

Classical Trajectory of Molecules in Electromagnetic Field: A Handy Method to Simulate Molecular Vibrational Spectra

Jiří Kessler and Petr Bourčák*



Cite This: *J. Chem. Theory Comput.* 2022, 18, 1780–1787



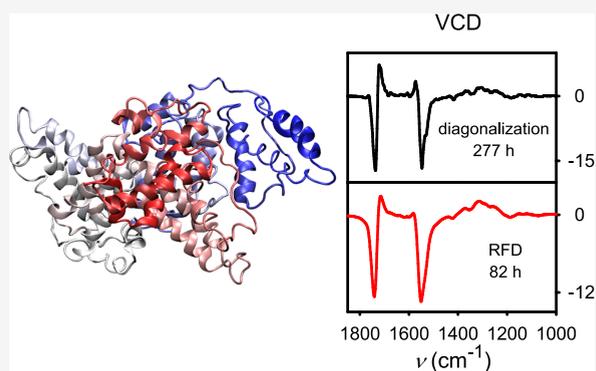
Read Online

ACCESS |

Metrics & More

Article Recommendations

ABSTRACT: Within harmonic approximations, molecular vibrational spectra are simulated in a standard way through force field diagonalization and following transformation of Cartesian to normal-mode tensor derivatives. This may become tedious for large systems of many thousands of atoms and also not necessary because of a limited resolution required to interpret an experiment. We developed an algorithm based on the real-time real-field molecular dynamics, effectively at zero temperature, invoked in a molecule by the electromagnetic field of light. The algorithm is simple to implement and suitable for parallel computing, and it can be potentially extended to more complicated molecular-light interaction modes. It circumvents the diagonalization and is suitable to model vibrational optical activity (vibrational circular dichroism and, to a lesser extent, Raman optical activity). For large molecules, it becomes faster than diagonalization, but it also enables the assignment of vibrational spectral bands to local molecular motions.



1. INTRODUCTION

Modern computational techniques make it possible to construct relatively precise harmonic force fields of fairly large molecular systems. This is extremely useful in the interpretation of vibrational spectra such as those sensitive to optical activity, adding more information to unpolarized spectroscopy. For example, the vibrational Raman optical activity (ROA) spectrum of the β domain of rat metallothionein of 400 atoms could be directly calculated using an efficient density functional method and the resolution-of-identity density-fitting technique.¹ Fragmentation methods such as ONIOM² or molecules-in-molecules (MIM)³ further shift the limitations; a MIM study was devoted to many proteins including the ITK kinase domain in a complex of 721 atoms.⁴ We have used the Cartesian coordinate-based tensor transfer (CCT) to construct the force field and intensity parameters of several proteins to analyze their vibrational circular dichroism (VCD)⁵ and ROA⁶ spectra, where the largest molecule, human serum albumin, had 9161 atoms.

For systems of thousands and tens of thousands atoms, complete diagonalization of the force field is tedious and may ultimately become the limiting factor. The most frequent full diagonalization algorithm based on the Householder reduction to a tridiagonal form⁷ is efficient, but the required time in case of the force field is roughly proportional to $(3N)^3$, where N is the number of atoms.⁸ The required computer memory is obviously proportional to $(3N)^2$, since at least the whole matrix needs to be stored for a fast processing. Usually, we are

interested in all eigenvalues and eigenvectors or at least in a significant part of the spectrum, which limits the use of iterative partial diagonalization methods such as those of Lanczos⁹ and Davidson¹⁰/Mitin¹¹ working within the Krylov subspace.¹² Special algorithms, such as the adaptive sampling CI method,¹³ may not be directly applicable to the vibrational problem.

In the present study, instead of the diagonalization, we determine vibrational band energies and intensities from a resonance during a real field dynamics (RFD), i.e., when the molecule is placed in the electromagnetic field of light. Time-dependent techniques have been already explored for generating the vibrational spectra many times in the past, albeit with a variable performance. Methods based on the Fourier transform and autocorrelation functions are popular¹⁴ as they can be used in both classical and ab initio molecular dynamics.¹⁵ They are immensely useful for studies of the properties of liquids;¹⁶ however, the resultant spectral pattern may converge slowly.¹⁷ For simulations of bigger molecules, we found similar problems when the accuracy was affected by an

Received: November 10, 2021

Published: February 9, 2022



uneven distribution of energy among vibrational normal modes.¹⁸ We also introduced an arbitrary time propagation to diagonalize a general matrix, which led to favorable diagonalization times, but the accuracy was limited.¹⁹

The method presented below does not seem to suffer from these problems. The algorithm describes a real experiment, such as light of defined polarization interacting with a polarizable molecule. In this way, we also obtain a didactic “toy” on which the interaction of a molecule with the light can be documented, e.g., for different light directions with respect to molecular orientation. In principle, it can be extended beyond the dipolar (but not the harmonic) approximation to model more complex experiments. Below, we provide the theoretical foundations and focus on practical simulations of infrared absorption (IR), vibrational circular dichroism (VCD), Raman, and Raman optical activity (ROA) spectra. The method clearly becomes relatively faster for larger systems.

2. METHODS

2.1. Molecule in the Electromagnetic Field of Light.

We suppose that the molecule is irradiated by a plane wave characterized by the intensities of the electric (**E**) and magnetic (**B**) fields dependent on time *t*. Figure 1 shows the

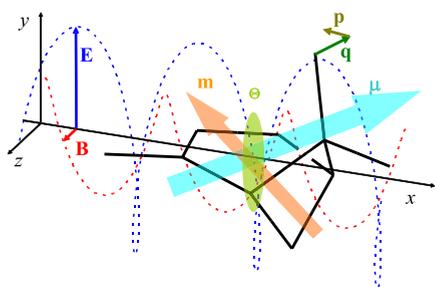


Figure 1. Molecule in the electromagnetic field of a plane wave. The field is characterized by the electric (**E**) and magnetic (**B**) intensities, the molecule by the electric (**μ**) and magnetic (**m**) dipole, and electric quadrupole (**Θ**) moments. The moment changes depend on the deviations of atoms from its equilibrium positions (**q**) or its momenta (**p**), which determines the system dynamic.

situation for circularly polarized light (CPL), where vectors **E** and **B** are perpendicular. The molecule interacts with the light through its electric **μ** and magnetic **m** dipole and electric quadrupole (**Θ**) moments. Classical molecular Hamiltonian *H* is a function of coordinates $d_{\lambda\alpha}$ (α -deviation of an atom λ from its equilibrium position) and momenta $p_{\lambda\alpha}$.²⁰ For brevity, we group the atomic indices into one ($d_{\lambda\alpha} \equiv d_{\lambda}$, etc.), so that

$$H = \sum_{\lambda} \frac{p_{\lambda}^2}{2m_{\lambda}} + V - \boldsymbol{\mu} \cdot \mathbf{E}(t) - \frac{1}{3} \boldsymbol{\Theta} \cdot \nabla \mathbf{E}(t) - \mathbf{m} \cdot \mathbf{B}(t) \quad (1)$$

where m_{λ} is atomic mass and *V* is the harmonic potential, $V = \frac{1}{2} \sum_{\lambda} \sum_{\mu} q_{\lambda} f_{\lambda\mu} q_{\mu}$, where $f_{\lambda\mu}$ are the force constants.

We use Hamilton's equations of motions²¹

$$\dot{p}_{\lambda} = F_{\lambda} - \gamma m_{\lambda} \dot{q}_{\lambda} \quad (2a)$$

$$\dot{q}_{\lambda} = \frac{\partial H}{\partial p_{\lambda}} \quad (2b)$$

where $F_{\lambda} = -\frac{\partial H}{\partial p_{\lambda}} - kq_{\lambda}$ is the force, the arbitrary term kq_{λ} stabilizes the dynamics performed in the Cartesian coordinates ($k = 10^{-3}$ atomic units), and the last friction term in eq 2a is dependent on the constant γ . The kq_{λ} term also allows for applications to nonoptimized geometries if the imaginary frequencies are not too large.

From eqs 1 and 2a,b, we get

$$-\frac{\partial H}{\partial q_{\lambda}} = \sum_{\beta} P_{\lambda}^{\beta} E_{\beta} + \frac{1}{3} \sum_{\beta, \gamma} Q_{\lambda}^{\beta\gamma} E_{\beta} E_{\gamma} - \sum_{\mu} f_{\lambda\mu} q_{\mu} \quad (3a)$$

$$\dot{q}_{\lambda} = \frac{p_{\lambda}}{m_{\lambda}} - \frac{2i\hbar}{m_{\lambda}} A_{\lambda, \beta} \dot{B}_{\beta} \quad (3b)$$

where $P_{\lambda}^{\beta} = \frac{\partial \mu_{\beta}}{\partial q_{\lambda}}$ are the dipole derivatives (atomic polar tensors), $Q_{\lambda}^{\beta\gamma} = \frac{\partial \Theta_{\beta\gamma}}{\partial q_{\lambda}}$ are the quadrupole derivatives, $\frac{\partial m_{\alpha}}{\partial p_{\lambda}} = \frac{2i\hbar}{m_{\lambda}} A_{\lambda\alpha}$, and $A_{\lambda\alpha}$ is the atomic axial tensor.²² Note that $\frac{\partial m_{\alpha}}{\partial q_{\lambda}} = 0$, which reflects the fact that nuclear motion, not just a geometry perturbation, is needed to generate the magnetic moment.²³ Various definitions of **m**, **A**, and **Θ** can be encountered in the literature; we follow conventions from ref 20.

Exploring the usual leapfrog scheme,²⁴ coordinates at time $t + \Delta$ were obtained as

$$q_{\lambda}(t + \Delta) = \left[2q_{\lambda}(t) - q_{\lambda}(t - \Delta) \left(1 - \frac{\gamma\Delta}{2} \right) + \frac{\Delta^2}{m_{\lambda}} (F_{\lambda} - 2i\hbar A_{\lambda, \beta} \dot{B}_{\beta}) \right] / \left(1 + \frac{\gamma\Delta}{2} \right) \quad (4)$$

where Δ is the integration time step.

2.2. Absorption and VCD Intensities. Let us suppose, for example, that at $t = 0$, we start to irradiate the molecule with right-CPL. For the geometry displayed in Figure 1, this could be achieved with intensities $\mathbf{E} = E_0 [0, \cos(\omega t), -\sin(\omega t)]$ and $\mathbf{B} = (1/c)[0, -E_z, E_y]$, where c is the velocity of light. Then, the absorption spectrum ε is proportional to energy w absorbed by the system

$$\varepsilon(\omega) = Kw \quad (5a)$$

where

$$w = \gamma \int \sum_{\lambda} m_{\lambda} \dot{q}_{\lambda} dq_{\lambda} = \gamma \int_0^T \sum_{\lambda} m_{\lambda} \dot{q}_{\lambda}^2 dt \quad (5b)$$

T is a time period, and *K* is a constant. VCD intensities can be calculated, for example, from energies obtained with and without the magnetic and quadrupolar components, $\Delta\varepsilon = \varepsilon_L - \varepsilon_R = 2K(w_{\text{without}} - w_{\text{with}})$, from runs with one and an opposite enantiomer, or from dynamics with left- and right-CPL. The last options, however, were accompanied by a large error due to the finite precision of the leapfrog integration.

The constant *K* can be determined by comparison to the exact quantum chemical result. The dipole strength of a $0 \rightarrow 1$ fundamental transition is²⁵

$$D = \sum_{\beta} \langle 0 | \mu_{\beta} | 1 \rangle \langle 1 | \mu_{\beta} | 0 \rangle = \frac{\hbar}{2\omega_1} \sum_{\beta} (P_{\beta}^{\beta})^2 \quad (6)$$

where \hbar is the reduced Planck constant, $P_I^\beta = \sum_\lambda P_\lambda^\beta s_{\lambda I} m_\lambda^{-1/2}$, $s_{\lambda I}$ are elements of the Cartesian-normal mode transformation matrix, and ω_I is the normal mode frequency. For the calibration, we neglect the arbitrary, magnetic, and quadrupole terms in eq 2a, getting $m_\lambda \dot{q}_\lambda = P_\lambda^\beta E_\beta - \sum_\mu f_{\lambda,\mu} q_\mu - \gamma m_\lambda \dot{q}_\lambda$. We can also replace the Cartesian coordinates by the normal ones, using²⁶ $\sum_{\lambda,\mu} \frac{s_{\lambda I} f_{\lambda,\mu} s_{\mu J}}{\sqrt{m_\lambda m_\mu}} = \omega_J^2 \delta_{IJ}$, $q_\lambda = \sum_I \frac{s_{\lambda I}}{\sqrt{m_\lambda}} Q_I$, and $\sum_\lambda s_{\lambda I} s_{\lambda J} = \delta_{IJ}$, so that the equation of motion for the normal mode coordinates is

$$\dot{Q}_I = P_I^\beta E_\beta - \omega_I^2 Q_I - \gamma Q_I \quad (7)$$

With $E_\beta = \text{Re}(E_{0\beta} e^{-i\omega t})$, we look for a stationary solution $Q_I = \text{Re}(\bar{Q}_I) = \text{Re}(Q_I e^{-i\omega t})$, which gives the amplitude as $Q_{I0} = \frac{P_I^\beta E_{0\beta}}{\omega_I^2 - \omega^2 - i\omega\gamma}$. If we rewrite eq 5Sa,b with the normal modes, we get $w(\omega) = \gamma T \omega_J^2 |Q_{0J}|^2$, and the integral over the absorption band is

$$\int w(\omega) d\omega \cong \frac{\pi T E_{0\beta}^2 \omega_I D}{\hbar} \quad (8)$$

When D is expressed in debye² and ϵ is in L mol⁻¹ cm⁻¹, we can write²⁷

$$D(\text{debye}^2) = 9.184 \times 10^{-3} \int \epsilon \frac{d\omega}{\omega} \quad (9)$$

Comparing eq 9 with eqs 8 and 5, we get $K = \frac{\hbar}{9.184 \times 10^{-3} \pi T E_{0\beta}^2}$

In the current implementation, the molecule is fixed in space. To model isotropic samples where all orientations are allowed, we irradiate the molecule along the three x , y , and z directions and average the signal.

2.3. Raman and ROA. For Raman scattering, the molecule is sensing the excitation light of frequency ω , and the scattered light of frequency ω' . In eq 1, we replace the multipole moments by the induced ones

$$\boldsymbol{\mu}(\omega') = \boldsymbol{\alpha} \cdot \mathbf{E}(\omega) + \frac{1}{3} \mathbf{A} \cdot \nabla \mathbf{E}(\omega) + \frac{1}{\omega} \mathbf{G}' \cdot \dot{\mathbf{B}}(\omega) \quad (10a)$$

$$\boldsymbol{\Theta}(\omega') = \mathbf{A} \cdot \mathbf{E}(\omega) \quad (10b)$$

$$\mathbf{m}(\omega') = -\frac{1}{\omega} \mathbf{G}' \cdot \dot{\mathbf{E}}(\omega) \quad (10c)$$

where $\boldsymbol{\alpha}$, \mathbf{A} , and \mathbf{G}' are molecular transition polarizabilities defined in ref 20 and \mathbf{E} and \mathbf{B} denote the excitation fields. These dipoles then interact with the scattered radiation characterized by the primed fields (\mathbf{E}' , \mathbf{B}'). Similarly as for IR and VCD, we get the equations of motion as

$$\begin{aligned} m_\lambda \dot{q}_\lambda &= -\frac{\partial V}{\partial q_\lambda} - \gamma m_\lambda \dot{q}_\lambda - k q_\lambda + \sum_{\alpha,\beta} \frac{\partial \alpha_{\alpha,\beta}}{\partial q_\lambda} E_\alpha E'_\beta \\ &+ \frac{1}{\omega} \sum_{\alpha,\beta} \frac{\partial G'_{\alpha\beta}}{\partial q_\lambda} (E'_\alpha \dot{B}_\beta - \dot{E}_\alpha B'_\beta) \\ &+ \frac{1}{3} \sum_{\alpha,\beta,\gamma} \frac{\partial A_{\alpha,\beta\gamma}}{\partial q_\lambda} (E_\alpha E'_\beta \gamma + E'_\alpha E_{\beta\gamma}) \end{aligned} \quad (11)$$

Many ROA experiments are possible depending on the geometry and polarization/detection scheme. In actual implementation, we mimic the scattered circular polarization (SCP) experiment in backscattering geometry.²⁵ Following the geometry in Figure 1, we use linearly polarized excitation light with nonzero complex field components $\tilde{E}_o = E_0 e^{i(-kx - \omega t)}$ and $\tilde{B}_\alpha = \epsilon_{\alpha o} \frac{E_0}{c} e^{i(-kx - \omega t)}$, where o is either y or z and k is the size of the wave vector. Because the sample is usually irradiated by unpolarized light in SCP, spectra obtained with the y and z polarizations were averaged. The scattered field is left/right-circularly polarized, with nonzero complex field components $\tilde{E}'_y = E_0 e^{i(k'x - \omega't)}$, $\tilde{E}'_z = \mp i E_0 e^{i(k'x - \omega't)}$, $\tilde{B}'_y = \pm i E_0 e^{i(k'x - \omega't)}/c$, and $\tilde{B}'_z = E_0 e^{i(k'x - \omega't)}/c$. For the dynamics, we obviously need the real parts only. The field products in eq 11 were approximated using

$$\begin{aligned} \text{Re}(U)\text{Re}(V) &\approx [\text{Re}(u)\text{Re}(v) + \text{Im}(u)\text{Im}(v)] \frac{\cos(\Delta\omega t)}{2} \\ &+ [\text{Im}(u)\text{Re}(v) - \text{Re}(u)\text{Im}(v)] \frac{\sin(\Delta\omega t)}{2} \end{aligned} \quad (12)$$

where $U = ue^{-i\omega t}$ and $V = ve^{-i\omega't}$ are two complex numbers and $\Delta\omega = \omega - \omega'$. In eq 12, terms dependent on the $\omega + \omega'$ sum were discarded as they are not relevant for the low-frequency vibrational motions.

2.4. Fourier Transform Extension. Theoretically, propagation of a δ -pulse following Fourier transform of the trajectories could speed up the computations. In practice, this procedure appeared to be too inaccurate to provide reasonable optical activity spectra. Nevertheless, it does provide accurate eigenvalues as indicated by approximate absorption/Raman intensities. The δ -pulse may thus be useful for fast qualitative estimates obtained in a fraction of computer time needed for more precise calculations.

To explore behavior of the Fourier path, the “ δ -pulse” was realized by a Gaussian function, $\delta(t) = \frac{1}{\tau\sqrt{\pi}} e^{-(t/\tau)^2}$, where $\tau \approx 2$ fs, so that $\mathbf{E}(t, \mathbf{r}) = \mathbf{E}_0 \delta(t - r/c)$, $\nabla \mathbf{E}(t, \mathbf{r}) = -\mathbf{r} \mathbf{E}_0 \dot{\delta}(t - r/c)/(rc)$, and $\dot{\mathbf{B}} = -\nabla \times \mathbf{E}$. Using the same Hamiltonian (eq 1) and propagation (eqs 4 and 11) as for the fixed-frequency calculations, we get the time dependencies of all coordinates and their Fourier transforms. In case of a complex fast Fourier transform (FFT), for example, we record the coordinates in N time steps, $q_{\lambda\alpha,k} = q_{\lambda\alpha}(t_k)$, $k = 0 \dots N - 1$, and obtain the transform as $\hat{q}_{\lambda\alpha,n} = \frac{1}{N} \sum_{n=0}^{N-1} q_{\lambda\alpha,k} e^{i\omega_n t_k}$, with $\omega_n = \frac{2\pi n}{T}$, $t_k = k\Delta$, and $T = N\Delta$. Realizing that $q_{\lambda\alpha,k} = \sum_{n=0}^{N-1} \hat{q}_{\lambda\alpha,n} e^{-i\omega_n t_k}$ and $q_{\lambda\alpha,k} = -i \sum_{n=0}^{N-1} \omega_n \hat{q}_{\lambda\alpha,n} e^{-i\omega_n t_k}$, the absorption spectrum at ω_n according to eq 5b, for example, is

$$w(\omega) = \gamma \omega_n^2 \sum_{\lambda,\alpha} m_\lambda \hat{q}_{\lambda\alpha,n}^2 \quad (13)$$

2.5. Atom-Absorbed/Dispersed Energy and Band Assignment. Dividing the absorbed/scattered energies (eq 5b) to individual atoms and integrating over a frequency interval (ω_1, ω_2), we obtain an atomic contribution as

$$w_\lambda = \gamma \int_{\omega_1}^{\omega_2} \int_0^T \sum_\alpha m_\lambda q_{\lambda\alpha}^2 dt d\omega \quad (14)$$

This makes it possible to link spectral intensities to particular molecular parts. For plotting, normalized atomic

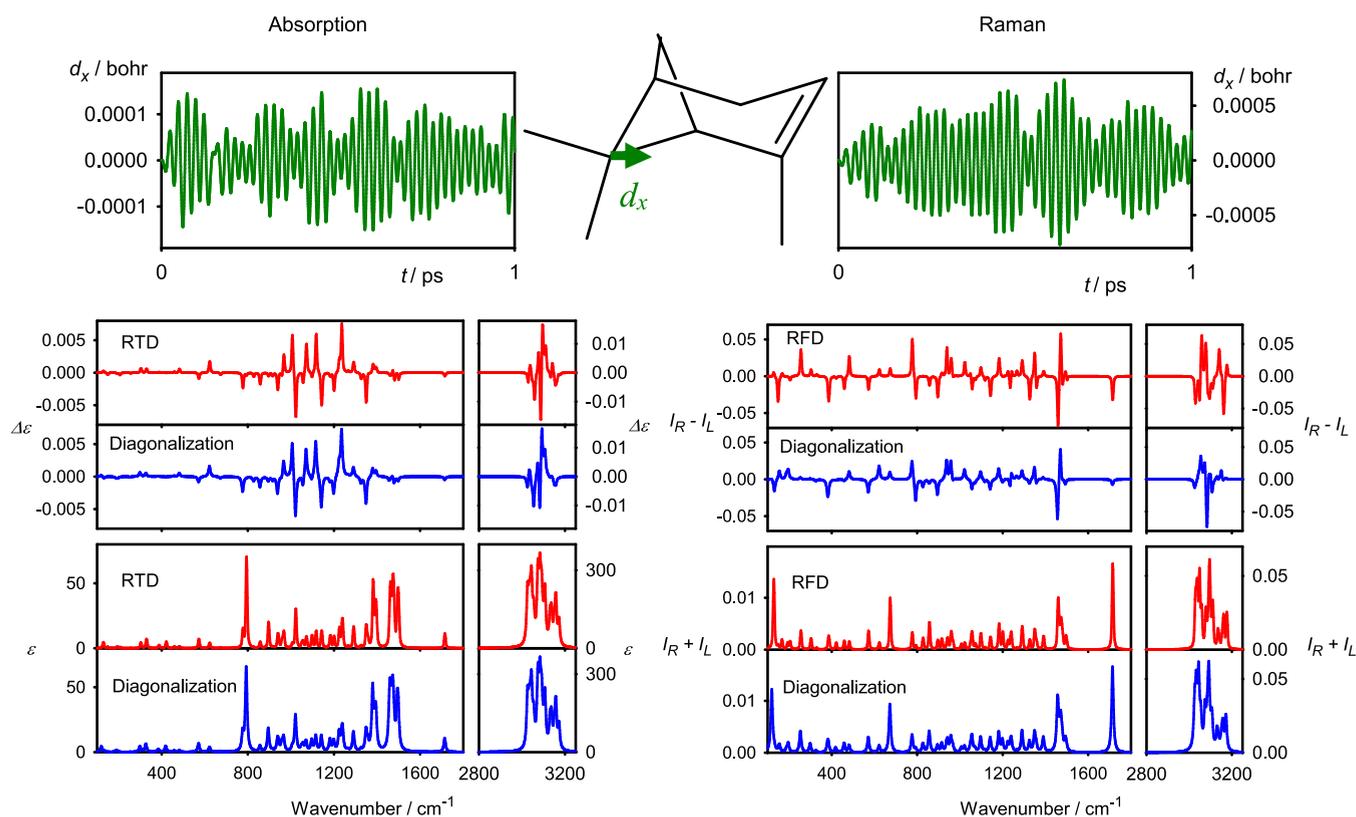


Figure 2. (1S)-(-)- α -Pinene, absorption (ϵ), VCD ($\Delta\epsilon$), Raman ($I_R + I_L$), and ROA ($I_R - I_L$) spectra simulated with real-field dynamics and by direct diagonalization. At the top, the x -coordinate of a carbon atom as a function of time is plotted, with an example for $\omega = 1600 \text{ cm}^{-1}$.

energies $w'_\lambda = (w_\lambda - w_{\min}) / (w_{\max} - w_{\min}) - 1/2$ expressed as false charges $q_\lambda = w'_\lambda - \langle w \rangle$ appeared to be particularly useful, where w_{\min}/w_{\max} is minimum/maximum of w_λ , and $\langle w \rangle = \frac{1}{N_{\text{at}}} \sum_{\lambda=1}^{N_{\text{at}}} w'_\lambda$. As a whole, the molecule is then “electrically neutral” ($\sum_{\lambda=1}^{N_{\text{at}}} q_\lambda = 0$) so that a positive “electrostatic potential” indicates spectrally active molecular parts. In the present study, the atoms were colored directly depending on the energy/charge magnitudes using the VMD software.²⁸

Alternatively, complex vibrational normal mode-like displacement was extracted for each frequency J as a Fourier transform of time-dependent coordinates, $\tilde{S}'_{\lambda J} = \int q_\lambda(t) e^{i\omega t} dt$. A dimensionless matrix $\tilde{s}'_{\lambda J} = \sqrt{m_\lambda} \tilde{S}'_{\lambda J}$ is obtained and normalized, $\tilde{s}_{\lambda J} = A_J \tilde{s}'_{\lambda J}$; A_J is a constant, so that $\sum_\lambda \tilde{s}_{\lambda J}^* \tilde{s}_{\lambda J} = 1$, and then it is transformed back to the Cartesian-normal mode matrix $\tilde{S}_{\lambda J} = \tilde{s}_{\lambda J} / \sqrt{m_\lambda}$. From the complex $\tilde{S}_{\lambda J} = \text{Re}[\tilde{S}_{\lambda J}] + i \text{Im}[\tilde{S}_{\lambda J}] = S_{\lambda Jx} + iS_{\lambda Jy}$, we obtain real displacements, $S_{\lambda J}$, when we calculate a 2×2 moment in the complex plane, $M_{ab} = \sum_\lambda S_{\lambda J a} S_{\lambda J b}^*$, diagonalize it so that $\sum_b M_{ab} v_{kb} = e_k v_{ka}$ with e_k and v_{ka} as eigenvalue and eigenvector elements, and make a projection

$$S_{\lambda J} = \sum_a v_{1a} S_{\lambda J a}, \quad (15)$$

where v_{1a} corresponds to the biggest eigenvalue e_1 . By this procedure, we implicitly extract the dominant phase for each frequency, $\varphi = \text{atan} 2(v_{12}, v_{11})$, and relate the atomic motions to it.

2.6. Calculations. The molecular property tensors (force field and dipole/quadrupole and polarizability derivatives)

were calculated by Gaussian software.²⁹ The equations of motion were coded in a Fortran program, enabling parallelization within an Open-MP environment.

α -Pinene was chosen for primary tests, where the force field, dipole and polarizability derivatives were obtained at the B3PW91/6-311++G** computational level. The CPCM solvent model with parameters for pentanoic acid was used to mimic neat α -pinene measurement. The real field dynamics was performed with a 0.25 fs time step, 20,000 steps in total, for each of 1675 frequencies evenly spread within the 100–3250 cm^{-1} interval (2 cm^{-1} increment). In atomic units, we set the values $\gamma = 1 \times 10^{-5}$, $E_0 = 1 \times 10^{-4}$, and $\delta = 2 \times 10^{-4}$. The field strength (within some reasonable interval) affects amplitudes of the atomic oscillations but not vibrational frequencies and spectral intensities, as these are given by the harmonic force field and dipole/polarizability derivatives.

For comparison, exact IR, VCD, Raman, and ROA intensities of α -pinene were calculated from the harmonic force field and dipole/polarizability derivatives.²⁵ For larger proteins (insulin, lysozyme, concavalein A, and human serum albumin), the vibrational parameters were taken from previous works^{5,6} as obtained by the CCT³⁰ method. To avoid problems with negative frequencies, these were projected out from the force fields. By default, 40,000 RFD steps were performed, otherwise with the same parameters as for α -pinene.

3. RESULTS AND DISCUSSION

3.1. Comparison of the Exact and Real Field Dynamics Spectra. For α -pinene, we compare the RFD spectra with those obtained by convolution of exact line intensities with Lorentzian functions of 10 cm^{-1} full width at

half maximum in Figure 2. The absorption (ϵ) and VCD ($\Delta\epsilon$) spectra obtained by the two methods are almost indistinguishable. However, RFD band frequencies were slightly dependent on the integration time step, similarly as documented in a previous study.¹⁸ The fast C–H stretching vibrations ($\sim 3000\text{ cm}^{-1}$) are affected the most. For example, increasing the time step from 0.25 to 0.5 fs shifted the C–H stretching bands by about 8 cm^{-1} farther from the exact results, whereas for low and mid-IR frequencies, the error ($\sim 1\text{ cm}^{-1}$) was rather negligible. A small frequency error ($\sim 3\text{ cm}^{-1}$) is also introduced by the stabilization constant (δ , cf. eq 2a). The stabilization seems to be the price to pay for the dynamics in the Cartesian coordinates, unlike for common MD simulations based on an internal coordinate force field. An introduction of internal coordinates, however, would lead to matrix multiplications scaled as $\sim (3N)^3$ and thus destroy a principal RFD advantage.

The RFD Raman spectra are also quite close to the exact results, whereas ROA intensities differ most, particularly for higher frequencies ($>2800\text{ cm}^{-1}$; right side of Figure 2). The intensity errors in the Raman and ROA cases can be partially attributed to the isotropic averaging. Whereas absorption and VCD averaging of the light propagation in all three x , y , and z directions provides the exact result, only an approximate average is obtained for Raman/ROA intensities comprising higher-order tensor sums.²⁰ For example, to average the dipole strength in eq 6, we need an average of the directional cosine elements i_β between laboratory and molecular coordinate axes, $\langle i_{\beta j} \rangle = \frac{1}{3}\delta_{ij}$, where $i, j = x, y, z$ are the molecular coordinates. The average strength is then $\langle D \rangle = \frac{1}{3}(D_x + D_y + D_z)$, the same as the average of dipole strengths for samples with the transition dipole oriented along the X , Y , and Z space axes, $\frac{1}{3}(D_X + D_Y + D_Z)$; $D_i = \langle 0|\mu_i|1\rangle\langle 1|\mu_i|0\rangle$. For Raman, the intensity is proportional to expressions such as $\alpha_{XX}\alpha_{XX}$,²⁰ where α_{XX} is the transition polarizability in the laboratory coordinates and cosine element averages, such as $\langle i_{Xj}k_{Xl} \rangle = \frac{1}{15}(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$, are needed. For this example, we get the isotropic results as $\langle \alpha_{XX}\alpha_{XX} \rangle = \frac{1}{15} \sum_j \sum_i (\alpha_{ii}\alpha_{jj} + 2\alpha_{ij}\alpha_{ji})$, which is not proportional to $(\alpha_{xx}\alpha_{xx} + \alpha_{yy}\alpha_{yy} + \alpha_{zz}\alpha_{zz})$ obtainable from the three oriented cases. The Raman/ROA results may be improved by taking into account a larger number of molecular orientations/light directions, which would, however, make the computations quite expensive.

Overall, however, compared to previous time-dependent methods (e.g., refs 18 and 19), the frequency and intensity errors for the present RFD seem to be quite small and controllable, and the implementation is simpler. The method is based on Newtonian molecular mechanics, mimicking actual experiment: inspecting the atomic paths (Figure 2, top), one can see as the molecule starts to oscillate when the light fields are applied, and the amplitude quickly stabilizes due to the friction constant (γ).

3.2. Computer Memory and Time Considerations. In our implementation of the algorithm (eqs 1–12), we calculate spectra intensities at each desired frequency ω_n , $n = 1\dots N_\omega$. For absorption and VCD spectra of the isotropic samples, we need six runs of the dynamics: three for x , y , and z light directions and two left and right circular polarizations. For Raman and

ROA, 12 runs are needed as separate simulations for two different incoming light polarizations are additionally needed.

The most time-consuming step is calculation of the force (eq 3a) from the force constants. This involves $3N_{\text{at}}$ multiplications for $3N_{\text{at}}$ atomic coordinates each time step t_k , $k = 1\dots N$. The total computer time is thus proportional to $6 \times (2) \times N_\omega \times N \times 3N_{\text{at}} \times 3N_{\text{at}}$.

The prefactor $6 \times (2) \times N_\omega \times N$ does not depend on molecular size, but is rather large and makes the procedure slow for small molecules. However, as the time ultimately scales as $\sim N_{\text{at}}^2$, the procedure becomes more convenient for large systems and eventually beats the direct diagonalization with the N_{at}^3 scaling. Additionally, force constants between distant atoms can often be neglected, which reduces the force computations and leads to nearly linear scaling! In our computations, we used the 10^{-3} (a.u.) threshold for the neglecting of the force constants.

The computer memory needed is given by the force field dimensions and grows only nearly linearly, too, although “private” copies are created for each thread in a parallel run. Only the fast Fourier transform-based procedure has significantly higher memory demands as all coordinates need to be kept in memory for all time steps.

The times obtained for the four proteins (Figure 3) confirm the theoretical expectations. The spectrum for the largest system (HSA) is generated much faster by RFD (82 h) than by direct diagonalization (277 h), and the RFD wall time can be further reduced by parallelization, lowering the spectrum resolution, or restricting the frequency interval. Spectra generated by RDT and complete diagonalization exhibit minor differences, which is irrelevant for a comparison to an experiment affected by experimental noise.⁵ The spectral quality, not deteriorating for large systems, is quite encouraging compared to the behavior of similar time-dependent methods.¹⁸

On the other hand, the RDT method becomes unreasonably slow for smaller molecules. Also, the method is not well-suited for nonpositive definite force fields. In this study, negative frequencies were simply projected out, which, however, needs computational time comparable to that for the direct diagonalization. Small negative frequencies ($< \sim 100\text{ cm}^{-1}$) can be handled by the $kq_{\lambda\alpha}$ term (eq 2a), but larger ones make the dynamics unstable, which prevents us from advertising RFD for nonequilibrium geometries or as a general method of matrix diagonalization.

From IR peaks obtained by the FFT method, we can get accurate eigenvalues; however, the intensities are only approximate, and the VCD spectrum does not remind the correct one at all (Figure 3, lower part). The chirality of the initial δ -pulse seems to be lost completely during the dynamics. Previously, for time-autocorrelation function methods, application of the δ -pulse has been more beneficial.³¹ However, in the present algorithm, the dynamics has to “remember” the chirality of the initial pulse all the time, which appears difficult for a great number of steps and complex molecules. The FT performance was somewhat better for α -pinene (not shown), and thus, the FFT-RFD combination could be perhaps further developed at least for medium-sized systems. This, however, goes beyond the scope of the present work.

3.3. Parallelization. The possibility of parallelization is a technically minor but practically important feature of the algorithm. In Figure 4, the speedup for 1–20 CPUs is plotted as obtained for the HSA molecule and using a do-cycle

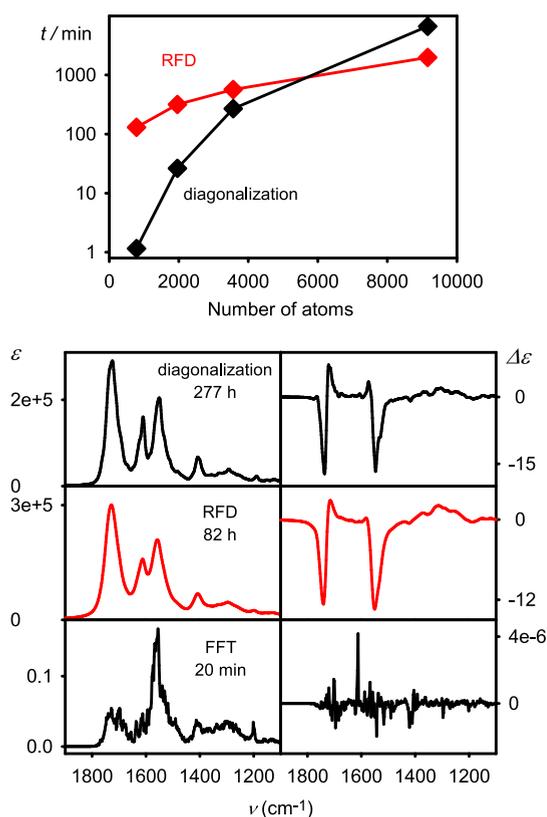


Figure 3. (Top) Times needed for generation of absorption (ϵ) and VCD ($\Delta\epsilon$) spectra from harmonic force fields of insulin (783 atoms), lysozyme (2021 atoms), concavalline A (3566 atoms), and human serum albumin (HSA, 9161 atoms) by RFD and conventional diagonalization. (Bottom) Spectra of HSA, where the results from the fast Fourier transform were added. Note that the FFT absorption intensities are plotted in a different scale, exhibiting large error, and the VCD is unusable. The times are for computations on one processor, RFD was run for 40,000 steps, 1950 frequencies within 100–4000 cm⁻¹, and 0.25 fs integration time.

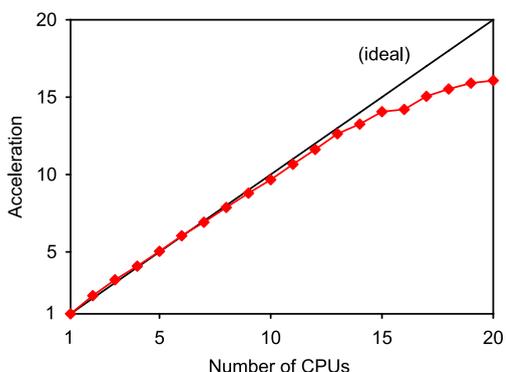


Figure 4. Acceleration of the HSA force field diagonalization achieved on a 20-processor station using an automatic do-cycle (OpenMP) parallelization.

parallelization (external cycle over vibrational frequencies) within the GFortran/OpenMP environment. As can be seen, the speeding becomes less effective but still significant when the number of processors approaches 20. This can be potentially improved in a different computer architecture or by code optimization; nevertheless, the parallelization already seems to be a significant advantage to common full

diagonalization routines, where the parallel formulation is either impossible or significantly limited (cf. refs 7 and 12; <https://nholmber.github.io/categories/parallel-programming/>). Another interesting feature is the acceleration for two and three CPUs in a minor way higher than the “ideal” one, which is probably caused by better memory management during the parallel processing.

3.4. Spectral Band Assignment. Examples of the atomic contributions to Raman/IR spectra based on the absorbed energies (eq 14) and normal mode displacement (eq 15) are provided in Figure 5. The 1014 cm⁻¹ Raman band can be clearly attributed to the three phenylalanine residues (two are visible in the figure). The absorbed energy of the 1716 cm⁻¹ amide I band is spread over more numerous amide groups and primarily localized on the carbonyl carbons. Note that the information from the energies/atomic pseudocharges and the displacements are not equivalent. The normal-mode like displacement of the phenylalanine ring mode, for example, (eq 15; right-hand side of Figure 5) shows that the hydrogens are moving as well, but the space filling/charge representation suggests that they scatter a minor part of the energy only. In both visualization methods, we recognize the trigonal symmetry of this mode (e_{1u} irreducible representation if related to the benzene ring).³²

3.5. Method’s Limitations and Relation to Autocorrelation Approaches. We have seen that within the harmonic approximation, the classical Hamilton’s mechanics may produce spectra practically identical to the quantum-chemical results at least for IR, Raman, and VCD. For small molecules, we did not find significant advantage of the present approach, but for bigger ones, the circumvention of the force field diagonalization led to significant savings of computer time. A limitation is the need of a positively definite force field, i.e., molecular geometry has to be optimized at least to the point that imaginary frequencies are small ($|\omega| < \sim 100$ cm⁻¹). This is problematic for large molecules where the force field is constructed from molecular fragments, but this can be perhaps addressed in the future by improved propagation or optimization schemes.

At this time, we also do not know how the method would behave for general molecular dynamic force fields and freely moving molecule/molecules or for ab initio molecular dynamics, where the forces are generated on the fly. The usual MD approach is based on the time (auto)correlation functions,³³ and although quantum corrections can be added,³⁴ the classical limit within the double-harmonic approximation is usually used. Therefore, we may suppose that the present method would behave similarly and provide similar results as the autocorrelation function approaches. For the case of the harmonic force fields and rigid geometries inspected in this study, autocorrelation appears inferior because of the slow convergence and uneven distribution of the energy among vibrational normal modes observed previously.¹⁸ The forced molecular dynamics invoked by the real electromagnetic fields presented above, in the current implementation effectively at zero temperature, avoids the need for excess averaging of molecular motions. The temperature can be added in principle, but this was not pursued in the current study.

4. CONCLUSIONS

We have formulated a procedure allowing us to efficiently generate vibrational spectra of large molecular systems based on classical Newtonian mechanics, Maxwell equations, and

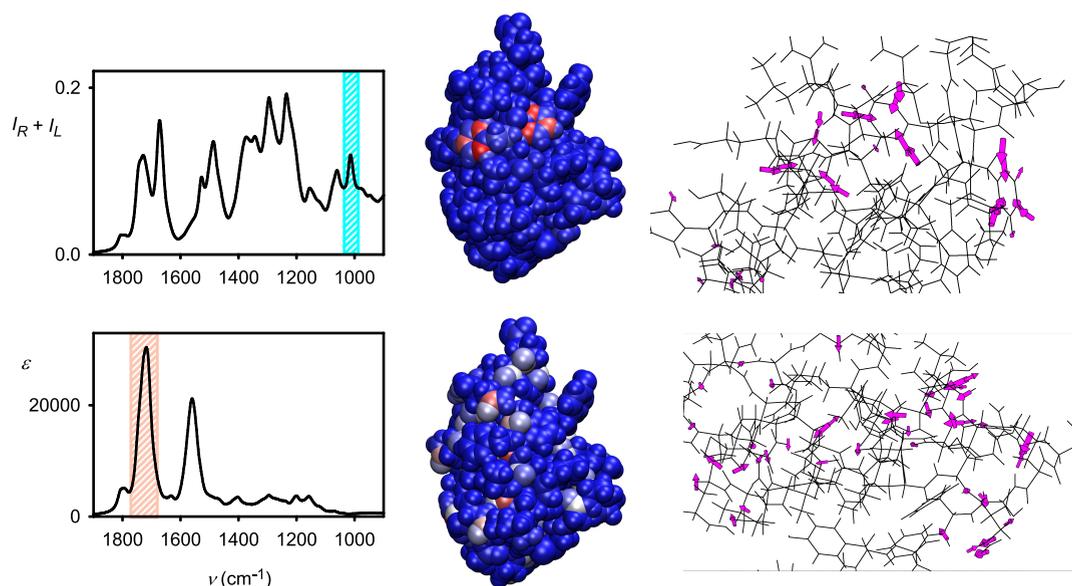


Figure 5. (Left) Example of the band assignment based on the energy scattered/absorbed on individual atoms of insulin (middle; eq 14). Red hues in the space filling model correspond to increased values. While the 1014 cm^{-1} Raman band can be sharply assigned to a ring vibration of phenylalanine side chains, the 1716 cm^{-1} amide I IR absorption band is more delocalized across most of the amide groups. (Right) Normal mode movement extracted from atomic trajectories (eq 15).

harmonic approximation. The method enables one to demonstrate the substance of the interaction of light with matter, including light polarization and vibrational optical activity, and potentially include phenomena beyond dipolar approximation and planar waves. In the present study, we investigated in detail the computer times needed to simulate the vibrational spectra of model protein molecules and the possibility to assign vibrational spectral bands to molecular motion based on the energies absorbed/scattered on particular atoms. Overall, the algorithm appears as a useful tool for simulation and understanding of molecular vibrational properties at the harmonic level.

AUTHOR INFORMATION

Corresponding Author

Petr Bouř – Institute of Organic Chemistry and Biochemistry, Academy of Sciences, 16610 Prague, Czech Republic;
 orcid.org/0000-0001-8469-1686; Email: bour@uochb.cas.cz

Author

Jiří Kessler – Institute of Organic Chemistry and Biochemistry, Academy of Sciences, 16610 Prague, Czech Republic;
 orcid.org/0000-0001-6307-4339

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.jctc.1c01138>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The work was supported by the Ministry of Education (CZ.02.1.01/0.0/0.0/16_019/0000729) and the Science Foundation (20-10144S) of the Czech Republic.

REFERENCES

- (1) Luber, S.; Reiher, M. Theoretical Raman optical activity study of the β domain of rat metallothionein. *J. Phys. Chem. B* **2010**, *114*, 1057–1063.
- (2) Vreven, T.; Morokuma, K. On the Application of the IMOMO (Integrated Molecular Orbital + Molecular Orbital) Method. *J. Comput. Chem.* **2000**, *21*, 1419–1432.
- (3) Jose, K. V. J.; Raghavachari, K. Molecules-in-molecules fragment-based method for the calculation of chiroptical spectra of large molecules: Vibrational circular dichroism and Raman optical activity spectra of alanine polypeptides. *Chirality* **2016**, *28*, 755–768.
- (4) Thapa, B.; Beckett, D.; Jovan Jose, K. V.; Raghavachari, K. Assessment of Fragmentation Strategies for Large Proteins Using the Multilayer Molecules-in-Molecules Approach. *J. Chem. Theory Comput.* **2018**, *14*, 1383–1394.
- (5) Kessler, J.; Andrushchenko, V.; Kapitán, J.; Bouř, P. Insight into vibrational circular dichroism of proteins by density functional modeling. *Phys. Chem. Chem. Phys.* **2018**, *20*, 4926–4935.
- (6) Kessler, J.; Kapitán, J.; Bouř, P. First-principles predictions of vibrational Raman optical activity of globular proteins. *J. Phys. Chem. Lett.* **2015**, *6*, 3314–3319.
- (7) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P., *Numerical Recipes in Fortran*. 2nd ed.; Cambridge University Press: New York, 1992.
- (8) Weiße, A.; Fehske, H., *Exact Diagonalization Techniques. In Lecture Notes in Physics*, Scalone, L., Ed. Springer: Heidelberg, Germany, 2007; Vol. 739, pp. 529–544.
- (9) Cullum, J. K.; Willoughby, R. A., *Lanczos Algorithms for Large Symmetric Eigenvalue Computations*. Birkhäuser: Boston, 1984; Vol. I&II.
- (10) Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* **1975**, *17*, 87–94.
- (11) Mitin, A. V. Iterative methods for the calculation of a few of the lowest eigenvalues and corresponding eigenvectors of the $AX = \lambda BX$ equation with real symmetric matrices of large dimension. *J. Comput. Chem.* **1994**, *15*, 747–751.
- (12) Saad, Y., *Numerical Methods for Large Eigenvalue Problems*. University Press: Manchester, 1992.
- (13) Tubman, N. M.; Freeman, C. D.; Levine, D. S.; Hait, D.; Head-Gordon, M.; Whaley, K. B. Modern Approaches to Exact

Diagonalization and Selected Configuration Interaction with the Adaptive Sampling CI Method. *J. Chem. Theory Comput.* **2020**, *16*, 2139–2159.

(14) McQuarrie, D. A., *Statistical Thermodynamics*. Harper and Row: New York, 1973.

(15) Fischer, S. A.; Ueltschi, T. W.; El-Khoury, P. Z.; Mifflin, A. L.; Hess, W. P.; Wang, H. F.; Cramer, C. J.; Govind, N. Infrared and Raman Spectroscopy from Ab Initio Molecular Dynamics and Static Normal Mode Analysis: The C–H Region of DMSO as a Case Study. *J. Phys. Chem. B* **2016**, *120*, 1429–1436.

(16) Praprotnik, M.; Janežič, D. Molecular dynamics integration and molecular vibrational theory. III. The infrared spectrum of water. *J. Chem. Phys.* **2005**, *122*, 174103.

(17) Inakollu, V. S. S.; Yu, H. A systematic benchmarking of computational vibrational spectroscopy with DFTB3: Normal mode analysis and fast Fourier transform dipole autocorrelation function. *J. Comput. Chem.* **2018**, *39*, 2067–2078.

(18) Horníček, J.; Kaprálová, P.; Bouř, P. Simulations of vibrational spectra from classical trajectories: Calibration with ab initio force fields. *J. Chem. Phys.* **2007**, *127*, No. 084502.

(19) Kubelka, J.; Bouř, P. Simulation of Vibrational Spectra of Large Molecules by Arbitrary Time Propagation. *J. Chem. Theory Comput.* **2009**, *5*, 200–207.

(20) Barron, L. D., *Molecular Light Scattering and Optical Activity*. Cambridge University Press: Cambridge, UK, 2009.

(21) Hentschke, R., *Classical Mechanics*. Springer: Cham, 2017.

(22) Stephens, P. J. Theory of Vibrational Circular Dichroism. *J. Phys. Chem.* **1985**, *89*, 748–752.

(23) Keiderling, T. A.; Bouř, P. Theory of Molecular Vibrational Zeeman Effects as Measured with Circular Dichroism. *Phys. Rev. Lett.* **2018**, *121*, No. 073201.

(24) Pastor, R. W.; Brooks, B. R.; Szabo, A. An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.* **1988**, *65*, 1409–1419.

(25) Nafie, L., *Vibrational optical activity: Principles and applications*. Wiley: Chichester, 2011.

(26) Papoušek, D.; Aliev, M. R., *Molecular Vibrational/Rotational Spectra*. Academia: Prague, 1982.

(27) Charney, E., *The Molecular Basis of Optical Activity*. Wiley-Interscience: New York, 1979.

(28) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(29) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. A.03*, Gaussian, Inc.: Wallingford, CT, 2016.

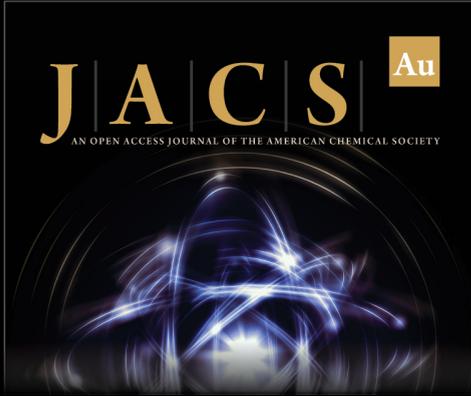
(30) (a) Bouř, P.; Sopková, J.; Bednářová, L.; Maloň, P.; Keiderling, T. A. Transfer of molecular property tensors in Cartesian coordinates: A new algorithm for simulation of vibrational spectra. *J. Comput. Chem.* **1997**, *18*, 646–659. (b) Yamamoto, S.; Li, X.; Ruud, K.; Bouř, P. Transferability of various molecular property tensors in vibrational spectroscopy. *J. Chem. Theory Comput.* **2012**, *8*, 977–985.

(31) Mattiat, J.; Luber, S. Vibrational (resonance) Raman optical activity with real time time dependent density functional theory. *J. Chem. Phys.* **2019**, *151*, 234110.

(32) Sakamoto, A.; Tasumi, M. Symmetry of the benzene ring and its normal vibrations: The “breathing” mode is not always a normal vibration of a benzene ring. *J. Raman Spectr.* **2021**, 2282.

(33) (a) Abbate, S.; Longhi, G.; Kwon, K.; Moscovitz, A. The use of cross-correlation functions in the analysis of circular dichroism spectra. *J. Chem. Phys.* **1998**, *108*, 50–62. (b) Scherrer, A.; Vuilleumier, R.; Sebastiani, D. Vibrational circular dichroism from ab initio molecular dynamics and nuclear velocity perturbation theory in the liquid phase. *J. Chem. Phys.* **2016**, *145*, No. 084101.

(34) Ramirez, R.; López-Ciudad, T.; Kumar, P.; Marx, D. Quantum corrections to classical time-correlation functions: Hydrogen bonding and anharmonic floppy modes. *J. Chem. Phys.* **2004**, *121*, 3973–3983.



JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Editor-in-Chief
Prof. Christopher W. Jones
Georgia Institute of Technology, USA

Open for Submissions

pubs.acs.org/jacsau ACS Publications
Most Trusted. Most Cited. Most Read.