

The Number of Canonical Structures of Each Degree of Excitation for an Unsaturated or Aromatic Hydrocarbon

G. W. WHELAND, *Gates Chemical Laboratory, California Institute of Technology*

(Received March 5, 1935)

A simple procedure is outlined for obtaining the number of canonical structures of each degree of excitation for an arbitrary hydrocarbon molecule. The results are applied to the approximate calculation of the energies of a number of such systems.

IN the quantum-mechanical treatment of the unsaturated and aromatic hydrocarbons, and of the free radicals,¹ it is frequently desirable to know the number of canonical structures of each degree of excitation.² These figures can be obtained, of course, by writing down all the structures, and by then determining the number of each type by actual count. This becomes quite impracticable, however, with complicated molecules,³ and some simpler procedure is necessary.

DESCRIPTION OF THE METHOD

The orbitals are first arranged formally in a circle in the usual manner.⁴ There is some arbitrariness in this step, since in general the final result will depend upon the way in which

the initial ordering is carried out. Thus, if the six orbitals in benzene are arranged in the sequence 1, 2, 3, 4, 5, 6, in which they come in the molecule itself, there are 2 unexcited and 3 first excited structures. On the other hand, if they are arranged in the sequence 1, 3, 5, 2, 4, 6, there are 0 unexcited, 1 first excited, 3 second excited, and 1 third excited structures. This causes no great difficulty, however, since there is usually one single order which is more logical than any of the others (for benzene, for example, the first of the two given above), and, in any case, the problem is uniquely determined when once the order has been established.

The general procedure will be to associate with each molecule, or grouping of atoms, a polynomial of the form

$$k_0 + k_1z + k_2z^2 + k_3z^3 + \dots + k_nz^n,$$

where k_i is the number of structures of the i th degree of excitation, and z is simply a parameter. Obviously, finding the polynomial which corresponds to a given molecule is completely equivalent to solving the problem as originally formulated, since the coefficients k_i are themselves the desired quantities.

I. THE MOLECULE CONSISTS OF A STRAIGHT CHAIN

The orbitals are arranged in the circle in the same order in which they come in the chain. The polynomial C_N , associated with the chain of $2N$ atoms, is given by the recursion formula

$$C_N = C_{N-1} + z \sum_{j=2}^N C_{j-1} C_{N-j},$$

which is derived in the following manner:

¹L. Pauling and G. W. Wheland, *J. Chem. Phys.* **1**, 362 (1933). L. Pauling and J. Sherman, *ibid.* **1**, 679 (1933). G. W. Wheland, *ibid.* **2**, 474 (1934); **3**, 230 (1935).

²The significance of the terms used here is as follows (cf. reference 1): In the method which is used by Pauling and his co-workers, and to which the conclusions reached in this paper are applicable, only the interactions between the $[p]_A$ orbitals are considered. These latter are arranged formally in a circle or polygon, and all structures are drawn in which each bond lies wholly within the figure, and in which no two bonds intersect. The structures thus obtained form the "canonical" set, and are the only ones which need be considered (cf. reference 4). The various bonds are of two types: "effective" bonds between orbitals that are actually adjacent in the molecule, and "ineffective" bonds between orbitals that are not adjacent. To the degree of approximation of the present treatment, the latter do not contribute to the stability of the molecule and are introduced solely in consequence of the mathematical formalism of the method. A structure in which all bonds are effective is called "unexcited"; a structure in which one bond is ineffective is called "first excited," and so on.

³This is seen from the fact that the number of canonical structures for a system of $2N$ orbitals, each occupied by a single electron, is given by the expression $(2N)!/N!(N+1)!$, which increases very rapidly with N . Thus, with $2N=14$, there are already 429 structures.

⁴G. Rumer, *Nachr. d. Ges. d. Wiss. zu Göttingen*, M. P. Klasse, p. 337, 1932. L. Pauling, *J. Chem. Phys.* **1**, 280 (1933).

We first number the orbitals in order, from 1 at one end of the chain to $2N$ at the other end. Now we consider the canonical structures which contain a bond between orbital 1 and some other definite orbital, say $2j$. This bond divides the molecule into three parts: (1) the orbitals 1 and $2j$ involved in the bond; (2) the orbitals 2, 3, \dots $2j-1$ lying on one side of the bond; and (3) the orbitals $2j+1, 2j+2, \dots$ $2N$ lying on the other side of the bond. (In the special cases in which $j=1$, or N , the second, or the third, part will be missing. We then simply treat the missing part formally as a chain of 0 orbitals, and associate with it the polynomial $C_0=1$.) If these three parts are considered separately, they can be represented by the polynomials J_1, J_2 and J_3 , respectively. Then $J_1=1$ if $j=1$ (i.e., if the bond between orbitals 1 and $2j$ is effective) and $J_1=z$ otherwise (i.e., if the bond is ineffective). Similarly $J_2=C_{j-1}$, and $J_3=C_{N-j}$. If the three parts are now considered as a whole, the corresponding polynomial, which represents the totality of structures with a bond between orbitals 1 and $2j$, is the product $J_1J_2J_3$. This follows from the fact that there can be no bonds between any two of the three parts, since such bonds would necessarily lead to noncanonical structures. The law for combining such noninteracting systems is then formally identical with that for multiplying polynomials. We now consider the canonical structures in which the orbital 1 is bonded with some further definite orbital, say $2k$. As in the previous case, the associated polynomial is equal to $K_1K_2K_3$, with $K_1=1$ or z , $K_2=C_{k-1}$, and $K_3=C_{N-k}$. Obviously none of the structures of this second group will occur also in the first group, and consequently the totality of structures in which the orbital 1 is bonded to either $2j$ or $2k$ will be represented by the sum $J_1J_2J_3+K_1K_2K_3$. This same procedure can now be repeated until all of the possible canonical structures have been taken into consideration. The resulting polynomial is then

$$C_N = \sum_{j=1}^N J_1 J_2 J_3,$$

which is just the recursion formula given originally.

Since we have adopted the convention that

$C_0=1$, and since we see at once that $C_1=1$, we can find the remaining polynomials in order. The first few members of the series are:

$$\begin{aligned} C_1 &= 1, \\ C_2 &= 1 + z, \\ C_3 &= 1 + 3z + z^2, \\ C_4 &= 1 + 6z + 6z^2 + z^3, \\ C_5 &= 1 + 10z + 20z^2 + 10z^3 + z^4. \end{aligned}$$

In general:

$$C_N = \sum_{j=0}^{N-1} \frac{N!(N-1)!}{j!(j+1)!(N-j)!(N-j-1)!} z^j \\ = \frac{1}{N} \sum_{i=0}^{N-1} \binom{N}{i} \binom{N}{j+1} z^j.$$

II. THE MOLECULE CONSISTS OF A RING

For a ring of $2N$ orbitals, we can carry through the same sort of procedure as in the case of the chain. The result is then

$$R_N = 2C_{N-1} + z \sum_{j=2}^{N-1} C_{j-1} C_{N-j}.$$

This can be expressed more conveniently as

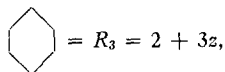
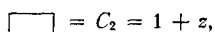
$$R_N = C_N + (1-z)C_{N-1},$$

in which form it is seen to be a special case of the general rule given below in III. The first few members of the series are:

$$\begin{aligned} R_2 &= 2, \\ R_3 &= 2 + 3z, \\ R_4 &= 2 + 8z + 4z^2, \\ R_5 &= 2 + 15z + 20z^2 + 5z^3. \end{aligned}$$


III. THE MOLECULE CONTAINS SEVERAL RINGS


It is convenient to represent the polynomial associated with a given molecule by means of a geometrical figure, which is constructed in the following way: After the orbitals are arranged in the circle, or polygon, lines are drawn connecting each pair of orbitals which are adjacent in the actual molecule, that is, each pair of orbitals for which the corresponding exchange integral is to be set equal to α . Thus we have, for example:



and so on.

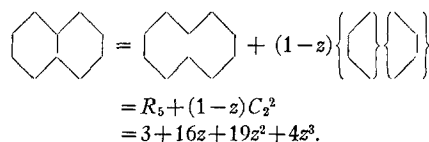
Now let us consider the relation between two figures which differ only in that one contains a single additional line which is absent from the other, as for

example: , corresponding to naphthalene,

and , corresponding to cyclo-

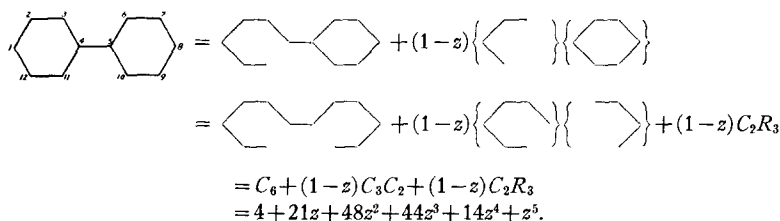
decapentane. For such cases, there is a general rule: If a figure, *A*, differs from a second, *B*, by having a single additional line, and if this line, together with its end points, divides *A* into two parts, *C* and *D*, then $A = B + (1-z)CD$. The

derivation of this equation follows from the same considerations as were used in I. Thus we consider the totality of the structures of each molecule in which there is a bond along this added line. In *A* these are represented by the product *CD*, and in *B* by the product *zCD*. Structures which do not contain this bond are not affected in degree of excitation by the added line. This completes the proof. The theorem, when applied to naphthalene, gives



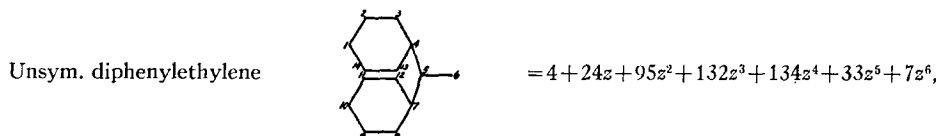
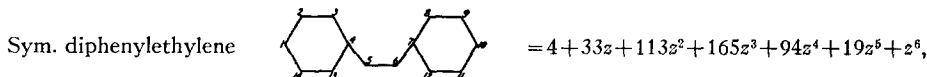
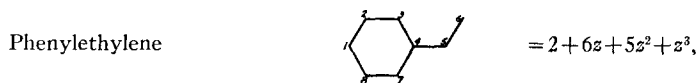
We find, then, that naphthalene has 3 unexcited, 16 first excited, 19 second excited, and 4 third excited structures, in agreement with the direct evaluation by Pauling and Wheland.⁴

A slightly more complicated example is afforded by biphenyl.⁵



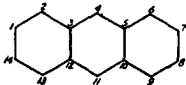
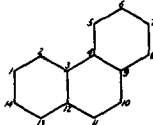
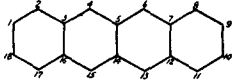
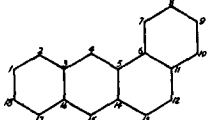
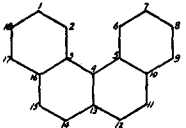
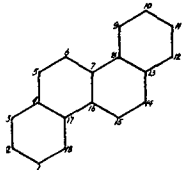
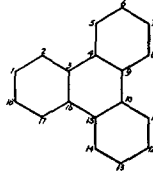
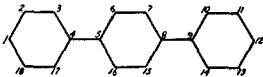
This result has been checked by writing down all 132 canonical structures, and counting the number of each degree of excitation.

A few further examples are given below:



⁵ Here, and in the following, we do not arrange the orbitals actually in a circle or polygon, but in whatever type of figure is most convenient for the purpose at hand. In cases in which there may be some question, the orbitals

will be numbered, to indicate the sequence which they would follow if they had been arranged in the polygon. It is absolutely essential that the same sequence be maintained throughout a given calculation.

Anthracene		$= 4 + 48z + 150z^2 + 163z^3 + 58z^4 + 6z^5,$
Phenanthrene		$= 5 + 47z + 148z^2 + 165z^3 + 59z^4 + 5z^5,$
Naphthacene		$= 5 + 110z + 649z^2 + 1556z^3 + 1646z^4 + 750z^5$ $+ 138z^6 + 8z^7.$
1,2 Benzanthracene		$= 7 + 112z + 642z^2 + 1551z^3 + 1654z^4 + 754z^5$ $+ 135z^6 + 7z^7,$
3,4 Benzophenanthrene		$= 8 + 116z + 636z^2 + 1542z^3 + 1663z^4 + 760z^5$ $+ 131z^6 + 6z^7,$
Chrysenes		$= 8 + 117z + 634z^2 + 1541z^3 + 1667z^4 + 759z^5$ $+ 129z^6 + 7z^7,$
Triphenylene		$= 9 + 117z + 645z^2 + 1557z^3 + 1659z^4 + 744z^5$ $+ 125z^6 + 6z^7,$
Terphenyl		$= 8 + 81z + 384z^2 + 1000z^3 + 1528z^4 + 1279z^5$ $+ 503z^6 + 76z^7 + 3z^8.$

In the resolution of several of the above figures, it is convenient to make use of a further rule which can easily be derived from the one above: Let $D_{p, q}$ be the polynomial corresponding to two straight chains, of p and q orbitals, respectively, which are not connected to each other. Then

$$D_{p, q} = C_{(p+q)/2} + (z-1)D_{p-1, q-1}$$

$$= \sum_{k=0}^{(p+q-1)/2} (z-1)^k C_{(p+q)/2-k}.$$

Thus

$$D_{7,5} = \begin{array}{l} \text{[Diagram: Octagon with a bond to the right]} \\ \text{[Diagram: Octagon with a bond to the left]} \end{array} = C_6 + (z-1)C_6 + (z-1)^2C_4 + (z-1)^3C_3 + (z-1)^4C_2 + (z-1)^5C_1, \\ = 12z + 32z^2 + 62z^3 + 20z^4 + 6z^5.$$

IV. THE MOLECULE CONTAINS AN UNPAIRED ELECTRON

This type of problem arises in the treatment of free radicals. The procedure is simply to introduce a "phantom orbital", X, containing a "phantom electron," which is formally bonded to the unpaired electron of the molecule.¹ The exchange integrals between this phantom orbital and the real orbitals are set equal to zero. As before, it is desirable first to obtain expressions for straight chains, and we find at once that the polynomial F_N , corresponding to the chain of $2N-1$ orbitals, is given by the expression

$$F_N = (1/z)D_{2N-1,1} = (1/z)[C_N + (z-1)C_{N-1}].$$

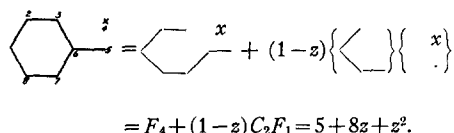
(The factor $1/z$ occurs since the bond to the phantom orbital X, although ineffective, is not considered to increase the degree of excitation.)

The first few members of the series are:

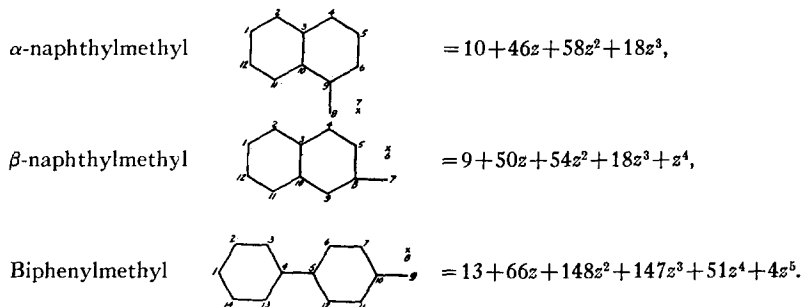
$$\begin{aligned} F_1 &= 1, \\ F_2 &= 2, \\ F_3 &= 3 + 2z, \\ F_4 &= 4 + 8z + 2z^2, \\ F_5 &= 5 + 20z + 15z^2 + 2z^3. \end{aligned}$$

It is interesting to observe that the coefficients k_i appearing in F_N are the same as those appearing in R_N , but that their order is reversed.

In the treatment of molecules containing rings, the same procedure can be used as before. For the phenylmethyl radical, for example, we have



A few other results are given below:



APPLICATION TO THE CALCULATION OF MOLECULAR ENERGIES

The author has shown recently⁶ that for a hydrocarbon molecule which does not contain any odd membered rings the energy which results from the simplified Heitler-London procedure of

⁶ G. W. Wheland, J. Chem. Phys. 3, 230 (1935).

Pauling and his co-workers¹ can be expressed as

$$W = Q + \frac{\sum R a_R g_R n_R}{\sum R a_R g_R},$$

where g_R is the number of structures of type R which appear with the coefficient a_R in the complete eigenfunction for the molecule, and n_R is

the number of effective bonds in any one of the structures R . Q and α are, as usual, the Coulomb integral and the single exchange integral, respectively. If all the structures of the i th degree of excitation are assumed to have an average coefficient a_i , then the above equation becomes

$$W = Q + \frac{\sum_i a_i k_i (b-i)}{\sum_i a_i k_i} \alpha,$$

where b is equal to the number of effective bonds in the unexcited structures. (If there are no phantom orbitals, $b = N =$ one-half the number of orbitals.)

If some method were available for estimating the coefficients a_i , the calculation of the energy of any molecule of the present type would be quite simple. The rule given in the previous paper,⁶ that $a_i = (0.3)^i$ is fairly satisfactory for small molecules, but it leads to too small values of the energy (in absolute magnitude) when N becomes greater than 5. It also fails to allow for variations of the coefficients between molecules of the same size, but of quite different structure. A rather better, but equally empirical rule can be put forward tentatively:

$$a_i = (0.04 - 0.93/N + 1.05B/N^2)^i,$$

where B is equal to the number of pairs of orbitals for which the corresponding exchange integral is α . Some results obtained by this method are given in Table I. It will be seen that they are really quite satisfactory. It is doubtful, however, if this procedure could be safely applied to molecules of widely different type from those considered here, since there is no assurance that the same rule for estimating coefficients would be at all satisfactory.⁷ In particular, the assumption

⁷ One readily sees, in fact, that the rule is no longer very good for the free radicals.

TABLE I.

Compound	a_1	Energy	
		Approx.	Correct ^a
Benzene	0.43	$Q + 2.61\alpha$	$Q + 2.61\alpha$
Cyclooctatetrane	0.33	3.31α	3.30α
Cyclodecapentane	0.27	4.05α	4.03α
Phenylethylene	0.33	3.30α	$3.31\alpha^c$
Sym. Diphenylethylene	0.23	5.66α	$5.59\alpha^c$
Unsym. Diphenylethylene	0.23	5.67α	$5.55\alpha^c$
Biphenyl	0.26	4.93α	$4.87\alpha^c$
Terphenyl	0.20	7.36α	$7.09\alpha^c$
Naphthalene	0.32	4.08α	4.04α
Anthracene	0.25	5.59α	$(5.53\alpha)^b$
Phenanthrene	0.25	5.65α	$(5.69\alpha)^b$
Naphthacene	0.21	7.15α	
1,2 Benzanthracene	0.21	7.21α	
3,4 Benzophenanthrene	0.21	7.24α	
Chrysene	0.21	7.24α	$(7.35\alpha)^b$
Triphenylene	0.21	7.26α	
Butadiene	0.36	1.74α	1.73α
Hexatriene	0.31	2.45α	2.48α
Octatetrane	0.27	3.21α	$3.22\alpha^c$

^a L. Pauling and J. Sherman, J. Chem. Phys. 1, 679 (1933). G. W. Wheland, J. Chem. Phys. 3, 230 (1935).

^b These values were calculated from the observed resonance energies (L. Pauling and J. Sherman, J. Chem. Phys. 1, 606 (1933)) with $\alpha = -1.5$ e.v. In view of the really surprising agreement which has been found, in all other cases, between the theoretical and the observed resonance energies (cf. ref. 1), this seems a safe method for estimating the order of magnitude of the former.

^c In these calculations, Pauling and Sherman used an approximate method of solving the secular equation. The procedure (for a justification of which see J. Sherman, J. Chem. Phys. 2, 488 (1934)) consisted in neglecting the second and higher excited structures and in arbitrarily equating coefficients. As a result, the coefficients of α appearing here are somewhat too low.

that a_i for a given molecule can be represented by the i th power of a constant quantity has little if any theoretical justification. The estimated coefficients may, therefore be very seriously in error, and the energy may come out approximately correct only in consequence of a cancelation of errors. Be this as it may, however, the method seems to be quite a satisfactory one from an empirical point of view, and can be applied with some assurance to the treatment of molecules which do not differ very greatly, either in size or in type, from those considered in Table I.