THE BAND GAP IN LINEAR POLYENES

D.R. YARKONY and R. SILBEY
Department of Chemistry, and Center for Materials Science & Engineering,
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 21 June 1976
Revised manuscript received 4 October 1976

The band gap in the spectrum of long polyenes is considered. It is shown, using a bond alternation model within a single configuration self-consistent field method, that the gap is determined largely by the difference in the exchange contribution to the ground and excited state energies. This model also predicts a $N^{-1}$ contribution to the gap, but of the incorrect sign (when compared to SCF calculations of moderate length polyenes). Secondly, it is shown that the exciton model, when multiple excitations are taken into account, gives a $N^{-1}$ contribution to the band gap, in contrast to the simplest exciton model.

1. Introduction

The nature of the lowest dipole allowed electronic transition in an extended polyene system has been a topic of considerable theoretical and experimental interest. The available experimental evidence [1] indicates that as a function of chain length the excitation energy ($\Delta E$) has the form

$$\Delta E(N) = \Delta E_0 + \Delta E_1 N^{-1},$$

where $\Delta E_1 > 0$. In this work we consider the analytic solution of simplified versions of two models, which have been solved numerically elsewhere, in order to understand the origin of the $\Delta E_0$ and $\Delta E_1$ contributions to $\Delta E$.

In section 2 we develop a simple single configuration approximation to the ground and lowest excited state wave functions. From this model we conclude that electron exchange plays a major role in determining $\Delta E_0$. In section 3, the exciton model of Simpson [2] is considered. There it is shown that if multiple excitations and the non-Bose–Einstein nature of exciton statistics are included in a simplified form of this model, a $\Delta E_1$ term emerges in $\Delta E(N)$. It had been previously supposed that such a term was lacking in exciton models. In section 4, we comment on the assumptions used in these model calculations and also consider the relationship of this study to previous work in this area.

2. Bond alternation method (BAM)

In this section, we develop a model of the infinite polyene which includes, in an approximate manner, the essential features of a single configuration wavefunction description of the system including electron–electron repulsion and exchange. This approach was motivated by the observation that in previous numerical self-consistent-field (SCF) calculations [3], the optimum one-electron orbitals showed bond alternation regardless of the interatomic spacing. It will be shown that the inclusion of electron-exchange is essential to the description of the band gap ($\Delta E(\infty)$), producing a gap of $\sim 2.5$ eV in a system with minimal bond alternation ($|\beta_L - \beta_S| = 0.12$ eV). In a subsequent section the results of this model calculation will be compared with previous work in this area including the exciton model considered in section 3.
The infinite polyene is modeled by a linear Bravais lattice with $2N + 1$ cells each containing two carbon atoms. The intercell (intracellular) spacing is $r_{L} (r_{S})$ with $r_{L} \gg r_{S}$. Periodic boundary conditions are assumed and integrals involving the $4N + 2$ localized $p$ orbitals are computed within the spirit of the Pariser–Parr approximation [4].

The evaluation of the band gap, the separation of the ground state and first excited state, proceeds as follows: (i) A set of appropriate one-electron orbitals is chosen to correspond to the eigenfunctions of a Hückel-type Fock operator. (ii) From these orbitals single configuration wave functions for the ground state and all singly-excited states with $k = \pi = 0$ (the same translational and spin symmetry as the ground state) are constructed and $\Delta E = E_{ex} - E_{gs}$ evaluated. (iii) Finally the excited state wave functions are symmetrized with respect to inversion and reflection through the origin of the chain and the final expression for $\Delta E(N)$ is determined.

The results of the procedure are summarized below with the details of the calculation relegated to the appendices. Let $p_{in} (i = 1, n = -N, N)$ denote the localized $p_{z}$ orbital on the $i$th carbon atom in the $n$th cell. Further, using the usual Hückel notation, let

\begin{align}
\langle p_{in} | h | p_{in} \rangle &= \alpha, \quad \text{if } i = i', n = n', \\
&= \beta_{S}, \quad \text{if } i \neq i', n = n', \\
&= \beta_{L}, \quad \text{if } n \neq n' \text{ and } n + i = n' + i'.
\end{align}

Finally, let

\begin{align}
X^{+}(j) &= \frac{1}{\sqrt{2}} \left[ \tilde{p}_{1}(j) + e^{ibj} \tilde{p}_{2}(j) \right], \\
X^{-}(j) &= \frac{1}{\sqrt{2}} \left[ e^{-ibj} \tilde{p}_{1}(j) - \tilde{p}_{2}(j) \right].
\end{align}

where $\delta = \pi / (2N + 1)$, and

\begin{equation}
\tilde{p}_{i}(j) = (2N + 1)^{-1/2} \sum_{n=-N}^{N} e^{i2\pi n / (2N + 1)} p_{in}, \quad -N < j < N.
\end{equation}

Then in the Hückel approximation the matrix elements of the Fock operator ($\langle X^{+}(j) | F_{H} | X^{+}(j) \rangle$) are given by (here $F_{H} = h$)

\begin{equation}
\begin{pmatrix}
\alpha - (\beta_{S} + \beta_{L}) \cos \delta j & ie^{-ibj}(\beta_{L} - \beta_{S}) \sin \delta j \\
-ie^{-ibj}(\beta_{S} - \beta_{L}) \sin \delta j & \alpha - (\beta_{S} + \beta_{L}) \cos \delta j
\end{pmatrix}.
\end{equation}

Note that $F$ is diagonal in $j$. Also observe that in the absence of bond alternation the $X^{\pm}(j)$ diagonalize $F$. For $\beta_{L} \neq \beta_{S}$ the matrix in (4) can be “rotated” to diagonal form. The resulting eigenvectors are

\begin{align}
\theta^{\pm}(j) &= \frac{1}{\sqrt{2}} \left( e^{-i\phi(j)} \tilde{p}_{1}(j) + e^{i\phi(j)} \tilde{p}_{2}(j) \right), \\
\theta^{-}(j) &= \frac{1}{\sqrt{2}} \left( -e^{-ibj} e^{i\phi(j)} \tilde{p}_{1}(j) + e^{ibj} \tilde{p}_{2}(j) \right),
\end{align}

and the corresponding eigenvalues

\begin{equation}
\langle \theta^{\pm}(j) | F_{H} | \theta^{\pm}(j) \rangle = \alpha \pm (\beta_{S}^{2} + \beta_{L}^{2} + 2\beta_{S} \beta_{L} \cos \phi)^{1/2},
\end{equation}

with (for $\beta_{S}, \beta_{L} < 0$)

\begin{equation}
tan 2\phi(j) = \frac{\beta_{L} - \beta_{S}}{\beta_{L} + \beta_{S}} \tan \delta j, \quad -\pi/2 \leq 2\phi(j) \leq \pi/2.
\end{equation}

From eqs. (2), (5), (6) we see that near the center of the Brillouin zone $X^{\pm}(j) \approx \theta^{\pm}(j)$ so that bond alternation has little effect on these eigenfunctions of $F_{H}$. However near the zone edges ($j = \pm N$), $2\phi(j) \rightarrow \pi / 2$ as $N \rightarrow \infty$ (for $|\beta_{L} - \beta_{S}| / (\beta_{L} + \beta_{S}) > 0$). Thus even for minimal bond alternation the $X^{\pm}(j)$ are poor approximations to the true eigenfunctions of $F_{H}$, the $\theta^{\pm}(j)$, for $|j|$ large.
Now according to step (ii) of the procedure, the ground electronic state of the polyene is approximated by

$$\psi_{gs} = A \prod_{-N < j < N} \theta^+(j) \alpha \theta^+(j) \beta. \quad (2.7)$$

The $2N + 1$ singlets corresponding to wave vector $k = 0$ are formed from the excitation $\theta^+(j) \to \theta^-(j)$ and are given by

$$\psi(j \to j^*) = A \prod_{m \neq j} \theta^+(m) \alpha \theta^+(m) \beta \theta^+(j) \beta^-(j) \frac{(\alpha \beta - \beta \alpha)}{\sqrt{2}}. \quad (2.8)$$

With $E(j \to j^*) \equiv \langle \psi(j \to j^*) | \hat{H} | \psi(j \to j^*) \rangle$ and $E_{gs} = \langle \psi_{gs} | \hat{H} | \psi_{gs} \rangle$, it is a standard result from Hartree–Fock theory [5] that

$$\Delta E(j) = E(j \to j^*) - E_{gs} \quad (2.9a)$$

$$\Delta E(j) = \langle \theta^-(j) | \hat{F} | \theta^-(j) \rangle - \langle \theta^+(j) | \hat{F} | \theta^+(j) \rangle - \langle J_{\theta^-}(j), \theta^+(j) - 2K_{\theta^-(j)}, \theta^+(j) \rangle, \quad (2.9b)$$

where $F$ is now the Fock operator corresponding to $\psi_{gs}$, $F = h + 2J - K$. In appendix A the requisite matrix elements in (9b) are evaluated and found to be

$$J_{\theta^-(j), \theta^+(j)} - 2K_{\theta^-(j), \theta^+(j)} = (2N + 1)^{-1} \sum_n \left[ C_4(n) - 2C_-(n) \right], \quad (2.10a)$$

$$\langle \theta^-(j) | J | \theta^-(j) \rangle - \langle \theta^+(j) | J | \theta^+(j) \rangle = 0, \quad (2.10b)$$

$$\langle \theta^+(j) | K | \theta^-(j) \rangle - \langle \theta^-(j) | K | \theta^-(j) \rangle = \sum_n C_{1,2}(n) \cos[2\delta j(n - \frac{1}{2}) - 2\phi(j)] \Phi(n), \quad (2.10c)$$

where

$$\Phi(n) = (2N + 1)^{-1} \sum_{0 < i < N} (2 - \delta_{i,0}) \cos[2\delta i(n - \frac{1}{2}) - 2\phi(i)] \quad (2.11a)$$

$$C_4(n) = \frac{1}{4} \sum_{i,j=1,2} C_{i,j}(n), \quad C_-(n) = \frac{i}{2} \sum_{i,j} (-1)^{i+j} C_{i,j}(n), \quad (2.11b,c)$$

$$C_{i,j}(n - n') = \langle p_{in} p_{jn} \rangle \langle 11_{12} \rangle \langle p_{in'} p_{jn'} \rangle \langle 2 \rangle. \quad (2.11d)$$

Inserting (10) into (9b) gives for $\Delta E(j \to j^*)$:

$$\Delta E(j) = 2(\beta_2^2 + \beta_2^4 + 2\beta_2 \beta_L \cos \delta j)^{1/2} - (2N + 1)^{-1} \sum_n \left[ C_4(n) - 2C_-(n) \right] \quad (2.12)$$

$$+ \sum_n C_{1,2}(n) \cos[2\delta j(n - \frac{1}{2}) - 2\phi(j)] \Phi(n).$$

From eq. (12) we see that $\Delta E(-j) = \Delta E(j)$. This degeneracy, which is a consequence of the periodic boundary conditions imposed in this model, can be lifted by two interrelated and essential equivalent approaches. In the first approach we observe that in the analogous linear system the true wave functions carry representations of the point group $C_{2v}$. If $P_R (P_L)$ represents the Wigner operator corresponding to reflection (inversion) through the origin, then it is easily shown (see appendix B) that

$$P_L \psi(j \to j^*) = P_R \psi(j \to j^*) = -e^{-2\delta j} \psi(-j \to -j^*). \quad (2.13)$$
Thus the appropriately symmetrized wavefunctions are

\[ \psi^+(j) = \frac{1}{\sqrt{2}} \psi(j \rightarrow j^*) + e^{i2\beta j} \psi(-j \rightarrow -j^*), \quad \psi^-(j) = \frac{1}{\sqrt{2}} e^{i2\beta j} \psi(j \rightarrow j^*) - \psi(-j \rightarrow -j^*), \] (2.13a, b)

where \( \psi^+(j)(\psi^-(j)) \) transforms like \( ^1B_u(1A_g) \).

The determination of the excitation energy (\( \Delta E_{\pm}(j) = \langle \psi^+(j) | H | \psi^\pm(j) \rangle - E_{\pm} \)) of these states requires the evaluation of the interaction matrix elements \( H_{j,\pm-j} = \langle \psi_{j^*}(-j^*) | H | \psi(-j \rightarrow -j^*) \rangle \). This calculation is performed in appendix C. The excitation energies now follows easily from eqs. (12), (13) and (C.5):

\[
\Delta E_{\pm}(j) = 2(\beta_S^2 + \beta_L^2 + 2\beta_S^2 \beta_L \cos 2\delta j)^{1/2} + \sum_n C_{1,2}(n) \cos [2\delta j(n - \frac{1}{2}) - 2\phi(j)] \Phi(n) \\
+ (2N + 1)^{-1} \sum_n [(2 \pm 2)C_-(n) - C_4(n)] \mp (2N + 1)^{-1} \sum_n \{\cos[4\delta j(n)]C_{1,1}(n) + C_{2,2}(n)\}/4 \\
+ \cos[\delta j(n - \frac{1}{2}) - 4\phi(j)] \{C_{1,2}(n) + C_{2,1}(-n)\}/4, \text{ for } 0 \leq j \leq N. \tag{2.14}\]

Equivalently eqs. (13), (14) could have been obtained by diagonalizing the \( 2 \times 2 \) matrix \( H(j) \) with matrix elements

\[ H_{j,\pm-j} = \Delta E(j), \quad H_{1,2}(j) = H_{2,1}(j). \]

The lowest excited state is obtained by minimizing eq. (14) with respect to \( j \) for fixed \( N \). However since we are dealing with an approximate model it is sufficient for our purposes to take \( j = N, \forall N \) and evaluate eq. (14) to order \( N^{-1} \) giving

\[
\Delta E_{\pm} = 2\beta_S + \beta_L + \sum_n C_{1,2}(n) \cos [2\delta j(n - \frac{1}{2}) - 2\phi(N)] \Phi(n) \\
+ (2N + 1)^{-1} \sum_n [(2 \pm 2)C_-(n) - (1 \pm 1)C_4(n)] \tag{2.15a}
\]

\[
= \Delta E_H + \Delta E_K + \Delta E_J. \tag{2.15b}
\]

This is the desired result. The first term in eq. (15) \( (\Delta E_H) \) is the usual Hückel expression for the band gap. The third term \( (\Delta E_J) \) which results from both Coulomb and exchange effects (see eq. 10 and appendix A) is easily seen to be of order \( \log N )/N \to N^{-1} \) since for large \( C_4(n) \approx 1/n. \) Thus for the \( N \to \infty \) limit only the first two terms need be considered. From eq. (10) it is seen that the second term in eq. (15) \( (\Delta E_K) \), i.e., the correction to the Hückel value of \( \Delta E \), is due entirely to the exchange interaction. It is important to note here that if \( X^2(j) \) had been used in eq. (9) rather than \( \theta^2(j) \), \( \Delta E_K \) (as well as \( \Delta E_H \)) would vanish.

Because of the complicated dependence of \( \Phi(n) \) on \( n, \Delta E_K \) (the exchange term) is most conveniently treated numerically. We consider two cases one with minimal bond alternation \( (\beta_S = -1.64 \text{ eV}, \beta_L = -1.52 \text{ eV}) \) [6] and a second with moderate \( (\beta_S = -2.39 \text{ eV}, \beta_L = -2.01 \text{ eV}) \) [7] bond alternation \( (r_B = 1.39 \text{ Å}, r_F = 1.42 \text{ Å} \) in both cases) [7]. The results (with \( C_{ij}(n) \) evaluated as in ref. [4]) are displayed in table 1. The salient feature of this table is that it is \( \Delta E_K \) rather than \( \Delta E_H \) which makes the dominant contribution to \( \Delta E \). Further while it is clear

<table>
<thead>
<tr>
<th></th>
<th>( \Delta E_K )</th>
<th>( \Delta E_H )</th>
<th>( \Delta E_{(\infty)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimal bond alteration</td>
<td>2.29</td>
<td>0.24</td>
<td>2.53</td>
</tr>
<tr>
<td>Moderate bond alteration</td>
<td>3.48</td>
<td>0.76</td>
<td>4.24</td>
</tr>
</tbody>
</table>

All energies are in electron volts. The relevant parameters for the two cases are given in the text. \( \Delta E_K \) extrapolated from data at \( N = 50, 100, 200 \).
from the table that $\Delta E$ can be made to agree reasonably well with experiment, it is not our purpose here to produce a quantitatively accurate model. Clearly the approximations involved in this calculation, the approximate evaluation of both 1- and 2-electron integrals, the use of approximate SCF orbitals and the neglect of configuration interaction argue against such a comparison. Rather what we hope to provide is an improved understanding of the origin of the band gap within a single configuration SCF (or SCF CI) [8] approach. The important point is that the alternant nature of the orbitals which is "inherent" in an SCF description of $\psi_{gs}$ leads to a pronounced change in the exchange contribution to the energy of the ground with respect to excited state (the Coulomb contributions are identical) which stabilizes $\psi_{gs}$ and hence the band gap (even for $\beta_L \approx 0$). Thus we conclude that (in a system where bond alternation can be expected to exist) the Hückel model cannot provide a realistic picture of the band gap since electron exchange plays an essential role in determining $\Delta E$.

3. Generalized exciton model

The exciton model of polyenes [2, 7] assumes that the polyene is made up of $N$ ethylene units ($n = 1, ..., N$) which in the lowest approximation are non-interacting. The excited state energies of the system are then just those of ethylene; for example, the first singlet state will be at an energy equal to that of the first singlet of ethylene, but will be $N$-fold degenerate. There will be, in addition, multiple excitation states, where two or more ethylenic units are excited. Let us represent the excited states of this "oriented gas" of ethylenic units by $|n_1, m_2, ...>/ = \phi_n^0 \phi_m^0 \phi_{p1}^0 \phi_{p2}^0 \ldots$ where $\phi_n^0$ represents the $nth$ excited state of the $n$th unit in the chain and $\phi_m^0$ represents the ground state of the $m$th unit in the chain. The total Hamiltonian will have matrix elements between states which differ by two excitations since the perturbation is a two body operator. In the simplest model of this system, we consider only one ethylene excited singlet state; then the Hamiltonian may be written in second quantized notation ($a_n^\dagger |0> = |n>$, etc.) as [9]

$$H = \epsilon_0 \sum_{n=1}^{N} a_n^\dagger a_n + \frac{1}{2} \sum_{n,m} H(n-m)(a_n^\dagger + a_n^\dagger)(a_m + a_m^\dagger)$$

(3.1)

where the fact that the ground state and double excitation states interact is explicitly taken into account by the term in $(a_n^\dagger a_m^\dagger + a_n a_m)$. We assume $H(0) = 0$. The usual approximation is to neglect "counter-rotating" terms in (1), i.e., those in $a_n^\dagger a_m$ and $a_n a_m^\dagger$. However, these are just the terms of interest to us. In the above, $\epsilon_0$ is the excitation energy of the ethylene unit and $H(n-m)$ is the interaction between units $n$ and $m$. Note that this model is highly simplified and is not expected to give accurate quantitative results. The model is only intended to show how terms in $N^{-1}$ can arise in the formula for the energy gap when multiple excitations are included.

The usual method of diagonalizing (1) is to assume that the operators $a_n$ and $a_n^\dagger$ are Boson operators. However, this is not exact and is in error to order $N^{-1}$. When one correct this and treats the counter-rotating terms correctly, terms in $N^{-1}$ arise in the energy of the first excited state.

First consider the statistics of the operators $a_n$ and $a_n^\dagger$. When the commutator $(a_n^\dagger a_n^\dagger - a_n a_n)$ operates on a state in which site $n$ is not excited we find

$$(a_n^\dagger a_n^\dagger - a_n^\dagger a_n)|0> = 0.$$

(3.2a)

However, when this commutator operates on a state in which $n$ is excited, we find

$$(a_n^\dagger a_n^\dagger - a_n^\dagger a_n)|n> = -|n>,$$

(3.2b)

since $a_n^\dagger|n> = 0$ (which means that two excitations cannot be on the same site). Thus, the commutator is not a pure number but is given by

$$[a_n, a_n^\dagger] = 1 - 2a_n^\dagger a_n.$$

(3.2c)
The other commutators in the system are Bose-like:

\[ [a_n, a_n^*] = 0, \quad n \neq m; \quad [a_n, a_m] = [a_n^*, a_m^*] = 0. \] (3.3a, b)

Thus, the statistics of these operators are neither Bose–Einstein nor Fermi–Dirac. In cases in which the average excitation number is small, the average of the commutator is given by

\[ \langle [a_n, a_n^*] \rangle = 1 - 2\langle a_n^* a_n \rangle = 1 + O(N^{-1}) \] (3.4)

Thus the usual approximation is to assume the excitations are Bosons. [In the case of nearest neighbor interactions, there is a transformation to Fermion operators. We discuss this below.] We will correct this to lowest order by transforming from excitation operators to Boson operators by a Holstein–Primakoff [10] transformation or, more generally, by a transformation introduced by Agranovich and Toscich [11]

\[ a_n = \left( \sum_{u=0}^{d_v} d_u b_n^{u*} b_n^u \right)^{1/2} b_n \approx (1 - b_n^{*2} b_n) \approx (1 - b_n^{*2}) \approx b_n^* (1 - b_n^{*2})^{1/2}, \] (3.5a, b)

where

\[ d_v = (-2)^v/(u + 1)! \] (3.5c)

The operators \( b_n, b_n^* \) satisfy Bose commutation rules and the usual methods can be used to treat the Hamiltonian. Substituting eq. (5) into eq. (1) results in

\[ H = H^{(0)} + H^{(1)} + H^{(2)}, \] (3.6a)

where

\[ H^{(0)} = \epsilon_0 \sum_n b_n^* b_n + \frac{1}{2} \sum_{n, m} H(n-m)(b_n + b_n^*)(b_m + b_m^*), \quad H^{(1)} = -\epsilon_0 \sum_n b_n^* b_n b_n^*, \] (3.6b, c)

\[ H^{(2)} = -\frac{1}{2} \sum_{n, m} H(n-m) b_n^* (b_n + b_n^*) b_m (b_m + b_m^*) + V, \] (3.6d)

where \( V \) is of sixth and higher order in the Boson operators \( (b, b^*) \). We will neglect \( V \) in what follows. We will now diagonalize \( H_0 \) and compute the first perturbation correction to the lowest states of \( H_0 \).

The zeroth order Hamiltonian, \( H_0 \), can be diagonalized by a transformation to new Boson operators after first introducing the wave vector representation (and assuming cyclic boundary conditions and \( N \) even)

\[ b_k = N^{-1/2} \sum_{n=1}^{N} b_n e^{ikn}, \quad k = 2\pi l/N, \quad l = -N/2, -N/2 + 1, \ldots, 0, 1, \ldots, N/2 - 1. \] (3.7)

Putting eq. (7) into eq. (6b) and noting that

\[ N^{-1} \sum_{n=1}^{N} e^{i(k-k')n} = \delta_{k, k'}, \] (3.8a)

\[ N^{-1} \sum_{n, m} H(n-m)e^{i(kn-k'm)} = \sum_{n-m} H(n-m)\cos k(n-m)\delta_{k, k'} \equiv \tilde{H}(k)\delta_{k, k'}, \] (3.8b)

we find

\[ H^0 = \sum_k \epsilon_0 b_k^* b_k + \frac{1}{2} \tilde{H}(k)(b_k + b_k^*)(b_{-k} + b_{-k}^*). \] (3.9)

We now introduce new Boson operators
\[ c_k = u_k b_k - v_k b_{-k}, \quad \tilde{c}_k = u_k b_k^* - v_k b_{-k}, \quad (3.10a, b) \]

where
\[ u_k^2 - v_k^2 = 1. \quad (3.10c) \]

By choosing \( u_k = \cosh \theta_k \), \( v_k = \sinh \theta_k \) and demanding that \( H_0 \) be diagonal in the new representation, we find
\[ \exp(-4\theta_k) = 1 + 2\tilde{H}(k)/\epsilon_0, \quad (3.11) \]

and the new energies \( \epsilon(k) \) given by
\[ \epsilon(k) = \epsilon_0(1 + 2\tilde{H}(k)/\epsilon_0)^{1/2}. \quad (3.12) \]

Note that to lowest order in \( \tilde{H}(k)/\epsilon_0 \) we have
\[ \epsilon(k) = \epsilon_0 + \tilde{H}(k), \quad (3.13) \]

which is the result of Simpson, quoted by Szabo et al. [1]. Thus, the inclusion of multiple excitation states changes the excitation energies in the Boson model in a simple manner. We have, in the new representation:
\[ H^0 = \sum_k \left[ \epsilon(k)c_k^* c_k + \frac{1}{2} \epsilon(k) + \frac{\epsilon_0}{2}(\epsilon(k) - \epsilon_0) \right], \quad (3.14) \]

and
\[ H^{(1)} = -(\epsilon_0/N) \sum_{k,l,m,n} b_k^* b_l^* b_m b_n \delta(k+l-m-n), \quad (3.15) \]
\[ H^{(2)} = -(2N)^{-1} \sum_{k,l,m,n} \tilde{H}(k)b_k^*(b_{-m} b_{+m})b_n(b_k + b_{-k})\delta(l+m-n-k). \quad (3.16) \]

At the level of \( H_0 \), the band gap is given by the minimum of \( \epsilon(k) \); there are corrections to this due to \( H^{(1)} \) and \( H^{(2)} \). Since \( H^{(1)} \) and \( H^{(2)} \) are out of order \( 1/N \), these corrections will, in general, contain terms of this order. Since the exciton model produces a band of closely spaced levels, first order perturbation theory is not appropriate. We will make no attempt to determine the value of the \( 1/N \) term within this model.

In the present model, the approximate eigenfunctions of the hamiltonian can be represented by wavefunctions in contrast to methods like the RPA. The transformation from the \( b \) operators to the \( c \) operators can be accomplished by a unitary transformation of the type \( c = \exp(S)b \exp(-S) \), where \( S \sim (b_k b_k - b_k^* b_k^*) \). The eigenstates of \( H^{(0)} \) are given by \( \exp(-S) \) operating on \( |0\rangle \), \( b_k^* |0\rangle \), \( \ldots \). The ground state of \( H^{(0)} \) then contains double, quadruple, etc., excitations of the \( b \) type (and hence of the \( c \) type), while the (odd) excited states contain single, triple, etc., excitations. These, then, represent wavefunctions of considerable configuration interaction. These wavefunctions, however, contain an unphysical feature: they allow multiple excitations on a site. This is removed by \( H^{(1)} \), \( H^{(2)} \) and \( V \) by introducing a "repulsion" between excitations. It is possible in one case to remove this effect completely. We turn to this next.

In the case in which \( H(n-m) \) is very short range (so that it is zero unless \( n \) and \( m \) are nearest neighbors), there is another procedure for treating the hamiltonian of eq. (1). This is to perform a Jordan-Wigner transformation [12] which is discussed in detail for this hamiltonian by Lieb, Schultz and Mattis [13]. By defining
\[ a_n = (-1)^{n-1} d_n, \quad a_n^* = (-1)^{n-1} d_n^*, \quad a_{n-1} = \sum_{j=1}^{n-1} a_j^* d_j, \quad (3.17) \]

we find that the \( d_n \) and \( d_n^* \) are Fermion operators. The hamiltonian becomes (in the case of periodic boundary conditions)
where
\[ a = \sum_{n=1}^{N} d_n^+ d_n = \text{number of excitations}. \quad J = H(\pm 1). \]

Transforming to running waves, \( \Delta k \) and \( \Delta k^* \), we find
\[
H = \sum_k (\epsilon_0 + 2J \cos k) d_k^+ d_k + \sum_k \sin k (d_k^* d_{-k}^* - d_k d_{-k})
+ (J/N) \sum_{k,k'} [d_k^+ d_{k'}^* e^{i(k+k'/2)} + e^{i(k-k')} d_k d_{k'}^* + \hbar c] [1 + \epsilon(1)\gamma].
\]

Neglecting the term proportional to \( J/N \) and taking multiple excitations into account by transforming to new Fermion operators, \( D_k \) and \( D_k^* \),
\[
D_k = \cos \phi_k d_k - \sin \phi_k d_k^*, \quad D_k^* = \cos \phi_k d_k^* - \sin \phi_k d_k,
\]
with \( \phi_k \) chosen so that the \( D_k \) operators refer to non-interacting Fermion modes to \( O(1/N) \) we find (aside from a constant term)
\[
H = \sum_k [(\epsilon_0 + 2J \cos k)^2 + 4J^2 \sin^2 k]^{1/2} D_k^* D_k + H^{(1)},
\]
with \( H^{(1)} \propto J/N \) in the space of even number of excitations (which includes the ground state). We thus have a closely spaced band of energy levels with the lowest energy (for \( J < 0 \) at \( k = 0 \)) corresponding to a band gap of \( \sim \epsilon_0 - 2|J| \). The corrections to the band gap are again of order \( 1/N \) (corresponding to corrections to the ground state energy).

We have presented the latter calculation (using a Fermion representation) to show that the unphysical property of the Boson representation (i.e., that more than one excitation can reside on a site) which is removed to low order by the Holstein-Primakoff transformation is not responsible for the \( 1/N \) terms. The Fermion representation does not have this unphysical feature and has a \( 1/N \) correction also. Unfortunately, the Fermion representation is impossible for longer range interactions, so it cannot be used in general.

The use of periodic boundary conditions is unnecessary in the Fermion representation. In fact, the free end boundary condition is a bit easier to work with.

In conclusion to this section, we reiterate our major finding. It is simply that treating the exciton model fully (i.e., correcting for statistics and multiple excitations) formally leads to a \( 1/N \) dependence on the band gap. We have not attempted to compute the numerical value of this correction.

4. Comparison with previous work

As mentioned previously, theoretical determinations of \( \Delta \varepsilon \) have been quite numerous. It is therefore important to juxtapose the simplified model calculations of this study with previous work in this area and with each other. We begin by establishing the relationship between the widely used Pariser-Parr (PP) integral approximation [4] (which was used in developing the BAM approximation) and the exciton model as described in section 3.

Consider a partitioning of the space of single excitons into vertical (denoted \( \psi(j \rightarrow j^*) \)) and charge transfer (denoted \( \psi(j \rightarrow j^*) \)) excitons. We now show that only the space of vertical excitons is non-interacting with the ground state (\( \psi_{gs} \)). Let
where in the spirit of the PP approximation

\[
\phi_i^\pm = \frac{1}{\sqrt{2}} (p_{1l} \pm p_{2l}), \quad \phi_i^- = \frac{1}{\sqrt{2}} (p_{1l} - p_{2l}).
\]

Then

\[
\langle \psi_{gs} | H' | \psi(j \to j^*) \rangle = \sqrt{2} \langle \phi_j^\pm | F | \phi_j^- \rangle
\]

where \( F \) is the Fock operator corresponding to \( \psi_{gs} \) (i.e., constructed from the set \( \{ \phi_i^\pm \}_{i=N}^{N} \)). For \( j = j' \), vertical excitons:

\[
\langle \phi_j^\pm | F | \phi_j^- \rangle = \frac{1}{2} (\alpha - \alpha + \beta_N - \beta_N) + \frac{1}{4} \sum_n [C_{1,1}(n-j) - C_{2,2}(n-j) + C_{1,2}(n-j) - C_{2,1}(n-j)]
\]

\[
- \frac{1}{4} [C_{1,1}(0) - C_{2,2}(0) + C_{1,2}(0) - C_{2,1}(0)] = 0,
\]

where we have assumed periodic boundary conditions, so as to cancel \( C_{1,2}(-n) \) with \( C_{2,1}(-n) \). However for \( j \neq j' \), charge transfer excitons:

\[
\langle \phi_j^\pm | F | \phi_j^- \rangle = \frac{1}{2} \beta_L + 0 - 0 = \frac{1}{2} \beta_L, \quad \text{if} \ |j - j'| = 1, \quad \frac{1}{2} \beta_L, \quad \text{otherwise}.
\]

So that only the space of vertical excitons is strictly non-interacting with the ground state.

We now compare the interaction of two single (vertical) excitons with that of the ground state and a double (vertical) exciton. Let \( \psi^I(j, j') \) and \( \psi^II(j, j') \) denote the two doubly excited singlets arising from the promotion \( (\phi_j^+ \to \phi_j^-, \phi_j^+ \to \phi_j^-) \) and \( (\psi^I(j), \psi^II(j)) \) representing the state corresponding to the coupling of two singlet (triplet) excitons. Then since the \( \phi^\pm \) form an orthonormal set:

\[
H^I(j, j') \equiv \langle \psi_{gs} | H' | \psi^I(j, j') \rangle = 2 \langle (\phi_j^+ \phi_j^-)(1) 1/2 | (\phi_j^+ \phi_j^-)(2) \rangle
\]

\[
- \langle (\phi_j^+ \phi_j^-)(1) 1/2 | (\phi_j^- \phi_j^-)(2) \rangle, \quad H^II(j, j') \equiv \langle \psi_{gs} | H' | \psi^II(j, j') \rangle = \sqrt{3} \langle (\phi_j^+ \phi_j^-)(1) 1/2 | (\phi_j^- \phi_j^-)(2) \rangle,
\]

\[
H^III(j, j') \equiv \langle \psi_{gs} | H' | \psi^III(j, j') \rangle = \sqrt{3} \langle (\phi_j^+ \phi_j^-)(1) 1/2 | (\phi_j^- \phi_j^-)(2) \rangle - \langle (\phi_j^- \phi_j^-)(1) 1/2 | (\phi_j^- \phi_j^-)(2) \rangle.
\]

Here \( \psi^I(j, j') \) is defined as in definition (2.8) with \( \phi_j^\pm \) replacing \( \theta^I(j) \).

Using the PP approximation to evaluate eq. (5) we find

\[
H^I(j, j^*) = 2C_{{-}}(j - j^*), \quad H^II(j, j') = 0, \quad H^III(j, j') = 2C_{{-}}(j - j'),
\]

so that only the double excitation arising from single excitons interacts with the ground state and it has the same interaction energy as the direct interaction of its constitutive single excitations. Finally note that for \( j - j^* \) large, \( C_{-}(j - j^*) \approx -1/j - j^* \) and therefore behaves like a dipolar interaction.

Thus we see that it is possible to understand the exciton hamiltonian, eq. (3.1), in terms of the PP approximation. It is also worth observing here that a more exact numerical calculation than that of section 3, taking into account charge-transfer excitons and treating exciton statistics exactly, has been performed within the PP approximation by Pople and Walsley [7]. However they did not investigate the \( N \) dependence of \( \Delta E \) and neglected the interaction of the charge transfer excitons with the ground state and the effects of multiple excitations.
In contrast to the vertical exciton model (VEM) the BAM method includes the effects of change transfer excitations since it considers the entire space of single excitations consistent with its basis set. However, unlike the situation in the VEM, it is easily verified that BAM excitons interact with the ground state through the exchange part of the Fock operator.

It is possible to circumvent the problem of BAM excitons interacting with $\psi_{ee}$ by substituting the true SCF orbitals for BAM orbitals. However, as was alluded to at the end of section 2, using ground state SCF orbitals to construct both ground and excited state prejudices the calculation in favor of $\psi_{ee}$, thus increasing $\Delta E$. Thus it is preferable, whenever possible, to reduce this prejudice by allowing for configuration interaction (CI) in the space of singly excited wavefunctions. These SCF CI calculations which, as we mentioned earlier, motivated the BAM calculation, have been performed numerically (using the PP approximation) by several authors [1,3]. These calculations exhibit the increase in $\Delta E$ above its Hückel value, which we have attributed to an exchange interaction. However, these calculations also illustrate a failure of the BAM. From eq. (15) it can be seen that the BAM predicts a $\Delta E(N)$ of the form $\Delta E(N) = \Delta E_0 - |C|N^{-1}$. The numerical work of Szabo et al. [1] on comparatively small, linear polyenes does not show this $N^{-1}$ dependence within the single configuration approximation. In fact, these authors show that only when a full CI is performed on the space of single excitations does an $N^{-1}$ dependence, with the experimentally observed positive sign, emerge. A possible explanation for this discrepancy is that the orbitals in the bond alternation model given above are not self consistent; in fact the first excited state has an exchange interaction with the ground state. When the orbitals are made self consistent, the ground state energy will be lowered, which may explain the inconsistency in the $1/N$ term between our model and the previous numerical work.

Finally we remark that the calculations considered here are basically extensions of the restricted Hartree-Fock [5] formalism. Thus the possible effects of spin density waves which have been discussed by other authors [14] have been ignored here.

5. Summary

We have considered the first strongly allowed dipole transition in an infinite polyene using simple exciton and single configuration SCF models. In the single configuration SCF model we demonstrated that the alternant nature of the SCF orbitals (i.e., the single bond–double bond character of the Mulliken population) leads to a large exchange contribution to $\Delta E_0$ even when the degree of bond alternation (as measured by $|\beta_g - \beta_{g'}|$) is minimal. In the exciton model we showed how the non Bose–Einstein statistics of the excitons could be included through the mechanism of the Holstein–Primakoff and Jordan–Wigner transformations. The use of these transformations when coupled with the inclusion of multiple excitation leads to a $\Delta E_1 \neq 0$ contribution to $\Delta E(N)$ which had been previously believed to be lacking in the exciton model. Finally we discussed the assumptions inherent in the models and considered this work in light of previous investigations in this area.

Appendix A

In this appendix the expressions given in eq. (2.10) are developed. The zero differential overlap (zdo) approximation, will be employed throughout, i.e.,

$$((p_{ia}p_{jb})(1)|1,2\rangle(p_{kc}p_{ld})(2)) \propto \delta_{ij}\delta_{kl}\delta_{ab}\delta_{cd}$$  \hspace{1cm} (A.1)

or equivalently $(p_{ia}p_{jb})(1) \propto \delta_{ij}\delta_{ab}$.

Thus from definitions (2.3), (2.5a, b):
\[
(\theta^+(m)\theta^-(j))(1) = (2N+1)^{-\frac{1}{2}} \sum_n e^{i2\delta(j-m)n} \times [e^{i[\phi(m)-\phi(j)]} e^{-i\delta j^2} p_{1n}^2(1) + e^{-i\delta m} e^{i[\phi(j)-\phi(m)]} p_{2n}^2(1)],
\]

(A.2a)

\[
(\theta^+(m)\theta^+(j))(1) = (2N+1)^{-\frac{1}{2}} \sum_n e^{i2\delta(j-m)n} \times [e^{i[\phi(m)-\phi(j)]} p_{1n}^2(1) + e^{i[\phi(j)-\phi(m)]} e^{-i\delta(j-m)} p_{2n}^2(1)],
\]

(A.2b)

\[
(\theta^-(m)\theta^+(j))(1) = (2N+1)^{-\frac{1}{2}} \sum_n e^{i2\delta(j-m)n} \times [e^{i[\phi(j)-\phi(m)]} e^{i[\phi(j)-\phi(m)]} p_{1n}^2(1) + e^{i[\phi(j)-\phi(m)]} e^{-i\delta(m-j)} p_{2n}^2(1)].
\]

(A.2c)

Thus

\[
J_{\theta^-(j)\theta^+(j)} = \langle(\theta^-(j)\theta^-(j))(1)1|l_{12}\rangle(\theta^+(j)\theta^+(j))(2)
\]

(A.3)

\[
= (2N+1)^{-\frac{1}{2}} \sum_{n,n'} \frac{1}{2} \langle [p_{1n}^2(1) + p_{2n}^2(1)] 1|l_{12}\rangle [p_{1n}^2(2) + p_{2n}^2(2)] = (2N+1)^{-\frac{1}{2}} \sum_n C_+(n),
\]

(K\theta^-(j)\theta^+(j) = \langle(\theta^-(j)\theta^+(j))(1)1|l_{12}\rangle(\theta^-(j)\theta^-(j))(2)
\]

(A.4)

\[
= (2N+1)^{-\frac{1}{2}} \sum_{n,n'} \langle [-p_{1n}^2(1) + p_{2n}^2(1)] 1|l_{12}\rangle [-p_{1n}^2(2) + p_{2n}^2(2)] = (2N+1)^{-\frac{1}{2}} \sum_n C_-(n),
\]

\[
\langle\theta^-(j)|\theta^-(j)\rangle = \sum_m \langle(\theta^-(j)\theta^-(j))(1)1|l_{12}\rangle(\theta^+(m)\theta^+(m))(2)
\]

(A.5)

\[
= (2N+1)^{-\frac{1}{2}} \sum_{n,m} C_+(n) = \sum_n C_+(n) = \langle\theta^+(j)|\theta^+(j)\rangle,
\]

\[
\langle\theta^+(j)|K\theta^+(j)\rangle = \sum_m \langle(\theta^+(j)\theta^+(j))(1)1|l_{12}\rangle(\theta^+(m)\theta^+(m))(2)
\]

(A.6a)

\[
= (2N+1)^{-\frac{1}{2}} \sum_{n,n'} \langle [e^{i2\delta(j-m)n}(n-n')] C_{1,1}(n-n') + C_{2,2}(n+n')
\]

\[
\pm C_{1,1}(n-n') e^{i2\delta(j-m)} e^{i\delta(j-m)} + C_{2,1}(n-n') e^{i2\delta(j-m)} e^{i\delta(j-m)}]
\]

(A.6b)

\[
= (2N+1)^{-\frac{1}{2}} \sum_{n,m} e^{i2\delta(j-m)n} [C_{1,1}(n) + C_{2,2}(n)]
\]

\[
\pm C_{1,2}(n) e^{i2\delta(j-m)} e^{i\delta(j-m)} + C_{2,1}(n) e^{i2\delta(j-m)} e^{i\delta(j-m)}
\]

(A.6c)

Therefore

\[
\langle\theta^+(j)|K\theta^+(j)\rangle - \langle\theta^-(j)|K\theta^-(j)\rangle = (2N+1)^{-\frac{1}{2}} \sum_{n,m} e^{i2\delta(j-m)n}
\]

\[
\times \{ e^{i2\delta(j-m)} C_{1,2}(n) + e^{i2\delta(j-m)} C_{2,1}(n) \}
\]

(A.7a)

\[
= (2N+1)^{-\frac{1}{2}} \sum_{n,m} \cos[2\delta(m-j)(n-\frac{1}{2})] + 2\phi(j) - 2\phi(m)] [C_{1,2}(n) + C_{2,1}(n)] / 2,
\]

(A.7b)
where the fact that \(-2\phi(-j) = 2\phi(j)\) and 
\(C_{ij}(n) = C_{ji}(-n)\). Finally, using the trigonometric identity for the sum of two cosines reduces \((A.7b)\) to

\[
(2N + 1)^{-1} \sum_{m \geq 0} \cos[2\delta(n - \frac{1}{2}) - 2\phi(j)] \cos[2\delta m(n - \frac{1}{2}) - 2\phi(m)] C_{1,2}(n)
\]

\[
= \sum_n C_{1,2}(n) \cos[2\delta(n - \frac{1}{2}) - 2\phi(j)] \phi(n),
\]

where \(\phi(n)\) is defined by eq. (2.11a).

\section*{Appendix B}

Here the transformation properties of \(\theta_s(j)\) and \(\psi(j \rightarrow j^*)\) under the operations of reflection and inversion through the origin of the chain are developed.

Let \(P_R(P_I)\) denote the Wigner operator corresponding to reflection (inversion). Then

\[
P_R P_{i,n} = P_{\ell,n}, \quad \text{where } \ell = 2, \ell' = 1; \quad P_1 P_{i,n} = P_{\ell,-n}.
\]

Therefore from definition (2.3)

\[
P_R P_{i} = P_{i} = -P_{I} P_{i} = P_{i} = -P_{I}.
\]

so that from definition (2.5)

\[
P_R \theta_s(j) = \frac{1}{2} \left[ e^{-i\theta(j)} p_2(-j) + e^{i\theta(j)} e^{i\theta(-j)} p_1(-j) \right]
\]

\[
= e^{i\theta(j)} \frac{1}{2} \left[ p_1(j) e^{i\theta(j)} + e^{-i\theta(-j)} p_2(-j) \right]
\]

\[
= e^{i\theta(j)} \theta^*(j) = -P_{I} \theta^*(j),
\]

where in going from (B.3a) to (B.3b) the fact that \(\phi(j) = -\phi(-j)\) was used.

Similarly we find

\[
P_R \theta^-(j) = -\theta^-(j) e^{-i\theta(j)} = -P_{I} \theta^-(j).
\]

From eqs. (B.3) and (B.4) the transformation properties of the \(\psi(j \rightarrow j^*)\) can be deduced. Eq. (2.8) shows that apart from fully occupied, and therefore totally symmetric, orbitals the spatial part of \(\psi(j \rightarrow j^*)\) can be written as

\[
\psi(j \rightarrow j^*) \sim \theta^+(j) \theta^+(j) \theta^-(j) \theta^-(j).
\]

so that

\[
P_R \psi(j \rightarrow j^*) \sim \theta^+(j) \theta^+(j) \theta^-(j) \theta^-(j)e^{i\theta(j)} \sim \theta^+(j \rightarrow j^*) \sim P_{I} \psi(j \rightarrow j^*).
\]

\section*{Appendix C}

Here the matrix element \((\psi(j \rightarrow j^*)|H|\psi(j' \rightarrow j'^*)) = H_{j,j'}\) as used in section 2 is evaluated. Since \(\psi(j \rightarrow j')\) has the form

\[
\psi(j' \rightarrow j^*) \sim \theta^+(j') \alpha \theta^+(j') \beta \theta^+(j) \theta^-(j \rightarrow j') (\alpha \theta - \beta \alpha) / \sqrt{2},
\]

(C.1)
and the spatial part of \( \theta^e(q) \) form an orthonormal set, one easily finds that

\[
H_{j,j'} = 2 \langle \theta^e(q) | \theta^e(q) \rangle (1) | j | r_{12} (| \theta^e(q) \rangle | \theta^e(q) \rangle (2)) - \langle i \theta^e(q) | i \theta^e(q) \rangle (1) | j | r_{12} (| \theta^e(q) \rangle | \theta^e(q) \rangle (2)) = 2I_1 - I_2
\]

(C.2)

\( I_1 \) and \( I_2 \) can be evaluated using eq. (A.2). In particular

\[
I_1 = (2N + 1)^{-1/2} \sum_{n,n'} \langle -e^{-i\theta/2} p_{1n}^2(1) + e^{i\theta/2} p_{2n}^2(1) | j | r_{12} [-e^{-i\theta} p_{1n'}^2(2) + e^{-i\theta} p_{2n'}^2(2)] \rangle
\]

\[
= (2N + 1)^{-1} e^{i\theta(\theta-j)} \sum_n \mathcal{C}_-(n)
\]

(C.3)

while

\[
I_2 = (2N + 1)^{-1} \sum_{n,n'} \langle e^{i\theta(\theta-j)(n-n')} \langle [e^{-i\theta} - e^{i\theta}] p_{1n}^2(1) + e^{i\theta} - e^{-i\theta}] p_{2n}^2(2) \rangle | j | r_{12} [e^{-i\theta} - e^{i\theta}] p_{1n}^2(1) + e^{i\theta} - e^{-i\theta}] p_{2n}^2(2) \rangle
\]

\[
= (2N + 1)^{-1} e^{i\theta(\theta-j)} \sum_n \mathcal{C}_+(n)
\]

\[
= e^{i\theta(\theta-j)} (2N + 1)^{-1} \sum_n \left[ \cos[2\theta(\theta-j)n] [C_{1,1}(n) + C_{2,2}(n)] / 4 \\
+ \cos[2\theta(\theta-j)n(n - 1/2 + 2n)] [C_{1,2}(n) + C_{2,1}(-n)] / 4 \right].
\]

Thus

\[
H_{j,j'} = e^{i\theta(\theta-j)} (2N + 1)^{-1} \sum_n \left[ 2C_-(n) - \cos[2\theta(\theta-j)n] [C_{1,1}(n) + C_{2,2}(n)] / 4 \\
- \cos[2\theta(\theta-j)n(n - 1/2 + 2n)] [C_{1,2}(n) + C_{2,1}(-n)] / 4 \right],
\]

which is the desired result.

References

[6] Values obtained from the equation: \( \beta = 68.5 \exp(-r/a) \) (in eV) where \( a = 0.3737 \) A using \( r_S \) and \( r_L \) in text: see L. Salem, Molecular Orbital Theory of Conjugated Systems (Benjamin, New York, 1966).