

Introduction to Computational Chemistry: Teaching Hückel Molecular Orbital Theory Using an Excel Workbook for Matrix Diagonalization

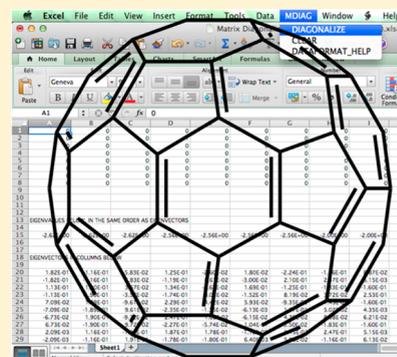
Joshua Litofsky and Rama Viswanathan*

Department of Chemistry, Beloit College, Beloit, Wisconsin 53511, United States

Supporting Information

ABSTRACT: Matrix diagonalization, the key technique at the heart of modern computational chemistry for the numerical solution of the Schrödinger equation, can be easily introduced in the physical chemistry curriculum in a pedagogical context using simple Hückel molecular orbital theory for π bonding in molecules. We present details and results of computations, including both the quintessential examples of polycyclic aromatic hydrocarbons discussed in text books and an interesting extension to a large molecule, C_{60} (buckminsterfullerene), the first member of the fullerenes to be discovered and synthesized, using a simple Excel spreadsheet-based VBA “application” that we have developed.

KEYWORDS: Upper-Division Undergraduate, Physical Chemistry, Computer-Based Learning, Computational Chemistry, MO Theory



In 2011, Johnson and Engel¹ published the results of a two-year effort at a major research university in integrating computational chemistry into the undergraduate physical chemistry curriculum. Their comprehensive report also reviewed earlier efforts at using computational molecular modeling in undergraduate education over the past two decades in the context of rapid advances in computer technology. In discussing their pedagogy to train students in using computational chemistry software, a key recommendation of the latest revised guidelines for undergraduate education² published by the American Chemical Society, Johnson and Engel¹ stated that (their) “course structure results in the software being used as a ‘black box’ for the first few exercises, which is difficult to address owing to the large amount of background material needed to understand the Hartree–Fock method.”

In 2009, we revised the physical chemistry curriculum at Beloit College in light of the 2008 ACS guidelines² that explicitly stipulated the integration of computational chemistry in the curriculum. Our approach has worked well with small classes in a significantly redesigned version of the traditional quantum chemistry course that now emphasizes molecular modeling, visualization, and computational chemistry (and is titled as such). In this article, we present details of a key element of our approach, namely an introduction to the principles of computational chemistry at an early stage of the course using Excel-based computer laboratory exercises assigned in the context of a detailed development of Hückel molecular orbital (HMO) theory.^{3–6} As is well known, HMO theory is conceptually very simple and is based on gross

approximations. However, MO energies and wave function coefficients, traditionally computed by solving polynomial equations using the method of secular determinants in linear variational theory,^{3–7} yield excellent qualitative insights into the behavior of hydrocarbons with π bonds, including predictions of delocalization, aromaticity, π bond order, etc. Thus, the theory has been an evergreen in introducing students to quantum chemistry, given its distinguished history as one of the first methods for computing such properties for conjugated hydrocarbons, long before the advent of modern electronic computers. It is also easy for students to connect the material with their prior knowledge and insights gained in other chemistry courses that they may have taken. (Typically, students taking physical chemistry at Beloit have either taken inorganic chemistry or organic chemistry, or are concurrently enrolled in one of those courses.)

Our pedagogical approach utilizes the modern computational technique for solving the Schrödinger equation using numerical matrix diagonalization⁸ of the Hamiltonian matrix.^{3–6} This eliminates the need for students to tediously solve polynomial equations of large degree (corresponding to the number of carbon atoms) based on secular determinants. A variety of custom software⁹ and scientific computing packages and libraries¹⁰ for matrix diagonalization are currently available and have been used in teaching HMO theory.^{3,4,11} What is novel about our approach is the systematic use of the ubiquitous Excel software environment and its associated spreadsheet cell (row, column) paradigm. In the context of

HMO theory, this allows for the direct input of Hamiltonian matrix coefficients into an Excel spreadsheet as two-dimensional arrays with N rows and N columns, where N typically is the number of carbon atoms. The matrix is then diagonalized by invoking a macro as a menu item from the menu bar. It should be noted that an Excel plug-in for matrix diagonalization using Dynamic Link Libraries (DLLs) was described¹² in this Journal in 1993, but the DLLs are incompatible with modern versions of the Microsoft Windows Operating System.

■ BACKGROUND: MATRIX DIAGONALIZATION IN HMO THEORY

In simple HMO theory, the basis set for a N -carbon system is the linear combination of N carbon 2p atomic orbitals (ao's), a gross approximation that however leads to great insights, both into the chemistry and (we emphasize) into the techniques of modern computational chemistry. Thus, the MO's ψ_j are given by

$$\psi_j = \sum_{i=1}^N c_{ij} \phi_i \quad (1)$$

where c_{ij} is the coefficient contribution of the carbon 2p basis function ϕ_i to the j^{th} MO ψ_j .

The matrix formulation of the HMO theory Schrödinger equation^{6,13} is then given by

$$HC = SCE \quad (2)$$

where H is the $N \times N$ Hamiltonian matrix composed of "energy" terms—the coulomb or resonance integrals H_{ij} , S is the $N \times N$ matrix composed of the overlap integrals S_{ij} , and C is the $N \times N$ matrix of coefficients c_{ij} (i indicates the basis function and j indicates the MO) with each column representing an eigenvector corresponding to a MO. E is the diagonal $N \times N$ matrix containing diagonal elements that are the MO energy levels, that is, eigenvalues.

In simple HMO theory for hydrocarbons, the coulomb integral H_{ii} is usually denoted by α , a negative number equal in magnitude to the ionization energy of a carbon 2p orbital electron. The resonance integral H_{ij} is denoted by β , also a negative quantity because it is a "bond" energy. Only neighboring atom interactions are considered, that is, $H_{ij} = 0$ for $i \neq j$ and $|i - j| \neq 1$. Resonance integrals for non-neighbor atom pairs are set to zero, as are all overlap integrals except for those involving a single atom, that is, $S_{ij} = 0$ ($i \neq j$), and $S_{ii} = 1$. S then becomes the unit matrix. Applying these conditions based on the simple HMO model, eq 2 reduces to

$$HC = CE \quad (3)$$

and the (unknown) MO energy levels can be determined by "diagonalization" of the known Hamiltonian matrix composed of coulomb and resonance energy terms. In essence, the diagonalization problem reduces to finding values for the matrix elements in C and E , given H .¹³ Mathematically, this can be achieved by premultiplying the left and right sides of eq 3 by the transpose matrix C^t , since

$$C^t HC = C^t CE = E \quad (4)$$

In the case of simple HMO theory, further simplification is possible because the matrix C is real and symmetric, and therefore $C^t = C$.

Computationally, the solution, like for many other problems amenable to numerical analysis, is to start with an initial guess

for the matrix C and then iteratively compute values for E in eq 4 or its equivalent, attempting to reduce the values of all off-diagonal elements to 0, a procedure described in detail in a later section. The results of the "converged" solution provide all the eigenvalues (energies that are diagonal values in E) and eigenvectors (columns in C that provide the coefficients contributed by each of the atomic orbitals to the MO wave function represented by the column) in terms of the specific finite basis set that was used. Because HMO theory is based on gross approximations with complete neglect of multiple atom overlap integrals and is expected to yield only qualitative results, we can dispense with using numerical values for the various integrals (matrix elements) and express all results in terms of symbolic values of the carbon 2p orbital coulomb integral α , and the nearest neighbor carbon-carbon resonance integral β . Thus, using scaled reference values of 0 for α and 1 for β , respectively, the Hamiltonian matrix in simple HMO theory is easily set up and diagonalized.

■ EXCEL VBA-BASED MATRIX DIAGONALIZATION SOFTWARE DEVELOPED AT BELOIT COLLEGE

Diagonalization of Real, Symmetric Hamiltonian Matrices

The energies (eigenvalues) and wave functions (eigenvectors) of the Hamiltonian matrix formulation of the Schrödinger equation (an eigenvalue equation) in eq 2 can be obtained by a number of sophisticated numerical techniques⁸ developed since the advent of the computer. However, many of these techniques are based on linear algebra and have been well known to mathematicians for centuries. A specific method, certainly not the most efficient but one that is easily understood by beginners studying numerical methods or students with limited exposure to linear algebra, is the Jacobi method for real, symmetric matrices^{5,6,13,14} such as the matrices used in Hückel MO and other quantum mechanical techniques. The Jacobi method was adapted to work with computers in 1960 by Greenstadt.¹⁵ Students are not required to know all the algorithmic details or the computer code implementation, except to understand that the basic principle involves iterative minimization of the magnitude of all off-diagonal elements of the matrix by a series of matrix multiplications involving suitably constructed transform matrices C and their transposes. The procedure is deemed to have converged when the largest off-diagonal element is less than a certain value (very close to zero) specified in the computer program. This termination value is selected based on round-off and number representation precision criteria for the programming language involved, and the resulting diagonalized matrix E then has the energies as the diagonal elements. Additionally, the final transformation matrix C provides the eigenvector (wave function) coefficients. Substantially more sophisticated and optimized techniques for matrix diagonalization⁸ form the core routines of all computer software currently available in computational chemistry,^{5,6} and therefore the exposure to these principles is of great value for all chemistry students today.

In the sections that follow, we present details of software developed at Beloit College to diagonalize real symmetric matrices using the Jacobi method, customized for HMO theory. For flexibility and use in diverse computing environments, we have developed MDIAG, an Excel application that uses standard Excel Visual Basic for Applications (VBA) macro code embedded into a workbook with a custom menu option for invoking the matrix diagonalization macro. We describe the

software and how it is implemented using the Hamiltonian matrix for naphthalene in simple HMO theory as an example, and then discuss an interesting contemporary and practical application of HMO theory in studying the properties of C_{60} (buckminsterfullerene), developed as a class project by the student coauthor (JL).

Matrix Diagonalization Using Excel VBA—the MDIAG Custom Workbook

The HMO Hamiltonian matrix for the ten-carbon bicyclic system naphthalene ($C_{10}H_8$) is easily set up (Table 1) using the Hückel rules specified above, with a numbering scheme for matrix element indices that begins with a bridging carbon as shown in the accompanying graphic (Figure 1):

Table 1. Simple HMO Hamiltonian Matrix for Naphthalene^a

α	β	0	0	0	β	0	0	0	β
β^b	α^b	β	0	0	0	0	0	0	0
0	β	α	β	0	0	0	0	0	0
0	0	β	α	β	0	0	0	0	0
0	0	0	β	α	β	0	0	0	0
β	0	0	0	β	α	β	0	0	0
0	0	0	0	0	β	α	β	0	0
0	0	0	0	0	0	β	α	β	0
0	0	0	0	0	0	0	β	α	β
β	0	0	0	0	0	0	0	β	α

^aThe rows and columns of the matrix are implicitly numbered 1 through 10; the matrix is real and symmetric. ^bNote that $\alpha = 0$ (reference) and $\beta = 1$ when entered into the spreadsheet.

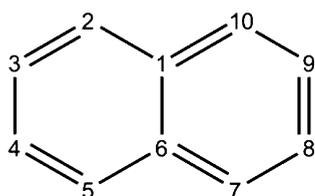


Figure 1. Structural formula for naphthalene, with carbon atom numbering corresponding to appropriate matrix elements in Table 1.

The actual computation of eigenvectors and eigenvalues for this system using the Excel spreadsheet displayed in the MDIAG workbook involves entering a label in row 1, column 1 (i.e., cell A1), the order of the matrix (in this case 10) in row 2, column 1 (i.e., cell A2), and then the Hamiltonian matrix elements corresponding to each of the coulomb ($\alpha = 0$) and exchange ($\beta = 1$) integrals, starting with row 3, column 1 (i.e., cell A3). Finally, the user selects the DIAGONALIZE custom menu item to perform the computation, which is typically completed in a fraction of a second for naphthalene. Figure 2 shows an actual screen capture of the Excel spreadsheet, showing the custom diagonalization menu added to the Excel standard menu bar, along with data and the results under the “EIGENVALUE” and “EIGENVECTOR” headers. Each specific eigenvalue is an HMO energy level and is displayed in units of β , referenced to α as the zero of energy. For example, the eigenvalue in the fifth column is actually $\alpha + 0.6180\beta$. Because α and β are both negative, the eigenvalue is actually smaller than the reference energy α corresponding to the carbon 2p ao. Thus, this eigenvalue corresponds to a bonding MO. Each eigenvector corresponding to an HMO energy level (a specific eigenvalue) is a column vector in that specific column, with elements corresponding to c_{ij} in eq 3. j is the

column number, corresponding to the j^{th} MO, and i is the (relative) row number starting with $i = 1$ for first row of the eigenvector coefficient listing.

In a typical 3 h computer laboratory session, students working in groups of two were systematically able to set up and diagonalize the Hamiltonian matrices for a series of open chain conjugated hydrocarbons, starting with the “reference” molecule with a single double bond, namely ethene, and working their way up to hexatriene and octatetraene. They then computed HMO energies and wave function coefficients for benzene and polycyclic aromatic hydrocarbons (PAHs), including naphthalene (Figure 2) and anthracene. A variety of homework assignments based on the computations were handed out in class. For example, these assignments included computation of delocalization energies based on the eigenvalues and electron densities and π bond orders based on the eigenvector coefficients. It is particularly straightforward to perform the tedious calculations for electron densities and bond orders, especially for large molecules, using the Excel formula bar feature and applying appropriate formulas to empty cells in the spreadsheets after the results of matrix diagonalization are displayed. The values for naphthalene are included and highlighted in Figure 2. There are four different values for the π bond orders computed for naphthalene, indicating that the carbon–carbon bonds are not identical. Students were asked to discuss the significance of the results by comparing them to the results for benzene, where all five π bond orders are identical. This can be rationalized on the basis of canonical resonance structures for naphthalene that are not equivalent, unlike the case for benzene. It is interesting to note that recent ab initio calculations and experimental data¹⁶ are in essential agreement with the predictions for naphthalene made on the basis of simple HMO theory. Many similar exercises and questions based on the structures of PAHs and other fused carbon ring systems are discussed in detail in various textbooks.^{3–7,13} We describe a novel application of our MDIAG Excel workbook software in computing the HMO theory-based properties of C_{60} , a system of great contemporary interest, in the section that follows.

APPLICATION IN A STUDENT PROJECT: C_{60} (BUCKMINSTERFULLERENE) HÜCKEL ENERGIES

The simple HMO eigenvalues for C_{60} (buckminsterfullerene), colloquially known as “buckyball”, are difficult to obtain analytically. In principle, the method of secular determinants that has been traditionally used to obtain eigenvalues would involve determinants of order 60. However, ingenious methods that make use of the inherent icosahedral symmetry of the molecule have been used to derive exact analytical solutions that can be computed using a hand calculator.^{17,18} An alternative is to set up the HMO Hamiltonian matrix for C_{60} and diagonalize it to obtain eigenvalues. This was done by one of us (JL) as a student final project for the Molecular Modeling, Visualization, and Computational Chemistry class at Beloit College in Fall 2012, using the Excel MDIAG workbook described above. The matrix was easy to set up because of the symmetry. Verifying that the matrix is correct posed an interesting problem. The three-dimensional cage structure of C_{60} (a truncated icosahedron with 20 hexagons and 12 polygons) made numbering and counting carbons and determining the nearest directly connected neighbors, required for filling in nonzero matrix elements, nontrivial and confusing. However, an easy method based on a graph theoretical analysis

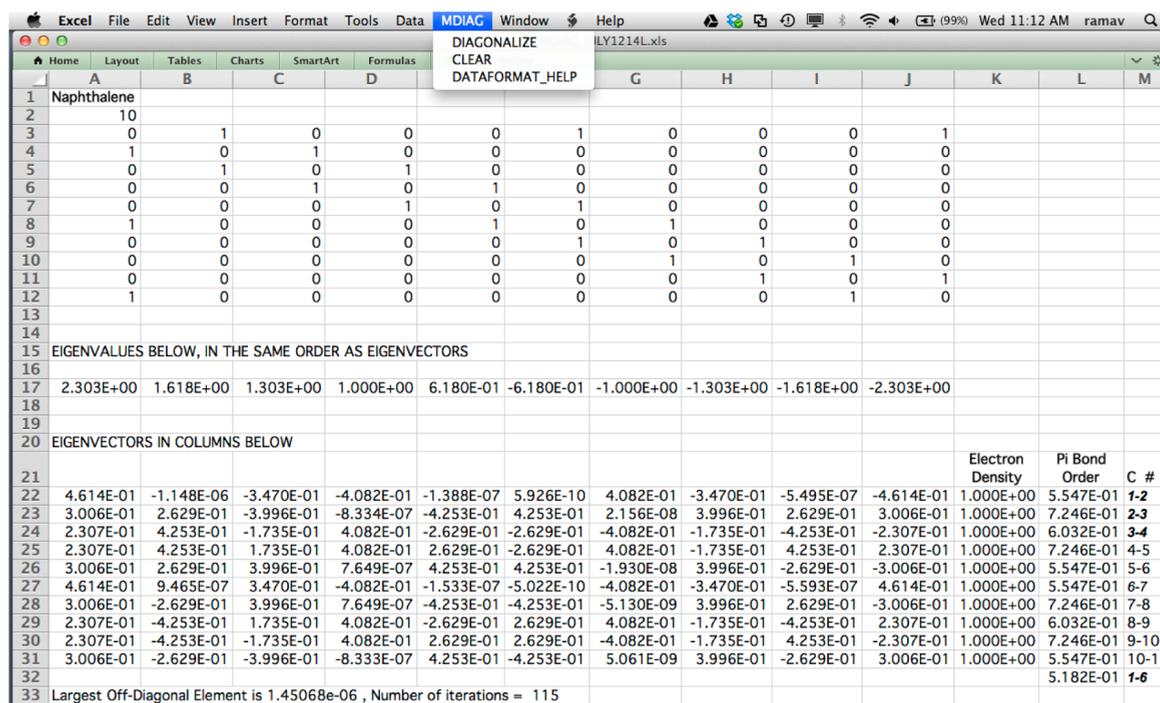


Figure 2. Screenshot (Mac OS X) of the HMO real symmetric matrix for $C_{10}H_8$ (naphthalene) set up in the MDIAG Excel workbook, along with the results—eigenvalues and eigenvectors. Note that the computation was “run” by selecting the DIAGONALIZE custom menu item.

of the Hückel matrix for C_{60} yields a simple two-dimensional “connection graph”¹⁹ that was used to number the carbons and easily determine the positions of the nonzero matrix elements.

The final matrix was also verified using the method of Vittal,²⁰ who created a simple two-dimensional paper model that can be used for the numbering and requires little more than a pair of scissors and tape to cut and fold to construct a three-dimensional C_{60} molecule. The results from our Excel-based computation, which took less than 10 s on a stock PC, are compared (Table 2) to the complete set of simple HMO theory results reported in the seminal review article by Kroto et al.,²¹ both to test the validity of the program and to demonstrate that it can be used for large molecules. The

Table 2. Eigenvalues for C_{60}

Computed (this work)	Literature (Kroto et al. ²¹)
3.000 ^a	3
2.757[3] ^b	2.757[3]
2.303[5]	2.303[5]
1.820[3]	1.820[3]
1.562[4]	1.562[4]
1.000[9]	1[9]
0.6180[5]	0.618[5]
−0.1386[3]	−0.1386[3]
−0.3820[3]	−0.3820[3]
−1.303[5]	−1.303[5]
−1.438[3]	−1.438[3]
−1.618[5]	−1.618[5]
−2.000[4]	−2[4]
−2.562[4]	−2.562[4]
−2.618[3]	−2.618[3]

^aThe numbers are in units of β , referenced to α , which is the zero of energy. ^bDegeneracy in square brackets.

agreement is excellent within the round off error limits and machine precision.

It is possible to compare experimental data²¹ for C_{60} against predictions made by simple HMO theory. The calculations are also historically significant because most of the theoretical studies^{18,19,21} published in the literature in the early and mid-1980s (when the formation of stable C_{60} was first experimentally verified) were HMO-theory-based, principally because of computer technology limitations. Students can compute bond orders and determine that there are two distinct values, corresponding to the five-membered and six-membered ring substructures in the truncated icosahedron. These values can be compared to values³ for various types of C–C bonds in terms of reactivity and aromaticity. The delocalization energy per carbon (0.5527β) can be calculated and is larger than the value (0.3333β) for benzene, the reference system for aromaticity. However, it is well known²¹ that C_{60} is not aromatic and has unique properties. Given that simple HMO theory assumes planarity, students can be asked to discuss the apparent anomaly on the basis of structure. It has also been determined that pure C_{60} is a perfect insulator²² because of the large “band gap” between HOMO and LUMO and is a weak electrophile,²¹ presumably because of its low-lying LUMO. Both of these properties can be predicted on the basis of MO energies computed using HMO theory.

CONCLUSIONS

We have extensively tested our approach using the Excel-based software and assignments that routinely involve computation of HMO energies and construction of HMO diagrams for a number of conjugated hydrocarbons, starting with the reference carbon–carbon system ethene with a single π bond, working with open chain conjugated systems through octatetraene, and then with benzene and PAHs, including naphthalene and anthracene. All these computational assignments can be easily

carried out in a single 3 h laboratory period. In addition, we have also assigned computations on C_{60} for extra credit, and the response and interest from students has been very good. Anecdotal feedback that we have received suggests that students taking the physical chemistry sequence over the past five years found the HMO calculations particularly interesting and challenging. Students have also indicated that it was very relevant in terms of understanding chemical bonding and structure and developing a better appreciation of the concepts of conjugation, resonance, and aromaticity in organic chemistry.

■ ASSOCIATED CONTENT

● Supporting Information

MDIAG, a custom Excel Workbook with VBA matrix diagonalization macros accessible from the menu bar, along with sample CSV data files for ethene and butadiene. This material is available via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ramav@beloit.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

J.L. wishes to thank Xincheng Deng, Beloit College class of 2014, for collaborating in the initial class project to fulfill requirements for the physical chemistry class in molecular visualization, modeling, and computational chemistry at Beloit College. The dedicated computer modeling and visualization laboratory where the Excel software was installed was partially funded by NSF CCLI Grant DUE-0633651.

■ REFERENCES

- (1) Johnson, L. E.; Engel, T. Integrating Computational Chemistry into the Physical Chemistry Curriculum. *J. Chem. Educ.* **2011**, *88*, 569–573.
- (2) *Undergraduate Professional Education in Chemistry: ACS Guidelines and Evaluation Procedures for Bachelor's Degree Programs (Spring 2008)*; American Chemical Society: Washington, DC, 2008; Section 4.3, p 7.
- (3) McQuarrie, D. A. *Quantum Chemistry*, 2nd ed.; University Science Books: Mill Valley, CA, 2008; pp 581–597. Also see pp 427–433 for an explicit mathematical discussion of matrix eigenvalue problems and matrix diagonalization. Readers are referred to specific computer programs and packages for performing numerical diagonalizations of large matrices.
- (4) Atkins, P.; de Paula, J. *Physical Chemistry*, 9th ed; W. H. Freeman: New York, NY, 2010; pp 395–407. Readers are referred to software for performing numerical matrix diagonalizations.
- (5) Rogers, D. W. *Computational Chemistry Using the PC*, 3rd ed; Wiley-Interscience: Hoboken, NJ, 2003; pp 169–230.
- (6) Lewars, E. G. *Computational Chemistry*, 2nd ed; Springer: New York, NY, 2011; pp 118–152.
- (7) Roberts, J. D. *Notes on Molecular Orbital Calculations*; Benjamin: New York, NY, 1962.
- (8) Golub, G. H.; Van Loan, C. F. *Matrix Computations*, 4th ed; Johns Hopkins University Press: Baltimore, MD, 2013.
- (9) Examples of custom interfaces and apps for matrix diagonalization include: WebMO, www.webmo.net, and Java applet SHMO, www.chem.ucalgary.ca/SHMO/ (accessed April 9, 2014).
- (10) Examples of scientific software packages that include matrix diagonalization are *Matlab*, www.mathworks.com/products/matlab/; *GNU-Octave*, www.gnu.org/software/octave/; and *Mathematica*, www.wolfram.com.mathematica/ (accessed April 9, 2014).
- (11) Ramakrishnan, R. A Simple Hückel Molecular Orbital Plotter. *J. Chem. Educ.* **2013**, *90*, 132–133 and references therein.
- (12) Kari, R. Dynamic Link Libraries in Windows 3.x. *J. Chem. Educ.* **1993**, *70*, 247–248.
- (13) Lowe, J. P. *Quantum Chemistry*, 2nd ed; Academic Press: San Diego, CA, 1993; pp 240–266. Also see pp 316–321 for details of the matrix formulations of eq 2, 3, and 4.
- (14) Bauer, H.; Roth, K. The Numeric Solution of Eigenvalue Problems. *J. Chem. Educ.* **1980**, *57*, 422–424.
- (15) Greenstadt, J. The Determination of the Characteristic Roots of a Matrix by the Jacobi Method. In *Mathematical Methods for Digital Computers*; Ralston, A., Wilf, H. S., Eds.; John Wiley & Sons: New York, NY, 1960; Vol 1, pp 86–91.
- (16) Baba, M.; Kowaka, Y.; Nagashima, U.; Ishimoto, T.; Goto, H.; Nakayama, N. Geometrical Structure of Benzene and Naphthalene: Ultrahigh-resolution Laser Spectroscopy and ab initio Calculation. *J. Chem. Phys.* **2011**, *135*, 054305–4.
- (17) Fowler, P. W.; Woolrich, J. π -Systems in Three Dimensions. *Chem. Phys. Lett.* **1986**, *127*, 78–83.
- (18) Brown, W. B. High Symmetries in Quantum Chemistry. *Chem. Phys. Lett.* **1987**, *136*, 128–133.
- (19) Manolopoulos, D. E.; May, J. C.; Down, S. E. Theoretical Studies of the Fullerenes: C_{34} to C_{70} . *Chem. Phys. Lett.* **1991**, *105*, 107. Refer to Figure 2 of the paper for the graph template.
- (20) Vittal, J. J. A Simple Paper Model for Buckminsterfullerene. *J. Chem. Educ.* **1989**, *66*, 282.
- (21) Kroto, H. W.; Allaf, A. W.; Balm, S. P. C_{60} : Buckminsterfullerene. *Chem. Rev.* **1991**, *91*, 1215. HMO Results reprinted from Davidson, R. A. *Theor. Chim. Acta* **1981**, *58*, 193–195.
- (22) O'Shea, J. N. Molecular Orbitals Tell the Story. *Science* **2005**, *310*, 453.