504. *The Prediction of the Spectra of Aromatic Hydrocarbons.*

By E. HEILBRONNER and J. N. MURRELL.

The abilities of the Hückel molecular-orbital theory, and of the Dewar approximation to this theory, to predict the position of the ${}^1L_a$ (p) band in the spectra of aromatic hydrocarbons, have been tested by statistical methods. The Hückel method is found to be the better on a 10% significance level. The regression line for the Hückel calculation is found to be

$$\Delta E/\beta = -0{\cdot}3736 + 0{\cdot}4558 \times 10^{-4}\nu_{\text{exp}}\ (\text{cm.}^{-1}),$$

and the variance about this regression $V(\Delta E/\beta) = 725 \times 10^{-6}$.

IN order to obtain a satisfactory interpretation of the electronic spectra of aromatic hydrocarbons, it is necessary to take into account the electron interaction terms in the Hamiltonian. A theory based on a one-electron Hamiltonian, such as the Hückel or the free-electron theory, fails even to predict the correct order of the excited states. However, Dewar[1] has shown that such a theory can be used to correlate the position of the first strong absorption band in this series of compounds: Platt's ${}^1L_a$ band[2] and Clar's p-band.[3]

Dewar used an approproximate method of calculating the Hückel excitation energies. The molecule is divided into two fragments, each containing an odd number of conjugated atoms. From the coefficients of the non-bonding orbitals of these fragments[4] he obtained, by first-order perturbation theory, an estimate of the separation of the highest bonding and the lowest anti-bonding Hückel orbital of the complete molecule. This is then taken to be proportional to the frequency of the ${}^1L_a$ band. Dewar stated that this approximate method predicted excitation energies better than the exact Hückel calculation, but he could not explain why this should be so. This statement has become established in the literature.[5]

Unfortunately Dewar's conclusion is due mainly to his determination of the magnitude of the resonance integral β from the benzene spectrum, which is just one case for which the one-electron theory breaks down, for reasons of symmetry. In this paper we give a re-appraisal of the Hückel and the Dewar method of predicting the position of the ${}^1L_a$ band, and by a statistical analysis of the calculated and the observed frequencies obtain an estimate of the accuracy with which the position of this band can be predicted for molecules whose spectra have not yet been determined.

The Table contains the experimental and theoretical results for all aromatic hydrocarbons which contain up to five condensed benzene rings, and some which are larger; it does not, however, include benzene and triphenylene which owing to their high symmetry are special cases (there is degeneracy in the highest bonding and the lowest anti-bonding orbital).

For the catacondensed hydrocarbons 1—19 we have made a linear regression of the calculated frequencies on the observed frequencies. We have assumed that once the ${}^1L_a$ band has been identified, then the errors in measuring these frequencies are negligible,

---

[1] Dewar, *J.*, 1952, 3532.

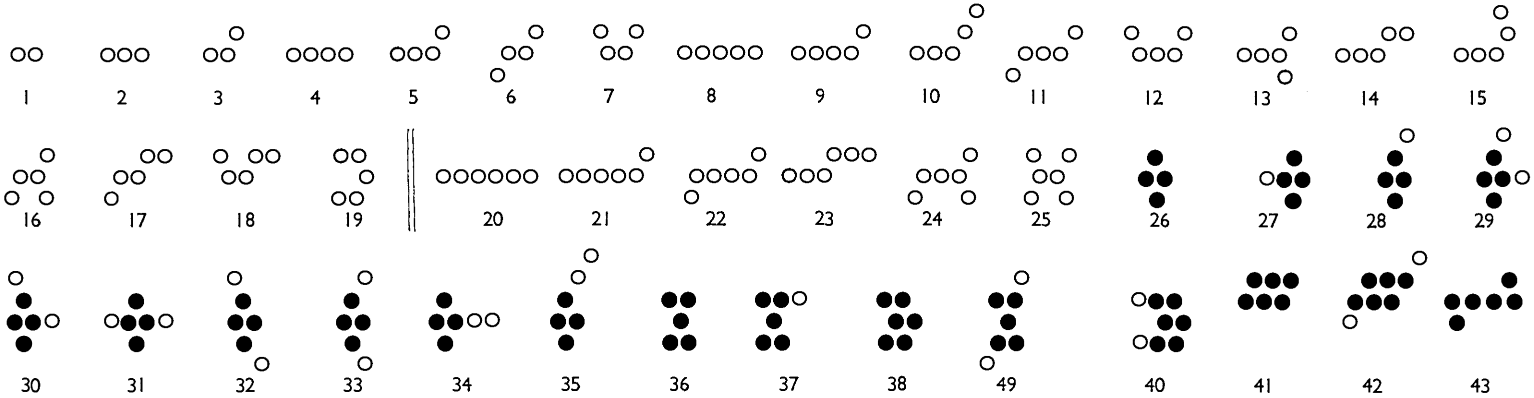
[2] Platt, *J. Chem. Phys.*, 1949, **17**, 484; Klevens and Platt, *ibid.*, p. 470.

[3] Clar, " Aromatische Kohlenwasserstoffe," Springer Verlag, Berlin, 1952, 2nd edn.

[4] Longuet-Higgins, *J. Chem. Phys.*, 1950, **18**, 265, 275, 283.

[5] *E.g.*, Daudel, Lefebvre, and Moser, " Quantum Chemistry," Interscience, New York, 1959, p. 90.

Comparison of the Hückel and the Dewar method.



					$10^{-4}\tilde{\nu}$	
No.**	Clar, page	$10^{-4}\tilde{\nu}$ exp. corr.	Hückel $\Delta E/\beta$	Dewar $\Delta E/\beta$	Hückel predict.	Dewar predict.
1	134	3·636	1·236	1·604	3·531	3·651
2	175	2·670	0·828	1·143	2·636	2·660
3	142	3·419	1·210	1·429 ‡	3·474	3·275
4	233	2·123	0·590 ‡	0·916	2·114	2·172
5 *	186	2·816	0·905	1·183	2·805	2·746
6	155	3·135	1·040	1·352	3·101	3·109
7	161	3·300	1·135	1·467	3·310	3·356
8 *	249	1·768	0·439	0·706	1·783	1·720
9 *†	240	2·240	0·654	0·972 ‡	2·254	2·292
10 *	213	2·816	0·874	1·166	2·737	2·709
11 *	199	2·879	0·947	1·200	2·897	2·782
12 *	204	2·989	0·983	1·302	2·976	3·002
13 *	196	3·006	0·998	1·313	3·009	3·025
14 *†	209	2·627	0·810	1·132	2·597	2·636
15 †	207	2·688	0·837	1·164	2·656	2·705
16 †	164	3·115	1·064	1·414	3·154	3·242
17 *	166	3·074	1·004	1·300	3·022	2·997
18 †	168	3·120	1·100	1·376	3·233	3·161
19 †	171	3·040	1·071	1·404	3·159	3·221
20 *	265	1·473	0·338	0·583	1·561	1·455
21 *‖	258	1·845	0·487	0·761	1·888	1·838

					$10^{-4}\tilde{\nu}$	
No.**	Clar, page	$10^{-4}\tilde{\nu}$ exp. corr.	Hückel $\Delta E/\beta$	Dewar $\Delta E/\beta$	Hückel predict.	Dewar predict.
22 *†	243	2·463	0·717	1·022	2·393	2·399
23 *†	225	2·481	0·696	1·094	2·347	2·554
24 *	205	3·033	1·045	1·377	3·112	3·163
25	—	—	1·023	1·198	3·064	2·778
26	324	2·999	0·901	1·156	2·796	2·688
27 †	333	3·016	0·994	1·287	3·000	2·969
28 †	336	2·750	0·742	0·926	2·448	2·193
29 *†	341	2·518	0·797	1·000	2·568	2·352
30 *‖	343	2·676	0·843	1·061	2·669	2·483
31 *‖	344	3·070	1·110	1·414	3·255	3·242
32 †	345	2·217	0·605	0·812	2·147	1·948
33 *†	346	2·549	0·684	1·088	2·320	2·541
34 *‖	351	2·929	1·011	1·295	3·038	2·987
35	—	—	0·740	0·912	2·443	2·163
36	284	2·304	0·695	1·053	2·344	2·466
37 ‖	291	1·950	0·530	0·750	1·982	1·815
38 †	301	2·581	0·878	1·216	2·746	2·817
39 †	—	—	0·711	—	2·380	—
40 †	—	—	0·588	—	2·110	—
41 *†	380	2·340	0·582	—	2·096	—
42 †	—	—	0·516	—	1·952	—
43 *†	389	1·638	0·397	—	1·691	—

* Spectrum taken in benzene, chloroform, or methylnaphthalene, 300 cm.$^{-1}$ is added to the observed frequency to correct for the solvent effect. † Hückel energies from Hall, *Trans. Faraday Soc.*, 1957, **53**, 573. ‡ Incorrect values given in reference 1. ‖ Hückel energies calculated by the authors. ¶ All Huckel energies from Coulson and Daudel, "Dictionaire des Grandeurs Théoriques Descriptives des Molecules," Fasc. II, except where indicated.
** Open circles denote rings attached by only one side; closed circles denote rings attached by more than one side.

so that these frequencies can then be used to test the usefulness of a theory. The calculated frequencies, on the other hand, are obtained from theories which involve a large number of assumptions. We assume that the differences between these calculated values and those which could be obtained by an exact calculation differ from compound to compound in an uncontrolled way, so that the calculated frequencies can be treated as our dependent variable in the statistical sense.

In the series 1—19 there are two compounds, nos. 5 and 14, for which the identification of the ^1L_a band is uncertain. In both cases the vibrational structure in the region in question is unusual and could possibly arise from the overlap of two electronic transitions. There are three vibrational peaks, any one of which might with justification be taken as the maximum of the ^1L_a band. For both molecules the first peak is slightly weaker than the other two. For no. 5 Clar chooses λ_1 as the wavelength of his *p*-band, for no. 14 he chooses λ_2. We have chosen λ_1 in both cases since it agrees much better with the wavelengths expected from the two theories.

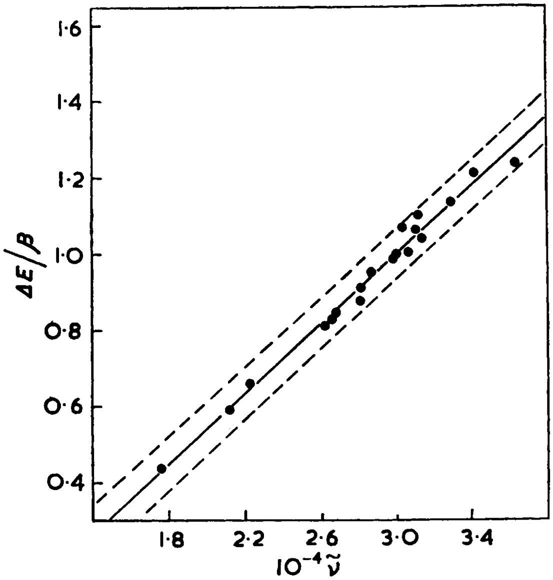


FIG. 1. Regression of $\Delta E/\beta$, calculated according to Huckel's method, on observed $\tilde{\nu}_{\text{corr.}}$ values of compounds nos. 1—19.

Regression: $\Delta E/\beta = -0{\cdot}3736 + 0{\cdot}4558 \times 10^{-4}\tilde{\nu}_{\text{corr.}}$.

Variance about the regression: $s^2 = 725{\cdot}4 \times 10^{-6}$ ($\phi = 17$).

Predicted $\tilde{\nu}_{\text{corr.}}$ values from: $10^{-4}\tilde{\nu}_{\text{corr.}} = 0{\cdot}8196 + 2{\cdot}1940\Delta E/\beta$.

Mean standard error of predicted values of $\tilde{\nu}_{\text{corr.}}$ of compounds nos. 1—19: $s = 0{\cdot}059 \times 10^4$ cm.$^{-1}$.

Confidence limits for predicted $\tilde{\nu}_{\text{corr.}}$ values at a 5% security level ($t = 2{\cdot}11$ for $\phi = 17$) are given by the broken lines (these limits are: $\pm 0{\cdot}128 \times 10^4$ cm.$^{-1}$ at $\overline{\Delta E/\beta} = 0{\cdot}933$, $\pm 0{\cdot}135 \times 10^4$ cm.$^{-1}$ at $\overline{\Delta E/\beta} \pm 0{\cdot}300$ and $\pm 0{\cdot}154 \times 10^4$ cm.$^{-1}$ at $\overline{\Delta E/\beta} \pm 0{\cdot}600$).

Compound no.	λ_1	λ_2	λ_3 (mμ)
5	359	344	329
14	385	367	349

Figs. 1 and 2 show the regression lines obtained for the Hückel and the Dewar calculations according to standard statistical procedures.[6] The linear regression line is the straight line obtained by the method of least squares. The scatter is measured by the variance of $\Delta E/\beta$ about the regression lines: Hückel $V(\Delta E/\beta) = 725 \times 10^{-6}$, Dewar

---

[6] Linder, " Statistische Methoden," Birkhäuser Verlag, Basle, 1957.

$V(\Delta E/\beta) = 1469 \times 10^{-6}$. The regression lines do not go through the origin, as the simple one-electron theories would predict. However, we are not in this paper testing the theoretical validity of these one-electron theories, which has been done many times before, but we are testing their usefulness for predicting spectra and this is increased if the line is not restricted to passing through the origin. Dewar does make this restriction, and this is a further reason why he finds the Hückel theory inferior to his own, since the intercept of the regression line on the y-axis is smaller for Dewar's calculation.

For any new molecule the Hückel or the Dewar energy can be calculated. Using the regression line we obtain the most probable value of the experimental frequency. We then require a measure of the uncertainty of this value. This is given by the confidence limits which are shown in Figs. 1 and 2, and which have been calculated according to Fieller's theorem [7] for a 5% security level. That is, there is a 19/20 probability that the true value lies within the limits shown by the broken lines, provided that the molecule belongs to the same class as those on which the regression line is based.

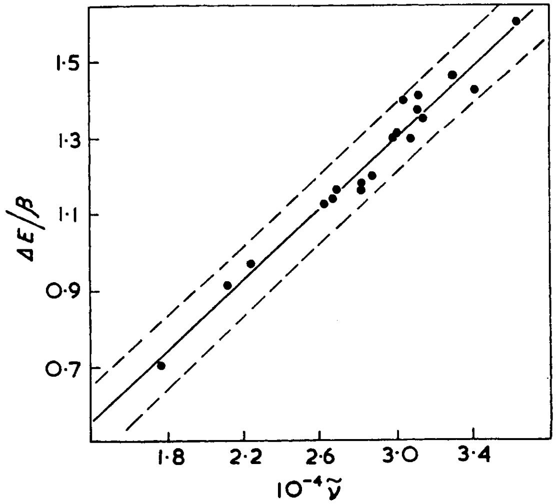


FIG. 2. Regression of $\Delta E/\beta$, calculated according to Dewar's method, on observed $\tilde{\nu}_{\text{corr.}}$ values of compounds nos. 1—19.

Regression: $\Delta E/\beta = -0{\cdot}0938 + 0{\cdot}4650 \times 10^{-4}\,\tilde{\nu}_{\text{corr.}}$.

Variance about the regression: $s^2 = 1469{\cdot}4 \times 10^{-6}$ ($\phi = 17$). Comparison of Hückel's and Dewar's methods: s^2 (Hückel)/s^2 (Dewar) $= 2{\cdot}03$. [For $P = 0{\cdot}05$, $F = 2{\cdot}3$; for $P = 0{\cdot}10$, $F = 1{\cdot}9$, with $\phi = 17$ (F single-tailed).]

Predicted $\tilde{\nu}_{\text{corr.}}$ values from: $10^{-4}\tilde{\nu}_{\text{corr.}} = 0{\cdot}2016 + 2{\cdot}1505\Delta E/\beta$.

Mean standard error of predicted values of $\tilde{\nu}_{\text{corr.}}$ of compounds no. 1—19: $s = 0{\cdot}082 \times 10^4$ cm.$^{-1}$.

Confidence limits for predicted $\tilde{\nu}_{\text{corr.}}$ values at a 5% security level ($t = 2{\cdot}11$ for $\phi = 17$) are given by the broken lines (these limits are: $\pm 0{\cdot}178 \times 10^4$ cm.$^{-1}$ at $\overline{\Delta E/\beta} = 1{\cdot}239$, $\pm 0{\cdot}188 \times 10^4$ cm.$^{-1}$ at $\overline{\Delta E/\beta} \pm 0{\cdot}300$, and $0{\cdot}221 \times 10^4$ cm.$^{-1}$ at $\overline{\Delta E/\beta}$ $\pm 0{\cdot}600$).

It appears from the Figures that the Huckel method is better for predicting spectra, and a statistical test shows that this is indeed so on a 10% significance level.

From Figs. 1 and 2 we have predicted the frequencies of the ^1L_a band for the cata-condensed hydrocarbons nos. 20—25 and for the pericondensed compounds nos. 26—43. Even though the pericondensed hydrocarbons are not of the same class as our calibration sample the results are still quite good. Pyrene (26) and 2,3-benzopyrene (28) show the greatest deviations from the predicted values of both the Hückel and the Dewar method.

---

[7] Fieller, *Quart. J. Pharm. Pharmacol.*, 1944, **17**, 117.

The standard errors for the predictions for compounds nos. 20—43 are bigger than those for the compounds nos. 1—19, but nevertheless the Hückel gives better results than the Dewar method.

It is hazardous to extrapolate the regression far beyond the range spanned by the compounds nos. 1—19. As we have pointed out, the regression has an intercept on the ν-axis. It is very likely that for large molecules the experimental frequency will not tend to zero with the Hückel energy, because of deviations from uniform bond length. This occurs, for example, with the linear polyenes. We have assumed that our approximate calculations differ from the exact calculation by random quantities. However, any systematic error, such as that introduced by an alternation in bond length, will be contained in the regression line. If we adopt Dewar's hypothesis that the regression line must go through the origin, then this also imposes certain restrictions on the nature of the uncontrolled error. The error must either tend to zero as we approach the origin, or it must become unsymmetrically distributed about the regression at this point. If one assumed the former, say that the variance is proportional to $\Delta E/\beta$, then the points can be fitted by a quadratic regression going through the origin. This is shown in Fig. 3 for the Hückel calculation. The exact treatment of such a heteroscedastic regression involves statistical difficulties, but a first approximation to the corresponding confidence limits can easily be obtained and these are shown as broken lines in the Figure. Inside the range covered by the compounds nos. 1—19 the results are very close to those obtained by the linear regression; however, outside the range things are very different owing to the different assumptions, and this again illustrates the danger of any extrapolation.

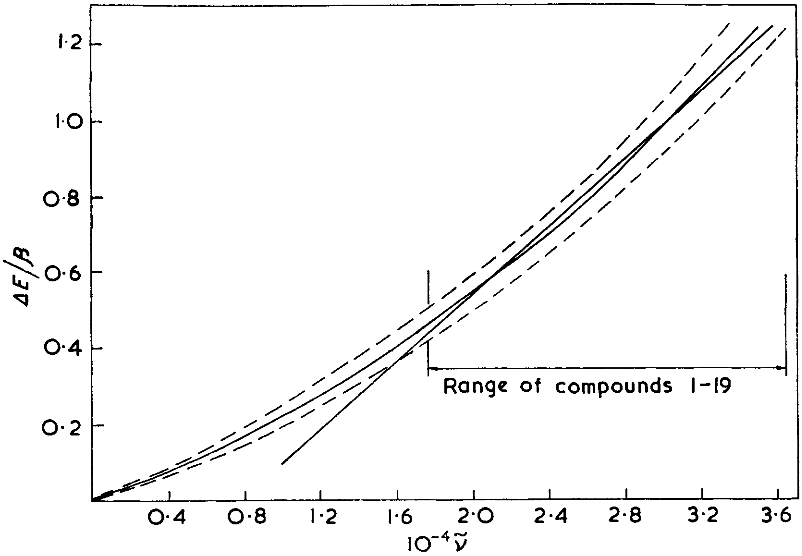


FIG. 3. Comparison of linear regression of $\Delta E/\beta$ calculated according to Hückel's method on observed $\tilde{\nu}_{\text{corr.}}$ values (Fig. 1) and the corresponding quadratic regression going through the origin.

Regression: $\Delta E/\beta = 0{\cdot}1616 \times 10^{-4}\tilde{\nu}_{\text{corr.}} + 0{\cdot}05586 \times 10^{-8} (\tilde{\nu}_{\text{corr.}})^2$.

The broken lines give a first approximation to the confidence limits for predicted $\tilde{\nu}_{\text{corr}}$ values at a 5% security level based on the above regression.

We thank the CIBA Aktiengesellschaft in Basle and the Schweizerische Nationalfonds for financial support.

ORGANISCH-CHEMISCHES LABORATORIUM DER EIDG. TECHNISCHEN HOCHSCHULE,
ZURICH, SWITZERLAND.
DEPARTMENT OF CHEMISTRY, THE UNIVERSITY, SHEFFIELD.
[*Received, October* 25*th*, 1961.]