NON-EMPIRICAL MOLECULAR ORBITAL CALCULATIONS

RELATED TO

AROMATIC AND HETEROAROMATIC SYSTEMS

by

JOHN D. NISBET

UNIVERSITY OF EDINBURGH

Ph.D. 1978

DECLARATION

I declare that I have written this thesis and performed the work described in it. Any result not obtained directly by me but used in the discussion has been clearly indicated.

John D. Nisbet

ABSTRACT

A large program of ab initio MO studies of conjugated molecules has been completed. A major point of interest was the molecular geometry in a number of planar and non-planar cases. The method was first applied to small molecules of known geometry, and then to large systems including cyclo-octatetraene, the longer annulenes, and various 7- and 9- membered ring heterocycles.

Contents

LAYOUT

For convenience in reading the text, where a particular Table is often cited in many places, the Tables are gathered at the end of each chapter, immediately beyond each set of references.

CHAPTER 1

INTRODUCTION

"Experiments are the only means of knowledge at our disposal. The rest is poetry, imagination."

Max Planck

The subject matter of this thesis can be regarded as being part of the branch of Theoretical Chemistry known as the 'Theory of Valence'[1]. However, it must immediately be emphasised that Theory and Experiment should never be considered separate domains of chemistry. Perhaps, the rôle of theory in chemistry is best described as the provision of a framework in which to organise meaningfully experimental knowledge and collate experimental facts[2]. Consequently, although the results presented in this work have been obtained mainly from theoretical considerations, relevant experimental findings have been included in subsequent discussion.

The interplay between theoretical ideas and experimental findings has been evident in the development of the subject of chemical valence. 'Classical' conceptions of the nature of chemical combination arose during the nineteenth century, with the atomic theory of matter as foundation; modern theories have replaced the old-established ones, with the development of modern physics, progress depending critically on the excellently laid foundations of 'pure' (experimental) chemistry. Current opinion originates from the birth of 'Quantum Theory' in physics in 1900, although the chemical aspect really dates from 1916 when the electronic theory of valence first made its impact through the efforts

of Kossel, Lewis, and Langmuir using contemporary ideas of atomic electronic structure[3]. Actually, Lewis' work, which forms the basis of the modern electronic theory of valence,was the culmination of numerous attempts to develop an electronic theory of the chemical bond, with the only really essential ingredient from physics being the existence of the newly-discovered electron (Thomson, 1895-97); complementary advances in theory in the early twentieth century, particularly Bohr's model of the electronic structure of the atom proposed in the interpretation of atomic spectra (1913-14) and deduced by combining Rutherford's concept of the nuclear atom (1911) with Planck's quantum theory, were not really helpful in elucidating the nature of the chemical bond. Thus, by about 1920, a chemical theory of valence had been produced[4], but it was only remotely connected to physicists' ideas of the structure of matter. The evolution of a single, unifying theory of chemical bonding had to await the development of Quantum Mechanics, which provided the theoretical framework used as a guide in understanding and predicting chemical phenomena.

Although the theory developed by Lewis, Langmuir and others coordinated a large body of chemical facts, it was a purely formal, "classical" description and was inadequate primarily because it lacked an intimate description of the dynamic behaviour of electrons and, consequently, of the chemical bond. With the advent of quantum mechanics, particularly in the wave mechanical formalism of the theory of the electron, the theory of valence made enormous progress. Thus, after 1926, quantum mechanica theory replaced the old quantum theory, associated with Planck, Einstein, Bohr, Sommerfeld and others, as the foundation of modern chemistry, in particular the electronic structure of atoms and molecules and the nature of chemical union. The account of molecular structure provided by Schrödinger's wave mechanics forms the basis of the theoretical method used in this work to produce results of chemical interest. It must be

emphasised that only a small part of the contribution of quantum mechanics to structural chemistry has been purely quantum-mechanical; the advances which have been made have been in the main the result of essentially chemical arguments - postulates are made and tested by empirical comparison with available chemical information, and used in the prediction of new phenomena. Thus, the position today is that the general principles of molecular structure and the nature of chemical valence are formulated on the basis of deduction from quantum mechanical theory, in contrast to induction from chemical facts, as was the position from the time of the proposal of the concept of valence in 1852 by Frankland up to 1926.

In modern physics, the fundamental discipline of Quantum Mechanics has become an extremely well-documented subject in the fifty years of its existence; numerous textbooks have been written, expounding all aspects of quantum theory, addressed mainly to physicists (and mathematicians)[5], but there are several excellent volumes intended for chemists[6]. Quantum Chemistry as a subject has evolved as a form of applied quantum mechanics[7], and is sometimes now regarded as being synonymous with Theoretical Chemistry[8]. Thus, quantum mechanics provides the physical principles and mathematical methods used in all theoretical work on the electronic structure and properties of matter[9]; of particular interest in this work is the application of quantum theory to valency problems in structural chemistry[10]. The basic principles of valence, which explain the details of molecular structure, in modern wave-mechanical terms are familiar to chemists, much of the subject matter having 'settled down' into a coherent theory[1], which forms an essential background to the discussion of results in this work. Similarly, nowadays, chemists have some acquaintance, at least, with the general concepts of the inherently mathematical discipline of quantum mechanics, normally with the picture portrayed by wave mechanics whereby

practical chemical problems are treated by application of the Schrödinger Wave Equation[11]. Thus, assuming the well-established principles of the quantum theory of valence and molecular quantum mechanics as detailed in various references cited above, there is presented in Chapter 2 a derivation of the particular theoretical method of calculation of electronic structures of molecules used in this work from the very general formal theory; relevant texts and references on this method of 'modern' quantum chemistry are relatively unfamiliar to chemists, but it must be noted that it is only the practical details of computation which are recent, the basis of the method being a long-standing one[12].

In summary, the aim of the project reported in this thesis was to perform a theoretical, quantum-mechanical examination of chemically interesting molecules to produce results to aid the understanding of chemical phenomena, using a specialised mathematical and physical model; in particular, the work is essentially in the area of physical organic chemistry.

References for Chapter 1

1. C.A. Coulson, "Valence", 2nd edition, Oxford University Press, London, 1961; J.N. Murrell, S.F.A. Kettle, and J.M. Tedder, "Valence Theory", 2nd edition, Wiley, London, 1970.

2. R. Hoffmann, Chemical and Engineering News, 52, 32-34 (1974), in a series of essays on the subject: "1774-1974, 200 Years Since Alchemy - Some Thoughts on Chemistry Today".

3. W. Kossel, Annalen der Physik, 49, 229 (1916); G.N. Lewis, Journal of the American Chemical Society, 38, 762-785 (1916); I. Langmuir, ibid, 41, 868-934 and 1543-1559 (1919).

4. G.N. Lewis, "Valence and the Structure of Atoms and Molecules", Chemical Catalog Co., New York, 1923 (Dover, New York, 1966); N.V. Sidgwick, "The Electronic Theory of Valency", Clarendon Press, Oxford, 1927.

5. R. McWeeny, editor, "Classical and Quantum Mechanics", topic 2 of The International Encyclopedia of Physical Chemistry and Chemical Physics, Pergamon, Oxford (editors D.D. Eley and F.C. Tompkins) - intended for 'chemical physicists'.

6. L. Pauling and E.B. Wilson Jr., "Introduction to Quantum Mechanics, With Applications to Chemistry", McGraw-Hill, New York and London, 1935; W. Heitler, "Elementary Wave Mechanics, With Applications to Quantum Chemistry", 2nd edition, Oxford University Press, London, 1956; H.L. Strauss, "Quantum Mechanics: An Introduction", Prentice-Hall, New Jersey, 1968; D. Rapp, "Quantum Mechanics', Rinehart and Winston, New York, 1971.

7. H. Eyring, J. Walter and G.E. Kimball, "Quantum Chemistry", Wiley, New York, 1944; K.S. Pitzer, "Quantum Chemistry", Constable, London, 1953; W. Kauzmann, "Quantum Chemistry: An Introduction", Academic Press, New York, 1957; F.L. Pilar, "Elementary Quantum Chemistry", McGraw-Hill, New York and London, 1968; I.N. Levine, "Quantum Chemistry", 2nd edition, Allyn and Bacon, Boston, 1974; R. Daudel, B. Pullman (editors), "The World of Quantum Chemistry. Proceedings of the First International Congress of Quantum Chemistry, Menton, France, 1973", Reidel, Boston and Dortrecht, Holland, 1974.

8. Specialist Periodical Report, "Theoretical Chemistry. Volume 1- Quantum Chemistry", The Chemical Society, London, 1974 (Volume 2, 1975); International Review of Science, "Theoretical Chemistry" (Physical Chemistry, Volume 1 - Series 2, 1975) , editors A.D. Buckingham and C.A. Coulson, Butterworth, London; "Modern Theoretical Chemistry", editor H.F. Schaefer, Plenum, New York, 1976.

9. J.C. Slater, "Introduction to Chemical Physics", McGraw-Hill, New York, 1939 (Dover Publication, 1971);

"Quantum Theory of Matter", 2nd edition, 1968;

"Quantum Theory of Atomic Structure", volumes I and II, 1960;

"Quantum Theory of Molecules and Solids", volume I, 1963, volume II, 1965, volume III, 1969, volume IV, 1974;

P.-O. Lowdin (editor), "Quantum Theory of Atoms, Molecules, and the Solid State", Academic Press, New York and London, 1966 (A Tribute to John C. Slater).

10. J.W. Linnett, "Wave Mechanics and Valency", Methuen, London, 1960; R.G. Parr, "Quantum Theory of Molecular Electronic Structure", Benjamin, New York, 1964; H.F. Hameka, "Quantum Theory of the Chemical Bond", Hafner, New York, 1975.

11. P.W. Atkins, "Quanta: a Handbook of Concepts", Clarendon Press, Oxford, 1974.

12. D.R. Hartree, Proc. Camb. Phil. Soc., 24, 89, 111, 426 (1927);
V. Fock, Z. Physik, 61, 126 and 62, 795 (1930);
J.C. Slater, Physical Review, 35, 210 (1930).

CHAPTER 2

THE THEORETICAL METHOD

"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble."

P.A.M. Dirac

Quantum mechanics in its well-established present form can provide an accurate description of the behaviour of electrons and nuclei in atoms and molecules, on which chemical properties ultimately depend. New fundamental changes in the framework of the theory may be required to rationalise 'high-energy' phenomena in physics; even so, conclusions on the atomic level are most unlikely to be affected. However, the assertion that quantum mechanics (non-relativistic theory) contains the solution of all chemical problems, which are reduced to unambiguous applied mathematical ones, is unverifiable because insuperable difficulties in practice arise from the lack of sufficiently powerful mathematical techniques. Moreover, there is an even greater handicap in that there is not sufficient correlation between quantum mechanical concepts and empirical chemical ones, such as valence and the chemical bond. The result is that these inherent limitations of both a practical and a theoretical nature, have rendered the application of quantum mechanics to chemistry only qualitative, in effect, and chemical theory embodies much empirical reasoning by necessity[1]. An extremely wide gap has opened between the descriptive theoretical chemistry used in chemistry textbooks and the computational quantum chemistry of present-day research; of particular importance here is the area of valency and

chemical bonding, and molecular electronic structure[2,3]. The aim of this chapter is to detail the practical method used to calculate molecular electronic structure and resulting properties, proceeding from extracts of some well known aspects of the general framework of quantum theory to less publicised ones, which have become of real use only in recent times; quantum mechanics is now becoming a useful tool in modern chemistry, as well as providing a language to systematically describe chemical phenomena. It is now intended to present an account of a procedure which has enabled the computation of molecular electronic wavefunctions to become a "routine" service, and then to show how such (approximate) wavefunctions can be analysed to produce results of chemical significance.

A. Calculation of the Electronic Wavefunction of a Molecule.

1. Quantum Theory of Atomic Structure

Although the electronic structure and properties of molecules is the essence of this work, it seems natural to consider atoms first; however, the presupposition of the existence of atoms is not necessary for the application of quantum mechanics to molecular structure but is certainly essential for any meaningful chemical information to be obtained.

(a) The Bohr Model of The Atom

The development of atomic theory gives an illustration of the interaction between experimental findings and theoretical ideas required to construct an accurate model. Atomic spectroscopy, since its inception

in 1861 by Kirchoff and Bunsen, has always provided a wealth of experimental information. One classic example of such investigation was the discovery, in 1885, of the series of lines in the spectrum of atomic hydrogen that now bears the name of Balmer, who formulated the following empirical relationship giving the positions of all the lines:

$$1/\lambda = R\left(\frac{1}{2^2} - \frac{1}{n_1^2}\right) \qquad n_1 = 3,4,5,..... \qquad - (1)$$

where λ is the wavelength of the observed line and R is a constant (the Rydberg Constant, value approx $1.1 \times 10^{8} m^{-1}$).

Atomic spectra were soon recognised as providing the key to the understanding of atomic structure, as long as their interpretation was possible. In the early years of this century, after the discovery of the electron by Thomson in 1897 and Rutherford's enunciation of the nuclear model of the atom in 1911, it was found that the application of classical physics to the problem of the electronic structure of atoms produced a picture with some gross inadequacies; in particular, the classical model of the atom predicts that atomic spectra are continuous, rather than discrete, as indicated by equation (1) for part of the hydrogen atom spectrum. To try to remove discrepancies between theory and experiment, in 1913 Niels Bohr postulated nonclassical "stationary states" of an atomic system - the combination of classical laws of motion to describe electron motion and certain arbitrary quantum conditions (idea based on Planck's hypothesis developed in another area of physics) led to the existence of discrete sets of electron energies (states). Thus, for the physical system consisting of an atom, the following expressions can be derived, assuming the quantisation of the angular momentum of the electron so that

$$p = n.\hbar \qquad -(2)$$

(p = angular momentum, $\hbar$ = Planck's constant/2π, n is an integer)

$$r = \frac{n^2\hbar^2}{me^2Z} \qquad -(3)$$

$$E_n = T + V = \frac{-mZ^2e^4}{2n^2\hbar^2} \qquad -(4)$$

(m = mass of electron, e = charge on electron, Z is atomic number of atom). 'r' is the radius of allowed electron orbits; 'E_n' is the total energy of the atom, the sum of kinetic (T) and potential (V) energies, only discrete values being allowed. The Bohr model, later refined principally by Sommerfeld, provided a completely satisfactory account of the spectrum of atomic hydrogen to a high level of precision; the spectrum can be summarised by application of equation (4) to describe electronic transitions between discrete states:

$$1/\lambda = \frac{2\pi^2me^4}{h^3c}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right) \qquad -(5)$$

(n_1,n_2 are integers).

With $n_2 = 2$, equation (5) yields the theoretical version of Balmer's empirical formula, equation (1); setting $n_2 = 1,3,4,5$ gives other series of lines which correlate with experimental findings.

Bohr's rationale outlined above was successful in some respects, but it failed the ultimate test for a theory in that it conflicted with experiment; it was also rather unsound to theoretical physicists. The model failed to explain the observed energy levels of any atom with two or more electrons. However, of more importance to chemistry was the failure to provide a basis for understanding bonding and valence.

Thus, the Bohr atom was significant because it was the first attempt at a model recognising the need for a departure from classical physics to describe the dynamic character of electrons in atoms; concepts and terminology from this model are still significant today, although the picture of the atom is now based on wave mechanics[4].

(b) Wave-Mechanical Model of the Atom

The calculation of the electronic structure of matter rests on Schrödinger's formulation of the system of mechanics known as "wave mechanics", a synthesis of Hamilton's formal analogy between mechanics and optics and of De Broglie's concept of wave systems associated with material particles (wave-particle duality). Thus, electronic and molecular motions are treated as waves on a basis of classical physics; the mathematical formalism of this treatment is also based on methods developed long before the advent of quantum theory. Wave mechanics is a specific case of a more general theory, quantum mechanics, and is best suited for application to chemistry. The Rutherford-Bohr picture of the atom is modified, rather than completely replaced, by application of wave mechanics which incorporates the lack of precision of the description of an electron's dynamical motion into a basically classical model. The new concepts in the model of the atom are the wave character of the electron and the statistical (probability) character of the knowledge of physical systems.

For the physical system consisting of an atom, or a molecule, the basic premise is that the state of such a system is completely described by a wave function, the mathematical properties of which can be translated into physical ones; more particularly, the electronic wavefunction is the chemist's goal. The starting-point in this aim is generally the celebrated Schrödinger Wave Equation of non-relativistic wave mechanics:

for a single-particle system,

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\underline{r},t) + \Phi(\underline{r})\Psi(\underline{r},t) = i\hbar\frac{\partial\Psi}{\partial t} \qquad -(6)$$

$\Psi(\underline{r},t)$ is the wave function (of space coordinates $\underline{r}$ and time t),

m is the mass of the particle (electron)

$\hbar$ is Planck's constant divided by 2π,

$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplace operator,

$\Phi(\underline{r})$ is the potential energy function for the system,

$i \equiv \sqrt{(-1)}$.

In the application to electronic structure calculations, the variation of the wave function in time is not connected to any change in the physical state of the system (atom or molecule), which is completely characterised by a function of "$\underline{r}$" only - often referred to as the wave function for the "stationary state". Thus, removing explicit time dependence gives the time-independent Schrödinger equation:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + \Phi(\underline{r})\right] u(\underline{r}) = Eu(\underline{r}) \qquad -(7)$$

$u(\underline{r})$ is the time-independent wave function,

E is a constant, identified with the total energy of the stationary state

More compactly, equation (7) becomes $\hat{H}u = Eu$ $\qquad -(8)$

where $\hat{H}$ is the Hamilton operator (Hamiltonian).

Quantities which are observables in classical mechanics correspond to operators in the mathematical formalism of quantum mechanics so that the classical Hamiltonian of a system, H = T + V (sum of kinetic energy, T, and potential energy, V), corresponds to the quantum mechanical Hamiltonian

$\hat{H} = \hat{T} + \hat{V}$, where $\hat{T}$ is the kinetic energy operator for the system $(-\frac{\hbar^2}{2m}\nabla^2)$ and $\hat{V}$ is the potential energy operator $(\Phi(\underline{r}))$. The 'system' described by $u(\underline{r})$ above is now simply the electron.

Proceeding from the very brief outline of abstract formalism given above, it is instructive to apply the foundation equation (8) to the case of the hydrogen atom, or one-electron atom in general; the solution of equation (8) for this physical situation is essential for the physical understanding of more complex, chemically significant cases.
Thus, for the hydrogenic atom of figure 1, equation (8) becomes

$$(-\frac{\hbar^2}{2\mu}\nabla^2 - \frac{Ze^2}{r})u(\underline{r}) = Eu(\underline{r}) \qquad -(9).$$

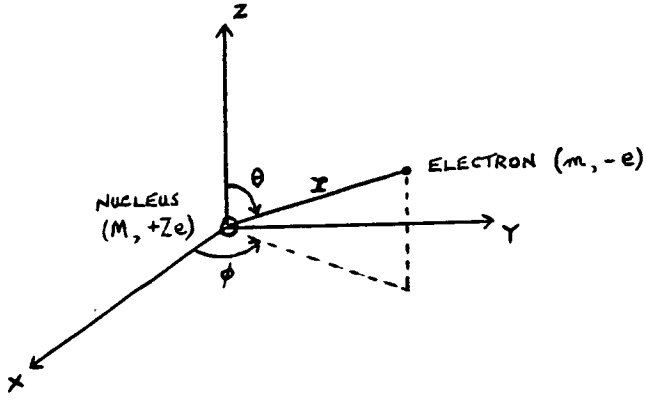


FIGURE 1

Quantities used in the discussion of the hydrogenic atom. The origin should be at the centre of mass, but little error is made if it is put at the nucleus.

For the system depicted in figure 1, $\hat{V} = -\frac{Ze^2}{r}$ and $\hat{T} = -\frac{\hbar^2}{2(M+m)}\nabla^2$. The form given in equation (7) for the time-independent wave equation is actually a particular one, although it is sufficient even for general molecular electronic wavefunction calculation. The complete $\hat{H}$ for the

hydrogenic atom is thus $-\frac{\hbar^2}{2(M+m)}\nabla^2 - \frac{Ze^2}{r}$, including the nuclear mass (M) and charge (Ze) and the electron mass (m) and charge (e). However, the resulting complete wave equation is separable into two equations, one describing the translational motion of the atom as a whole and the other describing the "internal motion"; these can be solved independently, and, in fact, only the latter is of real significance. The appropriate Hamiltonian operator is then

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla^2 - \frac{Ze^2}{r} \qquad -(10)$$

where $\mu = \frac{mM}{m+M}$ is the reduced mass of the system.

Thus, the single-particle equation (7) retains its significance, with μ replacing m as a very minor modification. Returning to equation (9) and figure 1, it is indicated that the natural coordinate system to use is one consisting of spherical polar coordinates, so that $\underline{r} = (r,\theta,\phi)$ and the Laplace operator ∇^2 is to be transformed from its cartesian representation given above[5]. With the problem restated in an internal coordinate system, equation (9) can be solved exactly, analytically, by well-established classical mathematical methods. The quantal characteristics of the solutions arise from constraints imposed from physical reasoning; these so-called "boundary conditions" for the problem are that the wave function should be everywhere finite, single-valued, and should have integrable squares, such restrictions arising from the probability interpretation inherent to wave mechanical descriptions of physical systems. Thus, for a solution representing a stationary state, $u(\underline{r})$ must tend to 0 as $r \to \infty$ and $\int_{\text{all space}} u^*(\underline{r})\,u(\underline{r})\,d\tau$ must exist ($u^*(\underline{r})$ is complex conjugate of $u(\underline{r})$). In particular, if $u(\underline{r})$ is a solution of equation (9), then so is any constant multiple of $u(\underline{r})$, and

so conventionally the solution is taken to be "normalised", meaning $\int_{\text{all space}} u^*(\underline{r})\,u(\underline{r})\,d\tau = 1$. Furthermore, wave functions usually encountered in practice are real ($u^* \equiv u$).

Equation (9) can be solved by straightforward algebraic manipulation, yielding a solution of the form: $u(\underline{r}) = R(r)T(\theta)U(\phi)$ -(11). Allowable wave functions $u(\underline{r})$ exist only for the conditions that

$$n^2 = -\frac{z^2\mu e^4}{2\hbar^2 E}, \quad n = 1,2,3,4..... \qquad -(12)$$

This means that the energy of the electron in a hydrogenic atom is restricted to the values

$$E_n = -\frac{z^2\mu e^4}{2\hbar^2}\,\frac{1}{n^2} \qquad -(13)$$

which is the same expression obtained from Bohr theory, equation (4) above (with μ replacing m for strict accuracy); quantisation in the model arises naturally in contrast to its arbitrary origin in the classical description.

The R functions of equation (11) are radial wave functions, whose detailed form depends on the values of the two quantum numbers n and ℓ ($\ell \leq n-1$). Some of these normalised functions are given in table 1.

Table 1

Description of function	Mathematical form
n = 1, K shell: ℓ=0, 1s	$R_{10}(r) = (\frac{Z}{a_o})^{3/2}.2e^{-\sigma}$
n = 2, L shell: ℓ=0, 2s	$R_{20}(r) = (\frac{Z}{2a_o})^{3/2}.(2-\sigma)e^{-\sigma/2}$
: ℓ=1, 2p	$R_{21}(r) = 3^{-\frac{1}{2}}(\frac{Z}{2a_o})^{3/2}.\sigma e^{-\sigma/2}$
n = 3, M shell: ℓ=0, 3s	$R_{30}(r) = {}^{2}/_{27}.(\frac{Z}{3a_o})^{3/2}.(27-18\sigma+2\sigma^2)e^{-\sigma/3}$
: ℓ=1, 3p	$R_{31}(r) = \frac{3^{-\frac{1}{2}}}{81}.(\frac{2Z}{a_o})^{3/2}.(6-\sigma)\sigma e^{-\sigma/3}$
: ℓ=2, 3d	$R_{32}(r) = \frac{3^{-\frac{1}{2}}5^{-\frac{1}{2}}}{81}.(\frac{2Z}{a_o})^{3/2}.\sigma^2 e^{-\sigma/3}$

$$a_o = \frac{\hbar^2}{\mu e^2}, \quad \sigma = \frac{2r}{a_o}$$

Plots of the radial probability density $R^2(r)$ and of the radial distribution function $4\pi r^2 R^2(r)$ are given in figures 2 and 3 to show the physical significance of the radial wave functions.

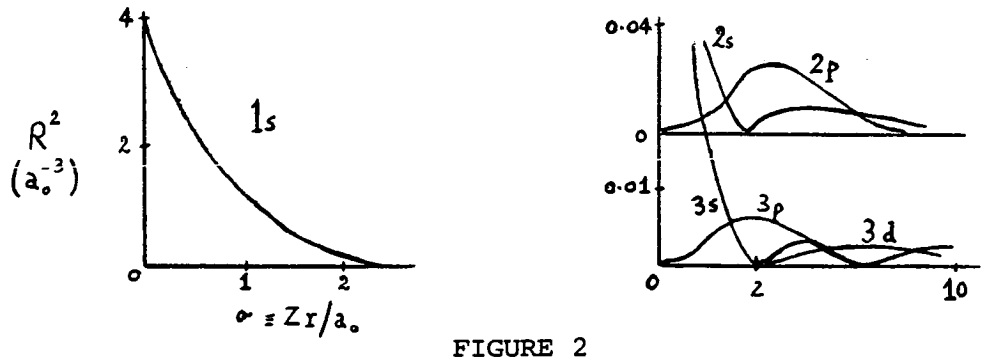


FIGURE 2

Plots of radial probability density as a function of r (θ,ϕ fixed);

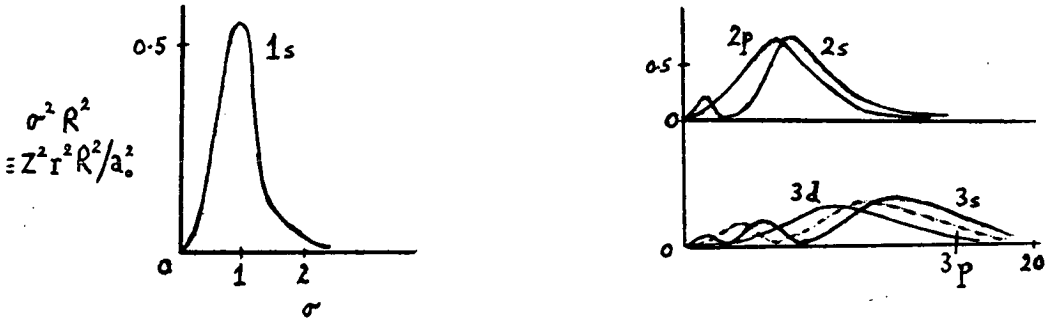


FIGURE 3

Plots of radial distribution function as a function of r.

The most probable value of r for an electron in a hydrogen 1s orbital is a_o (figure 3) - the radial distribution function gives the probability of finding the electron in a spherical shell of thickness dr at a distance r from the nucleus. As $a_o \equiv \frac{\hbar^2}{\mu e^2}$, this radius corresponds to the first Bohr radius (equation (3)) for hydrogen, although classical precision of meaning is not relevant any longer.

Returning to the solution $u(\underline{r})$ of equation (11) and its angular dependence, the product $T(\theta)U.(\phi)$ is found to be the well known series of mathematical physics functions called spherical harmonics, $Y(\theta,\phi)$, individual members being identified by two quantum numbers ℓ (as above) and $m(-\ell \leqslant m \leqslant \ell)$; the most important spherical harmonics for purposes here are given in Table 2.

Table 2

Description of Function	Mathematical Form (normalised)
ℓ=0, s function; m=0	$Y_{oo}(\theta,\phi) = \frac{1}{2\pi^{\frac{1}{2}}}$
ℓ=1, p functions; m=0	$Y_{1O}(\theta,\phi) = \frac{3^{\frac{1}{2}}}{2\pi^{\frac{1}{2}}} \cos\theta$
; m= $\pm$1	$Y_{1\pm1} = \frac{3^{\frac{1}{2}}}{2^{3/2}\pi^{\frac{1}{2}}} \sin\theta\, e^{\pm i\phi}$
ℓ=2, d functions; m=0	$Y_{2O} = \frac{5^{\frac{1}{2}}}{4\pi^{\frac{1}{2}}} (3\cos^2\theta-1)$
; m= $\pm$1	$Y_{2\pm1} = \frac{15^{\frac{1}{2}}}{2^{3/2}\pi^{\frac{1}{2}}} \sin\theta\cos\theta e^{\pm i\phi}$
; m=$\pm$2	$Y_{2\pm2} = \frac{15^{\frac{1}{2}}}{2^{5/2}\pi^{\frac{1}{2}}} \sin^2\theta e^{\pm i2\phi}$

The functions in table 2 are complex; in order to obtain wavefunctions which can be drawn in real space, appropriate linear combinations of spherical harmonics can be taken to yield real functions which are equally acceptable 'angular' solutions. Graphs of the angular wavefunctions can then be drawn in 3-dimensional space, i.e. as functions of θ and ϕ for constant r. Sections through such polar diagrams (surfaces) are shown in Figure 4 for some of the functions.

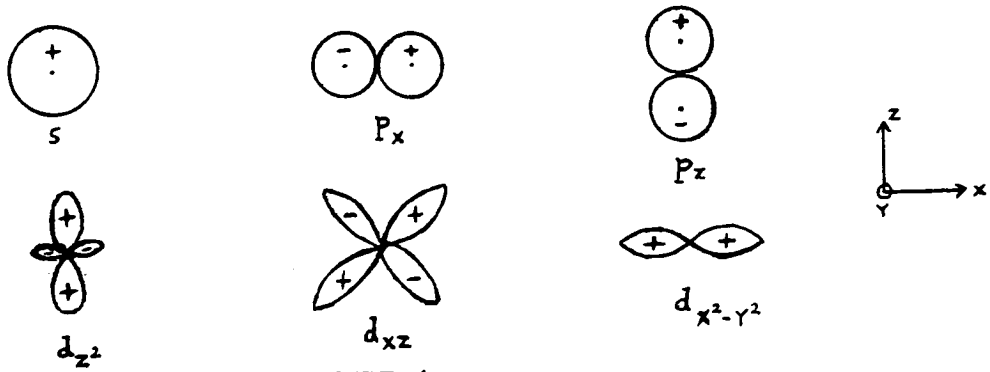


FIGURE 4

Polar diagrams for s,p,d functions (full surfaces generated by rotation around z-axis).

Although it is the square of the angular wavefunction that is related to a probability and hence physical significance, the actual wavefunctions depicted above (with sign included) have proved to be of use in valency arguments.

The total wave function for an allowed state of a hydrogenlike atom, described by the integral quantum numbers n,ℓ,m, is thus $u_{n\ell m}(r,\theta,\phi) = R_{n\ell}(r)Y_{\ell m}(\theta,\phi)$. It is not possible to give a pictorial representation of $u(\underline{r})$ which simultaneously shows the angular and radial variation, but the diagrams in figure 4 are a convenient, although approximate, representation. The allowed wavefunctions are called atomic orbitals, which are fundamental to all modern theory of valency.

The wave-mechanical model of the one-electron atom is consistent with a massive body of experimental evidence, and is also a more soundly based theory than Bohr theory (e.g. integral quantum numbers arise naturally from the solution of the Schrödinger Equation, in contrast to arising arbitrarily from a classical model). In addition, one-electron wavefunctions form the building blocks for the solution to more complex problems. The calculation of atomic structures in general has been developed to a very refined level[6], but, in this work, any required atomic wavefunctions are obtained as a special case of molecular

wavefunctions, which are calculated within a less refined model. The atomic orbital model forms the basis for molecular electronic wavefunction calculation, although, in practice, the precise mathematical form of hydrogenlike orbitals given above renders them unsuitable for use in a molecular environment.

It is convenient in quantum mechanical calculations to work in a set of units whose values are scaled to the fundamental physical quantities, the absolute numerical values of which disappear from the actual calculation. The following natural atomic units form a convenient set:

unit of mass = m_o, the rest-mass of the electron

unit of charge = $|e|$, the magnitude of the electronic charge

unit of length = $a_o = \frac{\hbar^2}{\mu e^2}$, the radius of the first Bohr orbit of the hydrogen atom

unit of action (energy x time) = $\hbar$

(unit of energy = $\frac{e^2}{a_o}$, consistent with above).

In atomic units, the time-independent Schrödinger equation for an electron moving in a potential V assumes the form: $(-\frac{1}{2}\nabla^2 + V)\Psi = E\Psi$ -(14)

For one-electron atoms, $V = -\frac{Z}{r}$ and the stationary states have wavefunction Ψ and energy $E = -\frac{Z^2}{2n^2}$.

In conclusion, the purpose of considering atomic electronic structure has been to emphasise the importance of the concept of the atomic orbital, which is the exact solution of the one-electron Schrödinger equation, in forming the chemically intuitive basis for describing the electronic structure of a molecule, for which a suitable wavefunction is necessarily approximate.

2. Quantum Theory of Molecular Structure

With a molecule as the physical system, the aim is to solve the appropriate Schrödinger equation to obtain a wavefunction describing electronic behaviour, which determines chemical properties. Before ultimately reaching a practicable method for the calculation of electronic wavefunctions for general molecules a series of approximations must be made, both in the physical and mathematical aspects of the wave-mechanical model. To illustrate the principles described below the particular example of the H_2 molecule will be used (Appendix 1).

(a) The Schrödinger Equation for a Molecule

Within the basic framework of quantum theory of the electronic structure and properties of molecules, solution of the Schrödinger equation

$$\hat{H}\,\Psi = E\,\Psi \qquad -(14a)$$

provides molecular electronic wavefunctions Ψ, and characteristic energy values E. Formally, Ψ is a function of the position and spin coordinates of all nuclei and electrons of a molecule. "Spin", or internal degree of freedom, of a fundamental particle is a concept which is a non-classical aspect of the wave-mechanical model; it arose naturally from general quantum mechanics and satisfies the test of experiment. The implications of spin are part of the general background which must be borne in mind in discussions of molecular structure, but fine details are not essential here so that is adequate to employ a nonrelativistic Hamiltonian $\hat{H}$, which neglects mathematical aspects of spin. Thus, for a molecular system of n electrons moving in the electrostatic field of N nuclei,

$$\hat{H} = -\tfrac{1}{2}\sum_{i=1}^{n}\nabla^2(i) + \sum_{i>j=1}^{n}\frac{1}{r_{ij}} - \sum_{\mu=1}^{N}\sum_{i=1}^{n}\frac{Z_\mu}{r_{\mu i}} - \tfrac{1}{2}\sum_{\mu}^{N}\frac{1}{m_\mu}\nabla^2(\mu) + \sum_{\mu>\nu=1}^{N}\frac{Z_\mu Z_\nu}{r_{\mu\nu}}$$

$$-(15)$$

where i,j run over electrons and μ,ν over nuclei and atomic units are assumed. In equation (15), the first term on the right-hand side is the quantum-mechanical kinetic energy operator which is applied to the coordinates of the electrons (as in section 1, $\nabla^2(i) \equiv \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2}$); the second term is the electron repulsion operator, with r_{ij} the distance between electrons i and j; the third term is the electrostatic nuclei-electrons attraction operator, with Z_μ the charge on nucleus μ and $r_{\mu i}$ the distance from nucleus μ to electron i; the fourth term is the nuclei's kinetic energy operator, with m_μ the mass of nucleus μ; the last term is the nuclear repulsion operator, with $r_{\mu\nu}$ the distance between nuclei μ and ν. As with the simple case of the hydrogen atom, as a consequence of the fact that the potential energy of the system of interacting particles depends only on relative inter-particle distances, it is possible to make equation (14) separable into translational and "internal" motion components which are independently soluble; this is done by using a system of coordinates based on the centre of mass of the system and internal coordinates. The $\hat{H}$ operator of equation (15) is for the complete motion[7].

Having reduced the most general situation to one where the time-independent, nonrelativistic Schrödinger equation is appropriate, the first really useful, simplifying approximation to be made is the Born-Oppenheimer Approximation (or Separation)[8]. Physically, the basis for this is that as a consequence of nuclei being much more massive than electrons so that the frequency of nuclear motions is much lower than that of electrons, it is possible to assume that electrons move in a potential provided by fixed nuclei. Mathematically, the nuclear kinetic energy term of $\hat{H}$ in equation (15) is separated off, and assuming that the total wavefunction Ψ of equation (14) can be expressed as

$$\Psi = \psi_e \psi_n \qquad -(16)$$

where the "electronic" wavefunction is defined by $\hat{H}_e \psi_e = E_e \psi_e$ -(17)

and the "nuclear" wavefunction by $(\hat{H}_n + E_e)\psi_n = E\psi_n$ -(18),

it is possible to obtain electronic wavefunctions, which are of interest chemically.

$$\hat{H}_n = -\tfrac{1}{2} \sum_{\mu}^{N} \frac{1}{m_\mu} \nabla^2(\mu)$$

and $\hat{H}_e = \hat{H} - \hat{H}_n$; ψ_e is a function of the electronic coordinates and the nuclear coordinates, whereas ψ_n is a function of the nuclear coordinates only. Thus, for a fixed position of the nuclei of the molecule, equation (17 can be solved (in principle, at least) to yield ψ_e and corresponding E_e, the electronic energy, which can then be taken as the potential energy determining the motion of the nuclei and inserted in equation (18) which is solved for ψ_n and E, the total energy of the molecule. Each electronic state described by ψ_e has its own set of nuclear wavefunctions ψ_n. There is clear evidence from spectroscopy that the separability of the total wave function of a molecule into an electronic and a nuclear part is in general a reliable assumption, which is actually the basis for the concept of potential energy curves and surfaces widely used in chemistry. The separation is valid provided the electronic wavefunction ψ_e is a slowly varying function of the nuclear coordinates, when a wavefunction of type in equation (16) is a good representation for the molecular system; this condition usually is satisfied in practice. Thus, it is the solution of equation (17) which is usually of central concern to chemists, so that the aim is to calculate an electronic wavefunction and corresponding electronic energy for a fixed nuclear configuration. The electronic energy E_e, which is not an observable quantity, is a function of the nuclear positions; the total energy E is the sum of E_e, evaluated at the particular configuration, and E_n, the nuclear energy.

All knowledge of the properties of a system of n electrons is contained in the electronic wavefunction $\psi_e(x_1, x_2, \ldots, x_n)$, where a typical electron i is described by its position in space, $(x_i, y_i, z_i) \equiv r_i$ in some coordinate frame, and the value of its spin S_i (effectively given by quantum number $+\frac{1}{2}$ or $-\frac{1}{2}$), so that $x_i \equiv (x_i, y_i, z_i, s_i)$. ψ_e is determined by the solution of $\hat{H}_e\psi_e = E_e\psi_e$, with

$$\hat{H}_e = -\tfrac{1}{2}\sum_{i}^{n}\nabla^2(i) + \sum_{\mu=1}^{N}\sum_{i=1}^{n} -\frac{Z\mu}{r_{\mu i}} + \sum_{i>j=1}^{n}\frac{1}{r_{ij}} \qquad -(19)$$

this Hamiltonian operator being effective within the Born-Oppenheimer Approximation, whereby nuclear coordinate variability is suppressed. H_e is often written as

$$\hat{H}_e = \sum_{i=1}^{n}\hat{h}(i) + \sum_{i>j=1}^{n}\hat{g}(i,j) \qquad -(20)$$

where $\hat{h}(i) = -\tfrac{1}{2}\nabla^2(i) + \sum_{\mu=1}^{N} -\frac{Z\mu}{r_{\mu i}}$ is called the "one-electron" Hamiltonian as it operates on the coordinates of electron i only,

and $\hat{g}(i,j) = \frac{1}{r_{ij}}$ is called the "electron repulsion" operator.

The primary object of calculations reported in this work is to find the eigenfunction ψ_e and eigenvalue E_e of the equation $\hat{H}_e\psi_e = E_e\psi_e$ with $\hat{H}_e$ given by equation (19), usually corresponding to the lowest energy state of the molecular system. As well as implicitly assuming from the outset the electronic Hamiltonian operator non-relativistic form is appropriate (electron distribution is determined predominantly by electrostatic effects, with magnetic effects really negligible) and that the fixed-nucleus approximation is valid, the determination of ψ_e is subject to much further approximation.

(b) The Pauli Principle

Formally, a molecular electronic wavefunction is determined by the solution of the Schrödinger equation, and simultaneously the equation

$$\hat{P}\,\psi_e(x_1,x_2,\ldots,x_n) = (-1)^p\,\psi_e(x_1,x_2,\ldots,x_n) \qquad -(21)$$

which is a mathematical formulation of the Pauli Principle for electrons: the wavefunction for a many-electron system must be anti-symmetric with respect to exchange of the coordinates x_i (space and spin) of any two electrons. $\hat{P}$ is a permutation operator which produces p transpositions of electron pairs[9]. More particularly, this means that in atoms no two electrons can have the same set of quantum numbers (n,ℓ,m,m_s) - n,ℓ,m arose in the H atom of the previous section and m_s is the spin quantum number, $+\frac{1}{2}$ or $-\frac{1}{2}$; this effect is fundamental to the "aufbau" principle of general atomic and molecular structure[10]. The electron spin coordinate S_i is involved essentially in a molecular electronic wavefunction even although the electronic Hamiltonian in the Schrödinger equation does not include spin explicitly; however, in practice, an approximate solution of $\hat{H}_e\psi_e = E_e\psi_e$ is sought and this must also satisfy equation (21) which is treated as a constraint on a trial ψ_e submitted to the Schrödinger equation. The constraint is satisfied quite easily by choosing a specific form for any function sought as an approximation to the true solution. Consequently, ψ_e is effectively treated as a function of space coordinates only, r_i, $\psi_e(x_i,y_i,z_i)$. Other constraints on ψ_e, as on wavefunctions in general mentioned in the case of the H atom, also fit in consistently with the usual choices of approximate ψ_e's.

(c) The Orbital Approximation

The molecular Hamiltonian of equation (19) shows that the actual analytical problem presented by the solution of the Schrödinger equation is a partial differential equation in 3n dimensions; in general, no further reduction into equations of smaller dimension is possible because of the presence of the operators $\hat{g}(i,j)$ (equation (20)). For one-electron systems, as shown in the previous section for hydrogenic atoms, there are no $\hat{g}(i,j)$ terms and the resulting three-dimensional equation can be solved exactly (n=1,N=1); additionally, for N=2 and n=1, the problem can be solved by classical analytical methods, operating within the Born-Oppenheimer Approximation. The molecule-ion H_2^+, described as the prototype of molecular structure[11], is the classical example of such a system. For N>2 and n=1, numerical solution is possible. However, these "exact" solutions, actually of an equation which is only approximate, are not of immediate interest in the solution of more practical problems. The numerical solution of the Schrödinger equation for n>2 and N>1 presents extreme computational difficulties, and it has been clear for some time that there is no hope of obtaining exact solutions for general molecules. From the chemist's point of view, this insolubility is rather irrelevant since numerical solutions, if they existed, would provide an extremely inaccessible form of chemical information. The purely mathematical approach to solution is not helpful; chemists must start from the idea of valence, that of considering the atomic structure of molecules, so that molecular wavefunctions are considered against a background of atomic wavefunctions. Consequently, although particular computational methods are based on rigorous mathematical principles, they are used within a model which incorporates a certain amount of empiricism.

The principal reason for the lack of exact solution for general molecules is that the motion of electrons is correlated, so that the accurate molecular wavefunction cannot be represented as a simple product of functions of individual electron coordinates; the type of separability, already exemplified above, cannot be employed. However, it has been found that the electronic motion is mainly determined by electron interaction with the nuclei and the mean mutual electronic interaction. This is reflected in the fact that atomic orbitals remain an appropriate language for the discussion of molecular electronic structure. An atomic orbital originally meant the solution (wavefunction) of any real or model single-electron Schrödinger equation, but current computational usage is more vague and is best summarised as any continuous function of three dimensions (space coordinates) whose square is integrable. The general physical interpretation of molecular wavefunctions is in terms of the basic units of atomic orbitals.

To obtain approximate wavefunctions, model approximations must be made. These consist of making simplifying assumptions about the nature of the physical system under investigation, a molecule, based on chemical and physical information. Mathematically, this means restricting the form of any function which is designed to approximate the molecular wavefunction. The molecular Hamiltonian of equation (20) is

$$\hat{H} = \sum_{i=1}^{n} \hat{h}(i) + \sum_{i>j=1}^{n} \hat{g}(i,j) \ .$$

If the electron repulsion terms are neglected, the simpler Schrödinger equation,

$$\hat{H}'\psi' = \left(\sum_{i=1}^{n} \hat{h}(i)\right)\psi' = E\psi' \qquad -(22)$$

can be reduced into n separate, one-electron equations (all identical):

$$\hat{h}(i)\phi_j(r_i) = \varepsilon_j\phi_j(r_i) \qquad i = 1, \ldots, n \qquad j = 1, \ldots, \infty \qquad -(23)$$

where $\phi_j(r_i)$ and ε_j are the eigenfunctions and eigenvalues of $\hat{h}(i)$. The required eigenfunctions and eigenvalues of the n-electron problem are then simply

$$\psi' = \phi_{j_1}(r_1)\phi_{j_2}(r_2)\phi_{j_3}(r_3) \ldots \phi_{j_n}(r_n) \qquad -(24)$$

$$\text{and } E = \varepsilon_{j_1} + \varepsilon_{j_2} + \varepsilon_{j_3} + \ldots + \varepsilon_{j_n} \qquad -(25)$$

($j_1, j_2, \ldots$ represent a selection of the possible solutions of equation (23)) In general, any linear combination of functions of type in equation (24) is an eigenfunction of equation (22):

$$\psi = \sum_i C_i\psi'_i \qquad -(26)$$

where the C_i are numerical coefficients.

When the full Hamiltonian, eqn. (20), is considered, such functions ψ are not eigenfunctions and hence not true molecular wavefunctions. However, if the orbitals $\phi_j(r)$ and/or coefficients C_i can be chosen carefully enough to compensate for the effect of the $\hat{g}(i,j)$, a suitable orbital model for systematic use can be formulated. It is possible to partition, approximately, the n-electron Hamiltonian into n separate one-electron Hamiltonians in order to obtain an approximate solution of the full Schrödinger equation given by a linear combination of products of orbitals defined by the one-electron Schrödinger equation. This approximate method of seeking solutions defines a whole class of orbital approximations;

the choice of orbitals $\phi_j(r)$, through the method of partitioning $\hat{H}$, defines a physical model, essentially comprising a set of building blocks for a molecular wavefunction ψ, which can be interpreted chemically in terms of the model one-electron orbitals $\phi_j(r)$.

The product form of ψ' in equation (24) does not satisfy the Pauli Principle; the spin coordinate does not appear in $\phi_j(r)$. The most direct way of introducing electron spin into the orbital model is to propose two spin "functions" which formally describe the two possible values of electron spin : spin factor $\alpha \equiv$ spin quantum no. $+\frac{1}{2}$, $\beta \equiv -\frac{1}{2}$. Thus, forming the product of a set of spatial orbitals $\phi_j(r)$ with the two possible spin factors defines a set of spin orbitals $\lambda_j(x)$, where

$$\lambda_k(x) \equiv \lambda_k(r,s) \equiv \phi_j(r)\alpha(s)$$

$$\lambda_{k+1}(x) \equiv \lambda_{k+1}(r,s) \equiv \phi_j(r)\beta(s) \qquad -(27)$$

giving twice as many spin orbitals as spatial orbitals. A particular electron is said to occupy spin orbital λ_j. Equation (24) can be rewritten as

$$\psi' = \lambda_{j_1}(x_1)\lambda_{j_2}(x_2) \ldots \lambda_{j_n}(x_n) \qquad -(28);$$

ψ' is still an eigenfunction of equation (22) as spatial operators $\hat{h}(i)$ have no effect on spin functions. A function which embodies the orbital model and satisfies the imposed Pauli Principle is given by

$$\Phi(x_1,x_2,\ldots,x_n) = \begin{vmatrix} \lambda_1(x_1)\lambda_1(x_2)\ldots\lambda_1(x_n) \\ \lambda_2(x_1)\lambda_2(x_2)\ldots\lambda_2(x_n) \\ \vdots \\ \lambda_n(x_1)\lambda_n(x_2)\ldots\lambda_n(x_n) \end{vmatrix} \qquad -(29)$$

where the well-established determinant notation for Φ has been used as a convenient representation of the detailed mathematical situation, which is not of interest here[12]. Within this model, the Principle reduces to: spin orbitals can be at most singly occupied, or spatial orbitals can contain either one electron or two electrons of opposite spin. The determinantal function Φ, intended as an approximate molecular wavefunction, satisfies the requirements mentioned previously; in fact, the component orbitals λ_j are interpretable as solutions of a Schrödinger equation so that constraints on Φ are normally met. In particular, Φ can be normalised by multiplication by N, with

$$N = (n!)^{-\frac{1}{2}} \begin{vmatrix} 1 & S_{12} & S_{13} & \cdots & S_{1n} \\ S_{21} & 1 & S_{23} & \cdots & S_{2n} \\ & & & & \\ S_{n1} & & & & 1 \end{vmatrix}^{-\frac{1}{2}}, \quad S_{ij} = \int \lambda_i(x)\lambda_j(x)\,dx \qquad -(30).$$

The full orbital model results from the existence of spin orbitals, formed from physically reasonable spatial orbitals, so that an approximate solution of the molecular Schrödinger equation is given by a linear combination of determinants of spin orbitals:

$$\psi = \sum_i C_i \Phi_i \qquad -(31).$$

The differential (Schrödinger) equation has been transformed so that the mathematical formulation of the problem becomes one where a set of functions of ordinary three-dimensional space is sought rather than one of 4n-dimensional space. This set, $\phi_j(r)$, and some numerical coefficients C_i are the basis for extracting chemical information from ψ. The problem is now in an area which is more amenable to systematic computation;

the orbitals ϕ_j are to be optimised, and/or the linear coefficients found, to approximate the true molecular wavefunction most closely.

(d) The Variation Method

The Schrödinger equation is not in a form readily solved by systematic use of approximation methods, and so it is transformed into an equivalent form particularly suited to computational implementation. It can be shown that an approximate wavefunction, $\tilde{\psi}$, has an expectation value for its energy, $\bar{E}$, which has the true energy , E_T, as its lower bound. Mathematically, the equation $\hat{H}\,\psi_T = E_T\psi_T$ has true solution ψ_T,

$$\text{and} \qquad E_T = \frac{\int \psi_T \hat{H} \psi_T d\tau}{\int \psi_T \psi_T d\tau} \qquad -(32)\ ;$$

if ψ_T is replaced by any approximate function $\tilde{\psi}$, then

$$E = \frac{\int \tilde{\psi} \hat{H} \tilde{\psi} d\tau}{\int \tilde{\psi} \tilde{\psi} d\tau} > E_T \qquad -(32a)\ .$$

This important and very general result suggests a practical procedure for the optimisation of any trial molecular wavefunction, particularly, the orbital model function. The approximate model wavefunction is substituted into the variational expression (32a) and the values of any adjustable parameters contained in the function (forms of orbitals, and linear coefficients) are varied until a minimum is found in the expression. The values of these parameters at the minimum define the best possible wavefunction of that particular functional form, i.e. the best possible description of the molecular electron distribution consistent with the limitations of the model. This variation principle, $\delta\bar{E} = 0$, applies strictly to only the lowest state of a given symmetry and is most widely used to determine ground state wavefunctions.

The transformation from a differential equation form to the minimisation of the value of an integrated expression has an "averaging" effect in that any approximate solution obtained will not, in principle, give a good description of molecular properties which depend on the value of $\tilde{\psi}$ at particular points in space.

In practice, the full optimisation of ψ of equation (31) is far too complex and time-consuming for many electron systems of chemical interest, and some restricted form of optimisation has to be effected.

(e) The Single-Configuration Model

Each electron in a complex polyelectronic system can be assigned in some way to its "own" spatial distribution according to the independent-electron model of electronic structure. "Independent" is best thought of as meaning independently assigned in the structure, and certainly not as independent of the motion (spatial distribution) of the other electrons. Within the orbital approximation scheme, the independent-particle model is realised by using a wavefunction consisting of a single determinant of spin-orbitals, a single orbital configuration (particular choice of n occupied spin-orbitals forming a single term of ψ in equation (31)). In this way, $\psi = \sum_i C_i \Phi_i$ has $C_1 = 1$ and $C_i = 0$ for $i>1$; all the computational effort is in choosing the best possible orbitals in the determinant Φ. This approach of optimising the $\phi_j(\lambda_j)$ is the basis of one of the two main model approximations of quantum chemistry, the Molecular Orbital method, which is actually implemented in this work. The other method, which uses fixed orbitals ϕ_j and optimises the coefficients C_i of an essentially multi-configuration wavefunction, is the Valence Bond method which has been used very widely by chemists in a qualitative way but is only now being applied quantitatively to chemical problems[13].

The single orbital configuration wavefunction is, assuming normalisation (eqn.30),

$$\psi = \Phi(x_1, x_2, \ldots, x_n) = \det\{\lambda_1(x_1)\lambda_2(x_2)\ldots\lambda_n(x_n)\} \quad -(33) .$$

Each spin-orbital λ_i can be occupied by one electron; each spatial orbital ϕ_i can be occupied by two electrons of opposite spin, "closed-shell" electronic system, or by one electron, "open-shell" system. The notation is usually contracted to

$$\psi = \det\{\phi_1\bar{\phi}_1\phi_2\bar{\phi}_2 \ldots, \phi_{n/2}\bar{\phi}_{n/2}\} \quad -(33a)$$

where $\lambda_1(x) = \phi_1(r)\alpha$, $\lambda_2(x) = \phi_1(r)\beta$, etc. and $\phi_1\alpha \equiv \phi_1$, $\phi_1\beta \equiv \bar{\phi}_1$ (closed-shell case). The determinantal nature of the function ψ has a rôle only in determining the form of the equations satisfied by the component spatial functions ϕ_i. Thus, working within the one-configuration model, the computational problem is the evaluation of the orbitals ϕ_i, and determinantal ideas recede into the background.

(f) The Molecular Orbital (MO) Method

Valence theory in chemistry is really concerned with changes in electron distribution on bonding. Ideally, accurate atomic wavefunctions should be used as units in molecular wavefunctions, but, within the intuitive independent electron model, atomic orbitals (AO's) have invariably been used as a basis for describing molecular electronic structure. The aim of the MO method is to find the best possible one-configuration (single-determinant) approximate solution of $\hat{H}\psi = E\psi$, where $\hat{H}$ is the non-relativistic, fixed-nucleus Hamiltonian of equation (20) and ψ is a determinant of spin-orbitals whose spatial components are

the MO's ϕ_i (equation (33)). The computational method is to assume a physically plausible form for the functions ϕ_i, which contain adjustable parameters, and optimise these parameters using the variation principle. For molecules in general, the approximate MO's ϕ_i, on both mathematical and chemical grounds, are conveniently expressed in the form of a Linear Combination of Atomic Orbitals (LCAO), of the constituent atoms of the molecule. This choice is a further approximation over and above the MO method, but it does lead to a convenient valence interpretation of the molecular wavefunction.

Thus, for a closed-shell molecule, which is the simplest and most common case, each MO ϕ_i in terms of the AO's χ_j is:

$$\phi_i = \sum_j T_{ji}\chi_j \qquad i = 1,2,\ldots,n/2 \qquad -(34) .$$

The elements T_{ji} of the matrix T are optimised using the variation principle, and the optimum matrix defines the structure of the MO's in terms of the AO's. The χ_j are expanded in terms of a set of basis functions η_K,

$$\chi_j = \sum_K C_{Kj}\eta_K \qquad -(35)$$

so that

$$\phi_i = \sum_j T_{ji}\left(\sum_K C_{Kj}\eta_K\right) = \sum_K \left[\sum_j C_{Kj}T_{ji}\right] \eta_K \qquad -(36),$$

giving an alternative method of computing MO's expanded directly in terms of the basis set of functions centred on each atom. A valence interpretation of this form of the MO method is possible by comparing the optimised coefficients of equations (35) and (36) to show the change in electron distribution on going from the atoms to the molecule.

Specifically, in this work, the LCAO MO method based on equation (34) is used, whereby chemical information is available most directly, as the elements of matrix T show the combination of AO's to form MO's; some mathematical flexibility is lost as the coefficients of the expansion in equation (35) are fixed, having been determined from an atomic calculation.

For a particular basis set η_K used to prepare the MO's (or AO's), the set of MO's that minimises $\overline{E}$, the expectation value of the energy of the molecule (equation (32)), are called SCF (Self-Consistent Field) orbitals, and $\overline{E} \equiv E_{SCF}$. A complete basis set would give the lowest possible value of $\overline{E}$ for a wavefunction of determinantal form, yielding "Hartree-Fock" orbitals and $\overline{E} \equiv E_{HF}$[14]. It has proved convenient to define the difference between E_T obtained by solving the nonrelativistic Schrödinger equation and E_{HF} to be the correlation energy:

$$E_T = E_{HF} + E_{corr} \qquad -(37) .$$

This definition and nomenclature is reasonable because the electrons move independently in the ψ_{HF} description, except for the correlation introduced by the Pauli Exclusion Principle[15].

(g) The Roothaan-Hartree-Fock Method

In summary, the mathematical problem is now:
assuming molecular electronic Hamiltonian,

$$\hat{H} = \sum_{i=1}^{n} \hat{h}(i) + \sum_{i>j=1}^{n} \hat{g}(i,j)$$

of equation (20) and wavefunction

$$\psi = \det\{\phi_1\bar{\phi}_1\phi_2\bar{\phi}_2\cdots\phi_{n/2}\bar{\phi}_{n/2}\}$$

of equation (33), find the optimum MO's ϕ_i by minimising the variational expression, $\int\psi\hat{H}\psi d\tau/\int\psi\psi d\tau$, with respect to the adjustable parameters T_{ij} of the assumed functional form,

$$\phi_i = \sum_j T_{ji}\chi_j \quad \text{(equation (34))}.$$

It is reasonable to suppose that the electronic structure of polyelectronic molecules can be based on the structure of the hydrogen (one-electron) atom. The forms of general atomic orbitals are expected to be similar to hydrogenic AO's, which have the general form:

(polynomial in r) x (spherical harmonic) x exp(-αr) (section 1(b)).

The central symmetry of atoms gives the Schrödinger equation a particularly simple form in spherical polar coordinates (r,θ,ϕ). Thus, the one-electron orbitals are linear combinations of the following type:

$$\text{(spherical harmonic)} \times \sum_i b_i r^i \,(\exp(-\alpha r) \qquad -(38)$$

where the coefficients b_i are fixed by the solution of the one-electron Schrödinger equation.

Formally, each MO ϕ_i is a linear combination of AO's χ_j (equation (34)); if n/2 MO's are required, the relation between the n/2 MO's and AO's (m in number, say) can be summarised in matrix notation as:

$$\phi = \chi T \qquad -(39)$$

where $\phi \equiv$ row vector $(\phi_1,\phi_2, \ldots \phi_{n/2})$

$\chi \equiv$ row vector $(\chi_1,\chi_2, \ldots, \chi_m)$

$T \equiv$ matrix of linear expansion coefficients to be optimised

$$\begin{bmatrix} T_{11} & T_{12} \cdots\cdots & T_{1\,n/2} \\ T_{21} & T_{22} \cdots\cdots & T_{2\,n/2} \\ & & \\ T_{m1} & T_{m2} \cdots\cdots & T_{m\,n/2} \end{bmatrix}.$$

To avoid over-complicating the derivation, it will be assumed that the AO's χ and the resultant MO's ϕ are orthogonal:

$$\int dr\ \chi_i(r)\chi_j(r) = \int dr\ \phi_i(r)\phi_j(r) = \delta_{ij} \qquad -(40)\ .$$

This restriction is easily removed when extending to the more general situation. The variational expression for the approximate energy $\overline{E}$ in terms of functions ϕ, $\overline{E} = \int d\tau\psi\hat{H}\psi / d\tau\psi\psi$, can be evaluated explicitly[16]; the result is

$$\overline{E} = 2\sum_{i=1}^{n/2}\{\int dr\phi_i(r)\hat{h}\phi_i(r)\} + \sum_{i>j=1}^{n/2}\{2(\phi_i\phi_i,\phi_j\phi_j) - (\phi_i\phi_j,\phi_i\phi_j)\} \qquad -(41)\ .$$

The standard notation for the energy integrals arising in molecular calculations has been introduced above:

$$(\phi_i\phi_j,\phi_k\phi_\ell) \equiv \int dr_1 \int dr_2 \phi_i(r_1)\phi_j(r_1)\hat{g}(1,2)\phi_k(r_2)\phi_\ell(r_2) \qquad -(42)\ .$$

In addition, $(\phi_i\phi_j,\phi_k\phi_\ell) \equiv G_{ijk\ell}$ and $G_{ijk\ell} = G_{k\ell ij} = G_{ij\ell k} = G_{ji\ell k} = G_{jik\ell}$. The integral of equation (42) is interpreted as the mean electrostatic repulsion energy between the charge density of electron $1(\phi_i\phi_j)$ and of electron $2(\phi_k\phi_\ell)$. To show the explicit dependence of $\overline{E}$ on the variables to be optimised, equation (34) is substituted in equation (41), yielding

$$\overline{E} = 2\sum_{i=1}^{n/2}\sum_{k,\ell=1}^{m} T_{ki}T_{\ell i}H_{k\ell} + \sum_{i=1}^{n/2}\sum_{k,\ell=1}^{m} T_{ki}T_{\ell i}G_{k\ell} \qquad -(43)\ .$$

The matrix H (elements $H_{k\ell}$) is defined by the integrals over AO's: $H_{k\ell} = \int dr\chi_k(r)\hat{h}\chi_\ell(r)$, and is usually referred to as the one-electron

Hamiltonian matrix. The matrix G is the "total electron interaction matrix" of Roothaan,

$$G_{k\ell} = \sum_{r,s=1}^{m} \{2G_{k\ell rs} - G_{kr\ell s}\} \, (TT^{\dagger})_{rs},$$

with $T^{\dagger}$ the transpose of T. $G_{k\ell rs} \equiv (\chi_k \chi_\ell, \chi_r \chi_s)$, known as the electron repulsion integrals or the two-electron integrals.

The i th column of matrix T, $T^{(i)}$, consists of the coefficients in the expansion of ϕ_i and these must be optimised by minimising $\bar{E}$ of equation (43) with respect to variations in the coefficients consistent with the normalisation and orthogonality of the functions ϕ_i, which is summarised by: $T^{(i)\dagger} T^{(j)} = \delta_{ij}$, as the χ_i have been assumed orthogonal. Variation of the coefficients defining the i th AO by $\delta T^{(i)}$ leads to the associated variation in $\bar{E}$, $\delta \bar{E}^{(i)} = 2\delta T^{(i)\dagger} H T^{(i)} + 2T^{(i)\dagger} H \delta T^{(i)}$

$$+ \delta T^{(i)\dagger} G T^{(i)} + T^{(i)\dagger} \delta G T^{(i)} + T^{(i)\dagger} G \delta T^{(i)} \qquad -(44)$$

dropping quadratic and higher terms in $\delta T^{(i)}$.

Now $\delta T^{(i)\dagger} H T^{(i)} = T^{(i)\dagger} H \delta T^{(i)}$ and $\delta T^{(i)\dagger} G T^{(i)} = T^{(i)\dagger} G \delta T^{(i)}$; also

$$(\delta G)_{k\ell} = 2 \sum_{r,s=1}^{m} \{2G_{k\ell rs} - G_{kr\ell s}\} \, \delta (T^{(i)} T^{(i)\dagger})_{rs} \text{ so that}$$

$$2\delta T^{(i)\dagger} G T^{(i)} = T^{(i)\dagger} \delta G T^{(i)} .$$

Thus, equation (44) simplifies to $\delta \bar{E}^{(i)} = 4\delta T^{(i)\dagger} H^F T^{(i)}$ $\qquad -(45)$

where $H^F = H + G$.

For a minimum in $\bar{E}$, each $\delta \bar{E}^{(i)}$ must vanish separately, consistent with variations in $T^{(i)\dagger} T^{(j)} = \delta_{ij}$, i.e. $\delta T^{(i)\dagger} T^{(i)} = 0$ (strictly, $\delta T^{(i)\dagger} T^{(j)} = 0$, but orthogonality of ϕ_i simplifies expression).

Consequently, $\delta T^{(i)\dagger}(H^F T^{(i)} - \varepsilon_i T^{(i)}) = 0$ -(46)

where ε_i is a Lagrange multiplier[17]. If (46) is to hold for arbitrary variations $\delta T^{(i)}$, then the contents of the bracket must vanish independently of $\delta T^{(i)}$:

$$H^F T^{(i)} = \varepsilon_i T^{(i)} \qquad i = 1,2,....,n/2 \qquad -(47)$$

or, in matrix notation, the whole set of n/2 equations is

$$H^F T = T\varepsilon \qquad -(48)$$

where ε is the diagonal matrix of the ε_i.

Thus, the original problem of the minimisation of $\bar{E}$ in equation (41) has been converted to an equivalent matrix eigenvalue problem - the computation of the eigenvalues and eigenvectors of the matrix H^F [17]. However, the problem is more complex than a single matrix diagonalisation since H^F contains T through the matrix G. Any method of solution of equation (48) must therefore be iterative; H^F and T must be found which both satisfy (48) self-consistently. This iterative nature of the solution gives the resulting MO's their familiar name: self consistent field molecular orbitals (SCFMO's). Equation (48) is a matrix form of the Roothaan-Hartree-Fock equations (RHF) and H^F is the RHF matrix.

To complement the preceding mathematical derivation, a brief mention of the physical interpretation of the RHF equations must be made[18]. In each equation $H^F T^{(i)} = \varepsilon_i T^{(i)}$, the column vector $T^{(i)}$ is simply the collection of optimum coefficients of the AO's χ_j defining the i th MO ϕ_i : $\phi_i = \chi T^{(i)}$. Assuming the RHF equations have been solved, the matrix H^F is the matrix representation of some "Hamiltonian" operator in the set of AO's. Physically, $H^F = H + G$; H is the one-electron Hamiltonian whose elements H_{ij} measure the mean kinetic energy and nuclear attraction energy of an electron with spatial distribution $\chi_i\chi_j$;

G contains all reference to electron repulsion integrals and represents the mean repulsion energy of an electron with density $\chi_i\chi_j$, and each matrix element $G_{k\ell}$ as defined above is the sum of two terms: the mean repulsion energy between $\chi_k\chi_\ell$ and the rest of the electronic system, and a term (exchange term) which corrects for the self-repulsion included in the first term and also includes the effect of the Pauli principle on the orbital model. This sum is the mean potential experienced by an electron having distribution $\chi_k\chi_\ell$ due to all other electrons. $H^F_{k\ell}$ $(= H_{k\ell} + G_{k\ell})$ is the total potential "seen" by such an electron. The method of solution of the RHF equations, based on making a guess at the $T^{(i)}$ and iteratively improving them has some physical parallel in seeking the stable arrangement of the electron distribution based on an assumed non-equilibrium starting point.

The eigenvalues ε_i of equation (47) have the dimension of energy, and are a measure of the energy of an electron associated with orbital ϕ_i. The precise meaning of these orbital energies will be given in the section on analysis of wavefunctions, where their importance will be emphasised.

In deriving equation (48), use has been made of the orthogonality of the functions χ: $T^{(i)\dagger}T^{(i)} = \delta_{ij}$. The usual, general situation is that the AO's χ are not orthogonal so that a so-called overlap matrix S can be generated:

$$S_{ij} = S_{ji} = \int dr \chi_i(r)\chi_j(r) \qquad -(49)$$

giving the overlap integrals between MO's ϕ_i and ϕ_j as $T^{(i)\dagger}ST^{(j)}$ and the orthogonality condition $T^{(i)\dagger}ST^{(j)} = \delta_{ij}$. The effect of variation in ϕ_i maintaining normalisation is $\delta T^{(i)\dagger}ST^{(i)} = 0$; the final RHF equation is then

$$H^F T = STε \qquad -(50)$$

where the overlap matrix plays a vital rôle in the definition of the MO's.

The optimum T matrix must satisfy the molecular RHF equation (50); effectively, in this work, the calculation of a molecular electronic wavefunction is the evaluation of optimum elements T_{ij}. The resulting MO's,

$$\phi_i = \sum_j T_{ji} \chi_j,$$

are not in general the best possible set of Hartree-Fock orbitals as they are limited by the capabilities of the AO functions χ_j, which are fixed by the nature of the atoms in the molecule. The solutions of (50) are the best possible MO's consistent with the limited set of AO's available. The RHF equation for the special case of spherically symmetric atoms, completely analogous to the above with a change in notation with the replacement of AO's ϕ by basis functions χ, can be solved in practice by special techniques; it is also possible to compute AO's directly from the differential equations by numerical methods. However, numerical AO's do not have direct application in molecular calculations[19]; the AO's which are solutions of the RHF equation are the basic units of any "valency" theory of molecular electron distribution. The general procedure in molecular calculations is to assume a set of AO's, expanded in terms of basis functions, and have the coefficient matrix T as the degree of freedom.

The development of the central ideas of orbital theories of atomic and molecular electronic structure outlined above has been rather general and abstract in summarising and systematising the ideas of qualitative quantum chemistry. More particular details of the computational aspect of the method of electronic wavefunction calculation are given in Appendix 1; an aid to understanding the above ideas is provided by considering a specific example of a molecular system.

B. Analysis of the Electronic Wavefunction of a Molecule .

Many chemical experiments can be interpreted in terms of expectation values of operators over unperturbed stationary states. "Theoretical" information can be provided by computing several expectation values of SCF wavefunctions, which approximate the molecular ground states. There is a selection of operators whose expectation values are now routinely calculable from a molecular wavefunction; the corresponding properties are readily measured in experiments and reveal basic aspects of the charge distribution.

The expectation value $\overline{P}$ of some operator P, where P can be written in terms of the coordinates and momenta of the nuclei and electrons of a molecule, for a normalised approximate wavefunction Ψ is given by

$$\overline{P} = \int \Psi P \Psi d\tau \qquad -(51) .$$

The properties of chemical interest may be broadly classed as one-electron properties and two-electron properties. The former correspond to operators that can be written as the sum of operators containing the coordinates of one electron (or none); the latter correspond to operators containing the interelectronic distance. Within the basis used, for the Hartree-Fock (SCF) wavefunctions the expectation value of the energy is stationary with respect to changes of the one-electron orbitals. It can be shown that the one-electron properties of molecules are expected to be represented rather well by expectation values of the approximate H-F wavefunction, whereas two-electron properties are expected to be represented more poorly [20].

(a) Energy Expectation Values

The following operators can be defined:

$$H_K = -\tfrac{1}{2} \sum_i \nabla^2(i) \qquad \text{Kinetic energy operator} \qquad -(52)$$

$$H_{NE} = -\sum_{\mu i} \sum z_\mu / r_{i\mu} \qquad \text{One-electron potential energy operator} \qquad -(53)$$

$$H_{EE} = + \sum_{i>j} 1/r_{ij} \qquad \text{Two-electron potential energy operator} \qquad -(54)$$

$$H_{NN} = + \sum_{\mu>\nu} z_\mu z_\nu / r_{\mu\nu} \qquad \text{Nuclear-repulsion operator} \qquad -(55)$$

The Hamiltonian operator of the molecular system has been partitioned into "classical" components. The total electronic energy operator is $H_E = H_K + H_{NE} + H_{EE}$ and the total energy operator is $H_T = H_E + H_{NN}$. Energy expectation values whose operators involve H_{EE} are two-electron properties; actually, such operators are special two-electron ones, being intimately involved in the determination of the optimum approximate wavefunctions, and general two-electron properties cannot be routinely evaluated and are not of interest in this work[21].

(b) General One-Electron Molecular Properties

Operators corresponding to such properties computed herein can be written in the form of a sum of one-particle operators for all nuclei and electrons in the molecule:

$$P = \sum_\mu z_\mu p'(\underline{r}_{\mu R}) + \sum_j ep(\underline{r}_{jR}) \qquad -(56)$$

with $Z_\mu \equiv$ nuclear charge, $e(=-1) \equiv$ electronic charge, $\underline{r}_{iR}$ is the vector from the reference position $\underline{R}$ to the position $\underline{r}_i$ of the ith particle ($\underline{r}_{iR} = \underline{r}_i - \underline{R}$) and the corresponding distance is $r_{iR} = |\underline{r}_i - \underline{R}|$;
$\underline{R} \equiv (R_x, R_y, R_z)$ and $\underline{r}_i \equiv (X_i, Y_i, Z_i)$.

For closed-shell molecules having the determinantal wavefunctions adopted in this study, the expectation value of a one-electron operator can be written in a simple form:

$$\overline{P} = \sum_\mu Z_\mu p'(\underline{r}_{\mu R}) + 2\sum_j e \left\langle \phi_j | p(\underline{r}_R) | \phi_j \right\rangle \qquad -(57) .$$

Here the first term gives the nuclear contribution and the second, the electronic contribution. For the above type of wavefunction, the electronic contribution is twice the sum over the occupied MO's of the expectation value resulting from one electron in each MO:

$$\left\langle \phi_j | p(\underline{r}_R) \mid \phi_j \right\rangle = \int \phi_j(1) p(\underline{r}_{1R}) \phi_j(1) d\tau \qquad -(58) .$$

(c) Moments of the Charge Distribution

The moments of the molecular charge are, like the energy expectation values, properties of the whole molecule. They characterise the size and shape of the charge distribution. The reference position $\underline{R}$ of equation (56) is the molecular centre of mass.

The first moment, or dipole moment, is a vector whose components x,y and z correspond to the one-electron operators $X_{\mu R} (\equiv X_\mu - R_x)$, $Y_{\mu R}, Z_{\mu R}$ for p' and X_{jR}, Y_{jR}, Z_{jR} for p; it is origin independent for neutral species.

The second moments have components x^2, y^2, z^2, xy, xz and yz corresponding to operators X^2_{jR}, Y^2_{jR}, Z^2_{jR}, $X_{jR}Y_{jR}$, $X_{jR}Z_{jR}$ and $Y_{jR}Z_{jR}$ for p; for an uncharged molecule with a permanent (nonzero) dipole

moment, these are origin dependent. The operator $r^2_{jR'} = X^2_{jR} + Y^2_{jR} + Z^2_{jR'}$ yields values for $\overline{r^{-2}}$ which, besides being a very good measure of the spatial extent of the wavefunction and charge distribution, also gives the diamagnetic part of the magnetic susceptibility.[22]

Corresponding to the operators $p = X^3_{jR'}$, $XY^2_{jR'}$, $XZ^2_{jR'}$, etc. are the third moments' components x^3, xy^2, xz^2, etc.; corresponding to $p = X^4_{jR'}$, $Y^4_{jR'}$, $X^2Y^2_{jR'}$, etc. are the even components of the fourth moments x^4, y^4, x^2y^2 containing only even powers of the coordinates, and to $p = X^3_{jR}Y_{jR'}$, $X_{jR}Y^3_{jR'}$, $X_{jR}Y_{jR}Z^2_{jR'}$, etc. are the odd components x^3y, xy^3, xyz^2 etc.

First and second, or dipole and quadrupole, moments are of particular importance in this work. There are extensive compilations of values of these for molecules obtained from experiment[23] so that the adequacy of calculated molecular wavefunctions can be discussed in terms of other than energy expectation values by comparison with empirical data. In addition, properties such as dipole moment and diamagnetic susceptibility are very useful in a general valence context.

(d) Expectation Values at Atomic Centres

In evaluating the required sums of equation (56), the nuclear centre $\underline{R}$ at which the property is being evaluated is dropped from the nuclear contribution sum.

The potential at nucleus μ due to all other electrons and nuclei is given by the operators $p = r^{-1}_{j\mu}$ and $p' = r^{-1}_{\mu'\mu}$. This potential is proportional to the diamagnetic part of the nuclear shielding in NMR[22]. More recently, the value of the quantum mechanical potential at a nucleus has also been shown to be a significant quantity in determining the chemical shift in X-ray photoelectron spectroscopy.[24]

The electric field at a nucleus μ is given by the gradient of the potential, with operators for the components $p = X_{j\mu}/r^3_{j\mu}, Y_{j\mu}/r^3_{j\mu}, Z_{j\mu}/r^3_{j\mu}$ (p' similarly). These components are identified with the Hellmann-Feynman part of the force exerted on an atom in a molecule.[25] In general, values of these operators at any nucleus should vanish only for a Hartree-Fock wavefunction and optimum (equilibrium) molecular geometry. Actually, at any geometry, the sum of all forces on all nuclei should vanish for a H.-F. wavefunction, so that any deviation from this result is attributable solely to the approximate nature of the wavefunction.[26] Values of the electric field have also been found useful in calculating the paramagnetic or high-field part of the nuclear shielding constant for the proton.[27]

The gradient of the electric field at nucleus μ is a tensor whose components, $(3x^2-r^2)r^{-5}, 3xyr^{-5}$, etc., correspond to the operators $(3X^2_{i\mu}-r^2_{i\mu})r^{-5}_{i\mu}, (3X_{i\mu}Y_{i\mu})r^{-5}_{i\mu}$, etc. This property is the prominent electronic structure quantity in nuclear quadrupole resonance spectroscopy.[28] Accurate values of electric field gradients can be used to deduce the value of nuclear quadrupole moments, and also to give a sensitive test of the electronic charge distribution in the immediate neighbourhood of the nucleus. The charge density at nucleus μ is given by the operator $\delta(r_{j\mu})$.[29]

(e) Population Analysis

The above mentioned one-electron properties are physical properties. In addition, there have been repeated attempts made to 'interpret' the wavefunction Ψ, or, to be more rigorous, Ψ^2 as it is $\Psi^2 dV$ which has physical meaning: when Ψ is a many-electron wavefunction, $\Psi^2 dV \equiv \Psi^2 d\underline{r}_1 ds_1 d\underline{r}_2 ds_2 \ldots d\underline{r}_n ds_n$ ($\underline{r}$ and s indicate space and spin variables of the n electrons) gives the probability of an instantaneous configuration

of all electrons. The most common physical properties, however, depend on the probability per unit volume of finding a single electron, no matter which, at a given point $\underline{r}$ in space; this is given by

$$P(\underline{r}) = n \int_{\underline{r}_1=\underline{r}} |\Psi|^2 ds_1 d\underline{r}_2 ds_2 \ldots d\underline{r}_n ds_n \qquad -(59) \ .$$

The factor n arises because of the indistinguishability of electrons. The quantity P is often referred to as the "electron density" since, for many purposes, the electron distribution may be treated as a smeared-out charge of density P (electrons per unit volume). Much investigation has sought to recover chemically useful information from Ψ, and to obtain reliable techniques to provide answers to questions posed by elementary descriptive valence theory. The most widely used method to characterise regions of space numerically, by the amount of electron density contained within the regions, is population analysis. Thus, the mathematical rigour of the Hartree-Fock model is supplemented by empirical reasoning to partition the electrons of a molecule among atoms or bonds. The most widely adopted definitions are those of Mulliken,[30] which are of relevance in this work.

The chemical information in the matrix T, introduced in equation (34), is made much more accessible by the definition of a new matrix R which defines the weights with which the various orbital products appear in the total electron density. The contribution from electrons in AO χ_i to the total electron density is $2(\sum_{j=1}^{n/2} T_{ij}^2)\chi_i^2$ for n/2 doubly-occupied spatial MO's. Considering the internuclear regions for bonding densities, the expression $2\sum_{k=1}^{n/2} T_{ik}T_{jk}\chi_i\chi_j$ gives the contribution of electrons occupying the region of overlap of χ_i and χ_j. In summary, $R = TT^{\dagger}$. The occupation number of any orbital product $\chi_i\chi_j$ is given by $2R_{ij}$;

the diagonal elements of R are called orbital charges, and the off-diagonal ones bond orders. R, or sometimes 2R, is known as the charge and bond order matrix.

With the one-determinant approximate function Ψ, it can be shown that $P(\underline{r})$ of equation (59) is given by $P(\underline{r}) = 2\chi(\underline{r})R\chi^{\dagger}(\underline{r})$ (matrix notation),

$$\text{or } P(\underline{r}) = 2 \sum_{i,j} R_{ij}\chi_i(\underline{r})\chi_j(\underline{r}) \qquad (60) .$$

The matrix R is a finite matrix representation of the total electron density in terms of the AO functions χ_i. The spatial dependence has been absorbed into the AO products $\chi_i\chi_j$, leaving the population numbers R_{ij} summarising the electron distribution. It is useful to note that the functional dependence of the molecular G matrix of equation (43) is on R rather than matrix T so that it is possible to regard R, not T, as the independent variable to be optimised in minimising the total electronic energy.[31]

The orbital model used here to obtain approximate solutions to the Schrodinger equation transfers much of the analysis of the electron density function (equation (59)) into the interpretation of matrix R.

$$\text{If } \underline{P} = 2R = 2TT^{\dagger}, \quad P(\underline{r}) = \sum_{i,j=1} \chi_i(\underline{r})\chi_j(\underline{r})P_{ij} \qquad -(61) .$$

To allow for non-orthogonality of the AO's χ, the orbital-product distributions can be re-normalised by division by the corresponding overlap term, S_{ij}, and re-defining the charges and bond orders gives

$$p_{ij} = \mathbf{P}_{ij}S_{ij} \text{ and } P(\underline{r}) = \sum_{i,j} \mathbf{p}_{ij}(\chi_i\chi_j/S_{ij}) \qquad -(62) .$$

Chemical interpretation is facilitated by confining attention to the electronic populations of atoms and interatomic regions. This rather coarse breakdown of the electron density yields

$$P(\underline{r}) = \sum_{\alpha=1}^{N} \{ \sum_{i,j\varepsilon\alpha} P_{ij}(\chi_i\chi_j/S_{ij}) \} + \sum_{\alpha\neq\beta=1}^{N} \{ \sum_{i\varepsilon\alpha} \sum_{j\varepsilon\beta} P_{ij}(\chi_i\chi_j/S_{ij}) \} \quad -(63)$$

where the notation $i\varepsilon\alpha$ means the summation runs over AO's χ_i centred on atom α. Thus,

$$P(\underline{r}) = \sum_{\alpha=1}^{N} P_{\alpha}(\underline{r}) + \sum_{\alpha\neq\beta=1}^{N} P_{\alpha\beta}(\underline{r}) \quad -(64)$$

emphasising the atomic and interatomic contributions. There is no physically satisfactory way of partitioning the charge associated with the interatomic regions among the atoms themselves - rather an artificial interpretation of the electron density. To formally associate electrons with individual atoms, Mulliken suggested that the electron density $P_{\alpha\beta}(\underline{r})$ be equally partitioned between α and β, yielding

$$Q_{\alpha} = \sum_{i,j\varepsilon\alpha} P_{ij} + \tfrac{1}{2} \sum_{j\varepsilon\beta} \sum_{i\varepsilon\alpha} P_{ij} \quad -(65)$$

as the overall charge on atom α in the molecule ($\sum_{\alpha} Q_{\alpha}$ = net charge on molecule). "Mulliken Population Analysis" is a particularly well-documented subject and it has been found that, despite limitations of the discrete analysis of electron density, it provides a very worthwhile way of summarising the mass of information in the molecular wavefunction and of correlating computational results with empirical chemistry. It is particularly simple to implement the scheme, but the indices of population calculated are basis set dependent; for formally balanced bases,[32] it can, however, give useful information when applied to a series of related molecules or when used in conjunction with other methods of analysis, but its predictive value is small when used in isolation on a single molecule. The general area of population

analysis has been extensively investigated;[33] of the alternative methods proposed, the Mulliken scheme is still the central one, having been thoroughly tested empirically.

(f) Molecular Orbitals and Eigenvalues

The molecular properties considered above are normally computed as expectation values of the ground-state, total electronic wavefunction. Individual components of the latter, molecular orbitals, are also of interest, particularly in recent times with the correlation which can be routinely made between calculated MO eigenvalues (orbital energies) and experimentally observed ionisation potentials from photoelectron spectroscopy.[34] In this context also, calculation of the electronic wavefunction of an individual molecule can be useful, whereas, in general, a comparative study of a series of molecules is the aim of the type of investigation reported here.

The RHF equation (48) when solved yields an MO ϕ_i and corresponding eigenvalue ε_i; the eigenvector ϕ_i is characterised by the column vector $T^{(i)}$ of coefficients in the expansion in terms of AO's. To find the precise meaning of ε, the effect on the total energy of the molecule of removing an electron from ϕ is considered. Equation (43),

$$\bar{E} = 2\sum_{i=1}^{n/2}\sum_{k,\ell=1}^{m} T_{ki}T_{\ell i}H_{k\ell} + \sum_{i=1}^{n/2}\sum_{k,\ell=1}^{m} T_{ki}T_{\ell i}G_{k\ell},$$

for the total electronic energy of a closed-shell molecule of n electrons, can be rewritten as

$$\bar{E} = 2\sum_{i=1}^{n/2} h_{ii} + \sum_{i=1}^{n/2}\sum_{j=1}^{n/2} g_{ij} \qquad -(66)$$

where $h_{ii} = \int \phi_i \hat{h} \phi_i dr$

and $g_{ij} = 2\int \phi_i \phi_i \hat{g} \phi_j \phi_j dr^2 - \int \phi_i \phi_j \hat{g} \phi_i \phi_j dr^2$

$= 2J_{ij} - K_{ij}$

(using previous definitions from Section A, equations (41) and (42)).

This states that the total electronic energy is the sum of:

(i) h_{ii}, the energy each electron would have alone in the nuclear framework, its kinetic energy plus nuclear attraction potential energy

(ii) J_{ij}, a coulombic repulsion between every pair of electrons

(iii) K_{ij}, an exchange interaction between every pair of electrons of the same spin.

Thus, the total electronic energy of the ion formed by removing an electron from the MO ϕ_k, by applying equation (66), is

$$\bar{E}^+ = 2 \sum_{i=1}^{n/2-1} h_{ii} + \sum_{i,j=1}^{n/2-1} g_{ij} + h_{kk} + \sum_{i=1}^{n/2} 2J_{ki} - K_{ki} \qquad -(67)$$

(using the property that $J_{ii} = K_{ii}$).

(The physical interpretation of equation (66) summarised above aids in this derivation). Therefore, the energy difference between the molecule and the ion is

$$\bar{E} - \bar{E}^+ = h_{kk} + \sum_{i=1}^{n/2} g_{ki} = \varepsilon_k \qquad -(68)$$

as the rightmost equality of (68) is simply a restatement of the k th equation of the set defined by equation (47).

In summary, for a closed-shell molecule the molecular "orbital energy" is equal to minus the ionisation potential of an electron from that orbital. This result is known as Koopmans' Theorem, or preferably Koopmans' Approximation.[35]

Equation (68) is an extremely significant illustration of the complementarity of theory and experiment, in particular, MO theory and the observation of ionisation potentials of individual electrons from molecules via the technique of Photoelectron Spectroscopy.[36] Koopmans' Approximation has been used extensively in the interpretation of photoelectron spectra; some examples arise in this work, and so some comment on this approach must be made.[37]

The result known as Koopmans' Approximation, for closed-shell molecules, is based on three approximations:

(i) the reorientation approximation - the derivation given above assumes that the MO's are unaltered when going from molecule to ion, i.e. no reorientation.

(ii) even if assumption (i) were valid, or if new orbitals were calculated for the ion by finding the coefficients $T_{ion}^{(i)}$ which optimised $\bar{E}^{+}$, there remain two errors; firstly, the relativistic error - Hartree-Fock Theory neglects relativistic effects so that Koopmans' Approximation effectively assumes that the relativistic energy is the same in both molecule and ion, which may be reasonable for ionisation by removal of outer electrons but certainly not for ionisation by removal of inner electrons.

(iii) there remains the correlation energy error inherent in H.-F. Theory, so that Koopmans' Approximation makes the assumption that the correlation energy must be the same in both molecule and ion -

generally, the correlation energy is less in the ion than in the parent molecule since correlation effects arise to a large extent from electron pair interactions; the correlation energy also varies with ionic state.

Additionally, for open-shell molecules, there are two more dangers in the indiscriminate use of Koopmans' Approximation:

(iv) the orbital energy ε_i found from Hartree-Fock calculations has no clear physical interpretation in the open-shell case .

(v) for open-shell states the wave functions may only be represented in multi-determinantal form.

Despite these objections, Koopmans' Approximation has proved to be extremely useful in the interpretation of photoelectron spectra, especially when applied to chemical species as considered in this work. Roughly speaking, the reason for this is that approximations (i) and (iii) above, which are usually the most significant, conveniently lead to errors which cancel if Koopmans' Approximation is applied to estimate an ionisation potential; effect (i) leads to the estimate being too large, but effect (iii) leads to too small a value. Conversely, when the reverse process of negative ion formation is considered, the use of the corresponding expression for estimating electron affinities has been shown to be unreliable.[38]

(g) Localised Descriptions of Electronic Structure

Two types of information can come from an approximate solution of the Schrodinger Equation for a molecular system. These may be classified as the qualitative and quantitative interpretation of the density function defined previously. In sections (a) - (d) above, the calculation of first-order properties of a molecule was shown to involve the computed

density functions and certain molecular integrals involving the relevant molecular operator. On the other hand, in section (e), the "chemical" information obtained from analysis of the density function was the important factor. Although a quantitative estimate of the electron density changes on bonding, such information is used in rather qualitative ways. The orbital model used here to obtain approximate solutions of the Schrödinger Equation throws much of the analysis of the density function into the interpretation of the charge and bond-order matrix R, defined in section (e). It is important to stress that the approximate molecular wavefunction and corresponding one-electron density function of equation (59) are continuous functions of space, whereas matrix R is a collection of coefficients which multiply spatial functions. This discrete analysis of the electron density has been found to be a very worthwhile way of summarising the mass of information in the molecular wave function and of correlating computational results with empirical chemistry.

Returning to more mathematical considerations, it is of interest to examine the effect, on a valence calculation, of changes in the orbital basis: in particular, the transformations induced by defining new AO's or MO's (or basis functions) in terms of linear combinations of the original functions. Thus, a new MO basis ϕ_L can be formed by taking linear combinations of the original MO set ϕ and the relevant coefficients collected in a square matrix L, so that, as $\phi = \chi T$, $\phi_L = \phi L = \chi TL = \chi T_L$ where $T_L = TL$. The equation $\phi_L = \chi T_L$ defines a new set of MO's in terms of the AO's, and therefore a corresponding R matrix can be defined, $R_L = T_L T_L^\dagger = T(LL^\dagger)T^\dagger$. If the matrix L is unitary, representing a transformation from an orthogonal set of MO's ϕ to a new orthogonal set ϕ_L, then $LL^\dagger = L^\dagger L = 1$, and so $R_L = R$.

As R defines the total electron distribution, it follows that unitary transformations among the (occupied) MO's do not change the total LCAOMO wavefunction.

In addition, the calculated value of any observable is unchanged by a linear transformation among the orbital basis functions;[39] in particular, a LCAOMO calculation gives a total electronic energy which is unchanged by non-singular linear orbital transformations, i.e. the linear independence of the orbitals is preserved.

It is possible to gain some insight into the nature of the valence electronic structure by using the invariance property of the one-determinant LCAOMO wavefunction with respect to linear transformations among the occupied MO's; a chemical interpretation of the molecular electronic distribution can be obtained.

The AO density matrix, which summarises the electron density, is not changed by transformations as above. This arbitrariness in the definition of the occupied MO's suggests that this implied freedom to choose a unitary transformation among the occupied MO's could be used to yield a set of MO's which are particularly adapted to chemical interpretation. This transformation would form a useful bridge from MO's to the familiar concepts of empirical chemistry. The MO's arising "naturally" as the solutions of the RHF equation are delocalised, as a typical MO has significant contributions (non-zero elements of T) from all the AO's of the molecule. This description of the molecular electron density is in marked contrast to the familiar chemical idea of localised electron-pair bonds and the invariance of bonds, lone pairs and inner shells with respect to changes in molecular environment. Examination of the valence electron density in a molecule does reveal that there are regions of high electron density in the bond and lone pair regions,

but this is the result of the superposition of the essentially delocalised contributions from the individual MO's. Thus, the unobservable individual MO contributions do add up to form a chemical picture. In contrast, Localised Molecular Orbitals can be defined, mathematically, as an alternative partitioning of the total electron density, which can then be viewed as the addition of spatially distinct regions of high electron density rather than the superposition of spatially overlapping densities.[40]

There are several well-defined criteria for the computation of localised MO's from the RHF delocalised MO's. One of the first, and most successful methods, is best introduced by considering the expression for the total energy of a one-configuration wavefunction in terms of integrals involving the MO's (equation (66)). The MO density matrix has a particularly simple form: $R_{ij} = \delta_{ij}$ and $P_{MO} = 2R_{MO} = 2I$. The energy expansion becomes[41]

$$\bar{E} = 2\sum_{i=1}^{n/2} \phi_i \hat{h} \phi_i \, dr + \sum_{i,j=1}^{n/2} \{2(\phi_i\phi_i,\phi_j\phi_j) - (\phi_i\phi_j,\phi_i\phi_j)\} \qquad -(69) .$$

The physical interpretation of integrals $(\phi_i\phi_i,\phi_j\phi_j)$ is the "classical" repulsion between an electron in MO ϕ_i and one in MO ϕ_j. Integral $(\phi_i\phi_j,\phi_i\phi_j)$, the exchange term, has no classical analogue, and is a consequence of the antisymmetry requirement.[42] Interpreting chemical effects in terms of a "semi-classical" electrostatic model, the following prescription for forming localised MO's can be given: the localised MO's (LMO's) for an electronic system are those for which the energy expression has the form (69) with minimum contribution from the exchange integrals.

Mathematically, this means the search for a set of orbitals ϕ_{LI} such that

$$\overline{E} = 2\sum_{i=1}^{n/2} \int \phi_{LI} \hat{h} \phi_{LI} dr + \sum_{i,j=1}^{n/2} (\phi_{LI}\phi_{LI}, \phi_{LJ}\phi_{LJ})(2-\delta_{IJ}) \quad -(70) .$$

The exchange integrals can be written as a matrix K with elements $K_{ij} = (\phi_i\phi_j, \phi_i\phi_j)$ - the diagonal elements are the "self-repulsion" coulomb terms. The unitary transformation which diagonalises the symmetric matrix K defines a set of orbitals for which the exchange integrals are small, i.e. the matrix L such that $L^{\dagger}KL$ = diagonal matrix has columns which express the ϕ_{LI} in terms of the original delocalised MO's ϕ_i. The LMO's are defined in terms of the AO's by $\phi_L = \chi TL$ ($\phi_L = \phi L$ and $\phi = \chi T$), in matrix notation.

This particular method of choosing LMO's has the advantage of being a simple diagonalisation. A more complete procedure would be to minimise the exchange contributions and the "off-diagonal" coulomb repulsion while simultaneously maximising the self-repulsion terms. There are technical difficulties associated with this maximum/minimum problem.[43] A variety of criteria can be used to define LMO's;[44] all methods produce "bond orbitals" which have chemically appealing localised character in line with chemical intuition.

The view can be taken that the LMO's are, for the chemist, more fundamental elements of molecular structure than the delocalised, or "canonical", MO's, and that the computation of the latter is just a numerical step in the computation of LMO's. It is possible to compute directly LMO's, by-passing the LCAO Method, as the validity of the RHF equations is independent of the use of AO's as basis functions, the choice of which affects the numerical accuracy of the resulting molecular wavefunction but not its formal validity as a variational trial function;

thus, functions which approximate LMO's can be chosen as basis - these could be based on general chemical knowledge of electron pairs and on the forms of LMO's obtained from LCAO MO's.[45] This type of approach does not involve an iterative calculation, the use of the variation principle for optimising the given form of the LMO's not yielding the reduction to a "pseudo-eigenvalue" problem as in the RHF method. In practice, the problem becomes computationally tractable if the LMO form is taken to be that of a spherical 1s Gaussian, $\exp(-\alpha r^2)$, centred at an appropriate point; this rather poor approximation, physically, has proved to be very useful, qualitatively and semi-quantitatively.[46]

The method of LMO's outlined above has important conceptual advantages over the simple LCAO MO scheme, but there is no numerical value in the process as the total molecular electronic wavefunction is invariant against the localisation procedure, which is, strictly, arbitrary. However, as the MO wavefunction is invariant against transformations in the AO basis, there is a localised picture which can be of conceptual and numerical value beyond the one-configuration model; regarding the constituents of the LMO's rather than the LMO's themselves as the invariant elements of molecular electronic structure, i.e. the hybrid AO's, it is possible to develop a method to optimise the hybrid AO basis to give the best possible pairwise bonding scheme, where the LMO's appear predominantly as linear combinations of pairs of hybrids.[47] Such considerations are essentially part of a Valence Bond approach to electronic structure, and, although chemically appealing, computational problems are very significant, in practice.[48]

Localised MO's form a bridge between chemical intuition and molecular quantum mechanics, and it is therefore useful to consider these when analysing a molecular electronic wavefunction to obtain chemically meaningful results.

C. The Hartree-Fock Method .

The electronic structure of molecules can be explored by analysing the electronic wavefunction, which is the exact description of electrons in the field of nuclei. The computation of a wavefunction is a "many-body problem", the general solution of which is not possible at present; the Hartree-Fock Method of section A is a technique, incorporating approximations as described therein, for such computation which is as general as feasible for the present and not of hindrance for likely future development - it is a formalism valid for atoms and molecules, and extendible in principle to solids and liquids, so that it forms a basis for the general description of quantised matter. The Hartree-Fock concept, namely, the hypothesis of "one-electron orbitals", retains an intuitive representation of the electronic structure of molecules, leading to an approximate wavefunction which, although yielding quantitative information of qualitative value only, is uniquely defined.[49] Formally, as well as in practice, the Hartree-Fock function can constitute the "zero order" function for exact functions.[50] It is the intention of this section to indicate the status of the Hartree-Fock wavefunction in relation to both theoretical and experimental considerations; it should be borne in mind that this work actually involves a very particular type of H-F wavefunction.

(a) Theoretical Considerations

Electronic wavefunctions of polyatomic molecules can be obtained as "Ab initio", or non-empirical, solutions of the Hartree-Fock Equations, and, from these, mean values of observables for general molecular species can be derived. Providing enough care is taken in choosing a good basis set and enough computer power is available to use large bases, the results

can approach the Hartree-Fock limit, i.e. the best possible solution to the H-F equations as would be obtained were they to be solved numerically or with an infinite basis set. In principle, the best possible electronic wavefunction within the limits of the H-F equations can be deduced, but these equations only approximately represent the reality of the molecular situation and are subject to some severe limitations, so that such a wavefunction is not exceptionally sophisticated, or, indeed, even highly satisfactory.

No mention is made of relativistic corrections in the derivation of the H-F equations. The virial theorem indicates that inner-shell electrons with the highest potential energy have the greatest kinetic energy, so that relativistic effects (variation of mass from "rest" value) may become important for these electrons, particularly in heavy atoms or molecules. Relativistic energies are certainly important contributors to the total energy of such species, but they are difficult to estimate.[51] Fortunately, however, quite frequently absolute energies are not of interest, but rather differences between energy levels; since inner-shell electrons are normally unchanged between electronic energy levels, the problem of relativistic energy may frequently be ignored.

Far more serious is the problem of the Correlation Energy. This is defined by equation (37) as the residual error in H-F calculations, even if relativistic effects are accounted for; equation (37) becomes in practice

$$E_T = E_{HF} + E_{corr} + E_{rel} \qquad -(71)$$

where E_T is the experimental total energy of the molecular species. The physical origin of this error is the representation of inter-electronic interaction by coulomb and exchange terms, each electron having a direct interaction with the averaged-out charge of all the others, but an exchange interaction only with electrons of the same spin.

In reality, an electron will have instantaneous interactions with all the others which will not be the same as the average interaction included in the SCF procedure, resulting in the correlation-energy error, which is partly accounted for in the case of electrons of the same spin by the exchange terms in the H-F equations.

Correlation energies can be estimated, essentially empirically, from the analysis of a H-F wavefunction;[52] functions can be added to the basis set to improve the wavefunction, although again empirical reasoning is used in choosing the explicit form of these to include correlation.[53] However, the correlation energy error is directly related to the incorrectness of the restrictive single-determinant of orbitals form for a molecular wavefunction. The inclusion of correlation effects requires a more sophisticated form of wavefunction, rigorously defined. Possibilities include: (1) the explicit introduction of inter-electronic coordinates, which leads to serious computational problems;[54] (2) valence bond constructions for the wavefunction and geminal or electron-pair formulations, such methods ("pair" theories) promising to make adequately correlated wavefunctions and energies for small molecules quite readily available[55] - the computational procedure involved is based on H-F LMO's associated with inner-shells, bonds and lone-pair (section B above) and the resulting wavefunction is closely related to classical chemical concepts. The formulation which has so far proved to be simplest to implement computationally is (3) the Configuration Interaction (CI) Method.[56] A CI construction can actually in principle, with the unattainable limit of a complete basis set, lead to an exact solution of the Schrödinger Equation (17). In this formulation, the molecular electronic wavefunction $\psi = \sum_i C_i \Phi_i$ (equation (31)) is a linear combination of determinants of spin orbitals. Excited configurations, Φ_i, are obtained by excitation from orbitals of the single determinant H-F

description to "virtual" MO's of the SCF solution - the leading term of equation (31) is the H-F solution. Thus, with the basis set of orbitals chosen, ψ is optimised by varying the coefficients C_i with the determinants Φ_i fixed. Alternatively, a limited number of different types of configurations can be assumed and the MO's forming ϕ_i optimised simultaneously with the coefficients C_i in the Multiconfiguration wavefunction (MCSCF procedure).[57] Thus, even in deriving more refined wavefunctions, the orbital model approach has been most fruitful - the simplest approximation is determined by this method (H-F), and this can be used to establish a starting-point for refinements in ψ.

In the CI method states arising from different spin-orbital configurations are mixed, so that the exact solution to the Schrödinger electronic equation is expanded in terms of the complete (infinite) set of determinantal wavefunctions, which in turn are constructed from some complete set of one-electron spin orbitals.[58] This obviously impracticable solution must be replaced by truncating the "full" CI wavefunction by selecting the most important configurations to express $\psi (= \sum_i C_i \Phi_i)$ as a linear combination of the determinantal wavefunctions Φ_i formed from configurations of identical symmetry and spin; the number of configurations which can be formed for a molecular system of n electrons is of the order m^n, using m basis functions in a H-F calculation to obtain n/2 doubly occupied MO's and $(m - \frac{n}{2})$ vacant or virtual orbitals for creating excited-state determinants by systematically promoting electrons from occupied orbitals of the ground-state to the virtual orbitals. The crucial computational problem in implementing this scheme in general is the very slow convergence of the CI expansion; this is because the virtual orbitals do not occupy the same physical space as do the occupied orbitals, as electrons in the former move in the field of all n electrons

and those of the latter in the field of (n-1) electrons. More promisingly, this last problem is overcome in the MCSCF procedure which improves the convergence of the CI expansion. In this method, the CI coefficients C_i of $\psi = \sum_i C_i \Phi_i$ and the MO coefficients T_{ij} of $\phi_i = \Sigma T_{ij} \chi_j$ are both varied to optimise ψ; a doubly iterative process is used, yielding the optimum orbitals and configurational mixing coefficients for the basis set used. Many fewer configurations are needed in the MCSCF method than in a conventional CI calculation in order to achieve the same results.

In this work, "large" molecules in the quantum-chemical sense (benzene forming the dividing line between small and large molecules) are considered. To obtain wavefunctions approaching the Hartree-Fock limiting form for such species entails the use of large basis sets, so that the calculations are, in fact, prohibitively expensive in general; the routine implementation of procedures as outlined above to obtain wavefunctions of superior quality is therefore some way off. Removing the approximation of the MO form of the molecular electronic wavefunction introduced in the H-F method is obviously desirable, but the extra effort required to produce a CI-type wavefunction is wasted if an excessively truncated basis set is used, as required by present computational techniques. Calculations on atoms and small molecules can now yield very refined wavefunctions, but the generalisation of methods used to larger, chemically interesting species is in the early stages of development.[59]

(b) Experimental Results

From the theoretical derivation of the H-F procedure for computing molecular electronic wavefunctions it is obvious that, because of the nature of the approximations made, such a wavefunction is unlikely to

approximate very closely the true one. However, the ultimate test of the adequacy of the H-F wavefunction is the comparison of results deduced from it with the corresponding experimental ones. It is intended now to summarise some of the well-established, general examples of such comparisons, and show where methods beyond the Hartree-Fock one are required to give an intimate description of chemical phenomena.

The assertion that single-determinant SCF theory (H-F level) is adequate for determining ground-state molecular properties is based on numerical experience as opposed to rigorous proof; exceptions do exist, and these can often be traced to physically significant effects not allowed for by such a theory. Some "counter-examples" are:-

(i) Atoms:- Figure 5 shows a plot of correlation energy ΔE, defined as the difference between the calculated total energy and observed total energy, against atomic number.[60] The correlation energy increases with Z; for carbon, the error is already quite large, 0.16 a.u. = 4.4 eV, and clearly in a molecular problem errors of this magnitude are tolerable only if they can be regarded as mainly errors in the absolute energy, and less significantly errors in relative energies compared to the separated atoms.

(ii) Ionisation and Excitation Energies:- there has often been found poor agreement between ionisation potentials observed experimentally (extensive research on atoms particularly in photoelectron spectroscopy[61]) and calculated values, either using computed orbital energies or separate calculations on neutral and ionic species; H-F ΔE values tend to be in very poor agreement with ultraviolet excitation energies observed.[62] It is important to note that it is questionable, on theoretical grounds, to obtain quantitative information on other states from ground-state wavefunctions alone;

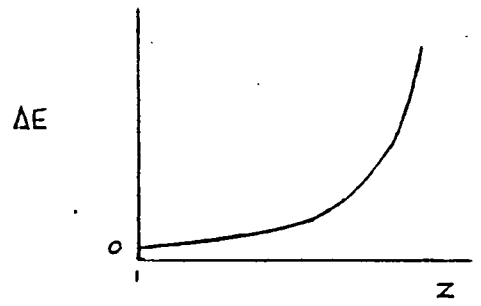


FIGURE 5 : CORRELATION ENERGY VERSUS ATOMIC NO.

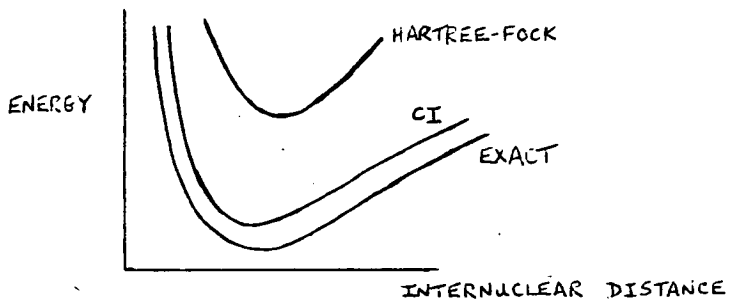


FIGURE 6 : DIATOMIC POTENTIAL CURVE WHEN DISSOCIATION IS TO OPEN-SHELL ATOMIC STATES

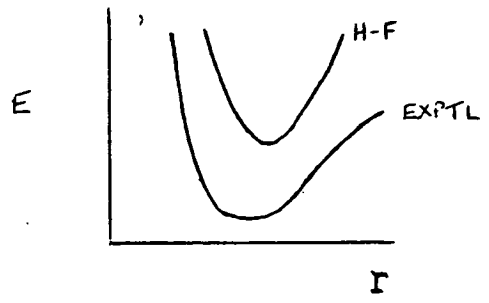


FIGURE 7 : H-F AND EXPTL POTENTIAL CURVES - APPROXIMATELY PARALLEL NEAR MINIMUM

however, as mentioned in Section B, quite good agreement often exists between orbital energies and ionisation potentials, but the reason is not that the H-F function is an adequate approximation, but rather a cancellation of errors occurs.

(iii) Dissociation Energies and Equilibrium Geometries of Diatomics:- Figure 6 shows a typical diatomic molecule potential curve behaviour when the dissociation is to open-shell atomic states.[63] The restricted H-F wavefunction dissociates incorrectly to a mixture of atomic states; dissociation energies calculated are frequently too small owing to correlation errors associated with the formation of additional electron pairs in the molecule.

In summary, H-F theory, or its approximate version LCAO SCF-MO theory as used in this work, provides a qualitatively correct account of many molecular properties, but the description given breaks down whenever two or more molecular states from different configurations have similar total energies and interact strongly, i.e. strong first-order CI when all the strongly interacting configurations must be included in the total electronic wavefunction. In the expansion $\psi = \Sigma C_i \Phi_i$, one term alone does not dominate, as happens when a H-F wavefunction is adequate. Thus, even for a closed-shell molecule whose MO's are well separated at the equilibrium nuclear geometry, first-order CI usually becomes important for large displacements of the nuclei such as those involved in dissociation or chemical reaction. More refined theories that are qualitatively correct for all molecular states and all nuclear geometries are well-established, but the problem of obtaining quantitative results, within chemical accuracy, for large molecules by a detailed analysis of correlation effects is still to be solved.

While recognising the deficiencies of the Hartree-Fock wavefunction, it is essential to consider, in contrast, its successes. Formally, a H-F wavefunction satisfies two very powerful theorems, the virial theorem[64] and the Hellmann-Feynman theorem, in addition to Brillouin's theorem,[65] as a result of which it can be shown that a molecular charge distribution derived from a H-F wavefunction and properties determined by it are predicted with surprisingly high accuracy (in fact, "one-electron" properties of Section B are correct to second order).[66] More particularly, a vitally important aspect is the calculation of potential energy surfaces. The concept of a potential energy surface is a consequence of the separation of the nuclear and electronic motions, resulting in solution of the Schrödinger equation within the Born-Oppenheimer Approximation.[67] The gain in conceptual simplicity afforded by this is enormous and its use underlies many chemical concepts - e.g. energy barriers, potential constants, frequencies of vibrational and rotational motions, bond lengths and angles as determined by an energy minimum. The Born-Oppenheimer, or clamped-nucleus, approximation, as mentioned in Section A, regards the form of the molecular electronic charge distribution as being determined by the electronic wavefunction evaluated for each static configuration of the nuclei; mathematically, the electronic function $\psi_i(\underline{X};\underline{R})$ is obtained by solving the electronic Schrödinger equation for a fixed nuclear configuration,

$$\hat{H}_e \psi_i(\underline{X};\underline{R}) = E_i(\underline{R})\psi_i(\underline{X};\underline{R}) \qquad -(72)$$

(cf. equation (17)). ψ_i depends explicitly on the electronic space-spin coordinates $\underline{X}$ and implicitly, as E_i, on the nuclear coordinates $\underline{R}$. The solution of equation (72) for all spatial arrangements of the nuclei generates an energy (hyper-) surface, which governs the motion of the nuclei.

Solutions to the full Schrödinger equation (14a) provided within this approximation are entirely adequate for most problems of chemical interest; breakdown occurs in the case of degeneracy or near degeneracy of the electronic states $\psi_i(\underline{X},\underline{R})$, resulting in the Jahn-Teller or Renner effects,[68] when the electronic wavefunction depends strongly on the nuclear motion. While non-negligible errors are found in the curvatures and shapes of Hartree-Fock calculated potential energy surfaces for polyatomic systems in the region of the equilibrium geometry, the heights of small energy barriers separating geometrical conformers are accurately predicted, acceptable values being obtainable even by calculations above the H-F limit.[69] Equilibrium geometries are predicted rather well by the H-F SCF procedure. Absolute energies obtained in H-F calculations, although mostly accurate to within 1% of the exact value, are in considerable error (correlation error), but for most problems of chemical interest, energy differences are important; it is now well established that the correlation error in the H-F energy remains relatively constant in the neighbourhood of the minimum or minima of the potential energy surface. Consequently, the H-F energy surface approximately parallels the true surface in these regions (Figure 7) and the ability of H-F theory to provide acceptable descriptions of polyatomic systems in the neighbourhood of equilibrium geometries is understandable. The correlation energy is relatively insensitive to changes in molecular conformation.[70] There are certain types of system for which the above approximate parallelism holds over the total surface; in particular, the H-F wavefunction shows the correct asymptotic behaviour. The rationalisation of this is that major changes in correlation energy occur only when the "reaction" process is accompanied by drastic change of electron configuration (number of electron pairs). Therefore, reactions involving closed-shell reactants

and products, with constant number of electron pairs implying small change in correlation energy, and where near-degeneracy (CI) effects are unimportant, can be treated very well by H-F theory.[71] This contrasts with the general situation (Figure 6) of the failure of a H-F wavefunction to describe properly the possible dissociation products of a system; the H-F potential surface increases too steeply along a coordinate leading to dissociation, also leading to a tendency to under-estimate bond lengths and over-estimate force constants.

In summary, LCAO MO SCF approximation to the H-F limit is a method capable of yielding semi-quantitative potential energy surfaces for closed-shell systems, with geometrical parameters characterising minima in the surface given to within 1 or 2%, energy barriers to ca. 4 kJ mol^{-1}, energies of reaction to $\pm$ 12 kJ mol^{-1} in favourable cases and usually to within $\pm$ 30 kJ mol^{-1}.

(c) The Role of "Ab Initio" Calculations

The construction of molecular electronic wavefunctions by so-called ab initio techniques is performed in practice by an imperfect formulation of the many-electron problem; the terminology ab initio has become associated with the fact that interactions, usually in the form of integrals, are evaluated accurately, and does not mean that the form of the wavefunction itself is not unduly constrained. In general, in each of the variety of calculations described as ab initio, there exist constraints imposed to achieve either tractability or simplicity and these in turn can have important physical consequences. The understanding of precisely how specific constraints affect a prediction is an evolving proposition, as also is the average level of sophistication of reported ab initio treatments. At the present time, some of the underlying uncertainties

are definitely known and these can often be understood in terms of a few examples and counterexamples. The availability of quantitative information from ab initio studies should also serve the important purpose of delineating the role of certain concepts in qualitative and semi-empirical descriptions of bonding.[72]

In particular, in the type of calculations reported in this work, the choice of basis set is effectively an empirical step, although the ensuing mode of calculation is non-empirical. Even in methods beyond the Hartree-Fock level, in performing a correlation energy calculation it is less a question of choosing the best formalism, than of using the best LCAO basis. In the Roothaan technique, an optimal wavefunction is obtained relative to the pre-determined basis set; although quantitative data of qualitative character is provided, it is this type of data which can be analysed to obtain some correlation between molecular structures. An exact molecular wavefunction provides a tool for deducing exact expectation values, but such data taken alone do not constitute understanding of the electronic structure of molecules. A characteristic of MO theory is that each progressive improvement or step has a natural physical explanation; the "Hartree-Fock limit" represents a well-defined plateau, in terms of its mathematical and physical properties, in the hierarchy of approximate solutions to the Schrödinger electronic equation. Estimates can be made of the type of chemical system, in terms of its electronic structure, to which each level of ab initio calculation may be expected to yield usable results, i.e. results with acceptable errors or with predictable bounds on errors. In recent times, there has been a welcome closing of the gap between the meaning given to the term "large molecule" by a quantum chemist and that given by an organic or inorganic chemist.[73]

References for Chapter 2

1. L. Pauling, "The Nature of The Chemical Bond, and The Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry", 3rd edition, Cornell University Press, New York, 1960 - presentation of empirical chemical theories within a framework obtained from the synthesis of quantum and valency theories.
2. Descriptive theoretical chemistry - references for Chapter 1, nos. 1, 7, 9, 10.
3. Present-day research - F. Herman, A.D. McLean, and R.K. Nesbet (editors), "Computational Methods for Large Molecules and Localised States in Solids", Proceedings of IBM Symposium (1972), Plenum, New York and London, 1973.
4. J.W. Linnett, "Wave Mechanics and Valency", Methuen, London, 1960; W. Heitler, "Elementary Wave Mechanics, With Applications to Quantum Chemistry", 2nd edition, O.U.P., London, 1956.
5. H. Margenau and G.M. Murphy, "The Mathematics of Physics and Chemistry", 2nd edition, D.Van Nostrand, Princeton, 1956 - chapter 5.
6. D.R. Hartree, "The Calculation of Atomic Structures", Chapman and Hall, London (and John Wiley and Sons, New York), 1957.
7. B.T. Darling and D.M. Dennison, Physical Review, 57, 128 (1940).
8. M. Born and J.R. Oppenheimer, Ann.Physik, 84, 457 (1927).
9. W. Pauli, Z. Physik, 31, 765 (1925).
10. Reference 1 for Chapter 1.
11. C.A. Coulson, "The Shape and Structure of Molecules", Clarendon Press, Oxford (1973).
12. J.C. Slater, "Quantum Theory of Molecules and Solids, Vol. 1: Electronic Structure of Molecules", McGraw-Hill, New York, 1963.
13. J. Gerratt, "Valence Bond Theory", a Specialist Periodical Report of the Chemical Society in Theoretical Chemistry, Vol. 1 (Quantum Chemistry), London, 1974.

14. J.I. Musher, Chemical Physics Letters, 7, 397 (1970); E.H. Lieb and B. Simon, Journal of Chemical Physics, 61, 735 (1974).

15. P.-O. Löwdin, Advances in Chemical Physics, 2, 207 (1959); E. Clementi, Journal of Chemical Physics, 38, 2248 (1963).

16. R. McWeeny and B.T. Sutcliffe, "Methods of Molecular Quantum Mechanics", Academic Press, New York, 1969.

17. Reference 5, chapters 6, 8, 10.

18. C.C.J. Roothaan, Reviews of Modern Physics, 23, 69 (1951); G.G. Hall, Proceedings of the Royal Society, A205, 541 (1951).

19. J.B. Mann, LASL Report Nos. LA-3690 and 3691 (1967).

20. C. Moller and M.S. Plesset, Physical Review, 46, 618 (1934).

21. G. Diercksen and R. McWeeny, Journal of Chemical Physics, 44, 3554 (1966); P.-O. Löwdin, Journal of Molecular Spectroscopy, 13, 326 (1964).

22. D.W. Davies, "The Theory of the Electric and Magnetic Properties of Molecules", John Wiley and Sons, New York, 1966.

23. A.L. McClellan, "Tables of Experimental Dipole Moments", W.H. Freeman and Co, San Francisco and London, 1963; W.H. Flygare and R.C. Benson, Molecular Physics, 20, 225 (1971).

24. T.K. Ha and C.T. O'Konski, Chemical Physics Letters, 3, 603 (1969); H. Basch, ibid, 5, 337 (1970).

25. J. Hellmann, "Einführung in die Quantenchemie", Deuticke and Co., Leipzig, 1937; R.P. Feynman, Physical Review, 56, 340 (1939).

26. P. Pulay, Molecular Physics, 17, 197 (1969).

27. S.I. Chan and T.P. Das, Journal of Chemical Physics, 37, 1527 (1962); H. Basch and A.P. Ginsberg, Journal of Physical Chemistry, 73, 854 (1969).

28. T.P. Das and E.L. Hahn, "Nuclear Quadrupole Resonance Spectroscopy", Academic Press, New York, 1958; R.L. Matcha, Journal of Chemical Physics, 47, 4595 (1967).

29. S. Fraga and G. Malli, "Many-Electron Systems: Properties and Interactions", W.B. Saunders Co., Philadelphia, 1968.

30. R.S. Mulliken, Journal of Chemical Physics, 23, 1833 (1955).

31. R. McWeeny, Reviews of Modern Physics, 32, 235 (1960).

32. R.S. Mulliken, Journal of Chemical Physics, 36, 3428 (1962).

33. P. Politzer and R.R. Harris, Journal of the American Chemical Society, 92, 6451 (1970); K. Jug, Theoretica Chimica Acta, 31, 63 (1973).

34. S.D. Worley, "Photoelectron Spectroscopy in Chemistry", Chemical Reviews, 71, 295 (1971); W.C. Price, "Photoelectron Spectroscopy", Advances in Atomic and Molecular Physics, 10, 1 (1974).

35. T. Koopmans, Physica, 1, 104 (1933).

36. H. Bock and P.D. Mollère, Journal of Chemical Education, 51, 506 (1974).

37. W.G. Richards, International Journal of Mass Spectrometry and Ion Physics, 2, 419 (1969).

38. Assuming $\overline{E} - \overline{E}^- = \varepsilon_K (= -E.A.)$, with MO ϕ_K unoccupied in neutral species, is not justified on mathematical grounds, in any case.

39. V. Fock, Z. Physik, 61, 126 (1930).

40. R. Daudel, M.E. Stephens, E. Kapuz and C. Kozmutza, Chemical Physics Letters, 40, 194 (1976).

41. J. Lennard-Jones and J.A. Pople, Proceedings of the Royal Society, A202, 166 (1950); H. Brion and R. Daudel, Compt. Rend.Acad.Sci. (Paris), 237, 567 (1953).

42. Reference 16, in particular p. 140.

43. C. Edmiston and K. Ruedenberg, Reviews of Modern Physics, 35, 457 (1963).

44. W. England, L.S. Salmon and K. Ruedenberg, Fortschr. Chem.Forsch., 23, 31 (1971); H. Weinstein, R. Pauncz and M. Cohen, Advances in Atomic and Molecular Physics, 7, 97 (1971).

45. A.A. Frost, Journal of Chemical Physics, 47, 3707 (1967); R.E. Christoffersen, Advances in Quantum Chemistry, 6, 333 (1972).

46. R.M. Archibald, D.R. Armstrong and P.G. Perkins, Journal of the Chemical Society-Transactions of the Faraday Society II, 70, 1557 (1974); P.H. Blustin and J.W. Linnett, ibid, 70, 274 (1974).

47. D. Peters, Journal of the Chemical Society, A, 644 (1966).

48. D. Peters, Journal of Chemical Physics, 51, 1559 (1969).

49. J.L. Whitten, Accounts of Chemical Research, 6, 238 (1973).

50. Reference 16, p. 76f.

51. I.P. Grant, Advances in Physics, 19, 747 (1970).

52. R. Colle and O. Salvetti, Theoretica Chimica Acta, 37, 329 (1975); H.O. Pamuk, ibid, 28, 85 (1972).

53. H.F. Schaefer, "The Electronic Structure of Atoms and Molecules", Addison-Wesley, Reading (Mass.), 1972 - so-called "polarisation" functions.

54. E.A. Hylleraas, Z. Physik., 54, 347 (1929) -original "correlated" wavefunction.

55. A.C. Hurley, "Electron Correlation in Small Molecules", Academic Press, New York, 1976.

56. E.A. Hylleraas, Z. Physik., 48, 469 (1928) - CI method first discussed; bibliography in reference 53.

57. T.L. Gilbert, Physical Review, A6, 580, 1972 - detailed bibliography.

58. Z. Gershgorn and I. Shavitt, International Journal of Quantum Chemistry, 2, 751 (1968).

59. B. Levy, Chemical Physics Letters, 18, 59 (1973).

60. A. Veillard and E. Clementi, Journal of Chemical Physics, 49, 2415 (1968).

61. D.A. Shirley (editor), "Electron Spectroscopy", North-Holland, Amsterdam, 1972.

62. G. Herzberg, "Electronic Spectra and Electronic Structure of Polyatomic Molecules", Van Nostrand, Princeton, 1967.

63. R.F.W. Bader and R.A. Gangi, in first entry of Reference 8 for Chapter 1 (Volume 2).

64. J.C. Slater, Journal of Chemical Physics, 1, 687 (1933); C.W. Kern and M. Karplus, ibid, 40, 1374 (1964); C.A. Coulson, Molecular Physics, 20, 687 (1971); J.C. Slater, Journal of Chemical Physics, 57, 2389 (1972).

65. L. Brillouin, Actualities Sci.Ind., 71, Secs. 159, 160 (1933-34).

66. R.F.W. Bader, in second entry of Reference 8 for Chapter 1.

67. P.G. Wilkinson and H.O. Pritchard, Canadian Journal of Physics, 47, 2493 (1969).

68. H.A. Jahn and E. Teller, Proc. Roy. Soc., A161, 220 and A164, 117 (1937); R. Renner, Z. Physik, 92, 172 (1934).

69. A. Veillard, in "Internal Rotation in Molecules " (editor W.J. Orville-Thomas), Wiley, London, 1974.

70. R.K. Nesbet, Journal of Chemical Physics, 43, 530 (1965) and Advances in Chemical Physics, 9, 321 (1965).

71. L.C. Snyder and H. Basch, Journal of the American Chemical Society, 91, 2189 (1969); W.J. Hehre, R. Ditchfield, L. Radom, and J.A. Pople, ibid., 92, 4796 (1970).

72. D.B. Cook, "Ab initio Valence Calculations in Chemistry", Wiley, London, 1974.

73. B.J. Duke, "Electronic Calculations on Large Molecules", a Specialist Periodical Report of the Chemical Society in Theoretical Chemistry, Vol. 2, London, 1975.

CHAPTER 3

PRACTICAL MOLECULAR WAVEFUNCTIONS

"In spite of the spectacular expansion of the computer capability, there is a lingering, if not growing, pessimism about the feasibility of nonempirical calculations for polyatomic molecules."

S. Huzinaga

The previous chapter has defined the chemical and mathematical nature of orbital theories in quantum chemistry; in this chapter, it is intended to examine the equations involved from a practical point of view. It will be shown that further approximations must be made to render the methods outlined previously feasible computational projects. Thus, from an implementation viewpoint, there must be considered the precise nature of the quantities involved and defined in the equations, the numerical techniques required for the calculation of such quantities, and the organisation of the whole computation. This chapter gives an examination of the effect of such considerations on the choice of basis functions and atomic orbitals; approximations of computational convenience are introduced into the model approximations presented previously.

A. Molecular Integral Considerations.

The definition of the RHF equation (48) of Chapter 2 and associated quantities shows the involvement of integrals including the AO's (or basis functions) and the one and two electron operators of the molecular Hamiltonian:

$$\int \mu_i(x_i)\hat{h}\mu_j(x_i)\,dx_i \qquad -(1)$$

$$\text{and } \iint \mu_i(x_i)\mu_j(x_i)\hat{g}\mu_k(x_j)\mu_\ell(x_j)\,dx_j\,dx_i \qquad -(2)$$

where the orbitals μ_i may be basis functions η_i or atomic orbitals χ_i. Integrals (1) are known as "one-electron" or "core" integrals, and (2) as "two-electron" or "repulsion" integrals. If there are m functions μ_i, then there are $\frac{1}{2}m(m+1)$ molecular integrals of type (1) to be evaluated, as a result of the symmetry properties of operator $\hat{h}$. The operator $\hat{g}(i,j)$ is simply a multiplying factor in the integrand and so the number of distinct molecular integrals (2) is $^1/8(m^4+2m^3+3m^2+2m) \equiv N_m$. Thus, for the specific example of the benzene molecule, with m = 36 (5 AO's centred on each carbon atom and 1 on each hydrogen atom), $\frac{1}{2}m(m+1) = 666$ and $N_m = 222,111$. There are a very large number of molecular integrals, particularly of type (2), to be evaluated during an orbital basis valence calculation, so that very efficient methods of computing these integrals are required, or the orbitals μ_i have to be chosen to ensure rapid integral evaluation.

(a) Approximate Atomic Orbitals.

In order to preserve the valence analysis of the molecular wave-function it is necessary to work with AO's χ_i, which, in practice, are expanded in terms of a larger set of basis functions η_i and the valence calculation is either performed directly in terms of the η_i or indirectly through the χ_i. Thus, m should be the number of basis functions - the primitive elements of the orbital model. Atomic calculations on elements in the first row of the periodic table have shown that the AO's can be adequately expressed in terms of about 18 basis functions of the Slater Type Orbital form (as in Chapter 2, equation (38)).[1] The benzene example now becomes rather more daunting, with 4656 one-electron integrals and an overpowering $N_m \sim 10^7$ electron repulsion integrals. These numbers show that two types of problem are presented by the molecular integrals:

(i) how are 10^7 integrals to be computed in reasonable time?

(ii) where and how can the computed integrals be stored for future use in the valence calculation?

The solution to the above requires two different techniques. One is a further approximation in the AO model; the other is a computational device.

Practical savings can be made by using AO's which are not full solutions of the atomic RHF equations.[2] One way to form approximate AO's is to reduce the expansion length of each AO χ_i in terms of the basis functions η_i. By carefully choosing the orbital exponent α in the STO basis functions, $\chi_i \sim r^{\nu} \exp(-\alpha r) \times$ (spherical harmonic), it is possible to express each AO approximately as just one term, when the distinction between basis functions and AO's disappears. The radial part of an exact AO and the best one term STO approximation to it are shown in Figure 1. It can be seen that the main features of the electron distribution are quite well reproduced by the approximate AO. STO's are particularly suitable for short expansions of AO's because of their similarity to hydrogenic orbitals. Unfortunately, this advantageous property of STO's, of furnishing physically reasonable AO's with few primitive functions, is counterbalanced by their giving rise to molecular integrals of types (1) and (2) which cannot be evaluated by standard analytical techniques and methods of direct numerical integration have to be used.[3] Numerical quadrature in many dimensions is particularly time consuming, and, for all but the largest and most powerful computing facilities (or the smallest molecular systems), the routine evaluation of molecular integrals using STO's is out of reach. Thus, to solve problem (i) above, the over-riding factor is the use of basis functions which are known to lead to analytically tractable molecular integrals.

(b) Gaussian Basis Functions for Many-Electron Molecular Wave Functions

Nowadays, as a result of the computational simplicity of Gaussian Type Function (Orbital), GTF(O), molecular integral expressions, the use of GTF's dominates the field of molecular valence calculations. The historical development of the use of GTF's is of particular interest in that it exemplifies the interaction of theoretical and empirical reasoning.

GTF's,which have the general form $r^{\nu} \exp(-\alpha r^2)$ x (spherical harmonic), do have the property that, when used as basis functions η_i, they define molecular integrals which are easily evaluated; in addition, the functional form of GTF's to some extent parallels that of STO's - the exponential "decay" at large r. Thus, GTF's would seem to hold out great promise for molecular calculations. However, after the initial proposal of GTF's by Boys in 1950,[4] numerous discouraging calculations using Gaussians were carried out over the next 10 years.[5] In the treatment of small molecular systems then, each AO was represented by a single Gaussian; it was found that the substitution of a single Gaussian for a single exponential function (STO) was a very poor procedure. Consequently, Gaussians were not extensively explored for use in atomic and molecular calculations. However, it was gradually becoming obvious that, although the effort to reach Hartree-Fock solutions by analytical expansion with STO's had been rewardingly successful in lighter atoms and diatomic molecules, no conspicuous change or breakthrough in mathematical analysis was apparent to ease the evaluation of general many-centre molecular integrals with STO's with adequate accuracy in reasonable time. Also, it was seen that a more adequate basis set than a single STO one was needed. Thus, in 1963, there appeared the first paper to report the results of an attempt to establish a systematic procedure for calculating molecular

wavefunctions in a basis of GTF's.[6] Since that time, extensive research has been carried out to obtain empirically sets of GTF's which provide practical, physically reasonable bases for the expansion of AO's.

The basic property of GTF's which is the crucial one in relation to use as a basis function in valence calculations on polyatomic molecules is as follows: the product of two Gaussian functions G_a and G_b centred on different points a and b is itself a Gaussian function, centred at e somewhere on the line joining these two points. Therefore a three- or four-centre integral may be reduced to a two-centre integral:

$$\int G_a G_b \left| \frac{1}{r_{12}} \right| G_c G_d \equiv \int G_e \left| \frac{1}{r_{12}} \right| G_f.$$ [7] This is the reason why the necessary multicentre integrals involved in a molecular calculation become much simpler to calculate when a basis of GTF's is used. Additionally, GTF's can form a complete set mathematically, meaning that any function can be expanded as a linear combination of appropriate GTF's. Offsetting the mathematical advantages of GTF's, is their physical form, which is outlined in Figure 2. A Gaussian does not resemble very closely the form of a real AO. In particular, the Gaussian lacks a cusp at the nucleus and hence the region near the nucleus is described rather poorly. The behaviour at large distances is also very different from that of an exact AO. When GTF's are used in atomic calculations, it has been found that good approximate AO's can be obtained, but the length of the expansion - the number of η's per χ - is greater than for the more physically realistic STO's, so that a basis of GTF's must be larger than an STO basis for AO's of the same quality by a factor of two to three, typically. Figure 3 illustrates a short expansion of an AO compared with the accurate AO. The main source of error in GTF expansions of AO's is in the region around the nucleus, which is not likely to be heavily involved in molecule formation, indicating that no gross errors in molecular electronic density changes are introduced by using AO expansions in terms of GTF's.

During the 1960's, a great amount of effort was put into deriving, empirically, basis sets of GTF's by performing calculations on atoms.[8] With a GTF expansion of an AO, $\chi_i = \sum_j C_j^{(i)} \eta_j$ -(3) the RHF equations for a particular atom, as given in Chapter 2, can be solved for the coefficients $C_j^{(i)}$, while simultaneously varying the orbital exponents α to obtain a minimum energy for the atom. The optimisation of the "non-linear" parameters α is a time-consuming task, as the orbital exponents are varied and for each set of α's the RHF equations are solved. The final optimum α's and expansion coefficients are those giving the minimum atomic energy. An alternative, and more widely used, procedure is to fit a linear combination of GTF's to a known AO (or approximate AO) by, for example, a least squares procedure. An approximate GTF expansion, $\tilde{\chi}$, of an AO χ can be written

$$\tilde{\chi} = \sum_{j=1}^{k} C_j^{(i)} \eta_j$$

where k is the expansion length. The minimisation of

$$\int | \chi - \tilde{\chi} |^2 \, dr$$

with respect to the coefficients $C_j^{(i)}$ and orbital exponents of the η_i optimises the expansion in the usual least squares sense. A very common "hybrid" of these two methods of using approximate GTF expansions of AO's is to write each AO as an optimum single STO and to expand this STO in terms of GTF's; this rather roundabout approach to approximate AO's is justified by computational convenience and historical development of quantum chemistry. In addition, the GTF expansion of STO's has a very valuable convenience property.[9] For a given expansion length and orbital type, the optimum coefficients in the expansion are independent of the STO orbital exponent (α), depending only on the functional form of the STO (chapter 2,equation (38)), and also the orbital exponents

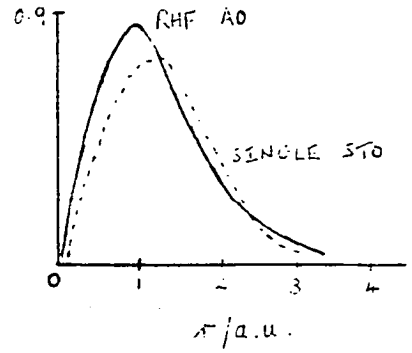


FIGURE 1 : 2p ORBITAL OF C

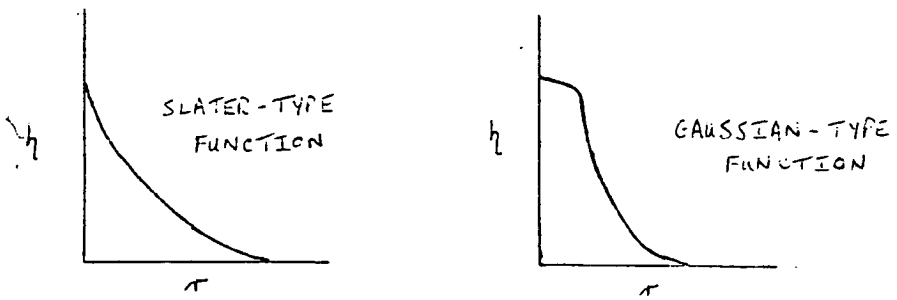


FIGURE 2 : GENERAL FORMS OF BASIS FNS.

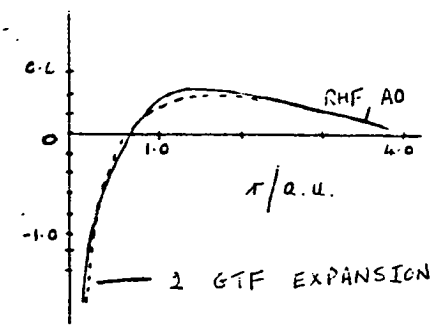


FIGURE 3 : 2s ORBITAL OF Be

of the GTF's η_i are essentially independent of α. Thus, the GTF expansion of STO's can be done once and for all for a given expansion length k. This expansion method has been carried through for the full range of STO functions using a wide variety of expansion lengths.

Atomic calculations using the RHF model are only a part of the development of GTF basis sets in valence calculations; in addition, calculations on molecules have contributed important principles.

(c) Contraction Techniques

Problem (i) of section (a) above can be solved by using basis sets of GTF's. Problem (ii) can be solved by a computational technique, as described below.

When each AO of a molecular system is expanded in terms of k GTF's

$$\chi_i = \sum_{j=1}^{k} c_j^{(i)} \eta_j,$$

then there are k^4 contributions to an AO electron repulsion integral

$$(\chi_i\chi_j,\chi_k\chi_l) = \sum_{r,s,t,u=1}^{k} c_r^{(i)} c_s^{(j)} c_t^{(k)} c_u^{(l)} (\eta_r\eta_s,\eta_t\eta_u) \qquad -(4)$$

to be computed in general. Similarly, there are k^2 contributions to the one-electron integrals. Equalities among $\chi_i,\chi_j,\chi_k,\chi_\ell$ may reduce their number in special cases, but the factor of k^4 is representative. As noted in section (a), if all these GTF integrals are computed and stored, huge amounts of storage space are required. For m AO's expanded in terms of k GTF's, $(m.k)N_m$ numbers must be stored. If, however, the m AO's are regarded as the essential degrees of freedom in the calculation, and not the mk basis functions, there is no point in storing the GTF integrals separately since only those combinations forming the AO integrals are

ever required. Thus, in practice, all the contributions in equation (4) are computed consecutively, the summation is performed and the whole AO integral is stored as one number. Storage space is then reduced to the requirement for $m.N_m$ electron repulsion integrals. Similar, but less spectacular, savings can be made in the computation of the one-electron integrals. This method of using fixed linear combinations of GTF's and storing only AO integrals is called "contraction", and, although mathematically trivial, it is of enormous computational value.[10] Many calculations on the electronic structure of large molecules would be quite impossible without contraction. The use of the contraction techniques excludes the MO method whereby the basis functions are used directly in the molecular RHF equation, as in Chapter 2, and so reduces the flexibility of the MO method.

Thus, the somewhat depressing conclusion, reached by solving problem (i) by introducing quite a large number of GTF's, that accurate calculations using GTF's might prove difficult in practice, is tempered by this method of reducing the number of variables in the SCF calculation with very little loss of accuracy. Instead of allowing all the coefficients of the basis function expansion to vary freely, certain coefficients are fixed relative to one another, thus forming groups of Gaussian functions. The MO is then expressed as

$$\phi_i = \sum_k c_k^{(i)} \gamma_k,$$

where γ_k is a small contraction of GTF's of the same type on the same centre,

$$\gamma_k = c_1' G_1 + c_2' G_2 + G_3' G_3,$$

for example.

In this way, a large basis set may be decomposed into a much smaller number of groups. In the variational calculation of the molecular wavefunction only the coefficient of the contracted Gaussian ($C_k^{(i)}$) is allowed to vary and not the relative proportions of the Gaussian within each group (C_i'). How much accuracy is lost as a result depends a great deal on the skill with which the initial basis of GTF's is contracted. The contraction process is largely a matter of using chemical intuition.[10]

At this point, all the necessary approximations involved in the MO model have been introduced (see Figure 4), and the outline of the development of a computationally feasible theory of molecular electronic structure is complete. Formally, in the calculations reported in this work, approximate MO's incompletely expanded in terms of approximate AO's are involved; physically, MO's are formed in a chemically realistic way using AO's which have all the main features of accurate atomic orbitals. The rather forbidding approximation "tree" culminates in the use in this work of Linear Combination of Gaussian Orbitals (LCGO) MO Calculations. It is now intended to describe the general strategy of such a computational method.

Figure 4

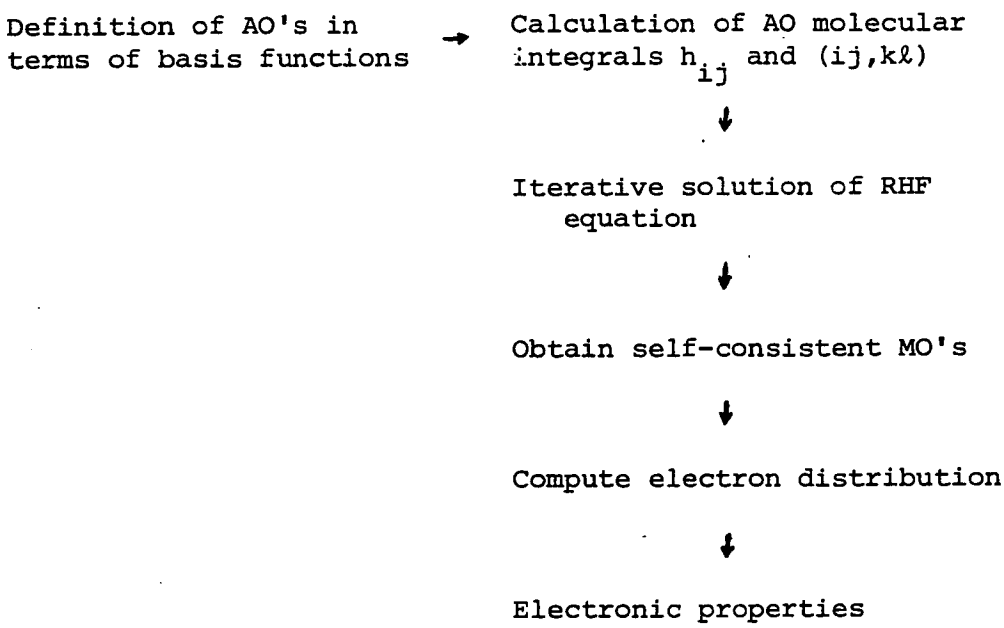
$\hat{H}\Psi = E\Psi$	Non-relativistic, fixed nucleus Schrödinger equation.
	Variational solution → "point properties" lost.
$\Psi = \sum_i C_i \Phi_i$	Orbital model - wavefunction expressed as a sum of determinants.
	Correlated motion of electrons lost here.
$\Psi = \det\{\Psi_1\overline{\Psi}_1 \ldots \Psi_{n/2}\overline{\Psi}_{n/2}\}$	One-configuration, "independent electron" model.
	Exact form of the optimum MO's lost here
$\Psi = \Phi T$ (matrix form)	LCAO expression for MO's.
	Exact form of optimum AO's lost here
$\Phi = \chi C$ (matrix form)	AO's expanded in terms of a set of basis functions.
	Correct asymptotic behaviour of the AO's lost here.
GTF basis	Forms of the basis functions chosen on computational grounds.

Summary of the Hierarchy of Approximations Involved in the MO Model.

B. The Modular Structure of Molecular Calculations.

The transition from the derivation of a formal set of equations to the working out of a strategy for their solution often requires a complete change of orientation. The procedures used for handling the numerical quantities involved in the solution of the "orbital model Schrödinger equation" seem very far from those associated with a partial differential equation, and from those which might be inferred from the chemical basis of the mathematics. Figure 5 shows the steps involved in performing molecular valence calculations by the MO method.[11]

Figure 5



The Modular Structure of the Computational Problem.

(a) Computation of Molecular Integrals

In Section A above, it was shown that the computational simplicity of the GTF molecular integral expressions is the over-riding factor in the choice of basis functions for use in calculations of approximate

molecular wavefunctions; this is summarised in the first step of Figure 5. Having chosen a suitable basis, the next step in the calculation is the computation and storage of the molecular integrals. For a chosen nuclear geometry, the same AO integrals can be used for any calculation on the electronic structure of the molecule or the related radicals and ions. In practice, this step is the most time-consuming in the calculation, so that efficient computation and flexible storage of the molecular integrals is involved. There has been a large amount of research in the two areas of rapid, accurate computation of integrals and the design of an efficient file structure for storing these for subsequent retrieval and use.[12] The electron repulsion integrals present the most difficulties from both points of view; in view of their relatively small number, the one-electron integrals are not of critical importance in this regard. In practice, the integrals are kept on a file external to the computer's main store; in Appendix 2, the general problems involved in molecular integral evaluation are outlined, and the main features of the GTF integral derivation and working formulae for each type of integral are given.

(b) The Matrix LCAO MO Equations

The molecular integrals are elements of the matrix H^F in the RHF equation (48) of Chapter 2; the linear variational problem, equation (32) of Chapter 2 was converted to the matrix diagonalisation problem - the computation of the eigenvalues and eigenvectors of the matrix $H^F = H + G(R)$. Thus, the RHF equation is solved iteratively, since G(R) is a function of the eigenvectors, yielding a "self-consistent" solution in terms of LCAO SCF MO's. The iterative process of solution must be started from an initial (physically realistic) guess at the

eigenvectors T, or the matrix R (= $TT^{\dagger}$). Using this guess at the electron distribution, the matrix H^F is formed from the stored integrals H_{ij} and (ij, kl). Computation of the eigenvectors of this matrix gives a new matrix T (and R) which defines a new matrix H^F. This process is repeated until the new T (or R) matrix does not differ from its predecessor by more than some tolerance decided (on physical grounds) in advance, i.e. T is self-consistent. Hopefully, in practice, the successive T(R) matrices converge to a final self-consistent solution rather than oscillating or diverging away from the true solution.

The computational problems in implementing this iterative scheme are:
(i) an algorithm for the efficient formation of G(R) from the stored electron repulsion integrals;
(ii) a matrix diagonalisation method;
(iii) matrix manipulation routines.
The generation of G(R) is an integral part of the method used for storing the electron repulsion integrals, and is discussed in Appendix 2. The computation of eigenvalues and eigenvectors of a symmetric matrix (as H^F is) is a well-researched problem of applied mathematics, and Appendix 3 outlines one "class" of methods for symmetric matrix diagonalisation; the matrix LCAO MO problem, $H^F T = T\varepsilon$, is equivalent to the problem of finding matrix U (unitary, as $UU^{\dagger} = 1$), m x m, containing T, m x $^n/2$, such that $U^{\dagger}H^F U = \varepsilon$, i.e. matrix H^F is transformed into diagonal form (ε is diagonal matrix of ε_i).

There are some complications which have been omitted in the treatment of the LCAO MO method given above. No allowance has been made for non-orthogonality of the AO's, which is actually the situation which exists in practice. Thus, strictly, the equation to be solved iteratively is

$H^F T = S T \varepsilon$ - equation (50) of Chapter 2, where S is the overlap matrix, defined by $S_{ij} = \int \chi_i \chi_j dr$. However, the formation of an orthogonal AO set is an orbital transformation similar to the computation of LMO's as given in Chapter 2, section B; it can be shown that no new numerical methods are required in the non-orthogonal case in LCAO MO calculations, and that the "overhead" resulting is simply three more matrix multiplications during each iteration[13] - significantly, transformation of H^F is required, but the transformation of the electron repulsion integrals to the orthogonal basis is not necessary.

Another complication is that the simple iterative scheme outlined above does not always converge to a self-consistent T matrix (oscillation, rather than divergence, is typical, often interpretable physically as the existence of two possible electronic distributions with very similar energies). Usually, as in this work, cases are "well-behaved", but extensive research has been done to derive mathematical procedures to induce convergence in difficult cases.[14] Although the contraction technique introduced in section A is rendered necessary by integral storage considerations, it also tends to facilitate the convergence to self-consistency as the dimensionality of matrix H^F is reduced by reducing the number of primitive basis functions to the number of AO's.

At the end of the iterative calculation, the final T matrix contains all the information about the molecular electron distribution within the LCAO model. The analysis of the electron density through the matrix R gives the physical and chemical information contained in the single configuration wavefunction defined by T. The diagonalisation of the m x m matrix H^F gives m eigenvectors; $(m - \frac{n}{2})$ eigenvectors contained in U are not part of T. These "unoccupied" columns of U are called Virtual Orbitals. They can be interpreted as follows:

the m orbitals of U, when placed in order of increasing ϵ_i, are all possible levels for the electrons of the molecule to occupy and the ones unoccupied in the ground state determinant are approximations to the excited orbitals, or, more strictly, a one configuration wavefunction using one of the virtual orbitals in place of one of the columns of T is an approximation to an excited state of the molecule. However, as H^F is a function of T with G(R) containing the electron repulsion integrals between each MO and all the other occupied MO's, the virtual orbitals are computed experiencing the electrostatic field of all the occupied orbitals, and so they are closer to the approximate excited orbitals of the negative ion of the molecule than to those of the neutral molecule. Although the variation principle does not act on them during the iterative solution of the RHF equation, the virtual orbitals do presumably have the general form of the excited orbitals for a many-electron system since one additional electron does not constitute a very severe perturbation to the molecule (cf. Koopmans' Approximation, Chapter 2, section B). The virtual orbitals will probably continue to be used as approximate excited orbitals since they have the all-important, computational advantage of being orthogonal to the occupied MO's as they are both part of the same U matrix.

The final steps of Figure 5, the analysis of the molecular wavefunction obtained in terms of SCF MO's by iterative solution of the RHF equation, are outlined in Chapter 2, section B.

C. Preliminary LCGO Calculations.

Having thus far outlined a practical procedure for obtaining molecular electronic wavefunctions, the final requirement for the

execution of it is the choice of a basis set of functions (Gaussian primitives). In practice, this step is crucial, and introduces a degree of empiricism into so-called non-empirical calculations. It is instructive, at this point, to give some consideration to calculations on atomic species. Molecular orbital theory, and RHF theory in particular, has the advantage of being conceptually based on atomic theory, with techniques that can be tested for atoms; the particular method used in this work to compute molecular wavefunctions is applicable to the limiting case of a single atom. The interest in atomic calculations here is that they can give indications of the adequacy of basis sets for molecular calculations, for which they also provide some background material. Nevertheless, in considering basis set construction for calculations on large molecules, it is necessary to use small molecular species as test cases.

(a) Calculations on Atomic Species

For atomic species of first- and second-row elements, at least, there is a wealth of experimental information, and "ab initio" calculations of the type in Chapter 2 can be carried out relatively easily so that extensive comparisons of data can be made. More particularly, in this work, calculations within the Hartree-Fock model are of importance. In this section some concrete examples are given to illustrate the abstract formalism thus far presented, and also the performance of the Hartree-Fock solution which represents the best attainable description of the electronic structure of a many-electron system in terms of the one-electron orbital approach. Numerical integration techniques may be used to solve the Hartree-Fock equations in the case of atoms by the iterative technique of Chapter 2, yielding true H-F orbitals.[15]

However, the lower symmetry of the nuclear field present in molecules necessitates the use of an expansion for the determination of the molecular orbitals. In Roothaan's method, it is assumed that each MO may be adequately represented by a linear expansion in terms of some (simpler) set of basis functions. Use of a complete (necessarily infinite) set of basis functions in the expansion would ensure absolute convergence to the well-defined H-F limit. In practice, only a finite number of basis functions can be employed, and the selection of such a set is of crucial importance in determining how closely the H-F solution is approximated.

The minimum size basis set which can be used in an SCF calculation includes one basis function for each occupied atomic orbital with distinct n and ℓ values as determined by the electronic configuration of the atom in question (with reference to hydrogenic AO's of table 1 of Chapter 2). Thus, the use of a single exponential basis function (Slater Type Function, equation (38) of Chapter 2) for each atomic shell leads to total energy errors in atomic calculations of approximately 1%.[16] A great improvement, over such a minimal or single-zeta basis set, in calculated energy and properties is obtained by doubling the size of the basis set, i.e. two STO's per occupied AO, yielding a double-zeta basis set.[17] In this way, atomic SCF calculations can yield results close to the H-F limit, for atoms from He to Xe.[18] For first-row atoms, a STO basis consisting of 6 s-functions and 4 p-functions is enough to give seven-figure accuracy in total energy for the ground states. It is instructive, especially with prospective molecular calculations in mind, to see how many Gaussian Type Functions (Section A) are necessary to achieve more or less the same accuracy in atomic calculations. Physically, Gaussians are inferior to Slater

functions, taken one for one, in representing AO's, and computationally it is found that between two and three times as many GTO's as STO's are required for comparable accuracy.[19] The number of integrals to be evaluated in a calculation is proportional to the fourth power of the number of basis functions; nevertheless, a calculation using GTO's is in general less time consuming than the corresponding STO one as the computation of a typical integral over primitive GTO's can easily be 100 times faster. In addition, in basis set considerations, the practical limitations involved in integral storage and SCF convergence render the use of the contraction technique necessary (Chapter 2),[20] in molecular calculations. Thus, a contracted GTO basis, yielding results of comparable accuracy to those from a double-zeta STO basis, can be obtained for first- and second-row atoms with only a few contracted functions required. However, in molecular wavefunction calculations on large molecules basis sets of more limited size than those mentioned above must be used. In Appendix 4, there are listed the basis sets of Gaussian primitives used in this work, for the selection of first- and second-row atoms considered here. These Gaussian basis sets, consisting of 7 s-type and 3 p-type functions for first-row atoms and 10 s-type and 6 p-type for second-row atoms, are those of Roos and Siegbahn,[21] obtained by optimising the Guassian function exponents in an atomic SCF calculation by minimising the total atomic energy; the basis for the hydrogen atom is that of Huzinaga,[22] consisting of 3 s-type functions obtained by fitting to a STO. There is also shown the contraction of these functions into a minimal basis set, used in most of the molecular calculations reported here. Tables 1 and 2 summarise the results of atomic calculations for the atoms of interest in this work, and show how the basis set used in this work compares with other ones. It is found that the SCF energy,

since it always represents an upper bound to the true energy, converges smoothly to the limiting H-F value as the size of the basis set is increased, but that the values of other properties may fluctuate erratically around the H-F value until this limit is very nearly attained.[22]

Calculations on atoms can be used to obtain the Gaussian primitives of the full uncontracted basis set. They can also be used in producing a contracted basis, i.e. in obtaining the contraction coefficients of each group of primitives. With the use of the contraction procedure, it is possible to retain nearly all of the accuracy of the original primitive set while substantially reducing the number of integrals which must be handled in the iterative solution of the matrix H-F equations. Contracted Gaussian basis sets suitable for use in molecular calculations can be obtained from atomic calculations if the nature of the molecular environment is taken into account; flexibility in the valence regions must be combined with an adequate description of the atomic inner shells.[23] Practically, the contracted functions are chosen through a careful analysis of the atomic expansion coefficients, by trying to find appropriate linear combinations of the primitive functions, and the results of extensive research are reported in the literature.[24] Tables 3 and 4 show, for the carbon atom, the effects of different possible contractions on calculations using the Gaussian primitives of Roos and Siegbahn. Such atomic calculations are of interest from a comparison point of view, since contraction is of practical interest only for molecular calculations.

Atomic calculations are significant because they can give useful insight into the physical nature of the mathematical basis sets used. Optimisation of finite basis sets to be used in general molecular computations is not feasible in practice as part of the molecular

calculation, but only is reasonable for the individual atomic basis sets. Even so, much s-p saturated sets which can yield H-F results for atoms up to scandium must be augmented with "polarising" functions (absent in the description of isolated atoms but whose presence is essential for a proper description of distortions in charge distribution caused by chemical bonding) in order to approach the H-F limit for molecules.[25]

(b) Calculations on Molecular Species

The properties of the Roos and Siegbahn Gaussian basis have been exhibited above by giving some expectation values calculated for atomic species. To give further information on the performance of such a "best-atom" basis when applied in a molecular environment, it is instructive to consider calculations on small test molecules. Of particular interest in this work, where the major class of molecules of interest is that of organic "pi" systems, are comparative results on ethylene, acetylene and methane; much research has been done on other types of small molecule,[26] but these three are regarded as prototypes in this work and serve to illustrate some general principles. At this point, the calculations reported are viewed with basis set considerations in mind; further significance of the results, chemically, for these and other small molecules is a consideration of the following chapter. Table 5 presents the results of electronic wavefunction calculations for the three molecules, and values for various molecular properties from different sources can be compared.

If "ab initio" MO treatments are to be extended to large molecules of chemical interest, it is necessary for tractable computation times,

as well as for convenient analysis of electronic wavefunctions in terms of traditional chemical concepts, to have suitable minimal contracted GTF bases, i.e. basis functions in one-to-one correspondence with the atomic core and valence electrons. In the approach adopted in this work, it is considered advantageous to have electronic wavefunctions for many molecules in the same basis set, in the hope that errors or inadequacies of the wavefunctions tend to remain constant from molecule to molecule and thus tend to cancel in comparative studies of properties. To finally obtain an optimised minimal basis set from the Roos and Siegbahn "best-atom" contracted basis, calculations on small molecules were employed. The final basis sets for the atoms of interest here are presented in Appendix 4; these are "scaled" minimal bases, for use in calculations on general molecules. It is of interest to describe the optimisation procedure, and ethylene is considered in illustration. The initial "best-atom" sets for carbon and hydrogen were obtained by consideration of atomic SCF calculations, firstly yielding optimised exponents and secondly yielding contraction coefficients. To obtain a minimal basis set optimised for ethylene, the best atom coefficients were retained, but the carbon 2s,2p(not 1s) and hydrogen 1s exponents were scaled until a set was obtained with the molecular SCF energy at a minimum.[27] The scale factors and resulting exponents are given in Table 6; calculated energies and properties are compared in Table 7. It is clear that reoptimisation of exponents in the molecule improves calculated values considerably. It is believed that the ethylene-optimised set is a suitable basis for large "sp^2-hybridised" hydrocarbons, and that the procedure of scaling a prototype molecule is generally useful. An atom may have several scaled sets, one for each type of

environment found in a molecule (e.g. corresponding to sp^2- and sp^3-carbon atoms). It is well established that the use of scaled functions, where the scaling factors are determined from molecular energy optimisation on related but smaller molecules and transferability to a range of larger molecules is assumed, leads to molecular energies substantially closer to the Hartree-Fock limit, for small basis sets. Physically, scaling of the valency shell AO's is a partial optimisation of the form of the AO's in a molecular environment, and partially compensates for the loss of flexibility in using a minimal basis set.

(c) The Standard Gaussian Basis Set

A molecular electronic wavefunction is obtainable by the solution of the Schrödinger electronic wave equation ((17) of Chapter 2). For many-electron systems, a direct attack on this differential equation has not led to sufficiently accurate approximations. This difficulty, however, can be resolved by translating the problem into an equally valid integral form, through the use of the Rayleigh-Ritz energy-variational principle, whereby approximate solutions may be obtained to any desired degree of accuracy for a many-electron system. Because of the formidable computational problems involved in deriving these solutions, chemical accuracy (implying errors of a few kJ mol^{-1} or less) has been achieved only in systems containing few electrons. In Chapter 2, it was shown that model approximations lead to the well-defined Hartree-Fock scheme as a practicable method for obtaining electronic wavefunctions. An efficient way to solve the H-F equations for polyatomic molecules is to expand the solutions in large sets of nuclear-centred Gaussian basis functions, which combine the two desirable, though apparently incompatible,

characteristics of tractability to the appropriate mathematical analysis and production of physically reasonable results. A further critical approximation, of a practical nature, must be made for computational convenience, that of using contracted sets of GTF's. In practice, as outlined in this Section, it is desirable to derive a finite basis set, only with a small sacrifice in flexibility yet saving considerably computationally, so that the true H-F solution can be closely approximated. The selection of an appropriate basis of GTF's for a molecular calculation presents somewhat of a dilemma. Basis sets with fully optimised exponents for the separate atoms at the H-F limit have been derived, but their use in general molecular calculations is precluded by their size, even if the contraction technique is used - the number of integrals stored is then reduced, but the number calculated remains the same. It is therefore necessary to use optimal Gaussian atomic bases of more limited size. In this work, the basis sets of Roos and Siegbahn for first- and second-row atoms, and Huzinaga for the H atom, are the standard ones chosen. For first-row atoms, there are 7s- and 3p- type primitive functions; for second-row atoms, 10s- and 6p-type; for H, 3s-type. This well-balanced type of basis (relative to adequacy of s- and p-type descriptions)[28] is of a size and nature decided upon by a compromise between accuracy of computation, relative to the H-F limit, and the expense of computer time for integral evaluation and SCF calculation (also, possibly, tractability of SCF calculation). Calculations on atomic and small molecular species are given above to exhibit the properties of these basis functions (primitives), and their behaviour when contracted into groups. Finally, a minimal basis of contracted functions, empirically chosen, is partially optimised for use in a molecular environment by a scaling technique, yielding the standard basis set used in the calculations on large molecules reported in the following chapters

(given for the relevant atoms in Appendix 4). Studies of small polyatomic systems have provided indications of the sensitivity of molecular properties to the choice of basis, and the reliability of the H-F approximation itself. Small basis set molecular treatments, as in this work, are on a more tenuous theoretical foundation. Generally, the successful calculation of a series of molecules is required to achieve confidence that the basis is sufficiently well balanced to describe the energetics and molecular properties even approximately.[29]

References for Chapter 3

1. J.C. Slater, Physical Review, 36, 57 (1930); E. Clementi, C.C.J. Roothaan and M. Yoshimine, Physical Review, 127, 1618 (1962).
2. Solutions of the RHF equations for atoms He through Xe are described and tabulated in E. Clementi and C. Roetti, Atomic Data and Nuclear Data Tables, 14, 177 (1974).
3. Direct numerical solution of the H-F equations is routine for atoms; partial numerical integration yields very good results for diatomic molecules - E.A. McCullough, Chemical Physics Letters, 24, 55 (1974).
4. S.F. Boys, Proceedings of the Royal Society, A200, 542 (1950).
5. L.C. Allen, Journal of Chemical Physics, 37, 200 (1962).
6. C.M. Reeves, Journal of Chemical Physics, 39, 1 (1963).
7. I. Shavitt, "Methods in Computational Physics", Volume 2, Academic Press, London and New York, 1963.
8. S. Huzinaga, D. McWilliams and B. Domsky, Journal of Chemical Physics, 54, 2283 (1971); J.D. Petke, J.L. Whitten and A.W. Douglas, ibid, 51, 256 (1969); D.R. Whitman and C.J. Hornback, ibid, 51, 398 (1969); A.J.H. Wachters, ibid, 52, 1033 (1970).
9. C.M. Reeves and R. Fletcher, Journal of Chemical Physics, 42, 4073 (1965); K. O-ohata, H. Taketa and S. Huzinaga, Journal of the Physical Society of Japan, 21, 2306 (1966); W.J. Hehre, R.F. Stewart and J.A. Pople, Journal of Chemical Physics, 51, 2657 (1969); R.F. Stewart, ibid, 52, 431 (1970).
10. E. Clementi and D.R. Davis, Journal of Chemical Physics, 45, 2593 (1966); J.M. Schulman, J.W. Moskowitz and C. Hollister, ibid, 46, 2759 (1967).
11. D.B. Cook, "Ab initio Valence Calculations in Chemistry", Butterworths, London, 1974.
12. The considerations of Appendix 2 are rendered operational in the large-scale molecular wavefunction programme "packages" now available - of particular relevance in this work is the ATMOL suite, Atlas Computing Division, RHEL (SRC, 1976).

13. P.-O. Löwdin, Advances in Quantum Chemistry, 5, 185 (1970).

14. V.R. Saunders and I.H. Hillier, "International Journal of Quantum Chemistry", 7, 699 (1973).

15. Numerical orbitals and energy quantities for all the atoms from Hydrogen to Lawrencium are given by J.B. Mann, LATR LA3690 and LA3691 (1962).

16. E. Clementi, D.L. Raimondi and W.P. Reinhardt, Journal of Chemical Physics, 47, 1300 (1967).

17. E. Clementi, R. Matcha and A. Veillard, Journal of Chemical Physics, 47, 1865 (1967).

18. C. Roetti and E. Clementi, Journal of Chemical Physics, 60, 4725 (1974).

19. S. Huzinaga, Journal of Chemical Physics, 42, 1293 (1965).

20. T.H. Dunning, Journal of Chemical Physics, 55, 716 (1971); J.M. Schulman, C.J. Hornback and J.W. Moskowitz, Chemical Physics Letters, 8, 361 (1971).

21. B. Roos and P. Siegbahn, Theoretica Chimica Acta, 17, 209 (1970); Reference 19.

22. P.E. Cade, Trans.Amer.Cryst.Assoc., 8, 1 (1972).

23. T.H. Dunning and P.J. Hay, "Gaussian Basis Sets for Molecular Calculations", in "Modern Theoretical Chemistry" (editor H.F. Schaefer), Plenum, New York, 1976 (Volume 3).

24. C. Salez and A. Veillard, Theoretica Chimica Acta, 11, 441 (1968); A. Veillard, ibid, 12, 405 (1968); S.S. Seung, M.C. Harrison and I.G. Csizmadia, ibid, 8, 281 (1967).

25. B. Roos and P. Siegbahn, Theoretica Chimica Acta, 17, 199 (1970). T.H. Dunning, Journal of Chemical Physics, 53, 3958 (1971).

26. C. Thomson, in "Theoretical Chemistry" (Specialist Periodical Report of the Chemical Society), Volume 2 (1975), Volume 3 (in press).

27. M.H. Palmer and A.J. Gaskell, Theoretica Chimica Acta, 23, 52 (1971); M.H. Palmer, A.J. Gaskell and M.S. Barber, ibid, 26, 357 (1972).

28. R.C. Raffenetti, Journal of Chemical Physics, 59, 5936 (1973).

29. H.F. Schaefer, Annual Reviews of Physical Chemistry, 27, 261 (1976).

Table 1

ENERGY COMPARISON (a.u.) OF CONTRACTED MINIMAL BASIS

WITH HARTREE-FOCK CALCULATIONS

	Hydrogen (^1S)		CARBON (^3P)	
	MIN.	H-F	MIN.	H-F
E_T	-0.4970	-0.5000	-37.6105	-37.6886
E 1s	-0.4970	-0.5000	-11.25443	-11.32552
E 2s			-0.67695	-0.70563
E 2p			-0.41311	-0.43334

	Nitrogen (^4S)		Oxygen (^3P)	
	MIN.	H-F	MIN.	H-F
E_T	-54.2754	-54.4009	-74.6121	-74.8094
E 1s	-15.54489	-15.62892	-20.56811	-20.66854
E 2s	-0.90633	-0.94523	-1.19083	-1.24428
E 2p	-0.53851	-0.56753	-0.58974	-0.63186

Table 2

ATOM ENERGIES IN SEVERAL BASIS SETS

	H	C	N	O
Present Uncontracted (7s3p)	-0.4998	-37.6563	-54.3390	-74.7000
Single-Zeta Slater	-0.5	-37.6224	-54.2689	-74.5404
Double-Zeta Slater	-0.5	-37.6867	-54.3980	-74.8043
Large Uncontracted GTO(11s7p)	-0.4999	-37.6881	-54.4001	-74.8081

Table 3

EFFECT OF CONTRACTION ON C ATOM ENERGIES

		E_T	E1s	E2s	E2p	Virial Ratio (-V/T)
I	Complete cont. (i.e. minimal)	-37.61049	-11.23490	-0.56111	-0.41311	1.971
II	2p only uncont.	-37.61059	-11.24559	-0.56502	-0.41483	1.972
III	2s only uncont.	-37.61942	-11.30292	-0.57572	-0.43194	1.987
IV	Valence Shell uncont.	-37.61944	-11.29845	-0.57390	-0.43138	1.987
V	1s only uncont.	-37.65241	-11.29782	-0.58640	-0.42127	1.998
VI	Double Zeta (Lowest Exponent Repeated)	-37.65348	-11.30739	-0.58492	-0.42	2.000
VII	Totally uncontracted	-37.65636	-11.30276	-0.58651	-0.42171	2.000

Table 4

EIGENVECTORS OF UNCONTRACTED FUNCTIONS

II	0.111	0.460	0.629	(2p)		
III	[-0.302(1s)]	0.449	0.649	(2s)		
IV	[-0.301(1s)]	0.448	0.650	(2s)		
	0.113	0.470	0.619		(2p)	
V	0.00483	0.0372	0.173	0.456	0.466	(1s)
VII	0.00482	0.0373	0.172	0.460	0.457	(1s)
	0.522	0.594				(2s)
	0.112	0.466	0.623			(2p)

Table 5

MOLECULAR ENERGIES IN SEVERAL BASIS SETS ($-E_T$/a.u.)

	Methane	Acetylene	Ethylene
Present Basis			
Unscaled (Best atom)	39.98584	76.44757	77.68925
Scaled (Ethylene)	40.10150	76.63130	77.82855
Scaled (Methane)	40.10323	-	-
DZ (Scaled)	-	-	77.93137
DZ (Unscaled)	-	-	77.93704
Large SP Basis	40.2016	76.8133	78.0160
Large SP Basis + POL.	40.2136	76.8482	78.0623
H-F Limit	40.225	76.86	78.08

Table 6

STANDARD MINIMAL BASIS SET (SCALED)

	Cont.Coeff.	Exponent	Scale Factor
C 1s	0.004813	1412.29	1.0
	0.037267	206.885	
	0.172403	45.8498	
	0.459261	12.3887	
	0.456185	3.72337	
2s	0.522342	0.557981	1.064
	0.594186	0.174021	
2p	0.112194	4.74919	1.135
	0.466227	0.966859	
	0.622569	0.226177	
H 1s	0.07048	6.99357	1.554
	0.40789	1.0587	
	0.64767	0.235235	

Table 7

COMPARISON OF SCALED AND BEST-ATOM BASES (C_2H_4)

	Best-Atom	Scaled	Extended (DZ+POLN.)
Total Energy (a.u.)	-77.68925	-77.82855	-78.0623
Orbital Energies (eV)			
$1a_g$	-310.6	-306.3	-305.625
$1b_{3u}$	-310.6	-306.2	-305.582
$2a_g$	-29.95	-28.07	-28.066
$2b_{3u}$	-22.76	-21.37	-21.778
$1b_{2u}$	-18.99	-17.78	-17.586
$3a_g$	-17.65	-15.99	-16.011
$1b_{1g}$	-15.11	-13.84	-14.006
$1b_{1u}$	-12.76	-10.86	-10.128
Populations H	0.80	0.86	0.87
C	6.40	6.28	6.26

CHAPTER 4

APPLICATION TO SMALL MOLECULES

"Without the experiences learned in the atomic, diatomic, and small molecule computational experimentation, we could not look forward to new advances in the chemical computation field."

E. Clementi

In the previous Chapter the results of calculations on some small molecules were presented in the context of considerations on basis sets. The intention of this chapter is to present calculations which can be used to determine the characteristics of the standard basis chosen; the small molecular species here can feasibly be studied by deriving more refined electronic wavefunctions than are obtainable with the rather modest standard minimal basis, but the aim is to examine such species more in the way of test cases so that in the application to larger molecules something can be said on the adequacy of the calculations in rationalising observed chemical behaviour and perhaps in making predictions. The preceding Chapters have described a method for calculating molecular electronic wavefunctions, indicating additional considerations involved in the implementation of it. Thus, armed with a basis set which describes the electron distribution of the free atoms constituting the molecule of interest, the final input requirement before commencing computation is the specification of the nuclear configuration, i.e. the geometrical structure of the molecule.

Within the Born-Oppenheimer Approximation, separating electronic and nuclear motions, the problem of solving the (time-independent)

Schrödinger Equation (14a of Chapter 2) for a molecule to obtain the electronic wavefunction reduces to the minimisation of the expression

$$E(X_1, X_2, \ldots, X_n) \equiv E(\underline{X}) = \int \Psi \hat{H} \Psi d\tau \qquad -(1)$$

subject to the constraint $\int \Psi \Psi d\tau = 1$. In expression (1), $\hat{H}$ is the fixed-nucleus electronic Hamiltonian of the molecule and Ψ is a trial wavefunction, which is an explicit function of the electronic variables and also a function of certain parameters denoted by the collective variables $\underline{X}$, with respect to which the minimisation is to occur. If further model approximations are made by regarding Ψ as being a (Slater) determinant of spin orbitals, ψ_i, each of which is an expanded function, $\psi_i = \sum_{i=1}^{m} C_i \phi_i$ with fixed expansion functions (AO's) ϕ_i and coefficients C_i as minimisation (variation) parameters, the constrained minimisation of E with respect to the C_i may be effected by solving

$$H^F C = SC\varepsilon \qquad -(2)\ ,$$

a matrix equation with $S_{ij} = \int \phi_i \phi_j d\tau$, C a column matrix of coefficients, H^F the Hartree-Fock matrix (Chapter 2) and ε the matrix of eigenvalues (MO energies). Thus, the electronic energy of a molecule, described by a closed-shell single-determinant function in the LCAO MO approach, is minimised with respect to the linear coefficients C_i (forming an (m x n) matrix for n occupied MO's); the non-linear programming problem of minimising the energy function E (equation (32) of Chapter 2) subject to $C^\dagger SC = I$ (unit matrix) is in practice transformed so that the resulting simultaneous equations are linear, producing the iterative SCF scheme of calculation according to the eigenvalue equation (2).[1] The individual optimal ψ_i, and hence Ψ, are determined in a single calculation as a result of the simple expression (2). In contrast, the

problem of minimising E with respect to the ψ_i (i.e. the AO orbital exponents, and contraction coefficients possibly) cannot be similarly transformed because of the complicated implicit dependence of E on ψ_i, so that in practice a sequence of calculations, each optimising E against some C_i, is required using discrete sets of exponents (contraction coefficients). The determination of the set that optimises E by this type of "brute-force" technique becomes impracticable for large molecules. As mentioned in the preceding Chapter, basis set optimisation is not feasible as a general procedure in calculations on large molecules.

The calculation of a molecular electronic wavefunction yields the corresponding eigenvalue, the total electronic energy of the molecule at the specified nuclear configuration. The total energy of the molecule is obtained by the addition of the potential energy term representing the nuclear repulsion, which is a constant for the "clamped" nuclei in the calculation. Internuclear distances occur in the Hamiltonian $\hat{H}$ of expression (1) and implicitly in Ψ; the minimisation of E with respect to these parameters again does not lead to a simple mathematical problem because of the implicit dependence. It is appropriate at this point to mention the problems associated with molecular geometry in this type of work, both from an experimental and a computational viewpoint.

A. Molecular Geometry.

(a) Experimental Considerations

There is now a general and lively appreciation of the limitations imposed by intramolecular vibrations on the observation and even on the definition of interatomic distances. Structure determination by gas electron diffraction[2] has been refined to the extent that the estimated standard deviation of a bond length in the most precise studies is about

0.001 Å. At this level of precision quite subtle effects become apparent, e.g. isotope effects on the vibrationally averaged value of the bond length. Furthermore, bond lengths derived from spectroscopic measurements may differ from electron diffraction bond lengths by an amount which is clearly outside experimental error. This arises because the spectroscopic and electron diffraction bond lengths are derived from observed quantities in which the effects of molecular vibrations are averaged in quite different ways. Another consequence of molecular vibrations is that the bonded and non-bonded interatomic distances observed by electron diffraction are not quite self-consistent, i.e. they do not correspond to a set of distances calculated from a rigid geometrical model, the so-called "shrinkage" effect.[3] In considering the inter-relation of molecular structures obtained from spectroscopic and electron diffraction data it is found that many different kinds of interatomic distance parameters can be defined; Table 1 presents the notation and brief definition of each of the important parameters.

The equilibrium structure of a molecule is conceived as the hypothetical vibrationless state, where all intramolecular modes of vibration are imagined as frozen at the minima of their potential energy curves; this concept depends on the Born-Oppenheimer Approximation. Internuclear distances in the equilibrium structure, denoted by r_e, are to a high degree of approximation (differences are of the order of 10^{-4} Å) isotopically independent,[4] making the equilibrium structure of central importance in molecular structure studies. All definitions of structural parameters in terms of some vibrationally averaged observed quantity are not isotopically independent, and this leads to problems in evaluating and in quoting structural parameters. Another reason for seeking the equilibrium structure is that it is this hypothetical vibrationless state which is fundamental to ab initio wavefunction calculations.

Other parameters of Table 1 are defined operationally. In an electron diffraction experiment, the data are obtained as an average over the vibrational motions of the molecule at a particular temperature. Two interatomic distance parameters are used in the reduction of the data. They are r_a (almost identical with the formerly used $r_g(1)$), derived directly from the experimental molecular intensity curve,[5] and r_g (synonomous with formerly used $r_g(0)$), obtained via the relation $r_g = r_a + u^2/r_e$ or, practically, $r_g \approx r_a + u^2/r_a$, where u is the mean amplitude of vibration for the given pair of atoms. The small correction u^2/r is typically about 0.002 Å. r_g has a particularly simple physical significance, being the average over the molecular vibrations of the internuclear distance; also, r_e can be derived from r_g in a straightforward manner, at least in principle, but a substantial improvement in experimental technique with respect to anharmonicity data determination is necessary before r_e values for polyatomic molecules will be readily available from electron diffraction work alone.

The operational definitions of bond length in rotational spectroscopy lead to the three parameters r_e, r_o and r_s, obtained from measured rotational constants directly without using vibrational corrections.[6] r_o and r_s have no simple physical interpretation for polyatomic molecules. The observed rotational constants for the ground vibrational state (or any other) are functions of the interatomic distances which are averaged in a complex and subtle way over the molecular vibrations. If the rotational constants can be extrapolated empirically to equilibrium ones, good r_e structures can be calculated from them, but this has so far proved possible only for simple polyatomic molecules.[7] If structural parameters are calculated from rotational constants for the ground

vibrational state, then the r_s structure is the most satisfactory; it is self-consistent and to a high degree of accuracy independent of isotopic species, although its physical significance is difficult to assess as it is defined only in terms of the operations needed to calculate it (Kraitchman's equation).[8] It appears that r_s is likely to be close to the desirable but often unattainable r_e for polyatomic molecules. r_s parameters cannot easily be compared with the physically well-defined r_g parameters. For polyatomic molecules, r_o structures' use should be tempered with caution.

It is possible, by applying corrections for some of the vibrational effects, to use the ground vibrational state rotational constants to calculate the r_z, or zero-point average, structure.[9] The r_z parameter corresponds to the mean positions of the nuclei so that it does have a precisely defined physical significance. It can be shown that, to a very good approximation,

$$r_z = r_e + \langle \Delta z \rangle_o \qquad -(3);$$

the two atoms concerned have equilibrium Cartesian coordinates (0,0,0) for atom i and $(0,0,r_e)$ for atom j, and by averaging the instantaneous displacements $\Delta x, \Delta y, \Delta z$ over the zero-point vibrations, the coordinates of the atoms' mean positions are $(\langle \Delta x_i \rangle_o, \langle \Delta y_i \rangle_o, \langle \Delta z_i \rangle_o)$ and $(\langle \Delta x_j \rangle_o, \langle \Delta y_j \rangle_o, r_e + \langle \Delta z_j \rangle_o)$, giving $\langle \Delta z \rangle_o = \langle \Delta z_j \rangle_o - \langle \Delta z_i \rangle_o$.[10]

r_g is the value of the instantaneous separation of a pair of atoms, averaged over a Boltzmann distribution of all possible energy levels of the molecular vibrations; it is found that

$$r_g = r_e + \langle \Delta z \rangle + (\langle \Delta x^2 \rangle + \langle \Delta y^2 \rangle)/2r_e \qquad -(4),$$

with the above notation. r_g is obtained at a particular temperature so that the vibrational average is taken over all vibrational states weighted by their Boltzmann factors; r_z is derived from a particular vibrational state, as indicated by subscript on average $\langle \Delta z \rangle_o$.[11] Because of the close parallel between r_z and r_g, it is convenient to define an electron diffraction parameter r_α by $r_\alpha = r_g - (\langle \Delta x^2 \rangle + \langle \Delta y^2 \rangle)/2r_e$ so that

$$r_\alpha = r_e + \langle \Delta z \rangle \qquad -(5),$$

where $\langle \Delta z \rangle$ is a Boltzmann-weighted average. The conversion of r_g into r_α can be performed if a harmonic force field is available for the molecule; in practice, the procedure required is well-established.[12] The r_α structure is geometrically consistent.

To complete the inter-relation of structural parameters from spectroscopy and electron diffraction, r_α is converted into an average for the zero-point vibrational state only by extrapolating to absolute zero, i.e.

$$r_\alpha^o = \lim_{T \to 0} (r_\alpha) \qquad -(6) .$$

The interatomic distances r_α^o are best calculated by first extrapolating r_g to absolute zero.[13] Thus, r_α^o and r_z represent the same physical quantity, the distance between the mean positions of a pair of atoms in the vibrational zero-point level of the molecule. In this way, it is possible to precisely compare the results from electron diffraction and spectroscopic experiments.

In practice, spectroscopic methods provide much more precise information on diatomic molecules (r_e) than electron diffraction measurements can at present supply; similarly, for triatomic and some small polyatomic molecules, electron diffraction studies serve mainly as a test of the technique and of the theory for interconversion of interatomic distance parameters.[14] In general, for moderately complicated molecules, it is often found that the molecular structure cannot be determined from electron diffraction data alone unless some assumptions are made in order to reduce the number of parameters refined; there are also difficulties in determining a physically defined structure, such as r_z, from spectroscopic data alone. When the data from the two sources is combined, it is possible to derive structural parameters with greater confidence than is possible with either technique separately, resulting in the r_{av} structure.[15]

There is now little doubt that for all but the simplest molecules the gas-phase molecular structure is most reliably determined by a simultaneous analysis of electron diffraction data and such rotational constants as are available. Even in cases where precise interconversion of such data is not performed, the spectroscopic data can be used to provide constraints among parameters in an electron diffraction refinement or as a check on the validity of an electron diffraction structure. Thus, a careful analysis can yield r_{av} parameters which may be converted approximately into r_e values by using the defining equations (3)-(6). It has been suggested that comparison of accurate molecular structures should be made in terms of the r_g bond length as common basis,[16] since r_g, the thermal average value of the instantaneous internuclear distance, represents the average value of the quantity of chemical interest. However, the use of the essentially equivalent r_α^o, r_z, r_{av} structures

may be preferable if interbond angles are considered, because the geometrical inconsistencies of the r_g structure makes angle definition difficult. The r_α^o type structure (the average of the projection of the instantaneous bond length on its equilibrium direction) is less satisfactory for bond lengths because of vibrational effects.[17] In principle, r_e parameters are the most desirable; for more complex molecules, approximate r_e structures are now being calculated from r_α^o and r_{av} structures using approximate anharmonicity constants,[18] and it may be that the corrections to be made to r_α^o or r_{av} to give r_e can be calculated with satisfactory accuracy on the basis of quite naïve models.

In summary, electron diffraction and spectroscopic techniques are capable of yielding results of high accuracy, but distance parameters obtained by different experimental techniques or different calculation procedures may differ by 0.01 Å or even more if the corrections, related to molecular vibrations, necessary to give the parameters the same physical meaning have not been made. If an electron diffraction investigation is combined with spectroscopic data, it is possible to obtain distance parameters for many compounds with standard deviations of the order of 0.001 Å or less; it is possible that improved techniques may reduce the errors still further and also make the method more useful for electron density and energy determinations (as in studies of dynamic aspects of molecule structure such as conformational equilibria, and thermodynamic properties) than it is at present.[19] In Table 2 there are shown experimental results on some types of molecule of interest in this work, illustrating the above considerations.

Gas-phase molecular structure determinations are, in principle, the most relevant to the calculations on isolated molecules considered in this

work. Intermolecular interactions may cause a difference between the structure of a free molecule and the molecule in a crystal. However, the results reported in electron diffraction and X-ray diffraction investigations may also deviate because of a difference in the definition of interatomic distances. In this case, no accurate formulae exist for correcting the distances obtained to the same distance type. X-ray crystallographic methods can tackle more general, complex molecular structures; electron diffraction usually gives more accurate determinations of interatomic distances.[20] The elegant combination of neutron and X-ray diffraction measurements is now beginning to provide experimental electron-density distributions properly corrected for the thermal motions of the atoms and sufficiently precise to be used as criteria for evaluating bonding theories.[21] Standard deviations in bond lengths of 0.01 Å or better are now reasonably commonplace in X-ray structure determinations. Crystal-structure analysis by neutron diffraction is a supplement, rather than an alternative, to X-ray diffraction, chemically speaking.[22] For some of the larger species considered in this work, experimental structure information is provided by crystal studies only. In crystallography it is thought that r_α bond lengths are derived, although there is confusion arising from systematic effects generally and the fact that the analysis of vibrational effects is less well understood because there are both inter- and intra-molecular components, leading to larger amplitudes; lower precision is thus achieved unless the study is carried out at low temperature. From the respective definitions, it is generally true that $r_\alpha < r_g$ so that crystallographically- derived lengths should usually appear shorter than electron diffraction ones. As well as considering the problem of non-equivalence of parameters, caution should be used in comparing results from studies of the two phases, gas and crystal,

as there may be conformational differences; the typical organic molecule, of relevance to this work, in the solid state exists as molecular units held together by weak van der Waals-type of intermolecular packing forces, i.e. non-ionic, resulting in the free molecule geometry of the vapour phase being largely retained. In some cases, e.g. biphenyl,[23] significant differences in molecular geometry between the phases arise if the molecular unit's conformation is determined by weak intramolecular forces; even so, it is likely that a change in dihedral angles only is involved, with bond lengths and angles hardly affected.

A further experimental technique, presently in the early stages of development, relevant to the determination of molecular structure is nematic phase nuclear magnetic resonance spectroscopy. The analysis of spectra of molecules partially oriented in a nematic or lyotropic liquid crystal solvent leads to coupling constants which are related to the inverse cube average function $r_d = \langle r^{-3} \rangle^{1/3}$.[24] The transformation of the effective structure parameters r_d leads to the vibrationally averaged molecular structure, r_α, enabling direct comparison of results with those from other techniques above. The electron diffraction r_α is identical with the NMR r_α if sample temperatures are equal; the limiting value r_α^o corresponds to the spectroscopic r_z. A complete structure determination by NMR alone is not possible; there are comparatively few molecules suitable for nematic phase NMR investigations, and a normal coordinate analysis of the vibrational spectrum of the species is required in order to apply the corrections for the harmonic vibrational motions to derive r_α. However, the method can be complementary to the gas-phase techniques above,[25] although it is uncertain that it is correct to say that interatomic distances in the species are practically invariant on changing from the vapour phase to the liquid crystal solvent.

(b) Calculational Considerations

Regarding molecular geometry as a property susceptible to calculation, it is possible, in principle, to compute the unambiguous equilibrium (r_e) structure of a molecule; comparison with experimental data, if available, may involve complications, as shown by the considerations in (a) above. For small polyatomic molecules, where refined wavefunctions are calculable, the elucidation of equilibrium structures and the investigation of multi-dimensional potential energy surfaces, which form the basis of models and theories of chemical reactions, are feasible projects.[26] A rigorous non-empirical treatment would compute the molecular wavefunction and properties as a function of nuclear configuration, without any empirical reasoning. However, in this type of work, bearing in mind the limitations of the single determinant approach and the lack of flexibility in basis sets used, the approach is less ambitious, basically because each nuclear configuration considered corresponds to a separate calculation so that general geometry optimisation, for example, is not computationally reasonable for polyatomics. In studying the electronic structure of a general molecular system, it is therefore usual to choose, empirically, a geometry, probably by using experimentally observed structural data, implying that the geometrical structure used is unlikely to be the actual equilibrium structure or that which would be calculated by the computational procedure in use, and adding a further approximation in practice to the computation. In general, a calculation on a particular molecular species is not critically dependent on having a precise geometrical structure in determining equilibrium properties. Some consideration of the effects of molecular geometry variation is made in this work; in particular, it is now the intention to discuss geometry optimisation as a computational procedure, in relation to small molecules,

where methods can be routinely applied, and also to larger molecules, where the aims in the present work will be mentioned.

The theory of finding minima of functions is just one aspect of the branch of mathematics called optimisation theory, which is rapidly developing.[27] Most modern minimisation techniques are designed to find local minima in the function by search techniques, characteristically assuming very little knowledge of the detailed analytic properties of the function to be minimised, other than the fact that a minimum exists. Just as in the non-linear variation problem encountered in proceeding beyond the Hartree-Fock limit (Section C of Chapter 2), there are inhibiting technical difficulties in attempting to solve non-linear programming problems, such as molecular geometry optimisation. It is possible to consider only unconstrained optimisation methods, as there are often convenient ways of introducing constraints into quantum chemical problems.[28] Of the methods which have so far found a use in quantum chemistry, the ones of particular relevance in this work are discussed below.

Of the category of non-derivative methods, where no explicit use is made of the derivatives of the objective function, one technique which has a long history in the context of nuclear position, and orbital exponent, variation is the Univariate Grid Search, or Cyclic Search, Method. This direct search method is based on the idea that the individual variables refer to coordinate axes $\underline{e}_1 = [1,0,0,\ldots,0]$ etc. in the n-space so that successive one-dimensional searches along each of the axes can be performed. The algorithm for the method is very simple; it is summarised below, indicating what computation is required.

(i) Select a point $\underline{a}$, set i = 1 .

(ii) Minimise $f(\underline{a} + \lambda \underline{e}_i)$ with respect to λ to obtain α_i, the value of λ that minimises the function along $\underline{e}_i$.

(iii) Replace $\underline{a}$ by $\underline{a} + \alpha_i \underline{e}_i$.

(iv) If $i \neq n$, set i = i + 1 and repeat from (ii); if i = n, set i = 1 and return to (ii) unless the α_i are less than some pre-set tolerance.

It is fairly obvious that if the variables of the problem are strongly dependent, then a minimum need not emerge from this process, and that it is only in the case of strictly independent variables that one can be sure of eventually reaching the required minimum. The crucial step in the procedure is undoubtedly minimising along the line $\underline{e}_i$; when the gradient of the function along the line is not readily available, one of the more efficient procedures (one-dimensional direct search) for locating the minimum is quadratic interpolation.[29] This conceptually simple approach is illustrated in this work in basis set scaling (Section C of Chapter 3), and in molecular geometry optimisation of acyclic species, which is not performed rigorously; the approach is to consider bond parameters as variables and empirical reasoning is used to reduce the dimensionality of the problem. Rigorous application of this type of alternating variable, or axial iteration, search method, or even of more sophisticated search techniques such as the Simplex Method,[30] is not feasible in practice as a general procedure in optimisation of ab initio wavefunctions because of the large number of function (i.e. total energy) evaluations required.

In addition in this work, some attempts at geometry optimisation are reported, adopting a more mathematical approach, in which the variables are the Cartesian coordinates of the atomic centres and a

derivative method of direct minimisation is used. The most well-developed of such methods are univariate in nature, i.e. they approach the minimum of the multivariate function along a sequence of lines (directions) in the many-dimensional space, and the problem is then the determination of an algorithm for the choice of these directions. These methods are thus based on a sequence of one-dimensional searches. Gradient methods, employing only first derivatives of the given function, are of relevance here. The particular method used here is part of a general class of variable-metric (or quasi-Newton) minimisation methods first suggested by Davidson[31] and subsequently discussed by Fletcher and Powell.[32] This class is generally regarded as the most powerful for minimising a smooth function for which the derivatives are readily available. Although minimisation of a n-dimensional function is liable to tax any optimisation method to the limit of its capabilities, as the energy functional, $E = \int \psi \hat{H} \psi d\tau$, is "pseudoquadratic", methods which are well-developed for the minimisation of quadratic functions are formally applicable in this case. Methods which guarantee to find the minimum of a quadratic in a specified number of iterations are said to be quadratically convergent; most objective functions can be well-approximated by quadratics in the neighbourhood of the minimum. The Fletcher-Powell method is basically a "conjugate direction" technique, attractive because of the characteristic finite termination properties and stability.[33] The algorithm is:

(i) Choose an initial point $\underline{a}$, and an initial matrix H_o which must be positive definite. Find $\underline{g}$ (gradient vector) and let $\underline{p} = -H_o \underline{g}$ (direction vector).

(ii) Minimise $f(\underline{a} + \lambda \underline{p})$ to yield a minimum at $\lambda = \alpha$; exit if this minimum is satisfactory, otherwise go to (iii).

(iii) Construct $\hat{\underline{a}} = \underline{a} + \alpha \underline{p}$, find $\hat{\underline{g}} = g(\hat{a})$.

(iv) Construct $\hat{H} = H + A + B$; $A = \alpha \underline{p}\underline{p}^{\dagger}/\underline{p}^{\dagger}\underline{y}$, $B = -H\underline{y}\underline{y}^{\dagger}H/\underline{y}^{\dagger}H\underline{y}$,

$\underline{y} = \hat{\underline{g}} - \underline{g}$;

find $\hat{\underline{p}} = -\hat{H}\hat{\underline{g}}$.

(v) Set $\underline{a} = \hat{\underline{a}}$, $\underline{p} = \hat{\underline{p}}$, $H = \hat{H}$ and return to (ii) .

This type of conjugate direction method uses function and gradient values, and makes estimates of the inverse of the matrix of second derivatives (H), choosing directions of descent according to $\underline{p}_i = -H_i \underline{g}_i$, where H_i is the estimate at the i th iteration; thus, p_i is determined by updating H and linear searches are used to produce conjugate directions.

In the implementation of optimisation schemes, there is always the problem of knowing when to stop a method; in practice, the user must determine convergence conditions which are sufficient for his needs. Traditionally, non-derivative methods of optimisation have been used in optimising non-linear parameters, such as nuclear parameters, in ab initio wavefunctions.[34] However, for geometry optimisation as a general procedure the more rigorous gradient type of method would be preferable; the analytic expressions for the required derivatives, subject to the constraints on the non-linear parameters and also to the condition that the constraints on the linear parameters continue to be bound during the variation of the non-linear parameters, have been known for some time,[35] but their use has been precluded by the computational expense of evaluation. Numerical approximation of gradients is sufficient for most derivative methods of optimisation, and this is the approach adopted in this work. It has been estimated that a good method of optimisation based on first derivatives should take of the order of $^1/n$ (n is the number of independent variables) of the time taken by a good non-derivative method,[36]

the former using more information about the objective function; however, the cost of this extra calculation of the gradients, using numerical approximation, in the procedure used in selected cases in this work renders the method impracticable as a general routine. Very recently, reappraisal of the situation regarding the evaluation of gradients analytically for ab initio wavefunctions has shown that a viable scheme for general geometry optimisation may be obtained by using a powerful, rapid variable-metric method of optimisation and calculating gradients analytically, as modern non-linear programming techniques (where much research is presently being done) can make the overall optimisation programme computationally feasible.[37]

(c) The Aims of Geometry Optimisation

Unlike other molecular properties of interest, the equilibrium configuration is exceptionally difficult to obtain, both by experimental and theoretical means. In the current work, it has not been feasible to effect geometry optimisation as a general procedure within the mode of calculation used; it may prove to be more reasonable in the future, as mentioned at the end of Section (b), but then it should be remembered that, within the usual restrictions of the single-determinant MO approach, it is useful to study trends within a series of molecules rather than isolated cases. Molecular geometry is obviously a special property because the calculated electronic wavefunction and derived properties depend upon the choice of nuclear configuration in the particular computation. The approach here has been to choose a physically reasonable geometry for the molecule under consideration, in general, which may be an experimentally derived one (generally, not r_e structure), thus hoping to closely approximate the true

equilibrium structure which could, in principle, be obtained by calculation. In selected cases, some form of geometry optimisation has been attempted. The aim then has been to deduce the gross features of the equilibrium geometry. In this way, the attractive, conceptually simple sequential univariate search, which is not a good method in the case of strongly dependent variables, can be adequate for locating a fairly approximate minimum in a function of dependent variables, although, in general, other methods are preferable for exact location of a minimum.

The calculated values of molecular properties vary with the geometry used. However, if the true equilibrium geometry is approximated sufficiently closely, meaningful results can be obtained. An extremely important part of the calculations reported here involves the molecular orbital energies, and their relationship to experimentally measured ionisation potentials (Section B of Chapter 2). The complementary nature of experiment and MO theory is well illustrated in this respect; MO considerations provide the best basis for interpreting experimental photoelectron spectra and the latter serve as an excellent test of MO methods. Photoelectron spectroscopy is the best available technique for directly measuring different molecular and atomic energy levels; approximate MO methods are well suited for understanding trends in measured data and in developping models without which single data are uninformative, so that bonding can be discussed in terms of chemical concepts.[38]

In this chapter, calculations on some small molecules are reported, with particular attention on molecular geometry and MO energies (photoelectron spectra), in which areas there is much information in the literature, both experimental and theoretical, for consideration with the results here.

(d) Examples of Molecular Geometry Optimisation

Small prototype molecules in this work, such as methane, acetylene, ethylene, butadiene, benzene, are considered at this point purely from the viewpoint of geometry optimisation, i.e. the minimisation of the total energy of the molecule with respect to positions of nuclear centres. For small molecules such as the above there is a wealth of data, both experimental and theoretical, for comparison with the results reported here. Thus, calculational procedures can be tested on small molecules, keeping in mind that application to large species is the real objective.

(I) Use of the Gradient of the Total Energy

The simplest molecule of interest as a reference in this work is Methane. With the assumption of a tetrahedral configuration of the nuclei, geometry optimisation becomes a problem in one dimension only; the total energy of the molecule is regarded as a function of the C-H bond length only. It is a straightforward task to fit a curve to a series of discrete points, corresponding to calculated total energies at various values of r(C-H), and hence to obtain the optimum r(C-H). Similarly, for optimisation of the structure of Acetylene, assuming linearity, and Benzene, assuming a regular structure, extension of the above procedure to the resulting two-dimensional problem is computationally feasible. Generalisation to higher dimensional problems, involving multi-dimensional surface fitting, is not really tractable. The above three molecules are considered in (ii) below. To illustrate the general situation here, the procedure described formally in (b) above is applied to ethylene and trans-1,3-butadiene, and then the results of some calculations on larger species are considered. Unlike the method of curve- or surface-fitting in which the total molecular energy is regarded as a function of geometrical parameters, this procedure involves the use of the gradient of the

total energy, along with a rapid, efficient method for using it to find the minimum energy configuration.

The problem of finding local minima on the many-dimensional potential energy surface of a molecule has been tackled by choosing a reasonable starting geometry, and then minimising the potential (total) energy function with respect to each internal degree of freedom, repeating the process until no further improvement is found. This procedure has generally proved to be satisfactory for molecules with a small number of degrees of freedom. Even if it were economical to study large molecules non-empirically by such a method, probably limiting the number of degrees of freedom by using symmetry constraints, there still remains a potentially serious problem; the magnitude of the potential energy function itself increases with increasing size of molecule, so that it becomes less sensitive to a given displacement of an atom. The minimum point is then more difficult to detect. The method of locating equilibrium geometries by focussing attention on the gradient of the potential energy function, rather than the function itself, largely eliminates the problems mentioned above. In addition, particularly in ring systems, there may be strong inter-dependence of nuclear coordinates, necessitating simultaneous minimisation with respect to several coordinates.

The first step of the algorithm of Section (b) above is the evaluation of the first derivatives of the total energy with respect to the nuclear coordinates of the initially chosen structure. Each derivative is the negative of the force on the particular atom in the particular direction. As mentioned in Chapter 2, one of the "one-electron" properties derivable from a Hartree-Fock SCF molecular wavefunction, in a straightforward manner, is the Hellmann-Feynman force on an atom in a molecule. Thus, from one

calculation, the set of forces on all the nuclei can be obtained. Unfortunately, these Hellmann-Feynman forces, computed using a limited basis set, are not physically reasonable for use in geometry optimisation.[39] Thus, in practice, much additional computation is required in obtaining the required gradients.

The total SCF energy of a molecular system, E^T, depends upon the nuclear coordinates, q, in two ways. Firstly, the expression for E^T contains integrals over the atomic basis functions which depend explicitly on the nuclear coordinates (basis functions move rigidly with the corresponding nuclei). Secondly, E^T depends on the wavefunction, or on the density matrix P, which also changes with the nuclear coordinates. This latter implicit dependence of E^T upon the q's need not be taken into account at the calculation of the derivatives dE/dq. This can be shown as follows: E^T as the explicit and implicit function of the q's can be written as

$$E^T = E^T(q, P(q));$$

$$\text{thus, } \frac{dE^T}{dq} = \frac{\partial E^T}{\partial q} + \frac{\partial E^T}{\partial P} \cdot \frac{\partial P}{\partial q} \quad .$$

The second term on the right-hand side must vanish because E^T has its minimum value at the SCF density matrix, leaving only the derivative of E^T with the explicit (non-linear) dependence on nuclear coordinates. In the ab initio procedure, the calculation of the forces by the differentiation of the basis function integrals is very time-consuming, lengthening a straightforward SCF calculation significantly; analytical formulae for the derivatives of Gaussian type basis functions with respect to atomic displacements have been derived[40], and, as stated in (b) above, the application of these may become more widespread.

The gradient g of the potential energy function E^T is a column vector, the components of which (in a cartesian coordinate system) are the partial derivatives of E^T with respect to each of the 3N cartesian displacement coordinates of the N atoms of the molecule. The magnitude of this vector is the magnitude of the internal force on the molecule, and its direction is that of steepest descent. Formally, g is defined by the relation:

$$\varepsilon\delta q^{\dagger} g(q^{o}) = E^{T}(q^{o} + \varepsilon\delta q) - E^{T}(q^{o}) \qquad -(7) .$$

In the limit of vanishing ε (a scalar), $g(q^o)$ becomes the gradient of E^T evaluated at the point q^o, which is a column vector whose components are the cartesian positions of the atoms in the molecule. The column vector $\delta q = q - q^o$ represents an arbitrary displacement. In this work, numerical differentiation (finite difference technique) is used to evaluate gradients of the total energy. This method, in general, involves two critical aspects; these are choice of step size, and numerical accuracy of the function involved (E^T). The formula for calculating a particular first derivative is:

$$\frac{\delta E^{T}}{\delta q_i} \approx [E^{T}(q + d_i \varepsilon_i) - E^{T}(q - d_i \varepsilon_i)]/2d_i \qquad -(8)$$

where ε_i is a unit vector, and d_i is the step size. d_i must be chosen sufficiently small to make equation (8) an adequate approximation (gradients are rigorously defined as limiting values), but large enough to avoid accumulation of round-off errors in the difference in calculated total energies. In the present work, step sizes of the order of 0.01 a.u. are used; the resulting estimated gradients are adequate for the present purposes.

The optimisation of the geomtry of ethylene is summarised by the results presented in Table 3. Within the constraint of a structure of D_{2h} symmetry, the geometry of ethylene is completely specified by

three parameters, namely the C=C and C-H bond lengths and the CCH (or HCH) angle. The initial structure assumed in this work was the equilibrium structure recently derived from experimental measurements,[41] shown in Table 3. It is possible to perform the optimisation in terms of internal coordinates (bond lengths and angles, dihedral angles) that are not necessarily uncoupled.[42] In this way, there is some similarity to the axial iteration technique, in that the gradient of the total energy with respect to a particular parameter is obtained by carrying out several calculations varying that parameter with others fixed. However, in the axial technique the new value of the parameter is derived by fitting some functional form, e.g. a quadratic, to the calculated points, assuming independence from other parameters. In the gradient technique, the calculated points are used to alter the parameter towards the optimum value, but the new value in the cycle is found simultaneously with all the others,in a cooperative manner, which is actually the essence of the procedure. Actually, in this work, the gradients of the total energy are found with respect to displacements of the atomic centres in the cartesian directions. This is equivalent to the above method; corresponding to the three unique bond parameters, three unique gradient components are required. With the notation of Figure 1, these are

$$\frac{\delta E^T}{\delta C(x)}, \quad \frac{\delta E^T}{\delta H(x)}, \quad \frac{\delta E^T}{\delta H(y)}.$$

All other gradient components are identically zero, by symmetry. When using these gradients, only one centre is shifted in each gradient calculation, so that computational savings can be made in molecular integral evaluation using the "merge" facility (Appendix 2). This may be regarded as an advantage of this type of gradient, in general, over that defined with respect to internal coordinates, where variation of a

bond parameter may lead to a shift of several centres, so that more "new" molecular integrals need evaluation. Thus, starting from the initial structure, the three gradient components are evaluated by shifting the appropriate centre in the appropriate direction by ± 0.01 a.u., and using equation (8).

Having evaluated the derivatives of the energy function, the variable-metric minimisation method used for the optimisation generates an improved vector q^{n+1}, giving the cartesian positions of the atom by the recursion relation

$$q^{n+1} = q^n - \alpha^n A^n g^n \qquad -(9) .$$

The method of determining α^n and A^n depends on the particular procedure used. In the original method proposed by Davidon, and discussed by Fletcher and Powell,[32] α^n is chosen as that value of λ which minimises the function $E^T(q^n - \lambda A^n g^n)$, so that the overall minimisation is carried out by a sequence of one-dimensional searches (optimum λ evaluation). This step involves examination of the potential energy function itself; computation of this quantity at several different conformations is required in order to obtain optimal λ, adding significantly to the computational expense of the overall minimisation. The first cycle is actually a simple steepest descent minimisation; A^n is set equal to the identity matrix. Thus, each centre is moved in a direction opposite to that of the resultant gradient vector at the centre, with the magnitude of the shift proportional to the magnitude of the resultant. Each cartesian coordinate is incremented by an amount equal to λ X negative of gradient with respect to the coordinate, yielding a new structure for each value of λ. A series of calculations is carried out by varying λ, with the optimum value of the latter found by fitting a quadratic, or a cubic, function to the set of points obtained. At the new structure, corresponding to atom shifts determined by optimum $\lambda(\alpha^n)$,

the gradient components are re-evaluated. The Hessian matrix A^n is updated using information from the previous cycle, via the algorithm

$$A^{n+1} = A^n + \alpha^n(-A^n g^n)(-A^n g^n)^T/(-A^n g^n)^T y - A^n y y^T A^n / y^T A^n y, \quad (y = g^{n+1} - g^n),$$

and used in equation (9) to modify the direction of the search vectors constructed from the gradients, in evaluating the new optimum $\lambda(\alpha^n)$. A set of increments of cartesian coordinates of the centres is generated, i.e. the product $A^n g^n$, and λ is optimised by another one-dimensional search. The cycle of gradient evaluation, coordinate increment determination, and optimum λ evaluation is repeated until convergence is obtained.

Formally, the above method is symmetry conservative,[43] which means that equilibrium geometries of a given symmetry can be found by simply specifying the initial geometry to be of that symmetry. The method essentially reduces the number of degrees of freedom to the number of symmetry adapted degrees of freedom, without requiring the often cumbersome transformation (and its inverse) to symmetry adapted coordinates. It can also be shown that since there are no external forces on the molecule, the gradient has no component corresponding to pure translations and rotations, so that there is no need to eliminate these six degrees of freedom. The 3N (N = number of atoms) cartesian coordinates can thus be used in a straightforward manner.

In Tables 3 and 4, the results of the geometry optimisation of ethylene, using the standard minimal scaled basis set and also using a double-zeta basis set derived from the best atom basis by splitting off the lowest exponent function of each AO of the minimal set, and trans-1,3-butadiene,using the standard basis, are presented. It is feasible to perform several iterative cycles in these cases; insight into the physical nature of the process can be gained with these simple structures.

If the energy function were exactly quadratic, convergence would be guaranteed within n cycles, where n is the number of variables being optimised. In each cycle, (2n+5) energy calculations are performed. The convergence of the procedure is dependent on factors such as the closeness of the initial structure to the optimum, the strength of the coupling of variables, and the degree to which the quadratic approximation is correct. In practice, two or three cycles only are computionally feasible for large molecules. The equilibrium structure should be a stationary point on the potential energy surface, i.e. the gradient vector g vanishes at this point. However, using numerical procedures, $g = 0$ identically is not a practical convergence condition. One possible condition is that the largest (in magnitude) component of the gradient should be less than a specified threshold. A threshold of 10^{-3} a.u. generally gives coordinates within $\pm$ 0.02 a.u. ($\pm$0.01 Å). There is a problem in simply using the gradient components because the λ value is combined with them in determining a "new" structure. The positive definite matrix A^n converges to the inverse of the force constants matrix; convergence could be facilitated if a good estimate of the latter were available. Calculation of the harmonic force constants, second derivatives of the energy with respect to cartesian displacements of nuclei, is prohibitively expensive, and they cannot really be estimated reliably, in general, although they can be in certain cases.[44] Effectively, the minimisation procedure determines an approximation to the force constant matrix from the consecutive forces; as well as deriving equilibrium structures, a theoretical vibrational force field can be determined using this type of approach.

Geometry Optimisation

In the following tables the symbols used are defined as:

g is component of gradient,

s is corresponding distortion amplitude (element of matrix product Ag);

α is the optimised value multiplying each distortion amplitude, yielding the cartesian coordinate increment to construct starting geometry for the next cycle;

S is the magnitude of the vector of distortion amplitudes, $= (\sum_i s_i^2)^{\frac{1}{2}}$ summing over all centres (i.e. including symmetry related centres to those in table);

σ is the product of α and S.

The unit of g is hartree $bohr^{-1} \times 10^{-2}$ (atomic units); the unit of s, S, σ is bohr $\times 10^{-2}$.

The above parameters are those of interest in the implementation of the optimisation procedure, and are collected in one of the tables of values for each molecule. The other table for each case shows the variation of total energy of the molecule and geometrical parameters with each cycle of the process. The total energy ($kJ\ mol^{-1}$) is given relative to a value of zero for the initial structure. The bond lengths are quoted in the chemically familiar $\mathring{A}$ units, and bond angles in degrees.

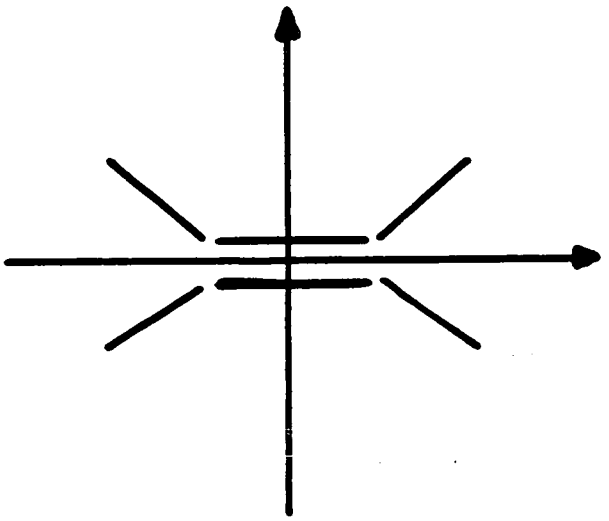
Further examples of the application of the systematic procedure of geometry optimisation to small molecules are provided by some five-membered ring heterocycles, with results given in Table 5[45]. The standard scaled minimal basis set was used in these studies; notation is given in Figure 1. The various species are treated to different extents, but they all illustrate the method. In the 1,2,3-Thiadiazole case, a computational saving was made in cycles 2 and 3. A significant fraction of the total number of basis functions used in calculating the molecular wavefunction is centred on the S nucleus, so that use of the "merge" facility in molecular integral evaluation is rendered ineffective for the cases where S is shifted; thus, the calculation of gradients at the S nucleus is particularly expensive. However, as the gradient is simply the negative of the force on the centre, the S gradients may be calculated if all the others are known, as a result of:

$$\sum_n F_x^n = \sum_n F_y^n = 0 \quad \text{(sum over all centres)}.$$

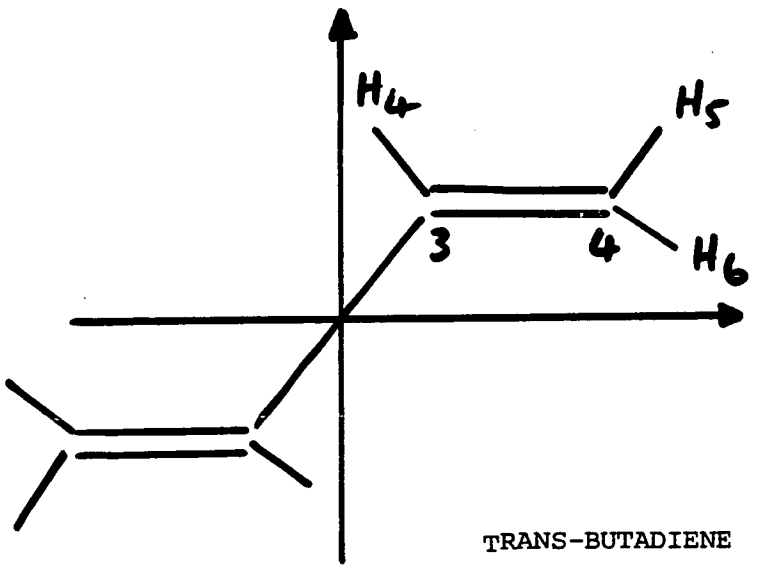
There is no net component of internal force in any direction acting on the molecule (F_x^n is X-component of force on centre n). In cycle 1, the S gradients were evaluated by separate calculations; the "sum of forces" result holds satisfactorily, only small residues occurring.

The calculated equilibrium structures of the molecules of Tables 3-5 can be compared with experimental geometric data. These test cases have not been treated typically in that the initial structures (cycle 0) are derived from experimental data, except for the unknown 1,2,3-oxadiazole; the structure of 1,3,4-oxadiazole was constructed by including the microwave geometry of 1,2,5-oxadiazole. For large species of interest in this type of work, if experimental geometries are available,

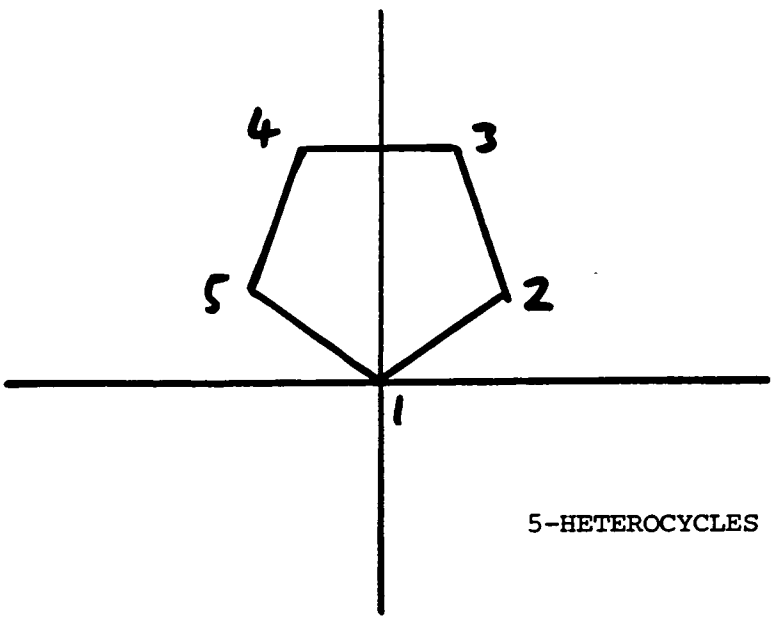
Figure 1 COORDINATE SYSTEMS



ETHYLENE



TRANS-BUTADIENE



5-HETEROCYCLES

there is usually no need for geometry optimisation. Thus, the starting-point for ethylene is a particularly good one, being a recently derived equilibrium structure from experimental measurements. A previous estimate of this structure is: r(C=C) = 1.330 Å, r(C-H) = 1.076 Å, $\hat{HCH}$ = 116.6° [46]. Thus, both the basis sets used yield calculated equilibrium structures in close agreement with experimental data. Of relevance also are other reported calculated structures; ethylene has been widely studied in various basis sets - rather wide range of structures have been obtained, e.g. r(C=C) from 1.30 to 1.36 Å. Allowing for variation in experimental results also, the equilibrium geometry found with the scaled minimal basis is probably in slightly better agreement; the double-zeta basis seems to underestimate the bond lengths, leading to a more compact molecule. The more flexible basis set gives a molecular total energy about 0.1 a.u. lower than that found using the scaled minimal basis. Scaling corresponds physically to a contraction of the constituent atomic orbitals in a molecular environment; with a basis of limited flexibility, some "aid" is given to the eigenvectors in determining the electron distribution. The scaled minimal basis leads to a molecular total energy about 0.15 a.u. lower than that from the best-atom basis. The more flexible double-zeta basis set, constructed from the same primitive functions, is unscaled; scaling actually leads to a marginal raising of the total energy for ethylene. The constituent AO's are contracted physically, the basis functions of larger exponent having extra weight. Overall, the double-zeta basis set leads to a more accurate electronic wavefunction and corresponding distribution; however, it seems to exaggerate the electron density between bonded centres (particularly, C=C) so that the greater screening of the nuclei leads to shorter bond lengths. This is consistent with the particular nature of the basis set, with the separating-off

of the higher exponent functions, corresponding to more diffuse electron distribution, and these make smaller contributions to the occupied MO's than in the minimal basis case. In general, with limited bases, there is no real correlation between actual calculated values of equilibrium geometrical parameters and quality of basis. In this case, it is fortuitous that the rather inflexible minimal basis yields such a good geometry. However, it has been found generally that limited bases do give calculated structures which are in reasonable agreement with experiment. In the application of a typical basis set to molecular calculations, there are two important factors which may lead to errors in computed geometries, namely the incompleteness of basis and the neglect of electron correlation inherent to the single-determinant MO approach. It appears that, from calculations on small molecules at (or near) the Hartree-Fock limit, where the latter factor only is operational, there is a tendency for single-determinant MO theory to underestimate bond lengths (Chapter 2). Physically, this can be explained by an over-estimation of the electron density between bonded centres, as a result of the neglect of correlation. Mathematically, an exact description of the molecular electronic wavefunction can be obtained by the admission of configuration interaction; the total wavefunction then contains contributions from electronic configurations in which "antibonding" MO's are occupied, such configurations corresponding to structures with increased bond lengths. Effectively, some electron density is removed from interbond regions, and the overall, averaged effect is to give a less compact molecular geometry geometry, in agreement with experimental data. The use of a limitee basis, as in this work, does not correspond to a well-defined situation, unlike the H-F limit case. Thus, one type of basis which has been used for systematic geometry optimisation of small molecules is the STO-3G one of Pople and co-workers[47]. This minimal basis is slightly less refined than

the standard one of this work; on comparing the results with experimental data, calculated geometries are in quite good agreement, but there is no definite systematic trend.[48] For the particular case of ethylene, the calculated equilibrium r(C=C) is 1.305 Å, and r(C-H) is 1.079 Å ($\hat{HCH}$ = 115.4°). Another study on small molecules has explored the use of an even less refined basis, STO-2G, in molecular geometry optimisation using the gradients technique. This rather crude basis yields calculated equilibrium geometries generally in reasonable agreement with experiment and STO-3G results, with a tendency to predict longer bonds than does the STO-3G basis. For ethylene, the STO-2G basis leads to r(C=C) = 1.318 Å, r(C-H) = 1.089 Å ($\hat{HCH}$ = 115.3°). The standard minimal basis here yields a longer r(C=C) and r(C-H). Allowing for variations in experimentally derived estimates, this r(C=C) is very close to the experimental equilibrium value; r(C-H) is probably overestimated, even allowing for the greater uncertainty in experimental measurements of r(C-H) and subsequent extrapolation to equilibrium (includes amplitudes of vibration which are larger for C-H). The $\hat{HCH}$ angle value, somewhat reduced below the "classical" 120°, is well reproduced by the calculations. The results for ethylene illustrate the variability of minimal basis results.

The more refined calculations using the double-zeta basis are nearer the H-F limit; the computed total energy is actually about 0.08 a.u. above the H-F limit. However, the calculated electron distribution is much closer to the H-F one than the minimal basis one, so that the more compact equilibrium geometry resulting is probably in close agreement with the H-F geometry, although less so with the experimental one. The 4-31G basis of Pople[47] is of similar quality to the double-zeta one here; the former has been used to study geometry optimisation of small molecules. For ethylene, the 4-31G equilibrium structure is

effectively identical to the one reported here, perhaps indicating that better quality basis sets give results close to the well-defined limiting situation. Of calculations beyond the H-F limit on ethylene, partial optimisation (with respect to r(C=C)) has been reported; as indicated above, the calculated equilibrium r(C=C) is lengthened, with r(C=C) = 1.36 Å with a double configuration SCF wavefunction, and r(C=C) = 1.34 Å using an extensive CI approach.[49]

In considering larger species such as butadiene and the 5-membered ring heterocycles, there are no equilibrium structures estimated from observed data, so that comparison between calculated and experimental is not direct. For butadiene, the initial structure (cycle 0) is a r_g structure (defined in Section(a) above). From experimental measurements, a r_{av} structure has also been derived; the geometrical parameters are very similar in the two cases.[50] The final structure reported for trans-1,3-butadiene in Table 4 is very similar to the starting one, except for r(C-C). The very close agreement with experiment is again rather fortuitous, as the simple direct comparison of structures excludes consideration of basis set effects and also definition of experimental geometrical parameters. An estimate of the equilibrium structure cannot really be made from experimental data. The $r_{av}(r_\alpha^\circ)$ structure of the vinyl group is effectively identical to that of ethylene. However, extrapolation to the equilibrium structure should include the effects of vibrational motion of the whole molecule, so that equilibrium parameters for the vinyl fragments cannot be derived separately. However, by using non-equilibrium structures as above, errors of about 0.01 Å in r(C=C) and r(C-H) are probably introduced. The geometrical parameter of most interest is r(C-C). The experimental value is 1.46-1.47 Å; the calculated value here is 1.51-1.52 Å.

The STO-2G and STO-3G similarly overestimate r(C-C), giving 1.498 Å, within a partial optimisation scheme, and 1.488 Å respectively; the r(C=C) length is calculated to be 1.327 Å and 1.313 Å respectively. There are no reported geometry optimisation studies using more extended basis sets. The minimal basis calculations yield a structure which is more "classical" than the real molecule appears to be, as judged by the length of the central, formally single C-C. The hypothetical conjugated C-C has been estimated to be 1.50-1.52 Å in length; the single bond between sp^3-hybridised C atoms in ethane is about 1.54 Å. In planar trans-1,3-butadiene, which is the predominant species, the C-C bond is thus somewhat shortened, indicating the effects of conjugation. It is of interest to note that the above-mentioned CI calculations on ethylene yield optimum r(C-C) = 1.47 Å for the "perpendicular" configuration obtained by twisting the CH_2 groups by 90°, and so breaking down the π-bonding; the resulting formal single C-C bond is not actually the same as the type in butadiene. For trans-1,3-butadiene, the standard scaled minimal basis set describes the double bonds well, as in ethylene. However, the C-C single bond is not described well, being lengthened by about 0.05 Å in the calculated equilibrium structure. By considering the C-H bonds as well, these probably also being overestimated compared to the expected equilibrium values, it seems that the adequate description of C=C is gained at the expense of the single bond descriptions. The calculated electronic distribution is probably exaggerated between doubly-bonded C atoms, and underestimated in the single bond regions, where the nuclei are then screened less efficiently, so that longer bond lengths are predicted. The calculated values of bond angles are in agreement with the experimental ones- the most significant is $\hat{CCC}$ which is increased a little beyond 120° - and similarly with the actual

equilibrium ones in all probability. The effect of lengthening single bonds appears to be a characteristic of the standard basis set; this may be a result of the scaling procedure used to partially optimise the best-atom minimal basis. The use of ethylene for deriving a scaled basis for conjugated systems may have led to a bias towards double bonds at the expense of formal single bond description. Although the calculated equilibrium structure for trans-1,3-butadiene corresponds to a "classical" structure based on experimental values of parameters, there is still evidence of conjugation according to the calculations, as shown below, so that planar butadiene is not described by the calculations as being completely non-conjugated, or classical. It is noteworthy that the basic difference between the initial and final structures is simply a lengthening of the central C-C bond in the latter by about 0.05 Å, and that the total energy improvement is only approximately 8 kJ mol^{-1}.

The experimental structures for the various five-membered ring heterocycles are derived from microwave spectroscopy, and are expressed in terms of r_s parameters (Section (a)). These are likely to exhibit some coupling so that even rough estimates of equilibrium values are rather difficult to make. However, allowing for uncertainties in the parameters concerned in the comparisons, the results of Table 5 show that the standard minimal basis quite adequately describes the molecular geometries, except for the rather severe overestimation of the lengths of formal conjugated single bonds between heavy centres. Such bond lengths in conjugated systems in general show much greater variation in length than double or triple bonds, indicating perhaps a shallower potential for single bond stretch-compression, so that the equilibrium length may be significantly different from an experimentally measured one, including vibrational effects. Nevertheless, the results of geometry optimisation of the

selected species above do indicate a systematic lengthening of single bonds, particularly those between heavy centres, when using the standard basis. For the sulphur heterocycles, where comparison of the experimental and calculated data is more meaningful, the total energy improvements are rather small (20-30 kJ mol^{-1}), indicating that the use of experimental structures can be a good approximation for medium-sized molecules, at least. In addition, the final structures are effectively obtained by lengthening the single bonds of the initial ones by significant amounts (greater than 0.05 Å), and yet the total energy changes are not exceptionally large for these substantial changes in geometry.

Having considered the calculated equilibrium geometries in the light of reported data to gain some insight into the use of the optimisation procedure, and standard basis, and their capabilities for extension to larger molecular species, it is instructive to examine the minimisation in more detail. The results shown in Tables 3-5 illustrate some points considered in Section(b) above. The rapidly convergent descent method for minimisation (Fletcher-Powell-Davidson) guarantees to find the minimum of a general quadratic function, calculable at all points and with analytically defined gradients , in n iterations (n is number of independent variables). In contrast, the total energy functional is not exactly quadratic in molecular geometry parameters, and gradients with respect to the latter are estimated in practice by a numerical difference method. Ethylene, with n = 3, shows that a precisely defined minimum in E^T is not obtained in 3 cycles, or even in 5 in the minimal basis case, although changes in E^T and geometrical parameters have become chemically insignificant by 3 cycles. In practice, the problem of deciding on the point to terminate the procedure arises, as, for typical species considered, n exceeds a

computationally reasonable number of cycles. One criterion of convergence commonly used in similar situations is that all components of the gradient vector should be less than some threshold, typically 10^{-3} a.u., corresponding directly to the definition of the equilibrium structure where all force components are zero. However, it is more appropriate in the procedure used here to use the distortion amplitudes derived from the gradient components. Thus, a criterion could be that the predicted absolute distance from the minimum, $S = \sum_i s_i^2$, be less than a prescribed amount, or that every component s_i be less than a given threshold. The latter corresponds to termination when the change in bond lengths between consecutive cycles is less than some prescribed value, typically 0.001 Å for small cases, but perhaps 0.01 Å for larger ones. In order to compare different examples, or to define a general level of convergence, the number of variables (n) can be introduced by defining the reduced gradient $\bar{g} = (\sum_i^n \frac{g_i^2}{n})^{\frac{1}{2}}$, or reduced distortion $\bar{s} = \frac{S}{n^{\frac{1}{2}}}$. Unfortunately, inherent to the procedure, is the empirically determined parameter α, a measure of the distance proceeded in the direction of search of the particular cycle. Thus, the parameter $\sigma (= \alpha S)$, or its components, should be used as the measure of convergence; σ_i is the actual increment of the ith variable at each cycle, giving effectively the change in bond length. Usually, α is of the order of unity, so that S and σ are closely related. Thus, the results of Tables 3-5 show the stability of the procedure, with E^T decreasing at each cycle, and S likewise. However, the influence of the α factor can lead to the E^T decrease not being smooth, and σ not decreasing continuously. To allow for a possibly large α factor, the final cycle should strictly be completed, even although the gradients appear to be small enough. For practical purposes, the test results indicate that when a reduced σ factor of 10^{-2} or less is reached, the bond lengths are accurate to $\pm$ 0.002 Å and

bond angles to $\pm$ 0.2°. In these medium and small molecules, at such a position, E^T is constant, within chemical accuracy. The qualitative behaviour outlined for these test molecules can be borne in mind when considering application of the procedure to larger species, where starting geometries may be rather arbitrarily chosen. The oxygen heterocycles of Table 5 are nearer this situation; the indications are that the method converges reasonably quickly even with poorer starting structures. As the process continues, the geometrical parameters tend to fluctuate about their equilibrium position; in addition, using the standard basis, there is the extra driving force tending to lengthen the single bonds, and this is a very significant effect in the cases here. In each cycle, the α optimisation averages out the effects of the individual atom shifts; at the resulting structure, new shifts are determined using information from each centre's current environment and also the previous one.

It appears that, when considering larger species, application of the above optimisation procedure may lead within two or three cycles to a molecular structure in which the geometrical parameters and total energy are correct within chemical accuracy, although not to full mathematical accuracy obtainable. Thus, gross features of molecular geometry may be highlighted at reasonable computational expense.

(II) Use of Non-Derivative Method

In this section, a few examples of attempts at geometry optimisation are presented, using a quite common, simple technique which is less precise, mathematically, than that of (I) above, and involves more chemical intuition. The particular molecules considered are methane, acetylene, butadiene (planar trans-form as in (I), planar cis-form and non-planar "perpendicular" form with a dihedral angle of 90° between π-systems), diazabutadiene, and benzene

There is a large amount of experimental information on the hydrocarbons (Table 2); there are even estimated equilibrium geometries for methane and acetylene. A recent spectroscopic investigation has yielded geometrical data for diazabutadiene, as shown in the Table giving the results of the calculations. In addition, there are pertinent results on these species from other theoretical investigations. Thus, as far as the optimisations go on these molecules, further characteristics of the standard basis can be deduced.

In general terms, as in section (b) above, the objective function to be minimised is the total energy of the molecule, and no explicit use is made of the derivatives of this function. A direct search is employed; the individual variables (geometrical parameters, e.g. r(C≡C) and r(C-H) in acetylene) refer to coordinate axes along which successive one-dimensional searches are performed. The algorithm for the method is given in section (b). More particularly, as in acetylene, for example, an initial structure is chosen (initial set of variables), and a calculation performed; r(C≡C) is then varied and the value that minimises the total energy is found (parabolic interpolation). Strictly speaking, the minimisation with respect to r(C-H) should be centred on the "new", optimised structure. However, it is possible to start from other than the minimum point obtained in the first series of calculations. This is really based on the physical nature of the system, in that it is known that r(C≡C) is the dominant parameter, and that the starting structure is likely to be a reasonable approximation to the minimum one. Similarly, chemical intuition can be used to

reduce the dimensionality of the problem, by concentrating on selected variables. The examples considered here have been examined only briefly, and a fairly approximate minimum found (although in some of them, at least, this simple method could locate the optimum structure accurately). The successful use of the simple method of optimisation, which is less rigorous mathematically than many algorithms which have now been designed to tackle the multivariate minimisation problem, illustrates that the latter ignore some basic issues in theoretical chemistry, in particular the specific practical problems encountered during molecular geometry optimisation. The first relevant issue is the scope of the geometry search. Rather than considering global searches of unch aracterised potential energy surfaces with unrestricted parameter variation, a notable increase in optimisation efficiency is obtained if energy surfaces are assumed quadratic in the geometrical parameters. Thus, as found in this work, the set of (computed total energy versus geometrical parameter) points can be approximated quite closely by a quadratic function (energy versus scaling parameter, as in Chapter 3, is similar). In practice, experimental data or computational experience on similar molecules permits estimation of bond lengths within 0.05 $\mathring{A}$ and of bond angles with 5° in an initial structure; within this margin around the optimum computed geometry, the energy surface is expected to be smooth and quadratic. Thus, the optimisation method is usually used for refinement of a structure for which a qualitative estimate is already available. A second issue relevant to the design of an optimisation algorithm is the status of present computational methods in theoretical chemistry; the energy corresponding to a particular coordinate arrangement is difficult to calculate non-empirically, so that a premium value is placed on each data point. Thus, in the univariate search, often only three points are considered along a particular axis so that an exact quadratic fit is involved; it is desirable to have a set of points which encloses the minimum point. If a greater number of points are computed, "least-squares" fitting can be used. In Table 6, the results obtained

here are presented; only one cycle of the procedure has been performed in each case. There is always a lowering of the total energy as each parameter is varied; the situation is more meaningful, physically, than the variation of the total energy with the parameter α involved in the gradients procedure in (I) above. However, smooth convergence is not necessarily found in either procedure. For the purposes of this work, partial optimisation using the simple univariate search method can lead to geometries which are close enough to the equilibrium one to ensure chemical accuracy, without involving great computational expense; the total energy is then only a few kJ mol^{-1} above that of the actual computed minimum. In addition, a univariate search can be performed to obtain an improved initial guess for the optimum gometry before performing a more refined technique, as in (I), in order to derive a precise equilibrium structure if this were required, or to tackle difficult cases (strong coupling among variables, or several important geometrical parameters involved).

Of the species considered here (results in Table 6), methane is the simplest possible case, assuming regular tetrahedral configuration, as there is only one geometrical parameter to be varied, so that univariate search is rigorous. The results given have been obtained using the standard minimal basis (ethylene scaled). There is a scaled minimal basis for methane itself (Appendix 4), for use with sp^3-hybridised carbons in larger molecules. The results using this basis are effectively identical with those obtained using the ethylene-scaled one (the total energy of methane using the methane-scaled basis is only about 3 kJ mol^{-1} lower than that with the standard basis). The calculated equilibrium bond length is close to that determined by extrapolation from experimental data (r(C-H) shows the largest relative variation of all bond lengths among the various definitions of "r"). The geometry used in the scaling procedure was rather different

from the calculated optimum one. This is of interest, in general, as there may be expected to be some coupling between actual geometrical parameters and the scale-factors, so that the latter may be particularly geometry-dependent, and conversely, the scale-factors may force the calculated equilibrium structure to be that at which scaling was performed. It so happens that the calculated equilibrium structure of ethylene in (I) above is close to that used during scaling. However, it is unlikely that the coupling effect is particularly significant here.

The optimisation of acetylene, assuming linearity, and benzene, assuming D_{6h} regularity, are exactly two-dimensional problems. In each case, a grid of points is considered, and straightforward curve-fitting in one-dimension can be performed for each variable to yield the values at the optimum structure. In addition, a surface-fitting exercise has been performed in these cases; a two-dimensional quadratic functional form is assumed ($E = ax^2 + bx + c$ extended to $E = ax^2 + bx + cy^2 + dy + exy + c$), and a least-squares fit performed[51]. This procedure is rather costly as a great number of points are required (full grid). The results for the equilibrium parameters are almost identical with those obtained assuming no coupling (e = o in equation above, effectively).

The butadiene structures have been optimised to a very limited extent. There are quite a large number of geometrical variables in each case. In a univariate search, only a selection of the points of the full grid is considered. In the trans-butadiene case, by simply considering two important parameters (r(C=C) and r(C-C)), an improved geometry is obtained with a total energy close to that found in (I) with considerably more computational expense. In addition, the gross features of the equilibrium geometry are obtained (particularly the elongated C-C bond).

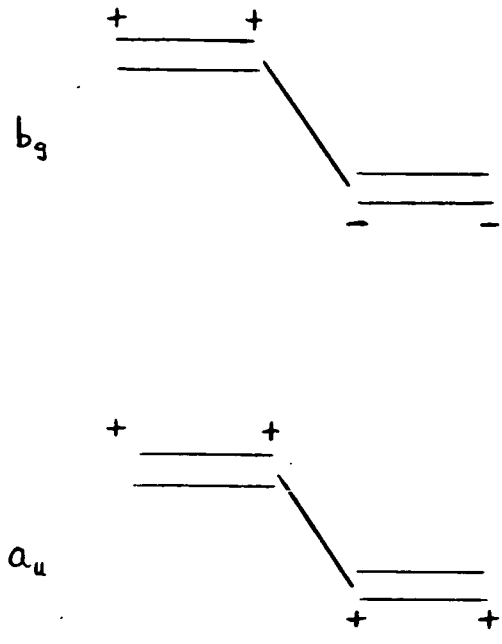
The rotational conformers of butadiene have been studied quite extensively theoretically, and experimentally, although geometry optimisations involved have not been complete (usually rigid rotation is assumed)[52]. The results obtained here are in general agreement with those obtained previously. The prominent feature is that the standard minimal basis here does not yield as great a shortening of the central C-C bond of the planar conformers from the value of the "classical" perpendicular form as is found elsewhere.

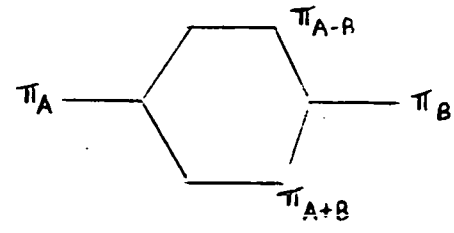
Overall, the performance of the standard minimal basis is seen by its yielding optimised molecular geometries in substantial agreement with experimental data and other theoretically derived data, with the greatest deviation being in its over-estimation of the length of formally conjugated single bonds. Trans-butadiene exemplifies this, and the same effect is found with planar trans-diazabutadiene where r(N-N) is somewhat elongated. With benzene also, the computed equilibrium r(CC) length is lengthened, although to a lesser extent, corresponding to the "intermediate" nature of the C-C bond.

B. Molecular Orbital Energies .

The selection of small molecules considered in Section A, and a few others, are of interest in their own right at this point. The five-membered ring heterocycles, treated from a geometry optimisation point of view above, have been studied closely, and are still being studied, by non-empirical calculations, employing the standard basis, with a view to correlation with experimental data.[53] The other small molecules relevant to this work are well-known species, both theoretically and experimentally. Many calculations, at different levels of refinement, have been performed on these species, and the nature of the chemical bonding has been described widely in the literature,[54] qualitatively and quantitatively. The results reported in this section are intended to show how the standard minimal basis set performs on the various species. Data for comparison, both from experiment and from other calculations, are provided. Some MO's of particular significance are illustrated schematically in Figure 2. The forms of these are of interest when considering larger molecular species. The computed molecular property of most significance here is the MO energy, with the correlating experimental property being the ionisation potential which is now routinely measured by the technique of photoelectron spectroscopy.

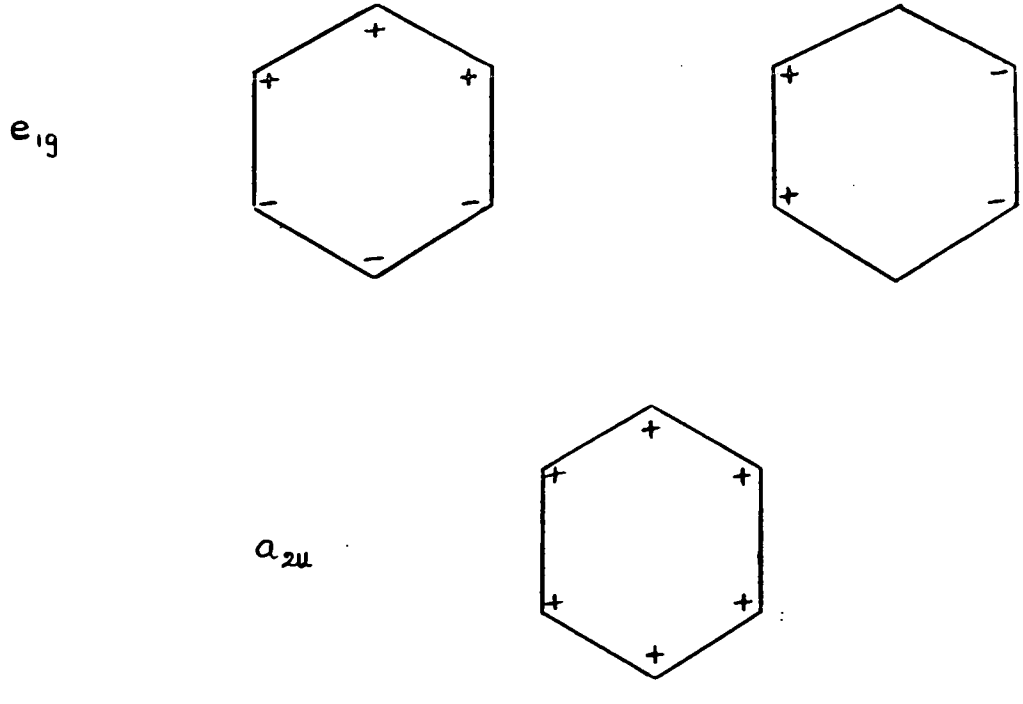
Following the pioneering work of Price, Potts and Streets, it has become a matter of considerable interest to obtain as many complete sets of valence ionisation energies of small hydrocarbons as possible.[55] These are the ionisation energies which correspond, within the usual Koopmans' Approximation (Chapter 2), to the ejection of a photoelectron from one of the occupied valence shell MO's. The contribution of the carbon 1s orbitals is neglected, because of their low basis energy (-290 to -295 eV), and for most practical applications, higher AO's than those occupied in the ground-state atoms and/or polarisation functions





ORBITAL INTERACTION.

(a)



(b)

FIGURE 2 : FORMS OF π-MO'S OF BUTADIENE AND BENZENE.

need not be taken into consideration. Such an extended set of data for closely related molecules is of prime importance for developing and testing theoretical models of varying degree of sophistication, currently used for the rationalisation and prediction of ionisation energies. Simple hydrocarbons (which most of the species involved here are) have the advantages of having all valence ionisation energies within an easily accessible region (ca. 8 to ca. 26 eV), and of having molecular σ-orbitals usually spread more or less uniformly over the whole of the molecule, with rare, symmetry-conditioned exceptions. The positive charge of the radical cation obtained by removing an electron from such delocalised orbitals is evenly distributed over the molecular frame, so that the effect of electron rearrangement is minimised. In addition, π-bands in the PE spectra are usually easily identified. An extensive compilation of experimental valence ionisation energies of 143 hydrocarbons has been published,[55] obtained by using He(Iα) (21.22 eV, λ = 58.4 nm) and/or He(IIα) (40.80 eV, λ = 30.4 nm) radiation for excitation. For smaller molecules, it is usually possible to deconvolute the overlapping band systems into well-defined, unambiguous individual bands. Data from other sources is included here.[56] The ionisation energies of the inner valence shell of some of the hydrocarbons (so-called C(2s)-shell) have also been measured, and extend the I.P. range of comparison. A conservative estimate of the uncertainty with which the positions of broad bands can be determined in the absence of strong overlap with other close-lying bands is $\pm$0.2 eV; occasionally, larger discrepancies are found in band positions from different sources. Whenever the ionisation energies are separated by less than about 0.5 eV and/or if the bands overlap strongly, an unambiguous assignment of the order of the radical cation states is not possible. The ionisation energies listed are the positions I_j^m of the band maxima; within experimental uncertainty, they are usually assumed equal to the vertical ionisation energies I_j^v for all practical purposes. The error limits are at least $\pm$0.1 eV in most cases.

From the first studies with the standard minimal basis set, it was found that scaling tends to lower valence shell binding energies and leads to better numerical agreement with photoelectron spectra when Koopmans' Approximation is used.[57] Scaling does not have any effect on orbital ordering; in fact, it is true to say that virtually all orbital orderings become stable once one passes the threshold of reasonable size in minimal basis sets. In general, from non-empirical calculations using the standard basis, a linear correlation (least-squares fit) between observed vertical I.P.'s and computed orbital energies for the valence shell can usually be made with reasonable precision, irrespective of whether this relationship is theoretically valid, i.e. I.P. (obs) = m I.P. (calc) + c. The fact that $m \neq 1$ (gradient) and $c \neq 0$ (intercept) is a representation of the errors due to incomplete basis, correlation energy, and the use of Koopmans' Approximation. The slope m usually has a value of about 0.7-0.9; calculated I.P.'s tend to be larger in magnitude than the corresponding observed ones, and this is characteristic of the standard basis, and is expected for a reasonably flexible and balanced basis used within Koopmans' Approximation. There is some similarity to the situation with molecular geometries (Section A) in comparing calculated and observed I.P.'s as the quantities involved are not defined in precisely the same way, although generally they are similar. The data presented in Table 7 for some small molecules of interest here show that there is a close correlation between calculated and observed MO energies (or I.P.'s), as has been found with a large number of smallish species, such as the azoles and azines, which are rather more complex.[53] With molecules where more than one geometry has been considered, the variation of computed orbital energies with geometry is of particular interest. Where differences are relatively small, such as comparing an experimentally deduced structure with one nearer the calculated equilibrium position,

variations in individual orbital energies tend to be small, although to varying extents; the net uncertainties in values probably are comparable with those from experiment. Any minor alteration in orbital ordering occurs only when the relevant orbitals are nearly degenerate, so that the individual I.P.'s are unlikely to be resolved experimentally. Thus, the choice of molecular geometry is unlikely to be critical when comparing with experimental data. In contrast, if geometry changes are the result of considering different molecular conformations, there are usually several MO's whose energies vary significantly, and this can be used for identifying purposes. The general effect of more refined basis sets, such as the ones currently being developed[58] (extensive double-zeta) and applied to trans-butadiene to yield a wavefunction very close to the Hartree-Fock limiting form, is to stretch the overall MO energy spectrum calculated at both ends, by a reasonably small amount. Cross-overs in orbital ordering tend to occur only when there is near degeneracy. In Table 7 there is shown a comparison of I.P.'s calculated by various methods for benzene - Koopmans' Approximation with the STO-3G minimal basis of Pople and co-workers, the standard basis, a DZ basis using the same primitive functions, and a very flexible basis (effectively H-F limit); a procedure which corrects the values from Koopmans' Approximation by going beyond the one-particle approximation; Δ SCF method, by including calculations on cationic species.[59] It can be seen that there are no changes of orbital order or groupings in any calculation.

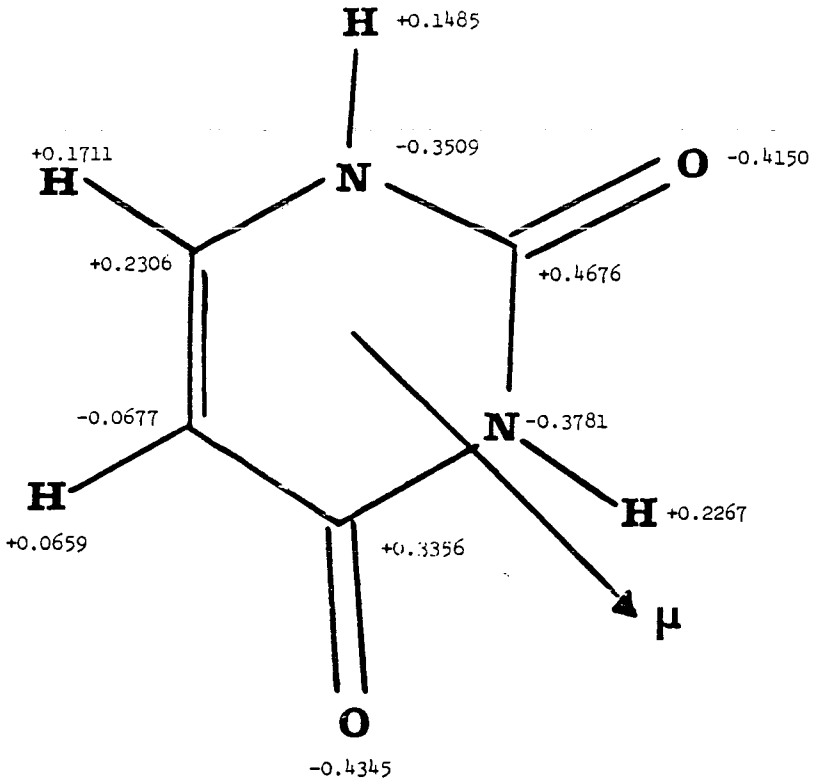
Overall, calculations on a large number and variety of small molecular species, for which PE spectroscopic data are available, have shown that the standard minimal basis performs well from the viewpoints of total energy and orbital energy, the latter also involving the application of Koopmans' Approximation. Thus, practical experience rather than rigorous theoretical background has shown the adequacy of the standard basis in these respects.

It has been found that the standard basis gives good agreement between calculated and experimental dipole moment absolute values, and also leads to quite good agreement with second moments except the quadrupole moment.[60] The signs of dipole moments are firmly established. It is noted that, for very polar molecules in particular, the dipole moment and atom populations tend to be sensitive to small changes in minimal basis sets (e.g. scaling parameters). Overall, the standard basis performs quite well; physically, electronic charge tends to be more concentrated in the vicinity of the atoms than is found with very flexible bases. Calculations on uracil, a rather complicated polar small molecule, using the standard basis and the refined double-zeta basis mentioned above, are included in Fig. 3 in order to give an extreme type of example, where discrepancies in calculations with various bases are likely to be greater. Uracil itself and related species are presently being studied.[58]

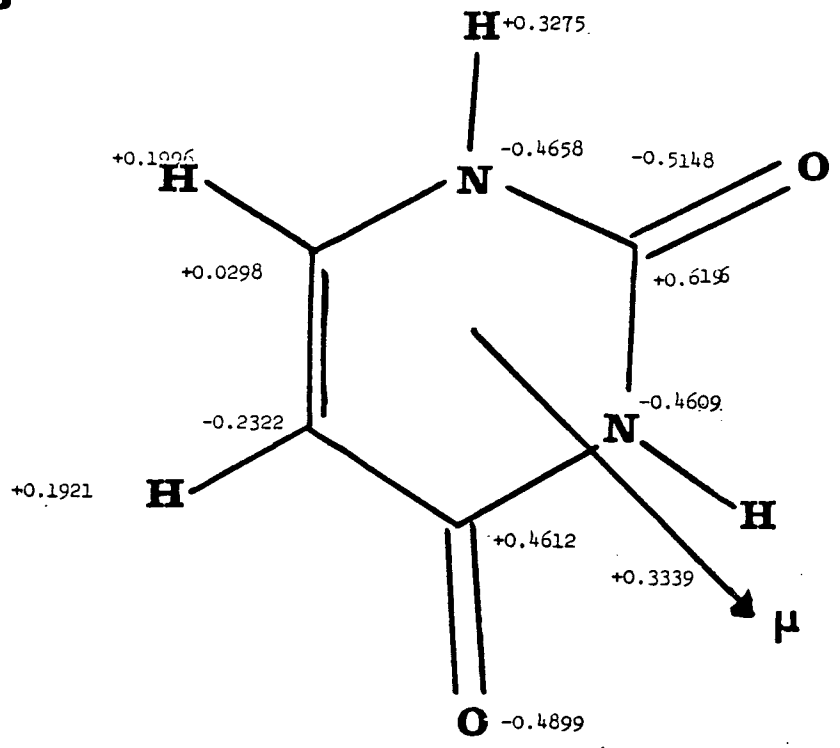
Figure 3 : CALCULATED AO POPULATIONS FOR URACIL.

C H A R G E S

D Z



M B



C. The Use of Small Molecules as Reference Structures.

If small molecular species such as those considered in the previous section are of interest in their own right, then it is reasonably easy to perform more refined calculations on them than those reported here. However, when considering larger species, where routine application of improved calculations is not feasible, the results from the small molecule computations can be invaluable in analysis of molecular electronic wave-functions of large molecules. Thus, using a common basis set, the forms of derived molecular orbitals of large cases can be viewed qualitatively in terms of those of smaller constituent units; the adequacy of the description provided in the latter cases can be determined by comparison with more accurate calculations. In this work, with interest in larger systems, an attempt to use the results from small species more quantitatively is made. The nature of "aromatic" systems is discussed in the following chapter; at this point, the use of small molecules as references is considered.

(a) Resonance Energy

The abnormal heats of hydrogenation and of combustion of benzenoid hydrocarbons have long been recognised as one of the features of aromatic compounds.[61] The difference between the value of each heat actually observed and that calculated on the basis of the classical system of non-interacting single and double bonds is the resonance or mesomeric energy. Although there is a considerable amount of thermochemical information on these resonance energies for hydrocarbons, there is much less data for heterocyclic systems.[62] Various studies using empirical or semi-empirical molecular orbital theory have given values for both carbocyclic and heterocyclic molecules, but more than one set is

available through the impossiblity of obtaining a unique set of parameters for the calculations.[63] This can be overcome by the use of non-empirical calculations. The correlation error, the difference between the experimental total energy and that from Hartree-Fock theory (Chapter 2), which is about 0.5-1.0% of the total for atoms from the first two rows of the periodic table, must be considered. Since in the determination of resonance energies, comparison is made with a "classical" structure of very similar geometry to the real one studied, the correlation error is largely eliminated.[64] The novelty of the approach used in this work is to carry out all of the following steps in reaching an estimate of the resonance energy of a molecule:-

(i) evaluate the energies of the individual bonds in "non-aromatic" systems;

(ii) then compute the energy of the classical analogue of the aromatic system;

(iii) compute the total energy of the molecule and take the difference between this and (ii).

Of course, it is essential that the basis set used in the determination of the bond and molecular energies is the same one.

In the evaluation of the bond energies, the energies of particular bonds are regarded as transferable between molecules, as in thermochemical bond dissociation energies. Thus, the total methane energy ($E^T(CH_4)$) is 4E(CH) (i.e. 4 times the C-H bond energy); similarly, the ethylene energy, $E^T(C_2H_4)$ is given by E(C=C) + 4E(C-H). To evaluate the single bond energy between sp^2 carbon atoms, the total energy of a twisted (gauche) buta-1,3-diene is used, where the π-systems are perpendicular to one another. Thus, $E^T(C_4H_6)$ = 6E(CH) + 2E(C=C) + E(C-C).

Similar procedures to the above establish bond energies for other CC bond types, such as E(C=C=C) from allene, and E(C≡C) from acetylene.

The energies of OH, NH,PH,SH,BH bonds can be established from the total energies of H_2O, NH_3, PH_3, H_2S, BH_3. In second-row element cases, two series of calculations can be performed, with and without the 3d atomic orbitals, so that the effect of these can be explicitly considered. The values for the C-O, C-N, C-P, C-S, C-B bond energies between sp^2 atoms (CX) are established using twisted vinyl derivatives CH_2=CH-XH, in which the XH bond lies perpendicular to the vinyl group; this again removes the π-electron conjugation between the centres, and, as in the buta-1,3-diene case, avoids the necessity to assume the resonance energy of the CH_2=CH-XH system (X=O, BH,PH,S,C_2H_3) is zero. In all of these cases the nuclear-nuclear repulsion is lower in the out-of-plane conformation, yet it is well established that conjugated olefins, ethers etc. prefer a planar configuration. It seems that the favourable electronic factor in the planar arrangement can be equated with a small amount of resonance energy, i.e. even in these "non-aromatic" species. The total energies of the model molecules using the standard basis set are shown in Table 8. It is worth noting at this point that the resonance energies of acyclic conjugated olefins are small, as expected, but not negligible, and this indicates that Dewar's recent adoption of a revised definition of resonance energy as the molecular energy in relation to the closest related polyolefin is probably not advisable. Experimentally, it has been found that the central double bond in 1,3,5-hexatriene is significantly different (longer) in length compared to the end double bonds, although the latter are similar to those of 1,3-butadiene; this also suggests that some interaction is occurring between the conjugated groups, and this is confirmed from the photoelectron spectra of conjugated olefins (Section B above).

The values of the bond energies derived from the model molecule total energies are presented in Table 9. These are applied in the following chapters. Some further points related to the use of small molecules as references in deriving the above simple scheme are considered below.

(b) Bond Energies

Although advances in computational ability have enabled the application of quantum mechanics to problems of molecular structure and properties to become reasonably straightforward, there has not been a corresponding advance in the methods of analysis and interpretation of molecular electronic wavefunctions, so that valuable information may be discarded unused. In general, an analysis is needed which will enable the good and bad features of wavefunctions to be characterised numerically, and will indicate which properties are most sensitive to the various approximations involved in calculations and which are insensitive. The interpretation of calculated results requires the introduction of concepts which can be defined precisely, and can also be used meaningfully to describe molecular properties or behaviour in a variety of contexts. One basic concept of chemistry in general is that of bond energy, much used in theoretical discussion; precise definition, on the basis of ab initio calculations, of this quantity seems reasonable.

The publication[65] of wavefunctions and properties for a large number of small molecules using a consistent set of atomic orbitals has opened up the possibility of comparative analyses of the results. One of the most obvious methods is the linear regression analysis of the total energy.[66] The simplest model of the total energy of a molecule is a sum of atomic energies:

$$E^T = hH + cC + nN + oO + fF + \ldots \qquad (10),$$

where h,c,n,... are the numbers of these atoms in the molecule and H,C,N,... the atomic energies. The atomic energies are determined by fitting the formula (10) to the calculated total energies for a large set of molecules. The results show that the energies can be reproduced to within about 0.025 atomic units; in terms of the total energies involved,

this accuracy of fit is high, but 0.025 a.u. is a very significant amount of energy, chemically. In the expression (10), the effects of bonding are being distributed between the atoms. In terms of classical thermochemistry, this formula is equivalent to the assumption that the energy of a heteropolar bond is the arithmetic mean of the energies of the related homopolar bonds. A more accurate model of the molecular total energy allows for the extra stability of the heteropolar bond by writing:

$$E^T = hH + ... + \alpha CH + ... + \beta HF + ... \quad -(11),$$

where α is the number of CH bonds and CH their additional energy or bond energy. This gives an improved fit to about 0.006 a.u. This dissection of E^T is uniquely defined. One result of formula (11) is that it gives the energy difference between the two sides of a chemical equation, provided all the species are neutral closed-shell molecules, in terms of the bond energies alone. Equation (11) refers to a "classical" molecule, with no extra stabilisation or destabilisation attributable to the existence of aromatic character or strain, for example. Species exhibiting such effects would be described by equation (11) with the addition of a further energy term on the right-hand side to account for these; such stabilisation, or destabilisation, energies could be evaluated for real molecules by referring to the appropriate, hypothetical species using fitted values for the parameters of equation (11).

As a particular application of the concept of bond energy, the physiochemical properties of conjugated organic molecules can be rationalised, i.e. by introducing the concept of resonance energy; there is general agreement as to the definition of this quantity as the actual bonding energy of a molecule less that "expected" for its most stable valence-bond structure composed of single, double and triple bonds, but no general agreement exists regarding how the bond energies for the reference

structure are to be evaluated. In this context, use of equation (11) is similar to the procedure of partitioning the binding energy of the reference molecules into bond energy contributions (binding energy = total energy - sum of free atom energies); the atomic energies of equation (11) are not exactly free atom energies. Thus, the resonance energy of a molecule could be estimated by calculating the difference in binding energy between real and "classical" species. However, it has been found more fruitful in this work to express the total energy as a sum of bond energy contributions, effectively incorporating atomic energy terms of equation (11) into appropriate bond energy terms. The resonance energy of a molecule is then the difference between the total energy and the sum of bond energy contributions, derived from reference molecules. The single, double and triple bond energies used are those appropriate to non-aromatic systems.

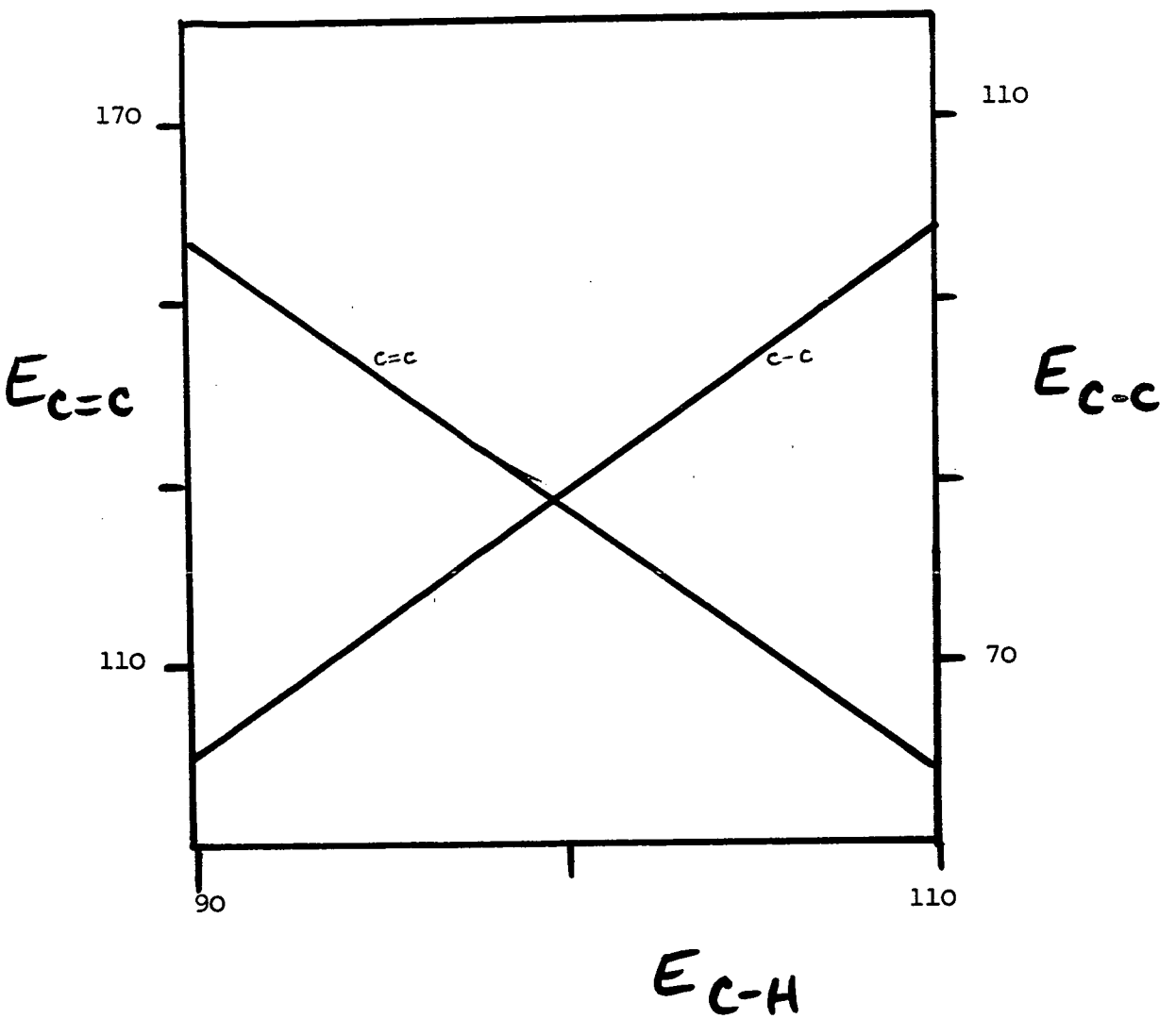
In practice, individual values for bond energies are derived as outlined in the previous part (a). Considering conjugated hydrocarbons composed of single and double bonds, bond energies E(C=C), E(C-C), E(C-H) are required. The natural route for evaluation of these involves the use of the acyclic polyenes, ethylene, butadiene, hexatriene, etc. Thus,

$$\begin{aligned} E^T(C_2H_4) &= E(C{=}C) + 4E(C{-}H) \\ E^T(C_4H_6) &= 2E(C{=}C) + E(C{-}C) + 6E(C{-}H) \\ E^T(C_6H_8) &= 3E(C{=}C) + 2E(C{-}C) + 8E(C{-}H) \end{aligned} \qquad -(12),$$

Unfortunately, any three equations referring to these polyenes are linearly dependent so that a unique solution for the three bond energy terms cannot be found in this way. From two equations, such as the first two of (12), individual values of E(C=C) and E(C-C) can be derived as functions of E(C-H); the relationships are illustrated in Figure 4 .

Figure 4: INTERRELATION OF BOND ENERGIES

($EXPTL, Kcal\,mol^{-1}$).



It is interesting to note that any value of E(C-H), and the corresponding values for the carbon-carbon energies, is consistent with the ground-state thermochemistry of all closed-shell, conjugated hydrocarbons. If unique bond energy terms are to be derived, it is necessary to resort to using some additional data. The approach in this work is to include methane as a reference molecule, using $E^T(CH_4) = 4E(C\text{-}H)$ to derive a unique value for E(C-H), which is assumed to be appropriate for use in the equations (12); thus, E(C=C) and E(C-C) are easily obtained. Generalising to other types of conjugated species, other bond energy values are required; unique values for these cannot be obtained by using small acyclic species as references, including the appropriate types of bonds. However, the inclusion of species such as ammonia, water, H_2S enables the energies of bonds to hydrogen to be evaluated and a unique set of bond energies to be finally obtained.

This "non-unique" situation can be expressed alternatively by noting that the resonance energy for a compound, defined in the above-mentioned way, is independent of the numerical values used for bond energies. Thus, for a completely unsaturated hydrocarbon C_nH_m (all m hydrogen atoms are bonded to the system of n unsaturated carbon atoms), it can be shown that: number of C-H bonds present = m

number of C=C bonds present = $^n/_2$

number of C-C bonds present = $n - {}^m/_2$.

For the non-aromatic, classical molecule, the total energy is

$$E^T = mE(C\text{-}H) + {}^n/_2\, E(C{=}C) + (n - {}^m/_2)E(C\text{-}C) \qquad -(13),$$

the summation of effective bond energies. Using the relevant reference molecules,

$$E(C{=}C) = E^T(C_2H_4) - 4E(C\text{-}H)$$

and

$$E(C\text{-}C) = E^T(C_4H_6) - 2E(C{=}C) - 6E(C\text{-}H)$$

$$= E^T(C_4H_6) - 2E(C_2H_4) + 2E(C\text{-}H),$$

so that the reference structure E^T can be expressed as

$$E^T = (m - {}^{3n}/_2)\, E^T(C_2H_4) + (n - {}^m/_2)\, E^T(C_4H_6) \qquad -(14).$$

Consequently, the total energy of the classical molecule is independent of the value of E(C-H), and depends only upon the total energy of the two reference molecules, ethylene and butadiene. Thus, any combination of bond energies which is consistent with the thermochemistry of the reference compounds must yield the same total energy for any molecule of interest. The expression (14) can be generalised to apply to any conjugated compound, by including appropriate reference structures. In estimating resonance energies the total energies of the reference molecules are the basic quantities; however, it is instructive to construct a table of unique values of bond energies as outlined above, and refer to this when calculating the resonance energy of a molecule. Although hypothetical classical structures of molecules appear in the model for deriving resonance energies, ultimately <u>real</u> molecules are used as the basis of evaluation.

(c) Energies of Reaction

The model described above for evaluating resonance energies of cyclic conjugated hydrocarbons leads to the practical equation (14); the total energy of the classical, non-aromatic structure is expressed as a combination of total energies of reference molecules (most simply, ethylene and 1,3-butadiene). An alternative, equivalent definition of resonance energy is the energy difference for the following reaction, involving real species only:

$$(n - {}^{m}/2)\ C_4H_6 \rightarrow C_nH_m + ({}^{3n}/2 - m)\ C_2H_4 \qquad -(15).$$

If classical C_nH_m were considered on the right-hand side, ΔE^T for the reaction would be zero, by definition. This approach to the problem of assessing stabilisation energies in cyclic conjugated hydrocarbons has recently been applied, using previously derived data from various sources.[67]

The calculated change in total energy of the appropriate reaction is directly comparable with the enthalpy change evaluated using experimental ΔH_f° data, thus establishing a consistent conceptual framework in which the experimental and the theoretical definitions of stabilisation energies involve the same real molecules as reference structures. The enthalpy changes for especially designed reactions, termed "homodesmotic", provide excellent parameters for characterising the stabilisation present in benzene and other cyclic conjugated hydrocarbons, as well as destabilisation (ring strain), which is typical of many saturated and non-conjugated ring systems.[68] This area, describing reactions between closed-shell species only, is one well-handled by Hartree-Fock theory (as in Chapter 2, Section C), and even, possibly, by approximations to it.

A homodesmotic reaction is one in which (1) there are equal numbers of carbon atoms in their various states of hybridisation in reactants and products, and (2) there is a matching of carbon-hydrogen bonds in terms of the number of hydrogen atoms joined to individual carbon atoms in reactants and products. Thus, the minimisation of extraneous energy contributions arising from changes in the types of C-C and C-H bonds singles out the energetic consequences of those structural features responsible for the stabilisation (or destabilisation). Thus, equation (15) is an example of such a reaction. Empirical resonance energies, evaluated traditionally from measurements on heats of combustion and hydrogenation, or ethane reduction,[69] actually include contributions from changes in hybridisation as well as changes in σ-bond compression energies in reactants and products, factors usually considered to be completely separate from electron delocalisation. Thus, although such resonance energies correspond to ΔH° values for chemical reactions involving real species only, the use of them has led to a confused situation. The use of

homodesmotic reactions yields a molecular index for gaugeing the enhanced stability of conjugated cyclic structures; this index is defined as a reaction parameter which facilitates direct comparison between theory and experiment, with the elimination,or minimisation, of extraneous energy contributions. Homodesmotic reactions constitute a special, rather restrictive subclass of isodesmic reactions,[70] extensively studied in non-empirical calculations of bond separation processes. Polycyclic hydrocarbons lead to a problem in that it is not possible to construct an appropriate homodesmotic reaction using only ethylene and butadiene; in fact, general equation (15) refers to an isodesmic reaction. A molecule such as divinyl-ethylene must be introduced as one of the references for polycyclic hydrocarbons. However, although the use of isodesmic reactions involves the conversion of different types of C-H bonds and the associated energy contribution, stability considerations based on reaction (15) are still a significant improvement over the conventional reactions which involve further changes in C-C bonds. Actually, strictly speaking, even the energy change in a homodesmotic reaction is not due solely to electron delocalisation (it is impossible to write down a chemical reaction involving cyclic and acyclic structures in which the energy difference between reactants and products is due only to electron delocalisation, as even the formal, "classical" bond types cannot be matched exactly), nor is it solely a property of the molecule being investigated. It is also worth remembering that, as far as the theoretical calculations are involved, it is not a trivial task to derive the optimum representation of the reference molecules. Obviously, in the last analysis, evaluations of stabilisation energies via the above model depend on the choice of comparison molecules.

The energy change in a homodesmotic reaction can serve as a quantitative measure of stability in conjugated organic systems, and can provide an excellent practical measure of those structural features that lead to molecular stabilisation. In contrast, traditionally used reactions are all isodesmic; in Table 10 are shown some relevant experimental data for benzene. In all reactions involved therein, the bonding is matched only in terms of formal type (single, double, triple), irrespective of its chemical nature. One consequence of this is a considerable variation in the magnitude of the stabilisation energy from the different sources. Furthermore, within the context of bond energy assignments, the very procedure of identifying the stabilisation energy as ΔH° for an isodesmic reaction carries with it the implicit assumption that the residue of mismatched bond energy terms can be equated to zero. Before putting the above considerations into practice, there is one final, very significant feature to be included. The determination of stabilisation energies relates the enhanced stability to the other reactant and product species that complete the reaction, i.e. the bonding in these species is implicitly adopted as a reference state. Considering benzene, the stabilisation energy is evaluated relative to the bonding energies in ethylene and 1,3-butadiene; the latter molecule, in reality, is not "classical", and is certainly stabilised to some extent by π-electron delocalisation. Thus, ΔH° for the homodesmotic reaction of Table 10 entry (5)) is relative to the stabilisation energy of three molecules of trans-1,3-butadiene (predominant form in reality). A "pure"single bond between sp^2-hybridised carbon atoms is required in the reference molecule. There is no unequivocal method for evaluating the stabilisation energy in trans-1,3-butadiene; some chemical structure, containing a formal

$C(sp^2)$ - $C(sp^2)$ single bond, would be required for reference, but the 1,3-dienes are themselves the simplest hydrocarbons possessing this structural element. The only type of reference structure that has the appropriate number of carbon atoms in the required hybridisation states, the appropriate number of each type of C-C bond, and the required number and type of each C-H bond is a homodesmotic isomeric form of butadiene; the rotational conformer (rotamer) of special interest in this context is the unique conformation of the molecule in which the $H_2C_1 = C_2H$ plane is twisted 90° relative to the $HC_3 = C_4H_2$ plane. This 90° - 1,3-butadiene, whose properties can be deduced from spectroscopic and other physical data,[71] is a real molecule, albeit an unstable one. This species is considered in Section B above; the energy difference between 90°- and trans-1,3-butadiene, the measure of the stabilisation of the latter planar structure, has been found experimentally to be 30 kJ mol^{-1}. The use of the 90°-conformer has a profound effect on the magnitude of the stabilisation energies compared to those evaluated using the stable trans isomer; this is exemplified for benzene in Table 10. This situation shows that it is possible to have a considerable variation in the magnitude of stabilisation energy determined from different homodesmotic reactions. However, this feature results from inherent stabilisation (or destabilisation) in the different reactant and product species used as reference structures, rather than from the mismatch of bond types.

In this work, cyclic conjugated species considered are not restricted to the simplest type; formal triple bonds and heteroatoms are introduced here. It is rather difficult in practice to preserve the rigorous homodesmotic conceptual framework in the general case; an accurate treatment of all the reference molecules required would be a formidable computational task. Thus, the approach used here, outlined in

part (a) above, involves the use of some simplifying assumptions within the basic model described in this section . It is reasonable to expect that the loss of rigour does not prevent meaningful results from being obtained. In particular, the benefits of using 90°-1,3-butadiene are still obvious; a unique reference structure of "localised" double and single bonds is obtained. One important result is it seems that the customary downplaying of the magnitude of the stabilisation energy in acyclic polyenes in general is questionable. Results for hexatriene are shown in Table 11 . Further applications of the approach for deriving resonance, or stabilisation, energies are reported in the following chapters. In Table 11 there are presented results obtained previously for aromatic-type molecules, using the Table 9 bond energies, which are on a very different energy scale from those of the thermochemist.[72]

References for Chapter 4

1. R. McWeeny and B.T. Sutcliffe, "Methods of Molecular Quantum Mechanics", Academic Press, London and New York, 1969.
2. A.G. Robiette, in Specialist Periodical Report of the Chemical Society on "Molecular Structure by Diffraction Methods", Vol. 1, 1973.
3. O. Bastiansen and M. Traetteberg, Acta Crystallographica, 13, 1108 (1960).
4. P.R. Bunker, Journal of Molecular Spectroscopy, 42, 478 (1972).
5. M. Iwasaki and K. Hedberg, Journal of Chemical Physics, 36, 2961 (1962).
6. D.R. Lide, Annual Reviews of Physical Chemistry, 15, 225 (1964).
7. Y. Morimo and E. Hirota, ibid., 20, 139 (1969).
8. C.C. Costain, Journal of Chemical Physics, 29, 864 (1958).
9. T. Oka, Journal of the Physical Society of Japan, 15, 2274 (1960).
10. K. Kuchitsu and S.J. Cyvin, in "Molecular Vibrations and Structure Studies", (editor S.J. Cyvin), Elsevier, Amsterdam, 1972.
11. Note the different averages involved in (3) and (4).
12. Y. Morimo and E. Hirota, Journal of Chemical Physics, 23, 737 (1955).
13. It is assumed that the atom pair acts as a zero-point Morse oscillator.
14. L.S. Bartell and K. Kuchitsu, Journal of the Physical Society of Japan, 17, Suppl. B II, 20 (1962).

15. K. Kuchitsu, T. Fukuyana and Y. Morimo, Journal of Molecular Structure, 1, 463 (1968).

16. K. Kuchitsu, "Gas Electron Diffraction", in MTP International Review of Science, Physical Chemistry Series 1, Vol. 2, MTP Co. Ltd., Oxford, 1972.

17. Vibrations causing atomic displacements perpendicular to the equilibrium bond direction may cause r_α° to differ substantially from r_g.

18. It is too early to judge the validity of the approximate r_e structures.

19. R.K. Bohn, in ref. 2 series, Vol. 5, 1977.

20. A.F. Cameron, in ref. 2 volume (I, 1973).

21. G.E. Bacon, "Neutron Diffraction", 3rd edition, O.U.P., 1975.

22. J.C. Speakman, in ref. 2 volume (I, 1973).

23. Biphenyl treated to a great extent in chemistry textbooks.

24. A. Saupe and G. Englert, Physical Review Letters, 11, 462 (1963).

25. K.C. Cole and D.F.R. Gibson, Journal of Molecular Structure, 28, 385 (1975).

26. H.F. Schaeffer, "The Electronic Structure of Atoms and Molecules", Addison-Wesley, Reading (Mass.), 1972.

27. R. Fletcher (editor), "Optimisation", Academic Press, London and New York, 1969.

28. Books (e.g. ref. 27) discuss the general theory of constrained optimisation.

29. L. Cooper and D. Steinberg, "Introduction to Methods of Optimisation", W.B. Saunders Co., Philadelphia, 1970.

30. J. Kowalik and M.R. Osborne, "Methods for Unconstrained Optimisation Problems", Elsevier, New York, 1968.

31. W.C. Davidson, A.E.C. Res. Develop. Rep., ANL 5990 (1959).

32. R. Fletcher and M.J.D. Powell, Computer Journal, 6, 163 (1963).

33. R. Fletcher, Computer Journal, 13, 317 (1970).

34. W.A. Lathan, L.A. Curtiss, W.J. Hehre, J.B. Lisle, J.A. Pople, Progress in Physical Organic Chemistry, 10, 175 (1973).

35. R. Moccia, Theoretica Chimica Acta, 8, 8(1967).

36. J.W. McIver and A. Komornicki, Chemical Physics Letters, 10, 303 (1971).

37. A. Komornicki, K. Ishida, K. Morokuma, R. Ditchfield, M. Conrad Chemical Physics Letters, 45, 595 (1977).

38. H. Bock and P.D. Mollere, Journal of Chemical Education, 51 506 (1974).

39. P. Pulay, Molecular Physics, 17, 197 (1969).

40. J. Garratt and I.M. Mills, Journal of Chemical Physics, 49, 1719 (1968).

41. J.L. Duncan, Molecular Physics, 28, 1177 (1974).

42. J.B. Collins, P. von R. Schleyer, J.S. Binkley, J.A. Pople Journal of Chemical Physics, 64, 5142 (1976).

43. The gradient (in cartesian coordinates) transforms as the totally symmetric representation of the point group of the initial structure.

44. P. Pulay, Molecular Physics, 18, 473 (1970).

45. G. Black, G. Howat, M.H. Palmer, and I. Simpson, unpublished results.

46. K. Kuchitsu, Journal of Chemical Physics, 44, 906 (1966).

47. M.D. Newton, W.A. Lathan, W.J. Hehre, J.A. Pople, Journal of Chemical Physics, 52, 4064 (1970); R. Ditchfield, W.J. Hehre, J.A. Pople, ibid, 54, 724 (1971).

48. W.A. Lathan, W.J. Hehre and J.A. Pople, Journal of the American Chemical Society, 93, 808 (1971).

49. M.H. Wood, Chemical Physics Letters, 24, 239 (1974).

50. W. Haugen and M. Traetteberg, Acta Chemica Scandinavica, 20, 1726 (1966) for r_g; ref. 15 for r_{av}.

51. C.A.C.M. Algorithm 176 (1963) used.

52. B. Dumbacher, Theoretica Chimica Acta, 23, 346 (1972); P.N. Skancke and J.E. Boggs, Journal of Molecular Structure, 16, 179 (1973).

53. R.H. Findlay, Ph.D. Thesis, Edinburgh, 1973.

54. The review articles given as References for Chapter 1 provide many sources of information.

55. G. Bieri, F. Burger, E. Heilbronner, J.P. Maier, Helvetica Chimica Acta, 60, 2213 (1977).

56. D.G. Streets and A.W. Potts, Journal of the Chemical Society, Faraday Transactions I, 1505 (1974); E. Lindholm, C. Fridh and L. Asbrink, Discussions of the Faraday Society, 54 (1972); R.M. White, T.A. Carlson and D.P. Spears, Journal of Electron Spectroscopy, 3, 59 (1974).

57. M.H. Palmer, A.J. Gaskell and M.S. Barber, Theoretica Chimica Acta, 26, 357 (1972).

58. I. Simpson, R.H. Findlay and M.H. Palmer, unpublished results.

59. W. Moyes and M.H. Palmer, unpublished results.

60. M.H. Palmer, R.H. Findlay and A.J. Gaskell, Journal of the Chemical Society, Perkin Transactions II, 420 (1974).

61. J.D. Cox and G. Pilcher, "Thermochemistry of Organic and Organometallic Compounds", Academic Press, New York, 1970.

62. M.H. Palmer, "The Structure and Reactions of Heterocyclic Compounds", Edward Arnold, London, 1967.

63. M.J.S. Dewar, "The Molecular Orbital Theory of Organic Chemistry", McGraw-Hill, New York, 1969.

64. R.K. Nesbet, Advances in Chemical Physics, 9, 321 (1965).

65. L.C. Snyder and H. Basch, "Molecular Wave Functions and Properties", Wiley, New York, 1972.

66. G.G. Hall, International Journal of Quantum Chemistry, Symposium No. 9, 279 (1975).

67. P. George, M. Trachtman, C.W. Bock and A.M. Brett, Theoretica Chimica Acta, 38, 121 (1975).

68. P. George, M. Trachtman, C.W. Bock, A.M. Brett, Tetrahedron, 32, 317 (1976).

69. N.C. Baird, Journal of Chemical Education, 48, 509 (1971).

70. P. George et al, Tetrahedron, 32, 1357 (1976). (as 68)

71. L.A. Carreira, Journal of Chemical Physics, 62, 3851 (1975).

72. M.H. Palmer, private communication.

Table 1

Bond-length Parameters Employed in Gas-Phase Molecular Structure Determination

r_e	distance between equilibrium positions of atoms
r_g	average value of interatomic distance (for a particular temperature)
r_a	distance most directly related to electron diffraction data
r_o	parameter which reproduces ground-state rotational constants (effective distance)
r_s	parameter calculated from ground-state rotational constants using Kraitchman's equations (isotopic)
r_z	distance between mean positions of atoms in ground vibrational state, calculated from spectroscopic data
r_α	distance between mean positions of atoms at a particular temperature, calculated from r_g
r_α^o	value of r_α^o extrapolated to temperature of OK - represents same physical quantity as r_z but differs in origin since it is derived from electron diffraction data
r_{av}	same physical quantity as r_z and r_α^o, but derived from simultaneous refinement of spectroscopic and electron diffraction data.

Table 2

EXPERIMENTAL DATA ON MOLECULAR GEOMETRIES

METHANE

r_g(C-H) = 1.106 Å

r_a(C-H) = 1.101 Å r_e = 1.085 Å

r_o(C-H) = 1.094 Å

ACETYLENE

r_o(C-H) = 1.057 Å r_e = 1.062 Å

r_o(C-C) = 1.209 Å r_e = 1.203 Å

ETHYLENE

r_g(C=C) = 1.337 Å	r_g(C-H) = 1.103 Å	HCH = 117.2°
r_o(C=C) = 1.338 Å	r_o(C-H) = 1.086 Å	HCH = 117.4°
r_z(C=C) = 1.338 Å	r_z(C-H) = 1.087 Å	HCH = 117.4°
r_α°(C=C) = 1.339 Å	r_α°(C-H) = 1.085 Å	HCH = 117.9°
r_e(C=C) = 1.334 Å	r_e(C-H) = 1.081 Å	HCH = 117.4°

BENZENE

r_a(C-C) = 1.397 Å	r_a(C-H) = 1.100 Å
r_g(C-C) = 1.399 Å	r_g(C-H) = 1.101 Å
r_α°(C-C) = 1.395 Å	r_α°(C-H) = 1.091 Å
r_z(C-C) = 1.396 Å	r_z(C-H) = 1.083 Å
r_{av}(C-C) = 1.396 Å	r_{av}(C-H) = 1.085 Å

BUTADIENE

r_g(C=C) = 1.345 Å, r_g(C-C) = 1.465 Å, r_g(C-H) = 1.094
CCC = 122.9°

r_α°(C=C) = 1.342 Å r_α°(C-C) = 1.463 Å, r_α°(C-H) = 1.093 Å
CCC = 123.6°

r_{av}(C=C) = 1.341 Å r_{av}(C-C) = 1.463 Å, r_{av}(C-H) = 1.090 Å
CCC = 123.3°

Table 3(a)

Ethylene - Scaled Minimal Basis

Centre/coordinate		Cycle 1		2	
		g	s	g	s
Cl	X	0.49075	-0.49075	-2.810775	0.55754
Hl	X	-0.73654	0.73654	-0.307455	-0.36424
Hl	Y	-0.75218	0.75218	-0.121285	-0.32087
α		1.22869		1.94959	
S		2.21692		1.25069	
σ		2.72391		2.43832	

3		4		5	
g	s	g	s	g	s
0.274365	-0.02293	-0.114475	-0.01336	-0.21841	0.059
-0.24153	0.35874	-0.06424	0.12007	-0.04516	0.17111
0.03179	0.09303	0.31169	-0.13866	-0.029805	0.00496
2.05475		8.81542		2.49746	
0.74192		0.36733		0.35238	
1.52446		3.23816		0.88007	

Geometrical Parameter	Cycle 0	1	2	3	4	5
r(C=C)	1.3340	1.3276	1.3391	1.3386	1.3374	1.3389
r(C-H)	1.081	1.089	1.091	1.094	1.092	1.093
$\hat{HCH}$	117.4	116.9	117.3	117.0	116.0	115.9
E^T	0	-0.790	-1.308	-1.401	-1.583	-1.602

Table 3(b)

Ethylene - Unscaled Double-Zeta Basis

Centre/coordinate		Cycle 1		2	
		g	s	g	s
Cl	X	2.29257	-2.29257	-0.20139	-0.07395
Hl	X	0.04866	-0.04866	0.21783	-0.19899
Hl	Y	0.408955	-0.408955	0.527345	-0.51335
α		0.70531		2.87244	
S		3.34518		1.10609	
σ		2.35939		3.17718	

3	
g	s
-0.008205	0.00613
-0.04081	0.04043
0.014985	-0.01519
10.10144	
0.08681	
0.87693	

Geometrical Parameter	Cycle 0	1	2	3
r(C=C)	1.3340	1.3169	1.3146	1.3153
r(C-H)	1.081	1.084	1.076	1.077
$\hat{HCH}$	117.4	116.5	116.3	116.1
E^T	0	-1.053	-1.520	-1.530

Table 4

Trans-1,3-butadiene-Scaled Minimal Basis

Centre/coordinate		Cycle 1		2	
		g	s	g	s
C3	X	-2.86765	2.86765	-0.04545	0.4933
C3	Y	-2.01725	2.01725	0.4286	-0.04014
C4	X	-0.285	0.285	-0.3204	0.31455
C4	Y	0.4284	-0.4284	-1.01345	0.78391
H4	X	0.3327	-0.3327	-0.33305	0.22716
H4	Y	0.0937	-0.0937	0.30195	-0.26869
H5	X	-0.5326	0.5326	-0.57665	0.56925
H5	Y	-0.53055	0.53055	-0.3761	0.40034
H6	X	0.6146	-0.6146	0.4926	-0.51161
H6	Y	-0.5391	0.5391	-0.6279	0.61336
α		0.84367		2.06446	
S		5.27455		2.09975	
σ		4.44998		4.33486	

3		4	
g	s	g	s
0.7046	-0.02565	-0.5486	0.16386
-0.5983	0.19816	1.02485	-0.13984
-0.23465	0.25759	0.5752	0.01536
0.94875	0.05501	-0.0198	0.02655
-0.4543	0.29949	0.6022	0.03422
0.34765	-0.27791	-0.3731	-0.09037
0.0005	0.30147	-0.1889	0.27507
-0.6422	0.45985	-0.30095	0.43035
-0.2224	-0.18421	-0.50445	0.01891
-0.98075	0.70493	-0.7520	0.74444
5.37275		5.58286	
1.48959		1.32068	
8.00319		7.37317	

Table 4 (contd.)

Geometrical Parameter	Cycle 0	1	2	3	4
r(C1=C2)	1.344	1.333	1.342	1.338	1.343
r(C2-C3)	1.467	1.498	1.507	1.512	1.515
r(C1-H1)	1.094	1.093	1.084	1.082	1.094
r(c1-H2)	1.093	1.094	1.087	1.092	1.091
r(C2-H3)	1.094	1.086	1.085	1.100	1.096
$\hat{C1C2C3}$	122.9	122.5	122.1	122.9	122.2
$\hat{C1C2H3}$	119.5	120.8	121.1	121.1	121.3
$\hat{C2C1H1}$	119.5	119.3	119.4	120.9	121.7
$\hat{C2C1H2}$	119.5	120.2	120.3	120.5	121.4
E^T	0	-3.046	-4.274	-6.129	-7.499

Table 5(a)

1,2,5-Thiadiazole

Centre/coordinate		Cycle 1		2	
		g	s	g	s
S	Y	8.83376	-8.83376	1.3746	-3.04079
N2	X	2.77082	-2.77082	0.19145	-0.75494
N2	Y	-1.47643	1.47643	-0.89505	1.06009
C3	X	2.20171	-2.20171	-0.4821	-0.07379
C3	Y	-2.51363	2.51363	2.3391	-1.39949
H3	X	0.06648	-0.06648	1.39225	-1.16918
H3	Y	-0.42682	0.42682	-2.12845	1.85738
α		1.31118		1.17735	
S		10.97519		5.11816	
σ		14.39045		6.02587	

3		4		5	
g	s	g	s	g	s
-0.09195	-0.37104	-0.1316	0.18779	0.006	
0.0857	-0.18708	-0.05035	0.05781	-0.0455	
0.811725	-0.55306	0.09555	-0.14883	-0.0092	
0.8141	-0.72348	-0.18745	0.14071	0.0141	
-0.24195	0.00352	0.08435	-0.03272	0.0173	
0.2069	-0.35497	0.0758	-0.05333	0.0005	
-0.5238	0.7347	-0.1141	0.08771	-0.0101	
2.21802		1.22102		-	
1.78826		0.38609		-	
3.96639		0.47142		-	

Table 5(a) (contd.)

Geometrical Parameter	Cycle 0	1	2	3	4
r(S-N)	1.631	1.692	1.713	1.713	1.711
r(N=C)	1.328	1.336	1.324	1.327	1.328
r(C-C)	1.420	1.451	1.451	1.468	1.467
r(C-H)	1.079	1.059	1.079	1.083	1.085
$\hat{NSN}$	99.6	96.7	95.6	95.8	95.9
$\hat{SNC}$	106.4	107.9	107.9	108.2	108.2
$\hat{NCC}$	113.8	113.8	114.3	113.9	113.9
$\hat{CCH}$	126.2	126.1	125.7	125.3	125.3
E^T	0	-20.68	-24.35	-25.20	-25.22

Table 5(b)

1,3,4-Thiadiazole

Centre/coordinate		Cycle 1		2		3	
		g	s	g	s	g	s
S	Y	3.0002	-3.0002	2.79565	-4.84130	1.3806	-1.52338
N3	X	6.70133	-6.70133	-0.08675	5.95499	1.34435	0.46325
N3	Y	-3.15920	3.15920	-4.79045	6.55494	-0.82995	0.84209
C2	X	3.06845	-3.06845	-0.92455	3.42488	1.5376	-0.67774
C2	Y	1.47690	-1.47690	2.71605	-3.44040	0.76225	0.63601
H2	X	0.85	-0.85	-2.1603	2.45122	-2.64275	2.15222
H2	Y	1.822	-1.822	0.6764	-0.69375	0.9019	-0.71643
α		0.96561		0.59828		1.52443	
S		12.24972		15.50511		4.02324	
σ		11.82845		9.27640		6.13314	

4	
g	s
0.167	-0.31512
1.65455	-0.72112
-0.6318	0.20123
-1.11339	0.68848
0.14415	0.36271
-0.3.88	0.56505
0.4039	-0.40623
-	
1.84392	
-	

Table 5(b) (contd.)

Geometrical Parameter	Cycle O	1	2	3
r(S-C)	1.721	1.738	1.749	1.758
r(N-C)	1.302	1.287	1.313	1.312
r(N-N)	1.371	1.442	1.480	1.487
r(C-H)	1.077	1.064	1.060	1.084
$\hat{CSC}$	86.3	86.8	87.1	86.0
$\hat{SCN}$	114.6	115.1	115.7	116.7
$\hat{CNN}$	112.2	111.6	110.7	110.3
$\hat{SCH}$	122.5	123.0	123.5	122.7
E^T	O	-12.592	-21.335	-24.870

Table 5(c)

1,3,4-Oxadiazole

Centre/coordinate		Cycle 1		2		3	
		g	s	g	s	g	s
O	Y	-9.87085	9.87085	3.7265	-0.05022	0.55955	
N3	X	7.9351	7.9351	-0.81305	0.53883	0.0013	
N3	Y	4.011	-4.011	1.35725	-1.44412	-0.45535	
C2	X	2.1671	-2.1671	0.1072	-1.72456	0.19165	
C2	Y	2.52733	-2.52733	-0.58255	0.48574	0.87145	
H2	X	2.925	-2.925	0.9535	-1.01781	0.3496	
H2	Y	-1.6029	1.6029	-0.96075	0.98322	-0.6952	
α		1.42301		1.39681		-	
S		17.31932		3.89610		-	
σ		24.64556		5.44211		-	

Geometrical Parameter	Cycle 0	1	2
r(O-C)	1.36	1.442	1.440
r(N-C)	1.30	1.317	1.332
r(N-N)	1.37	1.490	1.482
r(C-H)	1.077	1.091	1.098
$\hat{COC}$	106	100.9	101.2
$\hat{OCN}$	105	113.4	113.2
$\hat{CNN}$	112	106.2	106.2
$\hat{OCH}$	116	120.5	120.6
E^T	0	-51.357	-53.028

Table 5(d)

1,2,3-Thiadiazole

Centre/coordinate		Cycle 1		2		3	
		g	s	g	s	g	s
S	X	-0.08065	0.08065	1.6565	-0.13852	0.92825	-0.78988
S	Y	4.47335	-4.47335	1.9418	-2.2199	0.3037	-0.61304
N2	X	-5.2702	5.2702	-0.92945	1.40678	-0.7214	0.81805
N2	Y	4.7036	-4.7036	-2.27525	1.50708	0.4963	-0.1412
N3	X	-3.3292	3.3292	-1.96265	2.11259	0.5266	-0.08716
N3	Y	-6.08445	6.08445	1.69505	-0.83892	-0.049	-0.11901
C4	X	7.18995	-7.18995	-0.008	-0.78382	-0.33695	0.16644
C4	Y	-0.00535	0.00535	-0.0885	0.07934	-0.1879	0.1685
C5	X	1.13755	-1.13755	-0.23365	-1.09141	-1.10865	0.73391
C5	Y	-2.5330	2.5330	-1.50965	-0.09347	-0.1431	0.09498
H4	X	1.9855	-1.9855	1.08595	-1.14049	0.60035	-0.6278
H4	Y	-0.2564	0.2564	0.4182	1.37148	-0.4347	0.58995
H5	X	-0.08065	0.08065	0.3913	-0.36626	0.1118	-0.15373
H5	Y	4.47335	-4.47335	-0.18165	0.19457	0.0147	0.01989
α		0.99766		1.81627		-	
S		14.19335		4.42024		1.77751	
σ		14.16014		8.02835		-	

Table 5(d) (contd.)

Geometrical Parameter	Cycle 0	1	2
r(S-N)	1.692	1.711	1.746
r(N=N)	1.290	1.346	1.323
r(N-C)	1.366	1.421	1.449
r(C=C)	1.369	1.343	1.346
r(C-s)	1.689	1.719	1.740
r(C4-H4)	1.078	1.062	1.075
r(C5-H5)	1.081	1.078	1.071
$\hat{CSN}$	92.9	92.8	92.0
$\hat{SNN}$	111.2	111.1	111.5
$\hat{NNC}$	114.0	112.0	112.5
$\hat{NCC}$	114.1	114.6	114.3
$\hat{CCS}$	107.8	109.5	109.7
$\hat{H4C4C5}$	126.6	129.4	129.4
$\hat{H5C5S}$	122.6	121.2	120.9
E^T	0	-22.45	-27.07

Table 5(e)

1,2,3-Oxadiazole

Centre/coordinate		Cycle 1		2	
		g	s	g	s
O	X	-5.34615	5.34615	-0.7635	1.2481
O	Y	10.21475	-10.21475	2.32565	-3.158
N2	X	12.2899	-12.2899	-0.23075	-1.0977
N2	Y	7.9921	-7.9921	-5.83045	4.3534
N3	X	1.32795	-1.32795	4.25225	-3.9341
N3	Y	-20.6721	20.6721	1.0011	1.2995
C4	X	-3.26585	3.26585	-1.32045	1.5243
C4	Y	5.2582	-5.2582	-0.29165	-0.2975
C5	X	-5.6869	5.6869	-1.24405	1.7129
C5	Y	-0.76835	0.76835	3.5802	-3.1122
H4	X	2.17395	-2.17395	2.2743	-2.2594
H4	Y	-4.20765	4.20765	-3.1209	3.2303
H5	X	-1.4929	1.4929	-2.9678	2.8058
H5	Y	2.1804	-2.1804	2.33605	-2.3151
α		0.85461		2.12576	
S		29.62424		9.62587	
σ		25.31717		20.46229	

Table 5(e) (contd.)

Geometrical Parameter	Cycle 0	1	2
r(O-N)	1.36	1.47	1.51
r(N=N)	1.24	1.30	1.34
r(N-C)	1.42	1.48	1.49
r(C=C)	1.35	1.32	1.36
r(C-O)	1.36	1.41	1.40
r(C4-H4)	1.08	1.09	1.08
r(C5-H5)	1.08	1.07	1.07
$\hat{\text{CON}}$	106	104	104
$\hat{\text{ONN}}$	108	108	108
$\hat{\text{NNC}}$	114	111	109
$\hat{\text{NCC}}$	101	103	106
$\hat{\text{CCO}}$	111	114	113
$\text{H4}\hat{\text{C4}}\text{C5}$	111	115	116
$\text{H5}\hat{\text{C5}}\text{O}$	123	122	122
E^T	0	-82.1	-113.2

Table 6

GEOMETRY OPTIMISATION - RELATIVE ENERGIES (KJ mol^{-1})

(a) METHANE

r(CH)/Å	1.080	1.085	1.090	1.015	1.100
	0.5	0	0.3	0.5	0.6

$r(CH)^{MIN}$ = 1.087 Å

E_T = -40.1015 a.u.

(b) ACETYLENE

r(CH)/Å	1.05	1.10	1.15
r(CC)/Å			
1.190	4	1	21
1.210	4	0	20
1.230	9	5	25
1.250	20	15	35

$r(CC)^{MIN}$ = 1.200 Å, $r(CH)^{MIN}$ = 1.080 Å

E_T = -76.6297 a.u.

(c) BENZENE

r(CH)/Å	1.05	1.09	1.15
r(CC)/Å			
1.35	96	73	118
1.37	58	33	81
1.39	32	9	53
1.41	24	0	43
1.43	28	4	46

$r(CC)^{MIN}$ = 1.413 Å, $r(CH)^{MIN}$ = 1.093 Å

E_T = -230.1142 a.u.

Table 6 (contd.)

(d) Trans-Butadiene ("Minor" parameters fixed at exptl. values)

r(C=C)/Å		r(C-C)/Å	
1.330		1.420	16
1.344		1.467	5
1.400		1.500	0
		1.540	2
			1
			20

$r(C=C)^{MIN}$ = 1.515 Å, $r(C-C)^{MIN}$ = 1.340 Å

E_T = -154.5152 a.u.

at r(C=C) = 1.344, r(C-C) = 1.500:

CCC	110°	120°	130°	140°
E_T	54	0	6	64 → CCC^{MIN}=124°

(e) CIS-BUTADIENE

Almost identical to trans, except for CCC variation:

E_T	235	8	0	64(140°)→CCC^{MIN}=12€

E_T = -154.5106 a.u.

(f) 90°-BUTADIENE (1.300)

	24	
19	6	29
	2	
15	0	

$r(C=C)^{MIN}$ = 1.525 Å, $r(C-C)^{MIN}$ = 1.335 Å

E_T = -154.5060 a.u.

Table 6 (contd.)

(g) 2,3-DIAZABUTADIENE (METHANAL AZINE)

(TRANS)

	r(N-N)/Å	1.38	1.42	1.45	1.50	1.54
r(C=N) = 1.28, CNN = 105°:	E_T	39	28	25	27	35

→ $r(N-N)^{MIN}$ = 1.462 Å, E_T = 24

r(N-N) = 1.50, CNN = 105°: r(C=N)/Å	1.24	1.28	1.32	
E_T	58	27	25	

→ $r(C=N)^{MIN}$ = 1.303 Å, E_T = 23

r(N-N) = 1.50, r(C=N) = 1.28:

CNN/DEG	100	105	110	120
E_T	159	27	19	27

→ CNN^{MIN} = 114.4°, E_T = 4

OVERALL E_T^{MIN} = -186.3315

EXPTL: r(N-N) = 1.418, r(C=N) = 1.278

r(C-H) = 1.095, CNN = 111.4°, NCH = 121.1°

Table 7(a)

ORBITAL ENERGIES AND OBSERVED I.P.'s OF SMALL MOLECULES

	I.P.(OBS)/eV	
METHANE (T_d)	13.6	$1t_2$
	14.4	
	15.0	
	22.9	$2a_1$
ETHYLENE (D_{2h})	10.51	$1b_{3u}(\pi)$
	12.5	$1b_{3g}$
	14.8	$3a_g$
	15.9	$1b_{2u}$
	19.1	$2b_{1u}$
	23.7	$2a_g$
ACETYLENE ($D_{\infty h}$)	11.40	$1\pi_u$
	16.7	$3\sigma_g$
	18.7	$2\sigma_u$
	23.5	$2\sigma_g$
PROPYLENE (C_s)	9.73	2a"
	12.2	10a'
	13.1	9a'
	14.4	1a"
	14.4	8a'
	15.9	7a'
	18.2	6a'
	21.9	5a'
	23.7	4a'

Table 7(a)(contd.)

BUTATRIENE (D_{2h})	9.15	$1b_{3g}(\pi)$
	9.98	$2b_{3u}(\pi)$
	11.70	$1b_{2u}(\pi)$
	14.2	$1b_{2g}$
	15.0	$1b_{3u}$
	15.5	$5a_{g}$
	16.5	$4b_{1u}$
	20.6	$4a_{g}$
	23.0	$3b_{1u}$
VINYL ACETYLENE (C_s)	9.58	2a"(π)
	10.58	12a'(π)
	12.0	1a"(π)
	13.2	11a'
	15.2	10a'
	16.1	9a'
	17.4	8a'
	19.7	7a'
	22.9	6a'
BUTADIENE (C_{2h})	9.03	$1b_{g}(\pi)$
	11.46	$1a_{u}(\pi)$
	12.2	$7a_{g}$
	13.4	$6b_{u}$
	13.9	$6a_{g}$
	15.5	$5b_{u}$
	15.5	$5a_{g}$
	18.1	$4b_{u}$
	19.2	$4a_{g}$
	22.6	$3b_{u}$
	24.8	$3a_{g}$

Table 7 (b)

ORBITAL ENERGIES FOR BENZENE AS A FUNCTION OF BASIS SET (eV)

	STO-3 G	Minimal Basis	Double Zeta
$1e_{1g}$	7.57	9.82	8.96
$3e_{2g}$	11.79	13.51	12.89
$1a_{2u}$	12.33	14.46	13.48
$3e_{1u}$	14.54	16.27	15.71
$1b_{2u}$	14.99	17.07	16.48
$2b_{1u}$	15.77	17.34	17.19
$3a_{1g}$	18.00	19.38	19.12
$2e_{2g}$	20.87	22.35	22.20
$2e_{1u}$	25.94	27.61	27.54
$2a_{1g}$	29.57	31.41	31.33

Table 8

ANALYSIS OF MOLECULAR TOTAL ENERGY IN AROMATICS -

MODEL MOLECULES ($\bar{E}$NERGY, a.u.)

CH_4 (CH)	40.10180	NH_3 (NH)	56.0199	H_2O (OH)	75.7999
C_2H_4 (C=C)	77.83143	$CH_3CH=CH_2$ (C-C)	116.77453	$CH_2=CH-CH=CH_2$ Perpendicular (C-C)	154.50920
$CH_2=CH-NH_2$ Perp. (C-N)	132.70390	$CH_2=CH-OH$ Perp. (C-O)	152.46202	$CH_2=CH-F$ (C-F)	176.40569
CH_3CN (C≡N)	131.55927	$CH_2=O$ (C=O)	113.51009	C_2H_2 (C≡C)	76.60453

Table 9

BOND ENERGIES (a.u.)

C-H	10.02545	N-H	18.6733	O-H	37.89995
C=C	37.72963	C-C	18.89220	C-C (sp^2)	18.89724
C-N	27.55132	C-O	46.75609	C-F	108.59971
C≡N	101.48292	C=O	93.45919	C≡C	56.55363

Table 10

BENZENE RESONANCE ENERGY (EXPTL)

Reactions	ΔH^{o} = SE(RE)/kJ mol^{-1}
Isodesmic	
$C_6H_6+6CH_4 \rightarrow 3CH_2=CH_2+3CH_3-CH_3$	270 ± 5
$C_6H_6+3CH_3-CH_3 \rightarrow 3CH_2=CH_2+C_6H_{12}$	205 ± 4
$C_6H_6+2C_6H_{12} \rightarrow 3C_6H_{10}$	134 ± 4
$C_6H_6+3CH_3-CH_3+3CH_2=CH_2 \rightarrow 6CH_3-CH=CH_2$	138 ± 8
Homodesmotic	
$C_6H_6+3CH_2=CH_2 \rightarrow 3CH_2=CH-CH=CH_2$	91 ± 4

Table 11

RESONANCE ENERGIES (CALC) kJ mol^{-1}

C_6H_6	$C_{10}H_8$	PhF	PhCHO	PhOH	$PhNH_2$	C_6F_6
213	358	213	225	184	204	69

Furan	Pyrrole
89	88

Benzofuran	iso-Benzofuran	Indole	Isoindole	Dehydro[14]annulene
233	148	238	182	326

Trans-Butadiene	Cis,Cis-Hexatriene	Cyclooctatetraene
23	32	15

Barrelene
-480

CHAPTER FIVE

APPLICATION TO ANNULENES AND THEIR DERIVATIVES

"A number of larger ring annulenes have been synthesised in recent years, and a study of their behaviour has contributed considerably to our understanding of the concept of aromaticity."

F. Sondheimer

In this chapter, there are reported non-empirical calculations on some "large" molecules of organic chemistry, namely the Annulenes. A selection of these are considered, together with some of their derivatives, which are dehydro-annulenes and "bridged" annulene species. The unifying concept in the study of such compounds is that of Aromaticity, and so the first part of this chapter summarises the relevant background information, of both an experimental and a theoretical nature, on the idea of aromaticity in chemistry.

In "pre-electron" times, the designation "aromatic" was first used by chemists to categorise a specific class of substances (first half of the nineteenth century). During the ensuing 150 years, a considerable amount of research has been carried out in detailing, defining, and defending the notions to which the term has been applied. Although our understanding of the molecular and electronic structure of matter is far more sophisticated than that of the early 1800's, a precise and generally acceptable definition of aromaticity remains elusive. The earlier and still widely acknowledged viewpoint focuses attention on reactivity, or ground-transition-state characteristics of molecules as a reflection of aromatic properties; the alternative and increasingly popular view treats the physical (or ground-state only) properties as the

key to the issue. The present conceptual re-evaluation has been stimulated by a remarkable series of syntheses over the last twenty years. There has been considerable activity in the study of non-benzenoid compounds in particular, in the areas of theoretical, synthetic, and physical organic chemistry.[1] Before presenting results on specific species, some consideration must be given to experimental facts and theoretical ideas on aromaticity.

A. Aromaticity - Experimental Evidence.

The history of the concept of aromaticity is well-documented;[2] the general idea arose when it was observed that carbon compounds can be put into two groups, aromatic or aliphatic (non-aromatic), which differ in their physical and chemical properties to a sufficient extent to make the distinction a useful one. It has remained an inexact concept for which it is probably impossible to find a rigorous definition, but, nevertheless, this broad concept has been extensively used in chemistry. No generally acceptable experimental criterion of aromaticity exists at present; some attempts at definitions have been made, based on experimental facts such as those given below. All the properties of a molecule are potential sources of information about its aromaticity, although some are doubtlessly more sensitive to the "aromatic nature" of the electronic organisation than others. The following paragraphs survey briefly some of the molecular properties which seem likely to have a direct bearing on considerations of aromaticity.

(a) Molecular Geometry:- The geometry of a molecule can be direct and convincing evidence that it is like benzene, the prototype aromatic compound, and therefore aromatic. The aspect of bond length has been used

to class a molecule as aromatic if its carbon-carbon bonds are 1.36-1.43 Å in length, and as polyenic if it has alternating bond lengths of 1.34-1.356 Å for double bonds and 1.44-1.475 for single bonds.[3] A definition based on bond length has the disadvantages that a particular bond or bonds may be "unusual" (e.g. the transannular bond of azulene), that extension to hetero-atom systems is difficult, and that, most significantly, the bond lengths are unknown in the vast majority of cases. In addition, as in Chapter 4 Section A, the determination of bond lengths in practice is complicated by vibrational motion so that, in general, measurements do not yield results of sufficient precision. Another geometric feature which is relevant here is the planarity, or near planarity, of the ring atoms of the molecule. Quantitative measurement is less easy here. A severe problem is that there is very little geometrical information available for hydrocarbon ions. Thus, aromaticity defined in terms of molecular geometry is not a general, practicable criterion.

(b) Molecular Energetics:- Energetic information about the ground state of a molecule can be recorded thermodynamically as the standard heat of formation of a compound, ΔH_f° (temp), which is the energy required to form the molecule in its standard state from the elements in their standard states, at a particular temperature; the usual thermodynamic sign convention applies.[4] More useful for theoretical purposes is the atomisation energy, or heat of atomisation, H_a(temp) of the molecule. The latter quantity differs from ΔH_f° in that it uses as reference point the gaseous atoms and not the elements in their standard states. Heats of atomisation are always conventionally positive. The main advantage in using atomisation energies is that the resulting number is approximately the sum of the bond energies. Thus, using experimental thermochemical data, once the total energy of a number of molecules is known, comparison can be made to see if a given molecule has a total energy which is expected on the basis of electron organisation into simple "two-electron" bonds, i.e. whether the

atomisation energy is approximately the sum of the bond energies. The difficulty which arises is that reference molecules are generally hypothetical, such as using cyclohexatriene as a reference for benzene, and the estimation of energies involved is a controversial matter.[5] The estimation of "resonance" energy as the difference in energy between the real molecule and a hypothetical reference thus leads to various numerical values depending upon assumptions made in the calculation. Thus, combustion and hydrogenation experiments do not lead to rigorously defined resonance, or stabilisation energies; further complications arise with heterocyclic molecules, with ionic species, and with "strained" molecules.[6]

Further energetic information can be derived from equilibrium processes between two stable molecules, where only the energy difference between the two species is directly measurable so that care must be taken in inferring results.[7] Also, chemical-reactivity energetics, associated with the earliest notions of aromaticity, have been used in developping a criterion for aromaticity. However, there are interrelated (kinetic) aspects to be considered in this case, namely reaction rates and mechanisms, and the nature of the products of reaction, so that deductions from chemical reactivity using ground-state and transition-state information can be unreliable. Thus, "classical" experimental definitions of aromaticity, in terms of thermodynamics or chemical behaviour, are now not regarded as being particularly illuminating; however, in spite of the problems involved, the determinations of empirical resonance energies have provided one of the widest ranging and most often quoted series of data for assessing aromatic character on a quantitative basis.

(c) Spectroscopic Criteria:- Nuclear Magnetic Resonance Spectroscopy, which has developed into a very widely used tool in chemical investigations, has been applied extensively to relevant types of molecules; from the

model applied in analysing results, it has been suggested that the "ring current" deshielding of protons is sufficiently diagnostic of aromatic character for an aromatic compound to be defined as one which can sustain an induced ring current.[8] Further, it was proposed that the magnitude of the ring current might present a quantitative assessment of the aromaticity of a compound; measured proton chemical shifts in N.M.R. spectra have been used in considerations of the ring current criterion. There are complications with this attractive, straightforward picture of using proton N.M.R. results to assign degrees of aromaticity. Particularly relevant in certain annulene species,[9] the diamagnetic effect above is not dominant, and the N.M.R. results can be explained by invoking a dominating paramagnetic effect which reverses the ring current and consequently reverses the positions of the ring proton signals in the spectra. Attempts to make the ring current concept into a quantitative one have met with serious difficulties;[10] the ring current is not a simple function of the aromaticity of a molecule. There is a problem in choosing non-aromatic model compounds as references.[11] Further reasons for exercising care in interpreting N.M.R. data arise from purely experimental causes; again of relevance in annulene cases, N.M.R. spectra may vary markedly with temperature, and only at low temperatures is unambiguous information obtained.[12] A particularly deceptive situation arises when, at higher temperatures, different conformers interchange more rapidly than the time scale of the N.M.R. experiment can allow for, so that a very simple-looking spectrum emerges from a complicated physical situation. In spite of various difficulties, the ring current concept has been and remains for many authors a useful diagnostic test of aromaticity.

It is also widely felt that the various kinds of N.M.R. coupling constants can yield information about aromaticity. One suggestion is that the size of the $^{3}J_{HCCH}$ coupling constant of vicinal protons is a measure of aromaticity;[13] a large value of ^{3}J is connected with an olefinic bond and a smaller value with an aromatic bond. An important reservation is that the same type of cycle must be compared, and also the presence of heteroatoms adds complications. It can be argued that the ^{3}J coupling constants simply measure bond lengths rather than directly measuring aromaticity.[14] Vicinal coupling constants, and measurements of variation of chemical shifts with variation in solvents used,[15] are potentially useful for providing evidence for "bond fixation" or electron delocalisation, but so far their use has been restricted to a few cases.

Theoretical considerations show that there is some correlation between π-electron density and ^{13}C shifts measured in ^{13}C N.M.R. Spectroscopy.[16] The use of the directly bonded $J_{C-H}(^{13}C)$ coupling constant as a measure of diamagnetic anisotropy and so of aromaticity has also been recommended,[17] since there is said to be a relationship between this coupling constant and the chemical shift of the proton. ^{13}C N.M.R. data has not so far proved to be particularly useful in considerations of aromaticity.

Vibrational spectroscopy, both infra-red and Raman, can be carried out quite routinely. However, the results do not in general clearly show whether the electron organisation in the molecule studied is that characteristic of aromaticity. In so far as the vibrational spectrum is relevant to molecular geometry, in special cases, it can indirectly point to the aromaticity of a molecule, but in general it is not helpful with the aromaticity problem.

It has been suggested that electronic, or ultra-violet, spectroscopy can be used to indicate the aromaticity of a molecule.[18] However, there is the difficulty that two molecules, the ground state and the excited state, are intimately involved in an experiment, and a difference between these two, which differ drastically in electronic structure in general, is what is observed. Thus, it is difficult to draw firm conclusions about the ground state from observation of the U.V. spectrum; nevertheless, there has been frequent use of such observations in comparisons between molecules to indicate degrees of aromaticity.[19]

The measurement of ionisation potentials, via photoelectron spectroscopy or other techniques, is at first sight an attractive means of deciding whether a molecule is aromatic or not. It may seem that the electron energy levels of the molecule are given directly, but the actual observed quantities are differences in total energy of the neutral molecule and of the various states of the positive ion.[20] Thus, the information obtained refers to two species and not directly to one. The interpretation of experimental results is not particularly simple; the direct relevance of ionisation potentials to aromaticity is accordingly quite limited.

Electron spin resonance spectroscopy of radical species can also provide some limited information on charge distribution, and so on aromaticity of radicals.[21]

(d) Magnetic Effects:- Closely related to the dynamic N.M.R. experiment is the measurement of magnetic susceptibility of molecules.[22] Basically, molecules respond differently to a magnetic field along different directions so that the molecular magnetic susceptibility is anisotropic,

and the complete statement of the susceptibility requires the values of the three principal components, χ_1, χ_2, χ_3 of the susceptibility tensor along with the three principal directions along which these components lie. The magnetic susceptibility as commonly measured χ_m(obs) is the average of these three components and is defined by

$$\chi_m(\text{obs}) = \frac{1}{3}(\chi_1 + \chi_2 + \chi_3) \qquad -(1).$$

When dealing with planar molecules, it is generally true that the two in-plane components are about equal so that the anistropy may be defined as

$$\Delta K(\text{obs}) = \chi_3 - \tfrac{1}{2}(\chi_1 + \chi_2) \qquad -(2).$$

It has been appreciated for some time that, since aromatic molecules have marked anisotropies which may be due to their having large ring currents, the anisotropy / may well be a useful measure of aromaticity.[23] However, non-aromatic molecules may have significant anisotropies so that allowance must be made for "local" contributions with some kind of group additivity scheme.[24] Alternatively, the χ_m(obs) can be compared with a calculated value of this quantity, derived on some simple additivity assumption with correction terms. The outcome for aromatic molecules is that, because of the ring current, χ_3 is large and χ_m(obs) is larger than that calculated for a non-aromatic reference molecule.[25] The difference is called the susceptibility exaltation,

$$\Lambda = \chi_m(\text{obs}) - \chi_m(\text{calc}) \qquad -(3).$$

Broadly speaking, this measure of aromaticity does seem to be useful, since the common feature of electron organisation of aromatic molecules can

be regarded as giving rise to the anisotropy. However, a qualitative rather than quantitative criterion for aromatic character is provided as a result of the nature of the magnetic model used.

Other physical chemistry techniques in related areas, such as measurements of the Faraday Effect of the rotation of the plane of polarisation of light by transparent substances when placed in a magnetic field, within the ring-current model, have led to proposals of criteria for aromaticity;[26] experimental data is lacking at present for the test of such considerations.

(e) Miscellaneous Criteria:- The Dipole Moment is a property of the electronic ground state of a molecule, and so may seem to have some direct relevance to the aromaticity question. This hope is not realised in general. Dipole moments are poorly understood; there are several distinct major contributors to the total so that interpretation is fraught with difficulties.[27] Thus, it seems doubtful that much information about the aromaticity of molecules can be derived from this source.

The Nuclear Quadrupole experiment measures the electric field gradient at the quadrupolar nucleus and is a potentially valuable source of information about the electronic structure (charge distribution).[28] There is the disadvantage that few common nuclei have quadrupole moments; the most promising area of direct interest to the aromaticity question is in ^{14}N-heterocycles.

The Electron Affinity of a molecule, measured, for example, by reductive polarography, is similar to the ionisation potential in that it is not necessarily relevant to the aromaticity of the neutral molecule.

In summary, the criteria for aromaticity described above have been used rather unevenly and are of varying reliability and utility.

The three main sources of information on aromaticity are molecular geometry, atomisation energy, and N.M.R. spectra and the related diamagnetic anisotropy, commonly expressed in terms of the exaltation. On occasion, various other properties do give some insight, less straightforwardly and less decisively. Quantitative application is hindered in general. Basically, aromatic character can be viewed on the basis of electronic and physico-chemical properties, or on the development of studies on the thermodynamic stability and geometry of molecules, which is probably more enlightening for organic chemists.

In this work, the relevance of this section is that it shows the many areas of experimental research which can, and do, provide information to be considered when analysing the results of computations on "aromatic systems", in particular annulene-type species. Details of the results actually obtained by such methods as those above are presented when the computations on the particular species are considered below. As a general comment, existing experimental information needs to be collated and extended, with new quantitative investigations required, before rationalisation of aromatic character can be rigorously performed; interrelationship of the various available methods is not clear at present, and these methods need to be generalised.

B. Aromaticity - Theoretical Ideas.

Aromatic compounds have been the subject of quantum-mechanical studies for the last fifty years. At the present time, the "π-electron" molecular orbital description of the electronic wavefunction of benzene and related species is so familiar that it seems to have the status of a basic truth; however, as stressed in Chapters 2 and 3, no really accurate computation of the wavefunction of a molecule as large and complex as benzene is even remotely possible. At the very beginning of the quantum theory of molecules is found Hückel's application to the benzene ring.[29] The early treatments were characterised by complete neglect of the "σ" electrons and by neglect of explicit computation, accurate or approximated, of the electron-electron interaction. Further refinements were made in the theoretical procedures applied, but the basic assumption of a rigid separation between σ and π electrons was always imposed. Much of this work is of historical importance and represents the extent of quantum chemistry hope in the existence of a simple short-cut whereby simple calculations and far-reaching simplifying assumptions could lead to permanent solution of the very complicated theory of the electronic interaction in a molecular field. This work reports the results of non-empirical SCF calculations on aromatic molecules, with explicit consideration of all the electrons. It has been clearly demonstrated that the σ and π electrons interact so strongly that any approximation which does not give as much attention to the σ electrons as is given to the π electrons is on unsound grounds. Thus, much of the previous empirical and semi-empirical work on aromatic molecules is open to question, not only on its quantitative validity but even on its qualitative validity.

In principle, by calculation of the electronic wavefunctions of the relevant molecules, the problem of aromaticity could be attacked by studying the properties mentioned in Section A above. Criteria for aromaticity based on chemical activity, involving ground-state and excited-state species along with reaction mechanisms, cannot feasibly be studied theoretically by non-empirical methods of wavefunction calculation; such computations, of a very basic chemical nature, are now being carried out on small systems. For more typical aromatic systems, the physical viewpoint of aromaticity, underlining the properties of molecules in the ground state, is more amenable to theoretical considerations. In particular, the concept of resonance energy of conjugated systems is the most useful in practice in considering calculated results; calculated molecular geometries, within the limitations stated in Chapter 4, may provide additional information.

In the ideal situation of being able to compute the exact many-electron molecular wavefunction, it would be hoped that it would be possible to single out some characteristic quality or property which could be called the "aromatic quality" of molecules. It might then even be possible to find a quantitative measure of the degree of aromaticity of a molecule. The primitive example of this type of procedure of distinguishing within a molecular wavefunction some characteristic quality is the idea of the aromatic sextet of the benzene wavefunction;[30] the characteristic quality is a cyclic structure with three low-energy doubly-filled molecular orbitals (π-type). At the present time, within the self-consistent field, Hartree-Fock method of wavefunction computation, it is hoped that the aromatic quality of a wavefunction is already revealed by the one-determinant approximation. Having retreated to the Hartree-Fock level of approximation, the most important assumption

at this point is the orbital approximation (Chapter 2); the exact solution of the non-relativistic, fixed-nucleus Schrödinger equation, in terms of the configuration interaction expansion is

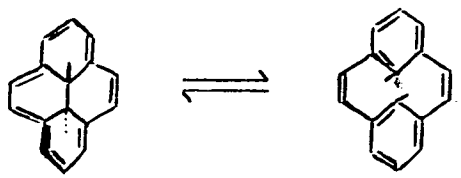
$$\Psi(\text{exact}) = \Psi(\text{HF}) + \sum_i C_i \Psi_i \text{ (excited)},$$

so that the artificiality of the separation into σ- and π-electrons is revealed. This effect of the one-determinant approximation disappears in the exact wavefunction; there is then no distinction between σ- and π-electrons so that the concept of aromaticity will be expressed in terms of all the electrons of a molecule. As a general observation, the results of the type of wavefunction computation of this work has shown the significance of the valence-shell σ-electrons of aromatic systems; previously, these have been assumed to be in simple two-electron chemical bonds, so that the σ-bonds of benzene and ethylene, for example, are considered essentially the same. Non-empirical all-electron calculations have shown an intimate mixing of the σ- and π-molecular orbitals on the energy scale, in general.

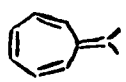
Before presenting the results obtained within the rigorous Hartree-Fock theory, it is of interest to consider the results derived using lower approximations and σ-π separability. Although the σ-electron distribution can drastically affect the π-electron distribution, and computations in which the σ-electrons are ignored can give misleading results, particularly when heteroatoms are involved, some basic concepts are introduced in the simple theories. Of the general theory, where σ-electrons are neglected, the oldest and simplest method is the Hückel Molecular Orbital theory for the calculation of the wavefunctions of π-electron systems.[31] The results relevant to aromaticity which follow from Hückel theory are well-known, and a summary is given below. The development of simple π-electron molecular orbital theory sparked off much interesting non-benzenoid and small ring carbon chemistry, so that aromaticity considerations were not confined

(a)  X = NR , O , S

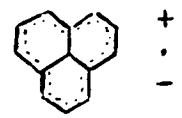
π-EXCESSIVE HETEROCYCLE HETEREPIN HETERONIN

(b) 

DIHYDROPYRENE ⇌ METACYCLOPHANEDIENE

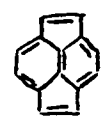
(c) 

HEPTAFULVENE

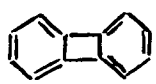
(d) 

PHENALENYL SYSTEM

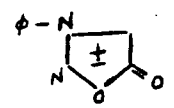
(e) DEHYDROANNULENES

(f) 

PYRACYLENE

(g) 

BIPHENYLENE

(h) 

SYDNONES AND MESOIONIC COMPOUNDS

FIGURE 1 : EXAMPLES OF NON-BENZENOID AROMATICS

to benzene and its homologues. This is exemplified by the species of Figure 1. More specifically, the series of monocyclic hydrocarbons represented by the formula C_pH_p ($p > 2$), the "Annulenes", have been basic to considerations on aromaticity, in a theoretical context. Benzene, the prototype aromatic molecule is a member of this series. Unfortunately, it is only very recently that experimental information on this series as a whole is becoming available.[32] Figure 2 shows the present situation on the members of the annulene series ($p = 3$ to $p = 9$), with neutral and ionic species; these annulenes are represented by polygons, which are usually regarded as being physically reasonable for these smaller members. The famous Hückel definition of aromaticity, summarised in the "(4n+2)" and "4n" rules, strictly applies only to planar monocycles such as these, but is often used in a wider context; simply, rings of p π atomic orbitals with 2,6,10,14,... (4n+2, n=0,1,2,3,....) electrons in a closed-shell ground-state are considerably lower in energy than are the rings with 4,8,12,....(4n,n=1,2,3,....) electrons. The associated "orbital energy" level diagrams, of Figure 3, are familiar representations of the situation. Thus, "4n+2" species are predicted to be stable, singlet aromatic molecules; "4n" species to be unstable, triplet, non-aromatic molecules. The simple rule has been modified by later workers who suggest that the stability of the "4n+2" cycles falls off as n increases until at n = 5 (approx.) the cyclic polyene with alternating single and double bonds is the stable form.[33] Recently, the traditional Hückel molecular orbital approach, which obviously is not based on the physical nature of the molecular problem, has been shown to be a special case of the application of mathematical topology to the problem.[34] The more rigorous approach of this work pays attention to the physical nature of the situation. In particular, as

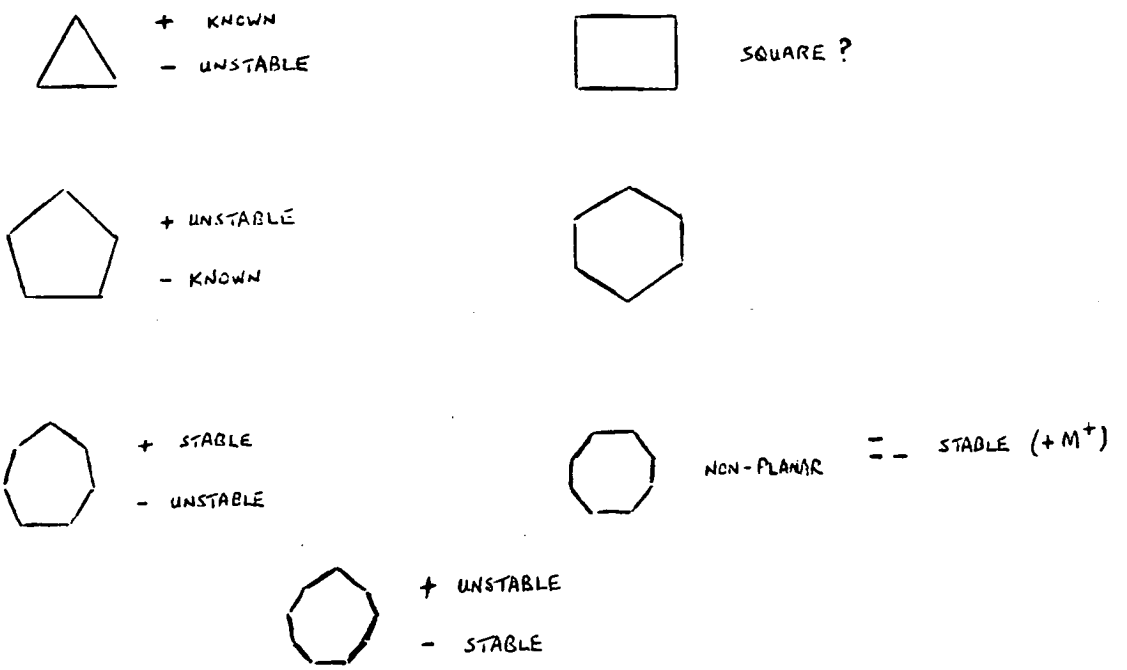


FIGURE 2 : HÜCKEL AROMATICS.

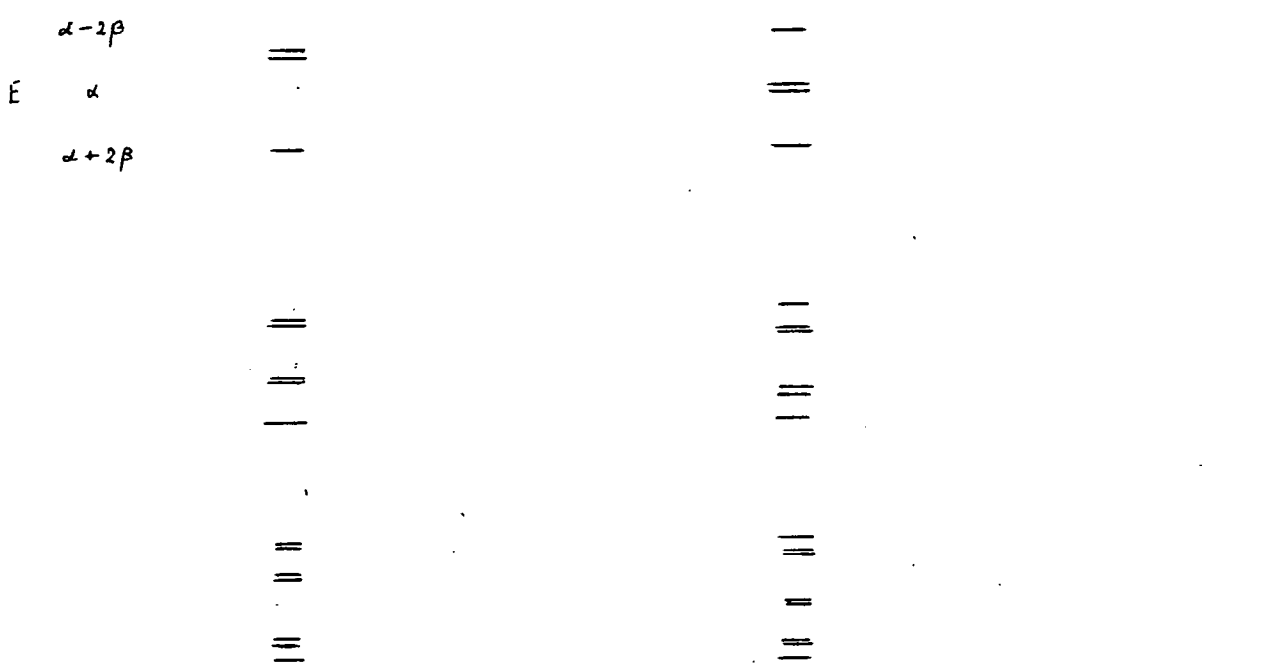


FIGURE 3 : ENERGY LEVEL SCHEME FOR CYCLIC POLYENES.

far as the annulenes are concerned, this is shown by the consideration of alternative conformations to polygons for the larger annulenes ($p > 9$); the effects of interelectronic and internuclear repulsions, and other terms neglected in semi-empirical methods, are included here.

Thus, with some indication of experimental and theoretical background on aromaticity, more specific annulenic species can now be considered.

C. Calculations on Some Even-Membered Annulenes.

Of the polymethines, compounds containing a single ring composed entirely of CH units, the ones with an even number 2n of such units can be represented as neutral cyclic polyenes $(-CH=CH-)_n$, while those with an odd number of CH units must occur as radicals or positive or negative ions. The former type are considered at this point.

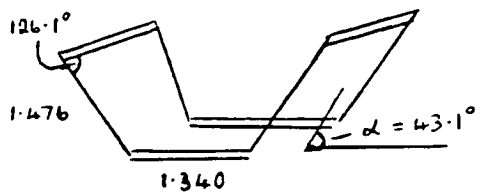
(a) Cyclooctatetraene.

The two smallest [4n] annulenes, cyclobutadiene and cyclooctatetraene, have been of particular historical and chemical interest. Since it was thought from an early stage that the special aromatic properties of benzene must be associated with the fact that two equivalent classical structures can be written for it, and since similar pairs of structures can also be written for cyclobutadiene and cyclooctatetraene, it was clearly of vital importance to chemical theory to see if the latter compounds would also prove to be aromatic. Interest in this field was intensified by the development of resonance theory, which predicted unambiguously that all three compounds should be aromatic. Willstätter in fact prepared cyclooctatetraene in 1911 and showed it <u>not</u> to be aromatic, by "classical" criteria (Section A)[35]. The synthesis was long and difficult; it was some time before it was repeated, but, in recent years, alternative routes have been discovered, enabling routine preparation[36]. Cyclobutadiene, on the other hand, long resisted many determined attempts at synthesis; only recently has it been prepared, and the chemistry of cyclobutadiene systems widely studied[37]. It is usually described as being "antiaromatic", in that its reactivity is such as to suggest that it must be less stable than a normal diene. Semiempirical calculations are said to account for the observed behaviour very well, as both cyclobutadiene and planar cyclooctatetraene are predicted to have negative resonance energies and so to be antiaromatic, unlike the predictions of simpler

resonance and Hückel MO theory[38]. Cyclooctatetraene can escape this fate by abandoning a planar geometry; the molecule does in fact have a tub-shaped structure, in which each double bond is almost perpendicular to its neighbours, the π overlap between them consequently being reduced. However, there is no such loophole for cyclobutadiene; it can escape only by undergoing some chemical reaction, and it is probably the most reactive diene known. Cyclobutadiene, the smallest of the even-membered annulenes, is small enough to have been studied theoretically with more refined methods than that of this work. Despite quite extensive research, including experimental studies, the nature of the electronic structure of the ground state of the molecule is still a controversial matter.[39] Calculations on this rather special case have been performed using the standard basis in related work.[40] Benzene, considered briefly in the preceding chapter, has also been extensively studied.[41] The next higher homologue cyclooctatetraene, is the first of the larger annulenes considered here.

Non-planar ("Tub") Structure.

1,3,5,7-Cyclooctatetraene has been a compound of fundamental importance in the understanding of conjugated π-systems. However, for the present, planar structures are ignored, and the rather accurately known molecular structure is used here.[42] As a typical large molecule, the molecular electronic wavefunction is computed, using an experimentally derived structure, of which the details are given in Figure 4. The geometry is derived from electron diffraction measurements; more recently, crystal structure data have yielded a similar geometry. Using the standard scaled minimal basis set, the non-empirical calculation led to the total energy quantities of Table 1 for the neutral ground state of cyclooctatetraene, at its experimental geometry. The components of the total energy of the molecule illustrate the general situation in which the total energy is the difference between two numerically large terms of opposite sign, namely the nuclear repulsion energy and total electronic energy, which itself is similarly composed of two opposing terms. Such considerations are particularly relevant when final energy differences, such as conformational energy differences, are considered. The resonance energy, as evaluated by the procedure of the preceding chapter, is small and positive. The reference structures' energies are nearly optimal so that the value of the resonance energy is a lower limit. Based on the typical behaviour of the basis set with respect to optimum geometry, it is likely that the main difference between the equilibrium structure which could be calculation and the experimental one used would be in the C-C single bond length; in particular, the value used here of 1.476 Å is probably somewhat shorter (by about 0.05 Å) than that which would be calculated. The experimental structure indicates that the cyclooctatetraene is basically an aggregate of essentially "classical", perpendicular 1,3-butadiene units.

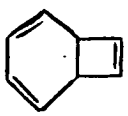


$r(CH) = 1{\cdot}100$ Å

$H\hat{C}C = 117{\cdot}6°$

FIGURE 4 : COT EXPTL. GEOMETRY

(a)

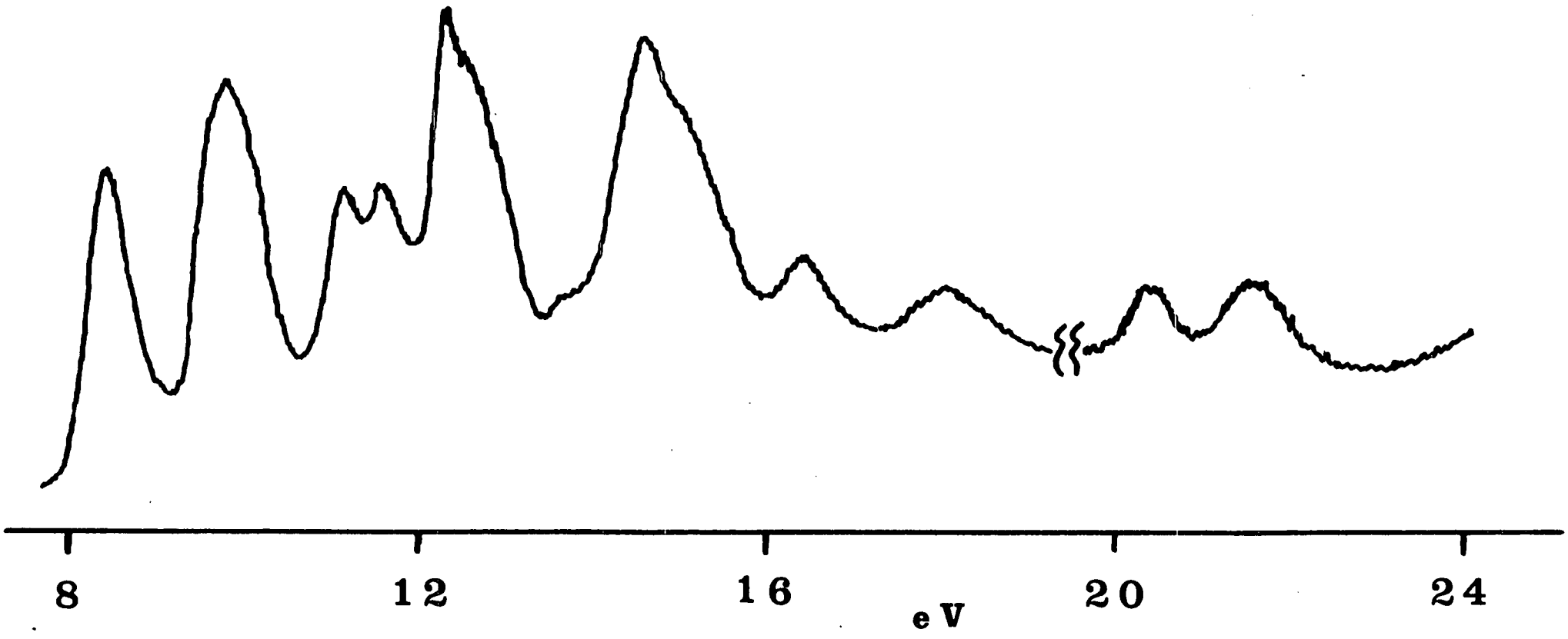


(b) BICYCLO [4.2.0] OCTA-2,4,7-TRIENE

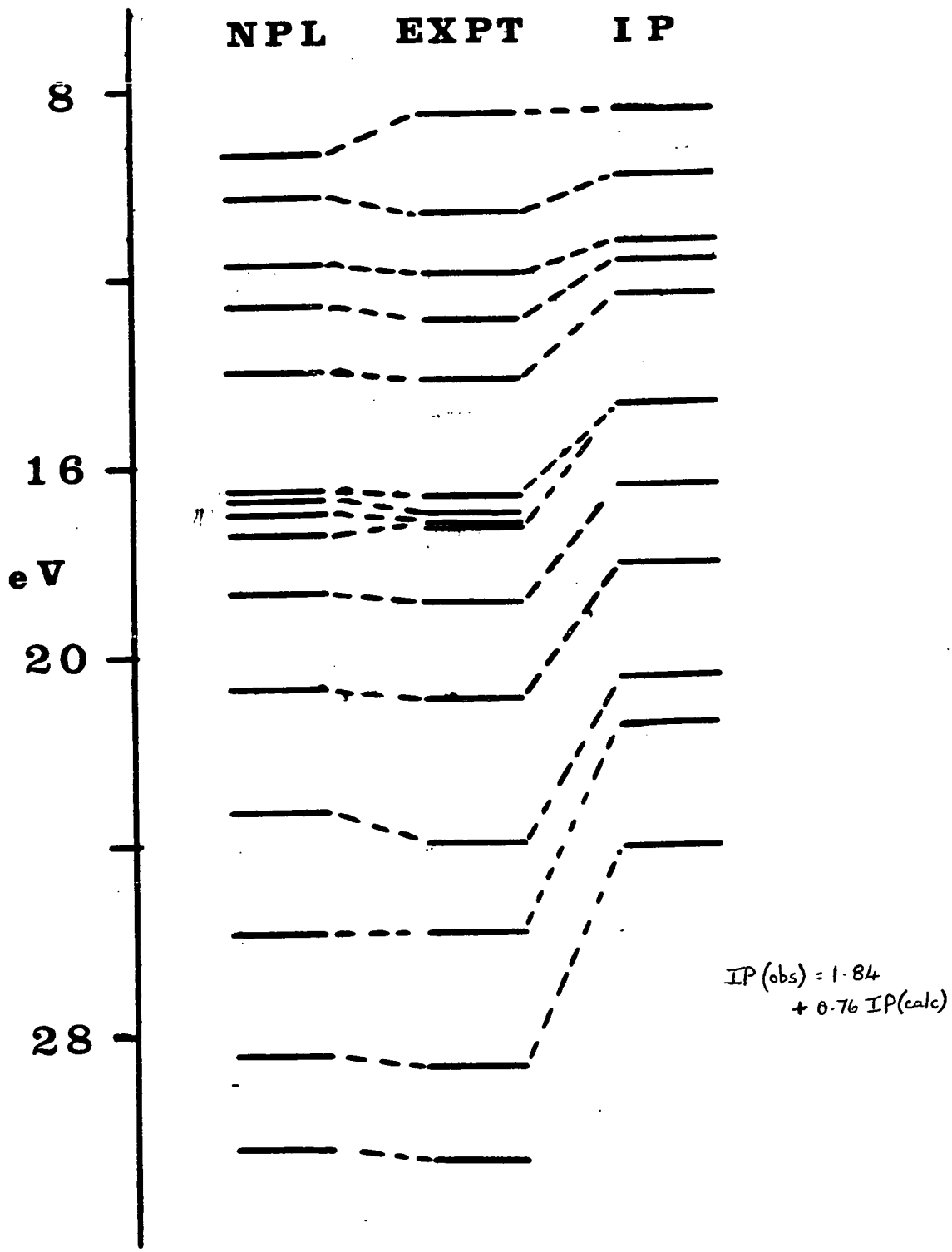
The resonance energy is actually evaluated with reference to such a molecule, and ethylene, so that the non-zero value indicates perhaps the presence of some π-type interaction in cyclooctatetraene leading to a lowering of total energy. There is very little of the usual interaction between neighbouring double-bonds as in planar butadiene. It seems unlikely that optimisation of the molecular geometry would yield a substantially lower total energy, and correspondingly higher resonance energy; in 1,3-butadiene, the corresponding improvement is only about 10 kJ $mole^{-1}$. Thus, the resonance energy probably remains at a value substantially lower than that calculated for benzene (210 kJ $mole^{-1}$). In relation to geometry optimisation, there are reported non-empirical calculations on cyclooctatetraene, using a scaled minimal basis which has fewer primitive functions than the standard one here; partial optimisation of the structure of cyclooctatetraene by separately varying the C=C-C bond angle, r(C=C) and r(C-C) leads to an energy minimum for a non-planar D_{2d} form, with CCC = 127.37°, r(C=C) = 1.340 Å, r(C-C) = 1.507 Å [43]. The total energy obtained is almost 1 a.u. higher than that in Table 1; however, the calculated structure is in reasonable agreement with experimental data, with the rather long C-C bond of particular interest, as the basis set was optimised using ethylene, as in this work.

The photoelectron spectrum of cyclooctatetraene is shown in Figure 5[44]. The experimental valence ionisation energies in this case have been obtained by using He(Iα)(21.22 eV) radiation for excitation. The band positions are tabulated in Table 2; in the above publication, only the first four bonds (lowest IP values) are considered in detail, and, for these, vertical and adiabatic IP's are given. The remaining

F I G . 5 P.E. SPECTRUM OF C_8H_8



F I G . 6 : C_8H_8 ORBITAL ENERGY CORRELATIONS.



band positions have been estimated from Figure 5, up to about 20 eV. Also in Table 2 are given the calculated orbital energies from the above-mentioned computation, and the assignment of the observed IP's is shown. The correlation between observed IP's (vertical) and orbital energies computed is illustrated graphically in Figure 6, where the parameters of the best-fitting line, in the least squares sense, passing through the points are shown. The overall fit is quite satisfactory, and follows the general trend found with the standard basis calculations on similar systems. It seems that the structure used, being that derived from experimental data, is a good approximation to the calculated equilibrium one. The previously reported non-empirical calculation, using a smaller basis set than the standard one here, shows some interesting effects of basis set on computed orbital energies. Using the experimentally derived structure as in this work, the first four orbital energy values (only these are reported) are similar to those here when the basis is the unscaled one. The experimental data are best reproduced by this basis, which yields the poorest total energy. Optimisation of the basis with respect to the total energy of reference compounds (scaling) shifts the IP's by 1.5 eV to 2.5 eV to lower values, without changing their differences by more than about 0.5 eV; methane scaled basis shows poorer agreement with experiment than ethylene scaled basis. In addition, in this study, a basis of the same size as the standard one here is used, yielding a slightly lower total energy than that obtained here; again, the calculated orbital energies, are rather low. One characteristic of the standard basis, quite distinct from total energy or equilibrium structure considerations, is that orbital energies are usually calculated satisfactorily; in particular,

values tend to be over-estimated, but this effect is inherent to Koopmans' Approximation. Systematic under-estimation of IP's is an indication of an unbalanced basis, which, nevertheless, may be satisfactory from a total energy point of view.

Ionisation energies up to approximately 26 eV can be measured by using a source of He(IIα)radiation for excitation. All valence IP's of simple hydrocarbons are found within the easily accessible region of 8-26 eV. The He(IIα)spectrum of cyclooctatetraene is also reproduced in Figure 5. A problem with the larger hydrocarbons is that bands in the spectrum tend to conglomerate, particularly when the molecular symmetry is low, and it becomes very difficult, if not impossible, to deconvolute the overlapping band systems into well-defined individual bands, let alone assign them unambiguously. The band positions have not been tabulated along with the published spectrum[44], with the intention of illustrating the possible precision of interpretation. In this particular case, the ionisation energy scale is calibrated in the region of 15-26 eV. Consequently, although the complete spectrum is reproduced, the scale on the low ionisation energy side, extending from 7-12 eV, is to be taken as a guideline only, and can be in error by approximately $\pm$ 0.2 eV. The observed He(IIα) IP's estimated from the spectrum are included in Table 2, and added to the data for correlation (Figure 6).

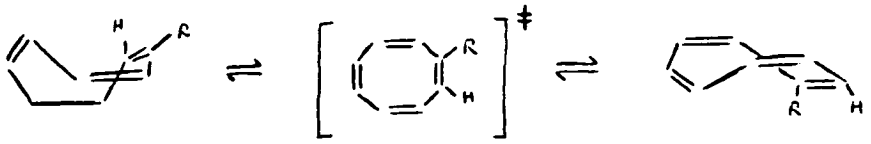
As indicated in Table 2, the first three bands in the PE spectrum of cyclooctatetraene correspond to ejection of an electron from one of the orbitals which are predominantly π-type in character. This conclusion is supported by results derived from a standard Hückel MO model[44], which gives surprisingly good agreement with experiment, since it assumes complete σ/π separation in a strongly non-planar system, thus neglecting considerable through-bond interactions which

are known to occur between the opposing pairs of two-centre π-orbitals[45]. In addition, through-space interaction between the opposite π-orbitals is likely to be of the same magnitude as that between the conjugated π-orbitals, as judged by the corresponding overlap integrals. However, it has been shown that the simple HMO treatment gives reasonable values for the positions and relative spacings of the π-orbital energies as a result of a fortuitous cancellation of interaction terms[44]. The HMO model treats the π-system as being independent of the σ-frame, and also assumes nearest-neighbour interactions only. A satisfactory parametrisation of the experimental results is obtained because of the compensation of the through-space interactions across the ring (stabilising effect) by through-bond interactions with lower lying σ-orbitals, i.e. by a breakdown of the σ/π separation. A more-refined semi-empirical model based on a MINDO/2 treatment, which includes all valence electrons, yields poorer agreement with experimental data, and also predicts the following sequence for the first four IP's (lowest first): $5a_1(\pi)$, $7e(\pi)$, $3b_1(\sigma)$, $4b_2(\pi)$[46]. The effect of yielding σ-orbital energies too close to the π-orbital ones is typical of MINDO/2; in addition, the relative values of the orbital energies is considered as further indication of the exaggeration of σ/π mixing. In the non-empirical calculation reported here, the relative orbital sequence (π,π,π,σ) is as observed; the four highest MO's, of a_1,e,b_2 symmetry, are composed almost exclusively of C(2p) atomic orbitals, and are effectively combinations of "local" π-orbitals. Thus, these four MO's can be considered to be the π-MO's of cyclooctatetraene. It is instructive to correlate these MO's with the corresponding true π-MO's of planar cyclooctatetraene, which is considered below.

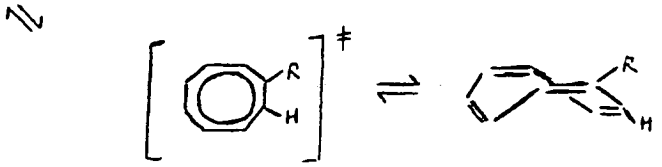
Planar Structures.

As well as having a rôle in purely theoretical considerations, planar cyclooctatetraene is of some physical significance. The dynamic physical properties of cyclooctatetraene include three fundamental processes, namely ring inversion, bond shift, and valence isomerisation. In recent years, many valence tautomers of the traditional monocyclic species, such as the annulenes, have been synthesised; these bicyclic and tricyclic species have been found to be of surprising stability and remarkable structure[47]. A valence tautomer of cyclooctatetraene is bicyclo[4.2.0]octa-2,4,7-triene (Figure 4); the latter is actually found to be of lower total energy, both experimentally (by 27 kJ mol^{-1}) and by non-empirical calculation[43]. The free energy of activation, $\Delta G^{\neq}$, for the ring closure conversion of cyclooctatetraene has been measured as 118 kJ mol^{-1} [48]; as normally prepared, this implies that 0.01% cyclooctatetraene exists as the bicyclic tautomer. The existence of such an energy barrier illustrates that absolute values of total energies of species are not the sole dominating factor in determining stabilities, and emphasises the two distinct aspects of kinetics and thermodynamics. There is some parallel with the situation of the heat of formation of cyclooctatetraene, which is a rather large positive quantity (about 300 kJ mol^{-1}). This is measured relative to the elements in their standard states; the energy of cyclooctatetraene is, however, about twenty times $|\Delta H^{o}_{f}|$ lower than the sum of the free C and H atom values (binding energy). As far as static electronic structure calculations are considered, the different valence tautomers are separate entities. In this work, only the monocyclic annulenic species are considered for the hydrocarbons C_NH_N; attention is centred on a region in the vicinity of the corresponding local minimum of the complete potential energy surface of C_NH_N species.

Unlike valence isomerisation, the other two rate processes are undergone by cyclooctatetraene / usually considered to involve ring flattening, and hence planar species as transition states. A number of experiments have been directed towards the understanding of the thermodynamics of cyclooctatetraene systems[49]; from low temperature NMR studies, a value of $\Delta G^{\neq}$ = 57.5 kJ mol^{-1} for ring inversion of cyclooctatetraene itself has been determined, and values from about 50-65 kJ mol^{-1} for simply substituted derivatives. Further studies involving substituted cyclooctatetraenes showed that there was another process occurring, namely bond shift. The two processes are illustrated in Figure 7. Ring inversion is depicted as involving a planar-alternate transition state; bond shift is shown to require a planar form of regular bond length. $\Delta G^{\neq}$ for bond shift has been determined at 72 kJ mol^{-1} for R = $-C(CH_3)_2OH$ in Figure 7, a value 10 kJ mol^{-1} greater than the measured ring inversion $\Delta G^{\neq}$ for the same compound. On the NMR time scale, ring inversion is a rapid process, whereas bond shift is rather slow. The conformational mobility of cyclooctatetraene involved in these processes exemplifies structural rearrangements of the one valence tautomer, occurring in a particular region of the energy surface. In the study mentioned above using non-empirical calculations, the most stable planar form of cyclooctatetraene was found to possess alternating bond lengths (D_{4h} symmetry), with practically the same values as found in the optimised "tub" conformation. From an energy profile for ring inversion, i.e. variation of total energy as a function of the $\hat{CCC}$ angle, the planar D_{4h} form (CCC = 135°) was found to be the energy maximum form; the flattening of the ring through the planar transition state was thus calculated to have a barrier ($\Delta H^{\neq}$) of 75 kJ $mole^{-1}$, the

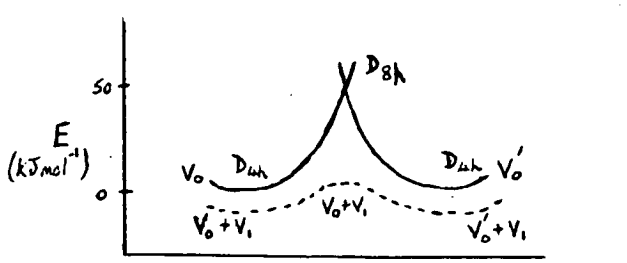


(a) COT RING INVERSION.



(b) COT BOND SHIFT

FIGURE 7.



V_G ≡ GROUND-STATE CONFIGURATION

V_1 ≡ EXCITED CONFIGURATION ($\equiv V_0$ FOR D_{8h})

FIGURE 8 : P.E. DIAGRAM FOR COT BOND SHIFT (PLANAR).

energy difference between optimised D_{2d} and D_{4h} forms. In Table 3, there are presented the total energy quantities, calculated using the standard basis, for a D_{4h} planar form with the same bond lengths as used for the non-planar conformation, thus corresponding to a "vertical" ring flattening. The energy values illustrate that the difference in total energy is much smaller in magnitude than the differences in the energy components arising in non-empirical calculations. As inversion proceeds from the non-planar form to the planar transition state, the nuclear and the bielectronic repulsions decrease markedly, but the monoelectronic attractions decrease in magnitude even more, so that the "tub" form is stabilised overall, by 37 kJ mol^{-1}. This calculated energy of activation is not very well defined, even as a fixed bond-length inversion value, as neither geometry is optimised. A few calculations performed on planar D_{4h} species indicate the variation of total energy with C-C bond length, which is the geometrical parameter susceptible to greatest variation in conjugated hydrocarbon systems, as shown in Table 4. Even in this cyclic system, the total energy is relatively insensitive to r(C-C) within about $\pm$0.05 Å of the equilibrium value, as found with 1,3-butadiene; it appears likely that the optimum r(C-C) for planar D_{4h} cyclooctatetraene is similar to that in 1,3-butadiene, with the former perhaps slightly shorter.

In Table 4, there are also presented the calculated total energies of regular planar (D_{8h}) cyclooctatetraene species. Cyclooctatetraene is a good candidate for rigorous molecular geometry optimisation, although a comprehensive treatment including various conformations and ionic species would be a very lengthy task. The results presented in Table 4 can be regarded as preliminaries in this respect, and can

perhaps give some rough ideas on the equilibrium geometries of various species. For the neutral closed-shell planar D_{8h} species, the optimum r(CC) bond length is 1.42 Å, very slightly longer than the value calculated for benzene (1.417 Å). The equilibrium regular structure corresponds to a local minimum on the planar species energy surface, and further complexities are introduced. Up to this point it has tacitly been assumed that only closed-shell, singlet species total energies have been of relevance. As is well-known from Hückel MO theory, the lowest total energy form for planar D_{8h} cyclooctatetraene is one which has a pair of degenerate MO's, each singly occupied, as the highest energy MO's, i.e. it is the "triplet" state. Actually, the restricted Hartree-Fock model does not include explicitly electron spin, so that there is no distinction made between the electron configuration in which the single electron spins are parallel (genuine triplet) and that in which they are opposed (singlet). The calculations reported in Table 4 show that the ground state of planar D_{8h} cyclooctatetraene is the triplet, which lies 500 kJ mol^{-1} below the singlet at each point, although the minimum in total energy occurs at effectively the same value of r(CC). In the alternating planar case, the degeneracy of the two MO's is removed, and the triplet species is of higher total energy than the corresponding singlet, although there is no parallel in the two variations of total energy with r(C-C), as in the D_{8h} case. As far as the bond shift energy barrier is concerned, the energy difference between planar D_{4h} and D_{8h} singlet species is required. From the present lowest energy calculations for each, a barrier of 450 kJ mol^{-1} is deduced; even if both forms were subject to complete geometry optimisation, it is very unlikely that there would be a significant change in the calculated barrier height. The original

experimental studies, referred to above, using variable temperature nmr, lead to a value of about 10 kJ mol^{-1} for the energy difference between alternating and symmetrical planar forms (= difference in free energies of activation of ring inversion and bond shift). More recent studies on the structure and bond shift kinetics of cyclooctatetraene using the technique of nematic phase nmr[50] have led to a value of 46 kJ mol^{-1} for the above energy difference. In this latter technique, ring inversion is not detectable, so that the effects of the two dynamic processes can be separated more easily than in the former case, where bond shift is not isolable. The calculated barrier is certain to be too high, because the situation under consideration is one in which the inadequacy of the single-determinant H-F approach is exposed. The singlet D_{8h} structure ground state is particularly poorly represented, as there are two configurations of identical energy resulting from the degeneracy of the pair of MO's which have to accommodate only one electron pair (the nonbonding π-orbitals in Hückel terms). At this particular structure, where the symmetry of the molecular geometry is higher than that of adjacent points on the energy surface, there are two closed-shell electronic configurations of identical energy, so that the ground state is particularly poorly represented by a single configuration wavefunction, and a combination of the two equally contributing configurations, at least, is required for a satisfactory representation. At alternating planar structures (D_{4h}), configuration interaction is less important, with one configuration being dominant in the ground state representation. Figure 8 illustrates the situation; the calculated value of the energy difference between alternating and regular planar forms is only an upper limit. On the basis of semi-empirical calculations (MINDO),

which involve single determinantal wavefunctions, the energy difference between completely optimised D_{4h} and D_{8h} planar forms was calculated to be about 60 kJ mol^{-1}, and it was concluded that this rather large value indicated that the transition state for bond shift is not the symmetrical form;[46] a symmetrical non-planar structure was proposed, such as the "crown" conformer considered earlier as a possibility for cyclooctatetraene itself, but no other evidence, experimental or theoretical, has supported this. Thus, although the triplet is still expected to be the lowest energy D_{8h} planar form, when improvement on the single configuration description is made, the closed-shell singlet is the relevant form for bond shift considerations; in this particular process, where total energies are of prime importance, the single configuration approach is not adequate for the regular planar structure, and the bond shifting is an example of a "reaction" poorly represented by H-F calculations.

Molecular Orbitals and Eigenvalues.

Since it became readily available by a straightforward synthetic method, from classical chemical studies on cyclooctatetraene it became progressively evident that this "8π-electron" system lacks significant resonance, or stabilisation, energy, which was estimated at between 10 and 20 kJ mol^{-1}. More recently, there has been a renaissance in cyclooctatetraene chemistry[51], and the view has been expressed that "the blatant non-aromaticity of cyclooctatetraene caused by π-electron instability, and favouring a D_{2d} tub conformation, coupled with obvious high levels of unsaturation, provides a molecule having unequalled facility for structural rearrangement", leading to an example of cyclic polyolefin chemistry. In a more theoretical vein, consideration of the electronic structure of planar cyclooctatetraene has led to contrasting viewpoints. Complementary to the concept of aromaticity as a particular aspect of bonding, the idea of antiaromaticity has been proposed, as an aspect of antibonding[52]. Thus, as far as cyclooctatetraene is concerned, it is possible that the barrier for ring inversion of the puckered neutral molecule involves the destabilising interaction of four double bonds in the planar conjugated system; in general, there is no strong evidence for conjugative destabilisation in 4n-annulene systems relative to "normal" compounds, in contrast to the enhanced stability of (4n + 2)-systems. With a definition of aromaticity based on thermodynamic stability, it is not possible to extract the effects of π-electron delocalisation. An alternative view to the above considers that π-electron delocalisation stabilises planar cyclooctatetraene, but that the tub form is more stable as a result of contributions by the σ-electrons to the total energy; cyclooctatetraene cannot exist in an unstrained planar form so that any gain in resonance energy due to π-interactions therein would be offset by the energy required to flatten out the ring.

On this basis, it could be concluded that cyclooctatetraene would be aromatic were it possible to avoid the complicating factor of ring strain. Semiempirical calculations reject this possibility; these, on considering π-electrons only, imply that planar cyclooctatetraene is antiaromatic, with a negative resonance energy and alternating bond lengths. More refined calculations, including all valence-electrons substantiate this[46].

In Table 3 are presented the calculated total energy quantities for the best regular planar (closed-shell) form of cyclooctatetraene; this is not completely optimal, but it is likely that even at the calculated equilibrium structure the resonance energy will still be significantly negative. This can be taken to mean that regular planar cyclooctatetraene is antiaromatic, although it is desirable to consider the nature of the individual molecular orbitals, particularly the π-type ones. The variation of orbital energies of corresponding MO's in the tub and planar conformations is shown in Figure 9. One obvious feature is that in the alternating planar form, three σ-MO's lie between the three highest and the lowest π-MO's, illustrating the general observation that the π-MO's as a group in conjugated systems do not lie at significantly higher energy than the σ-core. Considering the π-orbitals first, the forms of these are illustrated in Figure[10] for the planar conformations. All four π-orbitals in the alternating form are of lower energy than the corresponding ones in the regular form, with the overall stabilisation about 110 kJ mol^{-1}, which is approximately the difference in total energy of the two forms. The destabilisation (large negative resonance energy) of the regular planar closed-shell singlet conformation, without considering at this point the effect of configuration interaction in representing the actual ground-state of the molecular species, might be thought on this basis to be a result of π-electron destabilisation. The differences in orbital stability between corresponding pairs in

FIGURE 9 CYCLO-OCTATETRAENE

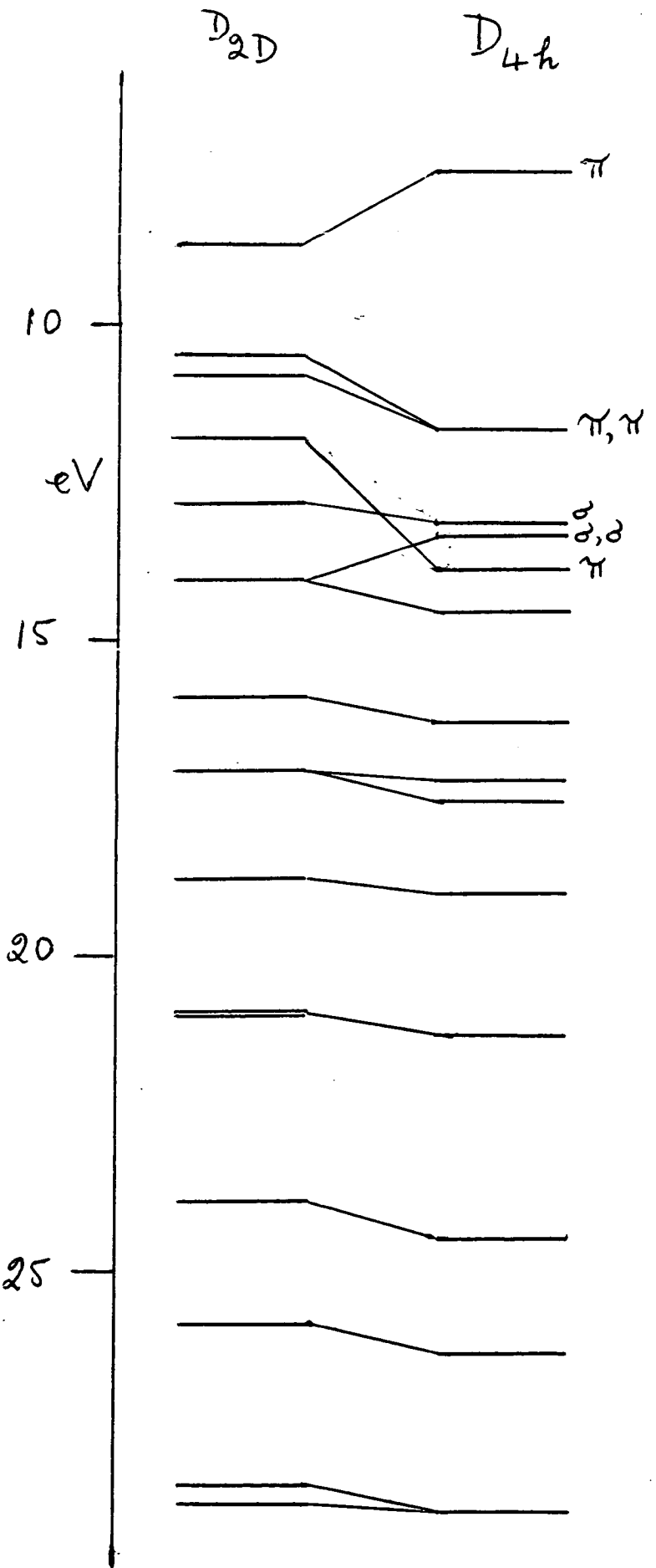
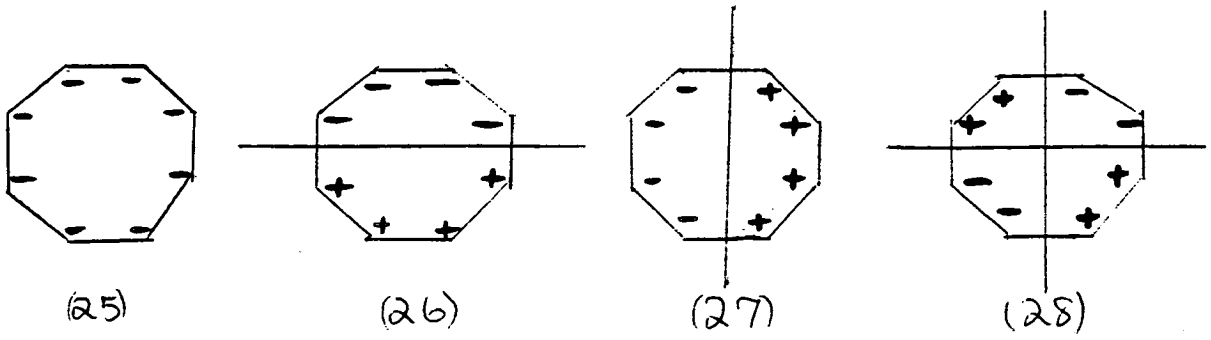
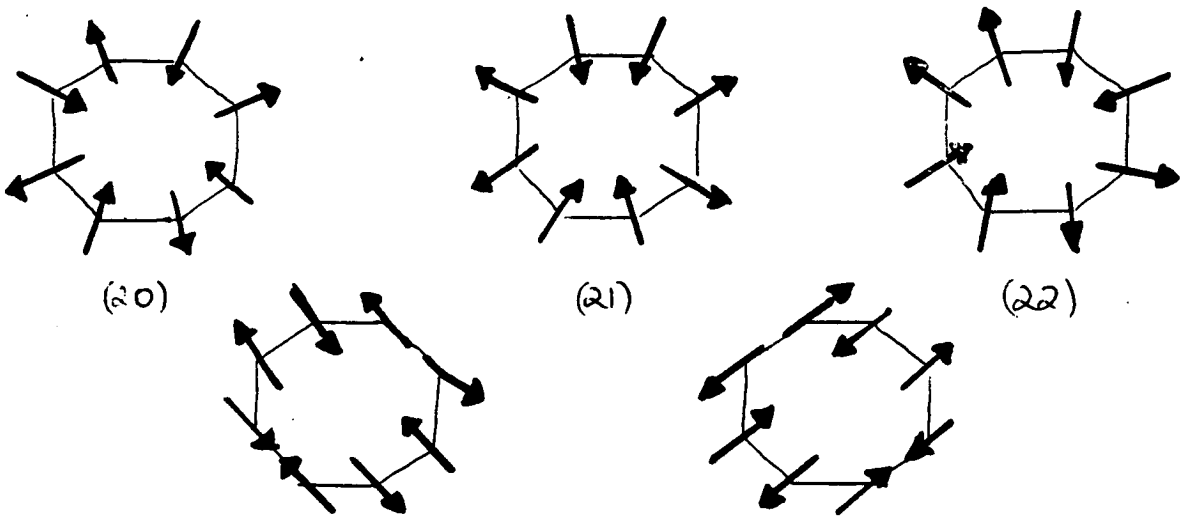


FIGURE 10



(25) (26) (27) (28)

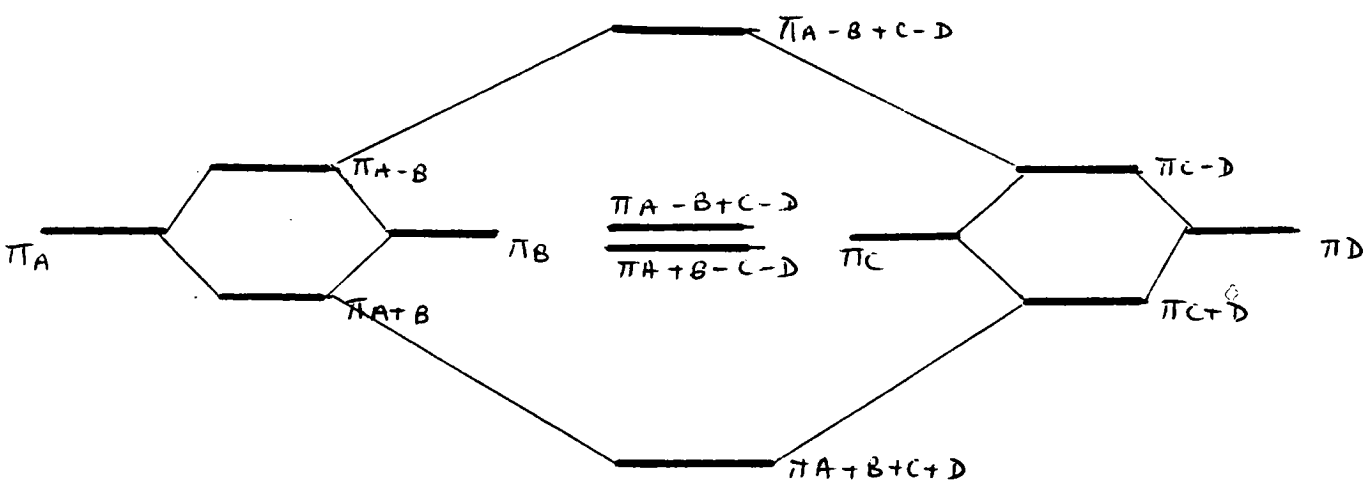
a) FORMS OF PLANAR COT π – MO'S



(20) (21) (22)

(23) (24)

b) FORMS OF PLANAR COT σ- M.O'S



c) ORBITAL INTERACTION DIAGRAM (NOT TO SCALE)

the two conformations, and the relative stability within each group of four, can be rationalised by simple qualitative considerations of the effects of atomic orbital (carbon $2p_z$) overlap, leading to bonding and antibonding contributions to the MO overall. In this way, the difference between the alternating and regular forms is basically a result of the bond length, or more generally interatomic distance, effect on the AO overlaps. Thus, the largest difference in stability is between the highest occupied π-orbitals of the two conformers; the substantial destabilisation of the regular π-orbital can be explained by noting the form of the MO, in which the bonding contributions (favourable p_z overlap) are enhanced in the alternating form as a result of the shorter interatomic distance (1.34 Å against 1.42 Å in the regular form), and the antibonding contributions (unfavourable p_z overlap) are less in the alternating form where the relevant interatomic distance is 1.50 Å (compared to 1.42 Å). As is usual, nearest-neighbour interactions are considered to be dominant. The other three π-orbitals (including a degenerate pair) are mainly bonding around the ring, and the bonding and antibonding effects are more balanced between the two conformers, so that there is little difference in stabilities of corresponding pairs. Examination of the orbital energies of the π-electrons in the computations on the other planar species reveals a trend consistent with the above qualitative considerations.

The four highest-energy MO's of the tub conformation can be regarded as π-type orbitals; the non-planarity of the molecular framework removes the distinction between σ- and π-electrons, but the forms of the MO's shows them to be of "local" π-character (Figure 10) and the correspondence with the actual π-orbitals of the planar species can be easily seen.

The calculated energies of these orbitals shows that the highest-energy one is substantially stabilised in the non-planar form (by about 130 kJ mol^{-1}), but that the other three are destabilised, especially the lowest-energy totally symmetric one, which is about 200 kJ mol^{-1} more stable in the planar alternating form. Again, the above considerations are consistent with the variation of orbital energy; in going from planar alternating to tub form, in this case, bond lengths are the same so that the nearest-neighbour interaction of AO's differs as a result of the different orientations of the out-of-plane (locally) p-orbitals. Thus, the twisting of the butadiene fragments by folding the molecule from the planar conformation renders much less favourable the overlap of AO's right round the ring in the totally symmetric π-orbital, which is destabilised compared to the planar case. On the other hand, the highest-energy orbital is stabilised in the non-planar conformer as a result of the decrease of the unfavourable overlap along the formal single bonds, while the bonding contributions between doubly-bonded centres is much the same in the two forms. The energy of the e-pair of π-type MO's is almost identical with that of the π-orbital of ethylene, and the pair of non-interacting π-orbitals of twisted 90^{o}-butadiene (Chapter 4). In the non-planar conformation, each of this pair of MO's is basically composed of a pair of local π-orbitals, diametrically opposed and hence there is little interaction between them, giving a similar type of effect to that obtained in the twisted butadiene case. In the e-pair of orbitals in the planar conformation, there is greater favourable interaction between adjacent π-orbitals (Figure 10) so that the e-pair is at somewhat lower energy. The relative orbital energies of the π-type MO's in the planar and non-planar conformers can also be rationised on the basis of simple orbital interaction diagrams (Chapter 4), as illustrated in Figure 10

for pairs of π-orbitals; "splitting" of combinations of orbitals is larger in the planar case as expected qualitatively.

On the basis of SCF orbital energies of π-type MO's, it probably would not be concluded that π-electron instability disfavours a planar eight-membered ring. However, there are difficulties in comparing directly non-empirical and semi-empirical results. In a Hückel, π-electron only, treatment, orbital energies are involved, and the total molecular energy is evaluated from the simple sum:

$$E^T = \sum_i n_i \varepsilon_i \qquad -(1)$$

where n_i is the orbital occupation number of orbital i of energy ε_i. In the Hückel one-electron model of annulene species, E^T thus evaluated as a sum of π-orbital energies is the basis of the distinction made between "4n" and "4n + 2" systems. This approach leads to the evaluation of a "delocalisation" energy by subtracting from E^T the total energy of the relevant number of ethylenes. Thus, for cyclooctatetraene, $E^T = 2.2\beta + 4.\sqrt{2}\,\beta + 2.0\beta = 9.656\beta$, and the delocalisation energy = $9.656\beta - 8\beta = 1.656\beta$; β is one of the parameters of the Hückel approach,[52] introduced as a measure of the interaction between two carbon $2p_z$ orbitals. Even for a "4n" system, the delocalisation energy is a large quantity, as β is usually estimated at about 80 kJ mol^{-1}; however, measured per π-electron, this energy is lower for "4n" systems - e.g. for benzene, delocalisation energy is 2β. Thus, in Hückel-type approaches, the lack of aromaticity of "4n" systems is really a relative effect, and it cannot be concluded that there is destabilisation (antiaromaticity). In fact, the delocalisation energy of non-planar cyclooctatetraene tends to zero as the twisting of the double bonds proceeds to give orthogonal π-systems. The results of more refined calculations, both semi-empirical and non-empirical as shown here, indicate that a regular planar cyclooctatetraene

configuration is destabilised (negative resonance energy using basic reference molecules), and that the lowest energy form has a stabilisation energy which is very small. This has been taken as substantiation of Hückel's rule. However, it is not possible to deduce a π-electron effect from the non-empirical calculations. The use of equation (1), with calculated SCF orbital energies, with benzene (chapter 4) and cyclooctatetraene does not parallel the simple Hückel case. Similarly, if calculated "one-electron orbital" energies, eigenvalues of the one-electron Hamiltonian (neglecting electron repulsion as in Hückel treatment) are used, no deduction on overall stability can be made. Calculated one-electron orbital energies for cyclooctatetraene are shown in Table 5 ; the SCF orbital energies for the π-electrons actually bear more resemblance to those of the Hückel approach than do the one-electron orbital energies.

The sigma electrons are an integral part of the all-electron calculations reported here. In Figure 9 there is also shown the variation of the calculated orbital energies of the valence σ-electrons as the molecular conformation changes. The results can be rationalised again in the same qualitative way as above with the π-type electrons; Figure 10 shows the forms of some of the σ-orbitals to illustrate this. The variation of orbital energies is such that some of the σ-orbitals lie in amongst the π-orbitals in the planar forms; this is mainly the result of the substantial stabilisation of the totally symmetric π-orbital in the planar species. One-electron orbital energies of the σ-orbitals are presented in Table 5 ; some of the values in a given conformation are substantially less negative than those of the π-orbitals. The sum of the one-electron orbital energies is given for each conformation in Tables[1,3] ; the sum of the one-electron and SCF orbital energies for a particular orbital is the electron energy, and the sum of the latter

over occupied orbitals is the total electronic energy. Thus, in any molecular conformation, the π-orbitals contribute significantly more to the total electronic energy than several of the σ-orbitals. The σ-orbital whose form is given in Figure 10 has a particularly low electron energy, although there are several orbitals above it in orbital energy sequence. This orbital is basically a C-H bonding orbital, but over the whole ring it shows some antibonding character; there is a similar type of orbital in butadiene and benzene showing an exceptionally small (in magnitude) one-electron orbital energy (Chapter 4). In considering the stability of a given MO and the variation of this with molecular conformation, it is instructive to examine the SCF orbital energy, as above. However, in deducing overall molecular stability, the simple additive relation (1), which has been invoked in the context of ab initio SCF calculations in discussing geometric effects in terms of orbital behaviour, is at variance with the rigorous relation:

$$E^T = \sum_i n_i \varepsilon_i + V_{nn} - V_{ee} \qquad -(2)$$

where V_{nn} is the total nuclear repulsion energy and V_{ee} is the total electronic repulsion energy; this relation arises from the fact that in $\sum_i n_i \varepsilon_i$ the electronic repulsions are counted twice and the nuclear repulsions are omitted. The use of (1) has been rationalised by suggesting that V_{nn} and V_{ee} largely cancel each other, i.e. that $V_{nn} - V_{ee} \approx 0$. However, in any of the cyclooctatetraene conformations, $V_{nn} - V_{ee} \approx 100$ a.u. ! (V_{ee} can be evaluated as the difference between the total one-electron energy and the total electronic energy given in the Tables). Recently, it has been shown that for molecules at, or near, their equilibrium conformations, the relation $V_{nn} - V_{ee} = \frac{1}{3} E^T$ is

approximately fulfilled, and that it can then be deduced that $E^T = \frac{3}{2} \sum_i n_i \varepsilon_i$ is approximately valid.[53] However, this approach is still unsatisfactory in the cyclooctatetraene case. Using the sum of π-orbital energies only yields the following ordering in total energy: alternating planar < regular planar < non-planar. If all the orbital energies are included, this changes to regular planar < alternating planar < non-planar. The calculated values reported here do not refer to the equilibrium geometries of the various conformations, but the results are likely to remain unaltered. It has been concluded that any quantity defined as a multiple of $\sum_i n_i \varepsilon_i$ cannot be used for the determination of molecular geometry, unless the orbital energies ε_i are defined in a different context from the Hartree-Fock orbital energies.[54] The one-electron orbital energies, which might appear to be in keeping with Hückel considerations, can be summed for the cyclooctatetraene conformations; these totals give the same ordering as the actual total energy values do. However, the total one-electron energy, or total electronic energy, are not satisfactory either because they are essentially inseparable from the total nuclear repulsion energy; thus, in non-planar cyclooctatetraene the nuclear repulsion energy is greater than in any planar species, but the one-electron orbital and electronic energies are more negative in the non-planar case (actually a complementary effect), so that the total energy differences involved are small compared to the contributing terms, and the total energy ordering does not necessarily follow the electronic energy one. In cyclooctatetraene, the SCF orbital energies can be used to deduce variation of MO stability with conformation, whereas the one-electron orbital energies (and electron energies which the latter dominate) cannot, as they all tend to vary in the same way; in going from planar to non-planar species, each of the one-electron orbital energies is decreased by roughly the same amount.

Thus, individual orbital behaviour on change of conformation can be rationalised, but the overall effect is difficult to explain. All the valence orbitals contribute to differing extents in determining the lowest-energy conformation, and the final result is determined by factors which are finely balanced. It has been stated that planar cyclooctatetraene is disfavoured by π-electron instability and not steric strain effects.[52] On the basis of non-empirical calculations, conformational effects do not appear to be due simply to the π-type orbitals. Alternating planar cyclooctatetraene is more stable than the regular configuration by a larger amount than simple Hückel theory predicts (about 10 kJ mol^{-1}). The calculated total energy difference between regular and alternating benzene structures is much smaller (about 25 kJ mol^{-1}) than the corresponding cyclooctatetraene value. A general, qualitative examination of the forms of all of the valence orbitals on the lines mentioned above shows that in the benzene case, within a given orbital, the bond length effect leads to an increase of both bonding and antibonding character (in different regions of the molecule); the orbital energies of corresponding orbitals in the D_{3h} and D_{6h} forms tend to be quite close. In the cyclooctatetraene case, there tends to be larger net stabilisation and destabilisation of MO's with variation in bond lengths (and also "twisting" in non-planar form). The calculation on the experimental "tub" form yielded orbital energies which showed reasonable correlation with experimental I.P.'s. The results of this calculation and others at the experimental structure yield a value for the ratio

$$\frac{\Delta_1}{\Delta_2} = \frac{\varepsilon(5a_1(\pi))-\varepsilon(7e(\pi))}{\varepsilon(7e(\pi))-\varepsilon(4b_2(\pi))}$$

which is somewhat larger than that observed (= 1). The value here is about 1.4. This effect has been attributed to an overemphasis of σ/π mixing in the calculations. It seems likely that the calculated equilibrium structure would differ from that used in the calculation mainly in a lengthening of r(C-C); on the basis of previous considerations, a lengthening of r(C-C) would be expected to lead to stabilisation of orbital ($5a_1$) and destabilisation of orbitals (7e) and ($4b_2$). This effect is evident in the planar species, although the energy shifts are rather small. A more marked shift is obtained by varying the dihedral angle, α (Figure 4). An increase in α, leading to an increase in non-planarity and twisting of neighbouring π-systems is expected to have the same effect on the π-type orbitals. Thus, another calculation on a "tub" conformation was performed, with r(C=C) = 1.34 Å, r(C-C) = 1.52 Å, $\alpha = 50^{\circ}$; the results are shown in Tables[1,2] and Figure 6. The ratio $\frac{\Delta_1}{\Delta_2}$ is now about 0.7. The total energy of this second non-planar conformation is about 5 kJ mol^{-1} lower than the previous one. It seems likely that both are close to the equilibrium structure. Including the variation of orbital energies and comparing with experiment, it seems that the calculated equilibrium value of α may be slightly greater than the experimental one. The second calculation is of lower total energy, even although most of the MO's are destabilised relative to the first one. The orbital energy of the highest occupied MO is particularly sensitive to geometry variation. As regards the fitting of the ratio $\frac{\Delta_1}{\Delta_2}$ to the observed value, the large calculated value does not seem to arise through an overemphasis of σ/π interaction as the structure is nearer to planarity than the equilibrium one. In the calculation performed at the experimental geometry, on comparing with the P.E. spectrum, the orbital energy of the highest occupied MO is the worst-fitting one, and the large value calculated

for the ratio $\frac{\Delta_1}{\Delta_2}$ is mainly due to the underestimated binding energy of this orbital. As mentioned above, the four highest energy orbitals are linear combinations of local π-orbitals, and correspond quite closely with the true π-orbitals of the planar form; in a sense, there is not much σ/π mixing, if σ contributions are considered as those corresponding to the σ-system of the planar form (in-plane bonding). The variation of the π-type orbital energies as α varies, and hence $\frac{\Delta_1}{\Delta_2}$, can be explained by considering only π-conjugative effects.

In the non-empirical Hartree-Fock approach, in contrast to the simple Hückel approach, all electrons and all MO's of the molecule are considered, and nuclear and electronic repulsion effects are included. In the problem of the electronic structure of cyclooctatetraene, and conformational effects, on going from a planar to non-planar form, the one-electron energy of all the orbitals is lowered, but electron repulsion effects (which again are overall effects of the MO framework) lead to orbital energies which vary in different ways. The overall electronic energy changes cannot be separated from the associated nuclear repulsion ones, yielding the overall total energy in a finely balanced situation; "strain" effects cannot be singled out.

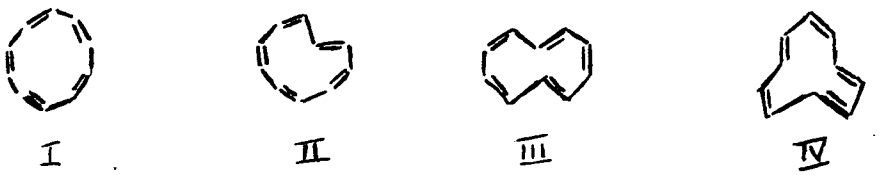


FIGURE 11 - 10-ANNULENE STRUCTURES.

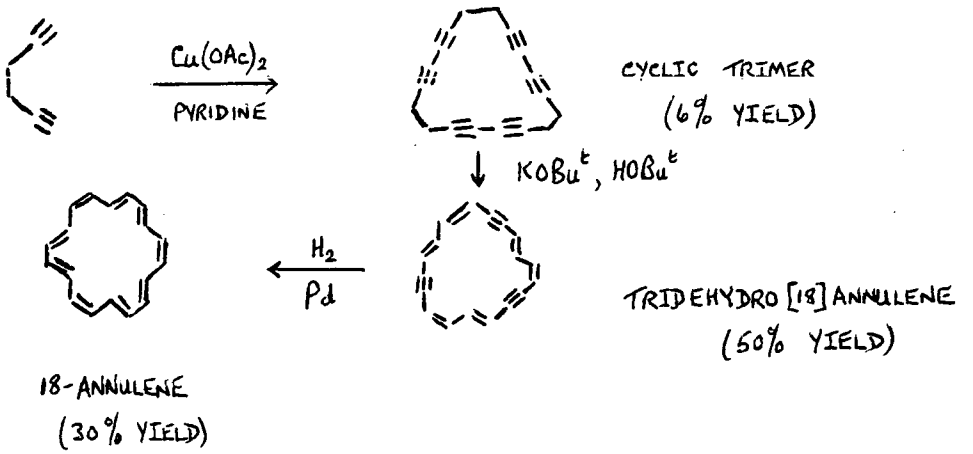


FIGURE 12 - SYNTHESIS OF 18-ANNULENE BY ACETYLENE COUPLING.

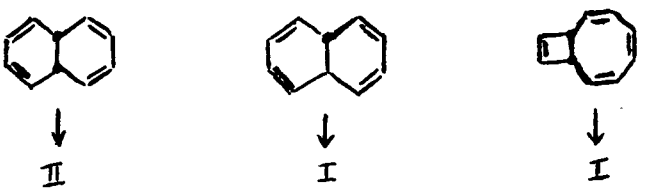


FIGURE 13 - KEY VB ISOMERS OF 10-ANNULENE.

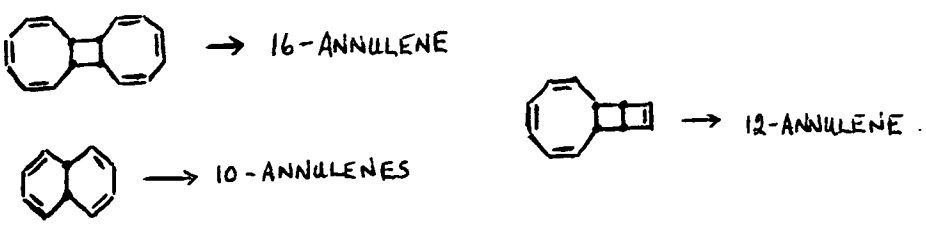


FIGURE 14 - PHOTOLYTIC RING OPENING SYNTHESES.

(b) 10-Annulene

Soon after the structure of benzene was proposed by Kekulé in 1865, chemists began to suspect that this C_6H_6 hydrocarbon might not be unique in its properties, but rather the first-discovered member in a series of cyclic conjugated polyenes similarly endowed with stability. As mentioned in (a) above, the synthesis of the next higher vinylogue, cyclooctatetraene, achieved in 1911, led to the test of this hypothesis and it was found that this homologue exhibited chemical reactivity like that expected of a linear polyene. In 1932, Hückel, employing a quantum-mechanical approach, developed a theoretical treatment which downgraded the so-called "aromatic sextet" from its exalted position of uniqueness, and repositioned it as one member of a series of structures expected to be stable by virtue of a closed shell of (4n+2) π-electrons. Hückel's theory (or, since its gain in currency, Hückel's Rule) was stated to be valid only for coplanar monocyclic polyolefins fully conjugated. The extent to which this rule has succeeded in permitting predictions of aromaticity or the absence thereof is now general knowledge; until relatively recently, the validity of the rule was only tested a little more widely than was possible when it was first formulated, by referring to the small ring species of Figure 2. As originally constituted, the rule applied only to that class of compounds now known as Annulenes, i.e. completely conjugated monocarbocyclic polyenes. For $p = 3$ to $p = 8$ in the C_pH_p series, NMR data on several species have shown that all the H nuclei are external to the C ring; the single line spectra are associated with structures represented by polygons. A number of larger ring annulenes (with p up to about 30) have been synthesised in recent years. These annulenes "proper" can be represented by various configurational formulae. In addition to polygons, more physically reasonable structures can be drawn; in Figure 11 are shown some examples

of 10-annulene structures. The most simple vinylogue of benzene for which Hückel predicted aromatic stability, cyclodecapentaene or 10-annulene, resisted the varied synthetic attempts of organic chemists for over 35 years and only recently and somewhat reluctantly yielded significant data relating to its nature.[55]

Cyclodecapentaene is the first of the "large" annulenes which have been studied by non-empirical calculations in this work. The smaller species of Figure 2, at least for p = 3 to p = 6, have been studied extensively, both experimentally and theoretically, with rather refined techniques. In this chapter, there are considered below some of the larger members of the annulene series, which are fundamental to organic chemistry, and some species formally derived from them; these rather elusive compounds are becoming more amenable to experimental techniques. The results of experiments in annulene chemistry in the last 15 years have shown that 10-annulene is a very different species from the other annulenes. Many investigators have attempted syntheses of cyclodecapentaenes. It will never be known how much of the published investigation of the chemistry of medium ring alkenes was originally directed towards the synthesis of cyclodecapentaene and, while failing in this respect, succeeded in leading to other worthwhile avenues of research. The first general method devised for the synthesis of the annulenes was developed by Sondheimer and co-workers; this involves the oxidative coupling of a suitable terminal diacetylene to a macrocyclic polyacetylene as the key step. The cyclic compound is then transformed to a dehydroannulene (i.e. a completely conjugated monocarbocyclic polyenyne), usually by prototropic rearrangement. Finally, partial catalytic hydrogenation of the triple to double bonds leads to the annulene. The synthesis of 18-annulene, the first higher annulene prepared, illustrates the method, as outlined in Figure 12[56]; the procedure is now well-documented[57]

The pure crystalline annulenes synthesised by this method now include all the even-membered rings from p = 14 to p = 24. It is also notable that the above route leads to dehydroannulenes of all ring sizes from p = 12 to p = 30 inclusive; as considered below, these species are also of theoretical interest, since criteria for aromaticity considered above for the annulenes apply also to the dehydro-compounds as far as the out-of-plane π-electrons are concerned. However, the C_{10} series has resisted attempts modelled after those syntheses so rewarding in the larger ring systems, based on acetylene coupling in a ring-closure approach. Of the various synthetic schemes devised for cyclodecapentaene, the only one whereby success has been claimed is the "valence bond isomer" or "valency tautomer" approach, a general synthetic plan in which synthesis is aimed not at cyclodecapentaene itself but at a well-chosen valence tautomer[58]. Synthesis of the polycyclic isomer, with the reasonable assumption that there exist means to overcome the energy barrier to interconversion, would be equivalent to direct construction of the 10-membered cycle. Naturally, in regard to thermal processes, the outcome of the synthesis is determined (assuming a negotiable energy barrier) by the equilibrium constant for interconversion of the two isomers. In the event of circumstances unfavourable for production of the cyclodecapentaene, rearrangement to the desired species might be promoted by means of photolysis or through the agency of suitable transition metal complexing agents applied under isomerisation conditions. The key valence bond isomers of cyclodecapentaene are shown in Figure 13, where there are illustrated allowed transformations (electrocyclic reactions), on the basis of the Woodward-Hoffmann selection rules[59]. Low temperature photolysis of trans-9,10-dihydronaphthalene ultimately led to the first claim for the physical detection and trapping of

10-annulene[60], when the evidence was rather circumstantial and it was stated that elucidation of the detailed nature of 10-annulene with regard to configuration, conformation, and electronic character would await more sophisticated experimental probes. A more convincing synthesis using cis-9,10-dihydronaphthalene as precursor has been described[61]. Experiments carried out provided strong evidence for the presence of 10-annulenes as transient intermediates, giving some information about their properties; isolation of the compounds necessary for definite proof was successfully performed, yielding pure crystalline samples at $-70^{\circ}C$. The $(CH)_{10}$ valence-bond isomer system has been extensively studied, revealing a complex intertwining of rearrangement processes[62]. As a generalisation of the above type of synthesis, a second route to the annulenes has been developed by the groups of Schröder and Oth, Masamune, and van Tamelen, involving the photolytic ring opening of a polycyclic valence isomer. 10-, 12-, 16-annulenes have been prepared by this method, as outlined in Figure 14[63]. Most of the higher annulenes and dehydroannulenes are highly coloured (red or brown) crystalline substances, relatively stable at room temperature in solution, but less stable in the solid state, with 10-annulene one obvious exception.

Apart from synthetic difficulties, there are problems posed by the cyclodecapentaene system in the consideration as a nonbenzenoid aromatic species. In fact, the failure of preparative attempts focused attention on factors other than the closed shell of 10 π-electrons. Approximate coplanarity, necessary for proper orbital overlap to ensure maximal electron delocalisation, gives rise to bond angle strain and/or nonbonded atomic repulsion in the interior of the carbocyclic structure, in a classical model. It was never certain whether the amount of energy involved in such factors was large enough to counterbalance the energy of aromatic stabilisation. That it might be, however, was indicated

FIGURE 15 : CIS-TRANS ISOMERS OF 10-ANNULENE.

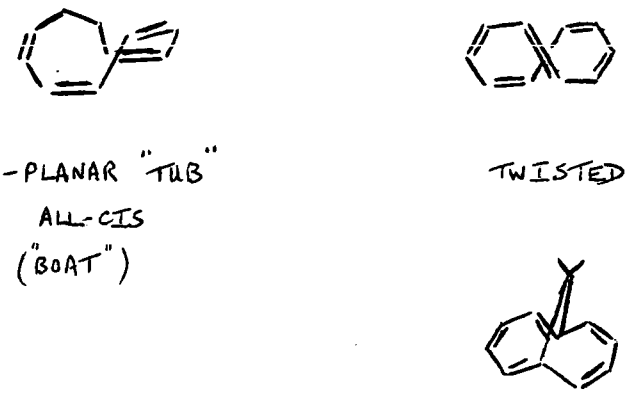


FIGURE 16 : NON-PLANAR SPECIES.

ALL-CIS (D_{10h}) $r(CC) = 1.398\,Å$, $r(CH) = 1.090\,Å$
$\hat{CCC} = 144°$, $\hat{CCH} = 108°$.

ALL-CIS (D_{5h}) $r(C-C) = 1.50\,Å$, $r(C=C) = 1.34\,Å$, $r(CH) = 1.10\,Å$
$\hat{CCC} = 144°$, $\hat{CCH} = 108°$.

TRANS-ISOMER (C_{2v}) (TCCCC) $r(CC) = 1.398\,Å$, $r(CH) = 1.090\,Å$.

CTTCT-ISOMER (C_{2v}) $r(CC) = 1.398\,Å$, $r(CH) = 1.090\,Å$

FIGURE 17 : 10-ANNULENE GEOMETRICAL DATA.

by the general tendency of medium-ring polyenes to undergo facile skeletal rearrangements and transannular reactions to relieve this strain. Furthermore, cyclodecapentaene, hypothetically, can exist in various cis-trans modifications, as exemplified by the isomers of Figure 15. The impact of this complicating factor becomes evident on trying to predict the nature or properties of cyclodecapentaene, or to interpret physical data, and it engenders appreciation for the special problems of synthesis and structure of this molecule. Qualitative estimation of the significance of steric destabilisation and angle strain has not led to definite conclusions. The cis, trans, cis, cis, trans-isomer can be considered to be benzene-like (free of angle strain); in this structure, there is strong interaction of the internal 1,6-hydrogen atoms and this demands distortion from planarity. There are groups of compounds, such as bulkily-substituted benzenes or paracyclophane species[64], which indicate that considerable aromatic stability can be retained even though nuclear atoms are forced out of plane, although the actual magnitude of the distortion which an aromatic nucleus can accommodate is unknown. The steric strain of cyclodecapentaene might be tolerated if aromatic stabilisation is sufficiently great, as there are known cases where a significant degree of distortion can be tolerated. In considering the problem of angle-strain destabilisation, the all-cis geometric isomer is the most important. The severe nonbonded interaction of H atoms interior to the ring is avoided, but significant in-plane distortion of bond angles is generated. Quantitative calculation of the energy expense involved in this distortion is precluded because of the size of the perturbation (strainless 120° to 144°); however, the all-cis isomer can be placed in context by comparing with the two species, cyclooctatetraene dianion and cyclononatetraenyl anion, both of which are 10 π-electronic species exhibiting physical properties characteristic

of aromatic, all-cis, planar compounds. However, the change in the number of angles being distorted and the non-linearity of the function describing the energy expense as a function of angle, mean it cannot definitely be concluded whether the strain energy of all-cis cyclodecapentaene is surmountable or not.

Since 10-annulene is so unstable, there is very little experimental information available on its physical and chemical properties. In particular, there are only conjectures about its geometrical structure based on the NMR and ultraviolet spectra and thermochemical data. Before presenting the results of non-empirical calculations on a few planar structures, the available background information is summarised.

The first successful experiments with regard to the synthesis of 10-annulene were the photolyses of the compounds of Figure 13 , where the conclusion was that each reaction proceeded via a common intermediate, which was most likely to be a 10-annulene or a mixture of 10-annulenes. After much further experimentation, it was found that the photolysate of cis-9,10-dihydronaphthalene near $-60^{\circ}C$ showed NMR absorptions at $\tau = 4.16$ (broad temperature dependent singlet) and $\tau = 4.34$ (sharp singlet) which were not due to any known $(CH)_{10}$ hydrocarbon, but which could be attributed to 10-annulenes. The data were most readily explained by attributing the $\tau = 4.16$ peak to trans-10-annulene and the 4.34 peak to cis-10-annulene (Figure 15). Although the first experiments provided strong evidence for the presence of 10-annulenes, isolation of the compounds was obviously necessary for definite proof. Ultimately, synthetically useful concentrations of the annulenes (10% of cis and 20% of trans) were obtained, and the two components were isolated in a pure state, crystalline at $-70^{\circ}C$. Both these compounds show NMR absorption only in the olefinic region. With the pure materials in hand, a firmer assignment of the geometry is possible on the basis of the NMR spectra, in conjunction with a conformational analysis. The

NMR spectrum of the τ = 4.16 component is temperature-dependent, showing a sharp singlet at -40°C but two peaks at -100°C, with no further change down to -150°C. The NMR (^{1}H and ^{13}C) spectra of the τ = 4.34 component are invariant over the whole temperature range. Only three 10-annulene isomers are usually considered as being geometrically possible, namely the all-cis-, trans-$(cis)^{4}$-, and trans-cis-trans-$(cis)^{2}$-isomers (Figure 15). Molecular models show that the molecules increase in rigidity in that order, so that equivalence of the protons should be most easily achieved by the all-cis-isomer. Specifically, on a qualitative molecular model basis, the lowest energy conformation of the all-cis-isomer appears to be the one having a tub-shaped portion attached to a near planar diene fragment (Figure 16). It can be shown that"pseudorotation" of this non-planar conformation leads to five equivalent (degenerate) forms, and provides a mechanism for making all the C atoms equivalent. This process requires only rotation about single bonds, and need not involve either single-double bond alternation or ring-flipping (with further expansion of the bond angles), the latter two processes requiring a higher activation energy than pseudorotation. The observation of a singlet in the NMR is consistent with the low-barrier pseudorotation equivalencing process. In addition, the singlet is consistent with a planar structure for the all-cis-isomer, although this is most unlikely in view of the low intensity UV absorption.

For the trans-$(cis)^{4}$-isomer the most stable conformation may be that shown in Figure 16 . A mechanism rationalising the observed equivalence of all nuclei (C or H) at -40°C involves bond alternation and rotation, causing the trans-double-bond to "migrate" round the ring, and this also explains the temperature variance of the spectrum, which indicates slower conformational inversion with the more rigid trans-$(cis)^{4}$-isomer than in the all-cis case. The interconversion of non-planar equivalent

conformers by this process requires a change of conformation and a bond alternation in the near planar conformation, of the trans-$(cis)^4$-isomer, so that it is quite reasonable that the nuclei become equivalent at a much higher temperature than for all cis-10-annulene.

In contrast to the above, there is no conceivable low energy process which makes all the nuclei of trans-cis-trans-$(cis)^2$-10-annulene equivalent. The lowest energy conformation is probably generated by conrotation or disrotation of the planar form to relieve the non-bonded repulsion between the internal hydrogen atoms. Bond alternation is likely to have a low activation energy, but this process still leaves three different kinds of proton, and further migration is most unlikely. The structures and stereochemistry of the two isolable 10-annulene isomers are thus reasonably secure.

On the question of the aromaticity of the 10-annulenes, there is a little more experimental information. The UV spectrum of the trans-isomer is similar to that of the bridged species, 1,6-methano[10]annulene (Figure 15). The approximate similarity might suggest that trans-10-annulene has electronic energy levels characteristic of cyclic electron-delocalised systems (the bridged species does possess an essentially planar 10 π-electron system[65]), so that the structure could be considered to be one with an approximately equal angle of twist between the π-orbitals on each adjacent C atom. In such a structure, a small deshielding diamagnetic ring current might be compensated by the shielding effect of the π-electrons on the opposite side of the ring, which might lead to the observed NMR absorption at $\tau = 4.16$. However, the D_2 symmetry of the structure implies that the low temperature ^{13}C NMR spectrum should show only three kinds of C atom; in fact, five ^{13}C peaks are observed so that this structure can be eliminated. Obviously, the similarity in the two UV spectra does not demand continuous overlap in the case of trans-10-annulene, and the intense absorption can simply

reflect the overlap of three double bonds in the non-planar structure. In the case of cis-10-annulene, the non-planar conformation mentioned above contains only two partially overlapping double bonds, which is reflected by the relatively weak UV absorption, which effectively excludes a planar conformation from consideration.

The photolysis experiments have led to the construction of a reaction scheme involving $(CH)_{10}$ isomers, including the two 10-annulenes, and rationalising the difficulty in practice of synthesising 10-annulenes, indicating the role of the dihydronaphthalenes[66].

From the theoretical stand-point, the all-carbon 10 π-electron monocycles are quite amenable to calculation of molecular electronic wavefunctions. It would not be unreasonable to study in some detail 10-annulene species by non-empirical calculations, so that some geometry optimisation, for example, could be performed, especially on planar species. However, in this work, the approach has been to perform some preliminary calculations on planar conformations of 10-annulene. It is unlikely that details of molecular and electronic structure will be forthcoming from experimental observation. Thus, non-empirical calculations using the standard basis set have been carried out on selected planar structures. Two geometries were constructed for the all-cis-isomer, one being a regular decagon and the other a classical (alternating) decagon; the geometrical parameters are given in Figure 17. For the trans-isomer, the geometry used is also presented there; a regular structure, based on benzene parameters was used. A planar geometry for the trans-cis-trans-$(cis)^2$-isomer is completely unreasonable, and so this was not considered. Another planar conformation, subject to rather severe internal hydrogen strain effects, was considered; this is based on the structure of cycl(3,2,2)azine (considered below) in which a

central nitrogen bridging atom notionally replaces the three internal hydrogens. There is experimental geometrical data on a simple derivative of the cyclazine, one important aspect being the planarity of the 10-annulene perimeter. The resulting C_{2V}-symmetry structure is not usually considered for 10-annulene itself, but a non-empirical calculation can be performed reasonably on such a species. In most cases in the literature, when calculations (empirical and semi-empirical) have been carried out on 10-annulene, only the planar decagonal conformation has been considered, and this is the one which is really relevant to the Hückel model. In this work, the few planar species considered are rather varied.

From a theoretical point of view, semi-empirical and empirical calculations have been performed almost exclusively on the planar decagonal conformation of 10-annulene, usually to derive a value of the resonance energy for this species; such estimates vary from substantially negative values, indicating destabilisation, to small positive ones[67]. A theoretical study of the geometry and stability of 10-annulene species has been reported[68]; five of the structures proposed in the literature, and mentioned above, were considered and studied by EHT, CNDO/2 and MINDO/2 methods. Partial optimisation of some of the molecular geometries was performed. The most stable member of the series of structures was found to be the boat form (Figure 16), even with its appreciably deformed valence angles. An estimate of the Dewar resonance energy of this most stable 10-annulene was made, giving a value of -155 kJ mol^{-1}, and this was taken as indicating antiaromatic character (corresponding resonance energies of benzene and cyclooctatetraene were +149 and -78 kJ mol^{-1} respectively). The second most stable form was found to be the non-planar helical conformation (Figure 16). The third most stable form was found to be the regular planar decagon, of total energy slightly lower than the

non-planar naphthalenic form. The relatively high stability of the planar decagonal structure was considered striking, bearing in mind the extremely deformed valence angles, and explained by compensation of a large part of the deformation energy by resonance energy in the perfectly planar system. The least stable form considered was the planar "trans" conformation. The fact that the planar trans form is less stable than the all-cis is surprising as, in the former structure, there is less valence angle deformation; the results of the calculations can be explained by considering the unfavourable non-bonding interaction in the trans form caused by one hydrogen atom directed into the molecule. The resonance energies of the two planar forms were calculated to be -230 and -300 kJ mol^{-1}. Overall, the energy differences involved in the above calculations were less than 150 kJ mol^{-1}, so that full geometry optimisation might lead to some significant changes.

Using the standard scaled minimal basis set, the total energy quantities reported in Table 6 were obtained for the particular structures of 10-annulene chosen (details in Figure 17). The planar decagonal structures are strictly the relevant ones in considering the Hückel model. The "classical" structure has a lower total energy than the regular one, in contrast to the prediction of the Hückel model, according to which there should be substantial aromatic stabilisation of the latter. The two total energy values computed are very close; both structures have significantly negative resonance energies (computed by the procedure outlined above), indicating substantial destabilisation of each (4n+2) system. The regular structure used here is very similar to that used in the semiempirical calculations mentioned above, but the estimated destabilisation energy is much less in the non-empirical calculations here. As with other annulene species considered in this work, the regular and alternating planar structures have simply been selected,

as reasonable estimates. Further calculations in the way of geometry optimisation would have to be performed before deciding on the relative stability of the D_{10h} and D_{5h} structures; however, the present conclusion is that the all-cis planar form, even if optimised completely, does not have a significant stabilisation energy (estimating from other optimisation calculations, it is unlikely that even a value of zero for the resonance energy would be attained. Carrying the estimating further, it is unlikely that the resonance energy of the best non-planar form will be substantially positive. Thus, the non-empirical calculations, including interelectronic and molecular geometry effects, have shown that the simple HMO model is not adequate when applied to 10-annulene.

The calculated orbital energies of the 10-annulene species are given in Table 7. The π-orbitals are the ones of immediate interest. In the planar decagonal forms, the π-orbital energies do conform in a general way to the familiar ones of Hückel theory, but there is not quantitative agreement (similar situation to that in benzene and cyclooctatetraene). In addition, there are σ-orbitals lying in amongst the π-orbitals, which is again typical of the type of system. The corresponding orbitals, both σ and π, of the two decagonal forms are very close in energy; this is similar to the situation in benzene, but distinct from that in cyclo-octatetraene. The forms of the π-MO's are shown in Figure 18; as in benzene, the symmetry of the nuclear framework in the polygon, and the resulting forms of all the MO's, leads to a situation in which variation of the internuclear distances produces counterbalancing stabilising and destabilising effects so that the individual orbital energies are almost independent of the r(cc) bond lengths used. As in the cyclooctatetraene case, if an experimental P.E. spectrum of 10-annulene becomes available, the relevant MO's for consideration will be those of a non-planar species (all-cis conformation being significant for 10-annulene also), but the results of the all-cis planar forms will be useful for reference.

Considering planar forms of 10-annulene with hydrogen atoms internal to the ring, the naphthalenic form is unreasonable because of the severe non-bonded interaction involving the two internal hydrogen atoms and also the carbons to which they are attached. This planar conformation has not been considered in this work. The cttct conformation of C_{2V} symmetry (Figure 17) has all the carbon atoms reasonably disposed, but there are three internal hydrogen atoms which are a source of strain in this structure so that it is not expected to be a reasonable conformation for 10-annulene itself. For the particular molecular geometry used here (regular, with respect to bond-length), this is confirmed; the total energy is about 0.9 a.u. higher than that of the all-cis planar forms, showing that the cttct form is destabilised by an enormous amount. Although the total energy value indicates that the structure is completely unreasonable, in other respects, the calculation is not dissimilar to those of the other planar forms. Thus, the calculated orbital energies are presented in Table 7 . The five π-MO's (forms shown in Figure 18) have energies which form a pattern roughly the same as those of the all-cis conformations; the completely symmetrical π-MO, which is the most stable, is somewhat lower in energy in the cttct form, where there is not exact degeneracy of the two pairs of π-MO's. As in the all-cis species, there are high-lying σ-MO's with energies in the π-MO range. However, in the cttct form, these σ-MO's are of a type unique to this conformation; they are predominantly internal C-H bonding orbital combinations. The high-lying C-C bonding σ-MO's of the cttct form are stabilised relative to those of the all-cis forms.

From the AO populations from the calculations on the 10-annulene species (Table 8), it can be seen that the internal hydrogens are distinguishable. The typical situation in hydrocarbon species calculations, using the standard basis set, is that in neutral systems the hydrogen AO population is practically invariant at 0.85 of an electronic charge, with

the total carbon population 6.15 e. The all-cis 10-annulenes conform to this behaviour, as do the external hydrogens and associated carbons of the C_{2V} form. However, the internal hydrogens have a significantly greater population (0.92, 0.95), with the carbons correspondingly of lower total population. There is a larger than normal accumulation of electronic charge on the interior hydrogens, qualitatively indicating some extra screening of the "unnaturally" disposed hydrogen centres.

The planar cttct conformation of 10-annulene, with three internal hydrogens, is clearly very unfavourable; considering also results on the similar conformation of 12-annulene considered below, it is unlikely that a near-planar conformation obtained by some distortion of the interior of the ring will be much more favourable, and very substantial deformation of the basic structure is likely to be necessary before a reasonable structure is achieved. At present, there is no experimental evidence for the existence of a conformation of 10-annulene with three internal hydrogens. This type of conformation has some relevance in considering cycl(3,2,2)azine (see below), which is formally derived from this 10-annulene species.

The remaining planar conformation considered here is the trans-form; the details of the molecular geometry chosen are given in Figure 17 . This particular structure, with a regular benzene-like carbon skeleton, is very similar to that referred to above in the semiempirical calculations. From the total energy values of Table 6 , it can be seen that the trans-form is calculated non-empirically to be the most stable. The semiempirical calculations have underestimated the stability of the all-cis and trans structures, particularly the latter. Of course, it is impossible to say at this point which of the two conformations is actually the lower in total energy, determined by complete geometry optimisation. However, it seems likely that the total energies of the two very different types of structures are very similar, and that even the most stable planar form is somewhat

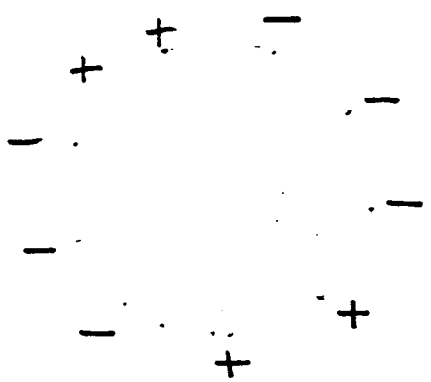
destabilised. It may be that a planar or near-planar trans-conformation, not being exceptionally destabilised, is involved in dynamic processes as an intermediate, as mentioned above.

In Table 7 are presented the calculated MO energies of the trans-form. The forms of the π-type MO's can be illustrated as in Figure 18. Allowing for the lack of exact degeneracy, the π-orbital energy is of the general form found for the other 10 π-species. That the single internal hydrogen is not unfavourably placed can be seen from the AO populations and the lack of the destabilised σ-orbital of internal C-H bonding nature, so that the interanl hydrogen is not noticeably distinguished from the external ones.

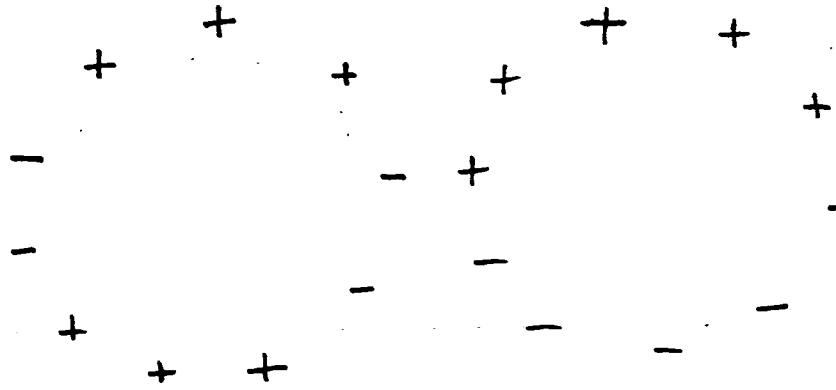
The calculations on the selected planar species here indicate how to proceed in a fuller investigation of the overall 10-annulene energy surface.

FIG. 18. [10]- ANNULENE M.O.'s

D_{10h}

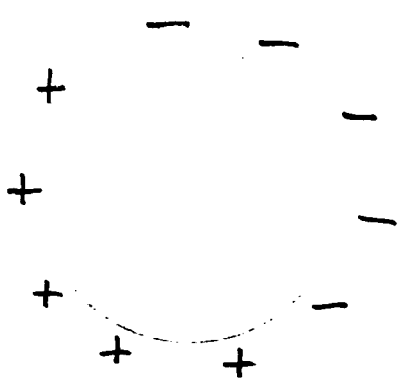


35.π (e)

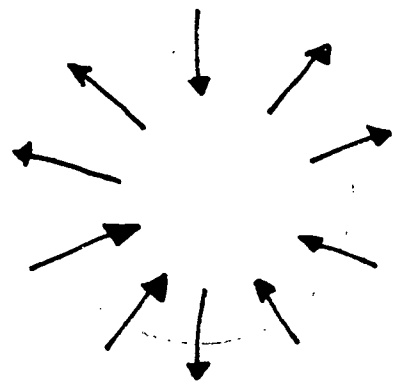


34.π(e)

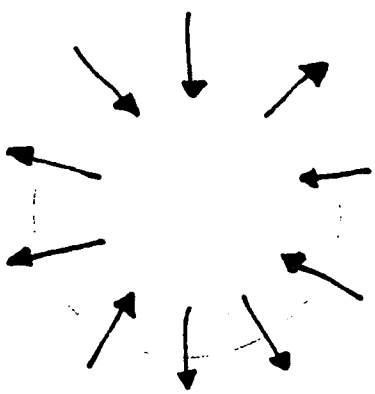
33.π (e)



32.π (e)

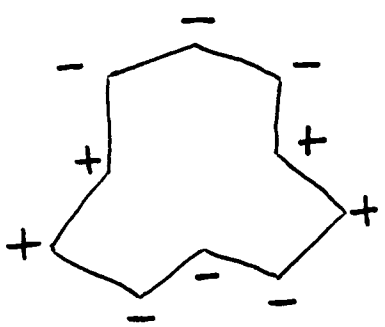


31.σ (e)

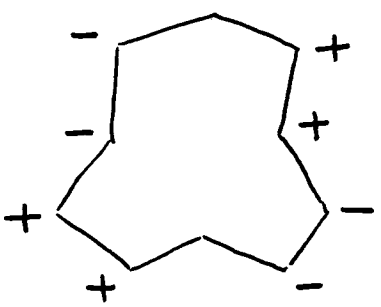


30.σ (e)

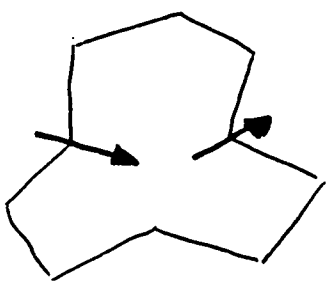
C_{2v}



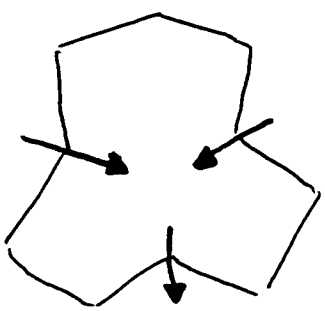
35.π



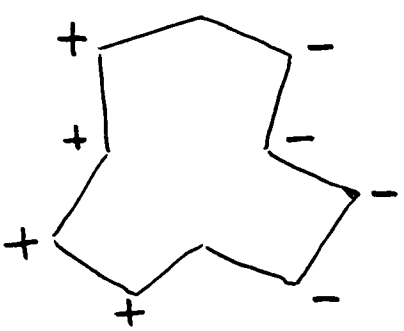
34π



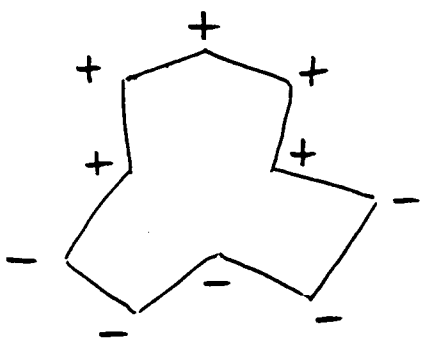
33σ (e)



32σ(e)



31π



30π

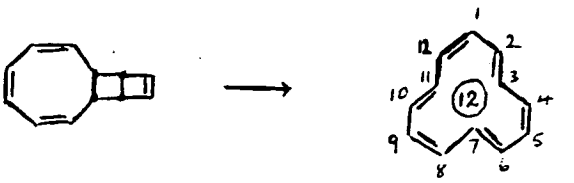


FIGURE 19 : SYNTHESIS OF 12-ANNULENE.

ALL-CIS (D_{12h})	$r(cc) = 1.398\,\text{Å}$, $r(CH) = 1.090\,\text{Å}$ $\hat{ccc} = 150°$, $\hat{cCH} = 105°$
ALL-CIS (D_{6h})	$r(c-c) = 1.50\,\text{Å}$, $r(c=c) = 1.34\,\text{Å}$, $r(CH) = 1.10\,\text{Å}$ $\hat{ccc} = 150°$, $\hat{cCH} = 105°$
CTCTCT (D_{3h})	$r(c-c) = 1.398\,\text{Å}$, $r(CH) = 1.090\,\text{Å}$ $\hat{ccc} = 120°$, $\hat{cCH} = 120°$
CTCTCT (C_{3h})	$r(c-c) = 1.50\,\text{Å}$, $r(c=c) = 1.34\,\text{Å}$, $r(CH) = 1.10\,\text{Å}$ $\hat{ccc} = 120°$, $\hat{cCH} = 120°$
CTTCTT (D_{2h})	$r(cc) = 1.398\,\text{Å}$, $r(CH) = 1.09\,\text{Å}$ $\hat{ccc} = 120°, 105°$; $\hat{cCH} = 120°, 127.5°$

FIGURE 20 : 12-ANNULENE GEOMETRICAL DATA.

(c) 12-Annulene.

Like 10-annulene, 12-annulene (cyclododecapentaene) has been synthesised only by the route involving the photolytic ring opening of a polycyclic valence isomer. The relevant precursor for 12-annulene, obtained from the dimer of cycloctatetraene, is shown in Figure 19 ; photolysis at -100°C yields 12-annulene[69], but it is not obtainable completely pure and can be investigated only in solution at low temperatures, since it rapidly undergoes thermal rearrangement to bicyclic valence isomers at room temperature. The extreme ease of such reaction in the case of 10- and 12-annulenes is responsible for the instability of these compounds. The only available experimental physicochemical information on 12-annulene has been obtained from NMR observations [69]. With 4n out-of-plane π-electrons, a planar 12-annulene system is expected to sustain a paramagnetic ring current according to the familiar model, a consequence of this being the deshielding of inner protons and the shielding of outer protons in the NMR spectrum. Thus, 12-annulene is considered antiaromatic and paratropic. In practice, the NMR spectra of annulenes are temperature dependent; exchange of protons occurs and at higher temperatures the spectra of all the relatively stable annulenes examined so far consist of a singlet. This difficulty can be overcome by lowering the temperature, although with 12-annulene some exchange still occurs at -170°C [69]. A small paratropic effect is observed, with inner proton bands at τ = 2.1 and outer at τ = 4.1. The dianion of 12-annulene, a (4n + 2)-π system, has been prepared, and shown to be diatropic, the NMR spectrum consisting of an inner proton band at very high field (τ = 14.6) and outer bands at very low field (τ = 3.0, 3.8). The data has been interpreted as showing that 12-annulene exists in a configuration with three inner and nine outer hydrogens, the usual representation of three-fold symmetry being shown in Figure 20 . Some non-empirical calculations on several configurations are reported in this work.

Planar Conformations.

The first 12-annulene species considered here are two planar duodecagonal conformations, corresponding to regular and alternating forms for the carbon skeleton. As with the smaller annulenes above, these forms are the relevant ones in Hückel model considerations. The bond parameters used are the standard ones in this work (Figure 20), so that, although the optimum structures are not determined, some general impressions can be obtained. Thus, using the standard scaled minimal basis, the total energy quantities given in Table 9 were calculated. A regular, benzene-like (D_{12h}) structure is found to be significantly destabilised, with a large negative resonance energy; this applies to the closed-shell singlet species. In accord with the simple Hückel model, the total energy of the triplet species is somewhat lower (by an amount which is only about a fiftieth of the destabilisation energy), this latter species being derived from the ground-state singlet by formally moving one of the pair of electrons of the highest occupied MO to the first unoccupied MO; the degeneracy of the two π-MO's involved is seen when both are occupied to the same extent, as in the so-called triplet species, this nomenclature implying only that there are two singly-occupied MO's. The energy difference between the singlet and triplet species is much smaller than in the cyclooctatetraene case above; it is therefore rather uncertain whether the non-empirical calculations really indicate that the triplet species is the ground-state, because the singlet is particularly inadequately described by a single-configuration wavefunction as a result of degeneracy effects. The alternating planar duodecagon is calculated to be of lower total energy than the regular form, as predicted by the simple Hückel model, with the closed-shell singlet being more stable than the lowest triplet species. It is remarkable that two really significantly different geometries of the planar duodecagonal structure are calculated to have very similar total energies.

Thus, it cannot really be concluded that the optimum geometry would still be an alternating ("classical") form. However, it is more certain that the optimum structure is likely to have a very substantially negative resonance energy, indicating significant destabilisation. Empirical and semi-empirical calculations have consistently produced the same qualitative result for planar 12-annulene [70].

The calculated orbital energies of the planar duodecagonal structures are listed in Table 9. The π-type of MO's (whose forms are illustrated in Figure 21) have energies which are in some form of agreement with those of the simple Hückel model; as with the annulenes considered above, there are several σ-MO's lying amongst the π-MO's so that the π-energies do not form an isolated band which lies above the σ-levels. By examining orbitals which correlate between the two species, it can be seen that there are several cases where there is a significant difference in orbital energy (about 1 eV), just as in cyclooctatetraene, but unlike the (4n + 2) benzene and 10-annulene, in which energies of correlating orbitals vary very little. The effect is observed in both the σ- and π-systems. As with the other polygonal forms of the carbon atom skeleton, the forms of the MO's are determined by the duodecagonal structure (topological effect), but the physical situation of the exact placing of the atomic centres on the duodecagon leads to the final orbital energy value. Some MO's have a form which is favoured energy-wise by a "classical" geometry, and others one which is disfavoured. As an example, it is reasonable that the highest occupied π-MO of the singlet is significantly stabilised in the alternating/structure structure (as in cyclooctatetraene). This is in qualitative agreement with semi-empirical calculations, with the perturbed (alternating) removing the degeneracy, by symmetry, of the pair of π-MO's, only one of which is occupied in the singlet, resulting in stabilisation of one.

Based on the calculations on the 4n-species, cyclooctatetraene and 12-annulene, the highest occupied π-MO has a very low orbital energy in the regular species. On qualitative grounds, it might be expected that singly-occupying each of the degenerate pair of π-MO's (triplet) would lead to a lower orbital energy for each, by removing the intra-orbital electron repulsion as in the singlet. An "ionisation-potential" significant energy can be defined for a singly-occupied MO; in cyclooctatetraene, the energy of each singly-occupied MO is somewhat lower than the doubly-occupied value, but the energy is practically invariant to single- or double-occupation in 12-annulene. This indicates that, in the Hartree-Fock model, the intra-orbital repulsion of the relevant MO when doubly occupied balances the "one-electron" energy contribution in 12-annulene, whereas the former is more dominant in cyclooctatetraene. In contrast to simpler models, sharing two electrons between the degenerate MO's of the regular species to form the triplet does have an effect on the energies of all the other MO's; however, the σ-MO energies are almost identical in the singlet and triplet species, and it is only the energies of the π-MO's which are shifted significantly (destabilised by about 1 eV in the triplet). This inter-orbital electron repulsion effect can be qualitatively understood by roughly considering the spatial distribution of the electrons of each π-MO.

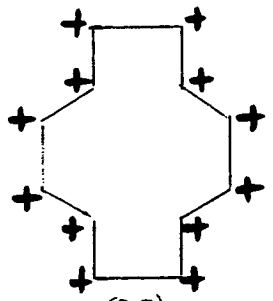
As the size of the annulene ring increases, the all-cis configuration is expected to become less significant physically. For 12-annulene, two other planar configurations are considered here, approaching more physically reasonable structures. One configuration is a (ctctct) one, of threefold symmetry with three internal hydrogens; this is the usual representation (Figure 20) of 12-annulene in the literature, although neither planarity or regularity is actually signified. Two geometries corresponding to this configuration have been considered; one is a regular form (D_{3h}) with benzene-like bond parameters, and the other a classical (alternating) form (C_{3h}) with

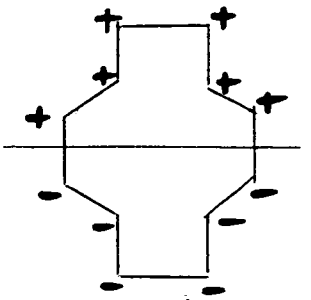
the usual selected parameters used in this work. The other configuration is a (cttctt) one, of twofold symmetry with four internal hydrogens (Figure 20). Only a regular geometry has been considered here (D_{2h}). Such planar conformations are not expected to have physical significance for 12-annulene itself, solely because of the internal hydrogen problem, but there is experimental evidence for these types of structure if bridging groups replace the hydrogens, as noted below. There is some improvement on the 10-annulene situation in that calculations on the planar 12-annulene species can be performed quite reasonably.

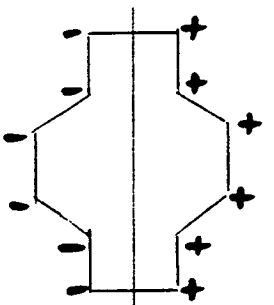
The calculated total energy quantities, using the standard minimal basis set, for the three planar 12-annulene species with internal hydrogens are given in Table 9 . All structures have computed resonance energies which are huge negative values, indicating that they are completely unreasonable physically. Previous semi-empirical and empirical calculations usually yield negative resonance energy values for 12-annulene structures, giving "chemical" values (up to a few hundred kJ mol^{-1}); the non-empirical calculations here include the expected effect of unfavourable nuclear repulsion quantitatively. The total energies of the two (ctctct) structures are very close; although significantly different geometries are involved, the two species are very similar "internally". As with the other annulenes considered in this work, regular and alternating geometries are considered; on a strict theoretical basis, the two structures should be relevant only in studying the all-cis polygonal conformations (highest symmetry forms). The optimal geometry of species such as the (ctctct) and (cttctt) forms of 12-annulene is likely to lie between the two extremes, with electron delocalisation leading to deviation from a completely classical form. However, optimisation will have negligible effect overall on the conclusions reached from the calculations already performed which have yielded information on the gross structures. The completely planar form of 12-annulene with internal hydrogens is very

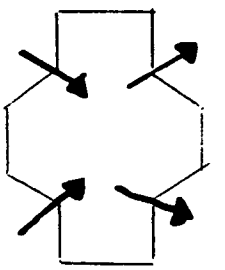
Figure 21. [12]-Annulene

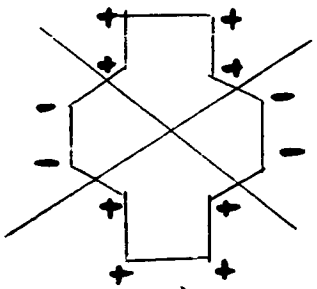
D_{2h}

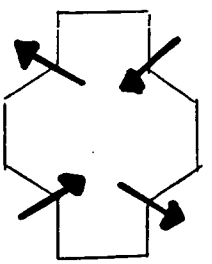

(35) π

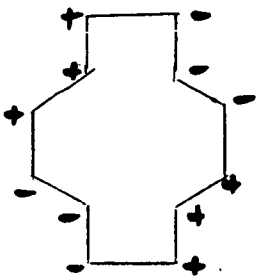

(36) π

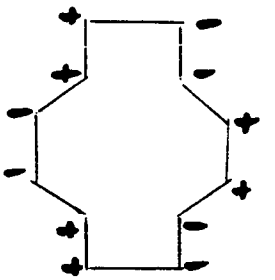

(37) π


(38) σ

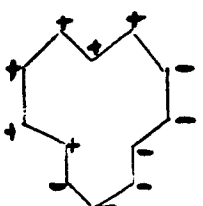

(39) π

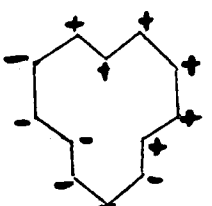

(40) σ

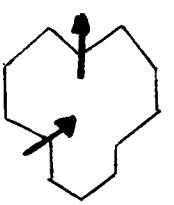

(41) π

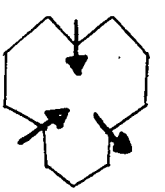

(42) π

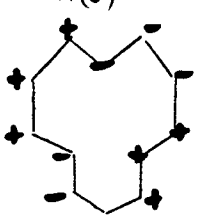
D_{3h}

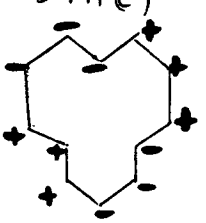

36 π(e)

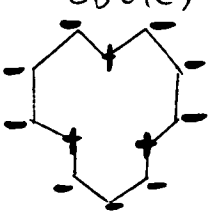

37 π(e)


38 σ(e)

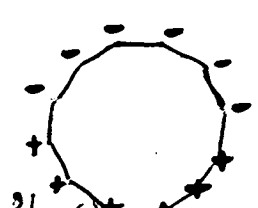

39 σ(e)

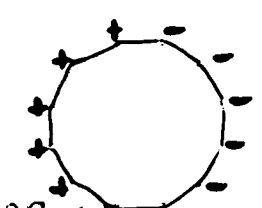

40 π(e)

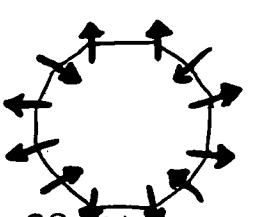

41 π e

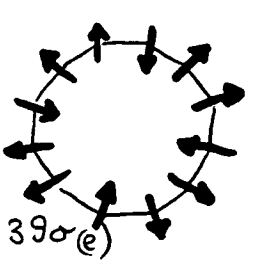

42 π

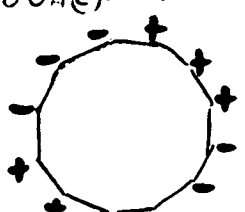
D_{12h}

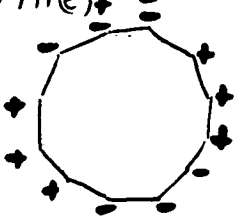

36 π(e)

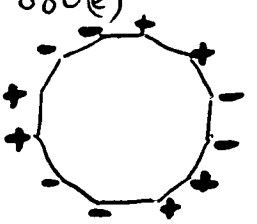

37 π(e)


38 δ(e)


39 σ(e)


40 π(e)


41 π(e)


42 π

heavily destabilised. The (cttctt) conformation is much less so, as anticipated qualitatively, as, of the four internal hydrogens, "interference" occurs only between two pairs separately, whereas, in the (ctctct) form, all three internal hydrogens are unfavourably disposed with respect to each other.

The computed orbital energies for the three conformations of lower symmetry are listed in Table 9. As usual, the π-MO's are probably of more interest. For the (ctctct) conformations, the forms of the π-orbitals are shown in Figure 21; degeneracy is removed in the classical (C_{3h}) structure. The energy level pattern of the π-MO's has the general appearance of the all-cis (Hückel) conformations, but is more stretched out. This indicates the greater interaction among the constituent π-systems of the (ctctct) form - more compact structure with more favourably disposed double bonds. In particular, the highest energy π-MO, the completely symmetrical and bonding one, is stabilised by about 1 eV over that of either all-cis form. The highest occupied π-MO is of low I.P. in both regular and alternating (ctctct) forms. In contrast, in the all-cis conformations, there is a significant stabilisation (by about 1.5 eV) of the highest occupied π-MO in the classical structure over the regular one, where one of a degenerate pair of π-MO's is occupied. In the (ctctct) structures the nature of the highest occupied MO shows that the exact form of the underlying nuclear centres (classical or regular) has little effect on the orbital energy, which will remain low in all geometries of this type. For both structures, the singlet closed-shell species is calculated to be more stable than the triplet; the highest occupied π-MO of the singlet is not one of a degenerate pair by symmetry, and so this result is expected. The energy difference is very small so that it cannot definitely be concluded that the actual lowest energy species of the planar (ctctct) form is the singlet.

Non-planar Conformations.

All the planar structures of 12-annulene considered above are destabilised, with reference to "classical" conjugated structures. Even if optimisation of the geometric configurations of the various symmetry types were performed, it is unlikely that significantly different conclusions would be reached. The planar all-cis duodecagonal conformation, whether of D_{12h} or D_{6h} symmetry, is not heavily destabilised, but the results here indicate that there is agreement with predictions of simplet models, although description of the destabilising effect in the same terms, such as the concept of anti-aromaticity, does not follow consequently. The more chemically appealing conformations, (ctctct) and (cttctt), are very heavily destabilised when completely planar. Thus, as expected by qualitative reasoning, planar 12-annulene cannot be considered as a realistic species. However, further investigation to find the equilibrium configuration and its total energy is desirable.

Proceeding to non-planar conformations, the (ctctct) and (cttctt) planar forms are useful starting-points. As a first step in finding a lower total energy structure, some improvement in the internal hydrogen situation seems reasonable. The particular geometry of the planar (cttctt) form used here is related to a geometry used for 14-annulene (considered below), and, in fact, the central part of the two geometries is the same (economical in practice). Thus, the unfavourability of the 12-annulene and 14-annulene planar forms is likely to be predominantly due to the internal hydrogen "crowding", which is the same in both; some movement of the internal hydrogens and resulting calculations are considered for 14-annulene below, and it is likely that similar results would be obtained with the (cttctt) form of 12-annulene. In addition, experimental evidence, constructed from NMR data and consideration of the nature of the synthetic route to 12-annulene, indicates that a conformation with <u>three</u> internal hydrogens is the only one of relevance, or, at least, one with nine hydrogens of one type and three of another. Thus, further calculations reported in this work deal with the (ctctct) conformation only.

All the simple, completely planar 12-annulene structures so far considered are very heavily destabilised. In contemplating optimisation of the geometry, the improvement in total energy required to attain even the reference level of zero stabilisation energy, corresponding to a hypothetical polyolefin, is so large that it is reasonable to conclude that if a 12-annulene structure of reasonable total energy exists, the molecular geometry will be far from any of the ones considered above. Optimisation of each of the above conformations within the constraint of planarity is likely to produce only relatively small total energy improvements so that conclusions will remain largely unaltered. Thus, any worthwhile geometry optimisation involves the introduction of non-planarity; none of the completely planar structures is a reasonable starting-point.

Application of the geometry otpimisation procedure including the gradients technique (Chapter 4) requires the construction of an initial molecular geometry. There is no direct experimental information. However, there is theoretical data, of a semi-empirical nature, available through the application of Allinger's "Molecular Mechanics" procedure.[71] This method for structure determination has been extended to hydrocarbons containing delocalised systems by including a quantum-mechanical (VESCF) π-system calculation in the iterative sequence, from which bond orders are obtained; force constants are calculated from the bond orders, and then the delocalised electronic system of the hydrocarbon molecule can be handled in the standard way of force field methods.[72] The procedure is applicable to non-planar systems as well as to planar systems. It has been applied to many simple compounds, e.g. butadiene, benzene, biphenyl, naphthalene, and to more complicated systems such as o-di-tert-butylbenzene, bicyclo 4.4.1 undecapentaene (methylene bridged 10-annulene). Insofar as experimental data are available, the agreement with experiment is generally good. In a few cases, structural predictions have been made, and one of

these is 12-annulene. Details of the geometry of 12-annulene calculated by Allinger's procedure are presented in Table 10 . The most stable conformation is a (ctctct) one, with torsional angles across the $C(sp^2)$-$C(sp^2)$ single bonds deviating 35-90° from planarity. As mentioned above, the NMR spectrum of 12-annulene has been interpreted as the rapidly interconverting conformation shown in Figure 22 , predicted to have 50-60° torsional angles about the C-C single bonds on the basis of models. The activation enthalpy for the interconversion was deduced via NMR to be 15.5 kJ mol^{-1}, with the suggested transition state shown in Figure 22 ; Allinger's calculation has the latter at an energy of 13.9 kJ mol^{-1} above the most stable conformation. This structure is not an unambiguous transition state, however; another structure (Figure 22) has a calculated relative conformational energy of 8.8 kJ mol^{-1}, but is discounted as a possible transition state or intermediate as it is a structural isomer with one fewer trans double bond than the lowest energy form, and considerable energy would be required to interconvert it, giving a much higher energy barrier.

It was considered that the geometry of the most stable conformation calculated by the above procedure would provide a reasonable starting-point for molecular geometry optimisation by non-empirical calculation. The data presented in Table 10 shows that all the CC bond lengths, CCC angles and dihedral angles are given. The calculated geometry does not possess rotational equivalence; it was decided to construct a rotationally equivalent (C_3) structure from the data by suitable averaging. In particular, the internal hydrogens are rendered equivalent. The structure involved is not a simple one; in order to construct a geometry, it was necessary to disect the structure into simpler parts and refer to a catalin model. Briefly, the 12-annulene molecule has a generally undulating C-C periphery; it possesses a C_3 propeller shape with C_1=C_2, C_4=C_5, C_8=C_9 as effectively

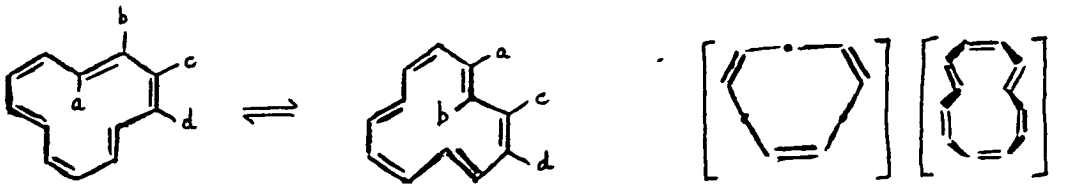


FIGURE 22 : INTERCONVERTING CONFORMATIONS OF 12-ANNULENE.

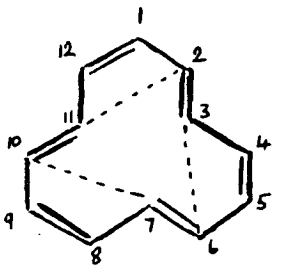


FIGURE 23 : 12-ANNULENE C_3 "PROPELLER"

the blades (Figure 23), each being part of a 4-carbon skeleton of a but-2-ene type. Even with the averaging of Allinger's data (and adding suitably situated H centres), it was a major task to construct a set of cartesian coordinates for the geometry of the molecule. There was the obvious advantage of using the cartesian system of the planar (ctctct) conformation and performing out-of-plane rotations, but this was outweighed by it being a long way off from the equilibrium structure. Having constructed a starting geometry, the aim is this work was simply to perform one cycle of the gradients procedure of geometry optimisation in order to obtain improvement in the gross structure.

Using the standard minimal basis, the results of the geometry optimisation calculations are summarised in Table 11 . The calculated total energy of the initial structure is very substantially more negative than that of any of the planar species, and is almost equal to that of the hypothetical reference. The gradients of the total energy are uniformly very small, bearing in mind the results of other optimisations; completing the cycle by finding the optimum α value does not nullify the tentative conclusions from considering the gradients alone, as the value of α (approximately 0.5) together with the variation of total energy with α shows that the initial geometry is near optimal. Bearing in mind the dimensionality of the optimisation problem, it is remarkable that there has been so little improvement in the geometry. The fact that the starting position has rather fortunately turned out to be so near to the optimum is particularly pleasing as even performing one cycle of the procedure is a large task. The total energy has improved by only about 15 kJ mol^{-1} over the cycle. Although the results from only one cycle cannot be completely conclusive, it does seem that there would not be significantly greater improvement on further optimisation. As expected

with such large 'flexible' systems (movement within shallow potential wells), the variation of geometrical parameters is somewhat more significant than that of the total energy. There is not really much change from the initial structure; there has been introduced some further alternation in the C-C periphery, and the C-H bond lengths have tended to lengthen, including the internal ones. It is concluded that the actual calculated equilibrium structure is essentially a classical one. All the bond lengths are typically classical (allowing for systematic elongation of C-C single bond lengths) and all bond angles are close to 120°. The internal hydrogens, as a result of the marked deviation from planarity, are situated at slightly more than 2Å from each other, giving a sterically favourable situation. The calculated stabilisation energy is effectively zero, in keeping with a classical structure (polyolefin). Thus, by adopting a far-from-planar conformation, 12-annulene can attain a 'respectable' total energy and avoid the heavy destabilisation attributed to steric effects.

As well as in total energy, other results of the calculations on the non-planar conformations show similarities to those of typical chemical species, unlike the situation with the planar conformations. Thus, the calculated AO populations (Table 12) of the non-planar 12-annulenes show a reasonably uniform distribution of electronic charge around the ring. In particular, the internal hydrogen population is that typically found in calculations on hydrocarbon species (about 0.85), and is effectively identical to that of the external hydrogens, in agreement with there no longer being a clear-cut distinction between the two types of hydrogen. The internal C-H dipoles are typical hydrocarbon ones; in the planar forms, the internal hydrogens have a greater AO population at the expense of the attached carbon so that the internal C-H dipole is larger than usual.

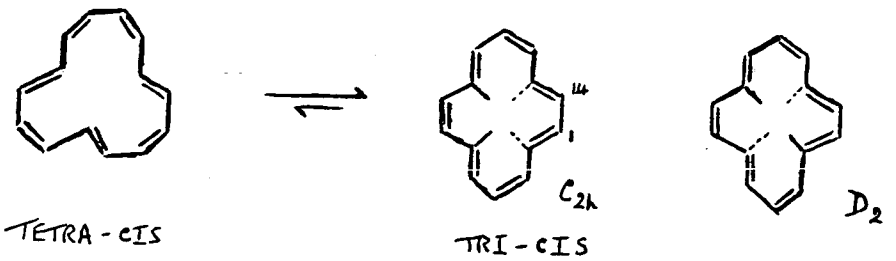
The calculated orbital energies of the non-planar conformations are given in Table 13 . In comparing with the results from the planar species, one obvious feature is that the highest occupied MO orbital energy in the non-planar form (corresponding to the first ionisation potential) has a value similar to that calculated in the smaller conjugated olefins (e.g. butadiene), unlike the HOMO orbital energy of the planar (ctctct) form which is particularly small in magnitude; this is consistent with a polyolefinic nature for the 12-annulene species. The non-planar conformation deviates so far from planarity that it is virtually impossible to correlate the orbital energies with those of the planar (ctctct) forms. However, by comparing much less precisely than usual, simply by noting the MO's of a- and e-type symmetry, it can be seen that the calculated spectra of the planar and non-planar forms are reasonably similar, with the exception of the top two levels (lowest IP). Apart from the obvious difference in values of the HOMO energies, there is a reversal of symmetry-types, so that the non-planar HOMO energy is in fact that of a degenerate pair; the least stable a-type MO in the non-planar form is thus a little more stabilised compared to that of the planar form. Thus, by far the most significant effect is the removal of the low I.P. a-type MO (one of a degenerate pair of π-MO's of 'alternating' character in the planar ring). At the same time, this also means that the low-lying lowest unoccupied MO of the planar form (the other of the degenerate pair of π-MO's) is removed so that it can be concluded, without explicit calculation and using the HOMO-LUMO energy separation from the calculation on the neutral singlet that the lowest energy triplet species of the non-planar form lies significantly above the ground-state singlet. From examining the forms of the MO's, it is seen that there is extensive delocalisation of the MO's. The distorted structure has not only removed any semblance of a σ-π separation,

but it is not really possible to examine the occupied MO's in terms of local π-systems. Apart from noting the classical nature of the non-planar macrocyclic structure, it is not really useful to describe it in terms of a simple model as it is so far removed from typical planar, or near-planar, aromatic-type systems. The internal hydrogen feature has been removed in the non-planar form; there are no high-lying σ-orbitals predominantly of internal C-H bonding character. This effect is obviously expected. However, the abnormally low I.P. of the planar forms has also been removed.

(d) 14-Annulene

Pure crystalline 14-Annulene can be synthesised by Sondheimer's method outlined above.[73] In contrast to 10- and 12-annulene, 14-annulene is relatively stable, and some more conventional chemistry can be performed with it, e.g. substitution reactions can be carried out. However, it is still some way from being easily handled, so that it is difficult to perform physical chemical measurements on it. As usual, the NMR spectrum has been obtained. At low temperature the spectrum of 14-annulene is typical of a diatropic (4n+2) system, consisting of inner proton bands at high field (τ 10.61) and outer proton bands at low field (τ 2.12). It was recognised early on that crystalline 14-annulene was converted to an equilibrium mixture of two stereoisomers on dissolution at room temperature. It seemed likely that these were conformational isomers, since cis-trans isomerism is not expected to occur under such mild conditions, by analogy with linear conjugated polyenes. However, after it was found that crystalline 16-annulene is converted in solution to an equilibrium mixture of two configurational isomers,[74] it was suspected that 14-annulene underwent the same type of configurational, rather than conformational, isomerism; this has been confirmed by determination of the NMR spectrum at very low temperatures (-155°C), which showed the presence of the two isomers of Figure 24. At -10°C, the equilibrium mixture consists of the "tri-cis" and "tetra-cis" structures in the ratio 92:8.[75] Thus, ready interconversion of different configurational isomers of certain annulenes is possible.

At room temperature, the "tri-cis" configuration is effectively the only one to be considered. The room temperature NMR spectrum consists of two sharp singlets at 4.42 and 3.93 τ in a ratio of about 6:1; the averaging effect of proton exchange is evident. In addition, this effect has been attributed to the existence of two conformational isomer (Figure 24).



TETRA-CIS

TRI-CIS

FIGURE 24 : CONFIGURATIONAL ISOMERISM OF 14-ANNULENE; ALSO, CONFORMATIONAL ISOMERISM OF TRI-CIS.

D_{2h} - REGULAR

$r(c\text{-}c) = 1.398\,\text{Å}$, $r(CH) = 1.090\,\text{Å}$

$c\hat{c}c = 120°$, $c\hat{c}H = 120°$

C_{2h} - ALTERNATING

$r(c\text{-}c) = 1.50\,\text{Å}$, $r(c=c) = 1.34\,\text{Å}$, $r(CH) = 1.10\,\text{Å}$

$c\hat{c}c = 120°$, $c\hat{c}H = 120°$

FIGURE 25 : 14-ANNULENE GEOMETRICAL DATA.

An early X-ray examination of the crystalline state indicated that, unless there is disorder in the crystal, the molecules are centrosymmetric, so that the C_{2h}-isomer was presumed the more stable isomer. Subsequently, an X-ray analysis of the crystal structure of 14-annulene has shown that the fast isodynamical transformation in solution at room temperature is "frozen out", and the "tri-cis" configuration only is found; moreover, the C_{2h}- conformer is the one in which 14-annulene exists.[76] The structure of the 14-annulene molecule in the crystal is described as substantially non-planar, with the "internal" C and H atoms about 0.21 and 0.77 Å respectively out of the best molecular plane. Details of the structure are given in Table [14]. The C-C bond lengths lie in the range 1.350-1.407 Å, but there is no significant chemical pattern to the values, such as bond alternation. Qualitatively, the steric overcrowding of the four internal hydrogens leads to distortion from planarity distributed throughout the molecule; torsion angles about the cis double-bonds have moved from the planar system 0° to $10\text{-}20^{\circ}$, whilst those about the trans from 180° to $160\text{-}170^{\circ}$. Valence CCC angles lie in the range $123\text{-}131^{\circ}$. Allinger force-field calculations (as with 12-annulene above) have been performed on the "tri-cis" configuration.[71] The results show that the C_{2h}-conformer is about 10 kJ mol^{-1} more stable than the D_2-conformer. The energy difference between the conformations is due primarily to the ring distortions caused by the internal hydrogen repulsions. In the C_{2h} form, the ring undergoes torsional deformations with an average deviation from planarity of 16.2° and bond angle deformations with an average ring angle of 125.2°; in the D_2 form, the ring is more planar (12.6° average deviation from planarity), but the average ring angle is expanded to 125.9°. The calculated geometrical parameters are presented in Table [14]. It is considered that the aromatic character of 14-annulene indicated by the low

temperature NMR spectrum is entirely compatible with the calculated structure (and crystal structure), with the approximately regular bond lengths of the carbon periphery a significant feature.

The above results are obviously of relevance in considering the calculations detailed below, these being performed without using the literature data.

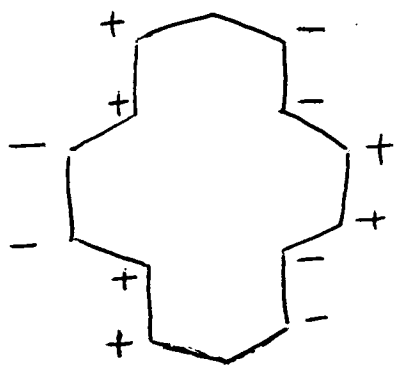
Planar Conformations

In the non-empirical calculations of this work only the "tri-cis" conformation of 14-annulene was considered. In previously reported empirical and semi-empirical calculations only this "chemical" species has been of interest, with planar polygonal species considered as being physically unreasonable. With the usual variety of theoretical procedures, the resonance energy of 14-annulene is usually calculated to be some significantly positive value, in keeping with the Hückel model, even although there is little similarity between the "tri-cis" and polygonal conformations.[77] Thus, two completely planar structures of 14-annulene were constructed, as usual,one being a regular form (all c-c bond lengths having the standard benzene value) and one being an alternating ("classical") form (Figure 25). Using the standard minimal basis set, calculations were performed and the total energy quantities of Table 15 were obtained. The calculated stabilisation energy of each planar species is very substantially negative (total energy is well short of the sum of individual bond energies). This is consistent with the qualitative prediction that internal hydrogen steric strain is dominant, so that aromatic stabilisation, if present, is masked. The heavy destabilisation is similar to that in the 12-annulene species with four internal hydrogens similarly situated. Another feature, which shows deviation from the prediction obtained by simply applying the Hückel MO model, is the greater stability of the "classical" structure, by rather a large amount (magnitude typical of resonance energies of aromatic species). In the smaller annulene species considered previously, the corresponding total energy difference is barely of chemical significance. In particular, in both classical and regular forms, the internal hydrogens are very similarly situated; the difference in total energy is an order of magnitude smaller than the destabilisation energy. However, as usual,

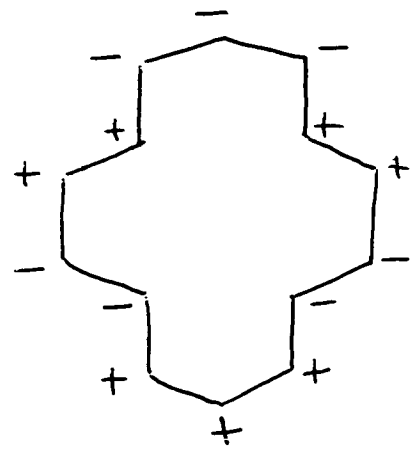
only particular geometries have been chosen, and it is therefore uncertain as to what the situation is at the lowest energy forms of the two structure types, although it is unlikely that geometry optimisation of the regular form will improve the total energy by the amount by which it lies above the classical form at present. It can be concluded with some certainty that the completely planar "tri-cis" conformation of 14-annulene is very heavily destabilised, and that the exact geometrical structure does not significantly affect the stability.

The calculated orbital energies of the two planar structures are presented in Table 16. Based on the results with these two rather different geometries, it is likely that nearly all the orbital energies are quite insensitive to the geometry of the carbon skeleton. This contrasts with the behaviour with the other (4n+2) species, benzene and 10-annulene; however, these latter two species were treated as regular and alternating polygons, and a polygonal 14-annulene species would probably show the same effect. With the constraint of the symmetry of the molecular geometry considered, the effect on orbital energies can be rationalised. The π-type MO's exemplify the situation, and the forms of these in the "tri-cis" 14-annulene are shown in Figure 26. Allowing for the lower symmetry of this conformation and the resulting lack of exact degeneracy, the pattern of π-MO energies is roughly similar to the Hückel model pattern. In the regular planar conformation the two highest occupied orbitals (both π-type) are of rather low orbital energy, referring to the corresponding values in the (4n+2) regular polygons considered above, lying in value between 4n and (4n+2) values. In the alternating planar conformation the corresponding orbitals are significantly stabilised, being the ones which show the greatest difference in the two planar forms. This effect is consistent with the forms of the orbitals. This behaviour is similar to that observed in the

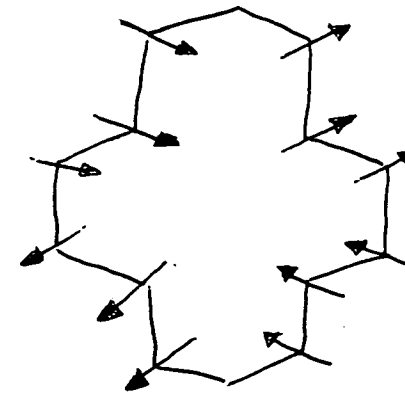
FIGURE 26. [14]-ANNULENE "OUTER" M.O.'s



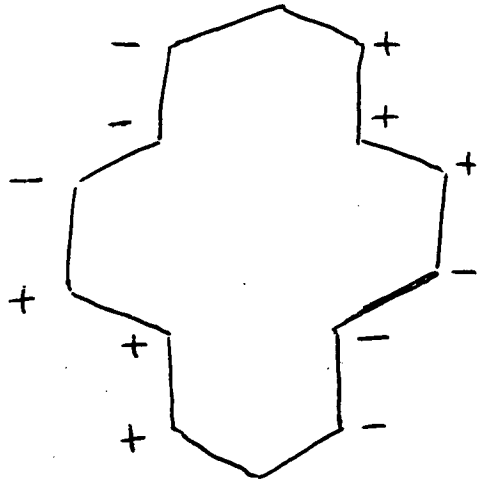
49π



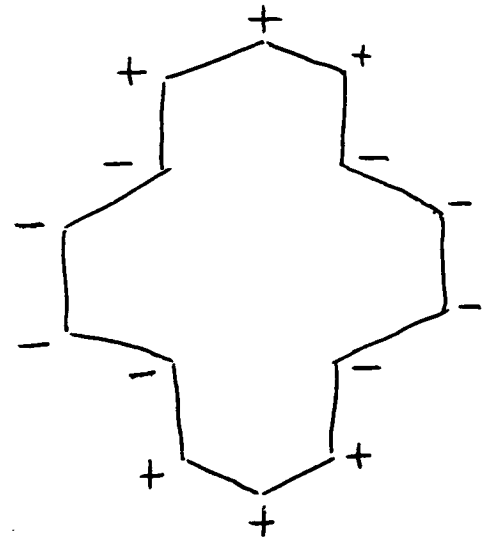
48π



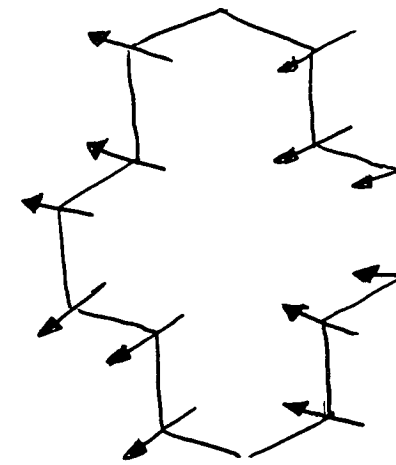
47σ



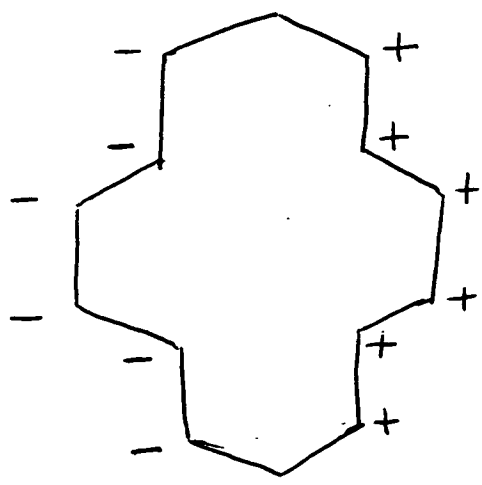
46π



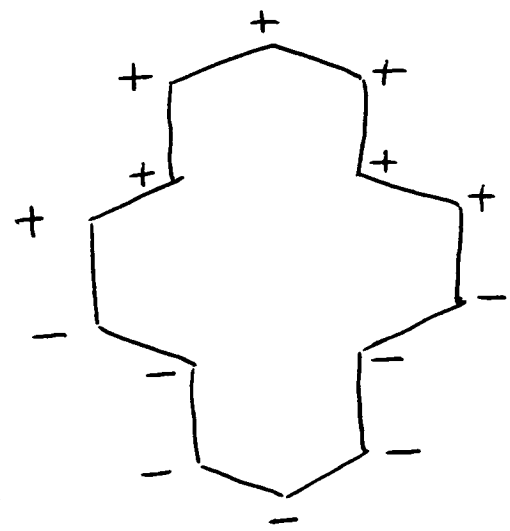
45π



44σ



43π



42π

4n species. In addition, the orbital energies of the lowest unoccupied MO's (π-type) are very small positive values; although these are not rigorously defined, this does indicate that there are relatively low-lying excited states of the species. Thus, in the "tri-cis" planar conformation, there is less similarity to (4n+2) behaviour with 14-annulene than with 6- or 10-annulene.

In considering the σ-type MO's, there are several of these lying amongst the π-type, with the ones of particular interest being those characteristic of the internal C-H bonds. There are two such orbitals, separated by about 1.3 eV (forms shown in Figure 26); the orbital energies are effectively the same in the regular and classical structures. As in the species with internal hydrogens considered above, the AO populations (Table 16 for 14-annulene) are typical of hydrocarbons, except for those of the internal hydrogens and associated carbons. The internal hydrogens have excess electronic charge (about 0.1) at the expense of the associated carbons, interpretable as a "screening"effect.

Non-planar Conformations

Before attempting some rigorous geometry optimisation, allowing a three-dimensional structure for 14-annulene, a simple gross change was made to the planar structure and a calculation subsequently performed as an indication of what to expect upon carrying out the rather expensive optimisation by the iterative technique of Chapter 3. Thus, considering the literature data on the geometrical structure mentioned above, the optimised structure of 14-annulene ("tri-cis" conformation) is not likely to deviate greatly from planarity, with only the internal hydrogens and associated carbons being distorted out of plane; this is in contrast to the situation with ctctct 12-annulene as shown above, as expected from the nature of the internal strain problem. Although the crystal structure data has been interpreted as showing a "significantly" non-planar structure, the deviation from planarity is not considered here to be very large. From the literature data, it is found that the internal hydrogens are rotated by about 15° from the mean plane of the molecule, with other small displacements of the other atoms, particularly the attached carbons. As a gross first approximation to the non-planar geometry, the regular planar structure constructed in this work was taken as basis and the four internal hydrogens were rotated by 15° out of plane, leaving all other centres intact. Two structures, of C_{2h} and D_2 symmetries, are two possible results, and these are the most physically reasonable as the internal hydrogen strain can be effectively completely removed in such conformations. Thus, calculations were performed on these two distorted planar structures; these were economical as only four hydrogen atom positions differ from the original basis calculation, so that these exploratory calculations did not incur much extra expense. The total energy values computed are given in Table 17. Both structures have considerably lower total energies than the planar ones.

The type of distortion of hydrogens carried out here, when done with more usual species, would tend to change (usually raise) the total energy by something of the order of 10 kJ mol^{-1} (typical "chemical" value). The improvement with 14-annulene by this rather arbitrary change is about 500 kJ mol^{-1}; this illustrates more quantitatively the internal strain effect, when the separation of unfavourably placed hydrogens leads to a very large improvement in total energy which completely masks the unfavourable local distortion of the hydrocarbon framework. However, the total energy has still to improve by the same sort of amount before a resonance energy of even zero is obtained for 14-annulene, and such improvement is likely to occur if the carbon skeleton can relax to counterbalance the hydrogen movements to remove the strain which has been added. The total energies of the C_{2h} and D_2 non-planar structures are very similar; the latter is very slightly lower, in contrast to the semi-empirical calculations' findings, but the difference is not significant here compared to the gross improvement made in the geometry in either form. The total energy values show that the very unfavourable completely planar geometry, whether regular or alternating, can be improved very substantially without distorting the carbon skeleton greatly from planarity (unlike 12-annulene), with effectively only the internal hydrogens being removed from the plane, and this can reasonably be done, yielding a near-planar structure for "tri-cis" 14-annulene.

The rather severe distortion of the 14-annulene has a drastic effect on the total energy, but it has only a minor perturbational effect on the rest of the molecule apart from the internal hydrogens and attached carbons, as judged by the computed orbital energies and AO populations (Tables 16,17). The total populations of the internal hydrogens and neighbouring carbons are typical hydrocarbon values, just like the atomic centres of the

remainder of the structure which have remained effectively constant in AO population. All the C-H bonds are similar, and the internal ones have lost their distinction. All the orbital energies, apart from two, have remained constant. The differences in corresponding values are really insignificant, being much smaller than the changes with C-C bond length in the planar form, and these tend to be very small. The two orbital energy values which change significantly on distorting the planar structure are those of the two unusually high-lying σ-orbitals, which are predominantly combinations of internal C-H bonding contributions. In the non-planar forms, both orbitals are stabilised by about 0.5 eV. From examination of the forms of the MO's, this can be explained by a decrease in the main antibonding contribution to the MO's by the relative movement of the two pairs of closest internal hydrogens (positions 3,13 and 6,10). This is the dominating effect, although there are effects on the bonding contributions and the lesser antibonding ones, and these vary with the orbital and also the structure. Thus, although both orbitals in each of the non-planar structures are stabilised by approximately the same amount so that the energy difference is nearly constant at 1.3 eV, there is a small splitting effect; the less stable MO of the pair is stabilised more by distortion, and the effect is slightly greater in the D_2-symmetry case. The overall observation is that the two σ-MO's behave as though they are "local" orbitals, and are really the only ones affected by the distortion of the internal hydrogens (whose AO's contribute very little to other MO's). However, the delocalised nature of the orbitals is emphasised by considering the total energy components. In going from the planar to the non-planar forms, the one-electron energy is reduced and the electronic energy is reduced by a smaller amount, indicating a decrease in electronic repulsion. The relatively large

decrease in nuclear repulsion leads to the substantial lowering of total energy. The two σ-MO's of interest are the main cause of the electronic energy changes; their individual one-electron energy terms are reduced, but electronic repulsion with other MO's is reduced by a greater amount.

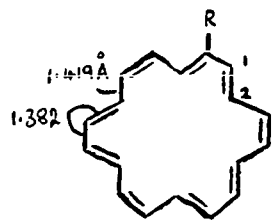
As a starting-point for more systematic geometry optimisation, the C_{2h} non-planar geometry was used. This had the advantage of a much lower total energy than any planar form, but, more importantly, the constraint of planarity was removed so that three-dimensional optimisation could be performed (impossible from initial completely planar form). This starting geometry had a regular C-C bond-length periphery, but this constraint was removable. The actual constraint on the optimisation performed was that of C_{2h} symmetry. This constraint on the relative internal hydrogen dispositions is unlikely to have a significant effect on total energy; more important is likely to be the effect of the constraint on the C-C periphery (regularity of bond-length relaxed, but alternation not completely so). The gradients approach of Chapter 4 has been used. The results obtained for two iterative cycles are shown in Table 18. After the first cycle, the total energy has been lowered by another substantial amount, but after the next, the improvement is quite small. The execution of the rather expensive procedure has substantially improved the position, but it is difficult to say whether significant further improvement is likely, i.e. whether convergence is not smooth or whether the total energy has almost converged. As with intermediate positions of the optimisation process, the actual final geometry obtained is rather distorted, although the indications are that a basically classical periphery prevails. One significant factor is the assumption here of a particular symmetry, and the relaxation of this extra constraint will provide further improvement in total energy by an unknown

amount. Thus, it seems likely that a nearly planar structure (with respect to carbon skeleton) is obtainable, without being very heavily destabilised. As the structure has varied throughout the optimisation, most of the orbital energies have remained nearly constant (e.g. π-type, in accordance with smallish deviation in planarity). The notable exceptions are the high-lying pair of σ-orbitals (C-H type), which are progressively stabilised until they have nearly reached a more normal level. The AO populations (Table 18) also indicate a more uniform electronic charge distribution, with the internal C-H dipoles practically indistinct from the others in the final structure.

(e) 18-Annulene

The first higher annulene to be prepared was 18-Annulene, and the synthetic method used by Sondheimer has been outlined above; the syntheis is now well documented.[57] Unlike the previously considered members of the higher annulenes which have been synthesised recently, 18-annulene is reasonably easy to examine and so several aspects of its chemical and physical properties have been noted. It has always been presumed that the molecule is reasonably planar, so that the question to be answered has been: is this (4n+2) species aromatic in the same way as benzene? The answer has been found to be that 18-annulene resembles benzene in some ways, but not in others. Thus, some "aromatic" properties of 18-annulene are summarised in Figure 27; by contrast, some "non-aromatic" properties are indicated in Figure 27. It has been shown that 18-annulene in its properties and reactions is certainly not as inert as benzene, and would not have puzzled the 19th century chemists the way benzene did.

18-annulene has been thoroughly studied theoretically; considerable ambiguity still exists regarding its structure and the factors affecting its stability. A structure of equal bond lengths, as originally predicted by simple bond order calculations, turns out upon examination from the valence bond point of view to be a saddle point.[78] Both VB and MO methods have been used to show that bond alternation is present in 18-annulene. The available experimental evidence, however, seems to contradict the alternation theory. The slightly distorted D_{6h} molecular symmetry found in the crystal structure[79] is claimed to rule out alternate long and short bonds, and instead shows a different type of distortion 12 inner (trans) bonds of mean length 1.382 $\pm$ 0.003 Å and 6 outer (cis) bonds of mean length 1.419 $\pm$ 0.004 Å, with ring carbon deviations of $\pm$ 0.085 Å from the mean plane of the ring. Semi-empirical calculations of π-electronic transitions led the investigators to conclude that the molecule does not have this



"AROMATIC" PROPERTIES.

ELECTROPHILIC SUBSTITUTION :
$R = NO_2$, COMe, CHO, etc.

X-RAY STRUCTURE : NEARLY PLANAR,
NO BOND ALTERNATION

^1H-NMR SPECTRUM : DIATROPIC
$-60°$ $\tau_I = 12.99$, $\tau_o = 0.72$
$+110°$ $\tau_{ALL} = 4.55$.

"NON-AROMATIC" PROPERTIES.

NOT PARTICULARLY STABLE,
SOLID DECOMPOSES ON STANDING.

RAPIDLY REACTS WITH Br_2, $KMnO_4$, etc.

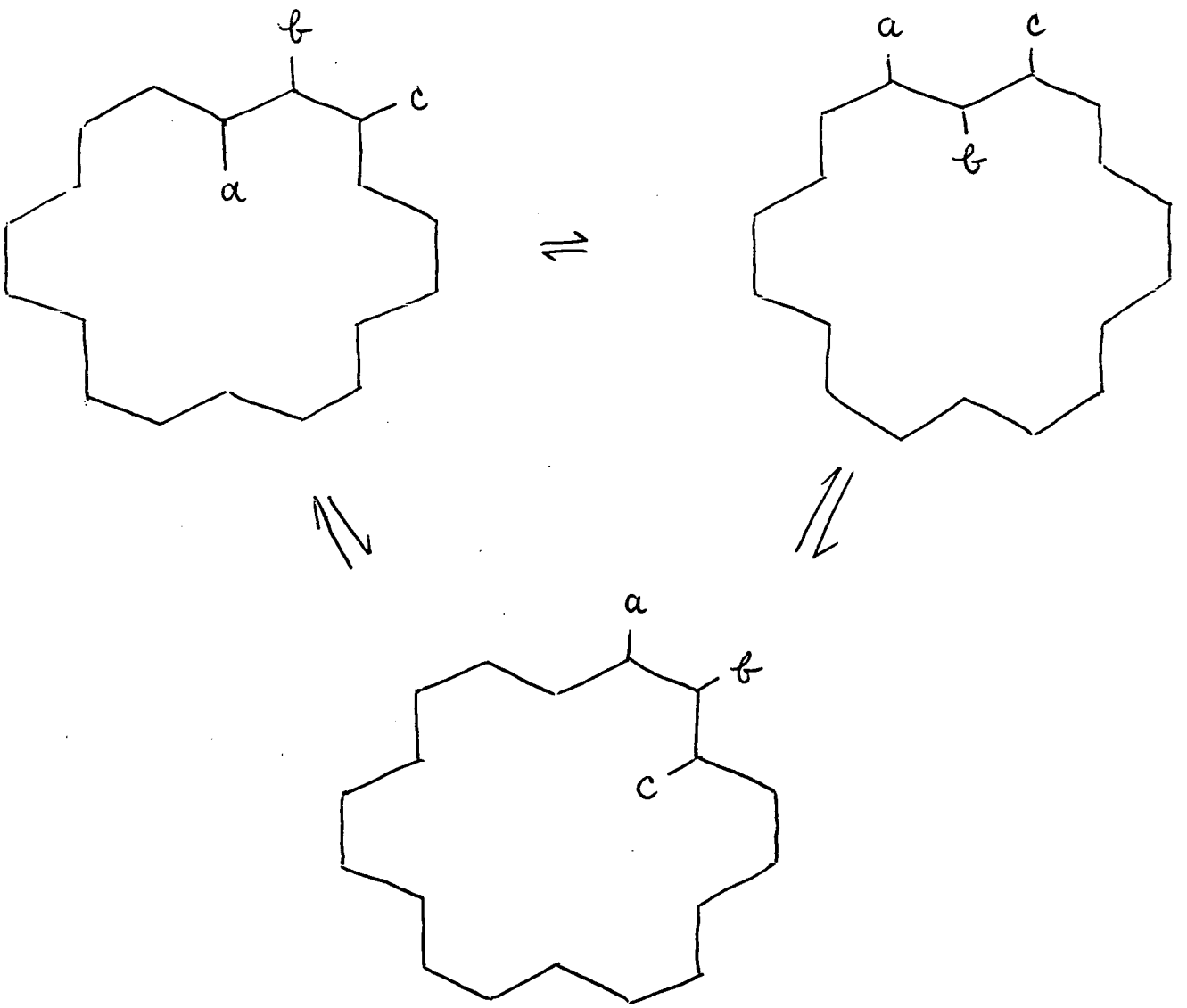
VALENCE ISOMERISATION → BENZENE +
BENZO-CYCLOOCTATETRAENE (90° THF).

FIGURE 27 : 18-ANNULENE PROPERTIES.

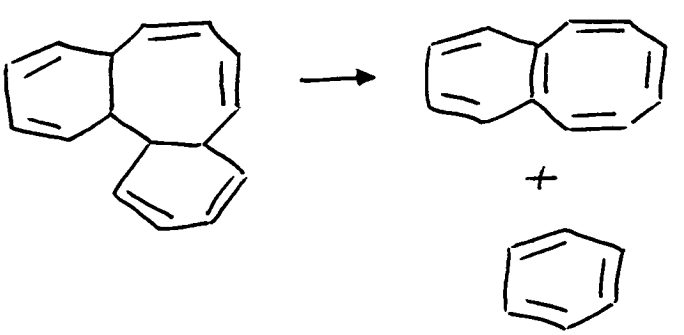
structure in solution,[80] but that a bond-alternated structure of D_{3h} symmetry gave a better fit with the observed electronic spectrum, and that a small amount of nonplanarity (D_3) might further improve the agreement. The "molecular mechanics" method of Allinger yields the D_3 structure as the most stable one, with a mean nonplanarity of 0.102 Å for the carbons[71] (valence parameters shown in Table 19), and this structure does lead to an even better fit between calculated and experimental UV spectra. The particular puckering arrangement of the carbons in the D_3 structure permits the inner hydrogens to move ± 0.568 Å out of the plane, and increases the distance between adjacent ones from 1.99 Å in the D_{3h} structure to 2.22 Å in the D_3. The problem of the discrepancy between this calculated geometry of the isolated molecule and the X-ray crystal structure remains to be explained. It is conceivable that crystal lattice forces (intermolecular van der Waals forces) could be sufficient to flatten out 18-annulene (cf. biphenyl where 20 kJ mol^{-1} is energy difference involved); it is calculated that 38 kJ mol^{-1} is required for this, and this amount might be available. Once the molecule is flat, the D_{3h} structure is calculated to be favoured over the D_{6h} by only 5 kJ mol^{-1}. Thus, it may be that the molecular structures in the gas phase and in the crystal are quite different.

To add to the confusion, in addition to the molecular geometry the stabilisation energy of 18-annulene has been estimated by various methods, both experimental and theoretical, and drastic discrepancies have occurred. Less refined theoretical models based on the topology of the planar conformation (cttcttctt) without structural details, usually yield a small positive value of the resonance energy, somewhat less than that calculated for benzene,[81] in keeping, qualitatively with a (4n+2) system. A recent examination of 18-annulene by the widely used more refined MINDO semi-empirical procedure yielded a value of 32 kJ mol^{-1} for the aromatic

Figure 28. Isodynamic Conformers
(a) OF 18-ANNULENE.



(b) [18]-ANNULENE $\xrightarrow[\text{THF}]{90^\circ}$



resonance energy.[81] This very small value refers to the stabilisation of the lowest energy form of 18-annulene. The molecular geometry was optimised by minimisation with respect to all 102 internal coordinates, and the planar D_{3h} form with alternating bond lengths was found to be most stable (structural details in Table 19). The results using the most refined all-valence-electron SCF MO method showed very close agreement with those from the simpler SCF π-approximation with σ,π separation assumed. The lowest energy D_{6h} form was calculated to lie 110 kJ mol^{-1} above the lowest D_{3h} form; the D_3 was estimated to lie above the D_{3h} by only a small amount, again in contrast to the results of the molecular mechanics calculations.

Experimental measurements have not led to a definitive value of the stabilisation energy. Qualitatively, the observed NMR spectrum and the thermal stability (greater than 4n systems of comparable size) have led to a substantial stabilisation energy associated with π-electron delocalisation being expected. Quantitative evaluation of the stabilisation energy has given the values of 420 kJ mol^{-1} (cf. benzene 150) via heat of combustion measurements, and 85 kJ mol^{-1} from NMR data.[82] The latter figure is derived from the interpretation of the activation enthalpy of conformational mobility (inside-outside proton exchange) derived from NMR spectra; three isodynamic conformers are considered to exist (Figure 28). To try to resolve the problem, a recent study has used reliable thermochemical measurements to furnish data, and a definition of stabilisation energy similar to the one used in this work in terms of reference structure used.[82] Thus, the stabilisation energy is defined as the difference between the enthalpy of formation (g at 298K) of a hypothetical planar Kekulé structure (alternating periphery) and that of 18-annulene itself (i.e. the heat of a special "reaction"). In practice, an appropriate sequence of reactions is required to yield the enthalpy of formation (ΔH^{o}_{f}) of 18-annulene, starting from 18-annulene and

producing molecules whose ΔH°_{f}'s are known or easily calculated, and measuring heats of reaction steps. The particular reaction chosen was the thermal rearrangement of 18-annulene (in solution above 90°C - Figure 28). Thermolysis using a differential scanning calorimeter to yield ΔH°_{f} was performed, and a value of 520 $\pm$ 20 kJ mol^{-1} was obtained. ΔH°_{f} for Kekulé 18-annulene was estimated at 675 $\pm$ 2 kJ mol^{-1}, by group-increment methods. Thus, the stabilisation energy of 18-annulene was calculated to be 155 $\pm$ 25 kJ mol^{-1}, which is very close to the benzene value. The much higher estimate from heat of combustion measurements is reckoned to be suspect because the great sensitivity of 18-annulene to oxygen is likely to have caused oxidation prior to combustion. The low estimate from examining conformational mobility may have resulted because it was assumed that all resonance stabilisation is lost in the transition state which contains "perpendicular" trans double bonds, thus reducing cyclic delocalisation; however, in actual fact, the transition state has three planar transoid butadiene fragments connected by"isolated" cis bonds so that if significant stabilisation of this species is assumed, the original resonance energy estimate can be adjusted to a value similar to the recently derived one.

In conclusion, the conflict at present is that experimental observations (NMR, UV, IR, PE) are best interpreted quantitatively using a D_{6h} π-delocalised structure for 18-annulene, whereas theoretical calculations regard this structure as much less favourable than the D_{3h} or non-planar D_{3} forms; estimates of the stabilisation energy also vary widely. Thus, non-empirical calculations should prove to be particularly beneficial in this area. However, it is only recently that such a large species has been amenable, and only one exploratory calculation has been performed and reported in this work. The structure chosen here is the simplest, namely a D_{6h} regular planar geometry using the standard benzene values for the C-C and C-H

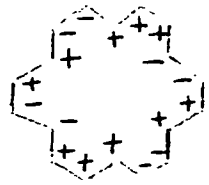
bond lengths; the resulting calculation, using the standard minimal basis, is still the largest single computation dealt with in this work.

The total energy quantities computed for the regular planar 18-annulene structure are presented in Table 20. It is difficult to estimate the total energy of the optimum planar conformation. In the smaller relatively unstrained annulene species, the energy lowering between the arbitrarily selected regular, "benzene" geometry and the optimum regular one, or the energy difference between regular and classical geometries, is rather small (about 30 kJ mol^{-1}). However, in the severely strained species, such as planar 14-annulene, the corresponding energy differences tend to be much larger (about 200 kJ mol^{-1}); this can be attributed to the effect of the internal hydrogens, with the total energy being very sensitive to the disposition of these so that relatively small shifts can lead to significant energy changes. Thus, in 18-annulene, as with the other species of similar conformation, geometries rather different in terms of the periphery (e.g. regular and alternating forms) have very similar internal hydrogen dispositions. As expected qualitatively, the internal hydrogens of 18-annulene do not appear to dominate geometrical considerations, so that optimisation of the planar form is likely to lead to an improvement in energy in keeping with that found in species without strained internal hydrogens. This means that the optimum planar form is likely to have a small stabilisation energy, unlikely to attain the benzene value, and almost certainly not to begin to approach a value three times that of benzene.

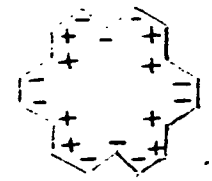
It is unfortunate that at present no published photoelectron spectrum of 18-annulene is available. However, unlike several other annulene species, it is probable that it will appear in the near future. In the absence of direct, rigorous geometry optimisation, the comparison of the observed P.E.

spectrum and computed orbital energies of certain chosen structure(s) could give some insight into the details of the geometrical structure, as is done with other species in this work. For the regular planar form, the calculated orbital energies are given in Table 21. As is usual, the π-orbital energies are distributed roughly in the Hückel pattern, even although this refers to a polygonal structure. A significant feature is the calculated first ionisation potential which is exceptionally low (applying Koopmans' approximation). From the forms of the π-orbitals (Figure 29), it is unlikely that the orbital energy of thehighest occupied degenerate pair will change greatly on varying the molecular geometry, e.g. to alternating D_{3h} form or slightly non-planar D_3 form (degeneracy will be removed), so that the calculated first I.P. of the optimum structure is likely to be very low. In addition, the large gap (about 4 eV) to the next I.P. is likely to remain (almost degenerate pair of π-orbitals). Although the value of the orbital energy of the lowest unoccupied degenerate pair of π-orbitals is not rigorously defined, there is an indication that these are low-lying MO's. Thus, without considering details of kinetics, it is expected that positive and negative ion formation in 18-annulene is much easier than in benzene; similarly, lower-lying excited electronic states of 18-annulene are expected (U.V. spectrum indicates the existence of excited states only a few eV above the ground state).

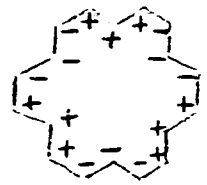
The three highest energy σ-MO's, lying in amongst the π-type, retain the "internal hydrogen" character observed in the smaller annulenes, but they are much more stable than those of the planar forms of the smaller species (non-planar 14-annulene, on optimisation, was yielding values approaching those of 18-annulene). There is very little distinction between internal and external hydrogens in these orbitals. The orbital energies of these characteristic orbitals is now only about 0.5 eV higher than the next



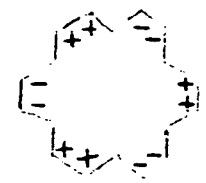
63 $e_{2u}(\pi)$



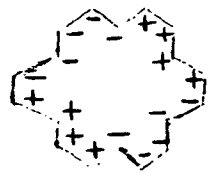
62 $e_{2u}(\pi)$



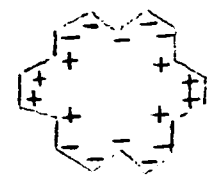
61 $e_{2g}(\pi)$



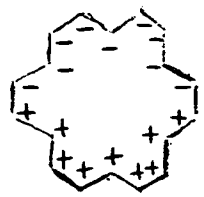
60 $e_{1g}(\pi)$



59 $e_{2u}(\pi)$



58 $e_{2u}(\pi)$



54 $e_{1g}(\pi)$

53 $e_{1g}(\pi)$

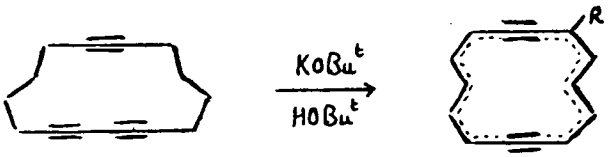
51 $a_{2u}(\pi)$

FIGURE 29 : FORMS OF 18-ANNULENE π-MO'S.

innermost σ-type, the start of the more typical ones (this fourth highest σ-MO is in fact nodal at the internal hydrogen). The computed total MO populations (Table 22) confirm the lack of distinct internal hydrogens; the values are effectively identical to those of any of the polygonal annulene structures (or any typical hydrocarbon, even), which have only "external" hydrogens.

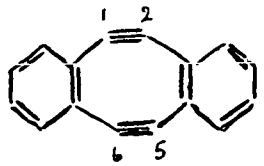
D. Calculations on Some Even-Membered Dehydroannulenes.

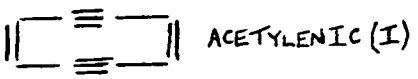
Dehydroannulenes are annulenes in which one or more of the double bonds have been replaced by acetylenic groups, i.e. sp-hybridised acetylenic carbon atoms are incorporated in the aromatic systems in place of sp^2-carbon atoms. The acetylenic carbon atoms placed in an annulene system contribute p-electrons in orbitals perpendicular to the molecular plane and also in orbitals in the plane towards the formation of a closed shell species. The first known dehydroannulene was 1,8-didehydro [14] annulene (Figure 30), prepared and extensively studied by F. Sondheimer and co-workers[83]. This classic example illustrates that the insertion of an acetylenic linkage in an annulene system increases the conformational stability of the resulting dehydroannulene system owing to the linear feature of the acetylenic bond. This situation is reflected in the NMR spectra of dehydroannulenes; the coalescence temperature of such spectra has been found, in general, to be much higher than that of the corresponding annulenes[84]. Consequently, dehydroannulenes have been regarded as suitable compounds for studies of aromaticity of macrocyclic conjugated systems; these completely conjugated monocarbocyclic polyenynes are involved as intermediates in Sondheimer's general method of annulene synthesis mentioned above, and, in this way, dehydroannulenes of all ring sizes from the 12- to the 30-membered inclusive have now been synthesised[56]. The dehydroannulenes are also of theoretical interest, since the criteria for aromaticity discussed for annulenes apply also to the dehydro-compounds as far as the out-of-plane π-electrons are concerned. Experimentally, dehydroannulene species have provided much more information than the parent annulenes. At present, the test of theoretical predictions regarding aromaticity of macrocyclic species has been performed in NMR experiments. The NMR spectra of most dehydroannulenes are essentially temperature independent, as the exchange of protons is much more inhibited than in the parent annulenes, so that, in this respect, the dehydroannulenes are preferable for the study of aromaticity in



(R=H)

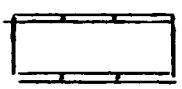
FIGURE 30 : SYNTHESIS OF 1,8-DIDEHYDRO[14]ANNULENE.





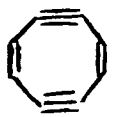
ACETYLENIC (I)

RECTANGULAR GEOMETRIES.



CUMULENIC (II)

$r(C\equiv C) = 1.205\ Å$
$r(C=C) = 1.340\ Å$
$r(C-H) = 1.090\ Å$



ACETYLENIC (III)

OCTAGONAL GEOMETRIES.

CUMULENIC (IV)

NON-CLASSICAL GEOMETRY (V)

$r(C=C) = 1.398Å$

FIGURE 31: CYCLOOCTADIENEDIYNE STRUCTURES.

conjugated macrocyclic systems. The main reason for the relative lack of proton exchange is presumably that moving the inner protons to "outer" positions by rotating of trans double bonds leads to non-equivalent conformers which are energetically less favoured. This is confirmed by the fact that the NMR spectra of certain dehydroannulenes, in which rotation does lead to equivalent structures, are in fact temperature dependent.

(a) Dehydro-[8]-Annulene.

The first Dehydroannulene considered in this work is not actually one of the traditional, strain-free species which have arisen quite often in the chemical literature. Of the many types of dehydroannulene species which can be formally generated from the parent annulenes, those which are synthetically accessible tend to be rather large. However, recently, in a study of 8-membered ring conjugated compounds, there has been reported the synthesis of dibenzo[a,e]-5,6,11,12-tetradehydrocyclo-octene (or sym-dibenzodidehydro[8]annulene, or sym-dibenzo-1,5-cyclo-octadiene-3,7-diyne)[85], a relatively stable species (Figure 31). This compound contrasts with the dibenzo-derivative of the parent cyclooctatetraene, which is a reasonably well-known compound. The cyclooctadienediyne derivative has been described as a presumably planar conjugated 8-membered ring compound, its NMR spectrum indicating a "paratropic" species. It is also reported that the parent, unsub-stituted compound might reasonably be expected to be relatively stable, but, unfortunately, so far it has not been synthesised. In moving towards the goal of this synthesis, very recently, the mono-substituted compound, 5,6,9,10-tetradehydrobenzocyclooctene has been prepared and described as the simplest known planar neutral conjugated 8-membered carbocycle, although it is very unstable[86]. Thus, this interesting, simple 8-membered dehydroannulene has been subject to some preliminary non-empirical calculations, in the hope that some experimental observations of interest may be reported in the near future. Experimentally, the dibenzo-compound has been found to be a relatively stable crystalline species, and its electronic spectrum has been interpreted as that of a highly conjugated compound. The contrast with the properties of the dibenzo-derivative of cyclooctatetraene might indicate that benzannelation is not a dominant effect and that the parent dehydroannulene could be a planar conjugated species.

The details of the molecular geometry of simple dehydroannulene species is of great interest; there is very little experimental information on this aspect. Rigorous geometry optimisation, by non-empirical calculation, of cyclooctadienediyne would be a quite reasonable task to perform, and would be very desirable if comparison of experimental and theoretical findings were to be done. However, at present, only a few selected geometries have been chosen, to provide information on the gross geometrical structure. As with other dehydroannulene species in this work, as well as a valence structure containing formal triple bonds (acetylenic species) one containing cumuleric units is considered. The dehydro[8]annulene species cannot be strain-free in a planar conformation; two extreme types of geometrical structure have been considered, one in which all the valence angle strain is formally in the ethylenic part of the system (or ends of the cumulenic units), and one in which this strain is solely around the acetylenic part (or centre of each cumulenic unit). Thus, two rectangular and two octagonal structures have been constructed (Figure 31). As far as bond-length is concerned, these four geometries are "classical", with standard values for the lengths as used in this work. Using the standard scaled minimal basis set, calculations at the above four geometries were performed, yielding the total energy quantities for the closed-shell, ground-state species given in Table 23 . There are very large variations in total energy, in chemical terms, and it seems likely that the octagonal acetylenic type of structure is the optimal one. The rectangular geometries are particularly unfavourable, as expected qualitatively, with the angle strain involved.

A further molecular geometry was considered along with the above. Bearing in mind the experimental information available for the dehydro [14] annulene considered below, a partly benzene-like structure was considered. This is most like the octagonal acetylenic one above, with formal triple bonds, but with the remaining parts of the structure of regular bond-length (benzene value). The total energy of this non-classical structure (Figure 31) is significantly higher than the corresponding classical acetylenic one. At this point, it might be concluded that the optimum structure is likely to be close to a classical acetylenic one, with the only significant deviation being the non-linearity of the acetylenic linkages, which have been considered to be reasonably "flexible" [87], although the distortion considered here is rather large. After the above calculations had been performed, the crystal structure of sym-dibenzo-1,5-cyclooctadiene-3,7-diyne was published [88]. It was concluded that the conjugated 8-membered ring in the crystalline state is substantially planar, and any small deviations observed probably result from intermolecular effects. The geometrical parameters derived are given in Table 24 . The cyclooctadienediyne ring has the "mixed" type of structure, with classical triple bonds but essentially delocalised (regular) ethylenic parts, although the latter have r(cc) somewhat longer than in benzene itself. The angles at the triply-bonded carbons are on average about 155°, which is very similar to the values observed in 1,5-cyclooctadiyne and cyclooctyne (gas phase); the distance between the triple bonds (about 2.6 Å) is also similar to that in cyclooctadiyne. The two outer benzene rings are planar within experimental uncertainty, and effectively coplanar with the 8-membered ring; this latter finding may be peculiar to the solid state, where "packing" inter actions dominate. In considering the direct experimental information available on the structure of the cyclooctadienediyne ring, it must be remembered that the data refers to the solid state and, more importantly here, that the benzannelation is likely

to have a very significant effect. Thus, another calculation was performed using a geometry based on the above experimental data (with hydrogens suitably placed - r(CH) = 1.10 Å and C-H bisecting exterior angle), and the computed total energy values are given in Table 23 . The total energy of this "experimental" structure is almost identical to that of the lowest energy structure above, the classical acetylenic one. Considering the results obtained from the selected geometries, it is obviously impossible to make precise conclusions on the equilibrium structure. However, examining the most realistic structures here, it does not appear as though the separation of the triple bonds is critical, so that interaction of in-plane π-type orbitals does not dominate the situation. Comparing the two "benzene-like" geometries, the very significant energy lowering in the "experimental" geometry is rather larger than found in other annulene species (e.g. cyclo-octatetraene) when varying C-C bond lengths by a relatively small amount. The other significant difference in the cyclooctadienediyne structures is in bond angle. This indicates that the valence angle values may be particularly critical in this strained species, and that both sets of angles deviate significantly from "natural" values.

In an attempt to further elucidate the details of the optimum geometry (e.g. classical, or partly regular), the characteristic orbital energy patterns computed for each of the above structures could be used, in conjunction with the observed photoelectron spectrum, if the latter were available. Unfortunately, this comparison cannot be made at present; recently, the He I P.E. spectrum of sym-dibenzodidehydro 8 annulene has been published[89], and that of the parent may yet be forthcoming. It was considered that the assignments of the resolved bands in the spectrum of the dibenzo-species was straightforward (Table 25), but that it was difficult to pin down the position of the π'-bands (in-plane π-type MO's) within a broad feature in the spectrum. The bonding and antibonding combinations of the in-plane

π'-orbitals were tentatively associated with the bands at 10.15 and 9.7 eV respectively. It was concluded that the substantial splitting due to the transannular "through-space" interaction of the π'-bonds is largely annihilated by further mixing with lower-lying σ-bonds. The overall form of the recorded spectrum is rather similar to that of another benzannelated dehydroannulene, and this may indicate that the benzene rings have a dominating effect and that the spectrum of the dibenzo-species is not particularly useful in considering the unsubstituted parent species.

All the structures of cyclooctadienediyne considered here are of the same symmetry (C_{2h}), and so the MO's of each are basically of the same form, and relative values of corresponding orbital energies computed (Table 26) can be understood in the usual way by examining orbital forms with bonding and anti-bonding contributions and the variation of these with relative nuclear positions. As usual, the π-type MO's are of particular interest. The forms have already been depicted in the case of cyclooctatetraene (Figure 10). Formally, the π-orbitals of the two 8-membered rings are identical in form, as the hydrogen atoms are irrelevant in this respect, and the octagonal structures of the dehydroannulene can be viewed as alternatives for cycloctatetraene itself. Thus, the various π-orbital energy values can be rationalised qualitatively by considering the various dispositions of the underlying carbon framework. The π-orbitals of the dehydroannulene forms are always stabilised compared to their cyclooctatetraene counterparts, and this can be attributed to the effects arising from the formally triply-bonded carbon AO interactions. The orbital energy patterns of the three most realistic geometries are rather similar (as in regular and alternating cyclooctatetraene). One distinguishing feature, however, is the significant stabilisation of the highest occupied π-MO in the classical structure compared to that of either regular one (by about 1 eV), as expected from the orbital form. Whatever the actual calculated equilibrium

structure is, it is likely that the predicted first I.P. will be relatively high for a planar species of this size. In addition, a further aid in comparison with an observed P.E. spectrum would be provided by the in-plane, pseudo-π-MO's, the energies of which vary significantly in the three structures. In particular, rather than absolute values, the energy split between the symmetric and anti-symmetric combinations of in-plane π-orbitals is of interest. The largest split is calculated to occur in the "experimental" structure (about 1 eV), which seems reasonable as the distance across the ring between acetylenic units is smallest at this geometry. In the constructed benzene structure, the split is effectively zero. However, in the classical structure, with acetylenic units furthest apart, the split is about 0.5 eV; this shows that distance between interacting units is not the only influence, but that interaction with AO's on the doubly-bonded carbons (and hydrogens) is an important factor - the splitting is not simply the result of interaction of two pure in-plane π-bonds, and the mid-point of the split is at an energy slightly above the "pure" acetylene π' value. In the other distorted dehydroannulene structures, the corresponding orbitals are not really predominantly π'-type and interaction across the ring is not the only main effect. This pair of orbitals constitute the only high-lying σ-MO's, and it is likely that they lie above all the π-type except one in the equilibrium structure.

In all the structures for cyclooctadienediyne considered, the orbital energy of the lowest unoccupied MO (π) is effectively zero. It seems that there is approximate, accidental degeneracy of the pair of π-MO's (HOMO-LUMO), which are exactly degenerate by symmetry in regular cyclooctatetraene, and split significantly in alternating cyclooctatetraene. Thus, even in the rather alternating dehydroannulene structure, there is a relatively low-lying triplet species (Table 23 gives total energy values for the various geometries). This effect is likely to persist even in the equilibrium structure, whether it is classical or partly regular. Thus, unlike the

situation in the annulene species with degeneracy by symmetry (4n "regular" species), the HOMO is relatively stable so that the HOMO-LUMO energy difference is relatively large.

The total AO populations computed at the various geometries are given in Table 27 . The values are effectively the same in all the structures. The hydrogens and associated carbons have total populations which are typical of ethylenic hydrocarbons. However, the formally sp-hybridised carbons have net population of zero, in contrast to the situation in open-chain acetylenes or cumulenes (net value of -0.2 to -0.3, typically). The electron distribution, according to the Mulliken analysis, of the dehydroannulene is similar to that of butenyne (reasonably uniform along carbon skeleton), but very different from that of butatriene (internal carbons markedly "negative"). Thus, even if the geometry is cumulenic, the electron distribution, and resulting chemical behaviour, is unlikely to be very similar to that of smaller cumulenes.

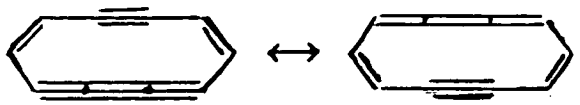
(b) Dehydro-[10]-Annulene

There has been longstanding experimental and theoretical interest in monocyclic 10 π-electron molecular systems[90]; those of greatest interest have been the all-carbon systems which have been most amenable to theoretical calculation. 10-Annulene species are considered elsewhere in this work, where difficulties which arise in the treatment of these, both experimental and theoretical, is mentioned. It was first suggested thirty years ago that the van der Waal's repulsions and angle strain which cause 10-annulenes to be regarded as non-aromatic species might, in principle, be reduced by replacing ethylene linkages with acetylenes[91]; it was proposed that 1,6-didehydro [10]annulene (Figure 32)might be a simple, aromatic compound. However, various unsuccessful attempts to prepare this species have been reported, and one is still in progress[92]. It has been concluded, qualitatively, that this dehydroannulene suffers from repulsion between the electrons of the two in-plane π-bonds, and this dominates the advantageous lack of any geometrical strain, so that the instability of the molecule can be understood. Formally, 1,6-didehydro[10]annulene is the smallest member of a series of didehydroannulenes which contain cumulenic and acetylenic linkages. In view of the difference in physical and chemical properties between such linkages, the preparation of dehydroannulenes containing cumulenic linkages has been of particular interest[91]; didehydro- and tetradehydro-annulene species, larger than the first member, the 10-membered ring compound, have been synthesised and studied. Members of these series of compounds cannot be represented by classical structural formulae which are either purely (acetylenic and ethylenic) or (cumulenic and ethylenic). Thus, as shown for didehydro[10]annulene in Figure 32, it is possible to depict the molecule as the resonance hybrid of two Kekulé-type structures, yielding a more adequate description in a symmetrical structure, with the formal

sp-hybridised carbon atoms of cumulenic linkage equivalent to those of acetylenic linkage.

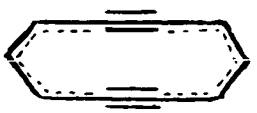
As there is no structural information on 1,6-didehydro[10]annulene, geometries have been constructed for this species. Two planar structures, differing very significantly in detail, have been considered, and non-empirical calculations using the standard minimal basis set have been performed. The molecular geometries are depicted in Figure 33 ; one is a "classical", unsymmetrical type of structure with standard bond parameters as used in this work, and the other is non-classical and symmetrical (based on the related didehydro[14]annulene species considered below, for which experimental data is available). The calculated total energy quantities are given in Table 28 . The total energy difference between the two forms is very small (barely of chemical significance). For two such widely different geometries, this could be an indication that the total energy of the species (lowest closed-shell state) is insensitive to geometry variation, although this is obviously uncertain at present. However, it does seem more likely that the calculated resonance energy (effectively zero for the two structures) will be increased by a significant amount on geometry optimisation, so that the benzene reference value (about 200 kJ mol^{-1}) will not be approached.

It so happens that the non-classical structure is of lower energy. As usual, it may be that the calculated optimum geometry does not follow the trend from considering isolated, extreme geometries. However, the calculated orbital energy spectra show significant variations so that, if experimental P.E. data were available, it might be possible to progress towards the equilibrium structure by comparing orbital energy data. The computed orbital energies for the two structures are listed in Table 29 . Overall, the two patterns are very similar, with the energies of correlating orbitals approximately the same (within 0.5 eV). The main exceptions occur at the low orbital energy end (in magnitude), the region experimentally better



(I)

FIGURE 32 : 1,6-DIDEHYDRO[10]ANNULENE VALENCE TAUTOMERS.

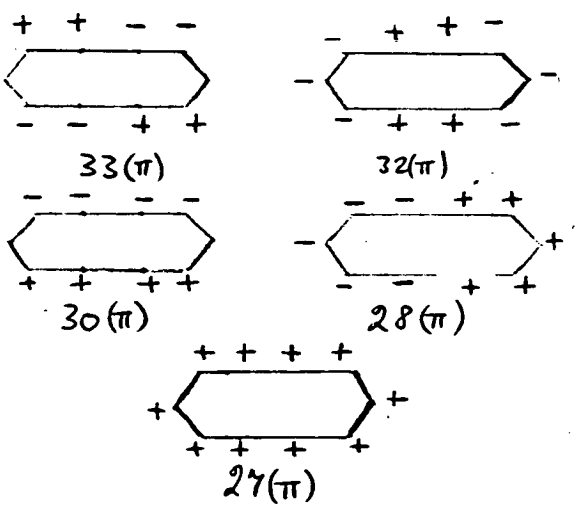


(II)

$r(c=c) = 1.398\,Å$, $r(c\equiv c) = 1.205\,Å$,
$r(c-H) = 1.090\,Å$.

FIG. 32 $\rightarrow r(c=c) = 1.340\,Å$

FIGURE 33 : MOLECULAR GEOMETRIES OF 1,6-DIDEHYDRO[10]ANNULENE.



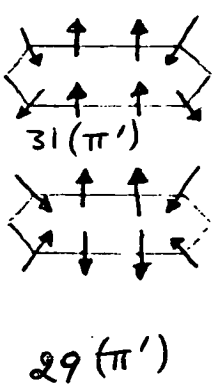


FIGURE 34 : FORMS OF π-TYPE MO's OF DEHYDRO[10]ANNULENE. π'-TYPE

resolved. In the "classical", cumulenic-acetylenic structure, there is calculated to be an almost exactly degenerate (accidental) pair of MO's as the HOMO's. One is a true π-type MO, and the other is basically the antisymmetric combination of the in-plane pseudo-π MO's of the acetylenic and cumulenic units. In contrast, the HOMO of the non-classical structure is calculated to be the true π-type one of the pair, and there is a gap of about 1.5 eV between it and the in-plane π'-type, which is almost degenerate with the next π-type. The forms of the π- and π'-type MO's are shown schematically in Figure 34 . The HOMO of the non-classical structure is really the only MO which is less stable than its counterpart in the classical structure; this is understandable from the consideration of the form of the MO. It is notable that each calculated first I.P. is very similar to that of the corresponding dehydro[8]annulene species (classical and non-classical),the value being relatively high for the type of species. The second highest occupied π-MO has essentially the same energy in the two structures, whereas the other three follow the trend of being slightly more stable in the non-classical structure. The other two MO's of most interest are the two σ-type which are basically the symmetric and anti-symmetric combinations of the in-plane pseudo-π MO's of the constituent acetylenic/cumulenic units. The energy splitting of the pair is effectively the same in both calculations, and is very large (3.8 eV), as expected, qualitatively, from the proximity of the interacting MO's of the units. In the more symmetrical non-classical form, the average position of the split pair is slightly higher in energy than the basic acetylene π'-MO. In the dehydro[10]annulene, the two MO's have some significant contributions from the in-plane AO's of centres other than the sp-carbons, so that the interaction of the triple bonds is not "pure"; further, smaller interaction of the pair of π'-type MO's with σ-type MO's of the other constituent units of the molecule can be invoked to yield

the overall effect. In the non-classical structure, although the magnitude of the splitting is the same as in the classical one, it is likely that the major interaction between the in-plane π'-type MO's is smaller; in the less symmetrical set-up the two reference levels are a normal acetylene one and an expanded acetylene one (corresponding to central cumulenic bond), which is of somewhat higher energy, so that although the distance involved is similar, there is a discrepancy in energy of the interacting levels.

The indication that the resonance energy of didehydro[10]annulene is small is consistent with it not being a species similar to benzene, and this has been attributed qualitatively to unfavourable electronic repulsion effects between the in-plane π'-MO's of this otherwise strain-free species. As with the nuclear repulsion effects considered to be dominant in planar 14-annulene, the MO model used here cannot be analysed in exactly these terms. However, the results of the non-empirical calculations can be examined in this light. The interaction between the in-plane π'-MO's is so large that the energy of the anti-symmetric combination is unusually low in magnitude for a σ-MO. The one-electron orbital energies of both π'-type MO's are significantly more negative than those of the π-MO's, and several of the other σ-MO's, so that their orbital energy values indicate that there are much larger electronic repulsion effects associated with these two MO's than with any others. However, the electron energies of these two are still more negative than those of several of the others. It must be stressed that the electron repulsion effects involve the interactions of each of the two MO's with all other occupied MO's; the electronic charge of the two π'-type MO's is relatively unfavourably distributed in relation to the distributions of all the other MO's. In contrast, MO's such as the totally symmetric π-type, or the σ-type which is basically C-H bonding of alternating character round the ring, have relatively low one-electron energy values,

but relatively high orbital energies as a result of lesser electron repulsion effects, as expected from the MO forms. If the total electronic energy of the species is lower than "expected" for a potentially aromatic system, and electronic repulsion is considered responsible, then it is an overall MO effect.

For the two structures considered, the total AO populations computed are listed in Table 30. The hydrogen AO populations in both are effectively the typical hydrocarbon value (ethylenic). The formally sp^2-hybridised carbons also have essentially the hydrocarbon value so that the C-H dipoles are typical. The sp-carbons, whether acetylenic or cumulenic as defined by geometrical parameters, have a total AO population of 6.0, i.e. no net charge, in contrast to the situation in acyclic acetylenes and cumulenes, where they have a significant net negative charge, particularly in the latter. The charge distribution along the multiply-bonded "sides" of the molecule is very different from that of a cumulene, even when the atom positions are chosen in accord with such a structure.

For the non-classical structure, a few additional calculations were performed; these made use of the unscaled minimal basis set, and the latter with some partial scaling. The calculated total energy quantities, orbital energies, and AO populations are given in Table 31.
Some indication of the variability of actual numerical values with basis set, when the latter is of limited form, is given for this relatively large molecular species, and illustrates some basis set effects to be borne in mind when considering the results of various calculations and the trends deduced from them.

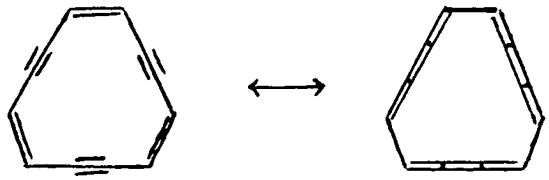
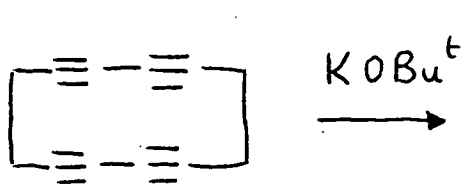


FIGURE 35 : 1,5,9 - TRIDEHYDRO [12] ANNULENE.



BRICK-RED,
m.p. 95°C.

STABLE IN ABSENCE OF O_2;
EXTREMELY REACTIVE WITH O_2.
NMR : $\tau = 5.58$
(SINGLET : -80° → 37°).

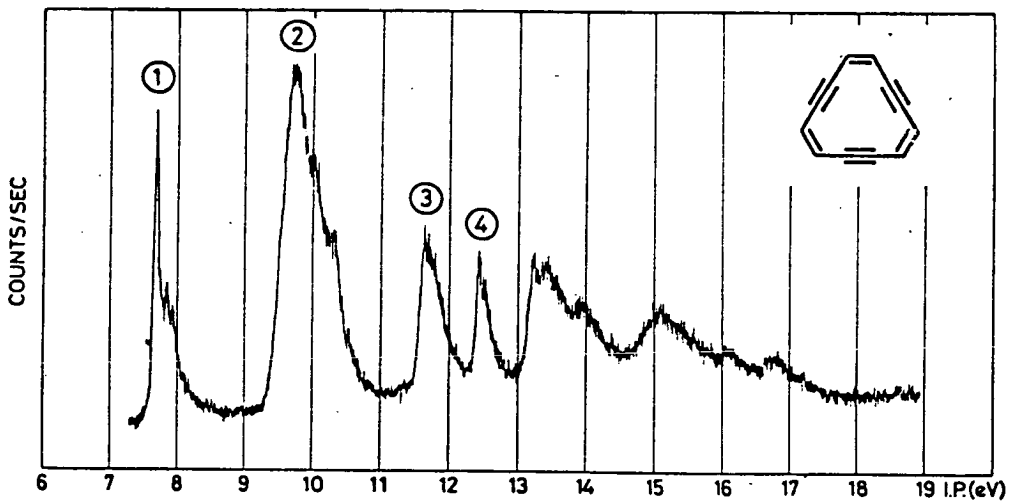
FIGURE 36 : SYNTHESIS OF TRIDEHYDRO [12] ANNULENE.

(c) Dehydro-[12]-annulene

The dehydro-12-annulene species considered here is a member of the series starting with cyclooctadienediyne, with succeeding members formed by incorporating a further double-bond triple-bond pair; thus, 1,5,9-cyclododecatriene-3,7,11-triyne, or 1,5,9-tridehydro[12]annulene is obtained (Figure 35). This species, unlike any other conceivable homologues, is completely strain-free, geometrically speaking. This 4n annulene species has been considered to be of particular importance because of the enforced planarity caused by the incorporation of the triple bonds. It has also been considered that the classically alternating ethylenic-acetylenic structure is appropriate for the equilibrium one, as π-bond localisation is enforced by the in-plane π'-bonds providing a bias favouring that valence structure with the greater number of formal triple bonds. The use of Benson's group additivity tables[93] leads to the prediction that the isomerisation reaction but-1-en-3 yne → 1,2,3-butatriene is endothermic by approximately 46 kJ mol^{-1}; thus, the acetylenic valence structure of tridehydro[12]annulene is predicted to be favoured over the cumulenic one (Figure 35) by approximately 138 kJ mol^{-1}, making it doubtful that the latter corresponds to a minimum on the potential surface. It has been concluded that the pernicious effects of cyclic conjugation in [4n]-annulenes predicted by MO theory are strongly tempered by bond alternation in this tridehydro[12]annulene, and that highly sensitive probes such as NMR chemical shifts must be used in order to reveal any special attributes of the 4n topology in the electronic ground state[93] . The open-shell character of such systems as this dehydroannulene does manifest itself by the ease of reduction to form stable radical anions and dianions, and by the long-wavelength electronic absorption spectra[93] (synthetic route to the tridehydro-annulene and available experimental information, which as usual is essentially physical rather than chemical, are summarised in Figure 36).

Very recently, the photoelectron spectra of 1,5,9-tridehydro[12]annulene and related compounds have been reported[89]. In the analysis of these, the tridehydro-annulene was taken to have an essentially planar equilibrium structure with alternating bond lengths. This presumption is established by X-ray structural data in the case of the tribenzo-derivative[94]; (cf. cyclooctadienediyne in (a)), and the vibrational spectra have been interpreted as indicating the presence of triple bonds. There is no angle strain in the molecule, and it has been concluded that the transannular interactions (distance between ortho-carbon atoms is 2.86 Å in tribenzo-species) are insufficient to produce a significant distortion of the triple bonds. The He I spectrum of 1,5,9-tridehydro[12]annulene is shown in Figure 37. Down to 13 eV ionisation energy there are four well-separated features, one or more of these resulting from several overlapping bands. The spectrum has been interpreted on the basis of Koopmans' Approximation using the simple LCBO (Linear Combination of Bonding Orbitals) model which has yielded accurate predictions of π-ionisation potentials for a large body of planar polyenes and polyynes[89]. The vertical I.P.'s are listed in Table 32, along with the results of the LCBO, SPINDO, MINDO, PPP models and of an Xα-calculation[89]. Idealised, standard geometries were used. The qualitative agreement between the various models is not considered to be surprising as the π-orbitals are largely determined by symmetry. It is concluded that the bands contained in the four prominent features are associated with the complete set of six out-of-plane π- and three in-plane π'-orbitals, placing the σ-orbitals at energies below -13 eV. The narrow shapes of the third and fourth bands lend support to their assignment to states of π-symmetry. The first ionisation band is narrow and structured, its first apparent vibrational peak being by far the most intense, indicating that the geometry (bond length alternation)

Figure 37: P.E. SPECTRUM OF $C_{12}H_6$.



largely persists in the ground state of the radical cation, presumably due to the presence of the in-plane π'-bonds. The observed P.E. spectrum can be viewed in the light of the results of the non-empirical calculations of this work.

In the absence of any experimental structural data for the tridehydro-annulene, two geometries have been constructed in the usual way. One corresponds to a classical acetylenic structure, with standard bond parameters as used in this work (Figure 38), and the other to the alternative mesomeric form, a classical cumulenic structure. These two distinct valence structures differ very significantly in the details of geometrical structure. Thus, using the standard minimal basis set, the computed results of interest for the two assumed structures are given in Tables 33-35. The total energy of the acetylenic form (cyclododecatrienetriyne) is significantly lower than that of the cumulenic form. It so happens that, using selected standard geometries, the total energy difference between the non-classical acetylenic forms and classical acetylenic-cumulenic ones is very small in the didehydro-annulenes, whereas that between acetylenic and cumulenic forms here and in (a) is very large. However, it is impossible to make conclusions about equilibrium geometries (above structures are really extreme cases of plausible ones), although it may be that the tridehydro[12] annulene and didehydro[8]annulene energies are more susceptible to geometry variation. Each of the tridehydro[12]annulene structures has its own pattern of orbital energies, which are significantly different, so that it is possible to make some further comment on which structure is nearer the actual one by comparing the computed orbital energies with the observed P.E. spectrum. The correlation between I.P. values using the acetylenic structure data is really quite close (especially for such a constructed geometry), whereas that using the cumulenic structure data is not close at all. The forms of the π-type and π'-type orbitals, which are nearly all

the ones for which the experimental measurements are reasonably precise, are illustrated in Figure 39 . Without performing the actual large-scale calculations, it is possible to predict roughly the variation of these MO energies, in particular, with geometry (bond-length) variation. Thus, it may be possible to obtain an even better correlation by using an acetylenic geometry which is slightly less classical with respect to the sp^2-carbons (i.e.moving a little towards the non-classical type of structure of didehydro[14]annulene in (d)), lengthening the formal double bond and shortening the formal single bond. This can be viewed as moving towards the other extreme geometry, the cumulenic one, but with the retention of a triple-bond r(CC) between sp-carbons. This is reasonably consistent with the experimental data on the tribenzo-derivative, where the dehydro-annulene ring has a classical triple bond, shortened single bond (compared to butadiene), and lengthened double bond, although the latter is probably particularly affected by the attached benzene ring (cf. also cycloocta-dienediyne case in (a)).

As qualitatively expected, on symmetry grounds, this 4n cyclic conjugated hydrocarbon has a closed-shell singlet ground state. However, the HOMO is not the same one in the two structures. There is a possibility of accidental degeneracy in the pair of π-MO's which form the HOMO-LUMO pair; in both geometries considered here, the orbitals are not nearly degenerate, but there is a reversal of relative stability, as expected from the orbital forms (bonding and anti-bonding contributions). In the acetylenic structure the first I.P. is calculated to be reasonably high, whereas that of the cumulenic structure is about 1.3 eV lower, and not really consistent with experiment. The HOMO-LUMO gap is rather large in both cases, and this indicates that the triplet species and other singlets lie well above the ground state in both. The second group of I.P.'s, as observed, is predicted in each calculation to arise from ionisation from a degenerate pair of π-MO's and a degenerate pair (+ one) of π'-MO's. The π-MO's of the

ACETYLENIC (D_{3h})

$r(c \equiv c) = 1.205$ Å , $r(c=c) = 1.340$ Å
$r(c-c) = 1.500$ Å , $r(c-H) = 1.090$ Å
ALL ANGLES = 120°, 180°.

CUMULENIC (D_{3h})

$r(c=c) = 1.340$ Å , $r(c-c) = 1.500$ Å ,
$r(c-H) = 1.090$ Å ; ALL ANGLES = 120°, 180°.

FIGURE 38 : TRIDEHYDRO[12]ANNULENE GEOMETRICAL DATA.

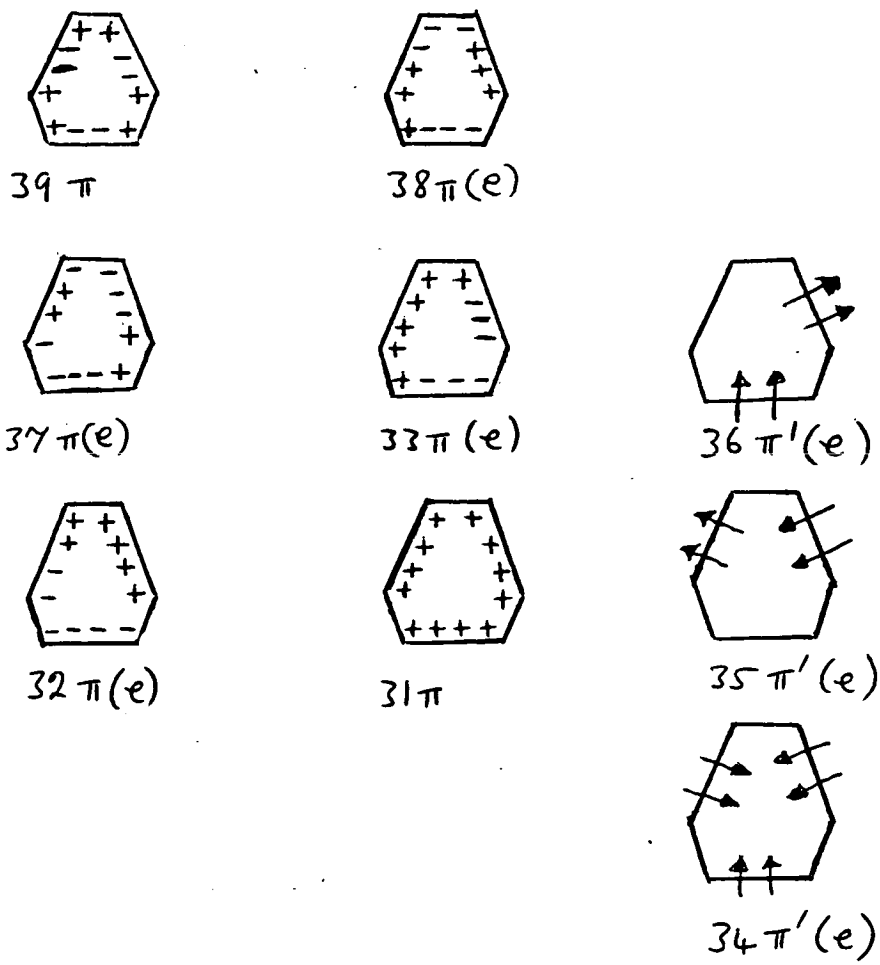


FIGURE 39 : FORMS OF π-MO's AND π'-MO's OF TRIDEHYDRO[12]ANNULENE.

acetylenic structure are slightly stabilised compared to those of the cumulenic, whereas the π'-MO's are significantly more stable in the acetylene case. Furthermore, the third π'-MO (a_1') is effectively of the same energy as the symmetry-equivalent pair in the cumulenic structure, and these three are slightly less stable than the π pair; in the acetylenic pair, the π'-MO's are split by about 0.65 eV, and all three are slightly more stable than the π pair. Thus, the five MO's of the acetylenic structure have energies which are more consistent with the observed values than those of the cumulenic form. As a result of overlap with the 2e" (π) band in the second feature of the observed spectrum, only an upper limit of $|\pm 0.6$ eV$|$ can be given with certainty for the difference between the ionisation energies of the in-plane π'-orbitals, $\varepsilon(5a_1)-\varepsilon(7e')$, although a value of about 0.3 eV is reckoned to be probable from the high-energy structure of the feature. In the acetylenic structure, it appears as though the interaction of the π'-orbitals is almost "pure", as the small calculated splitting, with the a_1 MO of lowest energy, is effectively centred around the isolated acetylene value; this is referred to as the "through-space" interaction. The calculated splitting in the cumulenic structure is zero, with the mean energy about 1.2 eV higher than the acetylene value. The energy of the individual interacting units of this structure is somewhat higher than the normal value, as the central cumulene bond can be regarded as an elongated triple bond. However, the π'-MO's lie even higher than this level (approximately 10.8 eV, estimated from ethylene values), and this indicates significant interaction with other σ-type orbitals (so-called "through-bond" interaction). The distances between the interacting π'-units are very similar in the two structures, so that the through-space interaction is expected to be similar in each, although the resultant orbital splittings are dissimilar. Of the semi-empirical calculations, in which individual components such as through-space and through-bond effects can be identified, only the SPINDO model results show close

agreement with those from the non-empirical calculations here; the calculated SPINDO overall split is -0.02 eV (a_1 energy lower than e') as a result of near cancellation of the two interactions, with the split from through-space alone being -0.54 eV. The so-called "through-bond" effect appears to depend mainly on the sp^2 part of the carbon framework. Completing the other main features of the spectrum, the remaining three π-MO's (e" and a_2"), which are of identical energies in the two calculations, are associated with the third and fourth bands, the narrow shapes of which lend support to their assignment to states of π-symmetry. The usual type of σ-orbitals all lie below the complete set of 6π- and 3π'-type of MO's, at energies below -13 eV; the cumulenic structure has rather a large gap between the most stable π-MO and least stable σ-MO's (e' and a_2' set), which are combinations of basically anti-bonding orbitals of the 1-12 bond etc. Thus, overall, the indications are that the tridehydro-annulene structure has a sp carbon-carbon bond which has formal triple-bond characteristics, although the remainder of the periphery may be less "classical", a cumulenic structure seems a less appropriate representation. The particular cumulenic geometry used here has three identical C=C bond lengths; shortening of the central bond, as observed in small open-chain systems, would probably lead to even worse agreement than found here. One obvious change, using such a slightly modified cumulenic structure, would be the near degeneracy of the HOMO and LUMO of the closed-shell ground-state species (the other orbitals most significantly effected would be the pair of σ-MO's which are the analogous, largely 2s AO combinations - MO's 18 and 19); this would lead to the existence of a low-lying triplet species (perhaps ground state). The P.E. spectrum of 1,5-didehydro[12]annulene (1,3,5,9-cyclododecatetraene-7,11-diyne), shown in Figure 37 , is practically identical with that of the tridehydro-annulene, so that the replacement of a triple bond with a double bond is a relatively minor perturbation, as far as orbital energies are concerned.

This is consistent with the structures of both dehydro-annulenes being largely classical.

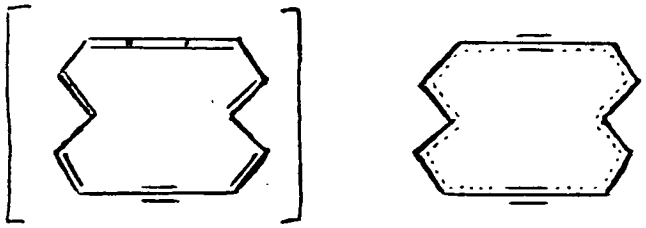
The computed AO populations for the two tridehydro-annulene structures are effectively the same, and follow the same trend as found with the other dehydro-annulenes; the C-H dipoles are typical of conjugated hydrocarbons, and the net population of the sp-carbons is zero. The electron distribution in each acetylenic (cumulenic) unit is markedly different from that in small open-chain systems.

(d) Dehydro-[14]-Annulene

Of all the annulene species which have now been synthesised and studied experimentally, 1,8-didehydro[14]annulene is one which occupies a special position. It was the first known member of the symmetrical didehydro-annulenes (cf. the 10-membered ring species in (b)); this dark red compound was prepared as shown in Figure 30, almost twenty years ago. The same criteria for aromaticity discussed for annulenes have been considered to apply to the dehydroannulenes, as far as the out-of-plane π-electrons are concerned; 1,8-didehydro[14]annulene was expected to be a relatively rigid planar 14 π-electron system, and therefore aromatic. In fact, it proved to be one of the "most aromatic" monocyclic non-benzenoid systems known. The relevant experimental findings are : it is relatively stable, undergoing electrophilic substitution reactions at the positions indicated in Figure 30: an X-ray crystallographic analysis has shown it to be planar and centro-symmetrical; the proton NMR spectrum indicates that it is strongly diatropic, the inner protons resonating at very high field (τ = 15.48) and the outer at low field (τ = 0.36, 1.46). Quite recently, Nakagawa has reported an efficient synthesis of such 1,8-didehydro[14]annulenes, and larger related didehydro-species [95] (all tend to have large, stabilising tertiary-butyl groups attached). All these species have been found to have essentially the same geometry, unlike the parent annulenes, so that it is possible to study experimentally the effect of increasing the value of n in aromatic (4n+2) π-electron systems keeping the geometry largely unchanged. So far, the proton NMR data indicates the progressive diminution of the diamagnetic ring current as n increases, although the ring current is still evident in the 26-membered ring compound.

No centrosymmetrical classical structure can be written for this dehydroannulene, and the most convenient structural representation is the one shown in Figure 40. Formally, the Kekule-type structure of this species contains cumulenic and acetylenic linkages, which differ widely in physical and chemical properties. However, the optimum symmetrical formula shows that the sp-hybridised carbon atoms of a cumulenic linkage can participate in the formation of an aromatic system in a manner equivalent to acetylenic carbon atoms. The results of NMR analyses, whereby "acetylene-cumulene" dehydroannulenes have been found to sustain larger ring currents than "acetylene" dehydroannulenes of similar ring size, have been interpreted as showing the decrease of bond-alternation in the former compounds. It is unfortunate that the photoelectron spectrum of this species has not been published, especially, as from its overall chemical behaviour, it would seem to be a reasonably straightforward exercise to obtain it. As far as this work is concerned, the spectrum would be particularly useful in providing a test of the adequacy of the type of calculation involved here for these annulenic species.

A single calculation has been performed on the didehydro[14]annulene, using the standard minimal basis set and a geometry based on the parameters from the X-ray crystal structure study (Figure 40); the experimental geometry was symmetrised, yielding a non-classical structure with a regular carbon skeleton (benzene-like) perturbed by the inclusion of two classical triple bonds. The computed quantities of interest are presented in Table 36 The calculated resonance energy is very substantial for this 14 π-electron species, and is about half as large again as the value for benzene. This estimate of the resonance stabilisation is a lower limit (presumably the geometry used is near optimal although not necessarily so), so that the very large resonance energy is consistent with the conclusion from experimental studies that the species is a clear-cut aromatic one. The calculated spectrum of orbital energies gives some points of interest. The first I.P., from a π-type orbital (π-MO's illustrated in Figure 41), is predicted to be rather low (6.8 eV calculated, expected to be somewhat larger than observed would be), and separated by about 1 eV from the next (another π-type ionisation). There is then a very large gap of about 3.5 eV before the next group of ionisations, which arise from four MO's which are practically degenerate, accidentally; there are two π-MO's, one of which is largely the symmetric combination of the two relatively distant acetylene units' π orbitals (orbital energy is effectively that of an isolated acetylene π-MO), and two π'-MO's. The latter show that there is very little interaction between the in-plane π-type orbitals of the acetylene units as the symmetric and anti-symmetric combinations are really degenerate, with an energy very slightly above the "pure" acetylene value. The anti-symmetric combination is actually calculated to be the more stable, and this is probably the result of the minor interaction with other lower-lying σ-orbitals, which is different in the two cases, as the π'-type MO's are not "pure". In contrast to the smaller dehydroannulene species, especially didehydro[10]annulene, there is no particularly high-lying σ-orbital. The remaining three π-MO's, separated by a further gap of about 2 eV,



$r(C \equiv C) = 1.208$
$r(C = C) = 1.398$
$r(C-H) = 1.085$
ANGLES = 120°, 180°.

FIGURE 40 : GEOMETRIC DATA FOR DEHYDRO [14] ANNULENE.

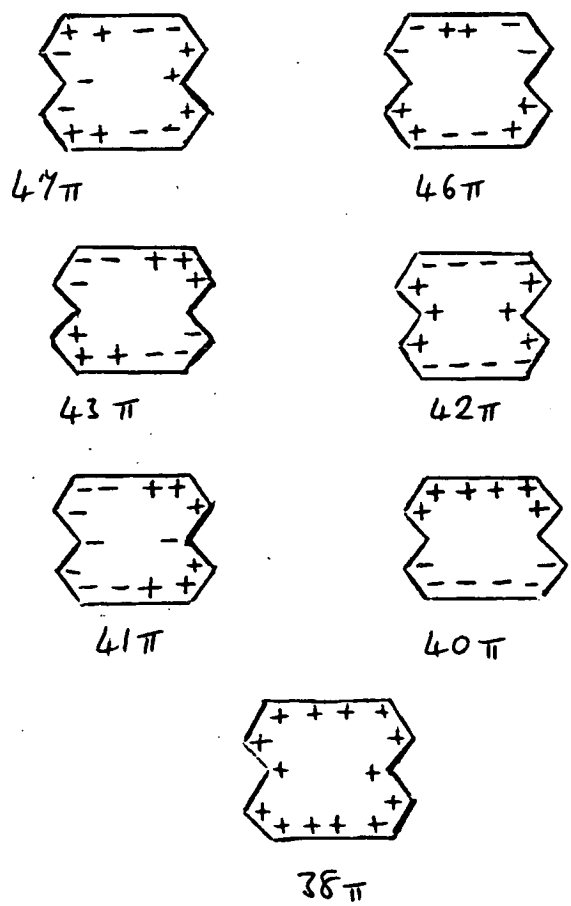


FIGURE 41 : FORMS OF π-MO'S OF DEHYDRO [14] ANNULENE.

complete a π-orbital pattern which is roughly similar to the Hückel one, with no symmetry grounds for this. Lying in among these most stable π-MO's are two σ-MO's of special character. This completely strain-free species does posses two "internal" hydrogens, and these σ-MO's are basically the symmetric and anti-symmetric combinations of internal C-H bonding orbitals. There is a splitting of about 0.5 eV, which indicates little interactions across the ring, although again the orbitals are not pure; the energies of the two orbitals are probably somewhat higher than that of the hypothetical isolated C-H bonding orbital, indicating interaction with other σ-orbitals. The orbitals are much more stabilised than the counterparts in the strained annulene species.

The population analysis illustrates that the internal hydrogens are not really distinct from the outer ones, confirming that they are not unfavourably placed. The carbon populations show the same trends observed in the other dehydroannulenes. The ethylenic carbons have effectively the same total population, with a value very similar to the typical hydrocarbon (conjugated) value. With the hydrogen populations also in the usual value range, the C-H dipoles are typical. The perturbing acetylenic carbons have a net population effectively zero. The small net excess negative charge on the carbon framework is thus reasonably evenly distributed along the regular ethylenic parts, drawn away from the triple-bond regions. The presence of the formal triple bonds does not disturb the typical C-H dipoles (as found in benzene, for example).

E. Calculations on Some Bridged Annulenes.

As mentioned earlier in this Chapter, the Hückel rule was originally formulated for monocyclic systems, the periphery of such a species comprising only one type of centre, with carbon the natural choice on physical grounds. Extensions to monocycles containing heteroatoms (e.g. N,O,S,P) seem to be very reasonable, and such considerations have been widespread in the chemical literature. Further, the rule has been widely applied to the periphery of polycyclic systems; in this way, cross-linking of the periphery is regarded as a minor perturbation. The formal generalisation of the Hückel rule[96] states that the presence of (4n+2)-rings will stabilise and the presence of (4m)-rings will destabilise the conjugated hydrocarbon. Recently, primitive Hückel MO theory has been shown to be a particular application of mathematical graph theory. With the underlying physical assumption that the ground state (thermodynamic) stability of conjugated systems is related to their π-electron energy, E^{π}, the above rule actually means that every (4n+2)-membered ring contained in the conjugated molecule has an increasing (positive) contribution to E^{π}, whilst the (4n)-rings have a decreasing (negative) contribution to E^{π}. Thus, E^{π} is a very important collective molecular parameter, which may be obtained either from direct calculations or closely approximated by means of topological factors reflecting the details of molecular structure (e.g. number of atoms, number of bonds, number and quality of rings, algebraic structure count). Such mathematical methods have led to a formal proof of the original Hückel rule.[96] However, physical reality is rather severely approximated; consequently, the introduction of heteratoms is one particular source of problems with this theoretical method. Thus, the approach of non-empirical calculations,

with a sounder physical basis, should prove useful in this context. Following on from the consideration of annulenes the particular polycyclic species considered below can be regarded as "bridged" annulenes. Owing to geometrical factors (ring "strain"), many of the annulenes, inherently involved with Hückel's rule and of a size reasonably handled computationally, adopt non-planar conformations. The constraint of planarity involved with theoretical considerations can be satisfied by studying the related dehydroannulenes, as in the previous section. In addition, the configurations of the usual representations of the parent annulenes can be retained, and the planarity constraint satisfied, by incorporating an internal atom(s) to replace the troublesome internal hydrogen atoms. Some examples related to the annulenes studied above are now considered.

Species with a Single Bridgehead Atom: Cycl[3,3,3]azine and Cycl[3,2,2]azine.

The Cyclazines are a class of compounds consisting of a cyclic π-electron system (π perimeter) bridged by a central sp^2-hybridised nitrogen atom.[97] The first representative, Cycl[3,2,2]azine (Figure 42(2)), can be regarded as an isoelectronic aza-analogue of the aceindylenyl anion, or as a bridged[10]-annulene.

Recently, further compounds of the series have become available;[97] in particular, Cycl[3,3,3]azine, which is of special interest, being structurally related to the isoelectronic phenalenyl anion, and the [12]-annulene (shown in Figure 42 (1)). Thus, the two species of particular interest here are the two cyclazines, formally regarded as being derived by incorporating an internal nitrogen atom in [10]-and [12]-annulenes, respectively; the two species have 4n and a (4n+2) π-electron peripheries, respectively. According to IUPAC nomenclature,

Cycl[3,2,2]azine is pyrido[2,1,6-cd]pyrrolizine, and Cycl[3,3,3]azine is pyrido[2,1,6-de]quinolizine.

(a) Background

Twenty years ago, before either of the cyclazines had been synthesised, they were studied theoretically using a simple π-electron MO theory.[97] The two heterocycles were considered to be of particular interest because the heteroatoms are not on the periphery of the π-electron system, and there are more π-electrons than conjugated atoms. The aim of the application of the simple MO method was to predict the orientation of electrophilic substitution of these electron-rich systems, assuming this could be treated by considering the π-electron distribution; the nitrogen atoms are in unusual environments in these compounds, and might be expected to have significantly different electronic properties from the much studied pyridine- and pyrrole-type nitrogens. The study concluded that both cyclazines would show considerably greater reactivity towards electrophiles, at certain ring positions (particularly position 1 in Figure 42) than benzenoid compounds; in addition, both should show considerable thermodynamic stability compared to related cyclic conjugated hydrocarbons, with the cycl[3,2,2]azine somewhat less so owing to σ-electron "strain" energy. However, following on from this, cycl[3,2,2]azine was synthesised and studied experimentally. Cycl[3,3,3]azine defied attempts at synthesis until more recently, and its physical and chemical properties were found to be in marked contrast to those of cycl[3,2,2]azine. In an attempt to rationalise this behaviour, a MO study using a Hückel-type (σ,π) approximation was carried out. The conclusion reached, on the basis of experimental and theoretical evidence, was that cycl[3,2,2]is an "aromatic" compound, whereas cycl[3,3,3]azine is actually "antiaromatic".

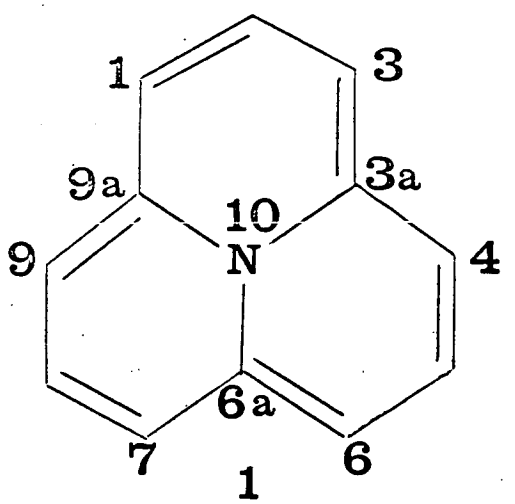
The approach in this work considers the cyclazines as natural extensions of the annulenes. Thus, formally, incorporation of the bridging nitrogen yields a physically reasonable planar annulene skeleton. Simple, qualitative considerations can rationalise the observed properties by describing the cyclazines with a model where the nitrogen substitution is a minor perturbation on the annulene framework, so that cycl[3,2,2]azine enables an aromatic[10]-annulene (4n+2) system to be attained, whereas cycl[3,3,3]azine is effectively a [12]-annulene (4n) system. However, this model of a perturbed perimeter of the cyclic π-electron system really appears to be inadequate, as there is likely to be extensive conjugation between the nitrogen lone pair and the π-electron system. This view emphasises the connection with the isoelectronic "nonalternant" aceindylenyl anion. The application of the Hückel rule as widely applied qualitatively leads to a reversal of the trend above; cycl[3,2,2]azine is a 12π-electron system, and cycl[3,3,3]azine a 14π-electron system, which seems inconsistent with experimental findings. Thus, the principal objectives of this study were to try to elucidate the electronic structure of these cyclazines using a more refined theoretical model, and to obtain and interpret the photoelectron spectra of the compounds; some investigation of the extent of the interaction of the nitrogen atoms with the hydrocarbon system is therefore possible by these means. Indolizine and its aza derivative (Figure 42(6),(7)) are included as reference molecules with bridgehead nitrogen atoms; some aza derivatives of cycl[3,2,2]azine, some very recently synthesised,[98] are given some consideration.

(b) Procedure

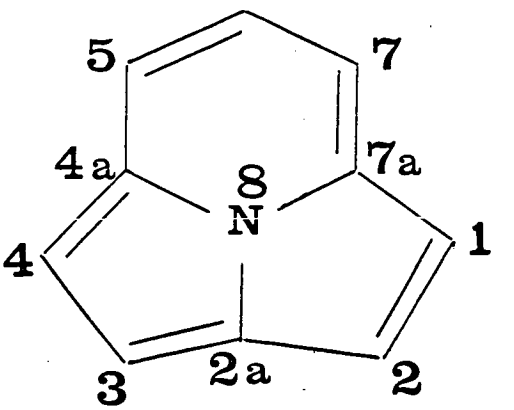
The two cyclazines and the two indolizines were prepared by previously reported procedures . . . . . The photoelectron spectra of these compounds, and some other related ones, were obtained with a Perkin Elmer PS18 modified by incorporation of an Helectros hollow cathode discharge lamp, which gives both He(I) (21.22 eV) and He(II) (40.81 eV) radiation. In all cases, the spectra were measured in admixture with the calibrants Xe($^2P_{3/2}$ 12.13, $^2P_{1/2}$ 13.44 eV), and Ar($^2P_{3/2}$ 15.75 eV) or N_2($^2\Sigma_g$ 15.58, $^2\Sigma_u$ 18.75 eV); further calibration was by means of He^+(^2S 24.59 eV) and He(I) β satellites of Xe.

The computational method was typical of the procedure used in this work. The standard Gaussian basis set was used, enabling direct comparison to be made with previous work on the photoelectron spectra of N-heterocycles.[99] There is very little experimental information available on the molecular geometry of the four heterocycles. For the two reference molecules, indolizine and an aza derivative, there are no experimental data; two attempts at geometry optimisation for indolizine within the semiempirical CNDO/2 framework have led to varying results, although one was close to a MINDO/2 result.[97] In the present work, a superposition of pyrrole or imidazole on 2-pyridinone was employed, yielding the structures shown in Figure 41 (8, 9). This method gives satisfactory agreement with experiment for other N-heterocycles. In practice, this constructed structure of indolizine is close to the average of the semi-empirical ones cited above. As far as the cyclazines themselves are concerned, there is an experimental geometry (X-ray crystal structure) for 1,4-dibromocycl[3,2,2]azine, and this is likely to be similar to that of the parent compound, notwithstanding an earlier

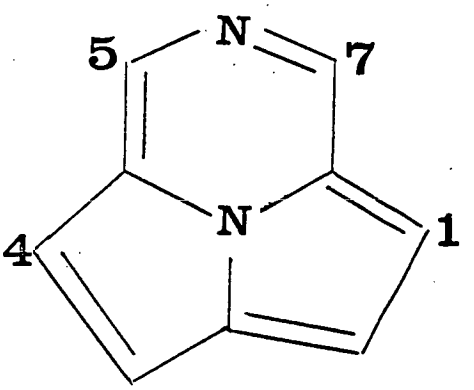
Figure 42: CYCLAZINES, AND RELATED SPECIES.



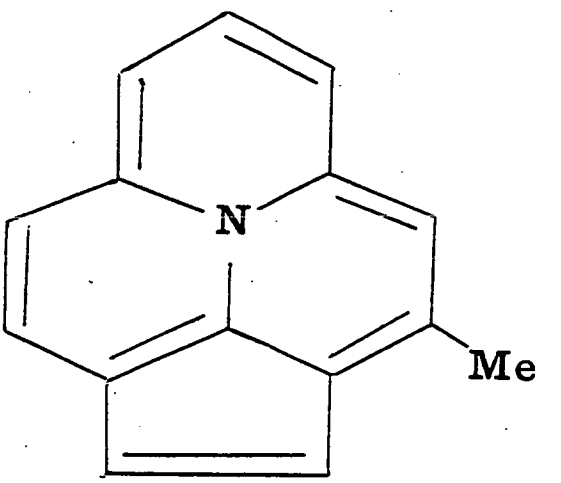
1



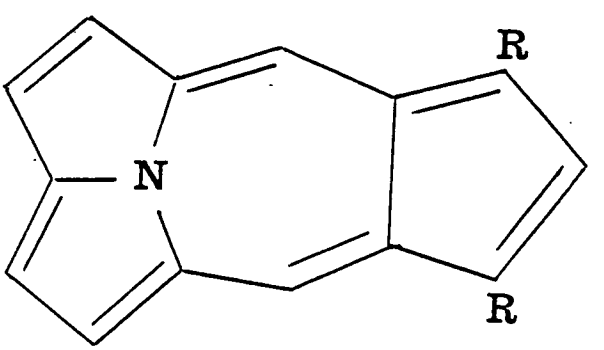
2



3

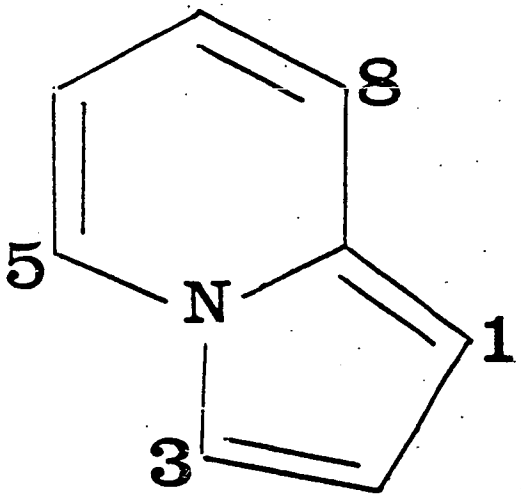


4

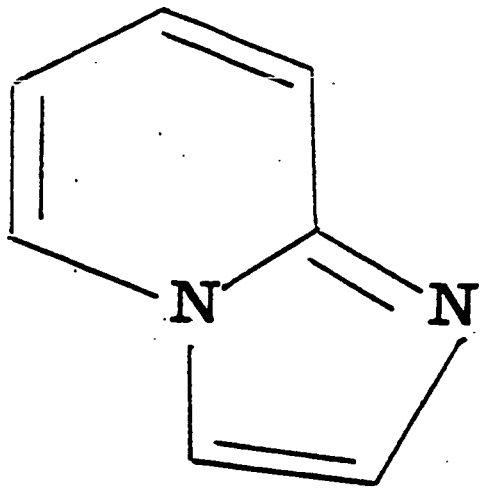


5 a, R = H ; b = Me

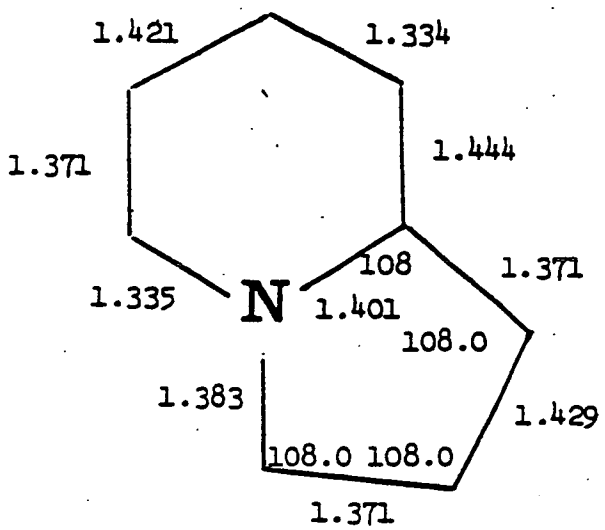
Figure 42 (cont.)



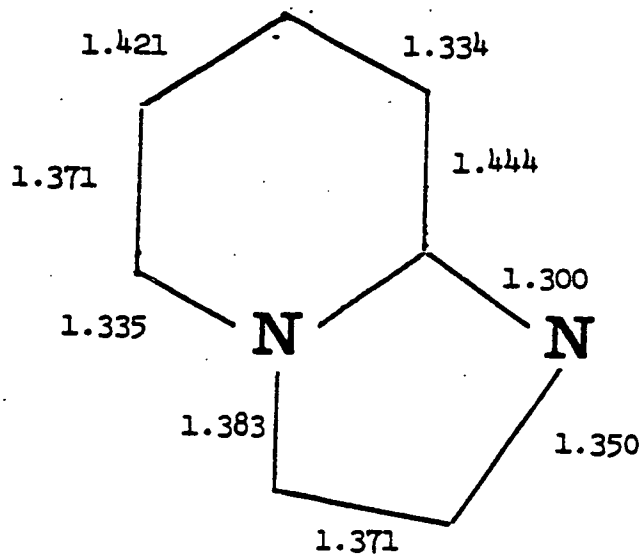
6



7



8



9

calculated geometry by semiempirical means.[97] In addition, in keeping with considerations from extended Hückel theory, a regular geometry was used for the [10]-annulene framework of the compound, in constructing an alternative structure for cycl[3,2,2]azine. Similarly, for cycl[3,3,3] azine, with no experimental data, two constructed planar geometries were considered, one corresponding to a regular structure for the [12]-annulene periphery, and the other to an alternating structure. In view of the size of the computations involved, it is impractable to consider optimisation of the geometries by ab initio methods. The aza-cycl[3,2,2]azines considered were assumed to have structures approximated by simply replacing the appropriate (C-H)'s by N's. The computed results and correlations with experimental data on these N-heterocycles given below.

(c) Results and Correlations

Total Energies

In Table[37](a) are shown the calculated molecular total energies for Cycl[3,2,2]azine in the two cases considered, i.e. using an "experimental" geometry (of a simple derivative) and a constructed regular geometry. The latter structure was simply that considered in Section C as a possible geometry for planar[10]-annulene with a nitrogen atom positioned in the centre of the ring replacing the three internal hydrogens. This regular structure is composed of a regular hexagon with benzenoid length of side, i.e. referring to Figure[42] r(CC) = r(CN) = 1.398 Å and all angles = 120° in ring (4a5677a8), and two pentagons of the same length of side, but not completely regular as the angle at N is constrained to be 120°, leaving the angle at each C in these rings as 105°. The experimental structure, actually derived by averaging the experimental parameters of the 1,4-dibromo-derivative to yield exact C_{2V} symmetry, has a [10]-annulene perimeter which is approximately the same as the regular one; the

major difference, in this respect, between the two geometries is in $r(C_5-C_6)/r(C_6-C_7)$ bond length, which is 1.441 Å in the experimental geometry. This lengthening may possibly relieve some of the strain from the 6,5,5-fused tricyclic system. In addition, the internal part of the structure in the neighbourhood of N is significantly different in the two cases. The experimental structure leads to a lower total energy, but only by 13 kJ mol^{-1}.

The calculated total energies for Cycl[3,3,3]azine are given in Table 37(a). Both structures considered are constructed basically from plausible planar [12]-annulene geometries. Thus, although the peripheries of the two structures are quite different, corresponding to a regular, benzenoid [12]-annulene (all bond lengths r(CC) = 1.398 Å, all angles CCC = 120°) and an alternating one (geometrical parameters in Figure 42), the internal parts are very similar. The alternating periphery form has a slightly lower total energy, by 30 kJ mol^{-1}, for the singlet state. The triplet benzenoid structure lies a further 129 kJ mol^{-1} above this; triplet species are of relevance in Hückel-type considerations, as they are predicted to be of lower energy than singlets for regular 4n annulenes, strictly for polygonal structures. The significance of the singlet difference here is uncertain, although it is in agreement with earlier semi-empirical calculations; indeed, the geometries used there are effectively identical to the two considered here. There is no experimental information for [12]-annulene itself, but the recent observation[97] that the low temperature ^{13}C magnetic resonance spectrum of heptalene, a further bridged [12]-annulene, shows non-equivalence of C1/C5 and of C2/C4 has been interpreted in terms of fast π-bond exchange at higher temperatures, in agreement with calculations. This could well point to a similar situation with

Cycl[3,3,3]azine, indicating a non-regular perimeter, but this must be regarded as unproven since other heptalene derivatives are known to be non-planar at normal temperatures.

In both Cyclazines, the orbital energies and most atomic populations are little affected by the relatively minor changes in geometry. The data from the computation of lower total energy are adopted in correlations with experiment.

For all geometries of the two Cyclazines, substantial resonance energies (RE) are calculated, based upon the bond energies of Chapter 4. The value for the alternating form of Cycl[3,3,3]azine is higher (by 30 kJ mol^{-1}) than the benzenoid form; this is an "anti-aromatic" characteristic, in agreement with the Hückel MO 4n-periphery predictions and semi-empirical calculations (ref. 97 gives a difference of 50 kJ mol^{-1}). However, the values for Cycl[3,3,3]azine are higher than those for Cycl[3,2,2]azine (Table 37), in contrast to the results of empirical and semi-empirical calculations, the latter showing the [3,2,2]azine to be more stabilised than classical [3,3,3]azine by 53 kJ mol^{-1}. The semi-empirical treatment calculates the σ- and π-energies separately as an approximation, and this does not occur in ab initio treatments. The present result can be attributed to a combination of lower σ-bond strain in the [3,3,3]azine, and to more efficient cross-linking of the ring in the symmetrical [3,3,3]system. The difference in reactivity, used as evidence in describing the [3,2,2]azine as aromatic and the [3,3,3] as anti-aromatic, is discussed below.

The RE's for indolizine and its aza-derivative (Figure 42) are much lower than in either indole (308 kJ mol^{-1}) or isoindole (250), as expected.

Photoelectron Spectra and their Assignment

Combined He(I) and He(II) spectra for the two Cyclazines, indolizine and imidazo 1,2-a pyridine are shown in Figure 43 ((a)-(d)). In addition, the He(I) spectrum for the 6-azacycl 3,2,2 azine (Figure 42(3)) is shown in Figure 43 (e). Spectra were obtained for the related compounds, 3-methylcyclopenta [cd]cycl[3,3,3]azine, cyclopenta[h]cycl[4,2,2]azine and its 6,8-dimethyl derivative (Figure 43(f)); these show little in the way of well-defined ionisation potentials beyond 10 eV, and so the low energy regions only are shown in Figure 43(f) . Observed I.P.'s for all the compounds are given in Table 37(b). Other than the first I.P., very few bands show fine structure (Figure 43(g)), and all of these are thought to arise from π-type ionisations[97].

Although a considerable number of I.P.'s have been identified (Table 37), multiple assignments have been necessary for certain portions of the envelope, where no separate maxima are observed. Assignment has been assisted by (a) previous experience with indole and isoindole[40] which are isomeric with the indolizines; (b) by variations in band intensity in the immediate neighbourhood, where cross-section differences should be small; (c) variations between I.P.'s both observed and calculated, and also spaces between groups of I.P.'s across the series of cyclazines and indolizines. Finally, the range of accessible I.P.'s in the present work is 5-25 eV; the restriction at the high energy end arises from the unfiltered nature of the He(I) plus He(II) radiation, and thus overlaying of the kinetic energy ranges from the two excitation sources. The $2S_N$ and $2S_C$ levels in pyrrole are at 29.5, and 23.8 plus 22.3 eV respectively;[97] some other $2S_C$ levels relevant to the present compounds are:

benzene 25.9, 22.7; pyridine 24.2, 23.3;

pyrimidine 24.4, 20.5; pyrazine 24.1, 21.0 eV

FIG. 43(a)

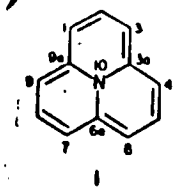
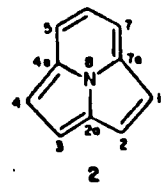
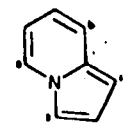


FIG .43(b)



2

FIG. 43(C)



6

He I

He II

He^+

6 10 14 18 22 26

eV

FIG. 43(d)

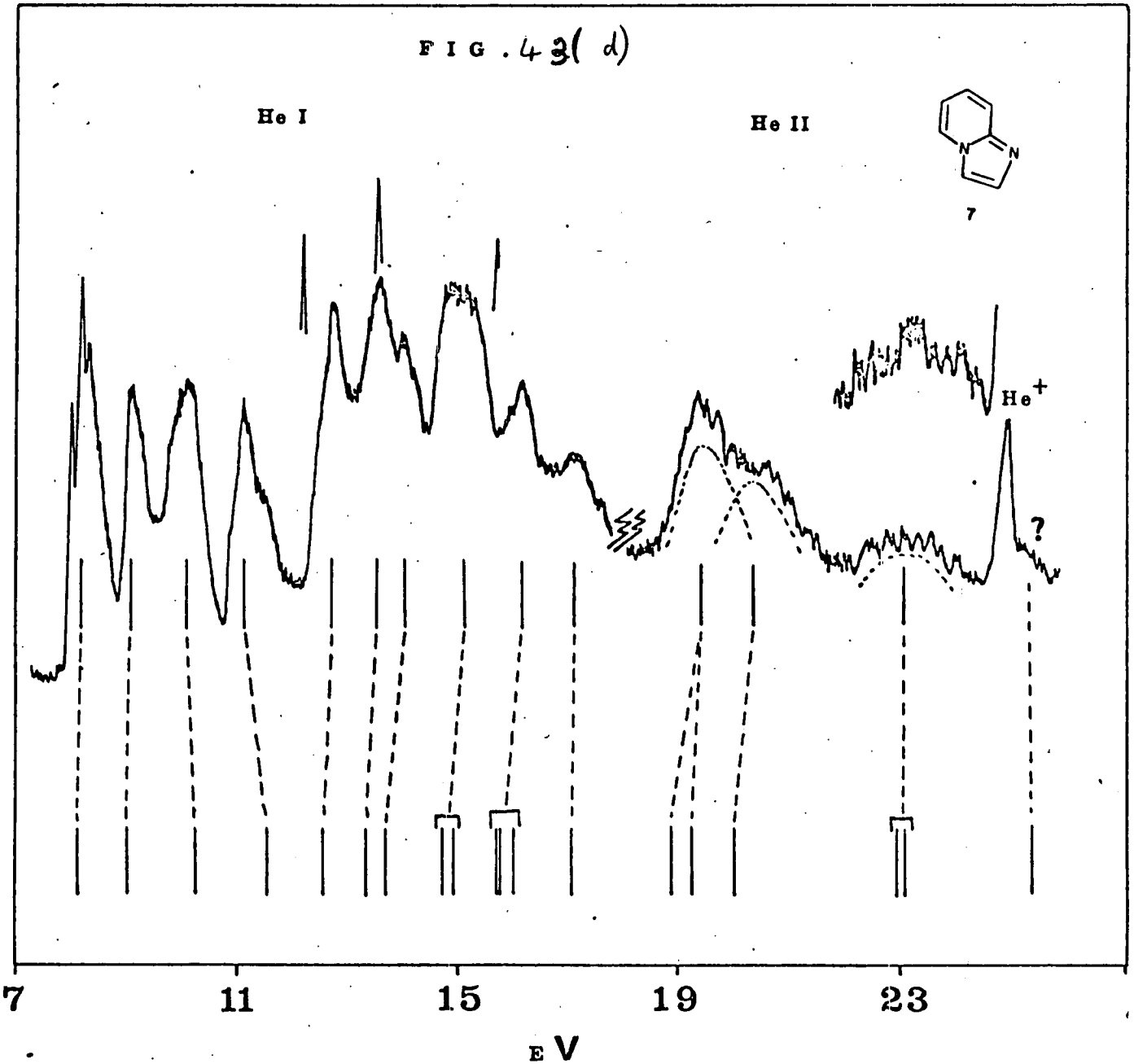
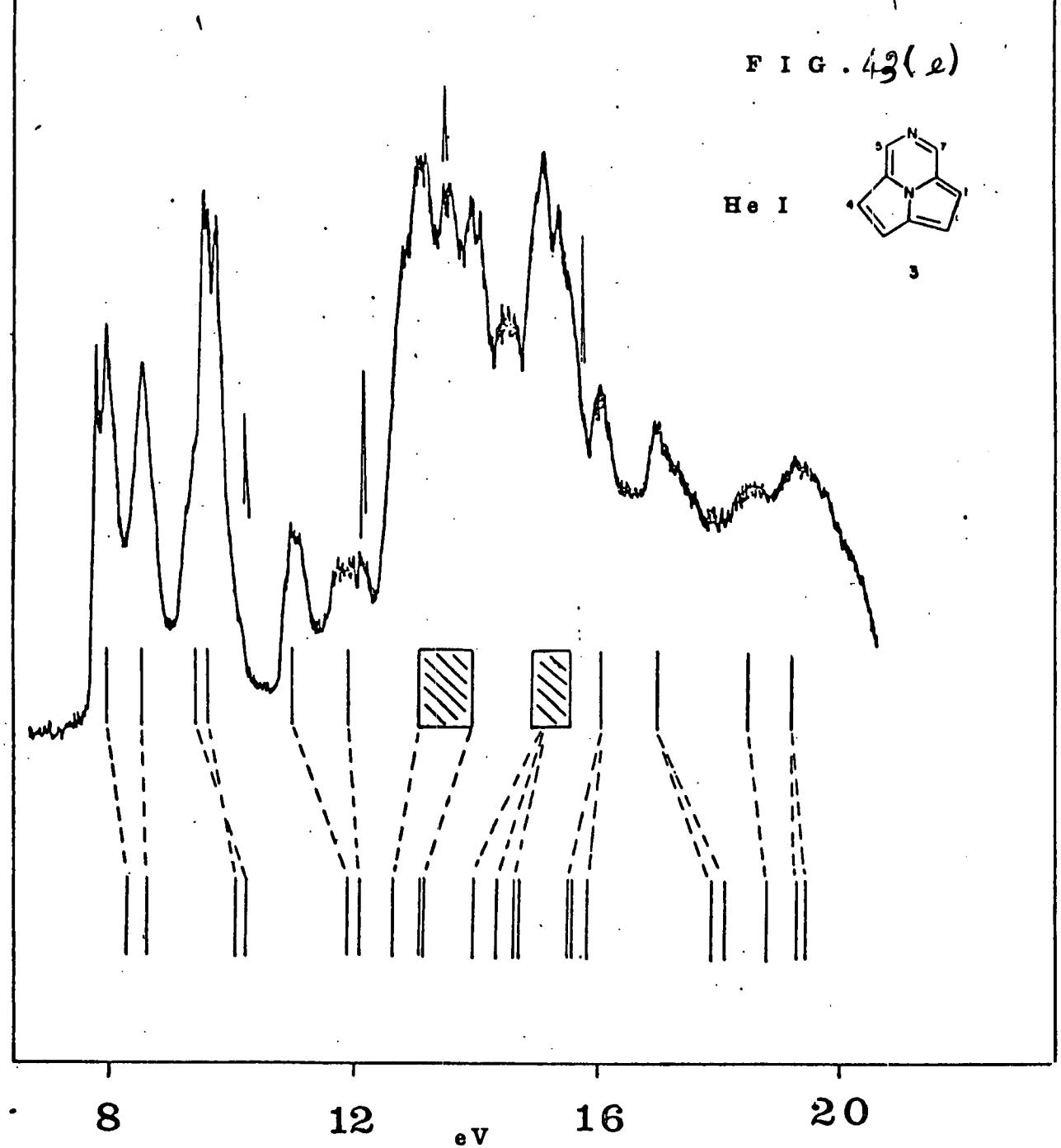
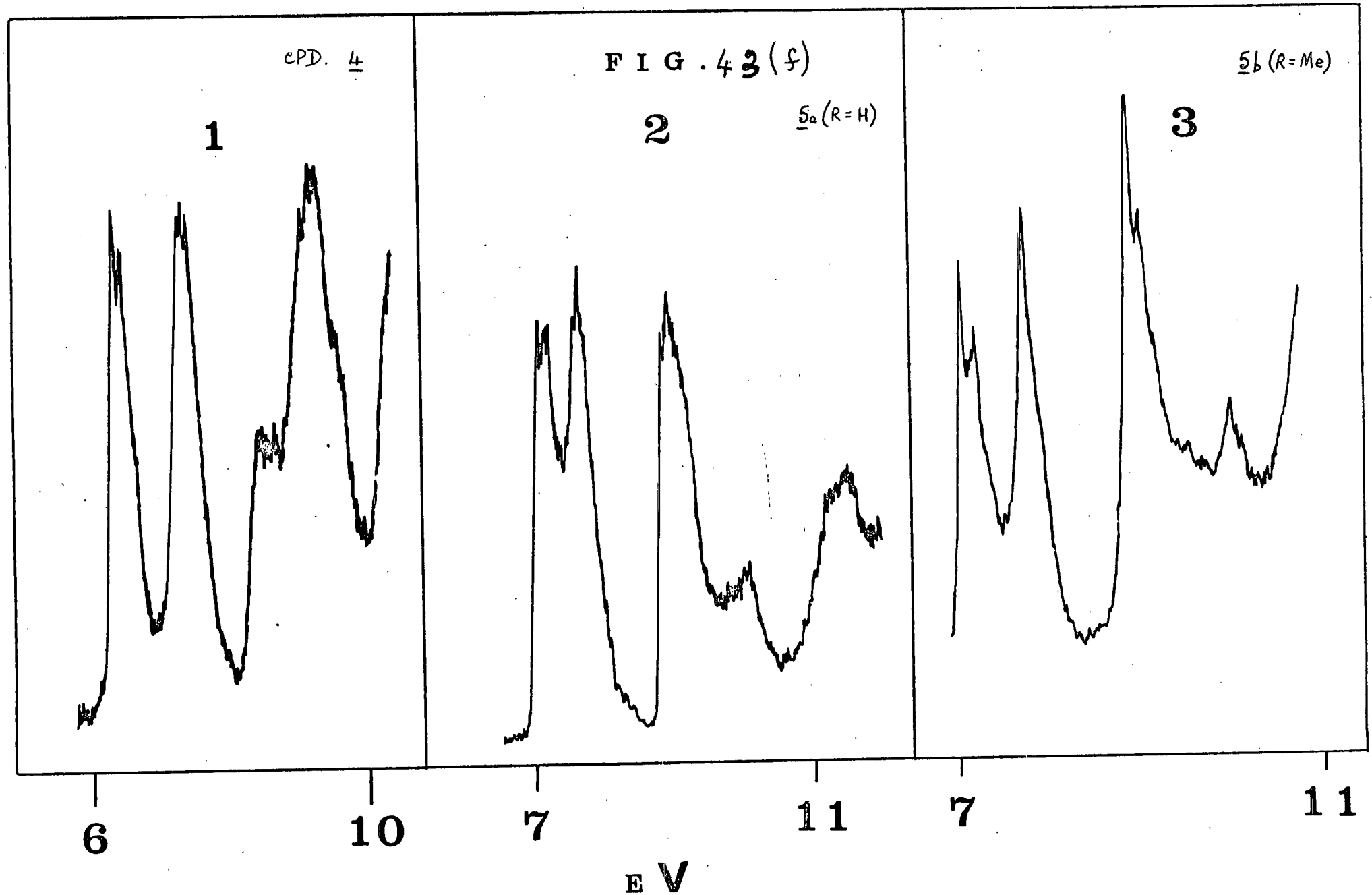


FIG. 43(e)



3

FIG. 43 (f)



Hence, it seems probable that at least 3 levels ($2S_N + 2 \times 2S_C$) cannot be observed in each of the compounds studied here, within the present I.P. range. With this in mind, the high energy end of the spectra has been assigned to account for variations across the series.

Some aza-substituted cycl[3,2,2]azines are considered below, but the calculation on 6-azacycl[3,2,2]azine can be considered at this point in relation to its photoelectron spectrum. With no experimental structural information available, calculation was made at a geometry derived by simply replacing C6-H6 by N6, thus yielding a first approximation to the actual structure, at least. Using the experimental structure of Cycl[3,2,2]azine, the resulting C5-N6 bond length of 1.441 Å can be regarded as one of the more unsatisfactory aspects of the aza-substituted structure. Thus, a minor alteration was considered, and the N6 position varied to produce r(C5-N6) = 1.399 Å; however, as shown in Table 37 the total energy of this second structure is marginally less negative. The computed orbital energies of either of these two very similar calculations are considered here, in the assignment of the measured photoelectron spectrum.

If all the low energy I.P.'s (5.8 - 11.5 eV) for the three cyclazines and two indolizines are placed in strict sequence, and compared with the calculated sequence using Koopmans' approximation (23 levels), then only 5 calculated levels are out of position; of these, 3 arise from the 6-azacycl[3,2,2]azine, whose geometry (and hence calculation) is probably less well-defined. Over this short range a linear correlation exists;

$$IP^{OBS} = 0.62\ IP^{CALC} + 2.6\ eV,$$

with standard deviations of slope, intercept and overall of 0.032, 0.353 and 0.342 respectively. This suggests that a correlation between calculated

and observed IP's can be constructed, the two ends being fairly firmly established. The more complex central region (12-18 eV) can then be assigned by consideration of IP groupings and variations in the series. In this way, an overall correlation exists:

$$IP^{OBS} = 0.771\ IP^{CALC} + 1.07\ eV,$$

with standard deviations in slope, intercept and overall of 0.011, 0.196 and 0.533 respectively. These correlations are shown in Figure 43. It is observed that the principal groupings and spacings are reasonably satisfactory, but less so for Cycl [3,3,3]azine and 6-azacycl [3,2,2] azine. Similar correlations have been obtained for related heterocycles,[40,99] and generally the slope increases as the range of I.P.'s is extended.

As far as previously reported calculations (semi-empirical and empirical only) are concerned, the only one reporting orbital energies is a CNDO/2 one.[97] It is well known that this computational method does not give satisfactory groupings of I.P.'s in the upper valence shell, and that the π-levels are often at too high a binding energy. This is evident in the only case, indolizine, in the present work where such semi-empirical orbital energies are available. Detailed comparison with the present data, and with the photoelectron spectrum, shows that it is not possible to arrive at a satisfactory explanation of the four I.P.'s at lowest binding energy, either in spacing relative to the main σ-system, or in order, where the CNDO/2 I.P. order is $\pi < \sigma < \pi < \sigma < \sigma$ etc.

Thus, having seen how the computed molecular wavefunctions fare in respect of orbital energy, the forms of the MO's are now examined.

Population Analyses.

In Table 37 are presented the total populations at the distinct centres of the five compounds. These atomic populations are further subdivided into 5 AO components, and p_σ, p_π components (for C and N). For all of the compounds nitrogen is an accepter from carbon, and carbon from hydrogen in total populations, as observed previously in other N-heterocycles.[40] For indolizine, imidazo[1,2-a]pyridine, and cycl-[3,2,2]azine, the total populations are very similar to those obtained by superposition of buta-1,3-diene (Chapter 4) with either pyrrole for indolizine and cycl[3,2,2]azine, or imidazole for imidazo[1,2-a]pyridine (Figure 42(n)). The population deficiencies at the bridgehead atoms in these three compounds, relative to C-2 in pyrrole or imidazole, arise from the absence of an attached H-atom. Most of the previous calculations on the relevant species cited above were π-electron, non-iterative HMO types, and clearly little direct similarity with an all-electron system can be expected. An exception to this for indolizine is the ω-technique iterative HMO method,[97] which yields both a wavefunction and π-electron distribution very similar to the present work. The valence shell semi-empirical calculations on indolizine produce an even closer similarity with the present work, particularly in the π-electron distribution, where often the total population orders are nearly identical. In the aza-indolizine, the CNDO/2 calculations[97] give the two nitrogen atoms, although differing in total density, nearly equal π-densities; in contrast, the carbon atoms are often highly polarised in the π-system. The present work suggests that the molecule is much closer to the classical structure; both nitrogen atoms are equal in total density, but widely differing in π-electron components.

Previous population analyses on the two cyclazines are much more fragmentary; the information given from some semi-empirical calculations is incomplete, but it is of interest to note that Dewar's π-distributions are not unlike those calculated for these two compounds here, except that N-10 in cycl 3,3,3 azine is a much stronger donor to the periphery in the present work (recovering much of this by σ-acceptance from the ortho- and para-positions, C-3a and C-2). This is perhaps not unexpected on the grounds that the molecule, by becoming meso-ionic (an extreme situation) can achieve the stable 14π periphery. There is considerable variation in populations with geometry change in cycl[3,3,3]azine, the charge separations being smaller in the "classical" alternating species. The same trends as in the pyridine populations (Figure 42(o)) are observed:-

σ, $C_\gamma > C_\beta > C_\alpha$; π, $C_\beta > C_\alpha > C_\gamma$, but the absolute values are significantly different.

The present calculations yield satisfactory values for the molecular dipole moments of indolizine and the aza-derivative; this is true also of the CNDO/2 results above, which, although leading to larger internal dipoles, these undergo mutual cancellation. A systematic difference evident in all the comparisons here is that the bridgehead N atom is a stronger π-donor and weaker σ-acceptor than in the present work. Thus, it is almost electrically neutral or even positively charged in the CNDO/2 calculations, whereas it is always significantly negative in the ab initio calculations.

One of the most obvious features of the I.P.'s of the series of N-heterocycles here is the extremely low first I.P.'s for the pair of cycle-[3,3,3]azines (5.87, 6.37 eV) when compared with either the cycl[3,2,2]- or cycl[4,2,2]-azines. These must be amongst the lowest values recorded for organic molecules; the limiting situation is perhaps represented by graphite whose first I.P. is 4.9 eV, but it is interesting to note that recent values for I.P.(1) in methano-bridged annulenes, and I.P.(1) for pyrene, are only slightly higher (6.7 - 7.9 eV). The size of the larger cyclazines has precluded the performing of calculations on these molecules, but the present results suggest that further P.E. studies of 4nπ-periphery condensed systems could lead to even lower I.P.'s, and possibly to compounds of interest as potential semi-conductors. The first I.P. of tridehydro[12] annulene of Section D is observed at 7.69 eV, and this might suggest that the bridging N atom is critical to the low I.P.'s above. The latter 4nπ system has LUMO of a_1'' symmetry, which is the type of the HOMO of cycl[3,3,3] azine. If the HOMO of the tridehydro-annulene ($2a_2''$) in the ground-state is replaced by the $1a_1''$ MO (excited state), and a calculation performed, the binding energy of the $1a_1''$ MO is only 3.9 eV. To add further evidence that the occupancy of the MO of a_1'' symmetry is responsible for low I.P.'s (bonding only between widely separated atoms - Figure 42), other molecules can be considered, of high enough symmetry. Thus, 12-annulene as a regular duodecagon has a calculated I.P. for such an orbital of 6.27 eV. In addition, the cyclic boron analogue, formed by combining a central boron atom with the hydrocarbon periphery, has a HOMO of 7.78 eV (LUMO is $1a_1''$). It seems unlikely that this (unknown) species would have an especially low first I.P. Ignoring the acetylenic in-plane (π') orbitals of the tridehydro[12]annulene, a direct comparison of the P.E. spectral data of the

latter and cycl[3,3,3]azine with the present assignments shows that the π-IP's are significantly shifted to lower binding energy in the azine when the N atom is nodal (1e" and 2e" at 10.3, 8.26 and 11.67, 9.5 respectively). The converse applies for the a" orbitals (14.52, 9.66 and 12.4, 7.69). From the population analyses, this arises because both a" orbitals of the azine have high lone-pair N character and the $(2p^{\pi})_N$ electron density arises from almost equal amounts of the two orbitals. Interaction of $2p_N$ (of higher binding energy) with linear combinations of $2p_c$ (of a" symmetry) lowers the inner combination ($1a_2$") and raises the antisymmetric outer combination ($2a_2$"). These resultant orbitals vary in energy depending upon the $2p_N/2p_c$ ratio. The assigned I.P.'s (and calculated data) are (eV):-

14.52 (17.17), 15.56 (18.54), 14.70 (17.63), 15.05 (18.95), 16.02 (18.55).

Cycl[3,2,2]azine (second value) lies out of position, perhaps because of a weakening of the σ-bonding in the strained system might possibly be offset by stronger π-interactions.

Thus, the present calculations suggest that the resonance energies of both cycl[3,2,2]azine and cycl[3,3,3]azine (whether classical or regular geometry) are high, and the difference in stability probably arises from the difference in binding energy of the HOMO rather than from the inherent instability of the [3,3,3]system in total. Thus, the very low I.P. of the latter suggests that oxygen is reduced to the radical anion (O_2^-) by it. The HOMO of $1a_1$" symmetry is only weakly bonding from its very nature. The present calculations and E.S.R. spectra support the assignment of an orbital of this type to the LUMO in cycl[3,2,2]azine. The close similarities of the two orbitals is consistent with the fact that the E.S.R. hyperfine splittings a_H are nearly equal in the cation of cycl[3,3,3] azine and the anion of cycl[3,2,2]azine at corresponding positions in the 6-membered rings.

Some further azacycl[3,2,2]azines have recently been synthesised, with N-substitution in the five-membered rings. Calculations can simply be performed by replacing a C-H with N, using the parent's geometry, at least as first approximations to the actual structures. There is no experimental information to compare with, but the calculations show that there is a "dilution" of the effect of N-substitution in these large molecules, in that the electronic structure is "rigid" as N replacing C-H hardly affects the forms of the MO's (as found in E.S.R. experiments, in general). Compared to cycl[3,2,2]azine insertion of a N atom in the periphery tends to raise all valence I.P.'s by a reasonably uniform amount (0.4-0.5 eV), irrespective of the position of substitution. Thus, the preliminary calculations on these species indicate that there is probably little else of further interest to be found with them.

References for Chapter 5

1. F. Sondheimer, Chimie, 28, 163 (1974).
2. P.J. Garratt, "Aromaticity", McGraw-Hill, London, 1971.
3. G. Leroy and S. Jaspers, Journal of Chemical Physics, 64, 470 (1967).
4. J.D. Cox and G. Pilcher, "Thermochemistry of Organic and Organometallic Compounds", Academic Press, London, 1970.
5. R.E. Davis and R. Pettit, Journal of the American Chemical Society, 92, 716 (1970).
6. D.L. Rodgers, E.F. Westrum and J.T.S. Andrews, Journal of Chemical Thermodynamics, 5, 733 (1973).
7. R. Breslow and J.T. Groves, Journal of the American Chemical Society, 92, 984 (1970).
8. J.A. Elvidge, Chemical Communications, 160 (1965).
9. J.A. Pople and K.G. Untch, Journal of the American Chemical Society, 88, 4811 (1966).
10. J.I. Musher, Journal of Chemical Physics, 46, 1219 (1967).
11. T. Nakajima and S. Katagiri, Molecular Physics, 7, 149 (1964).
12. F. Sondheimer et al, in "Aromaticity", Special Publication No. 21, Chemical Society, London, 1967.
13. P. Laslo and P. von R. Schleyer, Journal of the American Chemical Society, 85, 2017 (1963).
14. M.A. Cooper and S.L. Manatt, ibid, 91, 6325 (1969).
15. C.R. Kanekar and C.L. Khetrapal, Curr.Sci., 36, 67 (1967).
16. F.W. Wehrli and T. Wirthlin, "Interpretation of Carbon-13 NMR Spectra", Heyden, London, 1976.
17. J.H. Goldstein and G.J. Reddy, Journal of Chemical Physics, 36, 2644 (1962).
18. G.M. Badger, "Aromatic Character and Aromaticity", C.U.P., London, 1969.

19. G.Herzberg, "Molecular Spectra and Molecular Structure", Vol3, Van Nostrand, Princton, 1966.

20. D. Peters, Journal of Chemical Physics, 51, 1559 (1969).

21. A.L. and L.W. Bush, Journal of the American Chemical Society, 90, 3352 (1968).

22. J.H. Van Vleck, "The Theory of Electric and Magnetic Susceptibilities", O.U.P., London, 1932.

23. L. Pauling, Journal of Chemical Physics, 4, 673 (1936).

24. R.C. Benson and W.H. Flygare, Journal of Chemical Physics, 53, 4470 (1970)

25. H.J. Dauben, J.D. Wilson and J.L. Laity, Journal of the American Chemical Society, 91, 1991 (1969).

26. J.-F. Labarre, P. de Loth and M. Graffeuil, Journal Chim.Phys., 63, 460 (1966).

27. D. Peters, Journal of the Chemical Society, 2019 (1963).

28. E.A.C. Lucken, "Nuclear Quadrupole Coupling Constants", Academic Press, New York, 1969.

29. E. Hückel, Z. Physik, 70, 204 (1931).

30. J.W. Armit and R. Robinson, Journal of the Chemical Society, 127, 1604 (1925).

31. E. Hückel, "Grundzuge der Theorie Ungesattigter und Aromatischer Verbindungen", Verlay Chemie, Berlin, 1938.

32. F. Sondheimer, Accounts of Chemical Research, 5, 81 (1973).

33. L. Salem, "The Molecular Orbital Theory of Conjugated Systems", W.A. Benjamin, California, 1966.

34. W.C. Herndon, Journal of Chemical Education, 51, 10 (1974).

35. R. Willstätter and E. Waser, Ber. dtach.chem.Ges., 44, 3423 (1911).

36. A.C. Cope and G.G. Overberger, Journal of the American Chemical Society, 70, 1433 (1948).

37. G. Maier, Angewandte Chemie (Int.Ed.), 13, 425 (1974).

38. M.J.S. Dewar, "The Molecular Orbital Theory of Organic Chemistry", McGraw-Hill, New York, 1969.

39. G. Maier, H.-G. Hartan and T. Syrac, Angewandte Chemie (Int.Ed.), 15, 226 (1976); W.T. Borden, Journal of the American Chemical Society, 97, 5968 (1975).

40. S.M.F. Kennedy, Ph.D. Thesis, Edinburgh, 1977.

41. W. Moyes, Ph.D. Thesis, Edinburgh, 1977.

42. M. Traettenberg, Acta Chemica Scandinavica, 20, 1724 (1966).

43. G. Wipff, U. Wahlgren, E. Kochanski and J.M. Lehn, Chemical Physics Letters, 11, 350 (1971).

44. C. Batich, P. Bischof and E. Heilbronner, Journal of Electron Spectroscopy, 1, 333 (1972/73).

45. R. Hofmann, Accounts of Chemical Research, 4, 1 (1971).

46. N. Bodor, M.J.S. Dewar and S.D. Worley, Journal of the American Chemical Society, 92, 19; M.J.S. Dewar, A. Harget and E. Hasselbrach, ibid, 91, 7521 (1969).

47. E.E. van Tamelen, Accounts of Chemical Research, 5, 186 (1972).

48. G. Schröder, "Cyclooctatetraene", Verlag Chemie, Weinheim/Bergstr., 1965.

49. N.L. Allinger, J.T. Sprague and C.J. Finder, Tetrahedron, 29, 2519 (1973).

50. Z. Juz and S. Meiboom, Journal of Chemical Physics, 59, 1077 (1973).

51. L.A.Paquette, Tetrahedron, 31, 2855 (1975).

52. R. Breslow, Pure and Applied Chemistry, 28, 111 (1971).

53. K. Ruedenberg, Journal of Chemical Physics, 66, 375 (1977).

54. L.Z. Stenkamp and E.R. Davidson, Theoretica Chimica Acta, 30, 283 (1973).

55. S. Masamune and N. Darby, Accounts of Chemical Research, 5, 272 (1972).

56. F. Sondheimer, R. Wolovsky and Y. Amiel, Journal of the American Chemical Society, 84, 274 (1962).

57. K. Stöckel and F. Sondheimer, Organic Syntheses, 54, 1 (1974).

58. S. Masamune and R.T. Seidner, Chemical Communications, 542 (1969).

59. R.B. Woodward and R. Hoffmann, "The Conservation of Orbital Symmetry", Academic Press, New York, 1970.

60. E.H. van Tamelen and T.L. Burkoth, Journal of the American Chemical Society, 89, 151 (1967).

61. S. Masamune, K. Hojo, K. Hojo, G. Bigam, D.L. Rabenstein, ibid, 93, 4966 (1971).

62. S. Masamune and N. Darby, Accounts of Chemical Research, 5, 272 (1972).

63. J.F.M. Oth, H. Röttele and G. Schröder, Tetrahedron Letters, 61 (1970).

64. D.J. Cram and J.M. Cram, Accounts of Chemical Research, 4, 204 (1971).

65. S. Ito and Y. Fukazawa, Tetrahedron Letters, 1045 (1974).

66. A.V. Kemp-Jones and S. Masamune, in "Topics in Nonbenzenoid Aromatic Chemistry", Vol. 1 (editor T. Nozoe), John Wiley, New York, 1973.

67. H.P. Figeys, Tetrahedron, 26, 5225 (1970).

68. J. Laska and D. Loos, Journal of Molecular Structure, 21, 245 (1974).

69. J.F.M. Oth, J.M. Gilles and G. Schröder, Tetrahedron Letters, 67 (1970).

70. B.A. Hess and L.J. Schaad, Journal of the American Chemical Society, 93, 2413 (1971).

71. N.L. Allinger and J.T. Sprague, Journal of the American Chemical Society, 95, 3893 (1973).

72. N.L. Allinger and J.T. Sprague, ibid., 94, 5734 (1972).

73. F. Sondheimer and Y. Gaoni, Journal of the American Chemical Society, 82, 5765 (1960).

74. J.F.M. Oth and J.M. Gilles, Tetrahedron Letters, 6259 (1968).

75. J.F.M. Oth, Pure and Applied Chemistry, 25, 573 (1971).

76. C.C. Chiang and I.C. Paul, Journal of the American Chemical Society, 94, 4741 (1972).

77. M.J.S. Dewar and C. de Llano, ibid, 91, 789 (1969).

78. C.A. Coulson and W.T. Dixon, Tetrahedron, 17, 215 (1962).

79. J. Bregman, F.L. Hirschfeld, D. Rabinovich and G.M.J. Schmidt, Acta Crystallographica 19, 227 (1965).

80. F.A. Van-Cathedge and N.L. Allinger, Journal of the American Chemical Society, 91, 2582 (1969).

81. M.J.S. Dewar, R.C. Haddon and P.J. Student, Chemical Communications, 569 (1974).

82. J.F.M. Oth, J.-C. Bünzli and Y. de J. de Zelicourt, Helvetica Chimica Acta, 57, 2276 (1974).

83. F. Sondheimer, Y. Gaoni, L.M. Jackman, N.A. Bailey, R. Mason, Journal of the American Chemical Society, 84, 4595 (1962).

84. F. Sondheimer, I.C. Calder, J.A. Elix, Y. Gaoni, P.J. Garratt, K. Grohmann, G. Di Maio, J. Mayer, M.V. Sargent, R. Wolovsky, in "Aromaticity", Special Publication No. 21, The Chemical Society, London, 1967; R.M. McQuilkin and F. Sondheimer, Journal of the American Chemical Society, 92, 6341 (1970).

85. H.N.C. Wong, P.J. Garratt and F. Sondheimer, ibid, 96, 5604 (1974).

86. H.N.C. Wong and F. Sondheimer, Angewandte Chemie (Int.Ed.), 15, 117 (1976).

87. P. Politzer and R.R. Harris, Tetrahedron, 27, 1567 (1971).

88. R. Destro, T. Pilati and M. Simonetta, Journal of the American Chemical Society, 97, 658 (1975).

89. J. Wirz, Helvetica Chimica Acta, 59, 1647 (1976).

90. S. Masamune and N.T. Castellucci, Angew. Chem.Int.Ed.,3, 582 (1964).

91. T.J. Sworski, Journal of Chemical Physics, 16, 550 (1948); ref. 66.

92. C.B. Reese and A. Shaw, Chemical Communications, 331 (1972).

93. S.W. Benson, F.R. Cruickshank, D.M. Golden, G.R. Haugen, H.E. O'Neal, A.S. Rodgers, R. Shaw, R. Walsh, Chemical Reviews, 69, 279 (1969); R.C. Haddon, Tetrahedron, 28, 3613 (1972); ref. 32.

94. H. Irngartinger, L. Leiserowitz and G.M.J. Schmidt, Chem.Ber., 103, 1119 (1970).

95. K. Fukiu, T. Romoto, S. Nakatsuji and M. Nakagawa, Tetrahedron Letters, 3157 (1972).

96. J. Kruszewski and T.M. Krygowski, Canadian Journal of Chemistry, 53, 945 (1975); I. Gutman and N. Trinajstic, ibid, 54, 1789 (1976), and references cited therein.

97. M.H. Palmer, D. Leaver, J.D. Nisbet, R.W. Millar and R. Edgell, Journal of Molecular Structure, and references cited therein.

98. M. De Pompei and W.W. Paudler, Journal of Organic Chemistry, 41, 1661 (1976).

99. R.H. Findlay, Ph.D. Thesis, Edinburgh 1973.

Table 1

NON-PLANAR COT TOTAL ENERGY VALUES

	Exptl. Energy/a.u.	Assumed Energy/a.u.
Total Energy E_T	-306.716577	-306.718511
Electronic Energy V_{EL}	-631.147888	-631.318907
1-Electron Energy V_{IE}	-1055.131985	-1055.739663
Kinetic Energy T	310.106929	309.997524
Nuclear Repulsion V_{NN}	324.431311	324.600396
Binding Energy B.E.	1.85680	1.85874
Resonance Energy R.E. (kJ mol^{-1})	15	20
(E_T Triplet	-	-306.48083)

Table 2

NON-PLANAR COT ORBITAL ENERGIES

$-\varepsilon_i$(calc.)/eV		I.P.(obs)/eV-Vertical	Assignment
Exptl.	Assumed		
8.778	9.338	8.42	$5a_1(\pi)$
10.555	10.270	9.78	7e (π)
11.826	11.646	11.15	$4b_2(\pi)$
12.813	12.523	11.55	$3b_1$
14.074	13.974	12.3	6e
16.543	16.431	14.6	$3b_2$
16.947	16.698	14.6	$4a_1$
17.114	16.980	15.0	5e
17.124	17.405	15.0	$2a_2$
18.802	18.633	16.35	$3a_1$
20.883	20.639	18.0	4e
23.925	23.250	19.7(?), 20.4	$2b_1$
25.853	25.851	21.45	$2b_2$
28.667	28.429	~24.0	3e
30.598	30.470		$2a_1$
(-306.6)	(-306.5)		

Table 3

ALTERNATING PLANAR (D4H) COT TOTAL ENERGY VALUES

TOTAL ENERGY	ENERGY/a.u.
E_T	-306.69805
V_{EL}	-622.14278
V_{IE}	-1036.93215
T	310.02865
V_{NN}	315.44443
B.E.	1.83857
R.E.(kJ mol^{-1})	-35

Table 4

VARIATION OF PLANAR COT E_T WITH r(CC)-RELATIVE VALUES

D_{4H}	E_T/kJ mol^{-1}	82;160	12; -	0;206	1;266	Singlet;Triplet
	r(C-C)/Å	1.42	1.475	1.50	1.54	

(r(C=C) = 1.34 Å, r(CH) = 1.10 Å)

D_{8H}	E_T/kJ mol^{-1}	204; -	80;16	64;0	68-4	Singlet;Triplet
	r(CC)/Å	1.375	1.398	1.42	1.44	

(r(CH) = 1.090 Å)

E_T (Best D_{4h}) = -306.70258 (Singlet)

E_T (Best D_{8h}) = -306.67628 (Triplet)

Table 5

CALCULATED ONE-ELECTRON ORBITAL ENERGIES FOR COT (a.u.)

D_{8H}	D_{4H}	D_{2D}
-14.238	-14.330 (π)	-14.596 (π)
-14.178	-14.259 (π)	-14.636 (π)
-13.580	-13.801	-14.582 (π)
-14.170	-13.580	-14.126
-13.779	-14.247 (π)	-14.230
-13.761	-13.787	-15.184
-12.510	-12.632	-14.201
-13.772	-13.806	-14.007
-14.994	-15.069	-12.941
-13.703	-13.750	-14.021
-14.184	-14.279	-14.224
-14.543	-14.471	-14.644
-14.555	-14.839	-15.180
-14.990	-15.127	-15.650
-15.194	-15.321	-15.971
(-29.2)	(-29.3)	(-29.6)

(ordered in orbital energy sequence)

Table 6

TOTAL ENERGIES OF PLANAR 10-ANNULENES (a.u.)

	D_{10h}	D_{5h}	C_{2v}(cttct)	C_{2v}(ccccct)
E_T	-383.31734	-383.31832	-382.42739	-383.31953
V_{EL}	-825.20846	-819.16167	-868.17380	-849.95281
V_{IE}	-1391.04030	-1379.28704	-1476.23200	-1440.62576
T	387.58814	387.43371	389.96268	388.23135
V_{NN}	441.89112	435.84335	485.74642	466.63328
B.E.	2.24262	2.24360	1.35169	2.24481
R.E. (kJ mol^{-1})	-195	-192	-2622	-189

Table 7

ORBITAL ENERGIES OF PLANAR 10-ANNULENES (a.u.)

D_{10h}	D_{5h}	C_{2v}(cttct)		C_{2v}(ccccct)	
0.297(2) π	0.319(2) π	0.285 π		0.293(2) π	
0.461(2) π	0.451(2) π	0.292 π		0.455 π	
0.474(2)	0.441(2)	0.369 (CH)		0.463	
0.494(2)	0.489(2)	0.392 (CH)		0.466 π	
0.512 π	0.501 π	0.466 π		0.484	
0.561(2)	0.562(2)	0.472 π		0.495	
0.573	0.574	0.528		0.521	
0.655(2)	0.647(2)	0.543 π		0.526 π	
0.665	0.650	0.550		0.539	
0.695	0.684	0.570	0.811	0.574	0.861
0.765(2)	0.752(2)	0.589	0.849	0.608	0.873
0.891(2)	0.869(2)	0.632	0.957	0.627	0.991(2)
1.000(2)	1.002(2)	0.639	0.976	0.649	1.080
1.078(2)	1.070(2)	0.666	1.065	0.688	1.095
1.108	1.098	0.674	1.105	0.716	1.144
11.278(10)	11.276(10)	0.737	1.120	0.743	11.3(10)
		0.756	1.410	0.787	
			11.3(10)		

Table 8

AO POPULATIONS OF 10-ANNULENES

		C	H
D_{10h}	1s+2s	3.05	0.85
	2p(σ)	2.10	
	2p(π)	1.00	
D_{5h}	1s+2s	3.06	0.85
	2p(σ)	2.09	
	2p(π)	1.00	

C_{2V}(cttct)	C1	C4	C(other)	H1	H4	H(other)
	2.99	3.01	3.06	0.95	0.92	0.85
	2.02	2.06	2.10			
	1.03	0.99	1.00			
Total	6.04	6.06	6.16			

C_{2V}(cccct)	C(int)	C(ext)	H(int)	H(ext)
	3.05	3.05	0.80	0.85
	2.15	2.10		
	1.00	1.00		
Total	6.20	6.15		

Table 9

TOTAL ENERGIES OF PLANAR 12-ANNULENES (a.u.)

Singlet	D_{12h}	D_{6h}	D_{3h}	C_{3h}	D_{2h}
E_T	-459.80494	-459.86762	-459.11148	-459.2330	-459.48570
V_{EL}	-1033.81536	-1025.95955	-1107.78109	-1102.8761	-1092.58301
V_{Ie}	-1756.38861	-1740.95268	-1904.31766	-1894.3276	-1873.08203
T	465.37281	465.05495	467.64998	467.3637	467.11752
V_{NN}	574.01042	566.09192	648.66961	643.6432	633.09731
B.E.	2.51528	2.57796	1.82182	1.94334	2.19604
R.E. (kJ mol^{-1})	-714	-543	-2605	-2273	-1585
Triplet					
E_T	-459.81046	-459.75673	-459.10769	-	-459.50037
B.E.	2.52080	2.46707	1.81803	-	2.21071
R.E. (kJ mol^{-1})	-699	-845	-2615	-	-1545

ORBITAL ENERGIES OF PLANAR 12-ANNULENES (a.u.)

D_{12h}	D_{6h}	D_{3h}	C_{3h}	D_{2h}	
-0.230 π	-0.283 π	-0.222	-0.26	-0.213 π	
-0.366(2) π	-0.375(2) π	-0.342(2)	-0.35(2)	-0.340 π	
-0.458(2)	-0.450(2)	-0.400(2)	-0.41(2)	-0.351	
-0.474(2) π	-0.469(2)	-0.475(2) π	-0.47(2)	-0.374 π	
-0.492(2)	-0.4373(2)	-0.514(2)	-0.52(2)	-0.390	
-0.509	-0.489	-0.533 π	-0.53	-0.483 π	
-0.509 π					
-0.511	-0.500	-0.534	-0.53	-0.493 π	
-0.551	-0.535	-0.565	-0.57	-0.533 π	
-0.598(2)	-0.553(2)	-0.586(2)	-0.58(2)	-0.548	-0.699
-0.665(2)	-0.594(2)	-0.594	-0.60	-0.550	-0.735

Table 9 (contd.)

D_{12h}	D_{6h}	D_{3h}	C_{3h}	D_{2h}	
-0.673	-0.657	-0.687(2)	-0.68	-0.551	-0.744
-0.690	-0.659	-0.689	-0.69(2)	-0.581	-0.800
-0.749(2)	-0.680(2)	-0.700(2)	-0.69(2)	-0.618	-0.872
-0.858(2)	-0.737(2)	-0.804(2)	-0.80(2)	-0.626	-0.907
-0.955	-0.841	-0.877	-0.86	-0.631	-1.0(5)
-0.958	-0.907	-0.906	-0.91	-0.654	-1.2(2)
-1.1(5)	-0.987	-1.1(5)	-1.1(5)	-0.655	-11.3(12)
-11.28(12)	-1.1(5)	-1.345	-1.35		
	-11.28(12)	-11.30(12)	-11.3(12)		

Table 10

MOLECULAR MECHANICS GEOMETRY OF 12-ANNULENE

Bond	Length/Å	Angle	Deg	Angle	Deg
1-2	1.515	12-1-2-3	74.0	1-2-3	120.1
2-3	1.337	1-2-3-4	174.1	2-3-4	122.6
3-4	1.516	2-3-4-5	90.2	3-4-5	124.2
4-5	1.348	3-4-5-6	0.4	4-5-6	126.4
5-6	1.474	4-5-6-7	36.3	5-6-7	125.1
6-7	1.350	5-6-7-8	176.3	6-7-8	123.9
7-8	1.475	6-7-8-9	138.0	7-8-9	121.7
8-9	1.352	7-8-9-10	5.5	8-9-10	121.9
9-10	1.474	8-9-10-11	45.2	9-10-11	120.4
10-11	1.349	9-10-11-12	174.6	10-11-12	125.2
11-12	1.470	10-11-12-1	145.0	11-12-1	119.7
12-1	1.346	11-12-1-2	3.5	12-1-2	119.8

Table 11

12-ANNULENE GEOMETRY OPTIMISATION

Centre/Coord.		Cycle 1-Gradients (Hartree-Bohr^{-1} x 10^{-2})	
C1	X	4.36525	α = 0.555
	Y	1.119905	S = 15.53
	Z	0.7299	σ = 8.62
C2	X	-2.00435	
	Y	2.80245	
	Z	1.6291	
C3	X	0.98855	
	Y	-3.1403	
	Z	1.96885	
C4	X	-0.62145	
	Y	4.7917	
	Z	-1.99735	
H1	X	-0.40615	
	Y	-0.6222	
	Z	-0.27295	
H2	X	-0.3799	
	Y	-0.5117	
	Z	-0.05265	
H3	X	0.55950	
	Y	0.96005	
	Z	0.56980	
H4	X	-0.0833	
	Y	0.4472	
	Z	0.43	

Table 11 (contd)

Geometrical Parameter	Cycle 0
$r(C_1-C_2)$	1.490
$r(C_2=C_3)$	1.348
$r(C_3-C_4)$	1.486
$r(C_{12}=C_1)$	1.350
$r(C_1-H_1)$	1.079
$r(C_2-H_2)$	1.085
$r(C_3-H_3)$	1.085
$r(C_4-H_4)$	1.085
12-1-2	119.6
1-2-3	120.1
2-3-4	128.9
3-4-5	120.2
H1-1-2	120.1
1-2-H2	119.9
H3-3-2	120.0
3-4-H4	120.0
H1-1-12	120.3
H2-2-3	120.0
H3-3-5	110.6
H4-4-5	119.8
E_T	-460.04218 (α = 0)
	-460.04446 (α = 1)
	-460.00544 (α = 2)

α = 1; B.E. = 2.75480 a.u.; R.E. = -60 kJ mol^{-1}

Table 12

AO POPULATIONS OF 12-ANNULENES

D_{12h}	C	H
1s+2s	3.04	0.86
2p(σ)	2.14	
2p(π)	1.00	

D_{6h}	C	H
	3.06	0.86
	2.08	
	1.00	

D_{3h}	C1	C2	C3	H1	H2	H3
	3.02	3.09	2.99	0.87	0.85	0.95
	2.01	2.18	1.91			
	1.21	0.79	1.21			
Total	6.24	6.06	6.11			

C_{3h}	C1	C2	C3	H1	H2	H3
	3.05	3.07	3.02	0.86	0.85	0.95
	2.07	2.11	1.97			
	1.06	0.95	1.06			
Total	6.18	6.13	6.05			

D_{2h}	C1	C2	C3	H1	H2	H3
	3.06	3.02	3.05	0.84	0.90	0.85
	2.10	2.07	2.10			
	1.00	1.00	0.99			
Total	6.16	6.09	6.14			

C_3 (non-planar)	C	H
1s+2s	3.06	0.85
2p	3.09	

Table 13

ORBITAL ENERGIES OF NON-PLANAR 12-ANNULENE (a.u.)

$-\varepsilon_i$
-0.351(2)
-0.364
-0.416(2)
-0.447
-0.458(2)
-0.533
-0.548
-0.567(2)
-0.611(2)
-0.611
-0.649
-0.673(2)
-0.713(2)
-0.734
-0.799(2)
-0.866
-0.958
-1.01(2)
-1.10(2)
-1.13
-11.3(12)

Table 14

X-RAY STRUCTURE OF 14-ANNULENE, AND ALSO MOLECULAR MECHANICS GEOMETRY (C_{2h})

Bond/Å	Obs.(Av.)	Calc.
1-14	1.391	1.410
1-2	1.365	1.408
2-3	1.369	1.405
3-4	1.395	1.411
Angle/Deg		
14-1-2	128.4	124.2
1-2-3	124.8	126.9
2-3-4	129.9	123.7
3-4-5	124.2	126.9
2-1-14-13	0	0
14-1-2-3	162.3	158.5
1-2-3-4	163.8	163.4
2-3-4-5	14.3	18.7

Table 15

TOTAL ENERGIES OF PLANAR 14-ANNULENES (a.u.)

	D_{2h}	C_{2h}
E_T	-536.08249	-536.16157
V_{EL}	-1345.72124	-1336.37922
V_{IE}	-2328.14087	-2312.32514
T	544.89581	544.67624
V_{NN}	809.63874	800.21765
B.E.	2.57788	2.65696
R.E.(kJ mol^{-1})	-1805	-1589

Table 16

ORBITAL ENERGIES OF PLANAR 14-ANNULENES (a.u.)

D_{2h}			C_{2h}		
π	-0.245			-0.292	
π	-0.255			-0.298	
(CH)	-0.365	-0.603	(CH)	-0.367	-0.601
π	-0.393	-0.658		-0.394	-0.659
π	-0.406	-0.662	(CH)	-0.408	-0.659
(CH)	-0.413	-0.675	π	-0.410	-0.665
π	-0.489	-0.684		-0.479	-0.681
π	-0.501	-0.701		-0.497	-0.691
	-0.511	-0.747	π	-0.499	-0.738
	-0.519	-0.779		-0.502	-0.770
	-0.523	-0.828		-0.520	-0.810
π	-0.531	-0.847		-0.524	-0.826
	-0.551	-0.884		-0.539	-0.882
	-0.573	-0.955		-0.566	-0.951
	-0.575	-1.1(7)		-0.587	-1.1(7)
	-0.593	-11.3(14)		-0.591	-11.3(14)

Table 16(b)

AO POPULATIONS OF PLANAR 14-ANNULENES

	C(EXPT)	C(INT)	H(EXT)	H(INT)
1s+2s	3.06	3.04	0.85	0.91
2p(σ)	2.09	2.04		
2p(π)	1.01	0.99		
Total	6.16	6.07		

Table 17

TOTAL ENERGIES OF DISTORTED 14-ANNULENES (a.u.)

	C_{2h}	D_2
E_T	-536.38280	-536.38765
V_{EL}	-1344.69299	-1342.62996
V_{IE}	-2326.43216	-2323.11984
T	544.69605	544.32531
V_{NN}	808.31019	806.24231
B.E.	2.87819	2.88304
R.E.	-986	-973

"Internal" (CH) Orbitals : $-\varepsilon_i$ = 0.385, 0.428

AO Populations : H(Int) 0.85

C(Int) 3.06 2.09 1.0

Table 18

14-ANNULENE GEOMETRY OPTIMISATION

Centre/Coord.		GR Cycle 1	Cycle 2	
		g(= -s)	g	s
C1	X	0	0	0
	Y	-1.71	4.08	-3.3776
	Z	0	0	0
C2	X	5.15	-1.845	1.52899
	Y	0.01	-0.62	0.51483
	Z	-0.19	-1.5	1.2469
C3	X	-8.465	2.93	-2.38014
	Y	0.855	-5.73	4.75325
	Z	6.46	7.1	-8.04608
C4	X	-1.205	-7.64	6.35239
	Y	-0.045	2.4	-1.99286
	Z	-0.025	-0.965	0.80156
H1	X	0	0	0
	Y	0.69	-1.675	1.38671
	Z	0	0	0
H2	X	0.535	0.395	-0.3314
	Y	0.48	-0.22	0.17969
	Z	-0.315	-0.405	0.33832
H3	X	-22.81	-0.92	0.90725
	Y	-2.85	1.23	-1.00359
	Z	9.6325	-0.505	0.35891
H4	X	0.435	-1.235	1.0229
	Y	0.285	0.04	-0.03501
	Z	-0.71	-0.695	0.58164
	α	1·776		1·482
	S	55·4		24·7
	σ	98·4		35·1

Table 18 (contd.)

Geometrical Parameter	Cycle 0	1	2
$r(C_1-C_2)$	1.398	1.402	1.398
$r(C_2-C_3)$		1.410	1.376
$r(C_3-C_4)$		1.327	1.414
$r(C_4-C_5)$		1.399	1.368
$r(C_1-H_1)$	1.090	1.065	1.102
$r(C_2-H_2)$		1.087	1.074
$r(C_3-H_3)$		1.040	1.020
$r(C_4-H_4)$		1.072	1.045
14-1-2		118.5	121.2
1-2-3		124.6	122.1
2-3-4		125.2	124.3
3-4-5		121.2	121.6
E_T (KJmol^{-1})	0	-556	-667

"INTERNAL" (CH) ORBITALS : $-\varepsilon_i$ = 0·425, 0·468

AO POPULATIONS : H (Int) 0·85

C (Int) 3·05 2·09 1·01

Table 19

CALCULATED GEOMETRIES OF 18-ANNULENE

Bond/Å	Mol.Mech(D_3)	Mindo/3(D_{3h})
1-2	1.357	1.350
2-3	1.463	1.458
3-4	1.361	1.355
4-5	1.463	1.458
5-6	1.357	1.351
6-7	1.467	1.466
Angle/Deg.		
18-1-2	122.7	128.8
1-2-3	125.7	127.5
2-3-4	123.4	128.1
3-4-5	123.4	128.6
4-5-6	125.7	136.9
5-6-7	123.0	129.0

Table 20

TOTAL ENERGIES OF 18-ANNULENE (a.u.)-D_{6h}

E_T	-690.07935
V_{EL}	-1819.72083
V_{IE}	-3172.42983
T	698.37281
V_{NN}	1129.64148
B.E.	4.14486
R.E.(kJ mol^{-1})	56

Table 21

ORBITAL ENERGIES OF 18-ANNULENE (a.u.)-D_{6h}

	$-\varepsilon_i$			
e_{2u}	-0.222(2)	π		
b_{2g}	-0.360	π		
b_{1g}	-0.363	π		
e_{2u}	-0.454(2)	π		
b_{1u}	-0.481			
e_{2g}	-0.499(2)			
e_{1g}	-0.508(2)	π		
b_{2u}	-0.513			
a_{2u}	-0.526	π		
e_{1u}	-0.529(2)			
a_{2g}	-0.533		-0.764	a_{2g}
e_{2g}	-0.555(2)		-0.787(2)	e_{1u}
e_{1u}	-0.594(2)		-0.847	a_{1g}
a_{1g}	-0.594		-0.870(2)	e_{1u}
e_{2g}	-0.629(2)		-0.949(2)	e_{2g}
e_{1u}	-0.641(2)		-1.018	b_{1u}
b_{2u}	-0.684		-1.035	b_{2u}
b_{1u}	-0.688		-1.1(5)	
a_{1g}	-0.711		-11.3(18)	
e_{2g}	-0.744(2)			

Table 22

AO POPULATIONS OF 18-ANNULENE-D_{6h}

	C	H
1s+2s	3.05	0.86
$2p(\sigma)$	2.09	
$2p(\pi)$	1.00	

Table 23

TOTAL ENERGIES OF CYCLOOCTADIENEDIYNE SPECIES (a.u.)

	I	II	III	IV	V	"EXPTL"
$-E_T$	303.8653	303.9123	304.2654	304.1755	304.2277	304.2632
$-V_{EL}$	612.0680	607.3304	583.1643	583.3112	589.6373	586.6765
$-V_{IE}$	1017.5380	1008.7643	960.6907	961.2041	973.2487	967.7754
T	310.0170	309.2963	307.2606	307.2181	307.6321	307.3900
V_{NN}	308.2027	303.4182	278.8989	279.1357	285.4097	282.4133
B.E.	0.9934	1.0404	1.3935	1.3036	1.3558	1.3913
R.E. ($KJmol^{-1}$)	-1069	-941	22	-223	-81	16
$-E_T$ (triplet)	303.7125	303.9264	304.1472	304.1498	304.1898	304.2048

Table 24

X-RAY STRUCTURE OF SYM-DIBENZO-CYCLOOCTADIENEDIYNE

Length/Å	Angle/dec
r(C1-C2) = 1.197	1-2-3 = 155.8
r(C2-C3) = 1.443	2-3-4 = 114.2
r(C3-C4) = 1.426	

(8-membered ring only)

Table 25

P.E. SPECTRAL DATA FOR DIBENZO-CYCLOOCTADIENEDIYNE

I.P.(Vertical)/eV	7.76	8.74	9.3-9.8	10.15	10.94-11.0	12.0 →
Assignment	$3b_{1u}$	$2b_{2g}$	$1a_u, 2b_{3g}$, $7b_{3u}$	9ag	$2b_{1u}, 1b_{2g}$	$7b_{2u}$
	π	π	π,π π	π'	π,π	σ

Table 26

VALENCE ORBITAL ENERGIES OF CYCLOOCTADIENEDIYNES

$-\varepsilon_i$(CALC)/eV

I		II	III	IV	V	EXPTL	
8.94	π	7.08	8.76	7.55	7.86	7.98	π
9.78	π	9.41	10.9	9.96	11.1	10.7	π'
10.8	π'	9.61	11.5	10.4	11.2	11.7	π'
13.0	π	12.8	12.2	12.0	12.4	12.0	π
14.7		13.0	12.5	12.3	12.6	12.3	π
15.2	π'	16.2	14.7	14.7	15.1	14.8	π
15.5		16.2	15.1	15.9	15.9	15.7	
16.8		16.9	15.9	16.6	16.4	16.5	
17.3		17.0	17.5	17.0	17.7	17.2	
18.1	π	17.2	17.7	17.6	18.0	17.3	
20.4		19.8	18.8	18.5	19.1	18.9	
21.0		20.9	19.9	20.0	20.3	19.9	
22.0		21.1	21.9	22.2	22.6	22.3	
22.5		24.1	23.6	24.8	24.9	24.1	
25.2		25.7	26.8	25.7	26.4	26.4	
30.1		28.0	28.4	28.5	29.0	28.4	
32.6		32.7	29.7	29.3	29.8	29.4	
37.8		35.4	30.9	30.9	31.4	31.1	

(I and II are exceptional)

Table 27

AO POPULATIONS OF CYCLOOCTADIENEDIYNES

		C1	C2	H				
	1s+2s	2.98	3.08	0.83				
	2p(σ)	2.06	2.05					
	2p(π)	1.00	1.00					
	Total	6.04	6.13					
III	1s+2s	3.02	3.06	0.83	3.01	3.06	0.84	IV
	2p(σ)	2.03	2.06		2.03	2.06		
	2p(π)	1.00	1.00		0.96	1.04		
	Total	6.05	6.12		6.00	6.16		
I	1s+2s	2.96	3.10	0.84	2.98	3.10	0.84	II
	2p(σ)	2.04	2.06		2.03	2.05		
	2p(π)	1.03	0.97		1.02	0.98		
	Total	6.03	6.13		6.03	6.13		

Table 28

TOTAL ENERGIES OF DEHYDRO[10]ANNULENES (a.u.)

	I	II
E_T	-380.97088	-380.98128
V_{EL}	-793.34719	-802.59968
V_{IE}	-1329.53033	-1347.35455
T	384.69634	385.13867
V_{NN}	412.37630	421.61840
B.E.	1.88408	1.89448
R.E. (kJ mol^{-1})	25	53

Table 29

ORBITAL ENERGIES OF DEHYDRO[10]ANNULENES (a.u.)

I			II	
-0.313		π	-0.291	
-0.313		π'	-0.340	
-0.343		π	-0.342	
-0.455		π	-0.476	
-0.456	-0.688	π'	-0.480	-0.696
-0.479	-0.697	π	-0.487	-0.707
-0.537	-0.698	π	-0.557	-0.711
-0.547	-0.776		-0.563	-0.794
-0.558	-0.834		-0.574	-0.851
-0.585	-0.877		-0.585	-0.909
-0.589	-1.1(5)		-0.587	-1.1(5)
-0.648	-11.3(10)		-0.666	-11.3(10)

Table 30

AO POPULATIONS OF DEHYDRO[10]ANNULENES

	I					II				
	C1(5)	C2(4)	C3	H2	H3	C1	C2	C3	H2	H3
1s+2s	2.99	3.06	3.06	0.84	0.84	2.95	3.05	3.06	0.83	0.84
2p( )	2.05	2.07	2.07			2.09	2.08	2.09		
2p( )	0.97	1.05	0.97			0.98	1.04	0.97		
Total	6.01	6.18	6.10			6.02	6.17	6.12		

Table 31

EFFECT OF BASIS SCALING ON ACETYLENIC DIDEHYDRO[10]ANNULENE(II)

	Unscaled Basis		Partially Scaled Basis (Acetylenic C Unscaled)	
$-E_T$	380.62368		380.91350	
$-V_{EL}$	802.24208		802.53190	
$-V_{IE}$	1340.70994		1345.31032	
T	380.41091		383.48186	
$-\varepsilon_i$	π 0.362		π 0.309	
	π' 0.408		π 0.360	
	π 0.410		π' 0.365	
	π' 0.531		π, π' 0.494, 0.499	
	π 0.542	0.741	π 0.500	0.710
	π 0.546	0.755	π 0.572	0.724
	0.611	0.770	0.581	0.727
	0.623	0.857	0.590	0.812
	0.630	0.912	0.601	0.869
	0.640	0.978	0.602	0.930
	0.642	1.1(5)	0.688	1.1(5)
	0.735	11.4(10)		11.3(10)

AO POPULATIONS

	C1	C2	C3	H2	H3	C1	C2	C3	H2	H3
1s+2s	2.91	3.06	3.07	0.77	0.78	2.95	3.03	3.06	0.82	0.83
2p(σ)	2.12	2.14	2.15			2.14	2.07	2.10		
2p(π)	0.97	1.05	0.97			1.01	1.01	0.96		
	6.00	6.25	6.19			6.10	6.11	6.12		

Table 32

VERTICAL I.P.'s AND CALCULATED ORBITAL ENERGIES

FOR TRIDEHYDRO[12]ANNULENE

I_V(obs)	LCBO	SPINDO	MINDO/3	PPP	Xα	
7.69	8.44	8.35	7.49	7.51	6.79	$2a_2"(\pi)$
9.5-10.4	9.54	9.71-9.93	9.50-9.72	9.47	8.41-9.32	$2e",7e',5a_1^1$
11.67	11.96	11.63	12.17	11.78	10.27	1e"
12.4	13.06	12.39	13.29	12.52	10.98	$1a_2"$
13.2		13.06	13.23		12.44	6e'

Correlations: I.P. (Obs.) = 1.058 I.P. (SPINDO) - 0.869 eV

(Standard deviations in slope, intercept, overall are 0.023, 0.292, 0.211).

I.P. (Obs.) = 0.886 I.P. (Xα) + 2.186 eV

(S.D. : 0.037, 0.456, 0.404)

Table 33

TOTAL ENERGIES OF DEHYDRO[12]ANNULENES (a.u.)

	Acetylenic	Cumulenic
E_T	-456.52422	-456.40797
V_{EL}	-974.30703	-973.38612
V_{IE}	-1640.27800	-1638.62839
T	460.58360	460.46403
V_{NN}	517.78281	516.97815
B.E.	2.21643	2.10018
R.E. (kJ mol^{-1})	173	31

Table 34

ORBITAL ENERGIES OF DEHYDRO[12]ANNULENES (eV)

Acetylenic			Cumulenic		
π	8.528	$2a_2''$	7.178	π	$1a_1''$
π	10.757(2)		10.128	π'	
π'	11.050(2)		10.305(2)	π'	
π'	11.666		10.455(2)	π	
π	13.476(2)		13.476(2)	π	
π	14.491		14.498	π	
	15.334(2)		16.170(2)		
	16.006		16.723		
	17.452		17.222(2)		
	17.587(2)		17.443		
	19.007		18.828		
	19.728(2)		19.680(2)		
	22.450(2)		22.922(2)		
	23.545		24.974		
	26.849		25.729		
	27.779(2)		27.667(2)		
	29.956(2)		29.809(2)		
	30.448		30.399		
	308(12)		308(12)		

Table 35

AO POPULATIONS OF DEHYDRO[12]ANNULENES

	Acetylenic					
	C1	C2	H	C1	C2	H
1s+2s	3.07	2.98	0.82	3.06	2.98	0.83
	2.03	2.01		2.04	1.98	
	1.05	1.04		1.07	1.04	
	6.15	6.03		6.17	6.00	

Table 36

TOTAL ENERGIES OF DIDEHYDRO[14]ANNULENE (a.u.)

E_T	-534.35806
V_{EL}	-1252.68468
V_{IE}	-2143.47836
T	540.11902
V_{NN}	718.32662
B.E.	2.84146
R.E. (kJ mol^{-1})	325

ORBITAL ENERGIES

	ε_i(calc)/eV	
π	6.74	
π	7.81	
π'	11.14	
π'	11.21	
π	11.45	18.59
π	11.65	19.08
π	13.84	19.39
π	14.12	21.33
(CH)	14.46	21.72
π	14.78	22.12
	14.99	23.22
	15.01	24.99
	15.20	26.31
	16.07	27.03
	16.34	28.52
	17.25	29.19
	17.30	30.14
	18.58	30.80
		31.13
		307(12)

AO POPULATIONS

	C1	C2	C3	C4	H2	H3	H4
1s+2s	2.95	3.05	3.06	3.06	0.83	0.85	0.84
2p(σ)	2.08	2.09	2.09	2.09			
2p(π)	1.01	1.03	0.98	0.97			
Total	6.04	6.17	6.13	6.12			

Table 37

(a) MOLECULAR TOTAL ENERGIES AND DIPOLE MOMENTS

		(a.u.)	(Debye)
Cycl[3,3,3]azine		Alt. Geometry	Reg. Geometry
(a) Singlet	E_T	-512.72828	-512.71688
	V_{EL}	-1230.01436	-1235.59659
	V_{IE}	-2113.85563	-2124.45590
	T	518.34018	518.51420
	V_{NN}	717.28608	722.87971
	B.E.	2.65420	2.64280
	R.E. (kJ mol^{-1})	330	300
(b) Triplet	E_T	-	-512.66743
	B.E.	-	2.59335
	R.E. (kJ mol^{-1})	-	171
Cycl[3,2,2]azine		Exptl. Geometry	Reg. Geometry
	E_T	-436.01228	-436.00727
	V_{EL}	-994.95200	-988.96665
	V_{IE}	-1695.15400	-1683.62662
	T	441.52068	441.32672
	V_{NN}	558.93972	552.95938
	B.E.	2.15346	2.14847
	R.E. (kJ mol^{-1})	230	217
	Dipole moment	0.851	1.12
6-Aza-cycl[3,2,2]azine		Geometry A (r(C5-N6)=1.441)	Geometry B (r(C5-N6)=1.398)
	E_T	-451.90458	-451.90120
	B.E.	1.87765	1.87427
	R.E. (kJ mol^{-1})	195	186
	Dipole moment	1.646	1.548
		Indolizine	Imidazo[1,2-a] Pyridine
	E_T	-360.48442	-376.37263
	B.E.	1.84624	1.63651
	R.E. (kJ mol^{-1})	220	175
	Dipole moment	1.47	3.5

Table 37 (b) Experimental Ionisation Potentials and their Assignments (eV) with Calculated Orbital Energies

Cycl[3,2,2]azine

IP	$-\varepsilon_i$	Assignment (C_{2V})
7.63	8.12	$4b_1$
8.41	8.67	$2a_2$
9.26	10.70	$3b_1$ (LP_N)
10.77	13.18	$2b_1$
11.11	13.58	$1a_2$
12.85 (broad)	14.32, 14.75	$13b_2$, $18a_1$
	14.86, 15.13	$12b_2$, $17a_1$
13.59	16.24	$11b_2$
14.47	16.92, 17.16	$16a_1$, $10b_2$
15.15 (broad)	17.94, 18.38	$15a_1$, $14a_1$
15.56	18.54, 18.92	$1b_1$, $9b_2$
18.45	21.69, 22.11	$13a_1$, $8b_2$
	22.69, 23.85	$12a_1$, $7b_2$
22.15	26.64, 27.58	$11a_1$, $6b_2$
	28.33	$10a_1$

Table 37 (b) (contd.)

Cycl[3,2,2]azine

IP	$-\varepsilon_i$	Assignment (C_{2v})
7.63	8.12	$4b_1$
8.41	8.67	$2a_2$
9.26	10.70	$3b_1$ (LP_N)
10.77	13.18	$2b_1$
11.11	13.58	$1a_2$
12.85 (broad)	14.32, 14.75	$13b_2$, $18a_1$
	14.86, 15.13	$12b_2$, $17a_1$
13.59	16.24	$11b_2$
14.47	16.92, 17.16	$16a_1$, $10b_2$
15.15 (broad)	17.94, 18.38	$15a_1$, $14a_1$
15.56	18.54, 18.92	$1b_1$, $9b_2$
18.45	12.69, 22.11	$13a_1$, $8b_2$
	22.69, 23.85	$12a_1$, $7b_2$
22.15	26.64, 27.58	$11a_1$, $6b_2$
	28.33	$10a_1$

Cycl[3,3,3]azine

IP	$-\varepsilon_i$	Assignment (C_{3h})
5.87	5.86	a"
8.26	10.22	e"
9.66	11.84	a" (LP_N)
10.3	13.45	e"
10.69	13.51	e'
11.53	14.25	a'
12.01	15.25	e'
12.92	15.87, 16.67	a', a'
13.77	17.08	e'
14.52	17.70, 18.32	a", a'
15.53	19.15	e'
15.80	20.05	e'
19.91	23.52, 23.71	e', a'
22.68	28.02	e'
(23.58)	29.46, 30.80	a', e'

Table 37 (b) (contd.)

Indolizine

IP	$-\varepsilon_i$	Assignment (C_S)
7.24	7.53	a" (π)
8.50	9.14	a"
10.27	12.00	a"
10.96	12.99	a"
12.10	13.62	a' (σ)
12.75	14.31, 15.03	a', a'
13.44	16.36	a'
14.42	16.68, 17.05	a', a'
14.70	17.63, 17.96	a" (LP_N^{π}), a'
14.90	18.35	a'
15.68	19.65	a'
18.45	21.66, 22.08	a', a'
19.70	23.52	a'
21.76	26.24, 27.28	a', a'

Imidazo[1,2-a]pyridine

IP	$-\varepsilon_i$	Assignment (C_S)
8.19	8.62	a" (π)
9.08	9.77	a"
10.09	11.33	a' (LP_N^{σ})
11.07	13.07	a"
11.41	13.88	a"
12.64	14.41	a'
13.50	15.42	a'
13.94	15.87	a'
15.00	17.21, 17.48	a', a'
16.02	18.51, 18.55, 18.90	a', a", a'
17.03	20.29	a'
19.20	22.6, 23.09	a', a'
20.18	24.10	a'
22.73	27.98, 28.13	a', a'

Table 37 (b) (contd.)

6-Azacycl[3,2,2]azine

IP	$-\varepsilon_i$	Assignment (C_{2V})
7.65	8.59	b_1 ( )
8.51	9.07	a_2 ( )
9.30 (shoulder)	11.33	a_1 ( )
9.50	11.13	b_1 ( )
11.00	13.71	b_1
11.90	13.97	a_2
13.07	14.78	b_2 ( )
13.49	15.40	b_2
13.92	15.46	a_1
15.1	16.60, 17.15	b_2, a_1
15.35	17.66, 17.67	a_1, b_2
16.06	18.94, 18.95, 19.29	a_1, b_1, b_2
17.05	22.22, 22.49	a_1, b_2
18.02	23.48	a_1
19.28	24.24, 24.48	b_2, a_1

Ionisation Potentials

3-Methylcyclopenta[cd]cycl[3,3,3]azine

IP	6.37	7.35	8.41	9.15	9.55	10.33	12.1	14.0

Cyclopenta[h]cycl[4,2,2]azine

IP	7.06	7.63	8.90	10.05	14.47	12.58	13.7	15.07
	16.46	18.4						

6,8-Dimethylcyclopenta[h]cycl[4,2,2]azine

IP	6.99	7.57	8.82	9.9

Table 37 (c) Calculated Orbital Energies, by Symmetry Type ($-\varepsilon_i$)

Cycl[3,2,2]azine - Experimental geometry (C_{2V})

σ		π	
a_1	b_2	b_1	a_2
425.7	308.6, 307.5	18.54 (LP_N^{π})	13.58
308.6, 308.6	307.0, 306.4	13.18	8.67
307.5, 307.1	32.00	10.70	
307.0, 306.4	27.58	8.12	
37.67	23.85		
31.02	22.11		
28.33	18.92		
26.64	17.16		
22.69	16.24		
21.69	14.86		
18.38	14.32		
17.94			
16.92			
15.13			
14.75			

Cycl[3,2,2]azine - Benzene-like Geometry (C_{2V})

σ		π	
a_1	b_2	b_1	a_2
425.0	308.4, 307.4	17.88	13.18
308.4, 307.7	307.0, 306.3	13.45	8.56
307.4, 307.4	31.30	10.52	
307.0, 306.3	27.31	8.12	
36.70	23.80		
31.08	21.70		
28.24	19.02		
26.52	16.85		
22.64	15.95		
21.36	15.21		
18.36	13.81		
18.04			
16.76			
14.98			
14.38			

Table 37 (C) (contd.)

Cycl[3,3,3]azine - Alternating Geometry (C_{3h})

σ		π	
a'	e'	a"	e"
426.7	309.2, 307.3	17.70 (LP_N^{π})	13.45
309.1, 307.3	306.5, 305.8	11.84	10.22
306.5, 305.8	30.82	5.86	
36.06	28.02		
29.46	23.52		
25.03	20.05		
23.71	19.15		
18.32	17.04		
16.67	15.25		
15.87	13.51		
14.26			

Cycl[3,3,3]azine - Benzene-like Geometry (D_{3h})

σ			π		
a_1'	a_2'	e'	$a_1"$	$a_2"$	e"
427.9	305.9	309.6, 307.6	5.46	18.25	13.65
309.6, 307.6	24.80	305.9, 305.9		12.00	10.23
305.9	16.81	31.17			
36.92	14.25	28.19			
29.72		23.99			
24.24		20.29			
18.37		19.33			
15.96		17.08			
		15.29			
		13.75			

Table 37 (c) (contd.)

Indolizine - Constructed geometry (C_S)

σ		π	
a'	a'	a"	
426.4	22.08	17.63	(LP_N^{π})
308.6, 307.9	21.66	12.99	
306.8, 306.8	19.65	12.00	
306.7, 306.6	18.35	9.14	
305.9, 305.5	17.97	7.53	
36.46	17.05		
30.95	16.68		
29.72	15.36		
27.28	15.03		
26.24	14.31		
23.52	13.62		

Imidazo[1,2-σ]pyridine - Constructed geometry (C_S)

σ		π	
a'	a'	a"	
426.6, 422.9	23.09	18.55	(LP_N^{π})
309.7, 308.8	22.61	13.88	
307.6, 307.3	20.29	13.07	
307.3, 307.1	18.90	9.77	
306.9	18.52	8.62	
38.18	17.49		
33.77	17.21		
30.96	15.87		
28.13	15.42		
27.98	14.41		
24.10	11.33	(LP_N^{σ})	

Table 37 (C) (contd.)

Pyrazino[6,1,2-cd]pyrrolizine (6-Azacycl[3,2,2]azine)-

Assumed Geometry (C_{2V})

σ		π	
a_1	b_2	b_1	a_2
426.2, 423.5	309.0, 308.6	18.95 (LP_N^{π})	13.97
309.0, 308.6	307.3, 306.8	13.71	9.07
308.0, 307.3	32.41	11.13	
306.8	28.07	8.59	
38.13	24.24		
32.90	22.49		
29.33	19.29		
24.48	17.67		
23.48	16.60		
22.22	15.40		
18.94	14.78		
17.66			
17.15			
15.46			
11.33			

Table 37(d) - Population Analysis

Cycl[3,2,2]azine

	C-1	C-2	C-2a	C-4a	C-5	C-6	N-8
1s + 2s	3.0298	3.0581	2.9471	2.9524	3.0573	3.0682	3.3453
$2p_\sigma$	2.0538	2.0878	1.8451	1.8923	2.0852	2.0543	2.4085
$2p_\pi$	1.0996	1.0019	1.1318	1.0543	0.9990	1.0174	1.5413
TOTAL	6.1831	6.1479	5.9240	5.8989	6.1414	6.1399	7.2951
	H-1	H-2	H-5	H-6			
$1s_H$	0.8387	0.8498	0.8366	0.8482			

6-Azacycl[3,2,2]azine

	C-1	C-2	C-2a	C-4a	C-5	N-6	N-8
1s + 2s	3.0311	3.0585	2.9486	2.9539	3.0565	3.6002	3.3455
$2p_\sigma$	2.0566	2.0882	1.8464	1.8963	1.9605	2.6363	2.4083
$2p_\pi$	1.0899	0.9987	1.1277	1.0590	0.9998	1.0435	1.5407
TOTAL	6.1776	6.1446	5.9178	5.9103	6.0167	7.2799	7.2945
	H-1	H-2	H-5				
$1S_H$	0.8449	0.8342	0.8256				

Cycl[3,3,3]azine (D_{3h})

	C-1	C-2	C-3a	N-10
1s + 2s	3.0142	3.0792	2.9906	3.3791
$2p_\sigma$	1.9990	2.1561	1.9302	2.4240
$2p_\pi$	1.2012	0.8582	0.9112	1.4885
TOTAL	6.2144	6.0932	5.8319	7.2886
	H-1	H-2		
$1S_H$	0.8546	0.8399		

Cycl[3,3,3]azine (C_{3h})

	C-1	C-2	C-3	C-9a	N-10
1S + 2s	3.0388	3.0765	3.0214	2.9921	3.3690
2p	2.0267	2.1304	2.0043	1.9139	2.3663
2p	1.1099	0.9056	1.1824	0.9441	1.5737
TOTAL	6.1754	6.1125	6.2081	5.8501	7.3091
	H-1	H-2	H-12		
$1S_H$	0.8482	0.8442	0.8586		

Table 37(d)(contd.)

Indolizine

	C-1	C-2	C-3	C-5	C-6		
1s + 2s	3.0462	3.0620	3.0206	3.0110	3.0446		
$2p_\sigma$	2.0406	2.0820	1.8917	1.9760	2.0779		
$2p_\pi$	1.1322	1.0332	1.1480	0.9998	1.0635		
TOTAL	6.2191	6.1772	6.0603	5.9868	6.1860		
	C-7	C-8	C-8a	N-4			
	3.0421	3.0415	2.9762	3.3667			
	2.1197	2.1012	1.8392	2.4417			
	1.0038	1.0212	1.0724	1.5258			
	6.1656	6.1639	5.8879	7.3343			
	H-1	H-2	H-3	H-5	H-6	H-7	H-8
$1S_H$	0.8367	0.8409	0.8262	0.8184	0.8317	0.8344	0.8306

Imidazo[1,2-a]pyridine

	N-1	N-4	C-2	C-3	C-5		
1s + 2s	3.5521	3.3689	3.0038	3.0481	3.0039		
$2p_\sigma$	2.7162	2.4292	1.9409	1.9810	1.9567		
$2p_\pi$	1.1155	1.5374	1.1171	1.0234	1.0305		
TOTAL	7.3835	7.3355	6.0619	6.0525	5.9912		
	C-6	C-7	C-8	C-8a			
	3.0480	3.0355	3.0475	2.9021			
	2.0856	2.1049	2.1078	1.7319			
	1.0443	1.0220	0.9985	1.1115			
	6.1779	6.1624	6.1537	5.7455			
	H-2	H-3	H-4	H-5	H-6	H-7	
$1S_H$	0.8328	0.8004	0.8247	0.8292	0.8325	0.8163	

CHAPTER 6

APPLICATION TO SOME HOMOAROMATIC SYSTEMS AND RELATED SPECIES

"The original structural prerequisite for the 4q + 2 rule has also been subjected to extensive variations. Only the essentially pericyclic π-electron topology has been left more or less intact".

M.J. Goldstein and R. Hoffmann

In this Chapter there are reported calculations on some 7- and 9-membered ring compounds. Neutral hydrocarbons of this type cannot be completely conjugated systems. The concept of Homoaromaticity, advanced about twenty years ago[1], has proven inspirational to many chemists and has had the effect of directing much experimental and theoretical attention to this field[2]. The varied success encountered in the observation of homoconjugative delocalisation within certain ions and molecules of predesigned structure has also generated some consternation. At issue is the rather qualitative question: how extensively can π-electron delocalisation in (4n+2) conjugated polyenes be interrupted by one or more saturated atoms (sp^3 centres) with continued maintenance of some form of ring current? The idea of homoaromatic orbital overlap is illustrated by Figure 1. The familiar parallel alignment of π-orbitals characteristic of aromatic systems, pp-π overlap, is shown. This combination possesses a node in the plane of the bond axis, has angular momentum about this axis, and is doubly degenerate. When a (4n+2) cyclic array of these orbitals is fractured to accommodate an sp^3-hybridised centre, electron delocalisation can only be maintained if the two flanking π-atomic orbitals become canted, so that overlap becomes restricted to single lobes, the boundaries of which are limited to that surface of the molecule opposite to the bridging atom. The extent of interpenetration of such

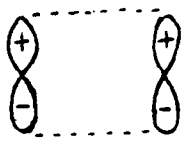
opposed p_z orbitals can be expected to depend heavily upon their mutual orientation and the internuclear distance separating the trigonal atoms. This orbital overlap lies intermediate between pp-π and pp-σ.

In this work, two of the simplest types of homoaromatic system are considered; these are neutral species, the monohomoaromatic 1,3,5-cycloheptatriene and the trishomoaromatic 1,4,7-cyclononatriene. Some heterocycles derived from these parents are also treated. Strictly speaking, only the two hydrocarbons can be classed as homoaromatic systems; derivatives obtained by simple substitution with heteroatoms can be usefully considered in the light of results obtained on the parent species, although they are technically classifiable as heteroaromatics. The designation of mono-, bis-, tris-homoaromatic is based on the number of sides of the ring where σ-skeleton is interrupted (removed or lengthened), and not on the number of methylene groups inserted on any particular side.

A. 7-Membered Rings.

(a) 1,3,5-Cycloheptatriene.

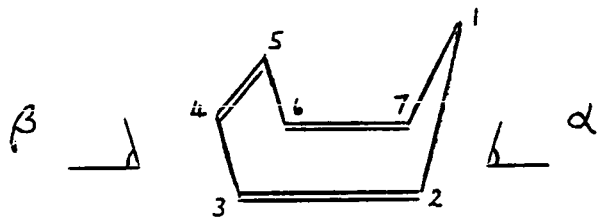
The first members of the completely conjugated annulene series considered in Chapter 5 are the 6- and 8-membered ring species. The intervening 7-membered ring compound, cycloheptatriene, is now considered; by necessity, this neutral hydrocarbon cannot be completely conjugated, and it incorporates a saturated carbon centre. This species is the classic (simplest) homoaromatic compound, formally derived from the (4n+2) conjugated polyene, benzene, by interrupting the carbon periphery with a bridging, saturated centre; it is monohomoaromatic as the σ-skeleton is interrupted on one side of the ring. Formally, there is expected to be typical conjugative interaction among the π-bonds, but some of the overall benzene effect is replaced by homoconjugative interaction between the π-bonds spatially separated by the sp^3-centre; this latter across-space interaction can lead to a closed loop over which all the π-electrons (6) can be delocalised. All-cis cycloheptatriene was synthesised some time ago[3], and has been studied extensively since, as it is a stable compound and reasonably easily handled experimentally. Some observed properties are given in FIG. 1. In view of its experimentally observed behaviour after its synthesis, there was much conjecture over its molecular and electronic structure, whether it has a planar carbon skeleton (annulene-like) or a nearly planar one (especially formally conjugated part of the molecule). Subsequently, an electron diffraction investigation of the geometrical structure was carried out, and this showed that the molecule in the gas phase adopts a "boat"conformation (Figure 1);[4] the geometrical parameters, bond lengths and angles, are effectively "classical". The boat conformation was confirmed by a microwave spectroscopic investigation;[4] there is not sufficient information provided to determine the geometrical parameters



"AROMATIC" pp-π OVERLAP

"HOMOAROMATIC" pp-π (LOCAL) OVERLAP

(a)



CYCLOHEPTATRIENE

"BOAT" CONFORMATION

$r(C_1-C_2) = 1.505$, $r(C_3-C_4) = 1.446$

$r(C_2=C_3) = 1.356$, $r(C_4=C_5) = 1.356$, $r(C-H) = 1.095$

$r(C-C^{ME}) =$

$C1\hat{C}2C3 = 121.8°$; $C2\hat{C}3C4 = 127.2$

E.D. $\alpha = 40.5$, $\beta = 36.5$

MWAVE $\alpha = 50$, $\beta = 29.5$

(b)

NMR : INVERSION BARRIER = 25 $KJmol^{-1}$.

DIHEDRAL ANGLE = 110°

(c) NORCARADIENE STRUCTURE.

(d) METHYLENE $C1(H_2)$ REPLACED BY N-H : 1H-AZEPIN

O : OXEPIN

S : THIEPIN

FIGURE 1 : HOMOAROMATIC DATA.

but, using the bond parameters derived from the electron diffraction study, it has proved possible to deduce the two dihedral angles (α and β) which characterise the boat. As shown in Figure 1, there is a significant difference between the two sets of values for the pair of dihedral angles. Neither the geometrical structure nor other observed properties (such as NMR data) are indicative of departure from a basically classical model for the structure of cycloheptatriene. Similarly, theoretical studies have led to the conclusion that the species is non-aromatic.[5] The experimental information which is particularly relevant in this work and for investigating effects such as interaction among double bonds is the photoelectron spectrum; it is particularly surprising that a satisfactory P.E. spectrum of cycloheptatriene has not been published and the results discussed, as the measurement is a straightforward exercise. Thus, the He I spectrum has been measured and is presented in Figure 2.

Several calculations using the standard minimal basis have been performed. Before considering the non-planar "boat" type of structure as found from experiment, structures with a planar carbon skeleton were constructed. One of these "planar" geometries is regular in C-C bond length and bond angle, which is rather unrealistic even for the optimum planar form ; the other geometries are alternating in C-C bond length (with lengths based on those found experimentally), with some variation in CCC bond angle considered. The internal angle of a regular heptagon is 128.6°. The optimum structure is expected to have a rather smaller value than this for the angle at the methylene carbon (unstrained value is about 110°). The results of the calculations on the planar structures are given in Table 1. Some idea of the variation of total energy with geometrical parameters is obtained; the partial optimisation indicates that the optimum angle at the methylene carbon is likely to be in the

FIG. 2a

P.E. SPECTRUM
OF C_7H_8

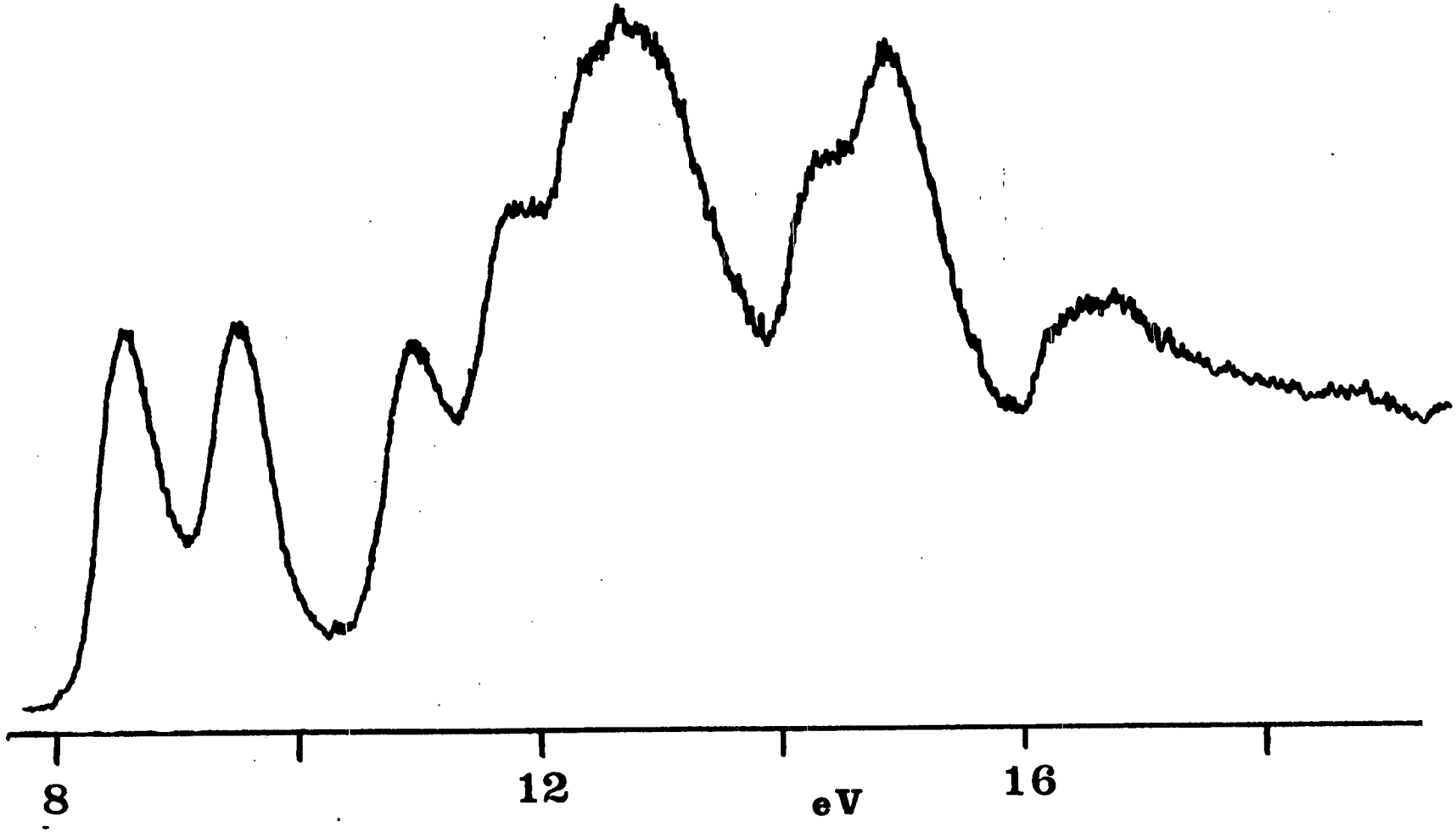
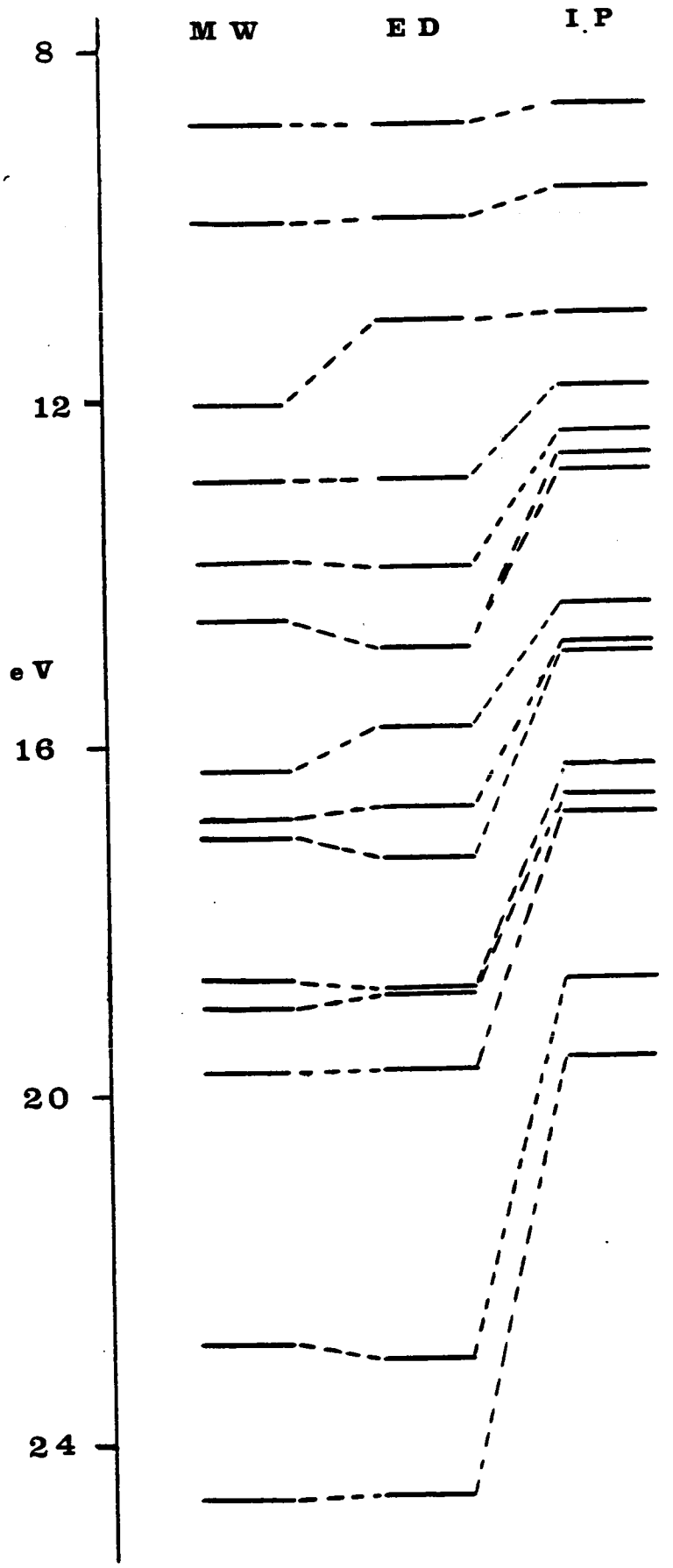


FIG. 2b : C_7H_8 ORBITAL ENERGY CORRELATION.



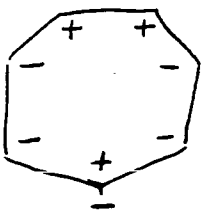
vicinity of 120°. As well as furnishing the basic forms of the MO's of the cycloheptatriene species when considering non-planar structures, the planar form does have physical significance. Under typical conditions, the boat form undergoes inversion, the simple mechanism involving a "planar" intermediate (cf. cyclooctatetraene in Chapter 5). From NMR data, the value of the barrier to inversion has been estimated at 25 kJ mol^{-1} [6].

Non-planar structures based on the two sources of experimental data mentioned above have been constructed, and calculations performed. The computed total energy quantities for the two geometries are presented in Table 2. The total energy of the microwave structure is significantly lower than that of the electron diffraction one; in fact, the latter is higher than the "classical" planar structures considered here. The two "boat" structures actually differ only in the values of the dihedral angles (α,β), although there is also variation in the bond angles which are not independent. It happens that the difference in total energy between the microwave structure and the best planar structure involved here is close to the experimentally observed inversion barrier; the calculations give a "vertical" inversion barrier height, meaning the bond lengths remain constant and only bond angles vary during the process. A complete theoretical examination of the inversion would involve optimisation of the boat and planar forms, as with cyclooctatetraene, and also some indication of the variation of the total energy of the molecular system with α and β.

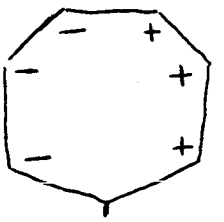
The computed orbital energies of the two boat structures are given in Table 3, along with the observed IP's derived from the measured HeI spectrum. The assignment of the spectrum is quite straightforward, and there is typical agreement between calculated and observed MO levels; the microwave structure correlates more closely with the observed IP's, although it is noticeable that the two calculated spectra are generally

very similar, with only a few pairs of corresponding levels differing significantly. On both a total energy and an orbital energy basis, it does seem likely that the values of α and β from the microwave study are closer to the actual equilibrium ones than those of the electron diffraction study. The variation of the orbital energies with geometry (planar and non-planar) can be understood in the usual qualitative way; on moving from planar to non-planar structures, there is an extension of the cyclooctatetraene situation as there are now two dihedral angles. As usual, the π-type MO's are of particular interest. Considering first the planar form, there are four π-MO's whose forms are illustrated in Figure 3 (there is some "contamination" of these MO's by the methylene hydrogens). The HOMO consists of basically four π-units (3 double bonds + methylene group), each bonded unit being anti-bonding with respect to its nearest neighbours; the calculated I.P. of this MO is quite low, and is separated by a very large gap (over 3 eV) from the next MO. The latter is almost completely based on two isolated doubly-bonded units (opposing double bonds, across the ring), with only a small contribution from the third double-bond centres, it is nodal at the methylene carbon. The calculated orbital energy is almost identical with that of an isolated ethylene molecule π-MO, consistent with the lack of interaction of the π-units of the 7-membered ring. The next three innermost MO's are effectively degenerate, accidentally. One is a π-MO, being completely bonding around the fully conjugated part of the ring, but antibonding between the latter and the methylene group. The orbital energy is substantially lower than that of the A_2 π-MO (by about 2 eV). These three π-MO's are formed predominantly from the AO's of the carbons of the conjugated part of the system. However, the energy splitting of the three levels is not a direct measure of the extent of interaction of the double bonds (so-called homoaromatic effect), as there is a significant contribution from the methylenic carbon AO in the

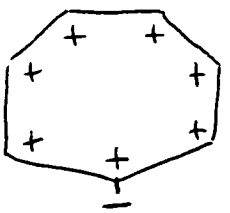
Figure 3. The π-M.O.'s of Cycloheptatriene (planar)



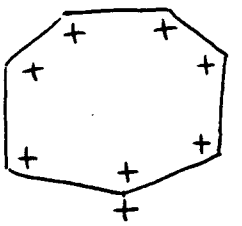
25 (b_1)



24 (a_2)



22 (b_1)



16 (b_1) (CH_2)

two B_2 π-MO's. The fourth π-MO is predominantly the methylenic unit π-MO, and lies well in amongst the σ-MO's. The B_2 π-MO's can be regarded as being formed overall from the interaction between the methylenic π-unit and the conjugated system π-units. A larger interaction is expected between the methylene and the completely bonding combination of the double-bond units, which are reasonably close in energy; the energies of these hypothetical entities are not known exactly, but can be estimated to be about 16 eV and 13 eV respectively. The symmetric combination of the two interacting units ($1B_2$) is stabilised then by about 0.5 eV, and the antisymmetric combination ($2B_2$) destabilised by about this amount. The HOMO ($3B_2$) is probably destabilised by a small amount also. The overall conclusion is that the π-MO energies do indicate that there is significant interaction among the double bonds, the extent being comparable to that found in benzene.

A plausible conformation for cycloheptatriene is the norcaradiene one (Figure 1), which has been found to be that of the seven-membered ring in some derivatives, although it is unknown for the parent species. A norcaradiene structure has been constructed, and a calculation performed, yielding the results in Tables 2,3. The total energy is not actually lower than that of boat cycloheptatriene. However, in reality, there is likely to be a large barrier to interconversion. The norcaradiene form is inaccessible experimentally, presumably as it is of higher energy than another tautomer, toluene, and conversion to the latter is not hindered. The computed orbital energy spectrum of the norcaradiene form does not resemble the experimental one at all, confirming the absence of the species.

(b) 1H-Azepin .

The seven-membered unsaturated heterocycles, 1H-azepin, oxepin, and thiepin (Figure 1) are of substantial intrinsic interest. This group of compounds is characterised by a cyclic array of 8 π-electrons and therefore can be expected to differ chemically from related monocyclic systems with (4n+2) π-electrons. Although planarity is feasible in such molecules, it would seem unlikely that maintenance of such a conformation would be energetically rewarding (anti-aromatic system) 1H-Azepin is unknown experimentally. A large variety of rather complex substituted species are known[7]. The geometry of some derivatives have been determined by X-ray diffraction, and the azepin ring has been found to adopt a boat conformation. Calculations on non-planar azepins have been performed, and also on planar species. The total energy of the latter always tended to be lower. As with other cases, relatively minor modifications in geometry (bond length and angle) lead to small changes in total energy. Variation of the dihedral angles of the boat tends to have greater effect on total energy. Some attempt at rigorous geometry optimisation of planar and boat forms has been attempted, using partially optimised initial structures. The results are summarised in Table 4. The planar species, after 3 cycles, is probably in a convergent situation. Only one cycle has been performed on the non-planar form, whose total energy is still above the best planar form, although it seems likely that the non-planar form total energy may fall below that of the planar form on further optimisation.

B. 9-Membered Rings .

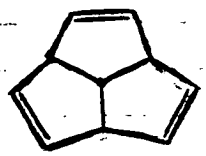
(a) Cis-cis-cis-1,4,7-Cyclononatriene .

This molecule is a classical example of a neutral system which satisfies the conditions regarded as necessary for the delocalisation of π-electrons among weakly interacting double bonds, the situation covered by the catchword "homo-aromaticity". Homoconjugative interaction among π-bonds, spatially separated from each other by sp^3-hybridised centres, should give rise to substantial first-order effects if the π-bonds are arranged in such a way that homoconjugation (across-space interaction) between nearest neighbours leads to a closed loop over which the π-electrons can be formally delocalised; in addition, if the Hückel rule is invoked, the number of interacting π-bonds must be 2n+1 (n=1,2,3,...) to give a favourable number of delocalisable π-electrons. All-cis cyclononatriene, synthesised first about fifteen years ago[8], is shown in Figure 4 ; the isolated double bonds of the 6π-electron framework are arranged in favour of homoconjugation. The molecular structure has been determined by X-ray analysis at -35°C, and that of the silver nitrate adduct at -125°C, with both determinations leading to practically the same results for the conformation of cyclononatriene, the "crown" of Figure 4[9]. Neither the interatomic distances (r(C=C) = 1.34 Å) nor the bond angles (C(CH_2)C = 108°) are indicative of any significant departure from a model which assumes three non-interacting double bonds. Furthermore, NMR data (proton chemical shifts) and measured heat of hydrogenation (ΔH = -323 kJ mol^{-1} as compared to -99 kJ mol^{-1} for cis-cyclononene) are regarded as confirming the absence of a significant interaction between the three double bonds. In addition, on the basis of a theoretical procedure based on molecular topology (generalisation of Hückel treatment) for analytically estimating Dewar-type resonance energies, it has been concluded that cyclononatriene is substantially

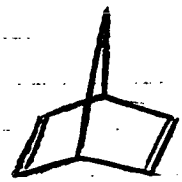
$r(CH) = 1{\cdot}09\,Å$

FIGURE 4 : "CROWN" CONFORMATION OF ALL-CIS CYCLONONATRIENE.

"EXPANDED" BENZENE



TRIQUINACENE.



NORBORNADIENE.

"EXPANDED" CYCLOBUTADIENE.

"SADDLE" CONFORMATION OF CYCLONONATRIENE.

non-aromatic, with a resonance energy of 3-4% that of benzene[10]. The reason for the failure of the experimental measurements to yield an indication of the presence, let alone the extent, of an interaction between the three double bonds in cyclononatriene has been illustrated by simple MO arguments, where the interaction between pairs of bonds is described by a resonance integral $\beta_{uv} = m\beta$ (u,v = 2,3; 4,5; 6,1 in Figure 4), where $0 \leqslant m \leqslant 1$. β is the typical Hückel parameter - to each of the double bonds a bonding (π_j) and an antibonding (π_j^*) linear combination of 2p-AO's is assigned, with orbital energies $\varepsilon_j = \alpha+\beta$ and $\varepsilon_j^* = \alpha-\beta$[11]. m=0 corresponds to three independent double bonds, and m=1 is the situation formally identical with that prevailing in benzene. Defining the total π-electron energy in the usual way (as a function of m) for the independent electron model, and the delocalisation energy, it follows that the delocalisation energy of cyclononatriene is less than 10 kJ mol^{-1} for reasonable values of m (about 8% of benzene value); this quantity is much too small to be detectable in view of the steric and conformational strain regarded as being present in the molecule. It is concluded that for such homoaromatic systems photoelectron spectroscopy is the method of choice for the direct determination of the magnitude of homoconjugative interaction, characterised by the parameter m in the simple model above.

The photoelectron spectrum of cyclononatriene is reproduced in Figure 5, and the observed ionisation potentials are given in Table 5, along with the orbital energies computed from a calculation using the standard minimal basis set and experimental geometric data (Figure 4). The linear correlation between observed and calculated I.P.'s, using Koopman' approximation, is shown in Figure 5. Assignment of the spectrum is reasonably straightforward.

In homoaromaticity considerations attention is concentrated on the first two P.E. bands (lowest I.P.). These correspond to ionisation

FIG. 5a

P.E. SPECTRUM OF C_9H_{12}

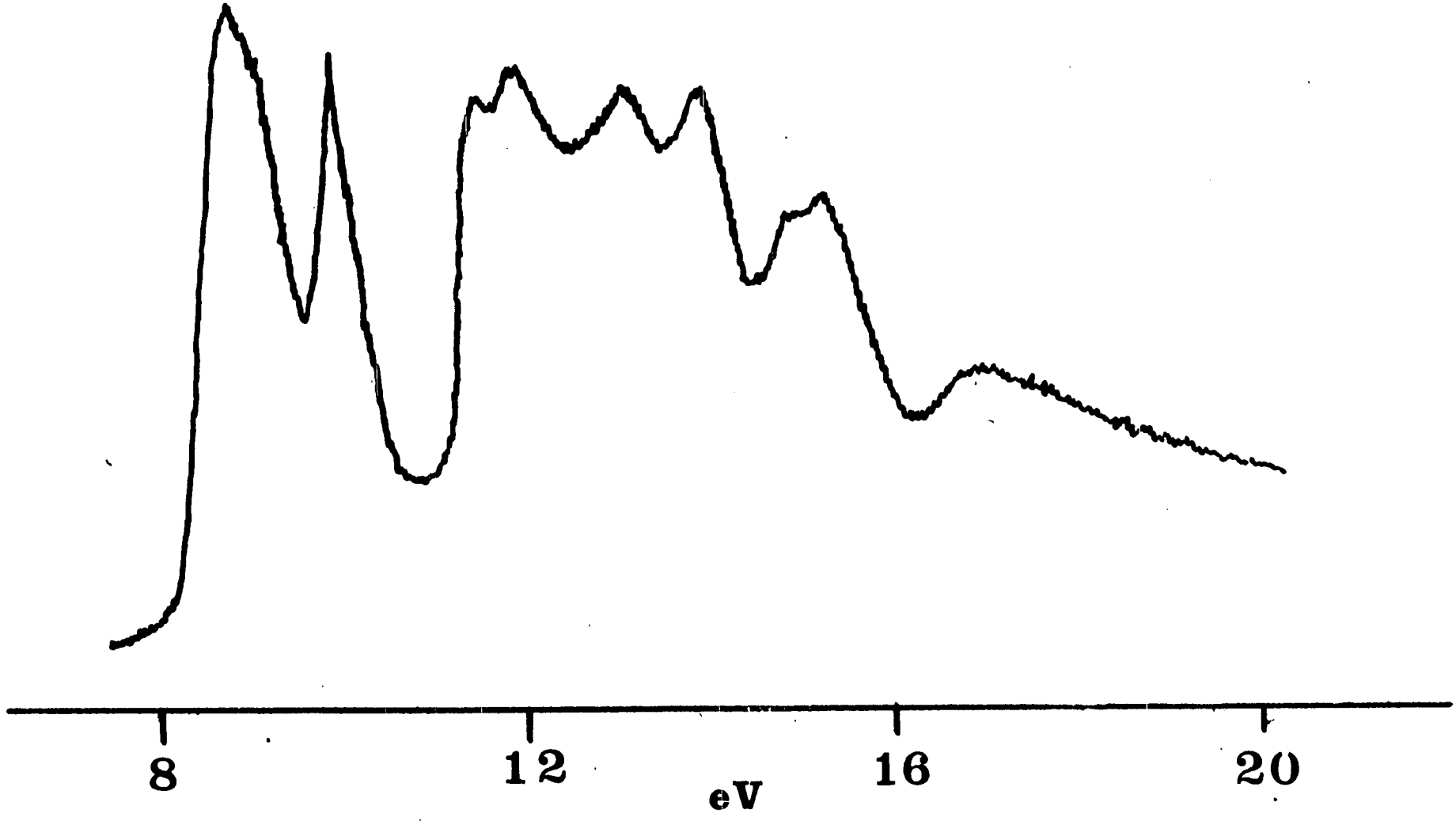
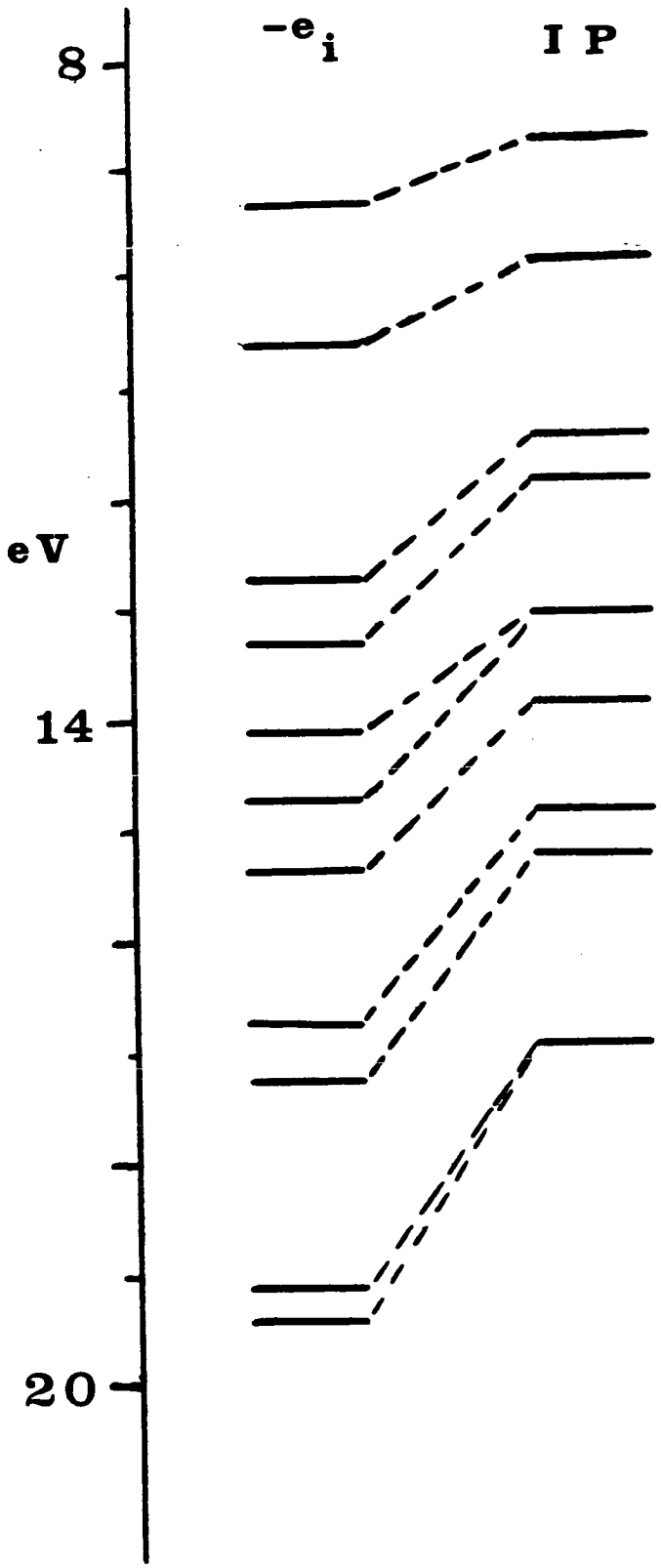


Fig. 5b : C_9H_{12} ORBITAL ENERGY CORRELATION.



processes for which the electron leaves either the degenerate e (π) orbitals (band 1) or the orbital a_1 (π) (band 2). Vertical ionisation potentials (positions of band maxima) are given in Table 5 . The complex structure of the first band, especially its shape with a flattened top, suggests strongly that the radical cation of cyclononatriene undergoes a Jahn-Teller splitting in its ^2E ground state; the first of the overlapping bands is located quite accurately at 8.77 eV, but the second is rather difficult to locate and the value given (8.9-9.0 eV) is somewhat uncertain. An average value of 8.9 eV is sufficient here where an exactly degenerate pair of orbitals is considered. The second band assigned to a π-type orbital is at 9.80 eV; this is ascribed to the cation in its first electronically excited $^2A_1$ state. Assuming the validity of Koopmans' approximation (and symmetrical Jahn-Teller split of the ^2E state), then

$$E^T(^2E) - E(^2A_1) = \varepsilon(e(\pi)) - \varepsilon(a_1(\pi)) \approx 0.9 \text{ eV} .$$

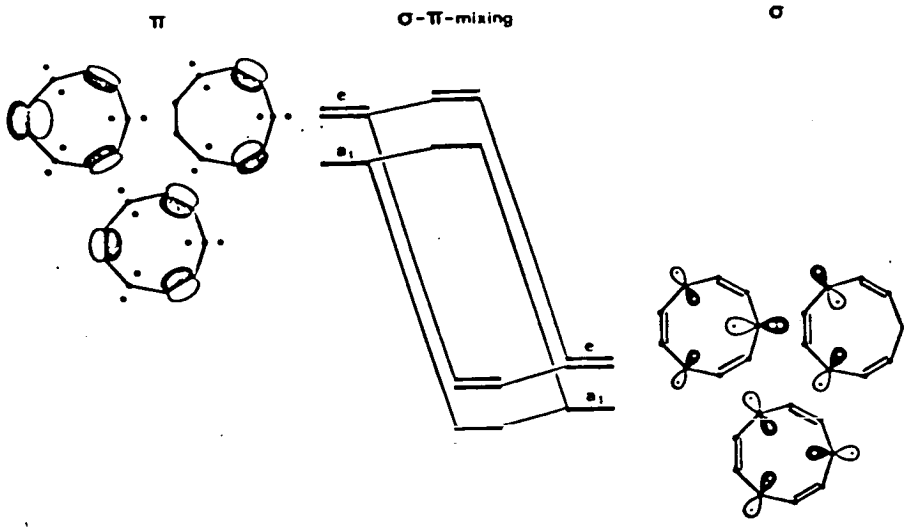
In benzene, the corresponding observed split is about 2.5 eV. According to a Hückel-type model, this value is equal to β; it can be shown that for the situation in cyclononatriene the split is given approximately by the quantity $3m\beta/2$. Thus gives a value for the parameter m of 0.24. Interpretation of the P.E. spectra of homoaromatic compounds such as cyclononatriene has been made on the basis of a simple Hückel model[11]. In cyclononatriene the interaction of the three double bonds is predicted to be almost pure homoconjugation, i.e. through-space as postulated in the naive treatment characterised by the expanded benzene structure of Figure 4 . In addition, the possibility that the resulting orbitals e(π) and $a_1(\pi)$ can mix with those lower-lying σ-orbitals of the methylene C-H bonds that belong to the same irreducible representations of C_{3V}. This type of mixing, which is inherent to the MO forms from non-empirical all-electron calculations, is usually given the label of through-bond

interaction (hyperconjugation), which does not seem to be particularly meaningful, even allowing for its qualitative nature. Figure 6 shows diagramatically the σ-π interaction; the orbitals $e(\sigma)$ and $a_1(\sigma)$ are linear combinations of the locally antisymmetric methylene orbitals, i.e. $a.2p(c) + b.[1s(H_{endo}) - 1s(H_{exo})]$. It is concluded, on the basis of the simple model, that the interaction of $e(\pi)$ with $e(\sigma)$ and of $a_1(\pi)$ with $a_1(\sigma)$ is small and does not result in a significant reduction of the original split $\varepsilon(e) - \varepsilon(a_1)$. It is concluded that the value of $m = 0.24$ is representative for the homoconjugative, through-space interaction of two π-bonds in a relative conformation such as present in all-cis cyclononatriene. By considering other possible relative conformations, it is also concluded that even the use of P.E. spectroscopy for the measurement of homoconjugative interaction is tied to special conditions: only when the basis orbitals π_j of the molecule under consideration are symmetry-equivalent or if their orbital energies are accidentally (almost) degenerate, will large first order changes be observable in the P.E. spectra of such compounds.

Considering the results of the non-empirical calculations reported here, separation into quantities corresponding to simple qualitative ideas cannot rigorously be effected. At the particular molecular geometry used, the calculated resonance energy is very small, being about -10 kJ mol^{-1}; it is unlikely that complete geometry optimisation would lead to a value which is much larger than this so that a figure of about 200 kJ mol^{-1} (as in benzene) probably would not be approached. The calculated orbital energy splitting, $\varepsilon(e(\pi)) - \varepsilon(a_1(\pi))$ is 1.26 eV. This is larger than the observed, but is in keeping with the trend of calculated orbital energies being larger than observed I.P.'s and the linear correlation between these; thus, the ratio of this split to that between $\varepsilon(a_1(\pi))$ and the orbital energy of the "σ onset" (2.15 eV) is quite close to the observed

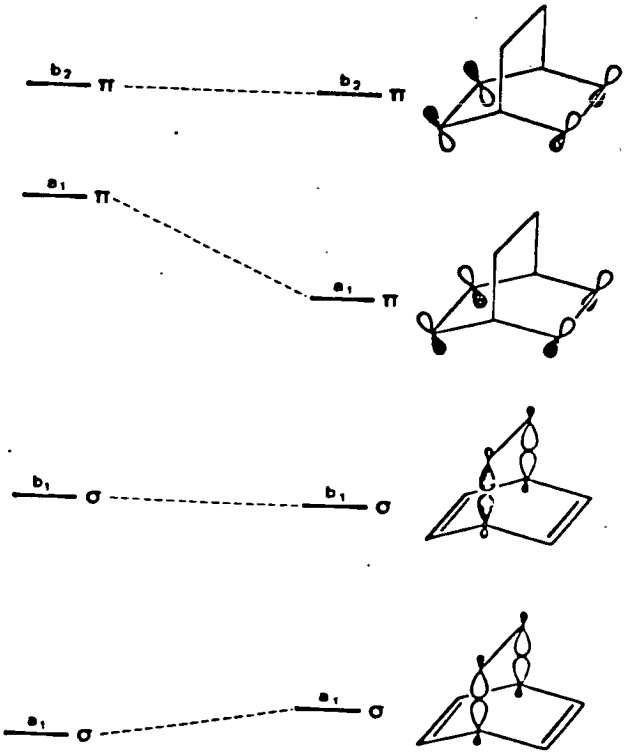
Figure 6.

(a)



σ-π-interaction diagram for cis-cis-cis-*1,4,7-cyclononatriene* (*1*)
The σ-orbitals shown correspond to the locally antisymmetric combinations of the C–H-σ-orbitals (the dots signify the hydrogen nuclei)

(b)



Schematic representation of the interaction of σ- and π-orbitals in bicyclo[2.2.2]*octadiene* (*14*)
Right: non-interacting orbitals. Left: orbitals after interaction. The symmetry classification is with respect to the group C_{2v}

NORBORNADIENE IS ANALOGOUS.

ratio (about 1:1.5). Similarly, the ratio of calculated $\varepsilon(e) - \varepsilon(a_1)$ splits in cyclononatriene and benzene is close to the observed one. In an attempt to interpret the cyclononatriene situation on a non-empirical calculational basis, the expanded Kekulé benzene structure of Figure 4 was considered. In "normal" benzene, or even in a Kekulé structure of reasonable alternating character (e.g. r(C=C) = 1.34 Å, r(C-C) = 1.50 Å), the e and a π-levels can be considered to arise from the interaction of the three double bonds according to the usual orbital splitting diagram (Chapter 4). From the calculations reported in Chapter 4, the effect of molecular geometry variation in benzene on the relative orbital energies can be easily seen and interpreted. As the alternating character of the structure is increased, the a π-orbital is destabilised (completely bonding round the ring so that favourable overlap of C 2p orbitals is reduced), and the e π-orbital pair is stabilised (less marked effect). The splitting is thus decreased. At the rather extreme alternating structure of the benzene part of the cyclononatriene structure, the splitting of the two π-levels is expected to be small, as the situation becomes closer to that of three isolated acetylene molecules; a separate calculation on this benzene molecule was performed to obtain a value for the orbital energy split. This is a non-rigorous attempt to separate out the effect of double-bond interaction in cyclononatriene ("pure" π-effect without the σ contamination). Actually, in the cyclononatriene structure, the double bonds are twisted and not coplanar; thus, a calculation on a twisted expanded benzene molecule was performed, simply by shifting the hydrogens out of plane in the structure of the first calculation into their cyclononatriene positions. The interaction between the double-bonds is thus further reduced (less efficient orbital overlap), and the orbital energy split is reduced. In cyclononatriene itself, the actual orbital energy difference (1.26 eV) between the π-type levels is greater than the value calculated for the double-bond interaction alone, as above. A further

contribution to the overall effect can be considered to come from the interaction of the conjugated double-bond system with the briding methylene groups. This is illustrated in Figure 6 . Only the central levels of the diagram are rigorously defined, and these are obtained in the cyclononatriene calculation; the forms of the relevant MO's derived can be seen to correspond to the synthesis illustrated diagrammatically. The upper set of orbitals can be represented as arising from the linear combinations, (C=C)-(CH_2), and the lower as (C=C) + (CH_2). Only the orbital energies of the actual levels can be calculated; the unperturbed constituents, according to the above model, are not rigorously defined, but the above calculation on the benzene structure can give an indication for the double-bond levels. On this basis, the calculated split of the orbital energies of the π-type levels of cyclononatriene can be considered to arise from two types of interaction. On the basis of the non-empirical calculations here, there is a significant contribution by the methylene groups' orbitals to the two highest occupied MO's of cyclononatriene. These π-type orbitals cannot be described as arising from almost pure homoconjugation among the double bonds; there is a conjugative effect right round the ring with the CH_2 groups intimately involved. The cyclic system of "crown" cyclononatriene may have some extra resonance stability (likely to be a small amount). The concept of homoaromaticity has been viewed as applying to a situation where a π-system is interrupted at one or more points by a saturated centre, but where the geometry still permits significant overlap of the p orbitals across the "insulating" gap, so that the physical continuity of the delocalised π-system is broken without affecting the system, and there is no participation in the prevention of the attainment of an aromatic system by the saturated centres. However, in cyclononatriene, such a model does not really describe the situation. The concept is only meaningful if the interaction of the

π-orbitals with the orbitals of the σ-bonds is negligible or the same for all π-orbitals. If the σ-π interaction is of the same magnitude as the homoconjugative π-π interaction and if, for reasons of symmetry, this interaction favours one of the linear combinations over the other, then homoconjugation as well as homoaromaticity are no longer important factors for the description of the electronic structure of such molecules.

A further example for the application of the above simple model, separating the overall σ-π interaction into components, is provided by the molecule of Triquinacene (Figure 4). This species is quite closely related to all-cis cyclononatriene; it is described also as an uncharged six-electron trishomoaromatic analogue of benzene, possessing three properly disposed double bonds for appreciable interaction to occur. The photoelectron spectrum of triquinacene has been obtained, although only the low I.P. region has been published[12]. Attention is focussed on the splitting of the first two occupied MO levels, as in cyclononatriene. The vertical I.P.'s measured are

e(π)	9.0 eV	(Jahn-Teller split leads to imprecise value: 9.0-9.3 eV)
a_1(π)	9.5 eV	
σ-onset	10.4 eV .	

In both cyclononatriene and triquinacene, the distances between the sp^2-hybridised carbon centres are similar, and the mutual orientations of inner p lobes also, so that the extent of interaction of the π-bonds in each species is expected to be much the same; this is the homoaromatic effect. In considering the photoelectron spectrum of triquinacene, it is assumed that the appreciable interaction of the π-bonds in cyclononatriene (0.9 eV split in P.E. spectrum) is due exclusively to the operation of through-space effects, which are dependent upon the extent of relevant orbital overlap and their energy separation ; the smaller observed split in triquinacene (0.5 eV) is attributed to the resultant of much the same through-space interaction of the π-levels along with an appreciable through-bond interaction (not present in cyclononatriene). This latter effect is more easily interpreted as an interaction between MO levels, in the same way as through-space interaction (π-levels), with the relevant MO's being of π- and σ-types. The situation in cyclononatriene is illustrated in Figure 6 . In triquinacene, the corresponding diagram differs in the relevant σ-orbitals. In both species, the a_1-e split of

the component π-orbitals is centred around the unperturbed ethylene value (about 10 eV on the observed I.P. scale). As far as the σ-orbitals are concerned, in cyclononatriene the a_1-e split of the combinations of C-H orbitals is likely to be centred around 14 eV (approximate value from observed first I.P. of methane). Although the size of this unperturbed split is uncertain, the respective energy separations lead to a larger $e(\pi)$-$e(\sigma)$ interaction than $a_1(\pi)$-$a_1(\sigma)$, so that the $a_1(\pi)$-$e(\pi)$ resultant split is larger than the "pure" π-level interaction. The resultant $a_1(\sigma)$-$e(\sigma)$ split, i.e. between MO's which are symmetric combinations of the basic π- and σ-levels, is calculated to be 1.9 eV, which is smaller than the unperturbed value. In triquinacene, the resulting $a_1(\pi)$-$e(\pi)$ split is reduced, mainly because the $a_1(\pi)$ level is destabilised relative to the corresponding level in cyclononatriene. There is no geometrical information available on triquinacene so that a satisfactory non-empirical calculation on this species is likely to be difficult to obtain. However, on the basis of semi-empirical calculations[12], there is an $a_1(\sigma)$ level (MO is composed mainly of contributions from the three central C-C bonds and the central C-H bond) which is closer in energy to the π-type levels than the $a_1(\sigma)$ level of cyclononatriene above. Thus, using the above model, the interaction between the unperturbed a_1 levels is larger than in cyclononatriene (smaller energy separation). Thus, in cyclononatriene and triquinacene, the observed splitting of the π-type levels is not explained by a model which includes the effects of π-type levels is not explained by a model which includes the effects of π-interactions alone (homoaromatic effect).

The importance of σ-π interaction is also shown by considering Norbornadiene (Figure 4), which can be described as an anti-homoaromatic system. The observed split between the orbitals $b_2(\pi)$ (highest occupied) and $a_1(\pi)$ (second highest) is 0.85 eV. In analogy with the model above,

assuming the interaction between the pair of double bonds to be a pure through-space interaction, the split is given by the expression

$$\varepsilon(a_1(\pi)) - \varepsilon(b_2(\pi)) = 2m\beta$$

leading to $m \approx 0.2$. This value is only about half of that obtained for cyclononatriene. Calculations on norbornadiene have been reported in a related work[13]. In addition, in this work, a calculation on the expanded cyclobutadiene of Figure 4 was performed in order to indicate the magnitude of the pure through-space interaction between the pair of double-bonds of norbornadiene. The orbital diagram (Figure 6) is a schematic representation of the interaction of the σ- and π-orbitals, where it is shown that the mixing of the $a_1(\sigma)$ orbital with the lower lying $a_1(\sigma)$ orbital is greater than that of the b_2 ones, so that the resultant π-level split is reduced from the value expected from the operation of π-interaction alone. The use of the observed I.P.'s gives a misleading value for the parameter m. In the above species, it is not a useful approximation to separate the system overall into σ- and π -components.

Returning to cyclononatriene, discussion of the remaining valence shell MO's and their orbital energies is deferred until trioxacyclononatriene is considered below, when a comparison of the two isoelectronic species' MO's is performed.

The all-cis crown conformation is considered to be the lowest energy form (trishomobenzene in some theoretical considerations). On the experimental side, there is some information on this form. Low temperature NMR studies yield a spectrum showing two separate methylenic proton resonances, whereas at higher temperature there is only one averaged line from these protons . The interpretation is that there is a dynamic inversion process occurring, and this is rapid (on NMR time scale) at high temperatures; Crown-crown interconversion, with an

observed free energy of activation of 41 kJ mol^{-1}, is the conformational process involved. In a study of molecular conformation of cycloalkatrienes by molecular mechanics strain energy calculations, twenty-one conformations of cyclononatriene were considered ; the all-cis crown conformation was found to be the most stable.[14] In considering the mechanism of crown-crown interconversion, it has been suggested that the transition state could be either planar (ammonia-like inversion) or saddle (boat-boat), which seems the most probable from molecular models. The above calculations indicate that the lowest energy planar form (C_{2V} symmetry) is too strained for consideration as a possible intermediate in the interconversion; a barrier which is an order of magnitude too high is obtained. It is concluded that the interconversion pathway from the crown form to its specular image implicates crossing through the saddle form which remains the most probable transition state. A non-empirical calculation was carried out on the saddle conformation of cyclononatriene, which can be formally easily derived from the crown by reflecting one of the methylene fragments in the base plane (Figure 4); the calculated total energy quantities are presented in Table 5 . The saddle form is less stable than the crown by over 200 kJ mol^{-1}. The rigid derivation of the saddle conformation from the crown is over-restrictive; there is an abnormally close approach of one of the methylenic hydrogens to the opposed olefinic group (H situated at 1.6 Å from each C atom), and this "strain" is relieved in the optimised molecular mechanics calculation by an opening out of the angles at the two symmetrically-related methylene carbons. This bond angle deformation (109.5° to 115°) reduces the unfavourable steric repulsion, and a reasonable estimate of the barrier to inversion can be obtained. Thus, it would probably be possible to derive a reasonable value for the inversion barrier non-empirically by performing optimisation of the saddle conformation (C_S symmetry), and the crown.

The saddle conformation of cyclononatriene considered is easily, and economically, derived from the crown. Although the resulting form is not very satisfactory if a near-equilibrium structure is required, it is of interest to consider the forms of the saddle MO's in the light of results obtained with the crown. In particular, as above, the π-type orbitals are considered. In a structure of lower symmetry, the e(π) degenerate pair of MO's of the crown conformation is split into an a' and a" combination in the saddle form. The two particular conformations used here are closely related; in the saddle conformation the basic expanded benzene fragment is derived from that in the crown by rotating one of the double-bond systems further out of the plane of the base of the crown. From the forms of the three π-MO's of benzene, it is expected that the overall interaction of the resulting π-type levels is reduced. The a_1 orbital is destabilised (a' in saddle) as favourable AO overlap is reduced; the e orbital component which correlates with the a' is stabilised as antibonding character is reduced while bonding character is much the same; the energy of the e orbital component which correlates with the a" remains approximately constant, being nodal at the unique double bond. This behaviour is confirmed by non-empirical calculation on this further expanded, distorted benzene structure; the situation is depictable schematically as in Figure 6 . In the actual calculation on saddle cyclononatriene, the calculated splitting of the π-type levels is somewhat greater than in the crown (orbital energies calculated for the saddle are given in Table 5). The three levels are centred around a value of about 9.2 eV; a "pure" π-type interaction, as in the calculations on completely planar species, leads to orbital energy level splitting about a value of 10.5 eV, which corresponds to the orbital energy of an isolated ethylenic double bond (calculated). This effect, as in the other species considered above, is indicative of further orbital interaction, which raises the energy of

the less stable component of the interacting pair of orbitals. In addition, the overall situation in saddle cyclononatriene

is dominated by the σ-π type of interaction discussed above, with the consequence that the sequence of the overall σ-π mixed orbitals is different from that of the basic π-type orbitals. The variation of the stabilities of the three component σ methylene orbital combinations on going from crown to saddle is likely to follow the same trend as the π-type orbitals, but the actual size of the effect is even less certain. The resulting effect is: the a' component orbitals are destabilised relative to the crown a_1's, and after mixing the final orbital is slightly destabilised compared to that in the crown, indicating a similar interaction; the a" orbital positions remain approximately constant and the combined level also; the a' orbital interaction (corresponding to one e component of crown) is significantly increased - both components are stabilised relative to crown, but the combined level is somewhat destabilised. The actual orbital energies of the cyclononatriene calculations are indicated in Table 5 . The lower-lying valence orbitals, complementary to the π-type considered above as they are formed from symmetric combinations of double-bond and methylene orbitals, show the effect of interaction (σ-π mixing) also, in an inverse sort of fashion; in particular, for example, the a" symmetric combination is significantly stabilised in the saddle compared to its crown counterpart.

Thus, in both conformations of cyclononatriene considered, the three highest occupied MO's, although largely of π-type character, are not simply formed as a result of π-type interaction alone (homoaromatic effect), and σ-π interaction is very significant (inherent to the structures); in fact, the latter effect dominates.

(b) Cis-cis-cis-1,4,7-trioxa-2,4,6-cyclononatriene.

This nine-membered heterocyclic species can be viewed formally as being derived from cyclononatriene of (a) above by substituting an oxygen atom for each methylene group (Figure 7). This molecule has been synthesised although studied experimentally to a lesser extent than cyclononatriene. There is no experimental geometric data available, although it is thought to exist in a "crown" conformation; in such a case, the species can be regarded as a heterocyclic analogue of a "homoaromatic" type of molecule. In addition, there is the alternative view of trioxa-cyclononatriene being of intrinsic interest as it can reasonably be considered as existing in a completely planar conformation, and so being in the class of heteroaromatic type of molecules. Actually, such a planar system is normally considered as a 12 π-electron system, and so is not expected to exhibit aromatic stabilisation. Thus, the planar conformation is considered first.

Planar Conformation.

Trioxa-cyclononatriene is rather a large molecule from the non-empirical calculation point of view. However, assuming the molecular structure to be highly symmetrical, it is not an unreasonable task to perform molecular geometry optimisation; this is also particularly desirable in this case as there is so little geometric information on species of this type, and none, in fact, for this particular species. Thus, to begin the optimisation procedure, an initial calculation was performed, using the standard scaled minimal basis, on the planar conformation comprising a regular nonagon for the heavy atom framework with C-H bonds bisecting the external angles at the carbon atoms; r(CC) and r(CO) were both chosen to be 1.36 Å, r(CH) was taken as 1.08 Å (COC and CCO angles = 140°, OCH and CCO = 110°) - set of parameters which seemed a reasonable compromise, after considering geometrical data from related systems. Geometry optimisation was effected

FIGURE 7 : ALL-CIS-TRIOXACYCLONONATRIENE "CROWN".

PLANAR CONFORMATION.

FORMS OF π-MO'S.

by using the gradient (Fletcher-Powell-Davidon) technique of Chapter 4. The symmetry of the nuclear configuration was assumed to be D_{3h}, so that the independent geometrical parameters to be optimised were r(CC), r(CO), r(CH), $\hat{COC}$, $\hat{OCH}$ (or any two of the four valence angles). Thus, at each cycle of the energy minimisation process, five gradient components were evaluated, by finite differences; these were the derivatives of the total energy with respect to the oxygen y-coordinate, carbon x- and y-coordinates, hydrogen x- and y-coordinates respectively (coordinate system of Figure 7). The details of the otpimisation are presented in Table 6 .

Three cycles of the iterative process were performed. The gradients evaluated at the initial estimated structure were reasonably small, and these decreased quite smoothly and rapidly. However, the "α factor" of the process must be considered in judging convergence of the process. In this case, the total energy did not decrease smoothly, mainly because of the quite large value of α obtained during the second cycle (approx. 5). Thus, after two cycles, the situation was uncertain; the gradient components were decreasing overall ($\underline{S}$ vector), but the α factor effect led a slightly diverging total energy sequence. However, after the third cycle, the position is more satisfactory; it seems as though convergence, within chemical accuracy, has been very nearly attained at this point. The variation of geometrical parameters with each cycle is also shown in Table 6 . The typical types of shifts are observed. The tendency of the basis set to yield elongated formal single bonds between heavy centres is a prime factor in determining the initial movements. The C-O bond length is increased during cycle 1, and the C=C bond length is decreased, so that an alternating nonagon is being formed, although it is somewhat distorted at this point; after only one more cycle, the optimised structure is closely approached. The final structure is quite similar to the initial estimate, with the main difference being the

introduction of alternating character into the heavy atom framework. It is likely that this effect is exaggerated by the basis set used. Nevertheless, this structure does not correspond to a completely "classical" one, and there does seem to be a slight lengthening of the C=C and shortening of the C-O; this effect is usually taken to be an indication of aromatic behaviour. However, on an energy basis, the resonance energy of the optimised structure is substantially negative, and is unlikely to change significantly on complete optimisation. The system does possess 12 π-electrons formally (4n), but, in addition, the bond angles at the carbon and oxygen atoms are rather unfavourable, especially as regards the latter. The stabilisation energy computed here is an overall value, inclusive of all qualitative effects such as aromatic stabilisation and strain. An indication of the significance of the σ-strain effect can be obtained by considering the results of calculations on planar 10- and 12-annulenes (Chapter 5). Although the particular molecular structures chosen were constructed without basis on experimental data, it is unlikely that geometry otpimisation would significantly alter the trend in calculated resonance energies, namely that the value for 10-annulene is about zero and that for 12-annulene is substantially negative, much more so than the value of trioxa-cyclononatriene. It is difficult to reconcile the calculated values of these three π-systems with the simple qualitative Huckel-type approach often used. In addition to the effect of the π-electrons alone, it is considered that σ-system strain effects should be significant in these systems. The substantial difference in stabilisation energy calculated for the two 12 π-electron (neutral closed-shell) systems is assumed to be consistent with the greater strain present in the 12-annulene structure as a result of bond angle deformation and inter-hydrogen steric effects. Bearing in mind the stabilisation energy of benzene (calculated at about 200 kJ mol^{-1})

as a reference figure, in 10-annulene the unfavourable strain counterbalancing the aromatic stabilisation is then substantial, yielding a net stabilisation energy of about zero; the strain in planar trioxacyclononatriene could be expected to be similar to that in 10-annulene so that the overall destabilisation of this 12 π-electron system would be expected to be somewhat greater than the calculated value (130 kJ mol^{-1}), unless aromatic destabilisation is insignificant. Thus, there is no quantitative consistency in an attempt to separate aromatic and strain effects, and this is in keeping with the all-electron approach.

Inspection of the MO's of planar trioxacyclononatriene is useful, both from the view of obtaining background information in examination of non-planar structures and from considerations related to those above on this planar 12 π-electron system. Thus the forms of the π-MO's are illustrated in Figure 7 , together with the orbital energy level pattern which can be compared with the familiar one of the parent 12-annulene (regular duodecagon, Chapter 5). In either alternating or regular duodecagonal 12-annulene, the 6 π-MO's (4 distinct energy levels) do not lie above the σ-core in energy; there are several σ-MO's lying between the most stable π-MO (completely symmetrical combination of carbon $2p_z$ AO's) and the others. There is a progressive effect through the simple polygonal structures of the annulenes in that the completely symmetrical π-MO moves "further in" as the size of the polygon increases, indicating a general stabilising effect. In a similar way, in trioxacyclononatriene there are four σ-MO's situated between the three highest occupied MO's (π) and the next pair of π-MO's, with a further two σ-MO's before the most stable π-MO is found. The trioxacyclonatriene (trioxonin) structure is obviously inherently different from the duodecagonal 12-annulene structure, topologically, even before its heterocyclic character is considered. In simple π-electron theory terms, the forms of the π-MO's and their

relative energies are deducible from the symmetry of the nuclear framework; in all-electron calculations, as reported in Chapter 5, the orbital forms are still symmetry-determined and the calculated orbital energies for the annulene polygonal structures follow broadly the pattern adopted by those of the simpler theory. In actual fact, only the occupied π-MO's have well-defined orbital energies in non-empirical calculations so that the exact positions of the π-MO's on the energy diagram are not all well defined. As an example, in a calculation on the neutral singlet, closed-shell regular 12-annulene species, where only one of a pair of degenerate π-MO's is occupied, the degeneracy is not actually shown by the computed orbital energies; only if both orbitals are occupied equally, as in the lowest triplet form, does the degeneracy become obvious in the calculation. Although only the neutral closed-shell singlet calculation on trioxacyclononatriene has been performed, it is obvious from the orbital energies that the highest occupied and lowest unoccupied π-MO's are not degenerate; despite the fact that virtual orbital energies do not have the same physical significance as the occupied ones, there is usually some indication of degeneracy, or near-degeneracy, effects from their calculated values. Thus, the occupied orbital energy pattern of trioxacyclononatriene deviates significantly from the Hückel-type pattern of a 12-annulene, or of a nonagonal structure (which cannot correspond to a neutral 9-annulene species). In Hückel terms, the antiaromaticity of a 12 π-electron system (strictly, duodecagonal) is essentially the result of the occupancy of a "non-bonding" MO (one of a degenerate pair). Qualitatively, the degenerate "non-bonding" pair of MO's is not present in the trioxacyclononatriene structure, (even in a regular bond-length structure, as the initial one here), and the π-MO energies tend to be of significantly lower energy than those of the parent annulene species (even aromatic ones). On this basis, this 12 π-electron system would not be classed as being destabilised. However, the overall

effect, including the σ-structure, is a structure with a significantly negative stabilisation energy, although it is not so unfavourable as duodecagonal 12-annulene.

Examination of the forms of the π-MO's emphasises the difference of the π-system of trioxacyclononatriene from that of a conjugated hydrocarbon. In a species such as 12-annulene, the π-energy levels can be viewed as the result of interacting "pure" ethylene π-levels, illustrated by the familiar "splitting" diagram (Chapter 5); the actual MO coefficients and corresponding forms show the nature of the symmetric and anti-symmetric combinations of π-orbitals involved in the occupied π-MO's. In trioxacyclononatriene, oxygen atoms replace double-bond units, and each contributes two π-electrons. Analogously to cyclononatriene considered above, the final π-MO structure can be regarded as synthesised from interacting π-orbitals (expanded benzene structure) on the one hand, and interacting oxygen atoms on the other. The situation is analogous to Figure 6. The benzene structure of the optimum planar form of trioxacyclononatriene is slightly more expanded than that of cyclononatriene; however, it is strictly planar. The size of the $a_1(\pi)$ - $e(\pi)$ split resulting is thus expected to be about 1.4 eV, centred around the ethylenic value. The interacting oxygens replace the methylene groups of cyclononatriene, and they are also strictly coplanar with the double-bond system. The overall π-type interaction is expected to be greater in this completely planar system than in cyclononatriene; an additional factor is the higher energy of the a, e pair of levels arising from combinations of oxygen $2p_z$ AO's (calculation on a system consisting of 3 oxygen atoms disposed as in trioxacyclononatriene yields a value of 15.6 eV for the energy of each of the two π-type levels - effectively no "splitting"). The calculated splitting of a (π) and e(π) in trioxacyclononatriene is 2.7 eV, with an inversion in the ordering compared to the basic benzene structure.

This is consistent with a much greater interaction between a levels than between e levels, and this is expected, qualitatively, from the relative energies of the levels.

The Mulliken population analysis of the MO's of the near-optimal planar trioxacyclononatriene is summarised in Table 6 . As far as the π-system is concerned, each oxygen atom donates 0.13 of an electron to its two neighbouring carbon atoms; each carbon has approximately the same π-charge as in furan, where there is only one oxygen donating to five carbons. Each oxygen is a σ-acceptor so that the net charge on each is -0.5, almost identical to that of the oxygen atom in furan. The net charges on the carbons and hydrogens are also very similar to those of the α-carbons and hydrogens of furan. The π-system of trioxacyclononatriene shows only a small perturbation of the classical description of two π-electrons per oxygen and one per carbon, and this corresponds to the relatively small interaction between the double-bond system and oxygen π-system, mainly because of the size of the energy separation of the relevant MO levels, leading to little tendency for uniform π-electron delocalisation around the ring.

In the σ-system, the MO's of trioxacyclononatriene can also be viewed as being formed through the interaction, or mixing, of the basic components, which are a double-bond system (expanded benzene) and a triangle of oxygen atoms. As a result of the relevant energies of the interacting orbitals, and the favourability of overlap, there are some very marked effects. One of the highest-lying MO's is the b_2 one which is the antisymmetric combination of a mainly C-H bonding orbital and an orbital composed of oxygen 2p components. The former is the type present in many of the annulene species (alternating character around the ring) and lies at about 19 eV in a benzene-type system; the latter is a symmetric combination of oxygen 2p AO's which are tangential to the ring, and is

expected to lie at about 15.5 eV in an oxygen atom system as found in trioxacyclononatriene. Based on the results of the calculations on a system of three such oxygen atoms (expanded triangular ozone species), there is really negligible interaction between the centres so that there is insignificant splitting of a, e MO levels, which all lie at about 15.5 eV; there are three sets, namely combinations of $2p_z$ AO's (π-type, perpendicular to plane, and considered above), and combinations of in-plane radial and tangential 2p AO's. There is a large interaction between the in-plane tangential component (a_1 symmetry) and the C-H bonding level, leading to the high-lying MO mentioned above at 14.4 eV and the complementary symmetric combination of the two types lying at 24.3 eV. In contrast, the e-pair of orbitals formed from the tangential in-plane oxygen 2p AO's (and some 2s) interact to a small extent with an e-pair of C-C bonding orbitals, giving two pairs of e-levels at 21.5 eV and 16.6 eV. The remaining oxygen 2p AO combination is that of the radial orbitals; these correspond to the lone-pair electron orbitals, whereas the above σ-MO's are predominantly C-C and C-O bonding. The formal a_1 lone-pair oxygen orbital interacts more strongly with a C-C bonding orbital than does the e-pair so that the resultant formal trioxacyclononatriene lone-pair MO's comprise an e-type at 14.8 eV and an a_1-type at 13 eV; these high-lying σ-MO's are particularly rich in oxygen 2s and 2p (radial) character, and they lie somewhat above the level of 15.5 eV which is an estimate for the unperturbed oxygen system, so that some idea of the extent of sigma "lone" pair interaction with the rest of the system can be gained. The complementary levels (symmetric combination of lone-pair and C-C bonding orbitals) lie at 19.9 (a_1) and 18.6 (e) eV. In the interaction situation with the component levels here, the oxygen system levels are the higher energy ones, unlike the situation in all other cases; in the oxygen system sets

of a_1 and e levels are split insignificantly, whereas in the double-bond system there is always some splitting (a_1 lower in energy in e), so that in this situation the a_1 interaction is greater despite a slightly more unfavourable energy difference. The remaining six valence-shell MO's are formed from combinations of carbon and oxygen 2s AO's. The oxygen and carbon a_1-e splits are each increased by the mutual interaction, with the a_1 interaction being more significant so that the predominantly carbon 2s MO's have the e-pair at lower binding energy than the a_1. The 2s levels are each shifted by about 1 eV compared to unperturbed ones. The remaining σ-orbitals are essentially unperturbed 1s AO's.

During the geometry optimisation process, the variation of orbital energies can be understood by considering the forms of the MO's in the light of the variation of bond parameters at each iterative cycle. In the case of planar trioxacyclononatriene, the overall trend is a lengthening of the C-O bond; the other geometrical parameters undergo little change, although the C-C bond length does fluctuate significantly, shortening at first and then tending to lengthen. From the four "points" at which calculations are performed, some idea of the convergence of orbital energies, as well as total energy, can be gained. In this particular case, at the optimal structure most of the orbital energies have settled down, within reasonable accuracy, although one or two are varying by amounts greater than the total energy variation. In terms of the model of interacting systems, the effect of geometry optimisation is to increase the O-O distance, and to slightly expand the benzene-like double-bond system, i.e. increase r(C-C) with r(C=C) almost unaltered although it is significantly shortened at cycle 2. Most of the orbitals are actually destabilised as the total energy improves. One exception is the highest occupied orbital; examining the form of this π-MO shows that

the antibonding character is reduced as r(C-O) increases, so that there is a continuous increase in stability as the structure is optimised. The next highest occupied MO's (e(π)) are also stabilised, although this is a fluctuating effect; the structure at cycle 2 particularly enhances the bonding character, and decreases the antibonding character, of this MO so that the energy is particularly low at that point. The next highest MO is also slightly stabilised. This is the a_1 lone-pair MO; although the variation of O-O distance is likely to have a negligible effect on the "pure" lone-pair character, the antibonding character introduced by mixing with C-C bonding orbital is reduced as optimisation proceeds. The e-type lone-pair level is effectively invariant to variation of geometry, as the mixing is very small. Most of the remaining orbitals are destabilised as the C-O bonding character is reduced on geometry optimisation. The b_2 orbitals are both markedly destabilised as both have a significant amount of C-O bonding character as well as the C-H. The MO's which are composed mainly of 2s AO's illustrate once more the interaction effects of the two component systems. The a_1 and e combinations of O 2s AO's are almost degenerate, with energy about 37 eV; the corresponding C orbital combinations show an energy splitting of about 0.5 eV centred around 27 eV, with the a_1 level the lower-lying one. There is a greater interaction between the a_1 O and C orbital combinations so that the higher-lying antisymmetric combination, predominantly of C AO character, shows inversion with the a_1 MO at higher energy; a splitting of about 0.5 eV is introduced into the symmetric combination which is predominantly of O AO character. As the molecular geometry varies, the unperturbed oxygen levels are unaffected; the carbon levels' energies vary somewhat, and at the final structure these levels are likely to be destabilised compared to the initial structure. The overall interaction between the two sets of levels decreases as the geometry is varied so that the

resulting effect is a destabilisation of the predominantly oxygen 2s MO; again, the lengthening of the C-O bonds tends to decrease the bonding character of orbitals such as these.

In considering hetero-aromatic character in trioxacyclononatriene, it is useful to refer to the simpler oxygen heterocycle, furan, formally a 6 π-electron system. In a study of the electronic charge distribution and moments of five- and six-membered heterocycles[16], it has been concluded that some indication of the aromaticity of these systems is possible from the extent of delocalisation of the valency shell of 6 π-electrons as indicated from the average position of the π-electrons. In the five-membered heterocycles, which include furan, there is a variation in the degree to which these six electrons separate into a group of four and a pair of electrons, and this is consistent with the variation in aromatic character as estimated from other considerations. Thus, in furan, the π-sextet has a significant amount of separate diene and lone-pair characters. In trioxacyclononatriene, there is a significant departure, in a similar way, from a symmetrically localised 12 π-electron system, so that the system exhibits classical behaviour with two groups of π-electrons, of triene and oxygen lone-pair characters. This is consistent with the π-electron energy levels varying significantly from an annulene 12 π-electron arrangement; the arrangement of π-levels in trioxacyclononatriene is similar to that in furan, with the addition of a double-bond level and two lone-pair levels essentially unperturbed, whereas there is significant variation between benzene and 12-annulene arrangements. The interaction between the oxygen lone-pair π-orbital and the hydrocarbon π-system in furan is similar in magnitude to that between the corresponding a_1-orbitals of trioxacyclononatriene, but there is very little interaction between the e-levels. The overall conclusion is that the 12 π-electron system of trioxacyclononatriene does not closely resemble a Hückel-type

π-system, and that, qualitatively, there is not much indication of a destabilised π-system through electron delocalisation (antiaromaticity). Unlike furan, however, trioxacyclononatriene is substantially destabilised, and this can be attributed to unfavourable σ-system strain. From examination of the energies of corresponding σ-MO's, it is obvious that there are greater interactions in the trioxacyclononatriene system than in furan. In particular, the a_1 lone-pair orbital of the three-oxygen system interacts with the hydrocarbon system MO more strongly than the oxygen lone-pair of furan, so that the resulting σ-MO of predominantly lone-pair character lies at 14.1 eV in furan and at 12.7 eV in trioxacyclononatriene; the e lone-pair orbitals also interact to a smaller extent with the hydrocarbon system, in the latter. In furan, there is a very small interaction between the "tangential" oxygen p-orbital and the mainly CH-bonding orbital (alternating round the ring), whereas this type of interaction is particularly marked in trioxacyclononatriene. The orbital population analyses, as mentioned above, show the overall effects of the differing extents of σ- and π-type interactions in the two systems.

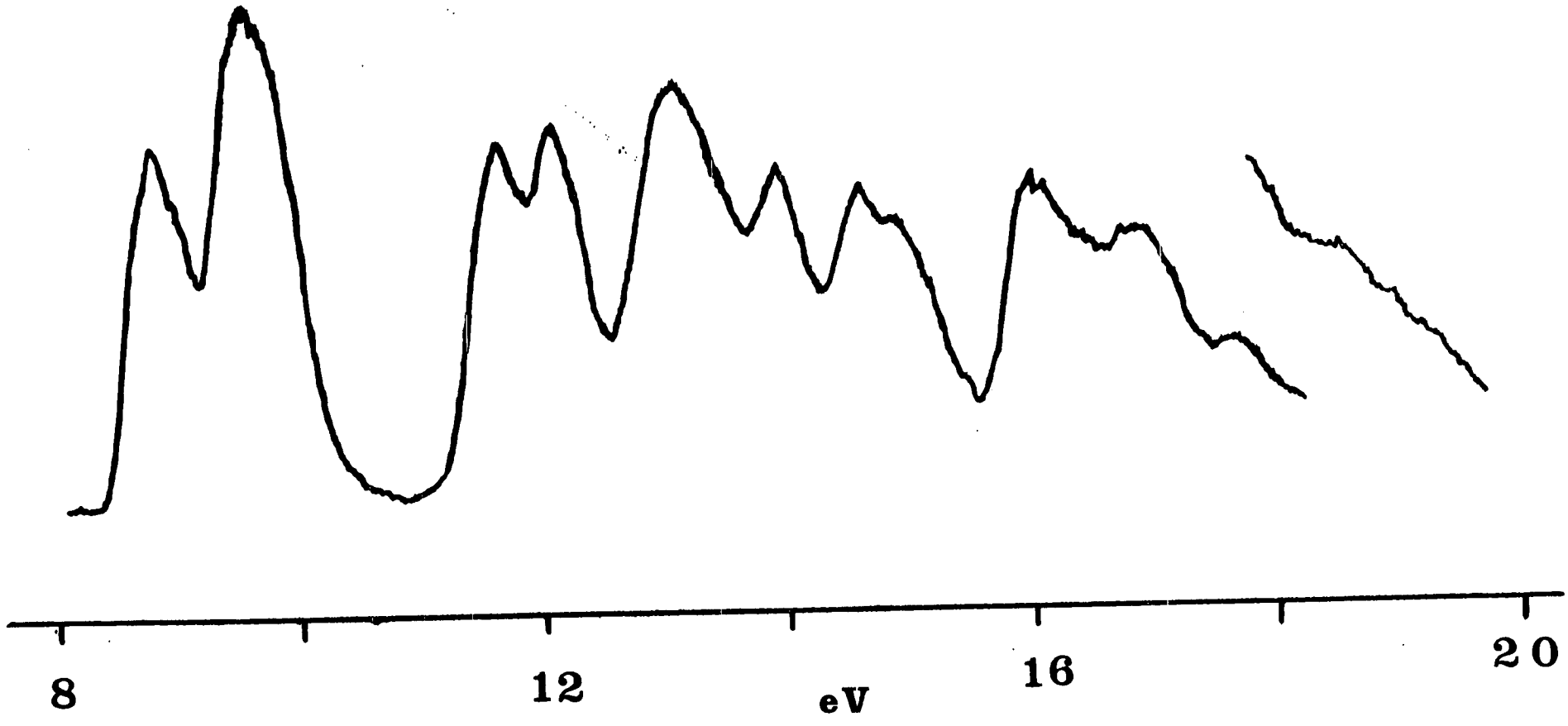
Non-planar Conformations

In contrast to cyclononatriene itself, there is little experimental information on the heterocyclic analogue, 1,4,7-trioxacyclononatriene. The most stable conformation of cyclononatriene has been found to be the "crown", and this is effectively the only form present under normal circumstances. Thus, for direct analogy, the crown form of trioxacyclononatriene should be considered. Recently, a method of preparation of this compound has been devised[15]. The thermal isomerisation of the available benzene cis-trioxide (3,6,9-trioxatetracyclo[6.1.0.$o^{2,4}$.$o^{5,7}$]-nonane), by pyrolysis in the gas phase in the temperature range 475-537 K, leads to the single product all-cis-1,4,7-trioxacyclononatriene, identifiable by its n.m.r. spectrum. Thus, having performed the above process, the photoelectron spectrum of the product was measured; this is reproduced in Figure 8 , and it is in agreement with the spectrum of trioxacyclononatriene obtained in an independent study[17]. It is really certain that the spectrum is of the all-cis isomer, which, bearing cyclononatriene in mind, is thought to adopt a crown conformation. There is no geometrical structure data available for trioxacyclononatriene, and there is even very little information on simpler related ether-type structures; divinyl ether is the most relevant species, but there is no relevant information. Thus, more so than usual, geometry optimisation is desirable.

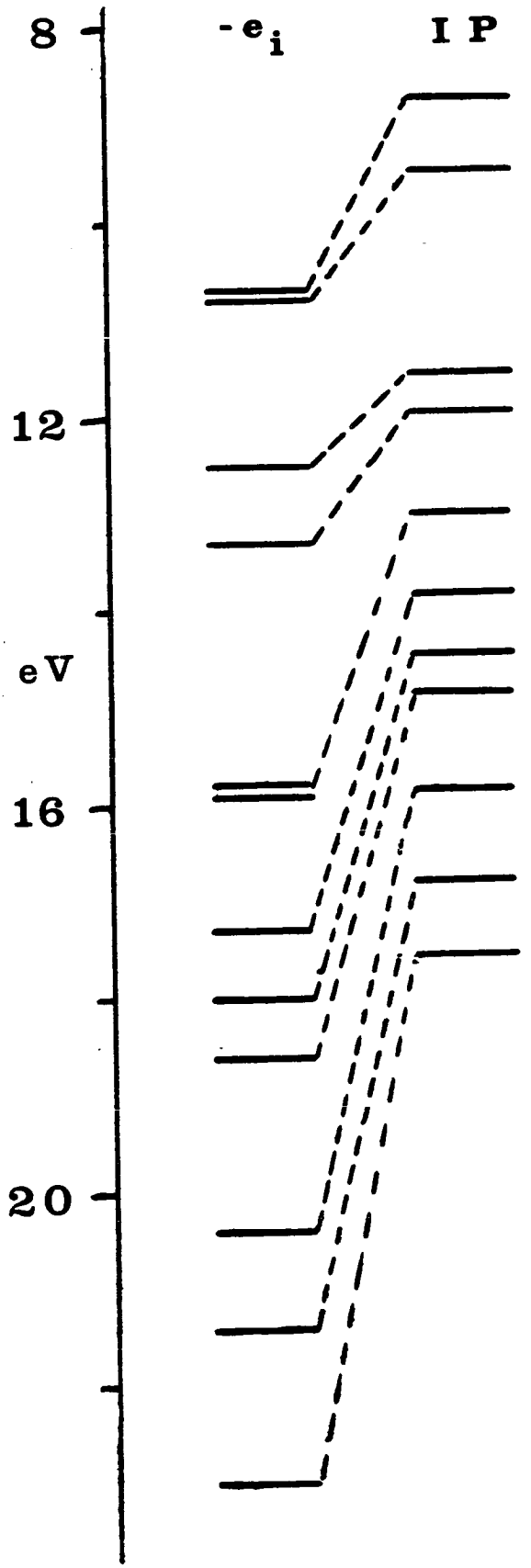
An initial structure for trioxacyclononatriene was constructed using bond parameters from ether-type molecules. The crown conformation for the all-cis isomer actually follows naturally from such a choice. The geometrical parameters and results of the calculation on this structure are reported in Table 7 . It can be concluded that the actual structure used is not close to the calculated equilibrium one. The computed total energy is substantially less negative than that of any of the completely planar structures considered above; in addition, the orbital energies

Fig. 8a

P.E. SPECTRUM OF $C_6H_6O_3$.



F I G . 8 b : $C_6H_6O_3$ ORBITAL ENERGY CORRELATION.



computed do not correlate at all well with the observed I.P.'s from the P.E. spectrum. One obvious feature of the spectrum is that, based on peak intensities, the lowest I.P. probably corresponds to ionisation by removal of an electron from an a-type orbital, whereas the next I.P. is probably that of an e-type orbital (contrast cyclononatriene with I.P.(e) < I.P.(a_1)); this ordering is reversed in the calculation where $|\varepsilon(e)| < |\varepsilon(a_1)|$.

Before performing geometry optimisation by the gradients procedure on this rather large case, it was decided that an improved starting structure was desirable. In the above structure, the parameters which are most doubtful (and critical) are r(CO) and $\hat{COC}$. Recent data on methyl vinyl ether and methyl allenyl ether (electron diffraction) indicate that the angle at the ether oxygen is significantly greater than the value for saturated ethers (about 105°), being nearer 120°. The CC and CO bond lengths vary somewhat, but the CO length tends to be slightly greater (as found in the optimisation of the planar species, allowing for the standard basis set lengthening of r(C-O)). Thus, the first step in the gross improvement of the geometry was to construct a crown with r(CC) = r(CO) = 1.36 Å, and $\hat{COC}$ = 105°. A partial, one-dimensional optimisation was then performed by varying $\hat{COC}$, and the results of the calculations are given in Table 7, where the variations of orbital energies are shown. The difference in geometrical structure between the original choice and the first of the three optimisation ones is in r(CC) and r(CO). Although a substantial improvement in total energy is achieved, the correlation of orbital energies with observed I.P.'s remains very similar, as the general trend is a small shift to higher orbital energies in the lower total energy calculation; in particular, the four highest orbital energies, corresponding to the most well-defined observed I.P.'s, are essentially unchanged. However, as the critical $\hat{COC}$ parameter

is varied through quite a large range, the total energy is lowered and there is more significant movement of orbital energies. Although in the structure with $\hat{COC} = 119^{\circ}$ (predicted minimum by parabolic fitting of the three computed total energies is 122°) the highest orbital energy is that of an a_1 orbital with the next highest an e-type, the overall fit of orbital energies and observed I.P.'s is still rather poor. Furthermore, the total energy is higher than that of the optimum planar form. Thus, otpimisation of the geometry as a whole is desirable.

Starting from the lowest total energy structure above, minimisation of the total energy by the gradient technique was performed. Two cycles were carried out, and the details of the calculations are presented in Table 8 , along with geometrical data. In the planar trioxacyclononatriene case, a third cycle was required to achieve convergence to chemical accuracy (total energy correct to within 10 kJ mol^{-1}, bond lengths to within 0.02 Å). Considering the results of the optimisation of the crown form, it is likely that within one more cycle, or two, the same level of convergence as with the planar form would be achieved. With the lack of a good initial structure, optimisation of the non-planar form has already been rather expensive. The total energy has been lowered significantly, and, assuming the likely further lowering if further iteration was performed, it is indicated that the stabilisation energy of the non-planar crown form of trioxacyclononatriene is nearly zero, implying a largely classical type of species. This conclusion is consistent with the geometrical parameters of the optimum structure so far, and these are unlikely to change dramatically on further optimisation. Although the geometry of divinyl ether is not known precisely, it is likely that the structure of trioxacyclononatriene can be regarded as an aggregate of vinyl ether units, in analogy to the situation in cyclononatriene.

It seems as though the current position is close to the optimum from the viewpoint of total energy or molecular geometry. However, the orbital energies have not wholly settled down, within reasonable limits. The variations of orbital energies on optimisation are included in Table 8 . At the present optimum structure, the fit of orbital energies to the observed I.P.'s is still not very satisfactory, particularly at the low I.P. end (Figure 8); the four highest orbital energy levels have been somewhat mobile throughout the optimisation calculations, and it is reasonable to expect that on performing one or two more optimisation cycles the orbital energies of these levels (and others) could shift to yield a better fit with experimental data, as variations of about 0.5 eV, or less, are all that would be required. During the optimisation procedure, the total energy of the molecular structure is always lowered, but the actual geometry of intermediate points may be rather distorted (seen quite clearly in the planar case), so that orbital energies may fluctuate before settling down at the optimum geometry value. At each cycle, the structure overall is altered to yield a lower total energy; there can be structural changes which are locally very significant (favourable or unfavourable) so that certain M.O.'s are particularly affected. As convergence is approached, structural variations are "smoothed" out, and become smaller.

In Table 8 , the orbital energies of the optimum planar form have been included; the correlation of MO's in the various non-planar forms is given, along with the planar form where the nature of MO's is most easily seen. The forms of the MO's of the "interrupted" conjugated system can be discussed in terms of the model whose basic components are the systems of three oxygen atoms and an expanded benzene-like structure, and the variation of orbital energy with variation in molecular geometry can be considered. In addition to the bond parameters of the various structures, the variation of certain "non-bonded" distances is also of

interest. On proceeding from planar forms to non-planar, the inter-oxygen distance is reduced from about 3.5 Å to 2.9 Å, and this value remains effectively constant in all the non-planar forms. By performing a calculation on the expanded ozone structure with this new "bond length", as done in considering the planar case above, there is an indication that the interaction between the oxygen AO's is becoming significant. In the non-planar form, the benzene-like structure component is less expanded, i.e. the C-C "bond length" is reduced from the planar trioxacyclononatriene value and is, in fact, similar to that in the cyclononatriene structure above as far as the optimum structure is concerned. The non-bonded distances are included in Table 8 , along with the usual geometric data. As in cyclononatriene, the benzene structure is actually distorted out of the base plane, with each local double-bond system being rotated out of this plane by about 40°. Each COC plane and the base plane include a dihedral angle of about 130°, which is similar to the corresponding value in cyclononatriene. Application of the model for describing the MO's is complicated by the effects of σ-π mixing in the non-planar conformation.

Conclusions

The studies were largely concerned with the calculation of equilibrium geometry for a number of non-planar and planar systems. The Fletcher-Powell procedure was used with the minimal basis set *ab initio* calculations developed previously. On a series of small molecules the calculations gave generally satisfactory agreement with the observed bond lengths and angles, although some trends were observed, e.g. C-C in C=C-C=C is systematically long. With these limitations in mind the method was then applied to a number of annulenes and related conjugated molecules; although some of the studies were not carried fully to optimal level, the main trends in the structures as found by the calculations were determined.

References for Chapter 6

1. S. Winstein, Journal of the American Chemical Society, 81, 6524 (1959).
2. S. Winstein, Quarterly Reviews, 23, 141 (1969).
3. E.S. Waight, "Rodd's Chemistry of Carbon Compounds", 2nd Edition, Elsevier, London, 1968, 349 (Ed. E.S. Coffey).
4. M. Trætteberg, J.Amer.Chem.Soc., 86, 4265 (1964) (electron diffraction); S.S. Butcher, J. Chem. Phys., 42, 1833 (1965) (microwave).
5. M.J.S. Dewar, "Molecular Orbital Theory of Organic Chemistry", McGraw-Hill, 1969, p.227.
6. F.A.L. Anet, J.Amer.Chem.Soc., 86, 458 (1964); J.B. Lambert, L.J. Durham, P. Lepoutere, J.D. Roberts, ibid., 87, 3896 (1965).
7. M.H. Palmer, International Review of Science, Organic Chemistry Series 2, Vol. 4, Heterocyclic Compounds, ed. K. Schofield, Butterworths, 1975, p.1.
8. P. Radlick and S. Winstein, Journal of the American Chemical Society, 85, 344 (1963).
9. J.D. Dunity, Perspectives in Structural Chemistry, 2, 1 (1968).
10. J. Aihara, J.Amer.Chem.Soc., 98, 2750 (1976).
11. P. Bischof, R. Gleiter and E. Heilbronner, Helvetica Chimica Acta, 53, 1425 (1970).
12. F. Tomas, J.I. Fernandez-Alonso, An.Quim., 72, 122 (1976).
13. R.H. Findlay, Ph.D. Thesis, Edinburgh, 1973.
14. G. Buemi, F. Zuccarello and D. Grasso, Journal of Molecular Structure, 42, 195 (1977).
15. D.E. Penny, Journal of the Chemical Society, Perkin II, 36 (1976).
16. M.H. Palmer, R.H. Findlay and A.J. Gaskell, Journal of the Chemical Society, Perkin II, 420 (1974).
17. E. Heilbronner, personal communication, to be published.

Table 1

RELATIVE ENERGIES OF "PLANAR" CYCLOHEPTATRIENES (kJ mol^{-1})

- VARIATION WITH CC(Me)C

Angle/Deg.	118.6	123.6	128.6	"REGULAR"
E_T	51	0	64	154

Energies/a.u.		
	E_T(Best Planar)	-268.97219
	V_{EL}	-535.15955
	V_{IE}	-888.48119
	T	271.96653
	V_{NN}	266.18736
	B.E.	1·72291
	R.E. (KJ mol^{-1})	2

Table 2

TOTAL ENERGY QUANTITIES OF "BOAT" CYCLOHEPTATRIENES

	E.D.	MWAVE	NORCARADIENE
$-E_T$	268.95186	268.98504	268·97631
$-V_{EL}$	540.57352	540.48592	545·82418
$-V_{IE}$	899.41758	899.11994	909·58534
T	272.14738	272.08673	272·00273
V_{NN}	271.62167	271.50088	276·84787
B.E.	1·70258	1·73576	1·72703
R.E. (KJ mol^{-1})	-52	35	12
INVERSION (kJ mol^{-1})		34	-

Table 3

I.P.'s AND ORBITAL ENERGIES OF CYCLOHEPTATRIENE

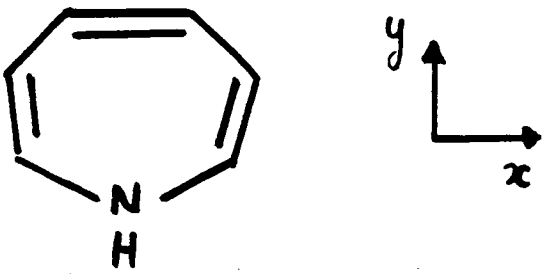
$-\varepsilon_i$/eV	E.D.	MWAVE	PLANAR		NORCARADIENE		I.P.(OBS)/ eV
a'(π)	8.86	8.85	7.86	$b_2(\pi)$	8.28	a"(π)	8.64
a"(π)	9.93	10.00	10.96	$a_2(\pi)$	11.20	a'(π)	9.60
a'(π)	11.02	12.08	12.75	b_1	12.50	a"	11.03
a"	12.93	12.95	12.86	$b_2(\pi)$	13.02	a'	11.89
a'	13.95	13.90	12.86	a_1	13.53		12.40
a'	14.85	14.53	14.55	b_1	14.04		12.67
a"	14.85	14.53	14.95	a_1	14.69		12.83
a'	15.79	16.29	16.42	b_1	16.22		14.38
a"	16.70	16.83	16.45	a_1	17.16		14.81
a'(π) Me	17.29	17.05	16.48	$b_2(\pi)$	17.33		14.91
a"	18.79	18.69	18.93	a_1	17.50		16.22
a'	18.84	19.01	18.98	b_1	19.51		16.55
a'	19.72	19.78	19.61	a_1	21.44		16.77
a"	23.01	22.84	23.41	b_1			18.70
a'	24.57	24.61	24.68	a_1			19.58(?)

Table 4

GEOMETRY OPTIMISATION OF PLANAR 1H-AZEPIN

CENTRE/CORD.		CYCLE 1	2		3	
		$g(= -s)x10^{-2}$	g	s	g	s
N	X	0	0	0	0	0
	Y	7.3284	2.74775	-4.29	-0.6401	-1.15
C1	X	-2.8014	-0.9473	1.57	0.674	-0.923
	Y	-1.5432	-0.6442	0.845	1.9591	-1.21
C2	X	5.4514	-3.7634	0.529	-2.3146	0.478
	Y	-2.7707	5.5152	-2.58	2.41555	-1.31
C3	X	4.04785	1.7159	-2.49	-1.6992	-0.321
	Y	-3.6199	-3.0554	3.21	-0.6689	0.172
HN	X	0	0	0	0	0
	Y	5.15635	-1.0818	-1.09	0.3663	-0.523
H1	X	0.9484	0.5471	-0.677	0.2319	-0.511
	Y	6.408	0.6258	-0.622	0.69175	-0.630
H2	X	1.44385	2.0183	-1.79	1.5616	-0.981
	Y	4.406	0.3457	-0.373	0.2559	-0.225
H3	X	0.2251	0.0714	-0.123	-0.2999	0.120
	Y	1.278	0.79375	-0.949	0.0099	-0.095
	α	0.4173	0.6944		1.667	
	S	19.10	9.132		3.784	
	σ	7.97	6.34		6.31	

Table 4 (contd.)



GEOMETRICAL PARAMETER	CYCLE 0	1	2	3
r(N-C1)	1.420	1.435	1.449	1.444
r(C1-C2)	1.360	1.359	1.345	1.346
r(C2-C3)	1.460	1.460	1.483	1.503
r(C3-C4)	1.348	1.330	1.311	1.325
r(N-HN)	1.05	1.04	1.03	1.04
r(C1-H1)	1.09	1.09	1.10	1.09
r(C2-H2)	1.09	1.07	1.09	1.08
r(C3-H3)	1.09	1.09	1.08	1.09
b-N-1	123.6	122.4	121.3	121.8
N-1-2	219.4	129.0	128.8	128.5
N-2-3	129.4	129.9	129.8	129.8
2-3-4	129.4	129.3	129.6	129.9
N-1-H1	115.3	115.0	115.5	115.4
1-2-H2	115.3	115.2	115.3	115.5
2-3-H3	115.3	115.9	115.7	115.7
E_T ($KJ\,mol^{-1}$)	0	-5.6	-12.6	-17.6

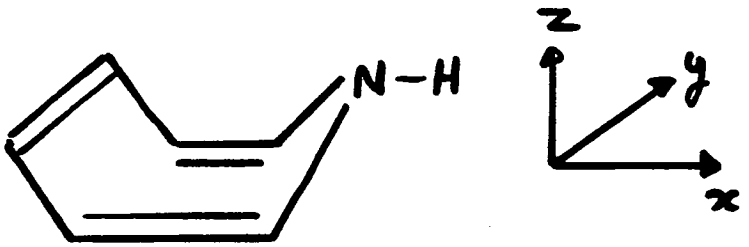
E_T(BEST PLANAR) = -284.91073 a.u. ; R.E. = 74 $KJ\,mol^{-1}$.

Table 4 (contd.)

GEOMETRY OPTIMISATION OF NON-PLANAR 1H-AZEPIN

CENTRE/COORD		CYCLE 1
		$g(= -s) \times 10^{-2}$
N	X	4.074
	Y	0
	Z	0.18835
C1	X	2.5752
	Y	0.8784
	Z	-0.4464
C2	X	-0.7474
	Y	4.9075
	Z	-2.74245
C3	X	-6.1546
	Y	-3.33557
	Z	2.41805
HN	X	-4.79965
	Y	0
	Z	-0.0602
H1	x	-0.20205
	y	-0.57655
	z	0.01375
H2	X	0.0829
	Y	-1.1739
	Z	0.1868
H3	X	0.3557
	Y	-2.7575
	Z	-0.5911
	α	0.7668
	S	15.77
	σ	12.09

Table 4 (contd.)



GEOMETRICAL PARAMETER	CYCLE	0	1	
r(N-C1)		1.420	1.431	Å
r(C1-C2)		1.360	1.352	
r(C2-C3)		1.460	1.471	
r(C3-C4)		1.348	1.335	
r(N-HN)		1.05	1.06	
r(C1-H1)		1.09	1.10	
r(C2-H2)		1.09	1.09	
r(C3-H3)		1.09	1.10	
6-N-1		113.8	114.5	
N-1-2		124.2	124.8	
1-2-3		124.0	123.5	
2-3-4		125.0	126.1	
E_T ($KJ\,mol^{-1}$)		0	-23.2	

E_T(BEST NON-PLANAR) = -284.9069 a.u. ; R.E. = 64 $KJ\,mol^{-1}$.

Table 5

TOTAL ENERGIES OF ALL-CIS-CYCLONONATRIENE (a.u.)

	"CROWN"	"SADDLE"
E_T	-346.85156	-346.76779
V_{EL}	-775.95292	-787.36343
V_{IE}	-1317.35531	-1340.32951
T	350.91363	351.40929
V_{NN}	429.10136	440.59564
B.E.	3·38733	3·30356
R.E. ($KJ\,mol^{-1}$)	-13	-233

ORBITAL ENERGIES AND OBSERVED I.P.'s

I.P. (OBS.)/eV	$-\varepsilon_i$/eV "CROWN"		"SADDLE"
8.77-9.0	9.334	e	8.39 , 9.25
9.80	10.594	a_1	10.1
11.3	12.740	a_2	12.5
11.85	13.288	e	13.4, 13.6
13.0	14.115	a_1	14.4
	14.699	e	14.7, 14.9
13.7	15.372	e	15.4, 15.5
14.7	16.751	a_1	16.0
15.2	17.262	e	17.3, 17.7
16.7	19.134	a_1	18.8
	19.425	e	
	20.322	a_2	
	23.569	a_1	
	25.618	e	
	28.860	e	
	30.464	a_1	

Table 6

GEOMETRY OPTIMISATION OF PLANAR ALL-CIS-

TRIOXACYCLONONATRIENE

CENTRE/COORD.		CYCLE 1	2		3	
		$g(= -s) \times 10^{-2}$	g	s	g	s
O1	X	0	0	0	0	0
	Y	5.07135	3.00425	-3.32	-0.30675	-1.25
C1	X	-5.6586	-0.0409	0.906	0.6106	-1.05
	T	-4.5348	2.42385	-1.35	-1.261	-0.523
H1	X	-0.6488	-0.4235	0.458	-0.95185	0.619
	Y	1.94085	1.44995	-1.52	1.53	-1.44
	α	0.6222	4.9297		1.6015	
	S	20.44	8.00		5.26	
	σ	12.7	39.5		8.43	

GEOMETRICAL PARAMETER		CYCLE O	1	2	3
r(C-O)	Å	1.36	1.388	1.429	1.431
r(C-C)		1.36	1.315	1.353	1.353
r(C-H)		1.08	1.086	1.082	1.084
COC	DEG	140°	138.1	138.2	138.0
CCO		140°	141.0	140.9	141.0
CCH		110°	111.4	112.1	111.8
E_T(KJ mol^{-1})		0	-34.6	-75.1	-79.4

E_T(BEST PLANAR) = -453.8331 a.u.; R.E. = -132 KJ mol^{-1}.

AO POPULATIONS OF PLANAR ALL-CIS TRIOXACYCLONONATRIENE

	O	C	H
1s+2s	3.76	3.03	0.82
2p(σ)	2.89	1.83	
2p(π)	1.87	1.06	
TOTAL	8.52	5.92	

Table 7

"CROWN" TRIOXACYCLONONATRIENE

INITIAL GEOMETRY (r(CO) = 1.325 Å, r(CC) = 1.344 Å, r(CH) = 1.08 Å;
COC = 105°, CCC = 120°)

TOTAL ENERGY VALUES (a.u.)

E_T	-453.73632
V_{EL}	-913.98974
V_{IE}	-1529.75953
T	461.01940
V_{NN}	460.25342
B.E.	1.25507
R.E. ($KJ\,mol^{-1}$)	-387
E_T(SADDLE)	-453.72091

SELECTED ORBITAL ENERGIES (a.u.)

-0.389	e
-0.417	a_1
-0.458	a_1
-0.502	e
-0.611	a_2

OPTIMISATION W.R.T. COC

	105°	112°	119°	COC(MIN = 122.3°)
E_T(kJ mol^{-1})	83.2	27.5	0	
E_i(a.u.)	-0.388 (e)	-0.393(a_1)	-0.396(a_1)	
	-0.413 (a_1)	-0.391(e)	-0.370(e)	
	-0.461	-0.455	-0.447	
	-0.500	-0.498	-0.495	
	-0.606	-0.606	-0.606	
E_T (SADDLE)	114.6	12.2	-29.6	

Table 8

GEOMETRY OPTIMISATION OF NON-PLANAR ALL-CIS

TRIOXACYCLONONATRIENE

CENTRE/COORD		CYCLE 1	2	3
		$g(=-s)\times10^{-2}$	g	s
O1	X	0	0	0
	Y	2.3161	0.14375	-0.544
	Z	-8.68	-3.8324	4.585
C1	X	-7.325	1.98325	-0.150
	Y	-2.6751	4.19855	-2.742
	Z	3.97585	0.83915	-1.392
H1	X	6.1776	-1.4768	-0.016
	Y	1.83135	2.5295	-2.297
	Z	1.0567	1.7329	-1.536

α	0.81939	1.89941	
S	30.95	12.91	
σ	25.4	24.5	

GEOMETRICAL PARAMETER		CYCLE 0	1	2
r(C-O)	Å	1.36	1.418	1.445
r(C-C)		1.36	1.308	1.360
r(C-H)		1.09	1.07	1.07
COC	DEG	119	116.2	112.5
CCO		125	126.5	127.0
CCH		120	119.4	119.0
E_T (KJ mol^{-1})		0	-76.8	-115
[E_T (SADDLE)		-29.5	-104	-122]

Table 8 (CONTD.)

SELECTED ORBITAL ENERGIES - VARIATION WITH GEOMETRY

$-\varepsilon_i$/a.u.	CYCLE 0	1	2	BEST PLANAR	BEST SADDLE
a_1	0.370	0.397	0.394	0.320	0.368
e	0.396	0.413	0.397	0.418	0.379
a_1	0.447	0.459	0.460	0.476	0.385
e	0.495	0.491	0.488	0.544	0.495
a_2	0.606	0.585	0.585	0.527	0.517
					0.531
					0.551

NON-BONDED DISTANCES (a.u.)

	BEST PLANAR	INITIAL NON-PLANAR	FINAL NON-PLANAR
r(O-O)	6.87	5.41	5.57
r(C-C)	4.98	3.97	4.54

(Corresponding Distances in C_9H_{12} are 5.74, 4.65)

TOTAL ENERGY QUANTITIES OF FINAL NON-PLANAR TRIOXA-CYCLONON-ATRIENES (a.u.)

$-E_T$	453.87045	453.87295
$-V_{EL}$	886.70004	897.11143
$-V_{IE}$	1477.26773	1498.14507
T	459.75093	459.88152
V_{NN}	432.82959	443.23847
B.E.	1.38920	1.39170
R.E. (KJ mol^{-1})	-34	-28

AO POPULATIONS (TOTAL)

O	C	H
8.45	5.91	0.83
8.52	5.93	0.83

Appendix 1

Example of MO Method : H_2 Molecule.

It is instructive to consider an actual example in order to see the significance of the formal expressions of MO theory of Chapter 2. The simple case of the neutral H_2 molecule can be used in this respect. The electronic wave equation, in atomic units, is

$$(-\tfrac{1}{2}\nabla_1^2 - \tfrac{1}{2}\nabla_2^2 - \frac{1}{r_{1A}} - \frac{1}{r_{1B}} - \frac{1}{r_{2A}} - \frac{1}{r_{2B}} + \frac{1}{r_{12}})\psi = E\psi$$

where subscripts 1 and 2 refer to the two electrons, and A and B to the two nuclei. $E_T = E + \frac{1}{R}$, R being the internuclear distance. The equation may be simplified by noting that apart from the r_{12} term the Hamiltonian is a sum of two H_2^+ Hamiltonians,

$$\text{i.e.} \quad (H_1^N + H_2^N + \frac{1}{r_{12}})\psi = E\psi .$$

If the electron repulsion were neglected and $H = H_1^N + H_2^N$, ψ could be replaced by a product of two one-electron functions χ_1 and χ_2, which are simply eigenfunctions of the equation $H_1^N \chi_1 = \varepsilon_1 \chi_1$ (H_2^+ case). Thus, building up molecular wavefunctions as products of one-electron solutions for H_2^+, the MO configuration of H_2 is $(\phi_1\alpha)(\phi_1\beta) \equiv \phi_1\bar{\phi}_1$, i.e. two electrons with opposite spins in the lowest orbital of the type obtained for H_2^+. As a Slater determinant, wavefunction ψ is

$$\psi = |\phi_1\bar{\phi}_1| \quad \text{, shorthand for } \frac{1}{\sqrt{2}}[\phi_1(1)\bar{\phi}_1(2) - \phi_1(2)\bar{\phi}_1(1)].$$

ϕ_1 can be expressed in LCAO form. The energy of the ground state (Singlet) is

$$\langle\psi|H|\psi\rangle = \tfrac{1}{2}\langle\phi_1(1)\bar{\phi}_1(2) - \phi_1(2)\bar{\phi}_1(1) \mid H_1^N + H_2^N + \frac{1}{r_{12}} \mid \phi_1(1)\bar{\phi}_1(2) - \phi_1(2)\bar{\phi}_1(1)\rangle$$

$$= \tfrac{1}{2}[\ \langle\phi_1(1)H_1^N\phi_1(1)\rangle .4 + \langle\phi_1(1)\phi_1(2)\ \frac{1}{r_{12}}\ \phi_1(1)\phi_1(2)\rangle .2]$$

$$= 2\ \varepsilon^N\ (\phi_1)\ +\ J(\phi_1\phi_1)\ .$$

The simplification of the expression uses the facts that the H^N are one-electron operators, space and spin parts are separable, and uses spin orthogonality. The final result is that the energy of H_2 is twice the sum of the energy if there were only one electron in the molecule, plus the coulomb interaction between the two electrons, i.e. a one-electron integral and a two-electron integral.

If one of the electrons is excited to the next unfilled orbital ϕ_2, and both electrons have the same spin, the resulting state is a Triplet. The wavefunction is $\psi = |\phi_1\phi_2|$ so that the energy is

$$E = \langle\psi|H|\psi\rangle = \tfrac{1}{2}\langle\phi_1(1)\phi_2(2) - \phi_1(2)\phi_2(1)|H_1^N + H_2^N + \frac{1}{r_{12}}|\phi_1(1)\phi_2(2) - \phi_1(2)\phi_2(1)\rangle .$$

In the expansion, no terms disappear owing to spin orthogonality ($\int\alpha\beta\ ds = 0$).

Thus, $E = \varepsilon^N(\phi_1) + \varepsilon^N(\phi_2) + J(\phi_1\phi_2)$ (as above)

$$- \langle\phi_1(1)\phi_2(2)\ \frac{1}{r_{12}}\ \phi_1(2)\phi_2(1)\rangle$$

$$= \varepsilon^N(\phi_1) + \varepsilon^N(\phi_2) + J(\phi_1\phi_2) - K(\phi_1\phi_2).$$

The two energy expressions obtained for the two states of H_2 illustrate the quite general form of energy expressions for molecules. The total energy is the sum of the one-electron energies (the energy each electron would have

were it the only electron in the molecule), plus a coulomb integral (interaction between charge clouds) for every pair of electrons, minus an exchange integral for every pair of electrons in the molecule which have the same spin. The first two types of term are purely classical electrostatic,but exchange is a purely quantum mechanical effect with no physical meaning, resulting from the Pauli Principle, which ensures that the wavefunction is antisymmetric. It is the integrals between different product terms in the sum of products which give rise to K integrals as shown above.

(Reference: W.G. Richards and J.A. Horsely, "Ab Initio Molecular Orbital Calculations for Chemists", Clarendon Press, Oxford, 1973).

APPENDIX 2

MOLECULAR INTEGRALS .

A. Molecular Integral Evaluation .

In Chapters 2 and 3 it was shown that the computation, storage and retrieval of integrals over basis functions is the most time-consuming part of any orbital basis valence calculation. A great deal of research has been done in the areas of rapid, accurate computation of these integrals, and the design of an efficient file structure for storing the computed integrals for subsequent use. In this Section, the briefest outline of the derivation of integral formulae is given, basically in order to show the computational advantages of Gaussian Type Functions. The complete derivations are lengthy, but well-known . The molecular integrals involve the basis functions, η_i, and the one- and two-electron operators of the molecular Hamiltonian:

$$\int \eta_i(\chi_i)\hat{h}(i)\eta_j(\chi_i)d\chi_i$$

$$\int \eta_i(\chi_i)\eta_j(\chi_i)\hat{g}(i,j)\eta_k(\chi_j)\eta_\ell(\chi_j)d\chi_i d\chi_j ,$$

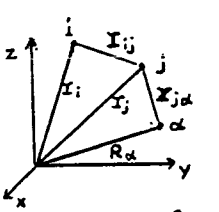


FIGURE 1.

where $\hat{h}(i) = -\frac{1}{2}\nabla^2(i) + \sum_{\alpha=1}^{N} -\frac{Z_\alpha}{r_{\alpha i}}$, $\hat{g}(i,j) = \frac{1}{r_{ij}}$

The notation used is defined in Figure 1, which is a diagram of an overall coordinate system to which all relative distances are referred. It is more convenient to represent the GTF's, η_i, of the general form $r^\nu \exp(-\alpha r^2)$ x (spherical harmonic), in the Cartesian-orientated form

$$\eta_i(\alpha,\ell,m,n,r_A) = N x_A^{\ell} y_A^{m} Z_A^{n} \exp(-\alpha r_A^2) \qquad -(1).$$

Here, the normalisation factor is

$$N = \left[\frac{2^{(\ell+m+n)+3/2}\,\alpha^{\ell+m+n+3/2}}{(2\ell-1)!!(2m-1)!!(2n-1)!!} \right]^{1/2}$$

$$((2\ell-1)!! = 1.3.5.....(2\ell-1)) \quad ;$$

$$x_A = x-X_A,\ y_A = y-Y_A,\ z_A = z-Z_A,\ r_A = r-r_A,\ r = (x^2+y^2+z^2)^{1/2}$$

and the function is centred at (X_A, Y_A, Z_A).

The Cartesian and spherical harmonic forms of the GTF's are very simply related; the s and p types are identical and the higher forms are simple linear combinations, e.g.

$$d(3z^2-r^2) = d(z^2) - \tfrac{1}{2}[d(x^2)+d(y^2)]$$

$$d(x^2-y^2) = \sqrt{3}/2\,[d(x^2)-d(y^2)] .$$

In order to retain the atomic orbital nomenclature, the GTF's are classified by the sum of the exponents of x_A, y_A, z_A:

$\ell+m+n = 0$ is an "s" type GTF

$\ell+m+n = 1$ is a "p" type GTF

$\ell+m+n = 2$ is a "d" type GTF, etc.

GTF's have the important property of allowing the various molecular integrals to be separated in Cartesian coordinates. The basis functions (and atomic orbitals) are centred on the various nuclei and therefore appear in the global coordinate system as functions of $(x_{i\alpha}, y_{i\alpha}, z_{i\alpha})$ and $r_{i\alpha}$, as in Figure 1. The centres of the Gaussians do not need to actually coincide with the physical nuclei of a molecular system; in fact, calculations of wavefunctions have used this fact to advantage.

The operations involved in the integral expressions are:

the overlap operator, which is simply 1;

the kinetic energy operator, $\tfrac{1}{2}\nabla^2$;

the electron-nucleus attraction operator, $1/r_{i\alpha}$

and the electron repulsion operator, $1/r_{ij}$.

In the following paragraphs, each class of molecular integral is considered and general expressions are given for the evaluation of molecular integrals using a GTF basis (equation (1)). Before giving the appropriate, purely mathematical formula in each case, some physical understanding of the type of integral involved is given by working with a particular example to illustrate the general principles.

(a) The Overlap Integral.

Considering as the integrand in this simplest type of integral the product of the two particular GTF's, a 1s orbital (e.g. on a H atomic centre) and a $2p_z$ orbital (e.g. centred on C nucleus), the overlap integral is

$$\int \eta(\text{H 1s})\,\eta(\text{C2p}_z)\,dr$$

$$= \int \exp(-\alpha^{H} r_H^2).Z_c \exp(-\alpha^{C} r_c^2)\,dr$$

(omitting the normalisation constants which play no rôle in the integration)

$$= \int (z-Z_c)\exp(-\alpha^{H} r_H^2 - \alpha^{C} r_c^2) \qquad -(2).$$

In equation (2), $r_H = |\, r-R_H \,| = [(x-X_H)^2 + (y-Y_H^2) + (z-Z_H^2)]^{\frac{1}{2}}$

and $r_c = |\, r-R_c \,| = [(x-X_c)^2 + (y-Y_c)^2 + (z-Z_c)^2]^{\frac{1}{2}}$,

where the H atom is at (X_H, Y_H, Z_H) and C is at (X_c, Y_c, Z_c).

The exponential term in the integrand can be simplified by the fact, familiar from statistics, that the product of the two normal (Gaussian) distributions is normal, i.e., the product of two 1s GTF's is a third

1s GTF centred on a point on the line joining the centres of the original GTF's. If

$$P = \frac{\alpha^H r_H + \alpha^C r_C}{\alpha^H + \alpha^C}, \quad r_p = r - P, \quad R = r_H - r_C;$$

then

$$\exp(-\alpha^H r_H^2 - \alpha^C r_C^2) = \exp\left[\frac{-\alpha^H \alpha^C |R|^2}{\alpha^H + \alpha^C}\right] \exp -(\alpha^H + \alpha^C)|r_p|^2 \qquad -(3).$$

The first exponential on the right-hand side of (3) is a constant and takes no further part in the analysis, and so is dropped. Since the remaining term in the integrand (2) can be expressed in terms of the components of r_p,

$$(z - Z_C) = Z_p + (P - R_C)_z = Z_p + c \quad (c \text{ is constant}),$$

the whole integrand can be expressed in terms of r_p as independent variable. Performing this transformation gives

$$\int (Z_p + c) \exp(-\alpha |r_p|^2) dx_p dy_p dz_p,$$

where $\alpha = \alpha^H + \alpha^C$. The integral can now be separated in Cartesian coordinates, giving

$$\int \exp(-\alpha x_p^2) dx_p \cdot \int \exp(-\alpha y_p^2) dy_p \cdot \int (Z_p + c) \exp(-\alpha z_p^2) dz_p \qquad -(4).$$

The integration limits for each of the coordinates x,y, and z are $\pm\infty$. Thus, for the GTF overlap integral, the basic integrals

$$\int_{-\infty}^{\infty} t^n \exp(-\alpha t^2) dt$$

have to be evaluated. Only the integrals with even n are non-zero, and

$$\int_{-\infty}^{\infty} t^{2n} \exp(-\alpha t^2) dt = \frac{(2n-1)!!}{2^n} \sqrt{\frac{\pi}{\alpha^{2n+1}}} \qquad -(5).$$

The overall value of the integral (2) can be obtained by applying (5) in expression (4).

The general overlap integral, using a basis of Slater Type Functions of the form $r^{\nu} \exp(-\alpha r) \times$ (spherical harmonic), can also be evaluated analytically. Although the three-dimensional integrand corresponding to that in (2) does not split into separate functions of x,y,z so that the multiple integral does not reduce to a product of three one-dimensional integrals, it is possible to transform to a coordinate system in which the multiple integrand can be reduced. This is the prolate spheroidal coordinate system, in which any STF overlap integral can be evaluated in terms of known standard integral forms.

The particular example considered above illustrates the main points in the derivation of the general formula for the overlap integral. With the GTF defined by α_i, ℓ_i, m_i, n_i with A the point of origin (equation (1)), the overlap integral between two GTF's centred at A and B is

$$(\alpha_1,\ell_1,m_1,n_1,A\ ;\ \alpha_2,\ell_2,m_2,n_2,B)$$

$$= N_1 N_2 \left(\frac{\pi}{\alpha_1+\alpha_2}\right)^{3/2} \exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}\overline{AB}^2\right) \times \sum_{(x)} G_i^x \sum_{(y)} G_j^y \sum_{(z)} G_k^z \qquad -(6)\,.$$

In equation (6), N_1 and N_2 are the normalisation constants (equation (1)); $\overline{AB}$ is the magnitude of the relative position vector, $|r_A - r_B|$; the symbolic summation signs $\sum_{(x)}$, $\sum_{(y)}$, $\sum_{(z)}$ emphasise the property of the reducibility of the molecular integral over GTF's into separate integrals over x,y,z - specification of one of these symbols is sufficient since the expressions for x,y, and z summations are precisely analogous - and

$$\sum_{(x)} = \sum_{i=0}^{L(\ell_1+\ell_2)/2}$$

with the symbol Li denoting the largest integer $\leq i$;

$$G_i^x = f_{2i}(\ell_1,\ell_2,\overline{PA}_x,\overline{PB}_x)\ \frac{(2i-1)!!}{2^i(\alpha_1+\alpha_2)^i}$$

with the auxiliary function

$$f_j(\ell,m,a,b) = \sum_{i=\max(0,j-m)}^{\min(j,\ell)} \binom{\ell}{i}\binom{m}{j-1} a^{\ell-i} b^{m+i-j}$$

defining the coefficients of powers of x in the expansion of $(x-a)\ (x-b)^m$, and $\overline{PA}_x$ is the component of the relative position vector $(r_p-r_A).\underline{e}_1$ ($\underline{e}_1$ is the unit vector in the x-direction).

(b) The Kinetic Energy Integral

The general analysis of the kinetic energy integrals is rather more involved because of the differentiations occurring in the integrand. It can be readily seen, however, that differentiating a GTF produces another GTF:

$$\delta/_{\delta x}(\exp -dr^2) = -2\alpha x\ \exp(-\alpha r^2)$$

Thus, the kinetic energy integrals are reducible to a weighted sum of overlap integrals, so that no new analysis is required in their evaluation. Again, the corresponding integrals over STF's can be evaluated analytically. In general, the kinetic energy integral over GTF's is:

$$(\alpha_1,\ell_1,m_1,n_1,A\ ;\ -\tfrac{1}{2}\nabla^2\ ;\ \alpha_2,\ell_2,m_2,n_2,B)$$

$$= N_1N_2[\alpha_2\{2(\ell_2+m_2+n_2)+3\}(\alpha_1,\ell_1,m_1,n_1,A;\ \alpha_2,\ell_2,m_2,n_2,B)$$

$$-2\alpha_2^2\{(\alpha_1,\ell_1,m_1,n_1,A\ ;\ \alpha_2,\ell_2+2,m_n,n_2,B)$$

$$+(\alpha_1,\ell_1,m_1,n_1,A\ ;\ \alpha_2,\ell_2,m_2+2,n_2,B)$$

$$+(\alpha_1,\ell_1,m_1,n_1,A\ ;\ \alpha_2,\ell_2,m_2,n_2+2,B)\}$$

$$-\tfrac{1}{2}\ell_2(\ell_2-1) \times (\alpha_1,\ell_1,m_1,n_1,A\ ;\ \alpha_2,\ell_2-2,m_2n_2,B)$$
$$-\tfrac{1}{2}m_2(m_2-1) \times (\alpha_1,\ell_1,m_1,n_1,A\ ;\ \alpha_2,\ell_2,m_2-2,n_2,B)$$
$$-\tfrac{1}{2}n_2(n_2-1) \times (\alpha_1,\ell_1,m_1,n_1,A\ ;\ \alpha_2,\ell_2,m_2,n_2-2,B)] \qquad -(7)\ .$$

This is the general expansion in terms of the seven possible overlap integrals generated by the differentiations implied by ∇^2. The expression looks unsymmetrical in that ℓ_1, m_1, n_1 appear in a form different from the appearances of ℓ_2, m_2, n_2; in fact, the integral shows correct Hermitian symmetry. If any (or all) of ℓ_2,m_2,n_2 are less than 2, then the relevant selection of the last three terms is omitted.

(c) The Nuclear Attraction Integral.

The analysis of the nuclear attraction, and electron repulsion, integrals is simplified by the property of GTF's expressed by equation (3). Further, the inverse distance operators can be expressed in a "Gaussian-like" form, at the expense of one additional integration, by the transformation:

$$r_{i\alpha}^{-1} = \pi^{-\frac{1}{2}} \int_o^\infty s^{-\frac{1}{2}} \exp(-sr_{i\alpha}^2)\,ds \qquad -(8)\ .$$

Remembering that $r_{i\alpha}^2$ splits into separate functions of the Cartesian coordinates, this transformation enables the nuclear attraction integrals to be split into three one-dimensional integrals over x,y,z and an additional "outer" integration over the transformation dummy variable s. The final integration over s reduces to the standard form:

$$\int_o^1 s^{2V} \exp(-ts^2)\,ds = F_v(t) \qquad -(9)\ .$$

This function $F_v(t)$ is related to the "error function" of statistics and methods for its evaluation are well known. Thus, even the most difficult molecular integrals reduce to known functions if GTF's are used, and so are analytically tractable. Conversely, with a basis of STF's,

a problem arises in the analysis of electron-nucleus attraction and electron repulsion integrals, when more than two centres are involved in the integrand. Thus, the "three-centre" nuclear attraction integral problem prevents the full analytical evaluation of even the one-electron Hamiltonian integrals for STF's as basis; this type of integral cannot be evaluated by standard analytical techniques and "brute-force" methods of direct numerical integration have to be used, and numerical quadrature in many dimensions is particularly time-consuming and not routinely applicable.

The general expression, using GTF's, for the nuclear attraction integral is:

$$(\alpha_1,\ell_1,m_1,n_1,A\ ;\ \frac{1}{r_c}\ ;\ \alpha_2,\ell_2,m_2,n_2,B)$$

$$= N_1N_2\left(\frac{\pi}{\alpha_1+\alpha_2}\right)\exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}\overline{AB}^2\right)$$

$$\times \sum_{(x)} A^x_{i,r,u} \sum_{y} A^y_{j,s,v} \sum_{z} A^z_{k,t,w}\ F_v(t) \qquad -(10)\ ;$$

in (10), $\sum_{(x)} = \sum_{i=0}^{\ell_1+\ell_2} \sum_{r=0}^{L_{i/2}} \sum_{u=0}^{L(i-2r)/2}$ etc.

and $A^x_{i,r,u} = (-1)^{i+u} f_i(\ell_1,\ell_2,\overline{PA}_x,\overline{PB}_x) \times \frac{i!\,\overline{CP}_x^{\,i-2(r+u)}\,\varepsilon^{r+u}}{r!u!(i-2r-2u)!}$

where $\varepsilon = \frac{1}{2}(\alpha_1+\alpha_2)$, $t = (\alpha_1+\alpha_2)\overline{CP}^2$,

$V = i+j+k-2(r+s+t)-u-v-w$ etc.

(d) The Electron Repulsion Integral

The electron repulsion integrals are physically and mathematically very similar to the nuclear attraction integrals; they represent the interaction between two electron distributions instead of between an

electron distribution and a point charge. Using the transformation:

$$r_{ij}^{-1} = \pi^{-\frac{1}{2}} \int_0^\infty s^{-\frac{1}{2}} \exp(-sr_{ij}^2) ds \qquad -(11) ,$$

an electron repulsion integral is simplified to six separate Cartesian integrations and the integral over s. With STF's, however, the general electron repulsion integral is analytically intractable; certain special cases can be treated analytically, but the normal situation requires numerical integration. The general polyatomic molecule, requiring numerical quadrature over 3,4 or 6 dimensions for the bulk of the STF integrals, cannot be routinely studied using a STF basis. Typically, the numerical computation of an integral yields a reliable value only after at least 10^3, 10^4 or 10^6 respectively computations of the integrand. There is a need to compute hundreds of thousands of the electron repulsion integrals, in general. Thus, the routine use of STF bases is restricted to atoms, small molecules and linear molecules. The basic problem is caused by the fact that distances, r_H and r_c, occur as such in the exponentials of the integrand, resulting in the irreducibility of the molecular integrals, in general.

The general expression for the electron repulsion integral is:

$$(\alpha_1,\ell_1,m_1,n_1,A\ ;\ \alpha_2,\ell_2,m_2,n_2,B\ ;\ \frac{1}{r_{12}}\ ;\ \alpha_3,\ell_3,m_3,n_3,C\ ;\ \alpha_4,\ell_4,m_4,n_4,D)$$

$$= N_1N_2N_3N_4 \frac{2\pi^2}{\gamma_1\gamma_1} \left(\frac{\pi}{\gamma_1+\gamma_2}\right) \exp\left(-\frac{\alpha_1\alpha_2\overline{AB}^2}{\gamma_1} - \frac{\alpha_3\alpha_4\overline{CD}^2}{\gamma_2}\right)$$

$$\times \sum_{(x)} B^x_{i_1,i_2,r_1,r_2,u} \sum_{(y)} B^y_{j_1,j_2,s_1,s_2,v} \sum_{(z)} B^z_{k_1,k_2,t_1,t_2,w} F_v(t) \qquad -(12).$$

In (12), $\sum_{(x)} = \sum_{i_1=0}^{\ell_1+\ell_2} \sum_{i_2=0}^{\ell_3+\ell_4} \sum_{r_1=0}^{Li_{1/2}} \sum_{r_2=0}^{Li_{2/2}} \sum_{u=0}^{u'}$ etc.,

where $u' = L((i_1+i_2)/2 - r_1 - r_2)$;

$$\gamma_1 = \alpha_1 + \alpha_2,\ \gamma_2 = \alpha_3+\alpha_4,\ t = \overline{PQ}^2/4(\gamma_1^{-1} + \gamma_2^{-1}),$$

$$V = i_1+i_2+j_1+j_2+k_1+k_2-2(r_1+r_2+s_1+s_2+t_1+t_2)-u-v-w;$$

$$B^x_{i_1,i_2,r_1,r_2,u} = (-1)^{i_2+u} f_{i_1}(\ell_1,\ell_2,\overline{PA}_x,\overline{PB}_x) f_{i_2}(\ell_3,\ell_4,\overline{QC}_x,\overline{QD}_x)$$

$$\times \frac{i_1!i_2!\gamma_1^{r_1-i_1}\gamma_2^{r_2-i_2}(2\delta)^{2(r_1+r_2)}}{r_1!r_2!(i_1-2r_1)!(i_2-2r_2)!(4\delta)^{i_1+i_2}}$$

$$\times \frac{(L(i_1+i_2-2r_1-2r_2))!\ \overline{PQ}_x^{\,i_1+i_2-2(r_1+r_2+u)}\delta^u}{u!(L(i_1+i_2-2(r_1+r_2+u)))!}$$

$$\delta = \tfrac{1}{4}(\gamma_1^{-1} + \gamma_2^{-1}) \quad \text{etc.}$$

The working formulae for evaluating molecular integrals in a basis of GTF's are presented in full generality above. It is instructive to mention some details of implementation of these mathematical expressions, and this is done in the following section.

B. Practical Details.

In the preceding Section the analytical expressions for the necessary molecular integrals are given; these formulae refer to single, uncontracted Gaussian functions, characterised by type (ℓ,m,n), which corresponds to the familiar representation of the angular component of the function, orbital exponent α, and centre A (equation (1)). In practice, the expressions are evaluated numerically (as opposed to numerical integration required with STF's), and this is very dependent on the values of ℓ, m, n of the functions in the integrand. Integrals involving s-type orbitals can be evaluated extremely rapidly; p and d orbitals involve integrals which are rather slow to evaluate and higher values of $(\ell+m+n)$ give molecular integrals which are prohibitively time-consuming for most applications. Many molecules of chemical interest involve bonding among s,p and d orbitals and the computation of molecular integrals for these molecules using a GTF basis is quite a straightforward task. Thus, from a practical point of view, it is wasteful to program the preceding expressions in full generality. Some examples of the more typical expressions required are given below. The normalisation constant for a single Gaussian (equation (2)) reduces to the simpler expressions:

$$N = \frac{\alpha^{3/2}}{(\frac{1}{2}\pi)^{3/4}} \qquad \text{for a s-type function,}$$

$$N = \frac{4\alpha^{5/2}}{(\frac{1}{2}\pi)^{3/4}} \qquad \text{for a p-type function} \qquad -(7)\,.$$

The overlap integral between two uncontracted Gaussians (equation (3)), for two of s-type, is:

$$(\alpha_1,0,0,0,A\ ;\ \alpha_2,0,0,0,B) = N_1N_2\ \left(\frac{\pi}{\alpha_1+\alpha_2}\right)^{3/2}\ \exp\left(-\frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2}\ \overline{AB}^{\,2}\right)$$

$$\equiv S^{OO}(\alpha_1,\alpha_2) \qquad -(8)\ .$$

For the combinations of s- and p- type functions giving rise to overlap integrals, simple expressions involving S^{OO} can be derived. The kinetic energy integral for any pair of uncontracted Gaussians (equation (4)) is expressible in the special formula for s-type functions only:

$$(\alpha_1,0,0,0,A\ ;\ -\tfrac{1}{2}\nabla^2\ ;\ \alpha_2,0,0,0,B) \equiv T^{OO}(\alpha_1,\alpha_2)$$
$$= K^{OO}(\alpha_1,\alpha_2)S^{OO}(\alpha_1,\alpha_2) \qquad -(9)$$

$$\text{where } K^{OO}(\alpha_1,\alpha_2) = \frac{3\alpha_1\alpha_2}{\alpha_1+\alpha_2} - \frac{2\alpha_1^2\alpha_2^2}{(\alpha_1+\alpha_2)^2}\ \overline{AB}^{\,2}\ .$$

The other special formulae for kinetic energy integrals, involving s- and p-type functions only, involve the sets of simple K- and S-type functions.

For a pair of s-type Gaussians, the nuclear attraction integral (equation (5)) is:

$$(\alpha_1,0,0,0,A\ ;\ \frac{1}{r_c}\ ;\ \alpha_2,0,0,0,B) \equiv V^{OO}(\alpha_1,\alpha_2)$$

$$= 2(\pi)^{-\frac{1}{2}}(\alpha_1+\alpha_2)^{\frac{1}{2}}S^{OO}(\alpha_1,\alpha_2)L^{OO}(C;\alpha_1,\alpha_2) \qquad -(10)$$

$$\text{where } L^{OO}(C;\alpha_1,\alpha_2) = F_o(t) = \int_o^1 \exp(-tu^2)\,du.$$

Each integral between a pair of functions is summed over the number of nuclei of the system (including nuclear charge) to give the total for the pair. All s- and p-type nuclear attraction integrals can be expressed in terms of S- and L-type functions. The expressions for the electron repulsion integrals over s- and p-type Gaussians are quite lengthy, but reasonably straightforward.

In defining the integrals over Gaussian functions, the functions $f_j(\ell,m,a,b)$ and $F_v(t)$ were used. The auxiliary function f_j, involving products of binomial coefficients, appears not only in the one-electron integrals but especially in the two-electron integrals. For cases where only s- and p-type Gaussians are involved, the f_j functions are easily evaluated. The $F_v(t)$ function, a form of the incomplete gamma function, is computed in a number of ways: by asymptotic expansion, by recursion formulae, or by tabular interpolation, depending upon the values of V and t, well-known algorithms of numerical analysis.

It is instructive to give some consideration to the physical interpretation and orders of magnitude of molecular integrals. The "coulomb terms" in the molecular Hamiltonian give rise to the orbital basis molecular integrals which have a direct electrostatic interpretation, and which suggest the physical interpretation of the kinetic energy integrals. Recalling the probability density interpretation of the orbital products (Chapter 2, Section B), $P_{ij}(r) = \eta_i(r)\eta_j(r)$ is the probability density of an electron confined to the region of overlap between η_i and η_j. The integral of $P_{ij}(r)$ over all space is a measure of how strongly the two orbitals overlap. The electron-nucleus attraction integrals,

$$\int \eta_i(r)\left(-\frac{Z_\alpha}{r_\alpha}\right)\eta_j(r)\ dr = -Z_\alpha \int \frac{P_{ij}(r)}{r_\alpha}\ dr,$$

are simply the coulomb energy of a negative charge distribution $P_{ij}(r)$ in the field of a positive point charge Z_α. The electron repulsion integrals,

$$\iint \eta_i(r_1)\eta_j(r_1)\,\frac{1}{r_{12}}\,\eta_k(r_2)\eta_\ell(r_2)\,dr_1 dr_2$$

$$= \iint P_{ij}(r_1)\,\frac{1}{r_{12}}\,P_{k\ell}(r_2)\,dr_1 dr_2,$$

represent the mean repulsion energy between two negative charge distributions $P_{ij}(r_1)$ and $P_{k\ell}(r_2)$. Rearrangement of the kinetic energy integralsinto this probability density form is not possible because of the essential operator nature of ∇^2. However, by analogy with the other integrals, the kinetic energy integral can be interpreted as the mean kinetic energy of a distribution of charge given by $P_{ij}(r)$.

This interpretation of the molecular integrals enables some qualitative consideration of some relative magnitudes to be made. If the centres of the two functions η_i and η_j are widely separated, then the exponential form of the η's ensures that the product $\eta_i\eta_j$ is small everywhere, so that any integrals in which $\eta_i\eta_j$ appears as a factor in the integrand are small. The overlap integral, the sum of the distribution, is small. The other molecular integrals are usually small when the overlap integral involving the same functions is small. It can therefore be expected that many of the electron repulsion integrals, involving two P_{ij} factors in the integrand, will be small since, in a typical molecule, there are many "distant" pairs of functions. The overlap integral may also be small, or zero, for reasons of symmetry; the positive and negative regions of overlap may cancel out. However, in general, in such cases nothing can be said about the sizes of the other integrals. The idea that the overlap density (differential overlap), or the overlap integral, governs

the relative sizes of the various electron repulsion integrals is the basis of many approximation methods within the MO framework. The size of the overlap is used as a criterion for the systematic neglect of classes of electron repulsion integral. Unfortunately, it is impossible to say, a priori, what effect their neglect has on a MO calculation.

In practice, in computing integrals between Gaussian functions of different types centred at different nuclei, three different situations can arise. Firstly, an integral may have a given absolute value, larger than a given threshold, such that it must be computed to full accuracy. Secondly, an integral may be smaller than a given threshold, but sufficiently large that it should be computed to some given accuracy. Thirdly, an integral may be so small that it can be entirely neglected, i.e. approximated by zero. In general, the one-centre integrals, of any category, are of the first type. The many-centre integrals can be of all three types. If the contraction technique is used, whereby an atomic orbital is expressed as a linear combination of basis functions, there is a further consideration of great significance; the number finally stored and used in a SCF calculation is the value of an integral over AO's, whereas integrals are evaluated over the component Gaussian basis functions, and then the relevant summations are made to yield the final AO value. For a two-electron integral over AO's,

$$G_{ijk\ell} = \int \chi_i(1)\chi_j(1)\ \frac{1}{r_{12}}\ \chi_k(2)\chi_\ell(2)dr_1 dr_2,$$

with

$$\chi_i = \sum_1^{N_i} \eta_r,\quad \chi_j = \sum_1^{N_j} \eta_s,\quad \chi_k = \sum_1^{N_k} \eta_t,\quad \chi_\ell = \sum_1^{N_\ell} \eta_u,$$

in order to compute

$$G_{ijk}\ ,\ (N_i \times N_j \times N_k \times N_\ell)$$

integrals over basis functions have to be evaluated. Thus, two

thresholds of integral value are used in practice. The first is referred to in the evaluation of an integral over Gaussian primitive functions; a critical aspect in determining overall speed of computation of integrals is the use of efficient algorithms for estimating the value of such an integral, so that the more laborious exact computation is only effected for those integrals which are determined to be larger than the threshold value. The second threshold refers to the value of an AO integral, obtained by summing the GTF contributions. In the non-empirical type of calculation used in this work all integrals involved are considered; only those (two-electron integrals) having a value below a certain threshold are neglected. Full accuracy in computation is fourteen decimal figures, so that a threshold of 10^{-13} for integral evaluation over primitives can be used. Allowing for the possibility of large groups of Gaussians representing the AO's (e.g. 6 primitive per group), the second threshold value of 10^{-10} corresponds to the above one. However, less than this full accuracy is adequate, especially for large molecule calculations; a value of 10^{-7} for the AO integral threshold is reasonable in practice. With the size of basis set used here, a value of 10^{-8} for the primitive integral threshold is adequate.

Some effort has been made to overcome the particular problem of the two-electron integrals - computation of a very large number of integrals, whose values cover a very wide range. It has not proved possible, by analysis of the distribution of values, to systematically approximate various types of integral involved, particularly the ones with small values. Other attempts to facilitate computation involve the use of special Gaussian basis sets, e.g. Gaussian "lobe" functions involving only s-type functions for easier integration and expressing higher functions as linear combinations of these. Easier computation of integrals in these cases

must be balanced against loss of accuracy in physical representation.

Having considered the problem of efficient computation of molecular integrals, it is necessary to develop an organisation of the file of integrals evaluated. In the design of an efficient, flexible file structure for storage, and retrieval, of AO integrals, the electron repulsion integrals again present the most difficulties. The storage of the one-electron integrals is not critical in view of the relatively small number of them; the matrix (two-dimensional array) of these integrals is conveniently stored. In practice, if there are m AO's involved, the distinct elements only are stored, requiring $\frac{1}{2}m(m+1)$ locations for each type of integral. The treatment of the four-index electron repulsion integrals, $G_{ijk\ell}$, is quite different. The use of a four-dimensional array, of m^4 elements, is not practicable in general; also, important equalities among the integrals $G_{ijk\ell}$ permit considerable space saving - $\frac{1}{8}(m^4 + 2m^3 + 3m^2 + 2m)$ in place of m^4 locations. This latter point is due to the permutation of indices and is independent of molecular symmetry. With complete generality,

$$G_{ijk\ell} = G_{jik\ell} = G_{ij\ell k} = G_{k\ell ij} = G_{\ell kij} = G_{k\ell ji} = G_{\ell kji} = G_{ji\ell k} \quad -(11).$$

Thus, only one of these integrals need be computed and stored. In using the stored integrals, and these equalities, it must be remembered that there can be 1,2,4, or 8 integrals with the same value as $G_{ijk\ell}$, depending on equalities among i,j,k,ℓ. Obviously, in deciding to work with only one of the equivalent AO integrals, an ordering convention among i,j,k,ℓ must be used to decide which one to store; when the computed integrals have been stored, they need to be identified for subsequent access in the SCF part of the calculation. The unique identification of an integral in the file is arranged by storing the values of the labels, i,j,k,ℓ together with the value of the integral $G_{ijk\ell}$. Despite the significant increase

in space required per integral in this method, there are the over-riding advantages that any integrals which are zero can be omitted, and molecular symmetry can be used to full advantage in the integral calculation by placing groups of integrals equal by symmetry together in the file as they are computed.

The most common use of a file of AO electron repulsion integrals is in the formation of the matrix G(R) during an LCAO MO calculation (Chapter 2, equation (43)). The definition of the elements of G(R) is

$$G_{ij} = \sum_{r,s,=1}^{m} R_{rs}[2(ij,rs) - (ri,js)] \qquad -(12).$$

In the formation of G(R), rather than considering each element G_{ij} in turn and locating the relevant component integrals, so that the file would have to be read a total of $\frac{1}{2}m(m+1)$ times for the whole matrix, the file is read only once (much more feasible, for technical reasons). Each integral on the file is examined, and is added to all the possible elements of G to which it contributes (multiplying by the appropriate element of R). The matrix G is formed by one "pass" of the integral file; this must be done in each cycle of the SCF calculation.

The computational time in molecular calculations can be reduced if there are common parts in molecular systems. For a sufficiently flexible basis set describing a molecular system, a large fraction of the integrals corresponding to the basis set can be used in a somewhat different system. If, in a system of N atomic centres, M of these have different cartesian coordinates with respect to the coordinates used in a previous computation, then only the integrals referring to the M centres, with the new coordinates, need to be computed. The file of integrals from the previous computation can be read and the common

integrals extracted, and re-labelled if necessary, to be combined with the newly computed ones. The nuclear charges of the common nuclei and the basis sets must be identical. In this way, e.g. in the study of related species, significant savings in computational expense can be made.

APPENDIX 3

SCF CALCULATION .

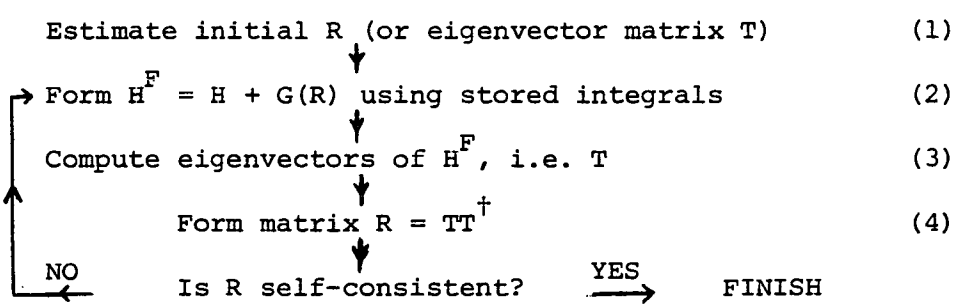
A. Matrix Diagonalisation.

A molecular electronic wavefunction, within the LCAO MO model, is obtained by solving the Roothaan-Hartree-Fock Equation, which in matrix form is:

$$H^F T = T\varepsilon \qquad -(1),$$

where T is the matrix of AO orbital coefficients (MO $\phi_i = \sum_j T_{ij}\chi_j$, as in Chapter 2), ε is the diagonal matrix of eigenvalues, and H^F is the RHF matrix, = H + G(R), matrices whose elements are the molecular integrals (Appendix 2). Orthogonality of AO's is assumed for the moment. Thus, the matrix eigenvalue problem (1) is to be solved. The computation of the eigenvalues and eigenvectors (columns of T) of H^F is basically a matrix diagonalisation problem; however, since H^F contains T through $G(R = TT^\dagger)$ any solution must be iterative, yielding a self-consistent solution (LCAO SCF MO's). The iterative process of solution is illustrated in Figure 1.

Figure 1



The computational problem of deriving an algorithm for the efficient generation of G(R) from the stored electron repulsion integrals was mentioned in Appendix 2. The computation of the eigenvalues and eigenvectors of a symmetric matrix (H^F) is a well-researched problem. An outline of one widely used "class" of methods is given below.

The matrix LCAO MO problem $H^F T = T\varepsilon$, with H^F (m x m) and $T(m \times {}^{n}/_{2})$ ($^{n}/_{2}$ occupied MO's, $^{n}/_{2} < m$), is an example of the more general matrix eigenvalue problem $H^F U = U\varepsilon$, where U is (m x m). The condition that the columns of U be normalised and orthogonal, $UU^{\dagger} = I$, means that U is unitary, $U^{\dagger} = U^{-1}$, and so $H^F U = U\varepsilon$ is equivalent to $U^{\dagger}H^F U = \varepsilon$. The problem is thus to compute a matrix U which transforms H^F into diagonal form via the relation $U^{\dagger}H^F U = \varepsilon$. Some light can be shed on the general problem by an analysis of the simplest case in which all matrices are (2 x 2); U is required such that

$$\begin{bmatrix} U_{11} & U_{21} \\ U_{12} & U_{22} \end{bmatrix} \begin{bmatrix} H^F_{11} & H^F_{12} \\ H_{21} & H^F_{22} \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} = \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix} \qquad -(2).$$

The condition of U being unitary reduces the number of independent elements in U from 4 to 1 since $UU^{\dagger} = I$ is, in full:

$$U_{11}^2 + U_{21}^2 = 1 \; ; \quad U_{12}^2 + U_{22}^2 = 1 \; ;$$

$$U_{12}U_{11} + U_{22}U_{21} = 0 \; ; \quad U_{11}U_{12} + U_{21}U_{22} = 0.$$

These equations can be given a convenient interpretation by considering U as the matrix defining a rotation in a two-dimensional space by an angle θ. Defining $c = U_{11} = \cos\theta$ and $s = U_{12} = \sin\theta$, then $U_{21} = -s$ and $U_{22} = c$ satisfy $UU^{\dagger} = I$ identically for all θ. Therefore, the matrix equation (2) can be re-written as:

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} H^F_{11} & H^F_{12} \\ H^F_{12} & H^F_{22} \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix} \qquad -(3)$$

using the fact that $H^F_{12} = H^F_{21}$. Equation (3) becomes

$$H^F_{12}(c^2-s^2) + sc(H^F_{22}-H^F_{11}) = 0 \qquad -(4);$$

this is a necessary and sufficient condition for the solution of (3).

Defining

$$C = c^2-s^2 = \cos 2\theta \quad \text{and} \quad S = 2sc = \sin 2\theta, \text{ and using}$$

$$c^2 = \frac{1+C}{2}, \qquad s^2 = \frac{1-C}{2}, \quad (4) \text{ becomes}$$

$$CH^F_{12} + \tfrac{1}{2}S(H^F_{22} - H^F_{11}) = 0, \text{ i.e. } \tan 2\theta = \frac{S}{C} = \frac{2H^F_{12}}{H^F_{11}-H^F_{22}},$$

This is the complete solution of the (2 x 2) problem since U is defined by θ through the above relations.

In general, with larger matrices, it seems obvious to try a method for deriving U involving the application of a succession of (2 x 2) "rotations" to the matrix H^F, and thus the elimination of the off-diagonal elements one at a time. Actually, this is not quite possible because the (2 x 2) rotations interfere with each other, i.e. on making an element "disappear", some elements earlier transformed to zero may reappear. Various methods have been devised, based on the rotation principle, to overcome this interference. The one used in the computations in this work is the Jacobi Method . This approach uses the optimistic idea that, if all the off-diagonal elements are transformed to zero in turn and sufficiently often, the cumulative effect of the whole set of transformations is to diagonalise the matrix. This proves to be the case; it is simply necessary to keep the process going until all the off-diagonal elements fall below some pre-set tolerance. The method is extremely easy to program, is very stable numerically and is very well-behaved at degenerate eigenvalues. It is actually the slowest of the commonly used diagonalisation methods, time consumption being proportional to the fourth power of the size of the matrix. Some important points in implementation in the LCAO MO SCF problem are given in the next section.

The analysis of the implementation of the LCAO MO method has been over-simplified. No allowance has been made for non-orthogonality of the AO's. In practice, AO's in a molecular calculation are not orthogonal, and the RHF equation is, strictly, $H^F T = STε$, where S is the overlap matrix ($S_{ij} = \int \chi_i \chi_j dr$). In fact, no numerical methods beyond matrix diagonalisation are required; simply, three more matrix multiplications during each iteration are necessary. An orthogonal set of AO's can be defined by a linear transformation among the original, non-orthogonal AO's:

$$\bar{\chi} = \chi X \qquad -(5)$$ (cf. discussion of localised MO's in Chapter 2, Section B).

The calculated value of any expectation value, in particular the total electronic energy, is invariant with respect to any (non-singular) linear transformation among the orbital functions. In equation (5), χ is the original AO set, $\bar{\chi}$ is an orthogonal AO set, X is the matrix of transformation coefficients. The transformed overlap matrix, $\bar{S} = X^\dagger S X = I$. Thus, the condition for the matrix X to generate an orthogonal set becomes $XX^\dagger = S^{-1}$, so that the overlap matrix over the original AO functions must not be singular if orthogonalisation is to be successful. The RHF equation, using non-orthogonal orbitals is:

$$H^F T = STε$$

$$\rightarrow H^F IT = SITε$$

$$\rightarrow H^F XX^{-1}T = SXX^{-1}Tε \qquad (XX^{-1} = I \text{ for any } X)$$

$$\rightarrow (X^\dagger H^F X)(X^{-1}T) = (X^\dagger SX)(X^{-1}T)ε$$

$$\rightarrow \bar{H}^F \bar{T} = \bar{T}ε \quad \text{since } X^\dagger H^F X = \bar{H}^F \text{ and}$$

$$X^\dagger SX = \bar{S} = I,\ X^{-1}T = \bar{T}.$$

The last equation is simply the RHF equation using the orthogonal basis $\bar{\chi}$. The logic of the implementation of the LCAO MO method using non-orthogonal AO's is illustrated by modifying Figure 1 as follows:-

before step (1), computation of matrix X from S;

before step (3), formation of $\bar{H}^F = X^\dagger H^F X$;

step (3), computation of eigenvectors of $\bar{H}^F$, i.e. $\bar{U}$

before step (4), formation of $U = X\bar{U}$, hence T (part of U).

The most important computational point in this scheme is that the transformation of the electron repulsion integrals to the orthogonal basis, which is an exceptionally demanding computational problem, does not have to be performed. The simple transformation of $\bar{H}^F$ is all that is required, so that no new procedures are required to handle non-orthogonal orbitals in LCAO MO calculations; the "overhead" due to non-orthogonality is the three matrix multiplications per iteration indicated above.

The orthogonalising matrix X of equation (5),given by the condition $X^\dagger X = S^{-1}$, is determined only to within a unitary transformation, as $XBB^\dagger X^\dagger = (XB)(XB)^\dagger = S^{-1}$ provided B is unitary ($BB^\dagger = I$). The method used in the computations in this work to determine X is the Schmidt Orthogonalisation Method, well-known in linear algebra; basically, the set of orthogonal orbitals is constructed by forcing successive members to be orthogonal to the existing functions, using a recursive formula:

$$\bar{\chi}_{i+1} = \chi_{i+1} - \sum_{j=1}^{i} \bar{\chi}_j S_{ij} \qquad -(6)$$

where $S_{ij} = \int \chi_i \chi_j dr$. The resulting X matrix, after carrying through this process is of "upper triangular" form, corresponding to each new orbital being a linear combination of the earlier ones. The calculation of such a matrix is typical of matrix inversion procedures (X is related to S^{-1}). The computational implementation of the method closely

parallels matrix inversion methods, and also suffers from the same difficulties with cumulative errors. Generation of X by (6) causes the building-up of rounding errors, particularly if some off-diagonal elements of S are large (i.e. S is almost singular). Numerical methods for the Schmidt procedure have been developed which avoid the build-up of errors. The use of orthogonal AO's is a computational convenience in solving the RHF equation rather than a factor of chemical importance. It is usually more meaningful, physically, to consider the derived MO's in terms of the non-orthogonal AO's, using the matrix of coefficients T.

B. Implementation of SCF Calculations .

In performing a closed-shell (restricted) Roothaan-Hartree-Fock calculation, before applying matrix diagonalisation techniques to H^F as in Section A, the first step is to provide an estimate of the initial matrix R (Figure 1); thus, the iterative process of solution is started from a physically realistic guess at the eigenvectors (matrix T), which define the trial set of molecular orbitals. There are three main ways of obtaining, in practice, trial eigenvectors to act as input to the first SCF cycle. One method is to diagonalise the "unscreened" one-electron Hamiltonian operator matrix ("core") (i.e. neglecting G(R) in $H^F = H + G(R)$), as though performing a calculation ignoring electronic interaction. This straightforward diagonalisation leads to a set of eigenvectors, of which the 'n' of lowest eigenvalue are deemed to define the trial set (assuming n doubly-occupied MO's in the case considered). A better mechanism for starting-up is to provide a set of numbers which are set to the negatives of the expected values of the diagonal elements of the H^F matrix at self-consistency. The utility of this procedure stems from the fact that diagonal RHF Hamiltonian matrix elements remain approximately invariant under change in molecular environment for a given basis set. It is possible to form a 'library' of appropriate parameters, after studying some model cases, which will be reasonably accurate in general. The trial eigenvectors are derived as follows:

H^F_{ii} = -A(I), where A(I) is the Ith element of the vector of input parameters - estimate of Ith diagonal element of converged H^F.

The matrix representations of the kinetic energy operator (T) and nuclear attraction operator (V) are available (molecular integrals over AO's). Thus, given some approximation for the coulomb-exchange

2-electron operator (G), an approximate H^F would be deduced from $H = T + V + G$.

The diagonal elements of G can be found from $G_{ii} = H_{ii} - T_{ii} - V_{ii}$, where the guessed values of H_{ii} are used. The off-diagonal elements of G and V can be expressed in the form of Mulliken approximations:

$G_{ij} = K(G_{ii} + G_{jj})$ and $V_{ij} = K(V_{ii} + V_{jj})$, where K is some unknown scalar. Then $G_{ij} = V_{ij}(G_{ii} + G_{jj})/(V_{ii} + V_{jj})$, so that

$H^F_{ij} = T_{ij} + V_{ij} + V_{ij}(G_{ii} + G_{jj})/(V_{ii} + V_{jj})$. This latter expression completes the guessed form of the converged H^F. The eigenproblem $H^F T = ST\varepsilon$ is solved (S is matrix of AO overlap integrals), and the resulting eigenvectors (T) are arranged in order of ascending eigenvalues (ε), thus defining trial MO's as linear combinations of the AO's.

The third method, obviously much more particular, is simply to use a set of eigenvectors derived in a closely related case as starting-points, e.g. similar geometries for the same species using the same basis.

Having provided a trial set of eigenvectors (trial electron distribution) matrix H^F is formed from stored matrices H and G. Diagonalisation of H^F, i.e. computation of a new T (or $R = TT^\dagger$), leads to a new definition of H^F. The process is repeated until the new R (one-particle spin-free density matrix) does not differ from its predecessor by more than some tolerance decided, or physical grounds, in advance. Thus, a wavefunction of any desired accuracy may be achieved; however, a practical SCF convergence threshold of 10^{-4} - 10^{-5} is a reasonable and worthwhile one. This means that the density matrix elements are converged to within absolute error $X(\sim 10^{-5})$, whilst the total electronic energy is correct within an absolute error X^2.

It is hoped that this SCF cycling leads to successive R matrices which converge to a final self-consistent solution, rather than oscillating or diverging away from the true solution. In this work, the cases considered are usually well-behaved. However, non-convergence of the iterative procedure to a self-consistent R (or T) sometimes occurs; typically, oscillation rather than divergence, is the case, usually interpretable physically as the current non-self-consistent T matrix corresponding to two possible electronic distributions of very similar energies. A method of inducing convergence in difficult cases is based on the use of damp factors and level shifters . An aspect of the physical nature of the problem is introduced into the mathematical matrix diagonalisation. Thus, in a molecular calculation in a basis of m linearly independent real AO's, for a 2n electron system, a set of n orthonormal doubly-occupied MO's (DOMOS) and a set of $n' = m-n$ virtual MO's (VMOS) are constructed, such that all m MO's form an orthonormal set. With χ, ϕ, ϕ' denoting row vectors of AO's, DOMOS, VMOS respectively, $(\phi \,\vdots\, \phi') = \chi(T \,\vdots\, T')$, where a column of matrix (TT') contains the AO coefficients of a given MO. The total wavefunction is invariant to mixing between DOMOS, so that only variations which mix DMOS with VMOS need be considered. It is possible to effect such mixing so that there is generated a perturbed wavefunction of lower energy than the unperturbed function; thus, convergence can be guaranteed, although not necessarily to the lowest energy stationary point. From a trial set of DOMOS, H^F can be constructed in the AO representation and diagonalised. A minor variant, yielding equivalent results, is to diagonalise H^F in the basis of trial MO's. The resulting eigenvectors, after some suitable ordering process usually based on the "aufbau" principle, define the iterated, or "improved", MO's as linear combinations of the trial MO's. Then, using these two representations, $H^F_{MO} = (T \,\vdots\, T')\, H^F_{AO} (T \,\vdots\, T')$

$$= \left[\begin{array}{c|c} T^{\dagger} H^{F}_{AO} T & T^{\dagger} H^{F}_{AO} T' \\ \hline T'^{\dagger} H^{F}_{AO} T & T'^{\dagger} H^{F}_{AO} T' \end{array}\right] \equiv \begin{bmatrix} A & B \\ C & D \end{bmatrix} .$$

If the eigenvectors of H^{F}_{MO} are partitioned in the form $(Q \vdots Q')$, then the iterated MO's may be expressed as linear combinations of AO's:

$$(\phi \vdots \phi') = \chi(T \vdots T')(Q \vdots Q').$$

The damp factor (λ) and level shifter (b) are used to define a modified Hamiltonian matrix:

$$H^{F}_{MOD} = \begin{bmatrix} A & \lambda B \\ \lambda C & D+bI \end{bmatrix}$$

where I is identity matrix of order n'.

The modifications are that the level shifter is added to each diagonal element of the block $T'^{\dagger} H^{F}_{AO} T$, whilst the elements of the off-diagonal blocks are multiplied by the damp factor. The procedure adopted is to diagonalise the modified matrix, and order the eigenvector columns based on the resulting eigenvalues. The first n columns define iterated DOMOS as linear combinations of the trial MO's. An analysis of this diagonalisation by first-order perturbation theory , in summary, shows that a sufficiently level-shifted procedure will give convergence to some stationary point on the energy surface, given any trial set of MO's. A damp factor setting of unity, and level shifter of zero, corresponds to the orthodox Roothaan procedure in performing SCF iterations. In practice, in difficult cases, experience has indicated that little is gained by choosing values of the damp factor different from unity; study of the effects of the level shift parameter on the rate of convergence (λ=1) has shown that a common feature of all convergence curves is that convergence sets in at some (generally unknown) critical value of the level shifter, with the optimum rate occurring at somewhat higher values.

Typically, b $\approx$ 1 often proves successful in forcing convergence.

As a result of the level-shifting technique for forcing convergence, the possibility arises that convergence to stationary points on the energy surface may occur, such that the converged MO's do not obey the "aufbau" principle, so that the resulting orbital configuration does not represent the ground state (MO ordering is based on eigenvalues of the level-shifted H^F - if large positive b's used, convergence to excited states with respect to the orthodox H^F is possible). To circumvent this difficulty, switching of VMOS and DOMOS can be used to force convergence to an "aufbau" configuration. Another practical consideration is that excessively high b's should not be used; convergence is then undoubtedly guaranteed, but at an unduly low rate.

If $T^{(k)}$ denotes the eigenvectors defining trial MO's for the kth cycle as linear combinations of the AO's, and $Q^{(k)}$ the eigenvectors resulting from the kth cycle, then

$$T^{(k+1)} = T^{(k)}Q^{(k)} = T^{(1)}Q^{(1)}Q^{(2)}Q^{(3)}.....Q^{(k)}.$$

This is not numerically satisfactory, since the build-up of round-off error is accumulated over k cycles. Round-off error can occur both in the indicated matrix multiplications and in the diagonalisations used to generate the Q matrices. To stabilise the method, the Schmidt procedure (Section A) is used to "refresh" the orthonormality of the columns of the T matrices after each cycle. A refreshed T matrix obeys the condition: $T^{\dagger}ST = I$ (S is AO overlap matrix). Such orthonormalisation has the side benefit of removing errors due to a non-orthonormal set of trial eigenvectors being supplied ($T^{(1)}$), although a linearly dependent set would still cause problems. The overall SCF "convergence procedure", in practice, is found to be very stable, and well suited to the Jacobi diagonalisation method, since H^F limits to diagonal form as convergence is approached.

As well as performing the SCF calculation in terms of the AO's of contracted basis functions as defined at integral evaluation time, it is possible to define a basis for the SCF stage using linear combinations of the original functions, with the restriction that the number of new functions is not greater than the original. Usually, this transformation of basis is done to create a set of "symmetry adapted" functions, so that each new function χ^{SA} constituting the MO's ϕ transforms according to one of the irreducible representations of the symmetry group to which the molecule considered belongs. Effectively, the original AO's (χ, perhaps contracted functions themselves) are formed into a new set of contracted functions (χ^{SA}) by combining AO's at different centres, and perhaps different types. It is not necessary to explicitly use the full symmetry present in a molecule in constructing the MO's ϕ. The SCF diagonalisation finds the proper configuration symmetry, even if not explicitly given as input, by combining the atom-centred χ functions into appropriate χ^{SA}'s. In practice, the use of symmetry adapted functions requires the extra step of transformation of the molecular integrals over the original AO's, but the SCF step is facilitated by "symmetry blocking" of the H^F matrix, thus reducing the dimensionality of the diagonalisation problem.

C. Open Shell Systems.

The methods and results of the LCAO MO procedure of Chapter 2 all depend on the applicability of equations limited to "closed shell" systems, in which the electronic structure is described by a single determinant of doubly occupied spatial orbitals. Mostly, in this work, species considered fall into this category; however, many important species have unpaired electrons or "open shell" electronic structure. In order to include these molecules in the computational scheme, the relevant LCAO MO open shell RHF equations have been derived. There are two possible approaches to unpaired electron systems within the LCAO MO framework, depending on the number of constraints to be imposed on the single determinant wavefunction. If the wavefunction is restricted only to be a variational solution of the Schrödinger equation and satisfy the Pauli principle (Chapter 2), the method is the "Different Orbitals for Different Spins" (DODS) procedure. If the approximate wavefunction is to represent a pure spin state, the Open Shell RHF equation is obtained.

In the DODS, or Unrestricted Hartree-Fock (UHF), method, the wavefunction is written as a single determinant of n^α orbitals, ϕ_i^α, which are occupied by electrons of α spin, and n^β orbitals, ϕ_j^β, in which electrons of spin β are placed:

$$\psi = \det \{\phi_1^\alpha \bar{\phi}_1^\beta \phi_2^\alpha \bar{\phi}_2^\beta \ldots \phi_{n^\alpha}^\alpha \bar{\phi}_{n^\beta}^\beta\} \qquad -(7)$$

where $n^\alpha \neq n^\beta$. The steps in deriving the variational equations for the ϕ_i^α and ϕ_j^β are completely analogous to those of Chapter 2 for the closed shell case. The spin orbitals of (7) are all orthogonal since the spatial orbitals ϕ_i^α are all eigenfunctions of the same effective Hamiltonian, and the same is true for the ϕ_j^β, although the spatial

orbitals do not have to be orthogonal. If a LCAO expansion for both ϕ_i^α and ϕ_i^β is assumed, the derivation of the equations satisfied by the linear coefficients is similar to that in Chapter 2; as each spatial orbital is occupied by only 1 electron, the factors of 2 disappear in the DODS energy expression, and, as $n^\alpha \neq n^\beta$, the symmetry of the expression for G(R) is destroyed, i.e. the effective potential is different for electrons of opposite spin. The LCAO MO is extended to cater for α and β quantities. Thus, $\phi^\alpha = \chi T^\alpha$ and $\phi^\beta = \chi T^\beta$, with corresponding charge and bond order matrices $R^\alpha = T^\alpha T^{\alpha\dagger}$ and $R^\beta = T^\beta T^{\beta\dagger}$. The set of coefficients, T^α, for the occupied α- spin MO's is obtained by solving the equation $H^{F\alpha}T^\alpha = ST^\alpha\varepsilon^\alpha$; T^β is obtained from $H^{F\beta}T^\beta = ST^\beta\varepsilon^\beta$. These two equations are coupled since

$$G_{ij}^\alpha = \sum_{r,s,=1}^{m} \{(R_{rs}^\alpha + R_{rs}^\beta)(ij,rs) - R_{rs}^\alpha(ir,js)\}$$

(G_{ij}^β similarly), so that $H^{F\alpha} = H + G^\alpha$ and $H^{F\beta} = H + G^\beta$ both depend on R^α and R^β. Provided $n^\alpha \neq n^\beta$, then $T^\alpha \neq T^\beta$ and the orbitals are indeed different orbitals for different spins; the electrons of opposite spins experience different mean electron repulsion through the second set of terms in G^α and G^β. In the case that $n^\alpha = n^\beta$, both equations above collapse into the closed shell RHF equation.

The computer implementation of the two coupled RHF equations involves no problems different in principle from the closed shell case. It is only necessary to carry out the iterative calculation for both sets of electrons in parallel, and require both T^α and T^β to be self-consistent. The description of molecular electronic structure given by a single configuration of these DODS orbitals has important applications in the interpretation and computation of electron spin properties. However, the use of a DODS wavefunction in describing magnetic systems gives a

conceptual picture rather different from the usual one of a set of doubly occupied spatial orbitals plus an "outer" unpaired electron (or electrons). This latter picture has its theoretical justification in the Open Shell LCAO MO method.

The object of the open shell procedure is to derive a single RHF equation, i.e. to define a single set of orbitals for the whole electronic system which are not associated with any particular electron spin, to define a single effective field for each electron, rather than two as in the DODS procedure. The MO wavefunction is

$$\psi = \det\{\phi_1\bar{\phi}_1\bar{\phi}_2\phi_2 \ldots \phi_{n^c}\bar{\phi}_{n^c}\phi_{n^c+1}\} \qquad -(8)$$

where a single unpaired electron in orbital ϕ_{n^c+1} (α spin) "outside" a closed shell of n^c doubly occupied orbitals is assumed. It is required to find an effective Hamiltonian which has eigenfunctions $\phi_1,\ldots,\phi_{n^c}$ and $\phi_{n^c+1}\ldots\phi_m$. Using the familiar LCAO MO expansion for all the orbitals, $\phi = \chi T = \chi(T_c, T_o)$, partitioning the matrix of MO coefficients, T, into closed shell columns T_c and open shell columns T_o. Corresponding density matrices for the two types of orbitals are $R_c = T_c T_c^{\dagger}$ and $R_o = T_o T_o^{\dagger}$. Paralleling the DODS case, the application of the variation principle gives the equations satisfied by T_c and T_o : $H^{FC}T_c = T_c\varepsilon_c$ and $H^{FO}T_o = T_o\varepsilon_o$ (orthogonal AO's);

$$G^c_{ij} = \sum_{r,s,=1}^{m} (R_{c_{rs}} + R_{o_{rs}}) \, [2(ij,rs) - (ir,js)] \quad \text{and}$$

$$G^o_{ij} = G^c_{ij} - \tfrac{1}{2} \sum_{r,s=1}^{m} R_{o_{rs}} (ir,js).$$

Additionally, the orbitals should be eigenvectors of the same effective Hamiltonian, parts of the same T matrix, for which $H^F T = T\varepsilon$, where H^F

is yet to be found. The matrix H^F has to combine the properties of H^{FC} and H^{FO}. Including virtual (unoccupied) orbitals, T_u, it can be shown that the following matrix H^F has eigenvectors T_c, T_o, T_u:

$$H^F = a(I-R_o)H^{FC}(I-R_o) + b(I-R_c)H^{FO}(I-R_c) + c(I-R_u)(2H^{FC}-H^{FO})(I-R_u) \quad -(9)$$

where a,b,c are arbitrary constants. This arbitrariness in the definition of H^F, which has the same eigenvectors as H^{FC} and H^{FO}, leads to a certain arbitrariness in the eigenvalues, which are not unique for the open shell RHF matrix. Since the eigenvectors are unique, an interpretation of the relation between the eigenvalues and the MO energies of photoelectron spectroscopy can be obtained by considering the "expectation values" of the eigenvectors with H^F, i.e. the quantities $T^{(i)\dagger} H^F T^{(i)}$, where $T^{(i)}$ is a column of T. The conversion of (9) into a form suitable for use with non-orthogonal orbitals gives equations which are cumbersome but quite straightforward to program. The simplicity of the orthogonal form (9) renders it useful in practice; non-orthogonal R matrices are used (together with molecular integrals over non-orthogonal AO's) to form H^{FC} and H^{FO}, which are then transformed to an orthogonal basis and orthogonal R's used to form H^F. The computer implementation scheme is a simple extension of the closed shell case (Section A), with all the additional steps being simple matrix multiplications and no new techniques involved.

It is usual to use the open shell method for the calculation of the electronic structure of unpaired electron systems, in which spin properties are not specifically sought. The DODS wavefunction does not represent a pure spin state, as it is not an eigenfunction of $\hat{S}^2$; it is therefore rather uncertain what values can be placed on computed spin properties using the DODS method. There are procedures for generating a spin eigenfunction from a DODS single-determinant wavefunction, but the

resulting function loses the property of being a variational solution of the Schrödinger equation. Any attempt to use the DODS method together with the spin eigenfunction constraint goes beyond the LCAO MO method into a MCSCF formalism.

The implementation of the open shell method is a reasonably straightforward extension of the closed shell procedure. In this work, calculations on unpaired electron systems are performed via the restricted Hartree-Fock SCF procedure for "half-closed" open shell cases, i.e. the minimisation of the energy of a single determinantal wavefunction constructed from n^D doubly occupied MO's (DOMOS) and n^O singly occupied MO's (SOMOS), the latter being of common spin factor. The basis set consists of m linearly independent real functions (possibly, the original set transformed, as in Section B), so that $n^V = m-n^O-n^D$ virtual MO's (VMOS) may be generated; the full set of m MO's are required to be orthonormal. The method used here is identical with that of Roothaan in the case where the state studied is not spatially (orbitally) degenerate. However, the situation is quite different for spatially degenerate states. The present procedure yields MO's which optimally describe only one component of the degenerate manifold, whilst Roothaan's procedures yield MO's which are used to construct the set of degenerate wavefunctions. In general, the energy of a degenerate state produced here will be lower than that given by Roothaan's procedure. Furthermore, discontinuities in the energy surface, observed with "symmetry equivalenced" procedures when Jahn-Teller distortions of molecular geometry are studied, do not appear when using "spatially unrestricted"methods. However, total wavefunctions produced by the latter may not be eigenfunctions of all the symmetry operators which commute with the total Hamiltonian.

In practice, trial MO's for the SCF procedure may be obtained in the same way as in the closed shell case. Convergence is generally more difficult to achieve in open shell cases; it is sometimes good practice to iterate some closely related closed shell system to a convergence threshold of 10^{-2} to provide a good starting point for the open shell iterations. In difficult cases, the same level shifting technique as used in the closed shell case can be employed, with appropriate extensions. With χ, ϕ^D, ϕ^O, ϕ^V denoting row vectors of AO's, DOMOS, SOMOS, VMOS respectively, then $(\phi^D \mid \phi^O \mid \phi^V) = \chi(T^D \mid T^O \mid T^V) = \chi T$; also, with closed and open shell density matrices $R^D = T^D T^{D\dagger}$, $R^O = T^O T^{O\dagger}$, the total wavefunction is invariant to mixing amongst the DOMOS and SOMOS respectively, being uniquely defined by R^D and R^O. A suitable effective Hamiltonian matrix is defined in the trial MO basis representation, and then diagonalised. The resulting eigenvectors, after suitable (usually "aufbau") ordering, define iterated MO's as linear combinations of the trial MO's. The general form of the single H^F_{MOD} is:

$$\left[\begin{array}{c|c|c} T^{D\dagger}H_{D}T^{D} & \lambda_{12}T^{D\dagger}H_{12}T^{O} & \lambda_{13}T^{D\dagger}H_{13}T^{V} \\ \hline \lambda_{12}T^{O\dagger}H_{12}T^{D} & T^{O\dagger}HST^{O}+xI & \lambda_{23}T^{O\dagger}H_{23}T^{V} \\ \hline \lambda_{13}T^{V\dagger}H_{13}T^{D} & \lambda_{23}T^{V\dagger}H_{23}T^{O} & T^{V\dagger}HVT^{V}+(x+y)I \end{array}\right]$$

introducing the first and second level shifters (x,y) and the damp factors $(\lambda_{12},\lambda_{13},\lambda_{23})$. If matrix H^F_{MOD} is diagonal, the conditions for a stationary energy (self-consistency) are satisfied, so that the total wavefunction is self-consistent. Furthermore, the DOMOS, SOMOS, and VMOS are canonical over Hamiltonians HD, HS, and HV respectively, the specification of which involves some arbitrariness. As in the closed shell case, an analysis of the iterative process by first order perturbation

theory shows that a "sufficiently" level shifted procedure guarantees convergence to a stationary point on the energy surface given any set of trial MO's, and will always proceed down the energy surface. It can be useful in practice, usually in studying excited states, to select iterated MO's on the principle of maximum overlap with the trial MO's, rather than using "aufbau" ordering, and so "lock" on to an electronic configuration different from that of the ground state. The element of arbitrariness in the definition of H^F_{MOD} is reflected in the various canonicalisation Hamiltonians which are possible for the DOMOS, SOMOS, VMOS. A set of MO's is canonical over a given Hamiltonian if that Hamiltonian is diagonal in the basis of the given MO's, with diagonal elements which define the MO "energies". The total wavefunction is invariant to the canonicalisation procedures adopted. A wide variety of such procedures have been adopted in open shell work, e.g. it is possible to choose Hamiltonians HD, HS to yield I.P. significant orbital energies. The Roothaan single H^F matrix is a particular form of H^F_{MOD} above, and so is a part of the general philosophy of the SCF minimisation procedure used here.

It is of interest to consider the ionisation of open shell systems, with respect to Koopmans-type considerations (Chapter 2, Section B). If a RHF-SCF calculation has been performed, yielding I.P. significant orbital energies for DOMOS and SOMOS, assuming that the common spin factor of the SOMOS is α, then for removal of an electron from a β spin DOMO (vertical ionisation), there is close similarity to closed shell systems. Thus, the IP of the process may be equated to the negative of the corresponding orbital energy, in the formation of an ion configuration of spin degeneracy one greater than that of the parent molecule. However, for removal of an electron from a SOMO, giving rise to an ion configuration

of spin degeneracy one lower than that of the parent molecule, the same Koopmans-type approximation involves different error considerations. The orbital relaxation energy of the ion tends to make the calculated I.P. too large, as for the closed shell case; the correlation energies of the parent molecule and the ion are similar since both have the same number of paired electrons, so that neglect of this error does not cause appreciable error, in contrast to the closed shell case where the error balances the relaxation one; it is also possible to form configurations of the same spin degeneracy by removal of an electron from a DOMO, so that interaction of such configurations tend to lower the energy of the open shell ionisation one, resulting in the calculated I.P. being too large. Thus, Koopmans' approximation for closed shell ionisation involves cancelling errors; however, the theory of open shell ionisation is accompanied by two errors both likely to make predicted I.P.'s too large, so that the resulting ion states are raised in energy relative to those from closed shell ionisation.

The implementation of the spin unrestricted Hartree-Fock (DODS) SCF procedure involves considerations basically the same as in the closed shell case; the minimisation of the energy of a single determinantal wavefunction constructed from n^{α} orthonormal α-spin MO's and n^{β} orthonormal β-spin MO's, yielding 2 sets of m orthonormal MO's, is effectively two separate closed shell problems (1 electron), and involves the diagonalisation of two H^{F} matrices. The procedure used here includes, as a special case, the method of Pople and Nesbet.

Appendix 4

Gaussian Basis Sets .

In the LCAO MO method, each molecular orbital is expanded in terms of a basis of atomic orbitals, each of which is represented by a linear combination of Gaussian type functions:

$$\phi_i = \sum_j T_{ij}\chi_j = \sum_j T_{ij}(\sum_k C_{jk}\eta_k) .$$

The expansion coefficients C form part of the input data of a molecular calculation, and are not obtained from the use of the variational principle. Thus, a bad choice of the C's can affect drastically the accuracy of the computation. In order to derive optimal ϕ's, the T coefficients are optimised by the Roothaan technique. The other method of optimising the ϕ's is by optimising the basis set of χ's, e.g. by orbital exponent optimisation; this involves "brute force" techniques, using trial and error, which is not computationally feasible in general. Thus, in the Roothaan technique, optimal ϕ's are obtained relative to a pre-determined basis set (χ). If the latter is poor, the resulting ϕ is equally poor.

The following tables present, for each atom involved in the molecular computations considered in this work, basis sets of contracted Gaussian type functions. These are in the Cartesian representation: each AO χ consists of a contraction of Gaussian primitives, η.

η is of the form $\exp(-\alpha r^2)(x^I y^J z^K)$.

In the tables the primitive functions are characterised by the orbital exponent of the radial factor, $\exp(-\alpha r^2)$; the angular factor, $(x^I y^J z^K)$, is represented by the familiar AO types (s,p,d). Primitives within a contraction are of common angular factor (and centre). The contraction coefficient of each primitive is also given.

HYDROGEN	COEFFICIENT	EXPONENT			
		I	II	III	IV
1s	0.07048	4.50037	7.37812	6.99357	8.38779
	0.40789	0.681277	1.11692	1.0587	1.26976
	0.64767	0.151374	0.24817	0.235235	0.282131

I : Unscaled, best-atom basis

II : Scaled methane basis (factor = 1.639)

III : Scaled ethylene basis (factor = 1.554)

IV : Scaled vinyl amine basis (factor = 1.864)

CARBON	COEFFICIENT	EXPONENT		
		I	II	III
1s	0.004813	1412.29		
	0.037267	206.885		
	0.172403	45.8498	DITTO	
	0.459261	12.3887		
	0.456185	3.72337		
2s	0.522342	0.524194	0.557981	0.557981
	0.594186	0.163484	0.174021	0.174021
2p	0.112194	4.18286	4.683331	4.74919
	0.466227	0.851563	0.953399	0.966859
	0.622569	0.199206	0.223019	0.226177

I : Unscaled Basis

II : Scaled methane Basis (2s Factor = 1.064, 2p Factor = 1.120)

III : Scaled ethylene Basis(2s Factor = 1.064, 2p Factor = 1.135)

NITROGEN	COEFFICIENT	EXPONENT	
		I	II
1s	0.004479	3038.41	
	0.034581	301.689	
	0.164263	66.4630	DITTO
	0.453898	17.8081	
	0.468979	5.30452	
2s	0.513598	0.764993	0.719093
	0.605721	0.234424	0.220359
2p	0.119664	5.95461	6.252341
	0.474629	1.23293	1.29458
	0.611142	0.286752	0.31009

I : Unscaled basis

II : Scaled vinyl amine basis (2s factor = 0.940, 2p factor = 1.050)

OXYGEN	COEFFICIENT	EXPONENT	
		I	II
1s	0.004391	2714.89	
	0.032764	415.725	
	0.158829	91.9805	DITTO
	0.454738	24.4515	
	0.48905	7.22296	
2s	0.477135	1.06314	0.967457
	0.583049	0.322679	0.293638
2p	0.129373	7.75579	7.57741
	0.481269	1.62336	1.58602
	0.604484	0.36503	0.356634

I : Unscaled basis

II : Scaled vinyl alcohol basis (2s factor = 0.910, 2p factor = 0.977).