2007

A graph-theory approach to global determination of octet molecules

Damir Vukičević

Ray Hefferlin

A graph-theory approach to global determination of octet molecules

Damir Vukičević[1] and Ray Hefferlin[2]

[1]Department of Mathematics, University of Split, Nikole Tesle 12, HR-21000 Split

[2]Southern Adventist University, Collegedale, Tennessee 37315, USA

Abstract

Set theory is used to forecast the existence of all possible octet-rule cyclic and acyclic molecules formed from main-group atoms and having ionic and/or covalent bonding with orders up to three

1. Introduction

The octet rule is found to be "obeyed" by most main-group molecular structures, the more so in organic chemistry. It should be possible to invert the problem—to assume the rule and then predict, using mathematical techniques, which molecules obey it. Such prediction has been done, using simple algebra, for molecules with two,[1] three,[2] and four[3] atoms in a way that can readily be extended to larger species.[3] It has also been done with a somewhat more abstract algebra [4] and (in principle) using category theory.[5] In the present paper, we build on previous work[6-10] and use set theory for the determination of *all possible* main-group molecules—cyclic or acyclic—having covalent and/or ionic bonding with orders 1, 2, or 3. We begin with singly-bonded acyclic structures and move on to doubly and triply bonded structures; then we take up molecules with cycles. A computer logic is included which should be more general than the extant computer programs that determine structures from libraries of fragments.[4,11,12]

2. Mathematical analyses

Let G be any connected graph with maximal degree at most 4. Denote by $\nu(G)$ a quadruplet of numbers (n_1, n_2, n_3, n_4) where n_i is the number of vertices of degree i in G.

Denote by $\Gamma(p_1,...p_k;q_1,...,q_l)$ the set of all simple graphs G (i.e. graphs without loops or multiple edges) such that the set of vertices of G is $V(G)=\{x_1,...x_k,y_1,...,y_l\}$; that $d_G(x_i)=p_i$ for each $i=1,...,k$; that $d_G(y_i)=q_i$ for each $i=1,...,l$; and that no two vertices x_i and x_j are adjacent. Denote by $\Gamma_c(p_1,...p_k;q_1,...,q_l)\subseteq\Gamma(p_1,...p_k;q_1,...,q_l)$ the set of connected graphs in $\Gamma(p_1,...p_k;q_1,...,q_l)$.

Let $X=\{x_1,...,x_k\}$ and $Y=\{y_1,...,y_l\}$. Let S_1 and S_2 be two disjoint subsets of $V(G)$. Denote by $E(S_1)$ the set of edges with both end-vertices in G and by $E(S_1,S_2)$ set of edges with one end-vertex in S_1 and other in S_2. $E(S_1)$ is a special case of $E(G)=E(V(G))$. Put $e(S_1)=card(E(S_1))$ and $e(S_1,S_2)=card(E(S_1,S_2))$.

Let us prove an arbitrary lemma:

Proposition 1. Let $p_1,...p_k,q_1,...,q_l\in N_0$ such that $p_1\ge p_2\ge...\ge p_k$ and $q_1\ge q_2\ge...\ge q_l$. Then

$$\Gamma(p_1,...,p_k;q_1,...q_l)\neq\emptyset$$

if and only if

$$p_1\le l \text{ and } q_{p_1}\ge 1 \text{ and } \Gamma\left(p_2,...,p_k;q_1-1,...,q_{p_1}-1,q_{p_1+1},...,q_l\right)\neq 0.$$

Proof: Suppose that $p_1\le l$ and $q_{p_1}\ge 1$ and $G\in\Gamma\left(p_2,...,p_l;q_1-1,...,q_{p_1}-1,q_{p_1+1},...,q_l\right)$, then it is sufficient to add to G one vertex and connect it with vertices of degrees $q_1-1,...,q_{p_1}-1$ in order to obtain graph in $\Gamma(p_1,...,p_k;q_1,...q_l)$.

Now suppose that $\Gamma(p_1,...,p_k;q_1,...q_l)\neq\emptyset$. Obviously, $p_1\le l$ and $q_{p_1}\ge 1$. Let $G\in\Gamma(p_1,...,p_k;q_1,...q_l)$ be graph such with maximal r such that x_1 is adjacent to all vertices $y_1,...,y_r$. Distinguish two cases:

CASE 1: $r=p_1$.
In this case $G-p_1$ is isomorphic to element of $\Gamma\left(p_2,...,p_l;q_1-1,...,q_{p_1}-1,q_{p_1+1},...,q_l\right)$ which proves the claim.

CASE 2: Suppose that x_1 is adjacent to $y_1,...y_r$, but not to y_{r+1}. Then x_1 is adjacent to some y_{r+1+t}, $t>0$. Since $d_G(y_{r+1})\ge d_G(y_{r+1+t})$ and x_1 is adjacent to y_{r+1+t}, but not to y_{r+1}, it follows that there is a vertex v such that $vy_{r+1} \in E(G)$ and $vy_{r+1+t} \notin E(G)$, but then $G - x_1 y_{r+1+t} - vy_{r+1} + x_1 y_{r+1} + vy_{r+1+t} \in \Gamma(p_1,..., p_k; q_1,...q_l)$, which is in contradiction with maximality of r in G. ■

Here, we restrict our analyses to the vertices with degrees $1,2,3$ and 4. The degree sequence $(p_1, p_2,..., p_k) = (4,...,4,3,...,3,2,...,2,1,...,1)$ can be compactly written as (m_1, m_2, m_3, m_4), which simply means that there are m_1 copies of value 1, m_2 copies of value 2, and so on. We write this correspondence $(p_1, p_2,..., p_k) \leftrightarrow (m_1, m_2, m_3, m_4)$ and analogously $(q_1, q_2,..., q_k) \leftrightarrow (n_1, n_2, n_3, n_4)$.

Here, we always assume that $p_1 \geq p_2 \geq ... \geq p_k$ and that $q_1 \geq q_2 \geq ... \geq q_l$. Also, we compactly write $\Gamma(p_1,..., p_k; q_1,..., q_l) = \Gamma((m_1, m_2, m_3, m_4);(n_1, n_2, n_3, n_4))$.

From Lemma 1 directly follows

Proposition 2. Let $m_1, m_2, m_3, m_4, n_1, n_2, n_3, n_4 \in N_0$ and let $n_4 \geq 1$ and let at least one of numbers m_1, m_2, m_3 and m_4 be different form 0. Then $\Gamma((m_1, m_2, m_3, m_4);(n_1, n_2, n_3, n_4)) \neq \varnothing$ if and only if one of the following holds:

1) $m_1 + 2m_2 + 3m_3 + 4m_4 \leq n_4$ and $\Gamma\left((0,0,0,0);\begin{pmatrix} n_1, n_2, n_3 + m_1 + 2m_2 + 3m_3 + 4m_4, \\ n_4 - m_1 - 2m_2 - 3m_3 - 4m_4 \end{pmatrix}\right) \neq 0$;

2) $p_1 \leq n_4 < m_1 + 2m_2 + 3m_3 + 4m_4$ and
$\Gamma\left(p_{j+1}, p_{j+2},..., p_k; \left(n_1, n_2 - n_4 + \sum_{i=1}^{j} p_i, n_3 + 2n_4 - \sum_{i=1}^{j} p_i, 0\right)\right) \neq 0$, where j is the smallest number such that $\sum_{i=1}^{j} p_i \geq n_4$;

3) $n_4 < p_1 \leq n_1 + n_2 + n_3 + n_4$ and $\Gamma\left(p_2, p_3,..., p_k; q_1 - 1, q_2 - 1,..., q_{p_1} - 1, q_{p_1+1}, q_{p_1+2},..., q_l\right) \neq 0$.■

Proposition 3. Let $m_1, m_2, m_3, m_4, n_1, n_2, n_3 \in N_0$ and let $n_3 \geq 1$ and let at least one of numbers m_1, m_2, m_3 and m_4 be different form 0. Then $\Gamma((m_1, m_2, m_3, m_4);(n_1, n_2, n_3, 0)) \neq \varnothing$ if and only if one of the following holds:

1) $m_1 + 2m_2 + 3m_3 + 4m_4 \leq n_3$ and $\Gamma\left((0,0,0,0);\begin{pmatrix} n_1, n_2 + m_1 + 2m_2 + 3m_3 + 4m_4, \\ n_3 - m_1 - 2m_2 - 3m_3 - 4m_4, 0 \end{pmatrix}\right) \neq 0$;

2) $p_1 \leq n_3 < m_1 + 2m_2 + 3m_3 + 4m_4$ and
$\Gamma\left(p_{j+1}, p_{j+2},..., p_k; \left(n_1 - n_3 + \sum_{i=1}^{j} p_i, n_2 + 2n_3 - \sum_{i=1}^{j} p_i, 0, 0\right)\right) \neq 0$, where j is the smallest

number such that $\sum_{i=1}^{j} p_i \geq n_3$;

3) $n_3 < p_1 \leq n_1 + n_2 + n_3$ and $\Gamma\left(p_2, p_3, \ldots, p_k; q_1 - 1, q_2 - 1, \ldots, q_{p_1} - 1, q_{p_1+1}, q_{p_1+2}, \ldots, q_l\right) \neq 0$. ■

Proposition 4. Let $m_1, m_2, m_3, m_4, n_1, n_2 \in N_0$ and let $n_2 \geq 1$ and let at least one of numbers m_1, m_2, m_3 and m_4 be different form 0. Then $\Gamma\left((m_1, m_2, m_3, m_4); (n_1, n_2, 0, 0)\right) \neq \varnothing$ if and only if one of the following three conditions holds:

1) $m_1 + 2m_2 + 3m_3 + 4m_4 \leq n_2$ and $\Gamma\left((0,0,0,0); \begin{pmatrix} n_1 + m_1 + 2m_2 + 3m_3 + 4m_4, \\ n_2 - m_1 - 2m_2 - 3m_3 - 4m_4, 0, 0 \end{pmatrix}\right) \neq 0$;

2) $p_1 < n_2 < m_1 + 2m_2 + 3m_3 + 4m_4$ and $\Gamma\left(p_{j+1}, p_{j+2}, \ldots, p_k; \left(n_1 + 2n_2 - \sum_{i=1}^{j} p_i, 0, 0, 0\right)\right) \neq 0$,

where j is the smallest number such that $\sum_{i=1}^{j} p_i \geq n_2$.

3) $n_2 \leq p_1 \leq n_1 + n_2$ and $\Gamma\left(p_2, p_3, \ldots, p_k; q_1 - 1, q_2 - 1, \ldots, q_{p_1} - 1, q_{p_1+1}, q_{p_1+2}, \ldots, q_l\right) \neq 0$.■

It can be easily seen that:

Proposition 5. Let $m_1, m_2, m_3, m_4, n_1 \in N_0$ and let $n_1 \geq 1$. Then $\Gamma\left((n_1, 0, 0, 0); (m_1, m_2, m_3, m_4)\right) \neq \varnothing$ if and only if $n_1 - m_1 - 2m_2 - 3m_3 - 4m_4$ is an even non-negative integer.■

The following theorem is given in Reference 7 (recall that the molecular graph is a simple connected graph):

Theorem A. Let n_1, n_2, n_3 and n_4 be nonnegative natural numbers such that $n_1 + n_2 + n_3 + n_4 \geq 3$. There is a molecular graph G such that $\nu(G) = (n_1, n_2, n_3, n_4)$ if and only if the following four conditions hold:

1) $n_3 + 2n_4 \geq n_1 - 2$;

2) $\dfrac{3n_3 + 4n_4 - n_1}{2} - n_2 \leq \dbinom{n_3 + n_4}{2}$;

3) $n_1 + n_3$ is an even number;

4) If $n_3 + n_4 = 1$ then $n_2 \geq 3n_3 + 4n_4 - n_1$.

From here we can deduce:

Proposition 6. Let n_1, n_2, n_3 and n_4 be nonnegative natural numbers such that $n_1 + n_2 + n_3 + n_4 \geq 3$. There is a simple graph G such that $\nu(G) = (n_1, n_2, n_3, n_4)$ if and only if the following three conditions hold:

1) $\frac{3n_3 + 4n_4 - n_1}{2} - n_2 \leq \binom{n_3 + n_4}{2}$;

2) $n_1 + n_3$ is an even number;

3) If $n_3 + n_4 = 1$ then $n_2 \geq 3n_3 + 4n_4 - n_1$.

Proof: Necessity can be proved completely by analogy to the proof of Theorem A. Let us prove sufficiency. Distinguish two cases:

CASE 1: $n_3 + 2n_4 \geq n_1 - 2$.

This case follows from Theorem A.

CASE 2: $n_3 + 2n_4 < n_1 - 2$.

Note that $n_1 - 2n_3 - 3n_4$ is an even number larger then 2. Let graph G consist of paths with $n_2 + n_3 + n_4 + 2$ edges and $\frac{n_1 - n_3 - 2n_4 - 2}{2}$ isolated edges. Pick arbitrarily n_3 vertices of degree 2 in the set A and n_4 vertices of degree 2 in B. Graph G' is created from graph G by adding a single pendant vertex to each vertex in A and 2 pendant vertices to each vertex in B. It can be easily seen that $G' \in \Gamma(0,0,0,0; n_1, n_2, n_3, n_4)$.■

Proposition 7. Let $m_1, m_2, m_3, m_4, n_1, n_2, n_3, n_4 \in N_0$ and let $\Gamma(m_1, m_2, m_3, m_4; n_1, n_2, n_3, n_4) \neq \varnothing$. Then $\Gamma_c(m_1, m_2, m_3, m_4; n_1, n_2, n_3, n_4) \neq \varnothing$ if and only if $m_3 + n_3 + 2m_4 + 2n_4 \geq m_1 + n_1 - 2$.

Proof: Suppose that $G \in \Gamma_c(m_1, m_2, m_3, m_4; n_1, n_2, n_3, n_4)$. Then number of edges in G is equal to $\frac{m_1 + n_1 + 2m_2 + 2n_2 + 3m_3 + 3n_3 + 4m_4 + 4n_4}{2}$, hence

$$\frac{m_1 + n_1 + 2m_2 + 2n_2 + 3m_3 + 3n_3 + 4m_4 + 4n_4}{2} \geq m_1 + n_1 + m_2 + n_2 + m_3 + n_3 + m_4 + n_4 - 1.$$

Which implies $m_3 + n_3 + 2m_4 + 2n_4 \geq m_1 + n_1 - 2$. Now suppose that $\Gamma_c(m_1, m_2, m_3, m_4; n_1, n_2, n_3, n_4) \neq \varnothing$ and that $m_3 + n_3 + 2m_4 + 2n_4 \geq m_1 + n_1 - 2$ but that $\Gamma_c(m_1, m_2, m_3, m_4; n_1, n_2, n_3, n_4) = \varnothing$. Let $G \in \Gamma(m_1, m_2, m_3, m_4; n_1, n_2, n_3, n_4)$ be a graph with the smallest number of components. From $m_3 + n_3 + 2m_4 + 2n_4 \geq m_1 + n_1 - 2$, it follows that at least one component in G contains a cycle. Let $v_a y_b$ (note that each edge has at least one vertex in y) be an edge of that cycle and let $v_c y_d$ be an edge of some other component of G. Note that

$$G - v_a y_b - v_c y_d + y_b v_c + y_d v_a \in \Gamma\left(m_1, m_2, m_3, m_4; n_1, n_2, n_3, n_4\right)$$

has a smaller number of components that G. This is a contradiction. ■

3. Algorithm

Propositions 1-7 yield a simple recursive algorithm which gives the answer to the question whether the set $\Gamma_c\left(m_1, m_2, m_3, m_4; n_1, n_2, n_3, n_4\right)$ is empty or not. We have the main function *IsEmpty*, the recursive procedure *RecIsEmpty* and the function *Collect* that transforms the ordinary sequence of numbers $\left(p_1,..., p_k\right)$ or $\left(q_1,..., q_l\right)$ into compact sequences $\left(n_1, n_2, n_3, n_4\right)$ and $\left(m_1, m_2, m_3, m_4\right)$.

IsEmpty $\left(m_1, m_2, m_3, m_3; n_1, n_2, n_3, n_4\right)$

1) If $m_3 + n_3 + 2m_4 + 2n_4 < m_1 + n_1 - 2$ then return *Set is empty*
2) Else return $\mathit{RecIsEmpty}\left(m_1, m_2, m_3, m_3; n_1, n_2, n_3, n_4\right)$

$\mathit{RecIsEmpty}\left(m_1, m_2, m_3, m_3; n_1, n_2, n_3, n_4\right)$

If $m_1 = m_2 = m_3 = m_4 = 0$ then
Begin

If $\left[\dfrac{3n_3 + 4n_4 - n_1}{2} - n_2 \le \dbinom{n_3 + n_4}{2}\right]$ and [$n_1 + n_3$ is an even number] and

[If $n_3 + n_4 = 1$ then $n_2 \ge 3n_3 + 4n_4 - n_1$] then return *Set is non-empty*

Else return *Set is empty*

End

Else if $n_2 = n_3 = n_4 = 0$ then
Begin

If $n_1 - m_1 - 2m_2 - 3m_3 - 4m_4$ is an even non-negative integer return *Set is non-Empty*

Else return *Set is empty*

End

Else if $n_3 = n_4 = 0$ then
Begin

If $m_1 + 2m_2 + 3m_3 + 4m_4 \le n_2$ then return

$$RecIsEmpty\left(0,0,0,0;n_1+m_1+2m_2+3m_3+4m_4,n_2-m_1-2m_2-3m_3-4m_4,0,0\right);$$

Else if $p_1 \le n_2 < m_1+2m_2+3m_3+4m_4$ then return

$$RecIsEmpty\left(Collect\left(p_{j+1},p_{j+2},...,p_k\right);n_1+2n_2-\sum_{i=1}^{j}p_i,0,0,0\right)$$

where j is the smallest number such that $\sum_{i=1}^{j}p_i \ge n_2$.

Else if $n_2 < p_1 \le n_1+n_2$ then return

$$RecIsEmtpy\left(\begin{array}{c}Collect\left(p_2,p_3,...,p_k\right);\\ Collect\left(q_1-1,q_2-1,...,q_{p_1}-1,q_{p_1+1},q_{p_1+2},...,q_l\right)\end{array}\right).$$

Else return *Set is empty*

End

Else if $n_4 = 0$

Begin

If $m_1+2m_2+3m_3+4m_4 \le n_3$ then return

$$RecIsEmpty\left(0,0,0,0;n_1,n_2+m_1+2m_2+3m_3+4m_4,n_3-m_1-2m_2-3m_3-4m_4,0\right)$$

Else if $p_1 \le n_3 < m_1+2m_2+3m_3+4m_4$ then return

$$RecIsEmpty\left(Collect\left(p_{j+1},p_{j+2},...,p_k\right);n_1-n_3+\sum_{i=1}^{j}p_i,n_2+2n_3-\sum_{i=1}^{j}p_i,0,0\right)\neq 0$$

where j is the smallest number such that $\sum_{i=1}^{j}p_i \ge n_3$;

Else if $n_3 < p_1 \le n_1+n_2+n_3$ then return

$$RecIsEmpty\left(\begin{array}{c}Collect\left(p_2,p_3,...,p_k\right);\\ Collect\left(q_1-1,q_2-1,...,q_{p_1}-1,q_{p_1+1},q_{p_1+2},...,q_l\right)\end{array}\right)$$

Else return *Set is empty*

End

Else

Begin

If $m_1+2m_2+3m_3+4m_4 \le n_4$ then return

$$RecIsEmpty\left(\begin{array}{c}0,0,0,0;n_1,n_2,n_3+m_1+2m_2+3m_3+4m_4,\\ n_4-m_1-2m_2-3m_3-4m_4\end{array}\right);$$

If $p_1 \le n_4 < m_1+2m_2+3m_3+4m_4$ then return

$$RecIsEmpty\left(Collect\left(p_{j+1},p_{j+2},...,p_k\right);n_1,n_2-n_4+\sum_{i=1}^{j}p_i,n_3+2n_4-\sum_{i=1}^{j}p_i,0\right)$$

where j is the smallest number such that $\sum_{i=1}^{j}p_i \ge n_4$;

If $n_4 < p_1 \le n_1+n_2+n_3+n_4$ then return

$$RecIsEmpty\left(\begin{array}{c} Collect\left(p_2, p_3, ..., p_k\right); \\ Collect\left(q_1-1, q_2-1, ..., q_{p_1}-1, q_{p_1+1}, q_{p_1+2}, ..., q_l\right) \end{array}\right)$$

Else return *Set is empty*

End

It can be easily seen that the recursion *RecIsEmpty* can be referenced at most five times during the execution of the program, hence the algorithm works very efficiently (in a constant time) for graphs of arbitrarily large size.

4. Molecules with a single covalent bonds and/or single ionic bonds

Suppose that we have eight types of atoms: metals with 1,2,3 and 4 electrons in the last shell (valences 1, 2, 3, and 4 as in Li, Be, B, and C) and non-metals with 7, 6, 5, and 4 electrons in the last shell (valences 7, 6, 5, and 4, that is, having 1, 2, 3, and 4 electrons missing in the last orbital shell, as in F, O, N, and C). Denote the numbers of metal atoms by me_1, me_2, me_3 and me_4 and the numbers of non-metal atoms with by nm_1, nm_2, nm_3 and nm_4.

Metal atoms are capable only of donating electrons to participate in ionic bonding. On the other hand non-metals are capable of accepting electrons to participate in ionic bonding, and also of forming covalent bonds. A molecule is stable if each atom has its last orbital shell filled with electrons. In ionic bonding, metal atoms obtain this stability by donating all electrons and non-metals obtain this stability by accepting the necessary numbers of electrons to fill their valence electron shell. A simple row-2 example is $[Li]^+[F]^-$.

In a molecule with only non-metal atoms, the atoms share outer-shell electrons such that there are eight electrons associated with each atom (counting both electrons in each bond). If the atoms are identical then the bond is strictly covalent; F_2 is a simple example. If the atoms differ then the bond is technically not a strict covalent bond but is often (as in this paper) referred to with that term. A row-2 example is F-O-F.

Van der Waals, hydrogen, and other bondings are not considered.

We are interested if there is molecule with only single covalent bonds and single ionic bonds that corresponds to numbers me_1, me_2, me_3, me_4, nm_1, nm_2, nm_3 and nm_4. We can solve this using our algorithm. We just calculate

$$IsEmpty\left(me_1, me_2, me_3, me_4, nm_1, nm_2, nm_3, nm_4\right)$$

and find out if there is molecule with the required properties. If we are interested only in molecules without cycles then we need to add the following condition:

$$me_1 + nm_1 = me_3 + nm_3 + 2me_4 + 2nm_4 + 2\,.$$

5. Acyclic molecules with a single and/or double covalent bonds and/or single and/or double ionic bonds

Denote $\nu(G)$ as in Section 1. Denote by Γ_{12} set of all connected graphs that have only single and/or double bonds and have no loops or nontrivial cycles (we assume that cycle is trivial if it consists of pair of vertices connected by double bond).

The following theorem is given in Reference 9:

Theorem B. Let $n_1, n_2, n_3, n_4 \in N_0$. Then, there is a graph $G \in \Gamma_{12}$ such that $\nu(G) = (n_1, n_2, n_3, n_4)$ if and only if $n_2 = n_3 = n_4 = 0$ and $n_1 = 2$; or the following five conditions hold:

1) $n_1 \equiv n_3 \pmod 2$
2) $n_1 \le n_3 + 2n_4 + 2$
3) if $n_3 = 0$, then $n_1 + 2n_2 \ge 4$
4) if $n_3 \ge 1$, then $n_1 + n_2 \ge 2$
5) if $n_1 = n_3 = 0$ then $n_2 = 2$.

Denote by $\Gamma_{12}(p_1,...p_k; q_1,..., q_k) = \Gamma_{12}((m_1, m_2, m_3, m_4); (n_1, n_2, n_3, n_4))$ set of all connected i.e. graphs without loops and nontrivial cycles and only with single and double bonds such that set of vertices of G is $V(G) = \{x_1,...x_k, y_1,..., y_l\}$ that $d_G(x_i) = p_i$, for each $i = 1,...,k$; $d_G(y_i) = q_i$ for each $i = 1,...,l$ and no two vertices x_i and x_j are adjacent.

Let us prove:

Theorem 8. Let $m_1, m_2, m_3, m_4, n_1, n_2, n_3, n_4 \in N_0$. Then,

$$\Gamma_{12}((m_1, m_2, m_3, m_4); (n_1, n_2, n_3, n_4)) \ne \varnothing$$

if and only if $m_1 = n_2 = m_2 = n_3 = m_3 = n_4 = m_4 = 0,\ n_1 = 2$ or the following six conditions all hold:

1) $m_1 + 2m_2 + 3m_3 + 4m_4 \le n_1 + 2n_2 + 3n_3 + 4n_4$

2) $n_1 + m_1 \equiv n_3 + m_3 \pmod 2$

3) $n_1 + m_1 \le n_3 + m_3 + 2n_4 + 2m_4 + 2$

4) if $n_3 = m_3 = 0$, then $n_1 + m_1 + 2n_2 + 2m_2 \ge 4$

5) if $n_3 + m_3 \ge 1$, then $n_1 + m_1 + n_2 + m_2 \ge 2$

6) $n_1 = m_1 = n_3 = m_3 = 0 \Rightarrow n_2 + m_2 = 2$

Proof: First, suppose that there is $G \in \Gamma_{12}\left(\left(m_1,m_2,m_3,m_4\right);\left(n_1,n_2,n_3,n_4\right)\right)$. Since no two vertices x_i and x_j adjacent, if follows that $m_1 + 2m_2 + 3m_3 + 4m_4 \le n_1 + 2n_2 + 3n_3 + 4n_4$ and the the remaining follows from the Theorem B. Now, let us prove the opposite implication.

If $m_1 = n_2 = m_2 = n_3 = m_3 = n_4 = m_4 = 0,\ n_1 = 2$, the claim is trivial. Hence, suppose that claims 1) – 6) follow. From Theorem B, it follows that there is graph $G \in \Gamma_{12}$ such that $\nu(G) = \left(n_1 + m_1, n_2 + m_2, n_3 + m_3, n_4 + m_4\right)$.

Denote by $\Gamma'_{12}\left(\left(m_1,m_2,m_3,m_4\right);\left(n_1,n_2,n_3,n_4\right)\right)$ the set of all connected i.e. graphs without loops and only with single and double bonds such that set of vertices of G is $V(G) = \{x_1,\ldots x_k, y_1,\ldots, y_l\}$ that $d_G(x_i) = p_i$, for each $i = 1,\ldots,k$; $d_G(y_i) = q_i$ for each $i = 1,\ldots,l$. Obviously, $\Gamma'_{12}\left(\left(m_1,m_2,m_3,m_4\right);\left(n_1,n_2,n_3,n_4\right)\right)$ contains graph isomorphic to G, hence $\Gamma'_{12}\left(\left(m_1,m_2,m_3,m_4\right);\left(n_1,n_2,n_3,n_4\right)\right) \ne \varnothing$. Let $G' \in \Gamma'_{12}\left(\left(m_1,m_2,m_3,m_4\right);\left(n_1,n_2,n_3,n_4\right)\right)$ be the graph in $\Gamma'_{12}\left(\left(m_1,m_2,m_3,m_4\right);\left(n_1,n_2,n_3,n_4\right)\right)$ with the smallest value of $e(X)$. If $e(X) = 0$, then $G' \in \Gamma_{12}\left(\left(m_1,m_2,m_3,m_4\right);\left(n_1,n_2,n_3,n_4\right)\right)$ and theorem is proved.

Hence, suppose that $e(X) \ne 0$. Since, $m_1 + 2m_2 + 3m_3 + 4m_4 \le n_1 + 2n_2 + 3n_3 + 4n_4$, it follows that $e(Y) \ne 0$, too. Denote $x_a x_b \in E(X)$ and $y_c y_d \in E(Y)$. Denote by $D(G)$ set of double bonds in G. Distinguish 4 cases:

CASE 1: $x_a y_c, x_b y_d \notin D(G)$.
Graph $G - x_a x_b - y_c y_d + x_a y_c + x_b y_d \in \Gamma'_{12}\left(\left(m_1,m_2,m_3,m_4\right);\left(n_1,n_2,n_3,n_4\right)\right) \ne \varnothing$ has smaller value of $e(Y)$, which is contradiction.
CASE 2: $x_b y_c, x_a y_d \notin D(G)$.
Graph $G - x_a x_b - y_c y_d + x_b y_c + x_a y_d \in \Gamma'_{12}\left(\left(m_1,m_2,m_3,m_4\right);\left(n_1,n_2,n_3,n_4\right)\right) \ne \varnothing$ has smaller value of $e(Y)$, which is contradiction.
CASE 3: $x_a y_c$, $x_b y_c \in D(G)$ or $x_a y_c$, $x_a y_d \in D(G)$ or $x_b y_d$, $x_b y_c \in D(G)$ or $x_b y_d$, $x_a y_d \in D(G)$.
In the first case y_c has degree 5; in the second x_a has degree 5; in the third case x_b has degee 5 and in the fourth case y_d has degree 5. Hence, in each case a contradiction is obtained.

All the cases are exhausted and the theorem is proved.■

Now a direct application to molecular structure will be made.

Denote the atoms by me_1, me_2, me_3, me_4, nm_1, nm_2, nm_3 and nm_4 as in Section 3. We are interested if there is molecule without rings, having only single and/or double covalent bonds and/or only single and/or double ionic bonds, and achieving stability for each atom. We just apply the Theorem 8, i.e. such a molecule exists if and only if one of the following three condition sets holds. The examples are all drawn from molecules with group-2 atoms, generalization to molecules with heavier atoms is not difficult.[1]

1) $me_1 = me_2 = me_3 = me_4 = nm_2 = nm_3 = nm_4 = 0,\ nm_2 = 2$ (as in O_2), or

2) $me_2 = me_3 = me_4 = nm_2 = nm_3 = nm_4 = 0,\ me_1 = nm_1 = 1$ (as in $[Li]^{+}[F]^{-}$), or

3) all of the following six relations hold:

3.1) $me_1 + 2me_2 + 3me_3 + 4me_4 \le nm_1 + 2nm_2 + 3nm_3 + 4nm_4$ (Li-O-Li and Li-O-F; O=C=O; Li-N=O and F-N=O; kite BeONLi)

3.2) $me_1 + nm_1 \equiv me_3 + nm_1 \pmod 2$ (Li-O-Li and Li-O-F; Li-N=O and F-N=O; kites OONLi and OONF; kites BeONLi and BeONF, where the beryllium atom donates electrons to the oxygen and nitrogen atoms)

3.3) $me_1 + nm_1 \le me_5 + nm_5 + 2me_4 + 2nm_4 + 2$ (Li-O-Li and Li-O-F; Li-O-O-Li and Li-O-O-F; $Li_2C{=}CLi_2$)

3.4) if $me_3 = nm_3 = 0$, then $me_1 + nm_1 + 2me_2 + 2nm_2 \ge 4$ (Li-O-Li, Li-O-F, and O=C=O satisfy the equality)

3.5) if $me_3 + nm_3 \ge 1$, then $me_1 + nm_1 + me_2 + nm_2 \ge 2$ (Li-N=O and F-N=O satisfy the equality;; kites OONLi and OONF and BeONLi satisfy the inequality)

3.6) $me_1 = nm_1 = me_3 = nm_3 = 0 \Rightarrow me_2 + nm_2 = 2$ (the simplest case is $[Be]^{2+}[O]^{2-}$)

6. Acyclic molecules with a single and/or double and/or triple covalent and /or ionic bonds

The following theorem is given in Reference 9, noting that the first inequality in line 3) is due to a missprint in Lemma 18 and hence that the following theorem contains $n_1 = 2$ instead of $n_1 + n_3 = 2$:

Theorem C. Let $n_1, n_2, n_3, n_4 \in \mathrm{N}_0$. There is a graph $G \in \Gamma_{123}$ if and only if one of the following four condition sets holds:

1) $n_1 = n_2 = n_4 = 0$ and $n_3 = 2$, or

2) $n_1 = n_3 = 0$ and $n_2 = 2$, or

[1] The examples provided for these statements should make it possible for the reader to supply examples for analogous statements in the following sections.

3) $n_1 + n_3 = 2$, $n_2 = 0$, $n_4 \geq n_3$ and $n_4 \equiv n_3 \pmod 2$, or

4) The fourth condition is that all of the following five relations hold:

4.1) $n_1 \equiv n_3 \pmod 2$

4.2) $n_1 \leq n_3 + 2 n_4 + 2$

4.3) if $n_3 = 0$, then $n_1 + 2 n_2 \geq 4$

4.4) if $n_3 \geq 1$, then $n_1 + n_2 \geq \max \{2 - n_4, 3 - n_3, 2 - (n_3 + 2 n_4 + 2 - n_1) / 4\}$

4.5) $n_1 + n_3 > 1$.

To make direct chemical application, this theorem can be utilized to give necessary and sufficent conditions on the numbers me_1, me_2, me_3, me_4, nm_1, nm_2, nm_3, nm_4 for the existence of molecules with only single and/or double and/or triple covalent bonds and only single and/or double and/or triple ionic bonds that corresponds to these numbers.
It is necessary and sufficient that the numbers satisfy

$$me_1 + 2me_2 + 3me_3 + 4me_4 \leq nm_1 + 2nm_2 + 3nm_3 + 4nm_4$$

and one of the following four conditions:

1) $me_1 = nm_1 = me_2 = nm_2 = me_4 = nm_4 = 0$ and $me_3 + nm_3 = 2$ or

2) $me_1 = nm_1 = me_3 = nm_3 = 0$ and $ne_2 + nm_2 = 2$

3) $me_1 + nm_1 + me_3 + nm_3 = 2$, $me_2 = nm_2 = 0$, $me_4 + nm_4 \geq me_3 + ne_3$ and $me_4 + nm_4 \equiv me_3 + ne_3 \pmod 2$

4) This fourth condition is that all of the following five relations hold:

4.1) $me_1 + nm_1 \equiv me_3 + nm_3 \pmod 2$

4.2) $me_1 + nm_1 \leq me_3 + nm_3 + 2me_4 + 2nm_4 + 2$

4.3) if $me_3 = nm_3 = 0$, then $me_1 + nm_1 + 2me_2 + 2nm_2 \geq 4$

4.4) if $me_3 + nm_3 \geq 1$ then

$$me_1 + nm_1 + me_2 + nm_2 \geq \max \left\{ \begin{matrix} 2 - me_4 - nm_4, 3 - me_3 - nm_3, 2 - \\ (me_3 + nm_3 + 2me_4 + nm_4 + 2 - me_1 - nm_1)/4 \end{matrix} \right\}$$

4.5) $me_1 + nm_1 + me_3 + nm_3 > 1$

7. Molecules that have bond orders up to three and that may contain cycles

The following theorems given in Reference 10 can be utilized in a completely analogous way for these molecules:

Theorem D. Let $n_1, n_2, n_3, n_4 \in N_0$. Then, $\Gamma_{12}^o(n_1, n_2, n_3, n_4) \neq \varnothing$ if and only if the following holds:

1) $n_1 \equiv n_3$ (mod 2)

2) $n_1 \leq n_3 + 2n_4 + 2$

3) $2\Delta \leq n_1 + 2n_2 + 3n_3 + 4n_4$

4) $2\Delta + 2\Delta' \leq n_1 + 2n_2 + 3n_3 + 4n_4 + 4$.

Theorem E. Let $n_1, n_2, n_3, n_4 \in N_0$. Then, $\Gamma_{123}^o(n_1, n_2, n_3, n_4) \neq \varnothing$ if and only if the following holds:

1) $n_1 \equiv n_3$ (mod 2);
2) $n_1 \leq n_3 + 2n_4 + 2$;
3) $2\Delta \leq n_1 + 2n_2 + 3n_3 + 4n_4$;
4) $2\Delta + 2\Delta' \leq n_1 + 2n_2 + 3n_3 + 4n_4 + 6$.

where Δ is maximal degree of the graph and Δ' is second maximal degree of the graph and Γ_{12}^o and Γ_{123}^o are defined analogously as above.

The results can be stated for in a chemically relevant way as follows:

CASE 1) There is molecule with only single and/or double covalent bonds and only single and/or double ionic bonds that corresponds to numbers me_1, me_2, me_3, me_4, nm_1, nm_2, nm_3 and nm_4 if and only if these five conditions hold:

1) $me_1 + 2me_2 + 3me_3 + 4me_4 \leq nm_1 + 2nm_2 + 3nm_3 + 4nm_4$

2) $nm_1 + me_1 \equiv nm_3 + me_3$ (mod 2)

3) $nm_1 + me_1 \leq nm_3 + me_3 + 2nm_4 + 2me_4 + 2$

4) $2\Delta \leq me_1 + nm_1 + 2me_2 + 2nm_2 + 3me_3 + 3nm_3 + 4me_4 + 4nm_4$

5) $2\Delta + 2\Delta' \leq me_1 + nm_1 + 2me_2 + 2nm_2 + 3me_3 + 3nm_3 + 4me_4 + 4nm_4 + 4$.

CASE 2) There is molecule with single and/or double and/or triple covalent bonds and single and/or double and/or triple ionic bonds that corresponds to numbers me_1, me_2, me_3, me_4, nm_4, nm_5, nm_6 and nm_7 if and only if

1) $me_1 + 2me_2 + 3me_3 + 4me_4 \le nm_1 + 2nm_2 + 3nm_3 + 4nm_4$

2) $nm_1 + me_1 \equiv nm_3 + me_3 \pmod 2$

3) $nm_1 + me_1 \le nm_3 + me_3 + 2nm_4 + 2me_4 + 2$

4) $2\Delta \le me_1 + nm_1 + 2me_2 + 2nm_2 + 3me_3 + 3nm_3 + 4me_4 + 4nm_4$

5) $2\Delta + 2\Delta' \le me_1 + nm_1 + 2me_2 + 2nm_2 + 3me_3 + 3nm_3 + 4me_4 + 4nm_4 + 6$.

8. Summary

The theory presented here gives the necessary and sufficient conditions for the existence of a molecule which has

- single and/or double and/or triple covalent bonds, and/or
- with single and/or double and/or triple ionic bonds, and
- exactly eight electrons in the valence shell of every atom (counting those that possessed before bonding and those that it gained or lost as a result of the bonding).

It is worthwhile to note what is *not* being claimed:

- That only octet molecules exist,
- That octet molecules are necessarily more stable under usual laboratory conditions than non-octet molecules,
- That any given octet molecule will be observable under usual laboratory conditions.

The synthesis of octet molecules, even if it requires working under unusual conditions, should be of interest for the following (rather esoteric) reasons:

- These working conditions, unusual or not, might after analysis be found to follow interesting patterns,
- The distributions of the molecules in their multidimensional spaces, known for diatomic[1] and triatomic[2] molecules, might be generalized to higher spaces.

References

1. B. James, K. E. Caviness, J. Geach, C. W. Walters, and R. Hefferlin, "Global Molecular Identification from Graphs. Neutral and Ionized Main-Group Diatomic Molecules," *J. Chem. Inf. Comp. Sci.* **42** (2002) 1-7.

2. J. Geach, C. W. Walters, B. James, K. E. Caviness, and R. Hefferlin, "Global Molecular Identification from Graphs. Main-Group Triatomic Molecules," *Croat. Chem. Acta* **75** (2002) 333-400.

3. C. W. Walters, K. E. Caviness, and R. Hefferlin, "Global Molecular Identification from Graphs. IV. Molecules with four Closed *p*-Shell Atoms and Beyond," *Croat. Chem. Acta* **77** (2004) 65-71.

4. M. Froyen and P. Herdewijn, "Correct Bond Order Assignments in a Molecular Framework using Integer Linear Programming with Applications where only Non-Hydrogen Atom Coordinates are Available," *J. Chem. Inf. Comp. Sci.* 45 (2005) 1267-1274.

5. L. Schulz, "Catagories in Chemistry and Chemical Computers," *J. Mol. Struct.* (Theochem) 187 (1989) 15-22.

6. D. Vukičević and A. Graovac, "On Molecular Graphs with Valencies 1, 2, and 4 with Prescribed Numbers of Bonds, Croat. Chem. Acta, **77** (2004) 313-319.

7. D. Vukicevic and N. Trinajstic, "On the Discriminatory Power of the Zagreb Indices for Molecular Graphs," *MATCH-Commun. Math. Comput. Chem.*, **53** (2005) 111-138.

8. D. Vukičević and A. Graovac, "Compact Valence Sequences for Molecules with Single, Double and Triple Covalent Bonds," *Croat. Chem. Acta* **78** (2005), 203-209.

9. D. Vukičević and A. Graovac, "Compact Valence Sequences for Molecules with Single, Double and Triple Covalent Bonds," *Croat. Chem. Acta*, 78 (2005) 203-209.

10. D. Vukičević and A. Graovac, "Compact Valence Sequences for Molecules with Single, Double and Triple Covalent Bonds II: Graphs with Non-trivial Cycles," *preprint*

11. S. Bohanec and J. Zupan, "Structure Generator GEN," *MATCH* No. 27 (1992) 49-85.

12. H. D. Luinge, "AEGIS, a Structure Generation Program in Prolog," *MATCH* No. 27, 175-189.