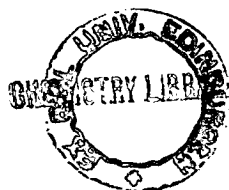


THE ELECTRONIC STRUCTURE OF SOME AROMATIC  
AND HETEROAROMATIC COMPOUNDS

by

Sheila M. F. Kennedy

Thesis presented for the Degree of Doctor of Philosophy



University of Edinburgh

1977

## Summary

Ab initio calculations using a linear combination of gaussian orbitals (LCGO) have been used to investigate the electronic ground state of several series of compounds.

The hydrocarbon series, benzene to pentacene with other related cyclic hydrocarbons, and the heterocyclic series, indole, benzofuran, and benzothiophen with the related hydrocarbons, indene, styrene, and naphthalene, and the corresponding quinoid series, isoindole, benzo(c)furan, benzo(c)thiophen, and 2H-indene have been investigated. The study has been extended to include some aza-, and diaza-analogues of the above heterocycles, and some heterocyclic N-oxides.

The LCGO calculations have been compared with Hückel and semi-empirical calculations where available, and in the case of the hydrocarbons have been used to determine values for the Hückel coulomb and resonance integrals. Calculated molecular orbital energy levels have been correlated within each series, and with experimental valence shell energy levels determined by photoelectron spectroscopy. Net atomic populations have been analysed in terms of bond moments, and dipole moments separated into sigma and pi components. The degree of aromaticity of these compounds has been estimated on the basis of their resonance energies with respect to that of benzene, and correlated with electronic structure.

A detailed analysis of the  $^1\text{H-N.M.R.}$  spectra of the mono-hetero compounds and some related compounds has been performed, and used to give additional information on aromatic character.

## CONTENTS

	Page
Chapter I: Theory	
Section 1: Quantum Theory	1
Section 2: Types of molecular orbital calculations	26
Section 3: Molecular and atomic properties	49
Section 4: Koopmans' theorem	52
Section 5: The ATMOL 2 programme	54
 Chapter II: Results and Discussion	
Section 1: Gaussian basis sets	61
Section 2: Molecular geometry	78
Section 3: Molecular energies	94
Section 4: The molecular orbital energy levels	119
Section 5: Photoelectron spectroscopy	132
Section 6: Dipole moments	170
Section 7: Population analyses and bond population moments	182
Section 8: The average position of $\pi$ -electrons	201
Section 9: Triplet states	206
Section 10: Hückel $\alpha$ and $\beta$ parameters	217
Section 11: Some heterocyclic N-oxides	226
Section 12: Correlation of XPS shifts with change in charge density, $\Delta q$ , as calculated by INDO	242
Section 13: Nuclear magnetic resonance spectroscopy	267
Section 14: Experimental methods	291
Appendix 1: Basis sets	299
Appendix 2: Molecular orbital energy levels	304
Appendix 3: Integral evaluation	335

Postgraduate courses attended and publications

CHAPTER I

SECTION I

Quantum Theory

## Early Quantum Theory

- "And God said, Let there be light, and there was light" -<sup>1</sup> whether or not light was created in this way, it was not until the early twentieth century that the physics of light, being studied from photoelectric experiments, brought a major breakthrough to quantum theory. Albert Einstein,<sup>2</sup> at this time, advanced the idea that energy was quantised, basing his postulation on the work of Max Planck,<sup>3</sup> who, early in 1900 announced a formula (1.1) for the energy density,  $\sigma$ , in black body radiation. For the first time, the experimental observations were completely represented mathematically.

$$\sigma dv = \frac{8\pi^2 h v^3}{c^3} \frac{1}{e^{hv/kT} - 1} dv \quad (1.1)$$

where  $v$  is the frequency,  $c$ , the speed of light,  $h$ , Planck's constant,  $k$ , the Boltzmann constant, and  $T$  the absolute temperature.

In deriving equations for black body radiation, the black body was considered as an empty cavity in equilibrium with its heated surroundings, containing radiation in the form of standing waves whose energies were represented by  $kT$ . This first led to Rayleigh-Jeans law for the energy density in the cavity (1.2).

$$\sigma dv = \frac{8\pi kT}{3c} v^2 dv \quad (1.2)$$

This failed to mimic the intensity distribution of the radiation and also implied infinite total energy in the cavity. This became known as the "ultraviolet catastrophe". Using Planck's formula, the energy density tended to zero at very high frequencies as was found experimentally. His assumption that the standing wave energy was equal to  $h\nu$ , was merely a tool to create the correct equation for the system. Einstein, however, quickly showed that the Planck relation,  $E = h\nu$ , implied that energy was quantised, and used it to explain the photoelectric effect on which Philipp Lenard, and R. A. Millikan had been experimenting. They

found that a metal plate with an applied potential under vacuum, and illuminated by a beam of monochromatic light, only emitted electrons when the frequency of the light was above a certain threshold value, and that the energy of the electron beam was proportional to the difference in the applied and threshold frequencies. The kinetic energy of the emitted electrons was found to be independent of the light intensity at any given frequency, and only the electron current varied in direct proportion to the intensity. Equation (1.3) describes the kinetic energy of the emitted electrons.

$$T = h ( \nu - \nu_0 ) + V \quad (1.3)$$

where  $\nu$  is the applied frequency,  $\nu_0$  the threshold frequency,  $V$ , the applied potential, and  $h$  the proportionality constant equivalent to that used by Planck in his formula for black body radiation density. These findings clearly fitted well with Einstein's theory of discrete, quantised energy, but it was not until Neils Bohr<sup>4</sup> announced his model for the hydrogen atom in 1913 that quantum theory began to be accepted. Bohr applied classical mechanics to a model of a hydrogen atom which consisted of a positive nucleus with an electron moving round it in a circular planar orbit. The angular momentum,  $p$ , of the electron was restricted to integral multiples,  $n$ , of Planck's constant (1.4).

$$p = n (h/2\pi) \quad (1.4)$$

He derived expressions for the energy (1.5) and radius(1.6) of the electron moving round the nucleus.

$$E = \frac{-2\pi^2 m e^4}{h^2} \cdot \frac{1}{n^2} \quad (1.5)$$

$$r = \frac{h^2}{4\pi^2 e^2 m} \cdot n^2 \quad (1.6)$$

Integer  $n$  is the principal quantum number, and for the hydrogen atom

$n = 1$ , giving the radius of the electron as (1.7).

$$r = \frac{h^2}{4\pi^2 m_e^2} \quad (1.7)$$

This value of  $r$  became known as the Bohr radius,  $a_0$ , and was set equal to one atomic unit, (1 a. u.).

Bohr assumed that energy was emitted or absorbed when an electron changed orbit, and by equating the energy change,  $\Delta E$ , to  $h\nu$ , he obtained equation (1.8).

$$\Delta E = h\nu = \frac{2\pi^2 m_e^4}{h^2} \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \quad (1.8)$$

where  $n_1$  and  $n_2$  are integers. By substituting  $\nu = c/\lambda$  in (1.8), (1.9) was obtained.

$$1/\lambda = \frac{2\pi^2 m_e^4}{h^3 c} \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right) = R \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \quad (1.9)$$

Equation (1.9) successfully predicted the atomic spectrum of hydrogen, the constant  $R$  being known as the Rydberg constant. It also accounted for other experimentally known series when  $n_1$  and  $n_2$  were allowed different integral values.

### The Schrödinger equation

Bohr's model of the hydrogen atom using classical mechanics was superseded in 1926 by Erwin Schrödinger's<sup>5</sup> wave equation, bringing quantum theory into the field of wave mechanics. In classical mechanics the energy of a system is given by (1.10).

$$E = T + V \quad (1.10)$$

Schrödinger suggested that the wave character of a system could be described by replacing  $T$  and  $V$  by linear operators to give a wave

equation (1.11).

$$(T + V)\Psi = E\Psi \quad (1.11)$$

where the solution to the wave function,  $\Psi$ , describes the spatial motion of all the particles in the system moving in the field of force specified by the potential energy operator,  $V$ .

The Schrödinger equation can be formulated in the following way.

Classically  $T = KE = \frac{1}{2}mv^2 = p^2/2m$  where  $p =$  linear momentum  $= mv$ . Each component of  $p$  in the  $x$ ,  $y$ , and  $z$  directions can be represented by an operator as follows.

$$p_x = \frac{h}{2\pi i} \frac{\partial}{\partial x} \quad p_y = \frac{h}{2\pi i} \frac{\partial}{\partial y} \quad p_z = \frac{h}{2\pi i} \frac{\partial}{\partial z} \quad \text{and } i = \sqrt{-1} \quad (1.12)$$

Consider a particle of mass,  $m$ , moving along the  $x$ -axis in a potential field  $V(x)$ ,

$$(T(x) + V(x))\psi(x) = E\psi(x) \quad (1.13)$$

$$T(x) = \frac{p_x^2}{2m} = -\frac{h^2}{8\pi^2 m} \frac{\partial^2}{\partial x^2} \quad (1.14)$$

The classical potential energy, and the quantum mechanical operator,  $V$ , are identical in this case giving (1.15).

$$\left[ -\frac{h^2}{8\pi^2 m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x) = E\psi(x) \quad (1.15)$$

Extending this to the general case for a single particle, (1.16) is obtained.

$$\left[ -\frac{h^2}{8\pi^2 m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V \right] \psi = E\psi \quad (1.16)$$



$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad - \text{ the Laplacian operator:}$$

$$\left[ -\frac{\hbar^2}{8\pi^2 m} \nabla^2 + V \right] \psi = E \psi \quad (1.17)$$

The linear operator  $(-\hbar^2/8\pi^2 m \cdot \nabla^2 + V)$  is known as the Hamiltonian operator, represented by H. If the Bohr radius,  $a_0$  (1.7) is used as the unit of length, (1.17) becomes:-

$$\left[ -\frac{1}{2} \nabla^2 + V \right] \psi = E \psi \quad (1.18)$$

Replacing  $[-\frac{1}{2} \nabla^2 + V]$  by H gives:-

$$H \psi = E \psi \quad (1.19) \text{ - the Schrödinger equation.}$$

For a system of particles (1.19) becomes:-

$$H \Psi = E \Psi \quad (1.20)$$

### Atomic orbitals

For any atom, the wavefunction  $\Psi$  in (1.20) will describe the nuclear and electronic motion. In order to simplify the situation the Born-Oppenheimer approximation is applied. This assumes that as the nuclear mass,  $m_n$ , is very much greater than the electronic mass,  $m_e$  at any one instant the nucleus is stationary with respect to electron movement. This separates off the nuclear kinetic energy in atoms, and also the nuclear repulsion terms in the case of molecules, from H, and leaves only those terms depending on the position of the nucleus, giving  $\Psi$  as an electronic wave function.

In the case of the hydrogen atom,  $\psi$  in (1.19) would represent a one electron wave function. These wave functions are called atomic orbitals, and to be physically meaningful must be single valued, continuous, and vanish at infinity. If  $\psi$  satisfies (1.19) then so does  $C\psi$ , where C is any number, and to fix the value of C a normalisation condition is

imposed (1. 21).

$$\int \psi^2 d\tau = 1 \quad \text{or} \quad \int \psi \psi^* d\tau = 1 \quad \text{if } \psi \text{ is complex} \quad (1. 21)$$

An additional property of the wavefunction is that if  $\psi_i$  and  $\psi_j$  are both solutions of (1. 19) then they are mutually orthogonal (1. 22).

$$\int \psi_i \psi_j d\tau = 0 \quad (1. 22)$$

These two conditions can be represented by a single equation (1. 23).

$$\int \psi_i \psi_j d\tau = \delta_{ij} \quad \text{- the Kronecker delta} \quad (1. 23)$$

$$\text{if } i = j \quad \delta_{ij} = 1; \quad \text{if } i \neq j \quad \delta_{ij} = 0$$

For the hydrogen atom various solutions of  $\psi$  correspond to discrete values of  $E$  which are determined by three integers or quantum numbers,  $n$ ,  $l$  and  $m$ . The wave function can be rewritten in terms of spherical polar coordinates, (1. 24) related to angles  $\Phi$  and  $\Theta$ , and to distance  $r$  of the electron from the nucleus, as shown in Fig. 1.

$$\psi_{nlm} = Y_{lm}(\Phi, \Theta) R_{nl}(r) \quad (1. 24)$$

where  $Y$  and  $R$  are angular and radial functions respectively.

The relationship between  $n$ ,  $l$ , and  $m$  is as follows:-

$n$ , principal quantum number	$1 \leq n < \infty$
$l$ , angular momentum quantum number	$0 \leq l \leq n-1$
$m$ , magnetic quantum number	$-l \leq m \leq l$

It is more usual to describe  $l$  values by letters, the following being the most frequently used,  $s$ ;  $p$ ;  $d$ ;  $f$ ; representing  $l = 0, 1, 2,$  and  $3$  respectively.

The angular and radial functions of the hydrogen like functions are shown in Tables 1 and 2. The radial part consists of a polynomial in  $r$  multiplied by an exponential term,  $\exp(-ar)$ , where  $a$  is known as the orbital exponent and is given by  $\frac{Z}{n}$ . Both  $Y$  and  $R$  must be normalised, and the overall

Table 1. Angular parts of the hydrogen like atomic orbitals

<u>l</u>	<u>m</u>	<u>orbital type</u>	<u>Y<sub>lm</sub></u>
0	0	s	$(\frac{1}{4\pi})^{\frac{1}{2}}$
1	1	p <sub>x</sub>	$(\frac{3}{4\pi})^{\frac{1}{2}} \sin \Theta \cos \Phi$
1	-1	p <sub>y</sub>	$(\frac{3}{4\pi})^{\frac{1}{2}} \sin \Theta \sin \Phi$
1	0	p <sub>z</sub>	$(\frac{3}{4\pi})^{\frac{1}{2}} \cos \Theta$
2	2	d <sub>xy</sub>	$(\frac{15}{16\pi})^{\frac{1}{2}} \sin^2 \Theta \sin 2\Phi$
2	-2	d <sub>x<sup>2</sup>-y<sup>2</sup></sub>	$(\frac{15}{16\pi})^{\frac{1}{2}} \sin^2 \Theta \cos 2\Phi$
2	1	d <sub>yz</sub>	$(\frac{15}{4\pi})^{\frac{1}{2}} \sin \Theta \cos \Theta \sin \Phi$
2	-1	d <sub>xz</sub>	$(\frac{15}{4\pi})^{\frac{1}{2}} \sin \Theta \cos \Theta \cos \Phi$
2	0	d <sub>z<sup>2</sup></sub> = d <sub>3z<sup>2</sup>-r<sup>2</sup></sub>	$(\frac{5}{16\pi})^{\frac{1}{2}} (3 \cos^2 \Theta - 1)$

Table 2. Radial parts of hydrogen like atomic orbitals

<u>n</u>	<u>l</u>	<u>orbital type</u>	<u>R<sub>nl</sub></u>
1	0	1s	$2a^{\frac{3}{2}} \exp(-ar)$
2	0	2s	$2a^{\frac{3}{2}} (1-ar) \exp(-ar)$
2	1	2p	$(\frac{4}{3})^{\frac{1}{2}} a^{\frac{5}{2}} r \exp(-ar)$
3	0	3s	$(\frac{2}{3}) a^{\frac{3}{2}} (3-6ar + 2a^2 r^2) \exp(-ar)$
3	1	3p	$(\frac{8}{9})^{\frac{1}{2}} a^{\frac{5}{2}} (2-ar) r \exp(-ar)$
3	2	3d	$(\frac{8}{45})^{\frac{1}{2}} a^{\frac{7}{2}} r^2 \exp(-ar)$

functions  $\psi_{nlm}$  are made orthogonal by the orthogonality of the functions in  $Y$ .

For the hydrogen atom the most probable value of  $r$  is when the electron density,  $\sigma$ , is at a maximum.

For  $\psi_{nlm} = Y_{lm}(\Phi, \Theta) R_{nl}(r)$

$$\begin{aligned} 4\pi r^2 \sigma &= \int_0^{\infty} 4\pi r^2 \psi^2 dr \\ &= 4\pi (Y_{lm}(\Phi, \Theta))^2 \int_0^{\infty} r^2 (R(r))^2 dr \end{aligned}$$

For a 1s orbital,  $R(r) = 2a^{-3/2} \exp(-ar)$

and for maximum  $\sigma$ ,  $\frac{d\sigma}{dr} = 0$

$$\therefore \frac{d\sigma}{dr} = \frac{d}{dr} (R(r))^2 r^2 = 0$$

$$\frac{d}{dr} (4a^3 e^{-2ar} \cdot r^2) = 0$$

$$8a^3 r e^{-2ar} - 8r^2 a^4 e^{-2ar} = 0$$

$$\div \text{ by } 8a^3 r e^{-2ar}$$

$$ar = 1$$

$$r = \frac{1}{a} = a_0$$

This is identical with the Bohr radius for the hydrogen atom (1.7), calculated on a classical mechanical basis.

The radial functions of the hydrogen-like orbitals, and their wave functions are shown diagrammatically in Figs 2 and 3. As can be seen from Fig. 2, the radial functions for  $n = l + 1$  eg 1s, 2p, always have the same sign. All other radial functions pass through the zero ordinate at least once. This crossing is called a node eg 2s, 3s, 3p all have radial nodes.

The Pauli exclusion principle<sup>6</sup> states that no two electrons in an atom can have the same set of quantum numbers, and as each atomic orbital accommodates two electrons, an additional quantum number,  $m_s$ , the

Fig. 4

The hydrogen atom,  $H$ .

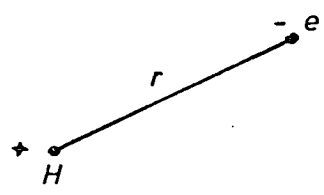


Fig. 5

The helium atom,  $He$ .

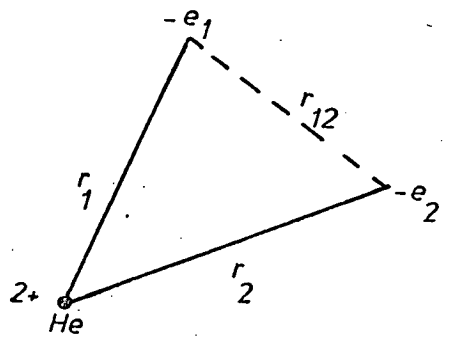
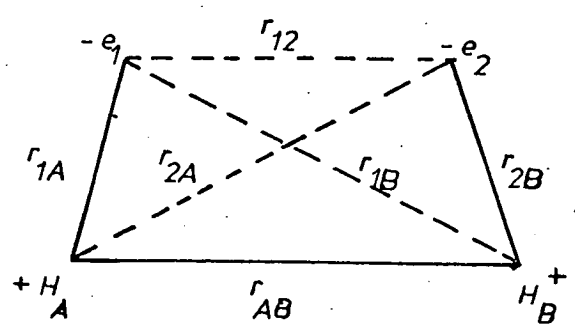


Fig. 6

The hydrogen molecule,  $H_A H_B$



spin quantum number is required to define each electron uniquely. The values of  $m_s$  are  $+\frac{1}{2}$  and  $-\frac{1}{2}$ , and are often referred to as  $\alpha$  and  $\beta$  spins. Each electron in an atom can now have a unique set of quantum numbers.

### Polyelectronic systems and the Slater Determinant

So far, only the hydrogen atom, mono electronic system, represented in Fig. 4 has been considered. This gives (1.18) as:-

$$-\left[\frac{1}{2}\nabla^2 + \frac{1}{r}\right]\Psi = E\Psi \quad (1.25)$$

For the helium atom, Fig. 5, containing two electrons, several other terms must be included in H.

$$\left[-\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}\right]\Psi = E\Psi \quad (1.26)$$

where  $-\frac{1}{2}\nabla_1^2$  = kinetic energy of electron 1

$-\frac{1}{2}\nabla_2^2$  = kinetic energy of electron 2

$\frac{2}{r_1}$  = coulomb attraction between the nucleus and electron 1

$\frac{2}{r_2}$  = coulomb attraction between the nucleus and electron 2

$\frac{1}{r_{12}}$  = coulombic repulsion between electrons 1 and 2.

Depending on the atom or molecule under consideration,  $\Psi$  will represent a varying number of electrons. Building up those electrons in hydrogen-like atomic orbitals will give the electronic structure. Taking the orbital energies as negative in the binding state, the lowest energy orbital i.e. greatest binding energy, is filled first, and as the quantum number increases, the binding energy decreases i.e. becomes less negative. The lowest energy orbital will then be with  $n = 1$  and  $l = 0$ , the 1s orbital, and will contain two electrons of opposite spin. The helium atom containing

two electrons, (1) and (2), will then be:-

$$\begin{aligned}\text{He} &= 1s(1) \alpha(1) 1s(2) \beta(2) \\ &= 1s(1) 1\bar{s}(2)\end{aligned}$$

This implies that  $\Psi$  in (1.26) could be represented by a product of two one-electron wave functions.

$$= \Phi_{1s}(1) \bar{\Phi}_{1s}(2) \quad (1.27)$$

However, the Pauli Principle states that the total wave function must be antisymmetric with respect to electron permutation. In (1.27), swapping electrons (1) and (2) would create a new wave function, and not the negative of the original. The problem is overcome by taking linear combinations of  $\Phi(1)$  and  $\Phi(2)$ , to give two equivalent combinations  $\Psi_s$  and  $\Psi_a$  which will be symmetric and antisymmetric respectively.

$$\Psi_s = \Phi_{1s}(1) \bar{\Phi}_{1s}(2) + \Phi_{1s}(2) \bar{\Phi}_{1s}(1) \quad (1.28)$$

$$\Psi_a = \Phi_{1s}(1) \bar{\Phi}_{1s}(2) - \Phi_{1s}(2) \bar{\Phi}_{1s}(1) \quad (1.29)$$

The Pauli principle excludes (1.28), and so the wave function for the helium atom when normalised becomes:-

$$\Psi_{\text{He}} = \frac{1}{\sqrt{2}} [\Phi_{1s}(1) \bar{\Phi}_{1s}(2) - \Phi_{1s}(2) \bar{\Phi}_{1s}(1)] \quad (1.30)$$

where  $\frac{1}{\sqrt{2}}$  is the normalising factor.

This can be represented in determinantal form by the Salter<sup>7</sup> determinant (1.31), the general form of which is shown in (1.32) for N electrons in N spin orbitals.

$$\Psi_{\text{He}} = \frac{1}{\sqrt{2}} \begin{vmatrix} \Phi_{1s}(1) & \bar{\Phi}_{1s}(1) \\ \Phi_{1s}(2) & \bar{\Phi}_{1s}(2) \end{vmatrix} \quad (1.31)$$

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \Phi_1(1) & \Phi_2(1) & \Phi_3(1) & \dots & \Phi_N(1) \\ \Phi_1(2) & \Phi_2(2) & \Phi_3(2) & \dots & \Phi_N(2) \\ \Phi_1(3) & & & & \\ \vdots & & & & \\ \Phi(N) & \dots & \dots & \dots & \Phi_N(N) \end{vmatrix} \quad (1.32)$$

The interchange of any pair of electrons corresponds to the interchange of two rows in the determinant, and has the effect of changing the sign but not the magnitude of the determinant. This satisfies the anti-symmetry condition of the wave function. If any two electrons have the same spin and the same orbital,  $\Phi_i$ , the determinant will have two identical columns, and thus vanish. The Pauli exclusion principle is therefore also satisfied.

#### The Variation Method and the LCAO procedure

The Schrödinger equation can only be solved exactly for one electron systems, and for many electron systems an alternative approach known as the variation method is used to give approximate solutions. If both sides of (1.20) are multiplied by  $\Psi$  and integrated over all variables, the equation can be rewritten as

$$\begin{aligned} E &= \frac{\int \Psi H \Psi \, d\tau}{\int \Psi \Psi \, d\tau} \\ &= \int \Psi H \Psi \, d\tau \quad \text{if } \Psi \text{ is normalised} \\ &= \langle \Psi H \Psi \rangle \end{aligned} \quad (1.33)$$

If a trial wave function  $\Psi_T$  is used in (1.33), an energy  $E_T$  can be evaluated. Several trial wave functions can be used until a minimum energy,  $E_{T\min}$ , is found. The true energy minimum,  $E_{\min}$  corresponding



to an exact solution of (1.33) will always be less than  $E_{Tmin}$ , but the nearer  $\Psi_T$  approaches to the exact wave function, the closer will  $E_{Tmin}$  be to  $E_{min}$ . Normally a wave function will be dependent on a number of variable parameters, for example, a finite set of linear parameters,  $c_1, c_2, \dots$ , and the energy will also be a function of these parameters. Since the Schrödinger equation gives stationary values of the energy, then mathematically, the energy is at a minimum when  $\delta E = 0$ .

$$\delta E = \delta \langle \Psi | H | \Psi \rangle = 0 \quad (1.34)$$

This corresponds to

$$\delta E (c_1, c_2, \dots) = \frac{\partial E}{\partial c_1} \delta c_1 + \frac{\partial E}{\partial c_2} \delta c_2 + \dots = 0 \quad (1.35)$$

Each term in (1.35) is zero and since  $\delta c_i \neq 0$ , then:-

$$\frac{\delta E}{\delta c_i} = 0 \quad (1.35a)$$

The energies of the stationary states can be approximated by solving (1.35a). The flexibility of the wave function will increase with increase in the number of variable parameters, and the solutions of  $E$  will be closer to  $E_{min}$ .

If  $\Psi$  can be written as a linear combination of fixed functions,  $x_1, x_2, \dots$  known as basis functions, then the variation theorem leads directly to a set of approximate energies and wave functions. The most common application is when  $x$  represents an atomic orbital, and the process is known as the linear combination of atomic orbitals (LCAO) procedure. This is widely used in molecular orbital calculations.

$$\Psi(c_1, c_2, \dots) = c_1 x_1 + c_2 x_2 + \dots \quad (1.36)$$

Substituting (1.36) in (1.20) gives

$$H \sum_i c_i x_i = E \sum_i c_i x_i \quad (1.37)$$

Multiplying both sides by  $\Psi_j$  where  $\Psi_j = \sum_j c_j x_j$ , and integrating gives

$$\left( \sum_j c_j x_j \right) H \left( \sum_i c_i x_i \right) = E \left( \sum_j c_j x_j \right) \left( \sum_i c_i x_i \right) \quad (1.38)$$

The general term in (1.38) is given by

$$c_i c_j (x_j H x_i - E x_j x_i) = 0 \quad (1.38a)$$

writing  $H_{ij} = \int x_i H x_j d\tau$ , and  $S_{ij}$  the overlap integral  $= \int x_i x_j d\tau$ , and using  $H_{ij} = H_{ji}$  and  $S_{ij} = S_{ji}$ , application of the variation method gives:-

$$\frac{\delta E}{\delta c_j} = c_i (H_{ij} - E S_{ij}) = 0 \quad (1.38b)$$

Applying the variation method to every term in (1.38) gives the general form of (1.38b) as

$$\sum_i c_i (H_{ij} - E S_{ij}) = 0 \quad (1.39)$$

These equations are known as the secular equations, and to equal zero simultaneously, the determinant multiplying the coefficients must vanish.

$$\left| H_{ij} - E S_{ij} \right| = 0 \quad (1.40)$$

Equation (1.40) is used in one form or another as a basis for molecular orbital calculations.

### The Hartree-Fock Method

Considering the hydrogen molecule shown in Fig. 6, (1.20) can be written:

$$\left[ -\frac{1}{2} \nabla_1^2 - \frac{1}{2} \nabla_2^2 - \frac{1}{r_{1A}} - \frac{1}{r_{2A}} - \frac{1}{r_{1B}} - \frac{1}{r_{2B}} + \frac{1}{r_{12}} \right] \Psi = E \Psi \quad (1.41)$$

Substituting  $H_i = (-\frac{1}{2}\nabla_i^2 - \frac{1}{r_{iA}} - \frac{1}{r_{iB}})$ , a one electron Hamiltonian, gives

$$[H_1 + H_2 + \frac{1}{r_{12}}] \Psi = E \Psi \quad (1.42)$$

Using (1.32),  $\Psi$  can be written as

$$\Psi = \frac{1}{\sqrt{2}} (1s(1) \bar{1}s(2) - \bar{1}s(1) 1s(2)) \quad (1.43)$$

assuming that the two electrons will be paired in the lowest energy molecular orbital, the form of which is given by  $\Psi$ , and naming this orbital the  $1s$  orbital.

$$\begin{aligned} \text{Using (1.34)} \quad E &= \langle \Psi | H_1 + H_2 + \frac{1}{r_{12}} | \Psi \rangle \\ &= \frac{1}{2} \langle 1s(1) \bar{1}s(2) - \bar{1}s(1) 1s(2) | H_1 + H_2 + \frac{1}{r_{12}} | 1s(1) \bar{1}s(2) - \\ &\quad \bar{1}s(1) 1s(2) \rangle \quad (1.44) \end{aligned}$$

Taking each term individually and separating into spin and spacial parts:-

$$\begin{aligned} \text{a) } &\langle 1s(1) \bar{1}s(2) | H_1 | \bar{1}s(1) 1s(2) \rangle \\ &= \int 1s(1) H_1 1s(1) dv_1 \int \alpha(1) \alpha(1) ds_1 \int 1s(2) 1s(2) dv_2 \\ &\quad \int \beta(2) \beta(2) ds_2 \\ &= \int 1s(1) H_1 1s(1) dv_1 = H_{1s} \end{aligned}$$

Since  $H_1$  operates only on electron (1) and as  $\alpha$  and  $\beta$  are orthonormal.

$$\int \alpha(1) \beta(1) ds_1 = 0 \quad \text{and} \quad \int \alpha(1) \alpha(1) ds_1 = 1$$

There will be four such terms ( $4 \times H_{1s}$ ), and all off diagonal terms will be zero as they involve both  $\alpha$  and  $\beta$  spins of the same electron.

$$\begin{aligned} \text{b) } &\langle 1s(1) \bar{1}s(2) | \frac{1}{r_{12}} | \bar{1}s(1) 1s(2) \rangle \\ &= \iint 1s(1) 1s(1) \frac{1}{r_{12}} 1s(2) 1s(2) dv_1 dv_2 \\ &\quad \int \alpha(1) \alpha(1) ds_1 \int \beta(2) \beta(2) ds_2 \end{aligned}$$

$$\begin{aligned}
 &= \iint 1s(1) 1s(1) \frac{1}{r_{12}} 1s(2) 1s(2) dv_1 dv_2 \\
 &= J_{1s 1s}
 \end{aligned}$$

There will be two terms of this type.

$J$  is called the coulomb integral, and represents the electron repulsion between the charge clouds of two electrons at a distance,  $\frac{1}{r_{12}}$ , from each other. This reduces (1.42) to

$$\begin{aligned}
 E &= \frac{1}{2} (4 \times H_{1s} + 2J_{1s 1s}) \\
 &= 2H_{1s} + J_{1s 1s} \quad (1.45)
 \end{aligned}$$

For a hydrogen molecule triplet state, for example, with one electron raised to the 2s orbital, and both electrons having the same spin. The molecular wavefunction will now be

$$\Psi = \frac{1}{\sqrt{2}} (1s(1) 2s(2) - 1s(2) 2s(1)) \quad (1.46)$$

The cross terms of the electron repulsion no longer disappear.

$$\begin{aligned}
 &\langle (1s(1) 2s(2) | \frac{1}{r_{12}} | -(1s(2) 2s(1)) \rangle \\
 &= - \iint 1s(1) 1s(2) \frac{1}{r_{12}} 2s(1) 2s(2) dv_1 dv_2 \int a(1) a(1) ds_1 \\
 &\quad \int a(2) a(2) ds_2 \\
 &= - \iint 1s(1) 1s(2) \frac{1}{r_{12}} 2s(1) 2s(2) dv_1 dv_2 \\
 &= -K_{1s 2s}
 \end{aligned}$$

$K$  is known as the exchange integral, and allows for the fact that the electron repulsion between electrons of parallel spin will be less than for electrons with paired spins due to the reduced probability of parallel spin electrons ever being close together.

The electronic energy of the hydrogen molecule triplet state is then given by

$$E = H_{1s} + H_{2s} + J_{1s 2s} - K_{1s 2s} \quad (1.47)$$

where  $H_{1s}$  and  $H_{2s}$  represent the energy of the separate electrons in the field of the bare nuclei. The generalised form of (1.47) is given by

$$\begin{aligned}
 E &= 2 \sum_i^n H_{ii} + \sum_i^n J_{ii} + \sum_i^n \sum_{j(i \neq j)}^n (2J_{ij} - K_{ij}) \\
 &= 2 \sum_i^n H_{ii} + \sum_i^n \sum_j^n (2J_{ij} - K_{ij}) \quad \text{since } J_{ii} = K_{ii} \quad (1.48)
 \end{aligned}$$

The total energy of the molecule is therefore twice the sum of the one electron energies over occupied molecular orbitals assuming doubly occupied orbitals, plus a coulomb repulsion integral for every pair of electrons, plus an exchange integral for every pair of electrons with the same spin. The expression for the energy has now been reduced to a set of three and six dimensional integrals, and forms the basis for the Hartree-Fock equations which were derived by Fock<sup>8</sup> based on earlier work by Hartree.<sup>9</sup>

Using (1.48), the energy,  $e_i$ , of a single electron in an orbital  $\psi_i$  can be written as (1.48a).

$$e_i = H_{ii} + \sum_j^n (2J_{ij} - K_{ij}) \quad (1.48a)$$

The quantity  $e_i$  is called the Hartree-Fock energy of the electron,  $i$ , in the orbital  $\psi_i$ .

Using (1.48) and (1.48a), the total electronic energy can be written as (1.48b).

$$\begin{aligned}
 E &= 2 \sum_i^n e_i - \sum_i^n \sum_j^n (2J_{ij} - K_{ij}) \\
 &= \sum_i^n (e_i + H_{ii}) \quad (1.48b)
 \end{aligned}$$

Using the variation method the best set of molecular orbitals, giving a minimum value for  $E$ , will be obtained by varying the one electron functions in the determinant representing the many electron wave function for the molecule. These molecular orbitals are known as self consistent or Hartree-Fock orbitals. As before, the stationary state corresponding

to the energy minimum must be found. The orbitals are restricted to being orthonormal throughout (1.23). The calculus of variations is used for this kind of constrained variational problem, and involves the minimisation of the function,  $G$ , given by

$$G = E - 2 \sum_i \sum_j e_{ij} S_{ij} \quad (1.49)$$

Using (1.48) this becomes

$$G = 2 \sum_i H_{ii} + \sum_i \sum_j (2J_{ij} - K_{ij}) - 2 \sum_i \sum_j e_{ij} S_{ij} \quad (1.50)$$

For minimum  $G$ ,  $\delta G = 0$ , and corresponds to a small change  $\delta\psi_i$  in  $\psi_i$ .

$$\delta G = 2 \sum_i \delta H_{ii} + \sum_i \sum_j (2\delta J_{ij} - \delta K_{ij}) - 2 \sum_i \sum_j e_{ij} \delta S_{ij} = 0 \quad (1.51)$$

where  $\delta H_{ii} = \int \delta\psi_i^*(1) \underline{H}(1) \psi_i(1) d\tau_1 + \text{complex conjugate} \quad (1.52)$

$$\delta J_{ij} = \int \delta\psi_i^*(1) \underline{J}_j(1) \psi_i(1) d\tau_1 + \int \delta\psi_j^*(1) \underline{J}_i(1) \psi_j(1) d\tau_1 + \text{complex conjugate} \quad (1.53)$$

$$\delta K_{ij} = \int \delta\psi_i^*(1) \underline{K}_j(1) \psi_i(1) d\tau_1 + \int \delta\psi_j^*(1) \underline{K}_i(1) \psi_j(1) d\tau_1 + \text{complex conjugate} \quad (1.54)$$

$$\delta S_{ij} = \int \delta\psi_i^*(1) \psi_j(1) d\tau_1 + \text{complex conjugate} \quad (1.55)$$

Here  $\underline{J}$  and  $\underline{K}$  are operators, defined as

$$\underline{J}_j(1) = \int \psi_j^*(2) \frac{1}{r_{12}} \psi_j(2) d\tau_2 \quad (1.56)$$

$$\underline{K}_j(1) \psi_i(1) = \left[ \int \psi_j^*(2) \frac{1}{r_{12}} \psi_i(2) d\tau_2 \right] \psi_j(1) \quad (1.57)$$

The complex conjugate parts in (1.52 - 55) can be ignored as the orbitals and their complex conjugates can be varied independently. The problem then deals with real functions, and real variations, and (1.51) becomes

$$\delta G = 2 \sum_i \int \delta \psi_i (1) (\underline{H} \psi_i + \sum_j (2\underline{J}_j - \underline{K}_j) \psi_i - \sum_j e_{ij} \psi_j) d\tau = 0 \quad (1.58)$$

For (1.58) to be zero  $(\underline{H} \psi_i + \sum_j (2\underline{J}_j - \underline{K}_j) \psi_i - \sum_j e_{ij} \psi_j)$  must be zero for every value of  $i$  giving

$$[\underline{H} + \sum_j (2\underline{J}_j - \underline{K}_j)] \psi_i = \sum_j e_{ij} \psi_j \quad (1.59)$$

This is often written

$$F \psi_i = \sum_j e_{ij} \psi_j \quad (1.60)$$

where  $F$  is the Fock operator, and represents a one electron Hamiltonian for an electron moving in a molecular environment.

In the form of (1.60) there exists for  $e_{ij}$  a complete set of values which is undesirable if the molecular orbitals are to be fixed uniquely, and arises from the property of determinants that any unitary transformation of the elements does not alter the value of the determinant. This means that the set  $\psi_i$  can be replaced by a new set  $\psi'_i$  where

$$\psi'_i = \sum_j T_{ij} \psi_j \quad (1.61)$$

provided that  $\sum_k T_{ik}^* T_{kj} = \delta_{ij}$  (1.62)

Substituting (1.61) in (1.60) merely replaces  $e_{ij}$  by a new set of constants  $e_{kl}$  where

$$e_{kl} = \sum_{ij} T_{kj}^* e_{ij} T_{jl} \quad (1.63)$$

However, as the  $e_{ij}$  form a hermitian matrix, a unitary transformation bringing the matrix to diagonal form such that  $e_{ij} = 0$ , unless  $i = j$ , can be performed giving (1.60) in the form of a standard eigenvalue equation:

$$F\psi_i = e_i\psi_i \quad (1.64)$$

These are known as the Hartree-Fock equations, and the resulting wave function is known as a restricted Hartree-Fock function as each molecular orbital contains two electrons.

### The Roothaan method for closed shell systems

Direct solution of the Hartree-Fock equations is difficult for molecules, and even for atoms it must be done by numerical integration, making it impossible to calculate properties which require the orbitals to be represented as analytical functions of the coordinates. It is necessary to find the best approximation to the Hartree-Fock orbitals in which the orbitals appear as analytical functions. The best method to date was introduced by Roothaan,<sup>10</sup> and involves representing each molecular orbital as a linear combination of atomic orbitals (1.36). Subscripts  $p, q, r, s$ , will be used for atomic orbitals.

$$\psi_i = \sum_p C_{pi} \phi_p \quad (1.65)$$

For the molecular orbitals,  $\psi_i$ , to form an orthonormal set, the number of basis atomic orbitals must be greater than or equal to the number of occupied orbitals. In the LCAO approximation this corresponds to

$$\sum_{pq} C_{pi} C_{qj} S_{pq} = \delta_{ij} \quad (1.66)$$

where  $S_{pq} = \int \phi_p(1) \phi_q(1) d\tau_1$  - the overlap integral

$$\text{between atomic functions} \quad (1.66a)$$



Writing  $H_{ii}$ ,  $J_{ij}$  and  $K_{ij}$  in terms of integrals over atomic orbitals, and substituting into (1.48) gives the total energy as

$$E = \sum_{pq} P_{pq} H_{pq} + \frac{1}{2} \sum_{pqrs} P_{pq} P_{rs} [(pq|rs) - \frac{1}{2}(pr|qs)] \quad (1.67)$$

$$\text{where } H_{ii} = \sum_{pq} C_{pi} C_{qi} H_{pq} \quad (1.68)$$

$$J_{ij} = \sum_{pqrs} C_{pi} C_{qj} C_{ri} C_{sj} (pq|rs) \quad (1.69)$$

$$K_{ij} = \sum_{pqrs} C_{pi} C_{qj} C_{ri} C_{sj} (pr|qs) \quad (1.70)$$

$$P_{pq} = 2 \sum_i^{\text{occ}} C_{pi} C_{qi} \quad (1.71)$$

where occ is the number of occupied orbitals, and

$$(pq|rs) = \iint \phi_p(1) \phi_q(1) \frac{1}{r_{12}} \phi_r(2) \phi_s(2) d\tau_1 d\tau_2$$

This is the general form of the two electron interaction integral expressed over atomic orbitals.

Applying the Hartree-Fock procedure, the small variation in  $\psi_i$  is given as:-

$$\delta \psi_i = \sum_p \delta C_{pi} \phi_p \quad (1.72)$$

and the stationary point of function  $G$  (1.50) becomes:-

$$\delta G = 2 \sum_i^{\text{occ}} \delta C_{pi} C_{qi} H_{pq} + \sum_{ij}^{\text{occ}} \sum_{pqrs} (\delta C_{pi} C_{qj} C_{ri} C_{sj} + C_{pi} \delta C_{qj} C_{ri} C_{sj}) \cdot$$

$$[2(pq|rs) + (pr|qs)] - 2 \sum_{ij} \sum_{pq} e_{ij} \delta C_{pi} C_{qj} S_{pq} \quad (1.73)$$

$$= 0$$

Again as  $\delta C_{pi}$  is arbitrary, (1.73) is zero if

$$\begin{aligned} \sum_q (C_{qi} H_{pq} + \sum_j^{\text{occ}} \sum_{qrs} C_{qj} C_{ri} C_{sj} (2(pq|rs) + (pr|qs))) \\ = \sum_j e_{ij} \sum_q C_{qj} S_{pq} \end{aligned} \quad (1.74)$$

Applying a unitary transformation of the form (1.61) gives (1.74) as

$$\sum_q (F_{pq} - e_i S_{pq}) C_{qi} = 0 \quad (1.75)$$

where the matrix elements of  $F$  are given by

$$F_{pq} = H_{pq} + \sum_{rs} p_{rs} [(pq|rs) - \frac{1}{2} (pr|qs)] \quad (1.76)$$

Equation (1.75) corresponds to the secular equations of (1.39) and again, for a set of non-trivial solutions

$$|F_{pq} - e_i S_{pq}| = 0 \quad (1.77)$$

The equations derived by the Roothaan method for the LCAO self consistent field molecular orbitals (LCAOSCFMO) are algebraic as distinct from the Hartree-Fock equations which were differential. These new equations were devised by Roothaan<sup>10</sup> and Hall<sup>11</sup> independently, and were first published in the same year.

### The Self Consistent Field

The variation method applied to the LCAO molecular orbitals in the Roothaan equations forms the basis of the self consistent field process. Initially a set of trial coefficients  $C_{pi}$  is used to generate a density

matrix, the elements of which are related to the elements of the Fock Hamiltonian by (1.76). This allows the Fock matrix to be set up, and using matrix algebra, a matrix  $E$ , the elements of which,  $e_{ij}$ , will be roots of (1.77), is obtained. This in turn from (1.75) produces a matrix  $C$ , of linear expansion coefficients  $C_{pi}$ . The process is then repeated using the new set of linear coefficients held in  $C$ , and successive iterations performed until the coefficients no longer alter within a set limit. This iterative process is known as the self consistent field method (SCF), and when applied to LCAO is known as the LCAOSCF method.

Open shell SCF methodsRoothaan's single Hamiltonian method.

For open shell molecules the Hartree-Fock equations (1.60) cannot necessarily be simplified by a unitary transformation, and off diagonal elements,  $e_{ij}$ , will remain. Roothaan<sup>12</sup> has modified the SCF equations to absorb the off diagonal elements by use of an effective Hamiltonian to obtain exact solutions. For a general open shell case containing doubly (c) and singly (o) occupied orbitals the energy will be given by

$$E = 2 \sum_k H_k + \sum_{kl} (2J_{kl} - K_{kl}) + f [2 \sum_m H_m + f \sum_{mn} (2aJ_{mn} - bK_{mn})] + 2 \sum_{km} (2J_{km} - K_{km}) \quad (1.78)$$

where  $k, l$  refer to closed shells,  $m, n$  to open shell,  $f$  is the fractional occupation, defined as the number of occupied spin orbitals outside the doubly occupied energy levels divided by the number of available open shell orbitals, and  $a, b$  are numerical constants depending on the coulomb and exchange interactions between unpaired electrons, and are determined by taking account of symmetry degeneracies. Using the variation method, and applying the Hartree-Fock procedure, the SCF equations obtained from (1.78) can be reduced to:-

$$F_c \phi_c = \eta_c \phi_c \quad \text{and} \quad F_o \phi_o = \eta_o \phi_o \quad (1.79)$$

In addition to the one electron operators and coulomb and exchange operators, the operators  $F_c$  and  $F_o$  involve expressions incorporating  $a, b$  and  $f$ . The solutions of the eigenvalue problem now obtained will produce a set of orbital energies for the doubly and singly occupied orbitals. This method is widely used for many open shell problems in the LCAO form. Convergence may be time consuming, but techniques exist to improve convergence rates, and these can be used in conjunction with this open shell method.

Nesbet's Method:- spin equivalence restrictions

Nesbet's<sup>13</sup> method also involves the use of an effective Hamiltonian, but the process is simplified by using only a single Hamiltonian to represent both the open and closed shells by restricting all orbitals of the same sub-shell to having the same radial part. This means that the spacial function can be combined with either the  $\alpha$  or  $\beta$  spin function to form two degenerate orbitals  $\psi$  and  $\psi'$ . This process is known as a spin equivalence restriction method. The total wavefunction is then built up from orthonormal doubly and singly occupied orbitals, and can easily be made to satisfy the necessary spin and symmetry requirements. An arbitrary averaging procedure is used to set up the effective Hamiltonian. For example for helium hydride,  $\text{He H}, (1s^2 2s^1)$  where superscripts represent the orbital occupancy, the effective Hamiltonian could be arbitrarily written as:-

$$[H + 2J_{1s} - K_{1s} + J_{2s} - K_{2s}] \quad (1.80)$$

This gives

$$\begin{aligned} e_{1s}^{\text{SCF}} &= e_{1s} + 2J_{1s1s} - K_{1s1s} + J_{2s1s} - K_{2s1s} \\ &= e_{1s} + J_{1s1s} + J_{1s2s} - K_{1s2s} \end{aligned} \quad (1.81)$$

$$\begin{aligned} e_{2s}^{\text{SCF}} &= e_{2s} + 2J_{1s2s} - K_{1s2s} + J_{2s2s} - K_{2s2s} \\ &= e_{2s} + 2J_{1s2s} - K_{1s2s} \end{aligned} \quad (1.82)$$

Using (1.48) the total electronic energy is given by

$$E = 2e_{1s} + e_{2s} + J_{1s1s} + 2J_{1s2s} - K_{1s2s} \quad (1.83)$$

$$\begin{aligned} \therefore (e_{1s}^{\text{SCF}} + e_{1s}) + \frac{1}{2}(e_{2s}^{\text{SCF}} + e_{2s}) &= 2e_{1s} + J_{1s1s} + J_{1s2s} - K_{1s2s} + \\ &\quad e_{2s} + J_{1s2s} - \frac{1}{2}K_{1s2s} \\ &= E - \frac{1}{2}K_{1s2s} \end{aligned} \quad (1.84)$$

The converged SCF energy will therefore differ from the Hartree-Fock energy in the general case, by an amount associated with the exchange integral terms involving the open shell electrons. In most cases the difference is small, but in all cases the wave function is sufficiently accurate to use to calculate the correction to the SCF energy required to give the energy value closer to the Hartree-Fock energy.

The accuracy of this method is dependent upon the averaging process used for the Hamiltonian, and it is often not clear as to what the best choice is, especially if the molecular wave function is not a single determinant.

### The use of Unrestricted Hartree-Fock orbitals in open shell cases

So far all methods have related to restricted Hartree-Fock orbitals in which two electrons of different spin have the same orbital function. In an unrestricted Hartree-Fock function, each spin orbital has a different orbital function as generalised in (1.32).

If a molecule has  $y$   $\alpha$ -electrons and  $z$   $\beta$ -electrons, two completely independent sets of molecular orbitals can be assigned to it, representing  $\psi^\alpha$  and  $\psi^\beta$ . As a restricted function is merely a special case of the unrestricted function, the latter leads in general to lower total energies. The SCF approach to unrestricted orbitals was first conceived by Slater,<sup>14</sup> and later expanded by Pople and Nesbet.<sup>15</sup> The theory follows the Hartree-Fock, and Roothaan procedures closely, and gives the total electronic energy as:-

$$E = \sum_i^y H_{ii} + \frac{1}{2} \left( \sum_i^{y+z} \sum_j^{y+z} J_{ij} - \sum_i^y \sum_j^y K_{ij}^\alpha - \sum_i^z \sum_j^z K_{ij}^\beta \right) \quad (1.85)$$

In the LCAO procedure  $\psi^\alpha$  and  $\psi^\beta$  are written as linear combinations of atomic orbitals.

$$\psi_i^\alpha = \sum_p C_{pi}^\alpha \phi_p \quad (1.86)$$

$$\psi_i^\beta = \sum_p C_{pi}^\beta \phi_p \quad (1.87)$$

This produces sets of  $\alpha$  and  $\beta$  density matrices as in (1.71).

$$P_{pq}^{\alpha} = \sum_i^y C_{pi}^{\alpha} C_{qi}^{\alpha} \quad (1.88) \quad P_{pq}^{\beta} = \sum_i^z C_{pi}^{\beta} C_{qi}^{\beta} \quad (1.89)$$

and the total density matrix is the sum of (1.88) and (1.89).

$$P_{pq} = P_{pq}^{\alpha} + P_{pq}^{\beta} \quad (1.90)$$

Substituting (1.86-87) into (1.85) gives

$$E = \sum_{pq} P_{pq} H_{pq} + \frac{1}{2} \sum_{pqrs} (P_{pq} P_{rs} - P_{pr}^{\alpha} P_{qs}^{\alpha} - P_{pr}^{\beta} P_{qs}^{\beta}) \cdot (pq|rs) \quad (1.91)$$

Applying Roothaan's method gives two sets of coupled equations.

$$\sum_q (F_{pq}^{\alpha} - e_i^{\alpha} S_{pq}) C_{qi}^{\alpha} = 0 \quad (1.92)$$

$$\sum_q (F_{pq}^{\beta} - e_i^{\beta} S_{pq}) C_{qi}^{\beta} = 0 \quad (1.93)$$

where the elements of  $F$  are given by

$$F_{pq}^a = H_{pq} + \sum_{rs} [P_{rs} (pq|rs) - P_{rs}^a (ps|qr)] \quad (1.94)$$

and  $a = \alpha$ , or  $\beta$ .

The LCAO expressions for the  $F$  matrix were given by Brickstock and Pople.<sup>16</sup> The solution of (1.92) and (1.93) proceeds as described for the SCF method by using sets of trial coefficients  $C^{\alpha}$  and  $C^{\beta}$  to construct the density matrices,  $P_{pq}^{\alpha}$  and  $P_{pq}^{\beta}$ .

## References

1. The Holy Bible, Genesis Ch. 1, v. 3.
2. A. Einstein, Ann. d. Phys., 1905, 17 (4), 132.
3. M. Planck, Ann. d. Phys., 1901, 4 (4), 553.
4. N. Bohr, Phil. Mag., 1913, 26, 1.
5. E. Schrödinger, Ann. d. Phys., 1926, 79, 361.
6. W. Pauli, Z. Physik., 1925, 31, 765.
7. J. C. Slater, Phys. Rev., 1930, 35, 509.
8. V. Fock, Z. Physik., 1930, 61, 126.
9. D. R. Hartree, Proc. Camb. Phil. Soc., 1928, 24, 89.
10. C. C. J. Roothaan, Rev. Mod. Phys., 1951, 23, 69.
11. G. G. Hall, Proc. Roy. Soc. (London), 1951, A205, 541.
12. C. C. J. Roothaan, Rev. Mod. Phys., 1960, 32, 179.
13. R. K. Nesbet, Proc. Roy. Soc., 1955, A230, 312, 322.
14. J. C. Slater, Phys. Rev., 1930, 35, 210.
15. J. A. Pople and R. K. Nesbet, J. Chem. Phys., 1954, 22, 571.
16. A. Brickstock and J. A. Pople, Trans. Faraday Soc., 1954, 50, 901.



CHAPTER I

SECTION 2

Types of molecular orbital calculations

## Non Empirical Calculations

### The Use of Slater Orbitals

The hydrogen-like orbitals, described in the previous *section*, have several disadvantages when used for molecules or large atoms. They are based on the assumption that each electron moves in a field of nuclear charge,  $Ze$ . This only applies for single electron systems, such as the hydrogen atom from which these functions are built up. In a molecule, or large atom, however, the electrons are on average at different distances from the nucleus at any one time, and the inner electrons partially screen the nucleus so that the outer electrons are affected by a nuclear charge of less than  $Ze$ . The functions, therefore, do not describe the atom or molecule accurately.

The second disadvantage of these hydrogen-like orbitals is that when obtained from an SCF calculation, the atomic orbitals are expressed in the form of a numerical table, and not as analytical functions. J. C. Slater<sup>1</sup> attempted to rectify these problems by devising approximate analytical functions which took account of the fact that each electron moved in a field of effective nuclear charge. Slater devised a set of rules<sup>2</sup> for calculating the effective nuclear charge,  $Z_{\text{eff}}$ , and called the difference between  $Z_{\text{eff}}$  and  $Z$ , the screening or shielding constant,  $s$ . The radial function  $R_{nl}(r)$  is then given by (2.1).

$$R_{nl}(r) = N_r^{n-1} e^{-ar} \quad (2.1)$$

where  $a = \frac{Z - s}{n}$  and  $N$  is the normalising factor.

The approximate nature of the Slater orbital is shown by the fact that the radial functions for the 2s, and 3s orbitals, which for the hydrogen-like functions have nodes, (see Figure 2), are now nodeless. The Slater orbitals for carbon,  $Z=6$ , are shown in Table 3, and a plot of these is shown in Fig. 7.

As can be seen from Table 3, the 2s Slater orbital has the same radial part as the 2p orbital, and is therefore not orthogonal to the 1s orbital.

Table 3 Slater orbitals for carbon

n	l	m	Orbital	Slater orbital
1	0	0	1s	$N_{1s} e^{-ar} = N_{1s} e^{-5.70r}$
2	0	0	2s	$N_{2s} r e^{-ar} = N_{2s} r e^{-1.625 r}$
2	1	1	2p <sub>x</sub>	$N_{2p_x} r \cos \Theta e^{-ar} = N_{2p_x} x e^{-1.625 r}$

Table 4 Ne atom: Uncontracted Gaussian Set, orbital exponents, and expansion coefficients

	orbital exponent	expansion coefficient		orbital exponent	expansion coefficient	
		1s orbital	2s orbital			2p orbital
S <sub>1</sub>	47870.2	0.00021	-0.00005	P <sub>1</sub>	129.802	0.00426
S <sub>2</sub>	7385.83	0.00162	-0.00038	P <sub>2</sub>	30.4192	0.03061
S <sub>3</sub>	1660.18	0.00863	-0.00206	P <sub>3</sub>	9.62151	0.11927
S <sub>4</sub>	460.539	0.03617	-0.00856	P <sub>4</sub>	3.54645	0.26912
S <sub>5</sub>	146.038	0.12134	-0.03097	P <sub>5</sub>	1.41435	0.35733
S <sub>6</sub>	50.4137	0.30702	-0.08388	P <sub>6</sub>	0.578893	0.33183
S <sub>7</sub>	18.7165	0.43944	-0.17194	P <sub>7</sub>	0.216044	0.16084
S <sub>8</sub>	7.39702	0.22518	-0.10947			
S <sub>9</sub>	2.6768	0.01554	0.37643			
S <sub>10</sub>	0.775195	-0.00230	0.57102			
S <sub>11</sub>	0.29176	0.00095	0.20449			

Fig. 7 Slater atomic orbitals for carbon along the X-axis.

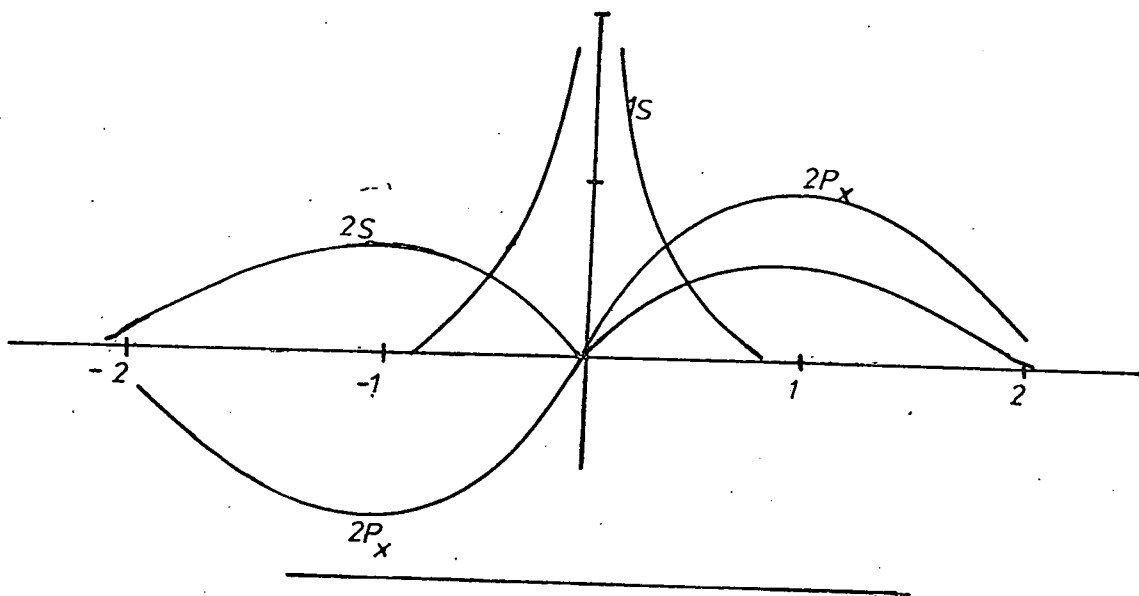


Fig 8

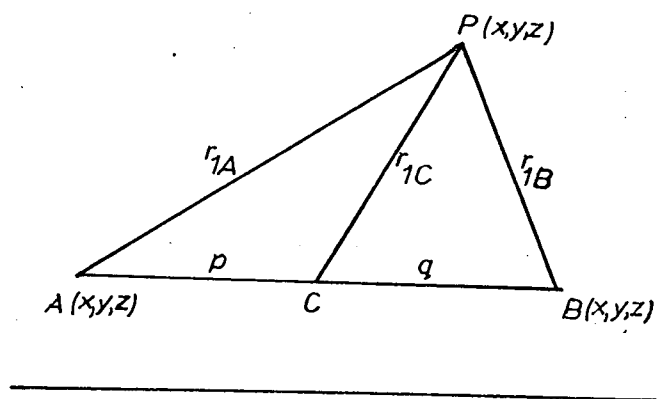
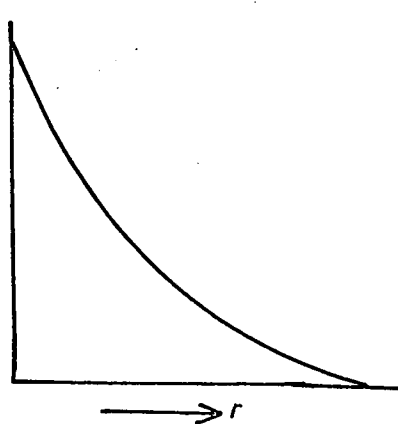
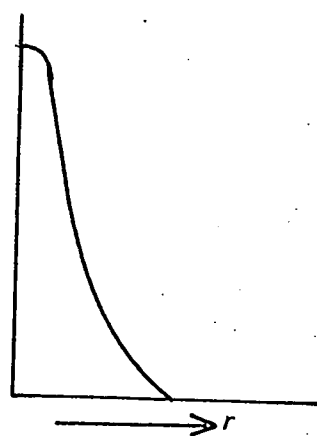


Fig.9



Slater type orbital.



Gaussian type orbital.

This can be overcome by creating a  $2s'$  orbital orthogonal to the  $1s$  orbital such that

$$\langle 1s | 2s' \rangle = 0 \quad \text{where} \quad \langle 1s | 2s' \rangle = \int_0^{\infty} 1s \, 2s' \, d\tau$$

putting  $2s' = \lambda 1s + 2s$  gives

$$\begin{aligned} \langle 1s | 2s' \rangle &= \langle 1s | \lambda 1s + 2s \rangle \\ &= \lambda \langle 1s | 1s \rangle + \langle 1s | 2s \rangle \\ &= \lambda + S_{1s \, 2s} \\ &= 0 \\ \therefore \lambda &= -S_{1s \, 2s} \\ \therefore 2s' &= 2s - S_{1s \, 2s} \, 1s \end{aligned} \quad (2.2)$$

where  $S_{1s \, 2s}$  is the overlap integral between the  $1s$  and  $2s$  orbitals. The  $2s'$  orbital is now orthogonal to the  $1s$  orbital, and is represented by a linear combination of  $1s$  and  $2s$  orbitals.

The values of  $\alpha$  are obtained from SCF calculations on atoms, the best  $\alpha$ 's giving the best total energies. They are then used for calculations on molecules containing these atoms.

For  $1s$  to  $3p$  functions, Slater orbitals provide a reasonable description of the atoms, giving better representation for the inner shells than for the valence shells. For higher functions they become inadequate, but a better overall radial function can be obtained by using a linear combination of Slater type orbitals (LCSTO) such that

$$R_{nl}(r) = \sum_i C_i \text{STO}(n_i, \alpha_i) \quad \text{where} \quad \text{STO} = \text{Slater type orbital} \quad (2.3)$$

The basis set of the STO's is fixed, and the coefficients,  $C_i$ , are optimised to give a minimum total energy using Roothaan's method. It has been found that in the LCSTO method, only a few STO's are required to produce energies within 0.001% of the Hartree-Fock minimum, and in general a linear combination of two STO's per atomic orbital is adequate to reproduce the atomic wave functions accurately. This is known as a double-zeta basis. A basis set consisting of one STO per occupied atomic

orbital is called a single zeta, or minimal basis set. For example, the helium atom,  $\text{He } (1s^2)$ , can be adequately represented by the following function:-

$$\Psi = 0.83415 \text{ STO } (1s: 1.44608) + 0.19060 \text{ STO } (1s: 2.86222) \quad (2.4)$$

The calculated energy using (2.4) is  $-2.8616700$  a. u., compared to  $-2.8616799$  a. u. given by the exact Hartree-Fock function. This was shown by Clementi.<sup>3</sup>

### The use of Gaussian type orbitals.

For diatomic molecules involving only two centres, the integral evaluation involved in the Roothaan procedure is relatively easy using STO's, but for a molecule with three or more atoms, the three centre one electron and two electron integrals are considerably more difficult and tedious to evaluate. Four centre two electron integrals are most difficult to evaluate as they cannot be expressed in analytical form, and must be evaluated by numerical integration.

The major breakthrough to this problem came when S. F. Boys<sup>4</sup> introduced a gaussian type orbital (GTO) of the form (2.5) to replace the STO.

$$G(r) = x^l y^m z^n \exp(-ar^2) \quad (2.5)$$

The product  $x^l y^m z^n$  represents the angular distribution of the function, and  $l$ ,  $m$ , and  $n$  can have any integral value. Using  $r^2$  in the exponential instead of  $r$ , enabled a four centre integral to be reduced to a two centre one. The four centre problem is shown diagrammatically in Fig. 8, and the overlap integral involving s-orbitals is evaluated below.

$\langle aA | bB \rangle =$  overlap integral, where  $a$ , and  $b$  are exponents centred on nuclei  $A$  and  $B$  respectively.

$$= \int_0^\infty \exp(-ar_{1A}^2 - br_{1B}^2) d\tau_1 \quad (2.6)$$

$$\text{Let } AB = R \quad \therefore \quad AC = p = \frac{b}{a+b} \cdot R \quad \text{and} \quad BC = \frac{a}{a+b} \cdot R = q \quad (2.7)$$

Using the cosine rule:-

$$r_{1A}^2 = p^2 + r_{1C}^2 + 2pr_{1C} \cos \theta \quad (2.8)$$

$$r_{1B}^2 = q^2 + r_{1C}^2 - 2qr_{1C} \cos \theta \quad (2.9)$$

Eliminating  $\cos \theta$ :-

$$\begin{aligned} qr_{1A}^2 + pr_{1B}^2 &= pq(p+q) + r_{1C}^2(p+q) \\ &= (p+q)(pq + r_{1C}^2) \end{aligned}$$

Substituting for  $p$  and  $q$  (2.7):-

$$\frac{a}{a+b} R r_{1A}^2 + \frac{b}{a+b} R r_{1B}^2 = R \left( \frac{b}{a+b} \cdot \frac{a}{a+b} \cdot R^2 + r_{1C}^2 \right)$$

Multiplying by  $\frac{a+b}{R}$  gives:-

$$ar_{1A}^2 + br_{1B}^2 = \frac{ab}{a+b} R^2 + r_{1C}^2 (a+b)$$

$$\therefore \langle aA | bB \rangle = \int \exp \left( \frac{-ab}{a+b} R^2 \right) \exp (-r_{1C}^2 (a+b)) dr_{1C} \quad (2.10)$$

The computational effort involved in using GTO's is therefore less than for STO's, but as the decaying exponential involves  $r^2$  rather than  $r$ , a GTO falls off more rapidly, and does not give a good representation of an atomic orbital especially at the nucleus and at large distances from the nucleus as shown in Fig. 9.

This defect can be rectified by using a large number of GTOs to represent an STO, and this method is known as the Linear Combination of Gaussian Orbitals (LCGO). The orbitals can then be represented as:-

$$\Psi = \sum_i C_i G_i \quad (2.11)$$

Huzinaga<sup>5</sup> found that more than twice as many gaussian functions, compared to Slater functions, were required to give a wave function of the same degree of accuracy. In spite of this, it has been shown by Whitten and Allen<sup>6</sup> that the use of GTOs will still substantially decrease the computational time used for molecular integrals.

### The contracted LCGO basis set

Using the LCGO method increases the number of basis functions used, and makes convergence in the SCF more difficult, and time consuming computationally. This difficulty can be overcome by reducing the number of variables in the SCF section, and this can be done by taking appropriate linear combinations of the gaussians representing the atomic functions. The molecular orbital can then be written as

$$\Psi_i = \sum_k C_{ik} G'_k \quad (2.12)$$

where  $G'_k$  is a contraction of gaussians of the same type on the same centre. For example,

$$G'_1 = c'_1 G_1 + c'_2 G_2 + c'_3 G_3 \quad (2.13)$$

where the coefficients  $c'$  are fixed, and in the SCF, only the coefficients,  $C_{ik}$  are allowed to vary. Consider the Ne (<sup>1</sup>S) atom as an example. Using Huzinaga's<sup>7</sup> standard optimal basis set consisting of 11 gaussian s-type functions, given as  $s_1$  to  $s_{11}$  in Table 4, and 7 gaussian p-type functions given as  $p_1$  to  $p_7$  in Table 4, the total energy for Ne(<sup>1</sup>S) is -128.5447 a. u. When the gaussians are contracted as shown in Table 6, the total energy is given by -128.54114 a. u. Table 5 shows the results of the contracted and uncontracted basis sets from calculations by Clementi and Mehl.<sup>8</sup> Using contracted functions the loss in accuracy in total energy is only  $3 \times 10^{-3}\%$ .



Table 5 Comparison of contracted and uncontracted sets,  $(Ne^1S)$

Type	Total energy	e(1s)	e(2s)	e(2p)	No. s basis	No. p basis
uncontracted	-128.5447	-32.772	-1.9300	-0.84999	11	7,7,7
contracted	-128.54114	-32.700	-1.9298	-0.84810	4	3,3,3

Table 6 Contracted functions for  $Ne(^1S)$  atom

$$S_1' = 0.00021 S_1 + 0.00162 S_2 + 0.00836 S_3$$

$$S_2' = 0.03617 S_4 + 0.12194 S_5 + 0.307012 S_6$$

$$S_3' = 0.43944 S_7 + 0.22518 S_8$$

$$S_4' = 0.37643 S_9 + 0.57102 S_{10} + 0.20443 S_{11}$$

$$P_1' = 0.00426 P_1 + 0.03061 P_2 + 0.11927 P_3$$

$$P_2' = 0.26912 P_4 + 0.35733 P_5$$

$$P_3' = 0.16084 P_6 + 0.33183 P_7$$

### Floating Spherical Gaussians

Using gaussian type functions to represent orbitals, only integrals of s-character are fast to evaluate, while those of p- and d-character take very much longer. To increase the speed of calculations the method of floating spherical gaussians was introduced by J. L. Whitten.<sup>9</sup> This method uses linear combinations of s-orbitals to represent p- and d-type orbitals as shown in Fig. 10, and gives the orbitals as

$$\begin{aligned}\psi_{p_y} &= \psi_{1s_1} - \psi_{1s_2} \\ &= \exp(-a_{1s_1} r^2) - \exp(-a_{1s_2} r^2)\end{aligned}\quad (2.14)$$

$$\begin{aligned}\psi_{d_{xy}} &= \psi_{1s_1} + \psi_{1s_4} - \psi_{1s_2} - \psi_{1s_3} \\ &= \exp(-a_{1s_1} r^2) + \exp(-a_{1s_4} r^2) - \exp(-a_{1s_2} r^2) \\ &\quad - \exp(-a_{1s_3} r^2)\end{aligned}\quad (2.15)$$

This enables the speed of integral evaluation to be increased at the sacrifice of the total energy which is usually not so good.

### Polarisation functions

In using either GTOs or STOs to represent atomic orbitals, the situation may arise where combinations of the occupied atomic orbitals to give molecular orbitals is not adequate to describe the molecule. Using lithium hydride, LiH, as an example, the molecular orbitals can be formed by combining the 1s, and 2s atomic orbitals of Li with the 1s orbital of H. Assuming that the lowest molecular orbital is almost purely  $\text{Li}_{1s}$ , and therefore non-bonding, the molecular orbitals can be given as:-

$$\psi_{1\sigma} = \phi_{1s}^{\text{Li}} \quad (2.16)$$

$$\psi_{2\sigma} = a \phi_{1s}^{\text{H}} + b \phi_{2s}^{\text{Li}} \quad (2.17)$$

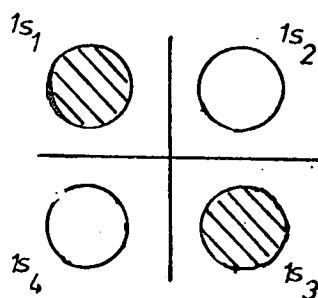
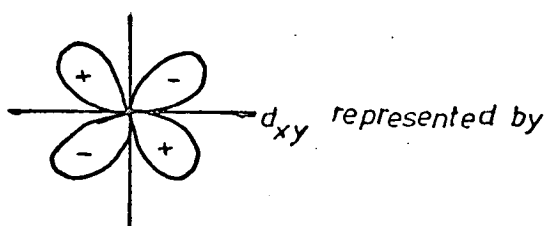
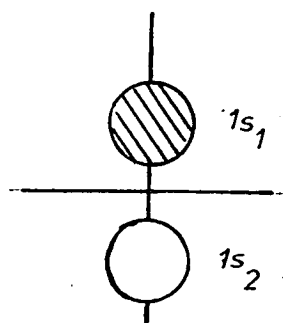
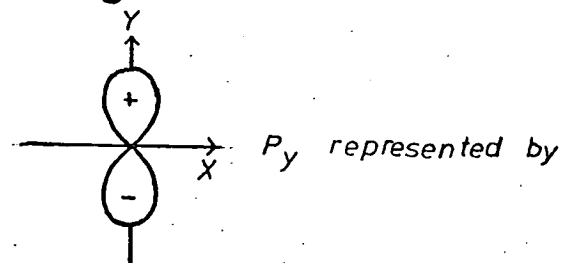
However, as written in (2.17) the bonding orbital is spherically symmetrical, and it is expected that the orbital will in fact be distorted from its spherical symmetry due to polarisation of the Li atom by the H atom. Orbital  $\psi_{2\sigma}$  is therefore grossly inadequate and the lithium  $2p_0$  function must be added to the basis set and is called a polarisation function. In 1960, B. J. Ransil<sup>10</sup> found that the orbitals for Li H were given by:-

$$\psi_{1\sigma} = 0.996 \phi_{1s}^{\text{Li}} - 0.017 \phi_{2s}^{\text{Li}} + 0.007 \phi_{2p_0}^{\text{Li}} - 0.006 \phi_{1s}^{\text{H}} \quad (2.18)$$

$$\psi_{2\sigma} = 0.141 \phi_{1s}^{\text{Li}} + 0.308 \phi_{2s}^{\text{Li}} + 0.211 \phi_{2p_0}^{\text{Li}} + 0.703 \phi_{1s}^{\text{H}} \quad (2.19)$$

This gave a total energy of -7.970 a. u. for LiH compared with -7.986 a. u. in the Hartree-Fock limit as calculated by Kavalas and Nesbet.<sup>11</sup>

Fig.10



## Empirical and semiempirical molecular orbital calculations

In the early days of molecular orbital theory, there were no computers able to cope with the extensive calculations required to give non-empirical molecular wave functions. The central feature of all molecular orbital calculations is the solution of the secular equations (1.39), and to enable these equations to be solved reasonably quickly approximate molecular orbital methods were devised. The earliest method, developed by Hückel,<sup>12</sup> was an essentially empirical one, whereby the elements of the secular determinant are given appropriate values. The obvious drawback of this type of method is that the final solution will always be strongly dependent on the input parameters.

The second type of approximate calculation introduces approximations for the atomic and molecular integrals which are used to determine the expressions for the elements of the secular determinant. The final solution is therefore still reached by an SCF approach, and these methods are known as approximate or semiempirical SCF methods, and were introduced as an improvement to the empirical approach of Hückel. Originally, both methods were applied only to  $\pi$ -electron systems of unsaturated planar molecules, and the  $\sigma$ -electrons were treated as a non-polarisable core. These methods were later extended to include all valence electrons.

### The Hückel Molecular Orbital (HMO) method

The Hückel method was the earliest one used for molecular orbital calculations, and was based on the LCAO treatment of  $\pi$ -electron systems.

The method was applied to unsaturated planar molecules composed of first row atoms in which each atom contributed one electron to the  $\pi$ -system. As the p-orbitals in the  $\pi$ -system are orthogonal to other orbitals on the same atom, the interaction between the  $\pi$ -electrons and the rest of the system should be relatively small, justifying the treatment of the  $\pi$ -system separately.

The method makes use of the secular equations (1.39), and the precise form of the Hamiltonian is not investigated. The method is heavily

parameterised by using standard values for  $H_{ij}$  and  $S_{ij}$ . The Hückel determinant for butadiene,  $c_1 = c_2 = c_3 = c_4$  is shown in (2. 20).

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} & H_{13} - ES_{13} & H_{14} - ES_{14} \\ H_{21} - ES_{21} & H_{22} - ES_{22} & H_{23} - ES_{23} & H_{24} - ES_{24} \\ H_{31} - ES_{31} & H_{32} - ES_{32} & H_{33} - ES_{33} & H_{34} - ES_{34} \\ H_{41} - ES_{41} & H_{42} - ES_{42} & H_{43} - ES_{43} & H_{44} - ES_{44} \end{vmatrix} = 0 \quad (2. 20)$$

In the Hückel method the following approximations are made:-

a) It is assumed that, in the case of hydrocarbons where all atomic orbitals are carbon p-orbitals,  $H_{ii} = \int \Phi_i H \Phi_i d\tau$  is constant, and is set equal to  $\alpha$ .

b) For  $i \neq j$   $H_{ij}$  is set equal to  $\beta$  in all cases where  $i$  and  $j$  are on directly bonded atoms, otherwise  $H_{ij} = 0$ .

c) The overlap integral,  $S_{ij} = \int \Phi_i \Phi_j d\tau$  is 1 for  $i=j$ , and is otherwise zero.

This simplifies (2. 20) to give (2. 21).

$$\begin{vmatrix} \alpha - E & \beta & 0 & 0 \\ \beta & \alpha - E & \beta & 0 \\ 0 & \beta & \alpha - E & \beta \\ 0 & 0 & \beta & \alpha - E \end{vmatrix} = 0 \quad (2. 21)$$

Dividing by  $\beta$  and setting  $\frac{\alpha - E}{\beta} = x$  gives (2. 22)

$$\begin{vmatrix} x & 1 & 0 & 0 \\ 1 & x & 1 & 0 \\ 0 & 1 & x & 1 \\ 0 & 0 & 1 & x \end{vmatrix} = 0 \quad (2. 22)$$

The determinant (2. 22) can be solved for  $x$ , and hence the values of  $E$  are obtained for the four  $\pi$ -orbitals of butadiene in terms of  $\alpha$  and  $\beta$ .

$$\begin{aligned}
 E_1 &= \alpha + 1.618\beta & E_3 &= \alpha - 0.618\beta \\
 E_2 &= \alpha + 0.618\beta & E_4 &= \alpha - 1.618\beta
 \end{aligned}
 \tag{2.23}$$

The linear equations corresponding to the secular determinant (2.22) are:-

$$\begin{aligned}
 C_1 x + C_2 &= 0 \\
 C_1 + C_2 x + C_3 &= 0 \\
 C_2 + C_3 x + C_4 &= 0 \\
 C_3 + C_4 x &= 0
 \end{aligned}
 \tag{2.24}$$

Knowing the values of  $x$  and using the normalisation conditions that  $\sum_r C_r^2 = 1$ , the coefficients  $C_1$  to  $C_4$  can be determined, and the  $\pi$ -orbitals,  $\psi$ , for the butadiene system can be written in terms of carbon p-orbitals,  $\phi$ , as follows:-

$$\begin{aligned}
 \psi_1 &= 0.372 \phi_1 + 0.602 \phi_2 + 0.603 \phi_3 + 0.372 \phi_4 \\
 \psi_2 &= 0.602 \phi_1 + 0.372 \phi_2 - 0.372 \phi_3 - 0.602 \phi_4 \\
 \psi_3 &= 0.602 \phi_1 - 0.372 \phi_2 - 0.372 \phi_3 + 0.602 \phi_4 \\
 \psi_4 &= 0.372 \phi_1 - 0.602 \phi_2 + 0.602 \phi_3 - 0.372 \phi_4
 \end{aligned}
 \tag{2.25}$$

The  $\alpha$  integral,  $\alpha_c$ , for the carbon atom is often taken as an arbitrary zero, and the  $\beta$  integral,  $\beta_{cc}$ , between two carbon atoms is taken as the unit of energy. For atoms other than carbon,  $\alpha$  and  $\beta$  are expressed in terms of  $\alpha_c$  and  $\beta_{cc}$  so that energies can be expressed in terms of  $\beta_{cc}$ . For a molecule  $xy$

$$\alpha_x = \alpha_c + h_{x,cc} \beta_{cc}
 \tag{2.26}$$

$$\beta_{xy} = k_{xy} \beta_{cc}
 \tag{2.27}$$

For qualitative fits of calculated and experimental data this is adequate

Table 7 Guide to bond indexing by Hess and Schaad

<u>Index</u>	<u>Bond type</u>	<u>Value in <math>\beta</math> units</u>
23	$H_2C = CH$	2.000
22	$HC = CH$	2.0699
22'	$H_2C = C$	2.000
21	$HC = C$	2.1083
20	$C = C$	2.1716
12	$HC - CH$	0.4660
11	$HC - C$	0.4362
10	$C - C$	0.4358

Table 8

<u>Compound</u>	<u>REPE <math>\times 10^{-3}</math></u>	<u>DEPE</u>
Benzene	65	0.333
Biphenyl	60	0.365
Naphthalene	55	0.368
Anthracene	47	0.380
Naphthacene	42	0.385
Pentacene	38	0.388
Azulene	22	
1,3-Butadiene *	2	
1,3,5-Hexatriene	-2	
Fulvene	-2	
Heptafulvene	-2	
Pentalene	-18	
Fulvalene	-33	

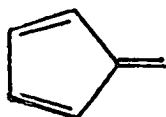
\* Values of the REPE for acyclic polyenes from ref. 13 lie in the range  $4 \times 10^{-3}$  to  $-2 \times 10^{-3} \beta$ .

but when quantitative values are required  $\beta_{CC}$  must be determined in conventional units of energy, and the value of  $\beta_{CC}$  varies with the method of measurement. For heteroatoms  $\alpha$ , and  $\beta$  are best determined as average values obtained for several different properties, where  $h_x$  and  $k_{xy}$  have been adjusted to give the best fit between calculated and observed properties of the system under consideration.

The biggest disadvantage of such a highly parameterised method is that the final answer will always be a function of the input parameters.

Despite the obvious limitations of the HMO method, it has recently been shown by Hess and Schaad,<sup>13</sup> that if a new quantity known as the resonance energy per  $\pi$ -electron (REPE) is used to determine stability and degree of aromaticity, instead of the old idea of delocalisation energies which proved unsuccessful in many cases, then the degree of aromaticity of a large number of benzenoid hydrocarbons, non-alternate hydrocarbons and acyclic polyenes can be predicted accurately. The REPE is defined as the difference between the total Hückel  $\pi$ -energy,  $E_{HMO}$  and the  $\pi$ -energy of the localised structure,  $E_{loc}$ , of the same compound, divided by the number of  $\pi$ -electrons. They first showed that  $E_{HMO}$  for an acyclic polyene could be determined additively from the  $\pi$ -double and single bond energies, but that to cover all types of polyenes, eight different bond types had to be defined, five C-C double bonds, and three C-C single bonds. Each bond type is indexed by two numbers, the first indicating bond order, and the second the number of hydrogens attached, as shown in Table 7. The bond energies are calculated in units of  $\beta$  and the  $E_{23}^{\pi}$  and  $E_{22'}^{\pi}$  are arbitrarily assigned the value of  $2\beta$ .  $E_{loc}$  is then calculated additively from the number of different types of single and double C-C bonds.

e. g. For fulvene



$$\begin{aligned}
 E_{loc} &= E_{22'}^{\pi} + 2E_{22}^{\pi} + E_{12}^{\pi} + 2E_{11}^{\pi} \\
 &= 2 + 4 \cdot 1.398 + 0.4660 + 0.8724 \\
 &= 7.4782 \beta
 \end{aligned}$$



$$RE = E_{\text{HMO}} - E_{\text{loc}} = (7.47 - 7.48)\beta = -0.01\beta$$

$$REPE = -0.01/6 = -0.002\beta.$$

Molecules with high REPE are predicted to be aromatic. As can be seen from the fulvene example, fulvene is predicted to be non-aromatic, this result being in agreement with its chemical behaviour. For benzenoid hydrocarbons, it was found that those with an REPE greater than  $0.05\beta$  do not undergo addition reactions with maleic anhydride, and those with REPE less than  $0.05\beta$  do, showing a tendency to be less aromatic and more polyolefinic in nature. Table 8 shows a representative example of hydrocarbons with their REPEs and delocalisation energy per  $\pi$ -electron (DEPE), where known. These are taken from the work of Hess and Schaad,<sup>13</sup> and show conclusively that the REPE is a better guide to aromaticity than the DEPE. It is obvious that this early theoretical method has not yet outlived its usefulness.

In the Hückel method electronic repulsion is not explicitly taken into account but must be indirectly involved in the choice of parameter values. An extension to the Hückel method was the Pölsner-Parr<sup>14</sup> method which adjusted the resonance integral,  $\beta$ , in connection with the choice of electronic repulsion integrals which made the method superior to that of Hückel.

Pople<sup>15</sup> also produced a  $\pi$ -electron only method which introduced the interaction between electrons and different cores and between electrons on different centres as the interaction of point charges, and extended his method to deal with heteroatoms with more than one  $\pi$ -electron.

### Extended Hückel Calculations

The next step forward was taken by Hoffman,<sup>16</sup> who introduced an all-valence electron treatment for molecules in which only the 1s electrons of carbon or a heteroatom were contained in the core, the other electrons being treated explicitly. This is known as the Extended Hückel Method (EHM). All electronic repulsion integrals are still ignored as in the

original Hückel method, but every valence overlap integral is included. The one centre integrals  $H_{ii}$  are semi-empirical quantities, but no longer equal. The valence ionisation potential for the particular orbital being considered is taken as the appropriate value of  $\alpha$ . The two centre integrals are taken directly proportional to the overlap, and are given as

$$H_{pq} = \frac{1}{2} K (\alpha_{pp} + \alpha_{qq}) S_{pq} \quad (2.28)$$

where the constant  $K$  is assigned the value 1.75 for use with STOs. The retention of the two centre integrals introduces a measure of long range interactions into the calculation. The absolute values of the orbital energies are not always accurate, but the qualitative ordering of energy levels is interesting and leads in many cases to a considerable overlap of  $\pi$  and  $\sigma$  levels.

#### Approximate SCF methods

With the advent of more efficient computers, came the advance in MO calculations of dealing with all-valence electron calculations on an SCF basis at a semi-empirical level. The time consuming and difficult part of an SCF calculation is the evaluation of the electron repulsion integrals, and approximate SCF methods were developed by reducing the number of electron repulsion integrals to be calculated. This was done by using the zero differential overlap (ZDO) approximation which sets

$$1) \quad S_{pq} = \int \Phi_p(1) \Phi_q(1) d\tau_1 = 0 \quad (2.29)$$

$$2) \quad (pq|rs) = (pp|rr) \delta_{pq} \delta_{rs} \quad (2.30)$$

The core integrals which also involve overlap distribution are not set equal to zero, but are treated semi-empirically to account for bonding effects due to overlap.

The extent to which the ZDO approximation is used for atomic orbitals leads to varying degrees of approximation in the semi-empirical SCF methods.

### Complete Neglect of Differential Overlap (CNDO)

The approximate SCF method which uses the ZDO approximations to the full extent is the CNDO method compiled by Pople, Santry and Segal.<sup>17</sup> The valence electrons are treated explicitly, and the inner shell is treated as a rigid core which affects only the one electron part of the Hamiltonian by modifying the nuclear potential. All monoatomic products,  $\Phi_p \Phi_q$  of different atomic orbitals are set equal to zero, but not the diatomic products. In addition the remaining two electron repulsion integrals are taken to depend only on the nature of the atoms A and B to which  $\Phi_p$  and  $\Phi_r$  belong, and not on the orbital type. This ensures the method to be rotationally invariant and gives (2.31).

$$(pp|rr) = Y_{AB}, \quad p \text{ on atom A, } r \text{ on atom B} \quad (2.31)$$

where  $Y_{AB}$  represents the average electrostatic repulsion between any valence electron on A and any one on B. When the interatomic distance  $R_{AB}$  is large,  $Y_{AB}$  is given by  $R_{AB}^{-1}$ . Using CNDO, the Roothaan equation (1.75) reduces to (2.32) and the corresponding matrix element of F (1.76) reduces to (2.33) and (2.34) when  $p=q$  and  $p \neq q$  respectively.

$$\sum_q F_{pq} C_{qi} = e_i C_{pi} \quad (2.32)$$

The general matrix element  $F_{pq}$  :-

$$F_{pq} = H_{pq} + \sum_{rs} P_{rs} [(pq|rs) - \frac{1}{2}(pr|qs)] \quad (1.76)$$

Using (2.30)

$$F_{pq} = H_{pq} + \sum_{rs} P_{rs} [(pp|rr) \delta_{pq} \delta_{rs} - \frac{1}{2}(pp|qq) \delta_{pr} \delta_{qs}] \quad (2.33)$$

$$\text{When } p=q \quad F_{pp} = H_{pp} + \sum_r P_{rr} (pp|rr) - \frac{1}{2} P_{pp} (pp|pp) \quad (2.34)$$

$$\text{When } p \neq q \quad F_{pq} = H_{pq} - \frac{1}{2} P_{pq} (pp|qq) \quad (2.35)$$

Using (2.34), (2.35) and (2.31) gives:-

$$F_{pp} = H_{pp} + \sum_B P_{BB} Y_{AB} - \frac{1}{2} P_{pp} Y_{AA} \quad (2.36)$$

$$F_{pq} = H_{pq} - \frac{1}{2} P_{pq} Y_{AB} \quad \Phi_p \text{ on A, } \Phi_q \text{ on B} \quad (2.37)$$

A second major approximation is applied to the one electron part of the Hamiltonian by adjusting the matrix elements of  $H$  to take account of the potential,  $V_A$ , due to the core of the atom  $A$  to which the electrons belong, and the potential,  $V_B$ , due to the core of any other atom  $B$ . The matrix elements of  $H$  are separated into one and two centre parts such that with  $\Phi_p$  on  $A$ :-

$$H_{pp} = U_{pp} - \sum_{B(\neq A)} (p | V_B | p) \quad (2.38)$$

where  $U_{pp} = (p | -\frac{1}{2}\nabla^2 - V_A | p)$  is the one centre part (2.39)

$$\Phi_p, \Phi_q \text{ on A gives:- } H_{pq} = U_{pq} - \sum_{B(\neq A)} (p | V_B | q) \quad (2.40)$$

If  $\Phi_p$  and  $\Phi_q$  are s, p, d type atomic orbitals  $U_{pq} = 0$  by symmetry, and  $(p | V_B | q)$  is taken as zero to be consistent with neglect of mono-atomic differential overlap, giving (2.40) equal to zero when  $\Phi_p$  and  $\Phi_q$  are on  $A$ . The term  $(p | V_B | p)$  is held constant for all  $\Phi_p$  on  $A$  to comply with invariance conditions and is set equal to  $V_{AB}$  where  $-V_{AB}$  is the interaction of any valence electron on atom  $A$ , with the core of atom  $B$ , and is given by  $R_{AB}^{-1}$  at large internuclear distances. When  $\Phi_p$  and  $\Phi_q$  are on different atoms  $A$  and  $B$ , differential overlap is not neglected and  $H_{pq}$  is given by (2.41).

$$H_{pq} = (p | -\frac{1}{2}\nabla^2 - V_A - V_B | p) \quad (2.41)$$

This represents the effect on the electronic energy due to the electron being influenced by the electrostatic field of two atoms at the same time,

and is known as the resonance integral,  $\beta_{pq}$ .  $\beta_{pq}$  is taken as proportional to the overlap integral.

$$\beta_{pq} = \beta_{AB}^{\circ} S_{pq} \quad (2.42)$$

$\beta_{AB}^{\circ}$  is a constant depending only on the nature of atoms A and B. The elements of the Fock matrix are now given by (2.43) and (2.44),  $\phi_p$  being on A and  $\phi_q$  on B.

$$\begin{aligned} F_{pp} &= U_{pp} - \sum_{B(\neq A)} V_{AB} + \sum_{B(\neq A)} P_{BB} Y_{AB} - \frac{1}{2} P_{pp} Y_{AA} + P_{AA} Y_{AA} \\ &= U_{pp} + (P_{AA} - \frac{1}{2} P_{pp}) Y_{AA} + \sum_{B(\neq A)} (P_{BB} Y_{AB} - V_{AB}) \quad (2.43) \end{aligned}$$

$$F_{pq} = \beta_{AB}^{\circ} S_{pq} - \frac{1}{2} P_{pq} Y_{AB} \quad (2.44)$$

The net charge on B,  $Q_B$ , can be written as :-

$$Q_B = Z_B - P_{BB} \quad (2.45)$$

where  $Z_B$  is the nuclear charge, and  $P_{BB}$  is the electron density, and (2.43) can be rewritten as (2.46).

$$F_{pp} = U_{pp} + (P_{AA} - \frac{1}{2} P_{pp}) Y_{AA} + \sum_{B(\neq A)} [-Q_B Y_{AB} + (Z_B Y_{AB} - V_{AB})] \quad (2.46)$$

Then  $(Z_B Y_{AB} - V_{AB})$  represents the difference in potential between the valence and core electrons of B, and is called the penetration integral. The term,  $-Q_B Y_{AB}$ , is the potential due to the net charge on atom B. The total energy can now be expressed as (2.47), the sum of the electronic and nuclear repulsion energies.

$$E = \frac{1}{2} \sum_{pq} P_{pq} (H_{pq} + F_{pq}) + \sum_{A < B} Z_A Z_B R_{AB}^{-1} \quad (2.47)$$

(2.47) is composed of monoatomic and diatomic energy terms and the total energy can be expressed as (2.48).

$$E = \sum_A e_A + \sum_{A < B} e_{AB} \quad (2.48)$$

$$\text{where } e_A = \sum_P^A P_{PP} U_{PP} + \frac{1}{2} \sum_p^A \sum_q^A (P_{pp} P_{qq} - \frac{1}{2} P_{pq}^2) \quad (2.49)$$

$$\text{and } e_{AB} = \sum_p^A \sum_p^B (2P_{pq} \beta_{pq} - \frac{1}{2} P_{pq}^2 \gamma_{AB}) + (Z_A Z_B R_{AB}^{-1} - P_{AA} V_{AB} - P_{BB} V_{BA} + P_{AA} P_{BB} \gamma_{AB}) \quad (2.50)$$

At large interatomic distances  $V_{AB}$ ,  $V_{BA}$ , and  $\gamma_{AB}$  approximate to  $R_{AB}^{-1}$ , and the second term of (2.50) becomes  $Q_A Q_B R_{AB}^{-1}$ . Two versions of CNDO exist known as CNDO/1<sup>17b</sup> and CNDO/2<sup>18</sup> which differ in their methods of parameterisation, the latter version CNDO/2 being a refinement of CNDO/1 to give better results, but both using the basic CNDO mathematics as already described. Both use Slater type atomic orbitals for the valence shells, CNDO/1 coping with atoms up to fluorine, and CNDO/2 having been extended to include second row atoms.<sup>18b</sup>

### CNDO/1

In CNDO/1 the electron repulsion integral,  $\gamma_{AB}$ , is calculated using valence s functions as a two centre coulomb integral.

$$\gamma_{AB} = \iint S_A^2(1) \frac{1}{r_{12}} S_B^2(2) d\tau_1 d\tau_2 \quad (2.51)$$

The valence S orbital on A is also used to calculate  $V_{AB}$ , the core of B being treated as a point charge at its nucleus.

$$V_{AB} = Z_B \int S_A^2(1) \frac{1}{r_{1B}} d\tau_1 \quad (2.52)$$

The matrix elements,  $U_{pp}$ , are derived semiempirically from the averaged experimentally determined ionisation potentials,  $I_p$  of the appropriate atomic states using (2.53).

$$-I_p = U_{pp} + (Z_A - 1) \gamma_{AA} \quad (2.53)$$

The resonance integral,  $\beta_{AB}^0$ , is given by (2.54).

$$\beta_{AB}^0 = \frac{1}{2} (\beta_A^0 + \beta_B^0) \quad (2.54)$$

where  $\beta_A$  and  $\beta_B$  are constants characteristic of atoms A and B, and are chosen to give the best overall fit between CNDO and non-empirical SCF calculations using a minimal basis set.

### CNDO/2

The CNDO/2 method calculates  $\gamma_{AB}$  and  $\beta_{AB}^0$  in the same manner as CNDO/1. In CNDO/1 the amount of penetration of electrons on one atom with the shells of another atom is overestimated. This leads to a net attraction, and consequently, predicted equilibrium distances for diatomic molecules are too small. This effect is related to the penetration integral,  $(Z_B \gamma_{AB} - V_{AB})$  and is corrected for in CNDO/2 by neglecting this integral completely. This is equivalent to setting  $V_{AB} = Z_B \gamma_{AB}$ . The potential  $V_{AB}$  is therefore no longer evaluated separately as in (2.52), but is related to the electron repulsion integral as in (2.55).

$$V_{AB} = Z_B \gamma_{AB} \quad (2.55)$$

This gives much improved results.

In estimating,  $U_{pp}$ , CNDO/1 used the ionisation potential, but the related electron affinity,  $A_p$ , could also have been used, (2.56).

$$-A_p = U_{pp} + Z_A \gamma_{AA} \quad (2.56)$$

In CNDO/2 the average of (2.53) and (2.56) is used giving (2.57).

$$-\frac{1}{2} (I_p + A_p) = U_{pp} + (Z_A - \frac{1}{2}) Y_{AA} \quad (2.57)$$

This gives the diagonal matrix element (2.43) as (2.58).

$$F_{pp} = -\frac{1}{2} (I_p + A_p) + [(P_{AA} - Z_A) - \frac{1}{2} (P_{pp} - 1)] Y_{AA} \\ + \sum_{B(\neq A)} (P_{BB} - Z_B) Y_{AB} \quad (2.58)$$

The off diagonal elements remain unchanged.

### Intermediate Neglect of Differential Overlap (INDO)

It was soon realised that using the CNDO methods, problems dealing with different electronic states arising from isoelectronic structures could not be assessed, as all the electron repulsion integrals for electrons on the same atom were approximated by  $Y_{AA}$  irrespective of the spins involved, and singlet and triplet states were therefore attributed with the same total electronic repulsion.

This was remedied by the introduction of an approximate SCF method which included the one centre exchange integrals. This was first introduced by R.N. Dixon<sup>19</sup> and was known as the Exchange Modified Zero Differential Overlap method (EMZDO). It was later elaborated by Pople, Beveridge and Dobosh,<sup>20</sup> and became known as the Intermediate Neglect of Differential Overlap method (INDO). This method is briefly described as follows.

If an s, p, d basis set is used, the off-diagonal core elements,

$F_{pq}$  ( $p \neq q$ ) vanish, and the only non-vanishing two electron integrals are  $(pp|pp)$ ,  $(pp|qq)$  and  $(pq|pq)$ . Using (1.76), (2.43) and (2.55), the F matrix elements become:-

$$F_{pp}^a = U_{pp} + \sum_r^A [P_{rr} (pp|rr) - P_{rr}^a (pr|pr)] \\ + \sum_{B(\neq A)} (P_{BB} - Z_B) Y_{AB} \quad (2.59)$$



$$F_{pq}^{\alpha} = (2P_{pq} - P_{pq}^{\alpha})(pq|pq) - P_{pq}^{\alpha}(pp|qq) \quad (2.60)$$

$p$  and  $q$  both on atom A, and  $p \neq q$ .

When  $p$  is on atom A, and  $q$  is on atom B,  $F_{pq}^{\alpha}$  is given as the  $\alpha$ -part of (2.44).

The one-centre two-electron integrals are determined using Slater type orbitals by means of the Slater-Condon<sup>21</sup> parameters,  $F^0$ ,  $G^1$  and  $F^2$ .  $G^1$  and  $F^2$  are found semiempirically to give the best fit with experimental atomic energy levels, while  $F^0 = Y_{AA}$  is calculated using Slater atomic orbitals as in the CNDO method. Table 9 shows the relationship between the integrals and the Slater-Condon parameters. The core integrals are again related to the electron affinity,  $A$ , and the ionisation potential,  $I$ , of the atomic orbitals, but are corrected to take account of the Slater-Condon parameters used.

The INDO method is restricted to first row atoms.

### The Spectroscopic - Potentials - adjusted INDO (SPINDO)

The INDO method was a great improvement on the CNDO method with the added advantage that the additional computation required was comparatively small. However, these methods were derived to reproduce as closely as possible, results which would be given by exact Hartree-Fock calculations. A method was now required which could be used to calculate the ground state properties of molecules. Such a method, known as the modified INDO method (MINDO) was devised by Baird and Dewar<sup>22</sup> to calculate ground state properties, and in particular heats of formation at a standard temperature. They selected a set of reference molecules, and chose parameters for the MINDO method to give good agreement with the observed properties of these molecules.

The values of  $G^1$  and  $F^2$  are retained as in INDO, but the elements  $U_{pp}$  (for  $p = s$  and  $p$  orbitals) and  $F^0$  are evaluated for each atom from an analysis of the atomic spectra of the first row atoms, taking each ground state as  $s^n p^m$  and using the experimental transition energies

Table 9

Two electron integral	Slater-Condon Parameter
$(ss \bar{s}s)$	$F^0 = Y_{AA}$
$(ss xx)^*$	$F^0 = Y_{AA}$
$(sx sx)$	$1/3 G^1$
$(xy xy)$	$3/25 F^2$
$(xx xx)^\dagger$	$F^0 + 4/25 F^2$
$(xx yy)^\dagger$	$F^0 - 2/25 F^2$

\* This assumes that the 2s, and 2p orbitals have the same radial part.

† If  $F^2$  is non-zero, then the interaction of electrons in different p-orbitals is distinguished.

Table 10

Atom	Parameter	MINDO	SPINDO
Carbon	$U_{ss}$	-49.659 ev	-48.289 ev
	$U_{pp}$	-41.159 ev	-41.159 ev
	$F^0$	11.089 ev	11.089 ev
	$F^2$	4.727 ev	4.727 ev
	$G^1$	7.285 ev	7.285 ev
	$\alpha$	1.62	1.925
Hydrogen	$U_{ss}$	-13.595 ev	-13.595 ev
	$F^0$	12.845 ev	12.845 ev
	$\alpha$	1.0	1.0

Table 10 (cont.)

Interaction	$f(R_{AB})$
(1s   1s)	0.13647
(1s   2s)	0.17832
(1s   2p $\sigma$ )	0.35100
(2s   2s)	$0.20187 + 0.09500/R^2$
(2s   2p $\sigma$ )	$0.27625 + 0.13000/R^2$
(2p $\sigma$   2p $\sigma$ )	$0.47000 + 0.2400/R^2 + 100 \exp(-5R)$
(2p $\pi$   2p $\pi$ )	$0.40375 + 0.19000/R^2$

found by Klopman<sup>23</sup> of the high spin states of configurations  $s^n p^{m+1}$ ,  $s^n p^m$ ,  $s^{m-1} p^m$ , and  $s^{n-1} p^m$ . The two centre electron repulsion integrals,  $Y_{AB}$ , are calculated using the Ohno-Klopman<sup>23</sup> expression (2.61).

$$Y_{AB} = -14.399 [R_{AB}^2 + (\rho_A + \rho_B)^2]^{-\frac{1}{2}} \quad (2.61)$$

$$\text{where } \rho_A = -7.1995/F_A^0 \text{ and } \rho_B = -7.1995/F_B^0 \quad (2.62)$$

and  $R_{AB}$  = internuclear distance between atoms A and B.

The expression (2.61) obeys the boundary conditions (2.63) and (2.64).

$$Y_{AB} \rightarrow \frac{1}{2} (F_A^0 + F_B^0) \text{ as } R_{AB} \rightarrow 0 \quad (2.63)$$

$$\text{and } Y_{AB} \rightarrow \rho^2/R_{AB} \text{ as } R_{AB} \rightarrow \infty \quad (2.64)$$

The resonance integral,  $\beta_{pq}$ , is set not only proportional to the overlap integral  $S_{pq}$  as in CNDO (2.42), but also to the mean of the valence state ionisation potentials,  $I_p$  and  $I_q$ , and to a factor of nuclear distance  $f(R_{AB})$  (2.65).

$$\beta_{pq} = S_{pq} (I_p^A + I_q^B) f(R_{AB}) \quad (2.65)$$

The MINDO method gave a good account of heats of formation of hydrocarbons, and good agreement with the first ionisation potential of many molecules as shown by Dewar, Worley,<sup>24</sup> and Haselbach.<sup>25</sup> This aspect was further investigated by Fridh, Åsbink and Lindholm<sup>26</sup> who were seeking a theoretical method which would reproduce all the valence ionisation potentials. They adjusted, and parameterised MINDO, and found that, for benzene, methane, ethane, ethylene and acetylene,<sup>27</sup> the new method, which they termed SPINDO, gave agreement between experimental and calculated IP.s of valence electrons within 0.3eV which, owing to the breadth of some of the bands in the experimental photoelectron spectra, was considered satisfactory.

The main difference between MINDO, and SPINDO is that in the latter method the resonance integrals,  $H_{pq}$ , and the overlap integrals,  $S_{pq}$ , are adjusted with respect to dependence on nuclear distance,  $R_{AB}$ . The dependence of  $S_{pq}$  on  $R_{AB}$  is adjusted by use of a new Slater exponent,  $\alpha$ , for carbon atoms, and the distance factor  $f(R_{AB})$ , which for MINDO has the same form for different types of interaction, is now adjusted with respect to size and distance,  $R_{AB}$ , for different types of interaction. Table 10 shows a comparison of MINDO and SPINDO parameters, and the factors  $f(R_{AB})$  for SPINDO.

## References

1. J. C. Slater, Phys. Rev., 1930, 36, 57.
2. C. A. Coulson, "Valence," Second Edition, Oxford University Press.
3. E. Clementi, IBM J. Res. Develop, 1965, 9, 2.
4. S. F. Boys, Proc. Roy. Soc. (London), 1950, A200, 242.
5. S. Huzinaga, J. Chem. Phys., 1965, 42, 1293.
6. J. L. Whitten, L. C. Allen, J. Chem. Phys., 1965, 43, 5170.
7. S. Huzinaga, J. Chem. Phys., 1965, 42, 1293.
8. E. Clementi and J. W. Mehl, "Quantum mechanical concepts and algorithms."
9. J. L. Whitten, J. Chem. Phys., 1963, 39, 349;  
ibid, 1966, 44, 359.
10. B. J. Ransil, Rev. Mod. Phys., 1960, 32, 329.
11. S. L. Kahalas and R. K. Nesbet, J. Chem. Phys., 1963, 39(3), 529.
12. (a) E. Hückel, Z. Physik, 1931, 70, 204.  
(b) E. Hückel, Z. Physik, 1932, 76, 628.  
(c) E. Hückel, Trans. Faraday Soc., 1955, 51, 600.
13. B. Andes Hess Jr., L. J. Schaad, (a) J. Amer. Chem. Soc., 1971, 93(2), 305; (b) ibid, 1971, 93(10), 2413.
14. R. Pairiser, R. G. Parr, J. Chem. Phys., 1953, 21, 466, 767.
15. J. A. Pople, Trans. Faraday Soc., 1953, 49, 1375.
16. R. Hoffmān, J. Chem. Phys., 1963, 39, 1397.
17. (a) J. A. Pople, D. P. Santry and G. A. Segal, J. Chem. Phys., 1965, 43, S 129.  
(b) J. A. Pople and G. A. Segal, J. Chem. Phys., 1965, 43, S 136.
18. (a) J. A. Pople and G. A. Segal, J. Chem. Phys., 1966, 44, 3289.  
(b) D. P. Santry and G. A. Segal, J. Chem. Phys., 1967, 47, 158.
19. R. N. Dixon, Mol. Phys., 1967, 12, 83.
20. J. A. Pople, D. L. Beveridge and P. A. Dobosh, J. Chem. Phys., 1967, 47, 2026.

21. J. C. Slater, "Quantum Theory of Atomic Structure," McGraw-Hill Book Company, Vol. 1, 339-342.
22. N. C. Baird and M. J. S. Dewar, J.Chem. Phys., 1969, 50, 1262.
23. G. Klopman, J. Amer. Chem. Soc., 1964, 86, 4550.  
(a) K. Ohno, Theoret. Chim. Acta, 1964, 2, 219.  
(b) G. Klopman, J. Amer. Chem. Soc., 1965, 87, 3300.
24. M. J. S. Dewar and S. D. Worley, J. Chem. Phys., 1969, 50, 654.
25. M. J. S. Dewar and E. Haselbach, J. Amer. Chem. Soc., 1970, 92, 590.
26. C. Fridh, L. Asbrink and E. Lindholm, Chem. Phys. Letters, 1972, 15, 282; ibid, 1972, 15, 408; ibid, 1972, 15, 567.
27. C. Fridh, L. Åsbrink and E. Lindholm, Faraday Discussions, 1972, 54, 127.

## CHAPTER 1

### SECTIONS 3 AND 4

3. Molecular and Atomic Properties.
4. Koopmans' Theorem



### Molecular and Atomic Properties

Every physical observable,  $m$ , and many related quantities can be characterised by a linear operator,  $M$ , which when used to operate on a wave function gives the quantum mechanical expectation value, QMEV, of the observable,  $m$ .

$$\text{QMEV} = \int \psi_i^* M \psi_i d\tau \quad (3.1)$$

Experimental measurements result in real values so that  $M$  must be hermitian i.e. has the property:-

$$\int \psi_i^* M \psi_j d\tau = \int \psi_j^* M \psi_i d\tau \quad (3.2)$$

A whole set of such operators exist, and when the QMEV of each is evaluated, a set of molecular and atomic properties is obtained. As many of these properties can be determined experimentally, the experimental results and the QMEVs can be compared to give a useful indication of the accuracy of the LCAO wavefunction being used.

#### Electron density

The electron density  $\rho(R)$  at a point  $R$  is given by the QMEV of the charge density operator,  $CD_{op}(R)$  as in (3.3).

$$\rho(R) = \int \psi_i^* |CD_{op}(R)| \psi_i d\tau \quad (3.3)$$

$$= N_i \sum_i^{occ} \psi_i(R) \psi_i(R) \quad (3.4)$$

where  $\sum_i^{occ}$  shows summation over occupied orbitals, and  $N_i$  is the orbital occupancy of  $\psi_i$ . For closed shell molecules (3.4) becomes (3.5).

$$\rho(R)_i = 2 \sum_i^{occ} \psi_i^*(R) \psi_i(R) \quad (3.5)$$

The total electron density will be equal to the total number of electrons ( $2n$ ) in the system, and corresponds to (3.6).

$$2n = \int \rho(R) dR = 2 \sum_i^{\text{occ}} \psi_i^*(R) \psi_i(R) dR \quad (3.6)$$

Expressing the molecular orbital,  $\psi_i$ , in LCAO form (1.65), gives (3.7) and (3.8).

$$2n = 2 \sum_i^{\text{occ}} \sum_{pq} c_{pi}^* c_{qi} \int \phi_p(R) \phi_q(R) dR \quad (3.7)$$

$$= \sum_{pq} P_{pq} S_{pq} \text{ using (1.71) and (1.66a)} \quad (3.8)$$

A detailed analysis of (3.8) will give gross atomic populations,  $N_A$ , and total orbital populations,  $N(\psi_i)$ . This technique was developed by Mulliken,<sup>1</sup> and is known as a population analysis.

From (3.8):-

$$N(\psi_i) = \sum_p c_{pi}^2 S_{pp} + \sum_{pq(p \neq q)} c_{pi} c_{qi} S_{pq} + \sum_q c_{qi}^2 S_{qq} \quad (3.9)$$

1
2
3

Mulliken defined terms 1 and 3 as net atomic populations, and term 2 as the overlap population. The gross atomic population,  $N_A$ , is then given by

$$N_A = \sum_i \left[ \sum_{p_A} c_{p_A i}^2 S_{p_A p_A} + \sum_{p_A(p_A \neq q)} c_{p_A i} c_{q i} S_{p_A q} \right] \quad (3.10)$$

where  $p_A$  is on atom A.

$$= \text{net atomic population of A} + \text{overlap population of A} \quad (3.11)$$

The total electron density (3.8) can also be expressed as 3.12.

$$2n = \sum_A N_A = \sum_i N(\psi_i) = \sum_{pq} P_{pq} S_{pq} \quad (3.12)$$

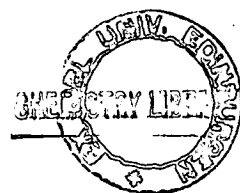
### The electric dipole moment

The electric dipole moment,  $\mu$ , is defined as the centre of gravity of positive (nuclear), and negative (electronic) charge. The total dipole is given by:-

$$\mu = \sum_n z_n r_n - \sum_i e_i r_i \quad (3.13)$$

$$= \sum_n z_n r_n - \langle \psi_i | \sum_j r_j | \psi_i \rangle \quad (3.14)$$

where  $z$  is the nuclear charge,  $n$  is the number of nuclei and  $r$  is the distance of the nucleus from the point at which the dipole is being calculated.



Koopmans' theorem in photoelectron spectroscopy

An ab initio SCF calculation provides a set of molecular orbital energies as specified in (1.48a) which can be correlated with the ionisation potentials of electrons in specific orbitals as found by the method of photoelectron (PE) spectroscopy, the beginnings of which were discovered by Lenard and Millikan as described on page 2. The correlation is made using Koopmans' theorem<sup>2</sup> by which an experimentally determined ionisation potential (IP) is set equal to the negative of the orbital energy,  $e_i$ , as shown in (4.1).

$$-e_i = \text{IP} \quad (4.1)$$

This is equivalent to setting the IP equal to the difference of two energy expressions, one for the molecule, and one for the resulting ion. There are three assumptions<sup>3</sup> made by (4.1). The first is that on removal of an electron from a molecule, there is no reorientation of the remaining electrons, so that the IPs can be expressed as the sum of the same integrals as in the orbital expression (1.48a).

The total experimental molecular energy would not be equal to the calculated energy even if true Hartree-Fock (HF) wavefunctions were used, due to the fact that HF theory does not take relativistic considerations into account. This effect is concerned mostly with inner electrons, which have large kinetic energies. However, as PE spectroscopy deals with valenceshell electrons, the second assumption in Koopmans' theorem, that the relativistic energy is the same for both ion and molecule, is reasonable.

The total experimental energy can be expressed as (4.2)

$$E_{\text{total expt.}} = E_{\text{HF}} + E_{\text{relativistic}} + E_{\text{correlation}} \quad (4.2)$$

The correlation energy arises due to electronic interaction in the HF procedure being an interaction of each electron with an averaged out charge of all the others. In reality, each electron will interact with

all other electrons instantaneously. This is partly corrected for by the exchange integral,  $K$ , for electrons of the same spin, but gives rise in general to a need for a correlation energy. The effect has been found to be related to the effect of electron pair interactions in the molecule and as such the correlation energy for a molecule, and its ions will differ, the latter being less. The third assumption made by (4. 1) is then that the correlation energy is the same for both molecule and ion, so that the use of Koopmans' theorem will normally produce energy differences, larger than the experimental IPs.

The use of Koopmans' theorem and its reliability are discussed more fully in Chapter II Section 4.

#### References

1. R. S. Mulliken, J. Chem. Phys. , 1955, 23, 1833, 1841;  
ibid, 1962, 36, 3428.
2. T. Koopmans, Physica, 1933, 1, 104.
3. W. G. Richards, Int. J. of Mass Spectrom. and Ion Physics,  
1969, 2, 419.

## CHAPTER 1

### Section 5

The Atmol 2 programme

## The Atmol 2 programme

Throughout this work, ab initio calculations of the LCGO SCF type have been carried out using the Atmol-2\* system of programmes on the IBM 360/195, 370/155 and the ICL 4/75 computers. The Atmol-2 system is composed of several packages which can be run concurrently or independently.

### The Integrals Package

This package computes and outputs onto a file the two electron integrals evaluated over the atomic orbital basis set. The integrals section tends to be time consuming, and two methods are available for reducing computing time. The first is the use of a SYMCEN directive which indicates local centres of symmetry in the molecular geometry, and enables two electron integrals, equal in absolute magnitude by symmetry, to be identified and evaluated only once. For example, a molecule such as that shown in Fig. 1 has two local symmetry centres, A and B, and these can be located by the programme via the SYMCEN directive. The programme is then able to determine centres related by symmetry, and will set the related two electron integrals equal. In ring A, the two electron integrals of the S-orbitals, given by  $\langle 11, 44 \rangle$  and  $\langle 22, 33 \rangle$  where the numbers correspond to electrons at positions 1 to 4 in Fig. 1, are set equal by symmetry. For the p-orbitals, a similar result is obtained in the z direction while in the x-direction  $\langle 11, 44 \rangle = -\langle 22, 33 \rangle$ . In ring B, similar relationships will exist, and the number of two electron integrals to be calculated is considerably reduced.

The second is the use of an accuracy factor which sets two threshold values for the two electron integral evaluations. In this work, these were set at  $10^{-7}$  and  $10^{-8}$ , indicating that if the absolute value of an integral is less than  $10^{-7}$ , it will not be output onto the storage file, and if the estimated value of an integral over primitives is less than  $10^{-8}$ , it is not taken into account in the total integral over contracted functions.

\* The Atmol 2 suite of programmes was compiled and mounted on the 360/195 IBM computer at the Atlas Laboratory by:- V. R. Saunders, M. F. Chiu, M. F. Guest.

All possible integrals of the form  $\langle f_i f_j | f_k f_l \rangle$ , where  $f$  denotes the basis function are evaluated and stored using the indices  $i, j, k$  and  $l$  whose values are given by the basis function ordering so that for each index, number of basis functions  $\langle \text{index} \rangle \geq 1$ . The storage of the two electron integrals is unimportant as they are stored with their reference indices. As integrals set to zero by the limit are not output, and these are often related by symmetry to other integrals, the related zero integrals are also ignored, and considerable filing space is saved. To take full advantage of this in minimising computer filing space and time, the molecules should be orientated such that the number of two electron integrals, whose values are likely to be zero by symmetry is maximised.

The one electron integrals are also evaluated and output onto a second file, known as the Dump file, and, again to save time and space, the symmetry of the matrices is used. For example, as the overlap integral  $S_{ij} = S_{ji}$ , then only  $n(n+1)/2$  elements of the matrix are required to define all unique overlap integrals, and are stored in the following order:-

$S_{ij}$	11(1)	12(2)	13(4)	14(7)	.....
matrix	21	22(3)	23(5)	24(8)	.....
	31	32	33(6)	34(9)	.....
	41	42	43	44(10)	....

The bracketed numbers give the storage order. The kinetic energy and total one electron integrals and the electronic contributions to the dipole moment are stored in a similar fashion. The nuclear repulsion energy and the nuclear contributions to the dipole moment are also evaluated in this section.

#### The closed shell restricted Hartree Fock (RHF) SCF package

In this package, before the iterative procedure starts, the facility exists of redefining the basis set as linear combinations of the basis



functions. This can reduce the time required for diagonalisation of the Fock matrix as the elements of the matrix are blocked according to the irreducible representation to which the molecular orbitals belong. The matrix is then composed of a set of submatrices which can be diagonalised separately. For example, a molecule of  $C_{2v}$  symmetry having molecular orbitals in the point groups  $A_1$ ,  $A_2$ ,  $B_1$  and  $B_2$ , will produce a Fock matrix composed of four submatrices corresponding to the molecular orbitals in each representation. Overall diagonalisation can be achieved by diagonalising each of the submatrices separately. In these cases the eigenvectors are output in symmetry blocks and orbital types can be identified more easily.

The integrals evaluated in the integrals section are over atomic orbitals, and must be converted in the SCF section to integrals over molecular orbitals. The elements of the Fock Hamiltonian are dependent on the atomic orbitals through the elements  $P_{pq}$  of the density matrix,  $P$ . The initial Fock matrix can be set up in three ways.

(a) It can be started with a set of trial vectors which have been obtained from a previous run on a closely related system. For example, the SCF section of the calculation on the triplet state of isoindole can be set up using the vectors obtained at convergence of the SCF section of the closed shell case. These are then used to set up a density matrix and hence the Fock matrix.

Normally this information will not be available for a completely new system, and one of the other methods must be used.

(b) The second method is to provide a set of input Fock matrix elements by setting the diagonal elements,  $F_{pq}$ , equal to their expected value at convergence. This information can be obtained from any previous run as the diagonal Fock matrix elements are almost constant for a particular atom in different molecular environments. The off-diagonal elements are then estimated by the programme taking the already available potential and kinetic energy integrals into account.

(c) The third method is used in the absence of any other suitable

data, and sets the diagonal Fock matrix elements,  $F_{pq}$ , equal to the one electron Hamiltonian elements,  $H_{pq}$ , by assuming that all  $P_{rs} = 0$ .

The Fock matrix is then diagonalised by the Jacobi method which zeros the off diagonal elements one at a time. For example, to diagonalise a matrix  $\underline{M}$ , a matrix  $\underline{A}$  is defined whose elements are given by  $a_{ij}$ . One element  $m_{ij} = m_{ji}$  is chosen to be zeroed, and the elements  $a_{ii}$  and  $a_{jj}$  of  $\underline{A}$  are set equal to  $\cos\Theta$  and  $-\cos\Theta$  respectively, and  $a_{ij}$  and  $a_{ji}$  are set equal to  $\sin\Theta$ . All other diagonal elements of  $\underline{A}$  are set equal to 1, and all other off diagonal ones to zero. The angle  $\Theta$  is then defined as:-

$$\tan 2\Theta = \frac{2m_{ij}}{m_{ii} - m_{jj}}$$

and the operation  $\underline{A}_{ij}^{-1} \underline{M} \underline{A}_{ij}$  yields  $\underline{M}'$ , a new matrix in which the elements  $m_{ij}$  and  $m_{ji}$  are now zero. Repeated operations zero other off diagonal elements, but can produce non zero values for previously zeroed elements. However, these values are always less than the original value of  $m_{ij}$ , and the process is repeated until all off diagonal elements are zero within a defined accuracy limit. In Atmol-2, this limit is  $10^{-10}$ , and leads to a convergence limit in total energy of  $10^{-5}$  a.u. If the matrix  $\underline{M}$  is given by:-

$$\underline{M} = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix}$$

then elimination of  $m_{12}$  would proceed as follows.  $\underline{A}$  is defined as:-

$$\underline{A} = \begin{vmatrix} \cos\Theta & \sin\Theta & 0 \\ \sin\Theta & -\cos\Theta & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

and  $\tan 2\Theta = \frac{2 \times 1}{0} = \infty$

$$2\Theta = 90^\circ$$

$$\Theta = 45^\circ \quad \cos \Theta = \frac{1}{2} \quad \sin \Theta = \frac{1}{2}$$

$$\underline{A}^{-1} \underline{M} \underline{A} = \begin{vmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} \begin{vmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{vmatrix}$$

$$\therefore \underline{M}' = \begin{vmatrix} 1 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -1 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{vmatrix}$$

The output eigenvalues giving the orbital energies are Aufbau ordered, but in certain cases it is possible for the programme to converge on an orbital ordering which does not represent the ground state of the molecule. In these cases, the "SWAP" procedure can be put into operation. This swaps any two molecular orbitals specified by the user, and presents a set of reordered eigen vectors to the SCF procedure which is then repeated to give convergence on the ground state. For example, if a calculation is expected to produce two  $\pi$  orbitals, but the output shows only one and gives a configuration  $\dots \sigma \pi \sigma \pi^*$ , then the lowest binding energy  $\sigma$  orbital can be swapped with the virtual  $\pi^*$  orbital to give  $\dots \sigma \pi \pi^* \sigma$ .

### Open shell packages

Two open shell procedures are available with Atmol-2, an RHF procedure

and an unrestricted Hartree Fock (UHF) procedure, the former using a single Fock Hamiltonian constructed from the doubly and singly occupied orbitals, and the latter using a double Hamiltonian procedure to represent electrons of opposite spin. In the RHF procedure, the elements of the Fock matrix can be evaluated in different ways depending on the results desired. For example, the matrix can be diagonalised in the fastest time by representing the matrix as follows:-

$$H = F + 2J(R_1) - K(R_1) + J(R_2)$$

where  $R_1$  and  $R_2$  represent the closed and open shell density matrices respectively. The exchange operator in the open shell is neglected. If the RHF procedure is used to reproduce results which would be obtained from a UHF procedure then the open shell exchange operator is included. The total wavefunction is invariant to the method of canonicalisation of the Fock matrix, and only the orbital energies are affected.

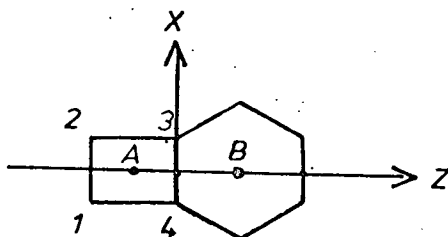
In open shell procedures it is often difficult to obtain convergence in the SCF section due to a swapping of doubly occupied molecular orbitals with singly occupied ones, and of singly occupied orbitals with virtual ones. The Atmol-2 RHF and UHF can force convergence by using level shifters which increase the energy separation between the doubly and singly occupied orbitals, and between the singly occupied and virtual molecular orbitals. This prevents large scale swapping of molecular orbitals and tends to force convergence. In the UHF procedure the level shifters increase the energy separation between the singly occupied  $\alpha$  or  $\beta$  spin orbitals and the corresponding virtual orbitals, and if the system monitors a divergent situation, then the programme automatically increases the level shifters in an attempt to force convergence.

### The Mulliken Population analysis

In this section, each molecular orbital is analysed, and a basis function population analysis is given for each orbital.

The atomic populations contributed by each molecular orbital are found by summation of the appropriate basis function populations, and are output under the heading "group populations." The total group populations are then found by summation of the group populations and are then doubled as each molecular orbital is analysed as if it contained one electron.

Fig. 1.



CHAPTER 2  
Results and discussion

## Gaussian Basis Sets

### The length of a basis set

It has already been mentioned (p 28. ) that calculations on molecular systems greater than diatomics are infinitely more manageable when GTOs are used to represent the atomic orbitals, due to the integration difficulties encountered when using STOs for anything other than atoms and diatomic molecules.

It is obvious from the literature that much work has been done on finding optimal basis sets, and many sets are in use. There are two methods practised for finding optimal sets. The first is to use only GTOs, and to minimise the total atomic energy for a particular number (N) of GTOs by a variational procedure. For example, for a 1s orbital,  $\psi_{1s}$  can be represented by a linear combination of GTOs.

$$\psi_{1s} = \sum_i c_i X_i \text{ where } X_i = \exp(-\alpha_i r^2)$$

The total atomic energy is then optimised using:-

$$E_{\text{atom}} = \langle \psi H \psi \rangle \text{ for all } c_i \text{ and } X_i$$

Since in any one atomic orbital eg 1s, all the GTOs are of the same symmetry, and the atomic orbital is represented by a linear combination of  $c_i X_i$ , the eigenvectors,  $c_i$ , can be used as fixed linear coefficients for that atomic orbital in a molecular calculation.

The second method is to use the known optimal values of STOs for atomic orbitals, and expand these in terms of a fixed number (N) of GTOs.

It has been shown that for first row atoms, <sup>1</sup> the Hartree-Fock (H-F) limit can be obtained to seven figure accuracy

Table 1 Number of GTOs (N) v energy of the hydrogen atom

N	E(a.u.)
1	-0.424413
2	-0.485813
3	-0.496979
4	-0.499277
5	-0.499809
6	-0.499940
7	-0.499976
8	-0.499991
9	-0.499997
10	-0.499999
Exact Value	-0.500000

Table 2 Atomic energy (a.u.) of C(<sup>3</sup>P) with respect to different s/p GTO basis sets

Ns \ Np	3	5	7	9	10
1	-35.4511	-37.0943	-37.2573		
2	-35.7322	-37.3899	-37.5573		
3	-35.7958	-37.4503	-37.6192		
5				-37.6852 <sup>2</sup>	
6					-37.6873 <sup>2</sup>



by using a basis set consisting of 6s and 4 p-type STOs. Therefore, the procedure of representing an atomic orbital by a single set of GTOs fitted to the best STO for the orbital by a least squares method is likely to produce a poor approximation.

First row atoms. By carrying out a variational procedure on first row atoms, Huzinaga<sup>2</sup> found expressions for the best single GTO to represent a particular STO, and also showed how the total energy for a hydrogen atom converges with increase in the number of GTOs. These results are shown in Table 1. It is clear that for the hydrogen atom a value of N equal to four is sufficiently accurate for most purposes. Similarly for the He (1s<sup>2</sup>)(1S)<sup>2</sup> state, the total energy and orbital energies are effectively at the Hartree-Fock (H-F) limit using five GTOs.

The accuracy of the orbital energies involved must also come into consideration as it is important that the basis set provides an accurate description of the orbitals in terms of energy as well as a good total energy. The orbital energies are controlled by the ratio of GTOs representing s and p type functions. The elements carbon, nitrogen, oxygen and fluorine were studied systematically by Csizmadia<sup>3</sup> et al in terms of s/p ratio, and total energies. The data for carbon, C(<sup>3</sup>P), total energies is shown in Table 2, and is supplemented by data from Huzinaga's<sup>2</sup> 9s/5p and 10s/6p basis set. It is worth noting at this point that for C(<sup>3</sup>P), the H-F limit of total energy is -37.68861 a.u., and with a double zeta STO basis set, it is -37.68668 a.u. Therefore, only the largest basis set in Table 2, is an improvement on

Table 3 Orbital energies (a.u.) for C(<sup>3</sup>P) with respect to various s/p GTO basis sets.

Basis Set	Type Size	3s/1p <sup>3</sup> 6	5s/2p <sup>3</sup> 11	7s/3p <sup>3</sup> 16	7s/3p <sup>4</sup> 16	9s/5p <sup>2</sup> 24	10s/6p <sup>2</sup> 28	H-F limit
1s		-10.0514	-10.7238	-10.7543	-11.3227	-11.3249	-11.3252	-11.3255
2s		+ 0.2145	- 0.1842	- 0.2254	- 0.7023	- 0.7051	- 0.7055	- 0.7056
2p		+ 0.0787	- 0.1945	- 0.2466	- 0.4217	- 0.4325	- 0.4331	- 0.4333

the double zeta STO result. On the basis of total energy, it would appear from Table 2 that the s/p ratio gives best results at approximately 2:1. Thus a 5s/2p basis set is better than a 7s/1p set, although the former has eleven functions ( $5s + (3 \times 2)p$ ) and the latter only ten functions ( $7s + (3 \times 1)p$ ). From the results it is also clear that the H-F limit is approached rapidly between a 7s/3p and a 9s/5p basis set. Roos and Siegbahn<sup>4</sup> later published a 7s/3p set giving C(<sup>3</sup>P) total energy as -37.656 a.u. which is even closer to the 9s/5p set of Huzinaga. However, Huzinaga did not claim to have truly optimised the orbital exponents due to the existence of multiple minima.

The orbital energies produced by the basis sets shown in Table 2, are listed in Table 3 with the additional information provided by the Roos and Siegbahn 7s/3p set and also the H-F limits. The inadequacy of the small 3s/1p set is aptly shown by the antibonding nature of the 2s and 2p orbitals. Even up to 7s/3p, the Csizmadia basis sets give almost degenerate energies for 2s and 2p but the Roos and Siegbahn set is virtually at the H-F limit. The larger sets by Huzinaga offer little additional advantage over the Roos and Siegbahn 7s/3p set considering the number of extra functions incorporated. For this reason the Roos and Siegbahn sets for first row atoms, and for hydrogen were used as a starting point for the basis sets used in this work.

Second row atoms. The variation method of gaussian fitting to atomic orbitals has also been used for second row atoms, sodium to argon, for varying numbers of GTOs. The atomic energy from basis sets produced by three sets of workers<sup>4-6</sup> for the sulphur atom, S(<sup>3</sup>P), are shown in Table 4. Only two

Table 4 Basis sets for S(<sup>3</sup>P)

Size	(10s/6p) <sup>4</sup>	(12s/9p) <sup>5</sup>	(14s/10p) <sup>6</sup>
Total energy	-397.450	-397.4978	
		-397.4818 C	-397.4931 C
H-F limit	-397.5049		

C = contracted basis sets

---

Table 5 C(<sup>3</sup>P)s contraction coefficients for the Roos Siegbahn (7s/3p) set.

1s	2s
+0.004813	-0.001020
+0.037267	-0.008141
+0.172403	-0.038437
+0.459261	-0.126098
+0.456185	-0.190474
+0.034215	+0.522342
-0.009977	+0.594186

---

of the sets shown are uncontracted. The uncontracted (12s/9p) set of Veillard<sup>5</sup> is very close to the H-F limit. However, it is important when choosing basis sets for different atoms that the sets chosen should be of comparable accuracy. As the (7s/3p) Roos and Siegbahn set was chosen for carbon in this work, giving 99.91% of the H-F limit, the most suitable basis set for S(3P) was the Roos and Siegbahn set (10s/6p) giving 99.98% of the H-F limit.

### Contraction

Calculations using, as for example in this work a (7s/3p) basis set for carbon still involve a vast number of integrals, which cause problems of storage during calculations, and are also time consuming due to their number, and the large number of variables which they produce during the SCF procedure. The process can be speeded up without a great loss in accuracy by contracting the functions in such a way that fixed sets of GTOs from the original uncontracted set are allowed to vary according to (2.12) for an atomic orbital.

$$\psi_i = \sum_k c_{ik}^l G_k^l \quad (2.12)$$

where  $G_k^l$  is a contraction of gaussians of the same type on the same centre, whose coefficients are fixed. The first step is to decide on the division of the uncontracted basis set of s and p functions into sub sets representing 1s, 2s and 3s orbitals, and 2p and 3p orbitals. To achieve a reasonable degree of orthogonality, each GTO must appear in only one orbital. Sub sets are chosen in such a way as to keep the eigenvectors of the neglected functions small in comparison to those of the retained functions. A further

Table 6 Coefficients for  $S(^3P)$  from the Roos and Siegbahn<sup>4</sup> (10s/6p) set.

1s	2s	3s
0.001546	-0.000429	0.000125
0.011973	-0.003315	0.000962
0.059943	-0.017197	0.005011
0.207528	-0.062990	0.018436
0.442977	-0.170675	0.051125
0.392193	-0.220562	0.068969
0.044891	0.406937	-0.152708
-0.011027	0.700363	-0.438273
-0.003370	0.055594	0.524282
-0.001458	-0.013599	0.655507

2p	3p
0.028414	-0.007411
0.175840	-0.045940
0.467398	-0.134234
0.485545	-0.120902
0.063619	0.450422
-0.010985	0.678482

decision to be made, is whether the subsets should be single, double or triple zeta. The latter produce flexibility in the calculation, but the extra variable can become time consuming in the SCF section. For this reason, the present work has been carried out using a minimal basis set i.e. single zeta. The C(<sup>3</sup>P)s contraction coefficients for the Roos and Siegbahn<sup>4</sup> (7s/3p) set are shown in Table 5, and the coefficients of the retained functions are enclosed. The two largest coefficients were grouped together to form the 2s orbital, the other five s functions being used for the 1s orbital. When the negative coefficient -0.190474 is included with the 2s pair of GTOs to represent 2s by three GTOs, a poorer result was obtained for the atomic energy<sup>7</sup> indicating that GTOs should not be mixed between orbitals. For this reason the 1s and 2s subsets were chosen with no overlap, and as far as possible no eigenvectors of large weight were omitted. Each p orbital was represented by three gaussians. A similar procedure was adopted for nitrogen, and oxygen using a (7s/3p) basis set and contracting to a minimal basis [5,2;3]. The Roos and Siegbahn<sup>4</sup> (10s/6p) set for sulphur was used as a guide for the sulphur basis set used in this work. The sulphur coefficients are shown in Table 6, the retained coefficients being enclosed. The s functions are split [6,2,2] but for the orbitals 2s and 3s, one gaussian is substantially present in both orbitals. However, to improve orthogonality it is ignored in the 3s orbital, giving a poorer approximation to the atom. The split of the 6p functions into [4,2] is straight-forward.

Veillard<sup>5</sup> studied various contractions of large basis sets,

Table 7 Total energies for S(<sup>3</sup>P) using contracted basis sets

Size	(10s/6p) <sup>4</sup>	(12s/9p) <sup>5</sup>	(14s/10p) <sup>6</sup>
Contraction	[6,2,2;4,2]	[6,2,1,1,1,1;3,3,1,2]	[4,4,3,3;4,3,3]
Energy a.u.	-397.400	-397.4818	-397.4913
	STO - 6G	STO - 9G	H-F limit
	-396.18	-397.184	-397.5048

Table 8 'Best Atom' Basis set contractions and energies

Atom	Contraction	Free atom energy(a.u.)	Hartree-Fock Limit(a.u.)	% of H-F limit
H( <sup>2</sup> S)	[3]	-0.4970	-0.5000	99.4
C( <sup>3</sup> P)	[5,2;3]	-37.6105	-37.6886	99.8
N( <sup>4</sup> S)	[5,2;3]	-54.2754	-54.4009	99.8
O( <sup>3</sup> P)	[5,2;3]	-74.6121	-74.8094	99.7
S( <sup>3</sup> P)	[6,2,2;4,2]	-396.7476	-397.5048	99.8



(12s/9p), and showed that the atomic energies varied substantially. In general, the calculations were not greatly contracted, and approximated to double zeta (DZ) calculations. Huzinaga's <sup>6</sup> larger sets (14s/10p), were more heavily contracted to 1½ zeta but produced, in general, better atomic energies. The least square fitted GTO to STO type of calculations, <sup>8,9</sup> which are minimal basis, give poorer results up to STO - 5G/6G, and must be of, at least, this magnitude to produce comparable results to the minimal basis Roos and Siegbahn sets. The total energies for S(<sup>3</sup>P) produced by these workers <sup>5,6,8,9</sup> are compared with the contracted Roos and Siegbahn <sup>4</sup> set in Table 7. The Roos and Siegbahn set compares favourably with all the much larger sets, and is only 0.1 a.u. from the H-F limit. It would appear that much is to be gained from using the smaller (10s/6p) set contracted to [6,2,2;4,2] for S(<sup>3</sup>P) which is of manageable size, and of reasonable accuracy. A summary of the 'best atom' basis sets used, the free atom energies produced by them, and their percentage of the H-F limit <sup>10</sup> is shown in Table 8.

#### The use of scaled basis sets.

A further adjustment to molecular calculations can be made by allowing the orbitals to expand or contract relative to those of the free atom to produce a more accurate molecular description. If the atomic orbital is given by:-

$$\psi_i = \sum_j c_j \exp(-\alpha_j r^2)$$

then it is possible to optimise  $c_j$  and  $\alpha_j$  for each atom in the molecule. Pople et al <sup>11</sup> used this method to simultaneously optimise the contraction coefficients, and exponents for the valence shell of atoms in various molecular

situations e.g.  $H_2O$ ,  $NH_3$ ,  $CH_4$  using the relationship:-

$$\psi^{mol}(r) = \xi^{3/2} \psi^{atom}(\xi r)$$

where  $\xi$  is the scaling factor, and  $\xi^{3/2}$  is the ratio of scaled to unscaled normalisation factors. The calculations were split-valence type, using different scale factors for the inner and outer valence shell. He showed that, for the outer valence shell in particular, the scale factors were considerably different from unity. For carbon and hydrogen, the molecular environment produced a contraction of the atomic functions i.e.  $\xi > 1$ , while the nitrogen and oxygen gave values less than unity indicating expansion of the orbitals.

The basis sets used in this work were scaled using a method devised by the M.H. Palmer<sup>12</sup> research group. Only the exponents are scaled and only one scale factor is used as the sets are minimal. The method uses ethylene as a starting point, and the hydrogen exponents are scaled as  $\exp(-k_H \alpha_H r^2)$ , where  $k_H$  is the scaling factor. The unscaled best atom set for carbon is used. The total energy is calculated over a range of  $k_H$ , and an optimum value determined. This scale factor is adopted as standard for hydrogen, and used to determine total energies for a range of scale factors,  $k_{2s}$  and  $k_{2p}$ , for carbon. This produces a grid of total energies with respect to  $k_{2s}$  and  $k_{2p}$ , and optimum values are determined by parabolic interpolation.

A final check was made by off-setting  $k_H$ , and calculating the total energy using  $k_{2s}/k_{2p}$  optimum values for carbon. It was found that the new  $k_H$  was effectively identical to

Table 9 Molecules used to produce scale factors

Molecule used	Atom Scaled	Scaling Procedure
CH <sub>2</sub> = CH-O-H vinyl alcohol	H attached to O	Using unscaled O to give $k_{H-(O)}$
	O	Using $k_{H-(O)}$ to give $(k_{2s}/k_{2p})$ for O
CH <sub>2</sub> = CH-NH <sub>2</sub> vinyl amine	H attached to N	Using unscaled N to give $k_{H-(N)}$
	N	Using $k_{H-(N)}$ to give $(k_{2s}/k_{2p})$ for N
CH <sub>2</sub> = N-O-H formaldoxime	H attached to O	Using unscaled O to give $k_{H-(O)}$
	O	Using $k_{H-(O)}$ to give $(k_{2s}/k_{2p})$ for O
	N	Using $k_{H-(O)}$ , $(k_{2s}/k_{2p})$ O to give $(k_{2s}/k_{2p})$ for N
CH <sub>4</sub> methane	H	Using unscaled C to give $k_H$
	C	Using $k_H$ to give $(k_{2s}/k_{2p})C$
CH <sub>2</sub> = S Thioformaldehyde	S	Using C,H scaled ethylene to give $(k_{3s}/k_{3p})$ for S

the standard  $k_H$  used to scale carbon. These scale factors were then used without further adjustment to scale other atoms in specific molecular environments. The scaled atoms used in this work, the molecules used for scaling, and an outline of the scaling procedures are shown in Table 9. For sulphur it was necessary to acquire an exponent for the 3 d orbitals. Each d orbital was represented by a single gaussian giving a coefficient of unity, and the exponent was scaled using optimum values for  $(k_{3s}/k_{3p})$  sulphur. It was found to be optimum at 0.54<sup>13</sup>. The basis sets used in this work are shown in appendix 1.

The use of scaled sets produces advantages and disadvantages. It is found that the use of scaled sets lifts the eigenvalues by about 2eV in the valence shell giving a better agreement with experimental orbital energies when compared using Koopmans' theorem. However, for the molecular properties which are dependent on individual atomic structures within the molecule, agreement with experiment is sometimes worsened considerably. One such property, which frequently suffers is the dipole moment. The molecular properties of furan, and pyrrole have been studied by M.H. Palmer et al using a minimal basis set,<sup>14</sup> and using molecular scaling procedures<sup>15</sup> on the minimal basis. In this work, and in others of M.H. Palmer<sup>16,17</sup> et al, binding energy (BE) or atomisation energy is calculated using the following equation:-

$$BE = E_{mol} - \sum_{atom} E_{atom}$$

where the molecular energy is calculated using scaled basis sets and the atomic energies are calculated using "best atom" basis sets to give reasonable values for binding energies.

Other workers <sup>11</sup> have used scaled sets for both the molecules, and the atoms, and as scaled sets are far from optimal for the atomic state, the binding energies produced by these calculations are abnormally large.

R E F E R E N C E S

1. E. Clementi, C.C.J. Roothaan and M. Yoshimine, Phys. Rev., 1962, 127, 1618
2. S. Huzinaga, J. Chem. Phys., 1965, 42 (4), 1293
3. S.S. Seung, M.C. Harrison and I.G. Csizmadia, Theoret. Chim. Acta(Berl), 1967, 8, 281
4. B. Roos and P. Siegbahn, Proc. Seminar on Computational Problems in Quantum Chemistry (Strassburg, 1969).
5. A. Veillard, Theoret. Chim. Acta., 1968, 12, 405
6. S. Huzinaga and C. Arnau, J. Chem. Phys., 1970, 53, 348
7. W. Moyes, private communication, Chemistry Dept., University of Edinburgh
8. R.F. Stewart and W.J. Hehre, J. Chem. Phys., 1970, 52, 5243
9. W.J. Hehre, R. Ditchfield, R.F. Stewart and J.A. Pople, J. Chem. Phys., 1970, 52, 2769
10. E. Clementi, "Table of Atomic Functions", Supplement to IBM J. Research and Development, 1965, Vol. 9, Page 2
11. R. Ditchfield, W.J. Hehre and J.A. Pople, J. Chem. Phys., 1971, 54, 724
12. a) M.H. Palmer and R.H. Findlay, Chem. Phys. Letters, 1972, 15, 416  
b) M.H. Palmer, A.J. Gaskell and M.S. Barber, Theoret. Chim. Acta, 1972, 26, 357

13. M.H. Palmer and R.H. Findlay, J. Chem. Soc., Perkin II, 1974, 1885.
14. M.H. Palmer and A.J. Gaskell, J. Chem. Soc., Perkin II, 1974, 420
15. M.H. Palmer and A.J. Gaskell, Theoret. Chim. Acta (Berl), 1972, 26, 357
16. M.H. Palmer and R.H. Findlay, Tetrahedron Letters, 41, 4165
17. M.H. Palmer and R.H. Findlay, J. Chem. Soc., Perkin II, 1974, 1885.

Fig. 1

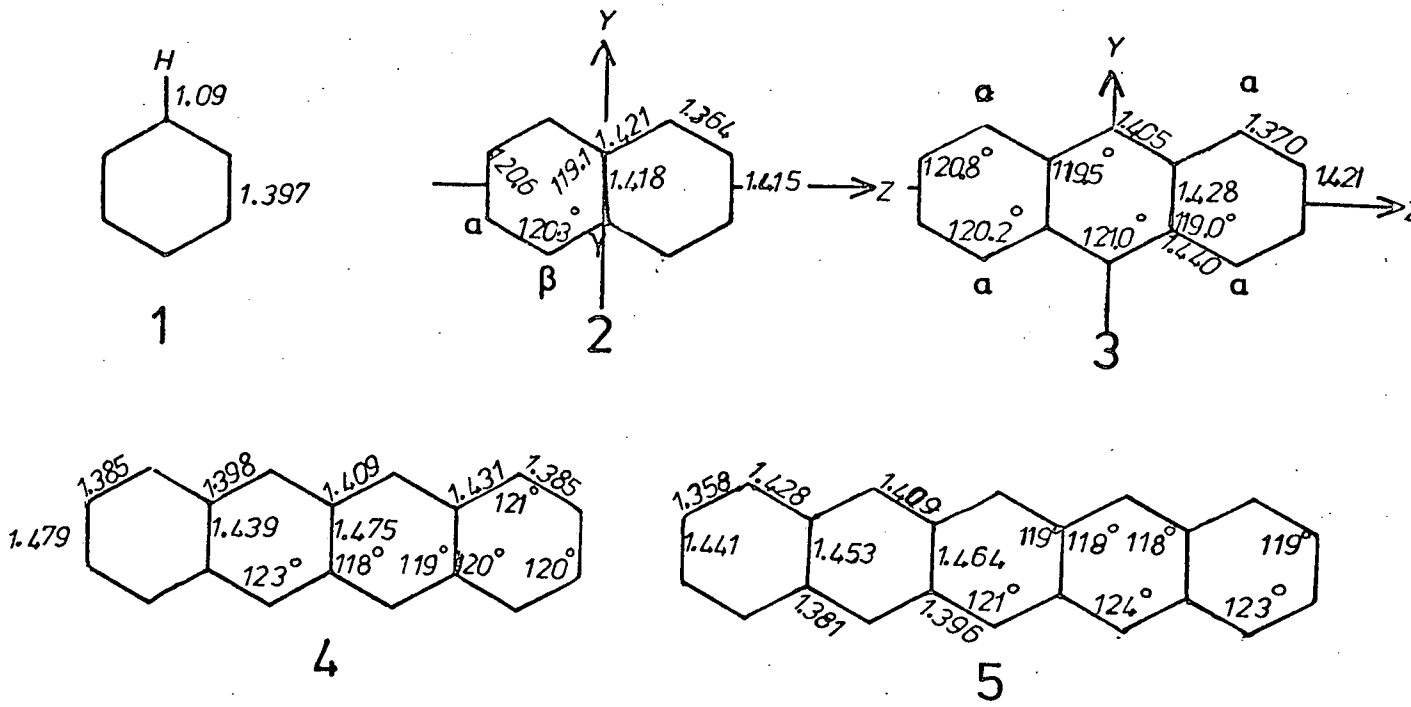
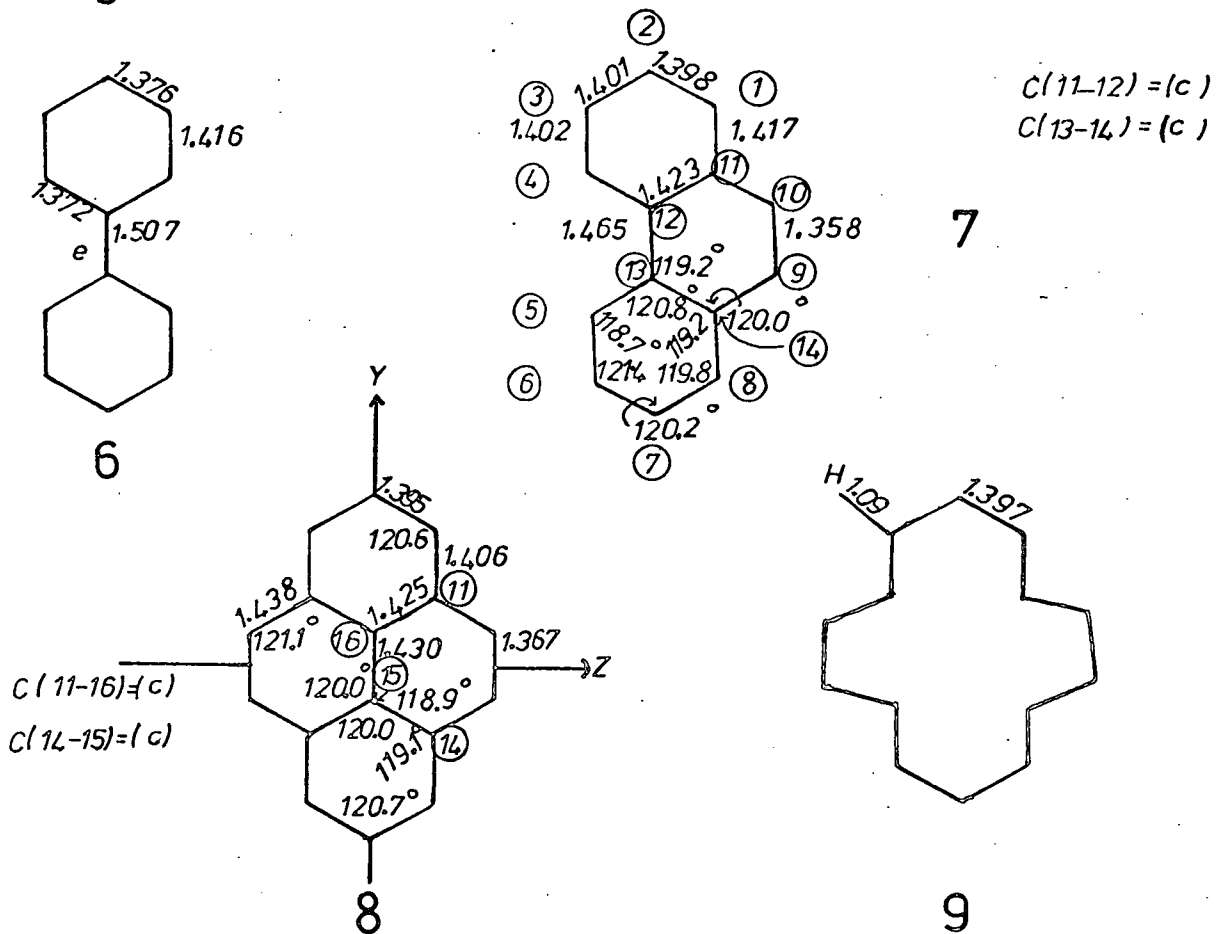


Fig. 2





## Molecular Geometry

The molecular geometries of the molecules involved in this work are shown, and described in the following section. In some cases, suitable geometric data already existed, and has been used, but in others, the data was not available, and so theoretically constructed geometries have been used. The molecular symmetries were used to predict the number of orbitals in each symmetry class which would be expected from the calculations.

Hydrocarbons. The simplest hydrocarbon studied in this section is benzene (1), and the experimental geometry,<sup>1</sup> as shown in Fig. 1, was used for the calculation.

Various X-ray and neutron diffraction studies of naphthalene now exist, but at the time the calculation was performed only X-ray data<sup>2</sup> and electron diffraction<sup>3</sup> data were available, the two sets being in close agreement ( $\pm 0.007\text{\AA}$ ). The X-ray crystal structure was used to give the geometry of naphthalene (2), shown in Fig. 1. As the number of rings present in the hydrocarbon increases, the available geometric data decreases, and so for the molecules, anthracene (3),<sup>4</sup> naphthacene (4)<sup>5</sup>, (often called tetracene) and pentacene (5)<sup>6</sup>, the X-ray crystal structures are rather old, and although bond lengths are quoted to three decimal places, the accuracy is probably not greater than  $\pm 0.02\text{\AA}$ .

The experimental geometries for 3, 4 and 5 are shown in Fig. 1, and were used for the calculations.

Biphenyl (6) is planar in the solid state <sup>7</sup> but non-planar in the gas phase having a dihedral angle of  $32^{\circ}$ . The planar geometry used in the calculation is shown in Fig. 2. The neutron diffraction structure of phenanthrene <sup>8</sup> (7) has been determined, and is shown in Fig. 2. If this geometry is compared with that of biphenyl, it can be seen immediately that the basic biphenyl geometry still exists but that the molecule is slightly distorted due to a short C(9) - C(10) bond which approximates to a classical double bond. Overall the molecule is biphenyl with a double bond attachment.

A recent neutron diffraction study of pyrene <sup>9</sup> (8) gave the geometry shown in Fig. 2, and was used in the LCGO calculation. Two Kekulé like structures can be devised for pyrene, and are shown as (a) and (b) in Fig. 3. The first, (a), has a [14]-annulene skeleton with a central double bond, while the second approximates to biphenyl with two extra double bonds placed symmetrically about the C(15)-C(16) bond. The neutron diffraction information clearly indicates the latter case as shown by the double bond nature ( $1.367\text{\AA}$ ) of the bonds at positions (f), and the single bond nature of the central C-C bond (e). In both (7) and (8), Fig. 2, the bonds at positions (c) are becoming elongated due to the extra double bonds which are stretching the molecule across positions 4, 11 and 5, 14 in (7) and across positions 11, 12 and 13, 14 in (8) in order to accommodate themselves as double bonds.

In going from biphenyl (6), to pyrene (8), the bond (e) is

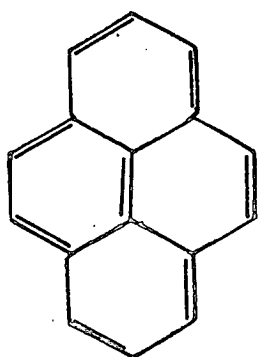
Table 1  $D_{2h}$  representations for  $1s_c$  and  $2p_{xc}$  for anthracene

	$\sigma$ ( $1s_c$ )		$\pi$ ( $2p_{xc}$ )
$a_g$	+ +	$a_u$	+ -
	+ +		- +
$b_{3g}$	+ -	$b_{3u}$	+ +
	- +		+ +
$b_{1u}$	+ -	$b_{1g}$	+ +
	+ -		- -
$b_{2u}$	+ +	$b_{2g}$	+ -
	- -		+ -

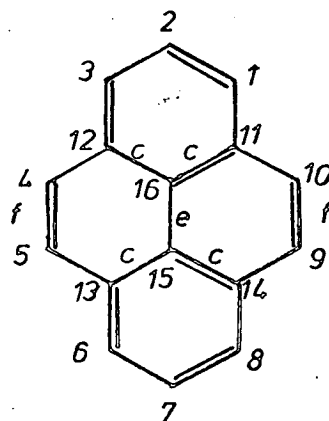
Table 2  $C_{2v}$  representations for  $1s_c$  and  $2p_{xc}$  of series B

	$\sigma$ ( $1s_c$ )		$\pi$ ( $2p_{xc}$ )
$a_1$	+ +	$b_1$	+ +
$b_2$	+ -	$a_2$	+ -

Fig. 3



a



b

becoming progressively shorter, again due to the accommodation of the outer double bonds. However, even in pyrene the bond (e) is still single in character at 1.430Å.

For comparison purposes, a calculation on [14]-annulene (9) was run using a planar geometry Fig. 2., adapted from the information available. This annulene is known to be non-planar in the crystalline state,<sup>10</sup> and using the classical structure with out-of-plane hydrogens within the ring, a lower total molecular energy was obtained compared with a planar structure, but the wave functions were not significantly different for each case. Using a planar structure provides a direct means of comparison of the molecule with pyrene (8). The simplest molecule, benzene (1), has the highest symmetry at  $D_{6h}$  while molecules 2-5 are  $D_{2h}$ . For comparison purposes of molecular orbitals in 1-5, it is convenient to "reduce" benzene to  $D_{2h}$  symmetry and to take the longer molecular axis (in 2-5) as the Z axis as shown in Fig. 1 for 2 and 3. There are eight representations in  $D_{2h}$  symmetry, four of which are sigma ( $\sigma$ ) representations, and four of which are pi ( $\pi$ ) representations. The maximum number of positions which can be related in  $D_{2h}$  symmetry is four. For anthracene (3) Fig. 1, the signs of the  $1s_c(\sigma)$  and  $2p_{xc}(\pi)$  wave functions for positions (a) in anthracene are shown in Table 1. The representations  $b_{1u}$  and  $b_{2u}$  are interchangeable and depend on which way round the molecular axes y and z are chosen. The same applies to the representations  $b_{1g}$  and  $b_{2g}$ . Planar biphenyl (6) and pyrene (8) also have  $D_{2h}$  symmetry, and for convenience the molecular axis of pyrene is chosen to relate to naphthalene as shown in Fig. 2. Phenanthrene, alone, of this series has a lower symmetry of  $C_{2v}$ .

Fig. 4

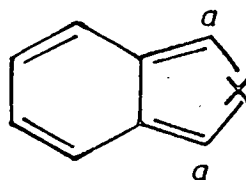
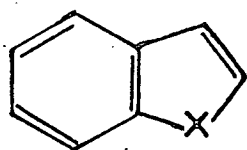
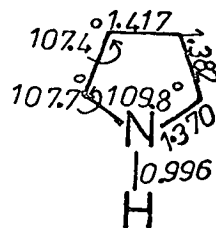
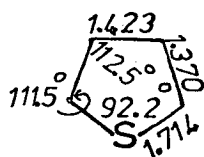
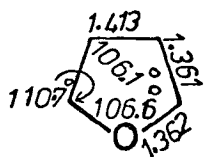


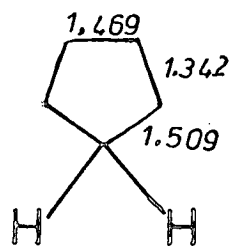
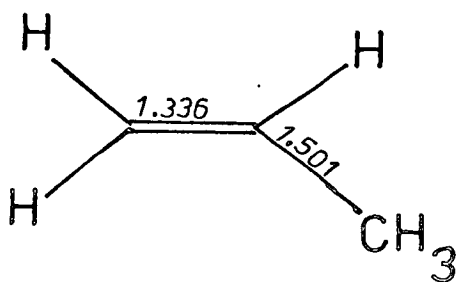
Fig. 5



10

11

12



13

14

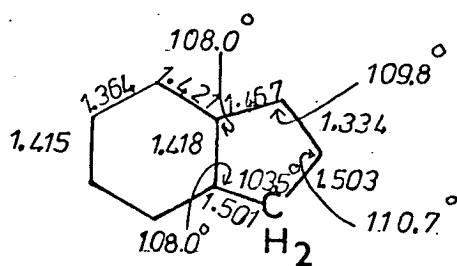
The bicyclic heterocycles I (: I = the series containing only one hetero atom and related hydrocarbons).

Geometries for the parent members of series A and B, Fig. 4, where X is CH<sub>2</sub>, O, S, or NH are not known at present. For X = O, S or NH, a theoretical geometry could well be constructed by the fusion of half of naphthalene (2) Fig. 1, with the corresponding five membered ring heterocycle, and for molecules with X = CH<sub>2</sub>, from half of naphthalene with propene to give A, X = CH<sub>2</sub> or with cyclopentadiene to give B, X = CH<sub>2</sub>. The geometries of furan(10)<sup>11</sup>, thiophen(11)<sup>11</sup>, pyrrole(12) propene (13)<sup>12</sup>, and cyclopentadiene (14)<sup>13</sup>, are shown in Figs. 1 and 5. The resulting geometries of the bicyclic compounds, thus constructed for X = CH<sub>2</sub>, O, S or NH are shown for series A in Fig. 6, 15 - 18, and for series B in Fig. 7, 20 - 23.

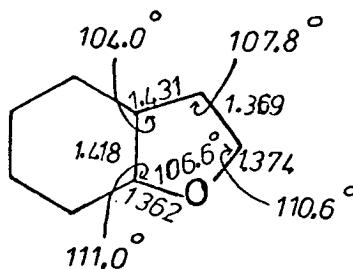
There are several 3 - substituted derivatives of indole with known geometry, 19 (a) - (d)<sup>14</sup>. On close investigation, it was obvious that 19(a) and (b) had identical bond lengths to within 0.01 - 0.02Å, while the other two were rather more different, one being a trinitro-benzene complex<sup>14d</sup>, and the other a disubstituted indole<sup>14b</sup>. It was decided therefore, to adopt 19(a)<sup>14a</sup> as a basis of the indole structure and its geometry is shown in Fig. 6.

It is apparent that the carbocyclic rings of 18 and 19a differ significantly, while the C(2) - C(3) bond of pyrrole (12) is significantly longer than the C(2) - C(3) bond of 19a, making the C(3) - C(3a) bond of 19a much longer than that in pyrrole. Clearly the geometries do not agree well, and although fusion of six- and five- membered rings may well produce significant changes in bond lengths, the effect of

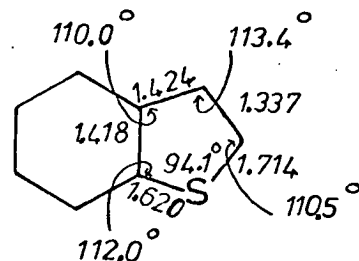
Fig. 6



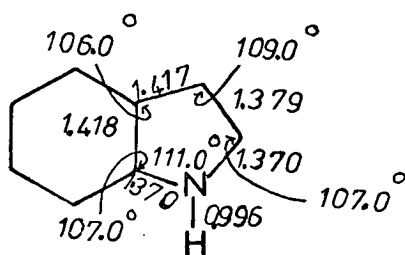
15



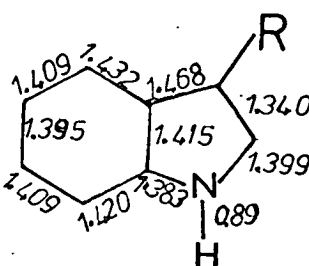
16



17



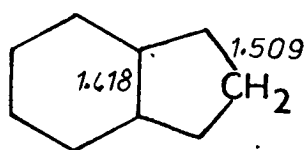
18



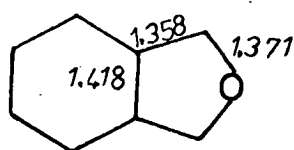
19

- a.  $R = \text{CH}_2\text{COOH}$
- b.  $R = \text{CH}_2\text{CH}(\text{NH}_3^+\text{Cl}^-)\text{COOH}$
- c.  $R = \text{CH}_2\text{CH}_2\text{NH}_3^+\text{X}^-$
- d.  $R = \text{CH}_3$ , *s*-trinitro benzene complex.

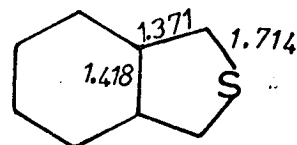
Fig. 7



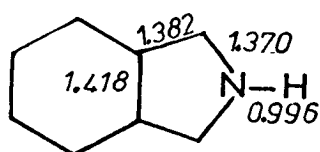
20



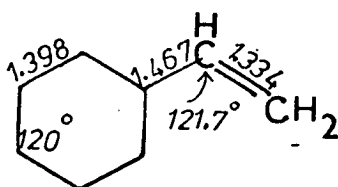
21



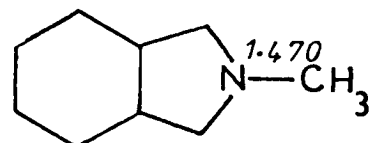
22



23



24



25

the very polar 3-acetic acid group must be considered.

It was decided to run two semiempirical INDO calculations on the indole molecule using:- a) the 3-acetic acid geometry, and b) the geometry produced by fusion of half of naphthalene with pyrrole. As INDO calculations are known to reproduce differences in geometric features fairly well, the comparison of total molecular energies for the two geometries would indicate the better geometry. The total energies produced were for a) the 3-acetic acid geometry, -71.2020 a.u. and for b) the fusion of half naphthalene with pyrrole, -71.2616 a.u.

It appears that the constructed geometry is the better one, and that the indole 3-acetic acid molecule may well be distorted relative to the parent molecule due to the highly polar nature of the 3-acetic acid group.

Bond lengths have been calculated for molecules 2<sup>15</sup>, 16 - 18<sup>16</sup> and 22<sup>17</sup> from self consistent bond length, ( $R_{ij}$ ) - bond order, ( $P_{ij}$ ) calculations using the relationship,  $R_{ij} = A - BP_{ij}$  where A and B are empirical parameters<sup>18</sup> based on molecules such as benzene, and ethylene. However, these calculations do not lead to bond angles, and are therefore insufficient for constructing geometries. However, calculations of this type on naphthalene are in reasonable agreement with experimental bond lengths although the  $C_{\alpha\beta}$  bond is always calculated to be slightly too long. Nevertheless, these calculations show that the carbocyclic parts of 16 - 18 are very similar to that of naphthalene, confirming the choice of the carbocyclic ring geometries of 16 - 18 to be that of half naphthalene, as a reasonable one.



The fusion of the geometries of benzene <sup>20</sup> with ethylene <sup>21</sup> produced a geometry for the related hydrocarbon, styrene (24) Fig. 7. The integrals for benzene <sup>†</sup> were already available and were merged with those of ethylene.

The last related molecule for which a geometry was required was N-methyl isoindole. This was constructed using the isoindole skeleton (23) with the addition of the methyl group, and is shown in Fig. 7, (25). The N-C bond length was taken to be that of N-methyl pyrrole whose structure has been determined by microwave spectroscopy <sup>22</sup>.

For the Kekulé series A, the molecules are planar and of  $C_s$  symmetry giving orbitals of  $a'$  and  $a''$  symmetry corresponding to  $\sigma$  and  $\pi$  type orbitals respectively. The Atmol 2 programme was allowed to find the ground state electron distribution but from theoretical considerations, five  $a''$  orbitals producing a  $10-\pi$  electron system were expected for indene (15), benzofuran (16), and indole (18), and six  $a''$  orbitals for benzothiophen (17), the additional  $a''$  orbital being due to the  $2p_x$  orbital of sulphur (if  $y-z$  is the plane of the molecule).

The related hydrocarbon, styrene is also of  $C_s$  symmetry and shows four  $a''$  orbitals as expected, three related to the benzene ring, and one to the ethylene grouping.

For the quinoid series B, Fig. 4, the molecules have  $C_{2v}$  symmetry, and calculations were expected to show orbitals of symmetry types  $a_1$ ,  $a_2$ ,  $b_1$  and  $b_2$ , the representations  $b_1$  and  $a_2$  giving the  $\pi$ -orbitals. The signs for the  $1s_C(\sigma)$  and  $2p_x C$  wave functions for positions (a) in series B, Fig. 4, are shown in Table 2.

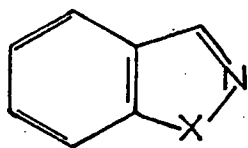
† Thanks to W. Moyes for providing the benzene integrals.

The geometry of indene (15) <sup>19</sup> has been studied by electron diffraction but insufficient details were published to give a complete geometry, making it necessary to use a constructed geometry. It was therefore decided to use the constructed geometries as shown in Fig. 6, 15 - 18 for series A. Series B has no derivatives of known geometry, and so geometries were constructed from half of naphthalene (1) with the corresponding heterocycle for 21 - 23 Fig. 7, and with cyclopentadiene for 2H-indene (20).

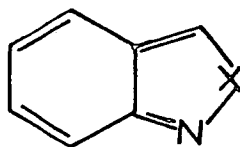
The use of a constant geometry for the carbocyclic ring in both series has the added advantage of saving computer time, in that many of the electron repulsion integrals, which take up much of the computational time, need only be calculated once, and can be reused in different calculations. The initial calculation was run on naphthalene,  $C_{10}H_8$ , consisting of 58 contracted basis functions. If half the naphthalene geometry,  $C_6H_4$ , is present in each calculation then the electron repulsion integrals from 34 contracted basis functions are preserved for further calculations.

The number of electron repulsion integrals is given by  $S = \frac{1}{2} (N^4/2 + N^3 + 3N^2/2 + N)$  where  $N$  is the number of contracted basis functions. Considering indole (18), which has 52 contracted basis functions, the total number of electron repulsion integrals is 950,131, those contributed by the carbocyclic ring ( $N = 34$ ) being 344,352. Roughly 2/5 of the computation is saved by merging the new integrals from indole (from  $C_2H_3N$ ) with those of the already computed half of naphthalene. As this can be done for each of the molecules 15 - 18 and 20 - 23 the saving overall is considerable.

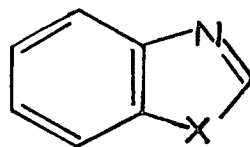
Fig 8



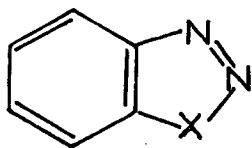
26 a X=NH  
b X=O  
c X=S  
d X=NCH<sub>3</sub>



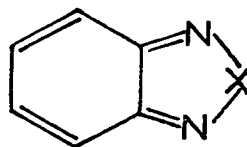
27 a X=NH  
b X=O  
c X=S  
d X=NCH<sub>3</sub>



28 a X=NH  
b X=O  
c X=S

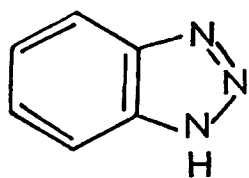


29 a X=NH  
b X=S  
c X=NCH<sub>3</sub>

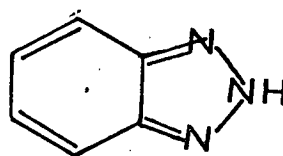
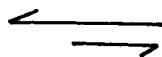


30 a X=NH  
b X=O  
c X=S  
d X=NCH<sub>3</sub>

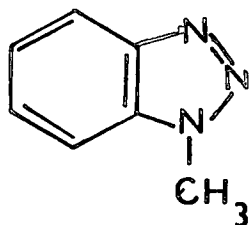
Fig.9



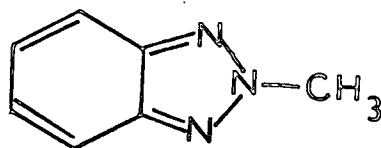
29 x=NH  
STABLE



30 x=NH  
UNSTABLE



29 x=NCH<sub>3</sub>  
FIXED KEKULÉ



30 x=NCH<sub>3</sub>  
FIXED QUINOID

It should be noted at this point that if the Atmol 2 programme did not produce the required symmetry state for the molecule, then the 'SWAP' facility was put into operation. This allowed the symmetry of virtual and occupied orbitals to be swapped, and the SCF procedure repeated, thus producing a molecule with the desired symmetry in the occupied orbitals. For N-methyl isoindole (25) the molecule was arranged to give  $C_s$  symmetry. However, the rotation barrier for methyl substituted olefins is known to be very low<sup>23</sup>, and normally less than 1 K cal. mole<sup>-1</sup> in energy., and therefore the choice of orientation of the methyl group is relatively unimportant for the purposes of the calculation.

The bicyclic heterocycles II (:II = the series containing two or more hetero atoms).

The natural extension of the bicyclic heterocycles was to proceed to similar systems containing two or more heteroatoms. The first series which was studied was the aza-derivatives of series A and B, Fig. 4, as shown in 26 and 27, Fig. 8, where X = NH, O, S, and N-Me.

This was further extended to include the series 28, where X = NH, O and S and the bicyclic molecules with three heteroatoms shown as 29, where X = NH, S, N-Me, and 30, where X = NH, O, S and N-Me, Fig. 8. In these compounds with X = NH, a tautomeric system is produced, and as only the 1H - series is stable unless there is a substituent, R, on the nitrogen, the calculations were extended to include an N-methyl substituent at both the 1H and 2H positions. This fixes the molecules in either a Kekulé or a quinoid like structure, an example of which is shown in Fig. 9, for 29, and 30 with X = NH, and N-methyl.

Fig.10

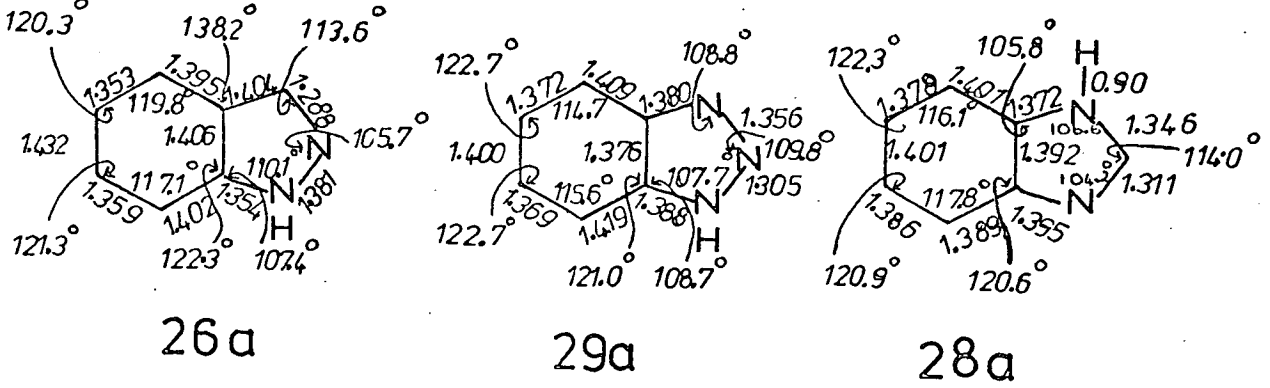
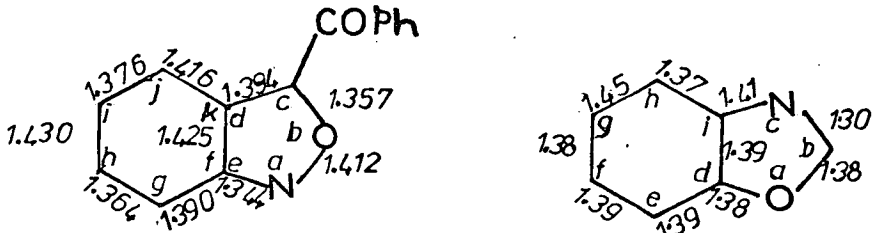


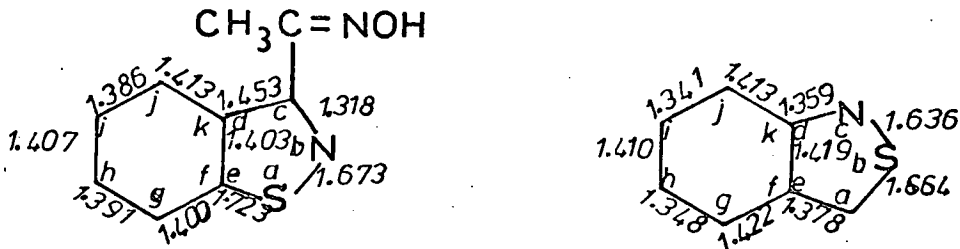
Fig.11



$\hat{a} = 104.6^\circ$   $\hat{b} = 110.5^\circ$   $\hat{c} = 108.5^\circ$   $\hat{d} = 104.4^\circ$   
 $\hat{e} = 111.9^\circ$   $\hat{f} = 121.7^\circ$   $\hat{g} = 117.0^\circ$   $\hat{h} = 123.0^\circ$   
 $\hat{i} = 120.0^\circ$   $\hat{j} = 117.9^\circ$   $\hat{k} = 120.1^\circ$

$\hat{a} = 104.6^\circ$   $\hat{b} = 112.7^\circ$   $\hat{c} = 107.4^\circ$   
 $\hat{d} = 126^\circ$   $\hat{e} = 112^\circ$   $\hat{f} = 124^\circ$   
 $\hat{g} = 121^\circ$   $\hat{h} = 115^\circ$   $\hat{i} = 121^\circ$

Fig.12



$\hat{a} = 94.5^\circ$   $\hat{b} = 111.5^\circ$   $\hat{c} = 114.8^\circ$   $\hat{d} = 110.3^\circ$   
 $\hat{e} = 109.0^\circ$   $\hat{f} = 122.9^\circ$   $\hat{g} = 117.2^\circ$   $\hat{h} = 120.6^\circ$   
 $\hat{i} = 122.1^\circ$   $\hat{j} = 118.1^\circ$   $\hat{k} = 119.1^\circ$

$\hat{a} = 109.4^\circ$   $\hat{b} = 97.7^\circ$   $\hat{c} = 107.1^\circ$   $\hat{d} = 116.3^\circ$   
 $\hat{e} = 109.5^\circ$   $\hat{f} = 120.3^\circ$   $\hat{g} = 117.3^\circ$   $\hat{h} = 122.8^\circ$   
 $\hat{i} = 121.0^\circ$   $\hat{j} = 119.1^\circ$   $\hat{k} = 119.5^\circ$

A number of X-ray crystal structures of 26-30 have been published in recent years, and detailed study of the individual molecules shows that in general they fall into two main categories:- (a) those with carbocyclic rings which are similar to benzene; (b) those with carbocyclic rings similar to naphthalene. Thus it was expedient to use either the benzene structure, 1, Fig. 1 or one half of naphthalene, 2, Fig. 1 to represent the carbocyclic rings in these molecules, and to 'MERGE' the appropriate structure with the new heterocyclic ring.

Considering first, molecules with only nitrogen heteroatoms, the crystal structures for 1H - indazole <sup>24</sup>, (26a, Fig. 8) and 1H - benzotriazole <sup>25</sup>, (29a, Fig. 8) show a considerable similarity to naphthalene in the carbocyclic ring whereas 1H - benzimidazole <sup>26,27</sup> (28a, Fig. 8) has a benzene like carbocyclic ring. The X-ray structures <sup>24,25, 26</sup> for these molecules are shown in Fig. 10.

For molecules containing nitrogen and oxygen in the heterocyclic ring, structures are only known for substituted compounds. For anthranil <sup>28,29</sup> (27b) only one structure <sup>28</sup> is of high precision and is shown in Fig. 11. The carbocyclic ring bears a close resemblance to naphthalene and the calculation was done by merging half of naphthalene with the heterocyclic ring structure. The only structure which exists for benzoxazole <sup>30</sup> (28b) is for a copper chelated compound and is not of high precision Fig. 11. The structure shows an abnormally long C(4) - C(5) bond of 1.45 Å which is inconsistent with the rest of the carbocyclic ring, and so was ignored. However, the structure is very similar to benzimidazole (27a) Fig. 8, and

so the calculation for benzoxazole was done by replacing the second nitrogen of benzimidazole by an oxygen atom. The geometry of 1,2 - benzisoxazole (26b) was taken to be that of indazole (26a) with N(2) replaced by O.

For the compounds containing nitrogen and sulphur, two geometries were available for a substituted compound of 26,  $X = S$ , but one <sup>31</sup> was more accurate than the other <sup>32</sup>, and was chosen as a basis for the heterocyclic ring. The structure is shown in Fig. 12 for the substituted benzisothiazole (26c). On inspection, it was clear that the carbocyclic ring bore a resemblance to benzene, and a geometry was constructed from benzene fused with the heterocyclic ring. Only one structure was available for 2, 1 - benzisothiazole <sup>33</sup>, (27c), and showed a carbocyclic ring, Fig. 12, close to naphthalene. A 'MERGE' of the integrals from half of naphthalene with the new integrals from the heterocyclic ring of 2, 1 - benzisothiazole formed the basis for the calculation.

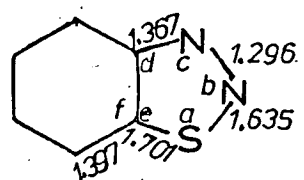
No reasonable geometries have as yet been published for 29,  $X = S$ , Fig. 8, and 30,  $X = O$  and  $S$ , Fig. 8, and so were constructed in the following manner. For 29,  $X = S$ , a constructed geometry of 1, 2, 3 - thiadiazole was fused with benzene as shown in Fig. 13, and for 30,  $X = O, S$ , the corresponding bicyclic heterocycles, anthranil (27b) and 2, 1 benzisothiazole (27c), were used as a basis, giving the geometries in Fig. 13.

In most cases the molecules were of  $C_s$  symmetry giving a  $\sigma/\pi$  separation of orbitals and  $10\pi$  - electrons in the valence shell. The higher symmetry molecules, 30,  $X = N-Me, O, S$  Fig. 8, were of  $C_{2v}$  symmetry giving orbitals in the symmetry classes  $a_1, a_2, b_1,$  and  $b_2$ , the  $\pi$  - orbitals coming from  $b_1$

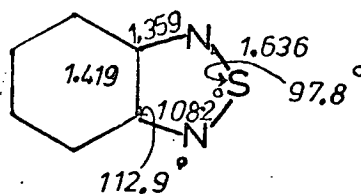
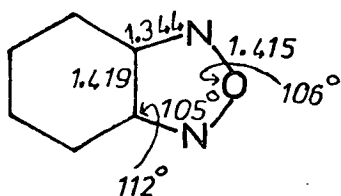
and  $a_2$ .

Where N-methyl compounds were used, the methyl geometry was taken from toluene and used with the N-methyl distance of N-methyl pyrrole<sup>22</sup>. The N-methyl group was made to lie along the original N-H line.

Fig.13



$$\begin{aligned} \hat{a} &= 90.3^\circ & \hat{b} &= 115.5^\circ & \hat{c} &= 114.0^\circ & \hat{d} &= 109.5^\circ \\ \hat{e} &= 111.0^\circ & \hat{f} &= 120.0^\circ & & & & \end{aligned}$$





R E F E R E N C E S

1. A. Almennigen, O. Bastiansen and L. Fernholt, Kgl. Nor. Vidensk. Selsk, Skr. No. 3 1958; A. Langseth and B.P. Stoicheff, Can. J. Phys., 1956, 34, 350.
2. D.J.W. Cruickshank and R.A. Sparks, Proc. Roy. Soc. A, 1960, 258, 270
3. A. Almennigen, O. Bastiansen and F. Dyvik, Acta Cryst., 1961, 14, 1056.
4. R. Mason, Acta Cryst., 1964, 17, 547; M.S. Lehmann G.S. Pawley, Acta Chem. Scand., 1972, 26, 1996
5. J.M. Robertson, V.C. Sinclair and J. Trotter, Acta Cryst., 1961, 14, 697.
6. R.B. Campbell, J.M. Robertson and J. Trotter, Acta Cryst., 1961, 14, 705
7. J. Trotter, Acta Cryst., 1961, 14, 1135.
8. M.I. Kay, Y. Okaya and D.E. Cox, Acta Cryst., 1971 27B, 26.
9. A.C. Hazell, F.K. Larsen and M.S. Lehmann, Acta Cryst., 1972, 28B, 2977.
10. C.C. Chiang and I.C. Paul, J. Amer. Chem. Soc., 1972, 94, 4741.
11. a) B. Bak, D. Christensen, W.B. Dixon, L. Hansen - Nygaard, J. Rastrup - Andersen and M. Schöttlander, J Mol. Struct., 1962, 9, 124.  
b) B. Bak, D. Christensen, L. Hansen - Nygaard and J. Rastrup - Andersen, J. Mol. Struct., 1961, 7, 58  
c) L. Nygaard, J Nielsen, J. Tormod, J Kirchheiner,

- G. Maltesen, J. Rastrup - Andersen and G.O. Sorensen, J. Mol. Struct., 1969, 3(6), 491.
12. D.R. Lide and D. Christensen, J. Chem. Phys., 1961, 35, 1374
13. L.H. Sharpen and W.H. Lavine, J. Chem. Phys., 1965, 43, 2765.
14. a) I.A. Karle, K. Britts and P. Gum, Acta Cryst., 1964, 17, 496.  
b) T. Takigawa, T. Ashida, Y. Sasada and M. Kakudo, B.C.S. Japan, 1966, 39, 2369  
c) I.L. Karle, K.S. Dragonette and S.A. Brenner, Acta Cryst., 1965, 19, 713  
d) A.W. Hanson, Acta Cryst., 1964, 17, 559
15. a) F. Momicchioli and A. Rastelli, J. Mol. Spectroscopy, 1967, 22, 310.  
b) M.J.S. Dewar and N. Trinajstic, J. Chem. Soc. (A), 1971, 1220  
c) M.J.S. Dewar and G.J. Gleicher, J. Chem. Phys., 1966, 44, 759  
d) M.J.S. Dewar, A.J. Harget, N. Trinajstic and S.D. Worley, Tetrahedron, 1970, 26, 4505  
e) J. Fabian, A. Mehlhorn and R. Zahradnik, J. Phys. Chem., 1968, 72, 3975
16. a) C. Aussems, S. Jaspers, G. Leroy and F. Van Remoortere, Bull. Soc. chim. belges, 1969, 78, 479  
b) N. Trinajstic and A. Hinchcliffe, Z. Phys. Chem. Neue Folge, 1968, 59, 271.  
c) L. Klasinc, E. Pop, N. Trinajstic and J.V. Knop, Tetrahedron, 1972, 28, 3465.

- d) A. Skancke and P.N. Skancke, *Acta Chem. Scand.*, 1970, 24, 23
- e) R.A. Sallavanti and D.D. Fitts, *Internat. J. Quant. Chem.*, 1969, 3, 33
- f) G. Favini and A. Gamba, *J. Chem. Phys.*, 1967, 64, 1443.
17. M.J.S. Dewar and N. Trinajstić, *J. Amer. Chem. Soc.*, 1970, 92, 1453
18. C.A. Coulson and A. Golebiewski, *Proc. Phys. Soc.*, 1961, 78, 1310
19. J.F. Southern, L. Schäfer, K. Brendhaugen and H.M. Seip, *J. Chem. Phys.*, 1971 55, 2418
20. G.E. Bacon, N.A. Curry and S.A. Wilson, *Proc. Roy. Soc.*, 1964, A, 279, 98.
21. L.S. Bartell, E.A. Roth, C.D. Hollowell, K. Kuchitsu and J.E. Young, *J. Chem. Phys.*, 1964, 42, 2683
22. L.V. Vilkov, P.A. Akishin and V.M. Presnyakova, *ZHUR. Strukt. Khim.*, 1962, 3, 3
23. a) G.J. Karabatsos and D.J. Fenoglio, *Topics in Heterochemistry*, 1970, 5, 187.
- b) S.L. Hsu and W.H. Flygare, *Chem. Phys. Letters*, 4, 317.
24. A. Escande and J. Lapasset, *Acta. Cryst.*, 1974, B30, 2009
25. A. Escande, J.L. Galigne and J. Lapasset, *Acta. Cryst.*, 1974, B30, 1490
26. C.J. Dik - Edixhoven, H. Schenk and H. Van der Meer; *Cryst. Struct. Comm.*, 1973, 2, 23

27. A. Escande and J.L. Galigne, *Acta Cryst.*, 1974, B30, 1647
28. M. Sundaralingam and G.A. Jeffrey, *Acta Cryst.*, 1962, 15, 1035; *ibid*, 1963, 16, A61.
29. M. Sax, J. Pletcher, D. Scholtz, R.M. Gerkin and J.L. Pinkus, *J. Chem. Soc., Perkin II*, 1976, 560.
30. P. Stenson, *Acta Chem. Scand.*, 1969, 23, 1514
31. A. Braibanti, M.A. Pellinghelli, A. Tiripicchio and M. Tiripicchio - Camellini, *Acta Cryst.*, 1973, B29, 43.
32. A. Gaetani, T. Vitali, A. Mangia, M. Nardelli and G. Pelizzi, *J. Chem. Soc., Perkin II*, 1972, 2125.
33. M. Davis, M.F. Mackay and W.A. Denne, *J. Chem. Soc., Perkin II*, 1972, 565.

### Molecular Energies

Calculated total energies,  $TE_{\text{calc.}}$ , tend to be large as a result of the constituent atom energies being fairly large. They are, therefore not in themselves very useful, and can only be used comparatively in limited cases. In instances, where experimental total energies,  $TE_{\text{expt.}}$  are available, they can be compared with  $TE_{\text{calc.}}$ . Determination of  $TE_{\text{expt.}}$  requires measurement of the heat of formation,  $\Delta H_f^0$ , of a compound from its elements in their standard state which is taken as gaseous at  $25^\circ\text{C}$  and one atmosphere pressure. This gives  $TE_{\text{expt.}}$  of, for example, hexane,  $\text{C}_6\text{H}_{14}$ , as:-

$$TE_{\text{expt.}} = \Delta H_f^0[\text{C}_6\text{H}_{14}(\text{g})] + 6\Delta H_f^0[\text{C}(\text{g})] + 14\Delta H_f^0[\text{H}(\text{g})]$$

The second term represents the heat of sublimation of carbon, and is not easily measured giving rise to differing values for  $\Delta H_f^0[\text{C}(\text{g})]$ . However, reliable values of  $\Delta H_f^0(\text{g})$  for most of the common elements<sup>1</sup> are now available, and in this work the recommended values in reference 1 are used.

For the thermochemist, the easiest value to determine for a compound is its heat of atomisation,  $\Delta H_a$ ; this is equated with the energy change involved when a molecule is converted into its constituent atoms, and can be used to give  $\Delta H_f^0$  for the compound using:-

$$\text{For hexane } \Delta H_a[\text{C}_6\text{H}_{14}] = 6\Delta H_f^0[\text{C}(\text{g})] + 14\Delta H_f^0[\text{H}(\text{g})] - \Delta H_f^0[\text{C}_6\text{H}_{14}(\text{g})]$$

As the  $TE_{calc.}$  is a function of the size of basis set used, the larger the basis set, the nearer  $TE_{calc.}$  will approach  $TE_{expt.}$  Few  $TE_{expt.}$  exist for the more complicated molecules of basic interest to organic chemists and so in these cases no comparison is available.

A use for direct use of  $TE_{calc.}$  has been found in the field of isomerism where the relative stability of isomers can be determined provided the calculations are performed using an identical basis set, and each isomer is studied at its equilibrium geometry.

The isomers must also be of similar structure and electron distribution e.g. naphthalene and azulene, Fig. 1 A, to give each a similar correlation energy. Isomeric pairs of very different geometries e.g. tetrahedrane and cyclobutadiene Fig. 1 B, are likely to have substantially different correlation energies, and therefore the calculated total energy difference is likely to vary considerably from the experimental one. It has also been shown that for isomers of very different structures<sup>1</sup>, the heats of formation of each are very different due to the very different nature of the chemical bonds involved. The isomer with the more negative energy will be the more stable. In some cases, the geometry of both isomers is not known, perhaps due to one species being unstable. The geometry of the dominant isomer is often well determined, and will lead to a relatively accurate total energy. The calculation of the unstable isomer must be performed using a constructed geometry, and may lead to a high (less negative) total energy. This could cause an overestimation

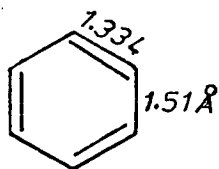
in the total energy difference between the isomers. In practice, where cyclic molecules are concerned, constructed geometries are highly constrained by internal bond angles, and the geometries are likely to be close to optimal. Often, several analogues of known geometry exist, and can be used as a basic guide. For cyclic molecules in particular, therefore, a good estimate can be obtained of the total energy without difficulty.

With lack in sensitivity of TE, it would seem more appropriate to compare  $\Delta H_a$  directly with some calculated quantity.  $\Delta H_a$  can be expressed<sup>1</sup> as

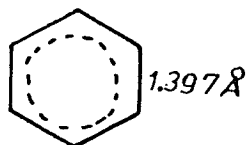
$$\Delta H_a = \left\{ \begin{array}{l} \text{Standard chemical} \\ \text{bonds} \\ (1) \end{array} \right\} - \left\{ \begin{array}{l} \text{Destabilisation} \\ \text{energies} \\ (2) \end{array} \right\} + \left\{ \begin{array}{l} \text{Stabilisation} \\ \text{energies} \\ (3) \end{array} \right\}$$

where (2) could arise from such factors as ring strain or steric hindrance, and (3) from delocalisation. As measurement of  $\Delta H_a$  produces a single value, it was thought that standard bond energies, found from molecules in which 2 and 3 did not exist, could be used in the above expression to determine a value of (2+3) for more complicated molecules. This would give a direct measure of stabilisation due to various effects in a molecule, and in particular, could be used to determine the degree of aromaticity in cyclic molecules. Comparison of experimental and calculated results could then be made on a much more sensitive scale. The combined effect of 2 and 3 became known as Resonance Energy, and led to much work being done in determining the standard bond energies to be used.

Fig. 2



*Kekulé*



*Real.*

Fig. 5

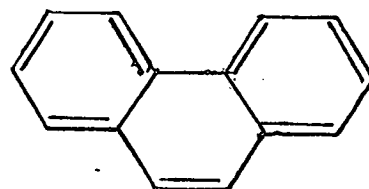
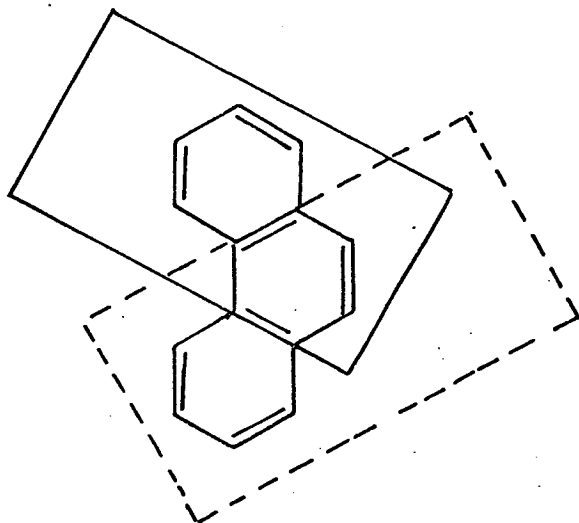
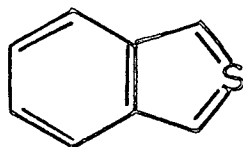


Fig. 6





### Resonance energy

The original work in this field was pioneered by Pauling<sup>2</sup> and Ingold<sup>3</sup> who defined resonance energy, RE, as the difference in the molecular energy and that based upon a total energy obtained from a set of classical bonds. For example, the RE of benzene was given by:-

$$RE = E_{\text{Total}} - ((6 \times E_{\text{C-H}}) + (3 \times E_{\text{C=C}}) + (3 \times E_{\text{C-C}}))$$

RE was quickly adopted as a basis for comparing aromaticity of molecules with benzene as the reference. However, it soon became apparent that refinements were required in the procedure to take account of the fact that the ideal classical geometry and the real molecule were often considerably different, as shown Fig. 2 by the Kekulé form of benzene which has three short and three long C-C bonds and the real structure which has six C-C bonds of equal length. It was, therefore, necessary when comparing real and ideal structures to take bond expansion and compression into account. This was first done by Dewar and Schmeising<sup>4,5,6</sup> who calculated the energy  $E_1$ , required to compress three C-C bonds (1.51Å) in a ring of six carbon atoms to 1.334Å to create a cyclohexatriene structure, and the energy,  $E_2$ , required to compress six C-C bonds (1.51Å) to 1.397Å to give benzene. They found  $E_1$  to be greater than  $E_2$ , and  $E_1 - E_2$  was designated the  $\sigma$ -strain relief energy, and indicated that benzene was stabilised relative to cyclohexatriene by this amount which they claimed could not be included in the resonance energy. A more recent revision of the definition of RE was also put forward by Dewar<sup>7</sup>, and states that the RE of a cyclic molecule is the difference in total energy between the cyclic species and the most closely related acyclic species.

This automatically assumes that all acyclic molecules have no RE, a fact which Dewar<sup>4</sup> had already concluded from his own results, by attributing any gain in energy of an acyclic species relative to a set of classical single and double bonds to stabilisation energy gained from changes in bond energy due to hybridisation. However, thermochemical studies<sup>1</sup> suggest that even buta-1,3-diene has a low RE. Also the central C=C bond, and the C-C bonds in hexatriene are respectively longer and shorter than expected on a classical basis, and this could well be due to RE. Yet this molecule has zero RE on Dewar's<sup>7</sup> scale, and is the reference point for benzene. Using Dewar's method, a very large annulene can have RE due to its cyclic nature yet the corresponding acyclic molecule must have zero RE. Yet one would imagine that if the ring was large enough, there would be little constraint on the molecular geometry in the cyclic form, and the open and closed species become basically similar.

In the present work RE<sup>8</sup> is defined on the basis of the original Pauling procedure as the difference between the molecular energy, and the sum of the classical bond energies for the molecule, where the latter are calculated from small non-conjugated acyclic molecules using identical basis sets for all calculations. For example, the C-H bond energy,  $E_{C-H}$ , is taken as one quarter of the molecular energy of methane,  $E_{CH_4}$ , and is assumed constant for all molecules.

$$E_{C-H} = \frac{1}{4} E_{CH_4}$$

Table 1 Molecular and derived bond energies (a.u.) from non-aromatic compounds

CH <sub>4</sub>	-40.10176	C <sub>2</sub> H <sub>4</sub>	-77.83154	(CH <sub>2</sub> =CH) <sub>2</sub>	-154.50920		
CH	-10.0254	C≡C	-37.7296	C <sub>sp<sup>2</sup></sub> -C <sub>sp<sup>2</sup></sub>	-18.8972		
CH <sub>2</sub> =CHOH	-152.46202	CH <sub>2</sub> =CHNH <sub>2</sub>	-132.54130	NH <sub>3</sub>	-56.01986	H <sub>2</sub> S(sp)	-397.84420
C <sub>sp<sup>2</sup></sub> -O	-46.7561	C <sub>sp<sup>2</sup></sub> -N	-27.5378	NH	-18.6733	SH(sp)	-198.9221
CH <sub>2</sub> =CHSH(sp)	-474.51823	CH <sub>2</sub> =CHSH(spd)	-474.61525	H <sub>2</sub> O	-75.79988	H <sub>2</sub> S(spd)	-397.94040
C <sub>sp<sup>2</sup></sub> -S	-207.7848	C <sub>sp<sup>2</sup></sub> -S(spd)	-207.8391	OH	-37.8999	SH(spd)	-198.9702
NH <sub>2</sub> SH(sp)	-452.68460	N <sub>2</sub> H <sub>4</sub> <sup>†</sup>	-110.85977	N <sub>2</sub> H <sub>2</sub> <sup>*</sup>	-109.63162		
NH <sub>2</sub> SH(spd)	-452.78239	N-N	-36.1662	N=N	-72.2850		
N-S(sp)	-216.4159						
N-S(spd)	-216.4655						

† N<sub>2</sub>H<sub>4</sub> geometry gauche-twisted.

\* N<sub>2</sub>H<sub>2</sub> experimental geometry, cis hydrogens.

Similarly,  $E_{C=C}$ , is calculated from the molecular energy of ethylene, and using  $E_{C-H}$ .

$$E_{C=C} = E_{C_2H_4} - 4 E_{C-H}$$

For  $E_{C-C}$ , the reference molecule is twisted buta -1,3-diene where the  $\pi$ -orbitals are perpendicular to each other, thus preventing conjugation.

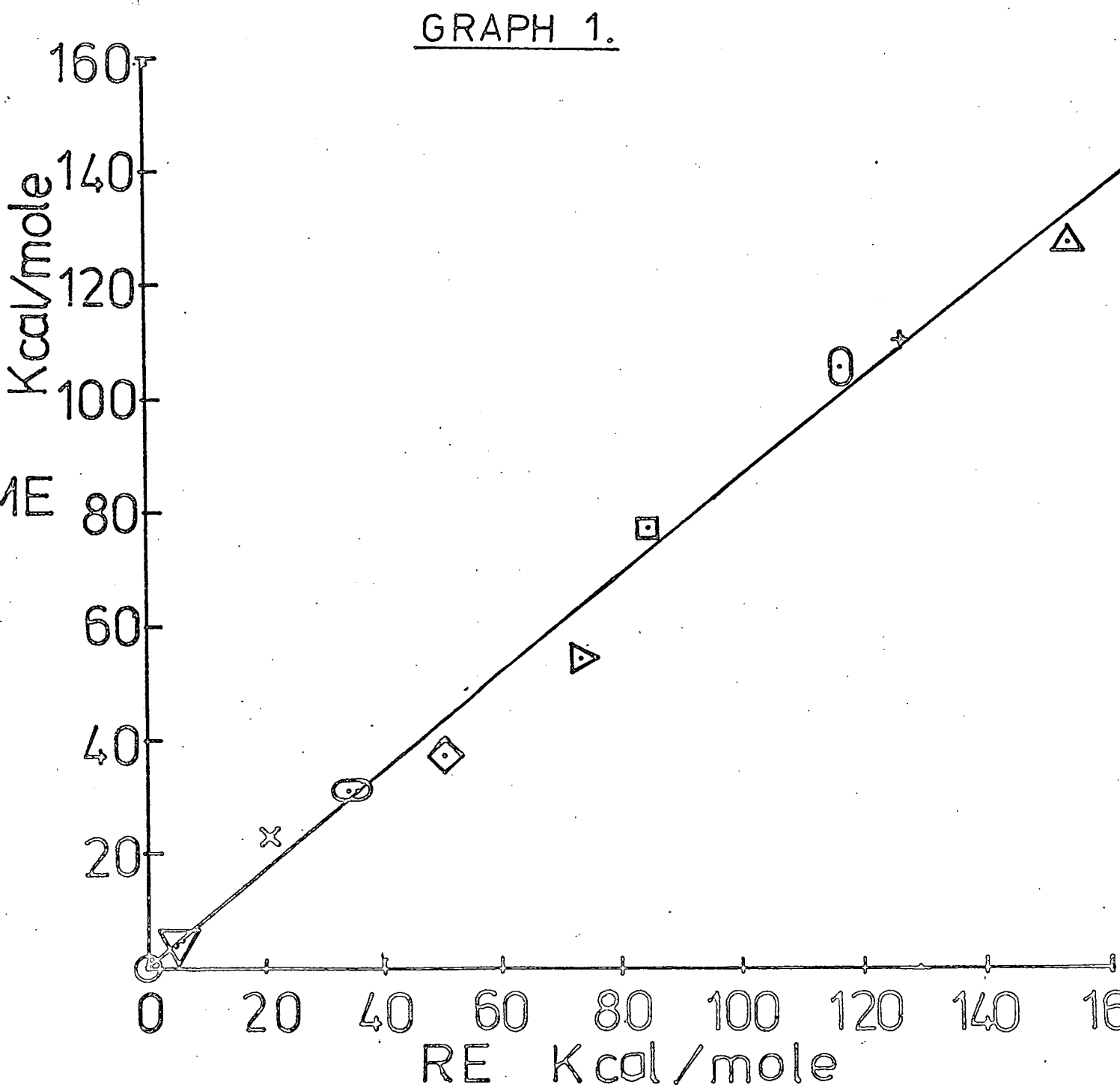
$$E_{C-C} = E_{C_4H_6} - 2 E_{C=C} - 6 E_{C-H}$$

Experimental geometries where known were used for the reference molecules, others were derived from known bond lengths and angles of related molecules. A list of bond energies used, and the molecules from which they were calculated are shown in Table 1.

The necessity to use non-conjugated molecules is shown by the difference in energies of planar and non-planar, vinylamine, ethylene thiol, and vinyl alcohol which are 37,14 and 10  $\text{kJ mol}^{-1}$  respectively. These are too large to be attributed to rotation energy barriers, and are therefore more likely to be caused by interaction of a lone pair with a  $\pi$ -orbital producing RE within the molecule.

A linear correlation is obtained between Pauling's thermochemical mesomeric energies <sup>2</sup> (ME), and REs as measured using the Palmer <sup>8</sup> method, and is shown on Graph 1, Fig. 3, for the molecules ethylene, cyclopentadiene, cyclo-octatetraene, furan, thiophen, pyrrole, indole, naphthalene, anthracene phenanthrene, and pyrene in order of ascending RE  $\text{kcal mol}^{-1}$ . Using a least squares fitting procedure of Pauling's and Palmer's data as shown on the graph gives  $ME = 0.874 RE - 4.47\text{kJ mol}^{-1}$  with standard deviation in slope, intercept and overall of 0.028, 1.823 and 3.876. Using

Fig.3



# GRAPH 2

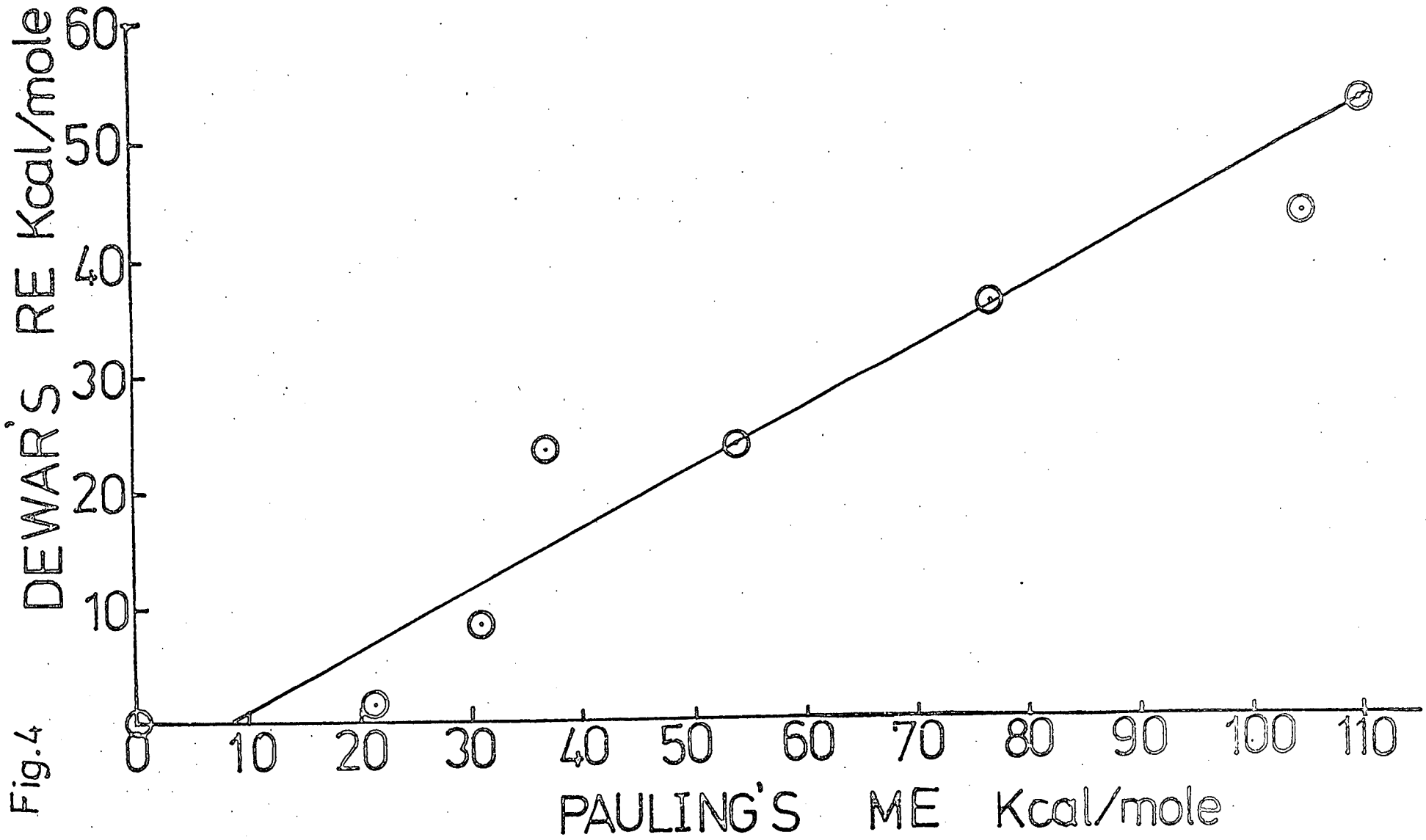
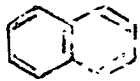


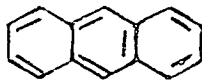
Fig.4



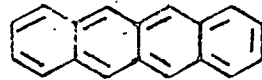
1



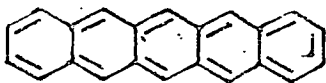
2



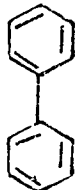
3



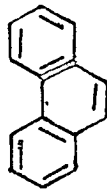
4



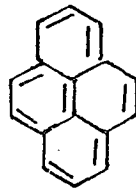
5



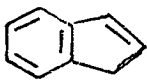
6



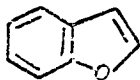
7



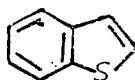
8



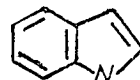
15



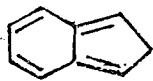
16



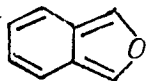
17



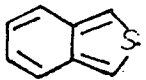
18



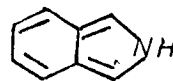
20



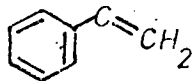
21



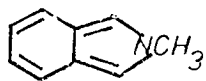
22



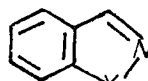
23



24

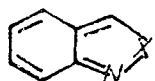


25



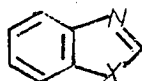
26

a, NH  
b, O  
c, S  
d, NCH<sub>3</sub>



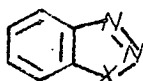
27

a, NH  
b, O  
c, S  
d, NCH<sub>3</sub>



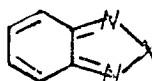
28

a, NH  
b, O  
c, S



29

a, NH  
b, S  
c, NCH<sub>3</sub>



30

a, NH  
b, O  
c, S  
d, NCH<sub>3</sub>

Dewar's definition of RE, both Dewar et al,<sup>9,10</sup> and Hess and Schaad<sup>11,12</sup> have used PPP- $\pi$  electron calculations, and HMO calculations respectively to calculate REs and their results are related as follows:-

$$\text{HMO} = 0.485(\text{PPP}) + 12.35 \text{ kJ mol}^{-1}$$

with standard deviation in slope, intercept and overall of 0.032, 0.991 and 1.464 respectively. Comparing Dewar's data with Paulings MEs gives:-

$$\text{PPP} = 0.494 \text{ ME} - 15.32 \text{ kJ mol}^{-1}$$

with standard deviations as above of 0.062, 4.40 and 5.467. These results are plotted in Graph 2, Fig. 4, for the molecules ethylene, furan, pyrrole, benzene, indole, naphthalene, anthracene and phenanthrene in order of increasing MEs. The HMO REs are approximately half the PPP- $\pi$  REs in magnitude, and the latter are approximately half Pauling's results in magnitude. Therefore the HMO REs are approximately a quarter the magnitude of the Palmer REs, although each set of calculations bears a linear relationship to the others, indicating that the sets are interconnected. The hydrocarbons. Table 2 shows the molecular energies and related properties for the hydrocarbon molecules (1)-(9) (Fig. 2, Section 1). A favourable comparison of total energies for 1, 2 and 6 is obtained with other workers<sup>13-15</sup>, while 3 and 4 show considerably better results than any of the other all electron calculations<sup>16</sup>. For naphthalene, 2, the present calculation using 184 GTOs gives a molecular energy of -382.3712 a.u. which is very close to the result of -382.7883 a.u., given by Buenker and Peyerimhoff<sup>17</sup> whose calculation used 290 gaussian lobe functions, and is 54 a.u. better than Christoffersen's<sup>18</sup> floating spherical



Table 2 Molecular Energies for Hydrocarbons 1 - 9

Energy (a.u.)	Benzene	Naphthalene	Anthracene	Tetracene	Pentacene
Total	-230.11426	-382.37127	-534.62322	-686.86791	-839.11286
Atomisation <sup>a</sup>	-1.469	-2.290	-3.106	-3.916	-4.725
Atomisation <sup>b</sup>	-1.782	-2.751	-3.716	-4.673	-5.630
Atomisation(expt) <sup>c,d</sup>	-2.101	-3.330	-4.555	-5.791	-
Resonance (kJ mol <sup>-1</sup> )	-212	-357	-489	-648	-805
Resonance (kJ mol <sup>-1</sup> ) (expt) <sup>c</sup>	-155	-314	-439	-535	(648)

Footnote to Table 1

- a. Relative to best atom basis
- b. Relative to scaled basis
- c. Reference 1
- d. Reference 19

Table 2 Contd.

Energy	Planar Biphenyl	Phenanthrene	Pyrene	[14]-Annulene
Total	-459.06935	-534.63964	-610.20596	-536.08249
Atomisation <sup>a</sup>	-2.774	-3.123	-3.468	-2.657
Atomisation <sup>b</sup>	-3.339	-3.732	-4.182	-3.387
Atomisation (expt) <sup>c,d</sup>	-4.029	-4.564	-5.108	-
Resonance (kJ mol <sup>-1</sup> )	-516	-532	-645	+1400
Resonance (kJ mol <sup>-1</sup> ) (expt) <sup>c</sup>	-	-460	-531	-

gaussian calculations.

Thermochemical studies of 1-4 indicate an atomisation energy increment between adjacent members of 1.230 a.u.  $1,19,20$  ( $3225 \text{ kJ mol}^{-1}$ ). It has been suggested that the experimental value of the atomisation energy of tetracene 4, is too high by about  $43 \text{ kJ mol}^{-1}$ . However, using the Laidler <sup>32</sup> bond energy scheme, this would lead to a single major discrepancy for the series. This scheme has proved successful for 1-4, and other hydrocarbon series, and it is therefore unlikely that 4 would produce a single major discrepancy. No experimental data exists for 5 but it seems reasonable to assume a similar increment between 4 and 5. The calculated atomisation energies for 1-5 using best atom basis sets show slightly diminishing increments ranging from 0.821 a.u. to 0.809 a.u. between adjacent members and represent approximately 70% of the experimental values. Other single configuration calculations using the H-F method <sup>24</sup> produce similar values. Using the same scaled basis set for the free atoms as for the molecule raises the atomisation energies to about 85% of the experimental values.

The REs of 1-5 show a steady decline in RE per ring added, the RE per ring being 212, 178, 163, 162, 161  $\text{kJ mol}^{-1}$  for 1-5 respectively. This is in agreement with the available thermochemical data, and suggests that naphthalene is less aromatic than twice benzene, and that the degree of aromaticity is declining from 1-5, and the order of aromaticity will be  $1 > 2 > 3 > 4 > 5$ .

Planar biphenyl, 6, shows a RE greater than twice that of benzene by  $92 \text{ kJ mol}^{-1}$  indicating a small degree of inter ring conjugation. When the Kekulé structures of phenanthrene, 7, are drawn, Fig. 5, it can be seen that the molecule may be regarded as either an extended naphthalene structure (a), or an ethylene bridged biphenyl (b). On the basis of REs, (a) might be expected to have a RE of approximately  $1\frac{1}{2}$  x RE of naphthalene and (b) a RE similar to that of biphenyl giving numerical values of approximately  $-471$  and  $-516 \text{ kJ mol}^{-1}$  respectively. The calculated RE is  $-532 \text{ kJ mol}^{-1}$  predicting (b) as the more likely structure, a point which is reinforced by the geometry data which suggests an ethylene bridged biphenyl. Pyrene (8), as already noted (Section 1) can be regarded as a [14]-annulene, with a central double bond or as a bridged biphenyl (Fig.3, Section 1). The geometry data suggested the latter, and as the RE of 8 is calculated to be  $-645 \text{ kJ mol}^{-1}$ , this tends to confirm that 8 is better considered as a bridged biphenyl, since 9 has a positive RE <sup>25</sup> ( $+1400 \text{ kJ mol}^{-1}$ ), showing destabilisation and lack of aromatic character in the planar state.

The molecule 9, however, is known to be slightly buckled owing to the internal hydrogens being out of plane to give maximum separation. In this conformation the RE would be expected to be less positive but would probably not alter substantially. However, when phenanthrene is considered as a singly bridged biphenyl, the RE increase over biphenyl itself is  $-16 \text{ kJ mol}^{-1}$ , whereas considering 8 as a doubly bridged biphenyl gives an increase in RE over biphenyl of

$-129 \text{ kJ mol}^{-1}$ . Therefore, it would appear on the basis of RE that 8 has obtained considerably more stabilisation energy than would be expected from a purely bridged biphenyl indicating a much larger degree of ring conjugation, and therefore increased aromatic stability.

### The bicyclic heterocycles 1

The total energies and related information for 15-24 are shown in Table 3. Indole, benzofuran and their quinoid isomers have been studied using an ST0-3G basis set <sup>33</sup>, and styrene has been studied by a GTO minimal basis calculation <sup>26</sup>, but in all cases the present work is 3-4.0 a.u. better. For benzothiophen, 17, and benzo (c) thiophen, 22, calculations were done with and without d-orbitals. Using LCGO calculations it is necessary to use six 3-d functions representing  $d_{x^2}$ ,  $d_{y^2}$ ,  $d_{z^2}$ ,  $d_{xy}$ ,  $d_{yz}$ ,  $d_{xz}$ . These were contracted <sup>27</sup> to the usual five 3-d functions,  $3d_{xy}$ ,  $3d_{yz}$ ,  $3d_{xz}$ ,  $3d_{z^2}$ ,  $3d_{x^2-y^2}$ , and a further s-type function  $3s'$ , corresponding to  $3d_{x^2+y^2+z^2}$ . Several runs of the SCF procedure were made using these functions selectively, and the total energies produced are shown in Table 4. As expected the largest basis set using  $S(\text{spd} + 3s')$  produced the best total energy. The addition of the  $3s'$  function leads to the largest single increment in total energy of 0.469 a.u. while the addition of a 3d function gives an increase of only 0.0154 a.u. This suggests that the angular properties of the 3-d functions are not used in forming strongly bonding orbitals but that the extra functions act merely as polarisation functions producing small adjustments in electron distribution. Properties which are directly related to electron distribution such as dipole, experience improvement.

Table 3 Molecular energies for the systems

	Naphthalene	Indene	Styrene	Indole
Total energy (a.u.)	-382.37127	-344.58111	-306.79044	-360.52655
Binding energy (a.u.)	-2.29027	-2.11071	-1.93044	-1.88815
Resonance energy(calc)(kJ mol <sup>-1</sup> )	357	225	208	308
Resonance energy(thermochemical) (kJ mol <sup>-1</sup> )	314		176	226
	Benzofuran	Benzothiophen (sp basis)	Benzothiophen (spd + 3s <sup>1</sup> basis)	Benzene
Total energy (a.u.)	-380.26093	-702.32203	-702.44606	-230.11426 -1.46945
Binding energy (a.u.)	-1.78283	-1.70843	-1.83246	212
Resonance energy(calc)(kJ mol <sup>-1</sup> )	232	242	283	155

Table 3 (Contd)

	2H-indene	Isoindole	Benzo (c) furan
Total energy (a.u.)	-344.52294	-360.50512	-380.2284
Binding energy (a.u.)	-2.05254	-1.86672	-1.68424
Resonance energy (calc)(kJ mol <sup>-1</sup> )	72	252	147

	Benzo (c) thiophen (sp basis)	Benzo (c) thiophen (spd + 3s <sup>1</sup> basis)
Total energy (a.u.)	-702.31240	-702.43085
Binding energy (a.u.)	-1.69780	-
Resonance energy (calc)(kJ mol <sup>-1</sup> )	217	259

Table 4 Total energies of benzothiophen and benzo (c) thiophen for various basis sets

Total energies (a.u.)

Basis Set	sp	sp + 3s <sup>1</sup>	spd	spd + 3s <sup>1</sup>
benzothiophen	-702.32203	-702.36890	-702.39913	- 702.44606
benzo (c) thiophen	-702.31239	-702.35956	-702.38369	-702.43085

Table 5 Variation in dipole with basis set

Dipoles (Debyes)

Basis Set	sp	sp + 3s <sup>1</sup>	spd	spd + 3s <sup>1</sup>	Expt
benzothiophen, 17	1.37	1.37	0.56	0.56	0.62
benzo (c) thiophen, 22	0.98	0.99	0.05	0.05	-



from the extended basis set. The dipoles produced for 17 and 22 using different basis sets are shown in Table 5, along with the experimental value for benzothiophen. Further confirmation of the role of 3d functions as polarisation functions is given by the very low eigenvectors ( $<0.1$ ) for orbitals associated with these functions. As seen from Table 4, the 3d orbitals lower the total energy of 17 by 0.12403 a.u., and in 22 by 0.11845 a.u. This smaller increase shown by 22 strengthens the conclusion that the 3d orbitals are polarisation functions as the quinoid isomer, 22, could exhibit an additional mode of delocalisation shown in Fig. 6, if the 3d orbitals participated in strong bonding. This would be expected to lead to a significant improvement in total energy but the calculations show that this is not the case.

Comparing the binding energies of the molecules in the Kekulé and quinoid series gives in all cases, those of the latter series as less negative than their Kekulé counterparts, showing them to be less stable to atomisation, as might be expected since the total energies show the quinoid series to be destabilised with respect to the Kekulé series.

Using previously determined REs for benzene, furan, thiophen, with and without d-orbitals and pyrrole of 212, 89, 144, 124 and 149  $\text{kJ mol}^{-1}$  it is possible to compare the REs of the bicyclic heterocycles of the Kekulé series with benzene plus the corresponding heterocycle. On this basis the REs for benzofuran, 16, benzothiophen ( $\text{spd} + 3s^1$ ), 17, and indole, 18, would be 301, 356, and 361  $\text{kJ mol}^{-1}$  respectively. The calculated values of 232, 283, and 308  $\text{kJ mol}^{-1}$  for 16 - 18 are all substantially less than this but maintain the same

order. This reduction in REs is in accord with the structures of these molecules which show a higher vinyl character in the C(2)-C(3) bond in the bicyclic than in the monocyclic heterocycles. Molecules, 16-18 are isoelectronic in the  $\pi$ -system with naphthalene but when compared with naphthalene, all are less aromatic. Both styrene, 24, and indene, 15, give REs close to that of benzene showing the C(2)-C(3) bond in indene to be an essentially localised double bond. Overall this produces an order in aromaticity of naphthalene > indole  $\approx$  benzo thiophen > benzofuran  $\approx$  indene > styrene. The order for the bicyclic heterocycles is the same as that for the monocyclic compounds, benzene > pyrrole  $\approx$  thiophen > furan in aromaticity. The general destabilisation of the quinoid isomers of 15-18 shown in the total energies is confirmed by relatively lower REs. Even the most aromatic isomer, isoindole, 23, has a substantially lower RE than naphthalene. Benzo (c) furan, 21, shows least aromatic character and has a RE less than that of benzene itself, but the general order of aromaticity is the same as for the Kekulé series, and is naphthalene > isoindole  $\approx$  benzo (c) thiophen > benzo (c) furan > 2H-indene.

Bicyclic heterocycles II The molecular energies and related properties are shown for 26-30 in Table 6. For the indazole isomers, 26a, and 27a, the total energies show that the 1-H isomer, 26a, is preferred in a tautomeric equilibrium. Experimental data found by uv,<sup>28</sup> and NMR<sup>29</sup> also support this. The 1-methyl isomer, 26c, is also favoured over the 2-methyl isomer, 27d, by 0.01475 a.u. Synthetic experience<sup>30</sup> has shown the conversion barrier between the two isomers to be low and interconversion occurs easily.

Table 6 Molecular energies for the bicyclic heterocycles II

	1H-Indazole	1,2-Benzisoxazole	1,2-Benzisothiazole		1-methyl Indazo
			(sp)	(spd + 3s <sup>1</sup> )	
Total energy (a.u.)	-376.40877	-396.28568	-718.36194	-718.48530	-415.36021
Resonance energy (calc.)(kJ mol <sup>-1</sup> )	278.1	196.2	620.0	671.0	-

	2H-Indazole	Anthranil	2,1-Benzisothiazole		2-methyl Indazol
			(sp)	(spd + 3s <sup>1</sup> )	
Total energy (a.u.)	-376.39168	-396.12733	-718.20391	-718.34339	-415.34546
Resonance energy (calc.)(kJ mol <sup>-1</sup> )	233.3	138.5	205.0	298.4	-

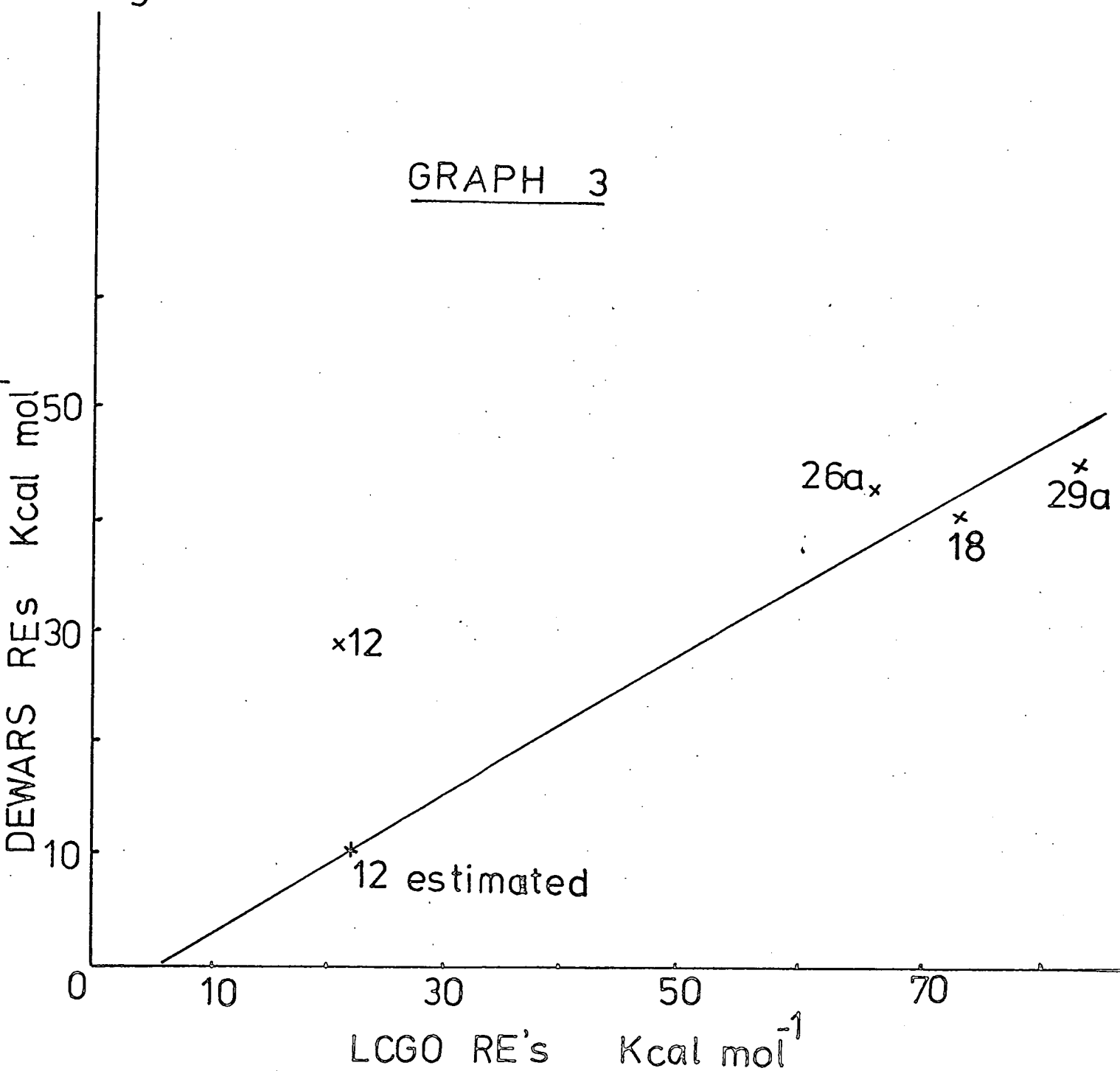
	Benzimidazole	Benzoxazole	Benzothiazole	
			(sp)	(spd + 3s <sup>1</sup> )
Total energy (a.u.)	-376.53486	-396.28568	-718.35372	-718.47191
Resonance energy (calc.)(kJ mol <sup>-1</sup> )	580.5	545.7	575.7	600.8

Table 6 (Contd)

	1H-Benzotriazole	1,2,3-Benzthiadiazole	1-methyl Benzotriazole
Total energy (a.u.)	-392.31581	-734.36021 (spd + 3s')	-431.15871
Resonance energy (calc.) (kJ mol <sup>-1</sup> )	350.9	659.2	-

	2H-Benzimidazole	2,1,3-Benzoxadiazole	2,1,3-Benzthiadiazole (sp)      (spd + 3s')	2-methyl Benzotriazole
Total energy (a.u.)	-392.28741	-412.01933	-734.11906    -734.26477	-431.15871
Resonance energy (calc)(kJ mol <sup>-1</sup> )	236.3	253.0	253.0      375.0	

Fig.7



	12	18	26a	29a
Dewar	28.6	39.8	42.0	43.9
LCGO	21.0	73.8	66.4	83.9

For the benzotriazole series, again the 1H-, 29a, and the 1-methyl, 29c, isomers are favoured, as expected, over the corresponding 2-isomers, 30a, d, which are quinoid in form. For the methyl isomers  $\Delta E_{\text{total}}$  is rather large (0.11057 a.u.) compared with that for the H-isomers (0.02839 a.u.) One would have expected the differences to be relatively constant as in the indazole series, 0.01709 a.u. between the H-isomers and 0.01475 a.u. between the methyl isomers. A constructed geometry was used for both the 2H- and 2-methyl compounds, and this large difference in total energy is perhaps the magnification of a slight distortion in the geometry of the latter.

The total energies of molecules containing sulphur atoms are given with and without d-orbitals. As expected the larger basis set leads to a better total energy, and again the additional orbitals act as polarisation functions. As previously shown (Graphs 1 & 2) a linear relationship exists between the present RE data and that of Dewar and Pauling. This is reinforced by Graph 3, Fig. 7, which shows the RE for molecules 12, 18, 26a and 29a plotted against Dewar's results. Only pyrrole, 12, is anomalous. RE is dependent on the bond energies chosen,<sup>31</sup> particularly so in cases such as pyrrole where bond alternation is pronounced, and the RE obtained with fixed bond lengths tends to be unreliable. In heterocycles where bond alternation is less pronounced, less variation in bond lengths is required, and experimental and calculated results will be in closer agreement.

The calculated REs for benzimidazole and related compounds,

28b, c, are all high relative to the corresponding heterocycle of Series A (Fig. 4, Section 1) yet the isomers with hetero atoms in the 1, 2 positions give REs close to the Series A molecules. However, as was noted previously, indazole, 26a, and benzotriazole, 29a, are similar to naphthalene in the geometry of the carbocyclic ring, whereas benzimidazole, 28a, has a benzene like carbocyclic ring. This is mirrored in their REs with 28a being closer to 2xRE of benzene, and 26a and 29a being closer to naphthalene.

In both 26b and 29b where an N-S bond appears in a Kekulé type structure the RE is high. A possible explanation of this is obtained by considering the molecules used to give the N-S bond energy. It was derived using the molecule  $\text{NH}_2\text{-SH}$  which is of unknown geometry, and was constructed using the geometries of  $\text{NH}_3$  and  $\text{H}_2\text{S}_2$ , with a staggered conformation. It is possible that the geometry used is not optimal, and will therefore lead to the molecular energy being less negative, and consequently the N-S bond energy will be too low. The sum of contributing bonds using this bond energy will be too low, and will lead to a RE which is too high.

REFERENCES

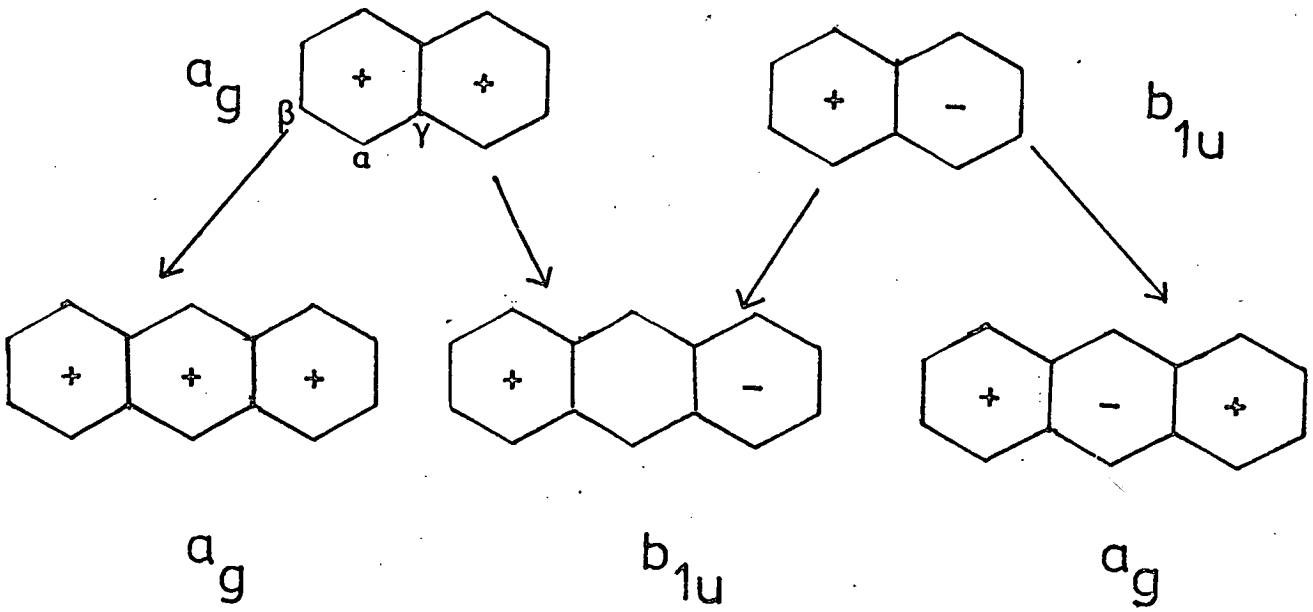
1. J.D. Cox and G. Pilcher, 'Thermochemistry of Organic and Organometallic Compounds', Academic Press, London, 1970.
2. L. Pauling, 'The Nature of the Chemical Bond', Cornell Univ. Press, New York, 1960, 3rd.edn., p.190.
3. C.K. Ingold, 'Structure and Mechanism in Organic Chemistry', Bell, London, 1969, p. 220.
4. M.J.S. Dewar and H.N. Schmeising, Tetrahedron, 1960, 11, 96.
5. M.J.S. Dewar and H.N. Schmeising, Tetrahedron, 1959, 5, 166.
6. M.J.S. Dewar and H.N. Schmeising, Colloques Internationaux du Centre National de la Recherche Scientifique, 1958, 82, 51.
7. M.J.S. Dewar, A.J. Harget and N. Trinajstic, J. Amer. Chem. Soc., 1969, 91, 6321.
8. M.H. Palmer and R.H. Findlay, J. Chem. Soc. Perkin II, 1974, 1885.
9. M.J.S. Dewar and C. de Llano, J. Amer. Chem. Soc., 1969, 91, 789.
10. M.J.S. Dewar, A.J. Harget, N. Trinajstic and S.D. Worley, Tetrahedron, 1970, 26, 4505.
11. B.A. Hess and L.J. Schaad, J. Amer. Chem. Soc., 1971, 93, 305.
12. B.A. Hess and L. J. Schaad, J. Amer. Chem. Soc., 1973, 95, 3907
13. P.J. Hay , and I. Shavitt, J. Chem. Phys., 1974, 60, 2865.



14. W.C. Ermler and C.W. Kern, J. Chem. Phys., 1973, 58, 3458.
15. J. Almlöf, Chemical Physics, 1974, 6, 135.
16. R.E. Christoffersen, Int. J. Quant. Chem. Symp., 1973, 7, 169.
17. R.J. Buenker and S.D. Peyerimhoff, Chem. Phys. Letters, 1969, 3, 37.
18. R.E. Christoffersen, J. Amer. Chem. Soc., 1971, 93, 4104.
19. W.C. Herndon, Thermochemica Acta., 1974, 8, 225.
20. H.J. Bernstein, Trans. Faraday Soc., 1962, 58, 2285.
21. F. Brogli, E. Heilbronner and T. Kofayashi, Hely. Chim Acta., 1972, 55, 274.
22. F. Brogli and E. Heilbronner, Angew. Chem. Int. Ed., 1972, 11, 538.
23. P.A. Clark, F. Brogli and E. Heilbronner, Hely. Chem. Acta., 1972, 55, 1415.
24. L.C. Snyder and H. Basch, J. Amer. Chem. Soc., 1969, 91, 2189.
25. Thanks to J. Nisbet. Chem. Dept. Edinburgh University for the calculation on [14] -annulene.
26. W.J. Hehre, L. Radom and J.A. Pople, J. Amer. Chem. Soc., 1972, 94, 1496.
27. A. Rauk and I.G. Csizmadia, Canad. J. Chem., 1968, 46, 1205.
28. (a) M. Barclay, N. Campbell and G. Dodds, J. Chem. Soc., 1941, 133.  
(b) V. Rousseau and H.G. Lindwall, J. Amer. Chem. Soc., 1950, 72, 3047.

29. (a) J. Elguero, A. Fruchier and R. Jacquier, Bull. Soc. Chim. Fr., 1966, 3041; *ibid*, 1967, 2691; *ibid*, 1969, 2064.  
(b) P.J. Black and M.L. Heffernan, Aust. J. Chem., 1963, 16, 1051.
30. This work - experimental section.
31. M.J.S. Dewar and G.J. Gleicher, J. Chem. Phys., 1966, 44, 759.
32. K.J. Laidler, Can. J. Chem., 1956, 34, 626.
33. J. Koller, A. Azman and N Trinajstic, Z. Naturforschm 1974, 29a, 624.

Fig.1



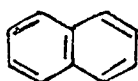
## The Molecular Orbital Energy Levels

From the LCGO calculations, an analysis of each molecular orbital was performed, and assigned to the relevant symmetry species of the symmetry group to which the molecule belonged. A list of molecular orbital energies is given in Appendix 2.

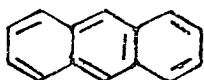
The hydrocarbons In this series, the molecules 1 - 5 are treated as having  $D_{2h}$  symmetry, and the wave functions of 3 - 5 can be regarded as the superposition of twice the wave function of the previous member of the series in either a symmetric or an antisymmetric combination, as shown in Fig. 1, for naphthalene and anthracene  $a_g$  and  $b_{1u}$  superpositions. This leads to the symmetry relationships for the  $\sigma$  and  $\pi$  - system as shown in Table 1. These symmetry relationships are used in 1 - 5 as an aid to orbital correlation. Fig. 2 shows the orbitals related by the  $a_g/b_{1u}$  relationship and Fig. 3 those related by the  $b_{2g}/b_{3u}$  relationship in the  $-(10-35)$ eV binding energy region. In the early  $\sigma$  - valence region of 4 and 5, the orbitals are close together and cannot be assigned unambiguously. The  $\sigma$  - valence region appears to be split into three sections. The highest binding energy section being  $2s_c$  levels related to  $2a_{1g}$  and  $2e_{1u}$  in benzene, producing a band structure centred at  $\approx -29.5$ eV with a range of  $\pm 4$ eV. The middle and upper sections are less distinct, but CH levels associated with  $2e_{2g}$ ,  $2b_{1u}$  and  $3a_{1g}$  in benzene, give a band centred at  $\approx -23.3$ eV, and the C-C  $\sigma$ -levels give a band centred at  $\approx -16.0 \pm 3$ eV. The  $\pi$  - levels for 1 - 5 are shown in Fig. 4



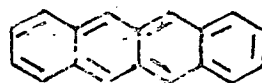
1



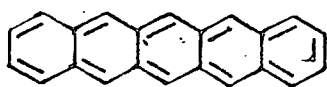
2



3



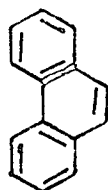
4



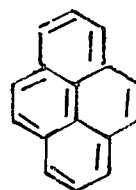
5



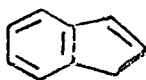
6



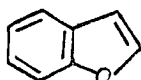
7



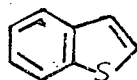
8



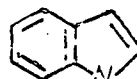
15



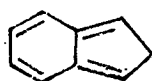
16



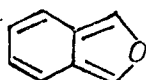
17



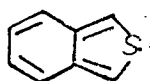
18



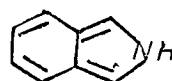
20



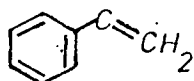
21



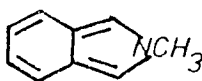
22



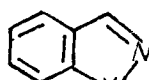
23



24

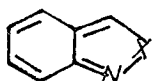


25



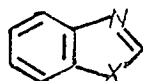
26

a, NH  
b, O  
c, S  
d, NCH<sub>3</sub>



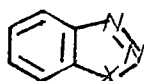
27

a, NH  
b, O  
c, S  
d, NCH<sub>3</sub>



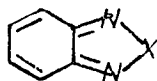
28

a, NH  
b, O  
c, S



29

a, NH  
b, S  
c, NCH<sub>3</sub>



30

a, NH  
b, O  
c, S  
d, NCH<sub>3</sub>

Fig. 2

$a_g, b_{1u}$

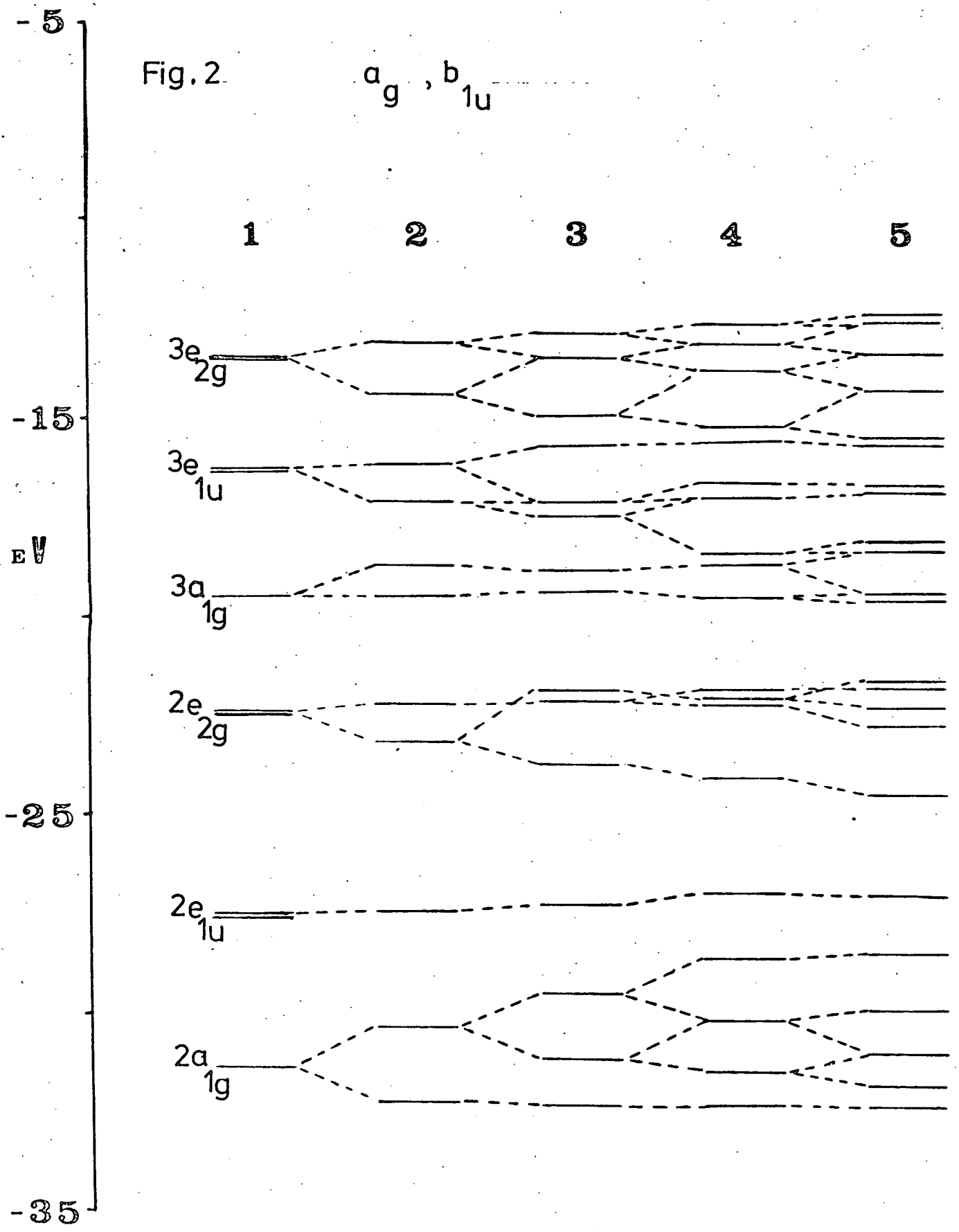
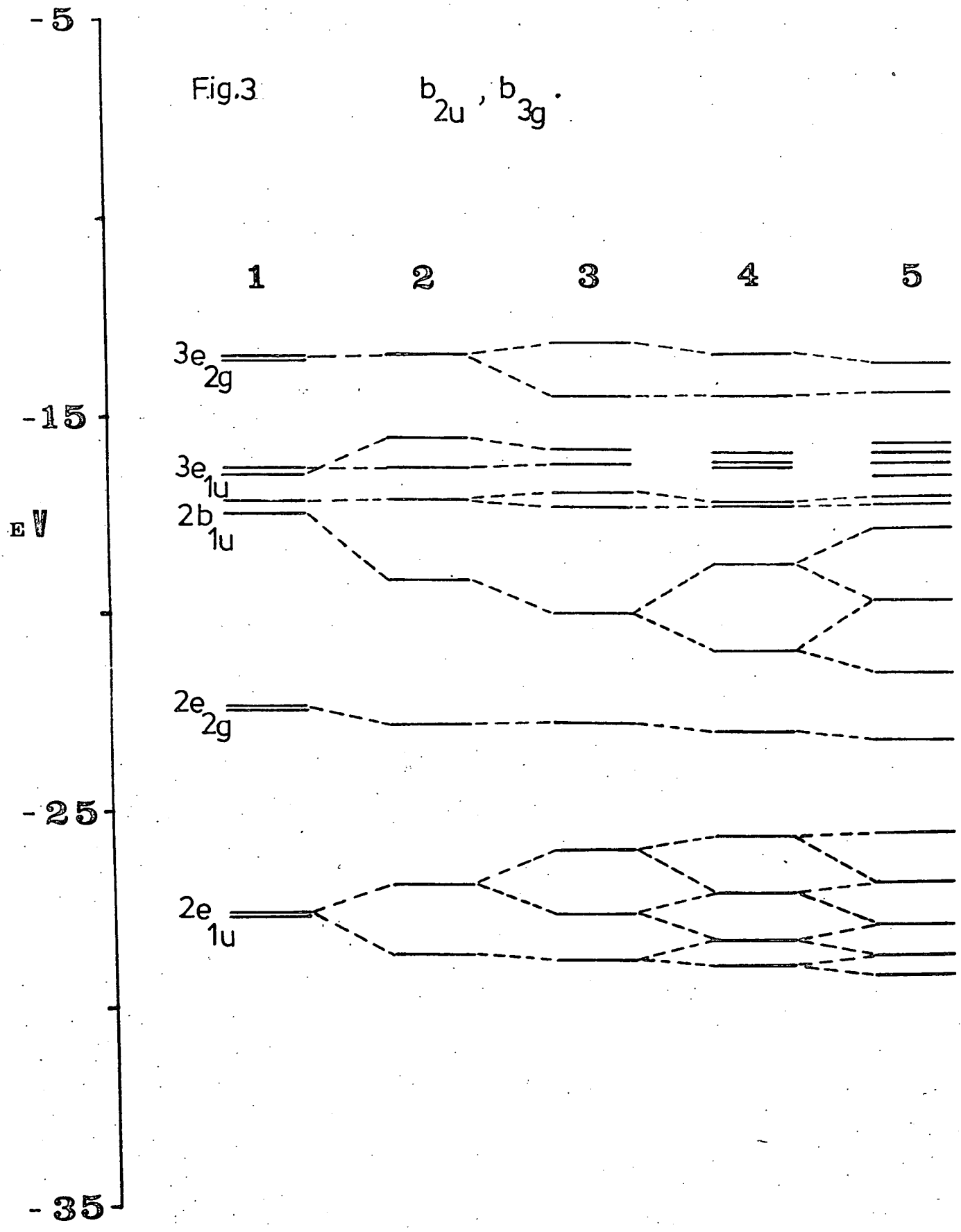


Fig.3

$b_{2u}, b_{3g}$







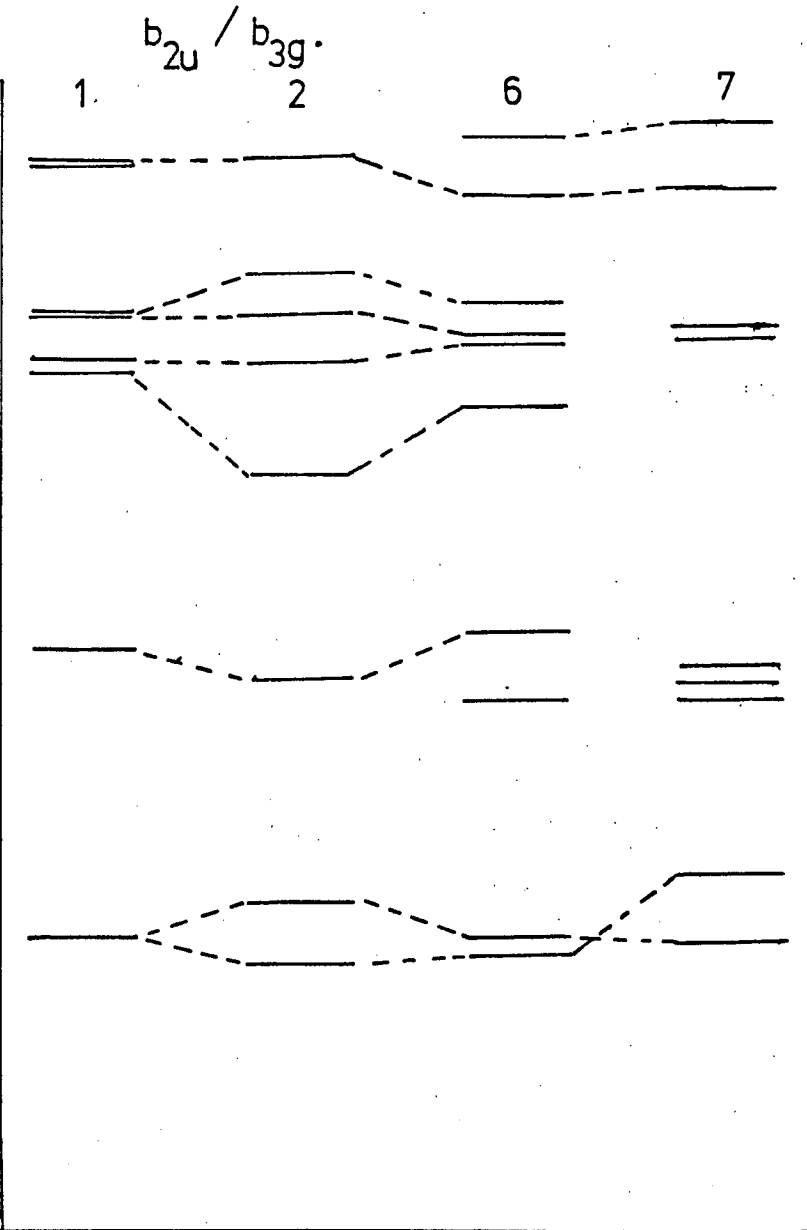
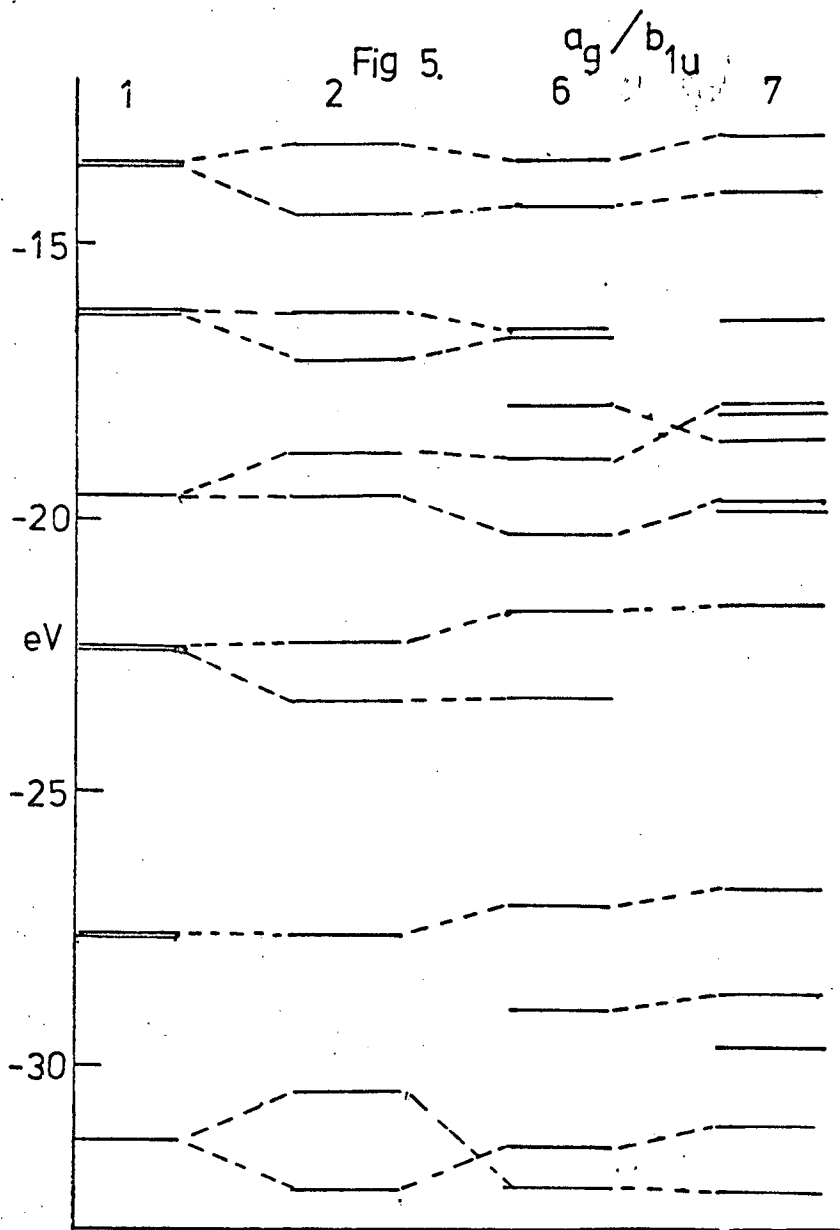


Table 1 Symmetry relationships in the  $\sigma$ - and  $\pi$ - system of the hydrocarbons using  $D_{2h}$  symmetry

---

$$\begin{array}{ll}
 a_g \pm a_g \rightarrow a_g; & b_{1u} \leftarrow b_{1u} \pm b_{1u} \\
 b_{2u} \pm b_{2u} \rightarrow b_{2u}; & b_{3g} \leftarrow b_{3g} \pm b_{3g} \\
 a_u \pm a_u \rightarrow a_u; & b_{1g} \leftarrow b_{1g} \pm b_{1g} \\
 b_{3u} \pm b_{3u} \rightarrow b_{3u}; & b_{2g} \leftarrow b_{2g} \pm b_{2g}
 \end{array}$$


---

Table 2 Calculated  $C_{1s}$  levels, atomic populations and experimental  $^{13}C$ -shifts.

---

Molecule	$C_{1s}$ Calculated	$^{13}C$ -shift	Atomic Population	
Indole	C(2)	-307.61	125.1	6.0224
	C(3)	-305.79	102.6	6.1904
	C(3a)	-306.47	128.7	6.0476
	C(4)	-306.53	121.3	6.1414
	C(5)	-306.28	122.3	6.1639

Table 2 (cont'd)

Molecule	$C_{1s}$ Calculated	$^{13}C$ -shift	Atomic Population	
	C(6)	-306.54	120.2	6.1508
	C(7)	-306.66	111.8	6.1645
	C(7a)	-307.85	136.1	5.8895
Benzofuran	C(2)	-308.89	144.8	5.9161
	C(3)	-307.13	106.5	6.1674
	C(3a)	-307.40	127.4	5.7998
	C(4)	-307.26	121.4	6.1371
	C(5)	-306.96	122.7	6.1551
	C(6)	-307.20	124.2	6.1421
	C(7)	-307.27	111.4	6.1534
	C(7a)	-309.38	115.0	6.0082
Benzothiophen	C(2)	-307.39	126.1	6.1662
(spd + 3s')	C(3)	-306.90	123.6	6.1585
	C(3a)	-307.20	139.4	6.0276
	C(4)	-306.99	123.5	6.1375
	C(5)	-306.79	124.0	6.1551
	C(6)	-306.96	124.0	6.1452

Table 2 (cont'd)

Molecule	$C_{1s}$ Calculated	$^{13}C$ -shift	Atomic Population	
	C(7)	-307.05	122.3	6.1429
	C(7a)	-307.71	139.6	6.0398
Benzimidazole*	C(2)		141.5	5.8672
	C(3a)		137.9	5.9047
	C(4)		115.4	6.0994
	C(5)		122.9	6.1008

†  $^{13}C$ -shifts are with respect to T.M.S. with  $SCS_2$  at 192.80 ppm from T.M.S.

\* Average values of  $C_{1s}$  and atomic populations across the Z axis are given as the results are average results for these positions.

Table 3 Hartree Fock orbital energies (eV)

Atom	2s	2p	3s	3p
C	19.2	11.8	-	-
N	25.7	15.4	-	-
O	33.9	17.2	-	-
S	-	-	23.9	11.9

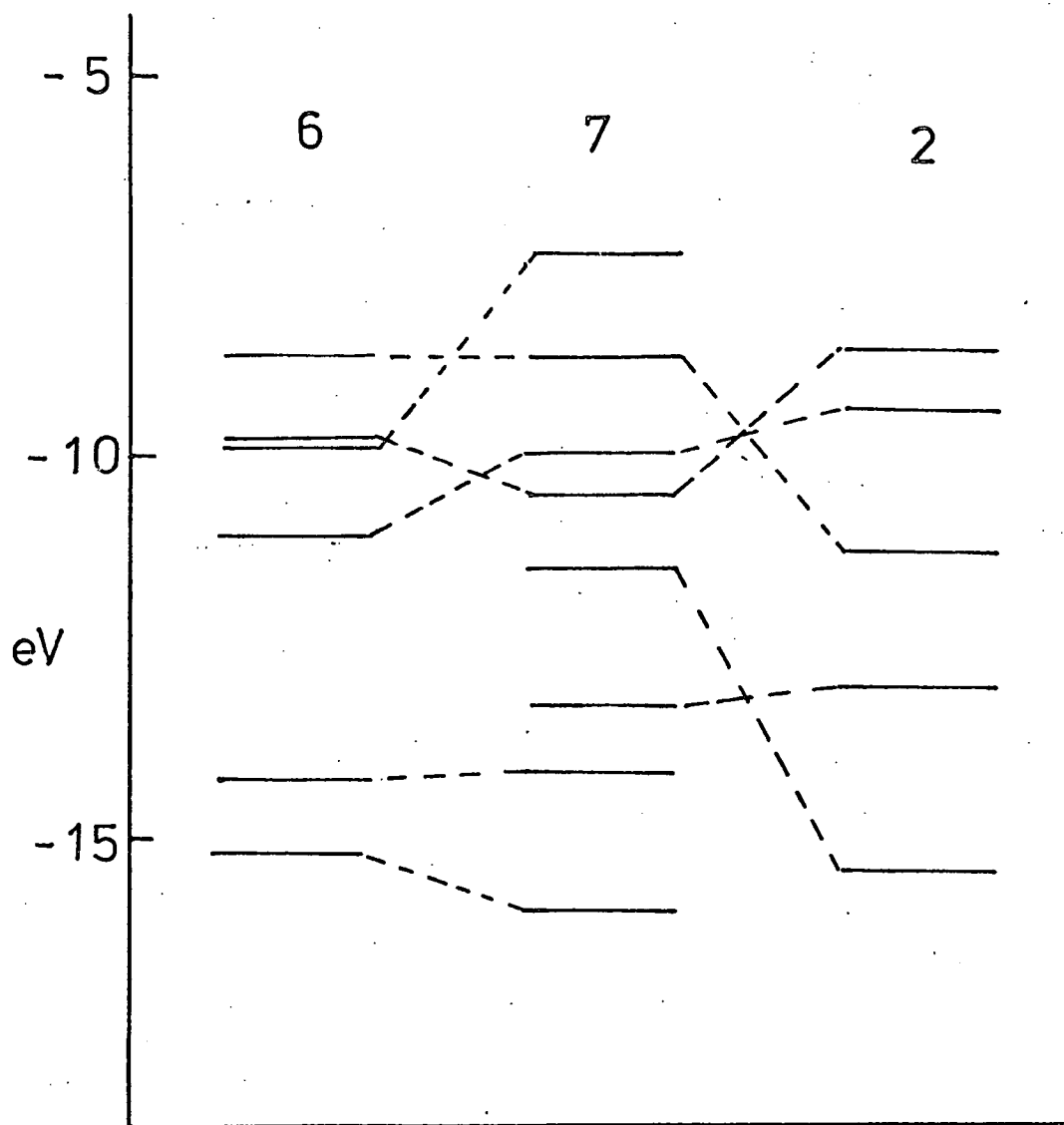
As the number of  $\pi$ -levels increases, those related by orbital type and symmetry move to higher binding energy. The  $\sigma$  orbitals of planar biphenyl, 6, are very similar to those of benzene, Fig. 5 with the missing CH level from the fusion of the two benzene rings being of mixed  $e_{1u}/b_{2u}$  symmetry. The total contribution of the  $\pi$  - system to the total energy of biphenyl compared with twice that of benzene give the former as 1.3eV larger. This is very small, and represents the degree of conjugation between the rings but it is obviously not sufficient to offset the steric effects within the molecule which lead it to be non-planar in the gas phase.

Although RE data and geometry suggest that phenanthrene, 7, is closer to biphenyl than to naphthalene, 2, an orbital correlation of  $\sigma$  - valence shell orbitals (Fig. 5) shows that there are a number of orbitals in 7 which have no direct counterpart in 2, or 6. Many orbitals can in fact be directly assigned to those in biphenyl, and are also of very similar energy. These orbitals are marked B in Appendix 2. It would seem that the idea of a bridged biphenyl for 7, gives a better description of its structure than the idea of an extended naphthalene structure, but is still not 100% successful.

For pyrene, 8, a number of  $\sigma$  orbitals have direct counterparts in 2 and 6, and are marked N and P respectively in Appendix 2. The related orbitals are also of very similar energy. In the  $\pi$  - system there is a 1:1 relationship between the orbitals of pyrene and biphenyl, and these can

Fig. 6

$\pi$ -ORBITALS.



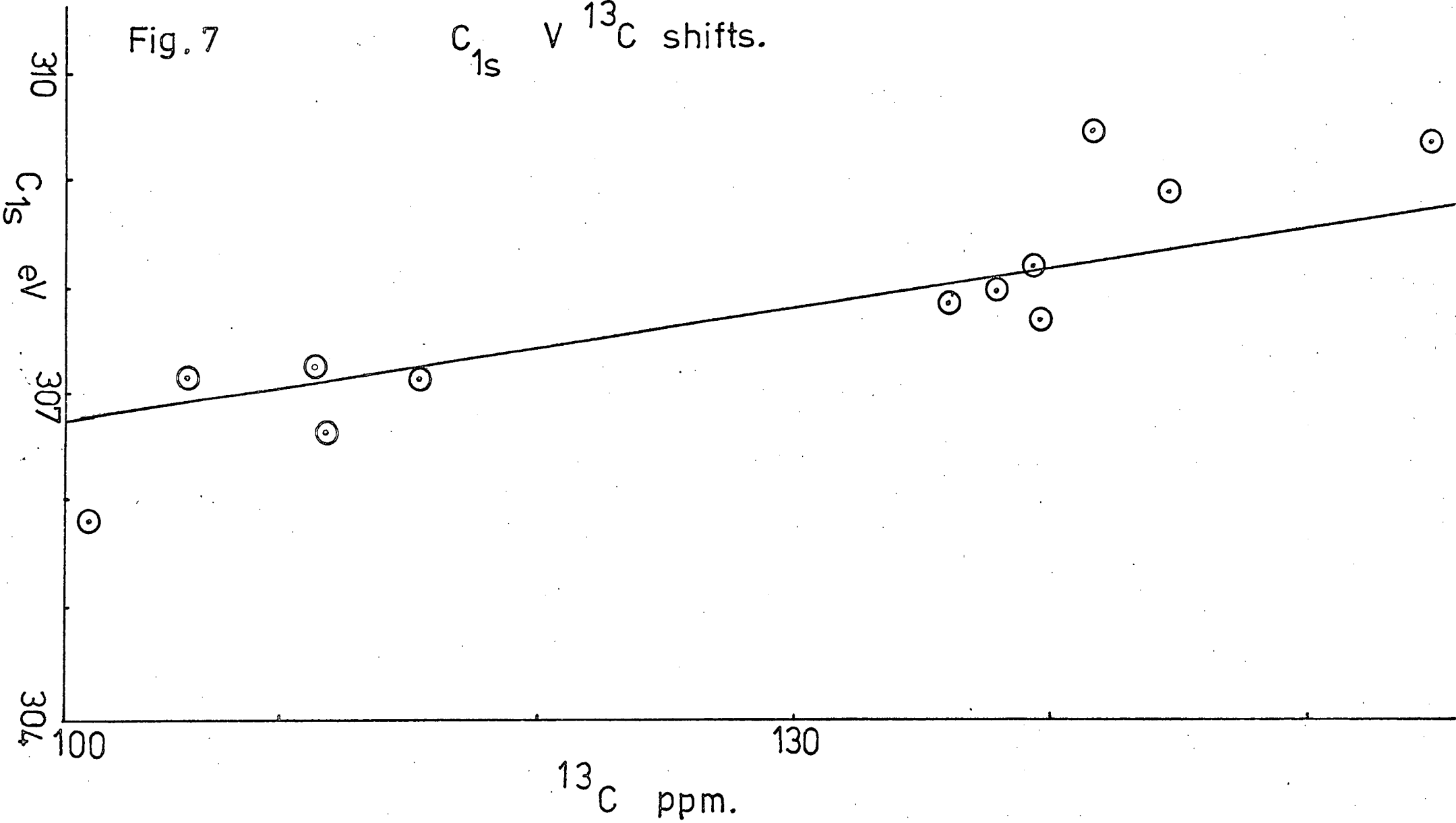
be correlated with naphthalene, Fig. 6.

The orbital energies of planar (14) - annulene, 9, are generally of lower binding energy than those of 8, or any of the other hydrocarbons. This agrees with its lack of RE in the planar state.

The bicyclic heterocycles I In all cases the  $1s$  orbital eigenvectors are large ( $>0.98$ ) as would be expected for the very localised core orbitals. For naphthalene, 2, indene, 15, and styrene, 24, the bridging carbon atoms are at highest binding energy. The XPS work on naphthalene<sup>1</sup> gives the order of core levels as  $1s_{\gamma C} > 1s_{\alpha C} > 1s_{\beta C}$  where  $\alpha$ ,  $\beta$  and  $\gamma$  are shown in Fig. 1, and for 2, the present calculation is in agreement. Of the heterocycles only 2H-indene and benzo(c)thiophen, 22, show this order, the others giving the C atoms  $\alpha$  to the hetero atom as the highest binding energy sites corresponding to regions of low electron density due to the electronegativity of the hetero atom, leaving the  $\alpha$  positions electron deficient. For 22, this indicates that it is closer to a hydrocarbon structure than the others. This is further strengthened by the small spread in  $1s_C$  levels (0.5eV) which resembles that of the hydrocarbons 1 - 5 (0.6eV), and is said to show a lack of charge separation in alternant hydrocarbons. Of the other heterocycles, only benzothiophen, 17, has a relatively small spread in  $1s_C$  levels of  $\approx 0.9$ eV, the others being in the 2eV range. This adds support to the theory that a sulphur atom<sup>2</sup> is electronically very similar to a  $CH=CH$  bond.

Fig. 7

$C_{1s}$   $V^{13}C$  shifts.





Several workers<sup>3,4</sup> have shown that a linear relationship exists between calculated  $C_{1s}$  energy levels and those obtained by XPS. It might be expected that a similar relationship should exist between calculated  $C_{1s}$  energy levels and  $^{13}C$  NMR shifts, as the  $C_{1s}$  energy levels are known to be sensitive to changes in the valency shell<sup>5</sup>. When the  $^{13}C$  shifts of indole<sup>6</sup>, benzofuran<sup>7</sup>, benzothiophen<sup>7</sup> and benzimidazole<sup>8</sup> are plotted against their  $C_{1s}$  levels, Graph 1 Fig. 7, evidence of a linear relationship is found but the scatter is fairly large. A least square fit of the data produces  $C_{1s} = 0.058^{13}C + 300.05eV$  with standard deviations in slope, intercept and overall of 0.009, 1.093 and 0.523. In general, the high binding energy  $C_{1s}$  levels are associated with the largest  $^{13}C$  shifts, and when the  $C_{1s}$  levels are compared with the atomic populations, a high binding energy is associated with a low population, Table 2. This seems reasonable as a low electron density around a nucleus will mean less electron-electron repulsion, and a stronger attraction of the outer electrons to the nucleus due to a reduced screening effect. The reverse relationship of binding energy and atomic population should give a relationship between  $^{13}C$  shifts and atomic populations, but it has been shown<sup>3</sup> that in certain cases high electron density is associated with high binding energy, and not with low binding energy as predicted. When a graph of  $^{13}C$  shifts against atomic populations for the molecules already

Fig. 8

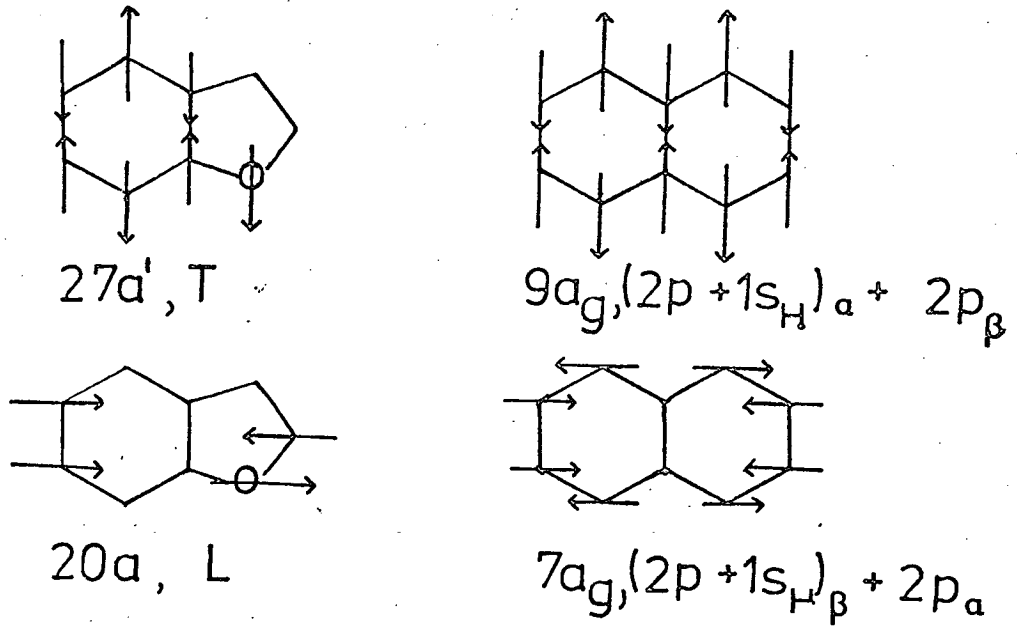


Fig. 9

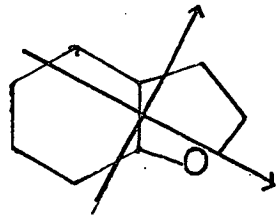
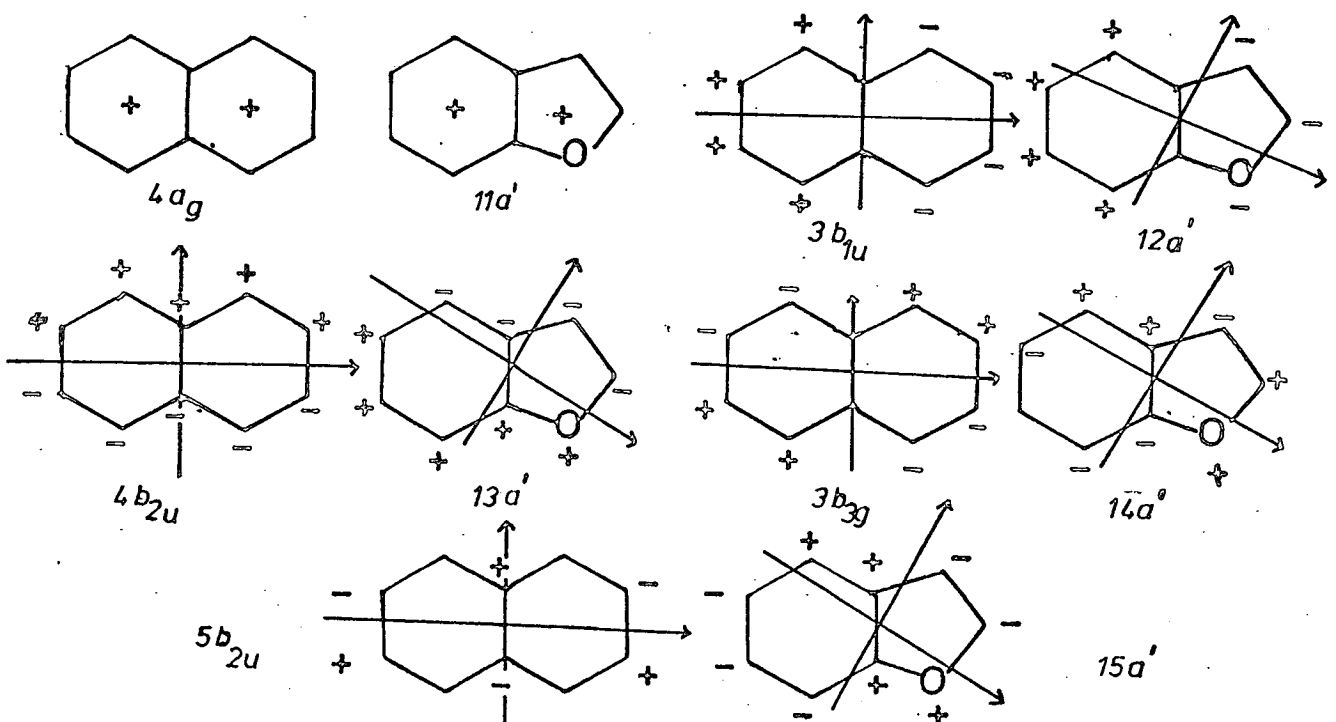


Fig. 10



mentioned, is drawn, no obvious relationship exists because of the small variations in many of the populations, especially for the benzenoid carbon atoms. This crowds the points together, and as insufficient data is available at the extreme of the graph no correlation could be made.

It is not possible, due to the lower symmetry of some of these molecules compared to naphthalene, to classify directly by symmetry species, but by a thorough analysis of the orbitals in these molecules it is possible in the  $\sigma$  - valence shell to introduce a further classification by dividing the 2p orbitals into two categories, those parallel to, or those perpendicular to the long molecular axis(Z) giving longitudinal(L), and transverse (T) orbitals.

The orbitals of 2, 15 - 18, and 20 - 24 have been analysed and the main character of each orbital is shown in Appendix 2, along with the orbital energies and symmetry species. Where appropriate, the corresponding naphthalene symmetry species is given. There are relatively few orbitals of mixed (L + T) type, and these are either purely radial or tangential with respect to the individual rings. Examples of L and T orbitals are shown in Fig. 8 with their naphthalene counterparts.

The molecules 15 - 18 which have only  $C_s$  symmetry show an "apparent" symmetry axis, Fig. 9 and orbitals are conventionally polarised longitudinally or transversely with respect to this axis, making correlation.

with naphthalene easier. An example of this is shown in Fig. 10 for the 2s orbitals of benzofuran and the corresponding naphthalene orbitals.

An orbital correlation diagram for the valence shell region of naphthalene with 15 - 18 and 24 is shown in Fig. 11, and one for naphthalene and the quinoid series in Fig. 12. In both cases the correlation diagram can be considered as three regions, the first containing orbital levels below -25eV, the second containing levels between -12.5 and -25eV and the third, those above -12.5eV. The first region corresponds to the 2s(C,N,O) and 3s(S) levels, and contains five levels, the highest binding energy one being associated in the heterocycles with the hetero atom. The N and O heterocycles show highly localised 2s (N or O) levels, but for the S heterocycles the 3s orbital is much more delocalised. This is shown definitively by the highest binding energy orbital in this region. The eigenvectors of this orbital in benzothiophen show a comparable contribution from C (2s) and S (3s), whereas in indole and benzofuran a much greater degree of localisation is shown, the effect being more pronounced in the latter, and the orbital showing O (2s) with an eigenvector of 0.76. This can be explained by considering the respective free atom orbital energies shown in Table 3. On inspection of the  $2s_x$  (X = N,O,C) and the 3s sulphur value, it is apparent that the orbitals closest in energy are C (2s) and S (3s), and those furthest apart are C (2s) and O (2s). Orbitals of

Fig.11

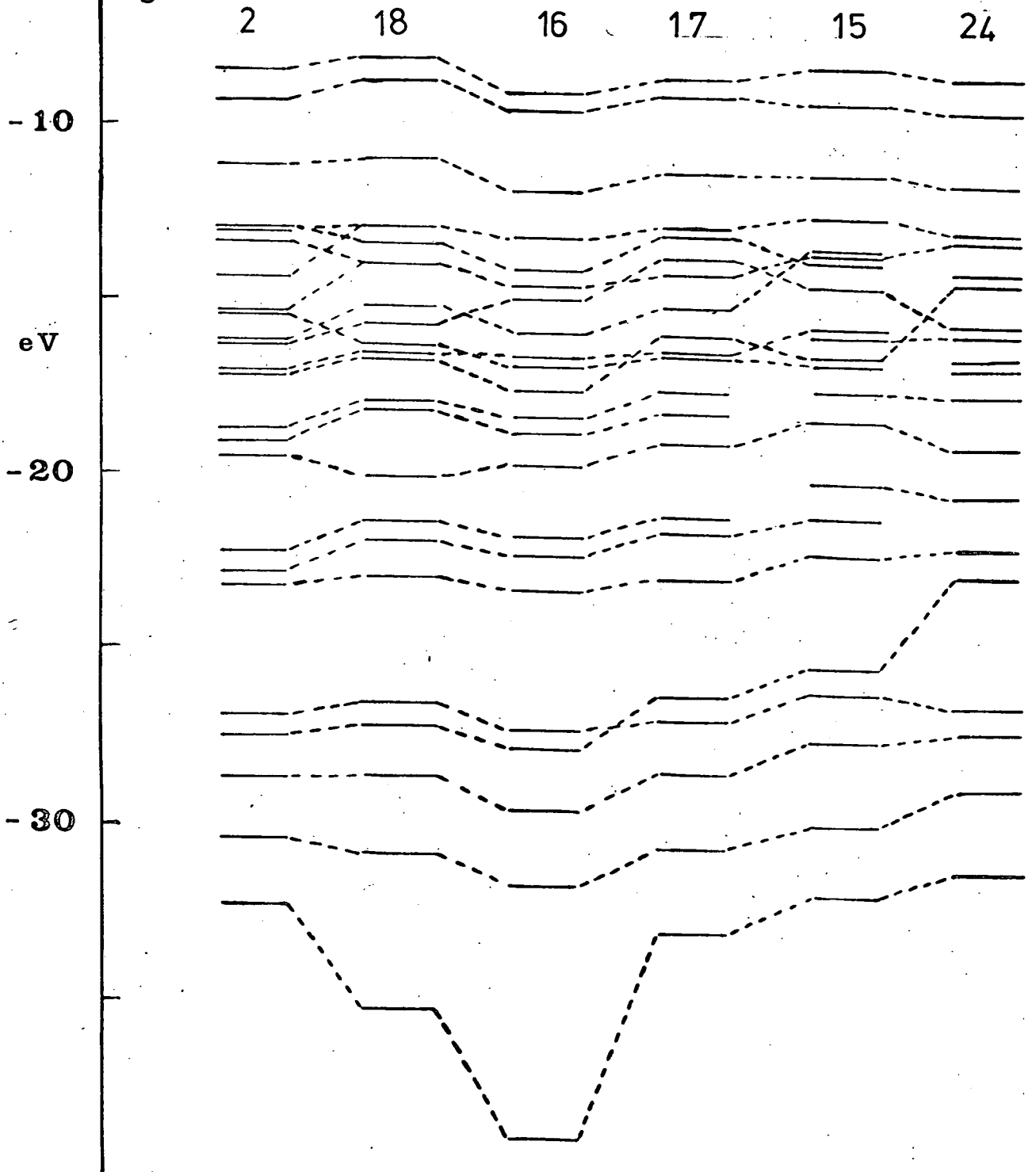
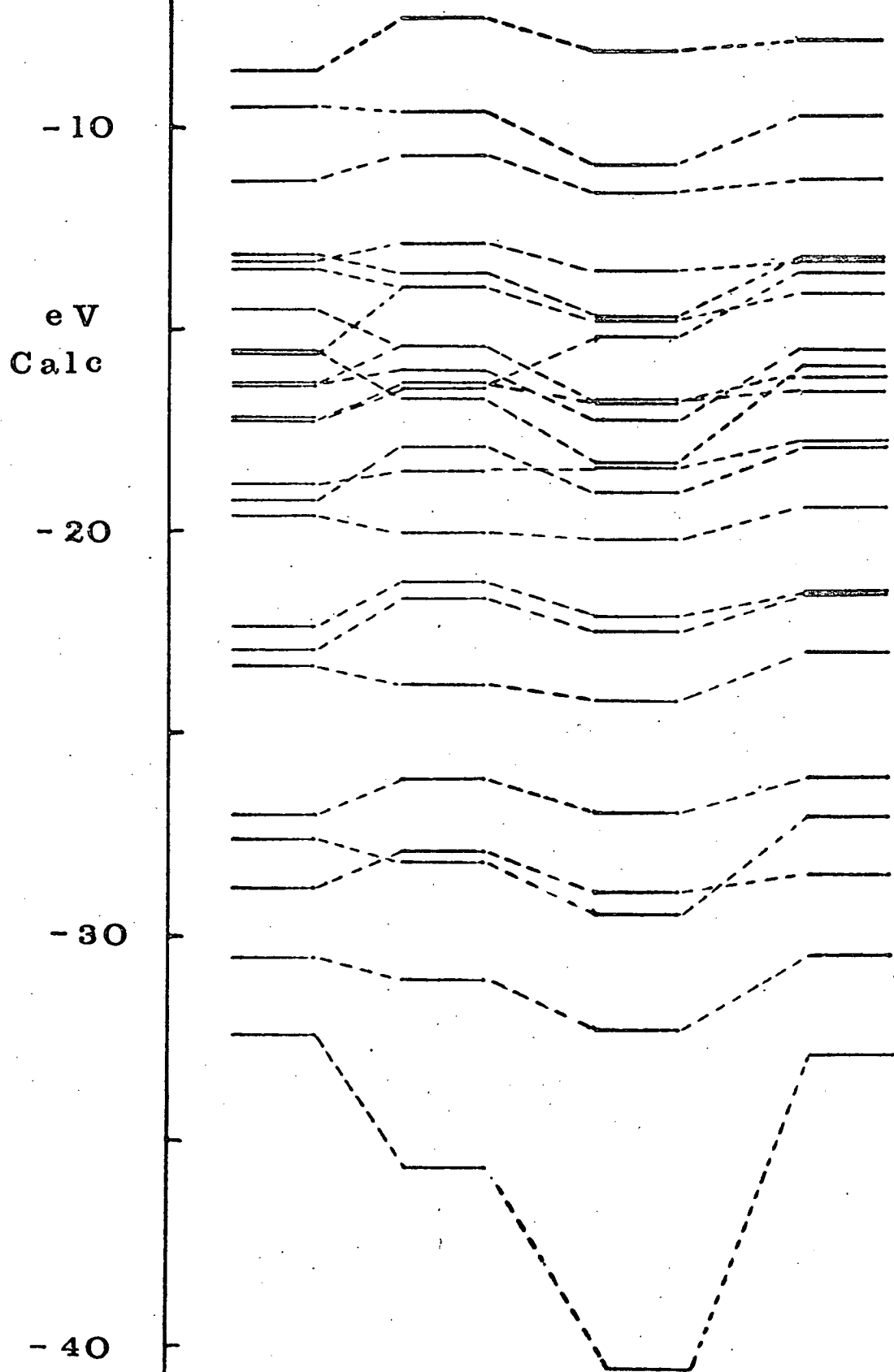
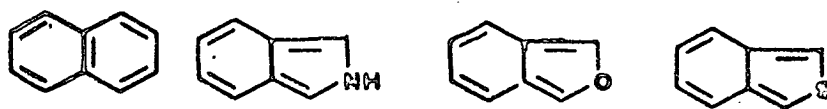


Fig.12



similar energy will mix best, and so molecular orbitals involving C (2s) and S (3s) will show most mixing and those of C (2s) and O (2s) will show least mixing. This is indeed the case as shown by the calculations, and is further reflected in the orbital energies of the sulphur heterocycles, 17 and 22, which bear a much closer relationship to the hydrocarbons than the other heterocycles do, again lending support to the idea of an S atom acting as  $\text{CH}=\text{CH}$ . For styrene the lowest binding energy 2s level is to be found much nearer to the centre region.

In the -12.5/-25eV region, the high binding energy end shows a subset of three largely  $2s_c$  levels with some 2p (C,N,O). This is found in all the molecules but the levels are more diffuse in styrene, indene and 2H-indene. The remaining levels in this region are outer valency shell  $\sigma$ -orbitals composed of combinations of 2p (C,N,O), 3p (S) and 1s (H), and  $\pi$ -orbitals at high binding energy.

In all cases, the low binding energy region shows three  $\pi$ -levels. With the exception of styrene where several orbitals have no direct counterpart in indene or naphthalene, and its structure better approximates to benzene and ethylene, the orbital correlation is good and few crossovers occur. These crossovers occur mainly in the outer valence shell and in most cases can be accounted for by differences in the free atom 2p/3p orbital energies of X (Fig. 4,

section 1) where these are as given in Table 3 at the HF limit.

In the quinoid series Fig.12, there is a marked variation in the separation of the first two  $\pi$ -levels compared with the Kekulé series Fig.11. It has previously been noted<sup>9</sup> that in the case of the monoheterocycles, and six membered ring azines<sup>10</sup>, that the separation of the two lowest binding energy  $\pi$ -levels can be related to the RE of the molecule, the latter varying in the reverse order to the separation, and indicating the degree of aromaticity of the molecule. On this basis, molecules of high RE will have the first two  $\pi$ -levels degenerate as in benzene, and those of lower RE will have a greater separation of the  $\pi$ -levels. Using this approach with the quinoid series and naphthalene, the order of aromaticity is given as naphthalene > benzo(c) thiophen > isoindole > benzo(c)furan > 2H-indene which is the order established by RE considerations. For the Kekulé series, the agreement with this method of determining aromaticity is less pronounced, giving the order as benzothiophen > benzofuran > indole > naphthalene > indene = styrene, but as near degeneracy of energy levels is often associated with an even electron distribution which can only arise in highly symmetrical molecules, the assymetry of these molecules may be the cause of a poor correlation. All other molecules considered by this method have been of  $C_{2v}$  symmetry. However, the separations for the Kekulé



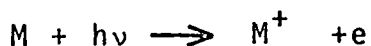
molecules indicate that less variation in REs should be shown than for the quinoid series, and this is in fact the case, Table 3, section 3.

REFERENCES

1. D.T. Clark and D. Kilcast, J. Chem. Soc. (B), 1971, 2243.
2. R.A.W. Johnstone and F.A. Mellon, J. Chem. Soc., Faraday II, 1973, 1155
3. M. Barber and D.T. Clark, Chemical Communications, 1970, 22; 23; 24.
4. D.W. Davies, J.M. Hollander, D.A. Shirley and T.D. Thomas, J. Chem. Phys., 1970, 52(6), 3295
5. C.A. Coulson and F.A. Gianturco, Molecular Physics, 1970, 18(5), 607.
6. R.G. Parker and J.D. Roberts, J. Org. Chem., 1970, 35(4), 996
7. N. Platzter, J.J. Basselier and P. Demerseman, Bull. Soc. Chim. Fr., 1974, 1, 905.
8. P.J. Pugmire and D.M. Grant, J. Amer. Chem. Soc., 1971, 93, 1880.
9. M.H. Palmer and R.H. Findlay, J. Chem. Soc., Perkin II, 1975, 974
10. M.H. Palmer, R.H. Findlay and A.J. Gaskell, J. Chem. Soc., Perkin II, 1974, 778.

Photoelectron Spectroscopy

The technique of photoelectron spectroscopy, PS, involves the irradiation of a closed shell molecule, M, with photons of energy,  $h\nu$ . The energies of the electrons emitted from M on formation of  $M^+$  are then measured.



The kinetic energy of the emitted electron,  $KE_e$ , is given by:-

$$\begin{aligned} KE_e &= h\nu - (E_{M^+} - E_M) \\ &= h\nu - BE, \text{ where } BE = \text{binding energy} \\ \text{or } BE &= h\nu - KE_e \end{aligned}$$

The measured BE involves not only the electronic energy but also any vibrational and rotational energy changes. In PS, rotational fine structure is not observed, and can be ignored. This gives:-

$$\begin{aligned} BE &= \Delta E_{M \rightarrow M^+} + \Delta E_{\text{vibrational}} \\ &= IP + \Delta E_{\text{vibrational}} \end{aligned}$$

where IP is the ionisation energy or potential. The ground state, M, will normally have ground state electronic and vibrational ( $\nu=0$ ) levels, and the transition  $M \rightarrow M^+$  where there is no vibrational level change ( $\Delta\nu=0$ ) will be given by:-

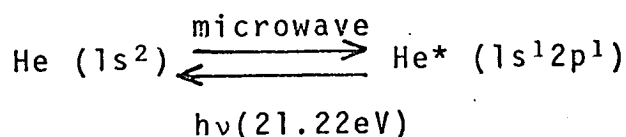
$$BE = IP_{\nu(0,0)}$$

This is known as the adiabatic IP. Transitions to other vibrational states are probable, and so each IP can consist of several vibrational lines forming a band. The probability, P, of a transition occurring is proportional to

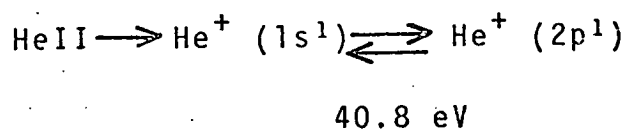
the square of the vibrational overlap integral of the states involved.

$$P \propto \left[ \int \psi_{\text{vib}}(M) \psi_{\text{vib}}(M^+) dr \right]^2$$

The most probable transition is called the vertical IP, and will be the most intense line in a particular PS band. There are two main regions in which photons are produced. One involves the use of uv region photons, and the other uses X-ray region photons. The commonest form of the first uses a HeI lamp as a source of photons. On irradiation with a micro wave discharge, helium gas gives an excited state which produces an intense emission line at 21.22 eV corresponding to the transition back to the ground state.



Using this source, only a limited number of valency shell IPs can be obtained. An extension to 40.8 e.V. can be obtained using a HeII source which gives this characteristic emission on changing ionic states.



However, it is difficult to produce a HeII source, free of HeI, and this causes difficulty in the overlap region. Most BEs are therefore measured using a HeI source.

The second irradiation method is the use of X-ray photons, which are produced either from a  $\text{MgK}_\alpha$  or  $\text{AlK}_\alpha$  source giving photons of 1254, and 1487 eV respectively. In this case, it is possible to determine the experimental IPs of the core electrons as well as those of the valence shell. This method known as X-ray photoelectron spectroscopy, X PS, is of limited use in the valence shell region because of the high

Fig. 1

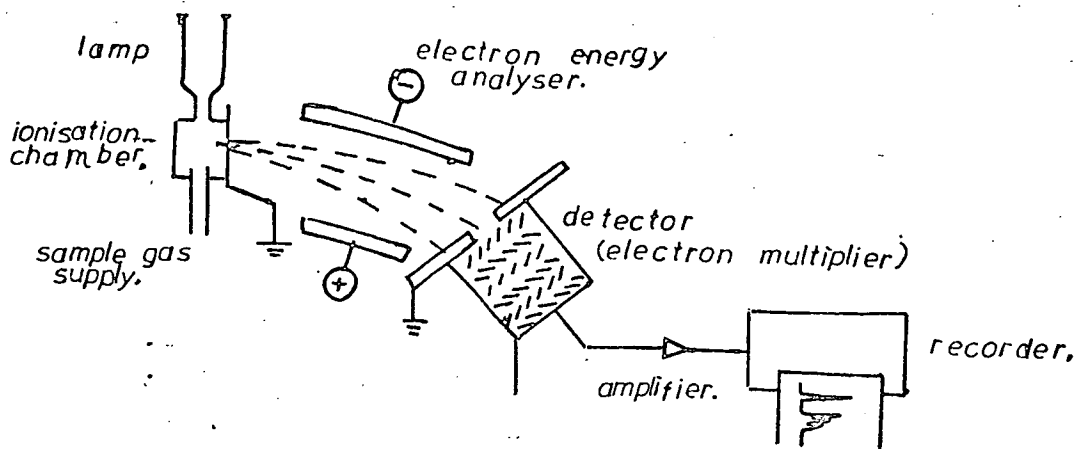


Fig. 2

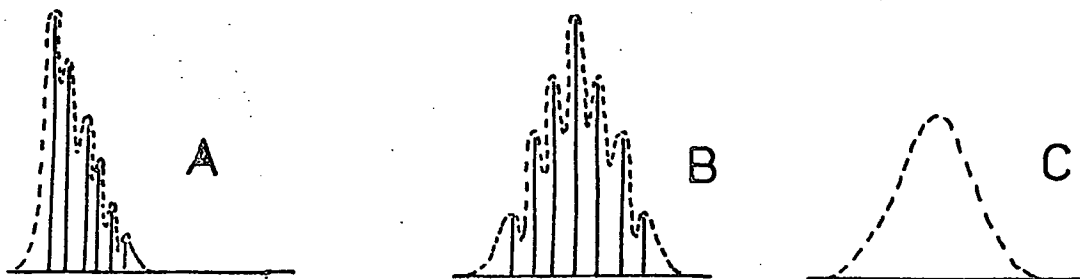


Fig. 3

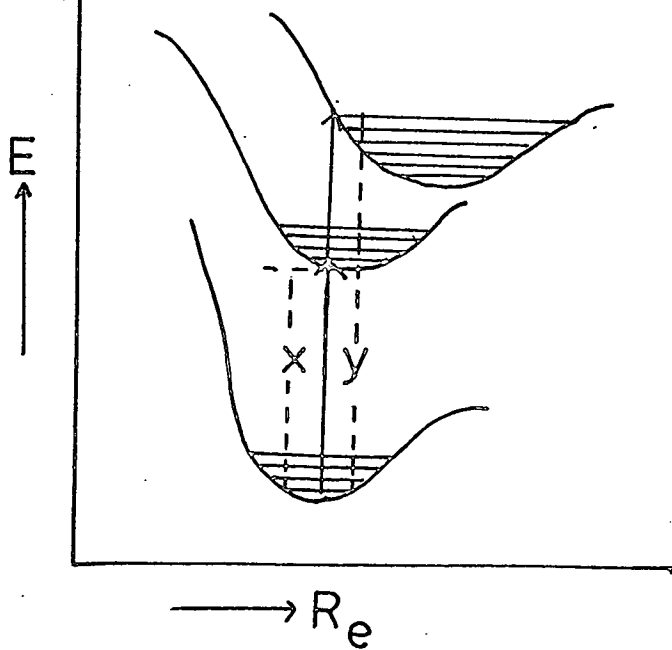
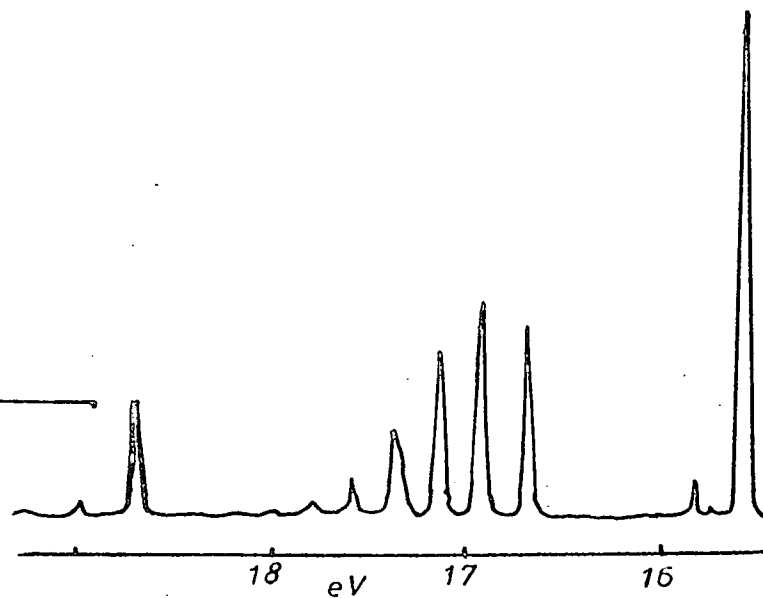


Fig. 4



line width at half height (approximately 1eV), which obscures any vibrational fine structure and overlap of closely placed IPs. Coupled with the low photoionisation cross section of the valence orbitals at these low wave lengths, this produces considerable loss in detail in comparison with the UPS (ultra violet) method which gives a line width at half height of about 0.02 eV. The high resolution possible with UPS enables many closely spaced IPs and also vibrational fine structure to be distinguished. A second disadvantage of XPS is that the samples must usually be solid whereas in UPS, the sample can be in any state initially although all are investigated in the gaseous state. This produces a greater degree of flexibility.

The KEs of the electrons produced on ionisation are measured by passing them through a magnetic field, and the relative numbers of each energy type are measured with an electron multiplier tube and plotted to give IPs of varying intensity. A block diagram of a UPS machine is shown in Fig. 1.

The shape of the bands produced by UPS is often a good indication of the types of orbitals involved in the ionisation processes. Three main band types A, B, C Fig. 2 are produced. Band types A and B both show vibrational fine structure. Type A corresponds to transitions in which the adiabatic IP has the highest probability. This corresponds to the process  $M \rightarrow M^+$ , where the equilibrium geometry is not much altered by the transition, as shown for X in Fig. 3. In these cases, the electron is being removed from a

relatively weakly bonding orbital, such as a  $\pi$ -orbital. In bands of type B, the vertical IP is the most intense transition corresponding to removal of an electron from a more strongly bonding orbital causing the equilibrium geometry to change considerably, as shown for y, Fig.3. These are often lone pair type orbitals. The third type, C, again shows removal of an electron from a strongly bonding orbital, usually a lone pair or  $\sigma$  orbital, and shows no fine structure.

The vibrational spacings also give clues to the type of transition. When a strongly bonding electron is removed, then  $\nu_{M^+}$  is normally less than the corresponding ground state vibrational frequency,  $\nu_{GS}$ , the decrease corresponding to the decrease in force constant. When  $\nu_{M^+}$  is greater than  $\nu_{GS}$ , this indicates removal of an electron from an orbital of antibonding nature. If the vibrational frequency is not altered to any great extent this indicates that the force constants for the ground and ionic state involved are similar, and therefore the electron was emitted from a non-bonding orbital. Using vibrational spacings together with band shapes gives a good guide to the identification of the orbital types involved in the ionisation processes. The UPS spectrum of  $N_2^2$  shows these features well (Fig. 4). The first IP is adiabatic with relatively little change from  $\nu_{GS}$  ( $2345 \text{ cm}^{-1}$ ) to  $\nu_{M^+}$  ( $2150 \text{ cm}^{-1}$ ), indicating removal of an electron from a weakly bonding orbital. The second band is of type B, and shows a considerable reduction in  $\nu$  ( $GS \rightarrow M^+$ ) suggesting electron removal from a strongly bonding orbital. The third band gives  $\nu_{M^+}$  ( $2390 \text{ cm}^{-1}$ ) greater than  $\nu_{GS}$ , indicating removal of an electron from an antibonding orbital

and is of type A. The ground state configuration of  $N_2$  is  $KK (\sigma_g 2s)^2 (\sigma_u^* 2s)^2 (\pi_u 2p)^4 (\sigma_g 2p)^2$ , and the first three bands correspond to removal of an electron from  $\sigma_g 2p$ ,  $\pi_u 2p$ , and  $\sigma_u^* 2s$  respectively.

In polyatomic molecules such as  $C_6H_6$ , no occupied orbitals are completely antibonding, so that the term 'antibonding' becomes less precise in definition. However, if the electron is expelled from a molecular orbital which shows strong bonding character, then the frequency of vibration associated with that bond, or set of bonds will be lowered. If the molecular orbital is largely nodal, then it has a certain amount of 'antibonding' character, and the vibrational frequency will be raised relative to the ground state frequency when an electron is expelled from the bond associated with the nodal character.

The experimental IPs determined using HeI can be correlated with the orbital energies produced by the calculations. This is done using Koopmans' theorem<sup>3</sup> as mentioned earlier, (Ch.1. p52) and approximates as follows:-

$$IP = -e_i \text{ (the orbital energy)}$$

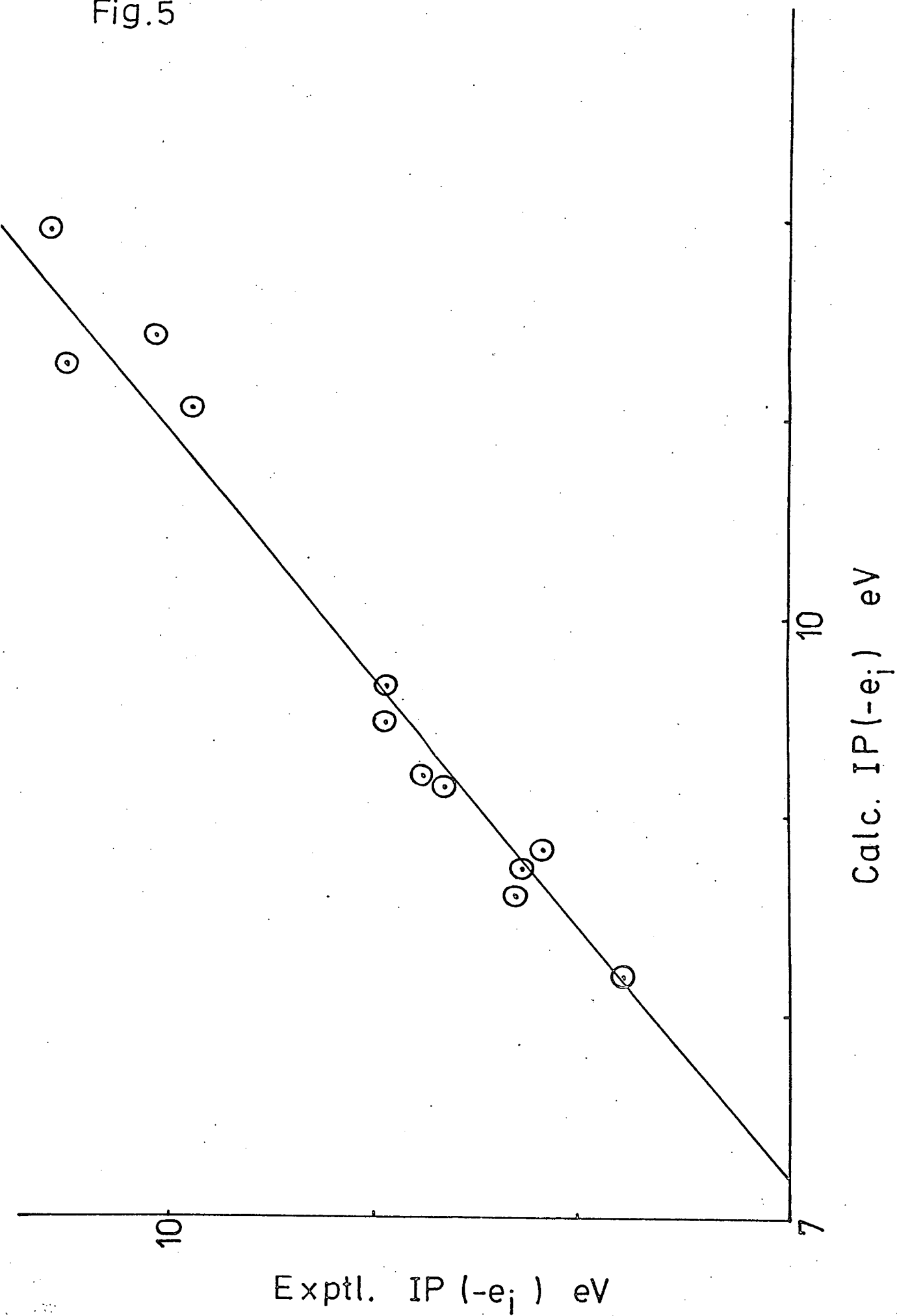
Ideally, a calculation should be performed on the ion, and the value of  $E_M - E_{M+}$  determined accurately for each ionisation process. For vertical IPs,  $E_{M+}$  would be required for several geometries to determine the equilibrium geometry for the state. An LCGO calculation on a molecule such as indole has approximately 900,000 integrals, and to perform similar calculations on all possible ions would take a considerable amount of computer time, and would be an extremely costly exercise. To contemplate such an exercise for several molecules of similar magnitude would be



impracticable, and it is therefore necessary to use some kind of approximation to obtain a correlation. Koopmans' theorem has been successful with small molecules, such as water, as shown by Meyer's <sup>4</sup> calculations on its ground state, and five possible ionic states. His results are shown in Table 1, along with the predicted IPs using Koopmans' theorem, and the experimental results. The IPs from Koopmans' theorem are all larger than the experimental results (Chapter 1 Section 4), but are consistent in the side on which they err, indicating that any linear correlation which exists, is valid and not fortuitous. With the molecules in this work, many of the first IPs are predicted almost exactly using Koopmans' theorem, and the well resolved IPs, normally the first three or four, all fall on the high side of the experimental values (Tables 5 and 9) indicating that a linear correlation is reasonable.

Eland <sup>5</sup> showed that a correlation of the vertical IPs of 14 diatomic, triatomic, and linear tetra-atomic molecules, with the H-F orbital energies using Koopmans' theorem, gave a linear relationship but with considerable scatter of, at best  $\pm 1$  eV, and greater for higher IPs. With such a high scatter in the line, Eland suggests that it does not merit an exact statistical analysis. This would appear to be true, but does not invalidate the use of Koopmans' theorem, as Eland's correlation was done using a variety of molecules with varying numbers of electrons. If a correlation is done using a series of similar molecules, then it would be expected to produce a line with less scatter, if Koopmans' theorem is valid. The graph obtained from plotting the first three experimental IPs of naphthalene, indole, benzofuran and

Fig.5



benzothiophen. against the calculated IPs using Koopmans' theorem is shown in Graph 1, Fig. 5, and gives a scatter of only  $\pm 0.25$  eV at worst. If only one molecule is being considered and IP  $V e_i$  plotted, then provided similar groupings of results occur between observed and calculated IPs, it would seem valid to perform a more exact analysis of results where Koopmans' theorem has been used. The correspondence in groupings of calculated and experimental IPs can be seen in Fig. 10 for naphthalene and related heterocycles.

The most studied example in the field of UPS  $V$  calculation, is that of benzene. The first IP at 9.3 eV is agreed by all workers to results from ionisation from the doubly degenerate  $1e_{1g}$   $\pi$  orbital. The second band is the controversial one whose intensity suggests three ionisations. The MO calculations<sup>6</sup> indicate these to be a doubly degenerate ionisation, and a single ionisation. All calculations give the ordering as  $3e_{2g}$ , a doubly degenerate  $\sigma$  orbital at  $\approx 11.4$  eV, and that at  $\approx 12.3$  eV as an ionisation from the  $1a_{2u}$   $\pi$  orbital. The band at 11.4 eV also shows a vibrational progression which Åsbrink et al.<sup>7</sup> attribute to the same degenerate vibration as that producing the progression on the band at 9.3 eV indicating the 11.4 eV band to be doubly degenerate and therefore corresponding to  $3e_{2g}$ . This agrees with the calculated ordering (low to high) of ionisation,  $1e_{1g}$  ( $\pi$ ),  $3e_{2g}$  ( $\sigma$ ),  $1a_{2u}$  ( $\pi$ ). However Price et al.<sup>8</sup> on an experimental basis proposed the ordering as  $1e_{1g}$ ,  $1a_{2u}$ ,  $3e_{2g}$ , on the grounds that  $\sigma$  bands do not normally show vibrational fine structure. Recently Price<sup>9</sup> has reconsidered his assignment, and now proposes that

Table 1 Calculated and experimental IPs for water

<u>Calculated IPs</u>	<u>IPs from Koopmans' theorem</u>	<u>Experimental IPs</u>
12.3	13.9	12.6
14.5	15.9	14.7
18.7	19.5	18.6
32.2	36.8	32.2
539.6	559.5	539.7

Table 2 Comparison of calculated IPs of benzene <sup>10</sup> (eV)

<u>Orbital Type</u>	<u>Minimal basis</u>	<u>Double Zeta</u>	<u>Open shell</u>
1e <sub>1g</sub> ( $\pi$ )	9.82	8.96	8.55
3e <sub>2g</sub> ( $\sigma$ )	13.51	12.99	13.14
1a <sub>2u</sub> ( $\pi$ )	14.46	13.48	13.99
3e <sub>1u</sub> ( $\sigma$ )	16.27	15.71	15.78
1b <sub>2u</sub> ( $\sigma$ )	17.07	16.48	16.46
2b <sub>1u</sub> ( $\sigma$ )	17.34	17.19	17.27

Table 3 Correlation data of observed and calculated IPs for

$$1 - 5 \text{ using } IP_{\text{expt.}} = aIP_{\text{calc.}} + b \text{ eV}$$

	<u>a</u>	<u>b(e.V.)</u>	<u>s.d.† in slope</u> <u>c</u>	<u>s.d. in</u> <u>intercept</u> <u>d</u>	<u>s.d.</u> <u>overall</u> <u>e</u>
1	0.798	+ 1.202			
2	0.755	+ 1.702	0.014	0.213	0.190
3	0.705	+ 1.850			
4	0.618	+ 2.863	0.021	0.210	0.089
5	0.614	+ 2.723	0.037	0.368	0.179

† s.d. = standard deviation

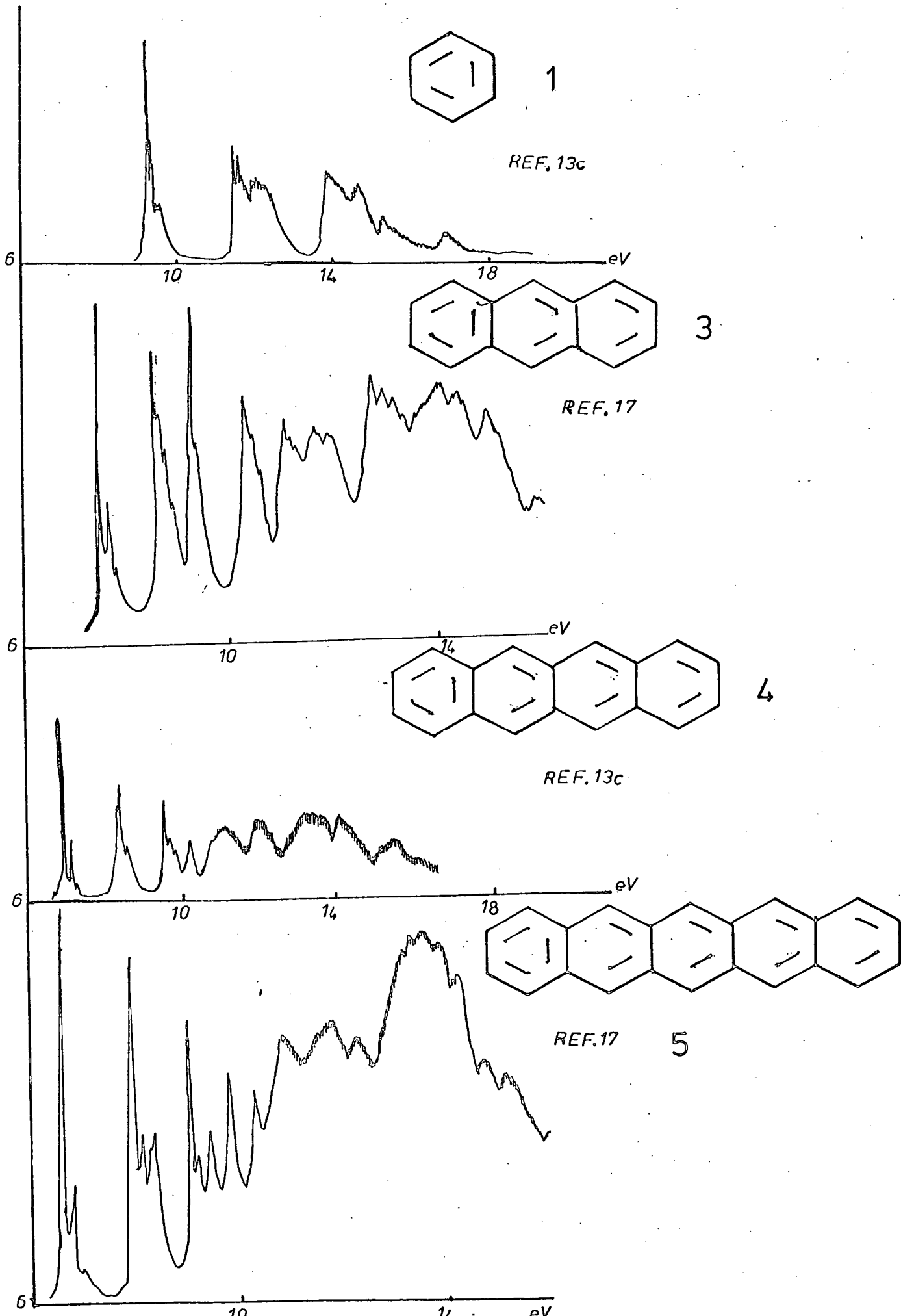
predicted by calculation using Koopmans' theorem. Minimal basis LCGO calculations on benzene, and its first six ionic states<sup>10</sup> predict the ordering of ionisations, using  $IP = E_{mol}^{SCF} - E_{ion}^{SCF}$ , to be  $1e_{1g}$ ,  $3e_{2g}$ ,  $1a_{2u}$ ,  $3e_{1u}$ ,  $1b_{2u}$ ,  $2b_{1u}$ . This is the same order as predicted using Koopmans' theorem. When a double zeta calculation on benzene<sup>10</sup> is used to give Koopmans' IPs, then the two sets of results are numerically in very close agreement as shown in Table 2. It can be concluded, therefore, that provided correlations performed using Koopmans' theorem are for related molecules or for a single molecule, the theorem can be used with confidence.

Results. The hydrocarbon molecules considered here show in their PE spectra bands of types A and C while the heterocycles show all three band types.

Only HeI spectra are reported along with a few XPS spectra where appropriate. The bands recorded using HeI are upper valence shell IPs, and tend to be mainly combinations of 2p levels with the 1s hydrogen levels. Some combinations of 2s with 2p occur in the HeI range, but IPs corresponding to purely s-type orbitals lie outwith this range. The calculated molecular orbital energy levels and their symmetries are shown in Appendix 2.

Reliable PE spectra exist for all the acenes considered in this work, and it was not thought necessary under the circumstances to record the spectra again. Discussion of spectra is therefore based on published results.

Fig. 6



For the bicyclic heterocycles I, PE spectra were recorded using a Perkin Elmer PS 16 spectrometer, and were calibrated with the argon doublet at 15.75 and 15.93 eV.

For the heterocycles II, the PE spectra were recorded on a Perkin Elmer PS 18 † and calibrated using the Xenon line at 12.13 eV and the first peak of the argon doublet at 15.75 eV.

XPS results were obtained using an A.E.I. † ES 200, and were calibrated with the acetone  $1s_{c\text{-methyl}}$  peak.

The spectra of the N-oxides were also recorded using the PS 16 spectrometer.

In the following discussion, y-z is taken as the molecular plane, and ionisation potentials are numbered such that  $IP_1$  is at lowest binding energy.

Hydrocarbons. Experimental PE spectra have been reported for molecules 1, 11, 12<sub>2</sub>, 11,13,14,15,16<sub>3</sub>, 11,12,13c,16,17,18<sub>4</sub>, 13b, 13c, 16<sub>5</sub>, 13c, 17<sub>6</sub>, 20, 21, 22, 7<sub>11</sub>, 12, 14b, 17 and 8<sub>12</sub>, 19. These are shown in Fig. 6 for 1 - 5 and Fig. 7 for 6 - 8. Only 1 and 2 have been discussed in terms of all electron calculations, and the others are assigned on the basis of HMO, PPP or other semiempirical methods. The correlation between calculated and experimental IPs was performed using data from reference 13c for 1 - 5, and from references 17 and 19 for 7 and 8 but in general the agreement on the experimental IPs is satisfactory between authors.

† Thanks are extended to the A.E.I. Scientific Apparatus division for use of their ES 200, and to Perkin Elmer Ltd. for the use of their PS 18.

Fig. 7

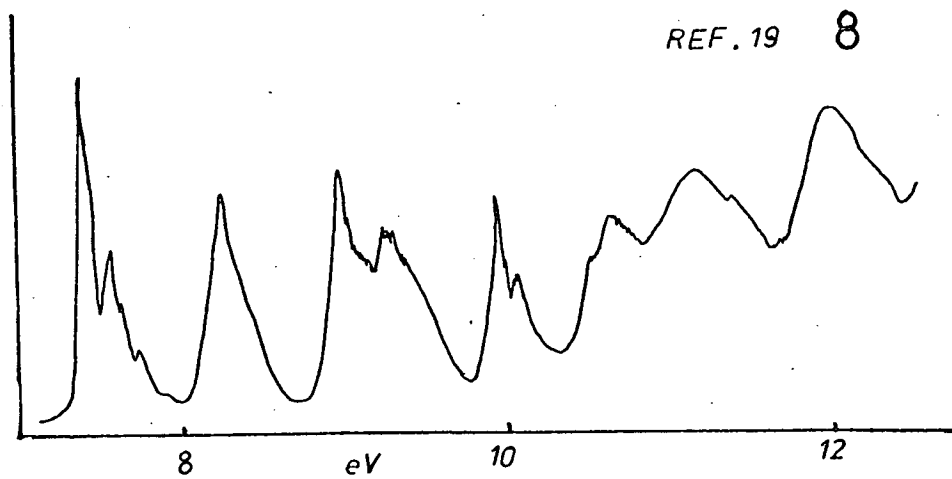
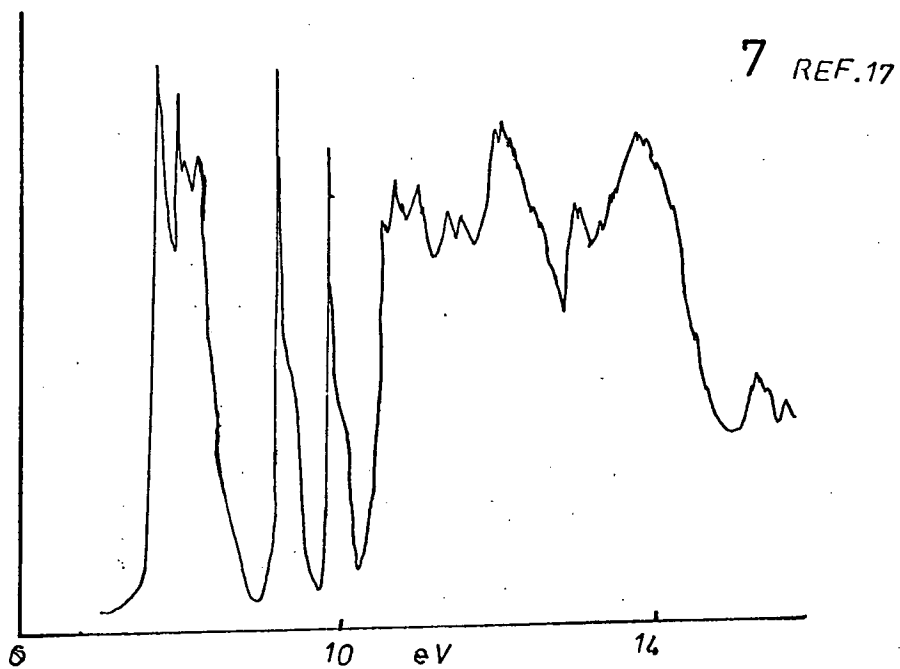
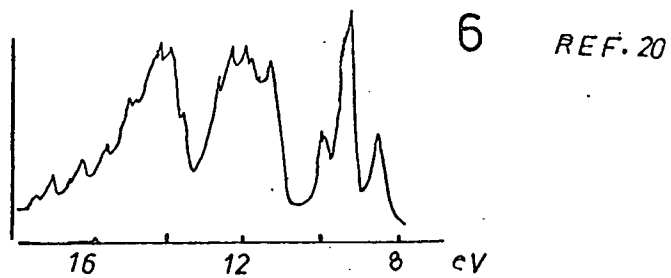
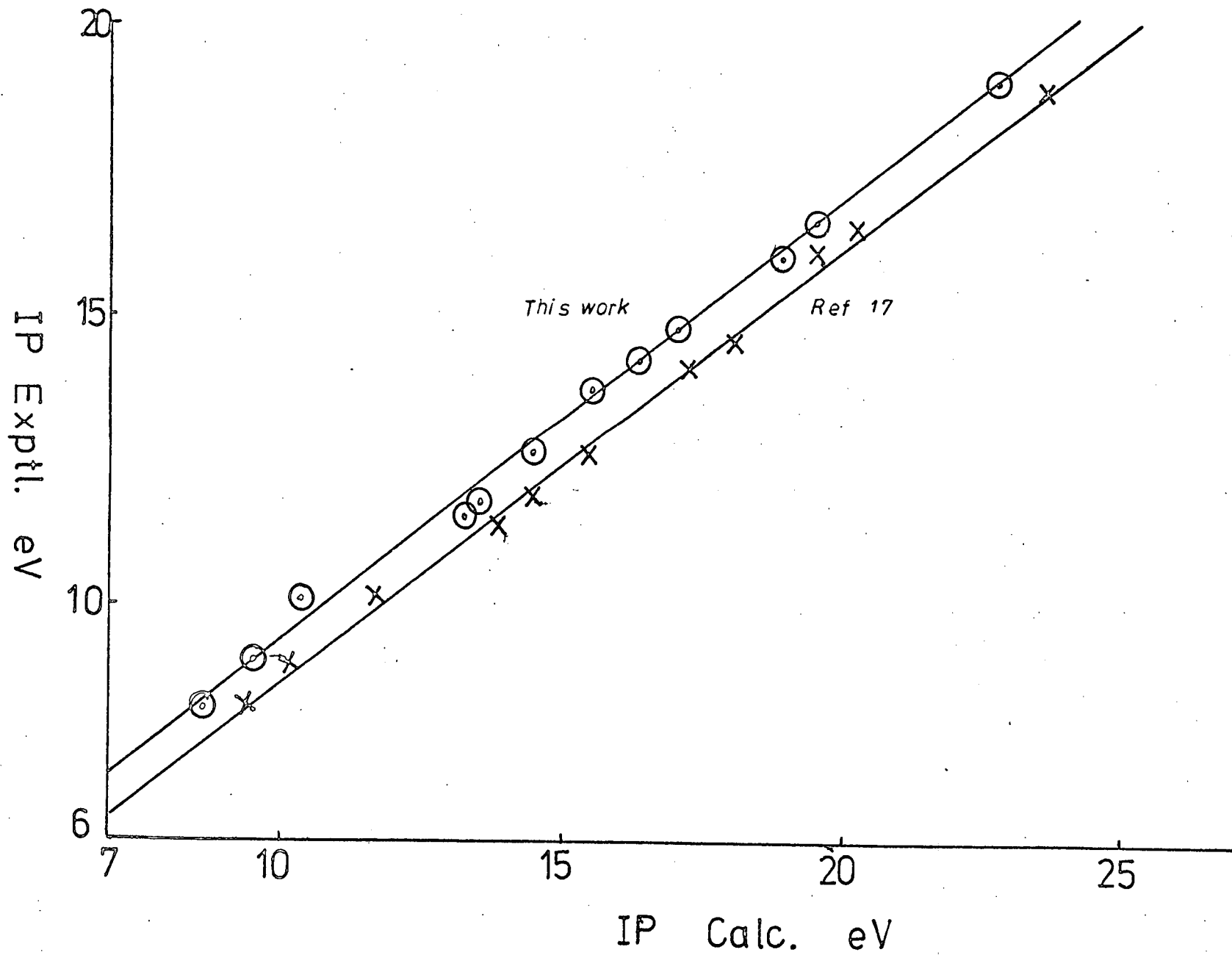




Fig.8 Graph 2



All correlations were done using the relationship

$IP_{\text{expt}} = a IP_{\text{calc}} + b \text{ (eV)}$  (c,d,e) where a and b are constants for a particular correlation corresponding to slope and intercept respectively, and (c,d,e) are the standard deviations in slope, intercept and overall. A correlation

based upon the lowest 63 IPs in 1 - 5, 7 and 8 yields a line represented by  $IP_{\text{expt}} = 0.717 IP_{\text{calc}} + 1.879 \text{ eV}$  (0.016, 0.209, 0.408) indicating a linear relationship between calculated and experimental results.

For naphthalene, a complete analysis of the PE spectrum has been performed, and when a correlation of observed and calculated IPs is carried out using the present data, and that from the all electron calculation by Buenker and Peyrimhoff, (molecular energies ref. 17) the correlation lines are very close, having the same slope but slightly different intercept as shown in Graph 2, Fig. 8. A similar exercise using the IPs of Lindholm et al plotted against the calculated orbital energies of this work gives the relationship  $IP_{\text{expt}} = 0.745 IP_{\text{calc}} + 1.85$ , which is very similar to the result of  $IP_{\text{expt}} = 0.755 IP_{\text{calc}} + 1.702$  obtained in this work.

The early bands of 1 - 5 are  $\pi$  levels as indicated by their shape, and all show fine structure. As the number of rings increases the first IP in each compound moves to lower binding energy (9.2 eV in benzene to 6.7 eV in pentacene) The onset of  $\sigma$ -levels also shifts to lower binding energy from 1 - 5 (11.6 to 10.5 eV) but the shift is not so pronounced as in the  $\pi$ -levels. This shift is well reproduced by the calculations, 9.82 to 6.58 eV for the  $\pi$ -levels and 13.51 to 12.5 eV for the  $\sigma$ -levels.

At the onset of the  $\sigma$ -levels it immediately becomes more difficult to assign bands on a one to one basis but for 1 - 3, the bands can be readily assigned on a group basis as shown for naphthalene, Fig. 10. For 4 and 5, the  $\sigma$ -levels are less well resolved and therefore more difficult to correlate unambiguously. However the spacings between minima are significant as they provide the guidelines for the placing of the centre position of a band, and relate well to those of calculated minima to give group correlations. Using this method it is possible, for example in 5, to assign the intensity maximum around 13.1 eV to the group of levels calculated at around 16.0 eV. This maximum corresponds to the group containing the most strongly bonding  $\pi$ -level,  $1b_{3u}$ . It is noticeable in these larger hydrocarbons that some of the  $\pi$ -levels are lost under the  $\sigma$ -envelope. It is also worth noting that a particular basis set using the H-F method produces a nearly fixed percentage (about 99.5%) of the experimental total and orbital energies. The absolute error in both total and orbital energies will therefore increase as the molecules become larger, and the slope of the correlation line is expected to decline on this basis. This is confirmed by the gradients shown in Table 3 for 1 - 5. The values of b, c, d and e are also shown in this table for 1 - 5.

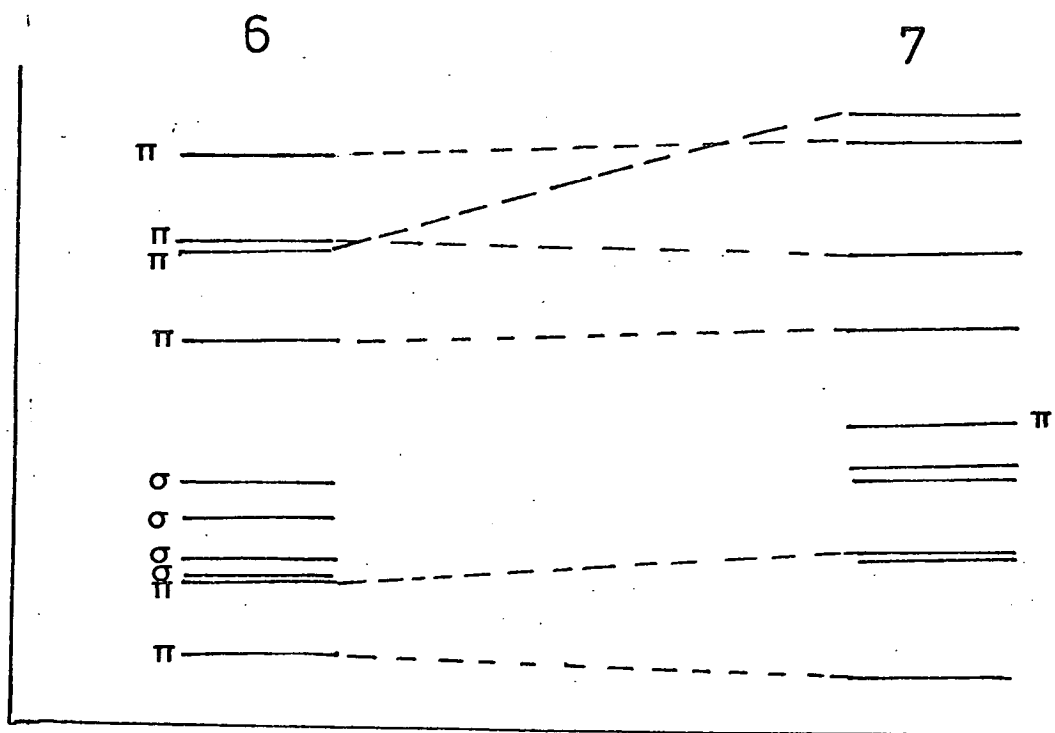
The LCGO calculations confirm the orbital assignments in references 13c, 17 and 19 for 1 - 5 as  $\pi$ -orbitals, and show that in most cases the HMO and PPP calculations have produced the correct ordering of orbitals, the exception being tetracene where the  $b_{3u}/b_{1g}$  and  $b_{2g}/a_u$  pairs are reversed. These

orbitals are labelled  $b_{2g}/b_{1u}$  and  $a_u/b_{3g}$  respectively in reference 13c due to a difference in axes. The authors admit that there is a degree of uncertainty in this assignment due to overlap with bands of higher intensity. Within each pair, the separation experimentally is less than 0.2 eV, and probably too close to produce an unambiguous identification using a semiempirical method. This would suggest that the LCGO calculation will have produced the correct ordering of orbitals for 4.

For biphenyl, 6, no direct correlation can be performed as the LCGO calculation uses a planar geometry while in practice the molecule has a dihedral angle of  $32^\circ$  in the gas phase. This will of course increase  $\sigma$ - $\pi$  interaction and decrease  $\pi$ - $\pi$  conjugation across the rings, and both of these factors are likely to affect the IPs. However, when the present calculation on planar biphenyl is compared with an all electron calculation by Almlöf<sup>23</sup> on twisted biphenyl, it is found that the orbital energies of each calculation agree to within 0.3 eV or less. It should be possible to correlate the experimental IPs with the calculated values using planar biphenyl. The resulting correlation line using the first four IPs is of the form  $IP_{\text{expt}} = 0.604 IP_{\text{calc}} + 3.06 \text{ eV}$  which is similar to that obtained for 4 - 5. The first five levels are shown by calculation to be  $\pi$  levels, and when compared with the PE data the calculation reproduces well the small difference between the second and third IPs ( $\Delta IP_{2 \rightarrow 3} = 0.10 \text{ expt.}, 0.13 \text{ calc}$ )

Both 7 and 8 from geometry and RE data show a similarity to bridged biphenyl, and would be expected, if this is the

Fig. 9



case, to show a PE spectrum similar to planar biphenyl, but perturbed by addition of the ethylene bridges.

Phenanthrene appears to conform to this expected pattern, its spectrum being very similar to that of fluorene<sup>20</sup> which has a planar biphenyl structure with a methylene bridge. The first band in the PE spectrum of 7 (Fig. 7) shows a new component at 8.12 eV when compared with fluorene, and from Hückel calculations<sup>20</sup> was shown to arise from the antisymmetric combination of the ethylene  $\pi$  orbital with the third  $\pi$ -orbital of fluorene. When an orbital correlation is performed between phenanthrene and planar biphenyl using the present results, a similar result is obtained (Fig. 9). The additional  $\pi$ -orbital in *phenanthrene* is  $\pi_5$ , and from calculation is shown to occur after  $\pi_4$ . This is confirmed by the PE spectrum which shows a sharp peak at 10.59 eV which is missing in fluorene<sup>20</sup>.

The orbital ordering for 8 as produced from semiempirical calculations<sup>19</sup> gives the first five levels as " $\pi$ -type with an order in symmetry species of  $2b_{2g}$ ,  $2b_{1g}$ ,  $3b_{3u}$ ,  $1a_u$  and  $2b_{3u}$  from  $D_{2h}$  symmetry. This is confirmed by the LCGO calculation, and a correlation of observed and calculated levels for 7 and 8 using early levels only gives  $IP_{obs} = 0.687IP_{calc} + 2.240$  eV (0.029, 0.295, 0.098) for 7 and  $IP_{obs} = 0.701 IP_{calc} + 2.036$  (0.057, 0.641, 0.328) for 8. The LCGO calculations have shown that the orbital ordering of the hydrocarbon molecular orbitals is generally predicted correctly by HMO or other empirical and semi-empirical calculations. However, these semi-empirical calculations are known to produce  $\sigma$ -levels which are too low in BE with respect to the  $\pi$ -levels and it is often difficult to state categorically the nature of an orbital near the onset of the

$\sigma$ -region of the spectrum where overlap of  $\sigma$  and  $\pi$  is likely to occur. For this reason most correlations are done for hydrocarbons using only the first few  $\pi$ -levels which can be identified unambiguously. With LCGO calculations it is possible with a good resolution spectrum to correlate well into the  $\sigma$ -region using line groupings, and to identify  $\pi$ -bands obscured by the onset of  $\sigma$ -bands.

The bicyclic heterocycles (I) These heterocycles are related to naphthalene by replacement of a CH=CH by NH, O or S. All have  $10\pi$  valence electrons, as does naphthalene.

The calculations suggest that most valency shell molecular orbitals are strongly delocalised, but when the total electron density is summed over the doubly occupied orbitals it is found that it corresponds closely to specific chemical groupings e.g. C-C, C=C, C=N,  $\geq$ N: . In some cases the contributions to this localised electron density can be quite high in individual molecular orbitals. This is in agreement with UPS data where, for example, ammonia,  $\text{NH}_3^{24}$ , shows a lone pair level at 10.8 eV indicating a localised orbital. In contrast, although the total electron density at nitrogen in pyrrole shows a population of 1.62 electrons in the  $\pi$ -lone pair, it is found on further investigation to be delocalised between the  $1b_1$ , and  $2b_1$ , orbitals which from calculation<sup>25</sup> occur at -11.65 eV and -17.69 eV respectively. The PE spectrum shows two low lying levels at 8.23 and 9.22 eV which correspond to  $\pi$ -levels in the calculation. A third  $\pi$ -level is assigned

at 14.7 eV. This is in keeping with a delocalised  $\pi$ -system. If the nitrogen lone pair had been localised, then the spectrum would be expected to resemble that of butadiene<sup>27</sup>, which shows its first two IPs at 9.09 and 11.55 eV with the addition of a nitrogen lone pair level at binding energy  $\approx$  10.8 eV as in ammonia<sup>24</sup>. Using Meyer's<sup>4</sup> data for water (Table 1) the  $\pi$ -localised oxygen lone pair, which is unable to mix by symmetry, is evidenced by the IP at 12.6 eV, and shows only weak vibrational structure as would be expected of a lone pair. In oxygen containing heterocycles, of which furan is the simplest, a similar effect might be expected but, in this case, the lone pair is able to mix by symmetry, and is delocalised as shown by three  $\pi$ -type orbitals giving IPs of 8.9, 10.3 and 14.4 eV.<sup>28</sup> The third IP at 14.4 eV is mainly associated with the oxygen  $p_x$  orbital, and is moved to lower binding energy with respect to that in water.

It is predicted that the bicyclic heterocycles will show characteristics of both naphthalene, and the corresponding monocycle in their PE spectra with variations depending on the degree of delocalisation between the rings. The spectra of 2,<sup>11,29,30,31,32</sup> 15 - 18,<sup>14a,33,34</sup> and 24<sup>35</sup> have previously been recorded but in all cases interpretation has been limited to the first few orbital levels calculated using Hückel<sup>14a</sup>, SPINDO/1<sup>11</sup>, EHT and PPP calculations<sup>34</sup> or by using the effects of perfluorination<sup>36</sup>.

The only all electron calculation<sup>43</sup> on benzofuran, indole



Table 4 Correlation of experimental and calculated IPs  
for heterocycles (I) and related hydrocarbons

$$IP_{\text{expt}} = aIP_{\text{calc}} + b$$

Molecule	a	b	c	d	e
2	0.755	+1.702	0.014	0.213	0.190
15	0.773	+1.479	0.020	0.295	0.237
16	0.779	+1.300	0.010	0.170	0.152
17(sp basis)	0.776	+1.235	0.010	0.153	0.144
(spd + 3s')	0.774	+1.370	0.011	0.166	0.158
18	0.794	+1.051	0.010	0.164	0.163
24	0.767	+1.622	0.010	0.184	0.200
21	0.769	+1.425	0.013	0.204	0.177
22	0.785	+1.341	0.015	0.238	0.230
25	0.786	+1.321	0.017	0.259	0.238

Table 5 Assignment of Vertical IPs to orbital energies  $\epsilon_i$  (eV) for 2, 15 - 18, and 24

(2) Naphthalene

IP	8.3	8.95	10.5	11.05	11.4	11.85	12.5	13.5
$-\epsilon_i$	8.57	9.47	11.29	13.09	13.17	13.44	14.48	15.48/15.54

IP	14.0	14.45	15.85	16.4	18.8	22.1
$-\epsilon_i$	16.24/16.30	17.12/17.14	18.78	19.18/19.54	22.31/22.88/23.27	26.94/27.56/28.73

(15) Indene

IP	8.2	8.95	10.35	11.58	12.0	12.9
$-\epsilon_i$	8.58	9.61	11.60	12.87	13.76/13.93/14.12	14.86

IP	13.6	15.0	16.1	18.0
$-\epsilon_i$	16.05/16.26	16.89/17.07	18.07	20.49/21.45/22.48

Table 5 (cont'd)

## (24) Styrene

IP	8.55	9.25	10.55	11.6	12.2	12.7		13.7
$-\epsilon_i$	8.88	9.84	11.93	13.33	13.56	14.5/14.82		15.93/16.26

IP	15	15.5		16.6	17.85	18.9	19.35	22.3
$-\epsilon_i$	16.93	17.20	/ 18.02	19.44	20.84	22.39	23.14	26.91/27.69

## (16) Benzofuran

IP	8.66	8.94	10.58	11.83	12.67	13.31	14.01	14.31
$-\epsilon_i$	9.16	9.67	11.98	13.30/14.21	14.66	15.06	16.00	16.68

IP	14.54	15.69	16.09	16.69	18.54	19.30	
$-\epsilon_i$	16.92/17.66	18.41	18.88	19.78	21.82/22.34	23.47	

Table 5 (cont'd)

## (17) Benzothiophen

Basis	IP	8.75	8.75	10.07	11.20	11.45	12.15	12.45	13.10
spd + 3s'	- $\epsilon_i$	8.76	9.21	11.45	12.97	13.32	13.87	14.32	15.28
sp	- $\epsilon_i$	9.00	9.44	11.67	12.84	13.51	13.96	14.44	15.36

	IP	14.25	15.20	15.80	17.98	19.27
spd + 3s'	- $\epsilon_i$	16.07/16.52/16.64	17.68	18.30/19.16	21.22/21.71	23.03
sp	- $\epsilon_i$	16.35/16.75/16.76	17.75	18.43/19.35	21.35/21.83	23.23

## (18) Indole

IP	7.79	8.18	9.88	11.12	11.56	12.24	13.26
- $\epsilon_i$	8.20	8.83	11.08	12.99	13.46	14.06	15.24/15.67
IP	13.77	14.28	15.40	17.03	18.51		
- $\epsilon_i$	16.36	16.60/16.73	17.96/18.22	10.23	21.98/23.01		

and their quinoid isomers, lists only the first ionisation of indole and benzofuran as data taken from references 33 and 42.

The present calculations have afforded a complete study of the HeI spectra of these compounds which clearly show a large number of well resolved IPs. The spectra are shown in Figs 10 and 11. It can be seen that relatively few multiple assignments are necessary, but are assigned where required on the basis of band intensity and groupings of IPs as found from calculation. In each case, a linear correlation of experimental and computed IPs was found, and the results are listed in Table 4 using the same formula as for the hydrocarbons (Table 3). In all cases the observed IPs were found to be  $\approx 0.80$  of the calculated ones, as was found for the hydrocarbons. This was expected as all the calculations were performed using the same basis set, but less variation is found in the a parameters here than for the hydrocarbons due to this series being close to one another in molecular energy with the exception of naphthalene and styrene. The correlation figures used are shown in Table 5, and multiple assignments are shown. Vertical correlation lines of observed and calculated IPs are shown on the spectra, Figs 10 and 11. In most cases vibrational fine structure can be seen on the early bands, and Fig. 12 shows a 5 x expansion of the early bands in naphthalene, indole and benzofuran. The vibration frequencies have been measured from the spectra, and compared with measured gas phase vibrations of the neutral molecule which are

Fig. 10

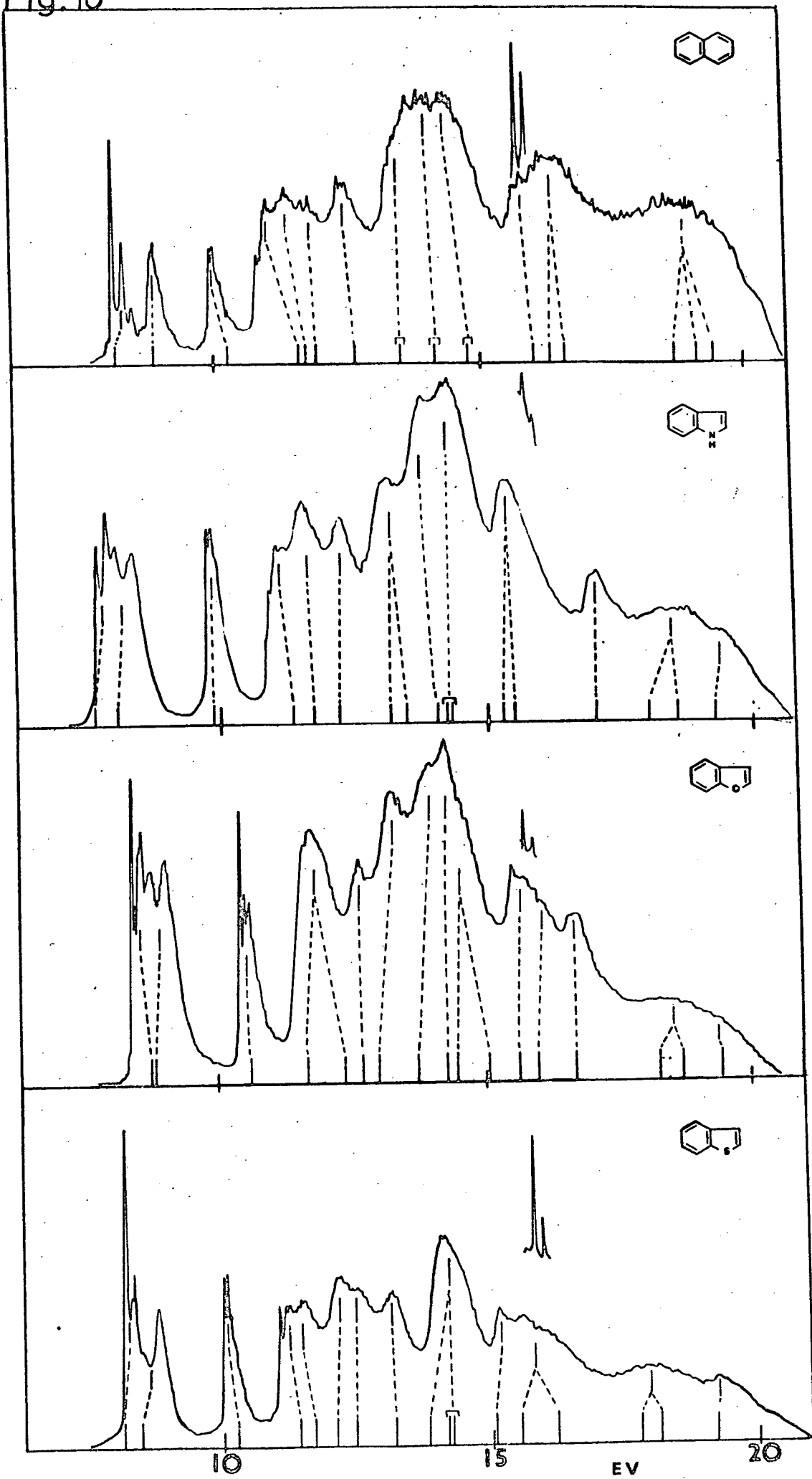


Fig 10 (cont'd)

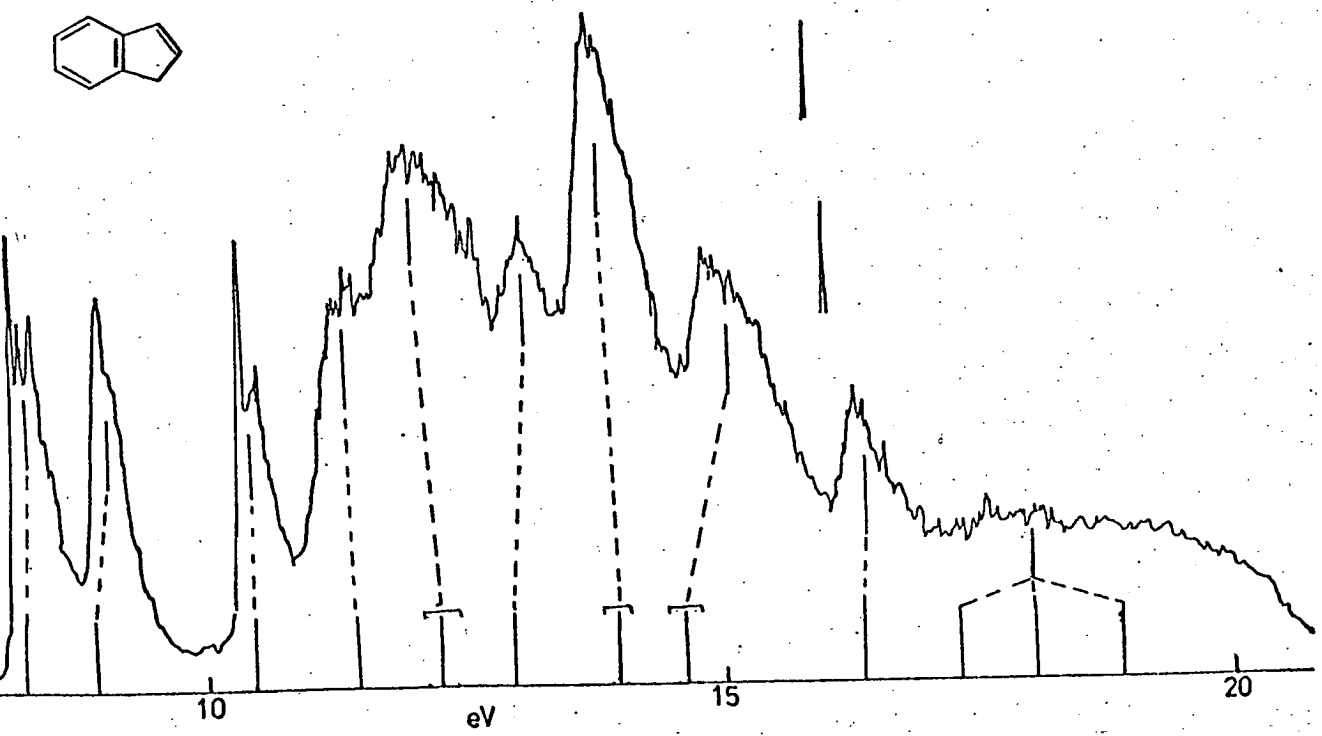
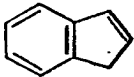
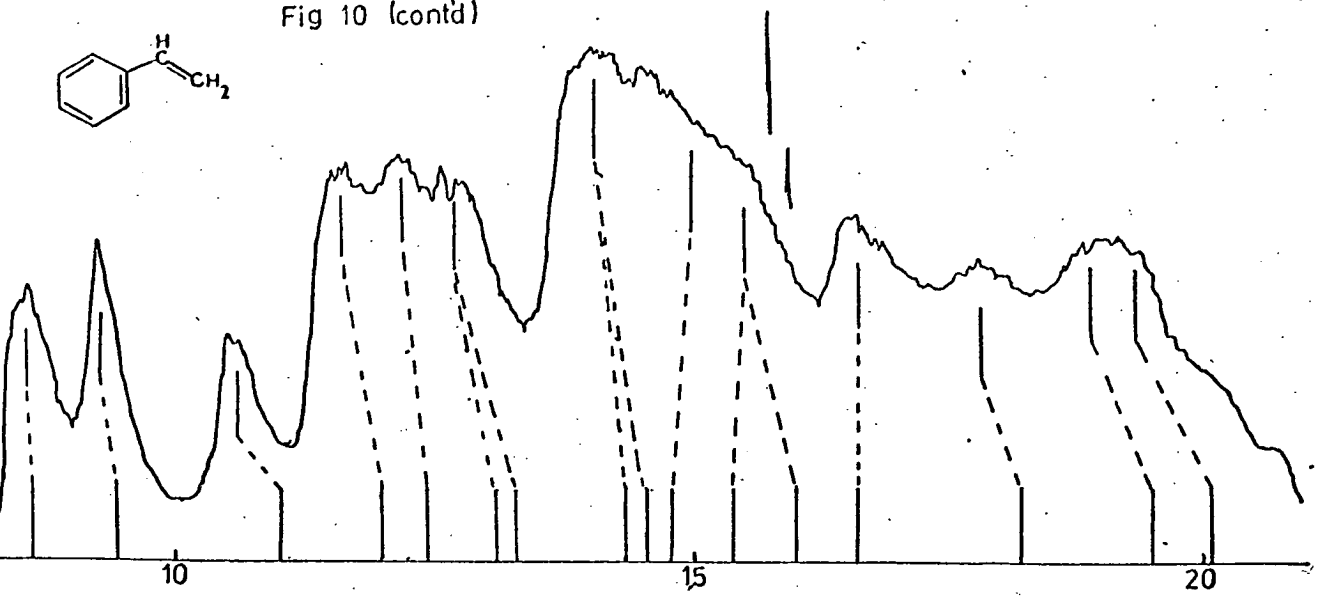
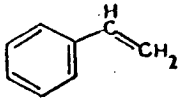
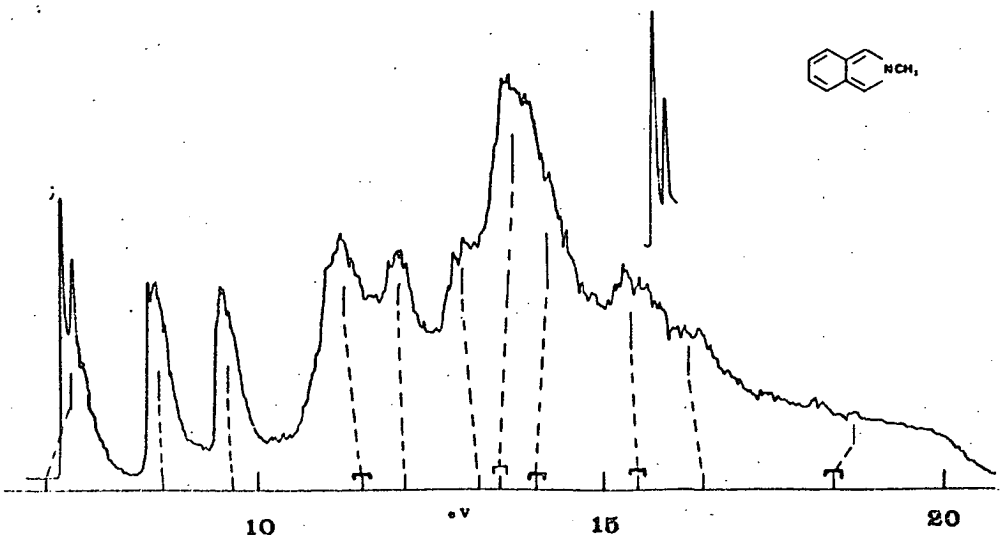
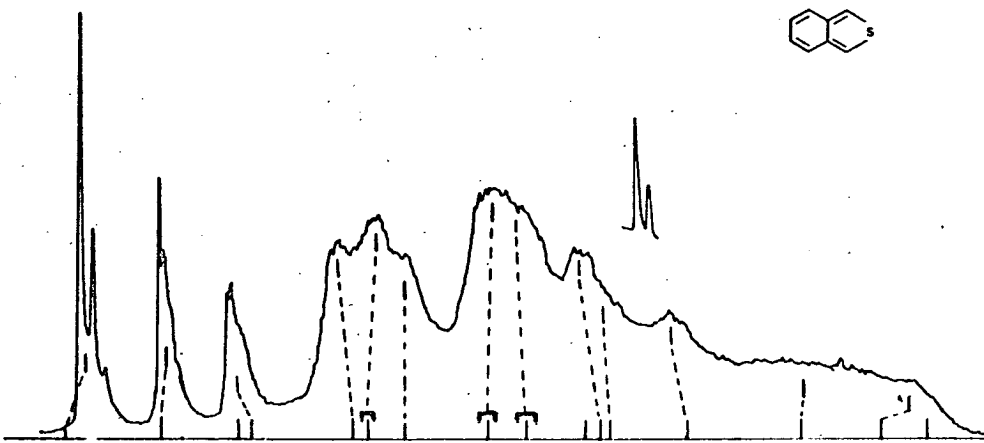
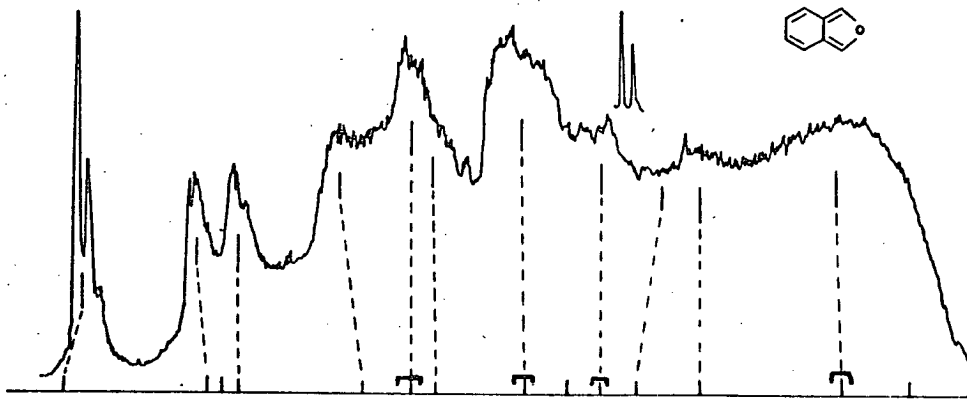
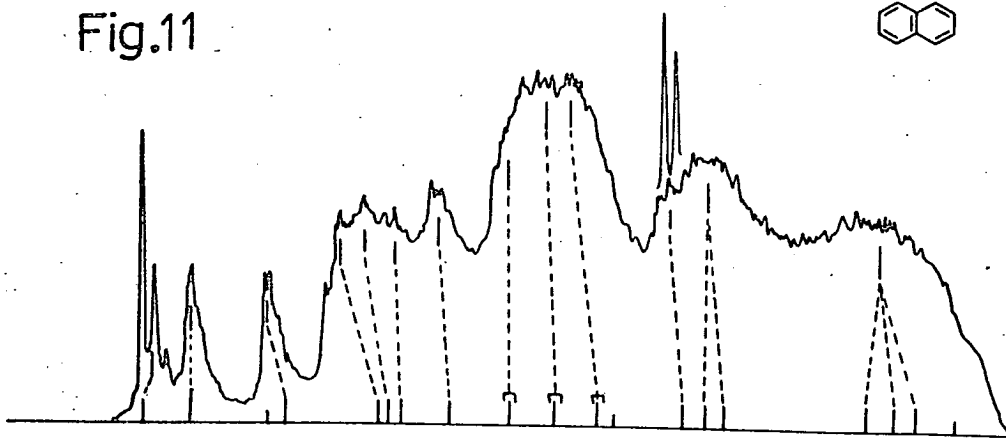


Fig.11



10

V

15

20



Fig.12

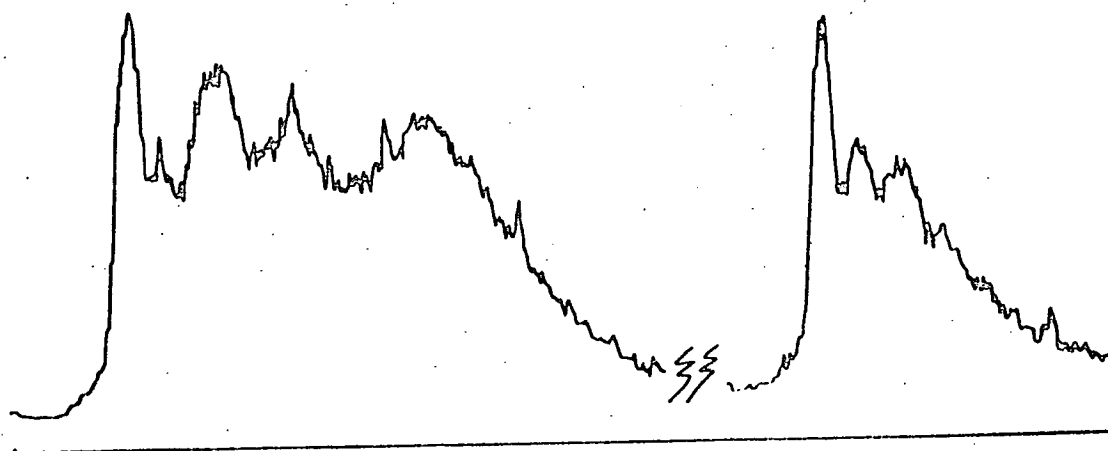
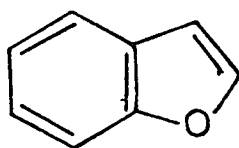
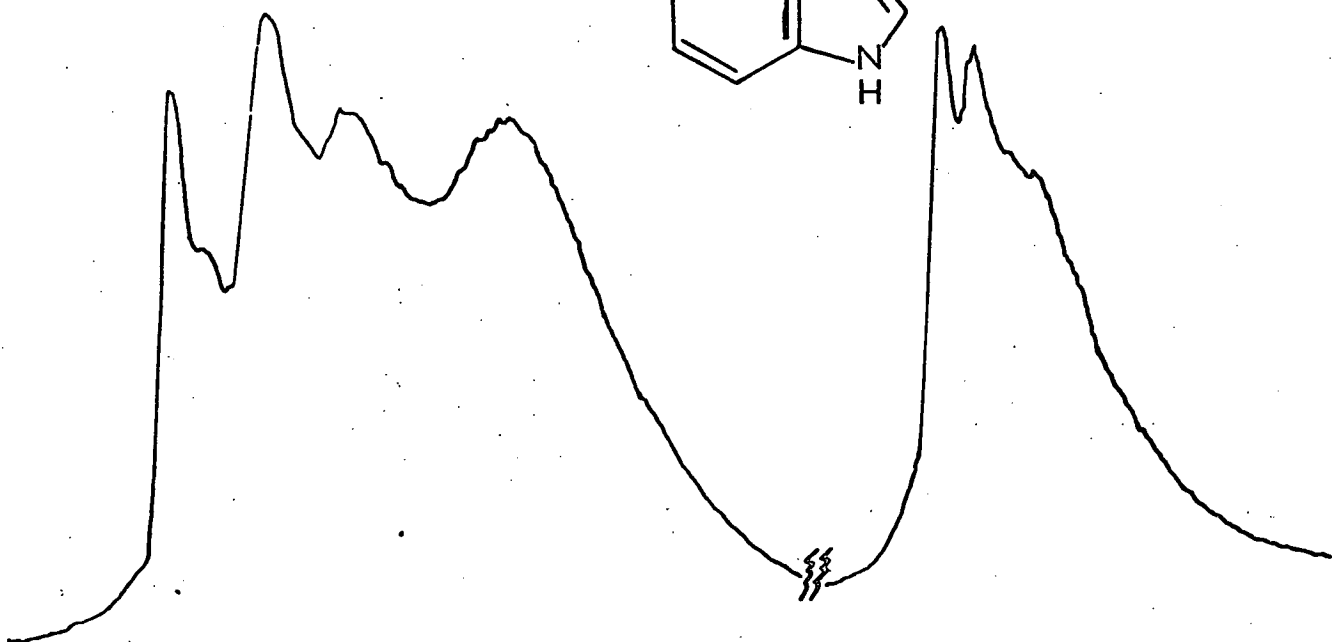
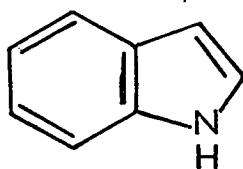
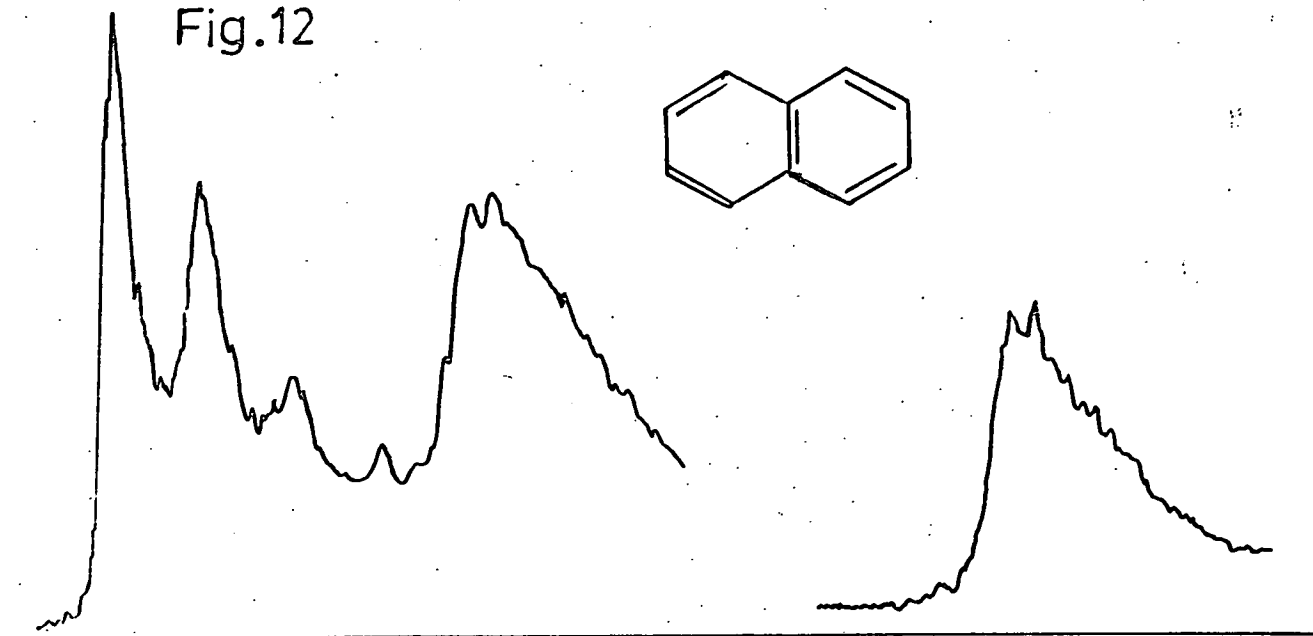
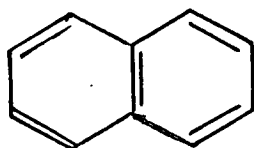


Table 6 Vibrational frequencies excited in the photoelectron spectra ( $\text{cm}^{-1}$ ) (neutral ground state modes in parentheses)

	IP <sub>1</sub>	IP <sub>2</sub>	IP <sub>3</sub>
Naphthalene	1400(1477)		1265(1393) 390(511)
Indole	1350(1454) 520(612)	2620	645(612)
Benzofuran	1480(1254) 580(540)		645(766)
Benzothiophen	1310(1497) 850(884)		

Table 7(a) Naphthalene ground state molecular vibrations and associated frequencies ( $\text{cm}^{-1}$ ).

<u>Vibration</u>	<u>(<math>\text{cm}^{-1}</math>)</u>	<u>nature of vibration</u>
$\nu_9$	511	C $\alpha$ $\gamma$ stretching
$\nu_2$	1477	(C $\alpha$ $\beta$ + C $\gamma$ $\gamma$ ) stretching + CH bending
$\nu_3$	1577	(C $\beta$ $\beta$ + C $\gamma$ $\gamma$ ) stretching + CH bending
$\nu_4$	1393	(C $\alpha$ $\gamma$ + C $\gamma$ $\gamma$ ) stretching

(b) Changes in bond population for naphthalene

<u>IP</u>	<u>change</u>
1	-0.030 e across C $\alpha$ $\beta$ +0.019 e across C $\beta$ $\beta$
3	-0.018 e across C $\alpha$ $\beta$ +0.009 e across C $\beta$ $\beta$ + 0.021e across C $\gamma$ $\gamma$

excited in the electronic (hot band) or fluorescence spectra<sup>29,30,37</sup>. For 16, 17 and 18 no detailed information is available on the molecular vibrations of the molecules, and only a tentative assignment can be made, Table 6.

For naphthalene, a normal co-ordinate analysis has been reported<sup>31</sup>, and leads to the information shown in Table 7a for the ground state molecular vibrations. The way in which ionisation will affect these vibrations can be studied by consideration of changes in bond population across the bonds involved in the vibration. If the population is decreased, the force constant for the vibration will be decreased, and the associated frequency will decrease. The opposite effect will occur for an increase in population. The ionisations  $IP_1$  and  $IP_3$  in naphthalene lead to the bond population changes shown in Table 7b. From these results, the frequencies which will be most affected are 2, 3 and 4 with 2 and 4 being lowered and 3 being raised relative to the ground state. The measured frequencies for naphthalene are assigned using this information.

The spectra, Figs 10 and 11, are all similar in structure with the first three IPs well defined in each case, the first two IPs being close together. Calculations give the first three levels as  $\pi$ -levels. This agrees with the spectral evidence, the first three bands being sharp, and showing vibrational structure. From calculation, the fourth level is  $\sigma$ , but for benzothiophen, the onset of the fourth band shows some vibrational structure suggesting another  $\pi$ -level. Of 15 - 18, the calculated energy difference

between levels four and five is least for benzothiophen (17) with a value of 0.35 eV, and greatest for benzofuran (16) at 0.91 eV. There is a possibility that for 17 the fourth and fifth orbital energy levels are too close together to produce an unambiguous ordering from the calculation.

It has been suggested<sup>11</sup> that a sulphur atom is electronically similar to a C=C bond, and if this is the case, in going from naphthalene to benzothiophen, the electronic structure should not be markedly affected. From the PE spectra of 2 and 17 this is indeed the case. For indole and benzofuran the first two IPs are much closer together than in naphthalene with the third IP isolated. In benzothiophen the IP separation in the first three IPs is  $\Delta IP_{1-2} = 0.58$  eV and  $\Delta IP_{2-3} = 1.32$  eV compared with 0.65 eV and 1.10 eV in naphthalene, and benzofuran and indole giving values of 0.28, 1.64 and 0.39 and 1.70 eV respectively. Clearly the PE of 17 resembles that of 2 more closely in the early region than those of 16 and 18.

Comparing the PE spectra of cyclopentadiene<sup>32</sup>(14), furan<sup>28</sup>(10), thiophen<sup>38</sup>(11) and pyrrole<sup>39</sup>(12) with those of the corresponding bicyclic compounds, several similarities can be seen. In the spectra of 10 - 12 and 14 the first two IPs lie close together but not so closely as in the bicyclic compounds. The third IP in the monocycles comes at much lower binding energy than that in the bicycles e.g. pyrrole, 3rd IP, 13.0 eV, indole, 3rd IP 9.88 eV. On closer examination of the nature of the molecular orbitals associated with the first three IPs of the monocycles, it can be seen that there is a 1:1

Fig. 13.

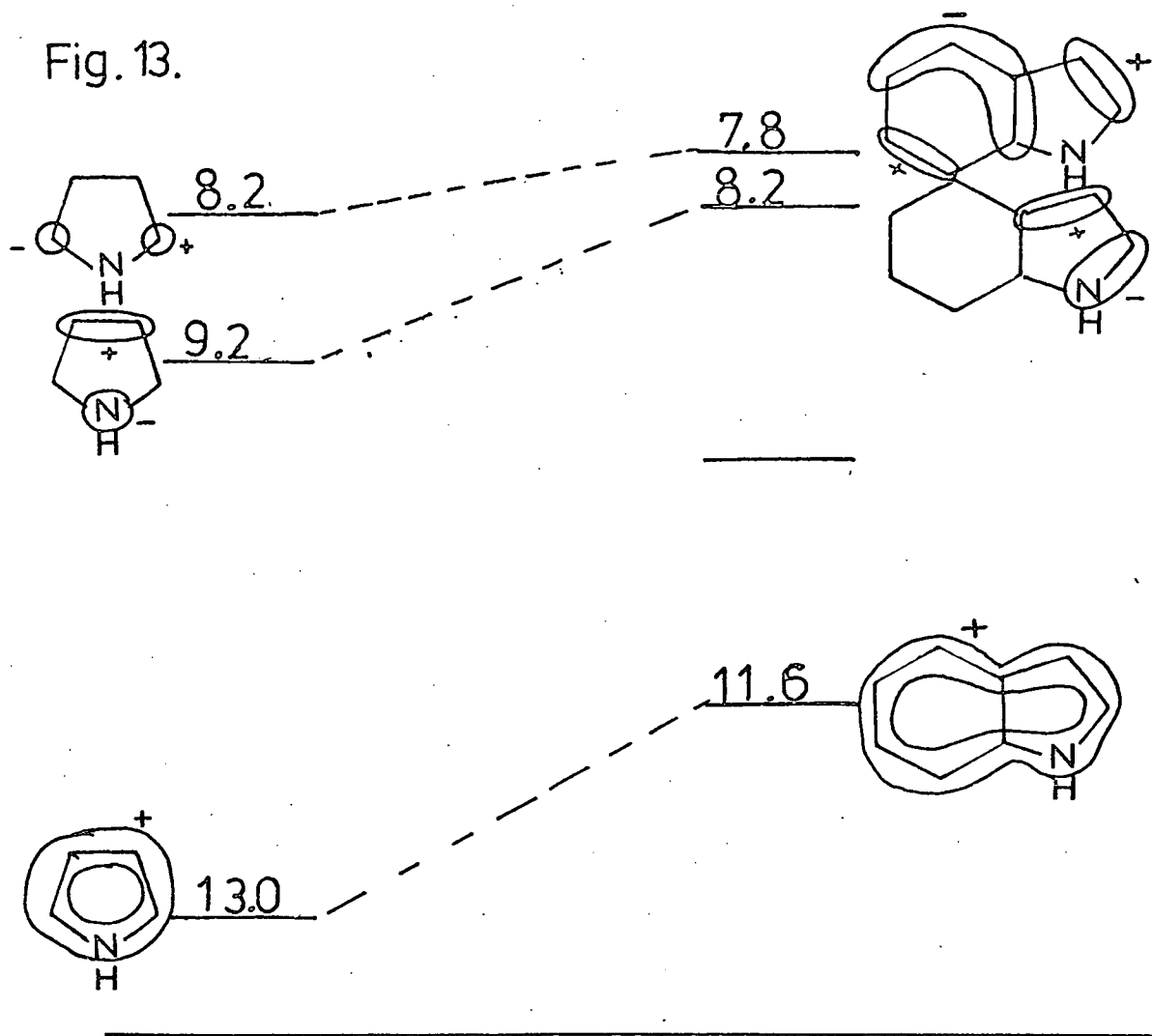
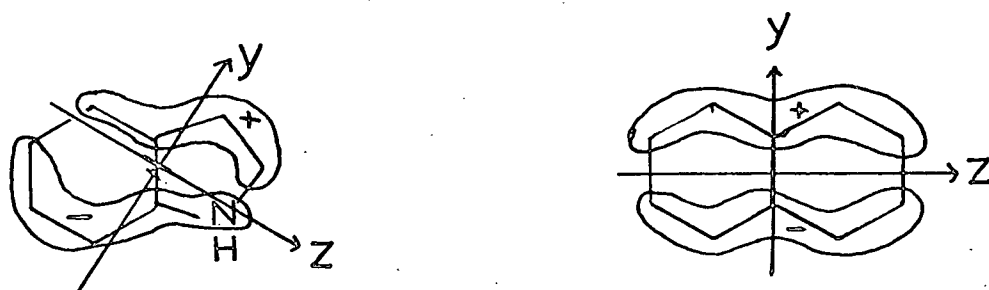


Fig.14



correspondence in the orbital types for 10 - 12 and 14, and when compared with the nature of the molecular orbitals in the bicycles, a similar orbital nature is found in the heterocyclic ring in orbitals 1( $\pi_1$ ), 2( $\pi_2$ ) and 5( $\pi_4$ ). These correspond to orbitals 1( $\pi_1$ ), 2( $\pi_2$ ), and 3( $\pi_3$ ) in the monocycles. A typical example of  $\pi$ -orbital correspondence is shown in Fig.13 for pyrrole and indole along with observed IPs. In each case, all three IPs are moved to lower binding energy in the bicycles, the first by  $\approx 0.6$ eV and the second and third monocyclic IPs by 1 - 2 eV. IP<sub>3</sub> in 15 - 18 is in all cases an addition to the spectra of the monocycles, and corresponds to the molecular orbital,  $\pi_3$ , with a nodal plane through the 'apparent' z - symmetry axis of the molecules as show in Fig.14 for indole.

For naphthalene (2), the first three IPs are more evenly spaced, but the overall spectrum is similar to that of the other molecules. The first three IPs are calculated to be  $\pi$ -levels, the fourth a  $\sigma$ -level, with the third IP,  $\pi_3$ , having a molecular orbital type corresponding to  $\pi_3$  of 15 - 18 but with the nodal plane through the true z-axis Fig. 14.

In the quinoid series: 20 - 23, and 25, only 22, benzo(c)thiophen has a published PE spectrum<sup>40</sup> other than the one shown here<sup>41</sup>, Fig.11, probably due to the difficulty in preparing these compounds, and their relative instability when prepared. No derivative of 20, suitable for UPS study could be made, those already existing being complexes of compounds such as iron nonocarbonyl<sup>13</sup>.

Isoindole (23) is not readily obtainable, but the N-methyl compound is accessible for study, and was used to give information on the structure of isoindole itself as calculations on each produce similar types of molecular orbitals as shown in Appendix 2. A correlation list of observed and calculated IPs is given in Table 8, and the results of a linear correlation of computed and experimental data are given in Table 4.

For 20, 21 and 23 the order of the first 5 levels from calculation is  $\pi_1, \pi_2, \pi_3, \sigma_1, \pi_4$  whereas benzo(c)-thiophen (22) gives  $\pi_1, \pi_2, \pi_3, \pi_4, \sigma_1$ . The orbital energy difference between levels four and five for 22 is only 0.13 eV compared with 1.09 eV for 21 and 0.75 eV for 23, and is perhaps too small to produce an unambiguous order of levels. The EHT and PPP calculations<sup>40</sup> produce an order of  $\pi, \pi, \pi, \sigma$ , but no information is given about the fifth level. The spectra of the quinoid series resemble the PE spectrum of naphthalene in the 7 - 12eV region, more than those of the Kekulé series, in the more even spacing of the first four bands. When the nature of the  $\pi$ -orbitals of 20 - 23 is compared with the corresponding monocycles, it is apparent from the placing of large eigenvectors that the nature of the first three orbitals of the monocycles is almost unaffected by addition of the carbocyclic ring, which can be attributed in part to a retention of  $C_{2v}$  symmetry, showing that the monocyclic structure is retained to a far greater extent, in these compounds than in the Kekulé series where, although the orbital types are maintained from monocycle to

Table 8 Assignment of Vertical Ips to Orbital Energies  $\epsilon_i$  (eV) For 21, 22 and 25

(21) benzo(c)furan

IP	7.92	9.58	10.21	11.66	12.72	13.02
$-\epsilon_i$	8.06	10.87	11.53	13.49	14.58/14.70	15.12

IP	14.30	15.45	16.32	16.92	18.92
$-\epsilon_i$	16.70/16.72/17.12	18.20/18.33	18.98	20.11	22.00/22.37/24.07

(22) benzo(c)thiophen

IP	7.80	8.99	10.00	11.45	12.01	12.49
$-\epsilon_i$	7.84	9.69	11.29	13.16	13.29/13.55	14.09

IP	13.70	14.02	14.94	15.28	16.27	18.22	19.75
$-\epsilon_i$	15.46/15.86	16.12/16.51	17.71	17.89	19.33	21.42/21.45	22.92



Table 8 (cont'd)

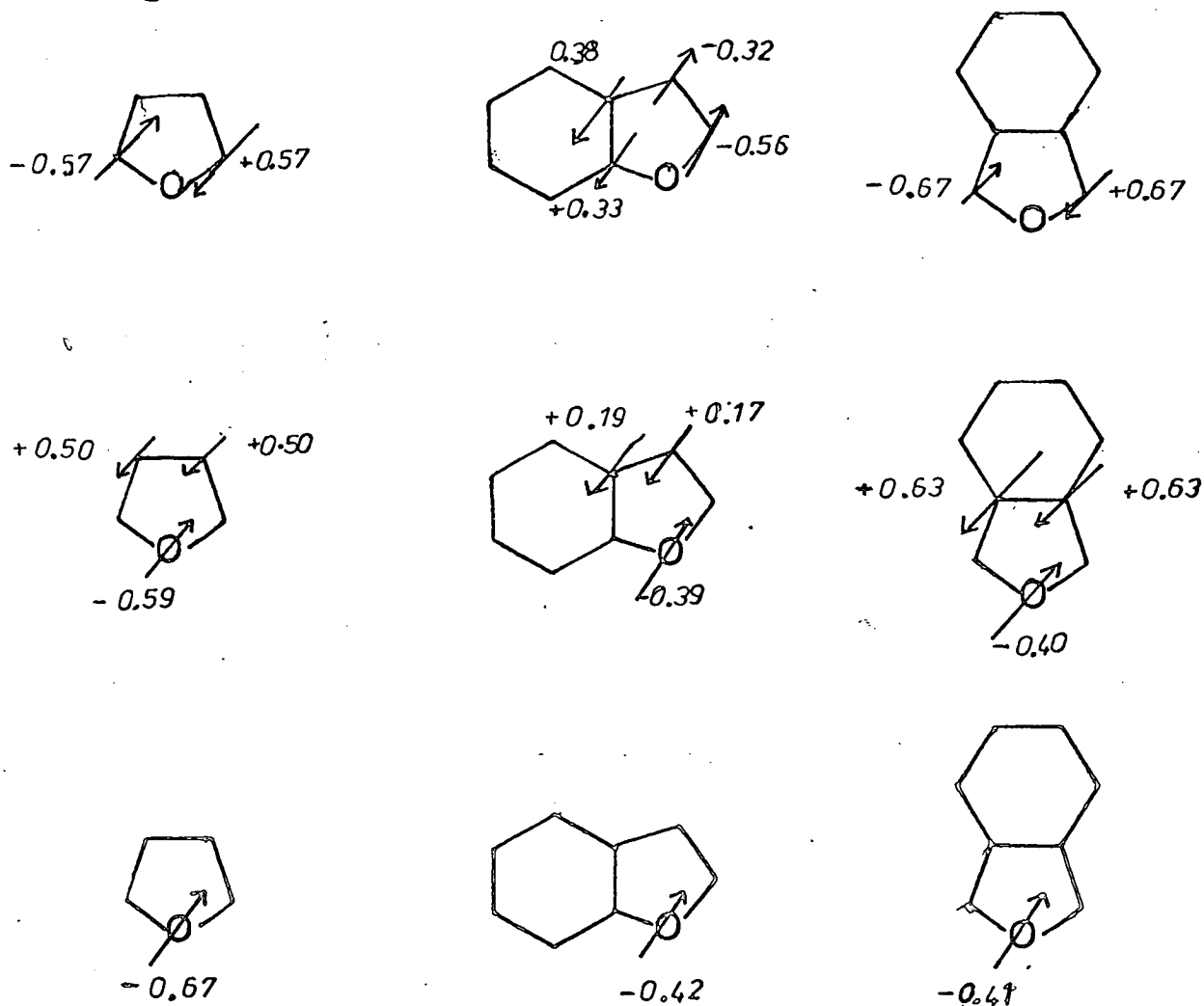
(25) N-methyl isoindole

IP	7.27	8.52	9.56	11.24	12.04	12.97	13.71
$-\epsilon_i$	7.14	9.26	10.55	12.71/13.16	13.74	15.12	15.42/15.60
IP	14.20	15.40	16.24	18.7			
$-\epsilon_i$	15.89/16.10/16.46	17.88/18.04/18.21	19.27	20.83/21.66/22.49			

Table 9 Difference in observed first and second IPs

	pyrrole	indole	N-methyl isoindole
$\Delta IP_{1-2}$ (eV)	1.00	0.39	1.25
	furan	benzofuran	benzo(c)furan
$\Delta IP_{1-2}$ (eV)	1.40	0.28	1.66
	thiophen	benzothiophen	benzo(c)thiophen
$\Delta IP_{1-2}$ (eV)	0.60	0.48	1.19
	cyclopentadiene	indene	
$\Delta IP_{1-2}$ (eV)	2.10	0.75	

Fig. 16



bicycle, a greater degree of delocalisation is shown. An example of this is given by comparing furan, benzofuran, and benzo(c)furan, Fig.16, giving  $p_{\pi}$  eigenvectors. Positions not showing eigenvectors are effectively nodal in that orbital.

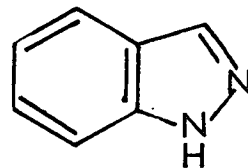
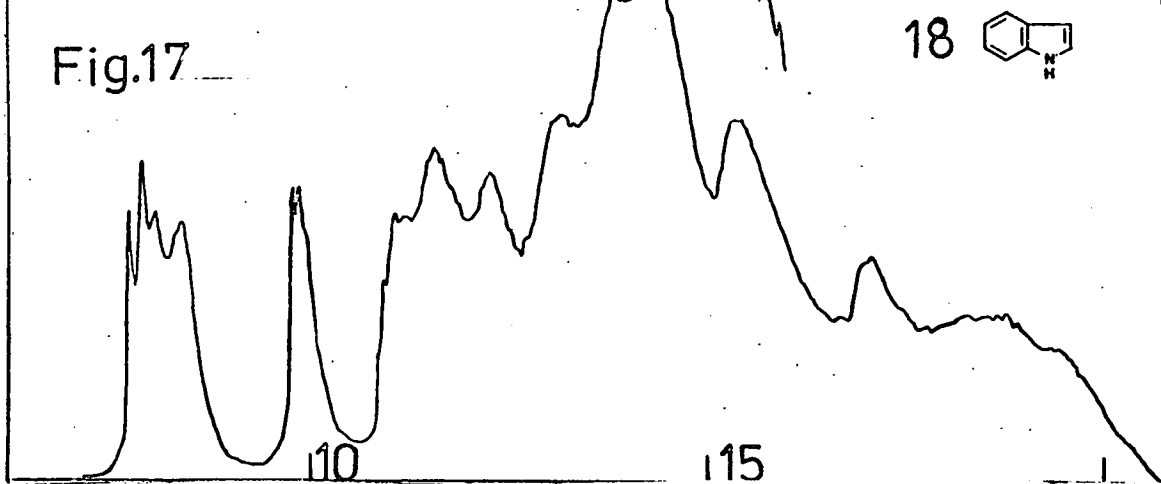
This idea is supported in part by the PE spectra which show the separation of the first two IPs to be similar in magnitude to those of the monocycles themselves, Table 9, indicating that in the early regions the PE spectra of the quinoid compounds are closer to those of the monocycles than are those of the Kekulé series.

Bicyclic Heterocycles II In this series PE spectra have been previously reported for only  $28c^{11,14b}$ ,  $29b^{11}$  and  $30b^{40}$  and  $c^{11,40}$ , most of which are sulphur containing heterocycles. Only  $28c^{14b}$  and  $30b,c^{40}$  have spectra available for comparison, and in all cases  $^{11,14b,40}$  only the first four or five IPs have been assigned. On comparison of the published spectrum of benzothiazole,  $28c^{14b}$ , with that in Fig.19, it is clear that better resolution has been obtained here in the 11 - 21 eV range, but the general form of the two spectra is the same. Both 30b, and c (Fig 20+22) show comparable resolution to the spectra of reference 40.

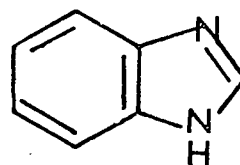
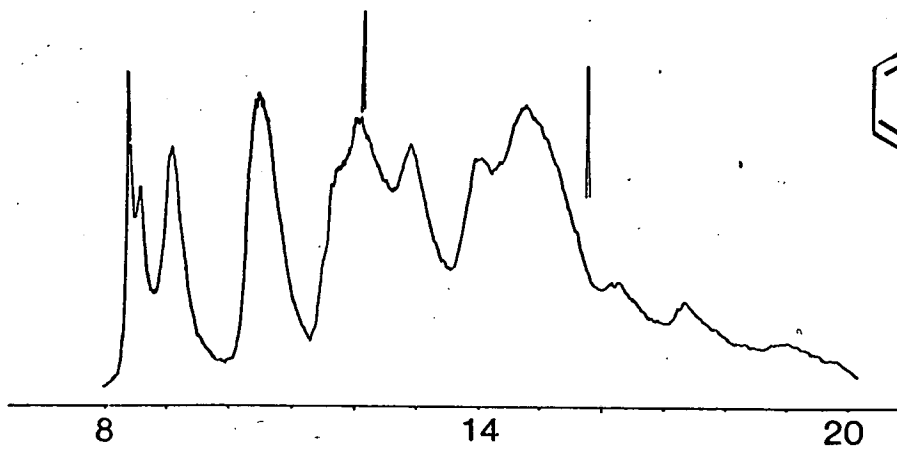
In the series 18, 26a, and 28a, shown in Fig.17 with their PE spectra, it is expected that replacement of  $\text{>CH}$  by  $\text{>N}$  will produce a new IP in the low energy region as the nitrogen lone pair will tend to be more localised than CH. The spectrum of indazole, 26a, shows the first two IPs to

Fig.17

18



26a



28a

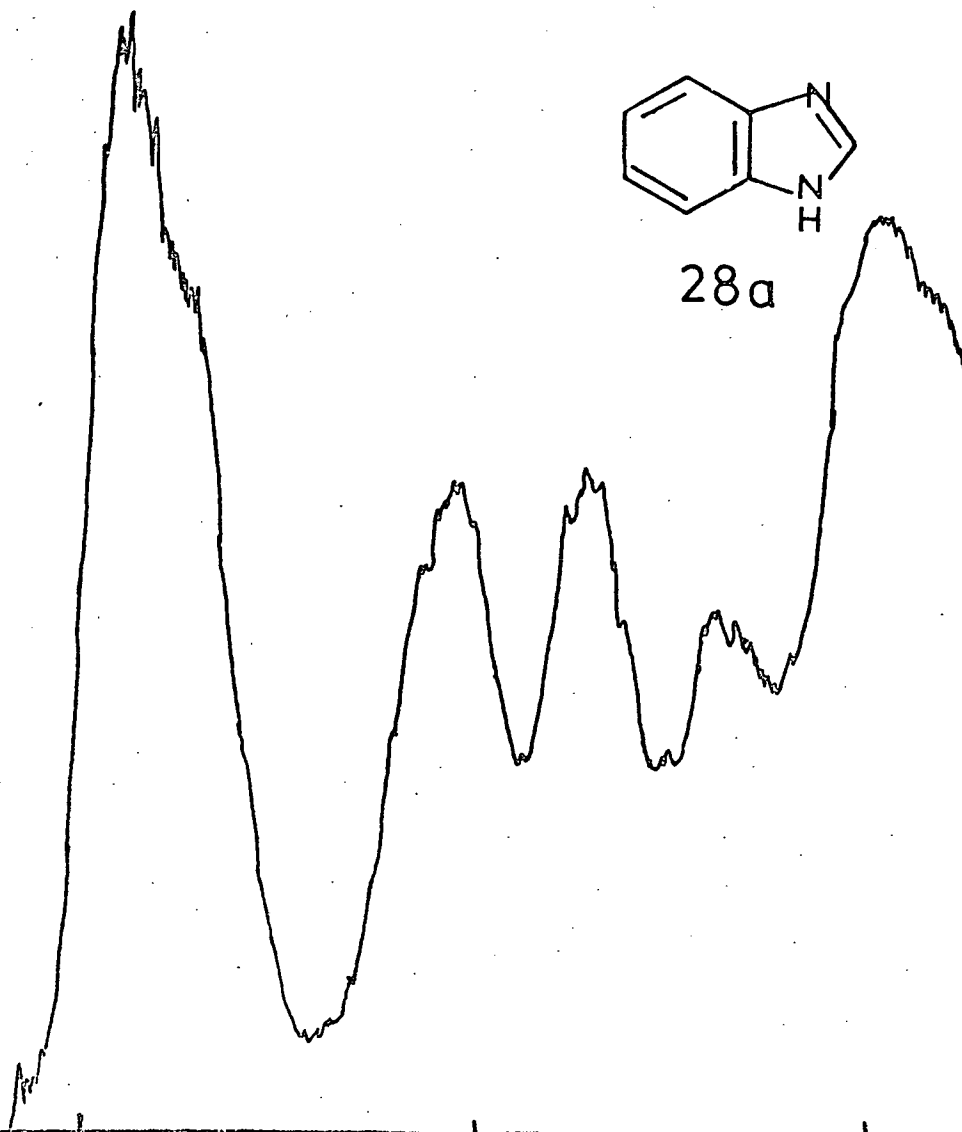
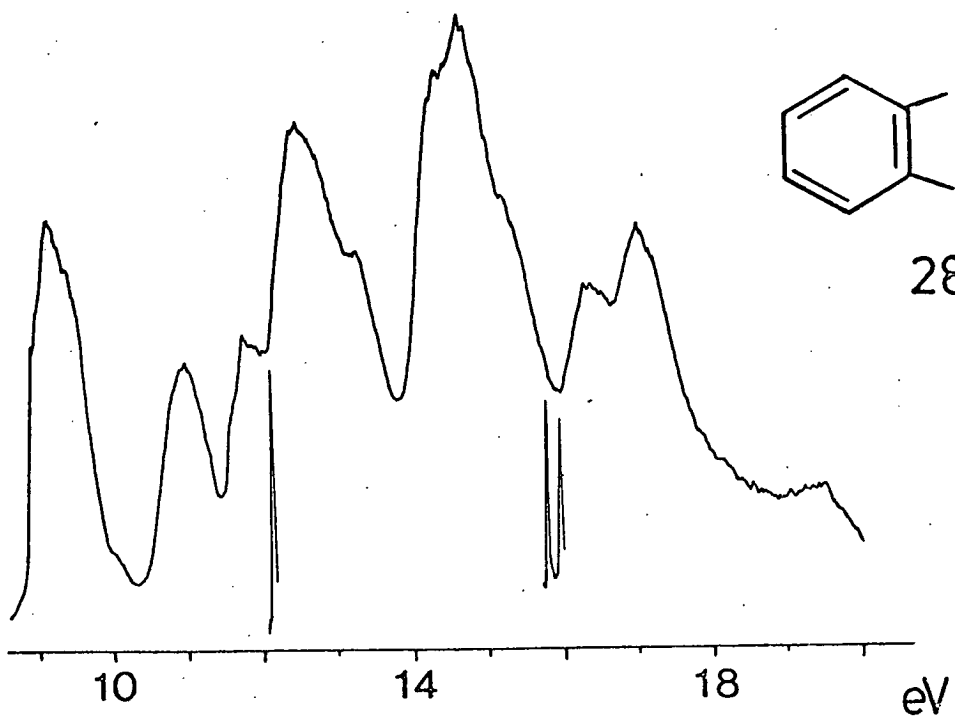
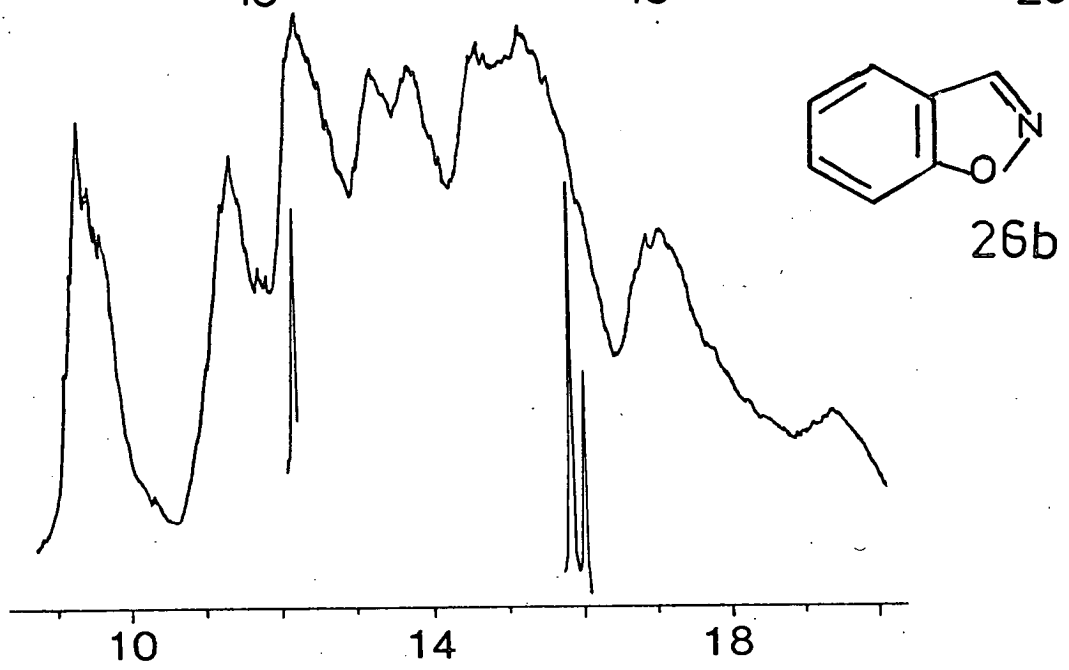
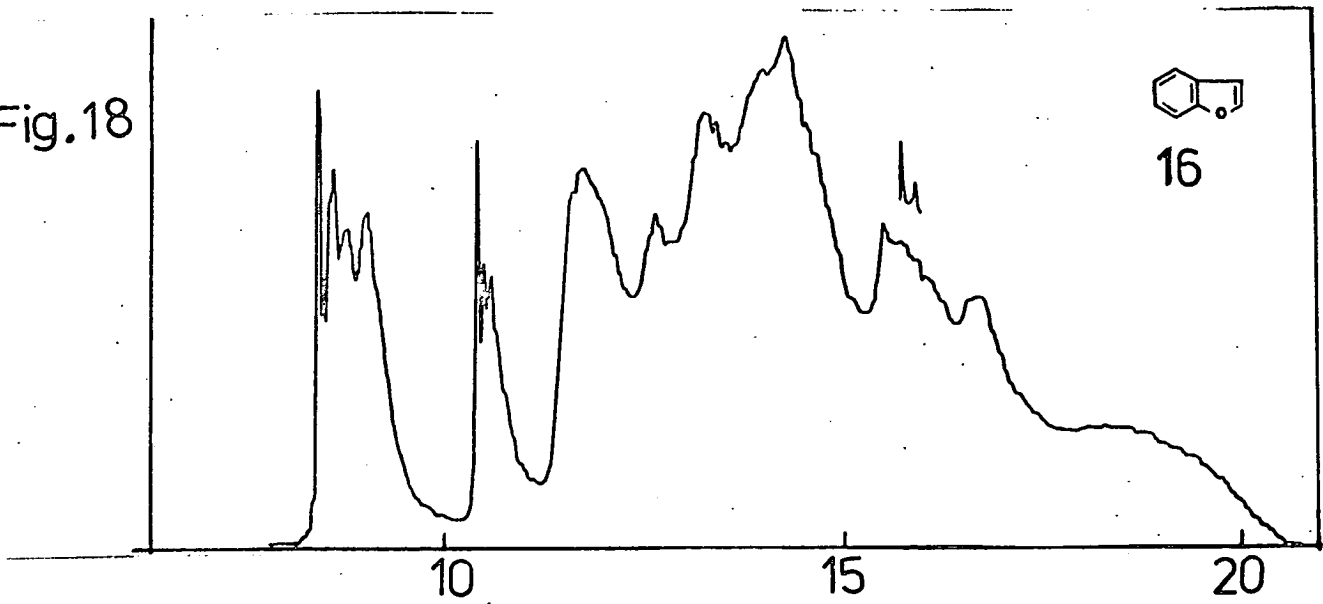


Fig.18



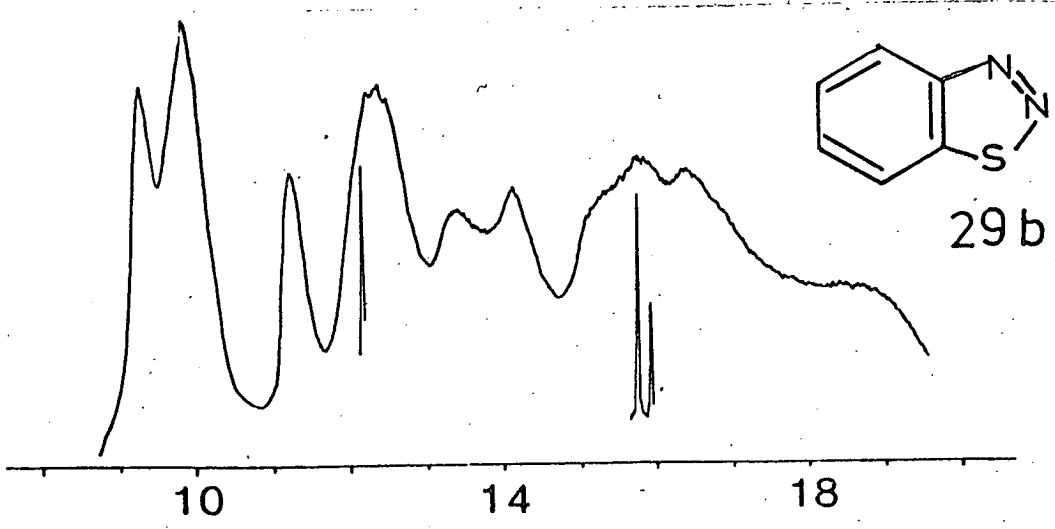
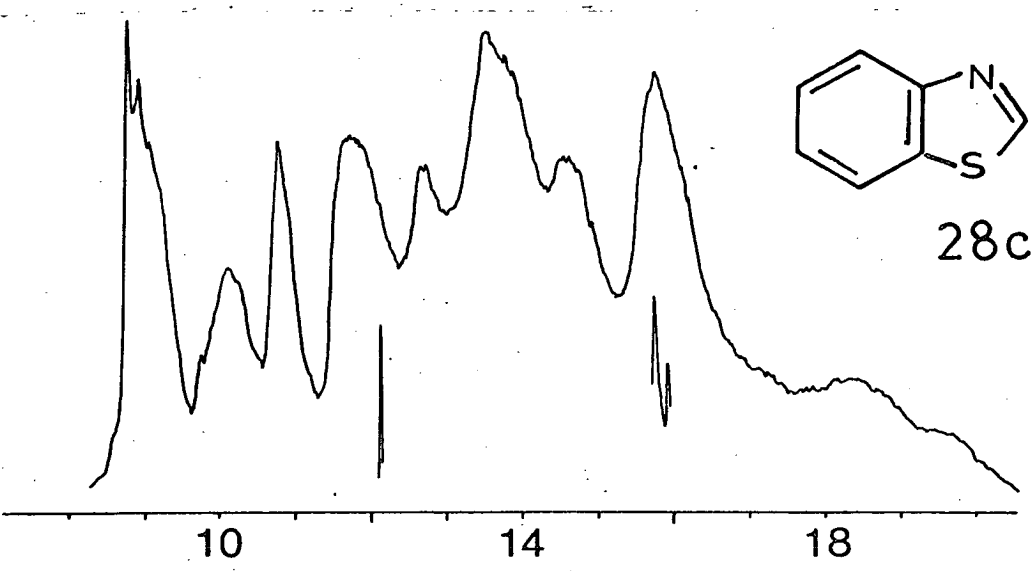
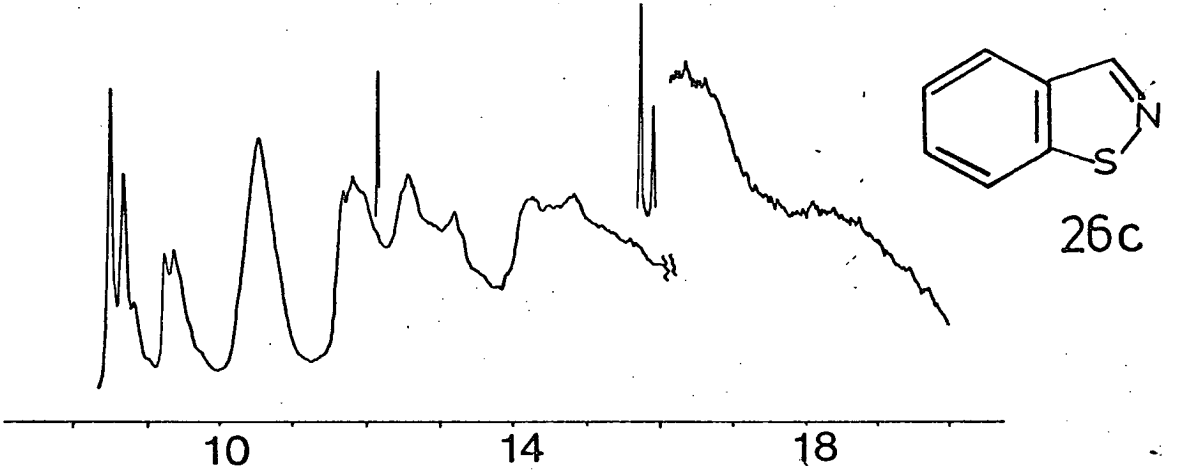
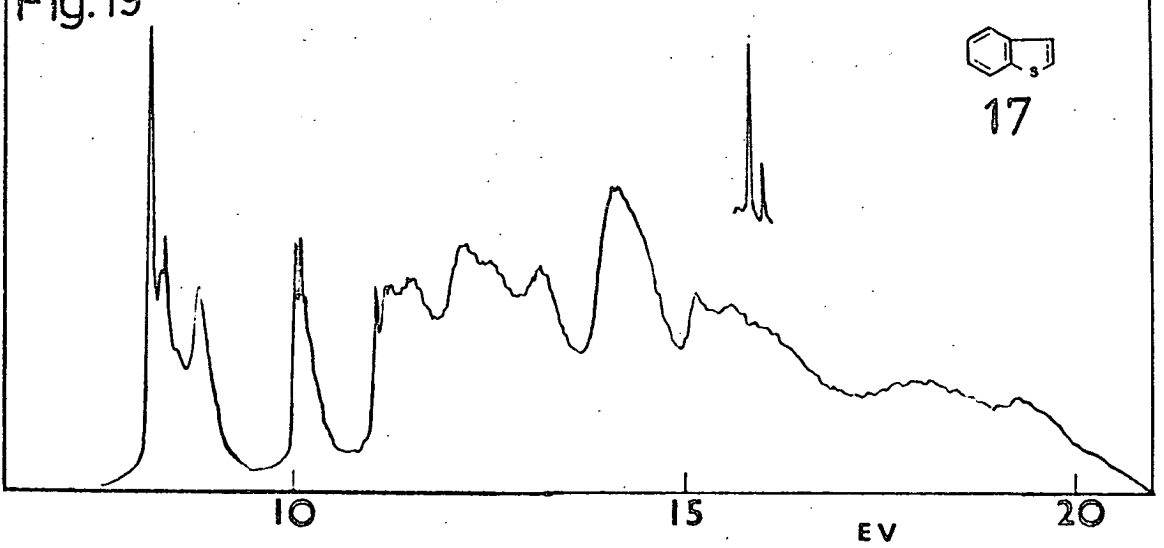
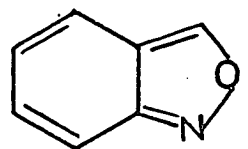
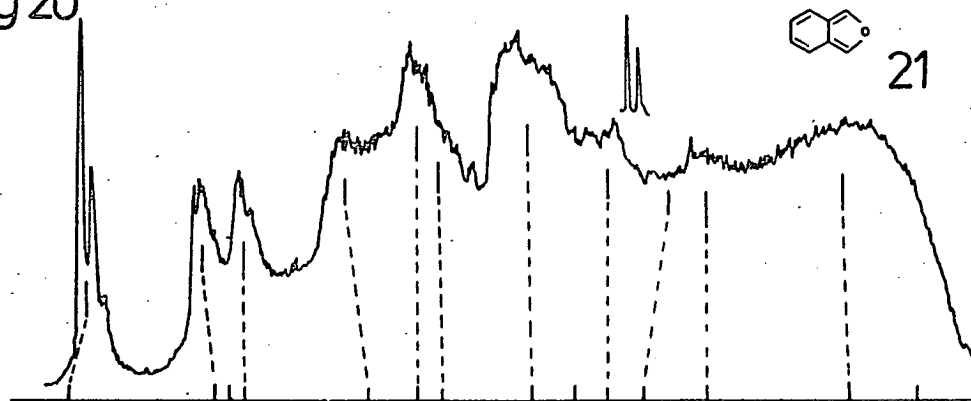
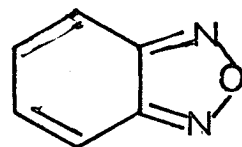
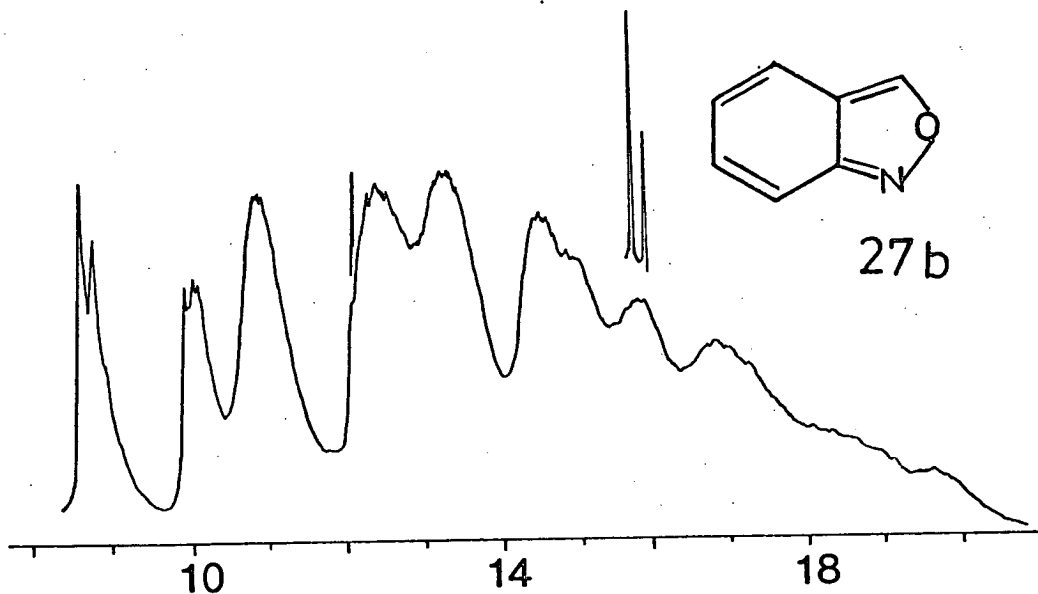


Fig 20



27b



30b

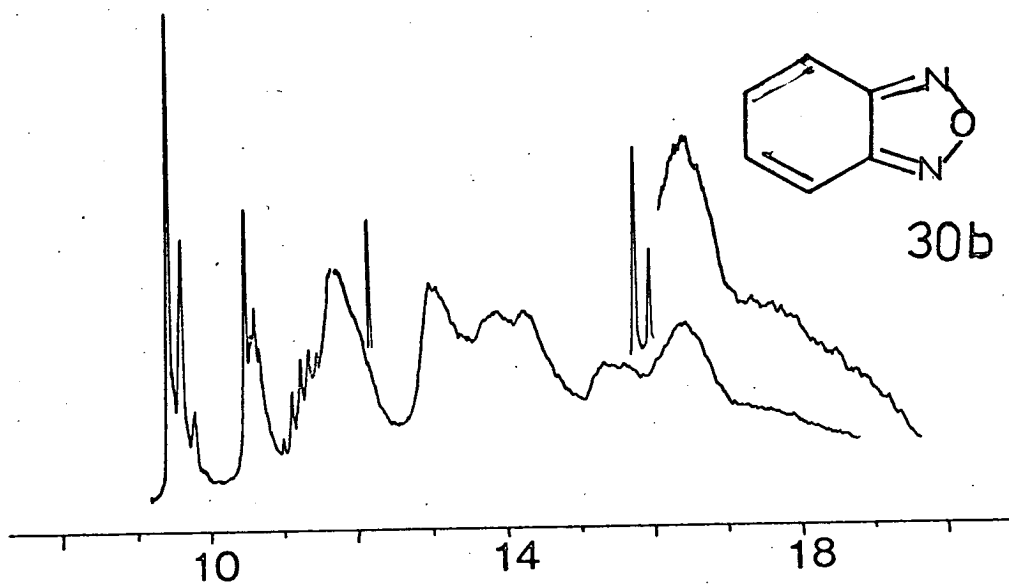
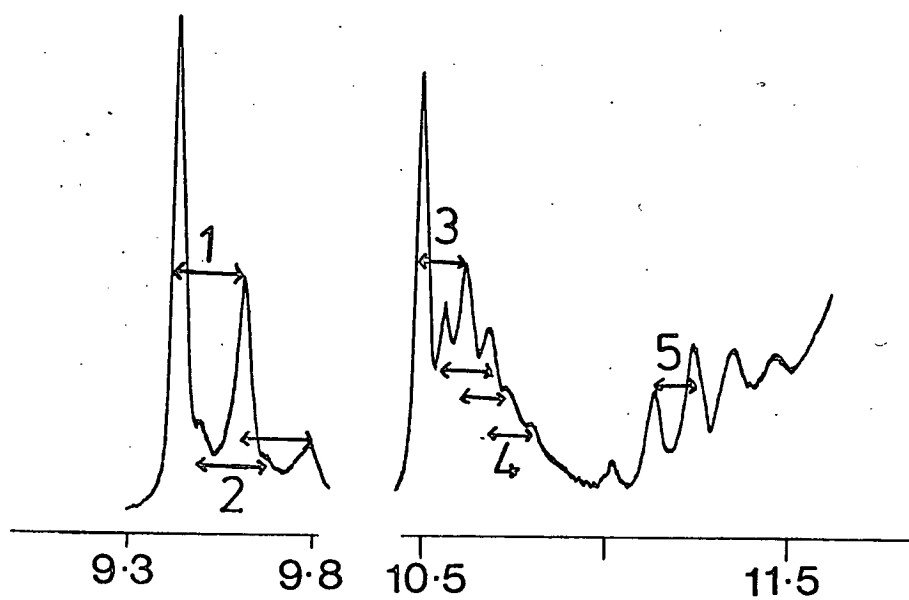
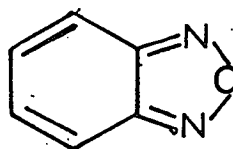


Fig. 21



- 1 = 1450      cm<sup>-1</sup>
- 2 = 1450
- 3 = 968
- 4 = 968
- 5 = 887



Fig. 22

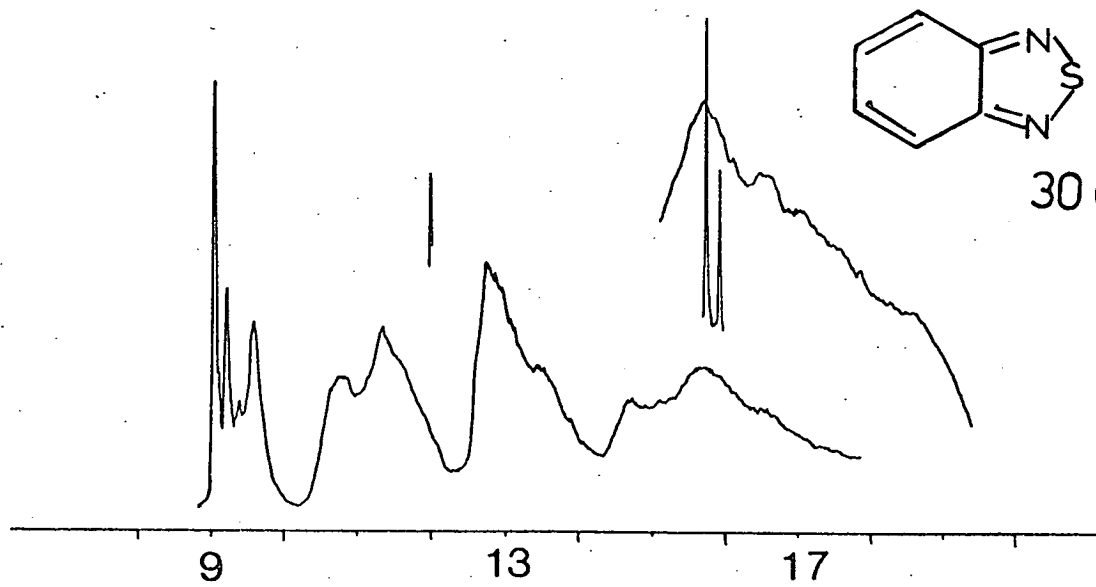
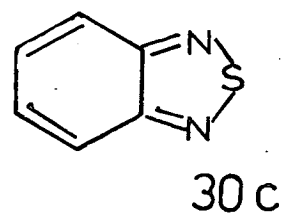
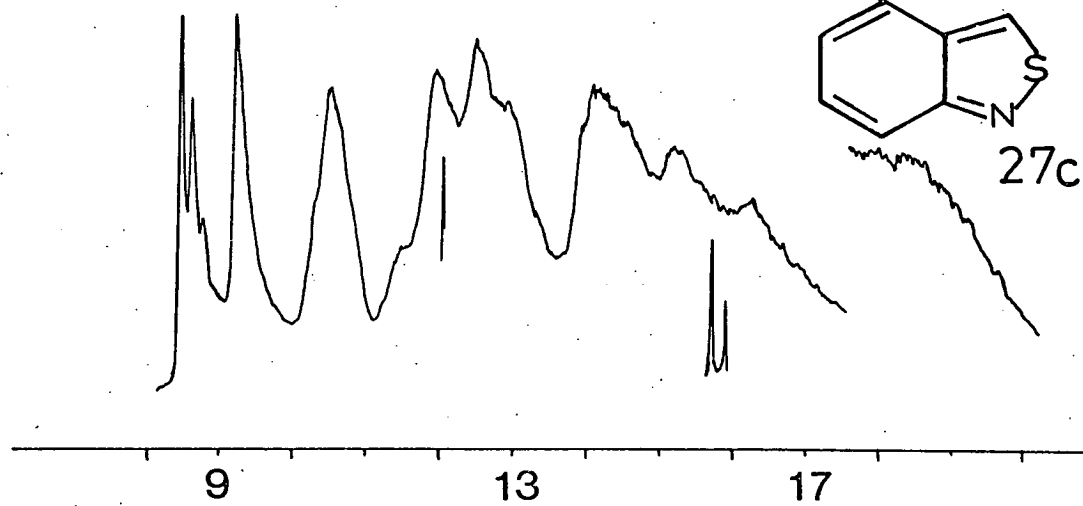
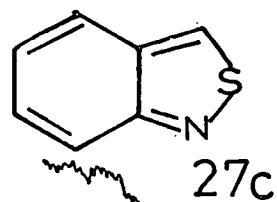
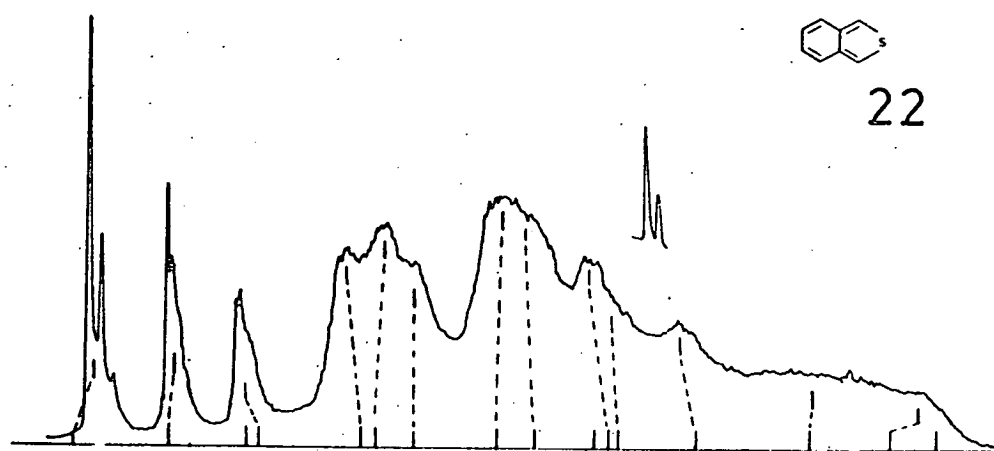
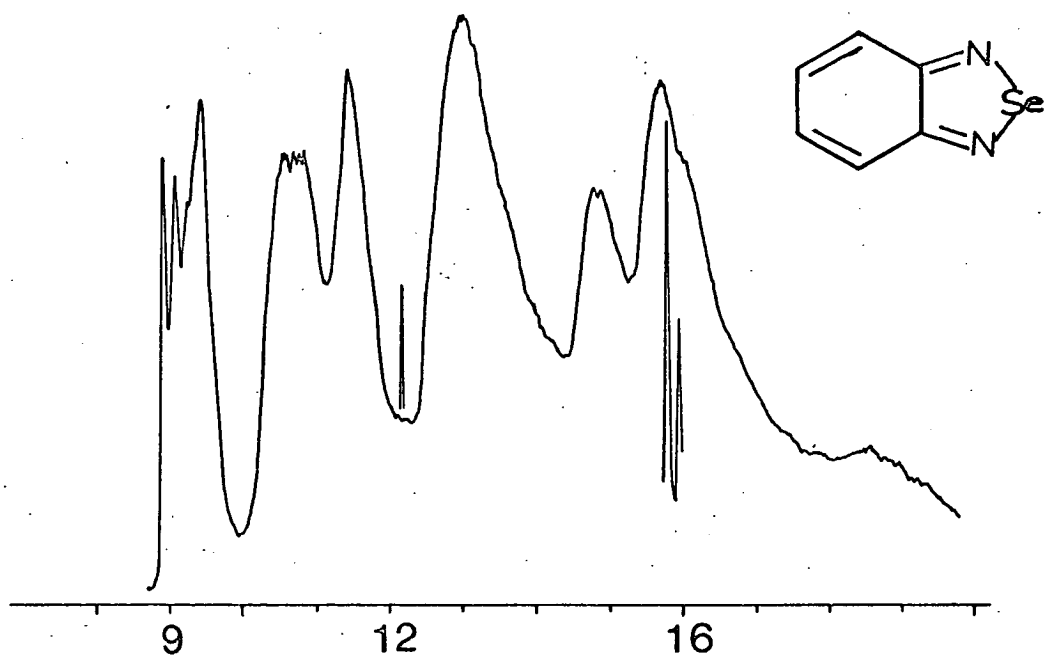
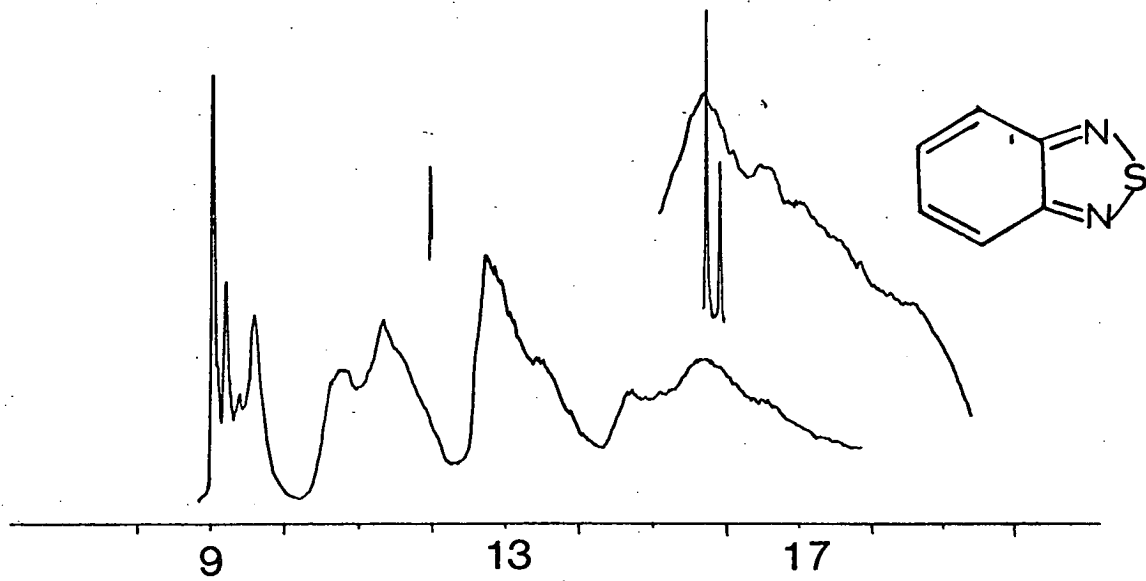
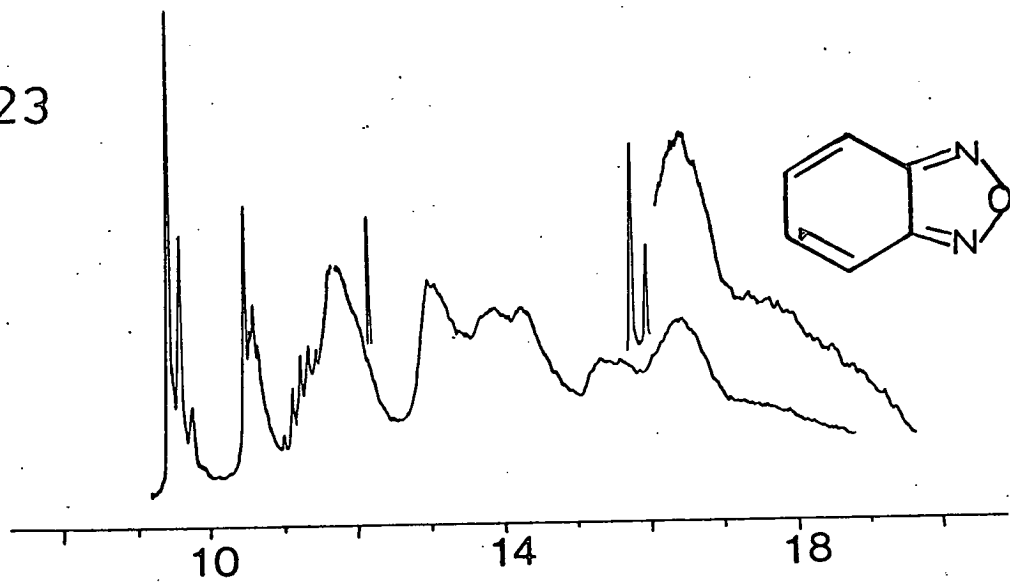


Fig. 23



be further apart than in indole, and the intensity, of the third IP suggests a degenerate pair. This is suggested by calculation to be a  $\pi$ -orbital and a heavily localised nitrogen  $p_y$  molecular orbital. The loss of CH has simplified the spectrum in the  $\sigma$ -region. The spectrum of 28a, benzimidazole, shows only the low energy region with the additional peak, shown by calculation to be associated with a nitrogen lone pair at 10.18 eV. The greater separation of the nitrogen atoms has led to less perturbation of the basic indole spectrum than in indazole where the nitrogen atoms are adjacent.

For the series benzofuran, 16, 1, 2-benzisoxazole, 26b, and benzoxazole 28b, the spectra are shown in Fig.18. The introduction of the nitrogen atom in 26b and 28b has made the first two IPs almost degenerate, and introduced an additional peak in the low energy region. The calculations suggest that the orbital ordering of the first four levels is  $\pi, \pi, \sigma, \pi$ , for 26b and 28b, giving the peaks at 11.4 and 10.95 eV in 26b and 28b respectively as the nitrogen lone pair.

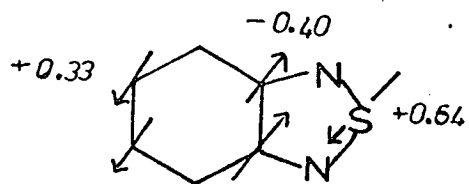
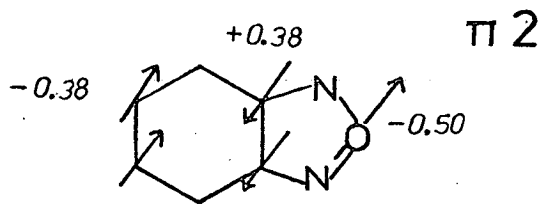
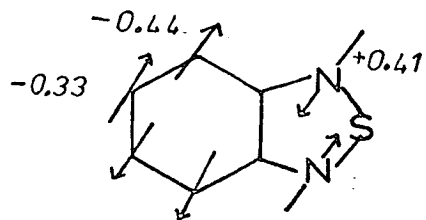
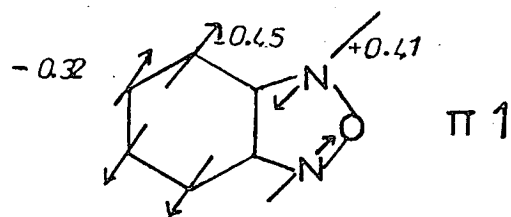
The spectra of 17, 26c, 28c and 29b are shown in Fig.19. The effect on the low IP region of having a nitrogen adjacent to sulphur is to separate the first two IPs, the largest separation being shown by 26c, of 0.77 eV compared with 0.48 eV in 17. The calculations suggest  $IP_3$  in 26c and  $IP_4$ , 28c to be a  $\sigma$ -orbital corresponding to the nitrogen lone pair. In 29c two additional peaks are expected in the early region compared with benzothiophen, and are shown by calculation to correspond to  $IP_3$  and  $IP_7$  as molecular orbitals localised on nitrogen  $p_z$  and  $p_y$ . In the spectrum  $IP_3$  is degenerate with  $IP_2$  ( $\pi$ ) at

9.87 eV, and  $IP_7$  comes at higher binding energy in the 12 eV region at the onset of the  $\sigma$ -bands. The spectrum of 26c and 29b are similar, the additional nitrogen showing up as an increase in intensity at  $\approx 12.3$  eV.

In the quinoid series, molecules related to benzo(c)furan, 21, are anthranil, 27b, and 2, 1, 3 - benzoxadiazole, 30b. Introduction of the nitrogen atoms appears to have little effect on the overall PE spectrum of 21, as shown in Fig.20. The calculations show  $IP_4$  to be a nitrogen lone pair level in 27b, and a symmetric combination of  $Np_y$  in 30b, in contrast with the Kekulé series where the nitrogen lone pair appeared as  $IP_3$ . The calculated difference in orbital energy between  $IP_3$  and 4 for 27b and 30b is 0.38 eV and 0.14 eV. The difference is small, and it could be argued that the calculation might be giving a reverse ordering. However, for 30b, at least,  $IP_3$  shows an interesting vibrational progression which would normally be expected with ionisation from a  $\pi$ -orbital, and not with a  $\sigma$ -lone pair type orbital suggesting that the calculated order is the correct one. It is noticeable in the spectrum of 30b that the first three IPs have been moved to higher binding energy compared with 21 or 27b, the first IP being most affected by a shift of -1.64 eV. This compound also shows remarkably good resolution, and vibrational progressions on the early IPs. These are expanded, and shown in Fig.21, along with the frequencies associated with each progression.

There are two compounds related to benzo(c)thiophen, 22, by introduction of one and two nitrogen atoms. These are

Fig. 24

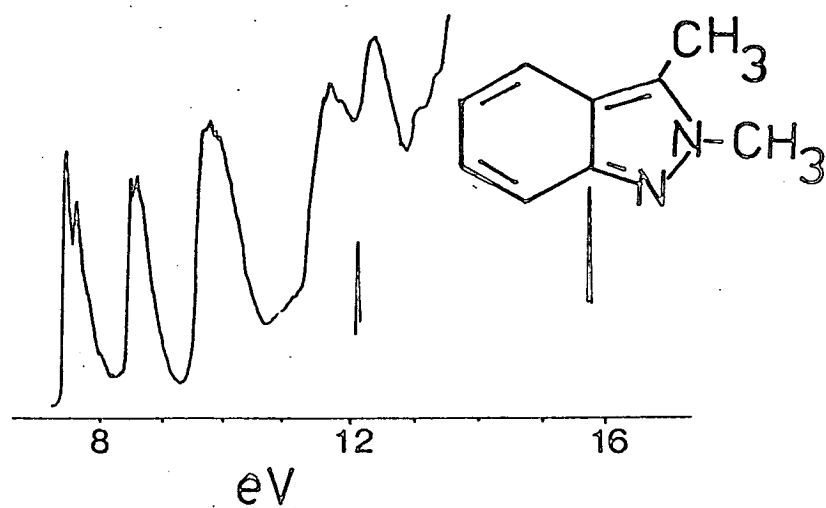
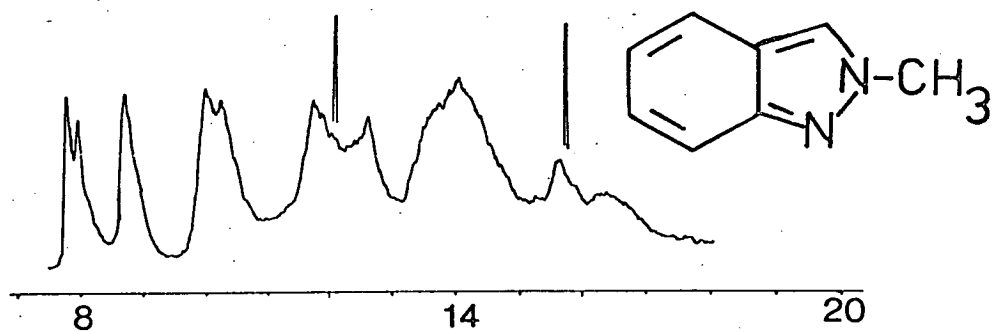
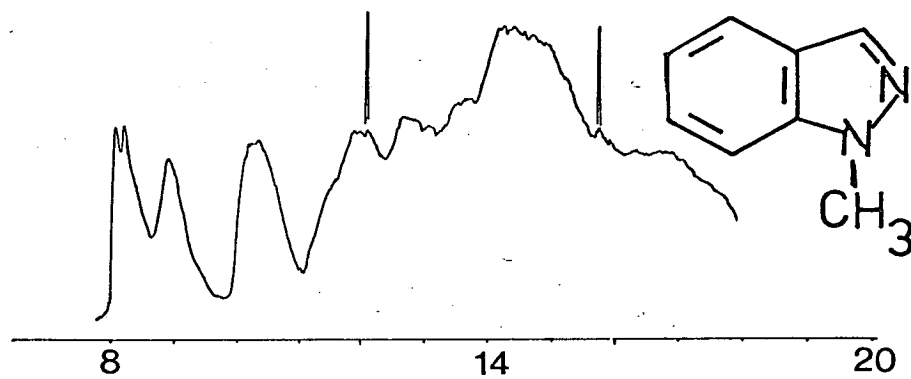
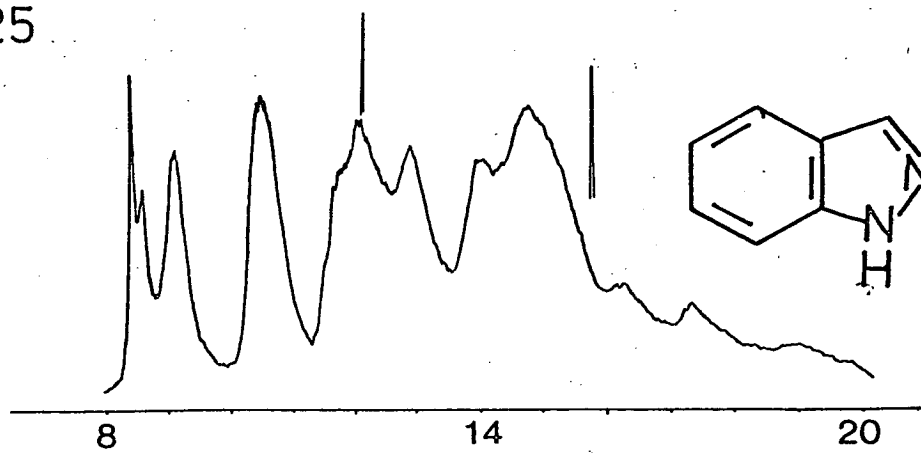


27c and 30c, and their spectra are shown in Fig.22. In 27c, the nitrogen lone pair level is degenerate with  $IP_3(\pi)$  at 10.6 eV, and in 30c the calculation indicates  $IP_3$  observed at 10.8 eV to be associated with the nitrogen lone pairs as a symmetric combination ( $N_{p_y} + N_{p_y}$ ). In all cases where an orbital is said to be associated with a nitrogen lone pair, the eigenvector for the appropriate nitrogen basis function is of the order 0.6 - 0.7 showing a large degree of localisation within the molecular orbital.

If the PE spectra of series 30 are inspected, where X is one of the group VI elements, O, S, or Se, the second IP is seen to move to lower binding energy, the first IP remaining in an almost constant position, Fig.23. For X = O, and S the calculations show the orbitals corresponding to  $IP_2$  to be of the form shown in Fig.24. The eigenvectors indicate that for X = S, the molecular orbital is more localised on the sulphur, thus moving the orbital to lower BE. No calculation exists for X = Se, but a similar effect may well occur. The forms of the lowest binding orbital show that the O and S atoms are effectively nodal, Fig.24, and the orbitals become similar in nature explaining the consistency of the first IP in each case.

The PE spectra of a series of methyl substituted indazoles are shown in Fig.25. The addition of a methyl group to the 1-position in indazole has little effect on the indazole spectrum other than reducing the resolution and increasing the intensity in the  $\sigma$ -region due to the

Fig.25



additional carbon atom. All IPs are moved marginally to lower binding energy in the substituted compound. In 2-methyl indazole, a Kekulé benzenoid structure can no longer be drawn, and must be replaced by a quinoid structure, as is also the case for 2,3-dimethyl indazole. The most pronounced effect on the spectra of the quinoid compounds, when compared with that of indazole, is a general shift in IPs to lower binding energy by about 0.5 eV in 2-methyl indazole and by an even larger shift in the same direction in 2,3-dimethyl indazole. This general molecular destabilisation with loss of a benzenoid ring is in accord with that shown by compounds such as indole and isoindole. The larger shift in 2,3-dimethyl indazole can probably be attributed to the introduction of the second methyl group. Methyl groups are electron donating, and will therefore destabilise the already electronegative nitrogen atoms, giving a shift in molecular orbitals to lower binding energies.

The core and inner valency shell orbitals can be examined experimentally by XPs but for the inner valency shell orbitals in particular, the peaks are broad and multiple assignment is necessary. The XPS spectra for benzofuran and benzothiophen are shown in Figs. 26 and 27. For benzofuran groupings in the inner valency shell are reasonably reproduced by the calculation. For benzothiophen only the early part of the spectrum is available, but the correlations of peaks in the 15 - 25eV region is good, and confirms the assignments made using UPS data, Fig. 10.



Fig.26

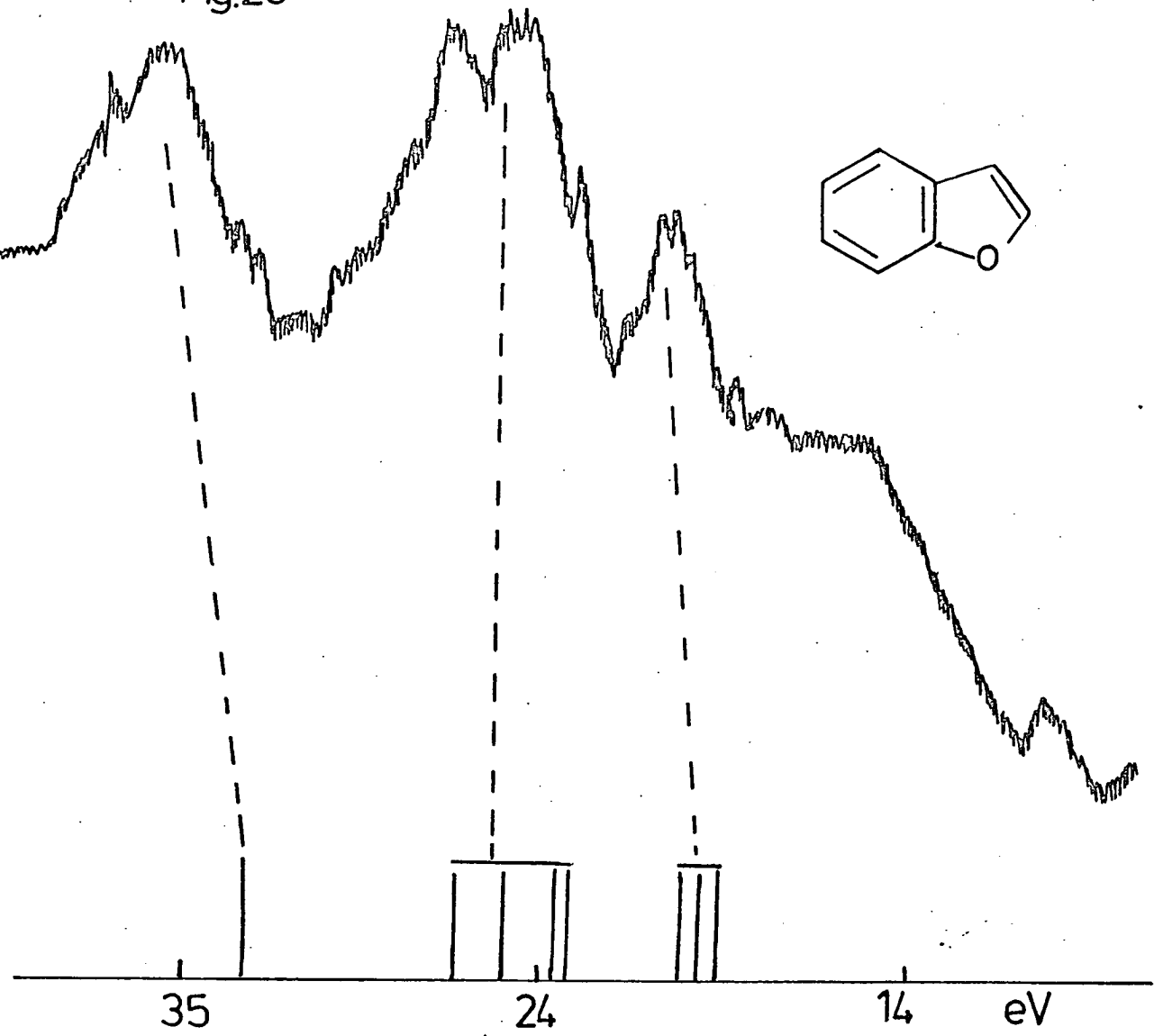
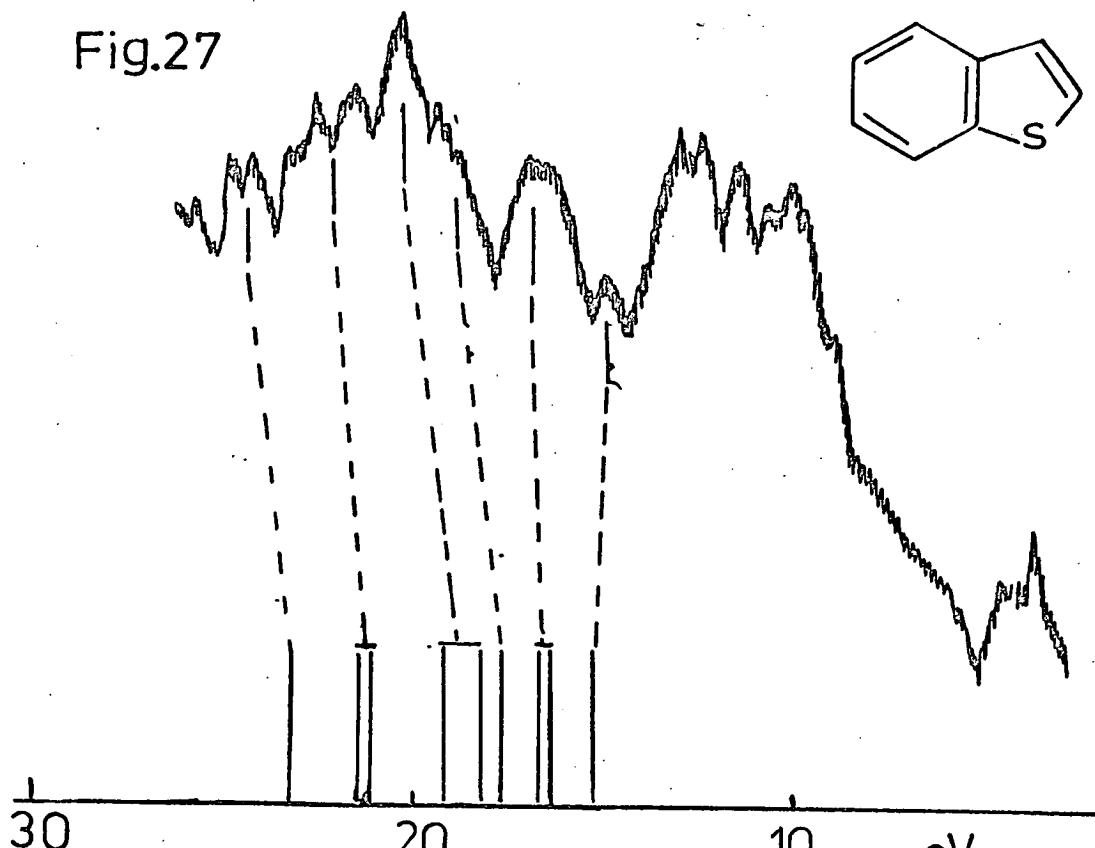


Fig.27



R E F E R E N C E S

1. A.R. Katritzky, 'Physical Methods in Heterocyclic Chemistry', 1974, Vol. VI, p9, Academic Press.
2. I.N. Levine, 'Molecular Spectroscopy', p 316, Wiley-Interscience, New York
3. T. Koopmans, Physica, 1933, 1, 104.
4. W. Meyer, Int. J. Quantum Chem. 1971, 55, 341
5. J.H.D. Eland, Photoelectron Spectroscopy, p90, London, Butterworths.
- 6(a) B.O. Jonsson and E. Lindholm, Ark. Fysik, 1969, 39, 65.  
(b) L. Prand, P. Millie and G. Berthier, Theor. Chim. Acta, 1968, 11, 169.
7. L. Åsbrink, E. Lindholm and O. Edqvist, Chem. Phys. Lett., 1970, 5, 609.
8. A.W. Potts, W.C. Price, D.G. Streets and T.A. Williams, Discuss. Faraday Soc., 1973, 54, 168.
9. W.C. Price, A.W. Potts and T.A. Williams, Chem. Phys. Lett., 1976, 37(1), 17
10. W. Moyes, Chemistry Department, University of Edinburgh, unpublished results.
11. R.A.W. Johnstone and F.A. Mellon, J. Chem. Soc. Faraday II, 1973, 69, 1155.
12. M.J.S. Dewar and D.W. Goodman, J. Chem. Soc. Faraday II, 1972, 68, 1784.
- 13(a) F. Brogli, E. Heilbronner and T. Kofayashi, Helv. Chim. Acta., 1972, 55, 274.  
(b) F. Brogli and E. Heilbronner, Angew. Chem. Int. Edn., 1972, 11, 538.  
(c) P.A. Clark, F. Brogli and E. Heilbronner, Helv. Chem. Acta., 1972, 55, 1415

- 14(a) J.H.D. Eland and C.J. Danby, *Z. Naturforsch.*, 1968, 23a, 355.
- (b) J.H.D. Eland, *Int. J. Mass Spectrom. Ion. Phys.*, 1969, 2, 471
- (c) *ibid*, 1972, 9, 214.
15. E. Lindholm, C. Fridh and L. Asbrink, *Disc. Faraday Soc.*, 1972, 54, 127.
16. F. Marschner, *Tetrahedron*, 1974, 30, 3159
17. R. Boschi, J.N. Murrell and W. Schmidt, *Disc. Faraday Soc.*, 1972, 54, 116.
18. D.G. Streets and T.A. Williams, *J. Electron Spectrosc.*, 1974, 3, 71.
19. R. Boschi and W. Schmidt, *Tetrahedron Letters*, 1972, 2577.
20. J.P. Maier and D.W. Turner, *Faraday Disc. Chem. Soc.* 1972, 54, 149
21. M.J.S. Dewar, E. Haselbach and S.D. Worley, *Proc. Roy. Soc. A*, 1970, 315, 431.
22. J.H.D. Eland and C.J. Danby, *Z. Naturforsch.*, 1968, 23a, 355.
23. J. Almlöf, *Chemical Physics*, 1974, 6, 135.
24. G.R. Branton, D.C. Frost, F.G. Herring, C.A. McDowell and I.A. Stenhouse, *Chem. Phys. Letters*, 1969, 3(8), 581
25. M.H. Palmer, R.H. Findlay and A.J. Gaskell, *J. Chem. Soc. Perkin Trans II*, 1974, 420.
26. S. Craddock, R.H. Findlay, and M.H. Palmer, *Tetrahedron*, 1973, 29, 2173.
27. C.R. Brundle and M.B. Robin, *J. Amer. Chem. Soc.* 1970, 92, 5550.
28. P.J. Derrick, L. Asbrink, O. Edqvist, B.-O. Jonsson and E. Lindholm, *Int. J. Mass Spectrom. Ion. Physics*, 1971, 6, 161.

29. D.P. Craig, J.M. Hollas, M.F. Redies and S.C. Wait, Phil. Trans. Roy. Soc. A, 1961, 253, 543, 569.
30. E.P. Krainov, Optical Spectroscopy, 1964, 16, 415.
- 31(a) G. Hagen and S.J. Cyvin, J. Phys. Chem., 1968, 72, 1446.  
(b) S.J. Cyvin, B.N. Cyvin and G. Hagen, Chem. Phys. Letters 1968, 2, 341.
32. P.J. Derrick, L. Åsbrink, O. Edqvist, B.O. Jonsson and E. Lindholm, Int. J. Mass Spectrom. Ion. Physics, 1971, 6, 203.
33. J.H.D. Eland, Int. J. Mass Spectrom. Ion. Physics, 1969, 2, 471
34. P.A. Clark, R. Gleiter and E. Heilbronner, Tetrahedron, 1973, 29, 3085.
35. Tsunetoshi Kobayashi, Katsuyuki Yokota and Saburo Nagakura, J. of Electron Spectroscopy and Related Phenomena, 1973, 3, 449.
36. C.R. Brundle, M.B. Robin, N.A. Kuebler and H. Basch, J. Amer. Chem. Soc., 1972, 94, 1466.
37. J.M. Hollas, J. Mol. Spectroscopy, 1962, 9, 138; Spectrochimica Acta, 1963, 19, 753.
38. P.J. Derrick, L. Åsbrink, O. Edqvist, B.O. Jonsson and E. Lindholm, Int. J. Mass Spectrom, Ion. Physics, 1971, 6, 177
39. Ibid. 1971, 6, 191.
40. P.A. Clark, R. Gleiter, and E. Heilbronner, Tetrahedron, 1973, 29, 3087.
41. M.H. Palmer, and S.M.F. Kennedy, J. Chem. Soc., Perkin II, 1976, 81.

42. M.J.S. Dewar, A.J. Harget, N. Trinajstic and S.D. Worley, Tetrahedron, 1970, 26, 4505.
43. J. Koller, A. Azman and N. Trinajstic, Z. Naturforsch, 1974, 29a, 624.

Dipole Moments

A molecule has a dipole moment when its centre of positive (+ $\delta e$ ) and negative (- $\delta e$ ) charge do not coincide. If these are separated by a distance  $r$ , then the dipole moment  $\mu$  is given by

$$\mu = (\delta e)r$$

and is a vector quantity.

The accuracy of a calculated dipole moment will depend on the accuracy with which the molecular geometry is known. Also, as calculated dipoles refer to individual molecules, it is only strictly correct to compare these values with experimental values determined in the gas phase. This is, of course, not always possible, the relevant data not always being available.

The calculated dipoles are separated into  $\pi$  and  $\sigma$  contributions, and in all cases the dipole direction is indicated by the angle of the dipole moment vector from the origin of the molecular axes, at the centre of the C(3a) - C(7a) bond for the bicyclic molecules and within the ring for the monocyclic molecules as shown in Fig. 1. The angles are quoted as positive anticlockwise or negative clockwise with respect to the long Z - axis, the negative end of the dipole being away from the origin as shown in Fig. 1.

The separation into  $\sigma$  and  $\pi$  components is somewhat arbitrary due to the difficulty of deciding on the nuclear contribution to each component. The separation

Fig.1

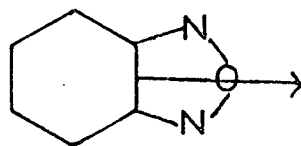
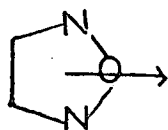
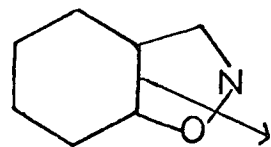
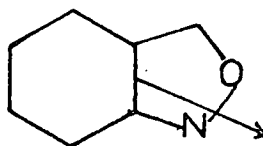
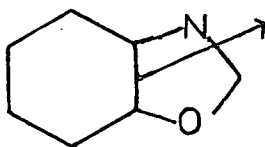
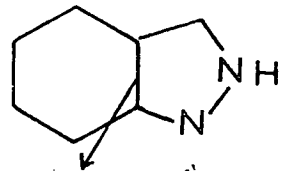
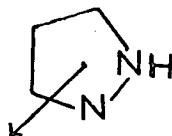
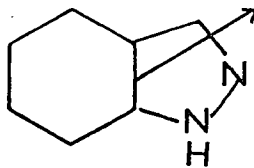
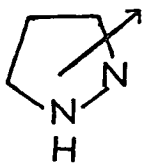
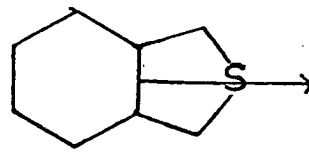
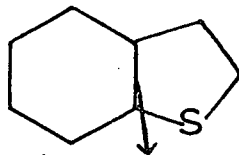
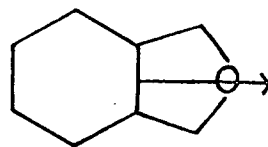
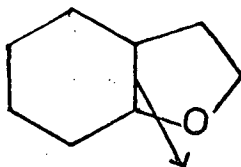
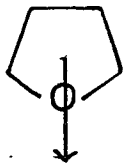
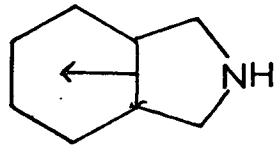
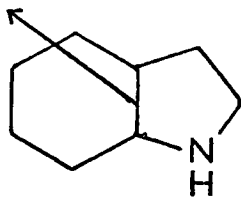
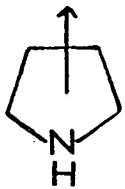
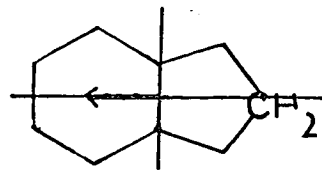
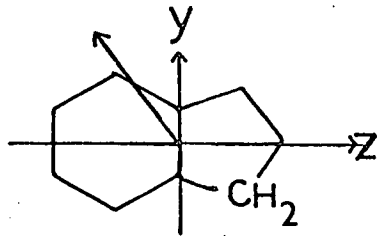
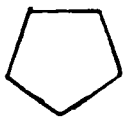


Fig 1 (cont'd.)

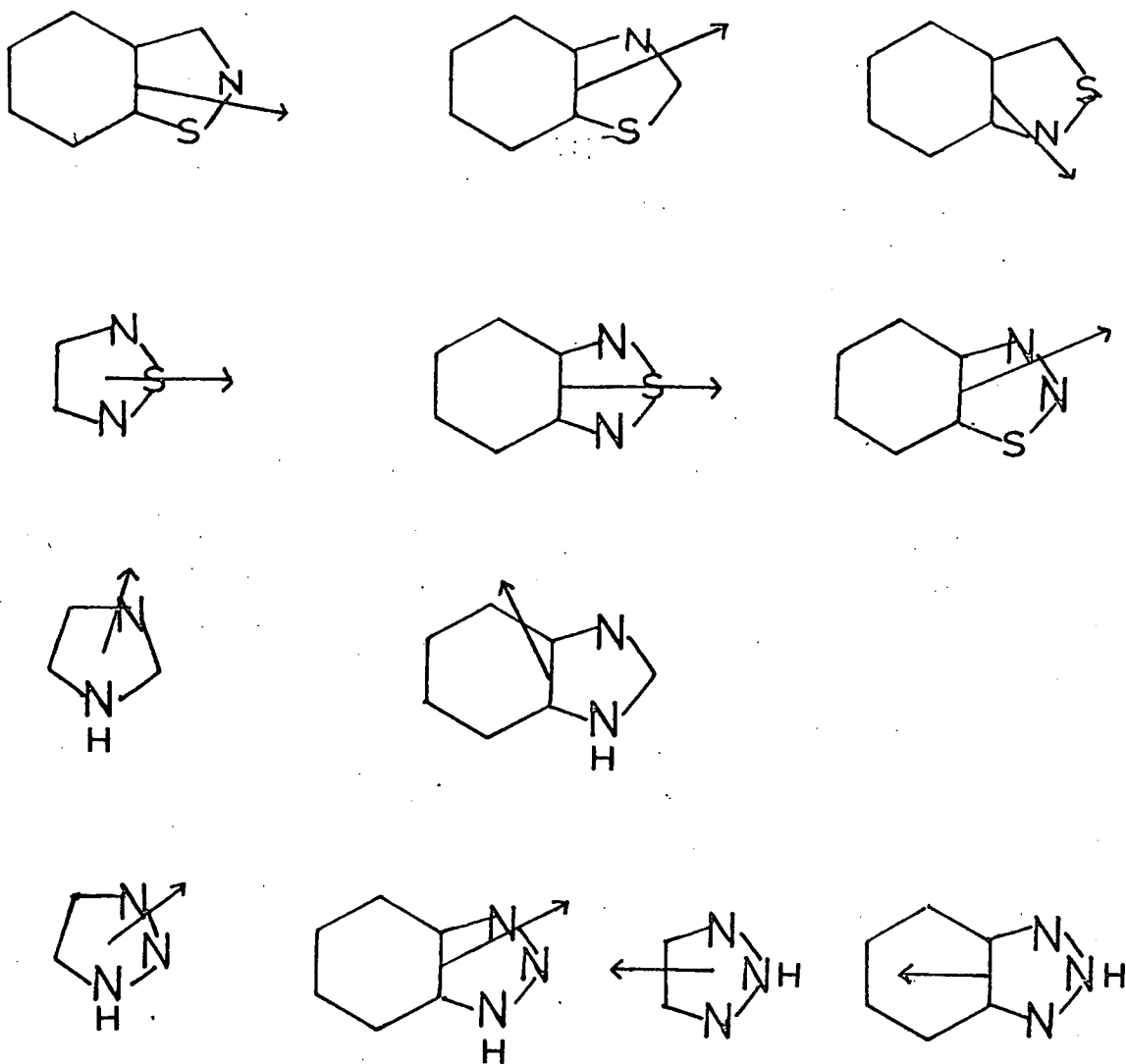
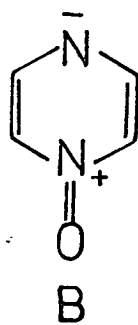
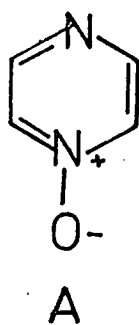


Fig.2



C	1 $\pi$	5 $\sigma$
N <sup>+</sup>	1 $\pi$	5 $\sigma$
N	1 $\pi$	6 $\sigma$
O <sup>-</sup>	2 $\pi$	7 $\sigma$

C	1 $\pi$	5 $\sigma$
N	1 $\pi$	5 $\sigma$
N	2 $\pi$	6 $\sigma$
O	1 $\pi$	7 $\sigma$



used here, is based on the method suggested by Berthier<sup>1</sup>, where the nuclear contribution to each system is defined in terms of the classical structure for the molecule. Considering pyrrole as an example, the classical electronic contribution to the  $\pi$  system from the N and C atoms is 2 and 1 electrons respectively. This is then assumed to give a nuclear contribution of +2 and +1 for N and C in the  $\pi$  system. Both N and C contribute five electrons to the  $\sigma$  system producing +5 as their nuclear contribution to the  $\pi$  system. There are two weaknesses in the theory, first, that it refers to a hypothetical structure and secondly the uncharged molecule could still have homopolar  $\sigma$  and  $\pi$  dipole moments of opposing sign. The weakness is exemplified by the separation of the dipole moment of pyrazine N-oxide which can be represented by the two canonical forms shown in Fig.2. The nuclear contribution to the  $\sigma$  and  $\pi$  system is given for each of the forms and can be seen to differ. These differences will produce different values of the  $\sigma$  and  $\pi$  dipole moments for A and B. This problem arises when the molecule can be represented by more than one classical structure but for the molecules under consideration here, this problem does not arise, and the separation procedure will yield only one value for each component.

The Bicyclic Heterocycles I      Experimental dipole moments exist for all the molecules in this series and references to their measurement have been obtained from reference 2, with the exception of benzofuran<sup>3</sup>.

Table 1 Dipole moments ( $\mu/D$ ) and directions

	Calculated		Observed		Calculated			
	$\mu$	$\theta$	$\mu$	$\theta$	$\mu_\sigma$	$\theta$	$\mu_\pi$	$\theta$
Styrene <sup>†</sup>	0.027	51	0.13 <sup>5</sup>	-				
			0.2 <sup>9</sup>	-				
			0.181 <sup>7b</sup>	75.8				
			0.43 <sup>6</sup>	-				
Indene	0.915	139	0.85 <sup>10</sup>	-				
Indole	2.31	138 (48)*	2.08 <sup>11</sup>	50	0.44	-44	2.75	129
Benzofuran	1.56	-82	0.79 <sup>3</sup>	-	3.00	-74	1.46	115
Benzothiophen (spd + 3s')	0.56	-88	0.62 <sup>12</sup>	-	2.48	-23	1.99	150

<sup>†</sup> Styrene is positive anticlockwise with respect to the C(1) - C(2) bond.

\* Angle with respect to the reference system used in reference 5.

Table 2 Dipole moments ( $\mu/D$ )<sup>†</sup> of the monocyclic compounds<sup>4</sup>

	$\mu_{\text{calc.}}$	$\theta_{\text{calc.}}$	$\mu_{\text{expt.}}$
pyrrole	2.01	90	1.80
furan	0.64	-90	0.67
thiophen (spd + 3s')	0.44	-90	0.53
pyrazole	2.85	52	2.21
imidazole	4.41	77	3.80
1H-1,2,3-triazole	4.50	59	-
2H-1,2,3-triazole	3.24	180	-
1,2,5-oxadiazole	2.96	0	3.38
1,2,5-thiadiazole	1.75	0	1.57

<sup>†</sup> "best atom" basis set used

The calculated dipole moment directions are considered to be correct as those of some five and six membered ring heterocycles<sup>4</sup> calculated using a similar "best atom" basis set to that used in the present work, gave good agreement between calculated and experimental results. Table 1 gives a summary of the calculated and experimental dipole moments along with the  $\sigma$  and  $\pi$  components for the Kekulé molecules. Good agreement is shown except for benzofuran. If the scaled basis set used here for benzofuran is used for furan, then a dipole moment of -1.01D is obtained, the minus sign indicating that the negative end of the dipole lies towards the hetero atom. This is high compared to the experimental value of 0.67D. The best atom basis set used in previous work<sup>4</sup> yielded a value of -0.64D, indicating that optimisation of the basis set on a total energy basis has not produced optimisation in this particular molecular property. It is possible that in the structure of furan the lone pair positions are critical to the dipole moment calculation, and in scaling the oxygen atom basis set it is possible that these lone pairs will have moved further from the atomic nucleus in an attempt to minimise the total energy. This increased charge separation will cause an increase in the dipole moment, and could account for the discrepancy between the experimental and calculated dipole moments for both benzofuran and furan. Tables 2, 3 and 4 show the calculated and experimental

dipole moments of the monocyclic molecules<sup>4</sup> and the bicyclic molecules. It can be seen from Fig.1 that the monocycles and the corresponding bicycles are polarised in a similar fashion in all cases. The majority of the dipole moments of the monocyclic molecules and of the bicyclic molecules cannot be compared directly due to the difference in basis set used for each series, but taking this into account it is obvious that the magnitudes are similar in most cases. The exceptions are those molecules containing three heteroatoms, all of which show a large increase in polarity towards the heterocyclic ring in the bicyclic compounds.

Only indole of the heterocycles has an experimental dipole moment angle estimated from substituted indoles<sup>5</sup>. The experimental works gives the dipole moments angles as positive with respect to the y-axis, and when the present angle is converted to this reference the calculated result is  $48^{\circ}$  compared with the experimental value of  $50^{\circ}$ . The estimated result is obtained by measuring the dipole moments for the substituted compound and the substituent, and adding the group moments vectorially. This procedure assumes that the substituent effect is always constant but in the case of indole the calculated and experimental results are in good agreement.

The calculated dipole moment for styrene is extremely low and in poor agreement with even the closest

experimental result<sup>5</sup>, but the range of values for this dipole moment<sup>6,7</sup> indicates a degree of uncertainty in its true value.

The calculated dipole moments for the quinoid compounds (Table 3) suggests that isoindole is more polar than indole, and that benzofuran and benzothiophen are less polar than their Kekulé counterparts. The net atomic charges in isoindole indicate as in indole and pyrrole that the dipole moment arises largely from the  $N^{\delta-} - H^{\delta+}$  bond.

From the data in Table 4 it is obvious from the sulphur containing heterocycles that both the magnitude and direction of the dipole moment are very sensitive to the size of basis set used, the larger basis set producing results in closer agreement to the experimental values.

A graph, Fig.3, has been drawn of experimental  $\mu$  vs. calculated dipole moments for a variety of molecules in series I and II of the bicyclic heterocycles. The values plotted are listed in Table 5. Of the fifteen molecules used, only three do not fit well. These are benzofuran, benzimidazole and benzotriazole. A reason for a high value of  $\mu_{calc.}$  for benzofuran has already been discussed. For the other two compounds the experimental measurement of the dipole moment was carried out in dioxane as solvent because of the insolubility of these compounds in less polar solvents. There is a possibility that there will be a degree of hydrogen

Fig. 3.

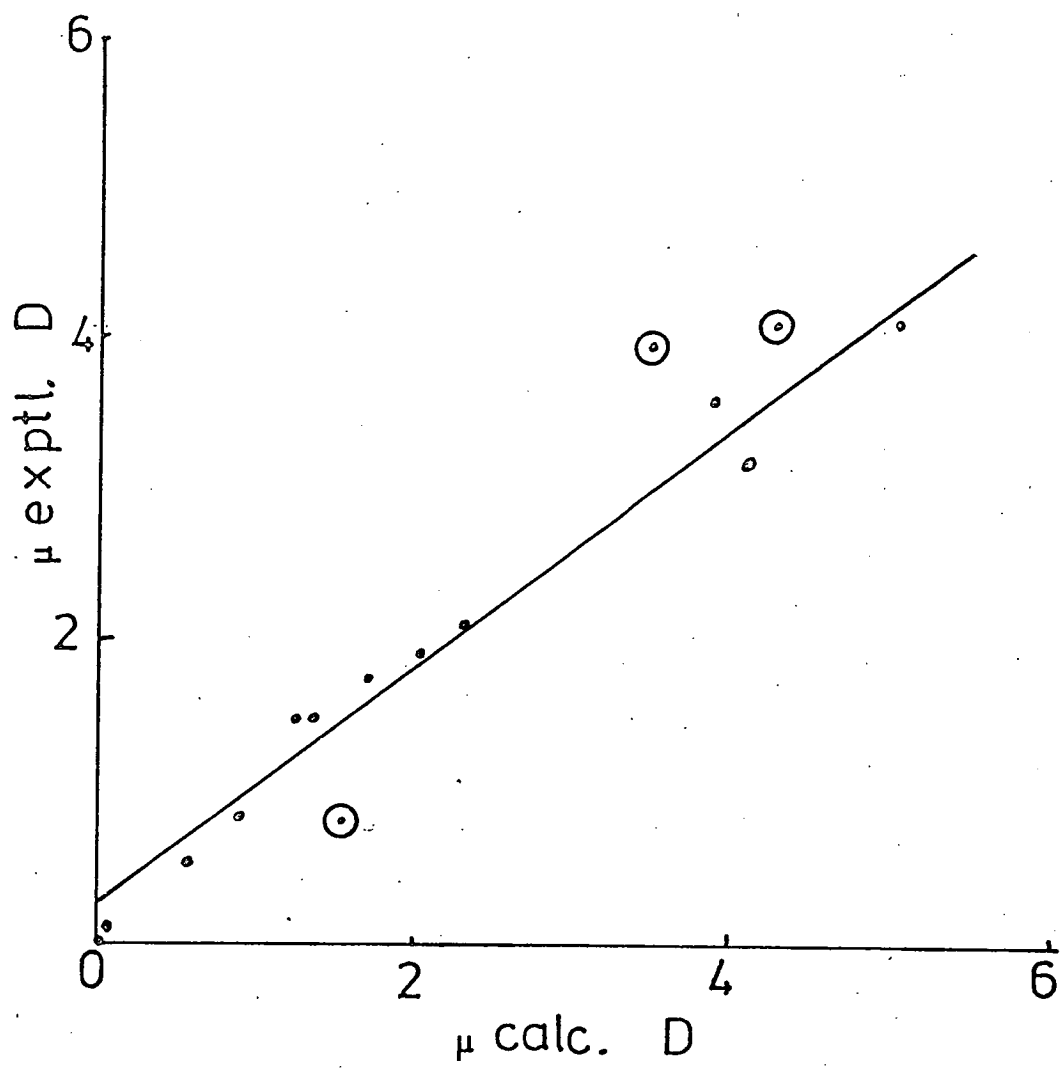
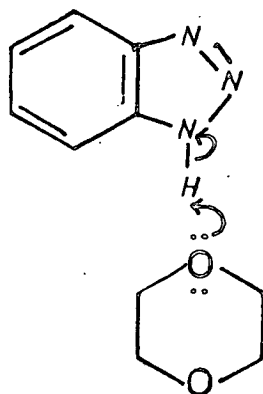


Fig. 4



bonding as shown in Fig.4 between the oxygen lone pair of the solvent molecule and the hydrogen of the N - H group which will be polarised  $N^{\delta-} - H^{\delta+}$ . This will tend to lengthen the N - H bond, increasing the charge separation, and consequently the dipole moment, leading to high experimental values. The scaled basis sets used may also lead to the average position of the N-lone pair in a  $\text{>N:}$  group to be too far out, as was suggested for benzofuran, leading again to an exaggerated charge separation and a high calculated dipole moment. A least squares fit of the data in Table 5 produces the relationship  $\mu_{\text{expt.}} = 0.846\mu_{\text{calc.}} + 0.176$ , with standard deviations in slope and intercept of 0.071 and 0.194, and an overall standard deviation of 0.398. The calculated dipoles are in general about 18% too high indicating that the calculations produce molecules which are too polar. Bearing in mind the sensitivity of the dipole moment to the basis set used, it must be remembered that the basis sets were scaled using total energy as the criterion for optimisation, and therefore may not be optimum with respect to this molecular property as already mentioned in the "Basis sets Section".

Table 3 Dipole moments ( $\mu/D$ )

	$\mu$	$\theta$
2H-indene	0.952	180
Isoindole	2.940	180
N-methyl isoindole	3.428	180
Benzo(c)furan	1.485	0
Benzo(c)thiophen (spd + 3s')	0.054	0

Table 4 Dipoles,  $\mu$ , for the bicyclic heterocycles II

	Calculated		Observed	
	$\mu$	$\theta$	$\mu$	solvent
Indazole (26a)	1.971	30 <sup>0</sup>	1.85 <sup>2</sup>	benzene
1-methyl indazole (26d)	2.092	49 <sup>0</sup>		
2H-indazole (27a)	2.572	-116 <sup>0</sup>		
2-methyl indazole (27d)	3.072	-133 <sup>0</sup>		
Benzotriazole (29a)	4.251	50 <sup>0</sup>	4.10 <sup>2</sup>	dioxane
1-methyl benzotriazole (29c)	4.575	60 <sup>0</sup>		
2H-benzotriazole (30a)	0.547	180 <sup>0</sup>		
2-methyl benzotriazole (30d)	1.574	105 <sup>0</sup>		
Benzimidazole (28a)	3.501	101 <sup>0</sup>	3.96 <sup>2</sup>	dioxane
Benzoxazole (28b)	1.238	32 <sup>0</sup>	1.48 <sup>2</sup>	
Anthranil (27b)	4.056	-41 <sup>0</sup>	3.09 <sup>2</sup>	benzene
1,2-benzisoxazole (26b)	4.096	-30 <sup>0</sup>		
2,1,3-benzoxadiazole (30b)	5.042	0 <sup>0</sup>	4.03 <sup>8</sup>	benzene



Table 4 (cont'd)

	Calculated		Observed	
	$\mu$	$\theta$	$\mu$	solvent
Benzothiazole (28c) (sp)	1.097	28 <sup>0</sup>	1.46 <sup>2</sup>	benzene
(spd + 3s <sup>1</sup> )	1.325	70 <sup>0</sup>		
1,2-benzisothiazole(26c) (sp)	3.211	-18 <sup>0</sup>		
(spd + 3s <sup>1</sup> )	2.494	-9 <sup>0</sup>		
2,1-benzisothiazole(27c) (sp)	2.858	-57 <sup>0</sup>		
(spd + 3s <sup>1</sup> )	2.342	-73 <sup>0</sup>		
2,1,3-benzthiadiazole(30c)(sp)	2.45	0 <sup>0</sup>	1.73 <sup>8</sup>	benzene
(spd + 3s <sup>1</sup> )	1.565	0 <sup>0</sup>		
1,2,3-benzthiadiazole(29b)	3.878	33 <sup>0</sup>	3.57 <sup>2</sup>	benzene
(spd + 3s <sup>1</sup> )				

Table 5

	<u>x-axis/<math>\mu</math> observed</u>	<u>y-axis/<math>\mu</math> experimental</u>
Naphthalene	0.00	0.00
Styrene	0.13	0.027
Indene	0.85	0.915
Indole	2.08	2.31
Benzofuran	0.79	1.56
Benzothiophen	0.62	0.56
Indazole	1.85	1.971
Benzotriazole	4.10	4.251
Benzimidazole	3.96	3.501
Benzoxazole	1.48	1.238

Table 5 (cont'd)

	<u>x-axis/<math>\mu</math> observed</u>	<u>y-axis/<math>\mu</math> experimental</u>
2,1,3-benzoxadiazole	4.03	5.042
Benzothiazole	1.46	1.325
2,1,3-benzthiadiazole	1.73	1.565
1,2,3,-benzthiadiazole	3.57	3.878
Anthranil	3.09	4.06

R E F E R E N C E S

1. G. Berthier, L. Praud and J. Serre, "Quantum Aspects of Heterocyclic Compounds in Chemistry and Biochemistry", eds. E.D. Bergmann and B. Pullman, The Israel Academy of Sciences and Humanities, Jerusalem, 1970, Academic Press.
2. A.L. McClellan, "Tables of Experimental Dipole Moments", Freeman, 1963.
3. R.D. Brown and B.A.W. Collier, *Theoret. Chim. Acta*, 1967, 7, 259.
4. M.H. Palmer, R.H. Findlay and A.J. Gaskell, *J. Chem. Soc., Perkin II*, 1974, 420.
5. H. Weiler - Feilchenfeld, A. Pullman, H. Berthod and C. Geissner-Prettre, *J. Mol. Structure*, 1970, 6, 297.
6. R.A.Y. Jones, A.R. Katritsky and A.V. Ochkin, *J. Chem. Soc.(B)*, 1971, 1795.
7. a) A.J. Petro and C.P. Smyth, *J. Amer. Chem. Soc.*, 1957, 79, 6142 : *ibid*, 1958, 80, 73.  
b) J.E. Plamondon, R.J. Buenker, D.J. Koopman and R.J. Dolter, *Proc. Iowa Acad. Sci.*, 1963, 70, 163 (C.A., 1964, 61, 11, 424a)
8. R.A.W. Hill and L.E. Sutton, *J. Chim Phys.*, 1949, 46, 244
9. Z. Yu. Kokoshko, V.G. Kiteava, Z.V. Pushkareva and V.E. Blokhin, *Zhür. obshchei. Zhim.*, 1967, 37, 58.

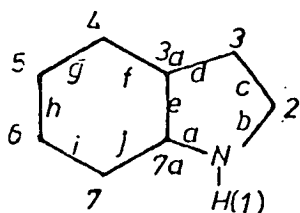
10. D.V.G.L.N. Rao, Indian J.Phys., 1955, 29, 398
11. E.F.J. Janetzky and M.C. Lebret, Rec. Trav. Chim., 1944, 63, 123
12. R.G. Charles and H. Fieser, J. Amer. Chem. Soc., 1950, 72, 2233.

Population Analyses and Bond Population Moments

Population analyses, giving total populations, and sigma ( $\sigma$ ) and pi ( $\pi$ ) components at each centre in the molecules, have been performed. The separation into  $\sigma$  and  $\pi$  components is used as an aid to identifying significant changes in electron density at a centre, as these changes may only show as small differences in the total population. The separation procedure also assists in the identification of the way in which a bond is polarised. The total populations give the overall polarisation but do not indicate whether this is due to dominance by the  $\sigma$  or  $\pi$  system, bond polarisation often not being compensatory in these systems.

M.H. Palmer<sup>1</sup> et al have devised a method of analysing net atomic populations in terms of of bond population moments relative to the unpolarised classical structure with completely separate  $\sigma$  and  $\pi$  systems. This method involves taking the net population at a centre in the  $\sigma$  or  $\pi$  system, and finding that centre's contribution to the system by subtracting the net population from the classical atomic population. A similar procedure is applied to all atoms in the molecule and the resultant population at each ring atom found. The bond population moments are then found by solving a series of homogeneous simultaneous linear equations for each ring atom. For a ring of  $n$  atoms,  $(n-1)$  equations are formed. The  $n$ th equation is formed by taking the vector sum of the components around the ring, and setting this equal

Table 1



## Net Populations

	N	C(2)	C(3)	C(4)	C(5)	C(6)	
$\sigma$	5.7891	4.9730	5.0847	5.1465	5.1283	5.1421	
$\pi$	1.6585	1.0494	1.1057	0.9949	1.0356	1.0087	
	C(7)	C(7a)	C(3a)	H(1)	H(2)	H(3)	H(4)
$\sigma$	5.1147	4.8486	4.9914	0.6656	0.8386	0.8576	0.8526
$\pi$	1.0498	1.0409	1.0562				
	H(5)	H(6)	H(7)				
$\sigma$	0.8564	0.8531	0.8579				

## Electron contributions from each centre to the bond moments

	N	C(2)	C(3)	C(4)	C(5)	C(6)	
$\sigma$	-0.7891	+0.0270	-0.0847	-0.1465	-0.1283	-0.1421	
$\pi$	+0.3415	-0.0494	-0.1057	+0.0051	-0.0356	-0.0087	
	C(7)	C(7a)	C(3a)	H(1)	H(2)	H(3)	H(4)
$\sigma$	-0.1147	+0.1514	+0.0086	+0.3344	+0.1614	+0.1424	+0.1474
$\pi$	-0.0498	-0.0409	-0.0562				
	H(5)	H(6)	H(7)				
$\sigma$	+0.1436	+0.1469	+0.1421				

Table 1 (contd.)

## Resultant contributions at each ring atom

	N	C(2)	C(3)	C(4)	C(5)	C(6)
$\sigma$	-0.4547	+0.1884	+0.0577	+0.0090	+0.0153	+0.0048
$\pi$	+0.3415	-0.0494	-0.1057	+0.0051	-0.0356	-0.0087
	C(7)	C(7a)	C(3a)			
$\sigma$	+0.0274	+0.1514	+0.0086			
$\pi$	-0.0498	-0.0409	-0.0562			

The equations are set up as follows, where the direction of the arrows and hence the signs in the equations are as shown in Figure 3 (page 194):-

 $\sigma$ -system

<u>1</u>	$a + b$	$= +0.4547$	<u>6</u>	$f - g$	$= +0.0090$
<u>2</u>	$b - c$	$= +0.1884$	<u>7</u>	$g + h$	$= +0.0153$
<u>3</u>	$c + d$	$= +0.0577$	<u>8</u>	$h - i$	$= -0.0048$
<u>4</u>	$d - e + f$	$= -0.0086$	<u>9</u>	$i + j$	$= +0.0274$
<u>5</u>	$a - e - j$	$= +0.1514$			

10  $a - b - c + d + e = 0$  (5-membered ring)

11  $e + f + g - h - i - j = 0$  (6-membered ring)

Substitution of equations (in sequence) into Equation 11 in the order:-

5, 8, (2x)7, (3x)6, (4x)4, (6x)10m leads to the  $\sigma$ -system values:

$a = 0.2238$ ,  $b = 0.2309$ ,  $c = 0.0425$ ,  $d = 0.0152$ ,  $e = 0.0344$ ,  $f = 0.0106$ ,  $g = 0.0097$ ,  
 $h = 0.0056$ ,  $i = 0.1104$ ,  $j = 0.0378$ .

$\pi$ -system

1     $a + b = +0.3415$

6     $-f + g = +0.0051$

2     $b - c = +0.0494$

7     $g + h = +0.0356$

3     $c - d = +0.1057$

8     $-h + i = +0.0087$

4     $d + e - f = +0.0562$

9     $-i + j = +0.0498$

5     $a - e - j = +0.0409$

10    $a - b - c - d + e = 0$

11    $e + f + g - h - i - j = 0$

This leads to the  $\pi$ -system values:

$a = 0.1868$ ,  $b = 0.1547$ ,  $c = 0.1053$ ,  $d = -0.0004$ ,  $e = 0.0728$ ,  $f = 0.0162$ ,  $g = 0.0213$ ,  
 $h = 0.0143$ ,  $i = 0.0230$ ,  $j = 0.0728$ .

The significance of the signs is such that following an initial assumed direction for the bond dipole, if a final sign for a bond dipole (a to j) is found to be negative, then the initially assumed direction for the bond dipole must be reversed.



Fig.1

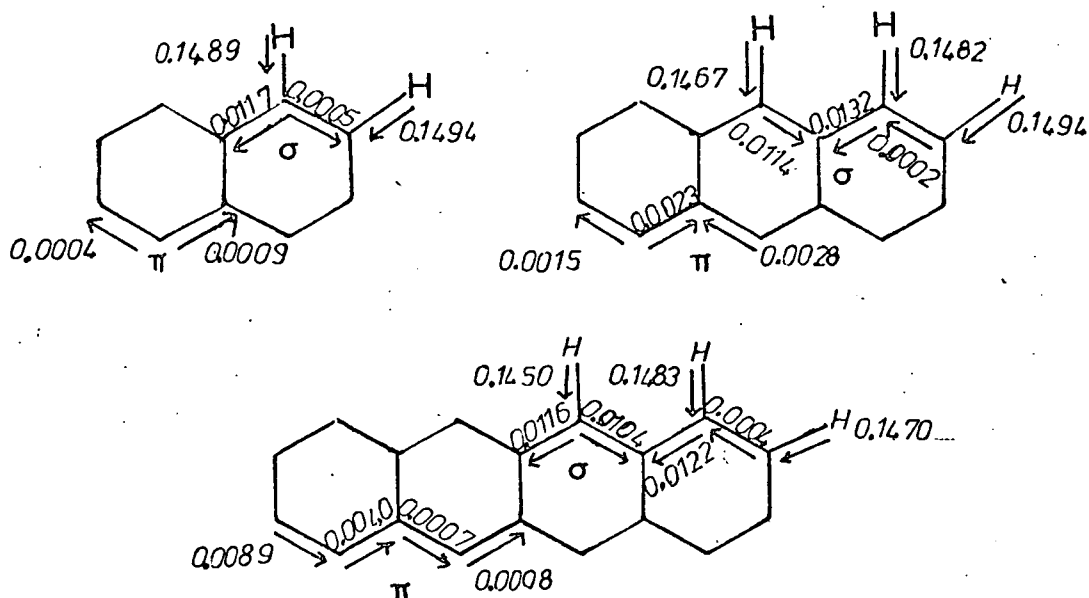
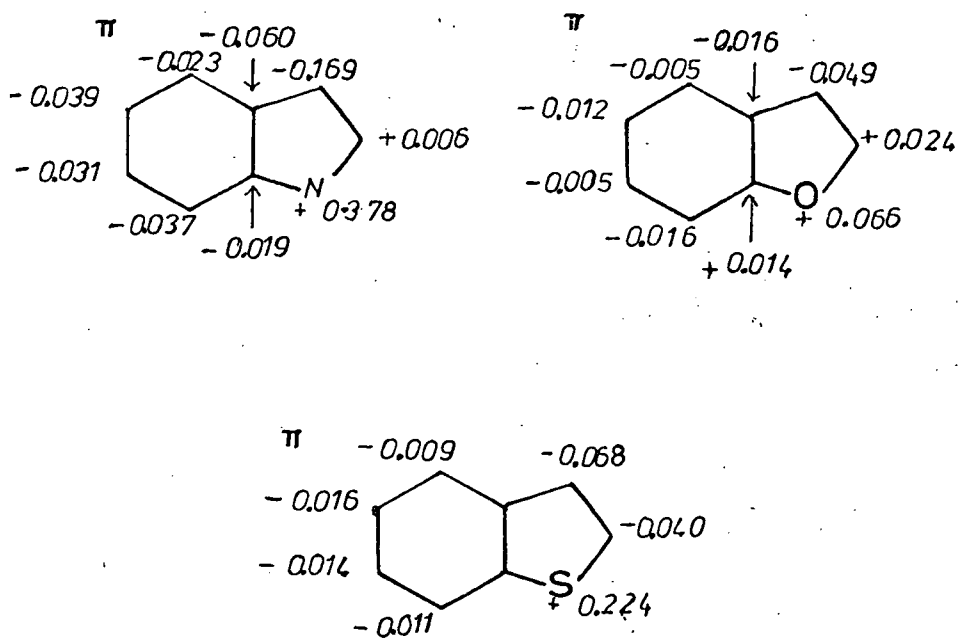


Fig.2



to zero as the process is relative to the neutral atoms. Table 1 gives the relevant data and procedure for the evaluation of the bond population moments in indole. For molecules with  $C_{2v}$  symmetry the process can be carried out by inspection starting with an atom on the  $C_2$  axis, and assuming a symmetrical split of the resultant population.

For the hydrocarbon series in this work, the overlap populations have also been considered as a means of obtaining a bond order,  $P_{rs}$ , bond length,  $R_{rs}$ , relationship, the bond orders being calculated from the LCGO calculations.

Hydrocarbons For the hydrocarbons (1-4), the  $\sigma, \pi$  and total populations at each centre are shown in Table 2. In each case the C - H bonds are polarised as  $C^{\delta-} - H^{\delta+}$  with a population separation of  $\approx \pm 0.15$  electrons. The hydrogen population is almost constant throughout the series, and the carbon populations vary by very small amounts, the bridgehead carbons having a small excess of population  $\approx 0.03$  electrons. A bond population moment analysis on the small  $\sigma$  and  $\pi$  separations of 2 - 4 yields the results shown in Fig.1. The  $\sigma$ -bond charge separation is in all cases small, and the  $\pi$ -moments negligible. This is consistent with early Huckel theory<sup>2</sup> which postulated a lack of charge separation in alternate hydrocarbons.

The overlap populations,  $N_{rs}$ , for these molecules are also shown in Table 2 and show significant variations. When separated into  $\sigma$ ,  $N_{rs}(\sigma)$  and  $\pi$ ,  $N_{rs}(\pi)$ , components

Table 2 Population Analysis

1. Benzene	C(1)	H(1)			
$\sigma$	5.1483	0.8517			
$\pi$	1.0000	-			
Total	6.1483	0.8517			
2. Naphthalene	C(1)	C(2)	C(4a)	H(1)	H(2)
$\sigma$	5.1381	5.1493	5.0206	0.8511	0.8511
$\pi$	0.9986	1.0004	1.0018	-	-
Total	6.1367	6.1497	6.0234	0.8511	0.8511
3. Anthracene	C(1)	C(2)	C(9)	C(1a)	
$\sigma$	5.1352	5.1492	5.1236	5.0246	
$\pi$	0.9962	1.0015	0.9946	1.0051	
Total	6.1314	6.1507	6.1182	6.0297	
	H(1)	H(2)	H(9)		
Total	0.8518	0.8506	0.8533		
4. Tetracene	C(1)	C(2)	C(5)	C(4a)	C(5a)
$\sigma$	5.1365	5.1466	5.1228	5.0226	5.0233
$\pi$	1.0049	0.9911	0.9997	1.0033	1.0016
Total	6.1415	6.1378	6.1225	6.0260	6.0249
	H(1)	H(2)	H(5)		
Total	0.8517	0.8530	0.8550		

Table 2 (contd)

(b) Overlap Populations ( $N_{rs}$ )

Benzene	1,2				
$\sigma$	0.3899				
$\pi$	0.1308				
Total	0.5207				
Naphthalene	1,2	2,3	1,9	9,10	
$\sigma$	0.3908	0.3889	0.3975	0.4082	
$\pi$	0.1691	0.0917	0.0879	0.1123	
Total	0.5599	0.4806	0.4854	0.5205	
Anthracene	1,2	2,3	1,10a	9a,10a	9,9a
$\sigma$	0.3921	0.3878	0.3963	0.3769	0.3997
$\pi$	0.1779	0.0782	0.0721	0.0915	0.1232
Total	0.5700	0.4660	0.4684	0.4684	0.5229

(Table 2) the variation is seen to be in the  $\pi$ -system, the  $\sigma$ -system remaining constant at  $0.390 \pm 0.01$  electrons. The  $\pi$ -system overlap population is seen to vary with the bond length,  $R_{rs}$ . The shorter the bond length the greater is the  $\pi$ -overlap populations,  $N_{rs}(\pi)$ . When a least squares fit is performed of  $R_{rs}$  with  $N_{rs}(\pi)$  using data for the molecules benzene, naphthalene and anthracene supplemented by data for ethylene, acetylene and buta - 1,3 - diene, giving a total of 35 points, the relationship is given as  $R_{rs} = 1.497 - 0.785 N_{rs}(\pi)$  with standard deviations in slope, intercept and overall of 0.058, 0.008 and 0.019 respectively. These results can be converted to a linear relationship in terms of Hückel  $\pi$ -bond order whose LCGO equivalent is  $P_{rs}$  by dividing the overlap population by the overlap integral. This produces  $R_{rs} = 1.491 - 0.197 P_{rs}$  with standard deviations as before of 0.007, 0.004 and 0.010 respectively using data in the range  $1.203 < R_{rs} < 1.467$ . All bond lengths in the hydrocarbon series are within this range, and the relationship of  $R_{rs}$  to  $P_{rs}$  gives the hydrocarbon bond lengths to within  $0.016 \text{ \AA}$ . This is comparable to accuracy obtained by experimental methods and the relationship could well be of value in constructing geometries of conjugated hydrocarbons where no experimental data is available. Similar relationships have been established using semiempirical SCF bond orders<sup>3</sup> but with less success. For example, these calculations tend to suggest that the  $C_{\alpha\beta}$  bond in naphthalene is greater than  $1.421 \text{ \AA}$ .

The Bicyclic Heterocycles I The total populations and  $\sigma$ -, and  $\pi$ -components are shown in Table 3. For styrene and indene, the total populations vary only marginally from centre to centre, as would be expected with hydrocarbons. In the  $\sigma$ - and  $\pi$ -populations, the C( $\sigma$ ) populations of the C atoms attached to H show an increase in population, giving hydrogen as a  $\sigma$ -donor. In indene where two hydrogen atoms are outwith the molecular plane but preserving the  $C_s$  symmetry of the molecule, the H(1) populations can be divided into  $\sigma$  and  $\pi$  components. Each of these hydrogens shows  $\sigma$  and  $\pi$  donation to the C(1) atom producing a total population at C(1) of 6.3188 electrons which is high compared with all the other ring C atoms.

The heterocycles show more variation in total populations, particularly in the heterocyclic ring where the C(2) and C(7a) positions adjacent to the heteroatom are most affected. The C(3) position shows a high total population which was predicted by earlier  $\pi$ -electron theories<sup>4</sup> to be due to increased  $\pi$ -population, and therefore should be a reactive position for electrophilic substitution. On inspection of the  $\sigma/\pi$  populations of the C(3) positions, it is apparent that the increased population is mainly due to an increase in the  $\pi$ -population.

The  $\pi$ -electron only calculations<sup>4</sup> of  $\pi$ -electron density produce in all cases a large difference in the C(2) - C(3) positions for indole compared with those of benzofuran and benzothiophen as shown in Fig. 2. From these results indole should show preferential electrophilic

Table 3 Population Analysis

Naphthalene	C(1)	C(2)	C(4a)	H(1)	H(2)				
	$\sigma$	5.1381	5.1493	5.0206	0.8511	0.8506			
	$\pi$	0.9986	1.0004	1.0018					
Total		6.1367	6.1497	6.0224					
Styrene	C(1)	C(2)or(6)	C(3)or(5)	C(4)					
	$\sigma$	5.0306	5.1353	5.1499	5.1473				
	$\pi$	1.0028	0.9996	0.9970	1.0012				
Total		6.0334	6.1349	6.1478	6.1485				
	$C_{\alpha}$	$C_{\beta}$	$H_{\alpha}$	$H_{\beta}(cis)$	$H_{\beta}(trans)$	H(2)or(6)	H(3)or(5)	H(4)	
	$\sigma$	5.1462	5.2667	0.8541	0.8588	0.8632	0.8541	0.8517	0.8518
	$\pi$	1.0026	0.9979						
Total		6.1488	6.2646						
Indene	C(1)	C(2)	C(3)	C(4)	C(5)	C(6)			
	$\sigma$	5.1766	5.1429	5.1393	5.1415	5.1435	5.1426		
	$\pi$	1.1422	0.9892	1.0224	1.0086	1.0089	1.0100		
Total		6.3188	6.1321	6.1617	6.1501	6.1524	6.1526		

	C(7)	C(3a)	C(7a)	H(1)	H(2)	H(3)	H(4)	H(5)	H(6)	H(7)
	5.0523	5.0198	5.0144	0.4409	0.8523	0.8500	0.8523	0.8528	0.8530	0.8543
	1.1011	1.0038	0.9900	0.4119						
Total	6.1534	6.0236	6.0044	0.8528						

Indole	N	C(2)	C(3)	C(4)	C(5)	C(6)	C(7)	C(7a)
$\sigma$	5.781	4.9730	5.0847	5.1465	5.1283	5.1421	5.1147	4.8486
$\pi$	1.6585	1.0494	1.1057	0.9949	1.0356	1.0087	1.0498	1.0409
Total	7.4476	6.0224	6.1904	6.1414	6.1639	6.1508	6.1647	5.8895

	C(3a)	H(1)	H(2)	H(3)	H(4)	H(5)	H(6)	H(7)
$\sigma$	4.9914	0.6656	0.8386	0.8576	0.8526	0.8564	0.8531	0.8579
$\pi$	1.0562							
Total	6.0476							

Benzofuran	O	C(2)	C(3)	C(4)	C(5)	C(6)	C(7)
$\sigma$	6.7050	4.8486	5.1343	5.1422	5.1355	5.1503	5.1226
$\pi$	1.7930	1.0675	1.0331	0.9949	1.0196	0.9918	1.0308
Total	8.4980	5.9161	6.1674	6.1371	6.1551	6.1421	6.1534



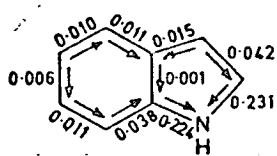
Table 3 (cont'd)

		C(3a)	C(7a)	H(2)	H(3)	H(4)	H(5)	H(6)	H(7)
	$\sigma$	4.7397	4.9940	0.8124	0.8423	0.8349	0.8450	0.8433	0.8478
	$\pi$	1.0551	1.0142						
Total		5.7948	6.0082						
Benzothiophen (spd + 3s' basis)	S	C(2)	C(3)	C(4)	C(5)	C(6)	C(7)		
	$\sigma$	12.1757	5.0973	5.1298	5.1447	5.1376	5.1483	5.1215	
	$\pi$	3.7839	1.0689	1.0287	0.9928	1.0175	0.9969	1.0214	
Total		15.9596	6.1662	6.1585	6.1375	6.1551	6.1452	6.1429	
		C(3a)	C(7a)	H(2)	H(3)	H(4)	H(5)	H(6)	H(7)
	$\sigma$	4.9845	4.9930	0.8301	0.8429	0.8481	0.8506	0.8479	0.8480
	$\pi$	1.0431	1.0468						
Total		6.0276	6.0398						

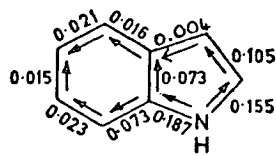
substitution in the 3-position, and the other molecules should show a slight preference for substitution at C(3) compared with C(2). Experimental evidence<sup>5</sup> shows that in indole the 3-position is the most reactive site for substitution, and that in benzothiophen, the 2 or 3 position is substituted depending on the electrophile used (e.g. nitration and halogenation give 3-substitution while alkylation with propylene or isopropanol gives almost entirely 2-substitution), but overall the 3-position is more readily substituted. Benzofuran, however, has its 2-position as the most reactive site. Attempts have been made to explain the experimental results in terms of atom localisation energies, but the results are unsatisfactory in that they depend on the electronegativity value chosen for the carbon atoms, the C(2) value being particularly sensitive.

Using the LCGO population data for the molecules, it can be seen that indole has by far the largest  $\pi$ -electron population at C(3), and only in this case is the C(3)  $\pi$  population greater than the C(2)  $\pi$ -population indicating that a definite preference should be shown at C(3) towards electrophilic substitution, as is the case. For benzothiophen and benzofuran, the C(2) $\pi$ -population is greater than that of C(3) by about 0.04 electrons in each case. This difference is only about half of that shown by indole, and in reverse indicating electrophilic substitution at the 2-position in these molecules. The anomaly here is benzothiophen which, although giving C(2) and C(3) substituted compounds,

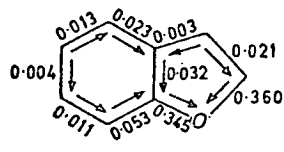
Fig.3



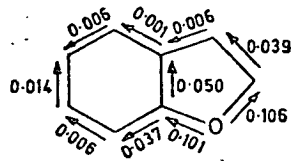
( $\sigma$ )



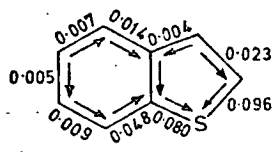
( $\pi$ )



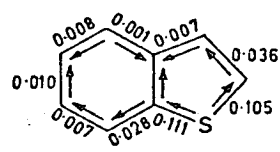
( $\sigma$ )



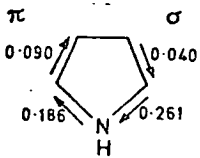
( $\pi$ )



( $\sigma$ ) (*spd + 3s'* basis)

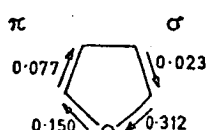


( $\pi$ ) (*spd + 3s'* basis)



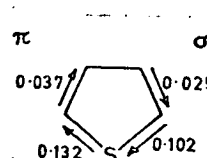
$\pi$

$\sigma$



$\pi$

$\sigma$

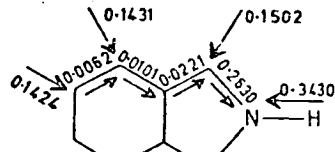


$\pi$

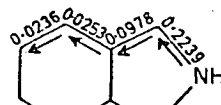
$\sigma$

(*spd + 3s'* basis)

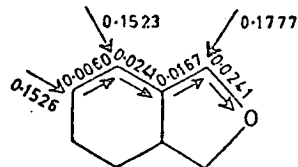
Fig.4



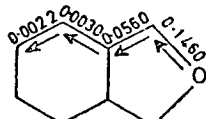
( $\sigma$ )



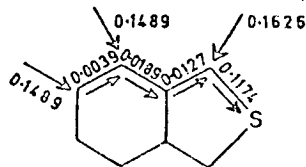
( $\pi$ )



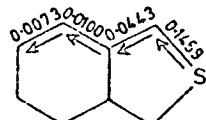
( $\sigma$ )



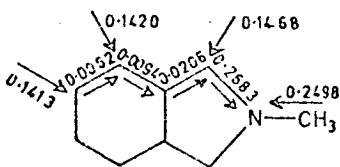
( $\pi$ )



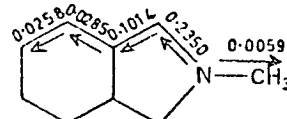
( $\sigma$ )



( $\pi$ )



( $\sigma$ )



( $\pi$ )

seems to prefer C(3) under mild conditions. This can perhaps be attributed to the postulated similarity in behaviour of an S atom to C=C. Replacing the S atom by  $\text{CH=CH}$  gives naphthalene, and naphthalene is more reactive to electrophilic substitution at the  $\alpha$  position which would be equivalent to the C(3) position in benzothiophen. When the heterocyclic molecules are analysed in terms of bond population moments the results are as shown in Fig. 3. Noting the direction of movement of  $\pi$  electrons in the heterocyclic ring shows indole with a large  $\pi$ -donation to C(3), and benzofuran and benzothiophen with a large  $\pi$ -donation at C(2) as expected from the total  $\pi$ -population at each centre. The small and relatively constant  $\pi$  bond moments in the carbocyclic ring indicate that the effect of the heteroatom is very much confined to the heterocyclic ring, but the long range effect of the heteroatom is reflected in the C(5) position as the terminal point in the  $\pi$ - and the starting point in the  $\sigma$ -ring currents. In all cases the directions of the  $\sigma$  and  $\pi$  electron movement are in opposition, and the heteroatoms act as  $\sigma$ -acceptors and  $\pi$ -donors. In the carbocyclic rings, the C(5) position is where the  $\pi$  electron movement centres, accepting  $\pi$ -electrons from adjacent carbon atoms (4 and 6), all other carbon atoms in this ring being  $\pi$ -electron donors and acceptors. When the heterocyclic ring contains a de-activating constituent, subsequent substitution occurs in the benzenoid ring, and using  $\pi$ -bond population moments, it would be predicted to occur at the C(5) position where the  $\pi$ -

electron movement terminates. However the net  $\pi$ -population at this centre must also be considered, as the highest  $\pi$ -population centres are given as C(5) and C(7) in this ring. When 3-ethoxy carbonyl indole is brominated the product is a mixture of the 5 and 7 mono-bromo substituted compounds<sup>6</sup> showing that the highest  $\pi$ -population centres have been preferentially attacked. So although C(7) is a  $\pi$ -donor and acceptor its donor nature is small compared to its acceptor nature and the resultant is a high  $\pi$ -population at that centre. The bond moments for the corresponding monocyclic compounds are also shown in Fig.3, and it can be seen that for the S and N heterocycles the  $\pi$  and  $\sigma$  components of the bi- and mono-cycles are of similar magnitude, but that for furan and benzofuran, the latter is a better  $\sigma$ -acceptor and a worse  $\pi$ -donor. This would tend to indicate that of the three molecules, benzofuran will be the least aromatic. This is indeed the case as shown by resonance energy considerations.

For the quinoid series, the populations are shown for the heterocycles in Table 4, and the bond population moments in Fig. 4. As with the Kekulé series, the  $\sigma$  and  $\pi$  systems are in opposite directions to compensate each other, and again the heteroatoms are  $\sigma$ -acceptors and  $\pi$ -donors, with the long range effect of the heteroatom being transmitted to the symmetrical C(5) and C(6) positions. The methyl group of N-methyl isoindole is a  $\sigma$ -donor to the nitrogen atom, and almost neutral in the  $\pi$ -system showing no hyperconjugative effect. The  $\pi$ -bond population moments show a mark

Table 4 Population Analyses

Isoindole		N(2)	C(1)	C(3a)	C(4)	C(5)	H(1)	H(4)	H(5)	H(2)
Population	$\sigma$	5.8691	4.9093	4.9880	5.1390	5.1362	0.8498	0.8569	0.8576	0.6570
	$\pi$	1.5522	1.1261	1.0725	1.0017	1.0236				
Total		7.4216	6.0354	6.0605	6.1407	6.1598	0.8498	0.8569	0.8576	0.6570
N-Methy- isoindole		N(2)	C(1,3)	C(3a,7a)	C(4,7)	C(5,6)	C(Me)	H(1)	H(4)	H(5)
Population	$\sigma$	5.7864	4.8991	4.9888	5.1388	5.1351	5.1237	0.8532	0.8580	0.8587
	$\pi$	1.5243	1.1336	1.0729	1.0027	1.0258	1.1647			
Total		7.3107	6.0327	6.0617	6.1415	6.1609	6.2884	0.8532	0.8580	0.8587
		H(Me)(av)		H(Me) <sub>ip</sub>						
	$\sigma$	0.4024		0.8217						
	$\pi$	0.4205								
Total		0.8229		0.8217						

Table 4 (cont'd)

Benzo(c)furan	O(2)	C(1)	C(3a)	C(4)	C(5)	H(1)	H(4)	H(5)
Population	6.7716	4.8086	5.0073	5.1342	5.1406	0.8223	0.8477	0.8474
	1.7079	1.0090	1.0530	1.0009	1.0022			
Total	8.4795	5.8986	6.0603	6.1351	6.1488	0.8223	0.8477	0.8474
Benzo(c)thiophen	S(2)	C(1)	C(3a)	C(4)	C(5)	H(1)	H(4)	H(5)
Population	12.2349	5.0579	5.0062	5.1340	5.1450	0.8374	0.8511	0.8511
	3.7082	1.1016	1.0027	1.0027	1.0073			
Total	15.9430	6.1595	6.1366	6.1366	6.1523	0.8374	0.8511	0.8511

increase in  $\pi$ -donation in the heterocyclic ring producing slightly larger bond population moments, in the carbocyclic ring when compared with the Kekulé compounds. Comparing the quinoid series with the corresponding monocyclic compounds gives an even closer relationship in bond population moments than for the Kekulé series. Even furan and benzo(c)furan are closely related. This supports the conclusions from PE data which showed the monocycles and quinoid compounds to have very similar molecular orbital make up in the valence shell.



R E F E R E N C E S

1. M.H. Palmer, R.H. Findlay, and A.J. Gaskell, J. Chem. Soc, Perkin II, 1974, 420.
- 2.(a) A. Streitwieser, "Molecular Orbital theory for Organic Chemists", Wiley, New York, 1961.  
(b) L. Salem, "Molecular Orbital Theory of conjugated systems", Benjamin, New York, 1966
- 3.(a) F.Momicchioli and A. Rastelli, J Mol. Spec., 1967, 22, 310.  
(b) M.J.S. Dewar and N. Trinajstic, J. Chem. Soc.(A), 1971, 1220.  
(c) M.J.S. Dewar and G.J. Gleicher, J. Chem. Phys., 1966, 44, 759.  
(d) M.J.S. Dewar, A.J. Harget, N. Trinajstic and S.D. Worley, Tetrahedron, 1970, 26, 4505.  
(e) J. Fabian, A. Mehlhorn and R. Zahradnik, J. Phys. Chem., 1968, 72, 3975.  
(f) M.J.S. Dewar and C. de Llano, J. Amer. Chem. Soc., 1969, 91, 789.  
(g) M.J.S. Dewar and N. Trinajstic, Coll. Czech. Chem. Comm., 1970, 35, 3136.
- 4.(a) O.E. Polanski and G. Derflinger, Montosh, 1961, 92, 1114.  
(b) R.D. Brown and B.A.W. Collier, Austral. J. Chem., 1959, 12, 152.  
(c) C.A. Coulson and A. Streitwieser Jr., "Dictionary of  $\pi$ -Electron Calculations", Pergamon Press, Oxford, 1965.  
(d) R. Zahradnik, C. Párkányi, V. Horák and J. Koutecky, Coll. Czech. Comm., 1962, 28, 776.

5. M.H. Palmer, "The Structure and Reactions of Heterocyclic Compounds", Edward Arnold, London, 1967, p 315.
6. B.E. Leggeter and R.K. Brown, Canad. J. Chem., 1960, 38, 1467.

### The Average Position of $\pi$ -electrons

For a molecule to have high aromatic character, it is necessary for the  $\pi$ -electrons to function as a group rather than as pairs of electrons. The average positions ( $\bar{x}$ ,  $\bar{y}$ ,  $\bar{z}$ ) of the  $\pi$ -electrons from a fixed point should therefore be constant within the molecule when it is strongly aromatic such as in benzene, and the degree of departure from this towards  $\pi$ -electrons functioning as separate pairs will indicate the lack of aromaticity of the molecule. For the monocycles, thiophen, pyrrole and furan<sup>1</sup>, the average  $\pi$ -electron pair positions are closest in thiophen, the other two molecules tending to a quartet of  $\pi$ -electrons in the hydro-carbon part of the molecule, and a pair of electrons near the heteroatom. The separation of the quartet and pair is greater for furan than for pyrrole. This gives the order of aromaticity as thiophen > pyrrole > furan which is the order obtained from RE considerations.

For the bicyclic heterocycles I, the average  $\pi$ -electron pair positions are calculated relative to the mid-point of the C(3a) - C(7a) bond which is the y-axis. The average distance  $\bar{r}$  of the total  $\pi$ -charge with respect to this point has also been calculated as shown in Table 1.

The numbering of the  $\pi$ -electron pair orbitals is given in order of increasing binding energy,  $\pi_1$ , being at lowest binding energy. The nearer the centre of  $\pi$ -charge is to the reference point, the more aromatic will be . . .

Table 1 Average Position of  $\pi$ -Electron Pairs In Indole, Benzofuran and Benzothiophen

		$\pi_1$ (Å)	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$	$\pi_6$	$\bar{r}$ (Å)
Benzofuran	$\bar{y}$	0.0544	-0.0011	0.2381	-0.0745	-1.2867	-	1.3902
	$\bar{z}$	0.0869	-1.7504	0.7317	-1.9363	1.6309	-	
Indole	$\bar{y}$	0.1687	-0.0145	0.2455	-0.1428	-0.9406	-	1.1441
	$\bar{z}$	0.2546	-1.1962	-0.0136	-2.0017	1.5062	-	
Benzothiophen (spd + 3s')	$\bar{y}$	-0.3789	0.0006	-0.0022	-0.1485	-0.2769	-2.4818	0.5344
	$\bar{z}$	-0.2024	-0.2684	0.1393	-1.0415	0.5106	2.8419	
Naphthalene								0.0

Table 2 Average Positions of the  $\pi$ -Electron Pairs in the Quinoid Molecules

	$\pi_1$	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$	$\pi_6$
2H-indene	0.130	-0.267	-0.387	-0.221	2.101	-
Isoindole	0.200	-0.173	-0.183	-0.636	1.469	-
N-methyl Isoindole	-0.213	-0.387	-0.611	-0.614	1.403	-
Benzo(c)furan	0.044	-0.357	-0.094	-0.550	1.738	-
Benzo(c)thiophen (spd + 3s')	-0.285	-0.261	-0.439	-0.478	0.169	2.343

the molecule with reference to naphthalene which will have  $\bar{r} = 0.0$  using this system. The values of  $\bar{r}$  give the order of aromaticity as naphthalene > benzothiophen > indole > benzofuran which is consistent with that obtained from RES.

In benzothiophen the five pairs of  $\pi$  - electrons are scattered near the reference point with the S(2p $\pi$ ) orbital essentially localised in the region of the sulphur atom. At the other extreme benzofuran has very scattered electron pairs, one near the oxygen atom ( $\pi_5$ ), one near the centre of the heterocyclic ring ( $\pi_3$ ), one near the reference point ( $\pi_1$ ) and two pairs near the centre of the carbocyclic ring ( $\pi_2$  and  $\pi_4$ ). Indole is intermediate as expected with two pairs of  $\pi$  - electrons lying close to the reference point.

Table 2 gives the  $\pi$  - electron pair average positions with the centre of the C(3a) - C(7a) bond as origin for the molecules of the quinoid series. Again the S(2p $\pi$ ) level of sulphur is localised near the S atom. The smallest spread in  $\pi$  - electron pairs (0.67 $\text{\AA}$ ) is encountered in benzo(c)thiophen and the largest in 2H-indene (2.49 $\text{\AA}$ ). The order of aromaticity is given as naphthalene > benzo(c)thiophen > isoindole  $\approx$  benzo(c)furan > 2H-indene. This ordering is comparable to that found by other methods but isoindole is normally found to be considerably more aromatic than benzo(c)furan.

The non-coaxial nature of the  $\pi$  - electron pair positions in the Kekulé series makes it difficult to compare the

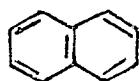
results for the Kekulé and quinoid series directly. This problem will be further aggravated for the more complex molecules of the bicyclic heterocycles II, and for this reason it was decided that the calculations of the  $\pi$ -electron pair positions would be of little use, and were, therefore, omitted.

R E F E R E N C E S

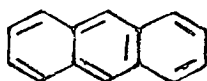
1. M.H. Palmer, R.H. Findlay and A.J. Gaskell,  
J. Chem. Soc., Perkin II, 1974, 420.



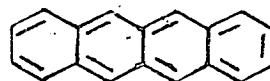
1



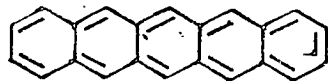
2



3



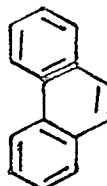
4



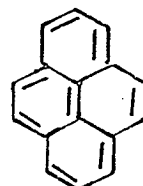
5



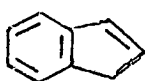
6



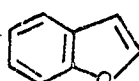
7



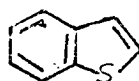
8



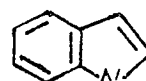
15



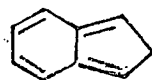
16



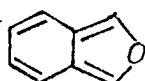
17



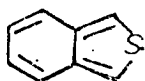
18



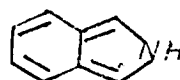
20



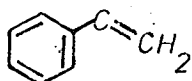
21



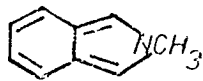
22



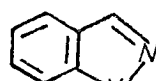
23



24

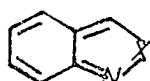


25



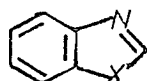
26

a, NH  
b, O  
c, S  
d, NCH<sub>3</sub>



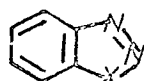
27

a, NH  
b, O  
c, S  
d, NCH<sub>3</sub>



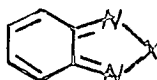
28

a, NH  
b, O  
c, S



29

a, NH  
b, S  
c, NCH<sub>3</sub>



30

a, NH  
b, O  
c, S  
d, NCH<sub>3</sub>



TRIPLET STATES

An investigation of the first excited triplet states of 2, 16 - 18 and 21 - 23 was performed using both RHF and UHF calculations, the transition taking place in each case by excitation of a  $\pi$ -electron in the highest occupied molecular orbital (HOMO) to the lowest unoccupied  $\pi$ -level (LUMO). Experimental data of low lying triplet state excitation energies exists for naphthalene<sup>1</sup>, indole<sup>23</sup> and benzothiopen<sup>4,5</sup>, but none exists for the quinoid series. The excitation energies were calculated using  $E_{\text{Total } M^*} - E_{\text{Total } M}$ , and the results are shown in Table 1. The excitation energies are also calculated using the semi-empirical INDO procedure. Where experimental data is available it can be seen that the UHF calculations produce the best agreement with experiment, but the magnitudes are  $\approx 0.5\text{eV}$  too large. The RHF and INDO procedures reproduce the trend in results fairly well, but not the magnitudes. The calculated excitation energies for the quinoid series and in particular for benzo(c)furan and 2H-indene are substantially lower than those of the Kekulé series. An analysis of the molecular orbitals along the lines of those of the closed shell molecules, showed that for each molecule, the closed shell molecular orbitals were mirrored in the open shell cases as shown for the valency shell orbitals of indole and isoindole, comparing RHF results, Fig. 1.

A list of open shell RHF orbital energies is given in

Table 1 Low Lying Triplet State Excitation energies (eV)

	Napthalene	Indole	Benzofuran	Benzothiophen
UHF	3.028	3.206	2.640	3.431
RHF	4.420	4.146	3.647	4.350
Expt.	2.665 <sup>1</sup>	3.06 <sup>2</sup> , 3.069 <sup>3</sup>		2.972 <sup>4</sup> , 2.965 <sup>5</sup>
INDO		6.549	5.896	4.134
	2H-indene	Isoindole	Benzo(c)furan	Benzo(c)thiophen
UHF	0.884	2.400	0.286	1.972
RHF	1.162	2.694	0.593	2.452
INDO		4.062	3.521	3.727

Fig.1

closed

$(\pm) = \sigma 2s$ ,  $\rightarrow = +p\sigma$ ,

open

$\leftarrow = -p\sigma$ ,

closed

$\rightleftharpoons = \pm p\pi$

open.

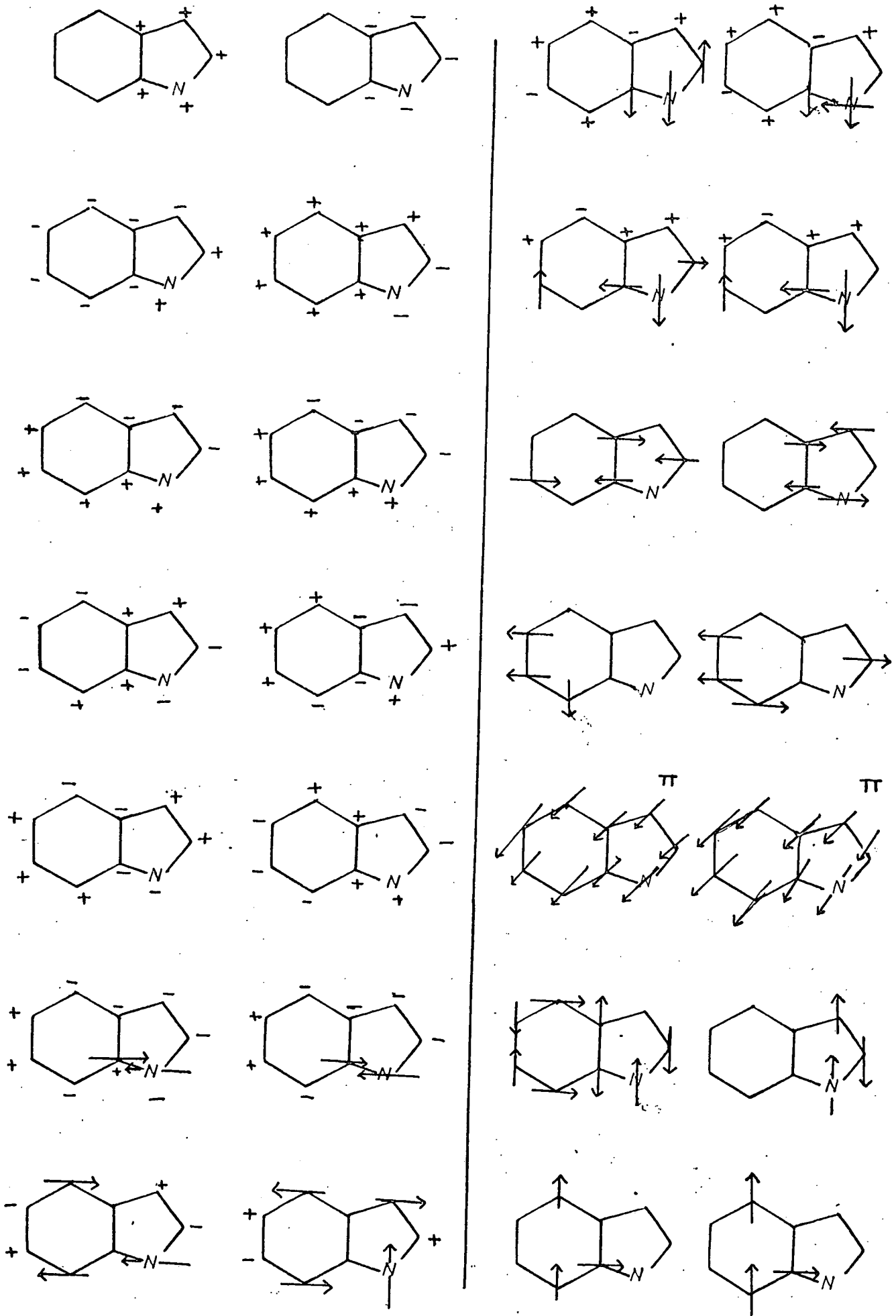


Fig 1 (contd)

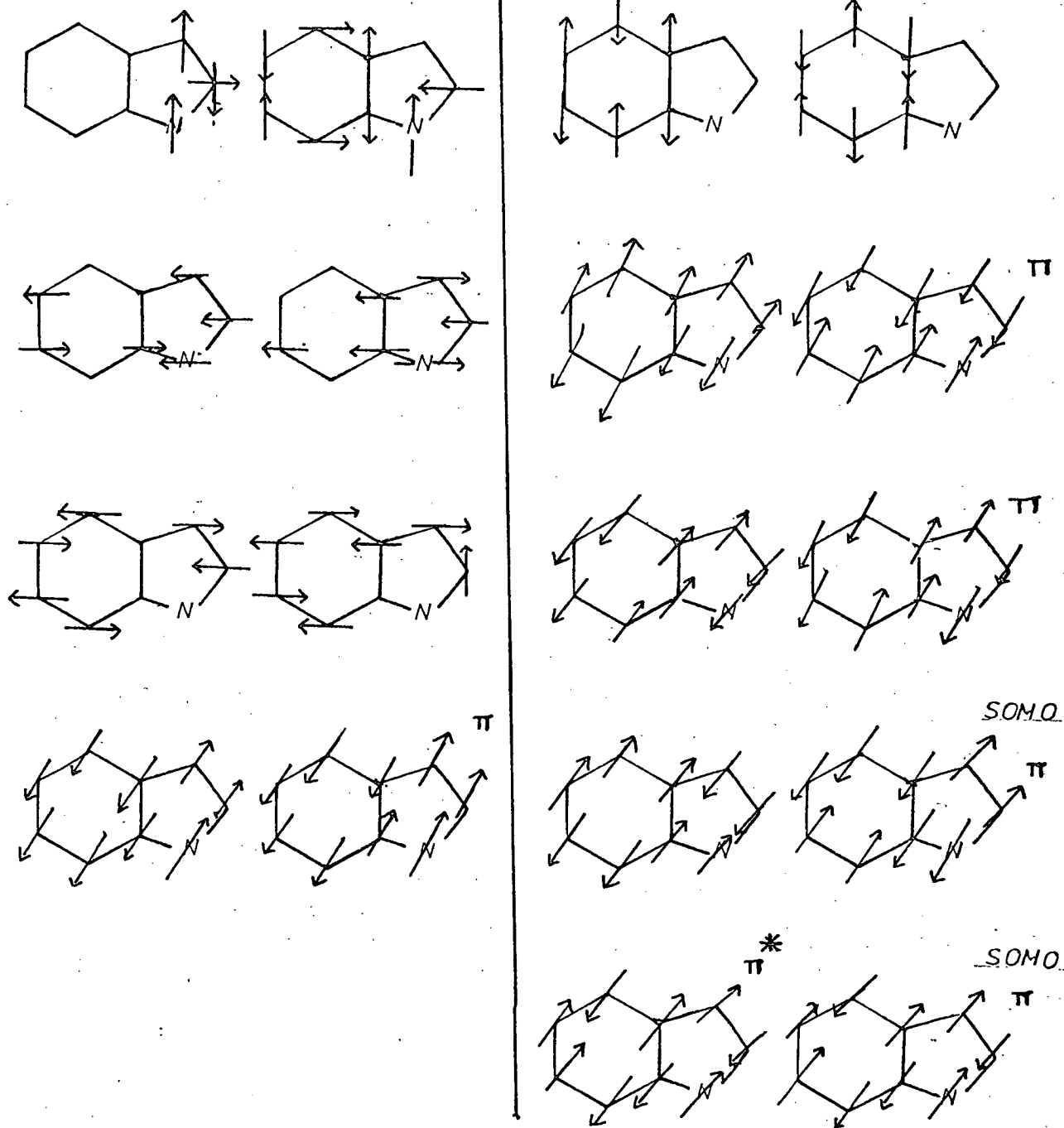
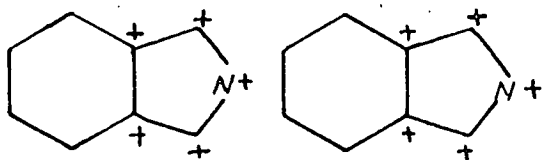


Fig.1 (contd.)

closed

open



closed

open

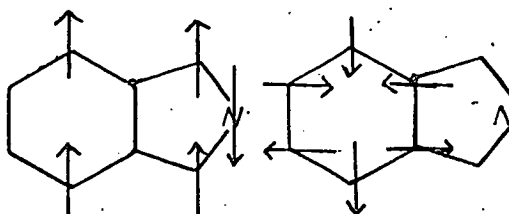
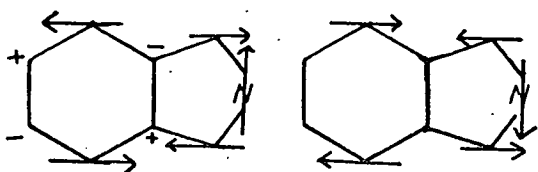
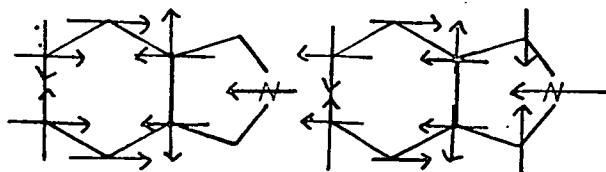
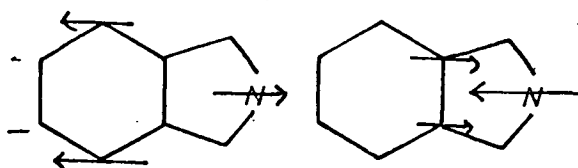
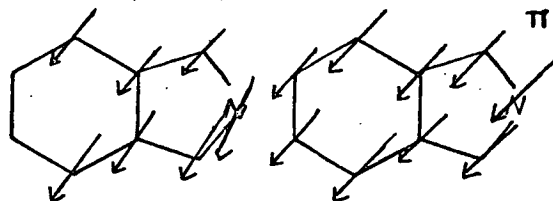
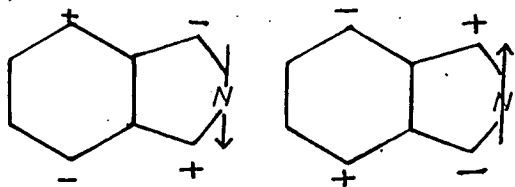
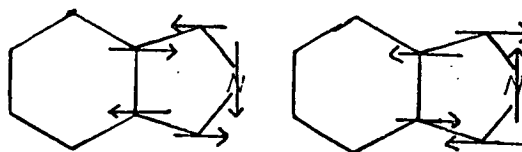
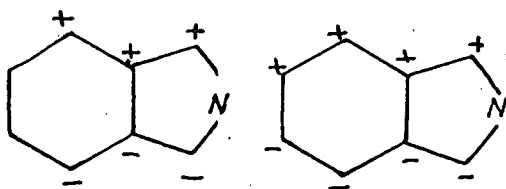
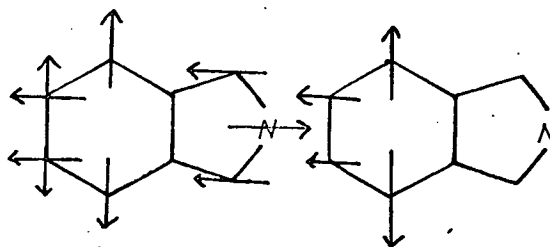
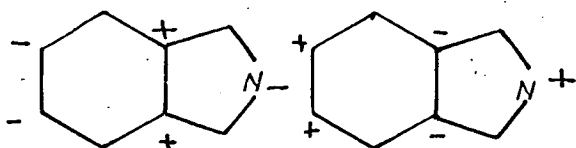
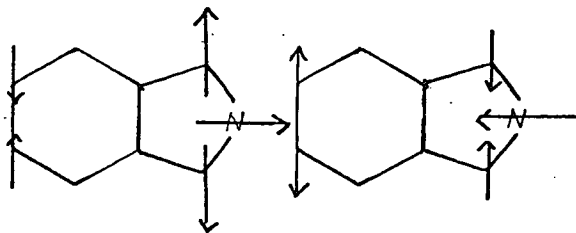
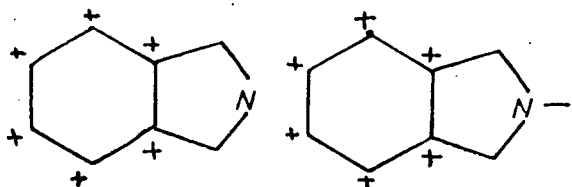
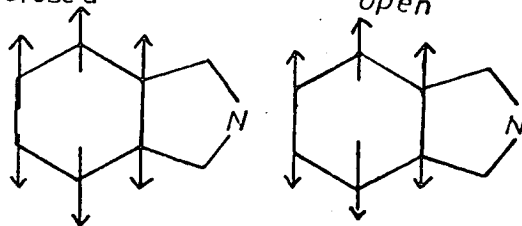


Fig.1 (contd.)

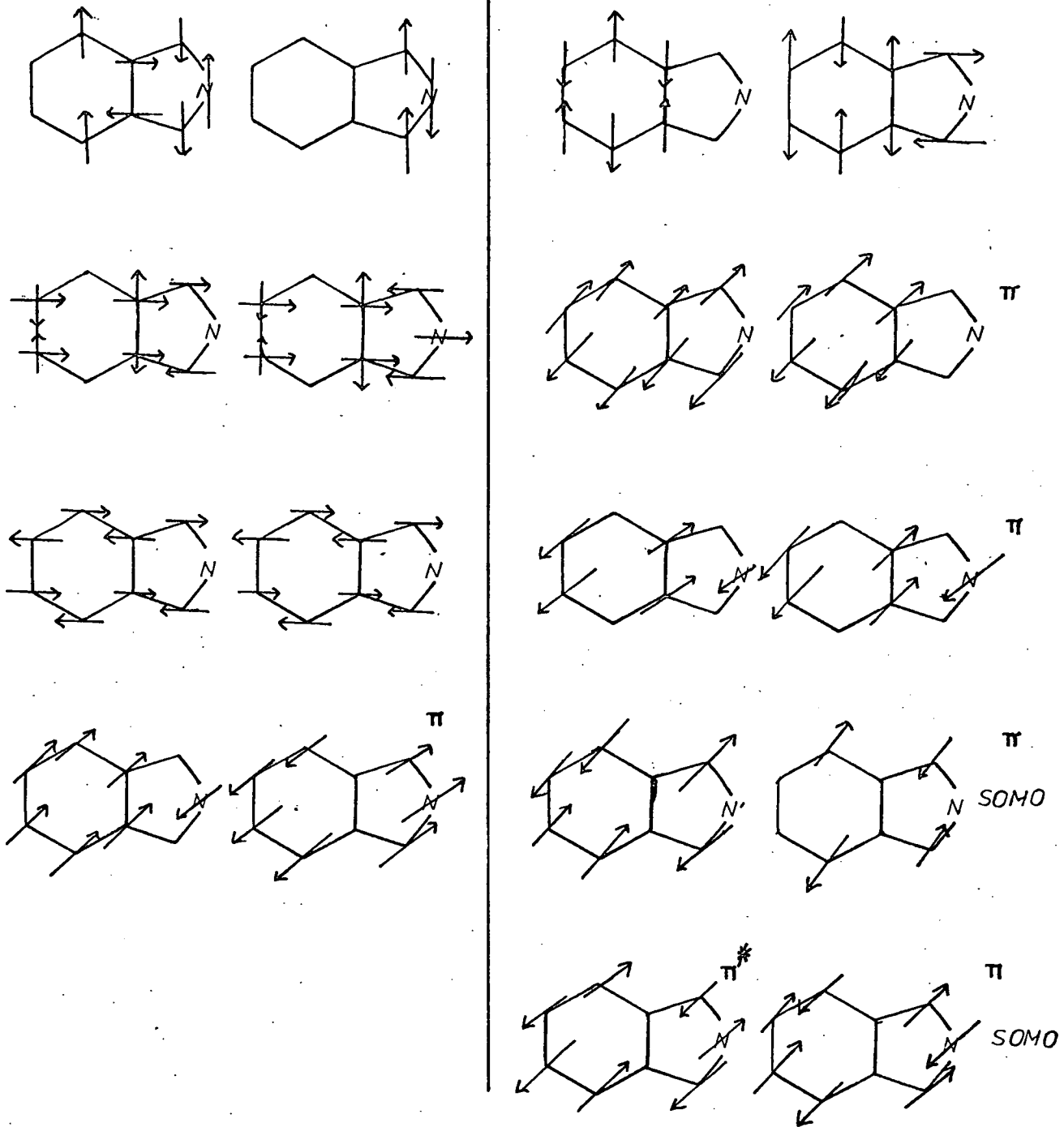
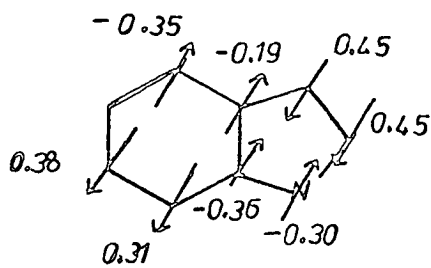


Fig. 2

HOMO  $\pi$



HOMO  $\pi^*$

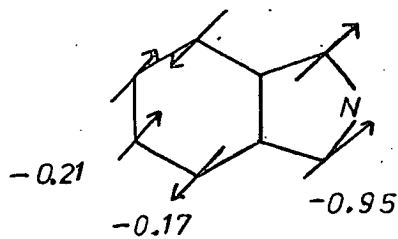
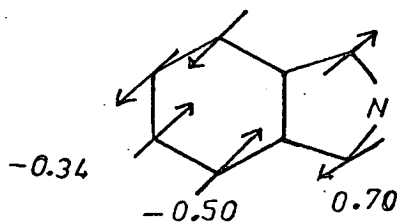
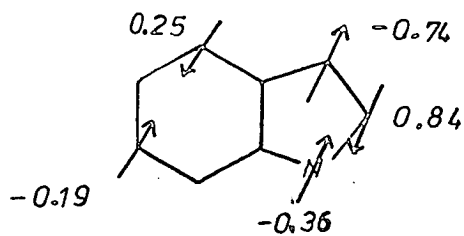
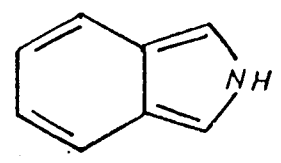
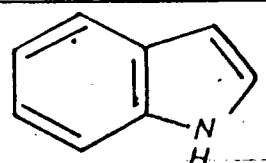
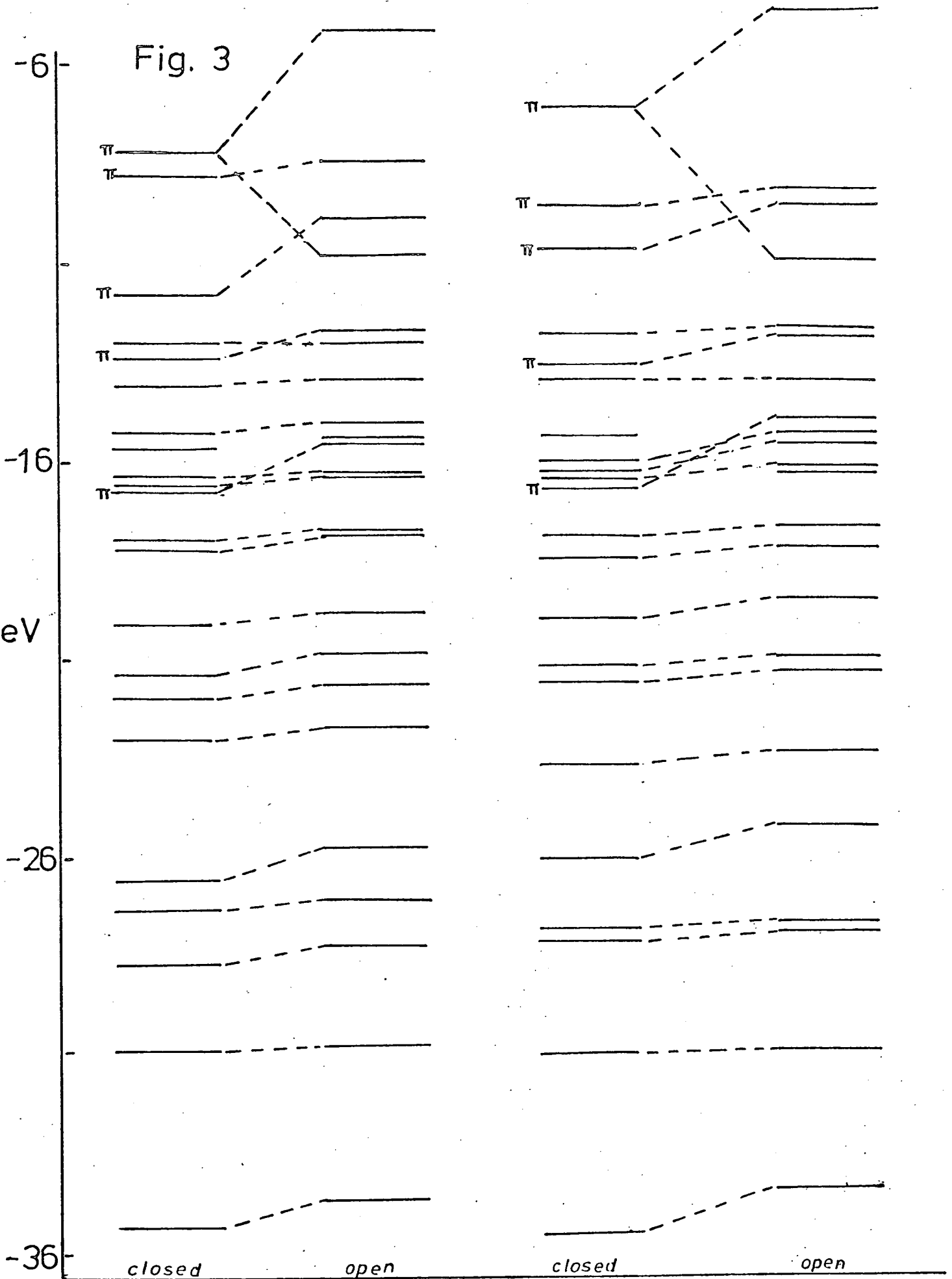


Fig. 3

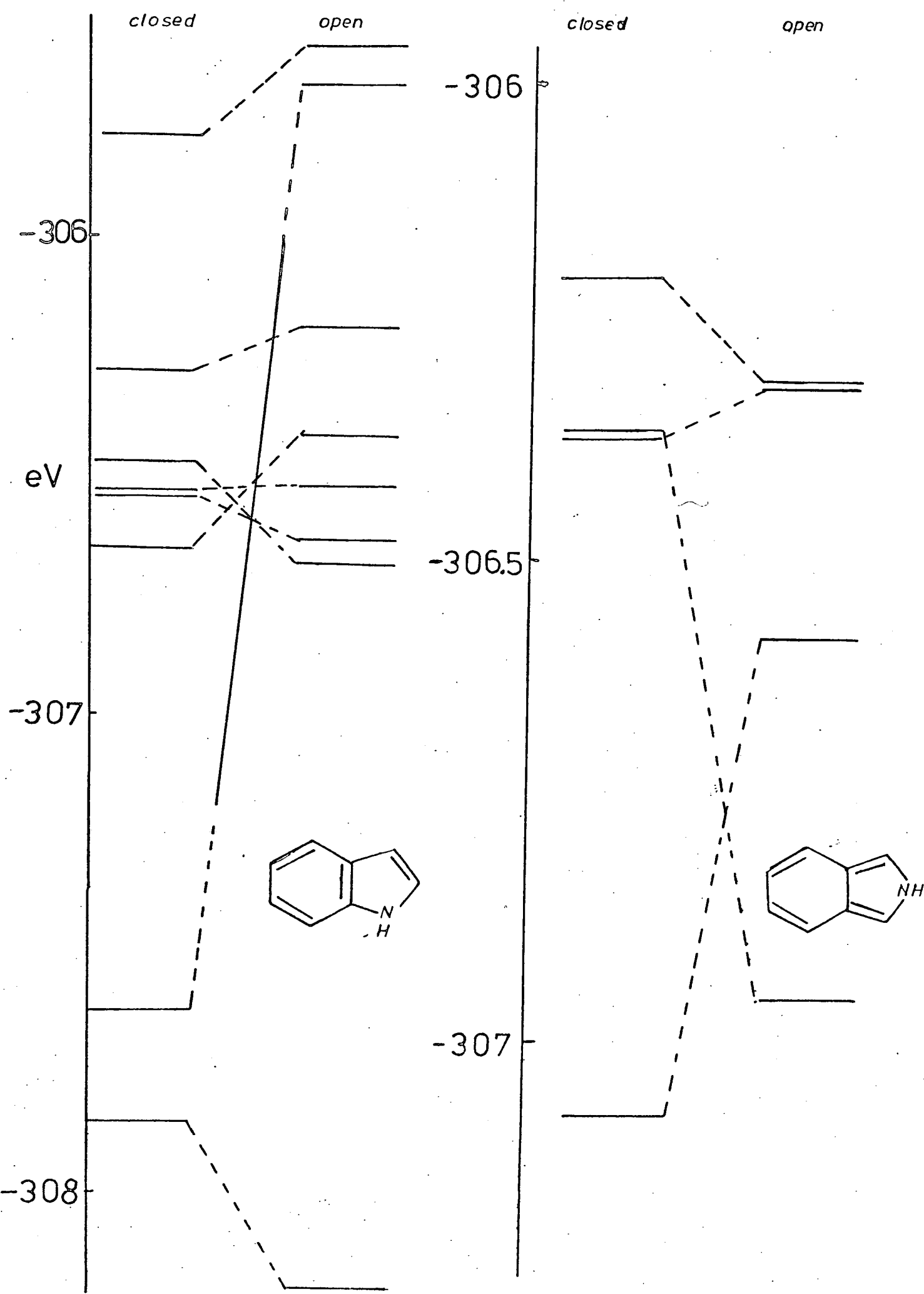




Appendix 2. Comparing RHF closed and open shell results shows a large scale destabilisation in the open shell ( $>1\text{eV}$ ), much of which can be explained by considering the nature of the highest occupied molecular orbital, HOMO  $\pi^*$ , in the open shell case, and that of the HOMO ( $\pi$ ) in the closed shell molecule. This is shown in Fig. 2 for indole and isoindole, the numbers indicating eigenvectors, and the arrows the direction of the  $p_\pi$  orbital concerned with reference to the molecular plane.

For both indole and isoindole there is a move of electron density from the hydrocarbon ring to the heterocyclic ring, concentrated at positions C(2), C(3) and N(1) in indole, and C(1) C(3) and N(2) in isoindole. The coulombic repulsion integral between the singly occupied molecular orbital, SOMO, and any other orbital in which these centres are dominant is likely to be increased causing destabilisation. In the valence shell where most of the orbitals are delocalised a general destabilisation is encountered. In indole, the most affected orbitals in the valence shell are the  $1a''$  and  $3a''$   $\pi$ -orbitals, the closed shell orbital energies being  $-16.73$  and  $-15.53$ , and the open shell  $-11.08$  and  $-9.91\text{eV}$  respectively. The  $1a''$  orbital is concentrated on  $N(p_\pi)$  and  $3a''$  on  $N(p_\pi) + C3(p_\pi)$ . In isoindole, the most affected orbitals are the  $1b_1$   $\pi$ -orbital centred on nitrogen  $p_\pi$ , and the  $6a_1$  orbital which is  $N(2s)/(C1 + C3)(2s)$ . The general destabilisation from closed to open shell molecular orbitals is shown in Fig.3 for isoindole and indole.

Fig. 4



This localisation of charge in the singly occupied  $\pi^*$  state has its most striking effect in the core orbitals where shifts of up to 2eV are witnessed. A core orbital energy diagram is shown for the same two molecules in Fig. 4 using RHF calculations. In indole the greatest shift is for the 3a' orbital which is the C2(1s) orbital, and shifts to lower binding energy by  $\approx 2\text{eV}$ . The only other molecular orbital which is markedly affected is the 2a' orbital which corresponds to the C3a(1s) bridging carbon atom 1s orbital. This moves to higher binding energy by  $\approx 0.5\text{eV}$ . The other 1s orbitals are also shifted in the open shell relative to the closed shell species but not so dramatically. In isoindole where the molecule has  $C_{2v}$  symmetry, the core orbitals are represented by linear combinations of the 1s orbitals across the long axis, each pair being degenerate. Again the two most affected orbitals are the corresponding ones to those in indole, the C(1s) orbitals  $\alpha$  to the hetero atom, and the bridging carbon orbitals. The  $\alpha$  C(1s) orbitals move to lower binding energy by  $\approx 0.5\text{eV}$  and the bridging carbon atoms to higher binding energy by the same amount. The core level shifts can be related to the change in electron density at the nuclei on excitation of an electron. If the  $\pi$ -electron densities at the affected positions are considered in the open and closed shell molecules, then it can be seen that in indole the  $\pi$ -populations at C(2) and C(3) arise almost entirely from the SOMOs centred at these positions, the contributions from the SOMOs being 0.8501e and 0.8177e, to a total of 1.1239e and 0.9847e respectively.

This increased electron density will increase the electron repulsion and result in a decrease in binding energy. The C(3a) position decreases in electron density in going from the closed to the open shell molecule, and results in the opposite effect of moving the core orbitals to higher binding energy. A parallel situation occurs at the bridging and C<sub>α</sub> positions in isoindole where the former π-density goes from 1.0725e to 1.0092e and the latter shows a contribution of 0.7515e from the SOMOs centred at these positions to a total of 1.0506e.

It is well known from XPS data that changes in the valence shell affect the core levels, and here the effect is quite marked especially for the Kekulé compounds.

Population analyses of the triplet states of the bicyclic heterocycles show an increase in population at the heteroatom, this increase arising from the π-system. This is not surprising as the triplet state is associated with the π-HOMO, and the π-LUMO. However, the total π-populations at each centre show only small differences between open and closed shell molecules. The populations for indole and isoindole, open and closed are compared in Table 2.

It is only in the past six years that the quinoid compounds have become synthetically available on a workable scale, but they are still relatively difficult to handle because of autoxidation and ease of polymerisation. It is obviously partly for these reasons that they remained illusive until the early 1970's although N-substituted isoindoles and benzo(c)thiophens were cited in the literature as early as

Table 2 Populations For Indole and Isoindole, Open and Closed Shell Cases

Indole - Closed

	N(1)	C(2)	C(3)	C(3a)	C(4)	C(5)	C(6)	C(7)	C(7a)
$\sigma$	5.7891	4.9730	5.0847	4.9914	5.1465	5.1283	5.1421	5.1147	4.8486
$\pi$	1.6585	1.0494	1.1057	1.0562	0.9949	1.0356	1.0087	1.0498	1.0409
Total	7.4476	6.0224	6.1904	6.0476	6.1414	6.1639	6.1508	6.1645	5.8895
	H(1)	H(2)	H(3)	H(4)	H(5)	H(6)	H(7)		
	0.6656	0.8386	0.8576	0.8526	0.8564	0.8531	0.8579		

Open

	N(1)	C(2)	C(3)	C(3a)	C(4)	C(5)	C(6)	C(7)	C(7a)
$\sigma$	5.7285	4.9142	5.1407	4.9892	5.1425	5.1189	5.1560	5.0889	4.9142
$\pi$	1.7498	1.1239	0.9847	1.0644	1.0001	1.0514	0.9845	1.0948	0.9464
Total	7.4783	6.0381	6.1254	6.0536	6.1426	6.1703	6.1405	6.1837	5.8606
	H(1)	H(2)	H(3)	H(4)	H(5)	H(6)	H(7)		
	0.6777	0.8543	0.8534	0.8534	0.8576	0.8517	0.8590		

Table 2 (cont'd)

## Isoindole - Closed

	N(2)	C(1)	C(3a)	C(4)	C(5)	H(2)	H(1)	H(4)	H(5)
$\sigma$	5.8695	4.9093	4.9880	5.1390	5.1362	0.6570	0.8498	0.8569	0.8576
$\pi$	1.5522	1.1261	1.0725	1.0017	1.0236				
Total	7.4217	6.0354	6.0605	6.1407	6.1598				

## Open

	N(2)	C(1)	C(3a)	C(4)	C(5)	H(2)	H(1)	H(4)	H(5)
$\sigma$	5.7085	4.9568	5.0252	5.1196	5.1396	0.6839	0.8490	0.8575	0.8561
$\pi$	1.7813	1.0506	1.0092	1.0345	1.0150				
Total	7.4898	6.0074	6.0344	6.1541	6.1546				

1963 (see Experimental section for details) Also many polyC-alkyl and C-aryl substituted compounds of the quinoid series were known. This difficulty in synthesis was postulated to be due to their lack of aromatic character<sup>6</sup>, and even anti-aromatic character in some cases, producing an unstable ground state. It is true that on the basis of REs that the quinoid series shows destabilisation relative to the Kekulé series, but all are shown to have some degree of aromatic character. Benzo(c)thiophen, and isoindole are calculated to have more aromatic character than benzo-furan on the basis of RE. It would seem that the RE results do not wholly agree with the practical results. It has been noted already that if the first excited triplet state of the Kekulé series is considered, then the UHF calculations (Table 1) give the best numerical results for excitation energies but are in general  $\approx 0.5\text{eV}$  too high. If it is assumed that a similar situation will exist for the quinoid series then it can be seen that, for benzo(c)thiophen and isoindole the excitation energy is of the order 1.5 - 2.0eV, and that for benzo(c)furan and 2H-indene, the excitation energy will be extremely small. Of the four compounds, benzo(c)thiophen is the most stable when prepared, then N-methylisoindole, followed by benzo(c)furan and last last 2H-indene which has never been isolated, except in the presence of metal complexing agents such as  $\text{Fe}_2\text{CO}_9$ <sup>13</sup>. This order is in agreement with excitation energies, and it is perhaps, therefore, this low excitation energy which is making the compounds appear

Fig.5

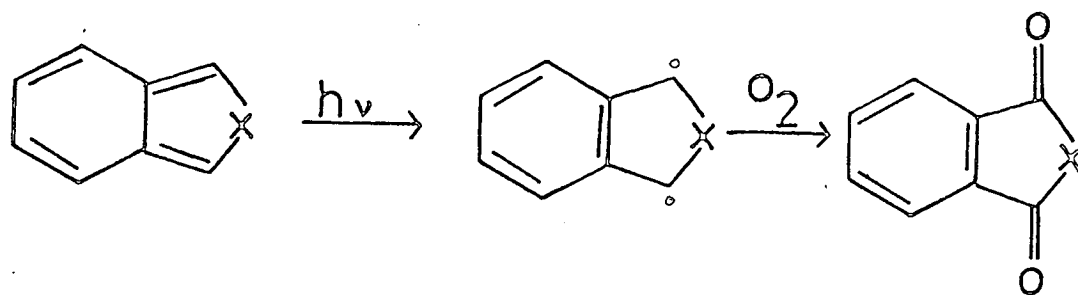


Fig.6

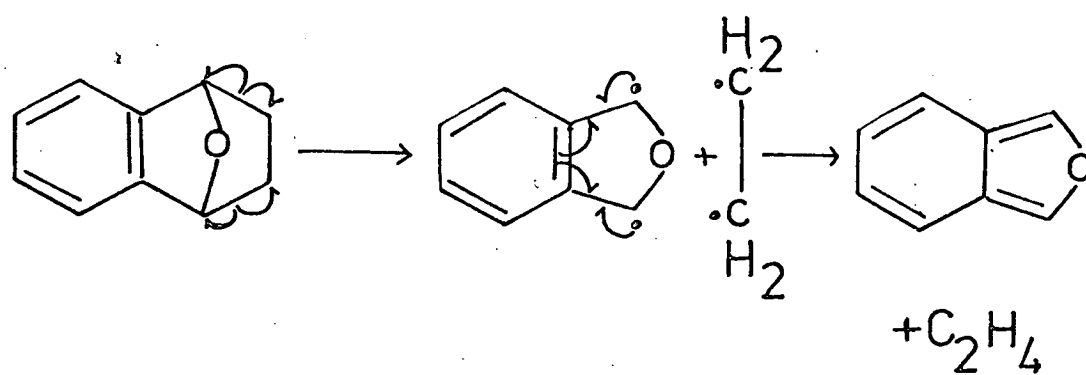
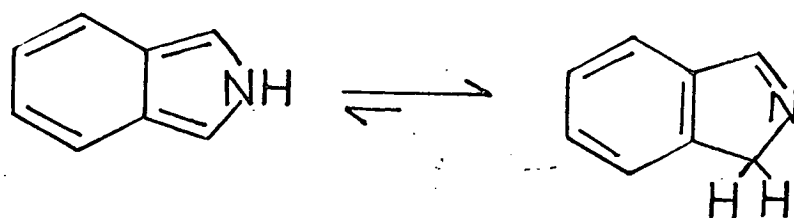


Fig. 7



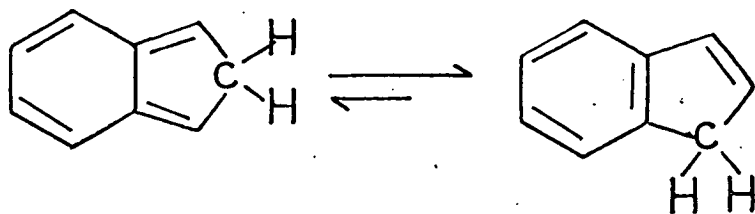


unstable, rather than instability of the ground state itself. This low excitation energy makes it possible for reactions such as autoxidation to take place merely in the presence of light, visible light being a source of energy in the region  $\approx 1.74 - 3.10\text{eV}$ . Autoxidation occurs in the 1,3 positions<sup>7-9</sup>, and therefore possibly via the triplet state which would produce the situation in Fig. 5. The autoxidation of benzo(c)furan<sup>8,9</sup> is the best documented of these reactions, which would be expected if the reaction proceeds via a low lying triplet state, as the excitation energy is extremely small, and the reaction will occur very easily. Its formation can also be induced by a retro Diels Alder reaction<sup>10</sup> which can also be explained in terms of a triplet state, Fig. 6.

Isoindole is also a fairly illusive compound but was reported as early as 1964<sup>11</sup> as being sensitive to oxidation and stable for only a few days under pure  $\text{N}_2$ . This can perhaps be attributed to the possible tautomeric forms which can exist, and which have been detected in solution<sup>12</sup>, Fig. 7. It is likely that tautomer B, Fig. 7 will be preferred as it has a benzenoid ring. N-methyl isoindole is much more accessible, no tautomeric equilibrium being available to it, and is less reactive than benzo(c)furan, as the excitation energies would suggest. Polyalkyl substituted isoindoles will also undergo autoxidation, the rate being up to several hours depending on the substituents, but the affected positions again being 1,3.

Benzo(c)thiophen is least reactive possibly because of a slightly higher excitation energy, and a lack of tautomeric equilibrium to remove the quinoid form. The compound 2H-indene is predicted on this basis to be extremely difficult to isolate due to a low excitation energy, and the possibility of a tautomeric equilibrium, Fig. 8, producing indene as the more stable tautomer. Hence, the need for a metal complexing agent<sup>13</sup> to prevent the transition from quinoid to Kekulé species taking place. The "instability" of these quinoid compounds is perhaps more effectively explained in terms of their ease of transition to a low lying triplet state, making them extremely reactive compounds. All of them possess a reasonable degree of aromatic character, and the calculations suggest, that in the absence of this low lying triplet state, that they would form stable compounds.

Fig.8



REFERENCES

1. H.B. Klevens and J.R. Platt, J. Chem. Phys., 1949, 17, 470
2. G.S. Kembrowskii, V.P. Bobrovich and S.V. Koneve, Zhur. Priklad. Spektroskopi, 1966, 5, 695.
- 3(a) R.C. Heckman, J. Mol. Spectroscopy, 1958, 2, 27  
(b) P.-S. Song and W.E. Kurtin, J. Amer. Chem. Soc., 1969, 91, 4892.
4. D.F. Evans, J. Chem. Soc., 1959, 2753..
5. M.R. Padhye and J.C. Patel, J. Sci. Ind. Res., India, 1956, 15B, 171, 206.
6. M.J.S. Dewar, A.J. Harget, N. Trinajstic and S.D. Worley, Tetrahedron, 1970, 26 (19), 4505.
7. L.J. Kricka and J.M. Vernon, J. Chem. Soc. (c), 1971, 2667.
8. R.H. Young and D.T. Feriozi, Chem. Comm, 1972, 841.
9. J. Olmsted and T. Akashah, J. Amer. Chem. Soc., 1973, 95, 621.
10. H.J. Hageman and U.E. Wiersum, Chemistry in Britain, 1973, 9 (5), 206.
11. J. Thésing, W. Schäfer and D. Melchior, Annalen, 1964, 671, 119.
12. R. Kreher and J. Seubert, Z. Naturforsch, 1965, 20b, 75.
13. W.R. Roth and J.D. Meier, Tetrahedron Letters, 1967, 2053.

### Hückel $\alpha$ and $\beta$ parameters

In HMO theory, the secular determinant for an alternant hydrocarbon of  $n$  atoms has  $n$  real roots of the form:-

$$\frac{\alpha - E_j}{\beta} = \pm m_{ij} \quad \text{where } j = 1 \rightarrow n$$

$$\text{or } E_j = \alpha \pm m_j \beta$$

The  $\pi$ -energy levels of the carbon atoms can therefore be represented as pairs of energy levels symmetrically placed above and below an energy zero taken as  $\alpha$ . Where a closed shell molecule is being considered, the electrons will be paired in the highest binding energy orbitals which will be given by:-

$$E_j = \alpha + m_j \beta$$

For linear polycyclic acenes,<sup>1</sup> the roots  $m_j$  are given by:-

$$m_j = \pm 1, \frac{1}{2} (1 \pm \sqrt{9 + 8 \cos \frac{j\pi}{n+1}})$$

where  $n$  is the number of rings in the system, and  $j = 1 \rightarrow n$ .

Using this expression, values of  $m_j$  can be obtained, and substituting orbital energies,  $e_i$ , found from LCGO calculations for  $E_j$ , in the expression  $E_j = \alpha + m_j \beta$  for the hydrocarbon series benzene to pentacene yields the relationship  $e_i = 4.228M + 5.61eV$  with standard deviations in slope, intercept and overall of 0.077, 0.118 and 0.339 respectively. The graph of this data is shown in Fig. 1, and the plotted data in Table 1. The HMO method tends to show a greater number degeneracies<sup>2</sup> than an SCF method, and these often correspond to  $M = 1.000$ . If these points are removed from the graph, the overall standard deviation improves to 0.270 while the slope (4.178) and intercept (5.729) remain virtually unchanged. The correlation is good, and it should be practical to use the above relationship to estimate LCGO  $\pi$ -orbital energies for

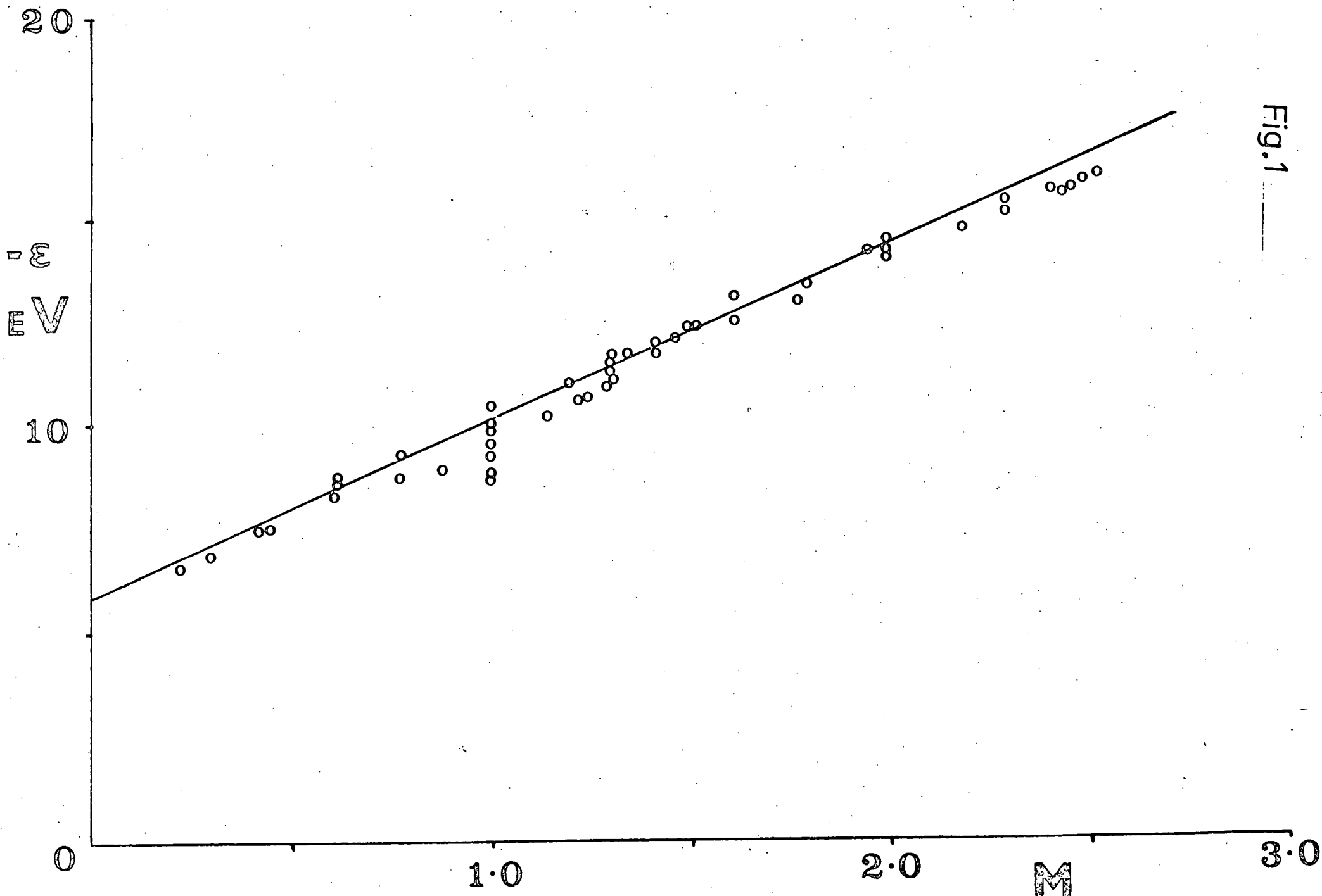


Fig. 1

Table 1. LCGO orbital energies -  $e_i$  (eV) and the corresponding  $m$  values.

	$-e_i$	$m$		$-e_i$	$m$
Benzene	9.82	1.00	Tetracene	6.86	1.180
	14.47	2.00		8.78	3.200
Naphthalene	8.57	0.618	9.25	3.920	
	9.47	1.00	10.88	5.870	
	11.29	1.302	10.99	5.880	
	13.09	1.618	12.07	7.290	
	15.48	2.302	12.98	8.820	
Anthracene	7.42	0.414	14.73	11.500	
	9.20	1.000	15.76	13.310	
	10.00	1.000	Pentacene	6.58	0.710
	11.70	1.414	8.65	2.300	
	11.96	1.414	8.70	3.060	
	14.22	2.000	10.40	5.150	
	15.70	2.414	10.56	5.290	
			11.64	6.800	
		12.37	7.660		
		12.49	7.920		
		14.06	10.320		
		15.19	12.290		
		15.95	13.570		

Table 2.  $-IP_1$  and  $m_1$  values for benzene to pentacene.

	$-IP_1$ (eV) <sup>7</sup>	$m$
Benzene	9.24	1.00
Naphthalene	8.15	0.618
Anthracene	7.40	0.414
Tetracene	7.01	0.290
Pentacene	6.64	0.212

polycyclic alternant hydrocarbons. It is unlikely that non-alternant hydrocarbons or Kekulé heterocycles would yield such a good correlation, especially the latter since the introduction of a heteroatom requires the addition of arbitrary parameters in the HMO calculations.

The slope and intercept of the above linear correlation yield values for the coulomb ( $\alpha$ ) and resonance ( $\beta$ ) integrals. Values for  $\alpha$  and  $\beta$  have been determined by several methods.<sup>3</sup> One method, which has received much attention, is the estimation of the  $\alpha$  and  $\beta$  integrals by equating them with the first IP of a molecule using the relationship:-

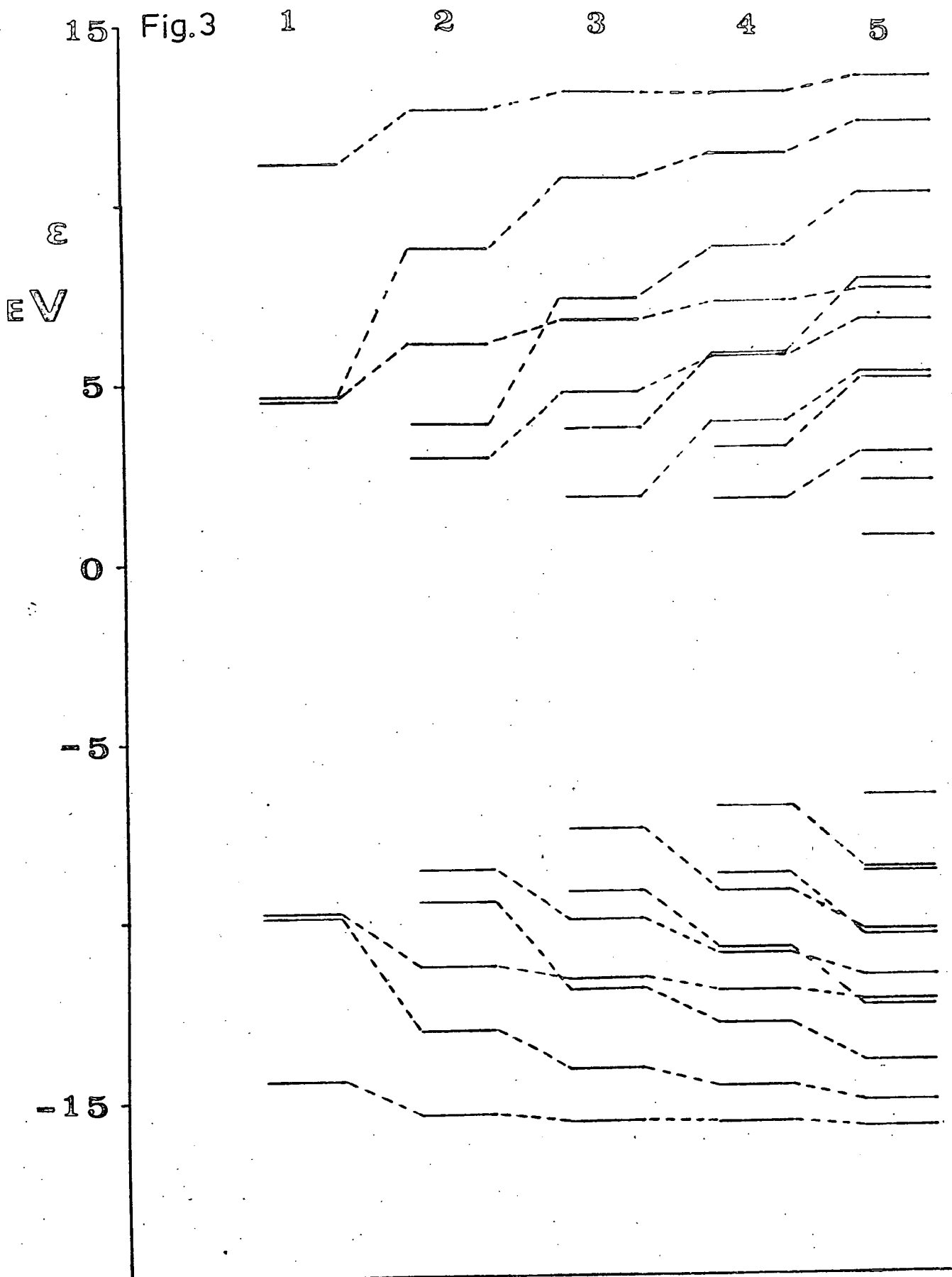
$$IP_1 = -(\alpha + m_1\beta)$$

where  $\alpha$  and  $\beta$  are the HMO integrals. For the series of polycyclic acenes, benzene to pentacene values of 7.07<sup>4</sup> and 6.24<sup>5</sup> for  $\alpha$  and 2.48<sup>4</sup> and 3.14 eV<sup>5</sup> for  $\beta$  are obtained. The  $\alpha$  value of 5.61 eV obtained here is in reasonable agreement, but the  $\beta$  value of 4.23 eV is rather different. However if a graph, Fig. 2, of  $-IP_1$  is plotted against the  $m$  values of the lowest binding energy occupied  $\pi$ -orbital, Table 2,  $\alpha$  is given as 6.15 eV and  $\beta$  as 3.08 eV indicating that the different  $\beta$  values are due to the incorporation of the higher binding energy IPs, as given by the LCGO calculations.

More recent work<sup>2, 6-8</sup> extends the correlation to include the first three or four IPs for molecules under consideration, and one such study on polycyclic hydrocarbons<sup>2</sup> gives  $\alpha$  as 5.937 and  $\beta$  as 2.944 eV. The linear relationship of  $e_1$  to  $m$  was obtained in reference 2 using observed IPs. This partly explains the very different  $\beta$  value to that obtained here. The work of Heilbronner et al<sup>7</sup> has been extended to include a correction in the calculated orbital energies to account for the change in equilibrium bond lengths on ionisation. Using this correction and HMO orbital energies,  $\alpha$  is given as 5.864 and  $\beta$  as 3.196 eV for the acenes, benzene to pentacene. This is closer to the values obtained here.

The HMO theory gives a symmetrical placing of occupied and virtual

Fig.3





$\pi$ -orbitals about an energy level,  $\alpha$ , and the average energies of an orbital pair is given by:-

$$(\alpha + \beta + \alpha - \beta)/2 = \alpha$$

The LCGO calculations show a spacing in occupied  $\pi$  and virtual orbital energy levels close to a mirror image, Fig. 3, with the symmetry axis at  $\approx 2.4$  eV. Omitting the highest occupied and lowest unoccupied pair of  $\pi$ -levels, the others obey the relationship,  $E = 1.00 m - 3.70$  eV. The total energy of each pair is therefore not constant using the LCGO results as it is using the Hückel method (2a).

#### References

1. C. A. Coulson and A. Streitwieser, "Dictionary of  $\pi$ -electron calculations," Pergamon Press.
2. R. Boschi, J. N. Murrell and W. Schmidt, Disc. Faraday Soc., 1972, 54, 116.
3. A. Streitwieser, "Molecular Orbital Theory for Organic Chemists," John Wiley and Sons Inc., New York and London.
4. D. P. Stevenson, private communication cited by ref. 3.
5. M. E. Wacks and V. H. Dibeler, J. Chem. Phys., 1959, 31, 1557.
6. J. H. D. Eland and C. J. Danby, Z. Naturforsch., 1968, 23a, 355.
7. P. A. Clark, F. Brogli and E. Heilbronner, Helv. Chim. Acta, 1972, 55, 1415.
8. R. Boschi and W. Schmidt, Tetrahedron Letters, 1972, 25, 2577.

Fig.1

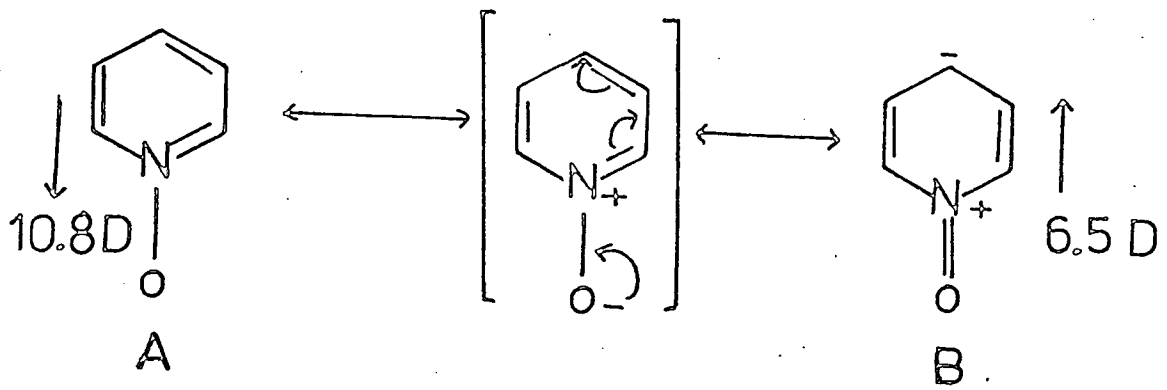


Fig.2

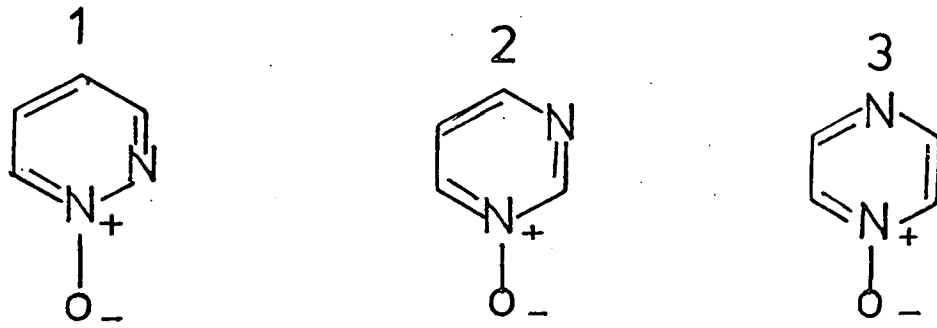
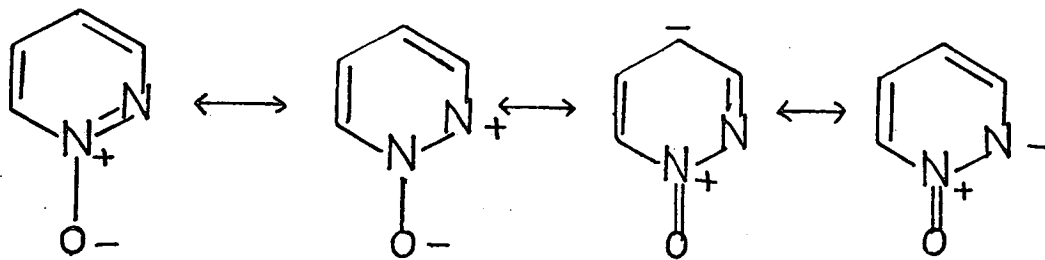


Fig.3



## Some Heterocyclic N-Oxides

The simplest heterocyclic N-oxide, pyridine N-oxide, has been studied extensively in the field of electronic absorption,<sup>1-6</sup> but in all cases the calculations which have been used for comparison were semi-empirical at best, and results in certain cases have been shown to be very sensitive to the parameters used.<sup>7</sup> Recently, interest has been shown in the PE spectrum of pyridine N-oxide,<sup>8,9</sup> but again assignment was made using the semi-empirical MINDO 2 method, and only the early bands were assigned. It therefore seemed appropriate to perform an LCGO calculation on the pyridine N-oxide system.

A further need for a good LCGO calculation on the system was brought about from studying the remarkably different reactions of pyridine and pyridine N-oxide. For example, pyridine undergoes electrophilic attack in the 3-position but only with great difficulty, whereas the N-oxide readily undergoes electrophilic substitution, the position depending on whether attack is on the protonated or unprotonated molecule. Attack on the unprotonated molecule yields the 2- or 4-substituted N-oxide but preference is often shown for the 4-position, probably due to the steric hindrance of the N-O group to the incoming electrophile to the 2-position. Nitration of pyridine N-oxide gives almost exclusively 4-nitropyridine N-oxide. The positive nature of the nitrogen atom also probably helps to direct the electrophile to the 4-position. However, mercuration of pyridine N-oxide under fairly mild conditions yields 2-acetoxymercuri pyridine N-oxide, while under strongly acid conditions 3-acetoxymercuri pyridine N-oxide is obtained probably via a protonated pyridine N-oxide molecule.

Nucleophilic attack on both pyridine and pyridine N-oxide occurs relatively easily in either the 2- or 4-position with a slight preference for the 2-position. Early studies of the dipole moment<sup>10</sup> suggested that these differing reactions could be the result of charge transfer as shown in Fig. 1. The extent to which B, Fig. 1, contributes will be a major factor

in determining the ease of electrophilic substitution, and could be estimated from an LCGO calculation. This would serve to shed light on the substitution process.

The phenomenon of increased reactivity at the atom para to the N  $\rightarrow$  O bond has also been noted in pyridazine N-oxide,<sup>11</sup> 1, Fig. 2, and should also occur in pyrimidine N-oxide, 2, Fig. 2. Semi-empirical methods have already indicated that the electron density at the para position to the N  $\rightarrow$  O is higher than at other positions. This could again be due to the electron transfer increasing the number of resonance forms available to the molecule as shown for 1 in Fig. 3. In the absence of a good calculation to throw light on this subject, it was decided to extend the LCGO calculations to include the diazine N-oxides.

At the time the work commenced no published structure existed for pyridine N-oxide, and earlier semi-empirical work had chosen the ring structure as a regular hexagon with an N-O bond length of 1.28 $\overset{\circ}{\text{A}}$ .<sup>5</sup> Some early data on substituted N-oxides,<sup>12</sup> and on the pyridinium ion<sup>12</sup> also existed, and in each case the ring geometry was found to be similar to that of pyridine. For this reason it was decided to adopt the ring structure of pyridine and to optimise the N-O length.

Molecular geometries. An earlier calculation on pyridine had been performed by the M. H. Palmer group<sup>f</sup> using an unscaled basis set, and was reused as the starting point for the pyridine N-oxide calculation. The N-O bond length was found to be optimal at 1.35 $\overset{\circ}{\text{A}}$  using CNDO calculations for optimisation.

A recent electron diffraction study of pyridine N-oxide<sup>14</sup> has shown the N-O length to be 1.29 $\overset{\circ}{\text{A}}$ . When a series of compounds containing N-O coordinate bonds are considered, a value of 1.35 $\overset{\circ}{\text{A}}$  does not in fact appear too anomalous as shown in Table 1. It would appear that as the degree

---

<sup>f</sup> Thanks are due to A. J. Gaskell who performed the calculations on unscaled pyridine and unscaled pyridine N-oxide.

Table 1. Comparison of N-O bond lengths

<u>Compound</u>	<u>N-O (Å)</u>
HCNO	1.199 <sup>15</sup>
N <sub>2</sub> O <sub>4</sub>	1.180 <sup>16</sup>
HNO <sub>3</sub>	1.218 <sup>17</sup>
Phenazine N-oxide	1.24 <sup>12</sup>
4. Nitropyridine N-oxide	1.26 <sup>12</sup>
4,4'-Trans-azopyridine, 1,1'-dioxide	1.283 <sup>12</sup>
2-Hydroxy methyl pyridine N-oxide	1.321 <sup>18</sup>
(CH <sub>3</sub> ) <sub>3</sub> NO	1.338 <sup>19</sup>

Table 2 Optimisation of the N-O bond length in pyrazine N-oxide using INDO calculations.

<u>N-O Length (Å)</u>	1.20	1.23	1.25	1.28	1.30
Total Energy (a. u.)	-70.21709	-70.23588	-70.23761	-70.23543	-70.23129
Optimum N-O length =	1.2536Å		T. E. = -70.23765 a. u.		

Table 3 Optimisation of the C-N length in pyrazine N-oxide using INDO calculations.

<u>C-N (Å)</u>	1.319	1.334	1.350	1.370
Total energy (a. u.)	-70.23510	-70.23765	-70.23539	-70.22558
Optimum C-N =	1.335Å			

Fig. 4

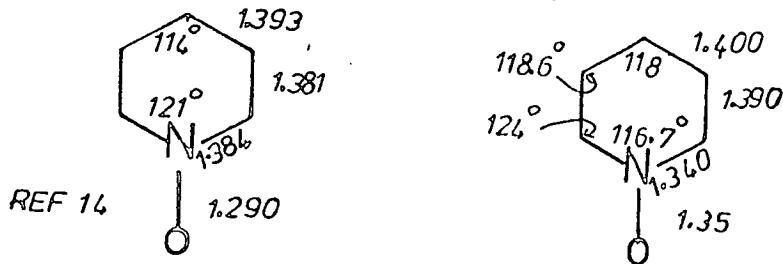


Fig. 5

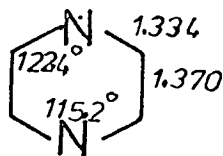


Fig. 6

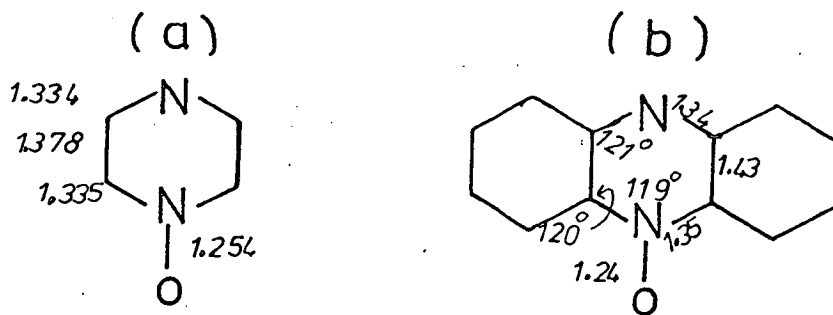


Fig. 7

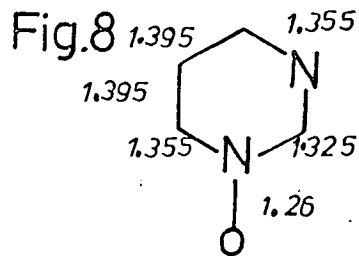
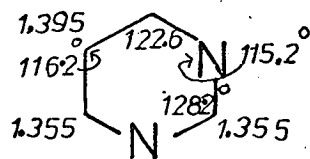


Fig. 9

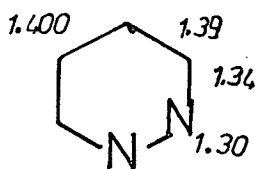
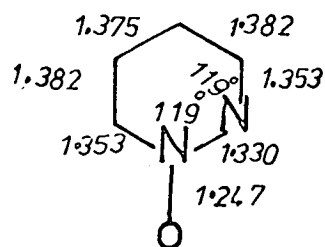


Fig. 10



of electron donation to carbon decreases, the N-O length increases, giving  $(\text{CH}_3)_3\text{NO}$ , where the methyl groups are electron donating, and will not accept electrons, a long N-O length. On this basis, it would be expected that the N-O length in pyridine N-oxide would be shorter than that in  $(\text{CH}_3)_3\text{NO}$  if a resonance such as that in Fig. 1 exists. However, the 2-hydroxy methyl pyridine N-oxide shown in Table 1 has a long N-O length, but this can perhaps be attributed to hydrogen bonding between the hydrogen of the OH and the oxygen of the N-O lengthening the N-O bond relative to that in the parent molecule. The geometries of pyridine N-oxide from reference 14 and that of pyridine + O are shown in Fig. 4. In the ring itself, the largest discrepancy is in the C-N length. Perhaps in making the ring too pyridine like the N and O have been unable to interact sufficiently, tending to a dissociation of the N and O atoms, and consequently a long N-O length.

No geometries exist for the diazine N-oxides, and again the problem of selecting a suitable geometry arises. To avoid the situation which had arisen with pyridine N-oxide, it was decided to attempt to optimise the ring geometry as well as the N-O bond length. The first molecule to be considered was pyrazine N-oxide, 3, Fig. 2, and the crystal structure<sup>20</sup> of the parent compound was used as the geometry basis, Fig. 5. The N-O length was then optimised using INDO calculations, the results of which are shown in Table 2. The optimum N-O length was found as the parabolic minimum to be  $1.254\text{\AA}$ . The next stage was to vary the ring geometry, keeping the N-O length fixed at  $1.254\text{\AA}$ . Comparing the geometries in Fig. 4 indicates that only the bonds adjacent to the N-O bond are markedly affected in forming the N-oxide. Bearing this in mind, it was decided to vary only the C-N(1) bonds of the ring. This would of course also vary the C-C bonds to accommodate any changes in the C-N lengths. The INDO results are shown in Table 3, and the optimum geometry obtained in Fig. 6a. When this geometry is compared with that of phenazine N-oxide,<sup>12</sup> Fig. 6b, which is a related structure, then the N-O lengths are in good agreement. The N-C(1) and C(1)-C(2) bonds are longer in phenazine N-oxide but this could be due to the molecule

having to adjust to accommodate the carbocyclic rings. The geometry in Fig. 6a was used in the LCGO calculation.

For pyrimidine N-oxide, 2, Fig. 2, the ring structure was taken as the crystal structure,<sup>21</sup> Fig. 7. The N-O and N(1)-C(2) bond lengths were then optimised as shown in Table 4, producing the optimum geometry shown in Fig. 8.

A similar procedure was used to construct a geometry for pyridazine N-oxide, 1, Fig. 2, using the crystal structure of the parent molecule,<sup>13</sup> Fig. 9, as the starting ring structure. The ring structure had already been optimised<sup>†</sup> leaving only the N-O bond to optimise as shown in Table 5, and the final geometry is given in Fig. 10. In all cases the N-O bond was set initially to bisect the external angle.

An additional effort to obtain an accurate representation for the molecules in the LCGO calculations was made by rescaling the oxygen basis set. The initial oxygen basis used was scaled using formaldoxime,  $\text{CH}_2=\text{NOH}$ , but it was felt that the coordinate N-O bond was a special case, and perhaps required a different basis set to represent it accurately. The scaling runs were performed using pyrazine N-oxide as the reference system, and the scaled formaldoxime basis set for oxygen was used initially. Representing the  $\text{O}_{2s}$  and  $\text{O}_{2p}$  exponents of the basis functions by the general formula  $\exp(-k\alpha r^2)$  where  $k$  is the scaling factor, yields the scaled formaldoxime values when  $k = 1$ . Table 6 shows the results. Using parabolic minimisation in two directions an optimum energy of  $-336.4371$  a. u. for  $k(\text{O}_{2s}) = 0.9728$ , and  $k(\text{O}_{2p}) = 0.9582$  was obtained, indicating that the molecular environment of oxygen in formaldoxime is remarkably similar to that in pyrazine N-oxide. Oxygen scaled pyrazine N-oxide was used in the LCGO calculations of the diazine N-oxides.

---

<sup>†</sup> Thanks are due to A. J. Gaskell for this information.



Table 4 Optimisation of N-O and N(1)-C(2) in pyrimidine N-oxide using INDO calculations.

<u>N-O (Å)</u>	1.23	1.25	1.28
Total energy (a. u.)	-70.23608	-70.23899	-70.23859
Optimum N-O =	1.260 Å		
<u>N(1)-C(2) (Å)</u>	1.300	1.326	1.355
Total energy (a. u.)	-70.22617	-70.24139	-70.23959
Optimum N(1)-C(2) =	1.325 Å		

---

Table 5 Optimisation of the N-O bond length in pyridazine N-oxide using INDO calculations.

<u>N-O (Å)</u>	1.215	1.250	1.280
Total energy (a. u.)	-70.31575	-70.31927	-70.31542
Optimum N-O =	1.247 Å	Total energy = -70.31931 a. u.	

---

Table 6. Scaling procedure for  $O_{2s}$  and  $O_{sp}$  using pyrazine N-oxide.

$k(O_{2s})$ \ $k(O_{2p})$		0.9	1.0	1.1
0.9	-336.42263357	-336.42841700	-336.41680859	
1.0	-336.43445099	-336.43581974	-336.42044481	
1.1	-	-336.41080413	-336.39242607	

Parabolic minimum  $k(O_{2p}) = 0.9728$   
 T. E. = -336.43701520 a. u.

Parabolic minimum  $k(O_{2s}) = 0.9582$   
 T. E. = -336.43728364 a. u.

Table 7. Total energies and dipole moments of the diazine N-oxides

	Total energy (a. u.)	Dipole Moment	
		Calculated	Experimental
Pyridine N-oxide scaled <sup>f</sup>	-320.57933	5.56	[ 4.24 <sup>22</sup> 4.19 <sup>7</sup>
" " unscaled	-320.29124	3.26	
Pyridazine N-oxide	-336.43064	5.12	5.21 <sup>5</sup>
Pyrimidine N-oxide	-336.44437	4.69	3.72 <sup>5</sup>
Pyrazine N-oxide	-336.43860	2.57	1.66

<sup>f</sup> Thanks are extended to W. Moyes and J. Nisbet for this information.

Fig. 10 a.

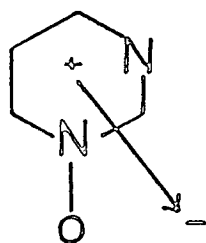


Fig10 b

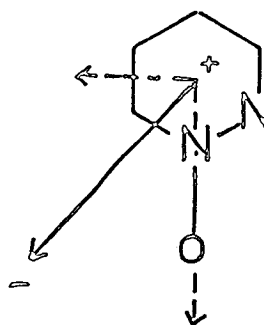


Fig11

Net  $\pi$ -populations.

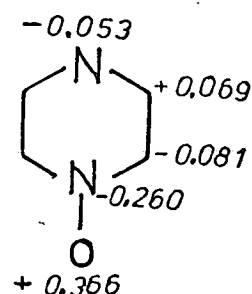
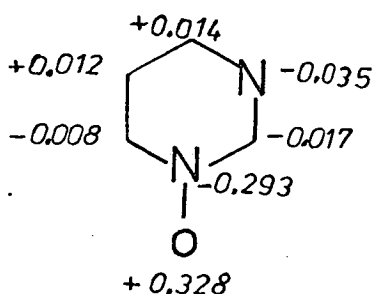
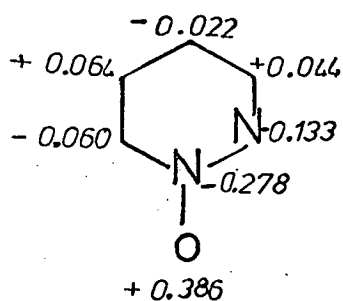
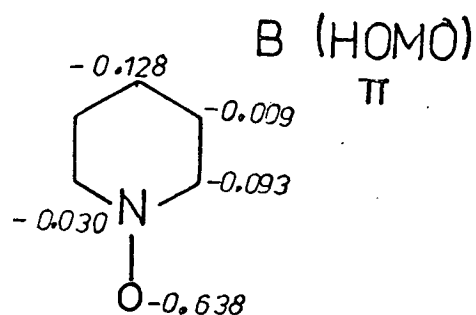
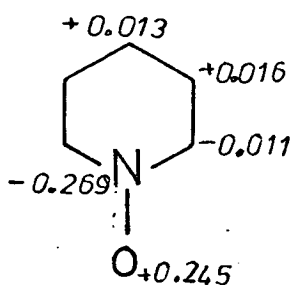
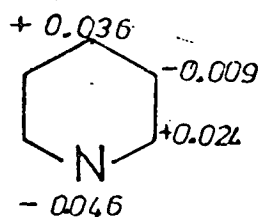
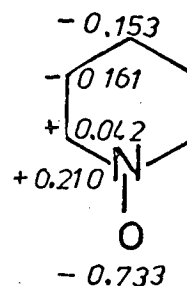
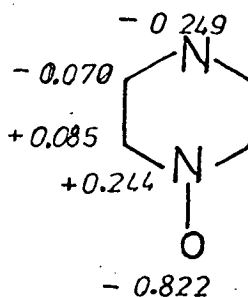
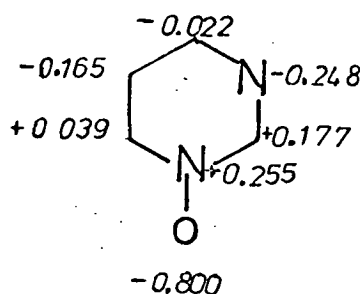
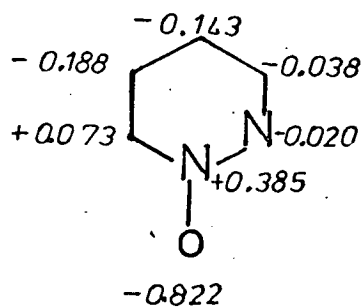


Fig.12

Net  $\sigma$ -populations.



Results. The total energies and dipole moments for the molecules under consideration are shown in Table 7.

The dipole moments of these compounds, and of pyridine N-oxide in particular, have been of considerable interest because of the canonical forms available to these molecules as shown in Figs. 1 and 3, and the resultant dipole will reflect the degree to which these occur.

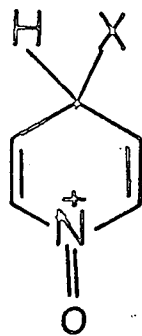
The electric dipole moments of the individual canonical forms in Fig. 1 have been estimated by one group of workers<sup>22</sup> to be 10.8D and 6.5D in the directions shown in Fig. 1. The resultant dipole, assuming a 50:50 mixture is 4.3D which is very close to the experimental value of 4.24D.<sup>22</sup> The LCGO calculations produce too large a dipole using a scaled basis set, and too small a dipole using the unscaled set. Bearing in mind that both furan and benzofuran, using a scaled basis set, have calculated dipoles which are too large, and that this could be due to the scaling procedure moving the lone pairs further away from the nucleus, then, if this effect were to arise here, the charge separation would be increased in favour of the  $N^+ - O^-$  form. This would produce a larger dipole than expected, as is the case.

The calculated dipoles of pyrimidine N-oxide and pyrazine N-oxide are also too large while that of pyridazine N-oxide is very close to the experimental value. The dipole moment is extremely sensitive to the molecular geometry and to the basis set, and the former probably significantly affects the dipole moment in cases, such as those considered here, involving charge separation. It is therefore not possible to place too much emphasis on the dipole moment information.

For pyridine N-oxide and pyrazine N-oxide the dipole is along the N-O bond as expected in these  $C_{2v}$  molecules, and the smaller magnitude of the latter dipole indicates that the contribution of the lone pair in the second nitrogen is in the opposite direction to that of the N-oxide as would be expected. In pyrimidine N-oxide the dipole is as shown in Fig. 10a and approximates to the resultant of the N-lone pair and the N-O bond contributions. In pyridazine N-oxide the dipole direction is as shown

Fig 13

a



b

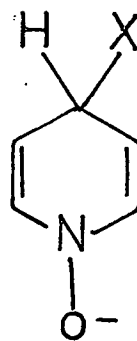
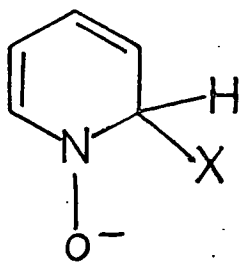
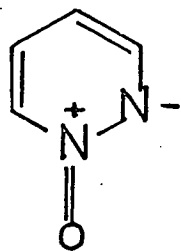


Fig.14



in Fig. 10b, and does not approximate to the resultant of a lone pair and N-O contributions as this would give a resultant dipole more in the direction of the N-N bond. The dipole is away from the N-N bond suggesting a contribution in the direction shown in Fig. 10b. This suggests that the nitrogen N(2) atom should be positive. Total populations show it to be negative but the centres C(4) and C(5) opposite the N-N bond are considerably more negative, Table 8, and will influence the dipole in this direction giving the resultant dipole as shown, Fig. 10b. A better indication, perhaps, of molecular polarisation can be obtained from the atomic populations. These are shown in Table 8.

As already mentioned the position of electrophilic substitution in pyridine N-oxide and in pyridine is 4- and 3- respectively, and the latter only with great difficulty. When the total populations of these two molecules are compared, then little difference is seen between them, and if anything, pyridine has slightly more negative ring centres. This does not seem to be compatible with the experimental evidence. However, if the net  $\pi$ -populations are considered, Fig. 11, then the 3-position in pyridine with a net population of -0.009 electrons is the one of highest electron density, agreeing with the experimental results of substitution. In pyridine N-oxide only the C(2)/C(6) positions are negative, and C(4) is only slightly less positive than C(3). It is perhaps a stabilised transition state of the form, shown in Fig. 13a, which activates the ring to electrophilic attack in the 4-position. The preference of C(4) over C(2) can be explained in terms of steric effects at C(2), and by the repulsion between the incoming electrophile and the "positive" nitrogen.

For pyridine N-oxide, this is not altogether a satisfactory explanation for electrophilic substitution in the 4-position. However, it is clear that electrophilic substitution in aromatic molecules occurs by attack upon the  $\pi$ -system, and unless the total electron densities differ dramatically at each centre, the determining factor for substitution will be the centre of highest electron density in the lowest binding energy  $\pi$ -orbital, as this orbital will be readily available to attack by the

Table 8.

## Atomic Populations

		N(1)	C(2)/C(6)	C(3)/C(5)	C(4)
Pyridine	$\sigma$	6.268	5.0304	5.1605	5.1678
	$\pi$	1.046	0.9760	1.0086	0.9640
	Total	8.314	6.0064	6.1614	6.1318
		O	N(1)	C(2)/C(6)	
Pyridine N-oxide	$\sigma$	6.7330	5.7900	4.9580	
	$\pi$	1.7550	1.2690	1.0110	
	Total	8.4880	7.0590	5.9690	
			C(3)/C(5)	C(4)	
	$\sigma$	5.1610	5.1530		
	$\pi$	0.9840	0.9870		
	Total	6.1450	6.1400		
		O	N(1)	N(2)	C(3)
Pyridazine N-oxide	$\sigma$	6.8216	5.6152	6.0196	5.0375
	$\pi$	1.6145	1.2780	1.1332	0.9558
	Total	8.4365	6.8932	7.1528	5.9933
			C(4)	C(5)	C(6)
	$\sigma$	5.1425	5.1882	4.9266	
	$\pi$	1.0217	0.9364	1.0604	
	Total	6.1642	6.1246	5.9870	

Table 8 (cont.)

		O	N(1)	N(3)	
Pyrimidine N-oxide	$\sigma$	6.8000	5.7454	6.2478	
	$\pi$	1.6723	1.2926	1.0352	
	Total	8.472	7.0380	7.2830	
		C(2)	C(4)	C(5)	C(6)
	$\sigma$	4.8226	5.0224	5.1645	4.9608
	$\pi$	1.0174	0.9856	0.9885	1.0082
	Total	5.8400	6.0080	6.1530	5.9690
		O	N(1)	N(4)	
Pyrazine N-oxide	$\sigma$	6.8220	6.7564	6.2485	
	$\pi$	1.6340	1.2600	1.0525	
	Total	8.4560	7.0164	7.3010	
		C(2)/C(6)	C(3)/C(5)		
	$\sigma$	4.9145	5.0697		
	$\pi$	1.0811	0.9301		
	Total	5.9956	6.0998		

Table 9.  $\sigma$  and  $\pi$  electron transfer by oxygen per electron

	$\sigma$	$\pi$
Pyridine N-oxide	$0.7330/6 = 0.1221$	$0.245/2 = 0.1225$
Pyridazine N-oxide	$0.8216/6 = 0.1367$	$0.3855/2 = 0.1928$
Pyrimidine N-oxide	$0.8000/6 = 0.1333$	$0.3277/2 = 0.1650$
Pyrazine N-oxide	$0.8820/6 = 0.1470$	$0.366/2 = 0.1830$



reagent. In pyridine, the highest occupied molecular orbital (HOMO) is of  $\pi$  type with  $a_2$  symmetry, indicating that it is nodal along the N-C(4) axis. This rules out substitution at the 4-position, and of the two remaining positions, C(3) has a slightly higher electron density, but is still almost neutral giving no activation towards electrophilic substituents. In pyridine N-oxide the highest occupied  $\pi$ -molecular orbital is of  $b_1$  symmetry, and is basically  $\pi(O)$  but attack at oxygen will not be substitution and will be reversible. Only attack in the ring itself will constitute a substitution reaction, and if the electron densities in the ring of the HOMO are considered, then C(4) is by far the most electronegative, Fig. 11 (B). The C(4) position carries a considerable negative charge, and will activate the molecule towards electrophilic substitution.

In nucleophilic substitution, the net  $\pi$ -atomic populations show that the 2 and 4-positions are positive in pyridine, and nucleophilic attack will be preferred at these positions. In pyridine N-oxide, the  $\pi$ -populations show position 3 to be the most positive, but the total populations give the C(2)/C(6) positions as being the only positive ones. It is perhaps the stabilised transition state, Fig. 13b, available when substitution is in the 2 or 4 position that activates the ring at these positions.

The chemical reactivity of pyridazine N-oxide in the 4-position at which it is readily attacked by electrophiles has been extensively studied.<sup>5</sup> In the  $\pi$ -populations C(6) is slightly more negative than C(4) but is overall a positive centre.

It is generally accepted that replacement of a ring C-atom by a more electronegative atom brings about a greater charge transfer from the N-O group to the ring.<sup>5</sup> Both the total and  $\pi$ -populations indicate that the nitrogen in the  $\alpha$ -position to the N-O group has the greatest effect in this manner. The most marked effects of charge transfer are normally found in the  $\pi$ -system, and the nitrogen in the 3-position shows the least ability to accommodate negative charge which is consistent with the lack of the resonance form D shown for pyridazine N-oxide in Fig. 3.

The total charge distribution and  $\pi$ -populations of pyridine N-oxide and pyrimidine N-oxide are remarkably similar. This is reflected in their dipole moments, Table 7, where both the calculated and the experimental values are similar. The total atomic populations of these two molecules, particularly at positions O, N(1) and C(4) suggest that the  $\text{>N}^{\dagger}=\text{O}$  structure is not significant here, and the lack of resonance form D, Fig. 3, in pyrimidine N-oxide is perhaps the reason for its more pyridine like structure.

The other two N-oxides, (2, 3, Fig. 2), show a fairly high population at C(4), and pyridazine N-oxide shows a reduced population at O compared with the other N-oxides in which the oxygen population varies little from molecule to molecule. In pyridazine N-oxide, the N(2) population is higher than for the other molecules, perhaps because of a greater contribution from the resonance form shown in Fig. 14. However, overall the N-O bond is polarised towards the oxygen but not to such a great extent as in any of the other molecules.

It is obvious that for the N-oxides, the oxygen acts strongly as a  $\sigma$ -acceptor and weakly as a  $\pi$ -donor, Figs. 11 and 12. In the  $\sigma$ -system the electron distribution approximates to A, Fig. 1. In the  $\pi$ -system, the back donation of electrons by the oxygen atom is a measure of the contribution of a resonance form of the type shown in B, Fig. 1. If the  $\sigma$ -acceptance per electron for oxygen, and the  $\pi$ -donation per electron for oxygen are calculated for these molecules, then the results shown in Table 9 are obtained. From these results it can be seen that the contribution of A, Fig. 1 for pyridine-N-oxide or of A and B, Fig. 2, for the diazine N-oxides is approximately 50% in all cases, the other 50% arising from resonance forms involving  $\text{>N}^{\dagger}=\text{O}$ . The diazine N-oxides show a slightly larger degree of back donation to the ring compared with pyridine N-oxide. This offsets the larger  $\sigma$ -donation to oxygen in the diazine N-oxides, and results overall in the oxygen atom in both pyridine N-oxide and the diazine N-oxides having a similar total population. However, although the effect per electron of the resonance forms is comparable in the  $\sigma$  and  $\pi$  systems, the

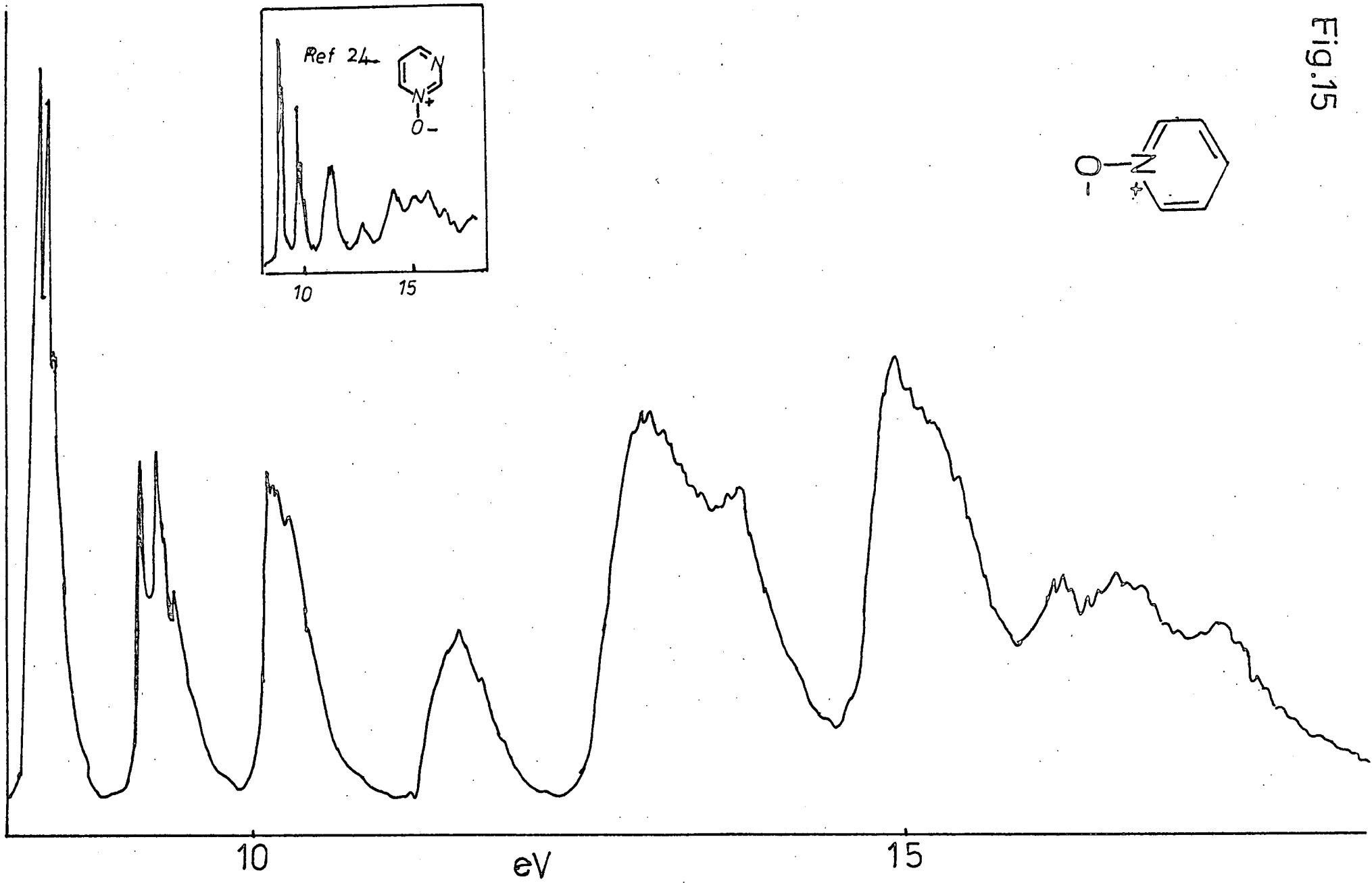


Fig.15

molecule overall is more influenced by the  $\sigma$ -system and each molecule is better represented by resonance forms of type A, Fig. 1 and A and B, Fig. 2, except for the fact that N(1) is always negative overall. Consideration of the  $\sigma$  and  $\pi$  populations indicates that this is due to back donation of the  $\alpha$  carbons to N(1) in the  $\pi$ -system, and this cannot be accounted for in resonance structures.

### Photoelectron Spectra of the N-Oxides

The PE spectrum of pyridine N-oxide, and several of its derivatives have been extensively studied,<sup>23</sup> producing good resolution spectra. The PE spectrum obtained in this work is shown in Fig. 15, and the experimental IPs are almost identical to those quoted in reference 23, Table 10. The IPs of reference 23 were assigned on the basis of information obtained from the IPs of benzene,  $\text{CF}_3\text{NO}$  and  $\text{CH}_3\text{NO}$ , and also using MINDO/2 calculations. The assignment of the first four IPs using MINDO/2 agrees with that found by the LCGO calculation, Table 11, but MINDO/2 assigns  $\text{IP}_5$  at 13.0 eV as  $\pi 1b_1$  whereas the LCGO calculation gives  $\text{IP}_5$  as orbital  $13a_1$ , a  $\sigma$  orbital, and the  $\pi(1b_1)$  orbital as  $\text{IP}_9$  at much higher binding energy.

At least squares fit of experimental and calculated IPs for pyridine N-oxide gives the relationship  $\text{IP}_{\text{expt}} = 0.736 \text{IP}_{\text{calc.}} + 1.89 \text{ eV}$ . The assignments are given in Table 12.

Due to the difficulty in finding a suitable experimental method for the preparation of pyrimidine N-oxide, it was decided to use published data<sup>24</sup> for comparison with the LCGO calculations. Comparing the spectrum of pyrimidine N-oxide with that of pyridine N-oxide confirms the similarity in molecular structure indicated by the populations. The two spectra are remarkably similar, and the general form is shown in Fig. 15. The effect of the second nitrogen is to move the first three IPs to slightly higher binding energy, implying that the orbitals have become more delocalised. The eigenvectors for these orbitals indicate this to be the case, especially in the HOMO which is  $\pi(\text{O})$  and the eigenvector on the oxygen is reduced in the pyrimidine N-oxide to 0.78 from 0.83

Table 10. Expt. IPs (eV) of pyridine N-oxide

IP	1	2	3	4	5	6	7
Reference 23	8.34	9.22	10.18	11.59	13.0	13.8	-
This work	8.40	9.15	10.15	11.60	13.1	13.7	15.0

Table 11. Assignment of IPs in pyridine N-oxide

Reference 23	$\pi(O)(3b_1)$	$\sigma(O)(b_2)$	$\pi(a_2)$	$\pi(2b_1)$	$\pi(1b_1)$
MINDO/2					
LCGO	$\pi(O)(3b_1)$	$\sigma(O)(b_2)$	$\pi(a_2)$	$\pi(2b_1)$	$\sigma(13a_1)$

Table 12 Assignments of IPs in pyridine N-oxide

IP	8.40	9.25	11.60	13.1	13.7	15.0
$-e_i$	8.68	10.60	13.76	14.52	15.61/16.80	17.87/17.99
IP	15.35	16.20	16.62			
$-e_i$	18.70	19.94	20.66			

Fig. 16

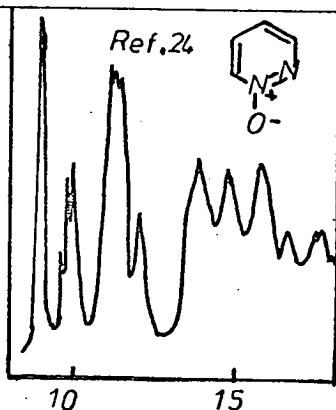
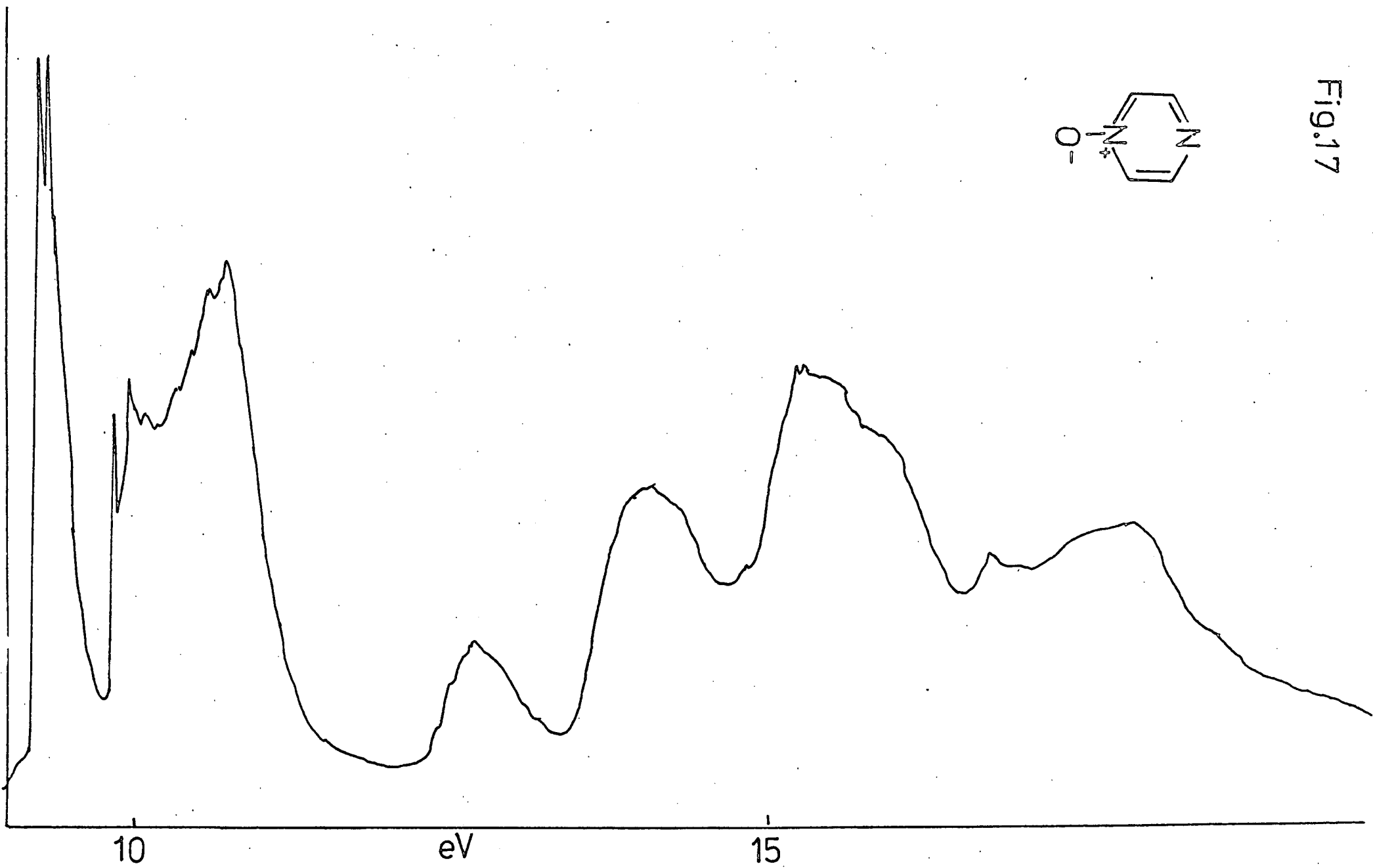
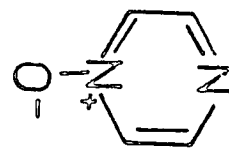


Fig.17



in the pyridine N-oxide. For this molecule (2), the CNDO/2 and MINDO/2 calculations give the assignments shown in Table 13. The LCGO calculation gives the order as  $IP_3$  ( $\pi$ ) and  $IP_4$  ( $\sigma$  (N/O)), whereas the semi-empirical calculations disagree on the ordering. The other assignments are in agreement, and the unassigned  $IP_6$  can be identified as a delocalised  $\sigma$ -orbital. The orbital associated with  $IP_2$  is predominantly  $\sigma$  (O) and not  $\sigma$  (O/N) as given by the semi-empirical calculations.

The spectrum of pyridazine N-oxide,<sup>24</sup> Fig. 16 shows four IPs below 13 eV, as in pyrimidine N-oxide, the main difference being the profile of  $IP_2$  which in the latter represents an adiabatic IP, and in the former a vertical IP. The LCGO eigenvectors associated with this orbital show it to be  $\sigma$  (O/N/C) in pyridazine N-oxide and  $\sigma$  (O) for the other, giving a more localised orbital, and hence the transition is (O  $\rightarrow$  O) corresponding to a vertical IP. The MINDO/2 ordering is shown in Table 13, and is the same as that produced by the LCGO calculation.

The PE spectrum of pyrazine N-oxide, Fig. 17, obtained in this work, shows better resolution on the first two bands than that of reference 24. Compared with the spectra of the other N-oxides,  $IP_1$  has moved to higher binding energy ( $> 8$  eV, others  $IP_1 < 9$  eV), and IPs 3 and 4 have moved to lower binding energy. For  $IP_1$  which is given as  $\pi$  (O) by the LCGO calculation, this agrees with the atomic population data which shows oxygen in this molecule to have a greater  $\pi$ -back donation effect than in pyrimidine N-oxide, producing less charge localisation on the oxygen atom, and a more delocalised  $\pi$  (O) orbital. The  $\pi$ -back donation of electrons to the ring is similar in pyrazine N-oxide and in pyridazine N-oxide, but the overall  $\pi$ -charge separation of N and O is larger for pyridazine N-oxide (0.6635 e) than for pyrazine N-oxide (0.6260 e) producing an overall delocalisation effect in pyrazine N-oxide.

A least squares fit of calculated and experimental data, Table 12, produces the relationship  $IP_{\text{expt}} = 0.68 IP_{\text{calc.}} + 2.35$  eV. The LCGO calculation reverses the order of assignments of  $IP_3$  and  $IP_4$  when

Table 13. Molecular Orbital assignments for the diazine  
N-oxides

	IP <sub>1</sub>	IP <sub>2</sub>	IP <sub>3</sub>	IP <sub>4</sub>
Pyridazine N-oxide <sup>24</sup>	8.89 [π(O)]	9.95 [σ(O/N)]	11.05/11.2 [σ(N/O); π]	
CNDO/2-MINDO/2				
LCGO	π(O)	σ(O/N/C)	π σ(N/C)	
	IP <sub>5</sub>	IP <sub>6</sub>		
CNDO/2-MINDO/2	12.0 (π)	13.8		
LCGO	π	σ		
	IP <sub>1</sub>	IP <sub>2</sub>	IP <sub>3</sub>	IP <sub>4</sub>
Pyrimidine N-oxide	8.80 [π(O)]	9.52 [σ(O/N)]	11.1/11.2 [σ(N/O); π]	
CNDO/2-MINDO/2				
LCGO	π(O)	σ(O)	π σ(N/O)	
	IP <sub>5</sub>	IP <sub>5</sub>		
CNDO/2-MINDO/2	12.60 (π)	14.0		
LCGO	π	σ		
	IP <sub>1</sub>	IP <sub>2</sub>	IP <sub>3</sub>	
Pyrazine N-oxide	9.17 [π(O)]	9.9/10.3 [σ(O); σ(N)]		
CNDO/2-MINDO/2				
LCGO	9.68 [π(O)]	12.05	12.11	
-e <sub>i</sub>	9.31	9.93	10.48	
	IP <sub>4</sub>	IP <sub>5</sub>	IP <sub>6</sub>	
CNDO/2-MINDO/2	10.68 (π)	12.72 (π)	13.85	
LCGO	12.49	15.49	16.65	
-e <sub>i</sub>	10.58	12.66	14.01	



Table 13 (cont.)

	$IP_7$	$IP_8$	$IP_9$	$IP_{10}$	$IP_{11}$
LCGO	17.25	18.44	19.76	19.84	21.04
$-e_i$	14.01	15.22	15.72	15.72	16.70
	$IP_{12}$	$IP_{13}$	$IP_{14}$		
LCGO	21.33	25.65	26.27		
$-e_i$	17.48	19.48	19.48		

compared with the semi-empirical calculation, and gives  $IP_6$  as  $\pi$ ,  
Table 12.

The LCGO calculations have been useful in removing the ambiguity  
of assignments in the PE spectra of the N-oxides where semi-empirical  
calculations have been used to identify the orbital types.

## References

1. T. Kubota, J. Chem. Soc. Japan, 1958, 79, 930; 1959, 80, 578.
2. T. Kubota, Bull. Chem. Soc. Japan, 1962, 35, 946.
3. M. Ito and N. Hata, Bull. Chem. Soc. Japan, 1955, 28, 260.
4. M. Ito and N. Hata, J. Chem. Phys., 1956, 24, 495.
5. T. Kubota and H. Watanabe, Bull. Chem. Soc. Japan 1963, 36, 1093.
6. E. M. Elveth, Theoret. Chim. Acta (Berlin) 1968, 11, 145.
7. S. Szoke, G. Jalsovsky and E. Dudar, Acta Chimia Academiae Scientiarum Hungaricae, 1972, 72(3), 241.
8. J. P. Maier and J. F. Muller, J. Chem. Soc., Faraday Transactions 2, 1974, 70, 1991.
9. J. P. Maier and J. F. Muller, Tetrahedron Letters, 1974, 35, 2987.
10. C. M. Bax, A. R. Katritsky, and L. E. Sutton, J. Chem. Soc., 1958, 1258.
11. M. Ogata and H. Kano, Chem. and Pharm. Bull., 1963, 11, 29 35.
12. E. Ochiai, "Aromatic amine oxides," Elsevier, Amsterdam 1967.
13. V. Drenzler and Rudolph, Z. Naturforsch., 1967, 22a, 531.
14. J. F. Chiang, J. Chem. Phys., 1974, 61(4), 1280.
15. H. K. Bodensch, Z. Naturforsch., 1969, 24, 1966.
16. D. W. Smith and K. Hedberg, J. Chem. Phys., 1956, 25, 1282.
17. P. A. Akishin, L. V. Vilkov, and V. Ya Roslovskii, Zh. Strukt. Khim., 1960, 1, 1.
18. R. Desiderato and J. C. Terry, J. Het. Chem., 1971, 18, 617.
19. A. Caron, G. Palenik, E. Goldish, and J. Donohue, Acta Cryst., 1964, 17, 102.
20. P. J. Wheatley, Acta Cryst., 1957, 10, 182.
21. P. J. Wheatley, Acta Cryst., 1960, 13, 80.

22. P. Hamm and W. v. Philipsborn, *Helv. Chim. Acta*, 1971, 54, 2363.
23. J. P. Maier and J. F. Muller. *Trans. Faraday Soc.*, 1974, 1991.
24. J. P. Maier, J. F. Muller, and T. Kubota, *Helv. Chim. Acta*, 1975, 58 1634.

Correlation of XPS shifts with change in charge density  $\Delta q$ ,  
as calculated by INDO.

Initial work on the series of molecules (Figs 1 and 2) considered here was done with a view to studying these molecules by LCGO calculations. Due to lack of information on the geometric structure of these molecules, it was decided to optimise the geometries semi-empirically using the INDO procedure. The large number of parameters to be optimised proved time consuming in several cases. The optimisation procedures are shown in Tables 1-8. In all cases optimisation was performed on only one parameter at a time, keeping all the others constant. In view of the optimisation difficulties, it was decided that the final geometries were possibly insufficiently accurate for LCGO work. As an alternative, the INDO calculations were used to study the molecules, in particular to study the effect of charge density changes on the core orbital energy levels which can be evaluated experimentally by XPS.

In XPS work,<sup>1,2,3</sup> attempts have been made to correlate the changes in XPS core electron binding energies, caused by changes in the valence electron structure, to the corresponding change in the charge densities of the atoms in the molecule. The difference in XPS binding energies of a core electron of, for example, an atom and the atom in a molecule is called the "shift." This shift,  $\Delta E$ , will be produced by a change in electrostatic potential due to the redistribution of valence electrons caused by a change in chemical environment.

The change in potential at an atom can be regarded as comprising of two components; a one-centre part representing the change in interaction of a core electron in an atom with a valence electron in the same atom, and a two-centre part representing the interaction of a core electron with the other atoms in the molecule. This gives the XPS shift, for a core electron on an atom A as:-

$$\Delta E_A = k \Delta q_A + \sum_{B \neq A} \frac{Z_B}{R_{AB}}$$

Fig. 1 (a) series - cationic molecules

- a(1) 2-amino cinnolinium cation  
a(2) quinolizinium cation  
a(3) N-amino pyridinium cation  
a(4) N-methyl pyridinium cation

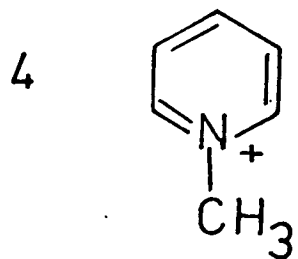
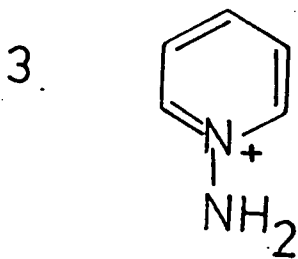
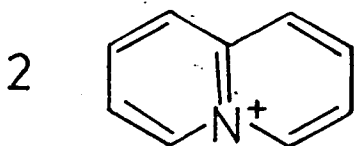
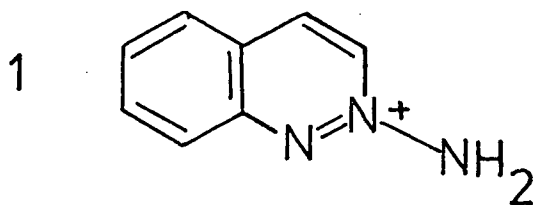


Fig. 2 (b) series - the neutral molecules

- b(1) cinnoline N(2)-oxide
- b(2) cinnoline N(1)-oxide
- b(3) cinnoline di-N-oxide
- b(4) 1,2,3-benzotriazine N(3)-oxide

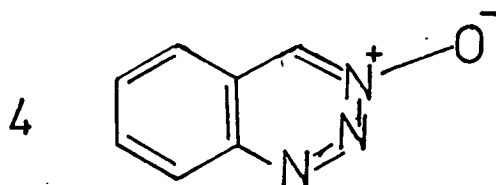
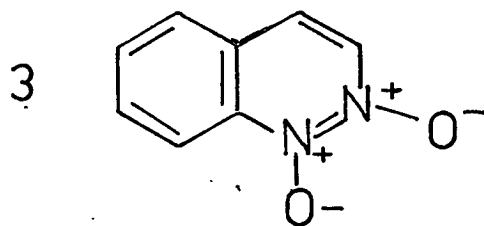
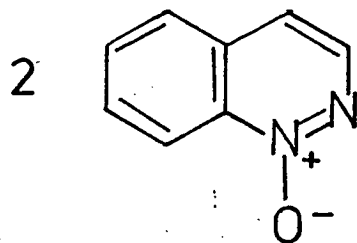
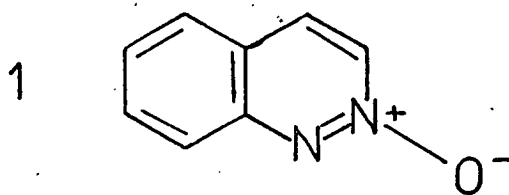


Table 9. Siegbahn's data used to evaluate  $k$ .

	E (eV)	$k q_{\text{CNDO}} + v + 1$
C(1s) H <sub>4</sub>	0	-0.1
C(1s) H <sub>3</sub> CH <sub>2</sub> OH	0.2	-0.0
C(1s) H <sub>3</sub> COCH <sub>3</sub>	0.5	-0.2
C(1s) H <sub>3</sub> COH	0.6	0.5
C(1s) H <sub>3</sub> COOH	0.7	0.4
C(1s) H <sub>3</sub> OH	1.6	2.1
CH <sub>3</sub> C(1s) H <sub>2</sub> OH	1.6	2.3
(CH <sub>3</sub> ) <sub>2</sub> C(1s) O	3.1	3.2
CH <sub>3</sub> C(1s) HO	3.2	3.3
CH <sub>3</sub> C(1s) OOH	4.7	4.5
C(1s) HF <sub>3</sub>	8.1	8.4
C(1s) F <sub>4</sub>	11.1	10.4



where  $\Delta q_A$  is the difference in the charge density of A compared with a reference system.

k represents the average interaction between a core electron of A and a valence electron in A.

Z is the nuclear charge of atom B.

$R_{AB}$  is the internuclear distance between A and B and  $Z_{B/R_{AB}}$  assumes the core electron of A as a point charge at A. To compensate for any systematic error in the evaluation of the reference state, a work function, l, can be introduced to give:-

$$\Delta E_A = k \Delta q_A + v + l \quad \text{where } v = \sum_{B \neq A} Z_{B/R_{AB}}$$

$$\Delta E_A - v = k \Delta q_A + l$$

The left hand side of the equation now represents the potential adjusted shift,  $\Delta E'$ .

$$\Delta E' = k \Delta q_A + l$$

The relationship  $\Delta E' - v - \Delta q_A$  can be plotted for an atom A in different chemical environments using the same reference state in each case to evaluate  $\Delta q_A$ . This will yield values of k and l, if the relationship is linear. The charge densities,  $\Delta q$ , are found for the atoms in the molecules by CNDO or INDO calculations, and  $\Delta q_A$  is evaluated from  $(q_{A \text{ std.}} - q_A)$  where  $q_{A \text{ std.}}$  is the charge density of atom A in its standard reference state.

Siegbahn<sup>1</sup> has shown that for the series of molecules listed in Table 9, the XPS C(1s) shifts using C(1s) of methane as a standard, correlate well, both for large shifts when the carbon atoms are in very different chemical environments e. g.  $\text{CH}_4$  and  $\text{H}_2\text{C}=\text{O}$ , and for small shifts when they are in similar chemical environments e. g.  $\text{CH}_4$  and  $\text{CH}_3 - \text{CH}_2\text{X}$ . The data used by Siegbahn is shown in Table 9, and yields the value of k as 21.9 eV. Other workers<sup>2</sup> have produced

Table 10.

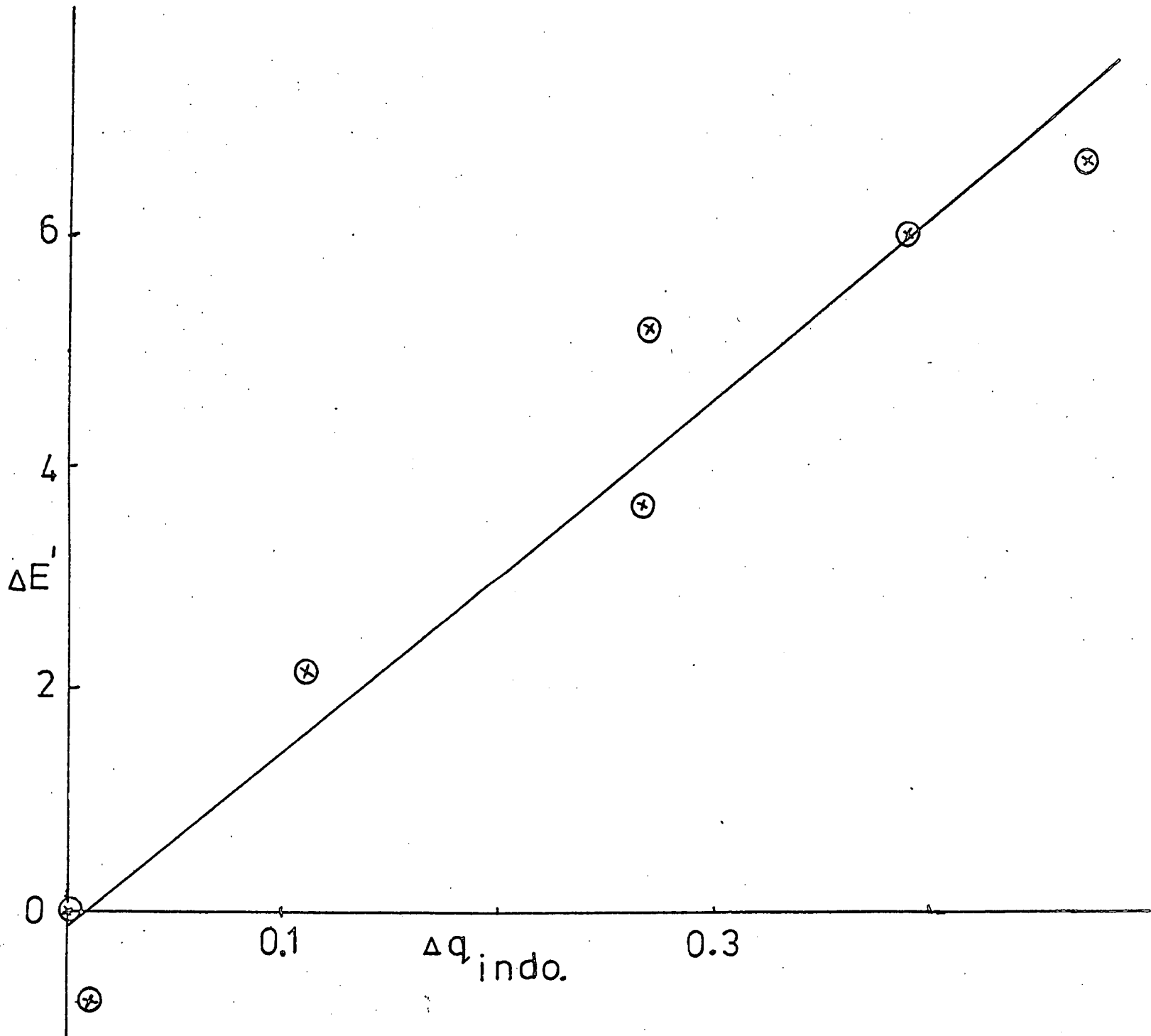
## Cationic Molecules

		XPS N(1s)(eV)	q INDO	V (eV)
a(1) 2-amino cinnolinium				
	cation			
1.	N <sup>+</sup>	400.0	0.3004	3.0501
2.	N-amine	400.0	-0.1698	9.6817
3.	N (1)	400.0	-0.0603	7.4395
a(2) quinolizinium cation				
4.	N <sup>+</sup>	402.1	0.1015	6.5882
a(3) N-amino pyridinium				
	cation			
5.	N <sup>+</sup>	401.4	0.2197	5.1079
6.	N-amine	398.9	-0.1606	9.3970
a(4) N-methyl pyridinium				
	cation			
7.	N <sup>+</sup>	401.0	0.0986	7.0464

Table 11.  $\Delta E'$  and  $\Delta q$  for the cationic molecules

Nitrogen No as in Table 9	$\Delta E'$	$\Delta q$
1	6.63	0.4702
2	0.00	0.0000
3	2.24	0.1095
4	5.19	0.2713
5	5.97	0.3895
6	-0.82	0.0092
7	3.63	0.2684

Fig.3



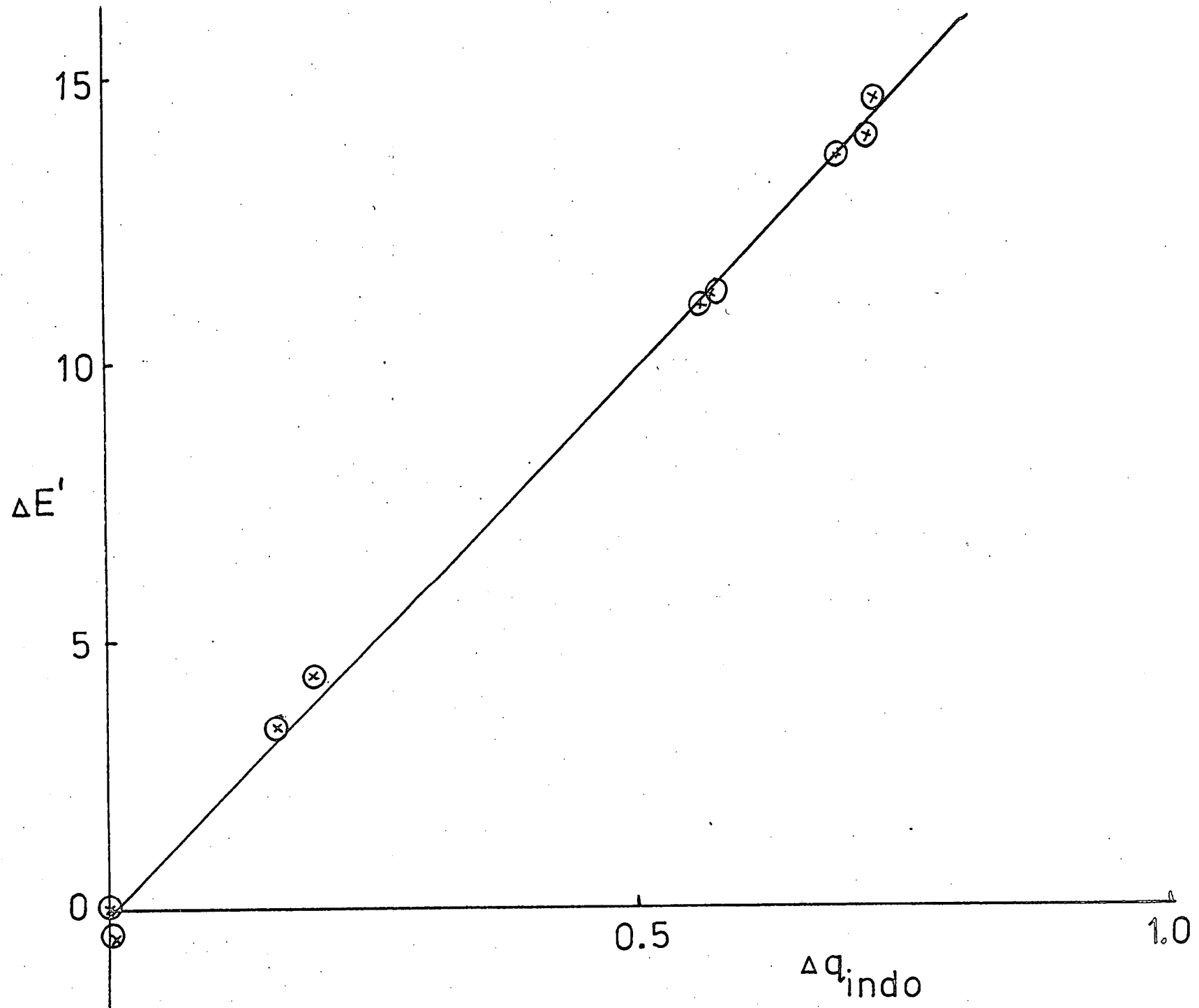


Fig.4

values of  $k$  for carbon in agreement with this.

It has also been shown<sup>4</sup> that different slopes of the graph corresponding to  $\Delta q$  v  $\Delta E'$  are obtained for the same atom depending on whether neutral or ionic molecular environments are being considered. For this reason it was decided to divide the molecules into two categories; (a) those containing nitrogen as a cation, and (b) those containing nitrogen in a neutral molecule, and to study the N(1s) XPS shifts\* in each series of molecules separately. The accuracy with which a correlation of  $\Delta q$  with  $\Delta E'$  for N(1s) levels can be performed will partly depend on the accuracy of the molecular geometry, and as these had been optimised, it was expected that, if a relationship did exist, it should be given fairly precisely.

For the cationic molecules, the reference XPS 1s nitrogen level was chosen to be that of the amino group in the 2-amino cinnolinium cation. Table 10 lists the XPS N1s levels,  $q$  and  $v$  associated with the N atoms in the cationic molecules, and Table 11 gives the values of  $\Delta E'$  and  $\Delta q$  to be plotted. A graph of  $\Delta E'$  v  $\Delta q$ , Fig. 3, for these atoms was drawn, and a least squares fit of the values obtained, yielding  $\Delta E' = 15.45 \Delta q - 0.088$  eV with standard deviations in slope and intercept of 1.68 and 0.46 respectively. Tables 12 and 13 show the equivalent data for the neutral molecules and a least squares fit of  $\Delta E'$  v  $\Delta q$  yields, using the N(1) atom of cinnoline N(2)-oxide as the standard,  $\Delta E' = 19.65 \Delta q - 0.008$  eV with standard deviations as before of 0.336 and 0.167. The graph of these results, Fig. 4, reflects the good linear relationship indicated by the least squares fit.

It can be seen that although a linear relationship exists for the cationic molecules, it is not so good as that for the neutral molecules. The results also indicate the variability of  $k$  for one particular atom. Therefore, for the relationship of  $\Delta E'$  to  $\Delta q$  to be valid only atoms

---

\* Thanks are extended to A. E. I. for use of their ES 200 spectrometer in obtaining the XPS spectra.

Table 12

## Neutral Molecules

		XPS N(1s)(eV)	q INDO	V (eV)
b(1) cinnoline N(2)-oxide				
1.	N <sup>+</sup>	403.3	0.4749	-6.5367
2.	N(1)	399.5	-0.2491	4.2485
b(2) cinnoline N(1)-oxide				
3.	N <sup>+</sup>	402.7	0.4697	-6.4066
4.	N(2)	399.0	-0.2462	4.1346
b(3) cinnoline di-N-oxide				
5.	N(1) <sup>+</sup>	403.3	0.3109	-2.8576
6.	N(2) <sup>+</sup>	403.3	0.3188	-3.0121
b(4) 1,2,3-benzotriazine				
N(3)-oxide				
7.	N <sup>+</sup>	403.0	0.4435	-5.8155
8.	N(1)	400.6	-0.0577	1.1035
9.	N(2)	400.6	-0.0872	2.0497

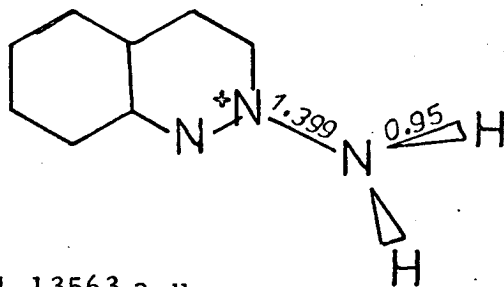
Table 13  $\Delta E'$  and  $\Delta q$  for the neutral molecules

Nitrogen No as in Table 11	$\Delta E'$	$\Delta q$
1	14.49	0.7240
2	0.0	0.0
3	13.86	0.7188
4	-0.38	0.0029
5	10.91	0.5600
6	11.06	0.5679
7	13.57	0.6926
8	4.25	0.1914
9	3.30	0.1619

in similar molecular environments can be compared. The cationic molecules, however, produce, in general, much smaller values of  $\Delta E'$  and  $\Delta q$  than the neutral molecules, and any error in the calculations will be magnified as a % error for these molecules in the  $\Delta E'$  and  $\Delta q$  values, tending to make the relationship between them worse.

Table 1.            2-Amino cinnolinium cation

The optimised geometry of cinnoline N(2)-oxide was used and the amino group was added bisecting the external angle. The amino group geometry was tetrahedral.



T. E. = -94.13563 a. u.

B. E. = -9.52663 a. u.

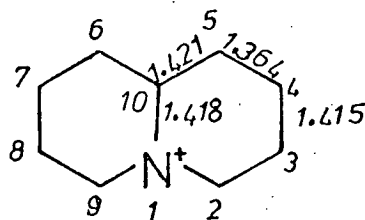
Dipole = 6.43772 D



Table 2.

Optimisation of the quinolizinium cation

Initial geometry - naphthalene.



T.E. = -78.6271819012  
 B.E. = -9.4295723716 a. u.  
 Dipole = 1.07230 D

1) C-N bonds reduced

	N(1) - C(10)	N(1) - C(2)	N(1) - C(9)	T. E. (a. u.)
1)	1.398	1.401	1.401	-78.6332199637
2)	1.388	1.391	1.391	-78.6342922883
3)	1.378	1.381	1.381	-78.6345346554
4)	1.368	1.371	1.371	-78.6336932784

Parabolic minimum of (2), (3), (4) gives:-

N(1) - C(10) = 1.381 Å    E = -78.63457605 (a. u.)  
 If N(1) - C(2) = 1.384 Å,    N(1) - C(9) = 1.384 Å  
 T. E. = -78.6344225687 a. u.  
 C(2) - C(3) = 1.372 Å

Table 2 (cont.)

2) C(3)/C(8) moved inwards  $\Delta x = \pm 0.01\text{\AA}$

First run C(2) - H(11) ] bisecting the external  
C(9) - H(18) ] angle

and C-H =  $1.084\text{\AA}$

T. E. = -78.6351811328 a. u.

<u>C(2) - C(3)</u>	<u>T. E. (a. u.)</u>
1) 1.372	-78.6351811328
2) 1.363	-78.6370666430
3) 1.354	-78.6377411728
4) 1.345	-78.6371986967

Parabolic minimum of (2), (3) and (4) gives

C(2) - C(3) =  $1.354\text{\AA}$  T. E. = -78.63774290 a. u.

Run (3) above corresponds to the best geometry.

Total increment change  $\Delta x = \pm 0.02\text{\AA}$  gives:-

C(3) - H ] =  $1.101\text{\AA}$   
C(8) - H ]

3) C(4)/C(7) moved inwards  $\Delta x = 0.01$

C(3) - H ] set at  $1.084\text{\AA}$  and  
C(8) - H ] posn. changed by  $\Delta x = \pm 0.02$

<u>C(2) - C(3)</u>	<u>C(3) - C(4)</u>	<u>T. E.</u>	<u><math>\Delta x</math></u>
1.354	1.41514	-78.63558	0.0
1.354	1.41503	-78.63491	0.01
1.354	1.41512	-78.63556	0.001
1.354	1.41516	-78.63559	0.001

Table 2 (cont)

- 4) C(3) - H } set at 1.084 Å and bisecting the external  
C(8) - H } angle.

T.E. = -78.6360930024 a.u.

- 5) Final geometry

C(3) - H, C(8) - H set equal to 1.101 Å, and  
bisecting the external angle

T.E. = -78.6384196853 a.u.

B.E. = -9.4408101557 a.u.

Dipole = 1.08938 D

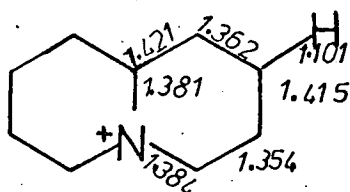
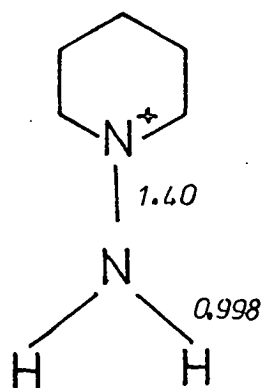


Table 3.Optimisation of N-amino pyridinium cation

An optimised geometry of the pyridinium cation was used.

The N-N bond length was set initially to  $1.40\text{\AA}$  and the N-H bond length to  $0.998\text{\AA}$ .

INDO calculation

Total energy =  $-61.6215577559$  a. u.

Dipole moment =  $0.97151$  D

molecule - planar

Table 3 (cont.)

1) optimisation of N-N bond length  
(molecule planar)

N-N	T. E. (a. u.)	$\widehat{\text{H}}\text{NH}$	N - H
1.40Å	-61.6215577559	120°	0.998Å
1.50Å	-61.5746436549	120°	
1.41Å	-61.6182654594	120°	
1.390Å	-61.6348700955	109° 40	1.057Å
1.380Å	-61.6175335736		
1.389Å	-61.6345033564		

Best energy corresponds to

N - N = 1.390Å

N - H = 1.057Å

$\widehat{\text{H}}\text{NH} = 109^\circ 40$

N - N = 1.390Å    N - H = 0.998Å     $\widehat{\text{H}}\text{NH} = 120^\circ$   
T. E. = -61.6245656441 (planar)

2) N-amino pyridinium cation non-planar

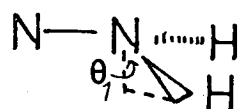
non-planar 1                      tetrahedral geometry

$\widehat{\text{H}}\text{NH} = 109^\circ 30$                       N - H = 0.998Å

T. E. = -61.6299388574 a. u.

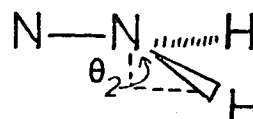
non-planar 2                      (increase  $\Theta_1$  by 5°)

(non planar 1)



$\Theta_1 = 69^\circ 11'$      $\Theta_2 = 74^\circ 11'$

(non-planar 2)



$\Theta_2$  : T. E. = -61.6255803653 a. u.

Table 3 (cont.)

non-planar 3 (decrease  $\Theta_1$  by  $2^\circ$ )

$$\Theta_3 = 67^\circ 11'$$

$$\text{T. E.} = -61.6278872462 \text{ a. u.}$$

non-planar 4

N - H bond length increased - tetrahedral geometry

N - H initially  $0.998 \text{ \AA}$

N - H length	T. E.
$0.998 \text{ \AA}$	-61.6299388574
$1.008 \text{ \AA}$	-61.6348741712
$.018 \text{ \AA}$	-61.6389869743
$1.028 \text{ \AA}$	-61.6422559077
$1.038 \text{ \AA}$	-61.6447208179
$1.048 \text{ \AA}$	-61.6464200350
$1.058 \text{ \AA}$	-61.6473904059
$1.068 \text{ \AA}$	-61.6476673954

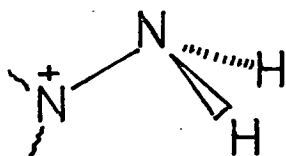
N-amino pyridinium cation best geometry

$$\text{N} - \text{N} = 1.390 \text{ \AA}$$

$$\text{N} - \text{H} = 1.068 \text{ \AA}$$

$$\text{HNH} = 109^\circ 30'$$

Tetrahedral



$$\text{T. E.} = -61.6476673954$$

$$\text{B. E.} = -6.1554332913 \text{ a. u.}$$

$$\text{Dipole} = 1.70852 \text{ D}$$

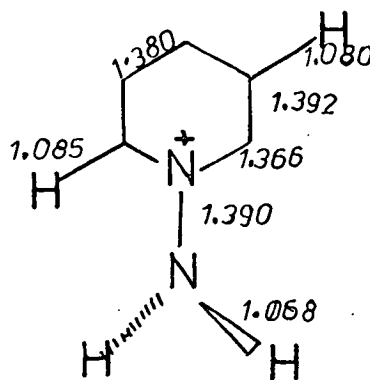
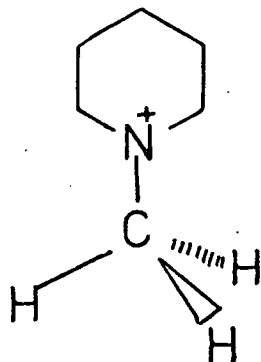


Table 4Geometry of the N-methyl pyridinium cation

Geometry previously optimised.

T. E. = -58.1007832937

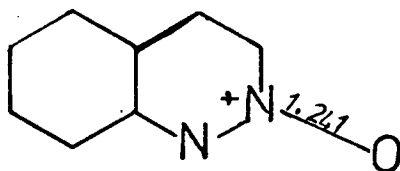
B. E. = -6.7080382424 a. u.

Dipole = 0.33168 D.

Table 5Optimisation of Cinnoline N(2) oxide

The optimised geometry of cinnoline N(1) oxide was used.

The N - O bond was set at 1.241 Å and bisecting the external angle.

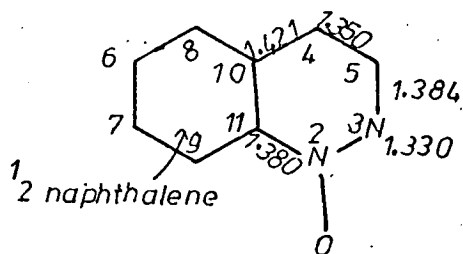


T. E. = -99.2992921305

B. E. = -9.3488387293 a. u.

Dipole = 5.74061 D.

Table 6.    Optimisation of cinnoline N(1)-oxide geometry



An optimal geometry of cinnoline as calculated in reference 5 was used as the basis for the N-oxide molecule.

The N-oxide bond length was set initially to 1.25 Å.

INDO calculation

T. E. = -99.2891363666 a. u.

B. E. = -9.3386829654 a. u.

Dipole = 5.47475 D.

1) Optimisation of N-O bond length

N-O length	Total energy (a. u.)
1.22 Å	-99.2880745958
1.24 Å	-99.2894241144
1.26 Å	-99.2880923325
1.242 Å	-99.2894026933

N-O bond length = 1.241 Å



Table 6 (cont.)

2) Optimisation of N(3) position

initial N - N length = 1.330 Å

N - O " = 1.241 Å

INDO total energy = -99.2894167890

Run I moving N(2) only  $\Delta x = +0.1$   $\Delta y = +0.1$ Run II moving N(2) only  $\Delta x = -0.01$   $\Delta y = -0.01$ Run III moving N(2) only  $\Delta x = -0.05$   $\Delta y = -0.05$ 

Final result on convergence of T.E. of 0.0001 a.u.

N-N = 1.303 Å

Total Energy = -99.2911056884 a.u.

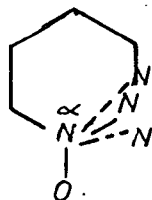
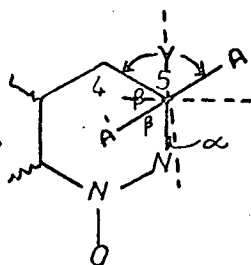
3) Optimisation of C(11)  $\hat{N}(2)$  N(3). $\hat{C}N(1) N(2)$  cinnoline =  $119^{\circ} 6'$  $\hat{C}N(1) N(2)$  optimised N(2) =  $124^{\circ} 0'$ a)  $\hat{C}N(1) N(2)$  put equal to  $121^{\circ} 0'$ b)  $\hat{C}N(1) N(2)$  put equal to  $125^{\circ} 0'$  $\alpha = 124^{\circ}$  T.E. = -99.2911056884 a.u. $\alpha = 121^{\circ}$  T.E. = -99.2858803590 a.u. $\alpha = 125^{\circ}$  T.E. = -99.2894349605 a.u. $\therefore \alpha$  optimised =  $124^{\circ} 0'$

Table 6 (cont.)

4) Optimisation of C(5)

a) moving C(5) only.

$$N(3) \hat{C}(5) C(4) = 120^{\circ} 35'$$

$$\alpha = 1^{\circ} 38'$$

$$\gamma = 57^{\circ} 47'$$

$$\beta + \gamma = \Theta = 61^{\circ} 56'$$

C(5) moved along the bisector of  $N(3) \hat{C}(5) C(4)$

$$\Delta x = -(\sin \Theta \times 0.1) = -0.088241$$

$$\Delta y = -(\cos \Theta \times 0.1) = -0.047050$$

initial increment along A-A' = 0.1 Å

Best total energy = -99.2921240016 a. u.

corresponding to an increment of 0.02 Å along (A-A').

5) Position of H(12)

The position of H(12) was finally fixed using the original bond length of 1.0844 Å and letting the C(5) - H(12) bond bisect the external angle

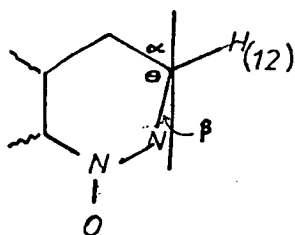


Table 6 (cont.)

$$\alpha = 57^{\circ} 2'$$

$$\beta = 0^{\circ} 55'$$

$$\Theta = (180 - 57^{\circ} 57') = 122^{\circ} 3'$$

External angle =  $237^{\circ} 57'$  giving the best geometry.

$$\text{T.E.} = -99.2956478172 \text{ a.u.}$$

$$\text{Dipole} = 5.49640 \text{ D}$$

$$\text{B.E.} = -9.3451944160 \text{ a.u.}$$

Table 7     Geometry of cinnoline di-N-oxide

The optimised geometry of cinnoline N(1)-oxide was used.

Each N-O bond was set equal to 1.241 Å, and in the same position as for the N(1) and N(2) oxides.

T. E. = -116.9281443889 a. u.

B. E. = -7.6856059227 a. u.

Dipole = 7.77806 C

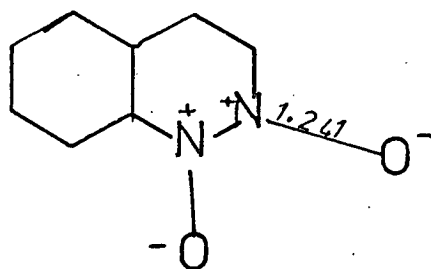
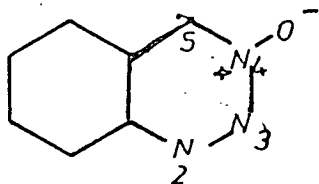
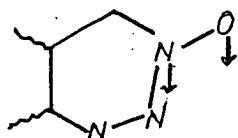


Table 81, 2, 3-Benzotriazine N(3)-oxide

Initial geometry cinnoline N-oxide



N-O = 1.241 Å bisecting the external angle.

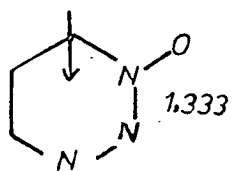
1) N(3) - N(4) shortened

$$\Delta y = -0.03 \text{ \AA}$$

	<u>C(3) - N(4)</u>	<u>T.E. (a. u.)</u>
1)	1.390	-102.845467
2)	1.361	-102.855431
3)	1.331	-102.857676
4)	1.301	-102.854662

Parabolic minimum of 2, 3, 4 gave:-

$$\text{T.E.} = -102.8576897 \text{ a.u. } \text{C(3) - N(4)} = 1.333 \text{ \AA}$$

2) C(5) moved

$$\Delta y = -0.03$$

	N(4) - C(5)	T. E. (a. u.)	$\Delta y$
1)	1.381	-102.857713	0.0
2)	1.365	-102.860850	-0.03
3)	1.349	-102.861186	-0.03
4)	1.333	-102.858863	-0.03

parabolic minimum of 2, 3, 4, gave :-

$$C(5) - N(4) = 1.355 \text{ \AA}$$

$$T. E. = -102.861385 \text{ a. u.}$$

$$B. E. = -8.811244 \text{ a. u.}$$

$$Dipole = 6.19036 \text{ D.}$$

Run (3) was taken as the most suitable geometry.

References

1. K. Siegbahn, C. Nordling and G. Johansson, "ESCA applied to free molecules," North Holland, Amsterdam, 1971.
2. D. W. Davis, D. A. Shirley and T. D. Thomas, J. Amer. Chem. Soc., 1972, 94, 6565.
3. M. Barber and D. T. Clark, JCS Chem. Comm., 1970, 23; ibid, 1970, 24.
4. F. A. Gianturo and C. Guidotti, Chem. Phys. Letters, 1971, 9(6), 539.
5. M. H. Palmer, A. J. Gaskell, P. S. McIntyre and D. W. Anderson, Tetrahedron, 1971, 27, 2921.

## Nuclear Magnetic Resonance Spectroscopy

NMR spectroscopy is used extensively to obtain information on molecular structure. In its commonest form, the structure is elucidated by observing the environment of the protons within the molecule. Two quantities, the chemical shift and the coupling constants of the protons provide useful information, the former giving information about the environment in which the proton exists, and the latter giving information on the arrangement of the protons within the molecule.

In a molecule, the spinning motion of a proton creates a magnetic field, which will interact with an applied magnetic field,  $H_0$ . The proton will be raised or lowered in energy in an applied field depending on the spin of the proton. If the proton spins with the field,  $\alpha$ -spin, then the energy will be lowered, and if it spins against the field,  $\beta$  spin, the energy will be raised. The energy difference,  $\Delta E$ , of these two levels will correspond to a frequency,  $\nu$ , and at  $\Delta E = h\nu$ , the proton with  $\alpha$  spin will be excited to the  $\beta$  spin energy level, and will absorb energy. This will correspond to a line on the NMR spectrum. Different absorptions are obtained for different protons due to the effective field,  $H$ , at the nucleus being different from the applied field,  $H_0$ , the difference arising because of the different chemical environments of the protons.

$$H = H_0 (1 - \sigma)$$

where  $\sigma$  is called the screening constant. The difference in frequencies of two nuclei A and B, is given by:-

$$\begin{aligned} \nu_B - \nu_A &= \nu_0 (1 - \sigma_B) - \nu_0 (1 - \sigma_A) \\ &= \nu_0 (\sigma_B - \sigma_A) = \nu_0 \delta_{AB} \end{aligned}$$

where  $\delta_{AB}$  is the chemical shift, and  $\nu_0$  is the operating frequency.  $\delta_{AB}$  is often expressed as

$$\delta_{AB} = \frac{\nu_B - \nu_A}{\nu_0} \times 10^6 \text{ ppm (parts per million)}$$

A standard such as tetramethyl silane (TMS) is usually used, and



Table 1 Spin combinations of  $\text{CH}_3$  and possible transitions

spin	$\text{CH}_3$ group	degeneracy	transitions
$1\frac{1}{2}$	aaa	1	
$\frac{1}{2}$	$\alpha\alpha\beta, \alpha\beta\alpha, \beta\alpha\alpha$	3	
$\frac{1}{2}$	$\alpha\beta\beta, \beta\alpha\beta, \beta\beta\alpha$	3	
$1\frac{1}{2}$	$\beta\beta\beta$	1	

Table 2 Expressions for  $H_{ii}$  for a two spin system

Spin state, i	$\psi_i$	F	$H_{ii}$
1	aa	1	$\nu_0 (1 - \frac{1}{2} \sigma_A - \frac{1}{2} \sigma_B) + \frac{1}{4} J$
2	$(\alpha\beta + \beta\alpha)/\sqrt{2}$	0	$\nu_0 (-\frac{1}{2} \sigma_A + \frac{1}{2} \sigma_B) - \frac{1}{4} J$
3	$(\alpha\beta - \beta\alpha)/\sqrt{2}$	0	$\nu_0 (\frac{1}{2} \sigma_A - \frac{1}{2} \sigma_B) - \frac{1}{4} J$
4	$\beta\beta$	1	$\nu_0 (-1 + \frac{1}{2} \sigma_A + \frac{1}{2} \sigma_B) + \frac{1}{4} J$

Table 3 Transition energies for an AB system

Transition	$\Delta F$	
3 $\rightarrow$ 1	1	$\frac{1}{2} J + C$
4 $\rightarrow$ 2	1	$-\frac{1}{2} J + C$
2 $\rightarrow$ 1	1	$\frac{1}{2} J - C$
4 $\rightarrow$ 3	1	$-\frac{1}{2} J - C$

Fig.1 AX

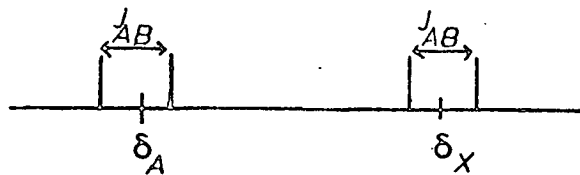


Fig.2  $A_3X$

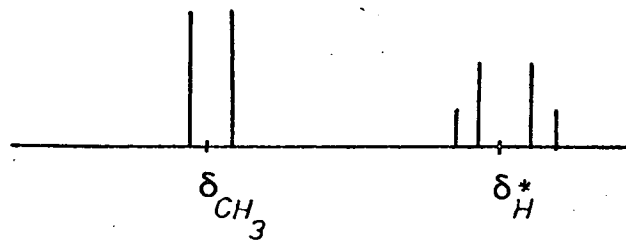
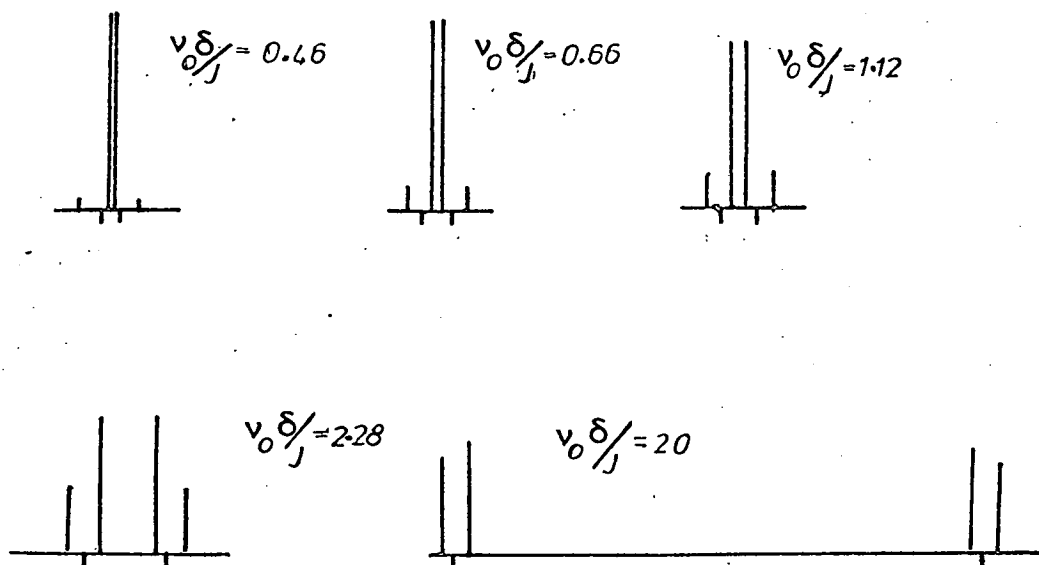


Fig.3



shifts are quoted with reference to the standard.

$$\delta_{A, \text{TMS}} = \frac{\nu_A - \nu_{\text{TMS}}}{\nu_0} \times 10^6 \text{ ppm}$$

Alternatively  $\nu_{\text{TMS}}$  is set at  $10\tau$  giving:-

$$\delta_{A, \text{TMS}} (\tau) = 10 - \delta_{A, \text{TMS}} \text{ ppm}$$

By using these field independent units,  $\delta$  or  $\tau$ , the chemical shift is field independent, and chemical shifts obtained at different operating frequencies can be directly compared.

Chemically non-equivalent sets of spins are usually labelled with letters which are well separated in the alphabet to denote large chemical shift separations. For example, the protons in  $\text{CH}_2\text{ClCHCl}_2$  produce an  $A_2X$  system, and those in  $\text{CH}_3\text{CH}_2\text{Br}$  give  $A_3X_2$ . Sets of spins with similar chemical shifts are labelled with letters close in the alphabet e.g. AB. The coupling constants,  $J$ , arise as a result of spin-spin interaction between nuclei of spin  $\pm\frac{1}{2}$ , and cause single absorption lines to be split into multiplets. The simplest case of spin coupling arises in an AX system. The proton, A, experiences two local fields at the nucleus depending on whether X is in an  $\alpha$  or  $\beta$  spin state, and taking account of the selection rule for a transition,  $\Delta F = \pm 1$  where  $F$  is the total spin quantum number, this causes the single absorption at the chemical shift to be split into two lines, symmetrically placed about the chemical shift. The spectrum will consist of two doublets of equal intensity, Fig. 1.

In a more complicated situation, for example for  $\text{CH}_3-\overset{*}{\text{C}}\text{HY}_2$ , the  $\overset{*}{\text{H}}$  proton experiences four local fields at its nucleus due to the spin combinations of the methyl hydrogens, Table 1. As the mixed  $\alpha/\beta$  states are three times as probable as the pure  $\alpha$  or  $\beta$  states, the ratio of lines in the multiplet is 1:3:3:1 for  $\overset{*}{\text{H}}$ . The methyl group experiences only two fields from the single  $\overset{*}{\text{H}}$  and will appear as a doublet. The general form of the spectrum is shown in Fig. 2 for this  $A_3X$  system. The

total areas under the doublet and quartet will be in the ratio 3:1.

In general, any proton affected by  $n$  other identical nuclei of spin  $\frac{1}{2}$  will produce a multiplet of  $(n+1)$  lines with an intensity ratio represented by the various points marked on Pascal's triangle. These simple multiplets are only obtained in cases where  $\nu_0 \delta \gg |J|$ , and the spectrum is said to be first order under these circumstances.

Second order spectra arise when  $\nu_0 \delta \leq |J|$ , giving overlapping multiplets which alter the line intensities and positions. The magnitude of  $\nu_0 \delta / J$  determines the degree of departure from first order spectra, relatively large values giving an almost first order spectrum. As the coupling constants,  $J$ , are independent of the applied field, a second order spectrum can be considerably simplified by increasing the operating frequency which increases  $\nu_0 \delta$ , and hence  $\nu_0 \delta / J$  becomes larger. This effect is shown in Fig. 3 for an AB system.

To aid in the interpretation of an NMR spectrum, it is possible to obtain a theoretical spectrum by solving the Schrödinger equation,  $H\psi = E\psi$ , where the Hamiltonian,  $H$ , represents the various magnetic interactions, and the  $E$ s will correspond to the energies of the various spin states. The  $\psi$ s are represented as linear combinations of the spin systems. For a two spin system, there are four combinations of spins:- (1)  $\alpha\alpha$ , (2)  $\alpha\beta$ , (3)  $\beta\alpha$ , (4)  $\beta\beta$ . Using the Pauli exclusion principle, the  $\alpha\beta$  pair is represented as linear combinations, (2)  $(\alpha\beta + \beta\alpha) / \sqrt{2}$ , and (3)  $(\alpha\beta - \beta\alpha) / \sqrt{2}$ , and the determinant to be solved is given by:-

$$\begin{vmatrix} H_{11} - E & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} - E & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} - E & H_{34} \\ H_{41} & H_{42} & H_{43} & H_{44} - E \end{vmatrix}$$

The expressions produced for  $H_{ii}$  are shown in Table 2. The only two non-vanishing off diagonal elements are  $H_{23} = H_{32} = \frac{1}{2} J$ .  $E_1$  and

Fig.4

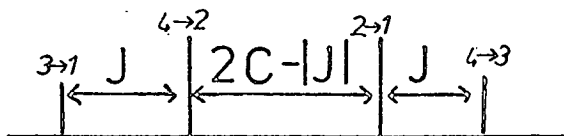


Fig.5

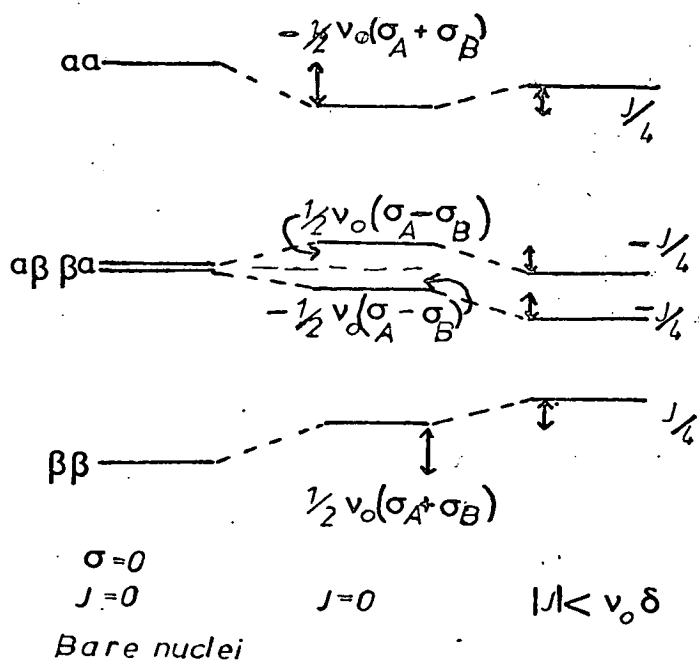
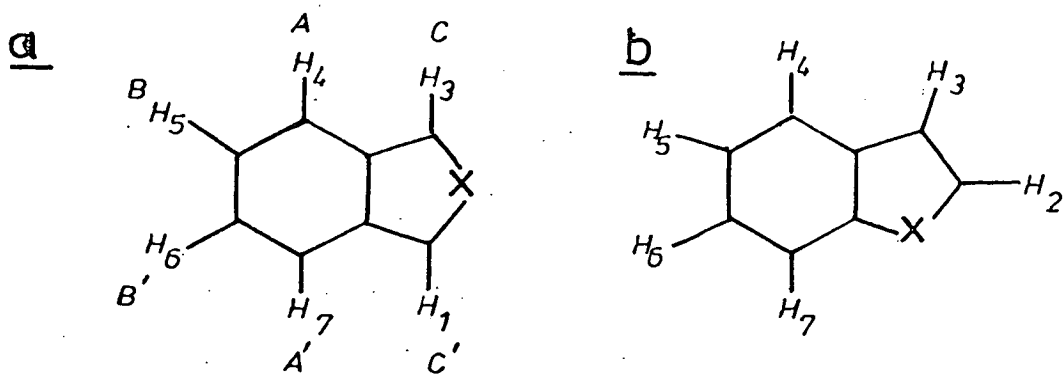


Fig.6



$E_4$  are given directly as:-

$$E_1 = C + 1/4 J$$

$$E_4 = -C + 1/4 J \quad \text{where } C = v_o [1 - (\frac{1}{2}\sigma_A - \frac{1}{2}\sigma_B)]$$

$E_2$  and  $E_3$  are found by solving the determinant

$$\begin{vmatrix} \frac{1}{2}v_o (1 - \sigma_A) & & & \\ & -\frac{1}{2}v_o (1 - \sigma_B) - \frac{J}{4} - E & & \frac{J}{2} \\ & \frac{J}{2} & & \\ & & -\frac{1}{2}v_o (1 - \sigma_A) & \\ & & & \frac{1}{2}v_o (1 - \sigma_B) - \frac{J}{4} - E \end{vmatrix}$$

to give:- 
$$E_2 = \frac{1}{2} [(v_o \delta)^2 + J^2]^{\frac{1}{2}} - \frac{J}{4}$$

$$E_3 = -\frac{1}{2} [(v_o \delta)^2 + J^2]^{\frac{1}{2}} - \frac{J}{4}$$

The transition energies relative to the mean given by  $E_1$ , and satisfying the selection rule  $\Delta F = \pm 1$  are shown in Table 3. These transitions are represented in Fig. 4 for an AB system. The absolute value of the coupling constant,  $|J|$ , is given by the separation of either pair of outside lines, and the separation of the inner lines gives  $2C - |J|$  from which  $\delta$  can be evaluated. The effect of perturbation on the energy levels of an AB system is shown in Fig. 5.

In this work, the main class of molecules under consideration are bicyclic compounds with six protons as shown in Fig. 6a. The method of analysis was by means of the LAMS<sup>†</sup> computer programme. The programme is presented with an initial set of parameters for coupling constants and chemical shifts, and from these generates a set of energy levels corresponding to the different combinations of proton spins by means of the theory already given. These are then used to generate a set of transition energies which are output as frequencies. The calculated frequencies are then assigned to experimental ones on a one to one basis, and the

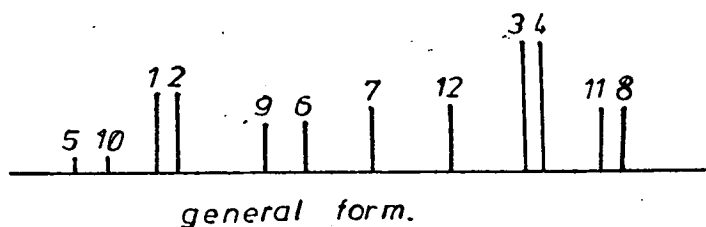
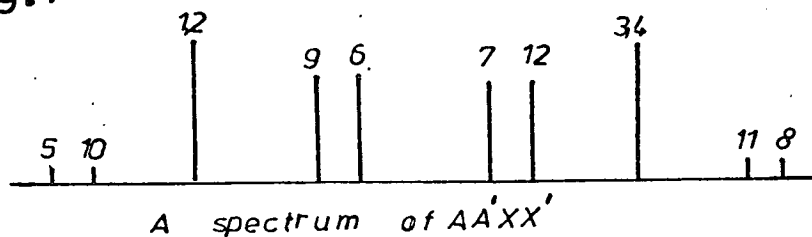
---

<sup>†</sup> Thanks are due to C. Haigh (Department of Chemistry, University College, Swansea) for provision of this programme

Table 4 Energies of A transitions for an  $A_2X_2$  system

Transition	Energy relative to $\nu_A$
1	$\frac{1}{2} N$
2	$\frac{1}{2} N$
3	$-\frac{1}{2} N$
4	$-\frac{1}{2} N$
5	$\frac{1}{2} K + \frac{1}{2} (K^2 + L^2)^{\frac{1}{2}}$
6	$-\frac{1}{2} K + \frac{1}{2} (K^2 + L^2)^{\frac{1}{2}}$
7	$\frac{1}{2} K - \frac{1}{2} (K^2 + L^2)^{\frac{1}{2}}$
8	$-\frac{1}{2} K - \frac{1}{2} (K^2 + L^2)^{\frac{1}{2}}$
9	$\frac{1}{2} M + \frac{1}{2} (M^2 + L^2)^{\frac{1}{2}}$
10	$-\frac{1}{2} M + \frac{1}{2} (M^2 + L^2)^{\frac{1}{2}}$
11	$\frac{1}{2} M - \frac{1}{2} (M^2 + L^2)^{\frac{1}{2}}$
12	$-\frac{1}{2} M - \frac{1}{2} (M^2 + L^2)^{\frac{1}{2}}$

Fig. 7



programme uses an iterative procedure to produce the best fit of experimental and calculated frequencies. The final set of frequencies is used to give values for the coupling constants and chemical shifts.

The basis procedure was to analyse the system initially as a four spin system of AA'BB' type, Fig. 6a, for protons 4 to 7 and then extend the system to a six spin one after a series of non-iterative calculations to establish acceptable values for the inter-ring coupling. An iterative procedure was then used on the AA'BB'CC' system. The basic spectrum consisted of 24 lines, 12 associated with transitions in A, and 12 with transitions in B, and each half of the spectrum being the mirror image of the other. In the special case where  $J_{BB'} > J_{AA'}$  and  $J_{AB} > J_{AB'}$  as is the case with the compounds considered here, Fig. 6a, the spectrum takes the form shown in Fig. 7, assuming  $\sigma_B > \sigma_A$ .

In the limit, where  $\nu_0 \delta \gg |J|$ , an  $A_2X_2$  system is obtained, and is also shown in Fig. 7. The transition energies are shown in Table 4, for an  $A_2X_2$  system, where K, L, M, and N are given as follows:-

$$K = J_{AA'} + J_{BB'}; \quad L = J_{AB} - J_{AB'}$$

$$M = J_{AA'} - J_{BB'}; \quad N = J_{AB} + J_{AB'}$$

To obtain trial values for the chemical shifts and coupling constants the experimental spectra were visually analysed as  $A_2X_2$  systems using the following relationships:-

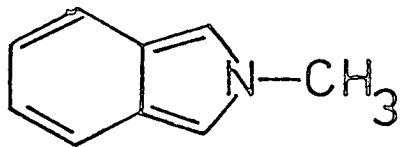
$$E_1 - E_3 = N; \quad (M^2 + L^2)^{\frac{1}{2}} = E_9 - E_{11}$$

$$[(\nu_0 \delta)^2 + N^2]^{\frac{1}{2}} = E_1 + E_3; \quad [(\nu_0 \delta + M)^2 + L^2]^{\frac{1}{2}} = E_9 + E_{11}$$

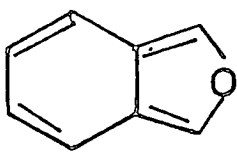
where the E values correspond to the transition frequencies of the numbered transitions in Fig. 7. From these the chemical shifts  $\nu_0 \delta_A$  and  $\nu_0 \delta_B$ , and the coupling constants  $J_{AB}$  and  $J_{AB'}$  can be evaluated.



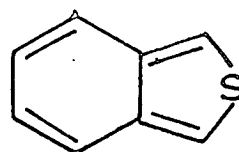
Fig.8



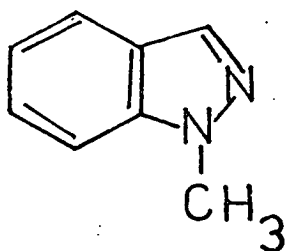
1



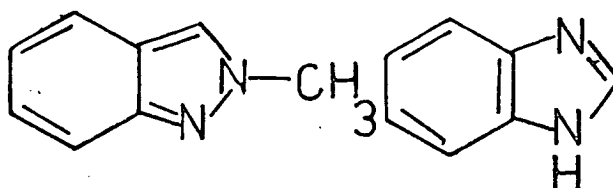
2



3

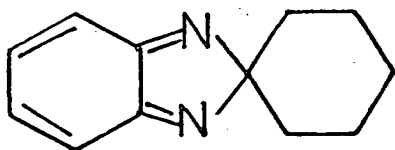


4

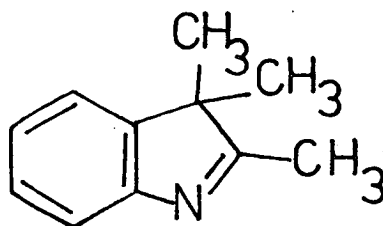


5

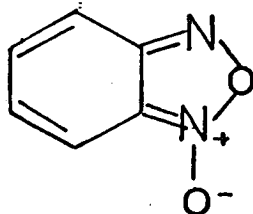
6



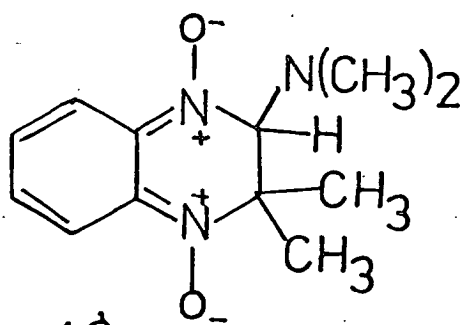
7



8



9



10

Since the operating frequency,  $\nu_0$ , always multiplies the chemical shifts, the coupling constants and the chemical shifts are expressed in cycles per second (c/s) or Hertz (Hz), rather than ppm.

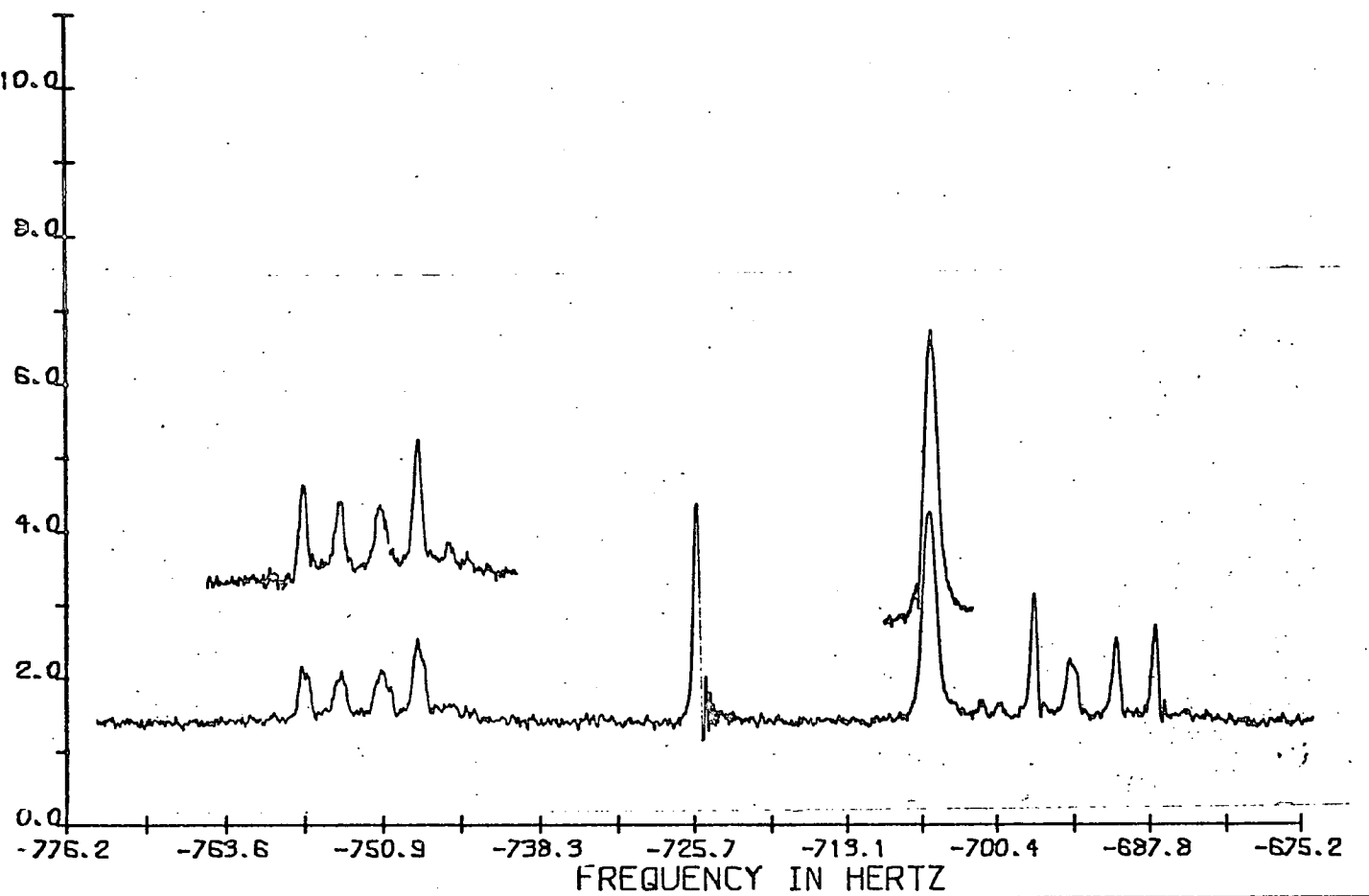
Results and discussion. The spectra were run wherever possible in  $\text{CDCl}_3$  as solvent to standardise any solvent effects on the chemical shifts, and were obtained on a Varian Associates HA 100 spectrometer. The molecules whose  $^1\text{H-NMR}$  spectra were analysed are shown in Fig. 8.

The computed and experimental spectra of the quinoid molecules of the bicyclic heterocycles I(1-3) are shown in Fig. 9-11. For N-methyl isoindole, Fig. 9, the experimental spectrum indicated a basically AA'BB' 24 line spectrum, the downfield set of frequencies being further split by inter-ring coupling with the C protons, Fig. 6a. These C protons appear as a broad singlet which becomes sharper on irradiation of the methyl protons, but the inter-ring coupling cannot be removed by irradiation of the C protons due to the proximity of their chemical shift to that of the B protons in the benzenoid ring. The spectrum could not therefore be further simplified, and was analysed initially as an AA'BB' system which was later extended to an AA'BB'CC' spectrum. The spectrum was run at  $0^\circ\text{C}$  and under nitrogen to prevent oxidation.

The spectra of benzo(c) furan and benzo(c) thiophen, Figs. 10 and 11, were run at  $0^\circ\text{C}$  and  $-20^\circ\text{C}$  respectively and under nitrogen. The former was analysed in two ways; first, as a four spin system which was extended to six spins; second, directly as a six spin system. The chemical shift of the C protons, Fig. 6a, appear well down-field of the others, and are sufficiently far removed ( $\approx 55$  Hz) to be decoupled allowing good estimations of coupling constants and chemical shifts for the other protons. The results of the two analyses are both acceptable within experimental limits, but the main difference between the two sets is in the meta,  $J_{\text{AB}'}$  and para,  $J_{\text{AA}'}$ , coupling constants in the benzenoid ring. In analysis 1, these are given as 0.798 and 0.837 Hz, while in analysis 2, they are given

Fig.9

N-METHYL ISOINDOLE EXPTL. NMR SPECTRUM



N-METHYL ISOINDOLE THEORETICAL NMR SPECTRUM

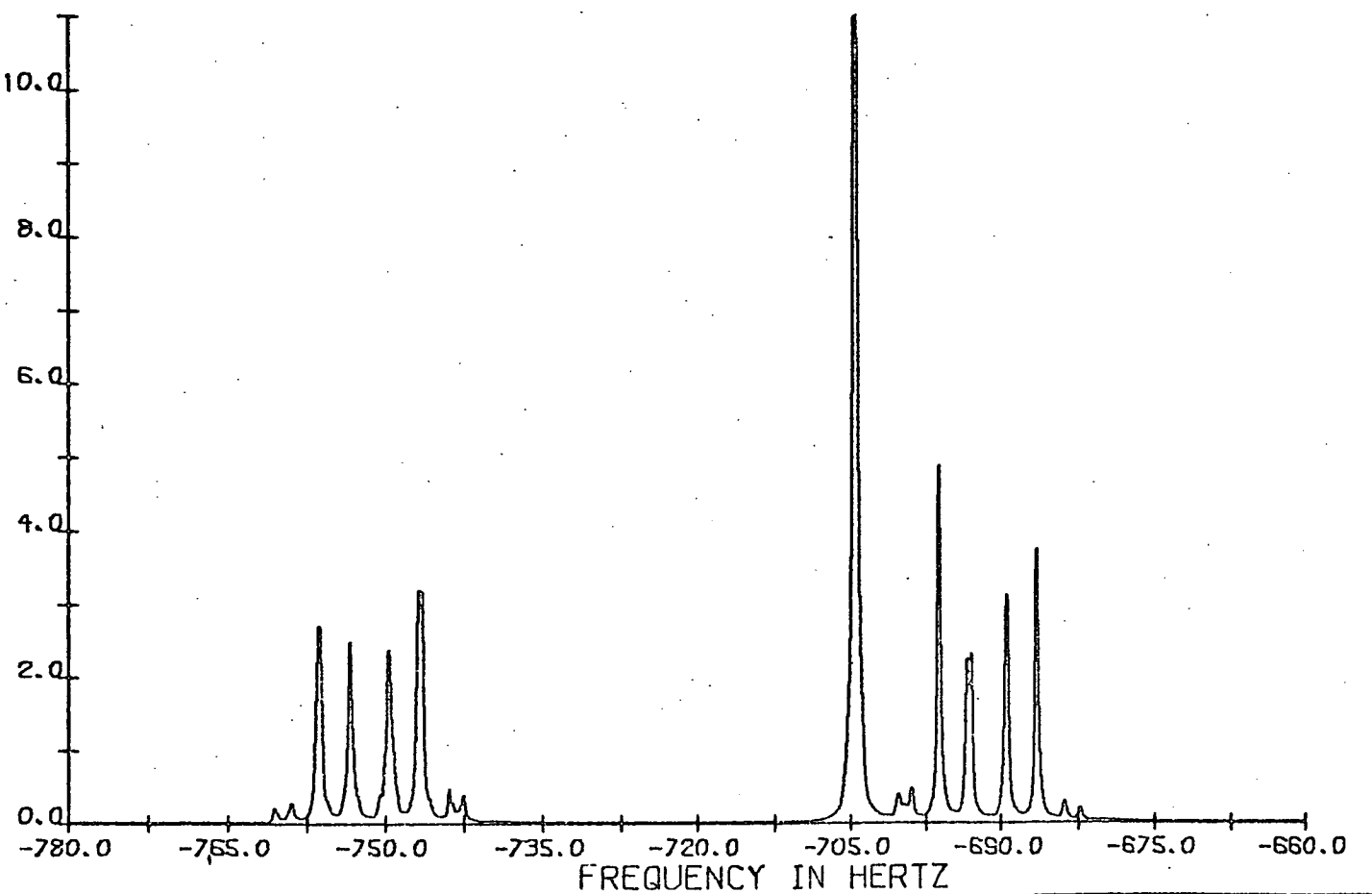
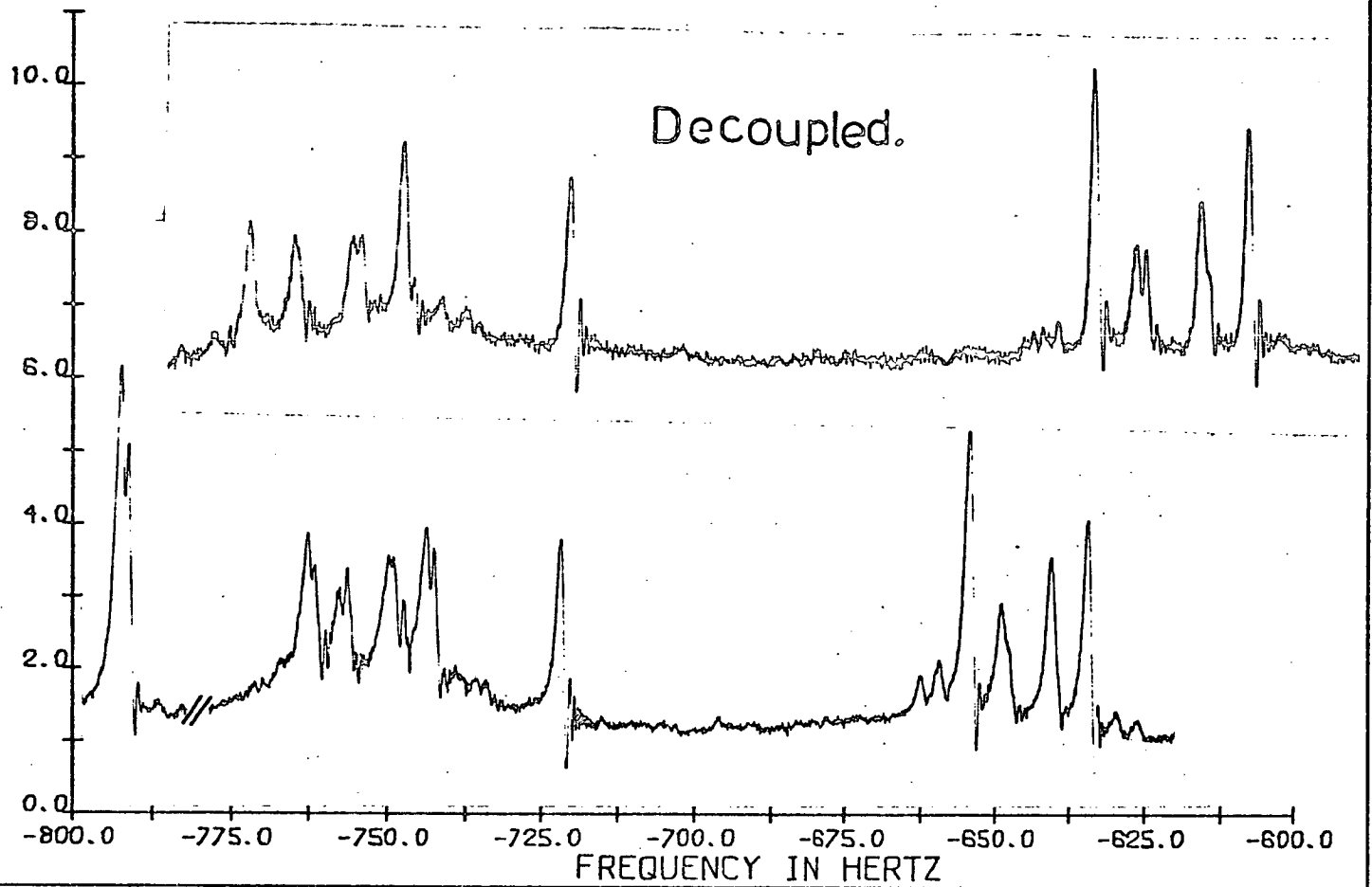


Fig.10

### ISOBENZOFURAN EXPT. NMR SPECTRUM



### ISOBENZOFURAN THEORETICAL NMR SPECTRUM

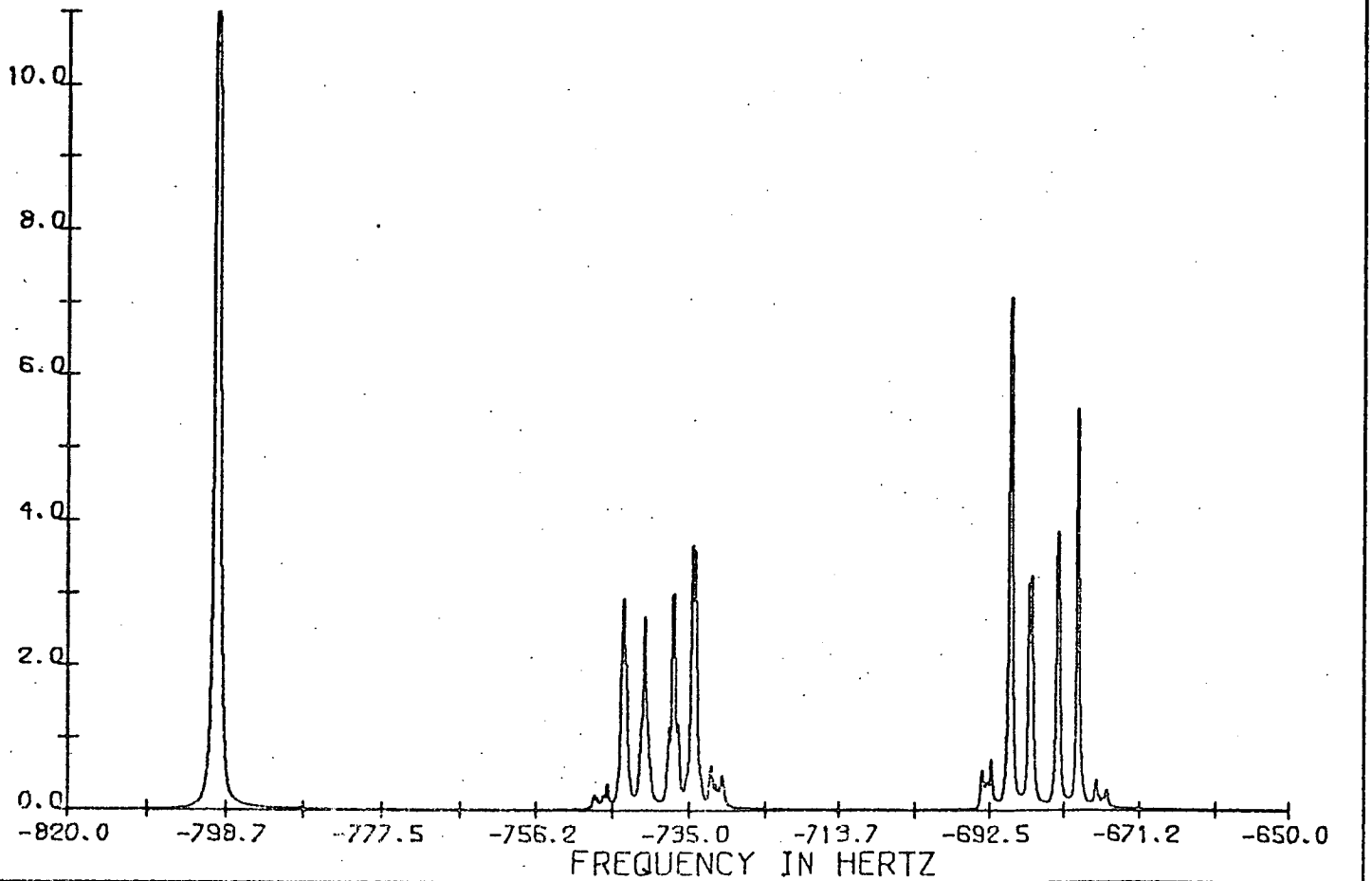
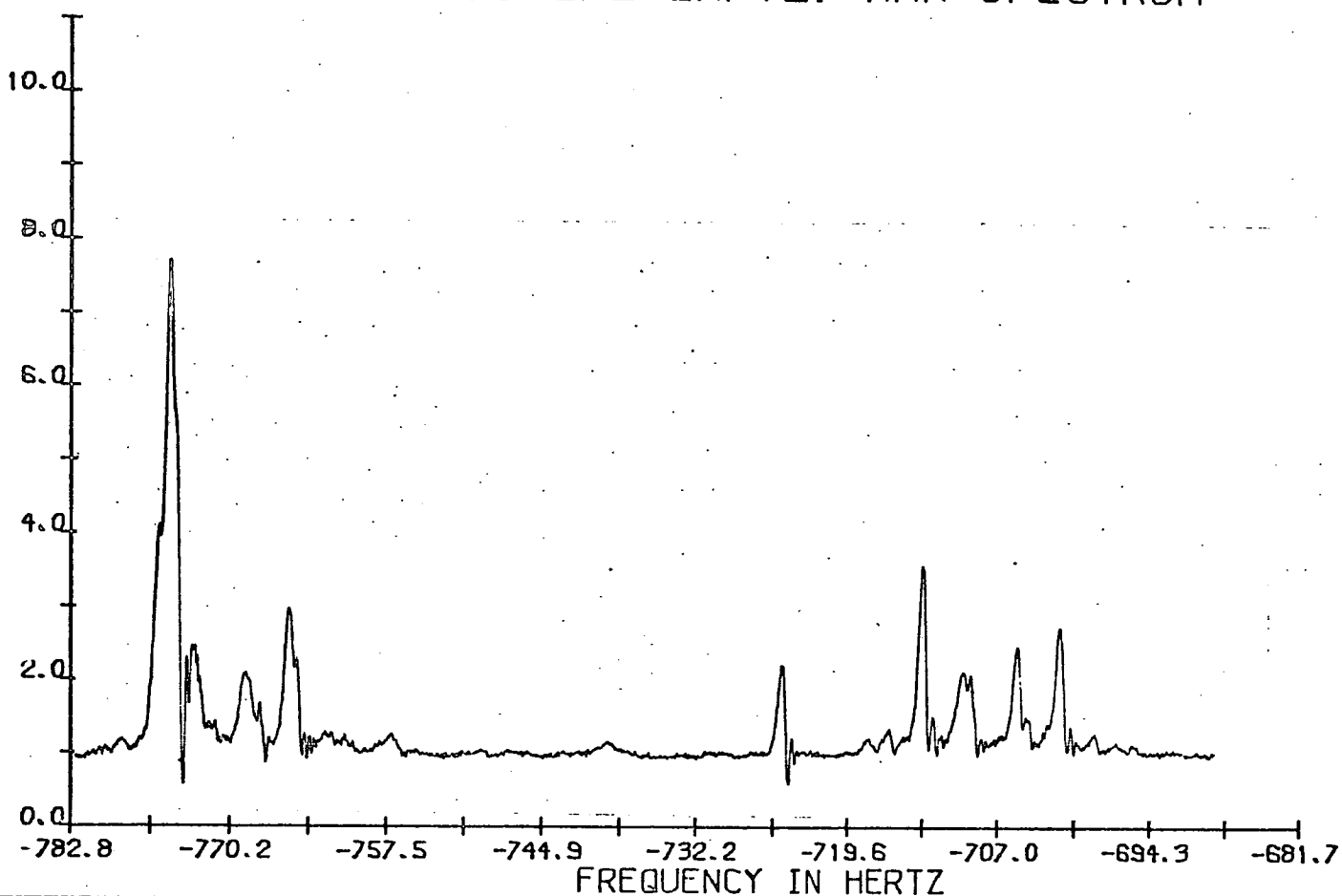
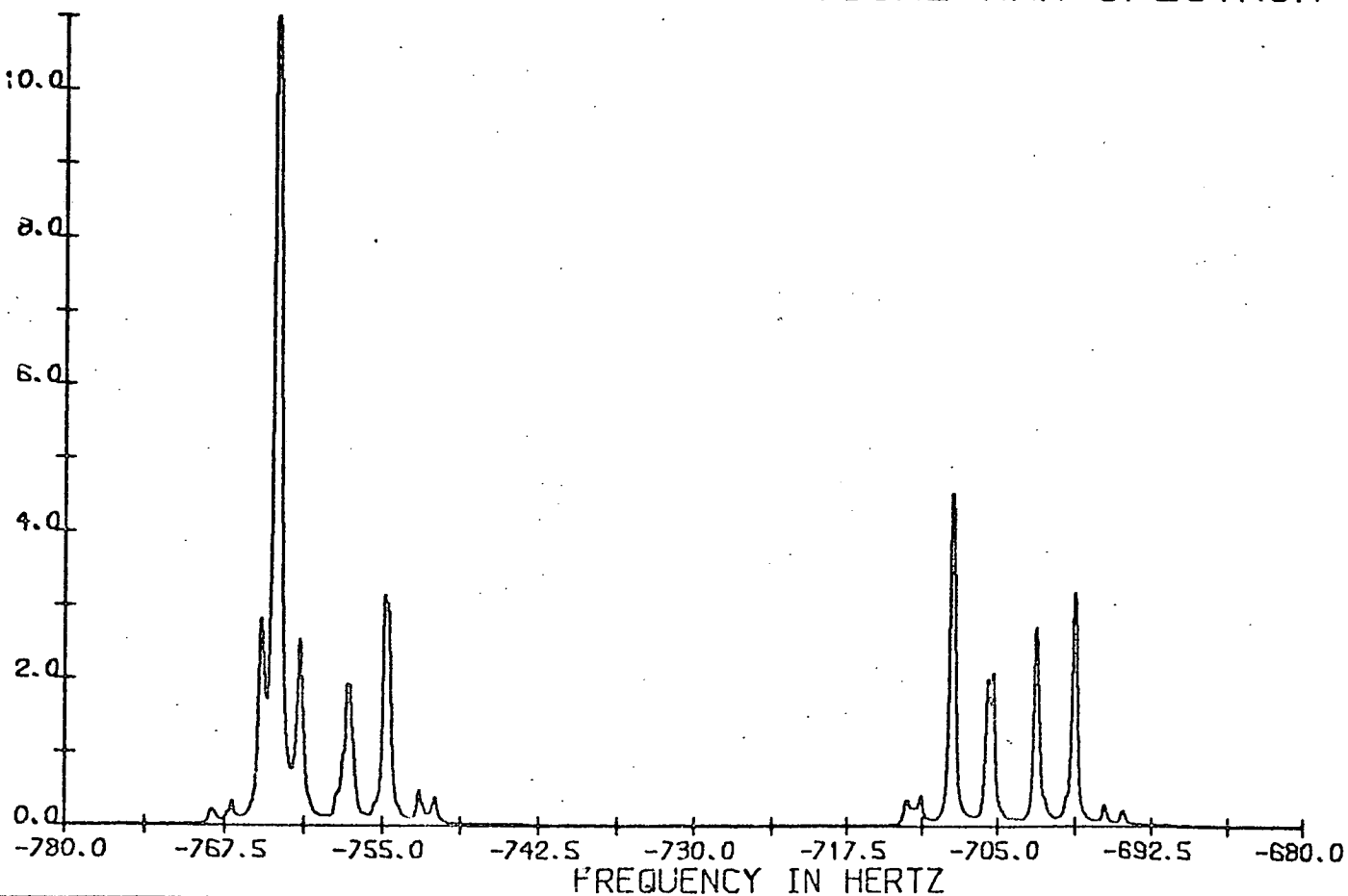


Fig. 11

## ISOBENZOTHIOPENE EXPTL. NMR SPECTRUM



## ISOBENZOTHIOPHENE THEORETICAL NMR SPECTRUM



by 1.008 and 0.570 Hz.

Benzo(c) thiophen again gave a basically AA'BB' spectrum with the A and C protons appearing in the same regions, making removal of the C proton signal by double irradiation impossible. Again inter-ring coupling is shown, and analyses was by means of a four and then six spin system. It was necessary with this compound to run the spectrum at  $-20^{\circ}\text{C}$ , as rapid decomposition took place in solution even at  $0^{\circ}\text{C}$ . An additional precaution of treating the  $\text{CDCl}_3$  with aqueous sodium dithionite solution to remove oxygen was also taken in this case to minimise decomposition by oxidation.

A summary of the final results obtained is given in Table 5, with the numbering system as in Fig. 6a

The chemical shifts of the C protons are  $-798.90$ ,  $-763.40$  and  $-704.54$  for  $X = \text{O}$ , S and N respectively. The very low shift of the C protons when  $X = \text{O}$  (Fig. 6a) can be explained by the high electronegativity of the oxygen atom. From electronegativity considerations of X, the expected chemical shift order of the C protons from low to high is  $\text{O} < \text{NMe} < \text{S}$ . The C protons when  $X = \text{NMe}$  appear to be out of place, but this is possibly due to the electron donating ability of the N-methyl group, resulting in more shielding of the protons. For the S compound, the chemical shifts of the C and A protons are very close indicating minimal deshielding by the S atom, and lending support to the theory that an S atom is electronically similar to  $-\text{CH}=\text{CH}-$ .

In the Kekulé series, benzofuran,<sup>5</sup> benzothiophen<sup>6</sup> and indole<sup>5</sup> have all been analysed at 100 MHz, and all show inter-ring coupling, especially across the 3, 7 positions, Fig. 6b. The quinoid series also show inter-ring coupling, and the values of  $J_{3,7}$  are shown for both sets of molecules in Table 6. In general the inter-ring coupling constant is smaller for the quinoid series as might be expected in systems where the delocalisation of  $\pi$ -electrons is likely to be peripheral compared with molecules which have a benzenoid ring.

In the bicyclic heterocycles II, the spectra of 1 and 2-methyl indazole

Table 5.  $^1\text{H}$  NMR parameters

$^1\text{H}$ Shifts	$\nu_1$	$\nu_4$	$\nu_5$
Benzo[c]furan <sup>a</sup> (i)	-798.90	-738.07	-683.90
(standard deviation)	(0.024)	(0.027)	(0.030)
Benzo[c]thiophen <sup>b</sup>	-763.40	-759.30	-704.16
(standard deviation)	(0.036)	(0.032)	(0.034)
N-Methylisoindole <sup>c</sup>	-704.54	-751.21	-691.72
(standard deviation)	(0.019)	(0.022)	(0.023)

$^1\text{H}$ Couplings	$J_{1,3}$	$J_{1,4}$	$J_{1,5}$	$J_{1,6}$	$J_{1,7}$
Benzo[c]furan <sup>a</sup>	-0.001	0.637	0.039	0.044	0.014
(standard deviation)	(0.035)	(0.039)	(0.046)	(0.046)	(0.039)
Benzo[c]thiophen <sup>b</sup>	0.017	0.422	0.030	-0.139	-0.086
(standard deviation)	(0.040)	(0.056)	(0.068)	(0.071)	(0.058)
N-Methylisoindole <sup>c</sup>	0.0	0.457	0.070	-0.070	-0.081
(standard deviation)	(0.040)	(0.035)	(0.042)	(0.042)	(0.035)

	$J_{4,5}$	$J_{4,6}$	$J_{4,7}$	$J_{5,6}$
	8.52	1.008	0.570	6.223
	(0.049)	(0.048)	(0.40)	(0.048)
	8.64	1.030	0.785	6.356
	(0.056)	(0.057)	(0.048)	(0.060)
	8.69	0.899	0.790	6.456
	(0.039)	(0.039)	(0.034)	(0.039)

a. Concentration  $7.7 \text{ gl}^{-1}$  in  $\text{CDCl}_3$  ( $273^\circ\text{K}$ ) under  $\text{N}_2$

b. Concentration  $7.3 \text{ gl}^{-1}$  in  $\text{CDCl}_3$  ( $253^\circ\text{K}$ ) under  $\text{N}_2$

c. Concentration  $16 \text{ gl}^{-1}$  in  $\text{CDCl}_3$  ( $273^\circ\text{K}$ ) under  $\text{N}_2$

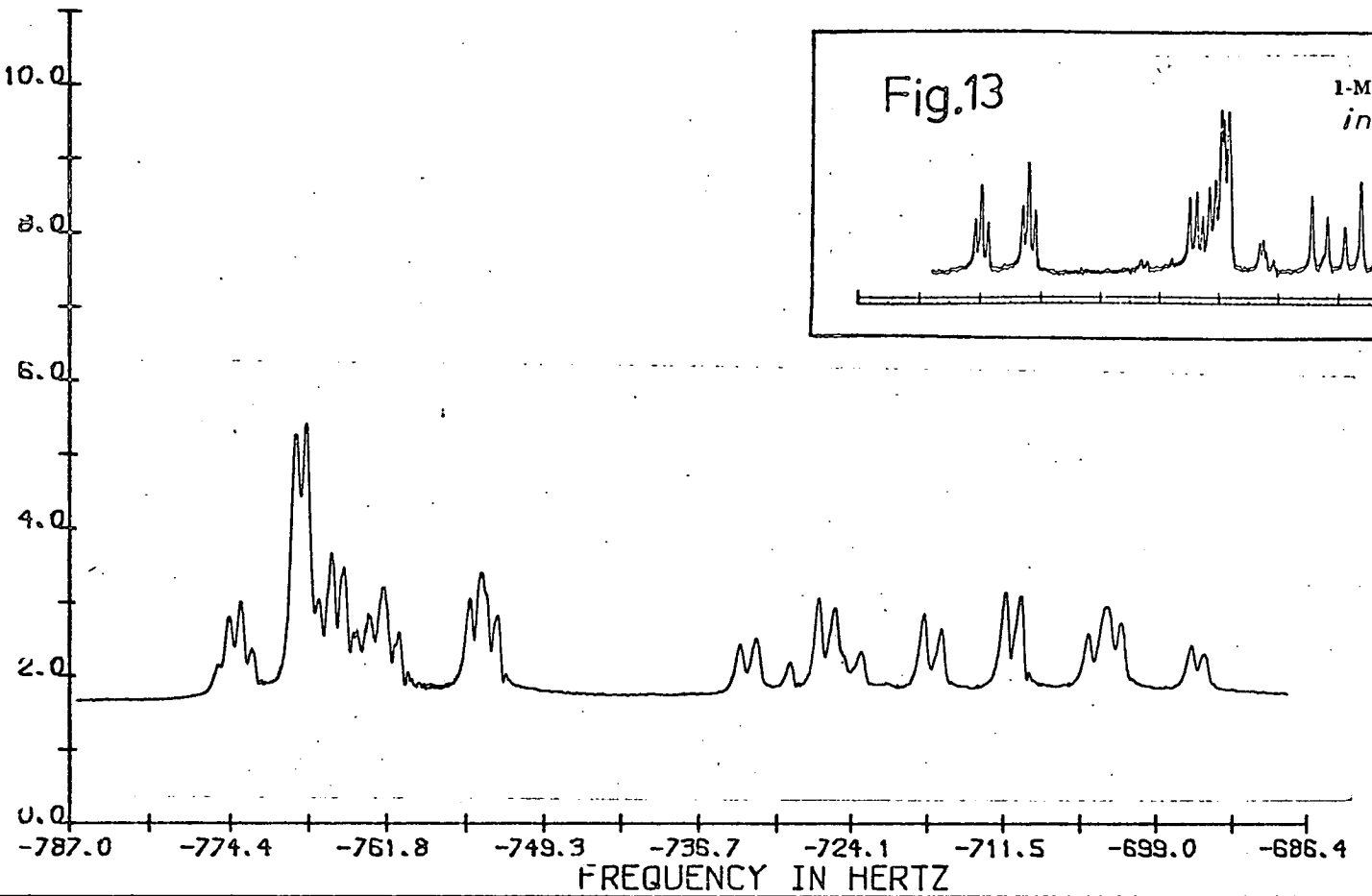
Table 6. Coupling constants for the quinoid and Kekulé molecules

	$J_{3,7}$	$J_{4,5}$	$J_{5,6}$	$J_{6,7}$
Indole	0.70	7.84	7.07	8.07
Benzofuran	0.87	7.89	7.27	8.43
Benzothiophen	0.86	8.09	7.22	8.06
N-Methyl isoindole	0.46	8.69	6.46	8.69
Benzo(c)furan	0.64	8.52	6.22	8.52
Benzo(c)thiophen	0.42	8.64	6.36	8.64



Fig. 12

2-METHYL INDAZOLE EXPTL NMR SPECTRUM



2-METHYL INDAZOLE THEORETICAL NMR SPECTRUM

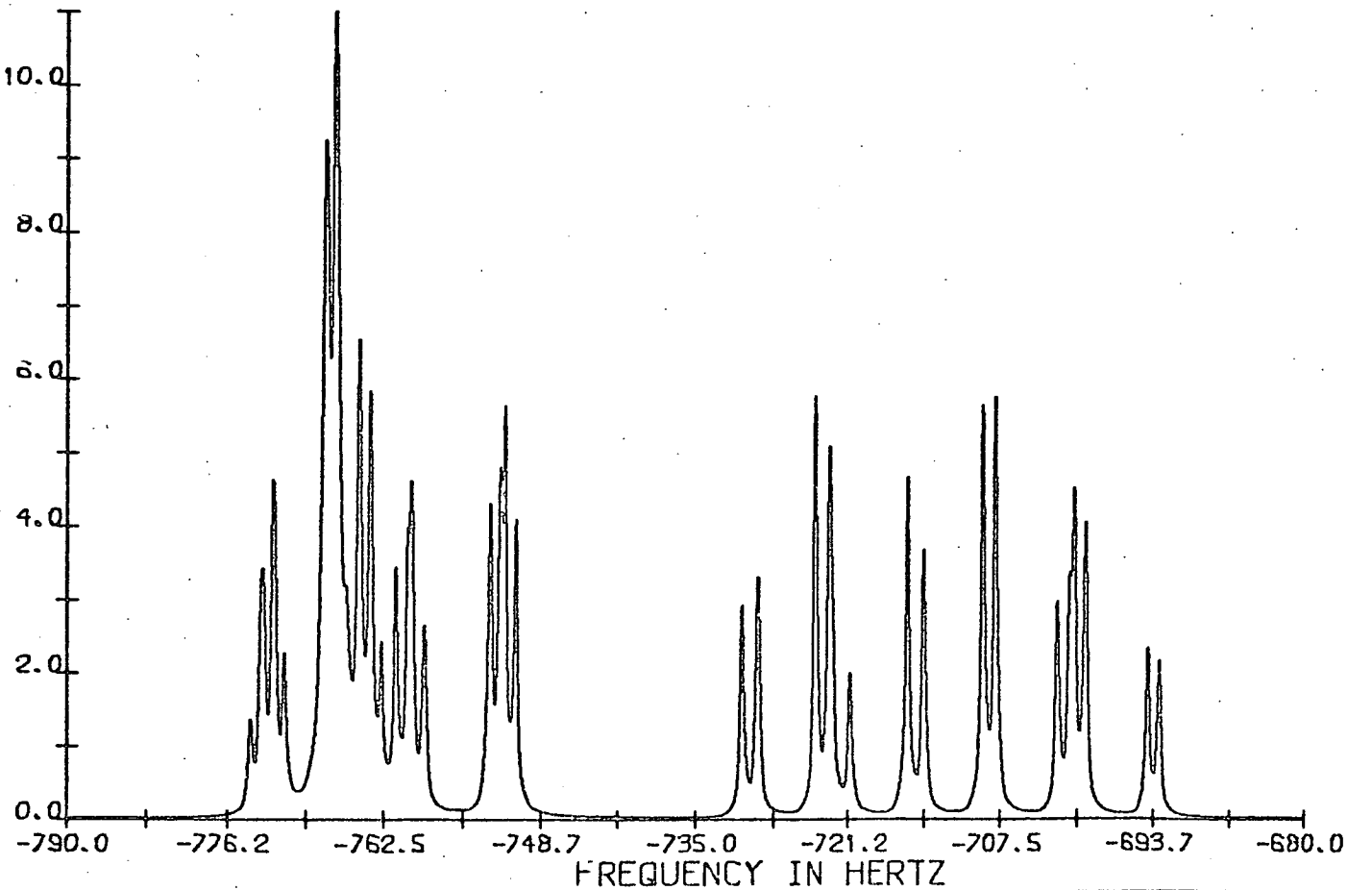


Fig.14

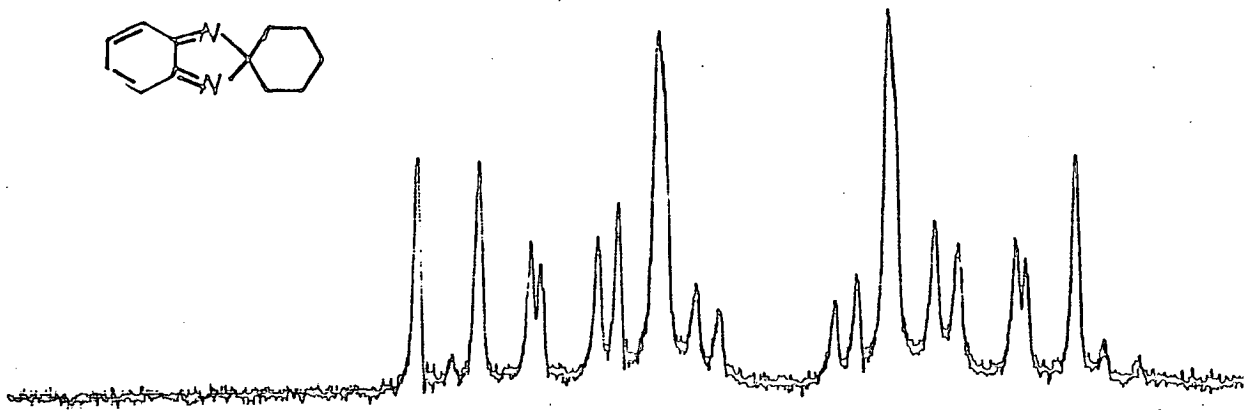
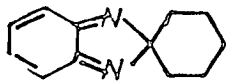


Fig.15

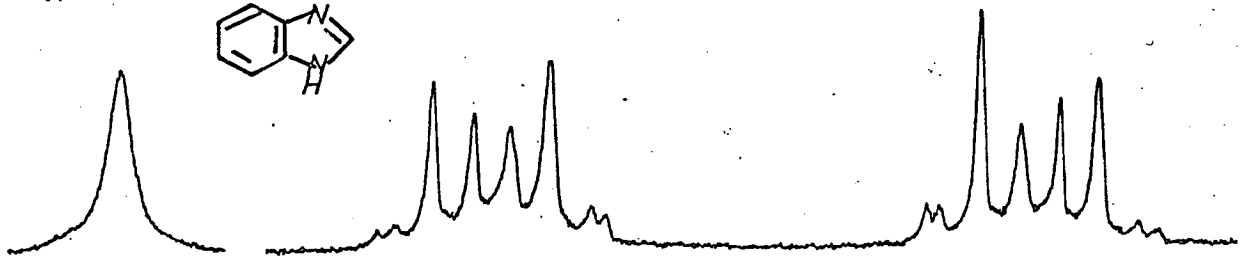
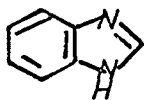
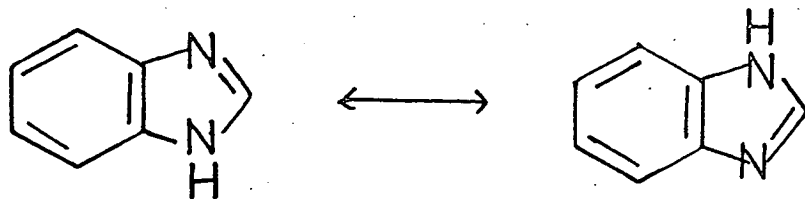


Fig.16



(4, 5) benzimidazole, (6), and 2-spirocyclohexyl benzimidazole, (7), have been analysed, Fig. 8. The computed data for these molecules is shown in Table 7.

The computed and experimental spectra of 2-methyl indazole are shown in Fig. 12, and the experimental spectrum of 1-methyl indazole in Fig. 13. Although the two spectra look very different, examination of the chemical shifts indicates that the main difference between the two is an upfield shift of H(3), and a downfield shift of H(7) in the 2-methyl compound. These shifts are caused by the nitrogen lone pair in this molecule giving deshielding at H(7), and a similar effect on H(3) in the 1-methyl compound, the latter effect being removed in the 2-methyl compound. The ortho coupling constants in the hydrocarbon ring of each molecule are similar. This is in agreement with resonance energy data which indicated that the two parent compounds were similar in their aromaticity with resonance energies of 66 and 56 kJ mol<sup>-1</sup> for the 1H and 2H compounds respectively. The coupling constants obtained here for these molecules are in good agreement with those already published.<sup>7</sup>

For benzimidazole and 2-spirocyclohexyl-2H-benzimidazole, the NMR spectra, Figs. 14 and 15, yield very different values for coupling constants, Table 7. It is interesting to note that for the spiro compound  $J_{\text{para}}$  is larger than  $J_{\text{meta}}$  by  $\approx 0.3$  Hz. In molecules with a high degree of aromaticity  $J_{\text{meta}}$  is usually larger than  $J_{\text{para}}$ , but molecules with obviously fixed alternating bond structures often show the opposite effect, Table 8. This suggests that the spiro compound structure consists of a high degree of alternant hydrocarbon type structure.

The spectrum of benzimidazole has been analysed by Black and Heffernan,<sup>9</sup> and on comparison with the analysis obtained here, Table 9, the only obvious discrepancy is in  $J_{4,6}$ , which is given by these workers as 1.4 Hz and is found here to be 0.95 Hz.

The spectra of both these molecules are AA'BB' in form. For benzimidazole this arises due to the averaging effect of the tautomers in solution, Fig. 16.

Table 7  $^1\text{H}$  - NMR parameters

$^1\text{H}$ shifts ( $\nu_{\text{O}} \delta$ )	$\nu_2$	$\nu_3$	$\nu_4$
(4) 1-Methyl indazole (0.03) <sup>†</sup>	-	-793.35	-770.71
(5) 2-Methyl indazole (0.04)	-	-767.2	-755.91
(6) Benzimidazole (0.04)	-	-	-765.11
(7) 2-Spirocyclohexyl-2H-benzimidazole (0.03)	-	-	-721.61
(8) 1, 2, 2'-Trimethyl indole (0.04)	-	-	-726.12
(9) Benzofuroxan (0.07)	-	-	-757.67
(10) 2, 2'-Dimethyl, 3-dimethyl amino, 2, 3-dihydro quinoxaline di-N-oxide (0.07)	-	-	-728.11

	$\nu_5$	$\nu_6$	$\nu_7$
(4)	-709.54	-733.74	-758.86
(5)	-702.03	-722.51	-768.07
(6)	-721.57	-721.57	-765.11
(7)	-699.62	-699.62	-721.61
(8)	-718.06	-728.44	-752.88
(9)	-746.06	-761.88	-782.25
(10)	-668.21	-671.36	-733.77

	Concentration ( $\text{gl}^{-1}$ )	Solvent
(4)	200 $\pm$ 20	$\text{CDCl}_3$
(5)	200 $\pm$ 20	$\text{CDCl}_3$
(6)	71.4	$(\text{CD}_3)_2\text{CO}$
(7)	71.4	$\text{CDCl}_3$
(8)	-	$\text{CDCl}_3$
(9)	135.2	$(\text{CD}_3)_2\text{CO}$ at $-60^\circ\text{C}$
(10)	-	

Table 7 (cont.)

<sup>1</sup>H coupling constants

	$J_{3,7}$	$J_{4,5}$	$J_{4,6}$	$J_{4,7}$	$J_{5,6}$	$J_{5,7}$	$J_{6,7}$
(4) (0.04)	0.86	8.81	0.89	0.97	6.87	0.86	8.38
(5) (0.04)	0.89	8.35	1.14	1.04	6.71	0.85	8.52
(6) (0.07)	-	8.37	0.95	0.78	7.02		
(7) (0.04)	-	9.40	0.99	1.28	6.07		
(8) (0.07)	-	7.38	1.17	0.67	7.48	1.15	7.48
(9) (0.09)	-	9.16	1.11	1.04	6.36	0.86	9.01
(10) (0.11)	-	9.24	1.20	0.87	6.46	0.99	9.26

<sup>+</sup> standard deviations given in brackets.

Table 8 Meta and para coupling constants of some hydrocarbons

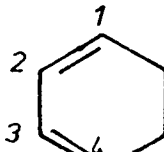
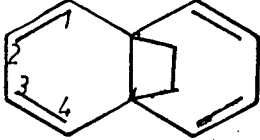
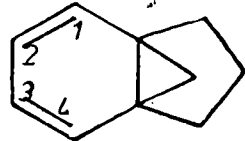
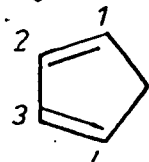
	$J_{1,3}^8$	$J_{1,4}^8$
	1.02	1.12
	0.72	1.09
	0.58	1.31
	1.09	1.93

Table 9 Comparison of  $^1\text{H}$ -parameters for benzimidazole

	$J_{4,5}$	$J_{4,6}$	$J_{4,7}$	$J_{5,6}$
This work	8.37	0.95	0.78	7.02
Reference 9	8.20	1.40	0.70	7.10

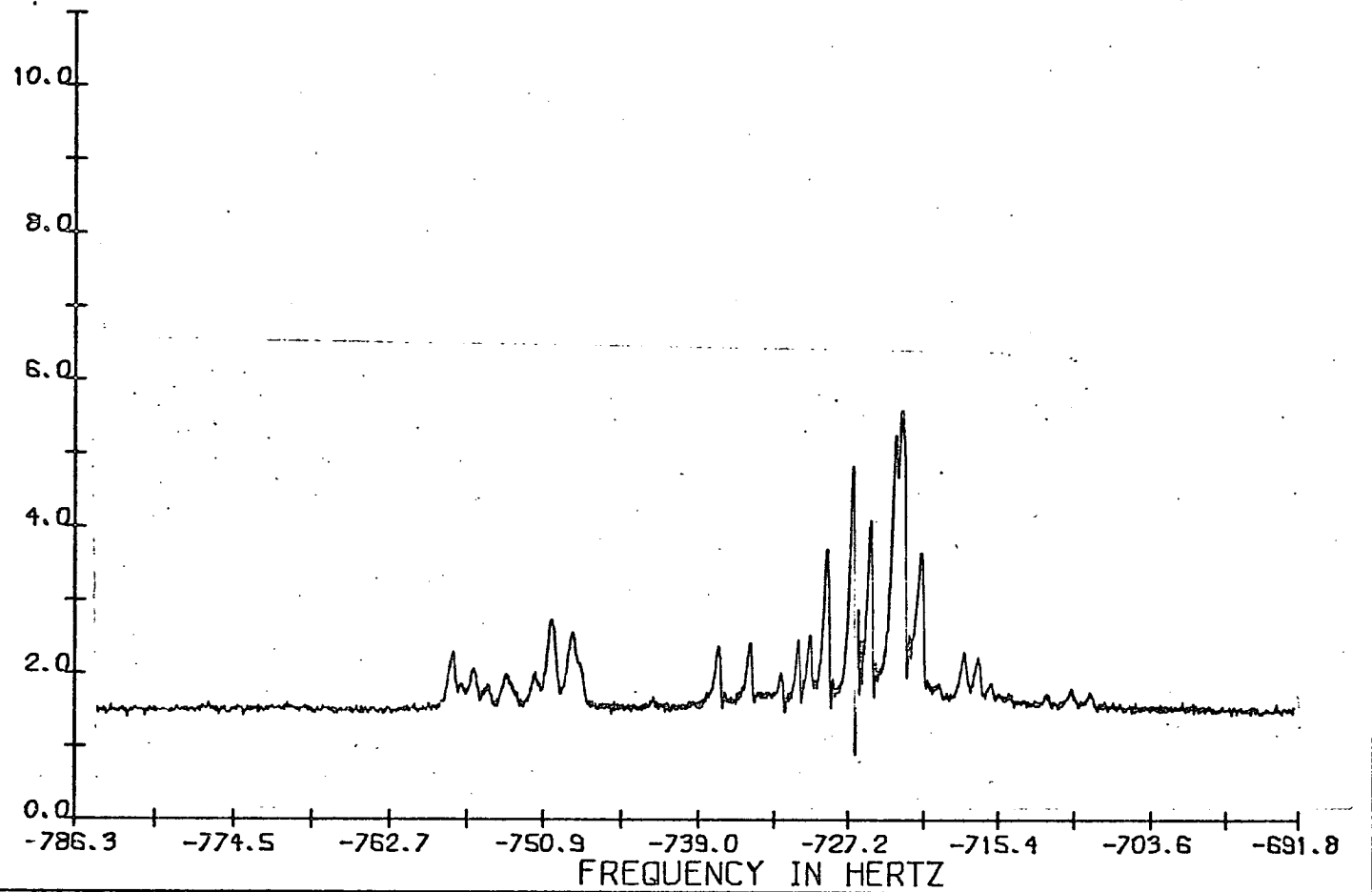
Table 10  $^1\text{H}$ -NMR parameters for 1, 2, 2'-trimethyl indole at 220 MHz

$^1\text{H}$ shifts	Standard Deviation
H-4 -1598.09	0.193
H-5 -1579.72	0.191
H-6 -1603.08	0.196
H-7 -1657.97	0.170

$^1\text{H}$ coupling constants		
J	Value	Standard Deviation
4, 5	7.53	0.293
4, 6	1.01	0.212
4, 7	0.38	0.295
5, 6	6.81	0.294
5, 7	0.88	0.256
6, 7	7.68	0.296

Fig. 17

2, 3, 3<sup>o</sup> -TRIMETHYL, 3(H)-INDOLE EXPTL. NMR



2, 3, 3<sup>o</sup> -TRIMETHYL 3(H)-INDOLE THEORETICAL NMR

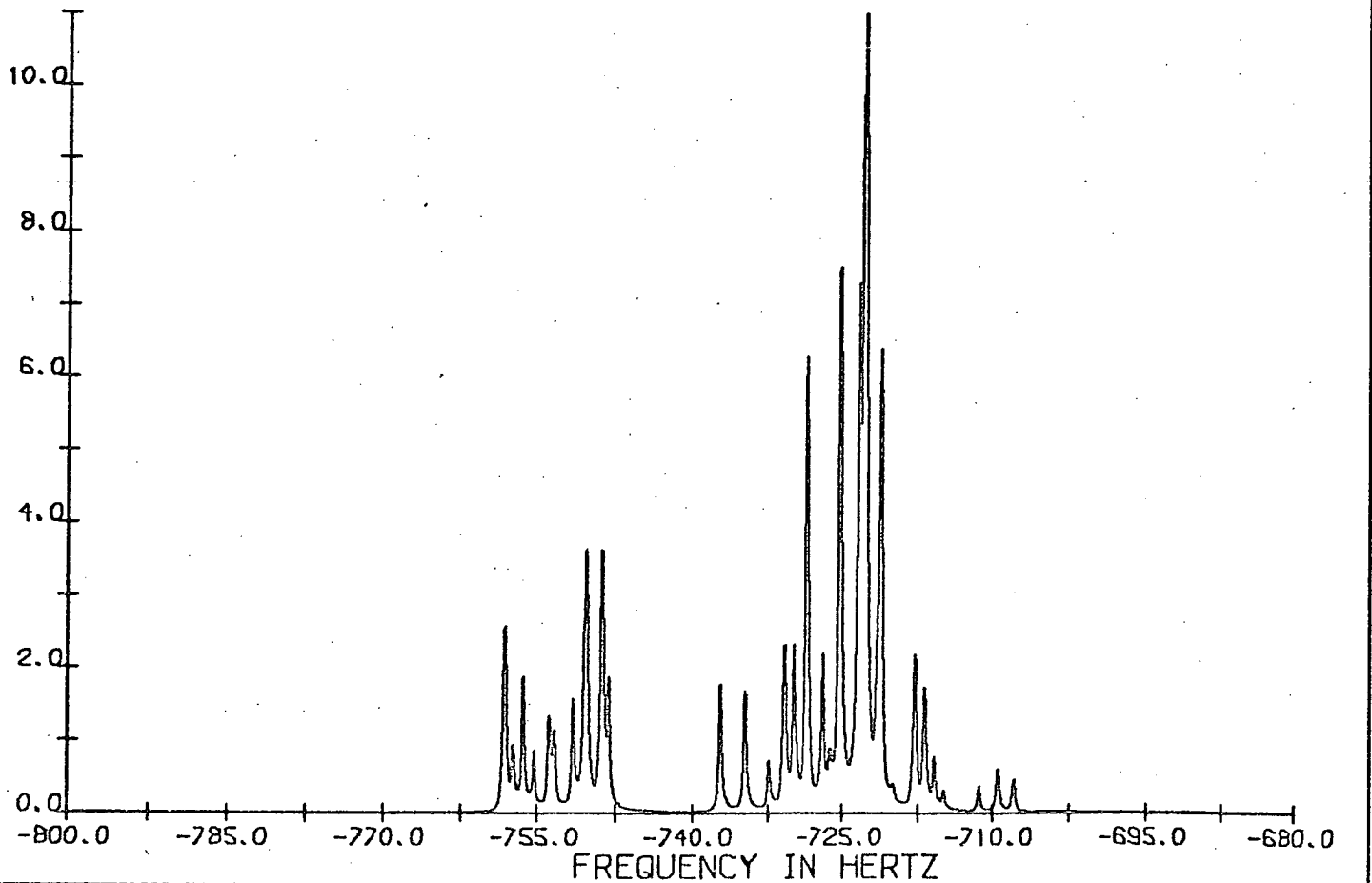


Fig.18

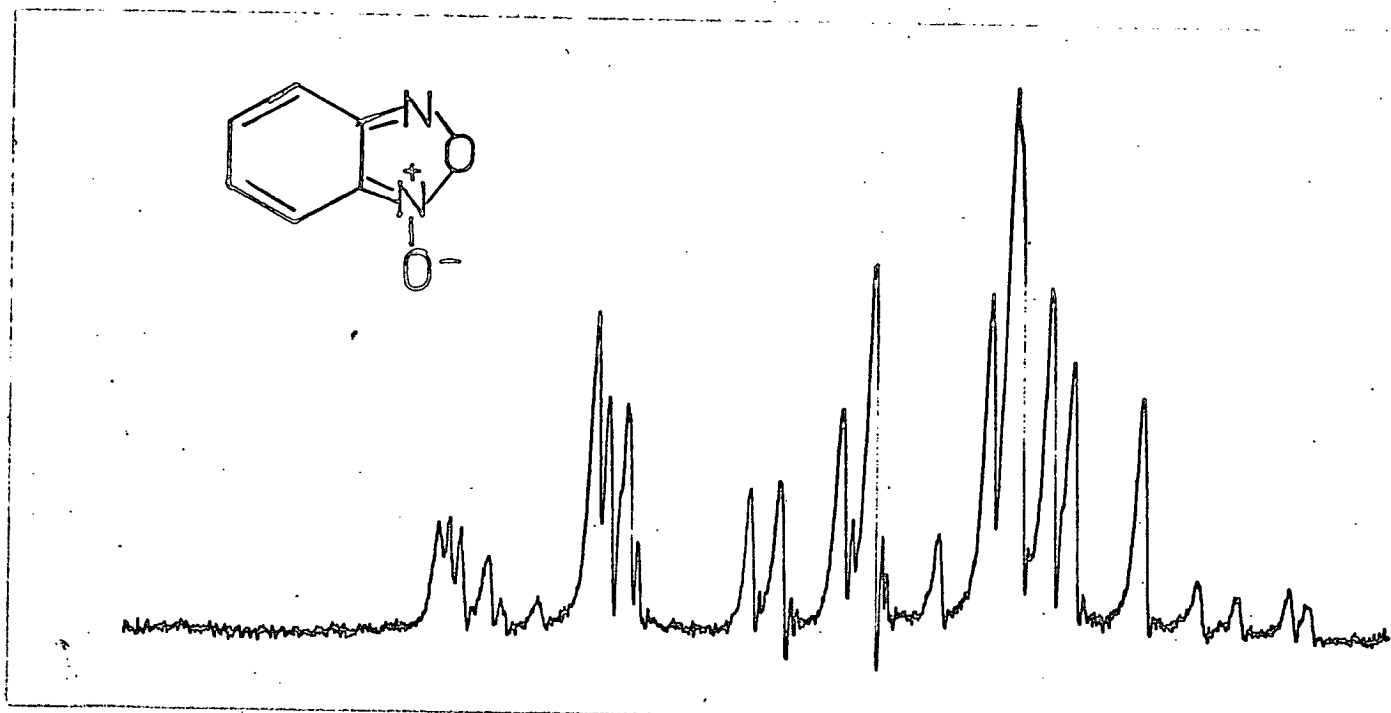


Fig.19

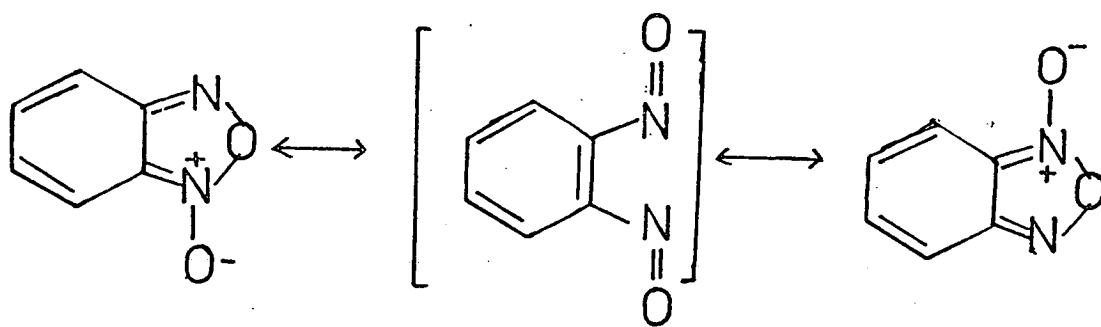
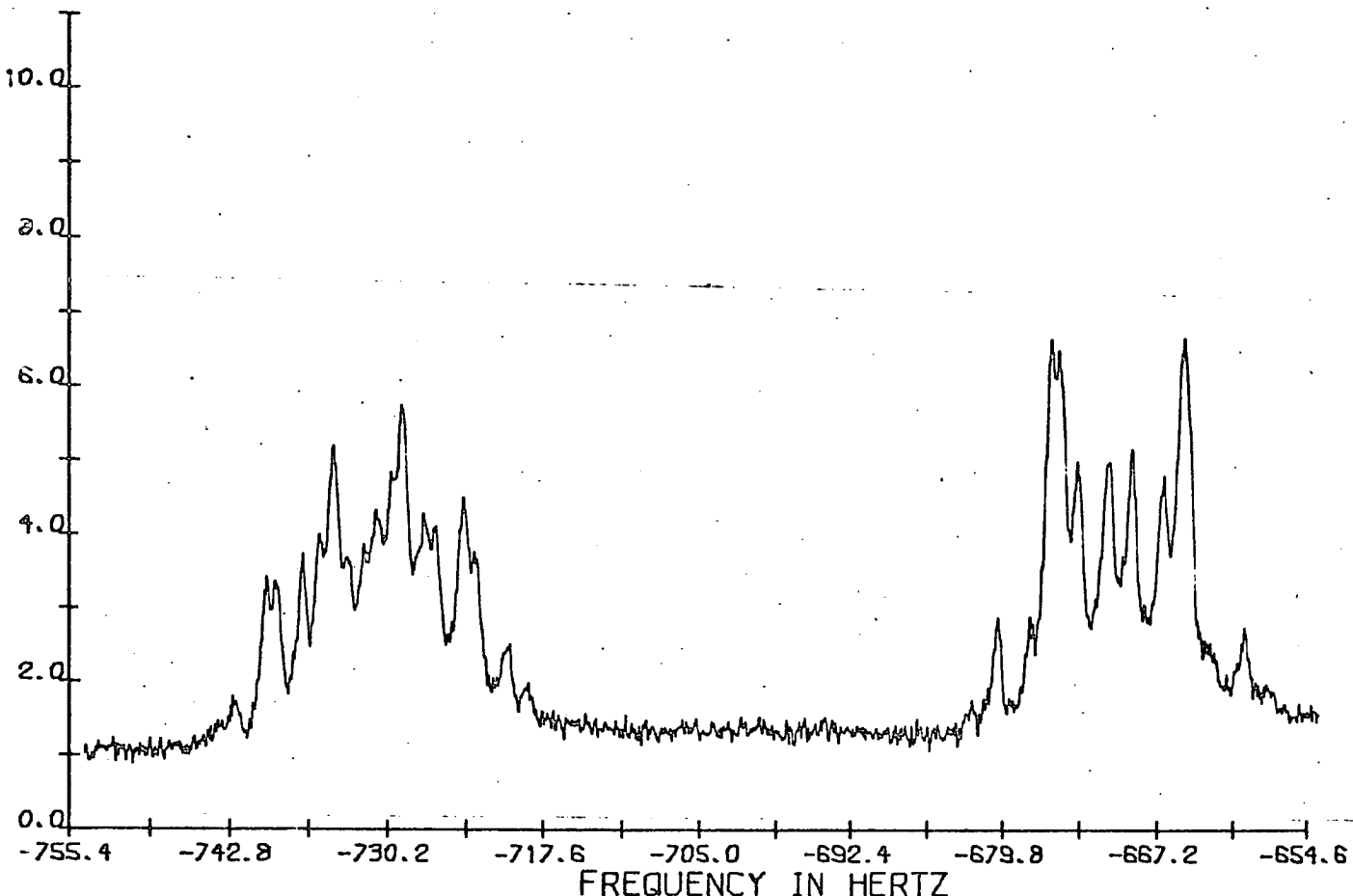


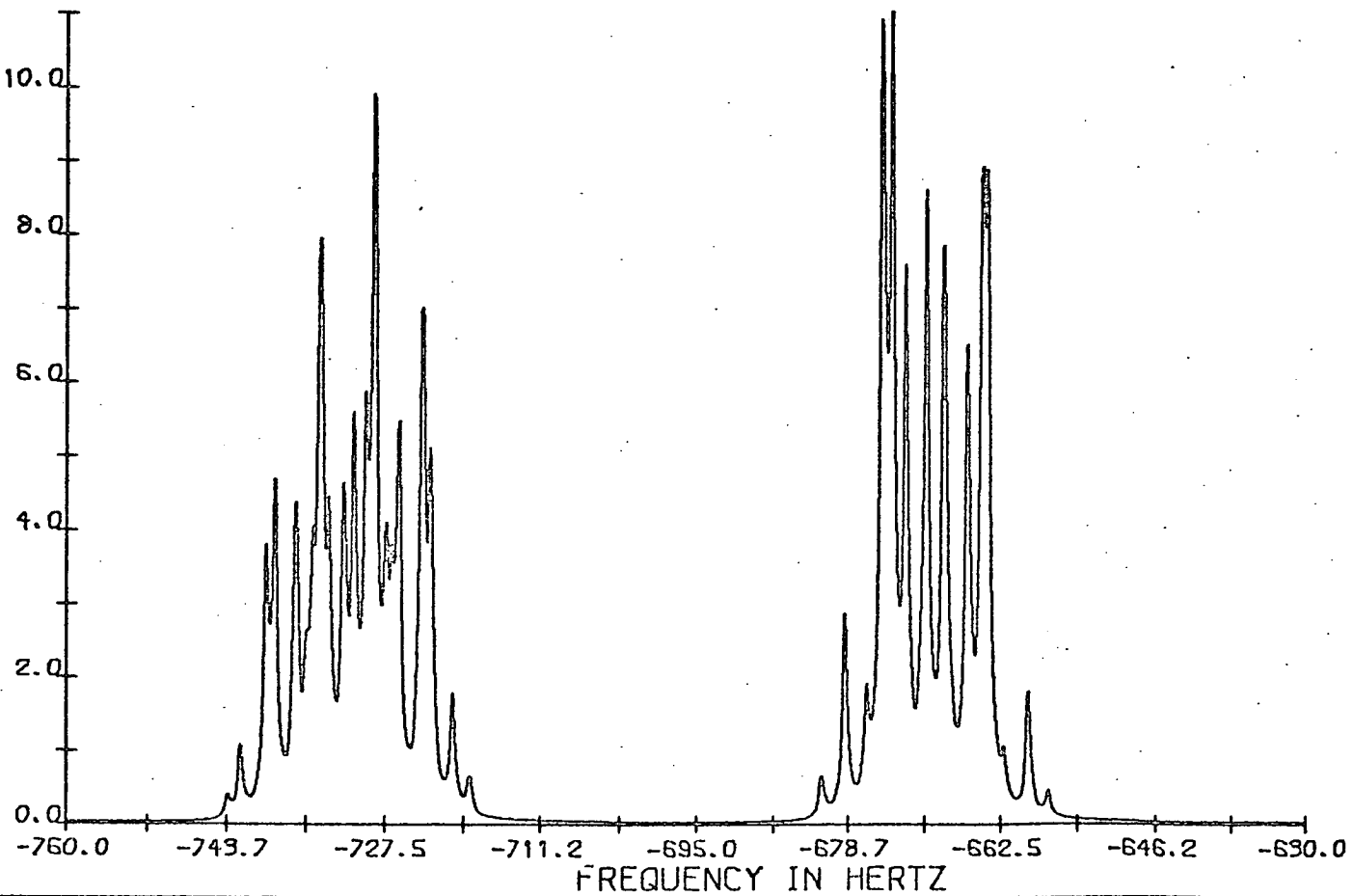


Fig. 20.

1,2-SUBSTITUTED QUINOXALINE DIOXIDE EXPTL. NMR



2-N(CH<sub>3</sub>)<sub>2</sub>, 3-(CH<sub>3</sub>)<sub>2</sub>, 2(H)-QUINOXALINE DIOXIDE



Three other compounds of interest were analysed, 2, 3, 3'-trimethyl-3H indole, (8), benzofuroxan, (9), and 2, 2'-dimethyl - 3-dimethyl amino — 2, 3-dihydroquinoxaline di-N-oxide, (10), Fig. 8. The first two gave very similar spectra of an ABCD type Figs. 17 and 18, with the H(7) proton in each case being considerably further downfield than the other three, Table 7. The trimethyl indole proved very difficult to analyse using the 100 MHz spectrum, and the compound was rerun at 220 MHz to produce a more first order spectrum. This proved more successful but even using the 220 MHz spectrum the iterative calculation produced unacceptable standard deviations in J values, Table 10. To improve these results, the chemical shifts calculated at 220 MHz were converted to shifts at 100 MHz, and used as input parameters along with the 220 iterated coupling constants. Using this procedure the standard deviations were reduced to an acceptable level, Table 7. The final results, Table 7, show the vicinal couplings to be almost constant (7.48, 7.48, 7.38), and to be very close to that in benzene, 7.54 Hz,<sup>10</sup> giving the molecular structure as a substituted benzene.

The NMR spectrum of benzofuroxan is well documented<sup>11, 12</sup>, and is known to give an ABCD spectrum at low temperatures, and an AA'BB' spectrum at higher temperatures indicating rapid tautomerism, Fig. 19. The spectrum obtained here was run at -60°C to give an ABCD type, Fig. 18. The J values of reference 11 were used as input parameters, and the final run of the iterative procedure yielded values of similar magnitude to those of reference 11, although these results are given as timed average results which makes  $J_{4,5}$  equal to  $J_{6,7}$ .

Inspection of the substituted quinoxaline di-N-oxide spectrum shows a distorted AA'BB' spectrum with the B protons showing a distinct twelve line spectrum, Fig. 20, and the A protons a more complicated multiplet. This proved to be a very difficult spectrum to analyse due to the proximity of the chemical shifts of each pair of protons, Table 7, and many non iterative runs were attempted before a reasonable assignment of transition frequencies could be made. Even in the final analysis the standard deviations were higher than normal but were considered acceptable.

The experimental spectrum of quinoxaline-di-N-oxide<sup>13</sup> has been recorded at 60 MHz, and the chemical shift of the carboxylic ring protons, H(5) and H(8) was given as 1.76  $\tau$ . The B protons of the substituted quinoxaline di-N-oxide appear furthest downfield at 2.72  $\tau$ , but are still considerably upfield of the parent compound indicating a less aromatic structure possibly due to a reduced ring current, or alternatively due to the N-O group back donating electrons into the carbocyclic ring. An ab initio calculation on the molecule would obviously throw some light on the subject, but unfortunately time did not permit this. The only coupling constants given by reference 13 is  $J_{4,5}$  at 6.8 Hz, whereas  $J_{4,5}$  in 10 is 9.26 again pointing to a much more localised structure in this case.

No published spectra exist for the quinoid molecules of the bicyclic heterocycles I, but the related molecule 1,3-diphenyl-N-methyl isoindole has been analysed. The vicinal coupling constants  $J_{4,5}$  and  $J_{5,6}$  are given for this molecule as 8.63 and 6.39 Hz, which is very close to the results obtained here. Several workers<sup>2,3</sup> have shown that a linear correlation can be drawn between changes in vicinal coupling constants with bond length, and the ratio  $J_{5,6}:J_{4,5}$  can be used to predict the degree of aromaticity of a compound with reference to benzene where C-C is constant at 1.397 $\text{\AA}$  and  $J_{4,5} = J_{5,6} = 7.54 \text{ Hz}$ <sup>10</sup> gives  $J_{5,6}:J_{4,5} = 1$ . The nearer the ratio is to 1, the higher will be the degree of aromaticity. For the quinoid molecules, Fig. 6a, X=O, S, and N-Me the ratios are 0.70, 0.72, and 0.74 respectively. The magnitudes of the ratio are very similar for this trio of molecules but if the trend is significant, then the order of aromaticity will be given as N-methyl isoindole > benzo(c)thiophen > benzo(c)furan which is consistent with the order obtained using resonance energies, and when compared with the ratios for naphthalene<sup>2,3b,4</sup>, 0.82, and cyclohexa-1,3-diene,<sup>2</sup> 0.52, the results are further enforced. However, when this analysis is applied to the Kekulé molecules of the bicyclic heterocycles I, Fig. 6b, certain difficulties arise due to the lower symmetry of these molecules.  $J_{4,5}$  no longer equals  $J_{6,7}$ , and so there is now more than one set of vicinal coupling constants. If  $J_{4,5}$  is used

in the ratio, as it is further removed from the heteroatom, then the ratio  $J_{5,6} : J_{4,5}$  becomes 0.93, 0.89 and 0.90 for X=O,  $S^6$  and  $NH^5$  giving an aromaticity order of O>N>S. These are all larger than the naphthalene ratio, 0.82, and are inconsistent with the resonance energy results which give the reverse order. It would appear that unless the symmetry of the molecule is high enough to provide a unique set of vicinal coupling constants, this method of determining aromaticity is not applicable. If the magnitude of the vicinal coupling constants of the Kekulé molecules are considered, Table 6, then it is obvious that if a relationship exists between  $J_{vic}$  and C-C length, then the C-C bond lengths in the benzenoid rings of these molecules show less variation than the corresponding bonds in the quinoid molecules, and are closer to benzene which is consistent with resonance energy results which give the Kekulé group as more aromatic than the quinoid group, and is therefore more benzene like.

For benzimidazole the  $J_{5,6} : J_{4,5}$  ratio is 0.84 whereas the ratio for the 2-spirocyclohexyl-2H-benzimidazole is 0.65, indicating that the latter has a markedly more alternating bond structure. For the substituted quinoxaline di-N-oxide,  $J_{5,6} : J_{4,5}$  is given by 0.69 which is very similar to that obtained for the spirocyclohexyl benzimidazole (7) suggesting a localised bond structure similar to that of (7). A value of 0.73 is obtained for  $J_{5,6} : J_{4,5}$  in benzofuroxan, and is similar to that obtained for the quinoid molecules, 1-3, Fig. 8.

Other molecules. The coupling constants obtained here for molecules 1-10 can be compared with published data. If the data is divided into two sets, quinoid molecules, Table 11, and Kekulé molecules, Table 12, then it is obvious on comparison of the two sets that there is strong evidence of a relationship between structure and coupling constants. In the Kekulé series the  $J_{4,5}$  and  $J_{6,7}$  coupling constants all lie in the range 7.80 to 8.4 Hz whereas the corresponding ones for the quinoid molecules lie in the range 8.6 to 9.6 indicating that the latter set correspond to a more localised structure as would be expected if the molecules approximate

Table 11

## Coupling constants (Hz) for Kekulé molecules

Molecule	$J_{3,7}$	$J_{4,5}$	$J_{4,6}$	$J_{4,7}$	$J_{5,6}$	$J_{5,7}$	$J_{6,7}$
Indole <sup>9</sup>	0.70	7.84	1.23	0.94	7.07	1.29	8.07
Benzofuran <sup>5</sup>	0.87	7.89	1.28	0.80	7.27	0.92	8.43
Benzothiophen <sup>6</sup>	0.86	8.09	1.16	0.73	7.22	1.17	8.06
Benzoxazole <sup>14</sup>	-	8.19	0.98	0.73	7.36	1.20	8.27
Benzothiazole <sup>14</sup>	-	8.25	1.13	0.65	7.24	1.10	8.22
2-Methyl-benzselenazole <sup>14</sup>	-	8.12	1.32	0.53	7.28	1.26	7.98
1, 2, 3-Benzothiadiazole <sup>15</sup>	-	8.35	1.04	0.78	7.05	1.00	7.88
Benzimidazole	-	8.37	0.95	0.78	7.02	0.95	8.37
Benzotriazole <sup>16</sup>	-	8.34	1.17	0.88	7.08	1.17	8.34
Naphthalene <sup>4</sup>	$(J_{1,5})0.85$	$(J_{1,2})8.28$	$(J_{1,3})1.24$	$(J_{2,3})6.85$			

Table 12

## Coupling constants (Hz) for quinoid molecules

Molecule	$J_{4,5}$	$J_{4,6}$	$J_{4,7}$	$J_{5,6}$	$J_{5,7}$	$J_{6,7}$
N-Methyl isoindole	8.69	0.90	0.79	6.46	0.90	8.69
Benzo(c)furan	8.22	0.80	0.84	6.17	0.80	8.82
Benzo(c)thiophen	8.86	1.03	0.79	6.36	1.03	8.86
2-Methyl benzotriazole <sup>5</sup>	8.59	0.97	0.98	6.77	0.97	8.59
2, 1, 3-Benzoxadiazole <sup>17</sup>	9.21	0.99	1.23	6.61	0.99	9.21
2, 1, 3-Benzthiadiazole <sup>17</sup>	8.83	1.12	0.87	6.59	1.12	8.83
2, 1, 3-Benzselendiazole <sup>17</sup>	9.09	1.20	0.94	6.51	1.20	9.09
Benzofuroxan	9.16	1.11	1.04	6.36	0.86	9.01
2-Spirocyclohexy-2H- benzimidazole	9.40	0.99	1.28	6.07	0.99	9.40
2, 2'-Dimethyl-3-dimethyl amino-2, 3-dihydroquinoxaline di-N-oxide	9.26	0.99	0.87	6.46	1.20	9.24
Cyclohexa-1, 3-diene	9.64	1.02	1.12	5.04	1.02	9.64

Fig. 21  $^{13}\text{C}$ -NMR. N-CH<sub>3</sub> ISOINDOLE

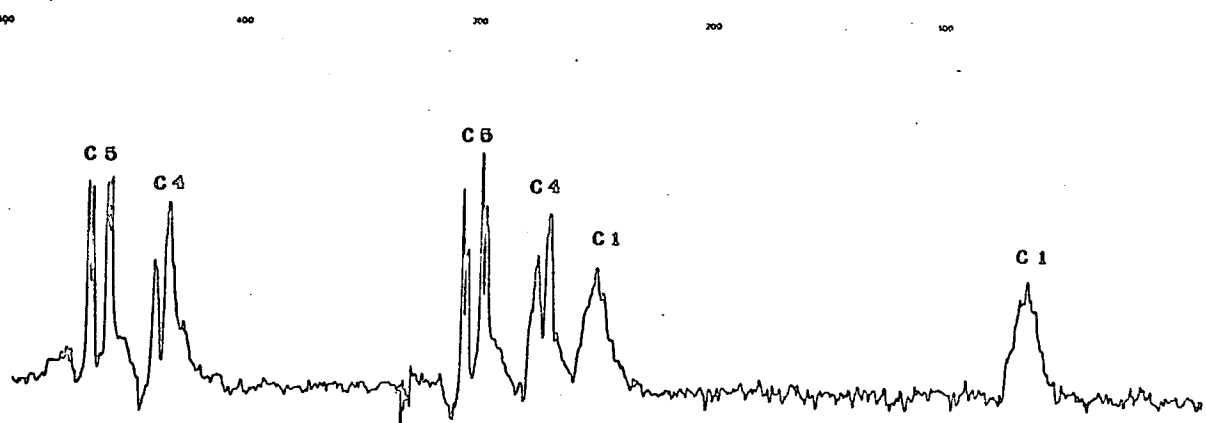
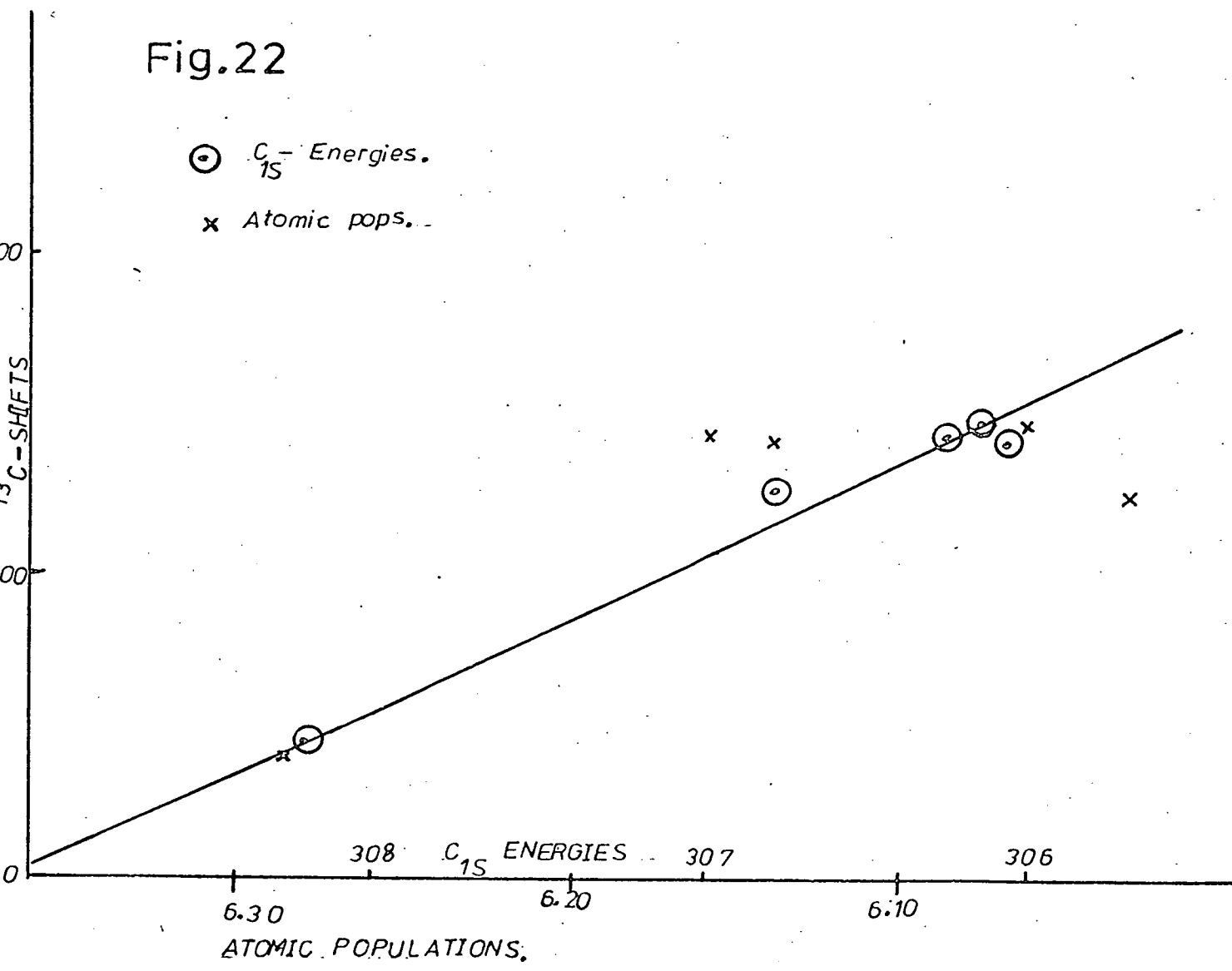


Fig. 22

⊙  $C_{1S}$  Energies.  
 x Atomic pops.



to the diagrammatic structures used to represent them. The  $J_{5,6}$  values also show that in the quinoid set, the C(5)-C(6) bond is much closer to a C-C single bond than it is in the Kekulé series, but in all cases the  $J_{5,6}$  values,  $\approx 6.5$  in the quinoid series when compared with  $J_{2,3}$  of cyclohexa-1,3-diene, 5.04 Hz, show some degree of multiple bond character. For the Kekulé series,  $J_{5,6}$  is remarkably constant, presumably due to its remoteness from the heterocyclic ring, and its magnitude  $\approx 7.2$  Hz, suggests that the bond is closer in nature to a double bond than to the single bond indicated by the classical structure.

$^{13}\text{C}$ -NMR. The only molecule in this series to be studied by  $^{13}\text{C}$ -NMR was N-methyl isoindole. The other quinoid molecules of this type, (2) and (3) Fig. 8 could not be studied in this way due to their instability in solution over extended periods of time. The  $^{13}\text{C}$ -NMR of N-methyl isoindole was obtained by the CAT method (Computer Accumulation Technique) whereby the sample is scanned many times, and the cumulative effect used to give the spectrum. The process can take several hours, and hence it was impossible to obtain spectra for (2) and (3). The chemical shifts obtained for N-methyl isoindole are shown in Table 13. The carbocyclic  $C_{\alpha}$  (4, 7) and  $C_{\beta}$  (5, 6) shifts are very similar but it was possible to confirm the assignments by off-resonance proton decoupling. The spectrum is shown in Fig. 21. It is commonly accepted that  $^{13}\text{C}$  shifts are related in the first instance to the electron density at each centre. When the chemical shifts are plotted against the atomic populations, then no obvious relationship is seen, Fig. 22, and if they are plotted against the  $C_{1s}$  orbital energies, Table 14, then the results approximate to a linear correlation, Fig. 22, as was found for the Kekulé molecules discussed in the atomic population section.

When the chemical shifts in the carbocyclic ring are compared with those of some Kekulé molecules, Table 15, then very little difference is shown although the N-methyl isoindole shifts are upfield of the others as is the case with the C(1)/C(3) shift compared with that of pyrrole or indole, Table 15.



Table 13 Chemical Shifts (ppm) of carbon in N-methyl  
isoindole

CH <sub>3</sub>	C(1)/(3)	C(4)/C(7)	C(5)/C(6)	C(3a)/(7a)
37.34	111.50	119.22	120.30	124.00

Table 14 Data plotted in the graph, Fig. 22.

N-methyl isoindole

Atomic population	C <sub>1s</sub> (-eV)	Chemical Shift
C(Me) 6.2884	308.2	37.34
C(1) 6.0327	306.8	111.50
C(4) 6.1415	306.3	119.22
C(5) 6.1609	306.1	120.30
C(3a) 6.0617	306.2	124.00

Table 15 <sup>13</sup>C-Chemical shifts for a series of molecules

	C <sub>α</sub>	C <sub>β</sub>
Naphthalene <sup>18</sup>	128.3	126.1
Indole <sup>19</sup>	121.3 (C-4)	122.3 (C-5)
Benzofuran <sup>19</sup>	121.1	122.7
Benzothiophen <sup>19</sup>	123.5	124.0
Indane <sup>20</sup>	124.0	125.9
Pyrrole	C(2) 118.7	
Indole	C(2) 125.1	

## References

1. P. Crews, R. R. Kintner, and H. C. Padgett, *J. Org. Chem.*, 1973, 38, 4391.
2. M. A. Cooper and S. L. Manatt, *J. Amer. Chem. Soc.*, 1969, 91, 6325.
3. (a) D. J. Bertelli and P. Crews, *Tetrahedron*, 1970, 26, 4717;  
(b) J. B. Pawliczek and H. Gunther, *ibid.*, 1970, 26, 1755.
4. R. W. Creceley and J. H. Goldstein, *Org. Magnetic Resonance*, 1970, 2, 613.
5. P. J. Black and M. L. Heffernan, *Australian J. Chem.*, 1965, 18, 353.
6. K. D. Bartle, D. W. Jones, and R. S. Matthews, *Tetrahedron*, 1971, 27, 5177.
7. J. Elguero, A. Fruchier, R. Jacquier and U. Scheidegger, *J. Chem. Phys.*, 1971, 68, 1113.
8. M. A. Cooper, D. D. Ellernan, C. D. Pearce, and S. L. Manatt, *J. Chem. Phys.*, 1970, 53, 2343.
9. P. J. Black and M. L. Heffernan, *Australian J. Chem.*, 1962, 15, 862.
10. J. M. Read, R. E. Mayo, and J. H. Goldstein, *J. Mol. Spectrosc.*, 1967, 22, 419.
11. R. K. Harris, A. R. Katritsky, and S. Oksne, and in part A. S. Bailey and W. G. Patterson, *J. Chem. Soc.*, 1963, 197.
12. P. Diehl and C. L. Khetrapal, *Colloque Ampere XV*, North Holland, Amsterdam, 1969.
13. K. Tori, M. Ogata, and H. Kano, *Chem. Pharm. Bull. (Tokyo)*, 1963, 11, 681.
14. F. L. Tobiason and J. M. Goldstein, *Spectrochim. Acta*, 1967, 23A, 1385.
15. W. H. Poesche, *J. Chem. Soc. B*, 1966, 568.
16. M. H. Palmer, R. H. Findlay, S. M. F. Kennedy, and P. S. McIntyre, *JCS Perkin II*, 1975, 1695.

17. N. M. D. Brown and P. Bladon, *Spectrochim. Acta*, 1968, 24A, 1869.
18. T. D. Alger, D. M. Grant, and E. G. Paul, *J. Amer. Chem. Soc.*, 1966, 88, 5397.
19. N. Platzter, J. J. Basselier and P. Demerseman, *Bull. Soc. Chim. Fr.*, 1974, 5-6, 905.
20. G. Jikeli, W. Herrig and H. Gunther, *J. Amer. Chem. Soc.*, 1974, 96, 323.

Fig.1

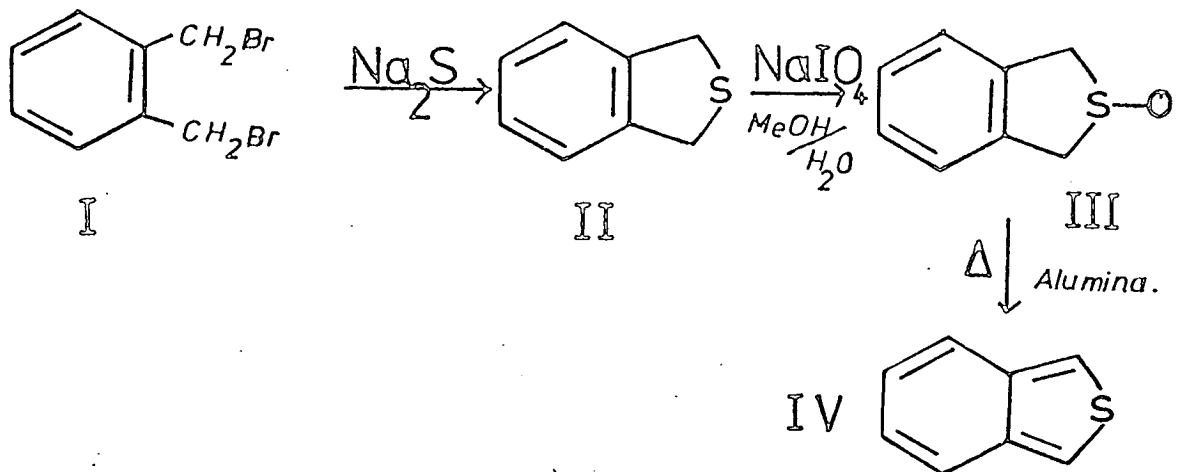
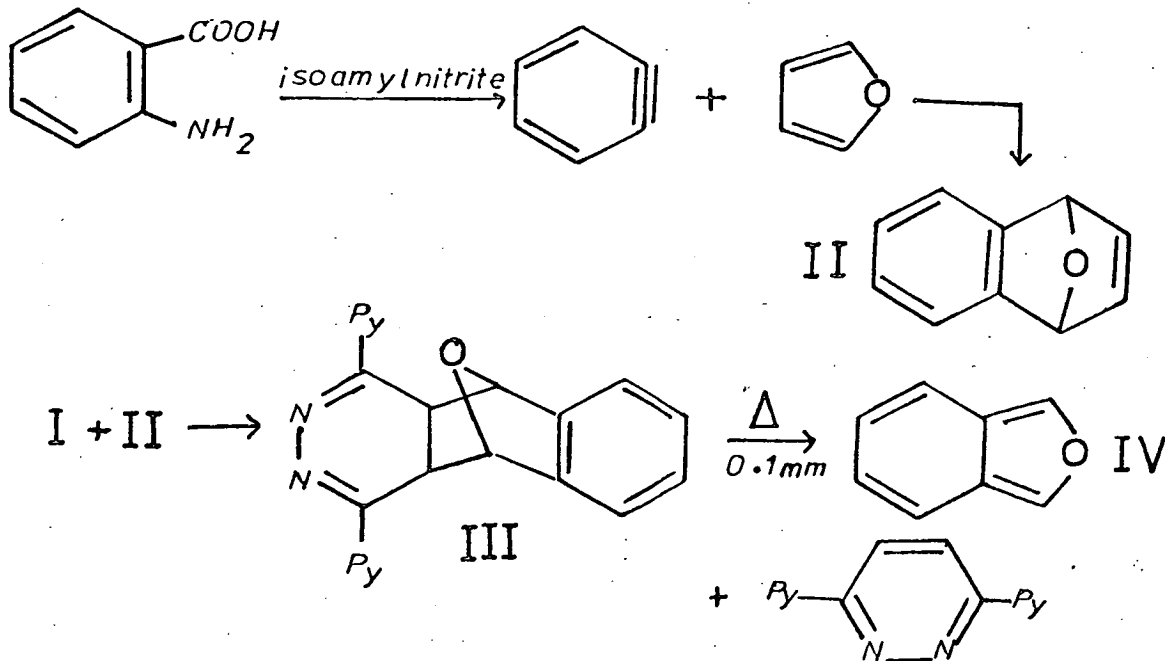
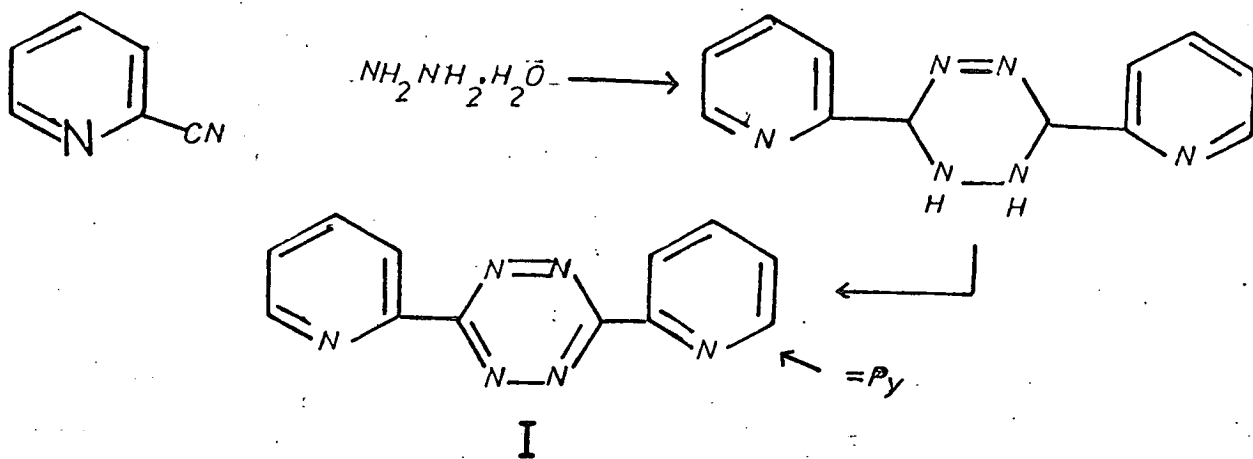


Fig.2

A



## Experimental Methods

The compounds of the Kekulé molecules of the bicyclic heterocycles series I, indene, benzofuran, benzothiophen and indole were all commercially available and were purified by either fractional distillation or vacuum sublimation. Naphthalene was also commercially available and was purified by vacuum sublimation.

In the quinoid series, the parent molecules were not readily accessible, and difficulty was encountered in several cases in obtaining pure samples.

The compound easiest to obtain was benzo(c)thiophen. Its preparation was first recorded in 1963,<sup>1</sup> and the method involved the dehydrogenation of 1,3-dihydrobenzo(c)thiophen. A similar preparation using 1,3-dihydro benzo(c) thiophen-2-sulphoxide was published in 1971,<sup>2</sup> and was used in this work. The reaction sequence is shown in Fig. 1.

### Sequence Fig. 1

- I. o-Dibromo xylene was prepared by the standard preparation. The product was a strong lachrymator and no recrystallisation was attempted:- E. F. Atkinson and J. F. Thorpe, J. Chem. Soc., 1907, 91, 1696.
- II. 1,3-Dihydro benzo(c)thiophen:- M. P. Cava and A. A. Deana, J. Amer. Chem. Soc., 1959, 81, 4266.
- III. 1,3-Dihydro benzo(c)thiophen-2-sulphoxide:- reference 2.
- IV. Benzo(c)thiophen:- reference 2.

Benzo(c)thiophen was observed to decompose rapidly in  $\text{CHCl}_3$  at room temperature to produce a yellow solution which eventually turned black, but only slow decomposition of the solid occurred at room temperature. To enable an NMR spectrum to be obtained, a sample was dissolved in  $\text{CDCl}_3$  which had been treated with sodium dithionite in  $\text{D}_2\text{O}$  to remove oxygen.

Fig.2

B

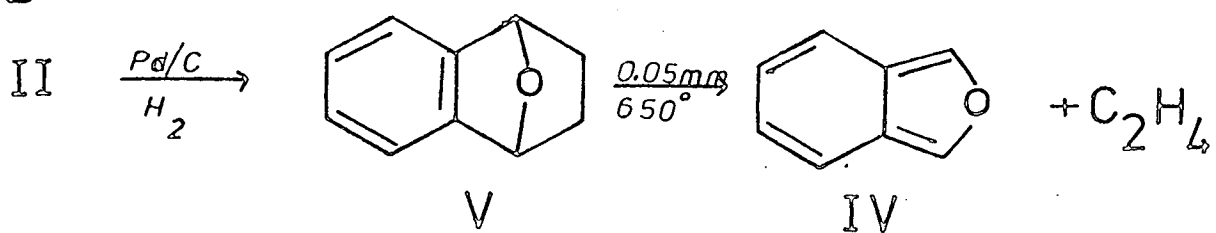


Fig.3

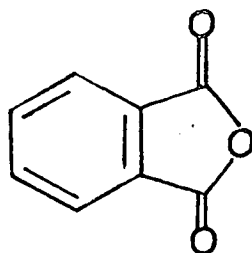
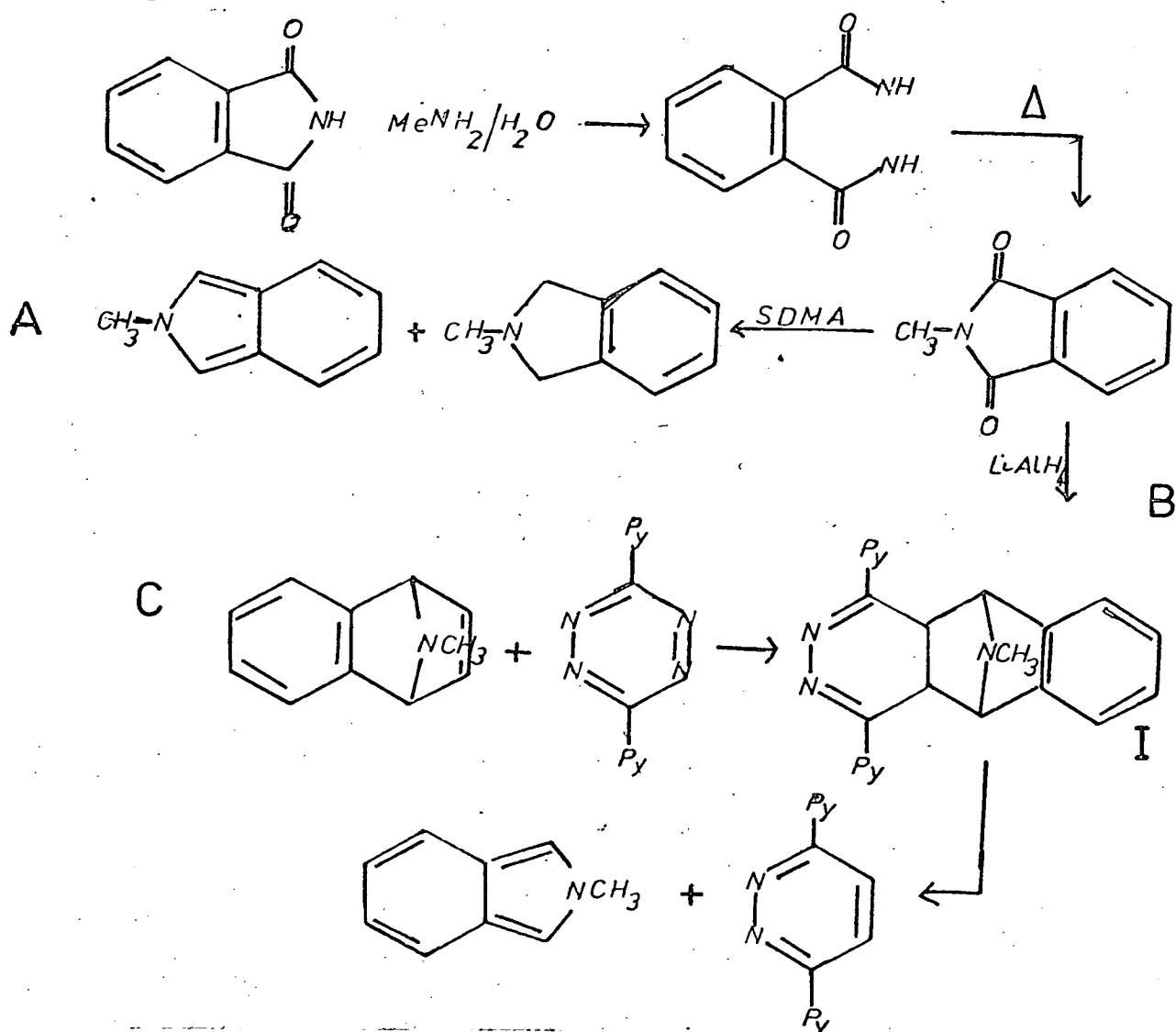


Fig.4



Benzo(c)furan was not isolated until the early 1970's, possibly due to its very reactive nature. Before this time only 1,3-disubstituted compounds,<sup>3</sup> the substituent often being a phenyl group, and more heavily substituted compounds<sup>4</sup> were known. Two methods were available for the preparation of the parent compound. Both methods were used and are shown in Fig. 2. Route A was the longer of the two, but was the one used initially in this work.

Sequence Fig. 2

A. I 3,6-Dipyridyl-1,2,4,5-tetrazine:- J. F. Geldard and F. Lions, J. Org. Chem. , 1965, 30, 318.

The 1,2-dihydro compound was prepared from 2-cyano pyridine and hydrazine hydrate, and was recrystallised from pyridine. It was then reacted with glacial acetic acid and sodium nitrite solution at 0°C. The product, I, was bright red in colour.

A. II 1,4-Dihydro-1,4-endoxy naphthalene:- L. F. Fieser and M. J. Haddadin, Canad. J. Chem. , 1965, 43, 1599.

The crude pale orange product was recrystallised from 60-80 petroleum ether.

A. III and benzo(c)furan IV:- R. N. Warener, J. Amer. Chem. Soc. , 93, 2346.

Intermediate III was formed by reacting I and II in nitromethane. A spontaneous reaction occurred, evolving nitrogen and precipitating a yellow solid, III.

The precipitate, III, was decomposed by heat, and sublimed at 0.1 mm through a glass wool plug onto a cold finger cooled by solid carbon dioxide. The 3,6-dipyridyl pyridazine remained in the pot.

B. V 1,2,3,4-Tetrahydro-1,2-endoxy naphthalene:- G. Wittig and L. Pohmer, Chem. Ber. , 1956, 89, 1334.

B. IV benzo(c)furan:- U. E. Wiersum and W. J. Mijs, JCSChem. Comm. , 1972, 347.

The product was prepared by flash vacuum pyrolysis. The compound V was passed at 0.05 mm through a silica tube preheated to 650°C, and the products collected in a trap cooled by liquid nitrogen. This solidified the benzo(c)furan and the ethylene. The products were kept under pressure, and the liquid nitrogen changed for dry ice/acetone as coolant. This allowed the ethylene to vapourise, and to be pumped off, leaving the benzo(c)furan as a white solid. This procedure is a retro Diels Alder reaction.

The parent compound, IV,<sup>5</sup> and some derivatives<sup>6</sup> are known to undergo rapid autoxidation at room temperature. For this reason the product was kept cool by solid carbon dioxide and under nitrogen. At higher temperatures the mass spectrograph of the compound showed a major peak at  $m/e = 148$ , possibly corresponding to the oxidised compound shown in Fig.3.

Preparations for many polysubstituted isoindoles exist,<sup>3a, 7-10</sup> and also some for N-substituted compounds,<sup>11-16</sup> but few references can be found for the parent molecule, isoindole itself.<sup>17-20</sup> The parent compound was first referred to in 1964,<sup>17</sup> and was recorded to be sensitive to oxidation and only stable for a few days under nitrogen. It was decided that as the N-methyl compound is more stable and better documented, it would be advisable to attempt this preparation first. Four methods were attempted but only one proved successful as shown in Fig. 4. Both methods C and D produced N-methyl isoindole but in C the yield was low because of the spontaneous decomposition of the intermediate I Fig. 4 C. This produced N-methyl isoindole in nitromethane solution from which it was extracted with 60-80 petroleum ether. The petroleum ether was evaporated, and the product vacuum sublimed to yield a small quantity of the pure product. Method D proved to be more satisfactory and gave a good yield of N-methyl isoindole.

#### Sequence Fig. 4 D

D.I N-(o-Chlorobenzyl), N-methyl amino acetonitrile:- B. Jacques and R. G. Wallace, JCS Chem. Comm., 1972, 397.



Fig 4

D

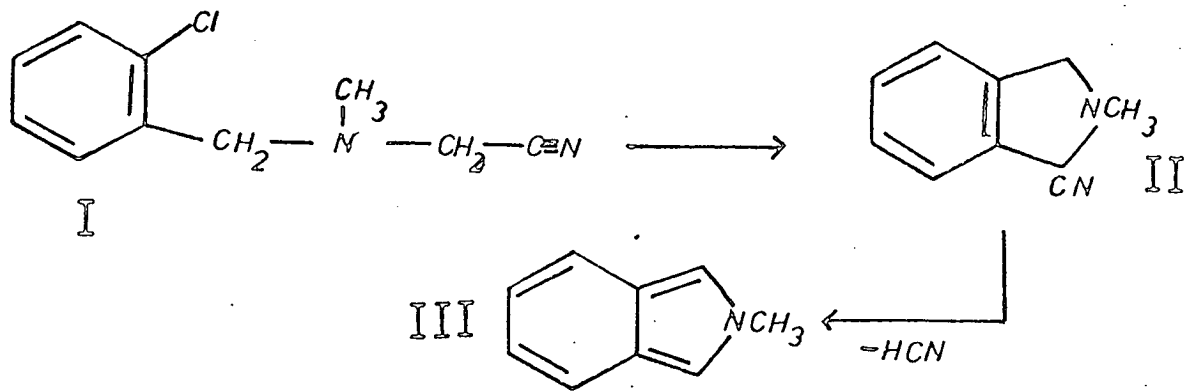


Fig.5

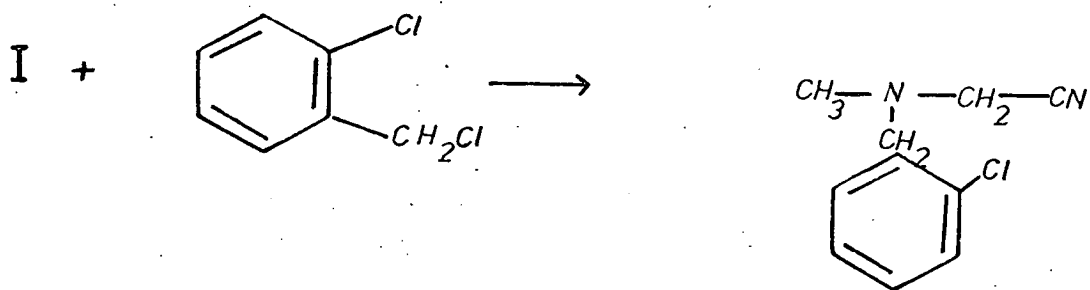
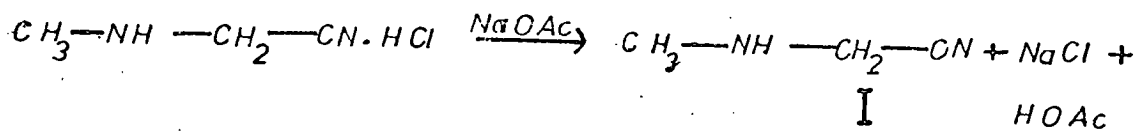
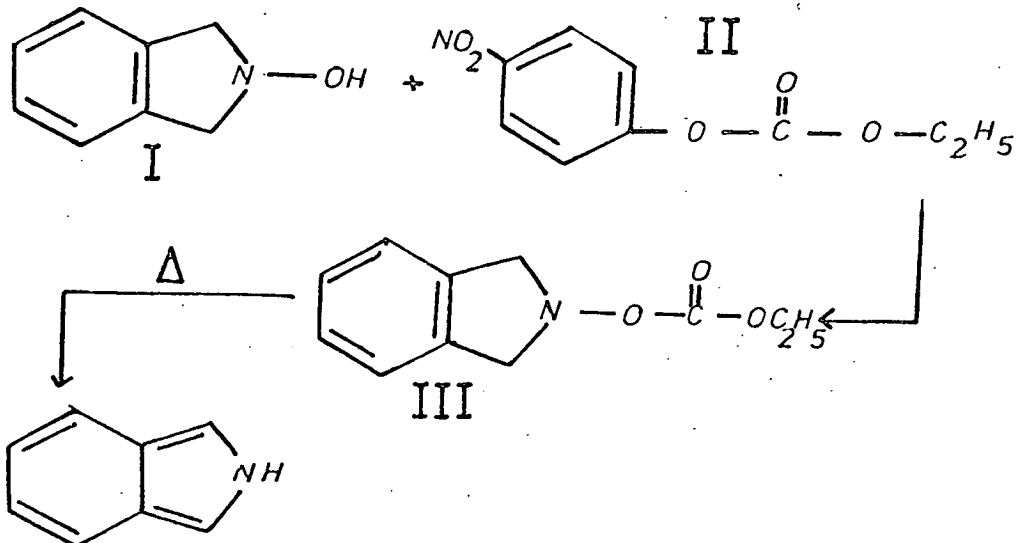


Fig.6



This compound was prepared as shown in Fig. 5 from N-methyl amino acetonitrile and o-chloro benzyl chloride. The product was a dark brown oil which even after two distillations was still contaminated. Purification was by spinning band column fractionation to yield the pure product.

D. II and III:- B. Jacques and R. G. Wallace, JCS Chem. Comm., 1972, 7, 397.

Intermediate II was not isolated, and was prepared from I reacted with potassamide in liquid ammonia. The reaction proceeded to completion within half an hour, and the reaction mixture was quenched with solid ammonium nitrate. The product was extracted with anhydrous ether. The product was kept under dry nitrogen and purified by vacuum sublimation to yield white crystals.

An attempt to prepare isoindole itself was made using the method of reference 19 shown in Fig. 6.

#### Sequence Fig. 6

I. Ethyl (p-nitrophenyl) carbonate:- J. M. A. Hoeflake, Rec. Trav. Chim., 1916, 36, 24.

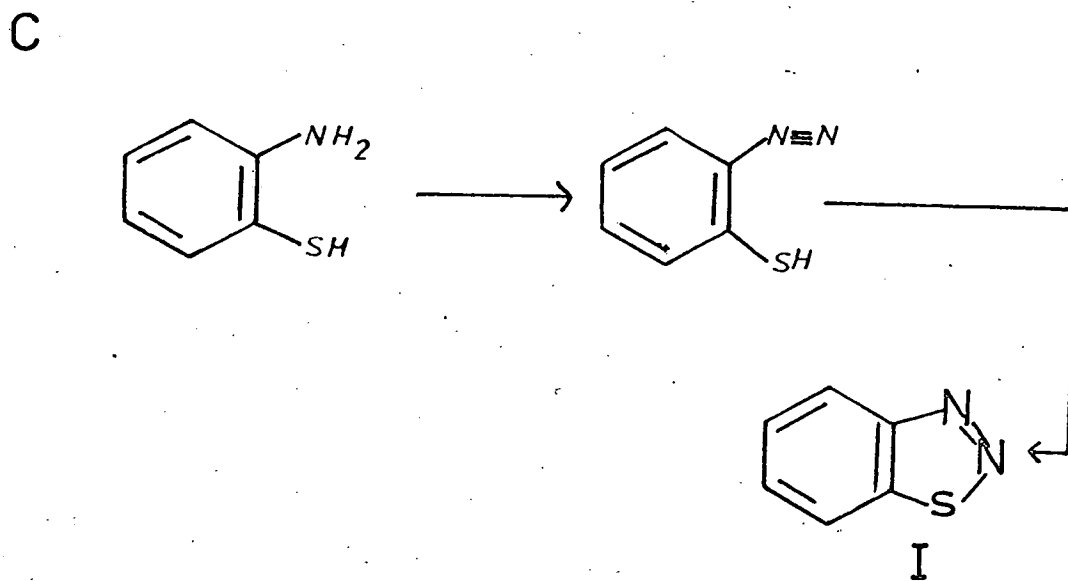
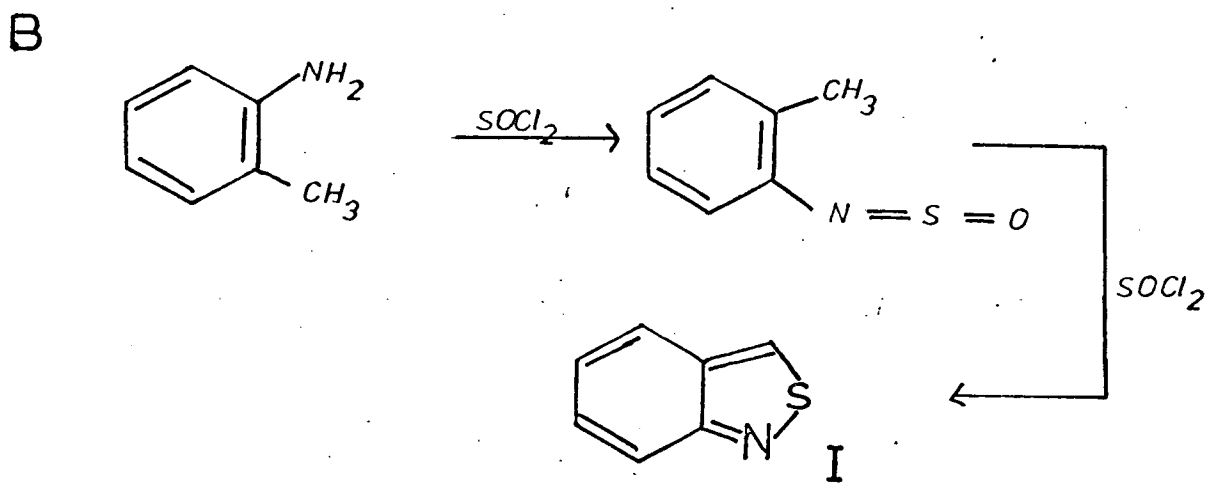
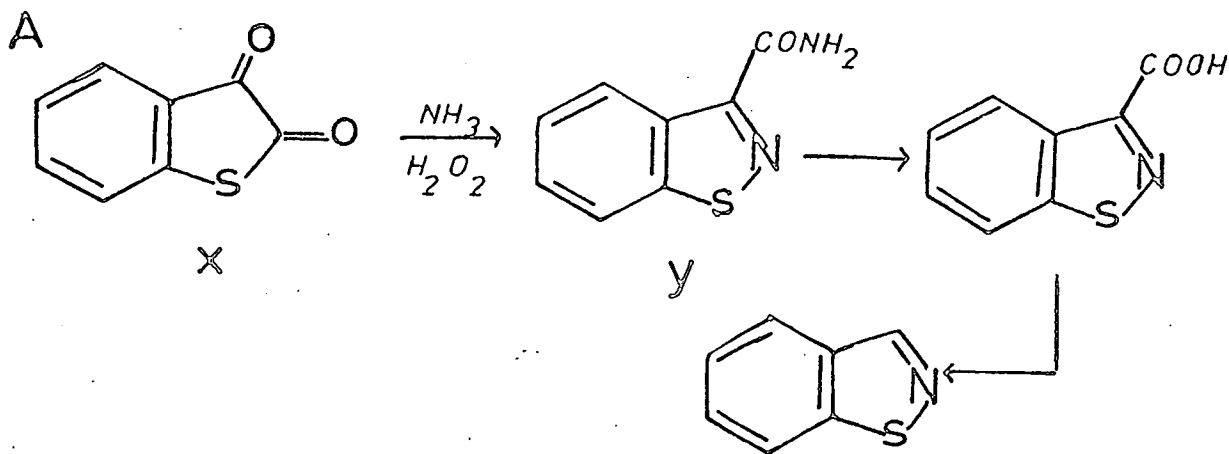
II and III:- R. Bonnett and R. F. C. Brown, JCS Chem. Comm., 1972, 7, 393.

Intermediate III was obtained fairly readily but pyrolysis even under nitrogen was accompanied by extensive decomposition, and a pure sample could not be obtained.

The compound 2H-indene has not yet been isolated in the free state but the related compounds 2, 2'-dimethyl 2H-indene iron tricarbonyl,<sup>20</sup> and 1, 3-diphenyl-2, 2'-dimethyl 2H-indene<sup>21</sup> can be prepared. However, neither of these would be suitable for providing information on the parent compound, so they were not prepared.

In the bicyclic heterocycles series II the following compounds were commercially available:- indazole, benzimidazole, 1, 2-benzisoxazole,

Fig. 7



1,3-benzoxazole, anthranil, 1,3-benzothiazole and 2,1,3-benzthiadiazole. A sample of 2,1,3-benzoxadiazole was kindly provided by Dr. P. Bladon of Strathclyde University. The other compounds of this series were prepared synthetically as shown in Fig. 7, by standard synthetic routes.

Sequence Fig. 7

A. X;- R. Stolle, Chem. Ber., 1914, 47, 1130.

A. I: 1,2-Benzisothiazole:- R. Stolle, Chem. Ber., 1925, 58, 2095.

The amide, Y, was treated with molar sodium hydroxide solution to produce the carboxylic acid which was then decarboxylated at 175°C to give I.

B. I: 2,1-Benzisothiazole:- M. Davis and A. W. White, J. Org. Chem., 1969, 34, 2985.

The crude product was purified by means of a picrate derivative.<sup>24, 25</sup>

The picrate was recrystallised from methanol, and the product, I, obtained by treating the picrate with dilute sodium hydroxide solution. Final purification was by distillation.

C. I: 1,2,3-Benzthiadiazole:- P. Jacobson and H. Janssen, Annalen, 1893, 277, 221.

The product was obtained directly from diazotisation of o-amino thiophenol.

In the diazine N-oxide series, only pyrazine mono-N-oxide was obtained pure, the pyridazine N-oxide being contaminated with pyridazine.

Pyrazine N-oxide was obtained by treatment of pyrazine with peracetic acid<sup>26</sup> at room temperature to yield a low melting solid which was purified by distillation.

Other compounds were synthesised with a view to obtaining their NMR spectra.

Fig.8

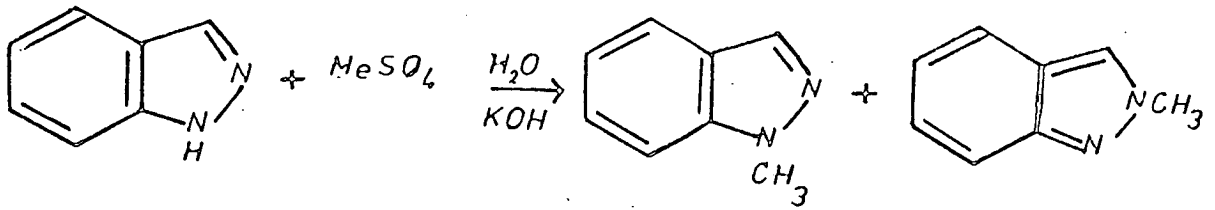


Fig.9

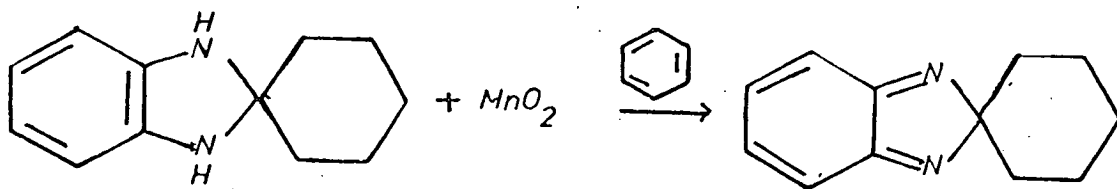


Fig.10

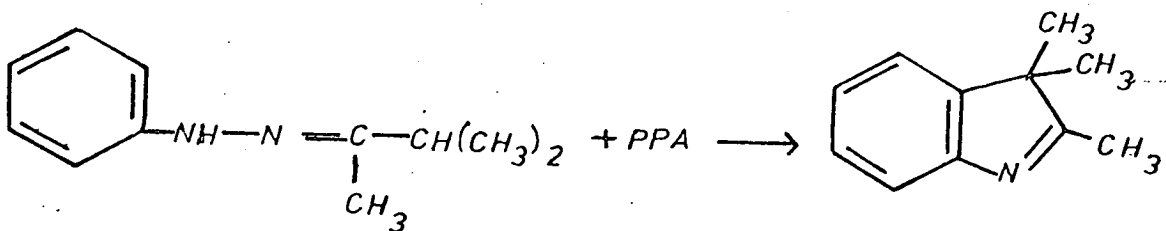
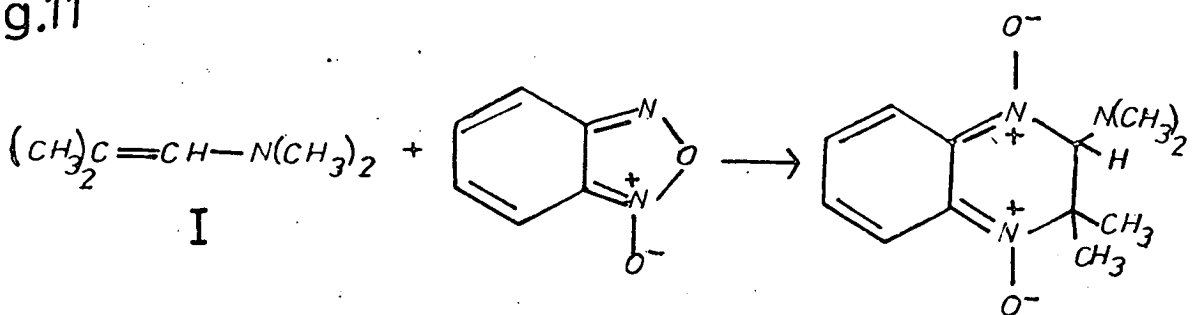


Fig.11



1- and 2-methyl indazole, Fig. 8

Methylation of indazole produced a mixture of the 1 and 2-methyl isomers which was separated by column chromatography using alumina. The 1-methyl compound was eluted first using a 20:80 mixture of ether and 60-80 petroleum ether. The eluent was increased in polarity but the 2-methyl isomer remained on the column until chloroform was used as eluent. Unreacted indazole was eluted from the column shortly after the 2-methyl indazole. The isomers were purified by recrystallisation with 60-80 petroleum ether.

2-Spirocyclohexyl-2H-benzimidazole,<sup>29</sup> Fig. 9

The 2-spirocyclohexyl isoindoline,<sup>27</sup> I, was reacted with activated manganese dioxide<sup>28</sup> in benzene to give the product<sup>29</sup> as a pale yellow solid which was recrystallised from 40-60 petroleum ether to produce off-white crystals.

2, 3, 3'-Trimethyl-3H-indole,<sup>30</sup> Fig. 10

The phenyl hydrazone was prepared from phenyl hydrazine and methyl isopropyl ketone, and the product was treated with polyphosphoric acid to give the 2, 3, 3'-trimethyl-3H-indole.

2, 2'-Dimethyl - 3-dimethylamino - 2, 3-dihydro quinoxaline di-N-oxide,<sup>31</sup> Fig. 11

The N,N'-dimethyl isobutenylamine,<sup>32</sup> was prepared from isobutyraldehyde and dimethylamine in the presence of anhydrous potassium carbonate to give a liquid. I was reacted with benzofuroxan to produce a dark red solid which was purified by column chromatography using florisil with a 50:50 mixture of benzene and chloroform as eluent. The solid obtained was recrystallised from an acetone/benzene mixture.

## References

1. V. R. Mayer, H. Kleinert, S. Richter, and K. Gewald, *J. Prakt. Chem.*, 1963, 20, 244.
2. M. P. Cava, N. M. Pollock, O. A. Mamer, and M. J. Mitchell, *J. Org. Chem.*, 1971, 25, 3932.
3. (a) J. D. White, M. E. Mann, H. D. Kirshentraum, and A. Mitra, *J. Org. Chem.*, 36, 1971, 1048.
4. (b) M. Avram, D. Constantinescu, I. G. Dinnlescu, and C. D. Nenitzescu, *Tetrahedron Letters*, 1969, 59, 5215.
4. W. Reid and K. H. Bonnighausen, *Annalen*, 1961, 639, 61.
5. R. H. Young and D. T. Feriozi, *JCS Chem. Comm.*, 1972, 841.
6. J. Olmsted and T. Akashah, *J. Amer. Chem. Soc.*, 1973, 95, 621.
7. J. C. Emmett, D. F. Veber, and W. Lwowski, *JCS Chem. Comm.*, 1965, 272.
8. R. Kreher and G. Vogt, *Angew. Chem. Int. Ed.*, 1970, 9, 955.
9. C. O. Bender and R. Bonnett, *J. Chem. Soc. (C)*, 1968, 3037.
10. M. E. Mann and J. D. White, *JCS Chem. Comm.*, 1969, 420.
11. J. K. Kochi and E. A. Singleton, *Tetrahedron*, 1968 24, 4649.
12. D. F. Weber and W. Lwowski, *J. Amer. Chem. Soc.*, 1964, 86, 4152.
13. R. Kreher and J. Seubert, *Angew. Chem. Int. Ed.*, 1964, 3, 639; *ibid.*, 1960, 5, 967.
14. H. Lund and E. T. Jensen, *Acta Chem. Scanda*, 1970, 24, 1867.
15. B. Jacques and R. G. Wallace, *JCS Chem. Comm.*, 1972, 7, 397.
16. D. L. Garmaise and A. Ryan, *J. Het. Chem.*, 1970, 7, 413.
17. J. Thesing, W. Schäfer, and D. Melchior, *Annalen*, 1964, 671, 119.
18. R. Kreher and J. Seubert, *Z. Naturforsch.*, 1965, 20b 75.
19. R. Bonnett and R. F. C. Brown, *JCS Chem. Comm.*, 1972, 7, 393.
20. W. R. Roth and J. D. Meier, *Tetrahedron Letters*, 1967, 2053.
21. K. Alder and M. Fremery, *Tetrahedron*, 1961, 14, 190.

22. G. Wittig, H. Tenhaeff, W. Schoch and G. Koening, *Annalen*, 1951, 572, 1.
23. G. Wittig, G. Closs and F. Mindermann, *Annalen*, 1955, 594, 89.
24. J. Goerdeler and J. Kondler, *Chem. Ber.*, 1959, 92, 1679.
25. M. Davis and A. W. White, *JCS Chem. Comm.*, 1968, 1547.
26. C. F. Koelsch and W. H. Gumprecht, *J. Org. Chem.*, 1958, 1603.
27. H. A. Staab and F. Vogtle, *Chem. Ber.*, 1965, 98, 2681.
28. J. Attenburrow, A. F. B. Cameron, J. H. Chapman, R. M. Evans, B. A. Hems, A. B. A. Jansen, and T. Walker, *J. Chem. Soc.*, 1952, (1), 1104.
29. O. Meth Cohn and H. Suschitsky, *Chem. and Ind.*, 1969, 443.
30. M. H. Palmer and P. S. McIntyre, *J. Chem. Soc. B*, 1969, 446.
31. J. W. McFarland *J. Org. Chem.*, 1971, 36, 1842.
32. K. C. Brannock and R. D. Burpitt, *J. Org. Chem.*, 1961, 26, 3576.



## APPENDIX 1

### Basis Sets

Standard Sets      From Roos and Siegbahns "best atom"  
calculations<sup>+</sup>

Carbon (5, 2; 3)

<u>Coefficient</u>	<u>Exponent</u>
0.004813	1412.290000
0.037267	206.885000
0.172403	45.849800
0.459261	12.388700
0.456185	3.723370
0.522342	0.524194
0.594186	0.163484
0.112194	4.182860
0.466227	0.851563
0.622569	0.199206

Nitrogen (5, 2; 3)

<u>Coefficient</u>	<u>Exponent</u>
0.004479	2038.410000
0.034581	301.689000
0.164263	66.463000
0.453898	17.808100
0.468979	5.304520
0.513598	0.764993
0.605721	0.234424
0.119664	5.954610
0.474629	1.232930
0.611142	0.286752

Oxygen (5, 2; 3)

<u>Coefficient</u>	<u>Exponent</u>
0.004324	2714.890000
0.032265	415.725000
0.156410	91.980500
0.447813	24.451500
0.481602	7.222960
0.504708	1.063140
0.616743	0.322679
0.129373	7.755790
0.481269	1.623360
0.604484	0.365030

Hydrogen (3)

<u>Coefficient</u>	<u>Exponent</u>
0.070480	4.500370
0.407890	0.681277
0.647670	0.151374

<sup>+</sup> B. Roos and P. Siegbahn, "Proceedings of the Seminar on computational Problems in Quantum Chemistry", Strassburg, 1969.

Sulphur (6, 2, 2; 4, 2, 2; 1)

<u>Coefficient</u>	<u>Exponent</u>	<u>Coefficient</u>	<u>Exponent</u>
0.001546	25506.3	0.028414	129.088
0.011973	3812.82	0.175840	29.6305
0.059943	860.556	0.467398	8.84715
0.207528	242.940	0.485545	2.85576
0.442977	79.0448	0.450422	0.626108
0.392193	27.5705	0.678482	0.175233
0.406937	6.49476	1.0	0.541000
0.700363	2.41078		
0.524282	0.469815		
0.655507	0.173396		

Scaled Sets

Hydrogen - scaled ethylene

<u>Coefficient</u>	<u>Exponent</u>
0.070480	6.993570
0.407890	1.058700
0.647670	0.235235

Carbon - scaled ethylene

<u>Coefficient</u>	<u>Exponent</u>
0.004813	1412.290000
0.037267	206.885000
0.172403	45.849800
0.459261	12.388700
0.456185	3.723370
0.522342	0.557981
0.594186	0.174021
0.112194	4.749190
0.466227	0.966859
0.622569	0.226177

Oxygen - scaled vinyl alcohol

<u>Coefficient</u>	<u>Exponent</u>
0.004324	2714.890000
0.032265	415.725000
0.156410	91.980500
0.447813	24.451500
0.481602	7.222960
0.504708	0.967457
0.616734	0.293638
0.129373	7.577410
0.481269	1.586020
0.604484	0.356634

Hydrogen - scaled vinylamine

<u>Coefficient</u>	<u>Exponent</u>
0.070480	8.387790
0.407890	1.269760
0.647670	0.282131

Nitrogen - scaled vinylamine

<u>Coefficient</u>	<u>Exponent</u>
0.004479	2038.410000
0.034581	301.689000
0.164263	66.463000
0.453898	17.808100
0.468979	5.304520
0.513598	0.719093
0.605721	0.220359
0.119664	6.252341
0.474629	1.294580
0.611142	0.310090

Oxygen - scaled formaldoxime

<u>Coefficient</u>	<u>Exponent</u>
0.004391	2714.890000
0.032764	415.725000
0.158829	91.980500
0.454738	24.451500
0.489050	7.222960
0.477135	0.967457
0.583049	0.293638
0.129373	7.709260
0.481269	1.613620
0.604484	0.362840

Scaled Sets (cont'd)Nitrogen - scaled formaldoxime

<u>Coefficient</u>	<u>Exponent</u>
0.004539	2038.410000
0.035048	301.689000
0.166481	66.463000
0.460026	17.808100
0.475311	5.304520
0.486157	0.719093
0.573358	0.220359
0.119664	6.252340
0.474629	1.294580
0.611142	0.301090

Carbon - scaled methane

<u>Coefficient</u>	<u>Exponent</u>
0.004813	1412.290000
0.037267	206.885000
0.172403	45.849800
0.459261	12.388700
0.456185	3.723370
0.522342	0.557981
0.594186	0.174021
0.112194	4.683310
0.466227	0.953399
0.622569	0.223029

Hydrogen - scaled methane

<u>Coefficient</u>	<u>Exponent</u>
0.070480	7.378120
0.407890	1.116920
0.647670	0.248170

Sulphur - scaled thioformaldehyde

<u>Coefficient</u>	<u>Exponent</u>
0.001575	25506.300000
0.012217	3812.820000
0.061166	860.556000
0.211761	242.940000
0.452013	79.044800
0.400193	27.570500
0.382204	6.494760
0.657797	2.410780
0.463545	0.412644
0.579568	0.152296
0.029069	129.088000

Scaled Sets (cont'd)Sulphur - scaled thioformaldehyde

<u>Coefficient</u>	<u>Exponent</u>
0.179893	29.630500
0.478170	8.847150
0.496736	2.855760
0.441037	0.650366
0.664344	0.182022

## APPENDIX 2

### Molecular Orbital Energy Levels

Molecular Orbital Energy Levels ( $-\epsilon_i$ /eV)

Benzene

$a_{1g}$	$b_{2u}$	$a_{2u}$	$e_{1u}$	$e_{1g}$	$e_{2g}$	$b_{1u}$
306.8	17.08	14.47	306.8	9.82	306.8	306.8
31.42	-	-	27.62	-	22.36	17.35
19.60	-	-	16.27	-	13.51	-

Naphthalene

$a_g$	Character	$b_{3g}$	Character
307.3	$1s_\gamma$	306.9	$1s_\beta$
306.8	$1s_\beta$	306.8	$1s_\alpha$
306.8	$1s_\alpha$	29.64	$2s_{\alpha+\beta}$
32.33	$2s_{\alpha+\beta+\gamma}$	19.18	$(2p+1s_H)_\beta$ (L)
27.56	$2s_{\beta-\gamma}$	16.30	$(2p+1s_H)_\alpha$ (T)
22.31	$(2p+1s_H)_\alpha$ (T)	13.44	$(2p+1s_H)_\beta$ (L)
18.78	$(2p+1s_H)_\beta$ (L)		
17.14	2p (L + T)		
13.17	$(2p+1s_H)_\alpha + 2p_\beta$ (T)		
$b_{3u}$	Character	$b_{1g}$	Character
15.48	$2p_{\alpha+\beta+\gamma}$	11.29	$2p_{\alpha+\beta+\gamma}$
9.47	$2p_{\beta-\gamma}$		



Naphthalene (Cont'd)

$b_{1u}$	Character	$b_{2u}$	Character
306.8	$1s_{\beta}$	307.3	$1s_{\gamma}$
306.8	$1s_{\alpha}$	306.8	$1s_{\beta}$
30.48	$2s_{\alpha+\beta}$	306.8	$1s_{\alpha}$
23.27	$2s_{\alpha-\beta} + 2p + (2p + 1s_H)_{\beta}$	28.73	$2s_{\alpha+\beta+\gamma}$
19.54	$(2p + 1s_H)_{\alpha+\beta}$	22.88	$2p(L + T)$
16.24	$2p_{\alpha-\beta+\gamma} (L)$	17.12	$(2p + 1s_H)_{\alpha} (T)$
14.48	$2p_{\alpha-\beta} (T)$	15.54	$(2p + 1s_H)_{\beta} + 2p_{\alpha} (L)$
$b_{2g}$	Character	$a_u$	Character
13.09	$2p_{\alpha+\beta}$	8.57	$2p_{\alpha+\beta}$

Anthracene

$a_g$	$b_{1u}$	$b_{2u}$	$b_{3g}$
307.3	307.3	307.3	307.3
306.8	306.8	306.8	306.8
306.8	306.8	306.8	306.8
306.7	31.31	306.7	27.78
32.44	27.42	28.92	22.86
29.61	21.98	26.11	17.36
23.85	19.54	20.04	16.25
22.26	17.20	16.99	13.16
18.92	15.80	15.89	-
17.57	13.60	14.51	-
15.04	-	-	-
12.91	-	-	-
$b_{3u}$	$b_{1g}$	$b_{2g}$	$a_u$
15.70	11.70	14.22	10.00
11.96	7.42	9.20	-

Tetracene

$a_g$	$b_{1u}$	$b_{2u}$	$b_{3g}$
307.4	307.4	307.4	307.4
307.4	306.8	307.4	306.8
306.8	306.8	306.8	306.8
306.8	306.8	307.1	306.8
306.8	31.65	306.8	28.49
32.43			25.80
30.32	24.22	27.30	21.08
27.12	21.99	23.14	17.29
22.40	18.87	18.84	16.27
22.20	18.53	17.26	15.99
19.64	15.74	16.37	13.50
17.14	15.36	14.53	-
16.77	13.24	-	-
13.92	-	-	-
12.74	-	-	-
$b_{3u}$	$b_{2g}$	$b_{1g}$	$a_u$
15.76	14.73	12.07	10.99
12.98	10.88	9.25	6.86
8.78	-	-	-

Pentacene

$b_{3g}$	$b_{2u}$	$b_{1u}$	$a_g$
307.4	307.4	307.4	307.4
307.4	307.4	307.4	307.4
306.9	306.9	306.9	306.9
306.9	306.9	306.9	306.9
306.8	306.8	306.8	306.8
28.90	306.8	32.01	306.8
27.07	29.41	30.08	32.60
23.40	28.16	27.23	31.19
19.80	25.79	22.93	28.69
17.13	21.66	21.99	24.70
16.29	17.96	19.56	22.45
15.79	17.32	18.27	21.82
13.74	16.60	17.08	19.69
-	16.02	15.88	18.51
-	14.51	14.47	16.84
-	-	12.77	15.64
-	-	-	13.56
-	-	-	12.54
$b_{3u}$	$b_{2g}$	$b_{1g}$	$a_u$
15.95	15.19	12.37	11.64
14.06	12.49	10.40	8.65
10.56	8.70	6.58	-

Planar Biphenyl

$a_g$	$b_{2u}$	$b_{1u}$	$b_{3g}$
307.4	307.4	306.9	306.9
306.9	306.9	306.8	306.8
306.9	306.9	-	27.65
306.8	306.8	27.98	22.03
32.19	31.50	23.28	16.73
28.94	-	-	-
23.28	27.06	17.95	16.03
20.23	21.66	-	13.05
17.91	-	-	-
16.54	18.90	16.69	-
13.51	16.69	14.04	-
-	14.30	-	-
$b_{3u}$	$b_{1g}$	$b_{2g}$	$a_{1u}$
15.33	14.35	10.02	9.89
11.21	8.76	-	-

Phenanthrene

$a_1$	$a_1$	$b_2$	$b_2$
307.3	23.31	307.3	21.57(B)
307.2	22.69	307.2	19.80
306.8	19.71(B)	306.8	17.92(B)
306.8	18.54(B)	306.8	16.62(B)
306.8	17.90(B)	306.8	16.10
306.8	16.67(B)	306.8	14.06(B)
306.7	16.32(B)	306.7	12.82(B)
32.31(B)	15.67	31.11(B)	-
29.66	14.98	27.80(B)	-
28.64(B)	14.09(B)	26.76(B)	-
26.57(B)	13.01(B)	22.99	-
	$b_1$	$a_2$	
	15.60	14.08	
	12.28	10.09	
	11.02	8.61	
	8.21	-	

Pyrene

$a_g$	$b_{1u}$	$b_{2u}$	$b_{3g}$
307.3	307.3	307.3	307.3
307.3	306.8	307.3	306.8
307.9	306.8	307.9	306.8
306.8	30.53(N)	306.8	28.32(N)
306.8	26.31(N,P)	306.8	23.26(P)
32.76(N,P)	22.59(N,P)	31.31(N,P)	19.35(N)
29.37(P)	18.94(N)	27.07(N)	16.77(N)
27.68(N)	16.74(N)	23.08(N)	15.06(P)
23.31(N)	15.33(N)	20.31(P)	12.96
19.63(P)	14.22	17.72(P)	-
18.46(N)	-	15.20(N)	-
16.64(P)	-	14.28(P)	-
16.07(N)	-	-	-
13.00(N)	-	-	-
$b_{3u}$	$b_{1g}$	$b_{2g}$	$a_u$
16.05	14.24	13.37	10.60
11.63	8.79	7.45	-
10.02	-	-	-

N = Naphthaline like orbitals

P = Phenanthrene like orbitals

## 14 - Annulene (benzene geometry)

$a_g$	$b_{1u}$	$b_{2u}$	$b_{3g}$
306.7	306.7	306.7	306.7
306.6	306.4	306.6	306.4
306.4	305.8	306.4	305.8
305.8	30.22	305.8	28.10
32.44	25.97	31.34	23.05
28.99	21.19	27.45	18.62
27.45	18.38	24.05	15.58
22.52	16.14	20.31	15.00
19.06	14.24	17.89	9.94
18.02	11.23	15.66	-
16.41	-	14.13	-
13.91	-	-	-
$b_{3u}$	$b_{2g}$	$b_{1g}$	$a_{1u}$
14.45	13.31	13.63	10.69
11.05	6.63	6.95	-



Benzofuran

a'	Character	a'	Character
561.4	1s <sub>1</sub>	19.78	2s <sub>C</sub> + 2p <sub>C+O</sub>
309.4	1s <sub>7a</sub>	18.88	C + CH (L)
308.9	1s <sub>2</sub>	18.41	C (L)
307.4	1s <sub>3a</sub>	16.92	C + CH (T)
307.3	1s <sub>7</sub>	16.68	C (L + T)
307.3	1s <sub>4</sub>	16.00	C + CH (L + T)
307.2	1s <sub>6</sub>	15.06	C + O (T)
307.1	1s <sub>3</sub>	14.66	C + O (L + T)
307.0	1s <sub>5</sub>	13.30	C + CH (T)
39.10	2s <sub>1</sub>		
31.76	2s <sub>C</sub>		
29.68	2s <sub>C</sub>	a''	Character
27.89	2s <sub>C</sub>	17.66	C + O
27.34	2s <sub>C</sub>	14.21	C + O
23.35	2s <sub>C</sub>	11.98	C + O
22.34	2s <sub>C</sub> + 2p <sub>C</sub>	9.67	C + O
21.82	2s <sub>O</sub> + 2p <sub>C</sub>	9.16	C

Benzothiophen

a'	Character	a'	Character
2495	1s <sub>1</sub>	21.71	2s <sub>C</sub> + 2p <sub>C</sub>
307.7	1s <sub>7a</sub>	21.22	2s <sub>C</sub> + 2p <sub>C</sub>
307.4	1s <sub>2</sub>	19.16	2s <sub>C</sub> + 2p <sub>C</sub>
307.2	1s <sub>3a</sub>	18.30	C + CH (L)
307.1	1s <sub>7</sub>	17.68	C + S (L + T)
307.0	1s <sub>4</sub>	16.64	C + CH (T)
307.0	1s <sub>6</sub>	16.52	C (L+T)
306.9	1s <sub>3</sub>	15.28	C (L)
306.8	1s <sub>5</sub>	14.32	C + CH (L)
238.1	2s <sub>1</sub>	13.87	C + S + CH (T)
180.7	2p <sub>z</sub> (1)	12.97	C + S (T)
180.6	2p <sub>y</sub> (1)		
33.13	2s <sub>C</sub> + 3s <sub>1</sub>	a''	Character
30.72	2s <sub>C</sub>	180.6	2p <sub>x</sub> (1)
28.58	2s <sub>C</sub>	16.07	C + S
27.10	2s <sub>C</sub>	13.32	C + S
26.37	2s <sub>C</sub>	11.45	C + S
23.30	2s <sub>C</sub>	9.21	C
		8,76	C + S

Indole

a'	Character	a'	Character
424.9	1s <sub>1</sub>	20.13	C + N (L + T)
307.9	1s <sub>7a</sub>	18.22	C + N + CH (L)
307.6	1s <sub>2</sub>	17.96	C + CH (L)
306.7	1s <sub>7</sub>	16.60	C + N (L + T)
306.5	1s <sub>6</sub>	16.36	C (T)
306.5	1s <sub>4</sub>	15.76	C + N + CH (T)
306.5	1s <sub>3a</sub>	15.24	C + N + CH (L)
306.3	1s <sub>5</sub>	13.46	C + CH (L)
305.8	1s <sub>2</sub>	12.99	C (T)
35.34	2s <sub>1</sub>		
39.90	2s <sub>c</sub>	a''	Character
28.69	2s <sub>c</sub>	16.73	C + N
27.26	2s <sub>c</sub>	14.06	C + N
26.20	2s <sub>c</sub>	11.08	C + N
23.01	2s <sub>c</sub> + 2p <sub>N+C</sub>	8.83	C + N
21.98	2s <sub>c</sub> + 2p <sub>N+C</sub>	8.20	C + N
21.42	2s <sub>c</sub> + 2p <sub>N+C</sub>		

Benzo(c)furan

$a_1$	Character	$b_2$	Character
562.7	$1s_2$	308.8	$1s_{1-3}$
308.7	$1s_{1+3}$	307.4	$1s_{3a-7a}$
307.4	$1s_{3a+7a}$	307.2	$1s_{4-7}$
307.2	$1s_{4+7}$	307.0	$1s_{5-6}$
307.0	$1s_{5+6}$	28.75	$2s_c (b_{2u})$
40.30	$2s_2$	26.87	$2s_c (b_{3g})$
32.15	$2s_c$	22.38	CO ( $b_{2u}$ )
29.34	$2s_c$	18.93	CO + CC (L + T, $b_{3g}$ )
24.08	$2s_c$	17.15	CH + CO (T, $b_{3g}$ )
22.00	CH + CC (T, $a_{1g}$ )	16.70	CH + CC (L + T, $b_{3g}$ )
20.11	CH + CC (L + T, $b_{1u}$ )	14.71	CH + CC (L, $b_{3g}$ )
18.33	CC (L, $b_{3g}$ )		
16.72	CC + CO (L + T, $a_{1g}$ )		
15.12	CC + O (L, $b_{1u}$ )		
13.49	CC (L, $a_{1g}$ )		
$b_1$	Character	$a_2$	Character
18.20	CO + CC ( $b_{3u}$ )	11.53	CC ( $b_{1g}$ )
14.58	CC ( $b_{2g}$ )	8.06	CC ( $a_{1u}$ )
10.87	CC ( $b_{3u}$ )		

Benzo(c)thiophen (spd + 3s')

$a_1$	Character	$b_2$	Character
2495	$1s_2$	307.2	$1s_{3a-7a}$
307.2	$1s_{3a+7a}$	307.1	$1s_{1-3}$
307.1	$1s_{1-3}$	306.8	$1s_{4-7}$
306.8	$1s_{4-7}$	306.7	$1s_{5-6}$
306.7	$1s_{5-6}$	180.9	$2p_2$
238.3	$2s_2$	28.41	$2s_c (b_{2u})$
180.9	$2p_2$	26.04	$2s (b_{3g})$
32.82	$2s_c + 3s_2 (a_{1g})$	21.43	CC (L, $b_{2u}$ )
30.49	$2s_c (b_{1u})$	17.89	CC (L, $b_{3g}$ )
26.99	$2s_c (a_{1g})$	16.51	CH <sub>4-7</sub> (T)
22.92	$2s_c + 2p_c (L, b_{1u})$	15.46	CH <sub>1-3</sub> (T)
21.45	CC <sub>3a,7a</sub> + CH (T, $b_{1u}$ )	14.09	CS + CC (L, $b_{3g}$ )
19.33	CH + CC (T, $b_{1u}$ )		
17.70	CH + CC (L)		
16.12	CH + CC (T, $a_{1g}$ )		
13.55	CS + CH + CC (L + T)		
13.29	CS + CC (L + T, $a_{1g}$ )		
$b_1$	Character	$a_2$	Character
180.8	$2p_2$	11.29	CC ( $b_{1g}$ )
15.86	CC + CS ( $b_{3u}$ )	7.84	CC ( $a_{1u}$ )
13.16	CC + CS ( $b_{2g}$ )		
9.69	CC ( $b_{3u}$ )		

Isoindole

$a_1$	Character	$b_2$	Character
425.7	$1s_N$	307.1	$1s_{4-7}$
307.1	$1s_{1+3}$	306.4	$1s_{3a-7a}$
306.4	$1s_{4+7}$	306.4	$1s_{5-6}$
306.4	$1s_{3a+7a}$	27.82	$2s_c (b_{2u})$
306.2	$1s_{5+6}$	26.04	$2s_c (b_{3g})$
35.49	$2s_N + 2s_c (a_{1g})$	21.59	CC + CN ( $b_{2u}$ )
30.98	$2s_c - 2s_N (b_{1u})$	17.87	CC + CH ( $b_{3g}$ )
28.12	$2s_c (a_{1g})$	16.24	CC + CH (T, $b_{3g}$ )
23.68	$2s_c + 2p_c (L, b_{1u})$	15.95	CC + CH (T, $b_{3g}$ )
21.19	CC <sub>3a,7a</sub> + CH (T, $a_{1g}$ )	13.90	CC (L, $b_{3g}$ )
20.00	CH <sub>1,3</sub> + NH ( $b_{1u}$ )		
18.44	CC + CH (L, $a_{1g}$ )		
16.41	CH + CN (L + T, $b_{1u}$ )		
15.34	CC (L + T, $b_{1u}$ )		
12.80	CC (L, $a_{1g}$ )		
$b_1$	Character	$a_2$	Character
16.68	CC + CN ( $b_{3u}$ )	10.64	CC ( $b_{1g}$ )
13.55	CN ( $b_{2g}$ )	7.23	CC ( $a_{1u}$ )
9.53	CC + N ( $b_{3u}$ )		

N-Methyl Isoindole. The symmetry in brackets gives the symmetry of the corresponding orbital in isoindole.

a'	Character	a'	Character
425.9	1s <sub>2</sub>	19.27	CH + CC (T, a <sub>1</sub> )
308.2	1s <sub>c(Me)</sub>	18.70	CN + CC (L + T, b <sub>2</sub> )
306.9	1s <sub>1</sub>	17.88	CC + CH
306.8	1s <sub>3</sub>	16.46	CH (T, b <sub>2</sub> )
306.3	1s <sub>4</sub>	16.10	CC + CN (L + T, a <sub>1</sub> )
306.3	1s <sub>7</sub>	15.89	CH + CC (L + T, b <sub>2</sub> )
306.3	1s <sub>7a</sub>	15.60	CN + CH <sub>Me</sub> (T, b <sub>2</sub> )
306.2	1s <sub>3a</sub>	15.12	CH + CC (L, a <sub>1</sub> )
306.1	1s <sub>6</sub>	13.73	CH + CC (L, b <sub>2</sub> )
306.1	1s <sub>5</sub>	12.71	CH + CC (T, b <sub>2</sub> )
35.68	2s <sub>2</sub> (a <sub>1</sub> )		
31.05	2s <sub>c</sub> (a <sub>1</sub> )	a''	Character
28.82	2s <sub>c</sub> (a <sub>1</sub> )	18.04	N + CH <sub>2</sub> (b <sub>1</sub> )
27.75	2s <sub>c</sub> (b <sub>2</sub> )	15.42	CC + CH <sub>2</sub> (b <sub>1</sub> )
26.03	2s <sub>c</sub> (b <sub>2</sub> )	13.16	CN + CC (b <sub>1</sub> )
25.93	CH <sub>Me</sub> (a <sub>1</sub> )	10.55	CC (a <sub>2</sub> )
22.49	CH + CC (T, a <sub>1</sub> )	9.26	N + CC (b <sub>1</sub> )
21.66	CN + CC (L + T, b <sub>2</sub> )	7.14	CC (a <sub>2</sub> )
20.83	CH + CC (T, a <sub>1</sub> )		

1H - Indazole

$a'(\sigma)$	$a'(\sigma)$	$a'(\sigma)$	$a''(\pi)$
425.7	37.42	18.90	17.58
424.8	31.68	18.03	14.01
308.8	31.00	17.51	12.21
307.9	28.32	16.88	9.55
307.3	27.79	16.38	9.25
307.3	23.66	14.77	
307.1	23.37	13.64	
307.1	22.22	12.40	
306.8	20.64		

2H - Indazole

$a'(\sigma)$	$a'(\sigma)$	$a'(\sigma)$	$a''(\pi)$
426.85	37.52	18.94	17.55
423.80	31.70	18.60	14.07
308.3	30.14	17.03	11.44
307.8	28.72	16.58	10.14
306.9	27.28	16.11	8.11
306.9	24.70	14.50	
306.8	22.19	13.84	
306.8	21.91	11.84	
306.7	20.58		



1H - Benzimidazole

$a'(\sigma)$	$a'(\sigma)$	$a'(\sigma)$	$a''(\pi)$
425.94	37.36	19.57	17.78
422.94	32.42	18.68	13.48
309.30	30.91	17.36	12.19
308.23	28.02	17.29	9.29
307.44	27.72	16.31	8.86
307.08	24.27	15.06	
307.04	23.57	14.71	
307.04	22.71	11.65	
306.78	21.08		

1H - Benzotriazole

$a'(\sigma)$	$a'(\sigma)$	$a'(\sigma)$	$a''(\pi)$
426.69	39.180	19.429	18.297
426.34	33.151	18.478	14.477
424.84	32.128	17.677	12.932
309.19	28.908	17.208	9.990
308.27	28.283	15.221	9.798
307.62	24.728	14.750	
307.62	24.065	13.703	
307.62	22.940	11.785	
307.23	21.067		

2H - Benzotriazole

$a'(\sigma)$	$a'(\sigma)$	$a'(\sigma)$	$a''(\pi)$
427.44	39.38	19.18	18.50
425.59	33.06	18.92	14.45
425.59	32.14	17.00	12.53
308.46	29.74	16.96	10.57
308.46	27.99	15.25	8.99
307.23	25.44	14.95	
307.23	22.84	13.22	
307.22	22.53	13.14	
307.22	20.44		

Benzothiazole (spd + 3s')

a' ( $\sigma$ )	a' ( $\sigma$ )	a' ( $\sigma$ )	a'' ( $\pi$ )
2495.2	38.627	19.003	16.582
423.78	31.307	18.995	13.518
309.08	29.445	17.561	11.973
308.30	28.136	16.656	9.714
307.89	27.419	15.877	9.151
307.56	24.202	15.259	
307.51	22.727	13.763	
307.45	21.864	11.196	
307.32	19.326		

1,2- benzisothiazole (spd + 3s')

a' ( $\sigma$ )	a' ( $\sigma$ )	a' ( $\sigma$ )	a' ( $\sigma$ )	a'' ( $\pi$ )
2495.6	238.66	24.25	14.95	16.28
424.26	181.25	22.73	13.65	13.60
308.40	181.22	21.76	12.15	11.64
308.00	181.14	19.71		9.93
307.63	34.99	19.04		9.10
307.59	31.51	18.41		
307.57	29.01	17.49		
307.46	28.61	16.99		
307.36	26.81	15.37		

2,1- benzisothiazole (spd + 3s')

$a'(\sigma)$	$a'(\sigma)$	$a'(\sigma)$	$a'(\sigma)$	$a''(\pi)$
2496.5	239.46	22.07	13.83	
423.49	182.06	21.96	12.02	181.97
308.32	182.02	19.50		16.69
307.77	35.17	18.16		13.72
307.59	31.40	18.09		11.87
307.17	29.57	16.79		10.01
307.11	27.65	16.56		8.59
307.06	27.30	14.65		
	23.31	14.17		

Benzoxazole

$a'(\sigma)$	$a'(\sigma)$	$a'(\sigma)$	$a''(\pi)$
562.88	41.461	20.243	19.099
424.19	34.218	19.292	14.376
310.79	31.589	17.945	13.167
309.63	28.675	17.262	10.110
308.35	28.421	16.508	9.625
307.86	24.728	15.734	
307.70	23.402	14.730	
307.70	22.882	12.594	
307.51	20.523		

1,2-benzisoxazole

$a'(\sigma)$	$a'(\sigma)$	$a'(\sigma)$	$a''(\pi)$
562.63	41.160	19.606	18.787
426.42	32.813	18.684	14.811
310.17	32.389	17.419	13.345
309.17	29.084	17.131	10.299
308.02	28.468	15.859	10.217
307.92	24.319	15.421	
307.92	23.172	14.077	
307.86	22.811	13.280	
307.51	20.395		

Anthranil

$a'(\sigma)$	$a'(\sigma)$	$a'(\sigma)$	$a''(\pi)$
563.48	41.257	19.416	18.695
425.38	32.979	18.484	15.044
310.17	31.141	17.418	12.344
308.98	29.529	17.233	11.292
307.99	28.015	15.581	8.896
307.78	24.736	15.247	
307.72	22.940	14.589	
307.56	22.656	12.729	
307.45	20.327		

1,2,3- benzthiadiazole (spd + 3s')

a' ( $\sigma$ )	a' ( $\sigma$ )	a' ( $\sigma$ )	a'' ( $\pi$ )
2496.65	38.185	19.620	17.430
425.71	32.004	18.894	14.109
425.44	30.430	17.975	12.529
308.70	29.220	16.585	10.378
308.38	28.310	15.820	9.597
308.08	24.815	14.639	
307.94	23.305	14.070	
307.29	22.480	11.181	
306.78	19.838		

2,1,3- benzthiadiazole (spd + 3s')

a' ( $\sigma$ )	a' ( $\sigma$ )	a' ( $\sigma$ )	a'' ( $\pi$ )
2497.50	36.573	18.771	17.297
2 x 424.27 (N1/3)	32.027	18.545	14.092
2 x 308.79	31.081	17.106	12.574
2 x 307.51	28.117	15.291	10.431
2 x 307.50	28.070	15.103	9.138
	23.864	14.654	
	22.581	12.842	
	22.356	12.098	
	19.167		

2,1,3- benzoxadiazole

$a'(\sigma)$	$a'(\sigma)$	$a'(\sigma)$	$a''(\pi)$
563.58	41.98	19.83	19.20
427.31	33.61	19.33	15.42
427.31	33.47	17.76	13.39
309.76	30.78	16.58	11.90
309.76	28.79	15.78	9.78
308.21	25.76	15.12	
308.21	23.47	14.08	
308.19	23.45	13.95	
308.19	20.11		



pyridine - N - oxide

$\sigma$	$\sigma$	$\sigma$	$\pi$
556.21	38.08	18.08	13.76
429.70	32.83	17.99	11.43
309.38	29.67	17.87	8.68
309.38	29.67	16.80	
308.42	24.89	15.61	
308.42	24.20	14.52	
308.26	20.66	10.60	
308.26	19.94		

pyrimidine - N - oxide

$\sigma$	$\sigma$	$\sigma$	$\pi$
557.98	40.36	21.18	19.73
431.41	36.15	19.24	15.18
425.62	32.42	17.81	12.51
310.59	30.73	17.63	9.34
310.02	26.63	16.01	
309.51	25.24	13.32	
309.02	21.35	11.58	

pyrazine - N - oxide

$\sigma$	$\sigma$	$\sigma$	$\pi$
558.5	40.93	21.04	19.84
431.9	36.22	19.76	15.49

## pyrazine - N - oxide (cont'd)

$\sigma$	$\sigma$	$\sigma$	$\pi$
425.0	32.85	18.44	12.11
309.8	30.61	17.25	9.68
309.8	26.28	16.50	
309.7	25.65	12.49	
309.7	21.33	12.04	

Open shell RHF orbital energies (-eV)

2H - indene RHF triplet

$\sigma$	$\sigma$	$\sigma$	$\pi$
306.94	32.42	17.69	16.07
306.94	29.98	17.23	14.12
306.56	27.85	16.22	10.05
306.56	26.37	15.54	9.68
306.53	24.92	14.47	
306.53	22.71	14.25	SOMO 10.79
306.37	20.80	13.53	SOMO 7.11
305.68	20.59	12.98	

Indene RHF triplet

$\sigma$	$\sigma$	$\sigma$	$\pi$
306.83	31.93	18.72	16.15
306.73	29.82	17.67	13.90
306.69	27.77	16.94	9.94
306.67	25.93	16.20	9.32
306.65	25.20	15.91	
306.65	22.28	14.57	SOMO 1.34
306.35	21.30	13.80	SOMO 8.99
305.06	19.99	13.42	
		12.79	

benzofuran RHF triplet

$\sigma$	$\sigma$	$\sigma$	$\pi$
560.32	39.10	18.53	13.09
309.42	31.64	18.17	10.44
307.91	29.00	16.83	9.34
307.60	27.77	16.61	
307.42	26.54	16.45	SOMO 13.90
307.19	23.07	15.78	SOMO 6.82
307.18	22.14	14.73	
307.15	21.24	14.41	
306.23	19.53	13.84	

benzo(c)furan RHF triplet

$\sigma$	$\sigma$	$\sigma$	$\pi$
560.51	39.17	18.61	16.34
308.17	32.04	18.13	10.24
308.16	29.15	16.89	10.18
307.68	28.41	16.54	
307.68	26.09	16.53	SOMO 11.75
307.11	23.87	14.65	SOMO 6.97
307.11	22.02	14.40	
307.02	21.70	14.14	
307.02	19.84	13.48	

benzothiophen RHF triplet

$\sigma$	$\sigma$	$\sigma$	$\sigma$	$\pi$
2495.40	238.40	22.41	14.05	181.00
307.42		21.39	13.72	15.42
306.93	180.97	20.97	12.83	11.61
306.80	180.87	18.92		10.20
306.68	32.80	18.03		8.95
306.66	30.15	17.37		
306.48	27.92	16.28		SOMO 12.27
306.16	26.70	16.20		SOMO 5.95
305.85	25.81	15.09		

benzo(c)thiophen RHF triplet

$\sigma$	$\sigma$	$\sigma$	$\sigma$	$\pi$
2493.69	236.92	21.24	13.56	
307.50	179.50	21.17	12.67	179.31
307.50	179.50	19.31		15.03
306.95	32.62	17.84		11.99
306.95	30.27	17.66		10.08
306.91	28.29	16.60		9.45
306.91	26.75	16.17		
306.60	25.45	15.26		SOMO 11.65
306.60	22.89	14.03		SOMO 6.87

Indole RHF triplet

$\sigma$	$\sigma$	$\sigma$	$\pi$
423.52	34.67	17.89	12.78
308.32	30.77	17.75	9.91
306.68	28.24	16.39	8.41
306.63	27.07	16.31	
306.52	25.83	15.53	SOMO 12.47
306.41	22.74	15.41	SOMO 5.25
306.19	21.65	15.09	
305.67	20.91	13.97	
305.60	19.87	13.01	

Isoindole RHF triplet

$\sigma$	$\sigma$	$\sigma$	$\pi$
422.84	34.38	18.20	14.97
306.96	30.89	17.68	12.69
306.96	27.88	16.33	9.55
306.58	27.64	16.14	9.13
306.58	25.20	15.60	
306.32	23.34	15.30	SOMO 10.91
306.32	21.34	13.92	5.36
306.31	20.97	12.94	
306.31	19.50		

APPENDIX 3

Integral Evaluation

Fig.1

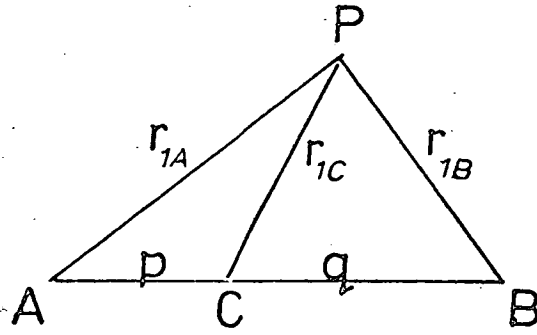


Fig.2

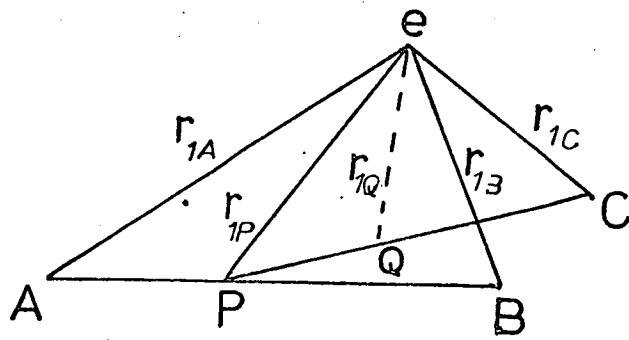
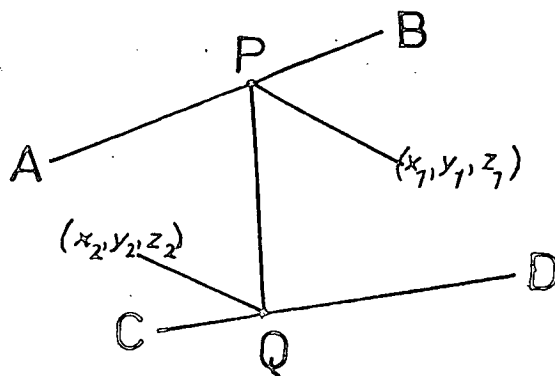


Fig.3





### Integral Evaluation

The basis of the LCGO calculations is the use of gaussians to represent the atomic orbitals. The general form of a gaussian is given by:-

$$G_i(r_A) = G(a_i; r_A) = x^l y^m z^n \exp(-a_i r_A^2)$$

where  $a_i$  is a positive parameter,  $r_A$  is the magnitude of a vector  $r$  which is the vector from a fixed point, A, to a variable point P (x, y, z), and  $x^l y^m z^n$  represents the angular distribution of the function with l, m and n having integral values. The gaussian  $G_i(r_A)$  can be expressed as follows.

$$G_i(r_A) = \exp[-a_i ((x - A_x)^2 + (y - A_y)^2 + (z - A_z)^2)]$$

This produces three simple gaussians with respect to x, y, and z separately.

It can be shown mathematically<sup>1</sup> that the product of two gaussians with different centres A and B can be represented by a third gaussian centred at point C which lies on the line AB.

$$G_i(r_A) \cdot G_j(r_B) = K G_k(r_C)$$

where K is a constant given by:-

$$K = \exp\left[\frac{-a_i a_j}{a_i + a_j} \cdot \overline{AB}^2\right] \text{ and } a_k = (a_i + a_j)$$

Thus for gaussians with fixed centres, the product of any number of gaussians will itself be a gaussian with a multiplying factor.

The integrals to be evaluated are as follows:- the overlap integral, the kinetic energy integral, the potential energy integral and the electron

1. I. Shavitt, "The gaussian function in calculations of statistical mechanics and quantum mechanics," from "Methods in Computational Physics." Ch. 1, Academic Press, 1963.

repulsion integral. All of these can be expressed in terms of the overlap integral producing a large saving in computer time.

### The overlap integral

The overlap integral can be represented as  $\langle aA | bB \rangle$  where  $a$  and  $b$  are the exponents of the gaussians centred at  $A$  and  $B$ , Fig. 1. The simplest integral of this type will be that for two  $s$ -type functions where  $l$ ,  $m$ , and  $n$  are zero in the gaussians.

$$\begin{aligned} \langle aA | bB \rangle = I &= \int_0^{\infty} \exp(-a r_{1A}^2) \exp(-b r_{1B}^2) d\tau \\ &= \int \exp(-a r_{1A}^2 - b r_{1B}^2) d\tau \end{aligned}$$

$AB (=R)$  can be partitioned into lengths  $p$  and  $q$ , corresponding to the ratio of exponents.

$$AC = \frac{b}{a+b} \cdot R = p; \quad BC = \frac{a}{a+b} \cdot R = q$$

Using the cosine rule:-

$$r_{1A}^2 = p^2 + r_{1c}^2 + 2pr_{1c} \cos\Theta$$

$$r_{1B}^2 = q^2 + r_{1c}^2 - 2qr_{1c} \cos\Theta$$

Eliminating  $\cos\Theta$  :-

$$\begin{aligned} qr_{1A}^2 + pr_{1B}^2 &= p^2q + qr_{1c}^2 + pq^2 + pr_{1c}^2 \\ &= r_{1c}^2(p+q) + pq(p+q) \\ &= (p+q)(r_{1c}^2 + pq) \end{aligned}$$

Substituting for  $p$  and  $q$ :-

$$\frac{a}{a+b} \cdot R r_{1A}^2 + \frac{b}{a+b} \cdot R r_{1B}^2 = R \left[ \frac{ab}{(a+b)^2} \cdot R^2 + r_{1c}^2 \right]$$

Multiply both sides by  $\frac{R}{a+b}$

$$ar_{1A}^2 + br_{1B}^2 = \frac{ab}{a+b} \cdot R^2 + r_{1c}^2 (a+b)$$

$$\therefore I = \int \exp\left(\frac{-ab}{a+b}\right) R^2 \exp(-r_{1c}^2 (a+b)) dr$$

$$\text{But } d\tau = 4\pi r_{1c}^2 dr_{1c}$$

$$\therefore I = \exp\left(\frac{-ab}{a+b}\right) R^2 \int \exp(-r_{1c}^2 (a+b)) 4\pi r_{1c}^2 dr_{1c}$$

Using a gamma function  $\Gamma$

$$\Gamma(n) = \int_0^{\infty} r^\lambda \exp(-ar^2) dr$$

$$= \frac{1}{2} a^{-(\lambda+1)/2} \Gamma\left(\frac{\lambda+1}{2}\right)$$

In I,  $\lambda = 2$  and  $a = a+b$  giving

$$\Gamma(n) = \frac{1}{2} (a+b)^{-\frac{3}{2}} \Gamma\left(\frac{3}{2}\right)$$

But  $\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\pi^{\frac{1}{2}}$  and I becomes:-

$$\begin{aligned} I &= \exp\left(\frac{-ab}{a+b}\right) R^2 \cdot 4\pi \cdot \frac{1}{2} (a+b)^{\frac{3}{2}} \cdot \frac{1}{2}\pi^{\frac{1}{2}} \\ &= \left(\frac{\pi}{a+b}\right)^{\frac{3}{2}} \exp\left(\frac{-ab}{a+b} \cdot R^2\right) \end{aligned}$$

The general overlap integral is denoted  $S_{ab}^{cd}$  where c and d are l, m or n of  $x^l, y^m, z^n$ . The overlap integral just evaluated will be denoted  $S_{ab}^{00}$  as l, m and n = 0.

Higher overlap integrals can be expressed in terms of  $S_{ab}^{00}$ . For example,

consider the integral  $\langle p_x | s \rangle$  represented by  $S_{ab}^{lo}$ , then as before:-

$$\langle aA | bB \rangle = I = \int_0^\infty x_A \exp(-ar_{1A}^2) \exp(-br_{1B}^2) d\tau$$

$$\text{But } \exp(-ar_{1A}^2) = \exp -a [(x_1 - A_x)^2 + (y_1 - A_y)^2 + (z_1 - A_z)^2]$$

where  $x_1, y_1, z_1$  are the coordinates of an electron at a variable point, P, and  $A_x, A_y$  and  $A_z$  are the coordinates of A.

$$\begin{aligned} \frac{\partial}{\partial A_x} (\exp -ar_{1A}^2) &= \exp(-ar_{1A}^2) 2a (x_1 - A_x) \\ &= \exp(-ar_{1A}^2) 2a x_A \end{aligned}$$

$$\therefore x_A \exp(-ar_{1A}^2) = \frac{1}{2a} \frac{\partial}{\partial A_x} (\exp -ar_{1A}^2)$$

$$\therefore I = \int \frac{1}{2a} \frac{\partial}{\partial A_x} [\exp(\frac{-ab}{a+b} \frac{AB^2}{AB})] \cdot (\frac{\pi}{a+b})^{\frac{3}{2}} d\tau$$

$$\frac{AB^2}{AB} = (A_x - B_x)^2 + (A_y - B_y)^2 + (A_z - B_z)^2$$

$$\therefore \frac{\partial}{\partial A_x} \exp(\frac{-ab}{a+b} \frac{AB^2}{AB}) = \frac{-2ab}{a+b} (A_x - B_x) \cdot \exp(\frac{-ab}{a+b} \frac{AB^2}{AB})$$

$$I = S_{ab}^{lo} = -\frac{b}{a+b} (A_x - B_x) S_{ab}^{oo}$$

The general form of the overlap integrals are given as follows for s and p orbitals:-

$$S_{ab}^{oo} = (\frac{\pi}{a+b}) \exp(\frac{-ab}{a+b} \frac{AB^2}{AB})$$

$$S_{ab}^{io} = \langle p_{ia} | s \rangle = -\frac{b}{a+b} (A_i - B_i) S_{ab}^{oo}$$

$$S_{ab}^{oi} = \langle s | p_{ib} \rangle = \frac{a}{a+b} (A_i - B_i) S_{ab}^{oo}$$

$$S_{ab}^{ij} = \langle p_{ia} | p_{jb} \rangle = [ \frac{-ab}{(a+b)^2} (A_i - B_i)(A_j - B_j) + \frac{\delta_{ij}}{2a(a+b)} ] S_{ab}^{oo}$$

The kinetic energy integral

This integral is represented by  $\langle aA | K | bB \rangle$  where  $K = -\frac{1}{2} \nabla^2$  and will take the form shown for two s-type gaussian functions.

$$I = \int_0^{\infty} (\exp(-ar_{1A}^2)) \left(-\frac{1}{2} \nabla^2\right) (\exp(-br_{1B}^2)) d\tau$$

where  $\nabla^2 = \frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} + \frac{\partial}{\partial z^2}$

and applies to  $\exp(-br_{1B}^2)$

Writing  $r_{1B}^2$  as  $x^2 + y^2 + z^2$  and  $\exp(-br_{1B}^2)$  as  $f$ , then

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial}{\partial x} (\exp(-b(x^2 + y^2 + z^2))) \\ &= -2bx \exp(-br_{1B}^2) = -2bxf \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= -2bf - (2bx \cdot -2bxf) \\ &= -2bf + 4b^2 x^2 f \end{aligned}$$

$$\therefore \nabla^2 = -6bf + 4b^2(x^2 + y^2 + z^2)f$$

$$-\frac{1}{2} \nabla^2 = 3b \exp(-br_{1B}^2) + 2b^2 r_{1B}^2 \exp(-br_{1B}^2)$$

$$\therefore I = b \int_0^{\infty} \exp(-ar_{1A}^2) \exp(-br_{1B}^2) (3-2br_{1B}^2) d\tau$$

Centering the gaussians at a point C gives:-

$$I = b \int (3-2br_{1B}^2) \exp\left(\frac{-ab}{a+b} \frac{\overline{AB}^2}{2}\right) \exp(-(a+b)r_{1C}^2) d\tau$$

Setting  $k = b \exp\left(\frac{-ab}{a+b} \frac{\overline{AB}^2}{2}\right)$  gives

$$I = k \int (3-2br_{1B}^2) \exp(-(a+b)r_{1C}^2) d\tau$$

Defining  $r_{1B}$  in terms of  $r_{1C}$  and  $\Theta$  as for the overlap integral produces

$$I = k \int \left( 3 - \frac{2ba^2}{(a+b)^2} \frac{\overline{AB}^2}{AB} - 2br_{1C}^2 \frac{-4b^2}{a+b} \frac{\overline{AB}}{AB} \cos \Theta \right) \exp(-(a+b)r_{1C}^2) dr_{1C} d\Theta$$

But  $d\tau = 4\pi r_{1C}^2 dr_{1C}$  and  $\int_0^{2\pi} A \cos \Theta = 0$ , giving

$$I = k \int 4\pi \left( \underbrace{3}_{\text{Term 1}} - \underbrace{\frac{2ba^2}{(a+b)}}_{\text{Term 2}} \frac{\overline{AB}^2}{AB} - \underbrace{2br_{1C}^2}_{\text{Term 3}} \right) \underbrace{r_{1C}^2 \exp(-(a+b)r_{1C}^2)}_{\text{Term 4}} dr_{1C}$$

$$\text{Term 1} \quad 12\pi b \exp\left(\frac{-ab}{a+b} \frac{\overline{AB}^2}{AB}\right)$$

$$\text{Term 2} \quad -8\pi \frac{a^2 b^2}{(a+b)^2} \frac{\overline{AB}^2}{AB} \cdot \exp\left(\frac{-ab}{a+b} \frac{\overline{AB}^2}{AB}\right)$$

$$\text{Term 3} \quad -8\pi b^2 r_{1C}^2 \exp\left(\frac{-ab}{a+b} \frac{\overline{AB}^2}{AB}\right)$$

$$\begin{aligned} \text{Using } \int_0^{\infty} r_{1C}^2 \exp(-ar_{1C}^2) dr_{1C} &= \frac{1}{2} a^{-3/2} \Gamma^{3/2} \\ &= \frac{(a+b)^{-3/2}}{2} \cdot \frac{\pi^{1/2}}{2} = \frac{(a+b)^{-3/2}}{4} \pi^{1/2} \end{aligned}$$

$$\text{Terms 1} \times 4 \quad 3b \left(\frac{\pi}{a+b}\right)^{3/2} \cdot \exp\left(\frac{-ab}{a+b} \frac{\overline{AB}^2}{AB}\right)$$

$$\text{Terms 2} \times 4 \quad -2 \frac{a^2 b^2}{(a+b)^2} \left(\frac{\pi}{a+b}\right)^{3/2} \frac{\overline{AB}^2}{AB} \cdot \exp\left(\frac{-ab}{a+b} \frac{\overline{AB}^2}{AB}\right)$$

$$\text{Terms 3} \times 4 \quad \int -8\pi b^2 \exp\left(\frac{-ab}{a+b} \frac{\overline{AB}^2}{AB}\right) r_{1C}^4 \exp(-(a+b)r_{1C}^2) dr_{1C}$$

$$= -8\pi b^2 \exp\left(\frac{-ab}{a+b} \frac{\overline{AB}^2}{AB}\right) \cdot \frac{(a+b)^{5/2}}{2} \cdot \frac{3\pi^{1/2}}{4}$$

$$= \frac{-3}{a+b} \left(\frac{\pi}{a+b}\right)^{3/2} \exp\left(\frac{-ab}{a+b} \frac{\overline{AB}^2}{AB}\right)$$

$$I = \left( 3b - \frac{2a^2 b^2}{(a+b)^2} - \frac{3b^2}{(a+b)} \right) \left( \frac{\pi}{a+b} \right)^{3/2} \exp \left( -\frac{ab}{a+b} \frac{AB}{AB} \right)$$

$$= \left( \frac{ab}{a+b} \right) \left( \frac{3-2ab}{a+b} \right) S_{ab}^{oo}$$

For the higher kinetic energy integrals, for example,  $\langle p_x | K | s \rangle$ , the integral becomes

$$\int_0^\infty x_A \exp(-a r_{1A}^2) \exp(-b r_{1B}^2) dr$$

As shown for an overlap integral

$$x_A \exp(-a r_{1A}^2) = \frac{1}{2a} \frac{\partial}{\partial A_x} \exp(-a r_{1A}^2)$$

and the integral as evaluated as

$$\langle p_x | K | s \rangle = \frac{1}{2a} \frac{\partial}{\partial A_x} \langle s | K | s \rangle$$

The general form of the kinetic energy integrals is as follows:-

$$T_{ab}^{oo} = K_{ab}^{oo} S_{ab}^{oo}$$

$$T_{ab}^{io} = K_{ab}^{io} S_{ab}^{oo} + K_{ab}^{oo} S_{ab}^{io}$$

$$T_{ab}^{oj} = K_{ab}^{oj} S_{ab}^{oo} + K_{ab}^{oo} S_{ab}^{oj}$$

$$T_{ab}^{ij} = K_{ab}^{ij} S_{ab}^{oo} + K_{ab}^{io} S_{ab}^{oj} + K_{ab}^{oj} S_{ab}^{io} + K_{ab}^{oo} S_{ab}^{ij}$$

where  $K_{ab}^{oo} = \frac{3ab}{a+b} - \frac{2a^2 b^2}{(a+b)^2} \frac{AB}{AB}$

$$K_{ab}^{io} = \frac{-2ab^2}{(a+b)^2} (A_i - B_i)$$

$$K_{ab}^{oj} = \frac{2a^2 b}{(a+b)^2} (A_j - B_j)$$

$$K_{ab}^{ij} = \frac{ab}{(a+b)^2} \delta_{ij} \quad \text{and} \quad \delta_{ij} \quad \text{is the Kronecker delta.}$$

The potential energy integral

This integral represents the coulombic attraction between a charge distribution and a point charge at a point C as shown in Fig. 2, and is given by  $\langle aA | V_c | bB \rangle$

$$I = \langle aA | V_c | bB \rangle = \int_0^\infty \exp(-ar_{1A}^2) \frac{1}{r_{1C}} \exp(-br_{1B}^2) d\tau$$

The gaussians at A and B are first centred at point P, as previously to give:-

$$\int \exp\left(\frac{-ab}{a+b} \frac{\overline{AB}^2}{r_{1C}}\right) \frac{1}{r_{1C}} \exp(-(a+b)r_{1P}^2) d\tau$$

$$= k \int \exp(-(a+b)r_{1P}^2) \frac{1}{r_{1C}} d\tau$$

$\frac{1}{r_{1C}}$  can be expressed as  $\pi^{-1/2} \int_0^\infty a^{-1/2} \exp(-ar_{1C}^2) da$ , giving:-

$$I = \frac{1}{\pi^{1/2} k} \int_0^\infty a^{1/2} \exp(-(a+b)r_{1P}^2) \exp(-ar_{1C}^2) da d\tau$$

The gaussians now centred at P and C are expanded about a point Q, Fig. 2.

$$I = \frac{1}{\pi^{1/2} k} \int_0^\infty \int_0^\infty a^{1/2} \exp\left(-\frac{a(a+b)}{a+b} \frac{\overline{PC}^2}{r_{1q}}\right) \exp(-(a+b)r_{1q}^2) da d\tau$$

$$d\tau = 4\pi r_{1q}^2 dr_{1q}$$

Using a gamma function for  $\int_0^\infty r_{1q}^2 \exp(-(a+b)r_{1q}^2) dr_{1q}$

gives this term equal to  $\frac{1}{2} (a+b) \Gamma\left(\frac{3}{2}\right)$

$$I = 4\pi^{1/2} \frac{1}{2} \pi^{1/2} k \int_0^\infty a^{1/2} (a+b)^{3/2} \exp\left(-\frac{a(a+b)}{a+b} \frac{\overline{PC}^2}{r_{1q}}\right) da$$

To simplify this expression a new variable  $\beta$  is defined as:-

$$\beta^2 = \frac{a(a+b) \overline{PC}^2}{a+b} \dots (1)$$



$$\therefore a = \frac{(a+b)\beta^2}{\beta^2 - (a+b)\overline{PC}} \quad \dots (2)$$

Then  $a=0$  gives  $\beta=0$ , and  $a = \infty$  gives  $\beta = \overline{PC} (a+b)^{1/2}$ . The integral limits can now be expressed in terms of  $\beta$  and the integration performed over  $\beta$ , if  $da$  is expressed in terms of  $d\beta$ . From (1) above:-

$$\overline{PC}^2 (a+b) a = (a+b+a)\beta^2$$

$$\overline{PC}^2 (a+b) \frac{da}{d\beta} = \beta^2 \frac{da}{d\beta} + (a+b+a) 2\beta$$

$$\left[ \overline{PC}^2 (a+b) - \beta^2 \right] da = 2(a+b+a) \beta d\beta$$

$$\left[ \overline{PC}^2 (a+b) - \frac{\overline{PC}^2 (a+b)a}{a+b+a} \right] da = 2(a+b+a) \beta d\beta.$$

$$\left[ \frac{\overline{PC}^2 (a+b)^2}{a+b+a} \right] da = 2(a+b+a) \beta d\beta$$

$$= 2(a+b+a) \frac{a^{1/2}(a+b)^{1/2} \overline{PC}}{(a+b+a)^{1/2}} d\beta$$

$$da = \frac{2(a+b+a)^{3/2}}{\overline{PC} (a+b)^{3/2}} a^{1/2} d\beta$$

$$\therefore a^{1/2} (a+b+a)^{3/2} da = \frac{2d\beta}{\overline{PC} (a+b)^{3/2}}$$

$$I = k\pi \int_0^{\overline{PC}(a+b)} \frac{2 \exp(-\beta^2)}{\overline{PC} (a+b)} d\beta$$

Using the integral transform  $F_0(t) = t^{-1/2} \int_0^t (\exp(-u^2)) du$

$$I = \langle aA | V_c | bB \rangle = \frac{2\pi}{a+b} \exp\left(-\frac{ab}{a+b} \overline{AB}\right) F_0\left(\overline{PC}^2 (a+b)\right)$$

The potential energy integrals are defined, using the following parameters:-

$$L_{c,ab}^{oo} = F_o(t) \text{ as above}$$

$$L_{c,ab}^{io} = L_{c,ab}^{oi} = (p_i - c_i) F_1(t)$$

$$\text{and } L_{c,ab}^{ij} = (p_i - c_i)(p_j - c_j) F_2(t) - \frac{2 F_1(t)}{a+b} \delta_{ij}$$

$$\text{where } F_m(t) = -\frac{d}{dt} F_{m-1}(t)$$

The integrals are then given by:-

$$V_{ab}^{oo} = \Theta \sum_c L_{c,ab}^{oo} S_{ab}^{oo}$$

$$V_{ab}^{io} = \Theta \sum_c (L_{c,ab}^{oo} S_{ab}^{io} + L_{c,ab}^{io} S_{ab}^{oo})$$

$$V_{ab}^{oi} = \Theta \sum_c (L_{c,ab}^{oo} S_{ab}^{oj} + L_{c,ab}^{oj} S_{ab}^{oo})$$

$$V_{ab}^{ij} = \Theta \sum_c (L_{c,ab}^{oo} S_{ab}^{ij} + L_{c,ab}^{oj} S_{ab}^{io} + L_{c,ab}^{ij} S_{ab}^{oo} + L_{c,ab}^{io} S_{ab}^{oj})$$

$$\text{where } \Theta = 2 \left( \frac{a+b}{\pi} \right)$$

### The electron repulsion integral

The general two electron repulsion integral can be represented by  $\langle aAbB \frac{1}{r_{12}} cCdD \rangle$  and is shown diagrammatically in Fig. 3. The overall integral may involve 1, 2, 3 or 4 centres. The simplest integral will be the one where both charge distributions are single centre.

$$I_1 = \langle \phi_A^2(1) \frac{1}{r_{12}} \phi_B^2(2) \rangle$$

Using this expression,  $I_1$  represents the coulomb integral.

When one charge distribution is single centred and the other two centred, the integral becomes

$$I_2 = \langle \phi_A^2(1) \frac{1}{r_{12}} \phi_B(2) \phi_C(2) \rangle$$

or 
$$I_2 = \langle \phi_A^2(1) \frac{1}{r_{12}} \phi_A(2) \phi_B(2) \rangle$$

The most complicated two electron integral is the one obtained when both charge distributions are two centred, and can take the following forms depending on the number of centres involved.

$$I_3 = \langle \phi_A(1) \phi_B(1) \frac{1}{r_{12}} \phi_A(2) \phi_B(2) \rangle \dots 2 \text{ centred}$$

$$I_4 = \langle \phi_A(1) \phi_B(1) \frac{1}{r_{12}} \phi_A(2) \phi_C(2) \rangle \dots 3 \text{ centred}$$

$$I_5 = \langle \phi_A(1) \phi_B(1) \frac{1}{r_{12}} \phi_C(2) \phi_D(2) \rangle \dots 4 \text{ centred}$$

Integrals of this type will be exchange integrals.

The integral evaluation is carried out along the following lines. The gaussians at A and B are collapsed to a single gaussian at P, and those at C and D to a single gaussian at Q, Fig. 3. The operator  $\frac{1}{r_{12}}$  is expressed as follows:-

$$\frac{1}{r_{12}} = \pi^{-\frac{1}{2}} \int_0^\infty a^{-\frac{1}{2}} \exp(-a r_{12}^2) da$$

The potential energy integral is then integrated with respect to  $a$ ,  $r_1$  and  $r_2$ . The integration is simplified by separation into x, y, and z components, and yields for  $I_1$ :-

$$I_1 = \frac{2\pi^{\frac{5}{2}}}{(a+b)(c+d)(a+b+c+d)^{\frac{1}{2}}} \exp\left(\frac{-ab}{a+b} \frac{AB}{AB}\right) \exp\left(\frac{-cd}{c+d} \frac{CD}{CD}\right) F_0(\beta)$$

where 
$$\beta = \frac{(a+b)(c+d) PQ}{(a+b+c+d)}$$

The general expressions for the two electron repulsion integrals are

given in terms of the overlap integral and an intermediate function, G.

In G, the upper integral limit is given as:-

$$\frac{S_1 S_2}{S_4} \frac{2}{PQ} \quad \text{where } S_1 = a+b \quad S_2 = c+d \quad S_4 = a+b+c+d$$

$$G^{oooo} = F_0(t)$$

$$G^{iooo} = G^{oioo} = \frac{S_2}{S_4} (P_i - Q_i) F_1(t)$$

$$G^{ooio} = G^{oooi} = \frac{S_1}{S_4} (P_i - Q_i) F_1(t)$$

$$G^{ijoo} = \frac{S_2}{S_4} \left[ \frac{S_2}{S_4} (P_i - Q_i) (P_j - Q_j) F_2(t) - \delta_{ij} \frac{1}{2S_1} F_1(t) \right]$$

$$G^{ooij} = \frac{S_1}{S_4} \left[ \frac{S_1}{S_4} (P_i - Q_i) (P_j - Q_j) F_2(t) - \delta_{ij} \frac{1}{2S_2} F_1(t) \right]$$

$$G^{iojo} = G^{iooj} = G^{oioj} = G^{oijo}$$

$$= \frac{-1}{S_4} \left[ \frac{S_1 S_2}{S_4} (P_i - Q_i) (P_j - Q_j) F_2(t) - \frac{1}{2} \delta_{ij} F_1(t) \right]$$

$$G^{ijko} = G^{ijok} = \frac{S_2}{S_4} \left[ \frac{S_1 S_2}{S_4} (P_i - Q_i) (P_j - Q_j) (P_k - Q_k) F_3(t) \right. \\ \left. - \frac{1}{2} (\delta_{ij} (P_k - Q_k) + \delta_{ik} (P_j - Q_j) + \delta_{jk} (P_i - Q_i)) F_2(t) \right]$$

$$G^{oijk} = G^{iojk} = \frac{S_1}{S_4} \left[ (P_i - Q_i) (P_j - Q_j) (P_k - Q_k) F_3(t) \right. \\ \left. - \frac{1}{2} (\delta_{ij} (P_k - Q_k) + \delta_{ik} (P_j - Q_j) + \delta_{jk} (P_i - Q_i)) F_2(t) \right]$$

$$\begin{aligned}
G^{ijkl} = & \frac{1}{S_4^2} \left[ \frac{S_1^2 S_2^2}{S_4^2} (P_i - Q_i) (P_j - Q_j) (P_k - Q_k) (P_l - Q_l) F_4(t) \right. \\
& - \frac{1}{2} \frac{S_1 S_2}{S_4} \left( \delta_{ij} (P_k - Q_k) (P_l - Q_l) + \delta_{ik} (P_j - Q_j) (P_l - Q_l) \right. \\
& + \delta_{il} (P_j - Q_j) (P_k - Q_k) + \delta_{jk} (P_i - Q_i) (P_l - Q_l) \\
& + \delta_{jl} (P_i - Q_i) (P_k - Q_k) + \delta_{kl} (P_i - Q_i) (P_j - Q_j) \left. \right) F_3(t) \\
& \left. + \frac{1}{4} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) F_2(t) \right]
\end{aligned}$$

The electron repulsion integrals can now be defined in terms of  $G$ ,  $S_{ab}$  and a multiplier  $M$ , where

$$M = 2 \left( \frac{S_1 S_2}{S_4} \right)^{1/2}$$

$$\langle S_A S_b | S_c S_d \rangle = M S_{ab}^{oo} G^{oooo}$$

$$\langle p_{ia} S_b | S_c S_d \rangle = M [ S_{ab}^{io} S_{cd}^{oo} G^{oooo} + S_{ab}^{oo} S_{cd}^{oo} G^{iooo} ]$$

$$\begin{aligned}
\langle p_{ia} S_b | P_{kc} S_d \rangle = & M [ S_{ab}^{io} S_{cd}^{ko} G^{oooo} + S_{ab}^{io} S_{cd}^{oo} G^{ooko} \\
& + S_{ab}^{oo} S_{cd}^{ko} G^{iooo} + S_{ab}^{oo} S_{cd}^{oo} G^{ioko} ]
\end{aligned}$$

$$\begin{aligned}
\langle p_{ia} p_{jb} | S_c S_d \rangle = & M [ S_{ab}^{ij} S_{cd}^{oo} G^{oooo} + S_{ab}^{io} S_{cd}^{oo} G^{ojoo} \\
& + S_{ab}^{oj} S_{cd}^{oo} G^{iooo} + S_{ab}^{oo} S_{cd}^{oo} G^{ijoo} ]
\end{aligned}$$

$$\begin{aligned}
\langle p_{ia} p_{jb} | p_{kc} S_d \rangle &= M [S_{ab}^{ij} (S_{cd}^{ko} G^{oooo} + S_{cd}^{oo} G^{ooko}) \\
&+ S_{ab}^{io} (S_{cd}^{ko} G^{ojoo} + S_{cd}^{oo} G^{ojko}) + S_{ab}^{oj} G^{iooo} + S_{cd}^{oo} G^{ioko}) \\
&+ S_{ab}^{oo} (S_{cd}^{ko} G^{ijoo} + S_{cd}^{oo} G^{ijko})]
\end{aligned}$$

$$\begin{aligned}
\langle p_{ia} p_{jb} | p_{kc} p_{ld} \rangle &= M [S_{ab}^{ij} (S_{cd}^{kl} G^{oooo} + S_{cd}^{ko} G^{ool}) \\
&+ S_{cd}^{ol} G^{ook} + S_{cd}^{oo} G^{ookl}) + S_{ab}^{io} (S_{cd}^{kl} G^{ojoo} \\
&+ S_{cd}^{ko} G^{ojol} + S_{cd}^{ol} G^{ojko} + S_{cd}^{oo} G^{ojkl}) \\
&+ S_{ab}^{oj} (S_{ab}^{kl} G^{iooo} + S_{cd}^{ko} G^{iool} + S_{cd}^{ol} G^{ioko} + S_{cd}^{oo} G^{iokl}) \\
&+ S_{ab}^{oo} (S_{cd}^{kl} G^{ijoo} + S_{cd}^{ko} G^{ijol} + S_{cd}^{ol} G^{ijko} + S_{cd}^{oo} G^{ijkl})].
\end{aligned}$$

## Postgraduate courses attended and publications

### Courses:

1. Introduction to Fortran  
Edinburgh Regional  
Computing Centre
2.  $^{13}\text{C}$  and Fourier Transform  
NMR  
Dr. D. Shaw  
Varian Associates Ltd.
3.  $^1\text{H}$ -NMR  
Dr. R. Lynden-Bell  
University of East Anglia
4. Introduction to ESCA and  
PE spectroscopy  
Dr. D. Whan and Dr. S. Cradock  
University of Edinburgh
5. Chemical applications of group  
theory  
Dr. K. Lawley  
University of Edinburgh
6. Basic molecular orbital theory  
for chemists  
Dr. M. H. Palmer  
University of Edinburgh

**The Electronic Structure of Aromatic Molecules. Non-empirical Calculations on Indole, Benzofuran, Benzothiophen, and Related Hydrocarbons**

By **Michael H. Palmer** \* and **Sheila M. F. Kennedy**, Department of Chemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ

Reprinted from

JOURNAL  
OF  
THE CHEMICAL SOCIETY

---

PERKIN TRANSACTIONS II

---

1974



## The Electronic Structure of Aromatic Molecules. Non-empirical Calculations on Indole, Benzofuran, Benzothiophen, and Related Hydrocarbons

By Michael H. Palmer<sup>o</sup> and Sheila M. F. Kennedy, Department of Chemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ

*Ab initio* calculations in a gaussian orbital basis (LGO) are reported for naphthalene, indene, styrene, indole, benzofuran, and benzothiophen. A bond energy analysis on small non-aromatic molecules yields energies for the classical Kekulé structures; the resonance energies derived for the real molecules, in close agreement with figures based upon thermochemical data, are: naphthalene 357, indene 225, styrene 208, indole 308, benzofuran 232, benzothiophen 183, benzene 212 kJ mol<sup>-1</sup>. The calculated dipole moments are in reasonable agreement with experiment, and are separated into  $\sigma$ - and  $\pi$ -components. The high values of the  $\pi$ -moments, and the degree of scatter of the  $\pi$ -electrons into isolated pairs are related to aromatic character. Although the molecules are not isoelectronic, extensive correlation of the orbitals in the series can be achieved, and the He<sup>I</sup> photoelectron spectra are interpreted in terms of the main groups of orbital energies.

PREVIOUSLY we reported non-empirical MO calculations of the ground states of various five- and six-membered monocyclic, potentially aromatic, systems containing

<sup>1</sup> M. H. Palmer and A. J. Gaskell, *Theor. Chim. Acta*, 1971, **23**, 52.

<sup>2</sup> M. H. Palmer, A. J. Gaskell, and M. S. Barber, *Theor. Chim. Acta*, 1972, **23**, 357.

<sup>3</sup> S. Cradock, R. H. Findlay, and M. H. Palmer, *Tetrahedron*, 1973, **29**, 2173.

<sup>4</sup> M. H. Palmer, A. J. Gaskell, and R. H. Findlay, *Tetrahedron Letters*, 1973, 4659.

<sup>5</sup> M. H. Palmer and R. H. Findlay, *Tetrahedron Letters*, 1974, 253.

one or more of the heteroatoms nitrogen, oxygen, phosphorus, or sulphur.<sup>1-9</sup> The benzo-derivatives of the five-membered ring heterocycles are of particular interest owing to the existence of the Kekulé-like and quinonoid

<sup>6</sup> M. H. Palmer and R. H. Findlay, *Tetrahedron Letters*, 1972, 4165.

<sup>7</sup> M. H. Palmer, A. J. Gaskell, and M. S. Barber, *J. Mol. Struct.*, 1972, **12**, 197.

<sup>8</sup> M. H. Palmer, R. H. Findlay, and A. J. Gaskell, *J.C.S. Perkin II*, 1974, 420.

<sup>9</sup> M. H. Palmer, R. H. Findlay, and A. J. Gaskell, *J.C.S. Perkin II*, 1974, 778.

series; in this paper we report results on the ground states of the former series, and compare them with some related aromatic hydrocarbons.

#### METHODS

**Geometries.**—The naphthalene calculation used the experimental <sup>10</sup> geometry (1); indene (2) was constructed from one half of naphthalene and fragments of propene,<sup>11a</sup> while styrene (3) was obtained similarly from the benzene <sup>11b</sup> and ethylene <sup>11c</sup> geometries. The simplest derivative of the indole system for which a geometry has been reported is the 3-acetic acid (4); <sup>12</sup> the carbocyclic rings of (1) and (4) differ significantly and the pyrrole portion of (4) has a significantly shorter C(2)—C(3) bond than that of pyrrole. Whilst the fusion of five- and six-membered rings may well produce these effects, the possibility that the ring system of the derivative (4) might be distorted relative to the parent molecule as a result of the 3-acetic acid group polarity could not be excluded. Furthermore, we needed to construct geometries for benzo-furan and -thiophen for which there are no reported experimental structures. Various groups of workers <sup>13-17</sup> have calculated bond lengths for (1) and (5)—(7) based upon self-consistent bond order ( $p_{ij}$ )-bond length ( $r_{ij}$ ) relationships by the equation  $r_{ij} = A - Bp_{ij}$  where  $A$  and  $B$  are empirical parameters based upon ethylene, benzene *etc.*<sup>18</sup> These calculations do not lead to values for the bond angles, and hence are insufficient for the present work; for naphthalene most groups <sup>13a,14,16a</sup> calculate lengths that are in substantial agreement with the various experimental values, except that the  $C_{\alpha\beta}$  bond is always calculated to be slightly too long. None the less the same series of calculations suggest that the carbocyclic parts of (5)—(7) are probably very similar to those of naphthalene (1).<sup>\*</sup> Thus it seems reasonable to use the experimental geometric parameters from naphthalene for much of the geometry of (5)—(7); furthermore this has the added computational advantage that many electron repulsion integrals, which form the bulk of the computation time, could be used in several calculations. It is well established that INDO semiempirical calculations of total energy reproduce differences in geometric features fairly well. Thus as a final check we carried out two INDO calculations of the indole system using (a) the 3-acetic acid geometry (4) and (b) the geometry (5) based upon the superposition of pyrrole and one half of naphthalene. The total energies were (4) —71.2020 and (5) —71.2616 a.u. Overall geometry (5) appears slightly better than (4); thus we chose to use (5) and constructed the corresponding ones (6) and (7), by the same procedure from furan and thiophen and naphthalene.

**The Gaussian Basis.**—The linear combination of *scaled*

\* Errors in these calculations are usually  $\pm 0.02$  Å, but are often better for selected bonds within each molecule.

† The best total energy for thioformaldehyde was —435.6818 (*spd* basis with the optimum *3d* exponent of 0.540). The best *sp* basis result was —435.5946 a.u.

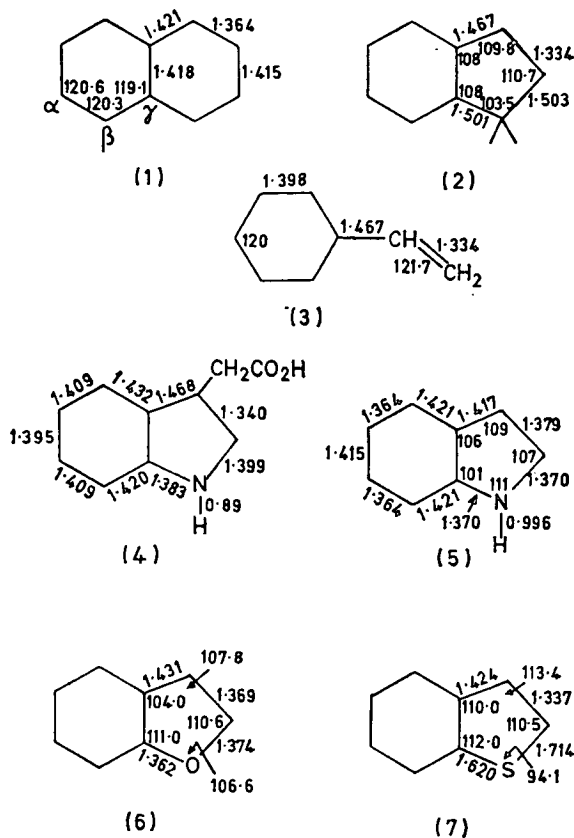
<sup>10</sup> A. Acmenningen, O. Bastiansen, and F. Dyvik, *Acta Cryst.*, 1961, **14**, 1056.

<sup>11</sup> (a) D. R. Lide and D. Christensen, *J. Chem. Phys.*, 1961, **35**, 1374; (b) G. E. Bacon, N. A. Curry, and S. A. Wilson, *Proc. Roy. Soc.*, 1964, **A**, 279, 98; (c) L. S. Bartell, E. A. Roth, C. D. Hollowell, K. Kuchitski, and J. E. Young, *J. Chem. Phys.*, 1964, **42**, 2683.

<sup>12</sup> I. A. Karle, K. Britts, and P. Gum, *Acta Cryst.*, 1964, 496.

<sup>13</sup> (a) F. Momicchioli and A. Rastelli, *J. Mol. Spectroscopy*, 1967, **22**, 310; (b) C. Aussems, S. Jaspers, G. Leroy, and F. Van Reemoortere, *Bull. Soc. chim. belges*, 1969, **78**, 479.

'best atom' gaussian functions (GTOs) consisted of five, two, three, and three GTOs for  $1s_X$ ,  $2s_X$ ,  $2p_X$  (where  $X = C, N,$  and  $O$ ) and  $1s_H$  respectively, while sulphur was represented by six, two, two, four, two, and one GTOs for  $1s$ ,  $2s$ ,  $3s$ ,  $2p$ ,  $3p$ , and  $3d$  respectively. The scaling procedures, based upon ethylene, vinyl alcohol, vinylamine <sup>2,6</sup> and thioformaldehyde,† are optimal in total energy for contraction



of this basis to the above orbitals. As before we define the heat of atomisation (binding energy) to be the difference between the total molecular energy using the *scaled* basis, and the atom sum using the *best atom* basis. The free atom energies (a.u.) using the unscaled basis, the Hartree-Fock limit,<sup>19</sup> and the percentage of the latter are: H (<sup>2</sup>S), —0.4970, —0.5000 (99.4); C (<sup>3</sup>P), —37.6105, —37.6886 (99.8); N (<sup>4</sup>S), —54.2754, —54.4009 (99.8); O (<sup>3</sup>P), —74.6121, —74.8094 (99.7); S (<sup>3</sup>P), —396.7476, —397.5048 (99.8). The computations were executed with the ATMOL-2 system of programs on an IBM370/195 computer, and the main results are shown in Tables 1—4.

<sup>14</sup> M. J. S. Dewar and N. Trinajstic, *J. Chem. Soc. (A)*, 1971, 1220; M. J. S. Dewar and G. J. Gleicher, *J. Chem. Phys.*, 1966, **44**, 759; M. J. S. Dewar, A. J. Harget, N. Trinajstic, and S. D. Worley, *Tetrahedron*, 1970, **26**, 4505.

<sup>15</sup> N. Trinajstic and A. Hinchcliffe, *Z. phys. Chem. Neue Folge*, 1968, **59**, 271; L. Klasinc, E. Pop, N. Trinajstic, and J. V. Knop, *Tetrahedron*, 1972, **28**, 3465.

<sup>16</sup> (a) J. Fabian, A. Mehlhorn, and R. Zahradnik, *J. Phys. Chem.*, 1968, **72**, 3975; (b) A. Skancke and P. N. Skancke, *Acta Chem. Scand.*, 1970, **24**, 23.

<sup>17</sup> (a) R. A. Sallavanti and D. D. Fitts, *Internat. J. Quant. Chem.*, 1969, **3**, 33; (b) G. Favini and A. Gamba, *J. Chim. Phys.*, 1967, **64**, 1443.

<sup>18</sup> C. A. Coulson and A. Golebiewski, *Proc. Phys. Soc.*, 1961, **78**, 1310.

<sup>19</sup> E. Clementi, 'Tables of Atomic Functions,' Supplement to IBM J. Research and Development, 1965, vol. 9, p. 2.

TABLE 1  
Molecular energies for the systems

	Naphthalene	Indene	Styrene	Indole
Total energy (a.u.) <sup>a</sup>	-382.37127	-344.58111	-306.79044	-360.52655
Binding energy (a.u.)	-2.29027	-2.11071	-1.93044	-1.88815
Resonance energy (calc) (kJ mol <sup>-1</sup> )	357	225	208	308
Resonance energy (thermochemical) (kJ mol <sup>-1</sup> )	314		176	226
	Benzofuran	Benzothiophen	Benzothiophen	Benzene
		( <i>sp</i> basis)	( <i>spd</i> + 3 <i>s'</i> basis)	
Total energy (a.u.)	-380.26093	-702.32203	-702.44606	-230.11426
Binding energy (a.u.)	-1.78283	-1.70843	-1.83246	212
Resonance energy (calc) (kJ mol <sup>-1</sup> )	232	242	283	155
Resonance energy (thermochemical) (kJ mol <sup>-1</sup> )				

<sup>a</sup> Energy conversion units (see 'Symbols, Signs and Abbreviations,' the Royal Society, London, 1969) are 1 a.u. = 2625.46 kJ mol<sup>-1</sup>, 1 a.u. = 27.211 eV, 1 eV = 1.6021 × 10<sup>-19</sup> J.

TABLE 2  
Molecular and derived bond energies (a.u.) from non-aromatic compounds

CH <sub>4</sub>	-40.10176	C <sub>2</sub> H <sub>4</sub>	-77.83154	(CH <sub>2</sub> =CH) <sub>2</sub>	-154.50920		
CH	-10.0254	C=C	-37.7296	C <sub>sp<sup>2</sup>-C<sub>sp<sup>2</sup></sub></sub>	-18.8972	H <sub>2</sub> S( <i>sp</i> )	-397.84420
CH <sub>2</sub> =CHOH	-152.46202	CH <sub>2</sub> =CHNH <sub>2</sub>	-132.54130	NH <sub>3</sub>	-56.01986	SH( <i>sp</i> )	-198.9221
C <sub>sp<sup>2</sup>-O</sub>	-46.7561	C <sub>sp<sup>2</sup>-N</sub>	-27.5378	NH	-18.6733	H <sub>2</sub> S( <i>spd</i> )	-397.94040
CH <sub>2</sub> =CHSH( <i>sp</i> )	-474.51823	CH <sub>2</sub> =CHSH( <i>spd</i> )	-474.61525	H <sub>2</sub> O	-75.79988	SH( <i>spd</i> )	-198.9702
C <sub>sp<sup>2</sup>-S</sub>	-207.7848	C <sub>sp<sup>2</sup>-S(<i>spd</i>)</sub>	-207.8391	OH	-37.8999		

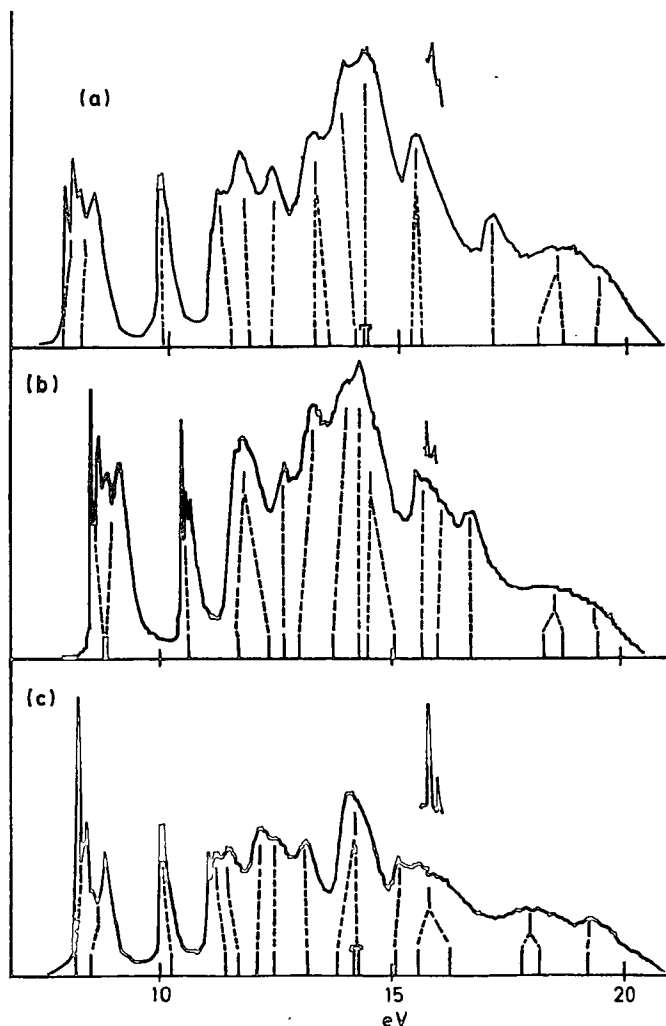


FIGURE 1 He I Photoelectron spectra of (a) indole, (b) benzofuran, and (c) benzothiophen

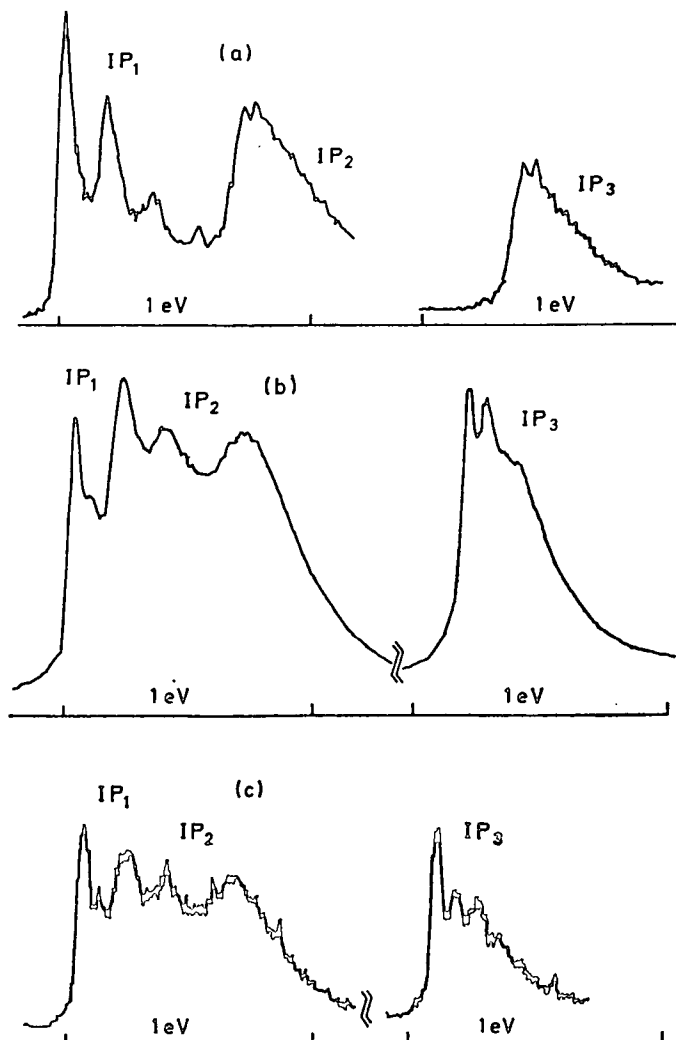


FIGURE 2 Expansion (×5) of He I photoelectron spectra of (a) naphthalene, (b) indole, and (c) benzofuran

## EXPERIMENTAL

The title compounds were purified by vacuum sublimation, except for benzofuran which was fractionally distilled. The photoelectron spectra using He<sup>I</sup> irradiation were ob-

acetone 1s<sub>0</sub> doublet; the valency shell spectra were generally poorly defined under these conditions, but that from benzofuran (Figure 3) allows identification of the principal 2s groupings.

TABLE 3

Molecular orbital energy levels (eV) of naphthalene							
$a_{1g}$	Character $a-c$	$b_{1g}$	Character $a-c$	$b_{2u}$	Character $a-c$	$a_{1u}$	Character $a,b$
-307.3	1s <sub>y</sub>	-306.9	1s <sub>β</sub>	-307.3	1s <sub>y</sub>	-8.67	2p <sub>a+β</sub>
-306.8	1s <sub>β</sub>	-306.8	1s <sub>α</sub>	-306.8	1s <sub>β</sub>		
-306.8	1s <sub>α</sub>	-26.94	2s <sub>a+β</sub>	-306.8	1s <sub>α</sub>		
-32.33	2s <sub>a+β+γ</sub>	-19.18	(2p + 1s <sub>H</sub> ) <sub>β</sub> (L)	-28.73	2s <sub>a+β+γ</sub>		
-27.56	2s <sub>β-γ</sub>	-16.30	(2p + 1s <sub>H</sub> ) <sub>α</sub> (T)	-22.88	2p (L + T)		
-22.31	(2p + 1s <sub>H</sub> ) <sub>α</sub> (T)	-13.44	(2p + 1s <sub>H</sub> ) <sub>β</sub> (L)	-17.12	(2p + 1s <sub>H</sub> ) <sub>α</sub> (T)		
-18.78	(2p + 1s <sub>H</sub> ) <sub>β</sub> (L)			-15.54	(2p + 1s <sub>H</sub> ) <sub>β</sub> + 2p <sub>α</sub> (L)		
-17.14	2p (L + T)						
-13.17	(2p + 1s <sub>H</sub> ) <sub>α</sub> + 2p <sub>β</sub> (T)						
		$b_{2u}$	Character $a-c$				
		-306.8	1s <sub>β</sub>				
		-306.8	1s <sub>α</sub>				
		-30.48	2s <sub>a+β</sub>				
		-23.27	2s <sub>a-β</sub> + 2p + (2p + 1s <sub>H</sub> ) <sub>β</sub>				
		-19.54	(2p + 1s <sub>H</sub> ) <sub>α+β</sub>				
		-16.24	2p <sub>a-β+γ</sub> (L)				
		-14.48	2p <sub>a-β</sub> (T)				
				$b_{1u}$	Character $a,c$		
				-15.48	2p <sub>a+β+γ</sub>		
				-9.47	2p <sub>β-γ</sub>		
				$b_{2g}$	Character $a,c$		
				-13.09	2p <sub>a+β</sub>		
$b_{2g}$	Character						
-11.29	2p <sub>a+β+γ</sub>						

<sup>a</sup> Allowable combinations of orbitals within the symmetry representation, where the positions are  $\alpha$ : 1, 4, 5, 8;  $\beta$ : 2, 3, 6, 7;  $\gamma$ : 4a, 8a. <sup>b</sup> CH and CC bonds of largely longitudinal character, *i.e.* parallel to the long axis of the molecule are labelled L; those of transverse character and parallel to the short in-plane symmetry axis are labelled T. <sup>c</sup> The molecule lies in the *yz* plane with the *z*-axis as the long axis. The  $D_{2h}$  symmetry nomenclature follows ref. 35.

TABLE 4

Molecular orbital energy levels (eV) for styrene, indene, indole, benzofuran, and benzothiophen

## Styrene

$a'$  ( $\sigma$ )  
-307.2, -306.8, -306.8, -306.8, -306.8, -306.8, -306.7, -306.2, -31.70, -29.31, -27.69, -26.91, -23.14, -22.39, -20.84,  
-19.44, -18.02, -17.20, -16.93, -16.26, -15.94, -14.50, -13.56, -13.33

$a''$  ( $\pi$ )  
-14.82, -11.93, -9.84, -8.88

## Indene

$a'$  ( $\sigma$ )  
-306.8, -306.8, -306.7, -306.7, -306.6, -306.6, -306.6, -306.5, -306.4, -32.25, -30.25, -27.88, -26.47, -25.72,  
-22.48, -21.45, -20.49, -18.70, -17.86, -17.07, -16.26, -16.05, -14.86, -13.93, -13.76, -12.87

$a''$  ( $\pi$ )  
-16.89, -14.12, -11.61, -9.61, -8.58

## Indole

$a'$  ( $\sigma$ )  
-402.9, -307.9, -307.6, -306.7, -306.5, -306.5, -306.5, -306.3, -305.8, -35.24, -30.90, -28.69, -27.26, -26.60,  
-23.01, -21.98, -21.42, -20.13, -18.22, -17.96, -16.60, -16.36, -15.67, -15.24, -14.06, -12.99

$a''$  ( $\pi$ )  
-16.73, -13.46, -11.08, -8.83, -8.20

## Benzofuran

$a'$  ( $\sigma$ )  
-561.4, -309.4, -308.9, -307.4, -307.3, -307.3, -307.2, -307.1, -307.0, -39.1, -31.76, -29.68, -27.89, -27.34,  
-23.35, -22.34, -19.78, -18.88, -18.41, -16.92, -16.68, -16.00, -15.06, -14.66, -13.30

$a''$  ( $\pi$ )  
-17.66, -14.21, -11.98, -9.67, -9.16

Benzothiophen (*spd* + 3s' basis)

$a'$  ( $\sigma$ )  
-2495, -307.7, -307.4, -307.2, -307.1, -307.0, -307.0, -306.9, -306.8, -238.1, -180.7, -180.6, -33.14, -30.72,  
-28.58, -27.10, -26.37, -23.03, -21.71, -21.22, -19.17, -18.30, -17.68, -16.64, -16.52, -15.28, -14.32, -13.87,  
-12.97

$a''$  ( $\pi$ )  
-180.55, -16.07, -13.32, -11.46, -9.21, -8.76

tained using a Perkin-Elmer PS16 spectrometer and were calibrated with the argon doublet at 15.75 and 15.93 eV (inset in Figure 1). Expansions ( $\times 5$ ) of the bands where appreciable fine structure is evident are shown in Figure 2. The band assignments in Figure 1 are based upon the calculated orbital energies using Koopmans' theorem. Condensed phase ( $-150^\circ$ ) studies of the electron energy levels were obtained using an A.E.I. ES200 (by courtesy of A.E.I. Scientific Apparatus Division) and were calibrated with the

## DISCUSSION

(a) *The Molecular Energies.*—The present calculation on naphthalene (Table 1) using 184 GTOs yields a total energy very similar to that of the 290 gaussian lobe function calculation by Buenker and Peyerimhoff,<sup>20</sup> but 54 a.u. better than the only other 'all electron' calculation,

<sup>20</sup> R. J. Buenker and S. D. Peyerimhoff, *Chem. Phys. Letters*, 1969, 3, 37.

using floating spherical gaussian orbitals.<sup>21,\*</sup> The present styrene calculation is *ca.* 3.0 a.u. better than an earlier GTO minimal basis calculation;<sup>22</sup> no *ab initio* calculations on the molecules (2) and (5)–(7) have been reported previously.

A study of the eigenvectors in the benzothiophen calculations with and without  $3d$  orbitals suggests that the role of the latter is that of polarisation functions rather than as strongly bonding orbitals, since in all cases the  $3d$  eigenvectors are very low ( $<0.1$ ). It is necessary to use six, rather than the usual five,  $3d$  functions in LCGO calculations; this allowed a clear differentiation of the roles of polarisation functions. The six functions

energy (and dipole moment):  $sp$  basis  $-702.32203$  (1.37),  $sp + 3s'$  basis  $-702.36890$  (1.37),  $spd$  basis  $-702.39913$  (0.56),  $spd + 3s'$  basis  $-702.44606$  a.u. (0.56 D); it is clear that a single  $3s'$  function leads to a much larger increment in total energy (0.0469 a.u.) than that for a  $3d$  function (0.0154 a.u.) (*cf.* refs. 6 and 8). Thus the angular properties of the  $3d$  functions are of little value except to lead to minor readjustments of the electron distribution, as evidenced by the improvement in the dipole moment (experimental 0.62 D).

The binding energy series (Table 1) is similar in the monocyclic and benzo-derivatives, benzene,<sup>24</sup> pyrrole,<sup>1,2</sup> thiophen ( $spd$ ),<sup>6</sup> and furan<sup>1,2</sup> having values 1.469, 1.060, 1.061, and 0.440 a.u. However, a more sensitive insight into the interaction of the rings and the degree of aromatic character is provided by the estimated resonance energies. The procedure is based upon partition of the *total* energy of selected molecules into bond contributions. In this way the correlation error, which leads to low bond energies on the usual thermochemical scale, is largely eliminated since the largest terms in the correlation energy are the intrabond pair correlation energies. The total energy of all the fragment molecules (Table 2) is calculated with an *identical* basis set; the contribution  $E_{C-H}$  is then  $\frac{1}{4}E_{CH_4}$ , and this figure inserted into the ethylene energy ( $E_{C_2H_4}$ ) yields  $E_{C=C}$ ; the problem of identifying the value of  $E_{C-O}$  between  $sp^2$ – $sp^2$  carbon atoms which is a fundamental problem with thermochemical studies is solved from inserting the above values  $E_{O-H}$  and  $E_{C-O}$ , into the total energy of twisted buta-1,3-diene where the  $\pi$ -orbitals are perpendicular and hence non-interacting. Similar procedures with ammonia, water, and hydrogen sulphide lead to  $E_{N-H}$ ,  $E_{O-H}$ , and  $E_{S-H}$ ; these data inserted into the computations on vinylamine, vinyl alcohol, and ethylenethiol where the protons are rotated out of plane to remove the  $\pi$ – $\pi$  lone pair–vinyl group interaction yield values of  $E_{C-X}$  where  $X = N, O, \text{ and } S$ . The bond and total energies of the fragment molecules are given in Table 2.

We now adopt the original definition<sup>25,26</sup> of resonance energy (RE), namely that it is the difference in total energy between the molecule (calculated or experimental) and that of the classical molecule with non-interacting olefinic and saturated systems, whose total energy is derived from the sum of the bond contributions.† The procedure yields a linear correlation with Pauling's<sup>28</sup> thermochemical mesomeric energies (ME) with the form

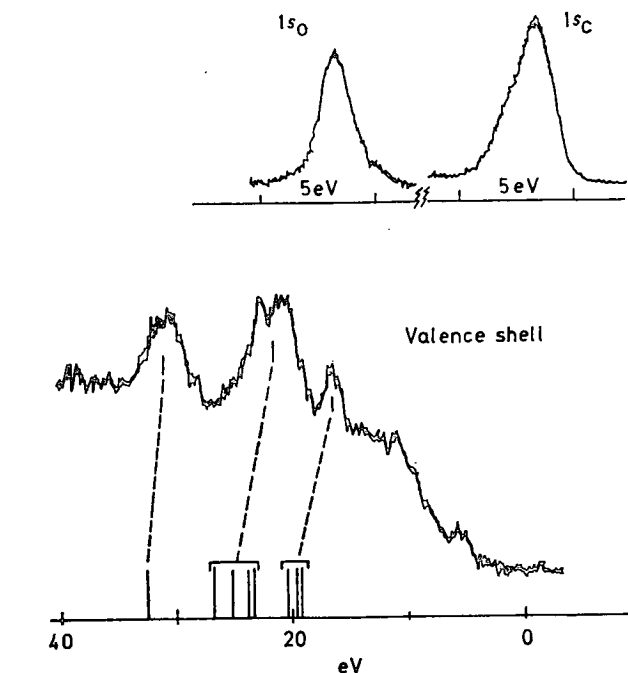


FIGURE 3 X-Ray photoelectron spectra of benzofuran

were contracted<sup>23</sup> to the usual group of five ( $3d_{x^2-y^2}$ ,  $3d_{xy}$ ,  $3d_{yz}$ ,  $3d_{zx}$ ,  $3d_{z^2}$ ), the  $spd$  basis, and a further  $s$ -type function ( $3s' \equiv 3d_{x^2+y^2+z^2}$ ). Selective use of these in the SCF procedure led to the following results for total

\* The orbital ordering (Table 3) for naphthalene is virtually identical in all three calculations, with only minor changes where near degenerate levels occur ( $5b_{2u}$  and  $4b_{1g}$ ,  $8a_{1g}$ , and  $6b_{2u}$ ). It is worth noting that for a wide range of aromatic and heterocyclic compounds the differences in orbital ordering between minimal and double zeta bases are insignificant.<sup>9</sup> This leads to confident assignments of the major groupings of the photoelectron spectra from these minimal basis calculations.

† The alternative definition introduced by Dewar<sup>27</sup> that the resonance energy becomes the difference in heat of atomisation between a given conjugated hydrocarbon and the value calculated for a corresponding classical (open chain) polyene by summing the appropriate bond energies, is very appealing in certain ways. It suffers from the disadvantage that (a) it assumes that polyolefinic systems have no resonance energy (*cf.* above), (b) no acyclic molecule can have any aromaticity, (c) the definition fails with cyclic systems where a saturated centre intervenes in the conjugation (as for example with cyclopentadiene). Clearly the latter could be used as an alternative to buta-1,3-diene in determining the energy of  $C_{sp^2}$ – $C_{sp^2}$ .

<sup>21</sup> R. E. Christoffersen, *J. Amer. Chem. Soc.*, 1971, **93**, 4104; R. E. Christoffersen, D. W. Genson, and G. M. Maggiora, *J. Chem. Phys.*, 1971, **54**, 239.

<sup>22</sup> W. J. Hehre, L. Radom, and J. A. Pople, *J. Amer. Chem. Soc.*, 1972, **94**, 1496.

<sup>23</sup> A. Rauk and I. G. Csizmadia, *Canad. J. Chem.*, 1968, **46**, 1205.

<sup>24</sup> M. H. Palmer and W. Moyes, to be published; M. H. Palmer and J. Nisbet, to be published.

<sup>25</sup> C. K. Ingold, 'Structure and Mechanism in Organic Chemistry,' Bell, London, 1969, p. 220.

<sup>26</sup> L. Pauling, 'The Nature of the Chemical Bond,' Cornell Univ. Press, New York, 1960, 3rd edn., pp. 190 *et seq.*

<sup>27</sup> M. J. S. Dewar, A. J. Harget, and N. Trinajstic, *J. Amer. Chem. Soc.*, 1969, **91**, 6321.

ME = 0.874RE - 4.47 kJ mol<sup>-1</sup>.<sup>\*</sup> The difference in energy between the planar and non-planar vinylamine (37), ethylenethiol (14), and vinyl alcohol (10 kJ mol<sup>-1</sup>) is sufficiently large for this not to be classified as a rotation barrier but rather as resonance energy. The data for the present molecules (Table 1) seem reasonable when compared with those for benzene (212), pyrrole (149), furan (89), and thiophen (no *d* orbitals 124, with *d* orbitals 144 kJ mol<sup>-1</sup>); thus naphthalene is rather less aromatic than twice benzene; styrene and indene are near to that for benzene, while the bicyclic heterocycles are substantially less than that of the monocyclic heterocycle plus benzene.† This last point is in accord with the even higher vinyl character in the C(2)-C(3) bond of the bi- than the mono-cyclic systems.

Many authors have proposed values for the resonance energies of simple benzenoid hydrocarbons and the heterocycles [e.g. (5)–(7)]; most of these values are either based upon thermochemical data<sup>25,26</sup> or calculations of an empirical or semiempirical type. Here we note that Dewar's definition of resonance energy,<sup>27</sup> when applied within the PPP  $\pi$ -electron framework<sup>28,29</sup> together with  $\alpha$ -bond energy estimates, or within a revised Hückel framework<sup>30–32</sup> where values for the  $C_{sp^2}$ - $C_{sp^2}$  single and double bond energies are included, leads to parallel results to the present work. Indeed these last two methods are related in a least squares fit ‡ by HMO = 0.485(PPP) + 12.35 kJ mol<sup>-1</sup> with standard deviations in slope, intercept, and overall of 0.032, 0.991, and 1.464, respectively. In the same way Dewar's data are related to Pauling's<sup>26</sup> by PPP = 0.494ME - 15.32 kJ mol<sup>-1</sup> with deviations (as above) of 0.062, 4.40, and 5.467 respectively.‡

(b) *The MO Energy Levels.*—The core  $1s_0$  levels are characterised by eigenvectors >0.98 (as are those of  $1s_N$ ,  $1s_O$  etc.) and are spread over a very small energy range. In all the bicyclic systems the bridging carbon atoms  $1s_0$  are at highest binding energy. For naphthalene the only case where experimental data are available, the X-ray photoelectron spectrum<sup>33</sup> order is again  $1s_{\alpha O} > 1s_{\alpha C} > 1s_{\beta O}$ ; since the LCGO method usually magnifies<sup>2,3,7,8,34</sup> the separation of  $1s_0$  levels, it seems like that the deconvolution technique of ref. 33 is leading to an artificially large separation. Even so the very small shift in  $1s_0$  levels in (1) and (3) is a substantial verification of the postulated lack of charge separation in alternant hydrocarbons.

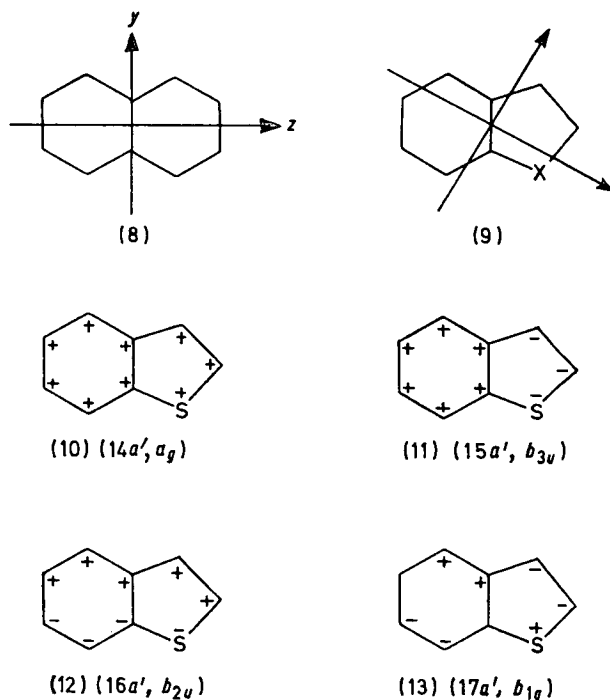
The inner valency shell orbitals (27–35 eV) in the hydrocarbons are linear combinations of  $2s_0$  orbitals, and all of (1), (3), and (5)–(7) show five such levels where one is largely N(5), O(6), or S(7) in the heterocycles. Above

<sup>\*</sup> Based upon data for benzene, naphthalene, anthracene, phenanthrene, cyclo-octatetraene, butadiene, ethylene, furan, pyrrole, thiophen, styrene, and indole; standard deviations in slope, intercept, and overall are 0.028, 1.823, and 3.876.

† An indication that these figures are not greatly effected by the basis set is afforded by similar results using the best atom basis<sup>1</sup> for the non-aromatic species (cf. Table 2) which lead to resonance energies for benzene (180), pyrrole (150), and furan (78 kJ mol<sup>-1</sup>) respectively.

‡ Based upon figures for benzene, naphthalene, anthracene, phenanthrene, furan, pyrrole, and indole.

this lies a further separate group of three largely  $2s_0$  levels (22–24 eV) which persist through the group (1), (5)–(7), but are more diffuse in (2), and are joined in (3) by a further  $2s$  level absent in the inner valency shell region. The outer valency shell region is largely combinations of  $2p_{C(N,O)}$  or  $3p_B$  mixed with  $1s_H$  and some loss of similarity in the spectra of (1)–(3) and (5)–(7) is apparent. However, a further classification of the valency shell  $\sigma$ -orbitals into those in which the constituent dominant atomic orbitals are either parallel to or transverse to the long molecular axis (8) is possible. We have previously noted that a similar classification into this longitudinal (L) and transverse (T) character can be done with many heterocycles.<sup>3</sup> In naphthalene,  $4b_{1g}$  for example consists of  $C_\alpha C_\beta + C_\alpha C_\gamma + C_\beta H_B$  bonding by orbitals parallel to the long ( $z$ ) axis (longitudinal polarisation), while  $6b_{3u}$  is largely  $C_\alpha C_\beta + C_\alpha C_\gamma$  bonding. Typical examples of orbitals of transverse (T) character in which the dominant  $2p_0$  orbitals are parallel to the transverse in-plane ( $y$ ) axis are  $6b_{2u}$  ( $C_\alpha H_\alpha$  bonding) and  $7b_{3u}$



( $C_\beta C_\beta$  bonding with rather less  $C_\alpha H_\alpha$  bonding). Comparatively few orbitals are of mixed L + T character (Table 3), and these are of entirely radial or tangential character with respect to the individual rings; examples

<sup>28</sup> M. J. S. Dewar and C. de Llano, *J. Amer. Chem. Soc.*, 1969, **91**, 789.

<sup>29</sup> M. J. S. Dewar, A. J. Harget, N. Trinajstic, and S. D. Worley, *Tetrahedron*, 1970, **26**, 4505.

<sup>30</sup> B. A. Hess and L. J. Schaad, *J. Amer. Chem. Soc.*, 1971, **93**, 305.

<sup>31</sup> B. A. Hess, L. J. Schaad, and C. W. Holyoke, *Tetrahedron*, 1972, **28**, 3657.

<sup>32</sup> B. A. Hess and L. J. Schaad, *J. Amer. Chem. Soc.*, 1973, **95**, 3907.

<sup>33</sup> D. T. Clark and D. Kilcast, *J. Chem. Soc. (B)*, 1971, 2243.

<sup>34</sup> U. Gelius, B. Roos, and P. Siegbahn, *Theor. Chim. Acta*, 1972, **27**, 171.

are  $5b_{3u}$  and  $8a_{1g}$  which are analogous to the radial and tangential orbitals  $b_{1u}$  and  $b_{2u}$  of benzene. The molecular

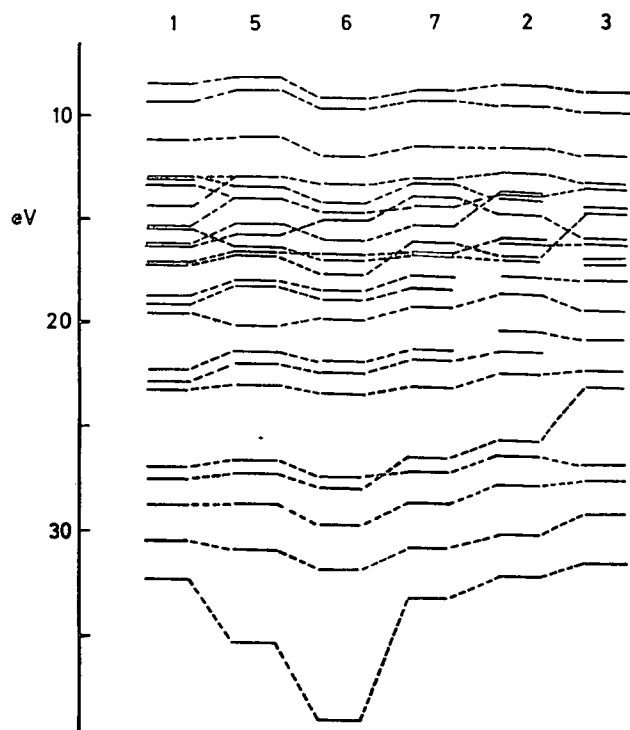


FIGURE 4 Correlation diagram for molecules (1)—(3) and (5)—(7)

orbitals of (2) and (5)—(7) show an apparent 'symmetry axis' as in (9), and are largely polarised longitudinally or

examples of these correlations are shown (10)—(13) for the  $2s_0 + 3s_8$  levels of benzothiophen with the corresponding naphthalene orbitals in parentheses. By the use of this property we are able to correlate many of the orbitals (Figure 4). Styrene again shows rather different properties, and is clearly closer to a combination of benzene and ethylene than to indene or the other compounds.

The comparatively small number of changes in orbital ordering in the series (1), (3), (5)—(7) is a reflection of the great similarity in electronic structure in the series. This is reflected in their electronic<sup>35,36</sup> and photoelectron spectra<sup>37-42</sup> also. Most of the cross-overs in Figure 4 can be attributed to changes in the free atom orbital energies, of which relevant ones in the Hartree-Fock limit are:  $2s$ : C, 19.2; N, 25.7; O, 33.9;  $2p$ : C, 11.8; N, 15.4; O, 17.2;  $3s$ : S, 23.9;  $3p$ : S, 11.9 eV. Thus the similarity of benzothiophen to naphthalene is also the most marked on this basis as well as in resonance energy (see also below).

The correlation of the  $\pi$ -orbitals in the series is the most marked. The simplified MO treatment of naphthalene in which it is treated as a ten-membered ring polygon,<sup>36</sup> and much used in the Platt method of analysis of electronic spectra<sup>35,36</sup> is not far from correct. The separation of the orbital pairs concerned ( $b_{1u}$ ,  $a_{1u}$ ), ( $b_{2g}$ ,  $b_{3g}$ ) is 0.9 and 1.8 eV respectively.

(c) *The Photoelectron Spectra. Assignments using the Orbital Energies.*—The  $\text{He I}$ <sup>37-43</sup> spectra of the compounds (1)—(6) have been recorded previously and this has been extended to  $\text{He II}$  for (1) and (3),<sup>43</sup> but interpretation has either been limited to postulated  $\pi$ -levels supported by Hückel,<sup>37,41</sup> SPINDO/1,<sup>43</sup> EHT, and PPP

TABLE 5  
Assignment of vertical IPs to orbital energies  $\epsilon_i$  (eV) for hydrocarbons (1)—(3)

Naphthalene									
IP	8.3	8.95	10.5	11.05	11.4	11.85	12.5	13.5	
$-\epsilon_i$	8.57	9.47	11.29	13.09	13.17	13.44	14.48	15.48/15.54	
IP	14.0		14.45	15.85	16.4		18.8	22.1	
$-\epsilon_i$	16.24/16.30		17.12/17.14	18.78	19.18/19.54		22.31/22.88/23.27		26.94/27.56/28.73
Indene									
IP	8.2	8.95	10.35	11.58	12.0		12.9		
$-\epsilon_i$	8.58	9.61	11.60	12.87	13.76/13.93/14.12			14.86	
IP	13.6		15.0		16.1		18.0		
$-\epsilon_i$	16.05/16.26		16.89/17.07		18.07		20.49/21.45/22.48		
Styrene									
IP	8.55	9.25	10.55	11.6	12.2	12.7	13.7		
$-\epsilon_i$	8.88	9.84	11.93	13.33	13.56	14.5/14.82		15.93/16.26	
IP	15	15.5	16.6	16.6	17.85	18.9	19.35	22.3	
$-\epsilon_i$	16.93/17.20		18.02	19.44	20.84	22.39	23.14	26.91/27.69	

transversely with respect to this axis. Of course the eigenvectors are not equal in magnitude across these 'symmetry axes' as they are in naphthalene. Typical

<sup>35</sup> H. H. Jaffe and M. Orchin, 'Theory and Applications of Ultraviolet Spectroscopy,' Wiley, New York, 1962, p. 294.

<sup>36</sup> J. R. Platt, *J. Chem. Phys.*, 1949, 17, 484.

<sup>37</sup> J. H. D. Eland and C. J. Danby, *Z. Naturforsch.*, 1968, 23a, 355.

<sup>38</sup> J. H. D. Eland, *Internat. J. Mass Spectroscopy Ion Phys.*, 1969, 2, 471.

<sup>39</sup> D. W. Turner, C. Baker, A. D. Baker, and C. R. Brundle, 'Molecular Photoelectron Spectroscopy,' Wiley-Interscience, London, 1970.

semiempirical calculations<sup>40</sup> or the effects of perfluorination.<sup>42</sup> The almost complete separation of  $\pi$ - and  $\sigma$ -levels observed in the present calculations (Table 3 and 4) offers a retrospective justification of the assignments based upon the Hückel method.

<sup>40</sup> P. A. Clark, R. Gleiter, and E. Heilbronner, *Tetrahedron*, 1973, 29, 3085.

<sup>41</sup> P. A. Clark, F. Brogli, and E. Heilbronner, *Helv. Chim. Acta*, 1972, 55, 1415.

<sup>42</sup> C. R. Brundle, M. B. Robin, N. A. Kuebler, and H. Basch, *J. Amer. Chem. Soc.*, 1972, 94, 1466.

<sup>43</sup> E. Lindholm, C. Fridh, and L. Asbrink, *Discuss. Faraday Soc.*, 1972, 54, 127.

TABLE 6  
Assignment of vertical IPs to orbital energies  $\epsilon_i$  (eV) for heterocycles (5)–(7)

Indole									
IP	7.79	8.18	9.88	11.12	115.6	12.24	13.26		
$-\epsilon_i$	8.20	8.83	11.08	12.99	13.46	14.06	15.24/15.67		
IP	13.77		14.28	15.40	17.03	18.51			
$-\epsilon_i$	16.36		16.60/16.73	17.96/18.22	10.23	21.98/23.01			
Benzofuran									
IP	8.66	8.94	10.58	11.83	12.67	13.31	14.01	14.31	
$-\epsilon_i$	9.16	9.67	11.98	13.30/14.21	14.66	15.06	16.00	16.68	
IP		14.54	15.69	16.09	16.69	185.4		19.39	
$-\epsilon_i$		16.92/17.66	18.41	18.88	19.78	21.82/22.34		23.47	
Benzothiophen									
Basis	IP	8.75	8.75	10.07	11.20	11.45	12.15	12.45	13.10
$spd + 3s'$	$-\epsilon_i$	8.76	9.21	11.45	12.97	13.32	13.87	14.32	15.28
$sp$	$-\epsilon_i$	9.00	9.44	11.67	12.84	13.51	13.96	14.44	15.36
	IP		14.25	15.20	15.80	17.98		19.27	
$spd + 3s'$	$-\epsilon_i$		16.07/16.52/16.64	17.68	18.30/19.16	21.22/21.71		23.03	
$sp$	$-\epsilon_i$		16.35/16.75/16.76	17.75	18.43/19.35	21.35/21.83		23.23	

In our previous work with monocyclic systems linear correlations between the calculated  $IP_{\text{calc}}$  and observed  $IP_{\text{obs}}$  of the form  $IP_{\text{obs}} = AIP_{\text{calc}} + B$  eV had  $A$  ca. 0.8,

TABLE 7

Vibrational frequencies excited in the photoelectron spectra ( $\text{cm}^{-1}$ ) (neutral ground state modes in parentheses)

	$IP_1$	$IP_2$	$IP_3$
Naphthalene	1400 (1477)		1265 (1393)
Indole	1350 (1454)	2620	390 (511) 645 (612)
Benzofuran	1480 (1254)		645 (766)
Benzothiophen	580 (540) 1310 (1497) 850 (884)		

whilst for a series of sulphur heterocycles, requiring rather larger bases than the present,  $A$  ca. 0.6 was obtained. Thus in the present work a value of  $A$  ca. 0.7 was

(1)–(6); thus comparatively few multiple assignments to a particular part of the band envelope are necessary, and these were done on the basis of band intensity and computed groupings and in the case of (1) and (3) follow the analysis of Lindholm *et al.*<sup>37</sup> The results are shown in Tables 5 and 6, and in Figure 1. The final correlation lines (with standard deviations in slope and intercept, and overall standard deviation) are: naphthalene,  $IP_{\text{obs}} = 0.755IP_{\text{calc}} + 1.702$  eV (0.014, 0.213, 0.190); indene,  $IP_{\text{obs}} = 0.773IP_{\text{calc}} + 1.479$  eV (0.020, 0.295, 0.237); styrene,  $IP_{\text{obs}} = 0.767IP_{\text{calc}} + 1.622$  eV (0.010, 0.184, 0.200); indole,  $IP_{\text{obs}} = 0.794IP_{\text{calc}} + 1.051$  eV (0.010, 0.164, 0.163); benzofuran,  $IP_{\text{obs}} = 0.779IP_{\text{calc}} + 1.300$  eV (0.010, 0.170, 0.152); benzothiophen ( $sp$  basis)  $IP_{\text{obs}} = 0.776IP_{\text{calc}} + 1.235$  eV (0.010, 0.153, 0.144), ( $spd + 3s'$  basis)  $IP_{\text{obs}} = 0.774IP_{\text{calc}} + 1.370$  eV (0.011, 0.166, 0.158). No special significance can be placed upon the small variations in slope and intercept.

TABLE 8  
Dipole moments ( $\mu/D$ ) and directions

	Total		Calculated				Observed	
	$\mu^a$	$\theta^b$	$\mu\sigma^c$	$\theta^b$	$\mu\pi^d$	$\theta^b$	$\mu$	$\theta^b$
Styrene	0.027	51					0.13, <sup>e</sup> 0.2, <sup>d</sup> 0.181, <sup>e</sup> 0.43, <sup>f</sup>	75.8
Indene	0.915	49					0.85, <sup>g</sup>	
Indole	2.31	48	−0.44	46	2.75	231	2.08, <sup>h</sup>	50
Benzofuran	−1.56	8	−3.00	196	1.46	25	0.79, <sup>i</sup>	
Benzothiophen ( $spd + 3s'$ basis)	−0.56	2	−2.48	247	1.99	60	0.62, <sup>j</sup>	

<sup>a</sup> The usual convention that the dipole moments of furan and pyrrole are negative and positive respectively, requires that the dipole moment is positive when the negative end lies away from the heteroatom and towards the C(3)–C(3a) bond; *i.e.* in the positive cartesian direction when the molecules are oriented as in (5)–(7). <sup>b</sup> Angles with respect to the C(7a)–C(3a) bond axis. The orientation is as in (5)–(7) and angles are positive anticlockwise. Styrene is positive anticlockwise with respect to the C(1)–C( $\alpha$ ) bond axis. <sup>c</sup> Ref. 49. <sup>d</sup> Ref. 53. <sup>e</sup> Ref. 52. <sup>f</sup> Ref. 50. <sup>g</sup> D. V. G. L. N. Rao, *Indian J. Phys.*, 1955, **29**, 398. <sup>h</sup> E. F. J. Janetzky and M. C. Lebet, *Rec. Trav. chim.*, 1944, **63**, 123. <sup>i</sup> R. D. Brown and B. W. Collier, *Theor. Chim. Acta*, 1967, **7**, 259. <sup>j</sup> R. G. Charles and H. Fieser, *J. Amer. Chem. Soc.*, 1950, **72**, 2233.

anticipated, since the scale factor  $A$  is dependent upon basis set size. A further feature assisting in the determination of the correlation line is the large number of IPs evident in the experimental spectrum in each of

Short progressions are evident in some of the ionisations from  $\pi$ -orbitals in (1) and (5)–(7) (Table 7); we have indicated gas-phase vibrations of the neutral molecule which are excited in the electronic (hot band) or



TABLE 9

Average positions <sup>a</sup> of  $\pi$ -electrons in indole, benzofuran, and benzothiophen

Benzofuran							Centre of <sup>a</sup> charge
	$\pi_1$	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$		
$\bar{y}^a$	-1.2867	-0.0745	0.2381	-0.0011	0.0544	-0.0487	
$\bar{z}^a$	1.6309	-1.9363	0.7317	-1.7504	0.0869	-0.499	
$\bar{r}^b$	2.3568	1.5207	1.2021	1.3394	0.5321		
Average $\bar{r}$ 1.3902							
Indole							Centre of charge
	$\pi_1$	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$		
$\bar{y}^a$	-0.9407	-0.1428	0.2455	-0.0145	0.1687	-0.0409	
$\bar{z}^a$	1.5062	-2.0017	-0.0136	-1.1962	0.2546	-0.4960	
$\bar{r}^b$	2.1612	1.5355	0.5409	0.7277	0.7550		
Average $\bar{r}$ 1.1441							
Benzothiophen ( <i>spd</i> + 3 <i>s'</i> basis)							Centre of charge
	$\pi_1^c$	$\pi_2$	$\pi_3$	$\pi_4$	$\pi_5$	$\pi_6$	
$\bar{y}^a$	-2.4818	-0.2769	-0.1485	-0.0022	0.0006	-0.3789	0.2668
$\bar{z}^a$	2.8419	0.5106	-1.0415	0.1393	-0.2684	-0.2024	-0.0697
$\bar{r}^b$	3.5092	0.4662	1.1061	0.3737	0.4794	0.2464	
Average $\bar{r}$ 0.5344							

<sup>a</sup> Relative to the midpoint of the C(3a)-C(7a) bond as origin and *y*-axis; distances in atomic units (1 a.u. = 0.529167 × 10<sup>-10</sup> m).  
<sup>b</sup> Relative to the centre of charge. <sup>c</sup> Localised 2*p<sub>π</sub>* on sulphur.

TABLE 10

Population analyses

Naphthalene	$\sigma$	C(1)	C(2)	C(4a)	H(1)	H(2)				
	$\pi$	5.1381	5.1493	5.0206	0.8511	0.8506				
Styrene		C(1)	C(2) or (6)	C(3) or (5)	C(4)					
	$\sigma$	5.0306	5.1353	5.1499	5.1473					
	$\pi$	1.0028	0.9996	0.9979	1.0012					
		C <sub>a</sub>	C <sub><math>\beta</math></sub>	H <sub>a</sub>	H <sub><math>\beta</math>(cis)</sub>	H <sub><math>\beta</math>(trans)</sub>	H(2) or (6)	H(3) or (5)	H(4)	
	$\sigma$	5.1462	5.2667	0.8541	0.8588	0.8632	0.8541	0.8517	0.8518	
	$\pi$	1.0026	0.9979							
Indene		C(1)	C(2)	C(3)	C(4)	C(5)	C(6)			
	$\sigma$	5.1766	5.1429	5.1393	5.1415	5.1435	5.1426			
	$\pi$	1.1422	0.9892	1.0224	1.0086	1.0089	0.0100			
	$\sigma$	C(7)	C(3a)	C(7a)	H(1)	H(2)	H(3)	H(4)	H(5)	H(6)
	$\pi$	5.0523	5.0198	5.0144	0.4409	0.8523	0.8500	0.8528	0.8528	0.8530
		1.1011	1.0038	0.9900	0.4119					0.8543
Indole		N	C(2)	C(3)	C(4)	C(5)	C(6)	C(7)	C(7a)	
	$\sigma$	5.7891	4.9730	5.0847	5.1465	5.1283	5.1421	5.1147	4.8486	
	$\pi$	1.6585	1.0494	1.1057	0.9949	1.0356	1.0087	1.0498	1.0409	
	$\sigma$	C(3a)	H(1)	H(2)	H(3)	H(4)	H(5)	H(6)	H(7)	
	$\pi$	4.9914	0.6656	0.8386	0.8576	0.8526	0.8564	0.8531	0.8579	
		1.0562								
Benzofuran		O	C(2)	C(3)	C(4)	C(5)	C(6)	C(7)		
	$\sigma$	6.7050	4.8486	5.1343	5.1422	5.1355	5.1503	5.1226		
	$\pi$	1.7930	1.0675	1.0331	0.9949	1.0196	0.9918	1.0308		
	$\sigma$	C(3a)	C(7a)	H(2)	H(3)	H(4)	H(5)	H(6)	H(7)	
	$\pi$	4.7397	4.9940	0.8124	0.8423	0.8349	0.8450	0.8433	0.8478	
		1.0551	1.0142							
Benzothiophen ( <i>spd</i> + 3 <i>s'</i> basis)		S	C(2)	C(3)	C(4)	C(5)	C(6)	C(7)		
	$\sigma$	12.1757	5.0973	5.1298	5.1447	5.1376	5.1483	5.1215		
	$\pi$	3.7839	1.0689	1.0287	0.9928	1.0175	0.9969	1.0214		
	$\sigma$	C(3a)	C(7a)	H(2)	H(3)	H(4)	H(5)	H(6)	H(7)	
	$\pi$	4.9845	4.9930	0.8301	0.8429	0.8481	0.8506	0.8479	0.8480	
		1.0431	1.0468							

fluorescence spectra<sup>44-46</sup> and which are plausibly assigned to these vibrations in the ion. Only in the case of naphthalene, where a normal co-ordinate analysis has been reported,<sup>47</sup> is there information about the nature of the ground state molecular vibrations. The frequency  $\nu_9$  ( $511\text{ cm}^{-1}$ ) in naphthalene has maximum amplitude for motion of  $C_\alpha C_\beta$  relative to  $C_\gamma$ , *i.e.*  $C_\alpha C_\gamma$  stretching; similarly  $\nu_3$  ( $1577\text{ cm}^{-1}$ ) corresponds to ( $C_\beta C_\beta + C_\gamma C_\gamma$ ) stretching + (CH) bending,  $\nu_2$  ( $1477\text{ cm}^{-1}$ ) to ( $C_\alpha C_\beta + C_\gamma C_\gamma$ ) stretching + (CH) bending,  $\nu_4$  ( $1393\text{ cm}^{-1}$ ) to ( $C_\alpha C_\gamma + C_\gamma C_\gamma$ ) stretching.<sup>48</sup> Ionisation from  $1a_{1u}$  ( $IP_1$ ) leads to a population decrease ( $-0.030e$ ) across  $C_\alpha C_\beta$  and an increase ( $+0.019e$ ) across  $C_\beta C_\beta$ , while for  $1b_{3g}$  ( $IP_3$ ) the corresponding figures are  $C_\alpha C_\beta$  ( $-0.018$ ),  $C_\beta C_\beta$  ( $+0.009$ ), and  $C_\gamma C_\gamma$  ( $+0.021e$ ). Thus in the ion  $\nu_3$  should be raised, while  $\nu_2$  and  $\nu_4$  should be lowered, and we assign the frequencies as in Table 7. Firm assignments for the heterocycles (5)–(7) must await more detailed information on the nature of the molecular vibrations.

(d) *The Molecular Dipole Moments.*—Comparison of the calculated and observed magnitudes (Table 8) shows good agreement for styrene, indene, indole, and benzothiophen, but the value for benzofuran is large by 1 D (assuming correct sign). The same basis set used here for oxygen also leads to a high value ( $-1.01\text{ D}$ ) for furan (expt.  $0.64$ ) whereas our earlier best atom basis gave  $-0.64\text{ D}$  for the latter.<sup>1</sup> Clearly optimisation in the scaling on a total energy basis reduced the optimisation in at least one molecular property. None the less, the present results and our earlier work<sup>9</sup> lead to unambiguous determination of dipole moment signs, and it seems likely that the higher dipole moment of benzofuran when compared with furan arises from a smaller level of  $\pi$ -back donation being counter balanced against the high  $\sigma$ -attraction; thus benzofuran is even less aromatic than furan (*cf.* resonance energies above). The data for benzothiophen and indole suggest that the ring systems are polarised in a similar fashion to the monocyclic species and thereby of somewhat higher aromatic character. The calculated angle that the dipole moment of indole makes with the molecular axes is in good agreement with that estimated from substituted indoles;<sup>49</sup> that for styrene differs significantly from that derived from one set of measurements,<sup>50</sup> but there appears to be some uncertainty in the value of this low moment.<sup>50-53</sup>

(e) *Average Positions of the  $\pi$ -Electrons.*—In our previous work<sup>8</sup> we introduced the concept that aromatic character as evidenced by a  $\pi$ -electron set, sextet, decet, *etc.* required that the electrons function as a group rather than as largely separate pairs. Thus their average

<sup>44</sup> D. P. Craig, J. M. Hollas, M. F. Redies, and S. C. Wait, *Phil. Trans. Roy. Soc. A*, 1961, **253**, 543, 569.

<sup>45</sup> J. M. Hollas, *J. Mol. Spectroscopy*, 1962, **9**, 138; *Spectrochimica Acta*, 1963, **19**, 753.

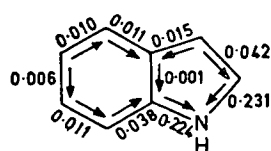
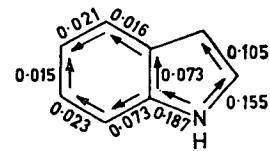
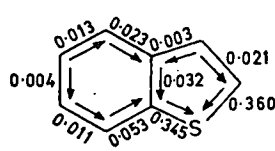
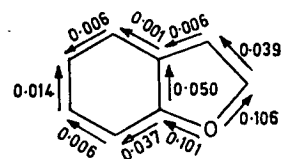
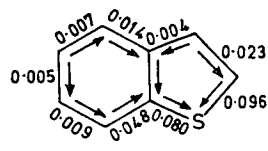
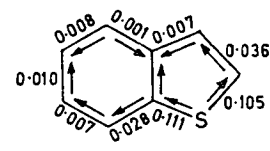
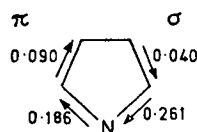
<sup>46</sup> E. P. Krainov, *Optical Spectroscopy*, 1964, **16**, 415.

<sup>47</sup> G. Hagen and S. J. Cyvin, *J. Phys. Chem.*, 1968, **72**, 1446; S. J. Cyvin, B. N. Cyvin, and G. Hagen, *Chem. Phys. Letters*, 1968, **2**, 341.

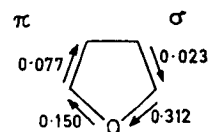
<sup>48</sup> G. S. Pawley, personal communication, *cf.* G. S. Pawley and S. J. Cyvin, *J. Chem. Phys.*, 1970, **52**, 4073.

<sup>49</sup> H. Weiler-Feilchenfeld, A. Pullman, H. Berthod, and C. Geissner-Prettre, *J. Mol. Structure*, 1970, **6**, 297.

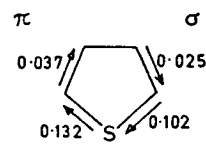
positions ( $\bar{x}$ ,  $\bar{y}$ ,  $\bar{z}$ ) from some fixed position should be (ideally) constant within the group. In some highly symmetrical molecules such as benzene this is automatically the case. In the case of the five-membered ring heterocycles  $C_4H_4X$  ( $X = S, NH, O, \text{ or } PH$ ) there was a progressive trend towards a quartet and a doublet, *i.e.* diene character. The average positions relative to the midpoint of the C(3a)–C(7a) bond axis for the molecules

(14) ( $\sigma$ )(15) ( $\pi$ )(16) ( $\sigma$ )(17) ( $\pi$ )(18) ( $\sigma$ ) ( $spd + 3s'$  basis)(19) ( $\pi$ ) ( $spd + 3s'$  basis)

(20)



(21)

(22) ( $spd + 3s'$  basis)

are reported in Table 9. These clearly show the increasing average distance ( $\bar{r}$ ) in the sequence naphthalene (0.0), benzothiophen, indole, and benzofuran, which together with the resonance energies (above) follow an acceptable sequence. The distances from the centre of nuclear change, which is close to the mid-point of C(3a)–C(7a) in

<sup>50</sup> R. A. Y. Jones, A. R. Katritzky, and A. V. Ochkin, *J. Chem. Soc. (B)*, 1971, 1795.

<sup>51</sup> A. J. Petro and C. P. Smyth, *J. Amer. Chem. Soc.*, 1957, **79**, 6142; 1958, **80**, 73.

<sup>52</sup> J. E. Plamondon, R. J. Buenker, D. J. Koopman, and R. J. Dolter, *Proc. Iowa Acad. Sci.*, 1963, **70**, 163 (*Chem. Abs.*, 1964, **61**, 11,424a).

<sup>53</sup> Z. Yu. Kokoshko, V. G. Kiteava, Z. V. Pushkareva, and V. E. Blokhin, *Zhur. obshchei Zhim.*, 1967, **37**, 58.

all cases, enables a connection between the  $\pi$ -dipole moment and these average  $\pi$ -electron positions to be established. In benzothiophen all five pairs of valency  $\pi$ -electrons are scattered near the centre of charge. In benzofuran on the other hand, the electrons separate into a quartet near the centre of the carbocyclic ring ( $\pi_2$  and  $\pi_4$ ), one pair at the centre of charge ( $\pi_5$ ), and pairs relatively close to the ring centre ( $\pi_3$ ) and the oxygen atom ( $\pi_1$ ).

(f) *Population Analyses*.—The hydrocarbons (1) and (3) show almost equal populations at the various centres, as required in the original concept of alternant hydrocarbons. In contrast the carbon  $\sigma$ -populations vary depending upon whether or not the atom is attached to hydrogen.

As expected the heterocycles (5)—(7) show rather larger population variations. In all cases (5)—(7) the heteroatom is a  $\pi$ -donor and  $\sigma$ -acceptor; the  $\sigma$ -acceptance is largely restricted to the immediately adjacent atoms C(2) and -(7a). Whilst the 3-position in (5)—(7) has a high total population, this arises more from  $\sigma$ - than  $\pi$ -polarisation, showing that the earlier  $\pi$ -electron theories produce the overall correct effect, but by the wrong mechanism.

As in our earlier work with monocyclic systems,<sup>8</sup> we have analysed the net atomic populations in terms of bond population moment contributions; full details of this method are given in ref. 8; the C-H/N-H bond contributions, immediately obvious from Table 10, are omitted as are those figures for the hydrocarbons. The carbocyclic rings (14)—(19) contain only very small bond population

moments, showing that the major effects of the heteroatom are restricted to the heterocyclic ring. However, there is a small series of shifts *from* ( $\sigma$ ) and *to* ( $\pi$ ) the 5-position in all cases; thus the heteroatom is exerting a long range effect on a donor-acceptor basis. Comparison of the bond moments with corresponding bonds in pyrrole (20), furan (21), and thiophen (22), shows that whilst the magnitudes ( $\sigma$  and  $\pi$ ) are similar in the sulphur and nitrogen compounds, they are distinctly different in benzofuran to furan. Thus the oxygen atom in (6) is a stronger  $\sigma$ -acceptor and poorer  $\pi$ -donor than in furan. These results are consistent with the varying levels of aromatic character estimated from the resonance energies above.

(g) *Conclusions*.—The *ab initio* studies reported for molecules (1)—(3) and (5)—(7) lead to satisfactory calibration of the main groups of photoelectron IPs and account for the dipole moments. The sequence of aromatic character naphthalene > benzothiophen = indole > benzofuran  $\simeq$  indene > styrene is obtained from both the resonance energies and an analysis of the average position of the  $\pi$ -electrons. The  $3d$  orbitals on sulphur in benzothiophen act primarily as polarisation orbitals and as such lead to modifications of the bonding, and as in chlorine trifluoride 'they improve but do not alter in character, the occupied MOs.'<sup>54</sup>

We are grateful to the S.R.C. and to the Director of the Atlas Laboratory for the provisions of computational facilities and for a grant to S. M. F. K.

[4/1204 Received, 20th June, 1974]

<sup>54</sup> A. Breeze, D. W. J. Cruickshank, and D. R. Armstrong, *J.C.S. Faraday II*, 1972, 1089.

**The Electronic Structure of the Quinonoid Bicyclic Heterocycles Isoindole, Benzo[c]furan, Benzo[c]thiophen, and 2*H*-Indene**

By Michael H. Palmer \* and Sheila M. F. Kennedy, Department of Chemistry, University of Edinburgh, Edinburgh EH9 3JJ

Reprinted from

JOURNAL  
OF  
THE CHEMICAL SOCIETY

---

PERKIN TRANSACTIONS II

---

1976

## The Electronic Structure of the Quinonoid Bicyclic Heterocycles Isoindole, Benzo[c]furan, Benzo[c]thiophen, and 2H-Indene

By Michael H. Palmer\* and Sheila M. F. Kennedy, Department of Chemistry, University of Edinburgh, Edinburgh EH9 3JJ

Non-empirical calculations for the electronic ground and lowest triplet excited states for isoindole, benzo[c]furan, benzo[c]thiophen, and *N*-methylisoindole show that some resonance energy (RE) is present in these systems, and that their instability can be attributed to a combination of a low RE and a low-lying excited state. The photoelectron spectra are reported for the last three compounds, and the ionisation potentials are assigned in the light of the calculations. Detailed analysis of the <sup>1</sup>H n.m.r. spectra of these compounds supports the conclusions of low aromatic character.

PREVIOUSLY we reported non-empirical MO calculations of the ground states of various five- and six-membered ring, potentially aromatic, systems containing one or more of the heteroatoms, nitrogen, oxygen, phosphorus, or sulphur.<sup>1-10</sup> The benzo-derivatives of the five-membered ring heterocycles are of particular interest owing

<sup>1</sup> M. H. Palmer and A. J. Gaskell, *Theor. Chim. Acta*, 1971, **28**, 52.

<sup>2</sup> M. H. Palmer, A. J. Gaskell, and M. S. Barber, *Theor. Chim. Acta*, 1972, **28**, 357.

<sup>3</sup> S. Cradock, R. H. Findlay, and M. H. Palmer, *Tetrahedron*, 1973, **29**, 2173.

<sup>4</sup> M. H. Palmer, A. J. Gaskell, and R. H. Findlay, *Tetrahedron Letters*, 1973, 4659.

<sup>5</sup> M. H. Palmer and R. H. Findlay, *Tetrahedron Letters*, 1974, 253.

to the existence of the Kekulé like (1) and quinonoid (2) series.

Recently we reported<sup>11</sup> calculations of the electronic structure of the ground states of the Kekulé series (1) and

<sup>6</sup> M. H. Palmer, A. J. Gaskell, and M. S. Barber, *J. Mol. Structure*, 1972, **12**, 197.

<sup>7</sup> M. H. Palmer and R. H. Findlay, *Tetrahedron Letters*, 1972, 4165.

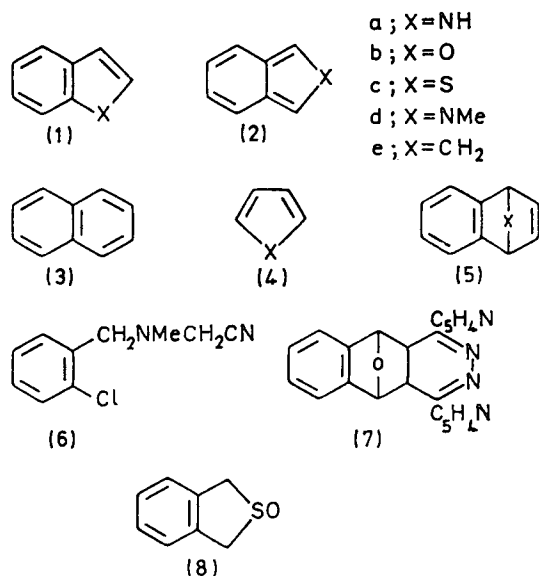
<sup>8</sup> M. H. Palmer, A. J. Gaskell, and R. H. Findlay, *J.C.S. Perkin II*, 1974, 420.

<sup>9</sup> M. H. Palmer, R. H. Findlay, and A. J. Gaskell, *J.C.S. Perkin II*, 1974, 778.

<sup>10</sup> M. H. Palmer and R. H. Findlay, *J.C.S. Perkin II*, 1974, 1885.

<sup>11</sup> M. H. Palmer and S. M. F. Kennedy, *J.C.S. Perkin II*, 1974, 1893.

now extend these results to the corresponding quinonoid compounds (2a—d) and compare them with both naphthalene (3) and the monocyclic systems (4a—c). Since



the relative instability of the quinonoid series could be attributed to either (a) a raised ground state energy (low resonance energy) or (b) a low lying excited state, we include calculations of the ground and first triplet excited state. We assign the photoelectron spectra of (2b—d) in the light of the ground state calculations. Finally we report a detailed analysis of the <sup>1</sup>H n.m.r. spectra of (2b—d) since relative values of the coupling constants have been used in discussions of aromatic character in related compounds (see below).

#### METHODS

**Preparations.**—We found the most satisfactory method for synthesis of isoindole (2a) was *via* (4; X = NCO<sub>2</sub>C<sub>4</sub>H<sub>9</sub>) and (5; X = NCO<sub>2</sub>C<sub>4</sub>H<sub>9</sub>),<sup>12</sup> but were unable to volatilise the material into the photoelectron spectrometer without extensive decomposition. *N*-Methylisoindole (2d) was prepared by a similar route, but conversion to (5; X = NMe) proceeded in very poor yield. The most successful route was from 2-chlorobenzyl chloride and methylaminoacetonitrile,<sup>13</sup> the intermediate (6) requiring purification from a range of by-products by spinning band column fractionation. Attempts at a reduction of *N*-methylphthalimide<sup>14</sup> always yielded a mixture of (2d) and the corresponding isoindoline. This same problem arose with attempts to dehydrate 2-ethylisoindoline *N*-oxide to 2-ethylisoindole.<sup>15</sup>

Benzo[*c*]furan was prepared from 1,4-epoxy-1,4-dihydronaphthalene (5; X = O) by two routes: (a) by interaction with 3,6-di-(2-pyridyl)-1,2,4,5-tetrazine<sup>16</sup> and pyrolysis<sup>17</sup> of the adduct (7), but which yielded (2b) contaminated with

3,6-dipyridylpyridazine and (b) by reduction of the epoxide (5) to the corresponding 1,2,3,4-tetrahydronaphthalene (Pd-C; MeOH; 1 atm.) followed by thermolysis<sup>18</sup> through a silica tube (650°; 0.05 mmHg) with a liquid nitrogen trap.

Benzo[*c*]thiophen (2c) was obtained by thermolysis of the sulphoxide (8) mixed with alumina (150°; 5 mmHg).<sup>19</sup>

**Photoelectron Spectra.**—A Perkin-Elmer PS16 spectrometer, calibrated with the argon doublet at 15.75 and 15.93 eV (insert in Figure 1), was used.<sup>11</sup> In general only the first

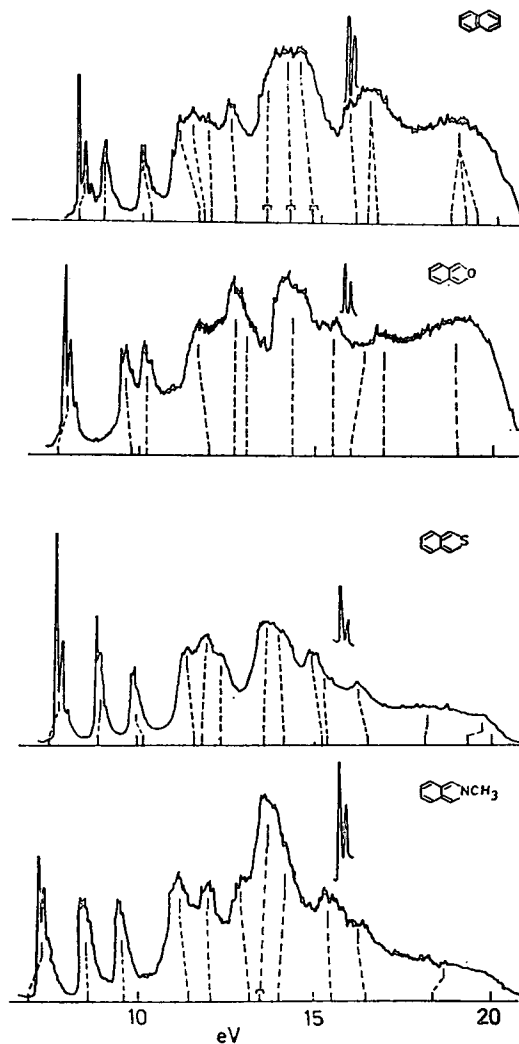


FIGURE 1 Photoelectron spectra (He<sup>I</sup>) of naphthalene and the quinonoid heterocycles (2b—d)

band in each compound showed appreciable fine structure, and this aspect was not pursued. The band assignments (Figure 1) are based upon the calculated orbital energies using Koopmans' theorem.<sup>11</sup> As in our work on the Kekulé

<sup>15</sup> R. Kreher and J. Scubert, *Angew. Chem. Internat. Edn.*, 1964, **3**, 639; 1966, **5**, 967; J. K. Kochi and E. A. Singleton, *Tetrahedron*, 1968, **24**, 4649.

<sup>16</sup> J. F. Geldard and F. Lions, *J. Org. Chem.*, 1965, **30**, 318.

<sup>17</sup> R. N. Warrener, *J. Amer. Chem. Soc.*, 1971, **93**, 2346.

<sup>18</sup> U. E. Wiersum and W. J. Mijs, *J.C.S. Chem. Comm.*, 1972, 347.

<sup>19</sup> M. P. Cava, N. M. Pollock, O. A. Mamer, and M. J. Mitchell, *J. Org. Chem.*, 1971, **36**, 3932.

<sup>13</sup> L. A. Carpino and D. E. Barr, *J. Org. Chem.*, 1966, **31**, 764.

<sup>14</sup> B. Jaques and R. G. Wallace, *J.C.S. Chem. Comm.*, 1972, 397.

<sup>15</sup> D. L. Garmaise and A. Ryan, *J. Heterocyclic Chem.*, 1970, **7**, 413; G. Wittig, H. Tenhaeff, W. Schoch, and G. Koenig, *Annalen*, 1951, **572**, 1; G. Wittig, G. Closs, and F. Mindermann, *ibid.*, 1955, **582**, 89.

series<sup>11</sup> we endeavoured to obtain X-ray photoelectron spectra of the valency shell orbitals, but the quality of the spectra from the solid samples is low, and only in the case of benzo[*c*]thiophen were well resolved peaks obtained at 1 221 and 1 229 eV kinetic energy corresponding to 3s<sub>g</sub> and 2s<sub>o</sub> groupings respectively (relative to Mg-K<sub>α</sub> 1 253 eV).

*N.m.r. Spectra and Analysis.*—<sup>1</sup>H and <sup>13</sup>C n.m.r. spectra were measured using Varian HA 100 and XL-100 spectrometers. In the former case the spectra of (2b–d) were run at various temperatures in CDCl<sub>3</sub> solution (at 273, 273, and 253 K respectively). The samples were run under nitrogen

respectively, while sulphur was represented by six, two, two, four, two, and one GTOs for 1s, 2s, 3s, 2p, 3p, and 3d respectively. The free atom energies (a.u.) using the unscaled 'best atom' basis, the Hartree–Fock limit,<sup>20</sup> and the percentage of the latter are: H(<sup>2</sup>S), –0.4970, –0.5000 (99.4); C(<sup>3</sup>P), –37.6106, –37.6886 (99.8); N(<sup>4</sup>S), –56.2754, –57.4009 (99.8); O(<sup>3</sup>P), –74.6121, –74.8094 (99.7); S(<sup>3</sup>P), –396.7476, –397.5048 (99.8). The computations were executed with the ATMOL-2 system of programs on an IBM 370/195 computer,\* and the main results are shown in Tables 1–5.\*

TABLE 1  
Molecular energies for the systems

	Naphthalene	2H-Indene	Isoindole	Benzo[ <i>c</i> ]furan	Benzo[ <i>c</i> ]thiophen ( <i>sp</i> Basis)	Benzo[ <i>c</i> ]thiophen ( <i>spd</i> + 3s <sup>1</sup> )
Total energy (a.u.) <sup>a</sup>	–382.371 27	–344.522 94	–360.505 12	–380.1234	–702.312 40	–702.430 85
Binding energy (a.u.)	–2.290 27	–2.052 54	–1.866 72	–1.684 24	–1.697 80	
Resonance energy (calc.) (kJ mol <sup>–1</sup> )	357	72	252	147	217	259

<sup>a</sup> Energy conversion units (see 'Symbols, Signs, and Abbreviations,' The Royal Society, London, 1969) are 1 a.u. = 2 625.46 kJ mol<sup>–1</sup>, 1 a.u. = 27.211 eV, 1 eV = 1.6021 × 10<sup>–19</sup> J.

TABLE 2

Low lying triplet state excitation energies (eV)

	Naphthalene	Indole	Benzo[ <i>b</i> ]furan	Benzo[ <i>b</i> ]thiophen
UHF	3.028	3.206	2.640	3.431
RHF	4.420	4.146	3.647	4.350
Expt.	2.665 <sup>a</sup>	30.6, <sup>b</sup> 3.069 <sup>c</sup>		2.972, <sup>d</sup> 2.965 <sup>e</sup>
INDO		6.549	5.897	4.134
	2H-Indene	Isoindole	Benzo[ <i>c</i> ]furan	Benzo[ <i>c</i> ]thiophen
UHF	0.884	2.400	0.286	1.972
RHF	1.162	2.694	0.593	2.425
INDO		4.062	3.521	3.727

<sup>a</sup> H. B. Klevens and J. R. Platt, *J. Chem. Phys.*, 1949, **17**, 470. <sup>b</sup> G. S. Kembrowskii, V. P. Bobrovich, and S. V. Koneve, *Zhur. Priklad. Spektroskopii*, 1966, **5**, 695. <sup>c</sup> R. C. Heckman, *J. Mol. Spectroscopy*, 1958, **2**, 27; P.-S. Song and W. E. Kurtin, *J. Amer. Chem. Soc.*, 1969, **91**, 4892. <sup>d</sup> D. F. Evans, *J. Chem. Soc.*, 1959, 2753. <sup>e</sup> M. R. Padhye and J. C. Patel, *J. Sci. Ind. Res., India*, 1956, **15B**, 171, 206.

or aqueous sodium dithionite [in the case of (2c)] to avoid air oxidation. In the <sup>13</sup>C spectra the assignments were confirmed by off-resonance decoupling, and are quoted in p.p.m. downfield from tetramethylsilane.

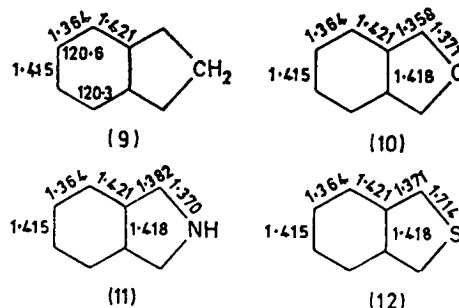
Computer analysis of the <sup>1</sup>H n.m.r. AA<sup>1</sup>BB<sup>1</sup>XX<sup>1</sup> spectra was carried out by means of the program LAMS (C. W. Haigh, Swansea). The analysis took place in two stages. (1) An AA<sup>1</sup>BB<sup>1</sup> analysis was completed to yield the lowest root mean square deviation in transition frequencies; a series of non-iterative studies of the AA<sup>1</sup>BB<sup>1</sup>XX<sup>1</sup> system were made with selected values of the additional coupling constants. (2) A full six-spin analysis was performed on the most promising set of parameters. In the case of (2b), the 1-H–3-H coupling was removed by double irradiation in stage (1), but this could not be done for (2c and d) owing to the proximity of these to the 4–7-H multiplets.

*Computations.*—We used the same basis of scaled 'best atom' gaussian functions (GTOs) as in our recent work.<sup>2,5,7,10,11</sup> These consisted of five, two, three, and three GTOs for 1s<sub>x</sub>, 2s<sub>x</sub>, 2p<sub>x</sub> (where X = C, N, or O), and 1s<sub>H</sub> res-

\* Additional data are contained in Supplementary Publication No. SUP 21618 (10 pp.). For details of Supplementary Publications see Notice to Authors No. 7 in *J.C.S. Perkin II*, 1975, Index issue.

† Various workers<sup>21,22</sup> have calculated bond lengths for (2a–c) based upon self-consistent bond order–bond length relationships; these calculations do not lead to values for the bond angles and are thus insufficient for the present work.

*Geometric Parameters.*—There is no experimental information for (2a–e).† We thus constructed geometries based upon the fusion of one half of naphthalene with the corresponding heterocycle, the actual values being shown in (9)–(12). This procedure had the additional computational



advantage that many electron repulsion integrals (relating to the naphthalene portion) could be used in several calculations. The orientation of the N-methyl group in (2d) was

<sup>20</sup> E. Clementi 'Tables of Atomic Functions,' in *IBM J. Res. Development*, 1965, **9**, 2.

<sup>21</sup> (a) M. J. S. Dewar, A. J. Harget, N. Trinajstic, and S. D. Worley, *Tetrahedron*, 1970, **26**, 4505; (b) M. J. S. Dewar and N. Trinajstic, *J. Amer. Chem. Soc.*, 1970, **92**, 1453.

<sup>22</sup> J. Fabian, A. Mehlhorn, and R. Zahradnik, *J. Phys. Chem.*, 1968, **72**, 3975.

chosen to have a  $\sigma$ - $\pi$  separation ( $a'/a''$  respectively in  $C_2$  symmetry).

TABLE 3

Molecular orbital energy levels (eV) for 2*H*-indene, isoindole, benzo[*c*]furan, benzo[*c*]thiophen, and *N*-methylisoindole

Energy	Character	Energy	Character
Isoindole			
$a_1$		306.4	1 $s_{4-7}$
425.7	1 $s_N$	306.4	1 $s_{3a-7a}$
307.1	1 $s_{1+3}$	306.2	1 $s_{5-6}$
306.4	1 $s_{4+7}$	27.82	2 $s_C(b_{2u})$
306.4	1 $s_{3a+7a}$	26.04	2 $s_C(b_{1g})$
306.2	1 $s_{5+6}$	21.59	CC + CN( $b_{2u}$ )
35.49	2 $s_N + 2s_C(a_{1g})$	17.87	CC + CH( $b_{1g}$ )
30.98	2 $s_C - 2s_N(b_{2u})$	16.24	CC + CH(T, $b_{1g}$ )
28.12	2 $s_C(a_{1g})$	15.95	CC + CH(T, $b_{1g}$ )
23.68	2 $s_C + 2p_C(L, b_{2u})$	13.98	CC(L, $b_{1u}$ )
21.19	CC $_{3a,7a} + CH(T, a_{1g})$	$b_1$	
20.00	CH $_{1,3} + NH(b_{2u})$	16.68	CC + CN( $b_{1u}$ )
18.44	CC + CH(L, $a_{1g}$ )	13.55	CN( $b_{2g}$ )
16.41	CH + CN(L + T, $b_{2u}$ )	9.53	CC + N( $b_{1u}$ )
15.34	CC(L + T, $b_{2u}$ )	$a_2$	
$b_2$		10.64	CC( $b_{2g}$ )
307.1	1 $s_{1-3}$	7.23	CC( $a_{1u}$ )
Benzo[ <i>c</i> ]furan			
$a_1$		307.4	1 $s_{3a-7a}$
562.7	1 $a_2$	307.2	1 $s_{4-7}$
308.7	1 $s_{1+3}$	307.0	1 $s_{5-6}$
307.4	1 $s_{3a+7a}$	28.75	2 $s_C(b_{2c})$
307.2	1 $s_{4+7}$	26.87	2 $s_C(b_{1g})$
307.0	1 $s_{5+6}$	22.38	CO( $b_{2u}$ )
40.30	2 $s_C$	18.93	CO + CC(L + T, $b_{1g}$ )
32.15	2 $s_C$	17.15	CH + CO(T, $b_{1g}$ )
29.34	2 $s_C$	16.70	CH + CC(L + T, $b_{1g}$ )
24.08	2 $s_C$	14.71	CH + CC(L, $b_{1g}$ )
22.00	CH + CC(T, $a_{1g}$ )	$b_1$	
20.11	CH + CC(L + T, $b_{2u}$ )	18.20	CO + CC( $b_{1u}$ )
18.33	CC(L, $b_{1g}$ )	14.58	CC( $b_{2g}$ )
16.72	CC + CO(L + T, $a_{1g}$ )	10.89	CC( $b_{1u}$ )
15.12	CC + O(L, $b_{2u}$ )	$a_2$	
$b_2$		11.53	CC( $b_{2g}$ )
308.8	1 $s_{1-3}$	8.06	CC( $a_{1u}$ )
Benzo[ <i>c</i> ]thiophen			
$a_1$		$b_2$	
2 495	1 $s_3$	307.2	1 $s_{3a-7a}$
307.2	1 $s_{3a+7a}$	307.1	1 $s_{1-3}$
307.1	1 $s_{1-3}$	306.8	1 $s_{4-7}$
306.8	1 $s_{4-7}$	306.7	1 $s_{5-6}$
306.7	1 $s_{5-6}$	180.9	2 $p_3$
238.3	2 $s_3$	28.41	2 $s_C(b_{2u})$
180.9	2 $p_3$	26.04	2 $s(b_{1g})$
32.82	2 $s_C + 3s_3(a_{1u})$	21.43	CC(L, $b_{2u}$ )
30.49	2 $s_C(b_{2u})$	17.89	CC(L, $b_{1g}$ )
26.99	2 $s_C(a_{1g})$	16.51	CH $_{4-7}$ (T)
22.92	2 $s_C + 2p_C(L, b_{2u})$	15.46	CH $_{1-3}$ (Y)
21.45	CC $_{3a,7a} + CH(T, b_{2u})$	14.09	CS + CC(L, $b_{1g}$ )
19.33	CH + CC(T, $b_{2u}$ )	$b_1$	
17.70	CH + CC(L)	180.8	2 $p_3$
16.12	CH + CC(T, $a_{1g}$ )	15.86	CC + CS( $b_{1u}$ )
13.55	CS + CH + CC(L + T)	13.16	CC + CS( $b_{2u}$ )
13.29	CS + CC(L + T, $a_{1g}$ )	9.69	CC( $b_{1u}$ )
	$a_2$	$a_2$	
	11.29	11.29	CC( $b_{2g}$ )
	7.84	7.84	CC( $a_{1u}$ )

TABLE 3 (Continued)

N-Methylisoindole		N-Methylisoindole	
Energy	Character	Energy	Character
$a_1$		19.27	CH + CC( $a_1$ )(T)
425.9	1 $s_3$	18.70	CN + CC( $b_2$ )(L + T)
308.2	1 $s_{C(Me)}$	17.88	CC + CH
306.9	1 $s_1$	16.46	CH( $b_2$ )(T)
306.8	1 $s_3$	16.10	CC + CN( $a_1$ )(L + T)
306.3	1 $s_4$	15.89	CH + CC( $b_2$ )(L + T)
306.3	1 $s_7$	15.60	CN + CH $_{Me}(b_2)$ (T)
306.3	1 $s_{7a}$	15.12	CH + CC( $a_1$ )(L)
306.2	1 $s_{3a}$	13.73	CH + CC( $b_2$ )(L)
306.1	1 $s_6$	12.71	CH + CC( $b_2$ )(T)
306.1	1 $s_8$	$a_{11}$	
35.68	2 $s_2(a_1)$	18.04	N + CH $_2(b_1)$
31.05	2 $s_C(a_1)$	15.42	CC + CH $_2(b_1)$
28.82	2 $s_C(a_1)$	13.16	CN + CC( $b_1$ )
27.75	2 $s_C(b_2)$	10.55	CC( $a_2$ )
26.03	2 $s_C(b_2)$	9.26	N + CC( $b_1$ )
25.93	CH $_{Me}(a_1)$	7.14	CC( $a_2$ )
22.49	CH + CC( $a_1$ )(T)		
21.66	CN + CC( $b_2$ )(L + T)		
20.83	CH + CC( $a_1$ )(T)		

## RESULTS AND DISCUSSION

*Molecular and Resonance Energies.*—Comparison of the quinonoid series total energies (Table 1) with the corresponding Kekulé isomers<sup>11</sup> shows that the former series

TABLE 4

Total dipole moments ( $\mu/D$ )  $\pi$ -electron average position ( $\bar{r}/\text{\AA}$ ) and second moments ( $\bar{r}^2$ )

	2 <i>H</i> -Indene	Isoindole	<i>N</i> -Methylisoindole
$\mu(D)$	0.952	2.940	3.428
$\pi_1^{a-c}$	0.130	0.200	-0.213
$\pi_2$	-0.267	-0.173	-0.387
$\pi_3$	-0.387	-0.183	-0.611
$\pi_4$	-0.221	-0.656	-0.614
$\pi_5$	2.101	1.469	1.403
$\pi_6$			2.526
Electronic component			
$\langle z^2 \rangle^{a,d,e}$	-190.04	-186.27	-302.90
$\langle y^2 \rangle$	-87.62	-85.51	-88.50
$\langle x^2 \rangle$	-13.68	-12.21	-15.17
		Benzo[ <i>c</i> ]thiophen ( $spd + 3s^1$ basis)	
$\mu/D$	-1.485	-0.054	
$\pi_1$	0.044	-0.285	
$\pi_2$	-0.357	-0.261	
$\pi_3$	-0.094	-0.439	
$\pi_4$	-0.550	-0.478	
$\pi_5$	1.738	0.169	
$\pi_6$		2.343	
Electronic component			
$\langle z^2 \rangle$	-182.98	-246.62	
$\langle y \rangle$	-84.10	-91.16	
$\langle x^2 \rangle$	-11.91	-13.61	

<sup>a</sup> All molecules lie in the  $yz$  plane with  $z$  as  $C_2$  symmetry axis, the heteroatom lies along the positive  $z$  cartesian direction. <sup>b</sup> Distance from the C(3a)-C(7a) bond, with the molecular orientation as in footnote a. This gives a strict comparison with the data from refs. 8 and 11. <sup>c</sup> The binding energy of orbital  $\pi_1$  increases with an increase in  $i$ . <sup>d</sup> In units of  $10^{-16}$  cm<sup>2</sup> (*cf.* ref. 9). <sup>e</sup> Values for benzene and naphthalene using the same orientation of the molecules are: benzene, -8.40 ( $\pi$ ), -60.87 ( $\sigma$ ), -60.87 ( $\sigma$ ); naphthalene -13.59 ( $\pi$ ), -107.47 ( $\sigma$ ), -239.85 ( $\sigma$ ).

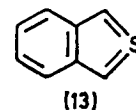
are considerably destabilised. The resonance energies (kJ mol<sup>-1</sup>) gave a simple direct measure of this, and data



from the monocyclic and Kekulé bicyclic compounds for comparison (with Table 1) are as follows: benzene, 212; pyrrole, 149; furan, 89; thiophen, 124 (*sp*-basis), 144 (*sp $d$*  +  $3s^1$  basis); indene, 225; indole 308; benzo[*b*]-furan 232; benzo[*b*]thiophen, 242 (*sp* basis), 283 (*sp $d$*  +  $3s^1$  basis).

In the heterocycles (2a—c), the resonance energy is substantially lower than either naphthalene, the sum of the monocyclic heterocycles and benzene, or the corresponding Kekulé-like heterocycle. In the case of benzo[*c*]furan the resonance energy is even lower than benzene itself. None the less the values are all significant in magnitude and as will be seen below this is consistent with their  $^1\text{H}$  n.m.r. spectra which do not show purely polyolefinic character. It is worth noting in passing that some Hückel<sup>23</sup> and semi-empirical calculations<sup>21,24</sup>

addition lowers the molecular energy by 0.118 45 a.u. marginally less than in the benzothiophen (0.124 03 a.u.). In view of the quinonoid form of the former, an increased  $3d_{\pi}$  participation might have been expected had the  $3d_{\pi}$  orbitals been important in the bonding for they offer an additional mode of delocalisation often expressed as



in (13). Only in the indene-2*H*-indene pair is a tautomeric equilibrium possible and the predicted energy difference favouring the Kekulé type is 153 kJ mol<sup>-1</sup>. Few derivatives of the 2*H*-indene series are known, and those that have been obtained are (a) 2,2-disubstituted<sup>25</sup> and

TABLE 5

Charge density at the nuclei and population analyses											
Isoindole	N(2)	C(1)	C(3a)	C(4)	C(5)	H(1)	H(4)	H(5)	H(2)		
Charge density	190.241	116.746	116.722	116.760	116.771	0.355	0.364	0.364	0.368		
Population $\sigma$	5.8691	4.9093	4.9880	5.1390	5.1362	0.8498	0.8569	0.8576	0.6570		
$\pi$	1.5522	1.1261	1.0725	1.0017	1.0236						
Total	7.4216	6.0354	6.0605	6.1407	6.1598	0.8498	0.8569	0.8576	0.6570		
<i>N</i> -Methylisoidole	N(2)	C(1,3)	C(3a,7a)	C(4,7)	C(5,6)	C(Me)	H(1)	H(4)	H(5)	H(Me)(av)	H(Me) <sub>1,7</sub>
Charge density	190.315	116.758	116.724	116.760	116.770	116.732	0.356	0.364	0.365	0.372	0.370
Population $\sigma$	5.7864	4.8991	4.9888	5.1388	5.1351	5.1237	0.8532	0.8580	0.8587	0.4024	0.8217
$\pi$	1.5243	1.1336	1.0729	1.0027	1.0258	1.1647				0.4205	
Total	7.3107	6.0327	6.0617	6.1415	6.1609	6.2884	0.8532	0.8580	0.8587	0.8229	0.8217
Benzo[ <i>c</i> ]furan	O(2)	C(1)	C(3a)	C(4)	C(5)	H(1)	H(4)	H(5)			
Charge density	292.162	116.798	116.754	116.750	116.785	0.343	0.360	0.360			
Population $\sigma$	6.7716	4.8086	5.0073	5.1342	5.1406	0.8223	0.8477	0.8474			
$\pi$	1.7079	1.0090	1.0530	1.0009	1.0022						
Total	8.4795	5.8986	6.0603	6.1351	6.1488	0.8223	0.8477	0.8474			
Benzo[ <i>c</i> ]thiophen	S(2)	C(1)	C(3a)	C(4)	C(5)	H(1)	H(4)	H(5)			
Charge density	2 666.875	116.877	116.752	116.773	116.784	0.351	0.361	0.362			
Population $\sigma$	12.2349	5.0579	5.0062	5.1340	5.1450	0.8374	0.8511	0.8511			
$\pi$	3.7082	1.1016	1.0027	1.0027	1.0073						
Total	15.9430	6.1595	6.1366	6.1366	6.1523	0.8374	0.8511	0.8511			

lead to substantially lower values for the resonance energy of (2a—c); however in this other work there is a change of definition of resonance energy. We have adopted the original one, namely that the resonance energy is the difference between the molecular total energy and the sum of the non-interacting bonds. In these other calculations, following Dewar *et al.*, the reference non-aromatic structure with zero resonance energy is defined as the corresponding acyclic polyolefin.

We have commented previously that the  $3d_{\pi}$  orbitals appear to fill the role of polarisation functions rather than strongly bonding orbitals in various formally covalent compounds; that is, the orbitals offer additional variational flexibility and have a refining role rather than changing the nature of the existing bonds. For benzo[*c*]thiophen this is also apparent, indeed  $3d_{\pi}$  orbital

(b) unstable<sup>26</sup> in the absence of metal complexing agents, e.g.  $\text{Fe}_2(\text{CO})_9$ .

*Triplet States.*—In the light of the significant amount of resonance energy still evident in (2a—c) we have investigated low lying triplet states to determine whether the instability of the compounds arises partly from low resonance energy and partly from a low lying excited state.

There is no experimental data available for triplet excitation in (2a—e), hence we also record data (Table 2) for the Kekulé-like series (1) where some experimental values are known.\* In all cases, the triplet state is from the highest occupied  $\pi$ -level (HOMO) to the lowest unoccupied  $\pi$ -level (LUMO); in the quinonoid series this corresponds to the transition  $a_2 \rightarrow b_1$ , and in naphthalene  $1a_u \rightarrow 2b_{2g}$ . Comparison with the experimental data

<sup>23</sup> B. A. Hess, L. J. Schaad, and C. W. Holyoke, *Tetrahedron*, 1972, **28**, 3657, 5299; B. A. Hess and L. J. Schaad, *J. Amer. Chem. Soc.*, 1973, **95**, 3907.

<sup>24</sup> M. J. S. Dewar and N. Trinajstić, *Theor. Chim. Acta*, 1970, **17**, 235.

<sup>25</sup> W. R. Roth and J. D. Meier, *Tetrahedron Letters*, 1967, 2053.

<sup>26</sup> H. Tanida, T. Irie, and K. Tori, *Bull. Chem. Soc. Japan*, 1972, **45**, 1999.

\* Using the same geometric parameters to those of the LCGO calculations, we also calculated the triplet excitation energies by the semi-empirical INDO procedure (Table 2). Although the method also shows lower values for the quinonoid than the Kekulé forms, the difference between the two series is much smaller, and the special position of benzo[*c*]furan is not seen. It seems unlikely that the INDO method is satisfactory for excitation energies of this type.

for the Kekulé series shows that the present calculations reproduce the experimental values but are generally too large by *ca.* 0.5 eV in the UHF procedure. The calculated excitation energies in the quinonoid series are all substantially lower, and particularly so with 2*H*-indene and benzo[*c*]furan. This is consistent with the ready decomposition of benzo[*c*]furan at room temperature, even in an evacuated system, and in the absence of air.

*The Molecular Orbital Energy Levels.*—Detailed analysis of the form of the MO wave functions showed that there is a marked similarity in type of orbital both through the

2*s*: 33.86, 2*p*: 17.20; S, 3*s*: 23.94, 3*p*: 11.90 respectively. That there are so few cross-overs is surprising in view of the range of aromaticity, as evidenced by the resonance energies from naphthalene to 2*H*-indene; however, the majority of the orbitals are from the  $\sigma$ -system, and this is likely to be largely unchanged from aromaticity considerations. In fact the  $\pi$ -orbitals are quite variable in position in (2*a*–*c*) and (3). As with the monocyclic heterocycles (4), no degeneracy of the  $\pi$ - (or indeed  $\sigma$ -) levels is required by symmetry, but we have noted<sup>27</sup> that the separation of the two lowest binding energy  $\pi$ -levels [*e.g.* 2*b*<sub>1</sub> and 1*a*<sub>2</sub> in (4; X = O or NH)] varies in an order similar to the resonance energy sequence for [4; X = PH (planar), O, NH, or S] and being nearly degenerate in the most aromatic cases; the same effect occurs in the six-membered azines.<sup>9</sup> In the present molecules (2*a*–*c* and *e*) the highest occupied pair of levels follow the same trend, the separations being in the order 2*H*-indene > benzo[*c*]furan > isoindole > benzo[*c*]thiophen > naphthalene. A consequence of degeneracy of energy levels is frequently an even (highly symmetrical) electron distribution, as for example in the cyclic system C<sub>*n*</sub>H<sub>*n*</sub>. We return to the question of aromatic character in this series of molecules in the electron distribution and n.m.r. spectral sections below.

*Photoelectron Spectra.*—In earlier work<sup>3,4,9–11</sup> we observed that the basis set used for C, H, N, O, and S leads to linear correlations of the former  $IP_{obs} = AIP_{calc} + B$  with *A ca.* 0.75 and *B ca.* 1–2 eV, when Koopmans' theorem is used to equate occupied orbital energies to ionisation potentials. It was not surprising therefore to obtain similar linear correlations for (2*b*–*d*) and (3) as follows (with standard deviations in slope, intercept, and overall in parentheses): naphthalene,  $IP_{obs} = 0.755IP_{calc} + 1.703$  eV (0.014, 0.213, 0.190); benzo[*c*]furan,  $IP_{obs} = 0.770IP_{calc} + 1.42$  eV (0.013, 0.204, 0.177); benzo[*c*]thiophen,  $IP_{obs} = 0.785IP_{calc} + 1.341$  eV (0.238, 0.229); *N*-methylisoindole,  $IP_{obs} = 0.786IP_{calc} + 1.321$  eV (0.017, 0.259, 0.238). In all cases the correlations were assisted by (a) previous experience with related molecules, (b) the 1:1 correspondence of experimental and calculated values at the low binding energy region, and (c) the similarity of the calculated levels for isoindole and its *N*-methyl derivative. The present calculations reproduce the correct order for the first three ( $\pi$ -electron) IPs for the molecules (2*b*–*d*) and (3) namely:  $(IP_1)_{(2a)} < (IP_1)_{(2c)} < (IP_1)_{(2b)} < (IP_1)_{(3)} < (IP_2)_{(2d)} < (IP_2)_{(3)} = (IP_2)_{(2c)} < (IP_3)_{(2d)} = (IP_3)_{(2b)} < (IP_3)_{(2c)} = (IP_3)_{(3)} < (IP_3)_{(2b)}$ , where the nomenclature  $(IP_n)_{(X)}$  refers to the *n*th IP of (*X*). The first four IPs of each of the series thiophen, benzo[*b*]thiophen, and isobenzo[*c*]thiophen have been reported;<sup>28</sup> the present calculations correctly order these 12 IPs with the single exception of  $(IP_3)_{C_8H_8S}$  and  $(IP_4)_{2O}$  which are reversed in order.†

<sup>27</sup> M. H. Palmer and R. H. Findlay, *J.C.S. Perkin II*, 1975, 974.

<sup>28</sup> P. A. Clark, R. Gleiter, and E. Heilbronner, *Tetrahedron*, 1973, 29, 3087.

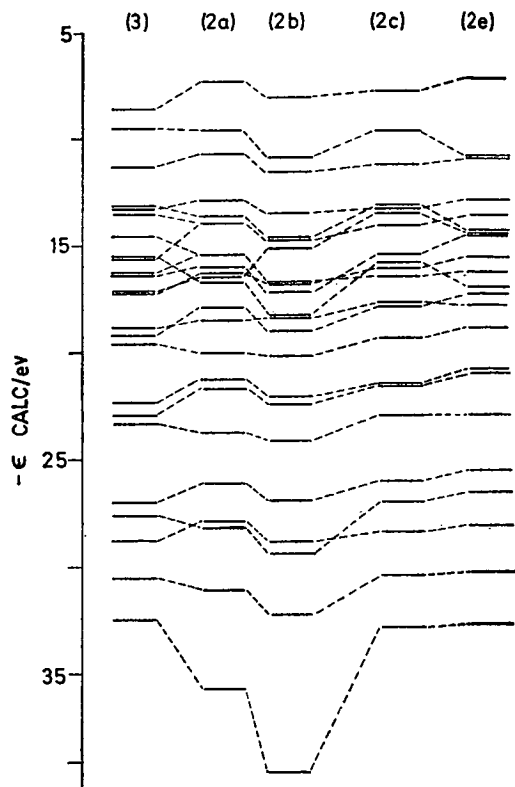


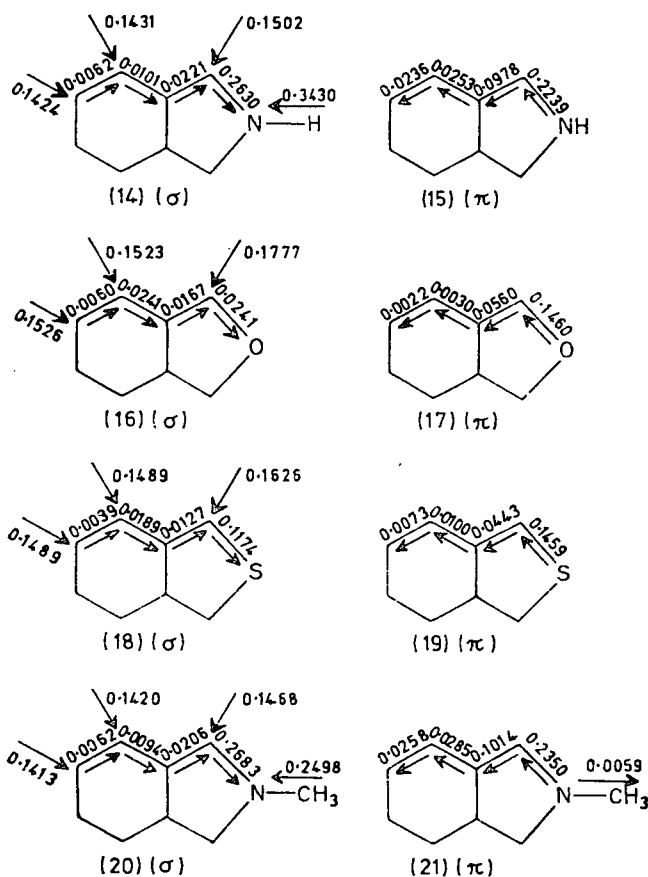
FIGURE 2 Correlation of the molecular orbital energy levels of isoindole (2*a*), benzo[*c*]furan (2*b*), benzo[*c*]thiophen (2*c*), 2*H*-indene (2*e*) with naphthalene (3)

series (2*a*–*e*) and also with naphthalene. The comparatively few cross-overs in energy between orbitals from (2*a*–*c* and *e*) (Figure 2) can be ascribed to differences in the free atom orbital energies of the group X in (2); relevant figures (eV) at the Hartree-Fock limit<sup>20</sup> are: C, 2*s*: 19.20, 2*p*: 11.79; N, 2*s*: 25.72, 2*p*: 15.44; O,

† The extended Hückel,<sup>28</sup> PPP,<sup>22,28</sup> and other<sup>21</sup> semi-empirical methods also lead to satisfactory correlations of the first three IPs, but the scatter is generally rather larger. There is also some ambiguity in the calculated data from ref. 28, where for both EHM and PPP calculations the text refers to considerably different correlation lines for the calculated and experimental IPs to those recorded in the caption to Figures 3 and 4 of ref. 28. It is of course possible that the calculations whilst reproducing the order correctly, do not attribute the correct symmetry type to each ionisation. At present there are few reliable ways of assigning the symmetry type experimentally beyond a  $\sigma$ - $\pi$  separation, and we feel that the substantial agreement between the various calculations (refs. 21, 22, 28, and this paper) is self supporting in all cases except where very closely spaced IPs are concerned.

**The Molecular Charge Distributions.**—The experimental dipole moment for (2a—d) have not yet been reported, but the present work (Table 4) suggests that isoindole should be more polar than indole (calculated 2.31,<sup>11</sup> experimental 2.08 D), while benzo[*c*]furan and benzo[*c*]thiophen should be less polar than either of the nitrogen compounds. Detailed analysis of the net atomic charges shows that much of the dipole moment in isoindole arises from the N<sup>δ-</sup>-H<sup>δ+</sup> bond contribution, as in indole and pyrrole.<sup>8</sup>

**Population Analysis (Table 5).**—As in the Kekulé series, the  $\sigma$ - and  $\pi$ -systems are polarised to the most distant part of the molecule (C-4, C-5), with the two systems operating in a compensating mode. In the *N*-methyl derivative, the CH<sub>3</sub> group is nearly neutral in the  $\pi$ -system (no hyperconjugative effect) but polarised towards nitrogen in the  $\sigma$ -system[(14)—(21)].



In the triplet states the heteroatom X in (2) receives additional population relative to the ground states for all cases (X = NH, O, or S); in view of the  $\pi$ -electron open-shell character of these triplet states it is not surprising that this arises primarily through the  $\pi$ -system, such that the heteroatom is a much weaker donor in the triplet state. However, detailed analysis of the  $\pi$ -MOs shows

<sup>20</sup> M. A. Cooper and S. L. Manatt, *J. Amer. Chem. Soc.*, 1969, **91**, 6325.

<sup>20</sup> H. Sofer and O. E. Polansky, *Monatsh.*, 1971, **102**, 256.

that very considerable reorganisation of the  $\pi$ -density occurs on formation of the triplet state; thus almost all the  $\pi$ -density at C-1 and -3 arises from the two single occupied orbitals rather than the double occupied ones. This is not at all evident from the small changes in total population.

**Second Moments.**—We have shown that the calculated second moments of the electronic charge distribution for a range of heterocycles are close to values obtained from microwave spectroscopy.<sup>8</sup> Again experimental values are awaited for (2a—d) and (3), but we report the calculated values here (Table 4).

The values for the  $\pi$ -system ( $x$ -axis) are of direct interest, and show that the  $\pi$ -electrons are much more diffuse in (2a—e) and (3) than in benzene, and that there is a trend to lower values as the electronegativity of the heteroatom increases. This is in part a result of an increasing level of localisation, and in part a smaller ( $2p_{\pi}$ )<sub>x</sub> atomic orbital radius.

**Average Position of the  $\pi$ -Electrons.**—We noted earlier that aromatic character is associated with a high tendency towards group (sextet, decet, etc.) behaviour, as opposed to pairwise  $\pi$ -electron behaviour.<sup>8,11,27</sup> This group character is best achieved by (ideally) coincident average electronic positions ( $\bar{x}, \bar{y}, \bar{z}$ ). For highly symmetrical hydrocarbons such as benzene this is automatically the case by symmetry and further development of the concept is required to distinguish between say benzene and naphthalene aromatic character in this way. However, for the heterocycles (2a—d) this problem does not arise (Table 4), and we record the average position of the  $\pi$  electrons along the C<sub>2</sub> symmetry ( $z$ -axis) with C(3a)—(7a) ( $y$ -axis) as origin. This enables direct comparison with both the monocyclic<sup>8</sup> and Kekulé bicyclic<sup>11,27</sup> systems reported previously. It is appropriate to omit the core  $\pi$ -electron ( $2p_z$ )<sub>s</sub> orbital ( $\pi_g$ ) from the analysis, and also to omit the antisymmetric CH bonding level ( $\pi_g$ ) of the CH<sub>2</sub> group in (2d). The remaining spread of values (Å) varies from naphthalene (0.0), through benzo[*c*]thiophen (0.76), isoindole (2.12), benzo[*c*]furan to 2*H*-indene (2.49 Å), a sequence again parallel to qualitative views of aromatic character, except for a possible inversion of order in (2a and c). Comparison with the Kekulé series is not straightforward owing to the non-coaxial character of the mean positions in the latter series.

**<sup>1</sup>H N.m.r. Spectra of (2b—d).**—The relative values of the vicinal coupling constants in adjacent bonds of carbocyclic systems have been widely used as a measure of relative bond orders<sup>29</sup> and through them relative levels of aromatic character.<sup>30-32</sup>

As was mentioned above, the molecules (2b—d) show six-spin coupled spectra, and these were analysed by an initial AA<sup>1</sup>BB<sup>1</sup> analysis of the 4—7-H portion, followed by a full 6-spin analysis after a series of non-iterative computations established acceptable estimates for the inter-ring

<sup>31</sup> J. D. White, M. E. Mann, H. D. Kirshenbaum, and A. Mitra, *J. Org. Chem.*, 1971, **36**, 1048.

<sup>32</sup> P. Crews, R. P. Kintner, and H. C. Padgett, *J. Org. Chem.*, 1973, **38**, 4391.

coupling constants. The final results of the AA<sup>1</sup>BB<sup>1</sup>CC<sup>1</sup> spectra yielded root mean square deviations of the transition frequencies around 0.1 Hz, and standard deviations in the coupling constants (Table 6) of *ca.* 0.05 (2b and c) and 0.04 Hz (2d). The values of the vicinal coupling constants  $J_{4,5}$  and  $J_{5,6}$  in *N*-methylisindole (Table 6) are very similar to those of the 1,3-diphenyl derivative ( $J_{4,5}$  8.63,  $J_{5,6}$  6.39 Hz) cited above; <sup>32</sup> furthermore, the near constancy <sup>32</sup> of the average vicinal coupling  $J_{av} = 0.5 (J_{4,5} + J_{5,6})$  of *ca.* 7.5 Hz can be extended to (2b—d). The ratio  $J_{4,5} : J_{5,6}$  is nearly constant for the latter trio (0.70, 0.72, 0.74 respectively), and lies in the expected position based upon aromatic character considerations, relevant figures for the ratio being: benzene

energy and coupling constant ratio indicating opposite aromatic character relative to naphthalene can be reconciled, if it is assumed that almost all the resonance energy in the Kekulé systems (1) is restricted to the carbocyclic ring; under these conditions the coupling constant ratio is expected to be near to that of benzene.

Finally we briefly report that the <sup>13</sup>C n.m.r. spectrum of (2d) yields chemical shifts as follows: CH<sub>3</sub>: 37.34; C-1 and -3: 111.50; C-4 and -7: 119.22; C-5 and -6: 120.30; C-3a and -7a: 124.00 p.p.m. respectively. The carbocyclic C<sub>α</sub> (4,7) and C<sub>β</sub> (5,6) shifts are very similar, and the assignment order was confirmed by off-resonance proton decoupling (above). The C<sub>α</sub>, C<sub>β</sub> shifts of naphthalene (128.3; 126.1),<sup>36</sup> indole (C-4, 121.3; C-5, 122.3),<sup>37</sup> indane

TABLE 6  
<sup>1</sup>H N.m.r. parameters

(a) <sup>1</sup> H Shifts	$\nu_1$	$\nu_4$	$\nu_5$	(a) H Shifts	$\nu_1$	$\nu_4$	$\nu_5$		
Benzo[ <i>c</i> ]furan <sup>a</sup> (i)	-798.90	-738.07	-683.90	<i>N</i> -Methylisindole <sup>c</sup>	-704.54	-751.21	-691.72		
(standard deviation)	(0.024)	(0.027)	(0.030)	(standard deviation)	(0.019)	(0.022)	(0.023)		
Benzo[ <i>c</i> ]thiophen <sup>b</sup>	-763.40	-759.30	-704.16						
(standard deviation)	(0.036)	(0.032)	(0.034)						
	$\nu_1$	$\nu_2$			$\nu_1$	$\nu_3$			
Naphthalene	-782.0 <sup>d</sup>	-745.7 <sup>d</sup>		Cyclohexa-1,3-diene <sup>f</sup>	-568.31 <sup>e</sup>	-582.89 <sup>e</sup>			
(b) <sup>1</sup> H Couplings	$J_{1,3}$	$J_{1,4}$	$J_{1,5}$	$J_{1,6}$	$J_{1,7}$	$J_{4,5}$	$J_{4,6}$	$J_{4,7}$	$J_{5,6}$
Benzo[ <i>c</i> ]furan <sup>a</sup>	-0.001	0.637	0.039	0.044	0.014	0.852	1.008	0.570	6.223
(standard deviation)	(0.035)	(0.039)	(0.046)	(0.046)	(0.039)	(0.049)	(0.048)	(0.40)	(0.048)
Benzo[ <i>c</i> ]thiophen <sup>b</sup>	0.017	0.422	0.030	-0.139	-0.086	0.864	1.030	0.785	6.356
(standard deviation)	(0.040)	(0.056)	(0.068)	(0.071)	(0.058)	(0.056)	(0.057)	(0.048)	(0.060)
<i>N</i> -Methylisindole <sup>c</sup>	0.0	0.457	0.070	-0.070	-0.081	0.890	0.899	0.790	6.456
(standard deviation)	(0.040)	(0.035)	(0.042)	(0.042)	(0.035)	(0.039)	(0.039)	(0.034)	(0.039)
Naphthalene	$J_{1,5}$	$J_{1,6}$	$J_{1,7}$	$J_{1,8}$	$J_{1,2}$	$J_{1,3}$	$J_{1,4}$	$J_{2,3}$	
Cyclohexa-1,3-diene <sup>f</sup>	0.85 <sup>e</sup>	-0.10 <sup>e</sup>	0.23 <sup>e</sup>	-0.45 <sup>e</sup>	8.28 <sup>e</sup>	1.24 <sup>e</sup>	0.74 <sup>e</sup>	6.85 <sup>e</sup>	
					9.64	1.02	1.12	5.04	
					9.55	0.89	0.88	5.35	

<sup>a</sup> Concentration 7.7 g l<sup>-1</sup> in CDCl<sub>3</sub> (273 K), under N<sub>2</sub>, AA<sup>1</sup>BB<sup>1</sup>CC<sup>1</sup> analysis. <sup>b</sup> Concentration 7.3 g l<sup>-1</sup> in CDCl<sub>3</sub> (253 K), under N<sub>2</sub>, AA<sup>1</sup>BB<sup>1</sup>CC<sup>1</sup> analysis. <sup>c</sup> Concentration 16 g l<sup>-1</sup> in CDCl<sub>3</sub> (273 K), AA<sup>1</sup>BB<sup>1</sup> analysis. <sup>d</sup> R. W. Creceley and J. H. Goldstein, *Org. Magnetic Resonance*, 1970, 2, 613. <sup>e</sup> M. A. Cooper and S. L. Manatt, *J. Amer. Chem. Soc.*, 1969, 91, 6325. <sup>f</sup> In C<sub>6</sub>H<sub>6</sub>.

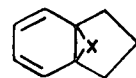
(1.00), naphthalene (0.82),<sup>29,33</sup> and cyclohexa-1,3-diene (0.52).<sup>29</sup> However, extension of the concept to bicyclic systems having lower than *D*<sub>2h</sub> symmetry brings certain problems, especially in the heterocyclic field; for example (a) there may be more than one pair of unique vicinal couplings and (b) the ratio may not yield information for the overall aromatic character, but merely for the one ring under consideration. The Kekulé bicyclic heterocycles (1) exemplify both these difficulties since  $J_{4,5} \neq J_{6,7}$ ; if the former is chosen, by virtue of being further removed from the heteroatom, the ratios  $J_{4,5} : J_{5,6}$  are 0.90 (1; X = NH),<sup>34</sup> 0.93 (1; X = O),<sup>35</sup> and 0.89 (1; X = S).<sup>34</sup> Both thermochemical data, and our earlier calculations,<sup>11</sup> suggest that the series (1; X = NH, O, and S) are less aromatic than naphthalene but more aromatic than benzene *overall*, *i.e.* have lower and higher resonance energies than these two reference compounds respectively. This anomaly, the resonance

<sup>29</sup> R. W. Creceley and J. H. Goldstein, *Org. Magnetic Resonance*, 1970, 2, 613, J. B. Pawliczek and H. Gunther, *Tetrahedron*, 1970, 26, 1755.

<sup>34</sup> K. D. Bartle, D. W. Jones, and R. S. Matthews, *Tetrahedron*, 1971, 27, 5177.

<sup>35</sup> P. J. Black and M. L. Heffernan, *Austral. J. Chem.*, 1965, 18, 353.

(124.0, 125.9),<sup>38</sup> other aromatic bicyclic compounds,<sup>38</sup> and non-aromatic conjugated compounds are similar [(22a), 119.2; (22b), 128.3, 126.3 p.p.m. respectively].<sup>39</sup> The most consistent feature of the *N*-methylisindole spectrum is the general upfield shift relative to most of the



(22) a; X = CH<sub>2</sub>

above by *ca.* 6 p.p.m. for the carbocyclic carbon atom, while the C-1 and -3 resonances are even further upfield than C-2 of pyrrole (118.7) or indole (125.2).<sup>37</sup> Whether this can be ascribed to a ring current effect remains to be determined.

**Conclusions.**—There seems little doubt that the quinonoid heterocycles (2) still retain some resonance energy (RE) in the ground state; however, the numerical values

<sup>30</sup> T. D. Alger, D. M. Grant, and E. G. Paul, *J. Amer. Chem. Soc.*, 1966, 88, 5397.

<sup>37</sup> J. B. Stothers, 'Carbon-13 NMR Spectroscopy,' Academic Press, New York, 1972.

<sup>38</sup> G. Jikeli, W. Herrig, and H. Gunther, *J. Amer. Chem. Soc.*, 1974, 96, 323.

<sup>39</sup> H. Gunther and G. Jikeli, *Chem. Ber.*, 1973, 106, 1863.

are low, and as long suspected, well below that of the Kekulé series (1). The lack of stability of the molecules is readily attributed to the low RE and to the low lying triplet state. The photoelectron spectra of (2b—d) are satisfactorily interpreted by the present calculations; the molecular orbitals are very similar to corresponding ones in naphthalene, but the  $\pi$ -levels disclose substantial shifts in energy between orbitals of similar symmetry across the series (2b—d) and (3). As a consequence of this the electron distributions and degree of electron delocalisation as evidenced by group behaviour is rather variable.

The  $^1\text{H}$  n.m.r. spectra of (2b—d) show strong alteration in magnitude of the vicinal coupling constants, and the ratio again suggests low aromatic character; however, the extension of this concept to heterocycles generally brings major difficulties, unless the aromatic considerations are restricted to individual rings, rather than overall aromaticity.

We are grateful to the S.R.C. and to the Director of the Atlas Laboratory for the provisions of computational facilities and for a grant to S. M. F. K.

[5/413 Received, 28th February, 1975]

---

**Reactivity of Indazoles and Benzotriazole towards *N*-Methylation and Analysis of the  $^1\text{H}$  Nuclear Magnetic Resonance Spectra of Indazoles and Benzotriazoles**

By Michael H. Palmer,\* Robert H. Findlay, Sheila M. F. Kennedy, and Peter S. McIntyre, Department of Chemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ

Reprinted from

JOURNAL  
OF  
THE CHEMICAL SOCIETY

---

PERKIN TRANSACTIONS II

---

1975

## Reactivity of Indazoles and Benzotriazole towards *N*-Methylation and Analysis of the <sup>1</sup>H Nuclear Magnetic Resonance Spectra of Indazoles and Benzotriazoles

By Michael H. Palmer,\* Robert H. Findlay, Sheila M. F. Kennedy, and Peter S. McIntyre, Department of Chemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ

Methylation of some simple 3-, 7-, and 3,7-substituted indazoles in alkaline solution leads to mixtures of both 1- and 2-methyl compounds, in which the 1-methyl isomer predominated in all except the 7-monomethylated cases. Steric effects are less marked in this series than in cinnolines. INDO Calculations of the electron density at the nitrogen were performed in an attempt to account for the relative reactivities. The <sup>1</sup>H n.m.r. spectra of the 1- and 2-methylindazoles are sufficiently different for this to be used as a diagnostic tool for the positions of methylation. A detailed ABCD analysis of various indazole and benzotriazole spectra has been carried out and the results assessed in terms of degree of aromatic character.

RECENTLY we reported<sup>1</sup> studies of the methylation of various 3- and 7-substituted cinnolines, in an attempt to test the relative importance of steric and electronic effects in the system. It was concluded that steric effects were dominant in the reaction.<sup>1,2</sup> In contrast von Auwers *et al.*<sup>3,4</sup> suggested that both isomers are formed in nearly equal amounts when indazole is methylated in alkaline solution. The contrast to the cinnolines appeared to be sharp, and worthy of re-investigation, since the indazole work<sup>3,4</sup> was based upon separation of the isomers by distillation or crystallisation. The virtue of studying the reactions in strongly alkaline solution is that the indazole should react as the anion, as also should benzotriazole chosen as a sterically similar species, but different electronic charge distribution (indazole and benzotriazole<sup>5,6</sup> have  $pK_a$  ca. 14 and 8.2 respectively). This avoids the problem of reaction on the neutral species being influenced by tautomeric N-H groupings.

We have recently reported<sup>7</sup> a number of non-empirical molecular orbital studies of the electronic structure of five-membered ring heterocycles and their benzo-derivatives. It is intended to extend these studies to the present compounds and their reactivity to methylation. However, the latter is not yet practicable in view of the large size of the computations. In these circumstances we have carried out semi-empirical studies (*cf.* ref. 2) of the electron density in the anions derived from indazole and benzotriazole.

Finally, during the course of the study it became apparent that the <sup>1</sup>H n.m.r. spectra for the 1- (1c and d) and 2-methyl- (2c and d) series showed systematic differences (Figure), and that the reported spectral analyses for some related compounds left much to be

desired. We thus undertook a detailed (iterative) analysis of various compounds (1a—d) and (2a—d).

### EXPERIMENTAL

**Synthetic Routes.**—Indazole (3a) and its 3- and 7-methyl derivatives (3b and g) were prepared from the *o*-alkylaniline.<sup>8</sup> 3-Carboxyindazole was obtained by diazotisation of isatin;<sup>9</sup> esterification with methanol-sulphuric acid and reduction of the product with methylmagnesium iodide yielded 2-indazol-3-yl propan-2-ol (3e). The 3-isopropenyl derivative (3f) was obtained by treatment of (3e) with phosphoric oxide in benzene.<sup>10</sup> The indazoles (3c, d, and j) were prepared by synthesis of the corresponding 3-substituted indole;<sup>10</sup> periodate cleavage<sup>11</sup> of the latter yielded the *o*-aminophenyl ketone, which was diazotised and then reduced (sodium metabisulphite) to the hydrazine which cyclised spontaneously to the indazole.<sup>12</sup> 7-Ethoxycarbonylindazole was obtained from methyl 3-methyl-2-nitrobenzoate by reduction to the amine, acetylation, and nitrosation with dinitrogen trioxide.<sup>13</sup> The 7-isopropenyl compound (3h) was prepared from this ester, as for (3f). Attempted reduction ( $H_2$ -Pd,  $H_2$ -Ni,  $N_2H_4$ -Ni) of the isopropenyl compounds (3f and h) to the corresponding isopropyl compound led to complex mixtures which did not contain the isopropylindazoles. The indazoles and their methylation ratios are given in Table 1.

**Typical Methylation Reaction.**—Indazole (3.00 g), potassium hydroxide (3.00 g), methyl iodide (9.00 g), and methanol (25 ml) were boiled for 4 h, cooled, and diluted with water (100 ml). The mixture was extracted ( $CHCl_3$ , 3 × 50 ml) and dried ( $MgSO_4$ ). A portion of the chloroform solution was evaporated and directly investigated by <sup>1</sup>H n.m.r. spectroscopy; the *N*-methyl resonances were generally well separated (Table 3)<sup>14</sup> and the mixture was analysed by integration; estimates based upon the height of the *N*-methyl resonance were inaccurate owing to long range coupling of the 2-methyl group to 3-H. N.m.r. investigation of the aqueous phase in the above separation showed

\* M. H. Palmer, A. J. Gaskell, and R. H. Findlay, *J.C.S. Perkin II*, 1974, 1893 and references cited therein.

<sup>8</sup> R. Huisgen and K. Bast, *Org. Synth.*, 1962, 42, 69.

<sup>9</sup> H. R. Snyder, C. B. Thomas, and R. L. Hinman, *J. Amer. Chem. Soc.*, 1952, 74, 2009.

<sup>10</sup> M. H. Palmer and P. S. McIntyre, *J. Chem. Soc. (B)*, 1969, 446.

<sup>11</sup> W. Dalby, *J. Amer. Chem. Soc.*, 1966, 88, 1049.

<sup>12</sup> K. von Auwers and P. Stodter, *Ber.*, 1926, 59, 529.

<sup>13</sup> R. Huisgen and H. Nakaten, *Annalen*, 1954, 583, 84.

<sup>14</sup> J. Elguero, A. Fruchier, and R. Jacquier, *Bull. Soc. chim. France*, 1966, 2075.

<sup>1</sup> M. H. Palmer and P. S. McIntyre, *Tetrahedron*, 1971, 27, 2913.

<sup>2</sup> M. H. Palmer, A. J. Gaskell, P. S. McIntyre, and D. W. W. Anderson, *Tetrahedron*, 1971, 27, 2921.

<sup>3</sup> K. von Auwers and K. Düesberg, *Ber.*, 1920, 53, 1179 and later papers.

<sup>4</sup> L. C. Behr, R. Fusco, and C. H. Jarboe, 'Pyrazoles, Pyrazolines, Pyrazolidines, Indazoles and Condensed Rings,' ed. R. H. Wiley, Interscience, New York, 1967.

<sup>5</sup> A. Albert, 'Physical Methods in Heterocyclic Chemistry,' ed. A. R. Katritzky, Academic Press, New York, 1963, vol. 1.

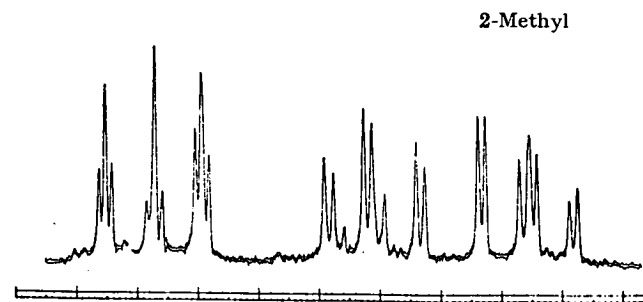
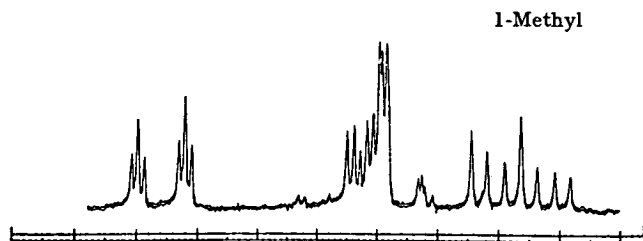
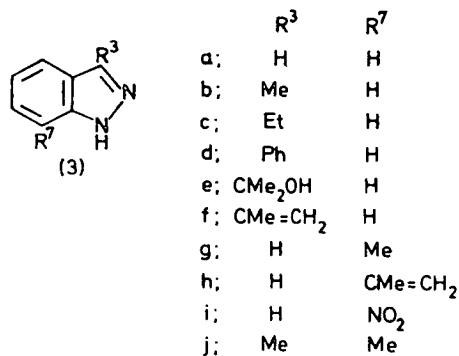
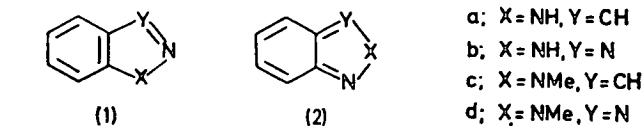
<sup>6</sup> J. E. Fagel and G. W. Ewing, *J. Amer. Chem. Soc.*, 1961, 73, 4360.

it to be free of indazole derivatives, so that no loss of material by quaternisation had occurred.

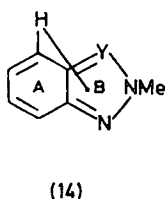
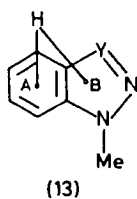
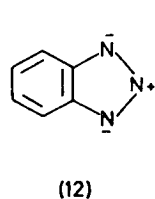
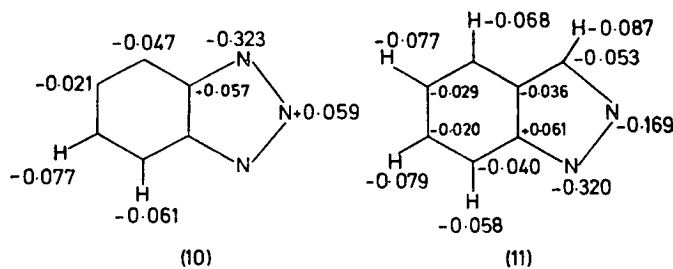
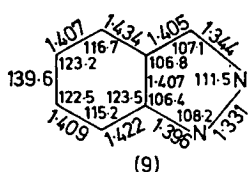
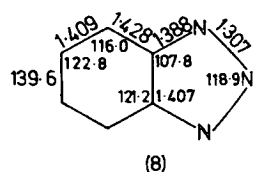
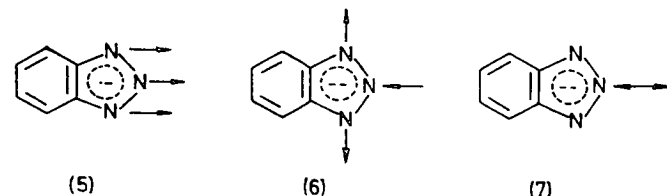
Evaporation of the bulk of the chloroform extract gave the crude methylindazole mixture. This was separated by

The physical properties of the indazoles obtained are recorded in Table 1. The assignment of structure to the isomeric *N*-methylindazoles was either by comparison with known samples<sup>4</sup> prepared by unambiguous methods, or by n.m.r. spectroscopy<sup>14</sup> (discussed in detail below) but the deshielding of the 7-H in the 2-methyl series is particularly evident (Figure).

The procedure for methylation of benzotriazole in alkaline solution was very similar. The products were again separated by chromatography after estimation of the mixture by n.m.r. spectroscopy. The characteristic AA<sup>1</sup>-BB<sup>1</sup> pattern of the 2-methylbenzotriazole spectrum contrasts with the ABCD type from the 1-methyl series;



<sup>1</sup>H N.m.r. spectra of 1- and 2-methylindazole



chromatography on alumina, where the 1-isomer was eluted preferentially by light petroleum (b.p. 60–80°) followed by benzene; isomerism of 2-methyl- to 1-methylindazole was observed when in prolonged contact with alumina.

<sup>15</sup> I. L. Karle, K. Britts, and P. Gunn, *Acta. Cryst.*, 1964, **17**, 496.

although the 1H-benzotriazole form of [(1a)  $\rightleftharpoons$  (2a)] is normally predominant in solution, unsubstituted benzotriazole shows an AA<sup>1</sup>BB<sup>1</sup> spectrum owing to rapid 1-H-3-H isomerism.

**INDO Calculations of the Charge Distribution in the Anions.**—The molecular geometry of the indazole and benzotriazole anions is not known. We started with the geometry of indole-3-acetic acid<sup>15</sup> for the carbocyclic system of the anions, and then proceeded to optimise the geometry of the five-membered ring. In our iterative process the total energy in the INDO semi-empirical method<sup>16</sup> is calculated for various geometric arrangements and allowed to achieve an optimal value by a sequence of one-dimensional parabolic minimisations. Thus certain modes of motion of the nuclei corresponding to vibrational modes, and as in (5)–(7) are performed in sequence which finally yielded (8) for the benzotriazole anion. By a similar but more lengthy process we arrived at (9) for the indazole anion.<sup>17</sup>

The large angle at N-2 of (8) is at first sight surprising, but is reasonable since much of the negative charge in the anion is localised in N-1 and N-3 of (10) thereby leading to substantial mutual repulsion by these centres; similarly the

<sup>16</sup> J. A. Pople and D. L. Beveridge, *Approximate Molecular Orbital Theory*, McGraw-Hill, New York, 1970.

<sup>17</sup> M. J. S. Dewar and G. J. Gleicher, *J. Chem. Phys.*, 1966, **44**, 769.



longer N-N distance in (9) than (8) can be accounted for in terms of charge repulsion (11). A break-down of the electron density at the nitrogen atoms, and conversion of the  $\sigma$ -system to radial and tangential components<sup>8</sup> at the ring atoms is shown in Table 3. The apparent positive charge density at N-2 in the indazole anion corresponds to a small contribution of the canonical form (12) to the overall electronic structure, and is often found in self-consistent field calculations on ions.<sup>18</sup>

## RESULTS AND DISCUSSION

**Methylation Ratios.**—The comparatively small change in the 1-:2-methylation ratio in indazole and its 3- and 7-Me derivatives (Table 1) shows only a weak steric

methylation procedure; thus the preponderance of 2-methylation from the 7-nitro-compound (3i) is probably a result of steric interactions.

As with indazoles,<sup>3,4</sup> methylation of benzotriazole leads to various proportions of the *N*-methyl compounds depending upon the reagent (Table 2); under strongly alkaline conditions the ratio is largely independent of the reagent (MeI or Me<sub>2</sub>SO<sub>4</sub>) and is also independent of the base concentration; the present results are in quantitative agreement with earlier work.<sup>20,21</sup>

In competition for a deficiency of methyl iodide (10 mol %), a reaction with equimolar amounts of indazole (IND) and benzotriazole (BT) after allowance for the

TABLE 1

Indazoles and their methylation ratios<sup>a</sup>

N-Substituent	C-Substituent										
	None (1a)	3-Me (1b)	3-Et (1c)	3-Ph (1d)	3-(CMe <sub>2</sub> OH) (1e)	3-(CMe=CH <sub>2</sub> ) (1f)	7-Me (1g)	7-NO <sub>2</sub> (1i)	3,7-Me <sub>2</sub> (1j)	5-NO <sub>2</sub> (1k)	6-NO <sub>2</sub> (1l)
H											
M.p. (°C)	147— 149 <sup>b</sup>	113	<i>c</i>	107— 108 <sup>d</sup>	167 <sup>e</sup>	84— 85 <sup>f</sup>	136— 137 <sup>g</sup>	190 <sup>h</sup>	174 <sup>k</sup>	208 <sup>b</sup>	180 <sup>b</sup>
1-Me											
M.p. (b.p.) (°C)	60— 61 <sup>i</sup>	35 <sup>i</sup>	(146—148 at 35 mmHg) <sup>i</sup>	80 <sup>i</sup>			52 <sup>j</sup>	104 <sup>k</sup>	51 <sup>l</sup>	154 <sup>k</sup>	126 <sup>k</sup>
2-Me											
M.p. (b.p.) (°C)	56 <sup>i</sup>	79— 80 <sup>i</sup>	(130 at 23 mmHg) <sup>i</sup>				(273— 275)	149 <sup>k</sup>	91 <sup>m</sup>	130 <sup>k</sup>	157— 158 <sup>k</sup>
Methylation ratio (1): (2) <sup>a</sup>	50:50	65:35	70:30	74:26	66:34	82:18	46:54	29:71	57:43	47:53	50:50

<sup>a</sup> Estimated accuracy  $\pm 3\%$ . <sup>b</sup> Commercial samples (Koch-Light) after purification. <sup>c</sup> Viscous oil which failed to solidify. <sup>d</sup> K. von Auwers and K. Hüttenes (*Ber.*, 1922, 55, 1112) give m.p. 107°. <sup>e</sup> Found: C, 68.5; H, 6.7; N, 15.1. C<sub>10</sub>H<sub>11</sub>N<sub>2</sub>O requires C, 68.2; H, 6.8; N, 15.1%. <sup>f</sup> Found: C, 75.7; H, 6.2; N, 17.5. C<sub>10</sub>H<sub>9</sub>N<sub>2</sub> requires C, 75.9; H, 6.3; N, 17.7%. <sup>g</sup> Found: C, 72.5; H, 6.0; N, 21.1. C<sub>9</sub>H<sub>8</sub>N<sub>2</sub> requires C, 72.7; H, 6.05; N, 21.2%. <sup>h</sup> Found: C, 73.8; H, 6.9; N, 20.0. C<sub>9</sub>H<sub>10</sub>N<sub>2</sub> requires C, 73.95; H, 6.85; N, 19.15%. <sup>i</sup> Ref. 4 gives the following m. or b.p.s for indazoles: 1-methyl, 61°; 2-methyl, 56°; 1,3-dimethyl, 36°; 2,3-dimethyl, 80°; 2-methyl-3-ethyl (389° at 760 mmHg); 1-methyl-3-phenyl, 80°; 1-methyl-6-nitro, 108—109°; 2-methyl-6-nitro, 159—160°. <sup>j</sup> Found: C, 73.95; H, 6.7; N, 19.25. C<sub>9</sub>H<sub>10</sub>N<sub>2</sub> requires C, 73.95; H, 6.85; N, 19.15%. <sup>k</sup> E. Noelting (*Ber.*, 1904, 37, 2576) gives 1-(or 2)-methyl-5-nitro, m.p. 128—129°; 1-(or 2)-methyl-7-nitro, m.p. 144—145°; 1-(or 2)-methyl-6-nitro, m.p. 155°. <sup>l</sup> Found: C, 75.15; H, 7.4; N, 17.3. C<sub>10</sub>H<sub>12</sub>N<sub>2</sub> requires C, 75.0; H, 7.5; N, 17.5. <sup>m</sup> Found: C, 73.9; H, 6.85; N, 19.1. C<sub>9</sub>H<sub>10</sub>N<sub>2</sub> requires C, 73.95; H, 6.85; N, 19.15%.

effect and contrasts strongly with the corresponding cinnolines.<sup>1</sup> However, as in the latter series (and indeed pyridazines<sup>19</sup>), the equal stereoelectronic effects of 3-phenyl and 3-ethyl substituents is apparent; it seems that the buttressing effect of the 4-H leads to a preferred out-of-plane conformation of the Ph group (3d) which shows an apparent size (thickness) similar to that of a CH<sub>3</sub> or CH<sub>2</sub> group. In contrast, the 3-isopropenyl group (3f) which exerts a larger effect on the alkylation ratio than (3d) must lie closer to co-planarity.

The high proportion of 2-methylation in the indazol-3-ylpropan-2-ol (3e) is curious, and suggests that stabilisation of the negative charge at N-2 may be achieved by some form of internal hydrogen bonding with the OH group. Since the N-1 : N-2 methylation ratios are little changed between the parent and 5- or 6-nitro-compounds, it seems that strongly polar groupings in the carbocyclic ring probably exert a minimal electronic effect on the

statistical 2 : 1 advantage at N-1 and N-3 in benzotriazole yielded the composition (*N*-Me-BT : *N*-Me-IND) 44.7 : 55.3; the relative proportions of the four products

TABLE 2

Reagent	Methylation of benzotriazole			
	CH <sub>3</sub> I- MeOH	CH <sub>3</sub> I- MeOH- NaOH	(CH <sub>3</sub> ) <sub>2</sub> SO <sub>4</sub> - NaOH- H <sub>2</sub> O	CH <sub>3</sub> N <sub>2</sub> - EtOH
1-Methylation	100	61	65	35
2-Methylation	0	39	35	65

were (1c), 25.8; (2c), 19.5; (2d), 36.4; (1d), 18.3% in essential agreement with the ratios when the reactions were carried out under separate conditions. Thus we find equal reactivity at N-1 in the anions (8) and (9), which are nearly equally sterically hindered.

The steric and electronic effects raise a number of

<sup>10</sup> L. Salem, 'Molecular Orbital Theory of Conjugated Systems,' Benjamin, New York, 1966.

<sup>11</sup> H. Lund, personal communication.

<sup>20</sup> F. Krollpfeiffer, A. Rosenberg, and C. Muhlhausen, *Annalen*, 1935, 575, 113.

<sup>21</sup> N. O. Cappel and W. C. Fernelius, *J. Org. Chem.*, 1940, 5, 40.

problems; it is pleasing to note that the calculated total electron densities at N-1 in the two systems (10) and (11)

TABLE 3

Charge distribution in the indazole and benzotriazole anions (10) and (11)

	N-1	N-2
Indazole anion		
2s	1.6068	1.5665
II	1.3725	1.0772
Radial 2p <sub>σ</sub>	1.3562	1.4681
Tangential 2p <sub>σ</sub>	0.9847	1.0570
3-Methylindazole anion		
2s	1.6065	1.5666
II	1.3553	1.0777
Radial 2p <sub>σ</sub>	1.3612	1.4633
Tangential 2p <sub>σ</sub>	0.9880	1.0499
7-Methylindazole anion		
2s	1.6065	1.5666
II	1.3699	1.0785
Radial 2p <sub>σ</sub>	1.3574	1.4667
Tangential 2p <sub>σ</sub>	0.9848	1.0564
Benzotriazole anion		
2s	1.6266	1.4750
II	1.3423	1.0195
Radial 2p <sub>σ</sub>	1.3464	1.4511
Tangential 2p <sub>σ</sub>	1.0169	0.9951

are identical, while that at N-2 in (10) and (11) is substantially less negative (Table 3). In heterocyclic

a CNDO-2 investigation of the methylation of the tetrazole anion by methyl chloride we have shown<sup>23</sup> that at both N-1 and N-2, σ-attack is preferred to either II or σ-II; however, an example of the latter must of course occur in the 3-methylation of the indole anion.<sup>23</sup>

It is well established from non- and semi-empirical<sup>2,7</sup> calculations that 2s<sub>N</sub> and 2s<sub>O</sub> orbital densities have a heavily localised character, and that the principal bonding between the atoms is between 2p<sub>N</sub>, 2p<sub>O</sub>, and 1s<sub>H</sub>; this led to the earlier proposition<sup>2</sup> that reactivity might be associated (in reactions which are mechanistically similar and where alternative sites exist) with the radial 2p<sub>N</sub> electron density in the heterocycle. For the cinnolines the reaction can unequivocally be assigned to pure σ-attack. For reaction at N-2 in both (8) and (9) again attack through the σ-system seems probable (Table 3) while for N-1 there is little to choose between pure σ- and pure II-attack. Since of course, it is possible to re-partition the electron density into mixed σ-II systems (equivalent to sp<sup>3</sup> hybridisation at N) it is possible that some at least of the N-1 reactivity occurs through either the latter or the II-route. This would account for the still substantial amount of attack on N-1 in (3i) in particular. This possibility of more than one mechanism for reaction in the indazole and benzotriazole anions calls for further experiments to test the hypothesis. It has an additional advantage that it does

TABLE 4  
Coupling constants (Hz) of indazoles

Indazoles (1a and c)	J <sub>3,7</sub>	J <sub>4,5</sub>	J <sub>4,6</sub>	J <sub>4,7</sub>	J <sub>5,6</sub>	J <sub>5,7</sub>	J <sub>6,7</sub>
Substituents							
None (0.08) <sup>a</sup>	0.89	8.22	1.00	1.12	6.97	1.04	8.48
None (0.10) <sup>b</sup>	0.99	8.15	0.95	1.04	6.89	0.82	8.36
1-Me (0.04)	0.85	8.81	0.89	0.97	6.87	0.86	8.38
3-Me (0.06) <sup>c</sup>		8.24	0.96	1.01	7.01	0.79	8.34
1,3-Me (0.07) <sup>c</sup>		8.29	1.02	1.08	6.96	0.85	8.46
7-Me (0.05) <sup>c,d</sup>		8.28	0.66		6.75		
3,7-Me <sub>2</sub> (0.10) <sup>c,d</sup>		8.11	0.91		6.99		
1,3,7-Me <sub>3</sub> (0.05) <sup>c,d</sup>		8.36	0.68		6.99		
3-Et (0.10)		8.26	0.91	1.01	6.97	0.90	8.44
1-Me, 3-Et (0.30)		8.13	1.05	0.84	7.14	0.80	9.17
1-Me, 5-NO <sub>2</sub> (0.40)	0.84		2.15	0.68			9.42
1-Me, 6-NO <sub>2</sub> (0.03)	0.95	8.83		0.76		1.97	
1-Me, 7-NO <sub>2</sub> (0.05)		8.02	0.91		7.87		
Indazoles (2a and c)							
Substituents							
2-Me (0.04) <sup>e</sup>	0.89	8.35	1.14	1.04	6.71	0.85	8.52
2,3-Me <sub>2</sub> (0.07) <sup>b</sup>		8.44	1.00	1.02	6.86	0.72	8.53
2,7-Me <sub>2</sub> (0.10) <sup>c,d</sup>		8.72	0.78		6.83		
2,3,7-Me <sub>3</sub> (0.05) <sup>c,d</sup>		8.39	1.00		6.71		
2-Me, 3-Et (0.09)		8.37	1.03	1.16	6.79	0.88	8.68
2-Me, 5-NO <sub>2</sub> (0.09)	0.84		2.15	0.75			9.65
2-Me, 6-NO <sub>2</sub> (0.07) <sup>f</sup>	0.84	9.23		0.80		2.03	
2-Me, 7-NO <sub>2</sub> (0.06)		8.02	0.91		7.87		

<sup>a</sup> Standard deviations in parenthesis. <sup>b</sup> Ref. 24. <sup>c</sup> Spin decoupling of 3- and 7-substituents in this analysis. <sup>d</sup> Order of coupling constants J<sub>4,5</sub> > J<sub>5,6</sub> assigned by analogy with other members of the series where no 7-substituent present. <sup>e</sup> J<sub>2,Me,3</sub> 0.5, J<sub>3,5</sub> 0.11 Hz. <sup>f</sup> J<sub>3,4</sub> 0.12, J<sub>3,5</sub> 0.12 Hz.

analogues of the cyclopentadienyl anion [e.g. (10) and (11)], it is possible for the reagent to attack purely in-plane (σ), perpendicular to the ring (pure II), or at any intermediate angle (σ-II). There is no certain knowledge of the mechanism in the cases (10) and (11); but in

not require the fortuitous balance of σ-steric and electronic factors at N-1 and N-2 in the indazole anion. It is

<sup>23</sup> M. H. Palmer and K. Mallen, to be published.

<sup>23</sup> M. H. Palmer, 'Structure and Reactions of Heterocyclic Compounds,' Arnold, London, 1967, p. 317.

of course clear that the 6/5 fusion in these bicyclic systems leads to lower steric effects in plane than the 6/6 fusion of the cinnolines.

*N.m.r. Spectra of the Indazole and Benzotriazoles.*—

(i) *Assignments.* As mentioned above the mixtures (1c)

In the 3- and 7-unsubstituted indazoles the ABCD analysis was assigned to the hydrogen atoms shown on the basis of the inter-ring coupling ( $J_{3,7}$ ) which is well known in bicyclic aromatic<sup>28</sup> and heterocyclic compounds.<sup>24,29</sup> In the 2-methyl series this was also

TABLE 5  
Chemical shifts ( $\delta$ ) of indazoles

Indazoles (1a and c) Substituents	N-Me	3-H	4-H	5-H	6-H	7-H
None (0.06) <sup>a,b</sup>		8.1030	7.7686	7.1163	7.3374	7.5886
1-Me (0.03) <sup>c</sup>	3.95	7.9335	7.7071	7.0954	7.3374	7.5886
3-Me (0.03) <sup>c,d</sup>		(2.63)	7.6436	7.0976	7.3167	7.3990
1,3-Me <sub>2</sub> (0.07) <sup>c,d</sup>	3.90	(2.53)	7.5682	7.0452	7.2865	7.2048
7-Me (0.03) <sup>c,d</sup>		8.110	7.5599	7.0385	7.1033	(2.56)
1,7-Me <sub>2</sub> (0.03) <sup>c,d</sup>	4.18	7.850	7.550	6.950	7.000	(2.62)
3,7-Me <sub>2</sub> (0.08) <sup>c,d</sup>		(2.54)	7.4805	7.0349	7.110	(2.47)
1,3,7-Me <sub>3</sub> (0.03) <sup>c,d</sup>	4.15	(2.55)	7.3770	6.9020	6.9615	(2.73)
3-Et (0.05) <sup>c</sup>		3.05, 1.42	7.6924	7.0928	7.3171	7.3940
1-Me, 3-Et (0.07) <sup>c</sup>	3.95	2.98, 1.30	7.6607	7.0771	7.3363	7.2944
5-NO <sub>2</sub> (0.03) <sup>b</sup>		8.370	8.810		8.260	7.770
1-Me, 5-NO <sub>2</sub> (0.03) <sup>c</sup>	4.15	8.1526	8.6662		8.2345	7.4362
6-NO <sub>2</sub> (0.03) <sup>b</sup>		9.200	8.000	8.960		9.270
1-Me, 6-NO <sub>2</sub> (0.03) <sup>c</sup>	4.18	8.0931	7.8128	7.9787		8.3581
7-NO <sub>2</sub> (0.03) <sup>b</sup>		9.250	9.320		9.190	
1-Me, 7-NO <sub>2</sub> (0.03) <sup>c</sup>	4.18	8.0600	7.9376	7.1436	8.0076	
Indazoles (2a and c) Substituents						
2-Me (0.04) <sup>c</sup>	3.80	7.672	7.5591	7.0203	7.2251	7.6807
2,3-Me <sub>2</sub> (0.08) <sup>c</sup>	3.85	(2.33)	7.5841	6.9336	7.1883	7.4014
2,7-Me <sub>2</sub> (0.09) <sup>c,d</sup>	4.10	(7.70)	7.4183	6.9508	6.9972	(2.62)
2,3,7-Me <sub>3</sub> (0.06) <sup>c,d</sup>	3.96	(2.42)	7.3092	6.8907	6.9763	(2.58)
2-Me, 3-Et (0.08) <sup>c</sup>	3.99	2.91, 1.30	7.5326	6.9702	7.2107	7.6096
2-Me, 5-NO <sub>2</sub> (0.08) <sup>c</sup>	4.26	8.2030	8.6663		8.0690	7.7160
2-Me, 6-NO <sub>2</sub> (0.06) <sup>c</sup>	4.30	8.0350	7.7308	7.8551		8.6320
2-Me, 7-NO <sub>2</sub> (0.06) <sup>c</sup>	4.16	8.2220	8.0339	7.1537	8.2817	

<sup>a</sup> Standard deviations in parentheses. <sup>b</sup> In (CD<sub>2</sub>)<sub>2</sub>CO, concentration 200 ± 50 g l<sup>-1</sup>. <sup>c</sup> In CDCl<sub>3</sub>, concentration 200 ± 20 g l<sup>-1</sup>.  
<sup>d</sup> Spin decoupling of the 3- or 7-substituent included in present analysis.

and (2c) and (1d) and (2d) were separated and subjected to detailed n.m.r. analyses. Of the present compounds only indazole<sup>24</sup> and benzotriazole and its 1- and 2-methyl derivatives<sup>25-27</sup> have been studied by iterative analyses; for indazole only the value of  $J_{5,7}$  lies outside the combined probable errors of the two analyses,<sup>24</sup> and in the light of the values for the other compounds (Tables 4 and 5) it seems that the earlier value may be more reliable, although the overall error in the present analysis is better. The results for benzotriazole (Table 6) and its 1- and 2-methyl derivatives are in good agreement with some of the previous analyses<sup>25,26</sup> but not others.<sup>27</sup>

<sup>c</sup> In studies of polycyclic hydrocarbons two steric factors resulting from *peri*-interactions have been identified,<sup>30</sup> namely naphthalene ( $S_N + 0.08$  Hz) and phenanthrene ( $S_{Ph} 0.30$  Hz). In fused 5,6-bicyclic systems the effect of a 3-methyl group is likely to lie between these two figures. Although comparison of (3a) with (3b), and their 1- and 2-methyl corresponding pairs does disclose an increase in  $J_{4,5}$  (ca. 0.10 Hz) this is not observed with all pairs when a 3-methyl substituent is introduced, particularly in the 7-methylindazole series.

<sup>24</sup> P. J. Black and M. L. Heffernan, *Austral. J. Chem.*, 1963, 16, 1051.

<sup>25</sup> (a) R. E. Rondeau, H. M. Rosenberg, and D. J. Dunbar, *J. Mol. Spectroscopy*, 1968, 23, 139; (b) H. Gunther, *Tetrahedron Letters*, 1967, 2967.

<sup>26</sup> A. J. Boulton, P. J. Halls, and A. R. Katritzky, *Org. Magnetic Resonance*, 1969, 1, 311.

consistent with assignment of the 7-H to the low-field position by analogy with early work on quinolines<sup>29</sup> and other heterocycles.<sup>14</sup> The assignments of the 3- and

TABLE 6  
Chemical shifts ( $\delta$ ) and coupling constants (Hz) for benzotriazoles

Substituent	4-H	5-H	6-H	7-H		
None (0.04) <sup>c</sup>	7.8728	7.3210	7.3210	7.8728		
1-Me (0.03)	8.0127	7.3356	7.4565	7.4751		
2-Me (0.03)	7.8306	7.3328	7.3328	7.8306		
Substituent	$J_{4,5}$	$J_{4,6}$	$J_{4,7}$	$J_{5,6}$	$J_{5,7}$	$J_{6,7}$
None (0.07) <sup>c</sup>	8.34	1.17	0.88	7.08	1.17	8.34
1-Me (0.02)	8.20	0.77	0.93	6.61	1.18	8.36
2-Me (0.05)	8.65	1.03	1.04	6.79	1.03	8.65

<sup>c</sup> Standard deviations in parentheses.

7-substituted compounds then followed on the assumption that a methyl substituent had an insignificant or small effect upon coupling constants,<sup>c</sup> but led to a

<sup>27</sup> P. J. Black and M. L. Heffernan, *Austral. J. Chem.*, 1962, 15, 862.

<sup>28</sup> R. W. Creceley and J. H. Goldstein, *Org. Magnetic Resonance*, 1970, 2, 613.

<sup>29</sup> C. W. Haigh, M. H. Palmer, and B. Semple, *J. Chem. Soc.*, 1965, 6003.

<sup>30</sup> M. A. Cooper and S. L. Manatt, *J. Amer. Chem. Soc.*, 1969, 91, 6325.

slight upfield shift on *ortho*-H or downfield shift with *peri*-H.<sup>29</sup> Coupling between the 7-methyl substituent (3g and j) and 4-, 5-, and 6-H was observed and the iterative analysis was assisted by spin decoupling of this substituent. The effects of a nitro-group on the adjacent coupling constants in the indazoles follow earlier studies of quinolines<sup>29,31</sup> and naphthalenes.<sup>32</sup>

(ii) *Comparison of the 1- and 2-methyl series (1) and (2).*

(a) Chemical shifts. Although the spectra (Figure) appear very different, this largely arises from the deshielding effect on the N-1 lone pair on H-7 in the 2-methyl series (1) and the consequent change in pattern from ABCD towards ABCX. Thus the 2-methyl series often show non-overlapping multiplet spectra (Table 5), in contrast to the 1-methyl series. It is well established that chemical shifts are more dependent upon solvent and concentration than the coupling constants; in the present work we used comparatively strong solutions in  $\text{CDCl}_3$ , which were chosen to be similar in strength in all cases, especially within 1- and 2-methyl isomeric pairs. Over five pairs of the 1- and 2-methyl series, the latter give signals *upfield* from those of the former in all cases, the differences being small and varying with the nucleus: 9.0 (4-H), 6.0 Hz (5-, 6-H). The 7-H difference is omitted owing to the change in environment noted above. At the present concentration range in  $\text{CDCl}_3$  the naphthalene resonances are at 782 (1-H,  $\alpha$ ) and 746 Hz (2-H,  $\beta$ ) respectively; whilst these are downfield of the position for the 4-( $\alpha$ ), 5-, and 6-H( $\beta$ ) in all the methyl-indazoles, the shift differences between the naphthalene and indazole resonances are small when the latter are compared<sup>30</sup> with those of cyclohexadiene (1-H; 568.3; 2-H, 582.9 Hz in  $\text{C}_6\text{H}_6$ ) even when possible differences in solvent are allowed for. The chemical shifts of the 3- and 5-H of 1-methylpyrazole are similar (730 and 736 Hz respectively);<sup>33</sup> by comparison the 3-H signals of the indazoles, which lie in the range 765–800 Hz, are significantly downfield in both 1- and 2-methyl series. Thus in the light of these effects upon both protons in the carbocyclic and heterocyclic rings we conclude that the shifts indicate ring current contributions from both rings (13) rather than from one (14) or no rings.

(b) Coupling constants. Comparison of the coupling

constants  $J_{a,b}$  (Table 4) between the 1- and 2-methyl series (five pairs of compounds) shows that  $J_{4,5}$  and  $J_{5,6}$  are larger in the 2-methyl series by *ca.* 0.20 and 0.23 Hz respectively; while  $J_{4,5}$  and  $J_{6,7}$  are also larger in the latter series the difference on average is much smaller (0.10 Hz). When interpreted through the well established vicinal coupling–II-bond order relationships<sup>30,31</sup> these suggest that the 4,5-bond order is slightly higher in the 2-methyl (1) series than the Kekulé-like series (2). However, the average values of the coupling constants in the quinonoid series  $J_{4,5}$  8.45 and  $J_{5,6}$  6.78 Hz are closer to those of the Kekulé-series heterocycles and naphthalene than to cyclohexa-1,3-dienes: some relevant figures ( $J_{4,5}$  and  $J_{5,6}$ ) are: indole: 7.84, 7.07;<sup>34</sup> benzofuran: 7.89, 7.27;<sup>34</sup> benzothiophen: 8.09, 7.22;<sup>35</sup> naphthalene: 8.30 ( $J_{1,2}$ ), 6.83 ( $J_{2,3}$ );<sup>30</sup> cyclohexa-1,3-diene: 9.64, 5.04;<sup>30</sup> *cis*-5,6-dimethylcyclohexa-1,3-diene: 9.46, 4.95 Hz,<sup>36</sup> respectively. These similarities again suggest considerable aromatic character in both the 1- and 2-methyl series.

*Conclusions.*—Methylation of some simple 3- and 7-substituted indazoles in alkaline solution leads to mixtures of both 1- and 2-methyl compounds. In most cases the 1-isomer is predominant and even when large 7-substituents are present, the 1-isomer is still formed in significant amounts. The electron density calculations on the indazole and benzotriazole anions explain the electronic preference for N-1, but in 7-substituted cases it seems probable that reaction at that centre occurs *via* an out-of-plane attack, in contrast to N-2 where  $\sigma$ -attack seems probable under all conditions.

The n.m.r. spectra of the simple indazoles show generally only very small differences between the Kekulé-like 1-methyl series and the quinonoid 2-methyl series. Detailed analyses suggest that the *additional* dienic character of the C-4—C-7 system of the latter is quite small, and that both series are much closer to naphthalene than to cyclohexa-1,3-diene.

We thank J. A. Scott (*née* Wightman) and I. B. Muiry who carried out some of the reactions.

[5/398 Received, 26th February, 1975]

<sup>31</sup> P. J. Black and M. L. Heffernan, *Austral. J. Chem.*, 1964, **17**, 558.

<sup>32</sup> P. J. Wells, *Austral. J. Chem.*, 1964, **17**, 967.

<sup>33</sup> L. M. Jackman and S. Sternhell, 'Nuclear Magnetic Resonance Spectroscopy in Organic Chemistry,' Pergamon, Oxford, 1969.

<sup>34</sup> P. J. Black and M. L. Heffernan, *Austral. J. Chem.*, 1965, **18**, 353.

<sup>35</sup> K. D. Bartle, D. W. Jones, and R. S. Matthews, *Tetrahedron*, 1971, **27**, 5177.

<sup>36</sup> (a) J. B. Pawliczek and H. Gunther, *Tetrahedron*, 1970, **26**, 1755; (b) H. Gunther and H. H. Hinrichs, *Annalen*, 1967, **706**, 1.

# Ground State Wavefunctions for Aromatic and Heteroaromatic Molecules

M.H.Palmer, A.J.Gaskell, R.H. Findlay, S.M.F.Kennedy, W.Moyes and J.Nisbet\*

Near to Hartree-Fock wavefunctions have been obtained for a wide variety of benzenoid hydrocarbons and benzene derivatives, as well as 5- and 6-membered ring heterocycles containing varying numbers of first and second row elements. Trends in the computed properties derived from the wavefunctions are described with particular reference to the correlation with magnetic susceptibility and photoelectron spectroscopy. The effect of basis set upon molecular properties and energies within this range of molecules is discussed.

We have been interested in the ground state electronic structures of aromatics and heterocycles, and in their reactivity for some years [1]. Thus it is appropriate that we should have studied a range of these molecules by non-empirical means. In the early days we had very limited computing facilities, but with large amounts of time available on an IBM 360/50, we endeavoured to arrive at a compromise between size of molecule that could be studied and basis set that it would be practicable to use. The obvious starting point was the work of Clementi *et al.* [2] on pyridine and pyrrole etc., and subsequently that of Berthier *et al.* [3] on benzene and related species. We took the view that a wide number of studies comparatively close to the Hartree-Fock limit would be more valuable than just a few molecules very near the limit. Furthermore there appeared to be no case for just trying to marginally leapfrog the work cited above by minor extensions of the basis. In fact most one-electron properties for these molecules are virtually unaffected by change in basis set beyond our present level. We used in our studies [4] on the 5- and 6-membered ring heterocycles a best atom basis consisting of 7 *s*-type and 3 *p*-type for first row elements [5], with three *s*-type for hydrogen. Subsequent studies [6] showed that the addition of polarisation functions [7] lowered the energy slightly, but scaling of best atom sets to better represent molecular environments became widespread and we contributed to this with our work on small model molecules where the molecular energy was optimised [6]. Certainly if one's only desire is to lower the molecular energy obtained, then scaling is more cost effective than the addition of extra basis functions; it also lowers the valency shell binding energy and leads to better numerical agreement with photoelectron spectra when Koopmans'

Theorem is used. Scaling does not have any effect upon the orbital ordering, but in some cases it can effect the one-electron properties. For example using our best atom bases and evaluating the dipole moments leads to generally very good agreement with experiment, see table 1 [8]. When scaled bases are used the dipole moments often show a significant change, and since our work usually leads to random scatter near the experimental value this can lead to poorer results. Examples are furan (experimental 0.6D) where we obtain 1.0D for the scaled basis but 0.6D for the unscaled one, similarly 1,2,5-oxadiazole is worse on scaling. In contrast pyrrole is improved both by scaling and by the addition of polarisation functions [6]. Many other 1-electron properties are less sensitive than dipole moments to scaling.

In none of the cases that we have studied with both best atom and scaled bases have the orbital orderings been significantly changed. Furthermore, in carbocyclic and heterocyclic aromatic compounds it is true to say that virtually all orbital orderings become stable once one passes the threshold of reasonable size in minimal basis sets. By this we mean that extended bases offer little additional change. For example, in the azines with over 70 correlations, the ordering is unchanged between minimal and double zeta bases [9] in all except two cases; in these cases near degeneracy is observed in any event. The computational cost of extended bases can be rarely justified at this point in time. Amongst others, Lipscomb *et al.* [10] found that minimal Slater based calculations were as reliable as extended sets for many operators on boranes. Thus we decided to work with a firmly based single basis set for a very wide range of molecules – the scaled best atom basis. We now do upgrading to

\* Department of Chemistry, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JJ

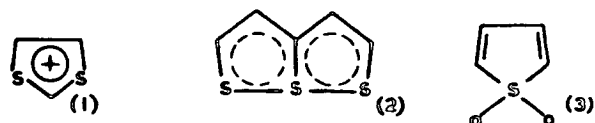
Table 1: Dipole moments ( $\mu/D$ ) and vector components in five-membered rings<sup>a,b</sup>

Name	Experimental Value	LCGO				
		Total $\mu$	$\mu_{\perp}$	$\mu_{\parallel}$	$\mu\sigma$	$\mu\pi$
Pyrrole	1.80	2.01	+2.01	0.0	-0.53	+2.54
Pyrazole	2.21	2.85	+2.23	+1.77	+1.65	+3.07
Imidazole	3.8	4.41	+4.31	+0.96	(-30.1°)	(+83.0°)
1H-1,2,4-Triazole	3.20	3.56	+3.50	+0.65	+1.46	+3.15
1H-1,2,3-Triazole					(+52.4°)	(+95.3°)
2H-1,2,3-Triazole					1.14	2.91
1H-1,2,3,4-Tetrazole		4.50	+3.26	+3.10	(+54.0°)	(+86.6°)
2H-1,2,3,4-Tetrazole	(2.30)	2.54	+2.24	+1.22	2.79	2.87
Furan	0.67	-0.64	-0.64	0.0	(+81.7°)	(+83.0°)
Thiophen ( <i>sp</i> basis)	0.53	-1.25	-1.25	0.0	1.57	2.71
( <i>sp</i> + 3 <i>s</i> ' + 3 <i>d</i> basis)		-0.44	-0.44	0.0	(-25.8°)	(+94.1°)
1,2,5-Oxadiazole	3.38	-2.96	-2.96	0.0	-2.77	2.13
1,3,4-Oxadiazole	3.04	2.75	+2.75	0.0	-3.22	1.97
1,2,4-Oxadiazole	1.2	1.18	-0.16	+1.17	-2.49	2.05
1,2,3-Oxadiazole		3.70	-0.22	+3.69	-4.46	1.50
					(-70.6°)	(+80.0°)
					-3.96	-2.16
					(-35.8°)	(+77.2°)
1,3,4-Thiadiazole ( <i>sp</i> basis)	3.28	3.38	3.38	0.0		
( <i>sp</i> + <i>d</i> basis)	3.28	4.23	4.23	0.0		
1,2,5-Thiadiazole ( <i>sp</i> basis)	1.57	-2.53	-2.55	0.0		
( <i>sp</i> + <i>d</i> basis)	1.57	-1.75	-1.75	0.0		
Phosphole (planar)		1.40	1.40	0.0	-0.70	2.23
Phosphole ( <i>PH</i> out-of-plane)		1.14	-0.54	1.00( $\mu_2$ )		
1,6,6a-Trithiapentalene ( <i>sp</i> basis)	3.01	-3.87	-3.87	0.0		
( <i>sp</i> + <i>d</i> basis)	3.01	-2.17				
1,6a,6-Dithiaoxapentalene ( <i>sp</i> basis)	3.78	-4.32	-3.78	-2.09		
( <i>sp</i> + <i>d</i> basis)	3.78	-3.70				
6a,1,6-Thiadioxapentalene ( <i>sp</i> basis)		-3.22	-3.22	0.0		
( <i>sp</i> + <i>d</i> basis)		-2.75				
1,6a,6-Dithia-azapentalene ( <i>sp</i> basis)		-4.23	-1.59	-3.92		
( <i>sp</i> + <i>d</i> basis)		-3.02				

(a) see [8]

(b) Angles with respect to positive direction of  $\mu_{\parallel}$  and measured anticlockwise; the sign of the dipole moment is taken with the negative end in the positive cartesian direction as positive

split valency shell or double zeta basis on a selective procedure in order to confirm that our hypothesis is reasonable, or to deal with particularly polar molecules. As a matter of practice all of our work with second row elements we include *d*-orbitals to see whether they are significantly populated in the ground state. This is uniformly not the case in all examples that we have studied where the element *S* or *P* is in a planar state, such as in the dithiolium salts (1) and trithiapentalenes (2) (thiathiophthens) [11]. The 3*d* orbitals are or course extensively utilised in the tetrahedral sulphur compounds, such as the thiophene mono- and di-oxides (3) [12].



Our interest in wavefunctions in the first instance has been to obtain:

- an interpretation of photoelectron spectra;
- to evaluate the moments and compare with experimental data;
- to evaluate the diamagnetic susceptibility and compare with experimental microwave data;
- to try to investigate that controversial subject in organic chemistry – aromatic character and resonance energy.

If it was possible to calculate the molecular *g*-value non-empirically then we would be able to calculate the total magnetic susceptibility. As we will see we have made progress on all of (a) to (d). We would also like to study the electronic spectra of the aromatic and heteroaromatic systems; we have done a little work on the first excited states, but again

within the single configuration procedure. One or two results of this are of particular significance, but in general the subject is not well described by this type of investigation so we will not cover this further.

Our work covers various systems. There is in fact surprisingly little work on conjugated olefins, particularly if floating spherical gaussian calculations are ignored. The justification for the latter view is not just based upon the fact that for molecules, such as hexatriene the FSGO energy is 33 au above our LCGO one, or that for naphthalene the difference is 54 au [13] but also that properties evaluated and (some) described below are unsatisfactory. We are working on both cyclic polyolefins including the annulenes, and acyclic systems. In the aromatic series we have an extensive study of molecules of type  $C_6H_4XY$  where  $X$  and/or  $Y$  are all the common substituents;  $H$ ,  $OH$ ,  $F$ ,  $NH_2$ ,  $NO_2$ ,  $CN$ ,  $CHO$  etc. As you will see this is an extensive triple triangular array since there are three isomers in each case. Although much of this work is complete, and we

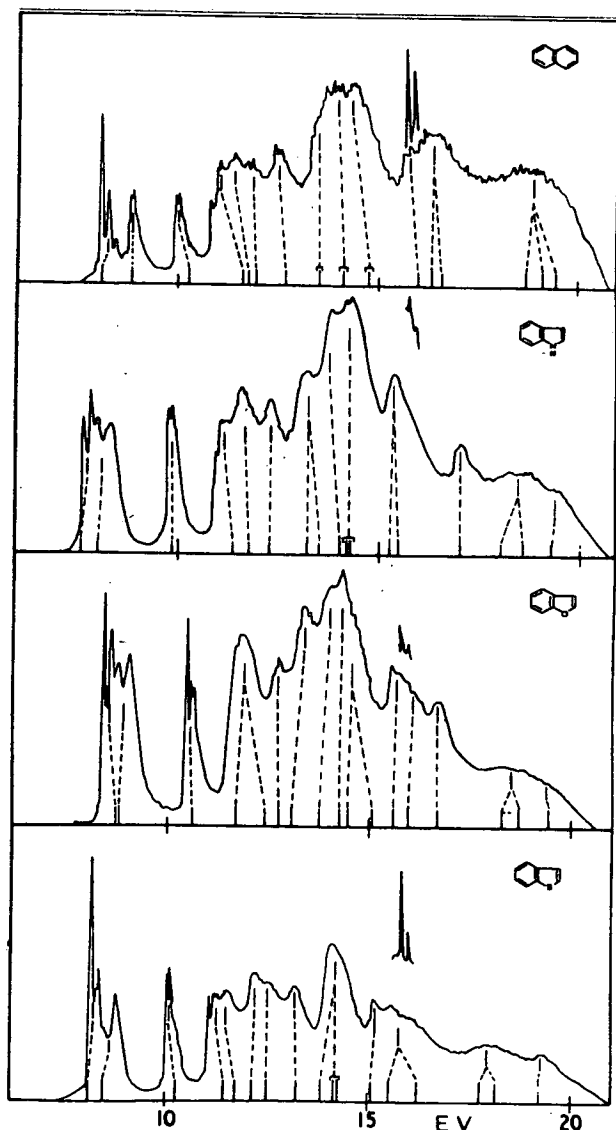


Figure 1

have assigned the photoelectron spectra – rather well as it turns out – we would not gain much from looking at extensive tables of data. From the fragmentary data, the correlation with experiment is not improved by larger basis sets, and the rather smaller bases by Pople and others [14] which lead to energies several au higher in energy are sure to be considerably poorer. Indeed we have repeated several of their computations and note that it is common to find eigenvalues in the 5-7 eV region. The other areas where we have made major studies are in the 6-membered ring heterocycles containing various numbers of nitrogen atoms with or without one atom of sulphur or phosphorus. Again the photoelectron data is very good. These studies are practicable with quite large systems as seen for indole, benzofuran, benzothiophene and naphthalene, see figure 1.

We have already referred to the dipole moment studies on heterocycles when discussing the subject of basis sets. In the substituted benzenes the conformations of unsymmetrical substituents ( $X$  or  $Y$ ) in the system  $C_6H_4XY$  bring about additional problems. We are active in this and have endeavoured to determine the most favoured geometry on the basis of energy, and then compare the dipole moment with the experimental one. Thus we find that in the case 1,2- $C_6H_4(NO_2)_2$  the best geometry is probably one in-plane and another perpendicular. These results are not complete and we don't want to commit ourselves to the final answer. For many simple aromatics we calculate a dipole moment relatively close to the experimental one. We reproduce trends within the isomers and the effects do show vector addition for substituents as has long been known experimentally. The results are not perfect and this can be attributed to several factors:

- the geometry of the ring; if not known from a gas phase determination this was taken as that of benzene;
- that the energy optimised scaling on oxygen and fluorine compounds generally leads to poorer dipole moments than best atom bases for these systems;
- that some of the dipole moments are anyway very high and this really calls for additional model compounds for the scaling.

One of the very controversial subjects of organic chemistry has been resonance and resonance energy. We feel that we may have made a contribution here. It is clear that one of the virtues of keeping to a standard basis is that small trends become apparent unexpectedly. Thus in trying to evaluate resonance energy one is in difficulty with a definition and a classical analogue on which to measure heats of hydrogenation etc. Strain energy is very important and so on. We felt that the results of Hartree-Fock studies are unlikely to lead to good heats of formation. We need to return to the classical definition of

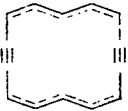
resonance energy if we are to study the subject at all. It was apparent to us that if we take a series of small molecules, the total energies are additive in bonds, table 2; so we propose to use the thermochemists approach, but *not* to equate the atom energies to zero. Thus the energy of the *C-H* bond  $E_{C-H}$  is  $E_{CH_4}/4$ , that of the  $E_{C=C}$  is obtained from  $E_{C-H}$  and the energy of ethylene. In this way, with a series of classical molecules we arrive at a set of bond energies, which are of course on a very different energy scale to those of the thermochemist. Firstly we observed that they work for a number of non-aromatic systems, and then we apply them, table 3, to conjugated molecules.

pyrrole is also reasonable, and the fusion of the rings with benzene even more so; thus while benzofuran (or indole) are only about the same in aromaticity as benzene, the iso series are markedly less. The result with hexa fluorobenzene is a direct result of trying to obtain the *C-F* bond energy from vinyl fluoride. We are unable to stop the  $\pi$ -systems mixing, and hence some interaction energy is being lost in  $C_6F_6$ . However, even using the *C-F* bond energy from  $CH_3F$  fails to alter the low resonance energy. Thus we must either accept the value, or question whether the *C-F* length (we used that from  $C_6H_5F$ ) was optimal in  $C_6F_6$ . All of the systems so far are nearly strain free; if we operate

Table 2: Analysis of molecular total energy in aromatics

Model Molecules (-Energy au)					
$CH_4$	40.10180	$NH_3$	56.0199	$H_2O$	75.79999
( <i>CH</i> )		( <i>NH</i> )		( <i>OH</i> )	
$C_2H_4$	77.83143	$CH_3CH=CH_2$	116.77453	$CH_2=CH-CH=CH_2$	154.50920
( <i>C=C</i> )		( <i>C-C</i> )		Perpendicular ( <i>C-C</i> )	
$CH_2=CH-NH_2$	132.70390	$CH_2=CH-OH$	152.46202	$CH_2CH-F$	176.40569
Perpendicular ( <i>C-N</i> )		Perpendicular ( <i>C-O</i> )		( <i>C-F</i> )	
$CH_3CN$	131.55927	$CH_2=O$	113.51009		
( <i>C-N</i> )		( <i>C=O</i> )			

Table 3: 'Resonance energies' (kcal/mole)

$C_6H_6$	$C_{10}H_8$	<i>PhF</i>	<i>PhCHO</i>	<i>PhOH</i>	<i>PhNH_2</i>	$C_6F_6$
50.8	85.4	50.7	53.7	43.9	48.6	16.6
Furan	Pyrrole					
21.2	21					
Benzofuran	Isobenzofuran	Indole	Isoindole			
55.6	35.2	56.7	43.3			
						
						77.7
<i>trans</i> Butadiene	<i>cis, cis</i> Hexatriene	Cyclooctatetraene	Barrelene			
5.5	7.8	(experimental geometry)	-114			
		6.1				

As is seen from the table we here quote just a few examples. The figures for the hydrocarbons, ethylene 0.0, butadiene 5.5, hexatriene 7.8, cyclopentadiene 16.3, benzene 50.8 kcal/mole follow a logical pattern. Naphthalene is less than twice benzene in agreement with thermochemical data. The substituted benzenes (with the exception of hexafluoro benzene to be described elsewhere) are normally in the 55-40 kcal region. In most cases adding  $\pi$ -electrons increases the resonance energy above that of benzene. The data for furan and

with strained systems then we must expect to obtain strange results. We have been studying other members of the  $C_nH_n$  series for various charged states; for cyclobutadiene the best square triplet (Hückel) state is preferred over the best square or rectangular singlet and neither has any resonance energy (i.e. resonance energy positive). Thus they are antiaromatic even before strain energy is included. Planar alternating cyclooctatetraene would have a resonance energy of -8 kcal/mole, so it is antiaromatic anyway, but the strain energy has not been



Table 4: Bond population moments

$\sigma$ System							
Azines ( $X^{\delta+} - Y^{\delta-}$ )							
$X-Y$	$H-C$	$C-N$	$N-N$	$C-C$			
	0.241	0.121	0.004	0.007			
Number of Points	29	30	7	11			
Maximum deviation	0.023	0.015	0.003	0.003			
Azoles $X^{\delta+} - Y^{\delta-}$							
$X-Y$	$C_{\alpha}-O$	$N_{\alpha}-O$	$C_{\alpha}-N(H)$	$N_{\alpha}-N(H)$	$C_{\alpha}-N_{\beta}$	$C_{\beta}-N_{\alpha}$	$N_{\beta}-N_{\alpha}$
	0.317	0.200	0.244	0.133	0.070	0.159	0.034
Number of points	7	4	8	6	6	6	4
Maximum deviation	0.010	0.070	0.015	0.016	0.018	0.026	0.019
$X-Y$	$H-N$	$C_{\beta}-N_{\beta}$	$N_{\beta}-N_{\beta}$				
	0.386	0.115	0.007				
Number of points	7	6	2				
Maximum deviation	0.050	0.004	0.007				
$\pi$ System Azines							
$X-Y$	$C-N$	$C-C$	$N-N$				
	0.004	0.008	0.012				
Number of points	30	11	7				
Maximum deviation	0.03	0.010	0.010				
Azoles ( $X^{\delta+} - Y^{\delta-}$ )							
$X-Y$	$O-C_{\alpha}$	$(H)N-C_{\alpha}$	$(H)N-N_{\alpha}$	$O-N_{\alpha}$	$C_{\alpha}-C_{\beta}$	$N_{\alpha}-C_{\beta}$	
	0.142	0.195	0.219	0.148	0.103	0.087	
Number of points	6	8	6	3	8	6	
Maximum deviation	0.014	0.023	0.030	0.030	0.026	0.020	
$X-Y$	$C_{\alpha}-N_{\beta}$	$N_{\alpha}-N_{\beta}$	$C_{\beta}-C_{\beta}$	$N_{\beta}-C_{\beta}$	$N_{\beta}-N_{\beta}$		
	0.107	0.089	0.0	0.006	0.0		
Number of points	6	4	1	6	1		
Maximum deviation	0.029	0.030	0.0	0.008			

included. In the nearly strain-free experimental geometry (which has a dihedral angle of about  $50^{\circ}$ ) there is still 6.1 kcal/mole of resonance energy showing that there is still an interaction between the olefinic systems, as is clear from photoelectron spectra of this type of system.

The Mulliken population analysis of heterocycles might be expected to lead to net atomic populations which vary markedly

- from molecule to molecule, and
- with basis set.

From a comparison of our minimal data with larger basis set calculations we find that this is far from the case; it is just the relative populations in the constituent atomic functions which differ. For example if our thiopene data is compared with Gelius *et al.* [15] the direct transference of our  $3d_{\sigma}$  to their  $3s$  leads to very similar results. We then noted that the populations at the atomic centres could be separated into bond population moments, see table 4, and that these are almost constant, with the sole exceptions of  $CH$  and  $NH$  which act as electron sinks for the whole system [8].

A further idea consistent with basic organic chemical notions is of the aromatic sextet behaving as a sextet i.e. having the same average position.

If the average positions of the electrons vary markedly then the  $\pi$ -system is not behaving as a unit, and this should lower the aromatic character. Studies of this both in mono and bicyclic systems are very revealing [8,12,13]. Of course there are difficulties in applying this approach to molecules with a centre of symmetry.

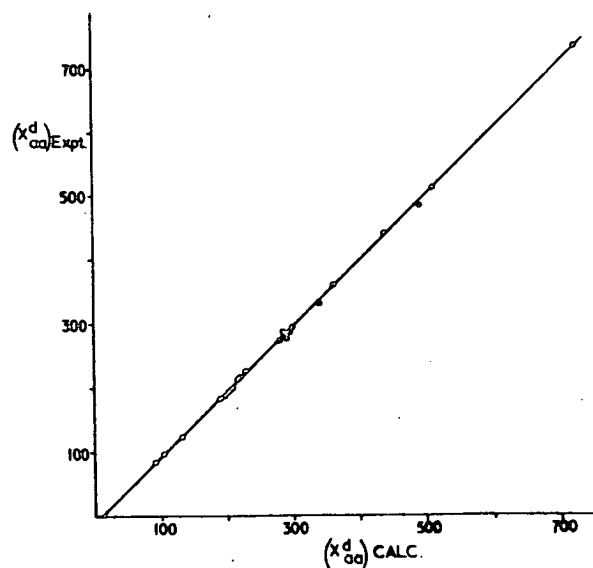


Figure 2: Diamagnetic susceptibility

Lastly we want to mention that we are able to compute the diamagnetic susceptibility, figure 2, for carbocyclic and heterocyclic aromatics within the experimental accuracy in almost all cases [16]. Whether the total magnetic susceptibility anisotropy can be regarded as a characteristic of aromaticity remains unproven. Certainly Flygare's own data, figures 3 and 4, shows that the out of plane term is related to the average in plane irrespective of whether the system is aromatic or not. Our work on the diamagnetic term in annulenes should provide some interesting results here. Preliminary results for some Hückel annulenes suggest that they are showing strong diamagnetic susceptibility anisotropy, but that the paramagnetic term less the  $g$ -factor portion already makes the sign of the sub-total opposite to that required by Flygare's proposal. There appears to be a linear correlation between the computed binding energy and the diamagnetic susceptibility anisotropy in isoelectronic series, see figure 5.

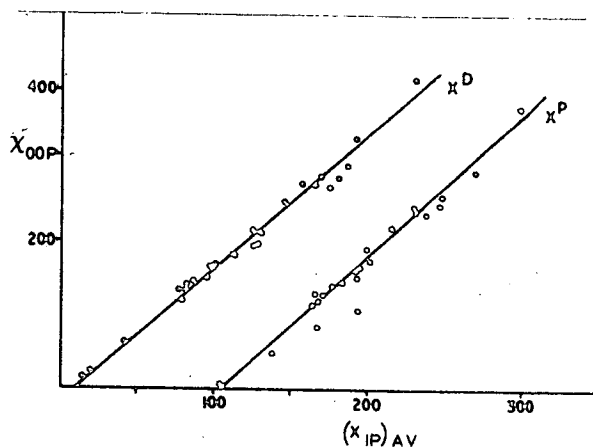


Figure 3: Magnetic susceptibility of acyclic systems (Flygare)  $X_p$  offset

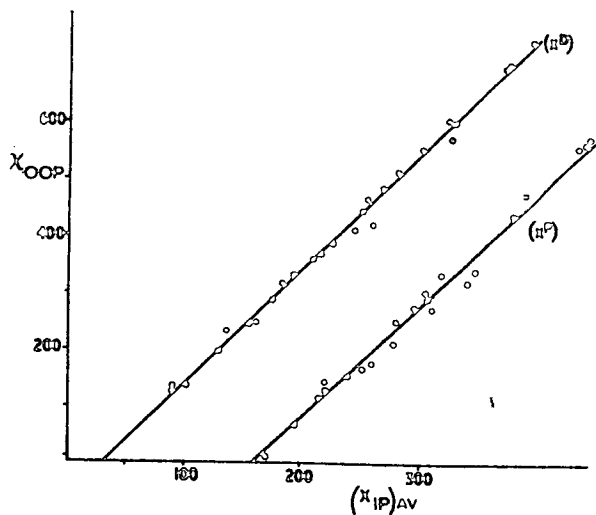


Figure 4: Magnetic susceptibility of cyclic systems (Flygare)  $X_p$  offset

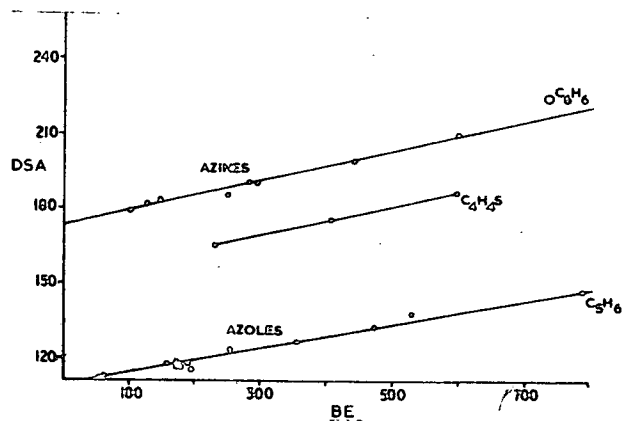


Figure 5: Diamagnetic susceptibility anisotropy and binding energy

### References

- [1] PALMER, M.H. (1967). *Structure and Reactions of Heterocyclic Compounds*, London: Arnold.
- [2] CLEMENTI, E. (1967). *J. Chem. Phys.*, 46, 4725, 2731, 2737.
- [3] BERTHIER, G., PRAUD, L. and SERRE, J. (1969). *Quantum Aspects of Heterocyclic Compounds in Chemistry and Biochemistry*, Israel Academy of Sciences and Humanities Symposium, 40, New York: Academic Press.
- [4] PALMER, M.H. and GASKELL, A.J. (1971). *Theoret. Chim. Acta*, 23, 52.
- [5] ROOS, B. and SIEGBAHN, P. (1970). *Theoret. Chim. Acta*, 17, 209.
- [6] PALMER, M.H., GASKELL, A.J. and BARBER, M.S. (1972). *Theoret. Chim. Acta*, 26, 357.
- [7] ROOS, B. and SIEGBAHN, P. (1970). *Theoret. Chim. Acta*, 17, 199.
- [8] PALMER, M.H., FINDLAY, R.H. and GASKELL, A.J. (1974). *J. Chem. Soc. (Perkin II)*, 420.
- [9] ALMLOF, J., ROOS, B., WAHLGREN, U. and JOHANSEN, H. (1973). *J. Electron Spectroscopy*, 2, 51.
- [10] PALMER, M.H., GASKELL, A.J. and FINDLAY, R.H. (1974). *J. Chem. Soc. (Perkin II)*, 778.
- [11] LAWS, E.A., STEVENS, R.M. and LIPSCOMB, W.N. (1972). *J. Am. Chem. Soc.*, 94, 4661.
- [12] PALMER, M.H. and FINDLAY, R.H. (1971). *Tetrahedron Letters*, 4165.
- [13] ——— and ——— (1974). *J. Chem. Soc. (Perkin II)*, 1885.
- [14] ——— and ——— (1974). *J. Chem. Soc. (Perkin II)*, (in the press).
- [15] PALMER, M.H. and KENNEDY, S.M.F. (1974). *J. Chem. Soc. (Perkin II)*, 1893.
- [16] HEHRE, W.J., RADOM, L. and POPLE, J.A. (1972). *J. Am. Chem. Soc.*, 94, 1496.
- [17] GELIUS, U., ROOS, B. and SIEGBAHN, P. (1972). *Theoret. Chim. Acta*, 27, 171.
- [18] PALMER, M.H. and FINDLAY, R.H. (1974). *Tetrahedron Letters*, 253.