Vector Representations of Kekule Structures of Benzenoids

Daouglas J. Klein

Texas A&M University @ Galveston, Galveston, TX 77553-1675, USA

* Corresponding author: E-mail: kleind@tamug.edu

Received: 17-06-2010

This paper is dedicated to Professor Milan Randić on the occasion of his 80th birthday

Abstract

Milan Randić’s contributions to resonance in benzenoids is recalled, with particular focus on his recent identification and exploration of his “algebraic Kekule structures” (in terms of a suitable sequence of digits). In particular, further such vectorial representations of Kekule structures are noted, are characterized a bit, and are compared to one another.

1. Setting the Stage

Milan Randić has worked on an immense variety of chemical graph-theoretic topics – often making seminal contributions. One frequent area of his work concerns Kekule structures (or perfect matchings) on a molecular graph (prototypically representing a homo-atomic π-network of a molecule). See, e.g., Milan’s review1 of aromaticity, especially of benzenoids. He references somewhere approaching 100 of his relevant works in the area, along with somewhere approaching 1000 other articles.

A finite graph \( B \) is termed benzenoid if it is planar, 2-connected, and all internal faces of a planar embedding may be taken as hexagonal such that a pair of rings either share a single edge or nothing. That is, \( B \) may be viewed as a planar network of edge-fused regular hexagons. For any benzenoid system \( B \), some common notation is useful. Let \( h, n, \) and \( e \) respectively denote the numbers of hexagons, vertices, and edges in \( B \). Also let \( n^\text{in} \) be the number of internal vertices (i.e., vertices belonging to 3 hexagons) in \( B \). Further presume that \( B \) is Kekulean, which means that \( B \) has at least 1 fully neighbor-paired resonance structure, termed a Kekule structure \( \kappa \). Then following Randić’s ideas2–7 denote:

- \(|\kappa|_\xi \) as the electron-count for \( \kappa \) in ring \( \xi \) of \( B \), obtained as the sum of: first the number of bonds of \( \kappa \) coinciding with an edge shared by \( \xi \) with a second ring; and second, twice the number of bonds of \( \kappa \) coinciding with an edge belonging to just \( \xi \).

- \( R(\kappa) \) as the \( h \)-component vector of numbers \(|\kappa|_\xi \), with \( \xi \) ranging over the \( h \) ring labels.

- \( R(\kappa) \) as the sum of the \(|\kappa|_\xi \) over all rings \( \xi \) of \( B \).

Indeed it was only recently that Randić introduced\(^2\) the vector \( R(\kappa) \) as a novel invariant, associated to a Kekule structure \( \kappa \) of \( B \). Milan called \( R(\kappa) \) an “algebraic Kekule structure”, but we describe such as the Randić vector representation – because \( R(\kappa) \) seems more like a vector than a member of an algebra. For contrast Milan described the usual pictorial representations of a Kekule structure as “geometric Kekule structures”. It seemed\(^2\) that the \( R(\kappa) \) are faithful for most cata-condensed benzenoids, in that each Kekule structure of a given \( B \) has a different vector \( R(\kappa) \), and indeed this was quickly shown\(^3\) to be true (so long as \( 2 \leq n < \infty \)). But otherwise \( R(\kappa) \) is sometimes unfaithful, with more than one Kekule structure giving the same vector \( R(\kappa) \). Obviously there is unfaithfulness for benzene. But for a more wide-spread sort of non-faithfulness see pyrene in figure 1. Indeed this unfaithfulness for pyrene essentially persists on appending rings in suitable ways – e.g., for the species of figure 2. Moreover, different aspects of Milan’s representation were quickly pursued, by several interested workers – as in \([3,4,5,6,7]\).

Here this notion of representations for Kekule structures is pursued from a broadened perspective, looking for other near-faithful, or even fully faithful vector representations of the various Kekule structures of a Kekulean benzenoid \( B \). Such representations continue to entail vectors (for each Kekule structure) with numerical values for components. After noting other such representations, we
indicate some characteristics, and make some comparisons.

Two distinct Kekule structures of pyrene, with the same indicated \( R(\kappa) = (5,3,3,5) \). These two Kekule structures also turn out to have the same Sahini symbol \( S(\kappa) = (3,2,2,3) \).

Two further species which manifest unfaithfulness. For the first species simply append rings on the left-most and right-most rings of pyrene (of figure 1). For the second species, take the two structures of figure 1 and make the top ring of the first structure coincide with the bottom ring in the second structure – and for a second Kekule structure with the same representation, make the bottom ring of the first coincide with the top ring of the second.

2. Other Vector Specifications for Kekule Structures

Before delving into these other vector representations for Kekule structures, a useful result is:

**Lemma 1** – Let \( B \) be a benzenoid with \( h \) hexagons and \( n_{in} \) internal vertices. Then \( n = 4h + 2 - n_{in} \) and \( e = 5h + 1 - n_{in} \), while also \( B \) is Kekulean = \( n_{in} \) is even & \( R(\kappa) = 4h + 2 - n_{in} \) is independent of \( \kappa \). The equations for \( n \) and \( e \) are widely recognized, and are here useful to express and relate dimensions of different Kekule-structure vector representations (such as \( R(\kappa) \)). The result concerning \( R(\kappa) \) follows from the realization that for each double bond (or equivalently for each \( \pi \)-bond), there is a net count of 2 electrons upon summation over all rings.

As a first further example, note that there is yet another vector representation of a Kekule structure described by Sahini in 1961, and also considered in Harary et al. in 1991. This Sahini representation \( S(\kappa) \) again has components in correspondence with the rings of \( B \), the component for a ring \( \xi \) taking a value which is the number of edges of \( \kappa \) in common with the edges of \( \xi \). This representation also is non-faithful for benzene and pyrene, as may be seen from figure 1. Again the unfaithfulness persists with suitable appending of extra rings. Sahini’s representation however has a narrower range for the values of its components: \{0,1,2,3\} for \( S(\kappa) \) vs. \{0,1,2,3,4,5,6,7\} for \( R(\kappa) \). Though \( S(\kappa) \) shares many general features with \( R(\kappa) \), evidently Milan’s did not entertain this index as he wished to partition the \( \pi \) electrons to rings – the idea of not “double counting” electrons goes back to Clar, with the idea being whole-heartedly endorsed (and even championed) by Milan, e.g., in his review.

As a second example, there is a common vector representation of Kekule structures, which we term geometric (since it is so intimately related to what Milan termed “geometric”). One introduces a vector \( G(\kappa) \) with components corresponding to the edges of \( \kappa \) such that the \{\( u,v \)\}th component of \( G(\kappa) \) is 1 if this edge \{\( u,v \)\} (between vertices \( u \) and \( v \) ) also occurs in \( B \) and otherwise is 0. An example structure is shown in figure 3. Perhaps Randić ignored this representation as “vector” because it is so directly related to the common “geometric” diagrams superimposed on \( B \) itself. As this geometric vector representation identifies the particular edges in a Kekule structure, it is clear that \( G(\kappa) \) is faithful. But it resides in a vector space of a dimension somewhat greater than that of Randić: \( e = 5h + 1 - n_{in} \) for \( G(\kappa) \) vs. \( h \) for \( R(\kappa) \). Still the components of the \( G(\kappa) \) are confined to a narrower range of values: \{0,1\} for \( G(\kappa) \) vs. \{0,1,2,3,4,5,6,7\} for \( R(\kappa) \).

A third connectivity vector representation of a Kekule structure, is also given as a vector \( C(\kappa) \), this time with a number of components corresponding to the number of starred sites of \( B \). [Here \( B \) is bipartitioned into starred and unstarred sites, such that each site of one set is adjacent only to others of the other set.] Then the component of \( C(\kappa) \) for such a starred vertex \( u \) corresponds to the direction which the edge of \( \kappa \) incident at \( u \) takes: say, 0 if this edge is in the vertical direction from \( u \), and + or – as it leaves from \( u \) to the right or left. Of course there are generally 6 choices for such an orientation of a molecular structure – and one might wish to take the standard IUPAC orientation. But also there are 2 choices for starred and unstarred vertices – though they of course give the same numbers of bonds for corresponding bond directions. Again see figure 3. Notably this vector representation is faithful, since it specifies the direction of each edge of \( \kappa \) from the (unique) “starred” site to which the edge is incident in \( B \). Here the dimension of the space in which these vectors reside is \( n/2 = 2h + 1 - n_{in}/2 \) vs. \( h \) for \( R(\kappa) \). On the other hand the range of values for the components is smaller: \{+,0,–\} for \( C(\kappa) \) vs. \{0,1,2,3,4,5,6,7\} for \( R(\kappa) \). Again this representation is faithful – indeed as is reflected in the use of much this sort of representation in different resonating VB computations.
We summarize:

**Theorem 2** – Both the Randić and Sahini vector representations sometimes are unfaithful, with each sometimes failing when the other does not, and sometimes both fail. The geometric and connectivity vector representations are faithful.

Here most of this has been already established – aside from the non-mutual failures of the Randić and Sahini schemes. Thus in figure 4 a case is shown where the Randić representation is faithful, while the Sahini scheme is not. And in figure 5 a case is shown where the Sahini representation is faithful, while the Randić scheme is not.

**3. Further Characteristics of Representations**

One might seek other general distinguishing characteristics of different vector representations, beyond the dimensions of the associated vector spaces and the ranges of the vector components. For instance, Randić points out that for his representation, every vector $R(\kappa)$ has a sum of its components independent of $\kappa$, and equal to the total number $n$ of electrons. Milan (and perhaps especially A. T. Balaban) advocate this vector as a chemically meaningful way to partition the $\pi$-electrons amongst the different rings – with this partitioning emphasized in the titles to their joint papers. The alternative sum

$$R_\xi = \sum_{\kappa} |\kappa|_\xi$$

is less explored but yields what evidently is an “aromatic content” for the ring $\xi$.

The Sahini representation sum over rings to generally give different values for each Kekule structure of a benzenoid $B$:

$$\sum_\xi S_\xi(\kappa) \equiv S(\kappa)$$

The sum here counts each interior $\pi$-bond twice (as each such bond is in 2 rings) whereas the $\pi$-bonds on the boundary are only counted once. Thus noting that there are $n/2$ $\pi$-bonds in all, we have $n - n/2 \leq S(\kappa) \leq n$, and further $n - S(\kappa)$ and $S(\kappa) - n/2$ respectively count numbers of interior and boundary $\pi$-bonds. Moreover if we count up the number of rings $\xi$ for which $S_\xi(\kappa) = 3$, then this is a conjugated 6-circuit count for the Kekule structure $\kappa$ – such a count $\partial_6(\kappa)$ arguably giving an “importance” of a Kekule structure $\kappa$. Thence $\partial_6(\kappa)$ leads to a plausible “size-consistent” weight $\omega^{\partial_6}(k)$ for $\kappa$ in a ground-state resonating valence-bond (VB) wave-function for $B$. Indeed such an Ansatz has been used, and there is further evidence in a more general context that such an Ansatz is reasonable. The alternative sum

$$\sum_\xi S_\xi(\kappa) \equiv S_\xi$$

(or the modified alternative $\sum_\xi |\kappa|^{\omega^{\partial_6}}_\xi$) could be a plausible weight measuring the “aromaticity” of a ring $\xi$, as was fairly early so suggested by Milan. This interpretation parallels that for the corresponding $R_\xi$-sum, though which might be better is presently uncertain.

For the (standard) geometric representation $G(\kappa)$, the sum $\sum_\xi |\kappa|^{\omega^{\partial_6}}_\xi$ gives the well-known Pauling bond order. The alternative sum $\sum_\xi \omega(\kappa)$ clearly gives a constant $\kappa$-independent value $n/2$.

Finally, the connectivity vector $C(\kappa)$, has simple sums $\sum_{\kappa} c_\xi(\kappa)$ or $\sum_{\kappa} c_{\kappa}(\kappa)$, with the first simply giving...
(again) the $\kappa$-independent value $n/2$. The second sum registers something about $\pi$-bond anisotropy at site $u$. But in place of this second sum, what is more fundamental is

$$\sum_n \exp \left\{ i \frac{2\pi}{3} C_u(n) \right\} \equiv C_u$$

(4)

This gives the direction on average for a bond from site $u$. A similar alternative to the first sum

$$2 \cdot n^{-1} \sum_u \left\{ \frac{2\pi}{3} C_u(x) \right\}$$

would also give a mean bond direction, now as averaged over the molecule for a given $\kappa$. These two complex sums depend on the orientational choices for $B$ – though in an especially simple way: just differing by an overall scalar factor $e^{i\frac{2\pi}{3}}$ or complex conjugation. The magnitudes are independent of these $6$ choices, whence it is natural to introduce

$$\left| 2 \cdot n \sum_u \exp \left\{ i \frac{2\pi}{3} C_u(\kappa) \right\} \right| \equiv \alpha(B)$$

(5)

to indicate the degree of anisotropy of double-bond orientations. Our choice of not labelling this with the Kekule-structure label $\kappa$ is made since this is $\kappa$-independent, as follows because the number of bonds in each orientation is$^{22-24}$ the same for each Kekule structure $\kappa$. That is, $\alpha(B)$ is an anisotropy invariant for $B$ – which if very high is an indicator of bond localization, with consequent chemical reactivities. The triple $(n_+, n_-, n_0)$ of the numbers of $\pi$-bonds in each direction has been a more frequent invariant – though when $C_u$ is combined with $n = 2(n_+ + n_- + n_0)$ this triple is recovered – and even from $\alpha(B)$ and $n$ one recovers the $n_+, n_-, n_0$ in unordered form.

The overall results of this section may be summarized in a pseudo-theorematic form:

Observation 3 – Each of the vector representations (Randić, Sahini, geometric, and connectivity) lead to different chemically meaningful component sums – sums either over $\kappa$ or over the components for a particular $\kappa$.

### 4. Efficiency of Vector Representations

In fact one may look at global comparisons of different vector representations. For a mode $X$ of vector expression where each Kekule structure is represented by an $m$-vector with $c$ choices for the values of each component of such a vector, one naturally defines an extraneous for $X$ as its potential information$^{25}$ content

$$x(X) \equiv \log_2 c^m$$

(6)

One might in general wish to subtract from this the information content for the least extravagant yet faithful measure, if we knew the value of the extravagance for this optimal mode of presentation. The bare extravagances for several such modes of vector representation of Kekule structures are:

<table>
<thead>
<tr>
<th>$X$</th>
<th>$x(X)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>geometric</td>
<td>$c \cdot \log_2 2$</td>
</tr>
<tr>
<td>Sahini</td>
<td>$h \cdot \log 4$</td>
</tr>
<tr>
<td>Randić</td>
<td>$h \cdot \log 7$</td>
</tr>
<tr>
<td>connectivity</td>
<td>$\left( \frac{3}{2} \right) \cdot \log_2 3$</td>
</tr>
<tr>
<td>minimal</td>
<td>$\max_n \log_2 K(B) \approx 0.5 \cdot h$</td>
</tr>
</tbody>
</table>

Here despite their extravagances relative to the geometric representation, the schemes of Randić and Sahini are not faithful (i.e., are incomplete designations of Kekule structures) – though the scheme of Sahini$^{9,10}$ is less extravagant. The connectivity designation is faithful, though with an extravagance comparable to that the Randić (unfaithful) scheme. The extravagance of this faithful connectivity scheme is less than that of the Randić scheme for $n \geq 0.46 \cdot h$, as occur for bulkier graphene-like structures, where the number of boundary sites $n - n_{in} - \n n$ and $n_{in} \rightarrow 2h$. Indeed this faithful (connectivity) scheme becomes less extravagant than Sahini’s (unfaithful) scheme for $n \geq 1.48 \cdot h$, though here there is less leeway before the graphenic limit. Finally the last extravagance is that for a hypothetical faithful scheme identifying the minimum extravagance which is possible, its value given in the table just being a rough estimate (from known numbers of Kekule structures of large graphenic fragments). There seems to be a notable excess extravagance to any of the $4$ concrete vector representations we have mentioned.

Indeed one may observe that there is a “quasi-ordering” of representations $X$ as mediated by $x(X)$. (A brief chemical discussion of partial orderings and the relation to quasi-orderings is available$^{26}$.) That is, given two vector representations $X$ and $Y$, we say that $X > Y$ if $x(X) > x(Y)$ for all choices of $B$. Such an ordering is neatly shown with a Hasse diagram which is minimal and such that when $x(X) > x(Y)$ there is a path (in this diagram) from $X$ to $Y$ with each step downward – as shown for our present circumstance in figure 6.

A related partial ordering might be introduced for the degree of faithfulness of different representations. That is, one says $X > Y$ if the $X$-representation is never unfaithful in a case (i.e., for a $B$) where $Y$ is unfaithful. The faithfulness Hasse diagram then is shown in figure 7. There also we have introduced a double representation $R \oplus S$ where the Randić and Sahini vectors are used together to simultaneously label Kekule structures. In figure 7, the feature which distinguishes a quasi-order is manifested, when the $3$ faithful representations are clustered together. (Perhaps also it should be noted that the extravagance of...
the double $R \oplus S$ representation would be $x(R \oplus S) = x(R) + x(S)$.

A further point is that granted a faithful representation in terms of $m$-vectors each of whose components takes $c$ possible values, then $c^m$ clearly provides an upper bound to the number $K(B)$ of Kekule structures of $B$.

Theorem 4 – If $X$ is a faithful vector representation, then $K(B) \leq 2^{x(X)}$.

In particular from the connectivity representation, we have $K(B) \leq 9^h \cdot 3^{1-n/2}$.

More generally for unfaithful representations, one might manage to introduce some sort of unfaithfulness measure $u(X)$ for representation $X$, and then develop a bound of the form $K(B) \leq 2^{x(X) - u(X)}$. There could be a more elaborate quasi-ordering, reflecting both the extravagance and degree of unfaithfulness.

5. Conclusion and Prospects

In summary, it has been shown (typically with illustrations):  
- that there are a number of (even natural) vector representations of Kekule structures;
- that some representations are unfaithful while others are faithful;
- that each of these natural representations is associated to meaningful graph invariants (or molecular descriptors);
- that some representations are less “extravagant” in terms of the potential realm of information contemplated; and
- that these ideas lead to quasi-ordering relations of representations.

It may be noted that the lesser extravagant yet faithful schemes provide a natural framework for resonating VB computations.

As a yet further comment note that a number of the ideas here developed in the context of benzenoids extend to grapheneic nanostructures embedded in suitable locally Euclidian surfaces other than the Euclidean plane. Particularly a torus or an extended cylinder can be covered by hexagons in a benzenoid-like fashion, so that a certain amount of what has been said here should extend to these circumstances. The formulas for $n$ and $e$ in lemma 1 no longer apply, but many of the ideas depending more explicitly on Kekule structures remain applicable. The 4 basic definitions for vector representations carry through: the Randić and Sahini representations work best for systems with all rings hexagons; the geometric representation works generally with little qualification; the connectivity representation applies for bipartite (or alternant) graphs. Fullerenes and nano-cones entail other sized rings, whence besides the mathematical restrictions mentioned, there can be questions about the chemical relevance and interpretation. Particularly in the cases with pentagonal rings, the systems are non-alternant (i.e., non-bipartite) whence Pauling’s Kekule-structure “phases” turn out to not so easily be handled and thence leads to problems of “phase consistency” (which at least in selected cases can be surmounted).

It may be noted that even at a modest level of generalization to infinite benzenoids some modifications occur. The formulas of lemma 1 no longer apply (at least in the simple form given) – and the result that the Randić $R(x)$ are faithful for cata-fusenes (with $n \geq 2$), turns out to be no longer true – as illustrated in figure 8. Of course real experimental benzenoids are never infinite, though it is emphasized that these extended systems provide a “proper” reference for large (e.g., nano-structural) systems – such as occurs in molecular orbital theory with “bands”. But also even in thinking of resonance theory and Kekule structures, useful results apply, e.g., as found elsewhere. Another extension entails vector representa-

![Figure 6](image-url)  
**Figure 6** – The Hasse diagram for extravagances of the 5 considered types of Kekule structure representations.

![Figure 7](image-url)  
**Figure 7** – The Hasse diagram for faithfulness of various types of Kekule structure representations.
ations for Clar structures – but also useful ideas emerge from this limit even when\textsuperscript{33} Clar structures are used in place of Kekule structures. Thus there remains much further to investigate about Kekule-structure representations – especially as concerns applications to the current high-interest carbon nano-structures.

The various results reported here then notably broaden the realm of discourse considered by Milan and the several other chemical graph theoreticians who have published on vector representations (again often called “algebraic” representations) of Kekule structures. It seems that there is much more to be done, with some of it of chemical import – as it is felt should interest Milan Randić, numerous chemical graph theoreticians, and perhaps also the more general chemical community.

Acknowledgement is made of support (via grant BD-0894) from the Welch Foundation of Houston, Texas.

Figure 8 – Three Kekule structures on an infinite (cata-condensed) phenacene chain. Each has $R(\xi) = 4$ for every ring $\xi$ in the chain. Here the Sahini representation distinguishes the first two (as well as the first and third), but it does not distinguish the last two from one another – so that even the combined $R \oplus S$ manifests unfaithfulness, on catafusenes.

6. References

Povzetek

Pregledali smo prispevke Milana Randića k študiji resonance benzenoidov, s posebnim poudarkom na njegovi nedavnii oznaki in raziskavi »algebrajskih Kekuléjevih struktur« (v smislu ustreznih zaporedij številk). Opažene so nadaljnje vektorske predstaviteve Kekuléjevih struktur, ki jih v tem prispevku delno okarakteriziramo in primerjamo med seboj.