

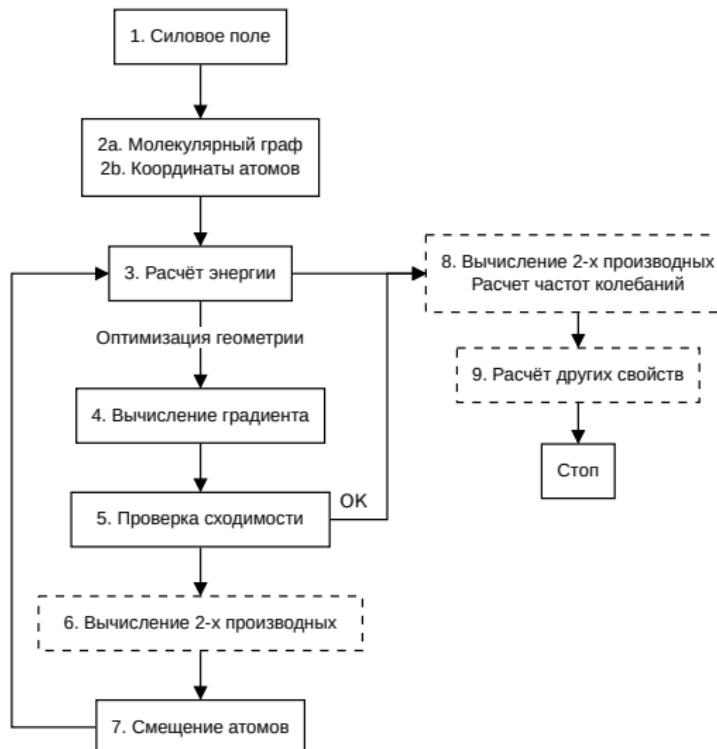
Этапы моделирования молекулярных систем

Иван Федягин

ВХК РАН

2020-04-09

Этапы расчёта



- Для малых молекул:
 - MM2, MM3, MM4 – Силовые поля, разработанные Н. Эллинджером¹
 - UFF: “Universal force field”², параметры задаются для атомов, остальные вычисляются
 - MMFF: MMFF94 – разработано в компании Merck на основе MM3²
 - GAFF: “General Amber Force field”, силовое поле AMBER, только более универсальное²
 - Для макромолекул (белки, липиды и т. д.)
 - GROMOS из пакета программ GROMACS
 - CHARMM
 - AMBER
 - Для расчёта энергии решётки кристаллов
 - UNI из программы Mercury



¹У. Буркерт, Н. Эллинджер, Молекулярная механика. Пер. с англ. - М.: Мир, 1986.

² Доступны в openbabel

MMFF94 — «Хороший выбор по умолчанию для органических молекул» (с сайта программы Avogadro)

$$V_b(r) = \frac{1}{2} 143.9325 k_b (r - r_0)^2 (1 + cs(r - r_0) + \frac{7}{12} cs^2(r - r_0)^2)$$

cs: кубическая константа

$$V_a(\theta) = \frac{1}{2} 0.043844 k_i (\theta - \theta_0)^2 (1 + cb(\theta - \theta_0))$$

сб: кубическая константа

$$V_{bg}(r_{ab}, r_{bc}, \theta) = 2.51210(k_{ab}\Delta r_{ab} + k_{cb}\Delta r_{bc})\Delta\theta$$

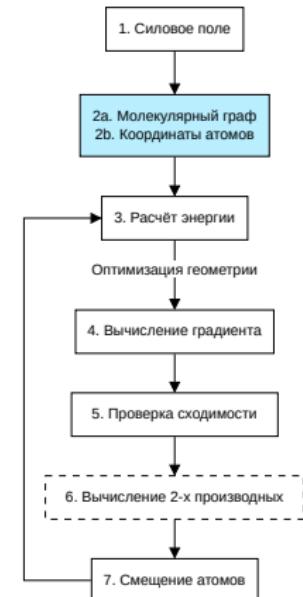
k_{ab} , k_{bc} : константы сгиба-растяжения



Координаты — обычно декартовы, откуда угодно.

SMILES: CC(C)Cc1ccc(cc1)[C@H](C)C(=O)O \Rightarrow

Молекулярный граф: матрица связанности: с типами связей.
Может вычисляться автоматически, но нужно проверять!



Если мы говорим о молекулярной механике, то энергия вычисляется по известной нам формуле:

$$V = V_h + V_g + V_t + (V_{it}) + V_e + V_{vdw} + \dots$$

Расчёте изолированных молекул в ММ — очень быстрый шаг: простые формулы

$$V_b(r) = \frac{1}{2}k_b(r - r_0)^2$$

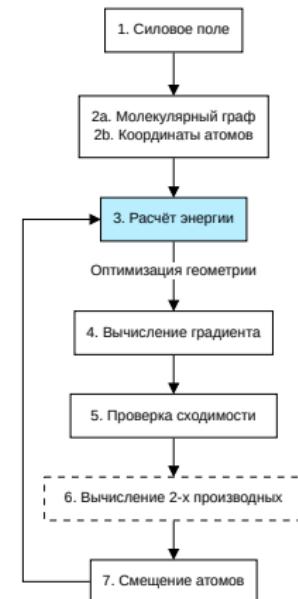
$$V_a(\alpha) = \frac{1}{2}k_a(\alpha - \alpha_0)^2$$

$$V_t(\omega) = k_{t,3}[1 - \cos 3(\omega - \omega_0)]$$

$$V_e(r) = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r_{12}}$$

$$V_{vdw}(r) = \varepsilon \left[\left(\frac{r_0}{r} \right)^{12} - 2 \left(\frac{r_0}{r} \right)^6 \right]$$

При этом нет проблем с V_e и V_{vdw} :

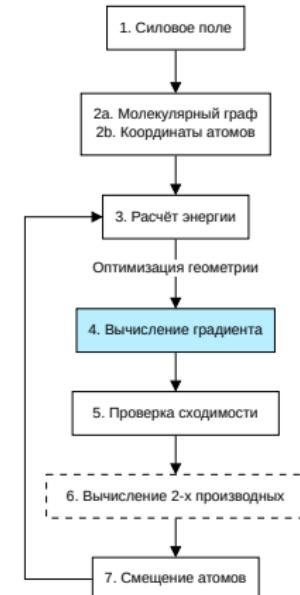
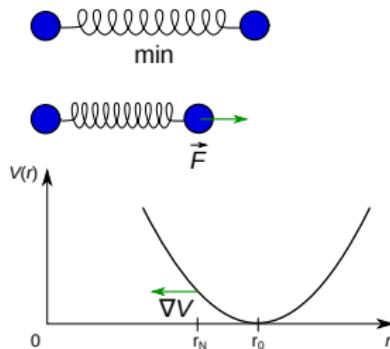


Градиент — векторная величина, показывающая направление возрастания функции.

$$-\nabla_i V = - \left(\frac{\partial V}{\partial x_i}, \frac{\partial V}{\partial y_i}, \frac{\partial V}{\partial z_i} \right) = \vec{F}_i$$

Сила = – градиент

Нам нужен **минимум** энергии, поэтому атомы смещаются в направлении сил.



6. Вычисление 2-х производных

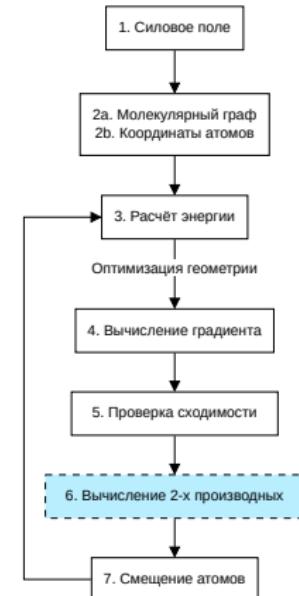
Матрица 2-х производных (матрица Гессе [Hessian matrix]; в просторечии часто просто Гессиан [Hessian]) используется некоторыми алгоритмами оптимизации для ускорения сходимости.

2-е производные быстро аналитически расчитываются в ММ.

В методах квантовой химии — долго и сложно программировать.

Используется приближённая матрица вторых производных: на 1-м шаге из силового поля (ММ пригодилась!), а дальше обновляется с учётом смещений атомов и градиента выполненного шага. Даёт хорошую сходимость при оптимизации.

Численное вычисление 2-х производных (мин. $3N$ вычислений градиента!) не используется при оптимизации геометрии.



Направление и величина смещения определяются алгоритмом оптимизации.

- Линейные методы требуют только знания градиента ∇V
 - Метод наискорейшего спуска [Steepest descent, SD]: хорошо сходится при больших значениях ∇V . Плохо вблизи пологих минимумов.

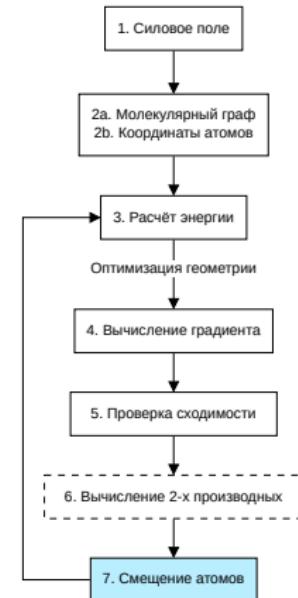
$$x = x_0 + c \nabla V(x_0)$$

- метод связанных градиентов [conjugate gradient, CG] — требует запоминания значений градиента для нескольких точек. Сходится намного быстрее SD.
 - Квадратичные методы требуют знания вторых производных

$$x = x_0 + \mathcal{V}''^{-1} \nabla V(x_0)$$

- Псевдо-квадратичные методы (quasi-NR), BGFS, RFO и др. Матрица \mathcal{V}'' или её производные оцениваются по предыдущим шагам.

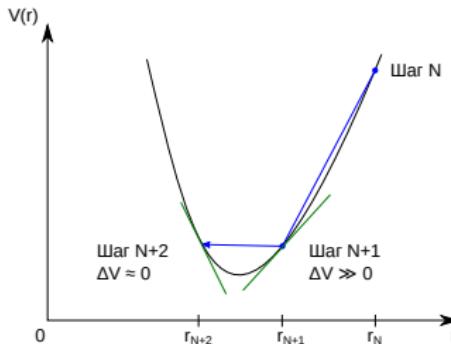
В сложных случаях или для поиска локальных минимумов может быть использован “trust radius” — максимально допустимая величина смещения $x - x_0$ (фиксированная или зависящая от других параметров или числа прошедших циклов оптимизации).



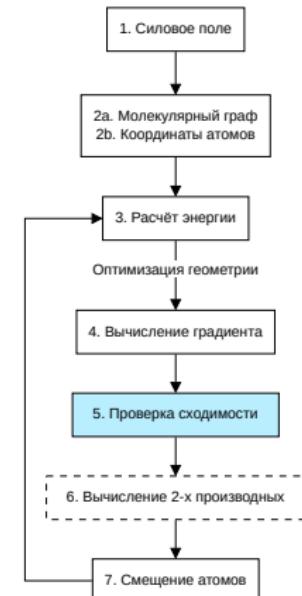
5. Проверка сходимости

Критерии сходимости

- Изменение энергии по сравнению с предыдущим шагом: плохой и ненадёжный критерий.



- Максимальные и среднеквадратичные значения градиента / сил на атомах.
- Максимальная и среднеквадратичная величина смещения.
- Число циклов оптимизации



Примеры предельных значений в различных программах:

● **obminimize (openbabel/Avogadro):**

по энергии 1×10^{-7} «единиц» или 2500 шагов

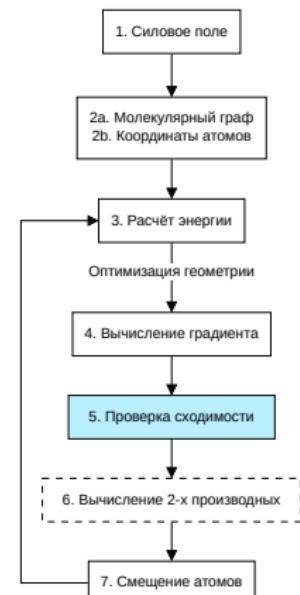
(из кода: ккал/моль или кДж/моль, зависит от силового поля + внутреннее ограничение на градиент в 1×10^{-2}

● **Gaussian**

Критерий	“Loose”	Обычный	“Tight”	“VeryTight”
Maximum Force	0.002500	0.000450	0.000015	0.000002
RMS Force	0.001667	0.000300	0.000010	0.000001
Maximum Displacement	0.010000	0.001800	0.000060	0.000006
RMS Displacement	0.006667	0.001200	0.000040	0.000004

● **Orca**

Критерий	“Loose”	Обычный	“VeryTight”
Energy change	0.000005		
RMS gradient	0.000100		
MAX gradient	0.000300		
RMS step	0.002000		
MAX step	0.004000		

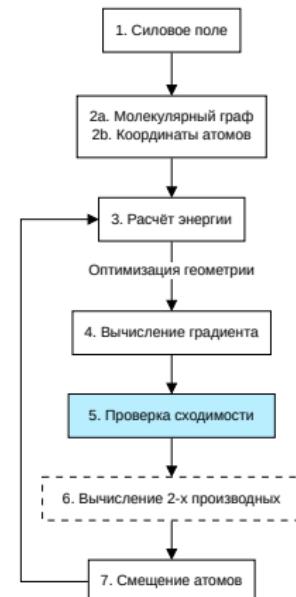


Примеры

Geometry convergence			
Item	value	Tolerance	Converged
Energy change	-0.0000071240	0.0000050000	NO
RMS gradient	0.0000817698	0.0001000000	YES
MAX gradient	0.0002887703	0.0003000000	YES
RMS step	0.0007282607	0.0020000000	YES
MAX step	0.0021850833	0.0040000000	YES

Max(Bonds)	0.0003	Max(Angles)	0.07
Max(Dihed)	0.13	Max(Improp)	0.00

Вывод комплекса программ ORCA



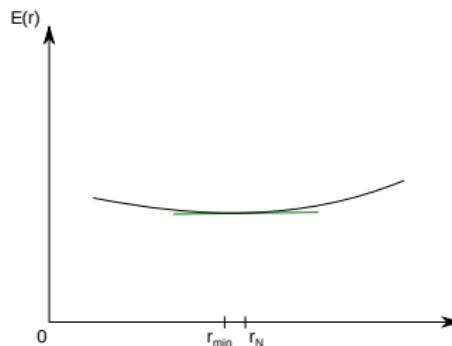
Примеры

```

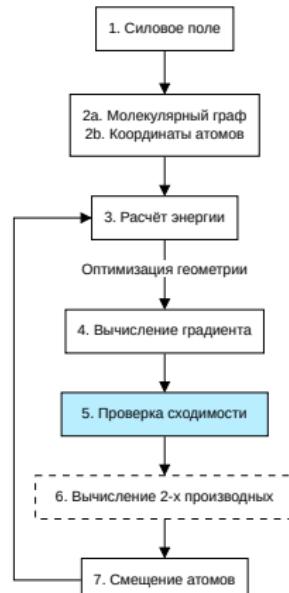
Item           Value   Threshold Converged?
Maximum Force    0.000000  0.000015  YES
RMS Force       0.000000  0.000010  YES
Maximum Displacement 0.001386  0.000060  NO
RMS Displacement 0.000209  0.000040  NO
Predicted change in Energy=-1.507965D-11
Optimization completed on the basis of negligible forces.
-- Stationary point found.

```

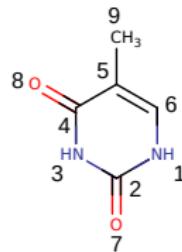
Вывод комплекса Gaussian: значения в а. е.



Пологая ППЭ: алгоритм «перепрыгивает» минимум



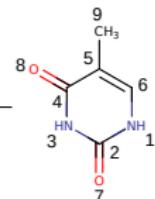
Чем плохо? Возможно наличие мнимых частот. Очень плохо? Нет.



Тимин

Программа: obminimize (<http://openbabel.org>)





Связь	ГЭД+МВ*	X-RAY**	Δ(%)	MP2(full)/cc-pVQZ*	Δ(%)	Δ(%)	PBE0-D3/def2-TZVPP
N1-C2	1.377(3)	1.353	-1.7	1.374	-0.2	1.377	0.0
C2-N3	1.378(3)	1.364	-1.0	1.374	-0.3	1.374	-0.3
N3-C4	1.395(3)	1.389	-0.4	1.392	-0.2	1.393	-0.1
C4-C5	1.466(9)	1.456	-0.7	1.452	-0.9	1.459	-0.5
C5=C6	1.344(16)	1.348	0.3	1.345	0.1	1.343	-0.1
N1-C6	1.372(3)	1.377	0.4	1.369	-0.2	1.369	-0.2
C2=O7	1.210(1)	1.238	2.3	1.212	0.1	1.209	-0.1
C4=O8	1.215(1)	1.226	0.9	1.217	0.1	1.212	-0.3
C5-C9	1.487(8)	1.496	0.6	1.486	-0.1	1.488	0.1
среднее абс.			0.9		0.3		0.2

Связь	ГЭД+МВ	MMFF94	Δ(%)	GAFF	Δ(%)	UFF	Δ(%)
N1-C2	1.377(3)	1.368	-0.7	1.352	-1.8	1.372	-0.4
C2-N3	1.378(3)	1.365	-0.9	1.351	-2.0	1.371	-0.5
N3-C4	1.395(3)	1.367	-2.0	1.351	-3.2	1.373	-1.6
C4-C5	1.466(9)	1.488	1.5	1.476	0.7	1.404	-4.2
C5=C6	1.344(16)	1.339	-0.4	1.38	2.7	1.401	4.2
N1-C6	1.372(3)	1.366	-0.4	1.389	1.2	1.401	2.1
C2=O7	1.210(1)	1.226	1.3	1.212	0.2	1.219	0.7
C4=O8	1.215(1)	1.226	0.9	1.215	0.0	1.22	0.4
C5-C9	1.487(8)	1.492	0.3	1.505	1.2	1.504	1.1
среднее абс.			0.9		1.4		1.7

* N. Vogt et al., *J. Phys. Chem. A*, 2008, 112, 7662–7670

** D. Braun et al., *Cryst. Growth Des.*, 2016, 16, 6, 3480-3496

7. Расчёт частот колебаний

Частоты колебаний

Зачем?

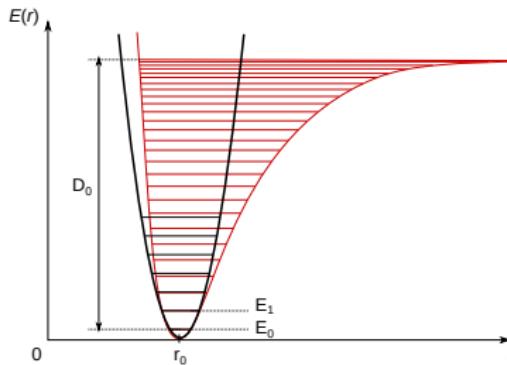
- Для установления типа точки на ППЭ (минимум, переходное состояние)
- Расчёт спектров ИК или КР, сравнение с экспериментом
- Расчёт энергии нулевых колебаний (ZPE)
- Расчёт термодинамических свойств (S, H, G)

Какие?

- Гармонические
- С учётом ангармонизма

Квантовый гармонический осциллятор

$$E = h\nu \left(n + \frac{1}{2} \right) = \hbar\omega \left(n + \frac{1}{2} \right)$$



Молекула HCl: гармоническое приближение и ангармонизм

E_0 — Энергия нулевых колебаний [“Zero-point energy”, ZPE]

D_0 — Энергия диссоциации

Матрица вторых производных

Частоты колебаний можно вычислить, зная геометрию и 2-е производные энергии по всем координатам (т.е. матрицу Гессе) \mathcal{V}'' .

$$\mathcal{V}'' = \begin{pmatrix} \frac{\partial^2 V}{\partial r_1^2} & \frac{\partial^2 V}{\partial r_1 \partial r_2} & \cdots & \frac{\partial^2 V}{\partial r_1 \partial r_n} \\ \frac{\partial^2 V}{\partial r_2 \partial r_1} & \frac{\partial^2 V}{\partial r_2^2} & \cdots & \frac{\partial^2 V}{\partial r_2 \partial r_n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \frac{\partial^2 V}{\partial r_n^2} \end{pmatrix}$$

7. Расчёт частот колебаний

Нормальные моды и частоты колебаний

Нормальные моды — коллективные движения атомов молекулы при возбуждении фотоном определённой частоты.

Нормальным модам соответствуют частоты колебаний.

Обычно вычисления проводятся в декартовых координатах, матрица V'' размерностью $3N \times 3N$ разбивается на блоки 3×3 , соответствующие координатам x_i , y_i и z_i соответствующих атомов.

$$V'' = \left(\begin{array}{ccc|ccc} \frac{\partial^2 V}{\partial x_1^2} & \frac{\partial^2 V}{\partial x_1^1 \partial y_1^1} & \frac{\partial^2 V}{\partial x_1^1 \partial z_1^1} & \frac{\partial^2 V}{\partial x_1^1 \partial x_2^1} & \dots & & \\ \frac{\partial^2 V}{\partial x_1^1 \partial y_1^1} & \frac{\partial^2 V}{\partial y_1^2} & \frac{\partial^2 V}{\partial y_1^1 \partial z_1^1} & \frac{\partial^2 V}{\partial y_1^1 \partial x_2^1} & \dots & & \\ \frac{\partial^2 V}{\partial x_1^1 \partial z_1^1} & \frac{\partial^2 V}{\partial y_1^1 \partial z_1^1} & \frac{\partial^2 V}{\partial z_1^2} & \frac{\partial^2 V}{\partial z_1^1 \partial x_2^1} & \dots & & \\ \hline \frac{\partial^2 V}{\partial x_2^1 \partial x_1^1} & \frac{\partial^2 V}{\partial x_2^1 \partial y_1^1} & \frac{\partial^2 V}{\partial x_2^1 \partial z_1^1} & \frac{\partial^2 V}{\partial x_2^2} & \dots & & \\ \dots & \dots & \dots & \dots & \dots & & \end{array} \right)$$

7. Расчёт частот колебаний

Масс-взвешенные координаты

Система *масс-взвешенных* координат:

$$x_1 = \sqrt{m_1} \Delta x_1,$$

$$y_1 = \sqrt{m_1} \Delta y_1,$$

$$z_1 = \sqrt{m_1} \Delta z_1,$$

$$x_2 = \sqrt{m_2} \Delta x_2 \text{ и т.д.}$$

В такой системе координат

$$\mathbf{F} = \mathbf{M}^{-\frac{1}{2}} \mathbf{v}'' \mathbf{M}^{-\frac{1}{2}}$$

$$\mathbf{M}^{-\frac{1}{2}} = \left(\begin{array}{ccc|cc} \frac{1}{\sqrt{m_1}} & 0 & 0 & 0 & \dots \\ 0 & \frac{1}{\sqrt{m_1}} & 0 & 0 & \dots \\ 0 & 0 & \frac{1}{\sqrt{m_1}} & 0 & \dots \\ \hline 0 & 0 & 0 & \frac{1}{\sqrt{m_2}} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{array} \right)$$

По правилам умножения матриц получается $f_{ij} = v''_{ij} / (\sqrt{m_i} \sqrt{m_j})$

7. Расчёт частот колебаний

От матрицы 2-х производных к частотам и нормальным координатам

$$\mathbf{F} = \mathbf{M}^{-\frac{1}{2}} \mathbf{v}'' \mathbf{M}^{-\frac{1}{2}}$$

Собственные значения матрицы \mathbf{F} связаны с частотами колебаний, а собственные вектора — направления и амплитуды смещений атомов.

Решение векового уравнения

$$|\mathbf{F} - \lambda \mathbf{I}| = 0$$

обычно находят диагонализацией матрицы.

Частоты ν связаны с собственными значениями λ как

$$\nu_i = \frac{\sqrt{\lambda_i}}{2\pi}$$

Собственные значения λ могут быть

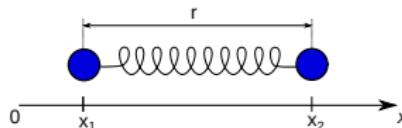
- Положительными: обычное колебание
- Нулевыми: трансляция и/или либрация
- Отрицательными: в переходном состоянии или седловой точке более высокого порядка (напр. $\nu = 50 \text{ i cm}^{-1}$; в программах часто записывается как -50 cm^{-1})

$$\left[E = h\nu \left(n + \frac{1}{2} \right) \right] \Rightarrow E_{ZPE} = \frac{1}{2}h \sum_{i=0}^N \nu_i$$

7. Расчёт частот колебаний

Двухатомная молекула

Для гармонического осциллятора (двуатомной молекулы с гармоническим потенциалом):



Поскольку $r = x_2 - x_1$

$$V(r) = \frac{1}{2}k(r - r_0)^2 \Rightarrow V(x_1, x_2) = \frac{1}{2}k(x_2 - x_1 - r_0)^2$$

$$\frac{\partial^2 V(x_1, x_2)}{\partial x_1^2} = \frac{\partial^2 V(x_1, x_2)}{\partial x_2^2} = k$$

$$\frac{\partial^2 V(x_1, x_2)}{\partial x_1 \partial x_2} = \frac{\partial^2 V(x_1, x_2)}{\partial x_2 \partial x_1} = -k$$

Двухатомная молекула

$$\mathcal{V}'' = \begin{pmatrix} k & -k \\ -k & k \end{pmatrix}$$

$$\mathbf{F} = \mathbf{M}^{-1/2} \mathcal{V}'' \mathbf{M}^{-1/2} = \begin{pmatrix} \frac{k}{m_1} & -\frac{k}{\sqrt{m_1 m_2}} \\ -\frac{k}{\sqrt{m_1 m_2}} & \frac{k}{m_2} \end{pmatrix}$$

$$|\mathbf{F} - \lambda \mathbf{I}| = \begin{vmatrix} \frac{k}{m_1} - \lambda & -\frac{k}{\sqrt{m_1 m_2}} \\ -\frac{k}{\sqrt{m_1 m_2}} & \frac{k}{m_2} - \lambda \end{vmatrix} = 0$$

$$\left(\frac{k}{m_1} - \lambda\right) \left(\frac{k}{m_2} - \lambda\right) - \frac{k^2}{m_1 m_2} = \lambda^2 - \lambda k \left(\frac{1}{m_1} + \frac{1}{m_2}\right) + \frac{k^2}{m_1 m_2} - \frac{k^2}{m_1 m_2} = 0$$

$$\lambda^2 - \lambda k \frac{1}{\mu} = 0$$

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \text{ — приведённая масса}$$

Двухатомная молекула

$$\lambda^2 - \lambda k \frac{1}{\mu} = 0$$

$\lambda_1 = 0 \Rightarrow \nu_1 = 0$: соответствует трансляции молекулы как целого

$$\lambda_2 = \frac{k}{\mu} \Rightarrow \nu_2 = \left[\frac{\sqrt{\lambda}}{2\pi} \right] = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}}$$

7. Расчёт частот колебаний

Учёт ангармонизма

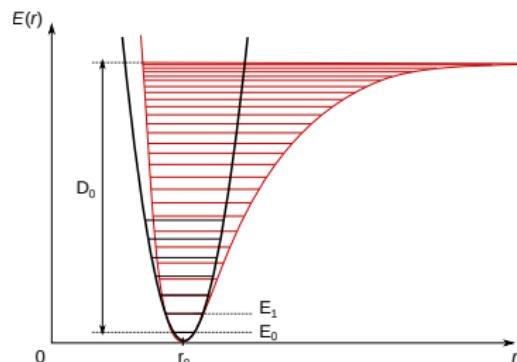
$$V(r) = D_0 \left(1 - e^{-a(r-r_0)}\right)^2$$

$$a = \sqrt{\frac{k}{2D_0}}$$

$$\nu = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}} \Rightarrow \sqrt{k} = 2\pi\nu\sqrt{\mu}$$

$$a = 2\pi\nu \sqrt{\frac{\mu}{2D_0}}$$

$$E = h\nu \left(n + \frac{1}{2}\right) - \frac{h^2\nu^2}{4D_0} \left(n + \frac{1}{2}\right)^2$$



Правила отбора для ИК-спектров в гармоническом приближении: $\Delta l = \pm 1$

Правила отбора для ИК-спектров в ангармоническом приближении: Δl — любой

Расчёт интенсивностей ИК спектров

Правила отбора для ИК-спектров: $\Delta\mu \neq 0$

Интенсивность I пропорциональна квадрату дипольного момента перехода:

$$I \propto |\langle \psi_{final} | \hat{\mu} | \psi_{initial} \rangle|^2$$

В ММ поглощение (интегральная интенсивность) описывается классически через

$$A_i = \frac{1}{nl} \int_{band} \ln(I_0/I) dv = C \left(\frac{\partial \mu}{\partial Q_i} \right)^2$$

Q_i — нормальные координаты колебания с частотой ν_i , n — концентрация, l — длина ячейки

($C = N_A \pi g_i / 3c^2$, g_i — степень вырождения колебания, c — скорость света, N_A —постоянная Авогадро)

$$\mu = \sum_{i=0}^N q_i \mathbf{r}_i$$