VIBRATIONAL SPECTRA OF DOUBLE HELICAL MOLECULES WITH C2 SYMMETRY (U)
Purdue Univ. Lafayette Ind. Dept. of Physics
Xin-Min Hua et al. 10 Mar 98 N00014-86-K-0252
Vibrational Spectra of Double Helical Molecules with $C_2$ Symmetry

X.M. Hua and E.M. Prohofsky

The existence of an additional symmetry element allows one to reduce the size matrix needed to calculate the vibrational normal modes. In this paper we detail how to reduce the matrix appropriate to an alternating DNA polymer which has a $C_2$ symmetry about the x-axis.
Vibrational Spectra of Double Helical Molecules with $C_2$ Symmetry

X.M. Hua and E.W. Prohofsky
Department of Physics
Purdue University
West Lafayette, IN 47907

Abstract

The existence of an additional symmetry element allows one to reduce the size matrix needed to calculate the vibrational normal modes. In this paper we detail how to reduce the matrix appropriate to an alternating DNA polymer which has a $C_2$ symmetry about the x-axis.
I. Introduction

A considerable number of observations both by infrared absorption and Raman scattering have been made on samples of double helical DNA. Many of these observations are on DNA polymers which have or come close to having structural symmetries. We have carried out helix lattice normal mode calculations for a number of repeating DNA polymers. The agreement between the theoretical predictions and experimental observations has been quite good and a reasonably good set of force has been refined for such calculations. Most such calculations have been on the simplest repeating homopolymers. There are a considerable number of observations on alternating homopolymers and we have also calculated the helix lattice modes for a number of these cases. The alternating polymers have a unit cell consisting of two base pairs which is twice as big as the simplest homopolymers. A true helix can be generated by applying a screw axis operation to the unit cell many times.

In applying the helix lattice formalism one gets the normal modes by diagonalizing a matrix which has dimension of 3 times the number of atoms in the unit cell. The matrix for alternating DNA is thus a square array twice as big in both columns and rows. The structure of the double helix is such that the alternating DNA has an additional symmetry, a $C_2$ operation, that is not an exact symmetry in the simpler homopolymers. The point of this paper is to detail how one can make use of this additional symmetry to reduce the size matrix needed to be diagonalized for the alternating homopolymers.
The method has been used to reduce the calculation for z-form poly(dG-dC)·poly(dG-dC) and it has led to a reduction in machine time by about an order of magnitude in the calculation of the entire normal mode spectrum. In an attempt to describe the method for incorporating the additional $C_2$ symmetry in as general a manner as possible, we do not use specific details of the internal force field of the alternating polymers. Many details specific to DNA will be left to a later publication.

The paper is arranged as follows: in section II we discuss the general features of helical structure and establish the dynamical matrix; Section III is devoted to a $C_2$ symmetry discussion and section IV develops scattering selection rules; conclusions are in section V.
II. General Properties of Helical Structure

Let us consider an infinite chain molecule with helical structure such that each unit cell is transformed geometrically into the next by the operation \((R|p)\), where \(R\) is the rotation through an angle \(\psi\) about the helical axis \(\hat{z}\):

\[
R = \begin{pmatrix}
\cos \psi & -\sin \psi & 0 \\
\sin \psi & \cos \psi & 0 \\
0 & 0 & 1
\end{pmatrix}, \quad (2.1)
\]

and \(p = pz\) is the translation along \(\hat{z}\). If the unit cell is the smallest repeating unit of the molecule, \(\psi\) is the pitch angle and \(p\) is the pitch.

The structure of the molecule is invariant under operation \((R|np)\), for if \(X^m_i\) is the equilibrium position of the \(i^{th}\) atom in the \(m^{th}\) unit cell, then

\[
(R^n|np)X^m_i = R^nX^m_i + npz = X^{m+n}. \quad (2.2)
\]

The choice of a unit cell is not unique, usually one takes the unit cell as the smallest repeating unit or as the primitive lattice cell of the corresponding space group. With the choice of a unit cell of \(S\) atoms one gets \(3S\) vibrational bands. Each band has vibrational modes that differ by having a phase shift from one unit cell to the next. The number of phase shifts is equal to the number of unit cells in the macroscopic sample, approaching a continuum of phase shifts as the number of cells in the sample approaches infinity. The different phase shifts are equivalent to different waves of varying wavelengths. Any modes with varying phase are infrared and Raman inactive. In cases where the original symmetry of the system is broken, inactive modes may become active.
In our helix lattice calculations the unit cell is chosen as the smallest repeating unit \(1,2,5,6\) as this gives rise to the smallest matrix for solution. In the harmonic approximation the potential energy of the molecule is

\[
U = \frac{1}{2} \sum_{\mathbf{mn}} \sum_{ij} q^m_i \cdot D_{ij} \cdot q^m_j ,
\]

(2.3)

where \(m,n\) label the unit cells and \(i,j\) label the atoms in a unit cell;

\[q^m_i = \sqrt{m} \, u^m_i\]

is the mass weighted cartesian displacement where \(m_i\) is mass of atom \(i\) and \(u^m_i\) is the displacement of atom \(i\) in \(m\)th cell from its equilibrium position \(X^m_i\). \(D_{ij}\) is the dynamical matrix containing the force constants between the atom at \(X^m_i\) and the atom at \(X^m_j\). The kinetic energy is

\[
T = \frac{1}{2} \sum_{\mathbf{m}i} \dot{q}^m_i \cdot q^m_i,
\]

(2.4)

And the eqns. of motion could be written as

\[
\gamma_{mn} q^m_i \cdot D_{ij} \cdot q^m_j = \omega^2 q^m_i ,
\]

(2.5)

where the harmonic behavior \(e^{-i\omega t}\) is assumed.

By using Born-von Karman boundary condition and taking into account the helical symmetry we impose a solution which is a mode of no phase shifts:

\[
q^m_i = \frac{1}{\sqrt{2\pi}} \int d\nu R^m q^m_i(\nu)e^{im\nu},
\]

(2.6)

where

\[
q^m_i(\nu) = \frac{1}{\sqrt{2\pi}} \sum_{m'} R^{m'} q^m_i e^{-im'\nu}.
\]

(2.7)
satisfies the secular eqn.

\[ \dot{j}\ D_{ij}^{\pm} q_j^{\pm}(t) = \omega_{ij} q_i^{\pm}(t), \]  

(2.8)

where

\[ \dot{j}\ D_{ij}^{\pm}(\theta) = F_{ij} R_{mn}^{\pm m \mp n} e^{-i(m-n)\theta}, \]  

(2.9)

provided that \( R_{mn}^{\pm} \) is only the function of \( m-n \) for a certain pair of \((i,j)\). In other words, \( R_{mn}^{\pm} \) is unchanged under operation \((E|p)\) where \( E \) is identity. The reality of \( q^{\pm m} \) implies

\[ q_i^{\pm}(0) = q_i^{\mp}(0), \]  

(2.10)

and the reality of the potential energy implies \( D_{ij}^{\pm} \) is hermitian:

\[ D_{ij}^{\pm} = D_{ji}^{\pm}. \]  

(2.11)

For the molecule possessing no symmetry operation which takes \( \theta \) to \( -\theta \), eqn(2.10) is simply the statement of double degeneracy due to time reversal.\(^7\)

It has been pointed out\(^1,2,5,6\) that the free center of mass motion of the helical molecule will occur at \( \theta = 0, \mp \psi \). The eigenvectors of the free motion can be obtained from eqn.(2.7) and are listed in table 1. Free center of mass motion is independent of force constants and gives restrictions on dynamical matrix \( D_{ij} \).
\[ \sum_j \sqrt{m_j} D_{ij} = 0, \quad \theta = 0 \quad (2.12a) \]

\[ \sum_j \sqrt{m_j} \left( D_{ij} x^j - D_{ij} x^j \right) = 0, \quad \theta = 0 \quad (2.12b) \]

\[ \sum_j \sqrt{m_j} \left( D_{ij} + D_{ij} \right) = 0, \quad \theta = + \psi \quad (2.12c) \]

where \( D_{ij} = \left( \overrightarrow{D}_{ij} \right)^{\alpha \beta} \), \( x^j = \left( \overrightarrow{x}^j \right)^{\alpha} \), and \( \alpha, \beta = x, y, z \); the factor \( \sqrt{m_j} \) comes from the fact that \( q_i \) is mass weighted.

From a classical point of view, the induced electrical dipole moment \( \mu \) of the molecule is given by

\[ \mu = \sum_{m,i} \frac{e_i u_i^{\text{m}}}{m_i} = \sum_{m,i} \frac{e_i q_i^{\text{m}}}{\sqrt{m_i}} \quad (2.13) \]

where \( e_i \) is the atomic partial charge \(^9\) assigned to atom \( i \). For the free center of mass motion, for example the longitudinal translation \( u_i^{\text{m}} = z \), there is no induced dipole and eqn.(2.13) implies \( \sum_i e_i = 0 \), i.e., the unit cell of the molecule must be neutral. It is well-known \(^{10}\) that the dissolved DNA molecule is not neutral. To allow free motion of DNA in solution, the counter ions must be included in the DNA to make the unit cell neutral. Otherwise the free motion of DNA in solution is not allowed.
III. The $C_2$ Symmetry

The standard z-form DNA poly(dG - dC) - (dG - dC) has a $C_2$ symmetry axis which is perpendicular to the helix axis and is usually taken to be the $\hat{x}$ axis in cartesian coordinates. In such a coordinate system an atom in the $m$th cell on one strand has equilibrium position $(x, y, z)$, then the corresponding atom in $-m$th cell on the antiparallel strand locates at $(x, -y, -z)$. For those DNA molecules without $C_2$ symmetry the backbone atoms still have approximate $C_2$ symmetry and $\hat{x}$ is picked in the same way.

For this $C_2$ symmetry discussion we rewrite the secular eqn. (2.8) as follows:

$$\begin{align*}
\sum_{i} \hat{D}_{ij} \hat{q}_{j} + \sum_{k} \hat{D}_{ik} \hat{q}_{k} &= \omega^2 \hat{q}_{i}, \\
\sum_{j} \hat{D}_{kj} \hat{q}_{j} + \sum_{l} \hat{D}_{kl} \hat{q}_{l} &= \omega^2 \hat{q}_{k},
\end{align*}$$

(3.1)

such that the dynamical matrix $\hat{D}$ is

$$\begin{pmatrix}
\hat{D}_{ij} & \hat{D}_{ik} \\
\hat{D}_{kj} & \hat{D}_{kl}
\end{pmatrix},$$

(3.2)

where $i, j = 1, ..., S_i$ label the atoms on one strand; $k = i + S_i, l = j + S_i$, and $k, l$ label the atoms on the antiparallel strand in such a way that atom $k$ is the same atom as atom $i$ but on a different strand. $S_i$ is the number of atoms in one strand. $2S_i - S$ - number of atoms in a unit cell. It has been pointed out by Linz' that the operations that do not leave invariant provide relationships between the corresponding matrix elements. In our case the $C_2$ operation inverts the orientation of DNA and hence takes $\theta$ to $-\theta$, then the following relations apply:
\[ \mathbf{D}^{-1}_{ij}(0) = R_{2x}^{-1} \mathbf{D}^{-1}_{k\ell}(0) R_{2x} = R_{2x}^{-1} \mathbf{D}^{-1}_{k\ell}(0)^* R_{2x}, \] (3.3)

\[ \mathbf{D}^{+}_{ij}(0) = R_{2x}^{-1} \mathbf{D}^{+}_{k\ell}(0) R_{2x} = R_{2x}^{-1} \mathbf{D}^{+}_{k\ell}(0)^* R_{2x}, \]

where

\[ R_{2x} = R_{2x}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \]

and \( \mathbf{D}^{-1}_{ij}(-\theta) = \mathbf{D}^{-1}_{ij}(0)^* \) follows immediately from (2.9).

Eqn. (3.3) implies that the solution of secular eqn. (3.1) takes the form of

\[ q_k(\theta) = R_{2x} q_i(\theta)^* e^{i\phi}, \] (3.4)

where \( \phi \) is a common arbitrary phase angle, and could be chosen to be zero, such that the real part of eigenvector is symmetric and the imaginary part is antisymmetric under \( C_2 \) operation. By imposing the solution (3.4) with \( \phi = 0 \) and separating Re \( \hat{q}_i \) and Im \( \hat{q}_i \) as independent variables the secular eqn. (3.1) becomes

\[ \mathbf{G}_{ij} \mathbf{q}_i + \mathbf{G}_{ij} \mathbf{q}_i = \mathbf{w}^{2\mathbf{R}} q_i, \] (3.5)

where

\[ \mathbf{G}_{ij} = \text{Re} \left[ \mathbf{D}_{ij}^{+} \right] + R_{2x}^{-1} \mathbf{D}_{ij}^{+} R_{2x}, \]

\[ \mathbf{G}_{ij} = -\text{Im} \left[ \mathbf{D}_{ij}^{+} \right] + R_{2x}^{-1} \mathbf{D}_{ij}^{+} R_{2x}, \]

\[ \mathbf{G}_{k\ell} = \text{Im} \left[ \mathbf{D}_{k\ell}^{+} \right] + R_{2x}^{-1} \mathbf{D}_{k\ell}^{+} R_{2x}, \]

\[ \mathbf{G}_{k\ell} = -\text{Re} \left[ \mathbf{D}_{k\ell}^{+} \right] + R_{2x}^{-1} \mathbf{D}_{k\ell}^{+} R_{2x}, \]

\[ q_i = \text{Re} \hat{q}_i, \quad q_i = \text{Re} \hat{q}_i, \]

\[ q_i = \text{Im} \hat{q}_i, \quad q_i = \text{Im} \hat{q}_i. \]
And the hermitian matrix $D$ (eqn. (3.2)) becomes a real, symmetric matrix $\mathcal{G}$:

$$
\mathcal{G} = \begin{pmatrix}
\Re\{D_{ij}^* + R_{2x}^{-1}D_{i2x}^* R_{2x}\} & -\Im\{D_{ij}^* - R_{2x}^{-1}D_{i2x}^* R_{2x}\} \\
\Im\{D_{kj}^* + R_{2x}^{-1}D_{k2x}^* R_{2x}\} & \Re\{D_{kj}^* - R_{2x}^{-1}D_{k2x}^* R_{2x}\}
\end{pmatrix}.
$$

At $\theta = 0$, $D$ is real and thus $\mathcal{G}$ is block diagonal. Eqn. (3.5) becomes two independent eqns., one for the symmetric and one for the antisymmetric eigenvectors. By checking table 1 we find that at $\theta = 0$ the acoustic longitudinal mode and torsional acoustic mode are both antisymmetric under $C_2$ operation. The restrictions on $\mathcal{G}_{kk}^*$ given by free center of mass motion are then

$$
\sum_j \sqrt{m_j} G_{ij}^{kk} = 0, \quad \theta = 0
$$

and

$$
\sum_j \sqrt{m_j} (G_{i2x}^{kk} x_j^x - G_{i2y}^{kk} x_j^y) = 0, \quad \theta = 0.
$$

At $\theta = \pm \psi$ the eigenvector of the transverse acoustic mode is $q_i^{\pm} = \sqrt{m_i} (1, \pm i, 0)$, and gives the restrictions on $\mathcal{G}$:

$$
\sum_i \sqrt{m_i} G_{ij}^{xk} e_i^x = 0, \quad \theta = \pm \psi
$$

and

$$
\sum_i \sqrt{m_i} G_{ij}^{yk} e_i^y = 0, \quad \theta = \pm \psi.
$$

Thus we see that the $C_2$ symmetry distinguishes the normal modes at $\theta = 0$ as symmetric and antisymmetric, at $\theta \neq 0$ the $C_2$ operation mixes the normal modes at $\theta$ and $-\theta$ to again form symmetric and antisymmetric modes. (See eqn. (3.4)). It is known that time reversal combined with an inversion element can often be used to make the dynamic matrix and its eigenvector real.
IV. Selection Rules

The infrared and Raman selection rules for helical molecules have been discussed by Higgs\textsuperscript{13} and the results are the following:

Infrared absorption can only occur at $\theta = 0, +\psi$, and Raman scattering, at $\theta = 0, +\psi, +2\psi$. In general these selection rules are applicable to double helical molecules. For the molecule possessing C\textsubscript{2} symmetry these selection rules are affected by the C\textsubscript{2} symmetry and the results are listed in Tables II and III where $x_+ = x + iy$, $x_0 = z$.

At $\theta = 0$ the symmetric modes produce no $y$ and $z$ components of induced dipole and the antisymmetric modes produce no $x$ component of induced dipole. At $\theta \neq 0$, the normal mode at 0 and $-\theta$ are mixed by the C\textsubscript{2} operation. The time reversal representations have been classified by Lax\textsuperscript{7} as of 3 types: real, pseudoreal and complex. In the vibrational mode problem a real representation means no double degeneracy due to time reversal, the subspace spanned by $q(\theta)$ and the subspace spanned by $q(-\theta)$ are identical. Thus transitions that occur at $\theta$ can also occur at $-\theta$. These results are reflected in Table II and III.

Furthermore the C\textsubscript{2} axis provides a special direction. For Z-DNA at the pitch angle $\theta = -60^\circ$ there are 6 such special directions in space. When incident light is polarized along these directions additional transition will occur at $\theta = 0$. These transitions are generally not found for a helical molecule without C\textsubscript{2} symmetry. The last line of Tables II and III lists these transitions.
V. Conclusions

Alternating polymers of DNA have an additional symmetry, a \( C_2 \) symmetry about the x-axis. This is in addition to the screw axis or helical symmetry found in all repeating DNA's. This additional symmetry can be used along with the helical symmetry and the ubiquitous time reversal invariance to greatly simplify helix lattice calculations for these alternating polymers. In this paper we have detailed the way to make use of this simplification. We show how to reduce the dimension of the dynamical matrix where possible, to greatly reduce the diagonalization time needed for solution. In practice we have reduced the calculation time for the modes of an alternating methylated Z-DNA by a factor of ten.\(^{14}\)

We have also examined the implications of the additional symmetry in changing the infrared and Raman selection rules. We find that a new absorption mode is possible in these polymers with \( C_2 \) symmetry. The additional absorption mode should cause greater total absorption for the alternating polymers than for simple homopolymers. The increase is strongest for light polarized transverse to the helix axis. In crystalline DNA where the six x-axis directions can be specified the selection rules may be verified directly by increased absorption at these special directions traverse to the helix axis.

Acknowledgements

This work has been supported in part by NIH grant GM 24443 and ONR grant N00014-86-K-0252.
References

14. X. N. Han and E. W. Prohofsky, "Normal Mode Calculation for Methylated Z-DNA polv(dG-m5dC) (dG-m5dC)", to be published in Biopolymers.
TABLE I. Unnormalized eigenvectors of free motion.

\[
\begin{align*}
\nu = 0, & \quad q^m_i = \sqrt{m_i} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad q_i = \begin{pmatrix} 0 \\ 0 \\ \sqrt{m_i} \end{pmatrix}, \quad \text{longitudinal acoustic mode} \\
\nu = 0, & \quad q^m_i = zx\sqrt{m_i} \hat{x}\hat{m}_i, \quad \hat{q}_i = \sqrt{m_i} \begin{pmatrix} -x^0 \\ x^0_i \\ y \\ 0 \end{pmatrix}, \quad \text{acoustic torsional mode} \\
\nu = \nu, & \quad q^m_i = \sqrt{m_i} \begin{pmatrix} x \\ 0 \end{pmatrix}, \quad \hat{q}_i = \sqrt{m_i} \begin{pmatrix} x + iy \\ y + ix \\ 0 \end{pmatrix}, \quad \text{x,y real and arbitrary}
\end{align*}
\]
Table II. Infrared Selection Rules

<table>
<thead>
<tr>
<th>Polarization of incident</th>
<th>Parameter 0 Without $C_2$</th>
<th>With $C_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>r.h. circular ($x_+$)</td>
<td>$+\psi$</td>
<td>$\pm \psi$</td>
</tr>
<tr>
<td>l.h. circular ($x_-$)</td>
<td>$-\psi$</td>
<td>$\pm \psi$</td>
</tr>
<tr>
<td>Linear parallel ($x_0$)</td>
<td>0</td>
<td>0, antisymmetric</td>
</tr>
<tr>
<td>Special direction $x$</td>
<td>No</td>
<td>0, symmetric</td>
</tr>
</tbody>
</table>

Table III. Raman Selection Rules

<table>
<thead>
<tr>
<th>Polarization</th>
<th>Parameter 0 Without $C_2$</th>
<th>With $C_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>incident</td>
<td>scattered</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$x_+$</td>
<td>$2\psi$</td>
</tr>
<tr>
<td></td>
<td>$x_-$</td>
<td>$\psi$</td>
</tr>
<tr>
<td></td>
<td>$x_0$</td>
<td>0</td>
</tr>
</tbody>
</table>

Special $x$

<table>
<thead>
<tr>
<th></th>
<th>Parameter 0</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0, symmetric</td>
</tr>
</tbody>
</table>
END
DATE
FILMED
DTIC
July 88