}(\#E \cdot (\#V + 2^r))$ where r is as in Corollary 3. In the applications, this is quite feasible, since r tends to be rather small.

4 Omniconjugated Graphs

Van der Veen et al. [18] have defined n-port ('n-terminal') π-conjugated molecules, with $n \geq 3$, as omniconjugated of type A, if direct linear conjugated pathways are present between all pairs of ports in all states of the ports. We now formalize the principles of such systems in a mathematical way.

We define a graph G with signature ε to be *omniconjugated* if $P = pG$ has at least two elements and the Kekulé cell of G is $Pow(P)_\varepsilon$ (i.e., maximal). Note that our concept omniconjugated represents what is called "omniconjugated of Type A" in [18,17].

Corollary 4. *Let G be omniconjugated. Let W be a Kekulé state of G. Let p and q be two different ports of G. Then G has an alternating curve (C, W), such that C is a simple path from p to q.*

Proof. Since $(W|P) \oplus \{p, q\} \in Pow(P)_\varepsilon$, and the graph is omniconjugated, the assertion follows from Theorem 2. □

The corollary means that all channels of an omniconjugated graph are always open (low resistance). Conversely, if graph G has at least one Kekulé state and all channels are always open, it is omniconjugated.

For every set P with at least two elements, omniconjugated graphs exist. This is shown as follows. Let A_2 be the graph that consists of one edge with its two nodes. The two nodes are ports. In this case, $\varepsilon = 0$ and $Pow(P)_\varepsilon$ contains the empty subgraph and the full subgraph which are both Kekulé states. Therefore A_2 is omniconjugated. More generally, the linear graph A_n of $n \geq 2$ consecutive nodes and $n - 1$ edges is also omniconjugated (the name A_n comes from the classification of Dynkin diagrams).

The smallest omniconjugated graph with ≥ 3 ports is the graph Δ_3 with 3 ports, and three internal nodes of degree 3 that form a triangle. More generally, let G_n for $n \geq 2$ be the complete graph with n nodes and $\binom{n}{2}$ edges. Let Δ_n be the graph obtained from G_n by attaching a port via an edge to every node of G_n. So, Δ_n has $2n$ nodes and $\binom{n}{2} + n$ edges. Note that $\Delta_2 = A_4$ and that Δ_4 is the graph used for K_5 in the proof of Lemma 2. We leave it to the reader to verify that all graphs Δ_n with $n \geq 2$ are omniconjugated. Conversely, we have:

Lemma 3. *Let G be a graph with $n \geq 2$ nodes. Let graph G' be obtained from G by attaching a port via an edge to every node of G. Assume that G' is omniconjugated. Then G is a complete graph.*

Proof. Since G has no isolated nodes, all nodes of G are internal nodes of G', and G' has n ports. Therefore, the signature ε of G' is n **mod** 2. Let P be the set of ports of G'.

It suffices to prove that every pair of different nodes of G is linked in G. Let $u \neq v$ be nodes of G. Let p and q be the unique ports of G' linked to u and v, respectively. Let g be the port assignment of G' given by $g = P \setminus \{p, q\}$. We have $g \in Pow(P)_\varepsilon$ since $(n - 2)$ **mod** $2 = \varepsilon$.

Since it is omniconjugated, graph G' has a Kekulé state W with $(W|P) = g$. Every edge e that contains a port $r \in g$ is in W. Since W is a Kekulé state of G', it follows that $e \notin W$ for every edge $e \in G$ that contains some node of G other than u and v. The two edges e of G' incident with p and q are not in W. Since W is a Kekulé state of G', it follows that u is element of a unique edge e of W. It follows that v is the other endpoint of e. This proves that u and v are linked in G. □

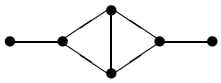
Section 3.3.1 of [17] presents a topological design program for omniconjugated models that have properties similar to those of our omniconjugated graphs. We transfer this idea to omniconjugated graphs in the following way.

Starting with an omniconjugated graph G, the question is: what kinds of operations on G are allowed that preserve its property of being omnicojugated? It turns out that operations like **Op.vi** and **Op.vii** of [17] are allowed because of Theorem 6. It is also easy to see that the operations **Op.ii**, **Op.v**, and **Op.iv** are preserved in the following sense:

Lemma 4. *Let G be an omniconjugated graph. Let G' be obtained from G by applying the following two types of operations:*
Op.ii. *Add arbitrary edges between internal nodes.*
Op.v. *Replace the edge to a port by two edges and a node of degree 2.*
Then G' is omniconjugated.

Lemma 5. ***Op.iv.*** *Let G' be a graph with port p' and let G'' be a graph with port p''. Let graph G be obtained from G' and G'' by removing the two ports and identifying their incident edges. Then G is omniconjugated if and only if G' and G'' are omniconjugated.*

For example, as indicated in [17, Scheme 3.1], we can apply **Op.ii** twice to the basic model A_6 to obtain the basic model **B**, which is therefore also omniconjugated.

B 

Starting with graph Δ_3, one can apply **Op.iv** to obtain omniconjugated graphs with arbitrary many ports. For this purpose, we therefore do not need the graphs Δ_n with $n > 3$.

In Theorem 6, we described a transformation from a graph G into a graph G' that has the same Kekulé cell. It follows that G' is omniconjugated if and only if G is omniconjugated. We conjecture that every omniconjugated graph in which all nodes have degree ≤ 3 can be obtained by starting with the basic models A_2 and Δ_3 and applying Theorem 6 and the Lemmas 4 and 5.

5 Functional Cells and Controlled Switches

In this section, we build upon the ideas of Chapter 6 of [17]. The idea of this chapter is that an alternating path between a pair of ports forms an open channel, and that the absence of any alternating path means closure of the channel.

A cell $K \subseteq Pow(P)$ becomes a *functional cell* by specifying an initial element $k_0 \in K$ and a linearly independent system of channels to probe and modify the current element of the cell. See the example given at the end of section 2.2.

5.1 A cell for the conjunction

Section 6.3 of [17] contains 15 of the 16 binary boolean operators implemented with pyracylene-based structures bearing six ports in various patterns. The only missing operator is the conjunction. Since $\neg A \wedge B$ is available in structure **6.8** (p. 140) of [17], the conjunction can be realized in pyracylene by taking a nonzero initial state.

We now sketch an approach that yields a simpler graph. We first design the functional cell. Let $k_0 = \emptyset$ be the initial element. In the style of [17], we choose two input channels A and B and a test channel T. Since we want to model the

conjunction of the inputs, test channel T must be open if and only if both input channels are signalled. Therefore, the cell K must contain the elements $\emptyset$, A, B, $A \oplus B$, and $A \oplus B \oplus T$. Since channel T must be closed in the states $\emptyset$, A, B, cell K must not contain the elements T, $A \oplus T$, and $B \oplus T$.

Since all eight port assignments mentioned should be different, the smallest candidate is a cell with 4 ports and 5 elements. The solution happens to occur as cell K_1 in the classification in the proof of Lemma 2 when we take $A = \{a, b\}$, $B = \{c, d\}$, and $T = \{b, d\}$.

A disadvantage of this solution is that both ports of test channel $T = \{b, d\}$ are used in the input channels. It may be difficult to independently signal A and B and probe T.

5.2 Reading output without disabling the switch

The example discussed above has an even more fundamental problem, which it shares with the example in section 2.2 and the results of [17]: if the test whether the output channel T is open is performed by signalling the channel, the cell moves to a state where it no longer satisfies its specification.

The point is discussed by [17] on page 149. She distinguishes signalling a channel by sending a soliton through it and reading a channel by means of a polaron. A polaron can be thought of as a double soliton [14]. In order to use the output of one cell as an input of another cell, we would like to read the output by means of solitons. This, however, requires that signalling the output channel leaves the cell in a state equivalent to the state before the output was read.

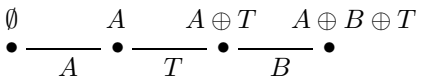
In the start state, and in all states reachable before the output is read, all inputs should be possible. In some but not all of these states, the output channel T should be open. This implies that signalling T when it is open, can bring the cell in a state where some of the input channels are not open. We therefore have to relax the requirement that input channels are always open, but since input must always be possible, we have to share inputs over more than one channel.

We therefore define an input *socket* to consist of a pair (A, B) of parallel channels of the form $A = \{p, r\}$, $B = \{q, r\}$ for ports p, q, and r, such that always either A is open and B is closed, or *vice versa*. When a soliton is sent to a socket, it will travel through an open channel of the two.

Remark 5. Note that we require that the two parallel channels of the socket share a port. In fact, if the parallel channels A and B are disjoint, a soliton sent through the pair can enter via a port of A and exit via a port of B, so that an unintended channel is signalled. In order to reckon with this possibility, one would have to impose additional requirements. □

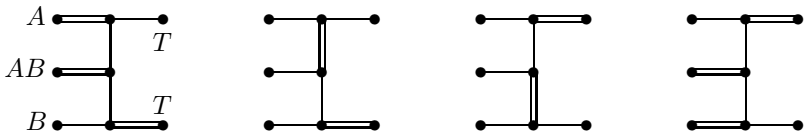
Example 4. The simplest case, called the *Y-cell*, has one input socket and one output channel. Let the socket consist of the channels A and B, and let T be the output channel. The vectors A, B, T are supposed to be linearly independent. We postulate that $K = k_0 \oplus L$ where $L = \{\emptyset, A, A \oplus T, A \oplus B \oplus T\}$.

The state space K is a translation of L. Either space consists of 4 states with three possible transitions. In the diagram, a transition indicated with channel c corresponds to the operation $x \mapsto x \oplus c$.



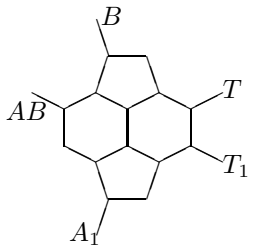
In all states, either channel A or channel B is available, and toggling this channel opens or closes the output channel T.

The simplest graph with such a Kekulé cell is a tree with 5 ports and 3 internal nodes. In this case, the vector space $Pow(P)_\varepsilon$ over $\mathbb{B}$ has dimension 4 (16 elements). The Kekulé cell has 8 elements. We use the initial state and the input socket (A, B) and the output channel T as depicted in the leftmost graph.



In the initial state depicted on the left, only channel A is open. If channel A is signalled, we get in state middle-left where the channels A and T are open. If T is signalled there, we get in state middle-right where the channels B and T are open. If B is signalled there, we get in the rightmost state where channel B is the only open channel.

This state of affairs can also be realized in a pyracylene derivative as shown below. The ports A_1 and T_1 start with a double bond in the initial state k_0.



Pyracylene with socket (A, B) and output channel T

In this case, $Pow(P)_\varepsilon$ has dimension 4. The Kekulé cell has 12 elements. From the start state k_0 only four of them are reachable by sending solitons over the channels A, B, T, but, e.g., one reaches the origin $\emptyset$ by sending a soliton from A_1 to T_1. The origin has four different Kekulé states, whereas k_0 and $k_0 \oplus A \oplus B \oplus T$ have only two different Kekulé states. This suggests that the origin is physically much more stable than the physical states corresponding to k_0 and $k_0 \oplus A \oplus B \oplus T$. The port assignments $k_0 \oplus A$ and $k_0 \oplus A \oplus T$ have only one Kekulé state. It is likely that such Kekulé states represent less stable physical states. □

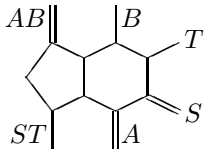
Remark 6. Is there a graph G with four ports p, q, r, t, such that the above structure can be realized by taking $A = \{p, r\}$, $B = \{q, r\}$, $T = \{t, r\}$? The

answer is no. This follows from Theorem 4 by taking $g = \emptyset$ and $g' = A \oplus B \oplus T$. □

5.3 Splitting for feedback

One may decide that, when an output channel T is read and toggled, a signal is sent back through T to reset the cell to its unread state. Since the output should also be sent forward to be used, we need a splitter: a cell with one input channel or socket and two output channels that are both opened when an input signal arrives.

We use an input socket (A, B) and output channels S and T. Let the origin $\emptyset$ be the initial state. Then the cell must contain $\mathbf{0}$, A, $A \oplus S$, $A \oplus S \oplus T$, $A \oplus S \oplus T \oplus B$. Since the channels S and T must be closed in states $\emptyset$ and $A \oplus S \oplus T \oplus B$, the cell must not contain S, T, $A \oplus S \oplus B$, and $A \oplus T \oplus B$. Since (A, B) is a socket, the cell must not contain B, $A \oplus B$, $S \oplus T$, and $S \oplus T \oplus B$. The cell may or may not contain the elements $A \oplus T$, $B \oplus S$, and $B \oplus T$.



Indene with socket (A, B)
and output channels S and T

The indene molecule has a graph with a Kekulé cell with these properties, as depicted here. The 9 interior nodes form a pentagon and a hexagon. In this case $Pow(P)_{\varepsilon}$ has dimension 5, and 32 elements. The Kekulé cell has 18 elements. We use the input socket (A, B) and the output channels S, T as depicted. Incidentally, S and T also form a socket. Initially, there are double bonds at both ports of channel A and at the eastern port of S. Indene has several solutions of this type, but all of these have initially three double bonds at ports, have output channels combined in a socket, and have the elements $A \oplus T$, $B \oplus S$, and $B \oplus T$ in the Kekulé cell.

6 Conclusions

In this paper, we presented a theoretical investigation into the computational possibilities of polycyclic hydrocarbons, based on connectivity via molecular paths with alternating single and double bonds. We confirmed and extended the results presented in [17]. Moreover, we introduced the new concepts of cell and functional cell in order to characterize the switching behaviour of single unsaturated polycyclic hydrocarbon molecules. We regard this as an interesting and fascinating subject, both from the point of view of new chemistry and new applications of mathematics, as well as for the speculations about their potential for molecular computation. We showed that it is, in principle, possible to build logical switches with π-conjugated systems of unsaturated polycyclic hydrocarbons, but we are fully aware of the tremendous technical problems in realizing

this potential. Anyway, the notion of functional cell emerging from our work is, in its own right, an interesting starting point for an alternative abstract model for computation. We showed that it encompasses logical switches; it would be interesting to investigate whether and how, e.g., data storage can be modeled with it.

So far, we have not yet found a connection between our investigations and quantum computing, although such a connection might very wel exist. The relevant properties of unsaturated polycyclic hydrocarbons are a direct consequence of the quantum mechanical behaviour of electrons, but there seems no role for entanglement here, which renders the connection with quantum computing still elusive.

References

1. F.L. Carter. *Molecular Electronic Devices*. Marcel Dekker, 1982.
2. C.A. Coulson. *Valence*. Oxford Press, London, 1961.
3. F. Dulić, S.J. van der Molen, T. Kudernac, H.T. Jonkman, J.J. de Jong, T.N. Bowden, J. van Esch, B.L. Feringa, and B.J. van Wees. One-way optoelectronic switching of photochromic molecules on gold. *Phys. Rev. Lett.*, 91:207402, 2003.
4. P.W. Fowler and A. Rassat. Symmetry and distortive π-electrons in two- and three-dimensional conjugated systems. *Phys. Chem. Chem. Phys.*, 4:1105–1113, 2002.
5. M. Grötschel, L. Lovász, and A. Schrijver. *Geometric algorithms and combinatorial optimization*. Springer V., Berlin, etc., 1988.
6. A.J. Heeger, S. Kivelson, J.R. Schrieffer, and W.-P Su. Solitons in conducting polymers. *Rev. Mod. Phys.*, 60:781, 1988.
7. H. Heitler and F. London. Wechselwirkung neutraler atome und homöpolare bindung nach der quantenmechanik. *Z. Phys.*, 44:455, 1927.
8. G.N. Lewis. The atom and the molecule. *J. Am . Chem. Soc.*, 38:762, 1916.
9. M. Mayor, H.B. Weber, J. Reichert, M. Elbing, C. Von Hänish, D. Beckmann, and M. Fisher. Electric current through a molecular rod-relevance of the position of the anchor groups. *Angew. Chem. Int. Ed.*, 42:5834, 2003.
10. S. Micali and V.V. Vazirani. An $\mathcal{O}(\sqrt{|v|} \cdot |E|)$ algorithm for finding maximum matching in general graphs. In *21st Annual Symposium on Foundations of Computer Science*, pages 17–27. IEEE, 1980.
11. A. Nitzan and M.A. Ratner. Electron transport in molecular wire junctions. *Science*, 300:1384, 2003.
12. S. Owre, N. Shankar, J.M. Rushby, and D.W.J. Stringer-Calvert. *PVS Version 2.4, System Guide, Prover Guide, PVS Language Reference*, 2001. `http://pvs.csl.sri.com`
13. L. Pogliani. From molecular connectivity indices to semiempirical connectivity terms: recent trends in graph theoretical descriptors. *Chem. Rev.*, 100:3827–3858, 2000.
14. S. Roth and H. Bleier. Solitons in polyacetylene. *Adv. Phys.*, 36:385, 1987.
15. S. Shaik and P.C. Hiberty. Myth and reality in the attitude toward valence-bond (VB) theory: Are its failures real? *Helv. Chim. Acta*, 86:1063–1084, 2003.
16. I.V. Stankevich, M.I. Skvortsova, and N.S. Zefirov. On a quantum chemical interpretation of molecular connectivity indices for conjugated hydrocarbons. *J. Molecular Structure*, 342:173–179, 1995.

17. M.H. van der Veen. *π-Logic*. PhD thesis, University of Groningen, May 2006.
18. M.H. van der Veen, M.T. Rispens, H.T. Jonkman, and J.C. Hummelen. Molecules with linear π-conjugated pathways between all substituents: Omniconjugation. *Adv. Funct. Mat.*, 14:215–223 and 742, 2004.
19. E.H. van Dijk, D.J.T. Myles, M.H. van der Veen, and J.C. Hummelen. Synthesis and properties of an anthraquinone-based redox switch for molecular electronics. *Organic Letters*, 8:2333–2336, 2006.
20. S.N. Yalirahi and M.A. Ratner. Interplay of topology and chemical stability on the electronic transport of molecular junctions. *Ann. N.Y. Acad. Sci.*, 960:153, 2002.