was achieved for the single bands and of ±0.1 eV for the shoulders.

Moments and the Energies of Solids

Jeremy K. Burdett[*1] **and Stephen Lee**

Contribution from the Department of Chemistry, The University of Chicago, Chicago, Illinois 60637. Received May 21, 1984

Abstract: The use of the moments method applied to Hückel-based tight-binding theory is described. By locating n-rings in the middle of an m coordinate tree and calculating the energy difference of this structure compared to the simple tree itself, it is shown how the energy difference between two structures is controlled by the first few disparate moments of the density of states. $\Delta E_r(X)$ curves are presented which describe the energetic difference between two structures whose first $r - 1$ moments are identical as a function of band filling (X) ($0 \leq X \leq 1$).

One of the fundamental problems associated with understanding the structure of solid-state systems is the connection between geometrical and electronic structure. Even in approximate theories, as simple as Hückel-based tight-binding theory, it is often difficult to correlate structural features with energetic effects. Much of this difficulty results, not from any inherent intransigence in the Hückel equations but from the complexity of the method of solution itself. As an example let us say we wish to determine the site preferences in an AB crystal associated with its two component atoms A and B, i.e., given a solid with two inequivalent sites, α and β, which out of A and B will occupy these two sites as a function of the number of valence electrons. The standard method requires determination of the charge density at these crystallographically inequivalent sites.[2] To do this we must solve the Hückel equations at various points in k space and generate a set of eigenfunctions and eigenvalues. We need to use a fine enough mesh in k space (a large enough number of points) so that by reordering the eigenvalues by energy we create a reasonable approximation to the energy density of states. We then fractionate each eigenfunction as to its composition in terms of α- and β-located orbitals, and integration via a population analysis leads to the charge density for a particular electron filling. Only then may we make predictions concerning the site preferences.

There are two arbitrary artifacts of this method of calculation, namely the use of k space and the generation of a large number of eigenfunctions, which have interposed themselves between the orbital topology of the system and the final structure prediction based on charge density. The development of new methods of solving the Hückel problem which obviates artificial concepts and relates structure and energy much more directly is certainly to be desired. Such a method has been in the physics literature for many years but more recently has been developed by several authors.[3-7] In this and the following two papers in this issue, we

shall use this technique to produce some very direct relationships between geometrical structure and electronic stability.

Method of Moments

First, we define here our use of the term "Hückel" Hamiltonian (**H**). In this, and in the following two papers in this issue, we will consider systems whose energy bands, within the framework of tight-binding theory, may be obtained by solution of the secular determinant

$$|H_{ij}(k) - S_{ij}(k) \cdot E| = 0 \tag{1}$$

where $H_{ij}(k)$ are the interaction and diagonal matrix elements of **H** linking the Bloch sums

$$\mu_i = N^{-1/2} \sum_j e^{ik \cdot R_{ij}} \chi_i(r - R_{ij}) \tag{2}$$

Here the χ_i are the atomic basis orbitals contained within the unit cell. The sum is over all unit cells of the crystal j, where the orbital χ_i is located on atoms residing at R_{ij} with respect to some arbitrary origin. The Coulomb integrals $H_{ii} = \langle \chi_i | \mathcal{H} | \chi_i \rangle$ are usually estimated[8] from the relevant ionization potential and the interaction (or resonance) integrals by the Wolfsberg–Helmholz (or Mulliken) relationship

$$H_{ij} = \langle \chi_i | \mathcal{H} | \chi_j \rangle = \frac{1}{2} K S_{ij}(H_{ii} + H_{jj}) \tag{3}$$

where K is a constant and $S_{ij} = \langle \chi_i | \chi_j \rangle$, the overlap integral linking the atomic orbitals χ_i and χ_j. Our use of the descriptor "Hückel" refers to a calculation using eq 1 where the overlap integrals involving the Bloch sums are put equal to zero unless $i = j$; i.e., $S_{ij}(k) = \delta_{ij}$. The overlap integral is retained in the evaluation of the interaction elements in eq 3. The calculation is therefore an extended Hückel one but with neglect of overlap. In molecules, of course, $S_{ij}(k) = S_{ij}$.

The method of moments is based on the following two observations.[3]

(1)

$$\sum_i E_i^n = Tr(\mathbf{H})^n \tag{4}$$

where **H** is the Hückel Hamiltonian matrix, E_i is the ith eigenvalue

(1) Camille and Henry Dreyfus Teacher–Scholar.
(2) See for example: Burdett, J. K. *Adv. Chem. Phys.* **1982**, *49*, 47.
(3) Cyrot-Lackmann, F. Thèse, Orsay, 1968.
(4) Ducastelle, F.; Cyrot-Lackmann, F. *J. Phys. Chem. Solids* **1970**, *31*, 1295; **1971**, *32*, 285.
(5) Cyrot-Lackmann, F. *J. Phys. Chem. Solids* **1968**, *29*, 1235.
(6) Cyrot-Lackman, F. *Surf. Sci.* **1969**, *15*, 535.
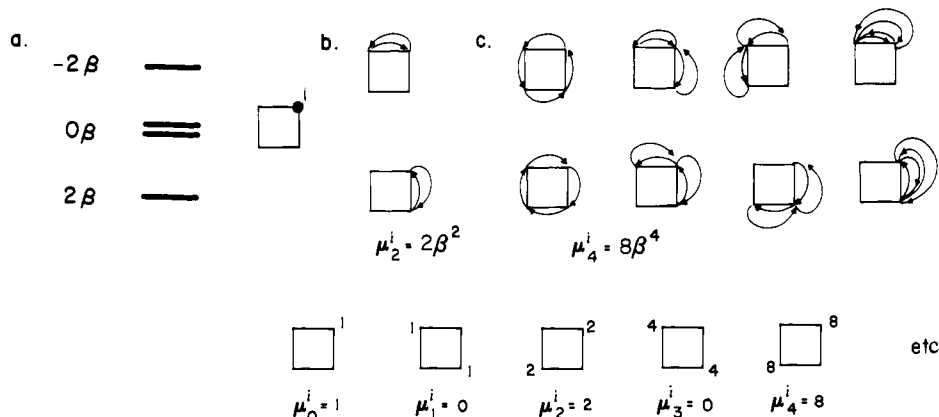(7) Gaspard, J. P.; Cyrot-Lackmann, F. *J. Phys. C* **1973**, *6*, 3077.

(8) E.g.; Burdett, J. K. "Molecular Shapes"; Wiley: New York, 1980.

a. -2β b.   c.

0β

2β

$\mu_2^i = 2\beta^2$ $\mu_4^i = 8\beta^4$

$\mu_0^i = 1$ $\mu_1^i = 0$ $\mu_2^i = 2$ $\mu_3^i = 0$ $\mu_4^i = 8$ etc

Figure 1. (a) Hückel energy levels of the pπ levels of cyclobutadiene referred to $\alpha = 0$. Notice $\mu_{odd} = 0$, $\mu_2 = 8\beta^2$, and $\mu_4 = 32\beta^4$. (b) Enumeration of the walks of length 2 and (c) of length 4. Since there are four symmetry equivalent orbitals and two walks of length 2 per orbital then $\mu_2 = 2 \times 4\beta^2 = 8\beta^2$. Similarly $\mu_4 = 8 \times 4\beta^4 = 32\beta^4$.

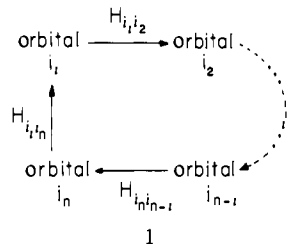of **H**, and the summation is over all eigenvalues generated by **H**.

Proof. As **H** is hermitian, there exists a unitary matrix **S** such that $S^{-1}HS$ is diagonal and whose diagonal elements are the eigenvalues of **H**. As $Tr(H)^n = Tr[(S^{-1}HS)^n]$ and $S^{-1}HS$ is diagonal with elements E_i, then $\sum_i E_i^n = Tr(H^n)$.

Anticipating later use, we note for infinite systems that $\sum_i E_i^n$ is replaced by $\int_{-\infty}^{\infty} E^n \rho(E)\, dE$ where $\rho(E)$ is the density of states (DOS) of the energy. The function $\int_{-\infty}^{\infty} E^n \rho(E)\, dE$ is generally called the *n*th moment or *n*th power moment. We shall represent it by $\mu_n(\rho)$ or simply μ_n if the function in question is clear.

(2) $Tr(H^n)$ may be interpreted geometrically. Recall that

$$Tr(H^n) = \sum_{i_1,i_2,\ldots,i_n} H_{i_1 i_2} H_{i_2 i_3} \ldots H_{i_n i_1} \qquad (5)$$

Let us examine a single term in this sum. The term which is the product of *n* elements of the **H** matrix may be viewed in a simple pictorial fashion (1).[3,9,10] This represents a path starting and

$$
\begin{array}{c}
\text{orbital} \xrightarrow{\;H_{i_1 i_2}\;} \text{orbital} \\
i_1 \qquad\qquad\qquad i_2 \\
H_{i_1 i_n} \big\uparrow \qquad\qquad\qquad \big\downarrow \\
\text{orbital} \xleftarrow[\;H_{i_n i_{n-1}}\;]{} \text{orbital} \\
i_n \qquad\qquad\qquad i_{n-1}
\end{array}
$$

1

ending with the orbital i_1. The weight associated with the path is the product of the $H_{i_a i_b}$ interaction matrix elements. In particular all paths which contribute to $Tr(H^n)$ must be paths which connect only interacting atoms (otherwise the product would contain a zero entry) and must also be closed, starting and ending at the same orbital. In a simple Hückel sense, the $H_{i_a i_b}$ will be equal to the Hückel β if the orbitals i_a and i_b are pπ orbitals located on adjacent atoms of a π-bonded network. In general these interaction matrix elements will be given by eq 3. Which values of the interaction matrix elements are set equal to zero in eq 5 allows inclusion of first nearest neighbors only or first and second nearest neighbors, etc., and in general allows the electronic model to be varied. In our initial explorations using this approach, it will be convenient to set the $H_{i_a i_a} = 0$ (i.e., Hückel $\alpha = 0$). This will remove from eq 5 all contributions from walks "in place".

Thus, the *n*th power moment is equivalent to the weighted sum over all closed paths of length *n* amongst the orbitals of the system. This powerful relationship will allow direct correlation between the DOS and geometrical structure. As an illustration of eq 5, we consider a simple molecular example in Figure 1. The pπ orbitals of cyclobutadiene have the eigenvalue spectrum shown

in Figure 1a (referred to $\alpha = 0$). Therefore, $\mu_2 = \sum E_i^2 = 8\beta^2$ and $\mu_4 = \sum E_i^4 = 32\beta^4$. In Figure 1b are drawn the closed paths of length 2 and of length 4 for a single orbital of the cyclobutadiene molecule. Since there are four such orbitals in the molecule, it is easy to see that there are eight walks of length 2 (so $\mu_2 = 8\beta^2$) and 32 walks of length 4 (so $\mu_4 = 32\beta^4$).

Construction of the Density of States

We see from **1** that there is a geometrical basis for the moments, attributed to that important function, the energy density of states. To use the "method of moments" in any computational fashion, we must first develop an efficient scheme to calculate the moments (walks) and second convert the moments into a DOS.

It turns out, at least in theory,[11,12] that one may always reconstruct the DOS from the sequence of its moments for the type of DOS which concern us. This is due to the fact that if two DOS, $\rho_1(x)$ and $\rho_2(x)$, both have the same sequence of moments, then $\rho_1(x) = \rho_2(x)$. Thus, no information has been lost in representing a DOS as a sequence of moments. This is proven in the Appendix sections I.1–I.4. In general, for a molecule or solid where not all the atoms are equivalent by symmetry, we may write the total density of states as a simple sum $\rho_{tot}(x) = \sum \rho_j(x)$ over all the orbitals *j* of the problem. $\rho_j(x)$ may be constructed from the set of moments derived by generation of the walks originating from orbital *j*. Appendix section I goes on to give the actual computational methods used in the reconstruction process. Two strategies are developed. The first assumes a trial function $\rho_{trial}(x)$ which may be corrected by the use of a power series

$$\rho(x) = \left(\sum_{i=0}^{\infty} \alpha_i x^i\right)\rho_{trial}(x) \qquad (6)$$

This approach is developed in the appendix sections I.8–I.11 and IV. A second method, employing a continued fraction is briefly described in Appendix section III. In general, the second method produces a more quickly converging result (and is the one we will use in quantitative applications). The power series approach will in our hands have the greater qualitative appeal. But, whichever method we choose, the results for crystalline materials (the only type of system we will study[13]) will be identical with those obtained in the more traditional fashion via evaluation of the energy levels associated with a mesh of *k* points. It will be in the interpretation of the results that the moments approach will demonstrate its utility.

(9) Sykes, M. F.; Fisher. M. E. *Adv. Phys.* **1960**, *9*, 315.
(10) See also methods associated with establishing isospectrality of graphs (Randic, M. *J. Comput. Chem.* **1980**, *1*, 380).

(11) (a) Vorobyev, Yu. V. "Method of Moments in Applied Mathematics". Gordon and Breach, New York, 1965. (b) Dalton, B. J.; et al. "Theory and Applications of Moments Methods in Many-Fermion Systems"; Plenum. Press: New York, 1970.
(12) Karlin, S.; Shapley, L. S. *Mem. Am. Math. Soc.* **1953**, *12*, 1.
(13) We describe the application of the method to molecules and some noncrystalline materials in: Burdett, J. K.; Lee, S.; Sha, W. C. *Croat. Chem. Acta* **1984**, *57*, 1193. And in: Burdett, J. K.; Lee, S.; Sha, W. C. *Nouv. J. Chim.*, in press.

In the remainder of this section, we consider the problem of finding a way to count the walks efficiently. In this article, we will develop a method which relies on the translational symmetry of the crystal. (Elsewhere we describe aspects of molecular problems.[13]) The approach is an extension of that of Montroll.[14] He considered the problem of a one-dimensional chain of atoms with a single orbital on each atom. Only orbitals on nearest neighbors are considered to have nonzero interaction integrals. As before, it will be convenient to set the Hückel α equal to 0 and put the Hückel β equal to -1. The problem is to calculate the number of different paths which start at a particular atom (orbital), move a certain number of steps, and finally return (in n steps altogether) to the original atom. Montroll noted that at each step, one has the choice of walking in the forward direction (an x step) or in the reverse direction (an x^{-1} step), shown in **2**.

2
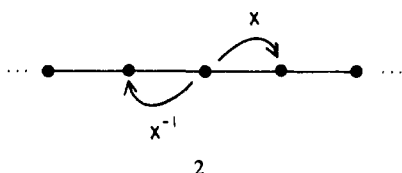
As we are interested in closed paths of length n, we need to examine sequences of x steps and x^{-1} steps such that the number of x's cancel the number of x^{-1}'s. Recalling the use of the binomial theorem in probability theory, we see that in the expansion of $(-1)^n(x + x^{-1})^n$, the coefficient of the term in x^0 ($\binom{n}{n/2}$ where n is even) represents all closed walks of length n and hence is equal to the nth power moment.

This idea may be extended to crystals with one orbital per unit cell. So for a square lattice

$$\mu_n = \text{coefficient of } x^0 y^0 \text{ term in}$$
$$(-1)^n(x + x^{-1} + y + y^{-1})^n$$

$$= \begin{cases} \binom{n}{n/2}^2 & n \text{ even} \\ 0 & n \text{ odd} \end{cases} \tag{7}$$

(Note that in both of these examples $\mu_{odd} = 0$. This is a direct result of the bipartite or alternant nature of both of the systems we have discussed. For nonalternant lattices then, μ_{odd} will not necessarily be equal to zero.) This notion may be extended further to crystals with more than one orbital per unit cell. To do this, we replace the single polynomial expressions above by matrices whose every element is such a polynomial expression. In particular let us call this matrix **M** with individual elements given by

$$M_{ij} = \sum_{n=(-\infty,-\infty,-\infty)}^{(\infty,\infty,\infty)} (-1)^{n_1+n_2+n_3} M_n(i,j) x^{n_1} y^{n_2} z^{n_3} \tag{8}$$

where $n = (n_1, n_2, n_3)$. $M_n(i,j)$ represents the interaction integral between the ith orbital of the home unit cell ($n = (0,0,0)$) and the jth orbital in the unit cell which lies at $n_1 a_1 + n_2 a_2 + n_3 a_3$ away from the home unit cell. The a_i are the direct lattice vectors. The dimension of **M** is the number of the different orbitals in the unit cell. As an example, consider the π out-of-plane orbitals in the "bathtile" net of **3**.

3
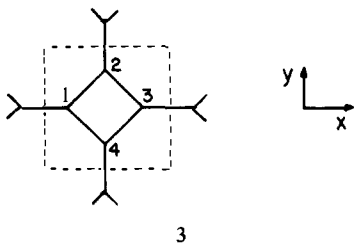
(14) (a) Montroll, E. W. In "Applied Combinatorial Mathematics"; Beckenbach, E. F., Ed.; Wiley: New York, 1950. (b) Montroll, E. W.; West, B. J. In "Fluctuation Phenomena"; Montroll, E. W., Lebowitz, J. L., Eds.; North Holland: Amsterdam, 1979.

The **M** matrix for this geometry is

$$M = \begin{pmatrix} 0 & 1 & x & 1 \\ 1 & 0 & 1 & y^{-1} \\ x^{-1} & 1 & 0 & 1 \\ 1 & y & 1 & 0 \end{pmatrix} \tag{9}$$

It serves the same purpose as the earlier one-orbital polynomial expression. In particular the constant coefficient of $(M^n)_{ii} = \mu_n$ for the ith orbital since

$$(M^n)_{ii} = \sum_{i_2 \ldots i_n} M_{ii_2} M_{i_2 i_3} \ldots M_{i_n i} \tag{10}$$

and the expression becomes identical with the one of eq 5.

The only difference between this and the earlier scheme is that we have reduced the number of orbitals from an infinite set to the collection of orbitals contained within a single unit cell, by making use of the fact that all orbitals i in every unit cell have the same connectivity. Therefore, using Montroll's method,[14] all we need keep track of is which unit cell has been reached by any given path. In the traditional Hückel calculation, an exactly analogous reduction may be made by Bloch factorization. From eq 1 and 2, the **H** matrix may be constructed. By use of a suitable similarity transform, it may be converted into one of the form of eq 11 where θ and φ contain the components of the wavevectors

$$H = \begin{pmatrix} 0 & 1 & e^{i\theta} & 1 \\ 1 & 0 & 1 & e^{-i\varphi} \\ e^{-i\theta} & 1 & 0 & 1 \\ 1 & e^{-i\varphi} & 1 & 0 \end{pmatrix} \tag{11}$$

k_x and k_y. Clearly by putting $e^{i\theta} = x$ and $e^{i\varphi} = y$, eq 9 and 11 become identical, a result expected of course from eq 5.

Now we are in a position to point out the connections between the traditional Hückel calculation and the moments approach.

(1) In the traditional method, the **H** matrix is diagonalized. As diagonalization is k-space-dependent, one must diagonalize at many points in k space and sum the results. In the moments method, H^n is evaluated and the constant coefficient of the trace recorded. (As this is independent of k, of course, multiplication is only performed once for each μ_n.)

(2) In the traditional method, eigenfunctions are calculated first. In the moments method, charge densities are computed and never eigenfunctions.

(3) The traditional tradiational method relies crucially on the simplifications provided by the translational symmetry in dealing with infinite solids. The moments method may use such symmetry if convenient, but, in general, the method may be applied equally well to any finite or infinite noncrystalline system.

(4) In the traditional method a sufficient number of k points need to be employed in the calculation to get a good approximation to the DOS. In the moments method, a sufficient number of moments need to be computed to arrive at a good approximation to the DOS.

As an example of how the moments method works, we consider the $p\pi$ out-of-plane network of a graphite sheet (**4**) with $p\pi-p\pi$ interaction integrals of β between nearest neighbors only. Figure

4
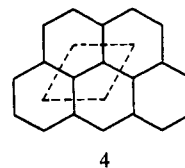
2 shows how, from a general starting point (Figure 2a shows a flattened semicircle whose center and width are determined by properties of the continued fraction[15]), an increasingly accurate DOS is generated with the addition of more moments. Nevertheless, as this example clearly shows, convergence to the final DOS can be slow. This is in keeping with the notion that it is

(15) Cyrot-Lackmann, F. *J. Phys. C* **1972**, 5, 300.

Figure 2. Continued fraction expansion for the π DOS of graphite using only 4, 6, 8, and 12 moments compared to the actual graphite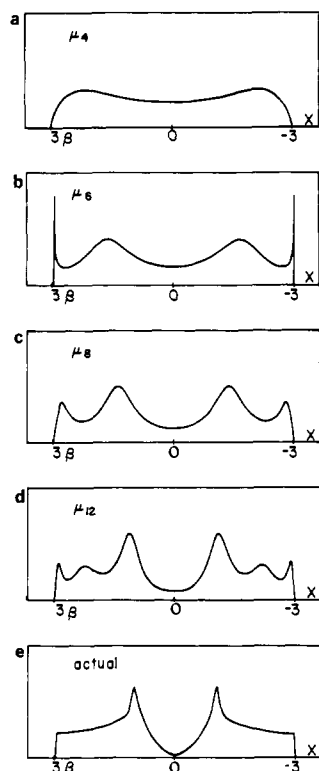 DOS. Only first nearest-neighbor interactions are considered. The correct DOS range of $\pm 3\beta$ is assumed throughout.

the higher moments which control the fine structure of the DOS. It is this feature of the method that has resulted in the moments approach being regarded unfavorably over the years by the physics community, with its general interest in the fine details of the density of states plot. Computationally it is probably easier to get a more accurate $\rho(E)$ by using the k-space approach than by inverting the moments. However, structural chemists are more interested in energies rather than in such details. In striking contrast to the above result, we find that the energy of stabilization, $\int_{-\infty}^{\epsilon_F} E\rho(E)\,dE$, as a function of electron count (and hence Fermi energy ϵ_F) exhibits the rapid convergence shown in Figure 3. The maximum error in the stabilization curve using the moments through μ_4 is 0.022β, through μ_8 is 0.0067β, and through μ_{12} 0.0013β. (The total graphite stabilization energy at the half-filled point is 0.786β).[16] This rapid energetic convergence is extremely important and will allow us to develop qualitative ideas of structure–energy relationships based upon the early moments of the energy density of states. Before proceeding, it is important to realize that with a very large number of k points via the traditional route employing Bloch orbitals or with a very large number of moments via the present route, identical densities of states will be generated. In other words, the moments approach is just another implementation of tight-binding theory. It could also be considered a cluster model once it is understood that the cluster must be large enough to contain all the walks of a given length present in the solid. In other words, the cluster is one which is perfectly embedded in the lattice. We show elsewhere[13a] by using the ideas of moments how the resulting density of states is more accurately computed by using a judiciously chosen set of special k points than by use of a periodic cluster or the "fragment within the solid" approach.[2] Here the ends of the cluster are wrapped around the edges and joined to the atoms on the other side.

---

(16) We assume prior knowledge that the π out-of-plane band extends from $+3\beta$ to -3β. These limits may be used to establish the a_{final} and b_{final} discussed in Appendix section III. Using the relations[15] $b_{\text{final}} + 2(a_{\text{final}})^{1/2} = +3\beta$ and $b_{\text{final}} - 2(a_{\text{final}})^{1/2} = -3\beta$, we find $b_{\text{final}} = 0$ and $a_{\text{final}} = {}^9/_4\beta^2$. Finally it should be noted that the band limits are $\pm 3\beta$ as the graphite system is 3-coordinate and bipartite.
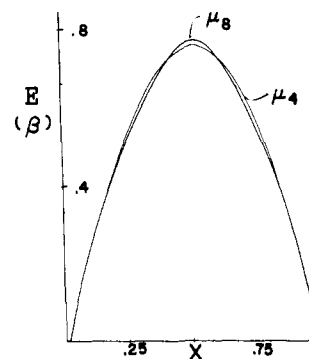

Figure 3. EBF of the fourth and eight moment curves of Figure 2 compared to the actual graphite π DOS of Figure 2. With the scale used in this graph, it is impossible to distinguish the μ_8 curve from the actual graphite curve.

Stabilization Energy

In this section we will examine properties of the stabilization energy and related functions. In doing this, we shall assume for convenience that the DOS lies purely within the interval -1 to 1, that is average value is 0, and that its area is equal to 2. The background to this choice is given in the Appendix sections I.1–I.2. We define

sym- bol	definition	constraints	name
$\varphi(x)$	$\varphi(x) = \int_{-1}^{x}\rho(t)\,dt$	$\varphi(-1) = -1$ $\varphi(1) = 1$	weighting function (WF)
$\epsilon(x)$	$\epsilon(x) = -\int_{-1}^{x}t\rho(t)\,dt$	$\epsilon(-1) = 0$ $\epsilon(1) = 0$ never negative	energy of stabilization as a function of Fermi energy (EFE)
$E(x)$	$E(x) =$ $-\int_{-1}^{\varphi^{-1}(t)}t\rho(t)\,dt$	$E(-1) = 0$ $E(1) = 0$ never negative	energy of stabilization as a function of band filling (EBF)

In the definition of $E(x)$, $x = -1$ corresponds to the empty band and $x = 1$ the full band. If we wish to compute the energies of two systems, we may either use $\Delta\epsilon(x)$, where the two systems will have the same Fermi energy but (invariably) a different number of electrons per formula unit, or $\Delta E(x)$, where the Fermi energies are different but the number of electrons per formula unit are the same. Of considerably greater chemical interest is the use of $\Delta E(x)$.

We may solve for these functions by calculating the DOS from the DOS sequence of moments and then converting the DOS itself into the desired new functions. More interestingly, we may instead convert the DOS sequence of moments into WF, EFE, or EBF sequences of moments and then reconstruct the appropriate function from such a sequence.

As $\mu_n(\varphi)$, $\mu_n(E)$, and $\mu_n(\epsilon)$ are all simply related to $\mu_n(\rho)$, this is quite practicable. Integration by parts leads to

$$\left.\begin{array}{l} \mu_n(\varphi) = \dfrac{1}{n+1}\,[2 - \mu_{n+1}(\rho)] \quad n \text{ odd} \\[2mm] \mu_n(\varphi) = -\dfrac{1}{n+1}\,[\mu_{n+1}(\rho)] \quad n \text{ even} \end{array}\right\} \tag{12}$$

$$\mu_n(\epsilon) = \frac{1}{n+1}\mu_{n+2}(\rho) \tag{13}$$

$$\mu_n(E) = \int_{-1}^{1}\frac{1}{n+1}[\varphi^{-1}(x)]^{n+1}x\rho(x)\,dx \tag{14}$$

Importantly we see that the nth moment of the DOS is proportional to the $(n-1)$th moment of $\varphi(x)$ and the $(n-2)$th moment of $\epsilon(x)$. The problem may be recast in a simpler form once we recall that we are seldom interested in the absolute magnitudes of φ, ϵ, and E but in values relative to some reference structure $\Delta\rho$, $\Delta\epsilon$, and ΔE where $\Delta\rho = \rho_1 - \rho_2$, $\Delta\varphi = \varphi_1 - \varphi_2$, etc., for two structures, 1 and 2. Expressing the above formulas as difference functions, we may write

$$\mu_n(\Delta\varphi) = \frac{1}{n+1}\mu_{n+1}(\Delta\rho) \qquad (15)$$

$$\mu_n(\Delta\epsilon) = \frac{1}{n+1}\mu_{n+2}(\Delta\rho) \qquad (16)$$

So if two DOS differ by a certain amount in their nth moment, then they differ by a scaled fraction of this in the $(n-1)$th moment of the WF difference function and by an analogously scaled fraction in the $(n-2)$th moment of the EFE difference function.

Let us therefore consider the following model problem. Let two systems have DOS, WF, EBF, and EBE of ρ_1, φ_1, E_1, and ϵ_1 and ρ_2, φ_2, E_2, and ϵ_2 respectively, and let

$$\mu_i(\rho_1) = \mu_i(\rho_2) \quad \text{for } i = 0, 1, ..., r-1 \qquad (17)$$

and define $\Delta\rho$, $\Delta\varphi$, etc., as above. We wish first to develop an intuitive feel for how $\Delta\varphi$, ΔE, and $\Delta\epsilon$ behave under such circumstances. Initially, consider the case for which $r = 3$ and assume that $\mu_3(\Delta\rho)$ is large. From our considerations above, we know $\mu_2(\Delta\varphi)$ is large, $\mu_i(\Delta\varphi) = 0$ ($i = 0$ and 1), and $\mu_1(\Delta\epsilon)$ is large, while $\mu_0(\Delta\epsilon) = 0$. Hence, we know that ϵ_1 and ϵ_2 are two curves


5

6
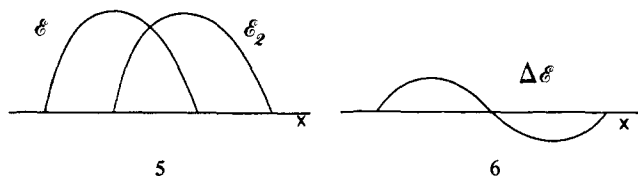
(always positive) which enclose the same area but have different means. This is shown schematically in **5**. $\Delta\epsilon$ may easily be seen to be of the form in **6**.

For $\Delta\varphi$, we know that $\mu_0(\Delta\varphi) = \mu_1(\Delta\varphi) = 0$ while $\mu_2(\Delta\varphi)$ is large. Hence we expect φ_1 and φ_2 to be centered around the same point but that φ_1 is "wider" than φ_2, a result shown schematically in **7**. $\Delta\varphi$ then has the functional form of **8**. If we further assume


7
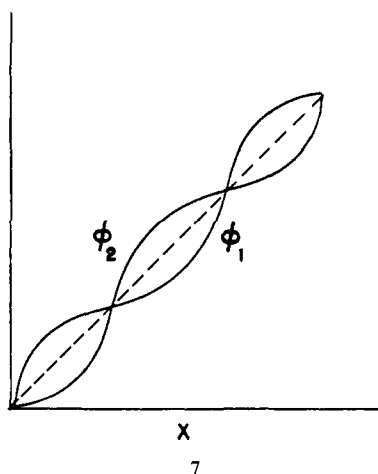
that each of the above two nodes occurs in different lobes of the $\Delta\epsilon$ curve, we may make a prediction about the shape of the ΔE curve. (This has an intuitive appeal to it, for why else would $\Delta\epsilon$


8
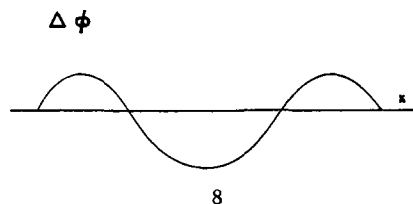
change sign if not for a node in $\Delta\varphi$?) To do this we note that whenever x is such that $\Delta\varphi(x) = 0$ then $\Delta E(x) = \Delta\epsilon(x)$. The ΔE curve will than have the same number of nodes as $\Delta\epsilon$ and will take the form shown in **9**, a curve similar to that for $\Delta\epsilon$.

In a similar intuitive spirit we may understand the curves which will arise for $r = 4$. Now $\mu_0(\Delta\epsilon) = \mu_1(\Delta\epsilon) = 0$ while $\mu_2(\Delta\epsilon)$ is large. Differences in the second moment are generally associated

9
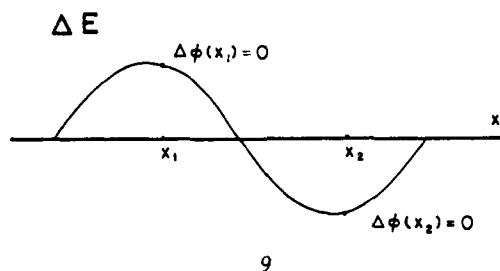
with the "width" of a function. Intuitively we expect ϵ_1 and ϵ_2 to have the shapes shown in **10**, which leads to a $\Delta\epsilon$ curve also shown in **10**. Using the same reasoning as above this will also be the approximate shape of the ΔE curve.

10
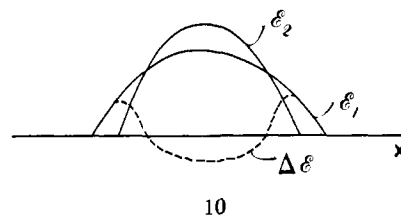
For the general case where $\mu_i(\rho_1) = \mu_i(\rho_2)$ for $i = 0,1,...,r-1$, we will state several general results given in the Appendix sections I and II where this problem is developed in a mathematical way.

(1) $\Delta\varphi$ has at least r lobes; $\Delta\epsilon$ has at least $r - 1$ lobes.

(2) To find out if there are exactly r and $r - 1$ lobes for $\Delta\varphi$ and $\Delta\epsilon$, respectively, we expand these two functions in terms of the Q polynomials given in Appendix sections I and II. When the first-order term dominates the Q expansion, there are exactly r and $r - 1$ lobes for $\Delta\varphi$ and $\Delta\epsilon$, respectively. (We exclude here for simplicity's sake the extra node described in Appendix section II.)

(3) When the first-order term in Q dominates, we also know the relative phase of each lobe. We divide this into two cases. Case 1. r is even. Here the lobe at the most bonding end of the spectrum is positive if $\mu_r(\Delta\rho)$ is positive. The signs of all the other lobes follow as a consequence of this. Case 2. r is odd. In this, the lobe at the most bonding end of the spectrum is positive if $\mu_r(\Delta\rho)$ is negative.

(4) In the case where the first-order term in Q dominates, we may also reconstruct the shape of the ΔE curve has the same number of nodes as the $\Delta\epsilon$ curve as well as the same relative phase for each lobe.

In the next section we show the general shape of such curves for $r = 2$–6.

Structure and Energetics

We now return to our initial question. How does the structure of a system influence the moments of the energy DOS and hence the DOS itself? We know that among the most important structural parameters are the atomic electronegativities (Coulombic values, H_{ii}, of the atomic orbitals), the coordination environment, and the strength of the bonds between atoms (related to the interaction integrals between the sets of atomic orbitals on the two atoms). Interestingly, it is these factors and these alone which determine the first and second moments of the energy density of states. By analogy, we expect the next most important structural features to manifest themselves in the third moment. There are two important considerations: (1) the electronegativity of atoms adjacent to any given atom (this will involve a walk "in place" at this atom before returning to the home atom) and (2) the number of atoms in the first coordination polyhedra which are bound to each other (this will involve a walk from one neighbor to another before returning to the home atom). In this section we will consider factors of the second sort. In effect we will study the more general problem of the effect of closed paths (paths in the sense of walking along bonds) of length n which involve n different atoms. We shall call such closed paths n-rings. Benzene for instance has one 6-ring.

Moments and the Energies of Solids

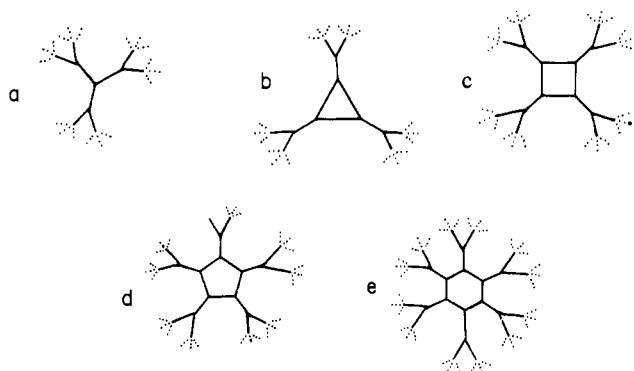*J. Am. Chem. Soc., Vol. 107, No. 11, 1985* 3055

Figure 4. (a) 3-tree; (b)–e n-ring in a 3-tree (n = 3–6). We compare in Figure 6 the $\Delta\varphi(x)$ and $\Delta E(x)$ curves for orbitals located at thee node indicated in (b)–(e) compared to that in (a).

To determine the effect of a given n-ring, we need to study systems where not only are extraneous effects such as electronegativity and coordination number eliminated but where the competing effects of other rings are removed as well. Figure 4 illustrates a sample set of such systems. Each net is a three-coordinate one; their only difference lies in the presence of a single ring in Figure 4b–e. Figure 4a, without any closed rings, is a 3-tree. We will take the obvious course and populate each node in these nets with a single orbital and will give all linkages equal weight (equal values of the interaction integral β). It is easy to see that Figure 4b–e shares, the same first two, three, four, and five moments, respectively, with Figure 4a, the 3-tree. They correspond therefore to increasingly smaller perturbations of the 3-tree itself. What is not yet clear is what sort of perturbations they are. A 3-ring for instance not only changes μ_3 but μ_5, μ_6, and higher moments as well. Some extra paths are shown in **11** and **12**. Hence, the effect of a single ring on $\varphi(x)$ and $\epsilon(x)$ will

length
5

length
6

11

12
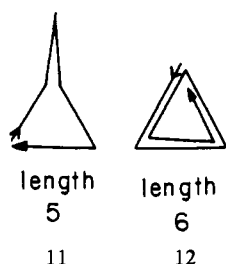
be a linear combination of several Q polynomials. Let us weigh these various effects.

A little counting shows that a single n-ring in an m-tree creates the following additional walks.

a

$$\mu_n(\Delta\rho) = 2(-1)^n$$

$$\mu_{n+1}(\Delta\rho) = 0$$

$$\mu_{n+2}(\Delta\rho) = 2[mn + m - n](-1)^{n+2}$$

$$\mu_{n+3}(\Delta\rho) = 0, \; n \neq 3$$

$$\mu_{n+3}(\Delta\rho) = 2, \; n = 3 \qquad (18)$$

b

$$\mu_n(\Delta\rho) = \left(\frac{1}{2(m-1)^{1/2}}\right)^n 2$$

$$\mu_{n+1}(\Delta\rho) = 0$$

$$\mu_{n+2}(\Delta\rho) = \left(\frac{-1}{2(m-1)^{1/2}}\right)^{n+2} 2(mn + m - n)$$

$$\mu_{n+3}(\Delta\rho) = 0, \; n \neq 3$$

$$\mu_{n+3}(\Delta\rho) = \left(\frac{-1}{2(m-1)^{1/2}}\right)^{n+3} 2, \; n = 3 \qquad (19)$$

Table I. Q Polynomial Expansion for n-Rings

n	β_{n-1}	β_n	β_{n+1}	β_{n+2}
	Where $\Delta\varphi = \sum\beta_i Q_i$			
3	2.94×10^{-2}	0	2.46×10^{-4}	-6.51×10^{-4}
4	-7.81×10^{-3}	0	-5.92×10^{-4}	0
5	2.21×10^{-3}	0	1.52×10^{-5}	0
6	-6.51×10^{-4}	0	4.07×10^{-6}	0

n	α_{n-2}	α_{n-1}	α_n	α_{n+1}
	Where $\Delta\epsilon = \sum\alpha_i Q_i$			
3	-4.42×10^{-2}	0	-5.92×10^{-4}	7.81×10^{-4}
4	1.04×10^{-2}	0	1.65×10^{-3}	0
5	-2.76×10^{-3}	0	-4.81×10^{-4}	0
6	7.81×10^{-4}	0	1.45×10^{-4}	0

In a, the Hückel β has been set equal to -1, and in b, the DOS has been confined to the interval -1 to 1.[17]

Solving for the coefficients of the appropriate Q polynomials (and weighting them by the magnitude of the Q polynomial itself) shows that for $\Delta\varphi(x)$, all terms in the Q expansion beyond the leading term are an order of magnitude smaller (Table I). Similarly, in the case of $\Delta\epsilon(x)$, only the first two terms in the Q expansion need be considered. The only exception to this is the case of the 3-ring itself which converges at a slightly slower rate. One interesting result from this calculation is that we are now able to associate a physical meaning to the Q polynomials. The Q polynomial of index s is a good approximation to the change in $\varphi(x)$ created by the presence of an $(s + 1)$-ring. We have also calculated $\Delta\varphi$ and ΔE by using the second-order approximation in Q. Figure 5 compares these results with the actual difference functions. The agreement can be seen to be quite good. Thus, the effect of the n-ring may be treated as a linear combination of the Q polynomials associated with $\mu_n(\Delta\rho)$ and $\mu_{n+2}(\Delta\rho)$. Furthermore, the $\mu_n(\Delta\rho)$ term is the dominating one. In Figure 5 we have defined a new index X which measures the fractional occupancy of the band (empty; $0 \leq X \leq 1$; full).

We wish to point out some interesting features associated with the $\Delta E_n(X)$ curves for various small m-rings displayed in Figure 5.

(a) At the left-hand side of these plots, near the empty band, all rings are stabilizing relative to the tree ($\Delta E > 0$). All rings generate a larger charge density than the tree at low fillings ($\Delta\varphi > 0$). This is a consequence of the fact that the leading term in the Q expression (via $\mu_n(\Delta\varphi)$) is $(2)(-1)^n(1/2(m-1)^{1/2})^n$ while Q itself alternates in sign.

(b) At the right-hand side of these curves, near the full band limit, all even rings are stabilizing and are associated with greater charge density, while all odd rings are destabilizing and are associated with reduced charge density. This again is a consequence of the form of the leading term of the expansion (as described in (a)) and that Q does not alternate in sign at the right-hand side of the diagram.

(c) At the half-filled band, rings of length $4n + 2$ are stabilizing, and those of length $4n$ are destabilizing, while those of length of $2n + 1$ have small effects. Similarly the charge density at the half-filled band is greatest for systems with rings of length $4n + 1$, smallest for systems with rings of length $4n - 1$, and unperturbed by rings of length $2n$. This again is a consequence of the leading term in Q and of the sign alternation in Q itself. In particular this means that there are more antibonding states than bonding states in a system with $(4n - 1)$-rings, more bonding states than antibonding states for systems with $(4n + 1)$-rings, and an equal number for systems with $2n$-rings.

---

(17) In order to transform our results, as we have done here, we need to calculate the band limits of an m-tree. Applying the continued fraction technique outlined in Appendix section III, it may be shown that for an unperturbed m-tree (whose Hückel β has been set to -1) $a_0 = 1$, $a_1 = m$, $a_j = m - 1$ where $j \geq 2$. This corresponds to band limits of $\pm 2(m-1)^{1/2}$. Finally it should be noted that the band limits are independent of the n-rings. This is in contrast to the system shown in Figure 8.
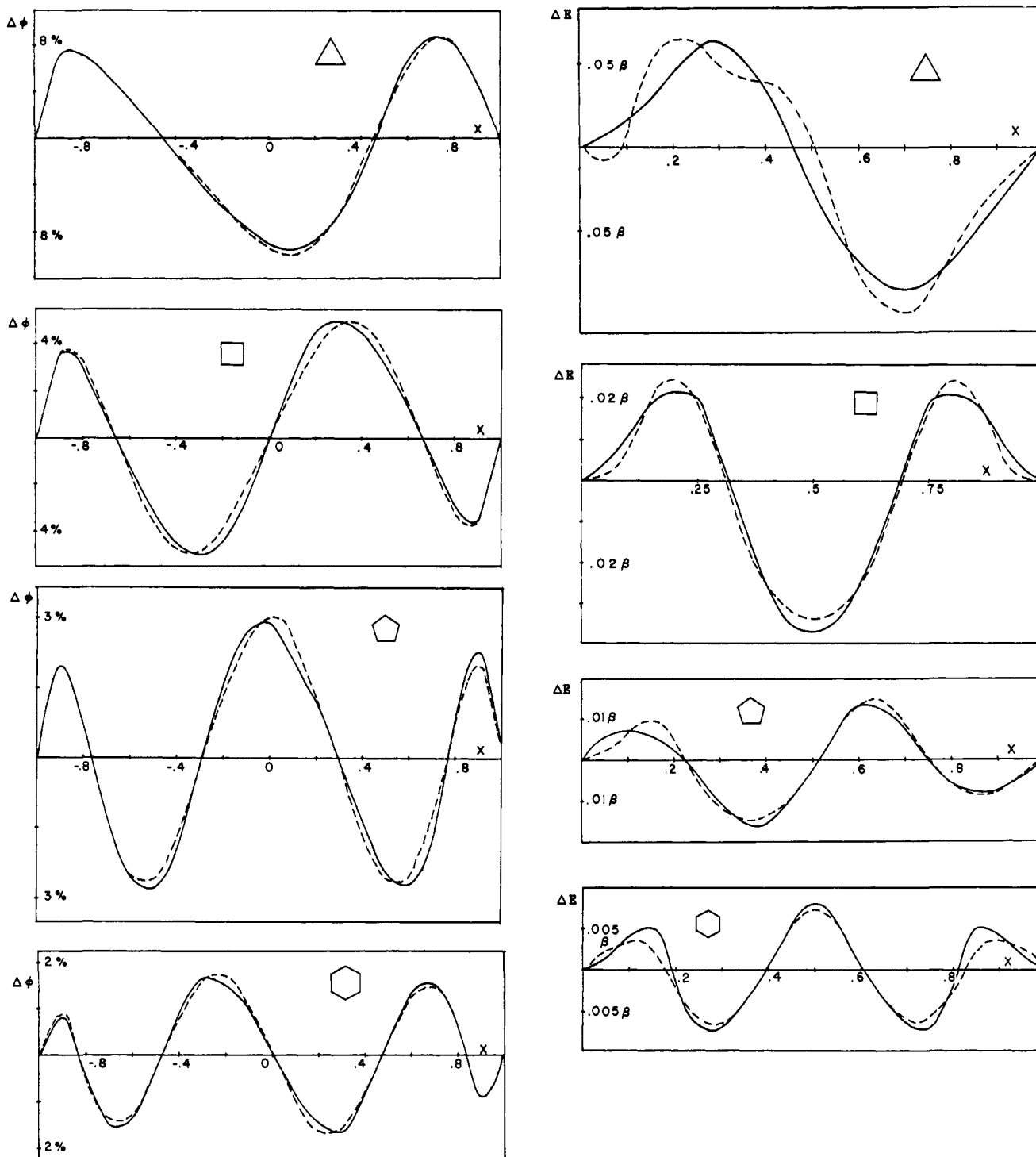
Figure 5. Actual $\Delta E(x)$ and $\Delta\varphi(x)$ curves, computed by using the continued fraction method, compared with estimates using the second-order Q polynomial approximation of Appendix section I. In all figures the dashed line represents the Q polynomial approximation and the solid line the results of the continued fraction method. The curves are all drawn with the understanding that when the ΔE and $\Delta\varphi$ curves lie in the upper plane of the graph, it is the n-rings which have the greater $\varphi(x)$ and $E(x)$ values. The ordinate in the $\Delta\varphi$ graphs are reported in terms of percentages of the area of a full band. Thus, if $\Delta\varphi(t) = x\%$, then $\int_{-\infty}^{t} \Delta\rho = x/100$, assuming $\mu_0(\rho_1) = \mu_0(\rho_2) = 1$.

(d) These comments are all consequences of the even functional nature of $\Delta E_n(x)$ and odd functional nature of $\Delta\varphi$ for $2n$-ring systems and the essentially odd nature of $\Delta E_n(x)$ and essentially even nature of $\Delta\varphi$ for $(2n + 1)$-ring systems. This is due to the fact that in the Q expansion for an n-ring, it is $\mu_n(\Delta\rho)$ and $\mu_{n+2}(\Delta\rho)$ which are important (the $\mu_{n+1}(\Delta\rho)$ and $\mu_{n+3}(\Delta\rho)$ terms are generally zero) and that these two are either both even or both odd. We note that the result for $2n$-rings may be found in their "alternant" nature by using traditional Hückel language.

(e) The number of crossings in ΔE and $\Delta\varphi$ for an n-ring system

is $n - 2$ and $n - 1$, respectively.[18]   This was discussed in the preceding section.

(f) The amplitudes of $\Delta\varphi$ and ΔE grow ever smaller as the ring size grows larger. This may be seen in the change in the coefficient of the leading term in the Q expansion $(1/2(m - 1)^{1/2})^n$. For n = 2 or larger (the only systems where rings occur), this factor

(18) The determination of the number of crossings has also interested earlier workers. See ref 4 and: Heine, V.; Samson, J. H. *J. Phys. F* **1980**, *10*, 2609.
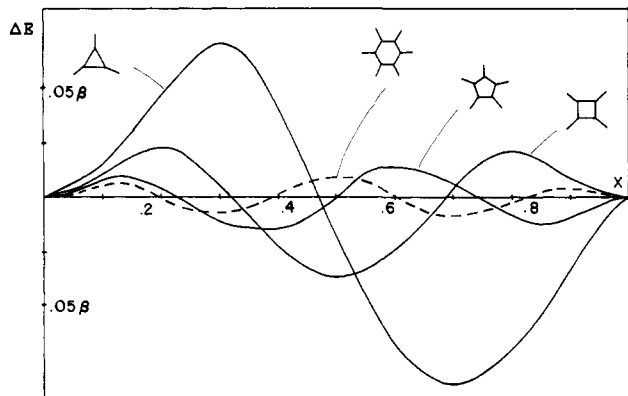
Figure 6. $\Delta E_n(x)$ curves for n-rings located in a 3-tree obtained by the continued fraction method. The curves are drawn with the same convention as Figure 5. One of the lines is dotted for clarity.
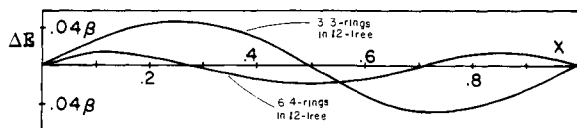
Figure 7. Retention of functional form for n-rings on a 12-tree (curves obtained by the continued fraction method.) As the relative effects of a single ring are small, we have embedded several vertex linked polygons. **16** illustrates the 3-rings. The six 4-rings are made in a similar fashion.

Table II. Ratio $\mu_n(\Delta\rho)/\mu_{n+2}(\Delta\rho)$ for an n-Ring in an m-Tree

	m				
n	3	4	6	8	12
3	1.125	1.083	1.050	1.036	1.023
4	1.375	1.333	1.300	1.2857	1.2727
5	1.625	1.583	1.550	1.536	1.523
6	1.875	1.833	1.800	1.786	1.773

overwhelms the approximate increase in the Q polynomials themselves on going from n to $n + 1$. This will be an important result since we will be able to concentrate on the energetic effects of the early disparate moments between systems.

(g) Collecting the $\Delta E_n(X)$ curves together for $n = 3$–6 as in Figure 6 leads to the following predictions for the approximate regions stability of m-membered rings: 3-ring most stable for 0%–46% band filling; 6-ring most stable for 46%–54% band filling; 5-ring most stable for 54%–70% band filling; and 4-ring most stable for 70%–100% band filling.

These results would be uninteresting were they to be applicable only to single rings lying on 3-coordinate trees. Let us loosen the constraints. First let us consider the interaction of a ring in different coordination environments. In keeping with the above discussion, let us place our $n$-ring in an m-tree where $m \neq 3$. We divide the changes this causes into two parts. The first is a change in the functional shape of $\Delta\varphi$ and ΔE. The second is a change in the amplitude of $\Delta\varphi$ and ΔE. We note that there would be little change in functional shape if the ratio of $\mu_n(\Delta\rho)$ and $\mu_{n+2}(\Delta\rho)$ stays constant since it is this ratio which determines the relative weight of the leading terms in Q. In general, we can see from eq 19 that the ratio is

$$\frac{mn + m - n}{4(m - 1)} = \frac{n}{4} + \frac{m}{4(m - 1)} \qquad (20)$$

This is tabulated in Table II for various m and n. The ratio is essentially constant. Since it is the n-ring size which determines which Q polynomials are included, there should be little change in functional shape with coordination number.

On the other hand, we should see a change in amplitude. To see this, we need only to consider the leading term in Q, namely $2(-1/2(m - 1)^{1/2})^m$. Thus, as m increases, the amplitude decreases. Figure 7 shows for various m-trees this change in amplitude, and the lack of change in functional shape, for various rings. Changes
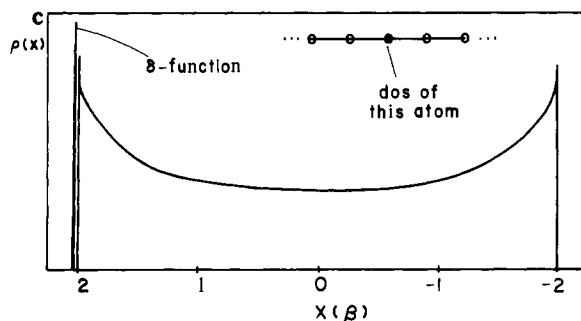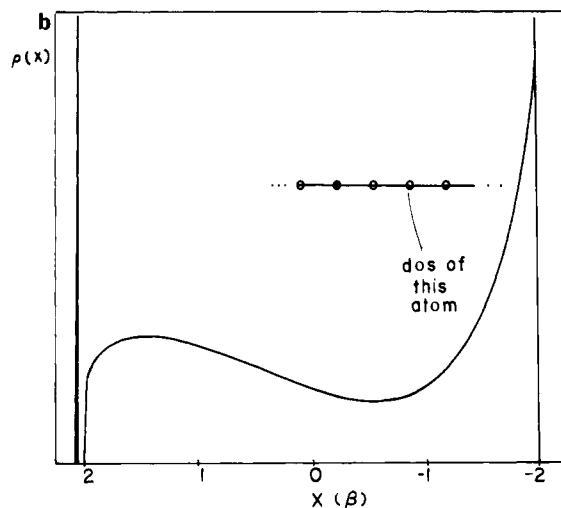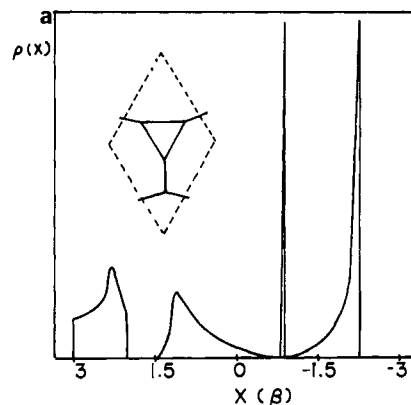
Figure 8. (a) Density of states for the net containing 3-rings and 9-rings. The unit cell of the net is given above (provided by T. Hughbanks). (b) and (c) Density of states for a linear chain of A atoms containing a single B atom (provided by W. C. Sha). We show here the DOS of two of the atoms. Note also the infinities which occur at the band edges. These are the well-known singularities which occur in one-dimensional crystals. In Hückel theory, such singularities can also occur in two dimensions. (All interaction integrals have been put equal to β.)

in amplitude do not alter the qualitative trends we have noted above—we may assume that the shape of the $\Delta\varphi$ and ΔE curves induced by the presence of an n-ring is approximately invariant to the coordination environment in which the m-ring lies.

A curious parallel exists between nonuniform coordination numbers and rings. For instance, the two orbitals indicated in **13** and **14** have identical moments to all orders.[19] Hence, the

13   14

presence of higher coordination for the first nearest neighbors in

---

(19) This was pointed out to us by W. C. Sha.

Figure 9. Nesting of the nodes of Q_{n+1} and Q_n. Shown here is an example for $n = 4$. N_i refer to the nodes of Q_4 and M_i refer to the nodes of Q_5. The shaded regions are areas where nodes are forbidden to occur.

14 produces a situation equivalent to a 4-ring.

Finally we note there is a simple linearity inherent in our ring analysis. Let us, for instance, consider two four-membered rings which share a common vertex as in **15**. The contribution to $\Delta\varphi$

2 4-rings	3 3-rings plus 3 6-rings	1 6-ring plus 1 5-ring
15	16	17

and ΔE made by the atom lying on both four-membered rings is identical with thaat of the $\Delta\varphi$ and ΔE functions for two separate four-membered rings added together, up through the eighth moment. In fact, only two vertex-sharing three-membered rings interacts via a moment less than or equal to the sixth moment. Thus, to a good approximation, we may consider the contributions of the rings in a system to be independent of each other. In **15-17** we show some examples of the rings which should be counted in some sample systems. Notice in **17** that the "rings" involve walks along the same linkage twice. Algebraically we may write this independence as the following expression for the energy difference between two systems

$$\Delta E(X) \simeq \sum_{\text{system 1}} N_n \Delta E_n(X) - \sum_{\text{system 2}} N_n \Delta E_n(X) \quad (21)$$

where N_n is the number of n-rings and the $\Delta E_n(X)$ are the curves of Figure 6 suitably modified to include their amplitude dependence on coordination number. For systems where the coordination number is the same, then

$$\Delta E(X) \simeq \sum \Delta N_n \Delta E_n(x) \quad (22)$$

The dominant term in this series will be the first for which ΔN_n is nonzero. The next few terms will be generally smaller in magnitude and less important and increasingly so as n increases beyond this value.

Appendix

In these four appendix sections, we describe in a simple fashion the mathematical basis behind the inversion of the moments to give the energy density of states.[20-23]

I. (1) The domain of any Hückel Hamiltonian DOS is bounded. Hence, all moments exist. The formula

$$\mu_j(\rho) = \int_{-\infty}^{\infty} x^j \rho(x) \, dx \quad (23)$$

may therefore be replaced by

$$\mu_j(\rho) = \int_l^L x^j \rho(x) \, dx \quad (24)$$

where l and L are the appropriate lower and upper bounds.

(2) In general there are three sorts of normalization:

(a) The area of the DOS ($\mu_0(\rho)$) can be set to any value.

(b) The Hückel $\alpha(\mu_1(\rho))$ may be shifted by any amount as energy translation is (or ought to be) invariant.

(c) All interaction integrals between pairs of atoms may be multiplied by a fixed constant. Such a transformation does not change the shape of the DOS. Often it is convenient to set $\max(l,L) = 1$. Other times it is convenient to set $\mu_2(\rho) = 1$.

(3) With Hückel Hamiltonians, one finds not only purely discrete and purely continuous DOS but mixtures of the two as well. Figure 8 illustrates two such systems. In Figure 8a one entire energy eigenvalue sheet of the Brillouin zone is entirely flat and hence degenerate. In Figure 8b and c, one eigenvalue is separate from all the others present in the system (such an eigenvalue which cannot occur in crystals is called a terminal state).

(4) The ordinary formulas for moments, i.e., $\sum_i x_i^j$ and $\int_l^L x^j \rho(x)$ dx, are therefore not always applicable to Hückel Hamiltonians. This is a familiar problem in quantum mechanics. Usually the problem is "resolved" by allowing $\rho(x)$ to include δ-functions. To conform with the notation used in the various works discussing moments, we do not use δ-functions but instead replace $\rho(x)$ by $\varphi(x)$ where $d\varphi/dx = \rho(x)$ for continuous values of $\rho(x)$ and where $\varphi(x)$ takes a discontinuous step proportional to the coefficient of the appropriate δ-function where $\rho(x)$ is discrete. Thus, $\varphi(x)$ is a nondecreasing function whose range has the width $\mu_0(\rho)$. Common intervals for the range are $[-^1/_2, ^1/_2]$ and $[0,1]$. When φ is used, the jth moment is represented as $\int_l^L x^j \, d\varphi(x)$ or simply $\int_l^L x^j \, d\varphi$. $\varphi(x)$ is called a weighting function.

(5) **Claim.** If two weighting functions ψ and φ are continuous at all but a finite number of points and if $\mu_j(\psi) = \mu_j(\varphi)$ for all j and if $\psi(-\infty) = \varphi(-\infty)$, then $\psi = \varphi$.

Proof. Assume false. Let $h(x) = \psi(x) - \varphi(x)$. Assume $h(x) \neq 0$ in part of the domain where neither ψ or φ is discontinuous. Let c be a point where ψ and φ are both continuous, not flat, and $h(c) \neq 0$. There is therefore a closed interval around c where $h(x)$ is either purely positive or purely negative. Introduce a polynomial series $\sum P_n(x)$ which converges uniformly to the function $P(x)$ shown in **18**.

18

(20) (a) The actual numerical method used in this paper has a kinship with the recursion method described by Heine and co-workers.[18b,c] There the a_i and b_i coefficients of Appendix section III are determined directly without the need to generate the moments themselves. The approach we have followed here, however, relies completely on a knowledge of the μ_i. (b) Haydock, R.; Heine, V.; Kelly, M. *J. Phys. C* **1972**, *5*, 2845. (c) Heine, V.; Haydock, R.; Bullett, D.; Kelly, M. *Solid State Phys.* **1980**, *35*, 1.

(21) Grant (Grant, W. J. C. *Physica* (Amsterdam) **1964**, *30*, 1433) explores some of the difficulties of the moment method. He points out thaat the knowledge of a finite number of moments (say the first n moments) in no way determines the DOS. One way of viewing this result is to recall that there exist functions whose first n moments are all zero. Thus, by including such functions, we can make arbitrary changes. Our view of the matter would be that although such effects can occur, they will only be observed if in the $n + 1$ walks a new and phenomenal structural feature also appears for the first time. If no such structural feature is apparent, we may ignore this possibility.

(22) Moments methods have been used for many years in several areas such as spectral densities associated with the vibrations of solids. See for example; Wheeler, J. C. *Phys. Rev. A* **1974**, *9*, 825.

(23) Wall, H. S. "Continued Fractions"; Van Nostrand: New York, 1948.

$$\int_I^L P(x) \, dh = \int_I^L \sum_n P_n(x) \, dh = \sum_n \int_I^L P_n(x) \, dh =$$

$$\sum_n \left\{ \int_I^L P_n(x) \, d\phi - \int_I^L P_n(x) \, dx \right\} = 0 \quad (25)$$

This is a contradiction. One can find a similar contradiction for the discrete portions of ψ and φ, QED.

(6) The above tells us in principle that it is possible to reconstruct a DOS from the moments of the DOS. Before showing an explicit reconstruction algorithm, let us first show how the moments of a DOS may be used to distinguish purely discrete DOS from other possible DOS.

(a) **Claim.** Let φ be a function composed of n steps as is shown in **19**. Let the steps be at $\lambda_i = 1,...,n$. Then there exists a

19
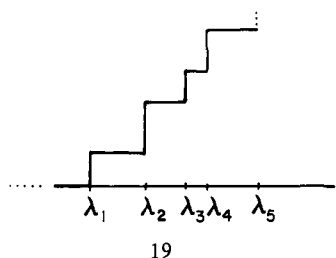
polynomial of order n, $P(x)$ such that for all bounded functions $f(x) = \int_{-\infty}^{\infty} P(x)f(x) \, d\varphi = 0$.

Proof. Let

$$P(x) = \prod_{i=1}^{n} (x - \lambda_i) \quad (26)$$

QED. We use this to show claim b.

(b) **Claim.** Let

$$M_m = \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_m \\ \mu_1 & \mu_2 & \cdots & \vdots \\ \mu_2 & \cdots & & \vdots \\ \vdots & & & \\ \mu_m & & & \mu_{2m} \end{pmatrix} \quad (27)$$

and let $\rho(x)$ be a purely discrete DOS. Let $\mathrm{Det}(M_n) = 0$ if $m \geq n$ where n is the number of nondegenerate eigenvalues in $\rho(x)$.

Proof. Consider the case $m \geq n$. As in (a) let $P(x) = \prod_{i=1}^{n}(x - \lambda_i) = \sum_{i=0}^{n} \rho_i x^i$. Consider a polynomial $R(x) = \sum_{i=0}^{m} r_i x^i$ of order m or less. We know

$$\int R(x)P(x) \, d\varphi = 0 \quad (28)$$

Written in matrix form

$$(r_1 r_2 \cdots r_m) M_m \begin{pmatrix} P_1 \\ \vdots \\ P_n \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0 \quad \text{for all } R(x) \quad (29)$$

As this is true for all $\bar{r}$ where $\bar{r}$ is the vector representation of $R(x)$, therefore, $M_m \bar{p} = 0$ where $\bar{p}$ is the vector representation of $P(x)$. Thus, $\bar{p}$ is an eigenvector of M_m with an eigenvalue of zero. Therefore, $\mathrm{Det}(M_m) = 0$.

On the other hand, let $m < n$. Assume $\mathrm{Det}(M_m) = 0$. As M_m is symmetric there exists an eigenvalue of zero whose eigenvector we will denote as $\bar{q}$. Let $Q = \sum_{i=0}^{m} q_i x^i$. But $\int Q^2(x) \, d\varphi \neq 0$ as for at least one $\lambda_j, Q(\lambda_j) \neq 0$. This is a contradiction. QED.

(7) As a corollary of (6) $\mathrm{Det}(M_m) \neq 0$ for any DOS with an infinite number of eigenvalues. In addition M_m has no negative eigenvalues.[24] Hence, for systems with an infinite number of

eigenvalues, M_m is always a positive definite hermitian matrix. Recalling "Babylonian Reduction", there exists a triangular matrix

$$P_n = \begin{pmatrix} P_{00} & P_{01} & \cdots & P_{0m} \\ & P_{11} & & \vdots \\ & & & \vdots \\ O & & & P_{mm} \end{pmatrix} \quad (30)$$

such that

$$P_m^{\dagger} M_m P_m = D_m \quad (31)$$

where D_m is a diagonal matrix. Let

$$P_j(x) = \sum_{i=0}^{j} P_{ij} x^i \quad (32)$$

These $P_j(x)$ have several interesting properties:

(a) $P_j(x)$ is always a polynomial of exactly the order j. Hence, $\{P_0(x), P_1(x),...,P_n(x)\}$ form a linearly independent basis set for the set of all polynomials of the order n or less.

(b) For every $\phi(x)$, there corresponds to it a sequence $\{P_j(x)\}$. These sequences are infinite sequences when $\phi(x)$ has an infinite number of different eigenvalues.

(c) $\int P_j(x)P_k(x) \, d\phi = w_j \delta_{kj}$ where $\{P_i(x)\}$ is the sequence corresponding to $\phi(x)$. One famous example of such sequences is the Legendre polynomials. The Legendre polynomials correspond to

$$\rho(x) = \begin{cases} 1/2 & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{or } \phi(x) = \begin{cases} 1/2 x & -1 \leq x \leq 1 \\ -1/2 & x < -1 \\ 1/2 & x > 1 \end{cases} \quad (33)$$

(8) We may use these Legendre polynomials, let us call them $L_j(x)$, to construct a specific algorithm to reconstruct a DOS from its sequence of moments.

Claim. For any Hückel DOS $\rho(x)$ contained in the interval $[-1,1]$, there exists a series expansion in $L_j(x)$

$$\rho_n(x) = \sum_{j=0}^{n} \alpha_j L_j(x) \quad (34)$$

such that $\lim_{n \to \infty} \rho_n(x) = \rho(x)$ and

$$\mu_i(\rho) = \mu_i(\rho_n) \quad \text{for all } i \leq n \quad (35)$$

Proof. Note that to prove the fact $\lim_{n \to \infty} \rho_n(x) = \rho(x)$, it suffices to show that $\mu_i(\rho) = \mu_i(\rho_n)$ for all $i \leq n$. The former follows from the latter by the result of (5). We now construct the appropriate sequence ρ_n. We do so by construction. First we set $\alpha_0 = 1/2 \mu_0(\rho)$. As $L_0(x) = 1$, $\int_{-1}^{1} \alpha_0 L_0 \, dx = \mu_0(\rho)$. This determines ρ_0. Now we set $\alpha_1 = (\mu_1(\rho) - \mu_1(\rho_0))/\mu_1(L_1)$. Thus, $\int_{-1}^{1} \alpha_1 L_1 + \alpha_0 L_0 \, dx = \mu_0(\rho)$ and $\int_{-1}^{1} (\alpha_1 L_1 + d_0 L_0)x \, dx = \mu_1(\rho) - \mu_1(\rho_0) + \mu_1(\rho_0) = \mu_1(\rho)$. This establishes ρ_1. It can been seen that in general by setting $\alpha_n = (\mu_n(\rho) - \mu_n(\rho_{n-1}))/(\mu_n(L_n))$, we can obtain the function ρ_n, QED.

Before continuing we would like to introduce the following notation. Whenever a sequence of polynomials is defined such that $\mu_n(P_n) = 1$, we shall call such a sequence properly normalized. We use the letter P_n here instead of L_n as in the next section; we shall see L_n is not the only sequence of polynomials which may be constructed which have a proper trigonal form.

(9) It is also possible to establish a sequence of polynomial functions $\{\rho_n\}$ approximating ρ such that not only are $\mu_i(\rho_n) = \mu_i(\rho)$ for $\lambda \leq n$ but they in addition enforce supplementary constraints.

(a) Let $\rho_n(-1) = \rho_n(1) = \rho(-1) = \rho(1) = 0$ where ρ need only be defined in the interval $[-1,1]$. Again we find a polynomial expansion which we shall call $\{Q_n\}$ which retains the proper triangular form given in sections 7 and 8; that is, $\mu_i(Q_n) = 0$ for $i < n$. Clearly the lowest order polynomial that may be used is one of second order and therefore we set

$$Q_0(x) = x^2 - 1 \quad (36)$$

---

(24) For were it to have a negative eigenvalue, there would then be an eigenvector $\bar{r}$ such that its corresponding polynomial $R(x)$ would have the following property: $\int R^2(x) \, d\phi < 0$. As ϕ is monotonically increasing, this leads to a contradiction.

where $\mu_0(\beta Q_0(x))$ is set equal to $\mu_0(\beta)$.

It is straightforward to calculate further Q_n, and we list here the first five.

unnormalized Q polynomials

$$Q_0(x) = x^2 - 1 = (x + 1)(x - 1)$$

$$Q_1(x) = x^3 - x = x(x + 1)(x - 1)$$

$$Q_2(x) = x^4 - \frac{6}{5}x^2 + \frac{1}{5} =$$
$$(x + 1)(x - 1)\left(x + \frac{1}{(5)^{1/2}}\right)\left(x - \frac{1}{(5)^{1/2}}\right)$$

$$Q_3(x) = x^5 - \frac{10}{7}x^3 + \frac{3}{7}x =$$
$$(x^3 - x)\left(x + \left(\frac{3}{7}\right)^{1/2}\right)\left(x - \left(\frac{3}{7}\right)^{1/2}\right)$$

$$Q_4(x) = x^6 - \frac{5}{3}x^4 + \frac{5}{7}x^2 - \frac{1}{21} = (x^2 - 1) \times$$
$$\left(x + \left(\frac{1}{3} + \frac{2}{3(7)^{1/2}}\right)^{1/2}\right)\left(x - \left(\frac{1}{3} + \frac{2}{3(7)^{1/2}}\right)^{1/2}\right)\left(x + \left(\frac{1}{3} - \frac{2}{3(7)^{1/2}}\right)^{1/2}\right)\left(x - \left(\frac{1}{3} - \frac{2}{3(7)^{1/2}}\right)^{1/2}\right) \quad (37)$$

We note the roots of each Q_r polynomial lie nested between the roots of the Q_{r-1} polynomial; for instance the roots of Q_3 at 0, $(3/7)^{1/2}$, $-(3/7)^{1/2}$ lie between the roots of Q_2 which are -1, $-(1/5)^{1/2}$, $(1/5)^{1/2}$, and 1. One may find the Q polynomial also by following the technique of Appendix section III. A noniterative method involves the following. In Q_{2n} polynomial, there are $n + 2$ parameters to be determined as $Q_{2n} = \sum_{j=0}^{n+1} c_{2j}x^{2j}$ and $n + 1$ constraints, i.e., $\mu_0, \mu_2, ..., \mu_{2n}$. Finally one of the c_j parameters may be arbitrarily fixed to equal one as a normalization condition. This results in an $n + 1$ variable, $n + 1$ simultaneous equation problem. A similar problem is found for Q_{2n+1} polynomials.

(b) We may consider other similar constraint problems. For example, say $g(x)$ is contrained so $g(-1) = -\frac{1}{2}$ and $g(1) = \frac{1}{2}$ where g is only defined from $[-1,1]$. Let $f(x) = g(x) - \frac{1}{2}x$. Then $f(x)$ can be approximated by the Q polynomials of section 10a and we let $g_n(x) = \frac{1}{2}x + f_n(x)$ where $f_n(x)$ are the Q expansions given above.

These ideas may be generalized to handle an arbitrary number of constraints.

(10) Such approximations are useful for at least the following two reasons:

(a) Let f and g be two functions defined on the interval $[-1,1]$ which share the first n moments. Then either $f = g$ or $h(x) \equiv f(x) - g(x)$ changes sign at least $n + 1$ times in $[-1,1]$.

Proof. By contradiction: As

$$\int_{-1}^{1} h(x)\, dx = \int_{-1}^{1} f(x)\, dx - \int_{-1}^{1} g(x)\, dx = \mu_0(f) - \mu_0(g) = 0 \quad (38)$$

Assuming $h(x) \neq 0$, then $h(x)$ changes sign at least once. Now assume it changes sign only r times where $r < n + 1$ at $x_1, x_2, ..., x_r$. Let

$$P(x) = \prod_{i=1}^{r} (x - x_i) \quad (39)$$

Therefore, $\int P(x)h(x)\, dx = 0$ as $P(x)$ is a polynomial of order r but $\int P(x)h(x)\, dx \geq 0$ by assumption. Therefore, we find a contradiction.

(b) The polynomial of order n which extremizes the integral

$$\int_{-1}^{1} (f - \sum_{i=0}^{n} \beta_i x^i)^2\, dx \quad (40)$$

is $g_n(x)$ where $g_n(x) = \sum_{j=0}^{n} \alpha_j(f)P_j(x)$ where the $\alpha_j(f)$ are defined in section 9.

Proof.

$$\frac{\partial}{\partial \beta_j} \int_{-1}^{1} (f - \sum_{i=0}^{n}\beta_i x^i)^2\, dx = 2\int_{-1}^{1}(f - \sum_{i=0}^{n}\beta_i x^i)x^j\, dx = 0 \quad (41)$$

Therefore

$$\int_{-1}^{1} f x^j\, dx = \int x^j \sum_{i=0}^{n} \beta_i x^i \quad (42)$$

Thus,

$$\mu_j(f) = \mu_j(\sum_{i=0}^{n}\beta_i x^i) \text{ for all } j \leq n, \text{ QED} \quad (43)$$

(11) It should be pointed out that these polynomial sequences do not in general converge very rapidly. In Appendix section III, we outline a much more efficient method. Our purpose here is to form a qualitative understanding of how moments affect a function's shape. We are especially interested in the following problem: Let two functions f and g defined from $[-1,1]$ have the same first $m - 1$ moments. Let $h(x) = f(x) - g(x)$. Let us have expanded $h(x)$ with respect to a properly normalized polynomial basis set $R_n(x)$. Set $h_n(x) = \sum_{i=0}^{n}\alpha_i R_i(x)$ where $\lim_{n\to\infty} h_n(x) = h(x)$.

As $\mu_j(h) = 0$ for $j = 0,1,...,m - 1$, therefore $\alpha_1 = \alpha_2 = ... = \alpha_{m-1} = 0$. We shall call calculating α_m and α_{m+1} determining $h(x)$ to first order with respect to R_n. Calculation via α_m, α_{m+1}, α_{m+2}, and α_{m+3} will be called second order.

In the case where $R_n(x)$ polynomials have the property of being purely even or odd, we find using the formulas of section 8

$$\alpha_n = \frac{\mu_n(\rho) - \mu_n(\rho_{n-1})}{\mu_n(R_n)} \quad (44)$$

Thus

$$\alpha_m = \mu_m(h)$$
$$\alpha_{m+1} = \mu_{m+1}(h)$$
$$\alpha_{m+2} = \mu_{m+2}(h) - \mu_m(h)\mu_{m+2}(R_m)$$
$$\alpha_{m+3} = \mu_{m+3}(h) - \mu_{m+1}(h)\mu_{m+3}(R_{m+1}) \quad (45)$$

II. (1) Here we will consider the following model problem. Let two systems have DOS, WF, EBF, and EFE given by ρ_1, ϕ_1, E_1, and ϵ_1 and ρ_2, ϕ_2, E_2, and ϵ_2, respectively, and let

$$\mu_i(\rho_1) = \mu_i(\rho_2) \quad \text{for } i = 0,1,...,m - 1 \quad (46)$$

defining $\Delta\rho = \rho_1(x) - \rho_2(x)$, $\mu_i(\Delta\rho) = 0$, when $i \leq m - 1$. (We note parenthetically that $\mu_j(f + g) = \mu_j(f) + \mu_j(g)$; i.e., the moments act in a linear fashion on functions.) Also

$$\mu_j(\Delta\phi) = \mu_j(\phi_1) - \mu_j(\phi_2) = 0 \quad j \leq m - 2 \quad (47)$$

where $\Delta\phi = \phi_1 - \phi_2$. And

$$\mu_k(\Delta\epsilon) = \mu_k(\epsilon_1) - \mu_k(\epsilon_2) = 0 \quad k \leq m - 3 \quad (48)$$

where $\Delta\epsilon = \epsilon_1 - \epsilon_2$. The result given in Appendix section I.9a deals with such a situation. It tells us that $\Delta\phi(x)$ must change sign in the interval $[-1,1]$ at least $m - 1$ times while $\Delta\epsilon(x)$ must change sign at least $m - 2$ times. This is a result with tremendous implications, as we will see in the following two papers in this issue.

(2) We may quantify and strength this result by using the Q polynomial expansion of Appendix section I.9. Here we consider only the first-order and typically dominant term in Q. Recall that this first-order term is $\alpha_n Q_n(x) + \alpha_{n+1}Q_{n+1}(x)$ where $\alpha_n = \mu_n$ and $\alpha_{n+1} = \mu_{n+1}$. In our example, $n = m - 1$ for $\Delta\phi$ and $n = m - 2$ for $\Delta\epsilon$. So

$$\Delta\phi \simeq -(1/m)\mu_m(\Delta\rho)Q_{m-1} - (1/(m + 1))\mu_{m+1}(\Delta\rho)Q_m \quad (49)$$

and

$$\Delta\epsilon \simeq (1/m - 1)\mu_m(\Delta\rho)Q_{m-2} + (1/m)\mu_{m+1}(\Delta\rho)Q_{m-1} \quad (50)$$

Thus to first order (Appendix section I.11), $\Delta\phi$ changes sign in the interval $[-1,1]$ $m - 1$ or m times while $\Delta\epsilon$ changes sign $m - 2$ or $m - 1$ times. This arises simply because $\alpha_n Q_n(x) + \alpha_{n+1}Q_{n+1}(x)$ is at most a polynomial of the order $n + 3$. Hence the first-order term for $\Delta\phi$ is at most a polynomial of the order

$m + 2$. Two of the nodes are fixed at ± 1. Thus, there are a maximum of m remaining nodes. Similarly we see for $\Delta\epsilon$ a maximum of $m - 1$ nodes.

(3) Of considerable importance to us will be the location of these nodes. Let us assume that α_n and α_{n+1} are of the same sign. [Note that if this is true for $\Delta\phi$, it is also true for $\Delta\epsilon$.] As noted in Appendix section I.10, the $n + 1$ nodes not located at ± 1 in Q_{n+1} are nested by the nodes in Q_n. We show this graphically in Figure 9. There are regions where nodes are forbidden to occur (the shaded regions) and others (unshaded) where nodes must occur. The only node unaccounted for is the possible "extra node" introduced by Q_{n+1}. It generally lies between the last node (M_5 in the diagram) and $+1$. In the case where α_n and α_{n+1} are of opposite signs, the shaded regions are now the regions where nodes must occur and the unshaded regions where they are forbidden. Now the extra node generally occurs between -1 and the first node of Q_{n+1} (M_1 in the diagram).

Using the first-order term in Q, we may also determine the sign of $\Delta\phi$ and $\Delta\epsilon$ from $[1,-1]$. To do so we note that if Q_n is an odd function Q_{n+1} must be even and vice-versa. Hence, only even Q_n have a nonzero value when evaluated at $x = 0$. But the number of nodes and the relative position of these nodes are fixed. By calculating $\Delta\phi(0)$ and $\Delta\epsilon(0)$ values on working outward from $x = 0$, the relative phases elsewhere may be calculated.

As can be seen from Appendix section I.9a, $Q_{4j+2}(0) < 0$ while $Q_{4j}(0) > 0$. Recalling eq 49 and 50, we find the sign of $\Delta\phi(0)$ and $\Delta\epsilon(0)$. These results are tabulated below and, as noted above, will determine the phase of $\Delta\phi$ and $\Delta\epsilon$ at points other than zero.

$m = 1$ $\mu_1(\Delta\rho) > 0 \Rightarrow \Delta\phi(0) < 0$, $\mu_2(\Delta\rho) > 0 \Rightarrow \Delta\epsilon(0) > 0$,
　　　$\mu_1(\Delta\rho) < 0 \Rightarrow \Delta\phi(0) > 0$, $\mu_2(\Delta\rho) < 0 \Rightarrow \Delta\epsilon(0) < 0$

$m = 2$ $\mu_3(\Delta\rho) > 0 \Rightarrow \Delta\phi(0) > 0$, $\mu_2(\Delta\rho) > 0 \Rightarrow \Delta\epsilon(0) > 0$,
　　　$\mu_3(\Delta\rho) < 0 \Rightarrow \Delta\phi(0) < 0$, $\mu_2(\Delta\rho) < 0 \Rightarrow \Delta\epsilon(0) < 0$

$m = 3$ $\mu_3(\Delta\rho) > 0 \Rightarrow \Delta\phi(0) > 0$, $\mu_4(\Delta\rho) > 0 \Rightarrow \Delta\epsilon(0) < 0$,
　　　$\mu_3(\Delta\rho) < 0 \Rightarrow \Delta\phi(0) < 0$, $\mu_4(\Delta\rho) < 0 \Rightarrow \Delta\epsilon(0) > 0$

$m = 4$ $\mu_5(\Delta\rho) > 0 \Rightarrow \Delta\phi(0) < 0$, $\mu_4(\Delta\rho) > 0 \Rightarrow \Delta\epsilon(0) < 0$,
　　　$\mu_5(\Delta\rho) < 0 \Rightarrow \Delta\phi(0) > 0$, $\mu_4(\Delta\rho) < 0 \Rightarrow \Delta\phi(0) > 0$

$m = 5$ $\mu_5(\Delta\rho) > 0 \Rightarrow \Delta\phi(0) < 0$, $\mu_6(\Delta\rho) > 0 \Rightarrow \Delta\epsilon(0) > 0$,
　　　$\mu_5(\Delta\rho) < 0 \Rightarrow \Delta\phi(0) > 0$, $\mu_6(\Delta\rho) < 0 \Rightarrow \Delta\epsilon(0) < 0$　(51)

Perhaps the most chemically interesting expression is $\Delta E(x) = E_1(x) - E_2(x)$. As shown above, the moments of $\Delta E(x)$ are not as easily calculated as those of $\Delta\phi$ and $\Delta\epsilon$. Yet $\Delta\phi$ and $\Delta\epsilon$ must to a large extent determine ΔE. For instance, say at a given Fermi level, ϵ_F, we already know $\Delta\phi(\epsilon_F)$ and $\Delta\epsilon(\epsilon_F)$ where as before we take these to represent $\phi_1(\epsilon_F) - \phi_2(\epsilon_F)$ and $\epsilon_1(\epsilon_F) - \epsilon_2(\epsilon_F)$, respectively. We see

$$\Delta E(\phi_1(\epsilon_F)) = \epsilon_1(\epsilon_F) - (\epsilon_2(\epsilon_F) + y) = \Delta\epsilon(\epsilon_F) + y$$

where

$$y = \int_{\epsilon_F}^{\phi_2^{-1}[\phi_1(\epsilon_F)]} t\rho_2(t) \, dt \qquad (52)$$

Hence, y represents the excess or deficiency of $\phi_2(\epsilon_F)$ compared with $\phi_1(\epsilon_F)$ [this is just $\Delta\phi$ itself] multiplied by the average energy that such a $\Delta\phi$ occupies on the DOS curve. This "average energy" is a function of ϕ_2. For instance, if beneath ϵ_F there is a band gap, then the "average energy" is much less than ϵ_F. Similarly if at ϵ_F the DOS has a sharp peak, then to a good approximation the "average energy" is just ϵ_F itself. Keeping in mind that we are only interested in qualitative results, we find it simplest not to concern themselves with the actual shape of ρ_2. Instead we assume ρ_2 to have a uniform rectangular distribution. Assuming $\mu_0 = 1$ and and upper and lower bound to ρ_2 of ± 1, then $\rho_2(\epsilon_F) = {}^1/_2$ and $\phi_2(\epsilon_F) = \epsilon_F/2$. Thus, the average energy is just $\epsilon_F - \Delta\phi$ and

$$\Delta E(\phi_1(\epsilon_F)) = \Delta\epsilon + (\epsilon_F - \Delta\phi)\Delta\phi \qquad (53)$$

Our expression for ΔE is now almost complete. The only difference between $\Delta E(x)$ and the $\Delta E(\phi_1(\epsilon_F))$ calculated above

is a stretching and contraction along the x axis caused by $\phi_1(\epsilon_F)$. To remove this warping along the x coordinate, we need evaluate ϕ_1. Unlike before, even for a qualitative answer, we cannot "assume" a shape for ϕ_1, and in practice, we use for ϕ_1 a function calculated via the moment method.

One interesting consequence of such an approach comes from the fact that to first order in Q, $\Delta\epsilon$ has n crossings while $\Delta\phi$ has $n + 1$ crossings where each of the $\Delta\phi$ crossings are nested between the crossings of $\Delta\epsilon$. But $\Delta E(\phi_1(x)) = \Delta\epsilon(x)$ whenever $\Delta\phi = 0$ and so $\Delta E(\phi_1(x))$ must have at least n crossings as must $\Delta E(x)$. Finally we would like to note that the "warping" in converting $\Delta E(\phi_1(x))$ into $\Delta E(x)$ has a significant effect on the placement of the nodes. This is particularly true when a large range of energy states is covered by very few energy levels. This happens in the following two cases: (1) There is a band gap in the middle of the DOS and (2) a large tail exists which trails off to one end of the energy spectrum. [Generally, in Hückel theory, the tail is at the most bonding limit, and in extended Hückel, the tail trails off to the most antibonding limit. In the latter, the reason for this is obviously due to the inclusion of overlap.]

In both of these cases in going from $\Delta E(\phi_1(x))$ to $\Delta E(x)$, the broad energy range collapses into a narrow band-filling range. Hence, all nodes are shifted toward the point of collapse.

We pay special attention to the effect due to tails and in the following paper in this issue include this effect in our qualitative rule.

III. (a) Construction of $P_j(x)$ Polynomials.[23] We know from Appendix section I that there exists for every nondiscrete $\phi(x)$ a complete basis set of polynomials $\{P_j(x)\}$ such that

$$\int P_j(x)P_i(x) \, d\phi(x) = k_j\delta_{ij} \qquad (54)$$

where $k_j > 0$ for all j and $P_j(x)$ is a polynomial of order j.

These $P_j(x)$ are only specified to an arbitrary constant. We use this constant to normalize $P_j(x)$ so that its x^j coefficient = 1.

We now show a recursive method to construct these polynomials.

Define the two sequences $\{a_i\}$ and $\{b_j\}$ such that[25]

$$\int x^r P_r(x) \, d\phi = a_0 a_1 a_2 \ldots a_r \qquad (55)$$

and

$$\int x^{r+1} P_r(x) \, d\phi = (-a_0 a_1 a_2 \ldots a_r)(b_1 + b_2 + \ldots b_{r+1}) \qquad (56)$$

Therefore

$$\int x^r P_r(x) \, d\phi = \int [P_r(x) + \sum_{i=0}^{r-1} c_i P_{r-1}(x)]P_r(x) \, d\phi = \int P_r^2(x) \, d\phi \qquad (57)$$

Claim.

$$P_m(x) = (b_m + x)P_{m-1}(x) - a_{m-1}P_{m-2}(x) \quad \text{where } P_{-1}(x) = 0 \qquad (58)$$

Before proving the above assertion, we note the polynomials so defined have the property

$$P_m(x) = x^m + (b_1 + b_2 + \ldots + b_{m-1})x^{m-1} + \text{lower order terms in } x \qquad (59)$$

Proof. We prove the claim by induction. First we show this true for $P_1(x)$. From our normalization constraint, we known $P_0(x) = 1$. Hence, from (2) $a_0 = 1$.[26] Similarly from (57), we know $\int x P_0(x) \, d\phi = -b_1$. Hence, as $\int P_1(x)P_0(x) \, d\phi = 0$, $P_1(x) = x + b_1$. This conforms to (58). Similarly we can show $P_2(x)$ also conforms to (58).

(25) Note $k_r = a_0 a_1 \ldots a_n$. As k_r is never zero, the above definitions of $\{a_i\}$ and $\{b_j\}$ involve no added assumption. This is so as $\int x^r P_r(x) \, d\phi = \int P_r^2(x) \, d\phi = k_r$. This identity is due to the fact that all polynomials of the order r or less may be expressed as a linear combination of $P_0(x), P_1(x), \ldots, P_r(x)$.

(26) We assume here that $\int d\phi = 1$ and therefore $\mu_0(\rho) = 1$.

Now let us assume the claim is true for the first $n - 1$ $P_j(x)$. All we need show is that $R_n(x)$

$$R_n(x) \equiv (b_n + x)P_{n-1}(x) - a_{n-1}P_{n-2}(x) \qquad (60)$$

has the property $\int R_n(x)P_i(x) \, d\phi = 0$ for $i = 0,1,2...,n-1$ to complete the proof. We prove this case by case.

Case 1. $i \le n - 3$.

$$\int R_n(x)P_i(x) \, d\phi = \int P_{n-1}(x)[xP_i(x)] \, d\phi = 0 \qquad (61)$$

as $xP_i(x)$ is a polynomial of the order $n - 2$ and hence can be written as a linear combination of $P_0(x),P_1(x),...,P_{n-2}(x)$.

Case 2. $i = n - 2$.

$$\int R_n(x)P_{n-2}(x) \, d\phi = \int (xP_{n-1}(x) - a_{n-1}P_{n-2}(x))P_{n-2}(x) \, d\phi \qquad (62)$$

$$= \int x^{n-1}P_{n-1}(x) - a_{n-1}(a_1 a_2...a_{n-2}) \, d\phi = 0 \qquad (63)$$

Case 3. $i = n - 1$.

$$\int R_n(x)P_{n-1}(x) \, d\phi = \int (b_n + x)P_{n-1}(x)P_{n-1}(x) \, d\phi =$$
$$b_n(a_1 a_2...a_{n-1}) + \int (x^n + (b_1 + b_2 + ... +$$
$$b_{n-1})x^{n-1})P_{n-1}(x) \, d\phi = 0 \qquad (64)$$

where the second to last equality follows from (56), QED.

(b) Relation of the $P_j(z)$ Polynomials to the Power Series $\sum_{j=0}^{\infty}\mu_j/z^{j+1}$. In the preceding section we showed how from a knowledge of the moments of a DOS we could construct an orthogonal polynomial basis set $\{P_j(z)\}$. In this section we construct a function $G(1/z)$ from the $\{P_j(z)\}$ basis set. This function $G(1/z)$ will be important for the actual reconstruction of the DOS from its moments. We define $G(1/z) = \sum_{j=0}^{\infty}\mu_j/z^{j+1}$. (We note parenthetically that this definition assumes the convergence of the above series.) Taking the product of $G(1/z)$ and $P_m(z)$, we find

$$G\left(\frac{1}{z}\right)P_m(z) = \sum_{i=0}^{\alpha}\sum_{j=0}^{\alpha}p_{jm}\mu_i z^{j-i-1} \qquad (65)$$

$$= \sum_{k=1}^{\alpha}z^{m-k}\left(\sum_{j=0}^{\infty}\mu_j p_{m+j-k+1,m}\right) \qquad (66)$$

where $P_m(z) = p_{m,m}z^m + p_{m-1,m}z^{m-1} + ... + p_{0,m}$ and where $p_{im} = 0$, $m < i$, or $i < 0$. Now we shall show that the coefficients in (66) corresponding to $z^{-1},z^{-2},...,z^{-m}$ are all zero. We note for $l = 0,1,2,...,m - 1$

$$0 = \int z^l P_m(z) = \sum_{i=0}^{\infty}\mu_{l+m-i}p_{m-i,m} \qquad (67)$$

Setting $j = l + m - i$

$$= \sum_{j=0}^{\infty}\mu_j p_{j-l,m} \qquad (68)$$

Comparing (68) and (66), we see when $m - k + 1 = -l$ for $l = 0,1,2,...,m - 1$, then the z^{m-k} coefficient is 0. This proves that the $z^{-1},z^{-2},...,z^{-m}$ coefficients are all zero. Let us now define another polynomial sequence $R_m(z)$ such that $R_m(z)$ contains all the nonnegative powers of z in the expansion (66).[27] Thus

$$G\left(\frac{1}{z}\right)P_m(z) - R_m(z) = c\frac{1}{z^{m+1}} + \text{lower powers in } z \qquad (69)$$

or

$$G\left(\frac{1}{z}\right) = \frac{R_m(z)}{P_m(z)} + c'\frac{1}{z^{m+1}} + \text{lower powers in } z \qquad (70)$$

where c and c' are undetermined constants. Thus, we see that $R_m(z)/P_m(z)$ expanded in powers of z has the same coefficients as $G(1/z)$ for z^j with $j = -1,-2,...,-m$. Hence, given any known sequence of moments, we may always produce a $R_m(z)/P_m(z)$

---

(27) Note $R_m(z)$ is a polynomial of the order $m - 1$.

which matches our known sequence.

Finally observe

$$G\left(\frac{1}{z}\right)P_m(z) - R_m(z) = G\left(\frac{1}{z}\right)[(b + z)P_{m-1}(z) -$$
$$a_{m-1}P_{m-2}(z)] - R_m(z) = c\frac{1}{z^{m+1}} + \text{lower powers in } z \qquad (71)$$

Thus, we may let $R_m(z) = (b_m + z)R_{m-1}(z) - a_{m-1}R_{m-2}(z)$. Both $R_m(z)$ and $P_m(z)$ therefore follow the same recursion relations.

Due to this recursion relation, both $R_m(z)$ and $P_m(z)$ may be calculated from a knowledge of $\{\mu_j\}$. Therefore $G(1/z)$ can also be derived from the moments.

(c) Hilbert–Stieljes Transform. $H(z)$, the Hilbert–Stieljes transform, is defined

$$H(z) = \int_{-\infty}^{\infty}\frac{d\phi(x)}{z - x} \qquad (72)$$

It is convenient to view z as a complex variable. For z sufficiently large, we may introduce the following convergent power series expansion:

$$H(z) = \int_{-\infty}^{\infty}\frac{1}{z}\frac{d\phi(x)}{1 - (x/z)} = \int_{-\infty}^{\infty}\sum_{n=0}^{\infty}\frac{x^n}{z^{n+1}} \, d\phi(x) = \sum_{n=0}^{\infty}\mu_n\frac{1}{z^{n+1}} \qquad (73)$$

Comparing this to the result of the preceding section, we see $H(z) = \lim_{m\to\infty}R_m(z)/P_m(z)$. As we have shown in the preceding section, $R_m(z)$ and $P_m(z)$ may both be calculated from a knowledge of the moments. All that remains is to invert $H(z)$ back into $\phi(z)$. To do so one uses Stieljes inversion formula

$$\frac{1}{\pi}\lim_{\epsilon\to 0}\int_s^t \text{Im}(H(z + i\epsilon)) \, dz = \phi(s) - \phi(t) \qquad (74)$$

It should be pointed out that in the case of a purely continuous DOS the above reduces to

$$\lim_{\epsilon\to 0}\frac{1}{\pi}\text{Im}(H(z + i\epsilon))|_s^t = \rho(s) - \rho(t) \qquad (75)$$

Similarly in the case of a purely discrete DOS

$$H(z) = \int_{-\infty}^{\infty}\frac{\sum_i \alpha_i\delta(x - x_i)}{z - x} \, dx = \sum_i\frac{\alpha_i}{z - x_i} \qquad (76)$$

where $\{x_i\}$ are the possible values of the distribution with respective weights of α_i. Thus, in the case of a discrete DOS, one knows directly what $H(z)$ is from the DOS. For worked examples we refer the reader to ref 13a. Finally we should like to note that the $\{a_j\}$ and $\{b_j\}$ sequences discussed above often form regular limiting patterns. In many cases $\{a_j\}$ and $\{b_j\}$ converge to a single limit point. In other cases they have two or more limiting points. In these cases the $\{a_j\}$ and $\{b_j\}$ march in cycles around each limiting point. In this, these sequences resemble decimal expansions of rotational numbers. Following the lead of Cyrot–Lackman, we use this fact by truncating (since we only usually know a finite number of moments) our a_j and b_j sequences to the "correct" limit points. We call these final values a_{final} and b_{final}. Of course there may be several a_{final} and b_{final} in a given structure. Finally as Cyrot–Lackman et al. have pointed out, these a_{final} and b_{final} control the limits of the DOS. The converse, of course, is true.

IV. The moment problem often occurs in the following form. We wish to reconstruct a DOS, ρ, from its sequence of moments, $\{\mu_n(\rho)\}$. In addition we know ρ resembles a known function which we call ρ^T. (T stands for trial.) For many ρ^T there exist well-cataloged methods to transform ρ^T into ρ. We illustrate this technique for $\rho^T(x) = (2\pi)^{1/2}\exp(-1/2 x^2)$.

The method is based on the observation

$$\int_{-\infty}^{\infty}H_n(x)H_m(x)(2\pi)^{1/2}\exp(-1/2 x^2) \, dx = n!\delta_{mn} \qquad (77)$$

where $\{H_n(x)\}$ are the Hermite polynomials. As the Hermite

polynomials form a complete basis set we may solve for $\rho(x)/\rho^T(x)$ in terms of them. Thus

$$\rho(x) = \sum_{j=0}^{\infty} c_j H_j(x)(2\pi)^{1/2} \exp(-\tfrac{1}{2}x^2) \qquad (78)$$

where $c_j = (1/j!) \int_{-\infty}^{\infty} \rho(x) H_j(x) \, dx$. This expansion though is a moment expansion.

Let us define

$$\rho_n(x) = \sum_{j=0}^{n} c_j H_j(x)(2\pi)^{1/2} \exp(-\tfrac{1}{2}x^2) \qquad (79)$$

Claim. $\mu_i(\rho_m) = \mu_i(\rho)$ for all $m \geq i$.

Proof.

$$\mu_i(\rho) = \int x^i \rho(x) \, dx = \int x^i \sum_{j=0}^{m} c_j H_j(x)(2\pi)^{1/2} \exp(-\tfrac{1}{2}x^2) \, dx \qquad (80)$$

for all $m \geq i$. In turn this

$$= \int x^i \sum_{j=0}^{i} c_j H_j(x)(2\pi)^{1/2} \exp(-\tfrac{1}{2}x^2) \, dx = \mu_i(\rho_m) \qquad (81)$$

The equivalence follows from the first footnote given in Appendix section III.

Hence, for arbitrary ρ^T, the problem reduces to finding the correct polynomial basis set. These though are nothing more than the $P_j(x)$ polynomials which correspond to the DOS ρ^T.

The Moments Method and Elemental Structures

Jeremy K. Burdett*[1] and Stephen Lee

Contribution from the Department of Chemistry, The University of Chicago, Chicago, Illinois 60637. Received May 21, 1984

Abstract: The ability of the method of moments to make qualitative predictions of structure is tested on the structures of main group and transition elements. Trends due to the presence of rings and bond angles of various sizes are discussed.

I. Introduction

In the previous[2] paper paper in this issue, we presented a simple method for estimating the electronic energy of solids by using a one-electron tight-binding model based on the Hückel approximation. This method in its simplest form expresses the energy difference between two structures in terms of the earliest disparate moment of their respective energy density of states (DOS). In Figure 1a we show the energy difference curves expected for a pair of structures as a function of band filling, X ($0 \leq X \leq 1$) where the first dominant moment difference occurs in μ_m ($m = 2$–6). Sometimes, if there is a "tail" on the bonding side of the density of states (arising perhaps via the inclusion of s and p orbitals in the bonding picture), then the nodes in these curves shift to smaller X. The convention used in Figure 1 is that when the curves have positive values, then that structure with the largest $|\mu_m|$ is the more stable. (We assume odd moments are negative in this statement.) Some generalizations are interesting to describe.

(1) Low moments are more important than higher moments in controlling the form of the energy difference curve. Real attention should be paid to μ_3 and μ_4.

(2) For two structures A and B, let us put $|\mu_3(A)| > |\mu_3(B)|$ and assume both $\mu_3(A)$ and $\mu_3(B)$ to be negative. Then, with reference to Figure 1, A will be the favored structure below the half-filled band and B will be favored above the half-filled band. The crossing point itself lies below the half-filled band for the $\Delta E_3(X)$ curves of the previous paper in this issue where we compared the energetics of the 3-ring within the 3-tree with the 3-tree itself. (In general the crossing points in the $\Delta E(X)$ curves will depend on the exact form of the densities of states for the two systems as discussed further in (5).)

(3) If $\mu_3(A) = \mu_3(B)$ but $\mu_4(A) > \mu_4(B)$, then B is favored at the half-filled band while A is favored at lower and higher band fillings.

(4) Systems with a large (negative) μ_5 are energetically stabilized just above the half-filled band. A large (positive) μ_6 stabilizes systems at the half-filled point itself.

(5) The presence of "tails" in the DOS which trail toward the bonding edge limit shift the crossing points of Figure 1a toward smaller X (Figure 1b). Such tails may be anticipated by the presence of large and negative values of μ_3. (If a large tail exists, then μ_4, μ_5, etc., should be large too of course. If μ_3 is large but μ_4 small, then this implies no strong tail.)

In this paper we wish to discuss the usefulness of these rules. Clearly we need to find out if the early moments are easy to determine (without the aid of a computer) and if the rules accurately predict Hückel energetics. The whole approach of course relies on the utility of one-electron, Hückel-based calculations themselves. There are many areas where such an approach does not accurately reflect electronic energy. Bond length changes and coordination number variations are two of them. We shall therefore in general avoid examples which are inappropriate for the Hückel or extended Hückel models.

II. Rings

The Hückel model of the π manifold in benzene (for example) is well-known.[3] It predicts a stabilization of benzene relative to other unsaturated carbon configurations—the so-called resonance energy. In terms of the moments approach[4] this result is due to both the fact that each carbon atom contributes one electron to the π system (a half-filled "band") and the presence in benzene of the six-membered ring. In the previous paper in this issue, we showed how at the half-filled point the six-membered ring was stabilized with respect to the simple tree. Similarly cyclobutadiene is viewed as an unfavorable arrangement by using traditional Hückel ideas. Analogously in the moments approach, 4-rings are destabilized at the half-filled point. However, as we will see frequently in this paper 6-rings are destabilized at other band

(1) Camille and Henry Dreyfus Teacher–Scholar.
(2) Burdett, J. K.; Lee, S. *J. Am. Chem. Soc.*, preceding paper in this issue.

(3) (a) Hückel, E. *Z. Phys.* **1931**, *70*, 204; **1932**, *76*, 628. (b) Streitweiser, A. "Molecular Orbital Theory for Organic Chemists"; Wiley: New York, 1961. (c) Heilbronner, E.; Bock, H. "The HMO Model and its Application"; Wiley: New York, 1976.
(4) We derive the Hückel eigenvalue spectra of some small conjugated hydrocarbons in: Burdett, J. K.; Lee, S.; Sha, W. C. *Croat. Chim. Acta* **1984**, *57*, 1193.