Компьютерные методы в фармакологии

В.Б.Сулимов

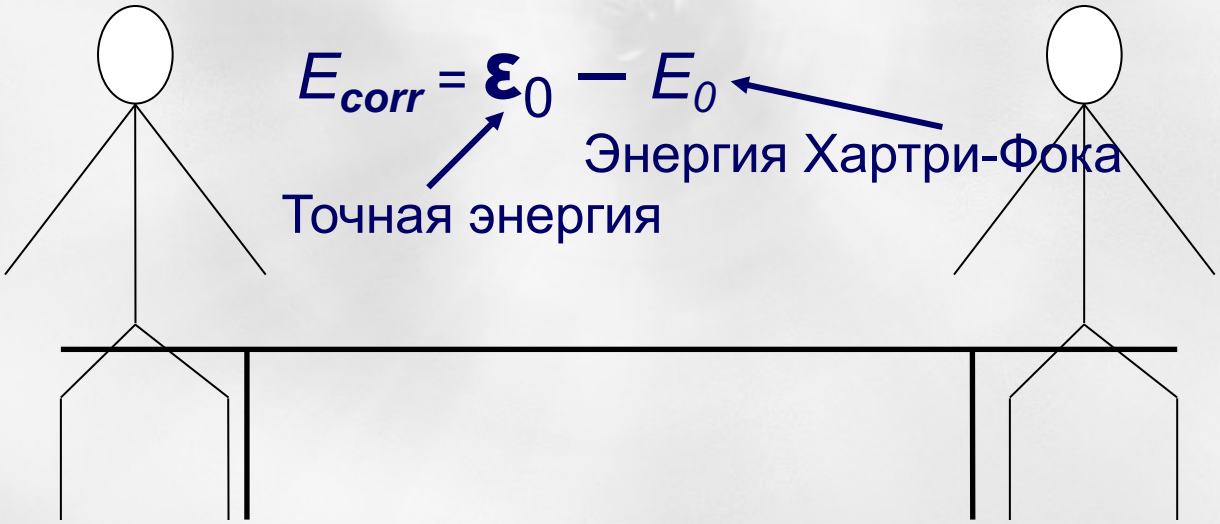
НИВЦ МГУ

Лекция № 10

Методы квантовой химии III

post-HF МЕТОДЫ
ЭЛЕКТРОННАЯ КОРРЕЛЯЦИЯ

В каждый момент времени электрон в молекуле движется, испытывая отталкивание от остальных индивидуальных электронов в системе, а не от усредненного поля электронов. В этом заключается эффект **электронной корреляции** и его учет и является целью более точных методов расчета.

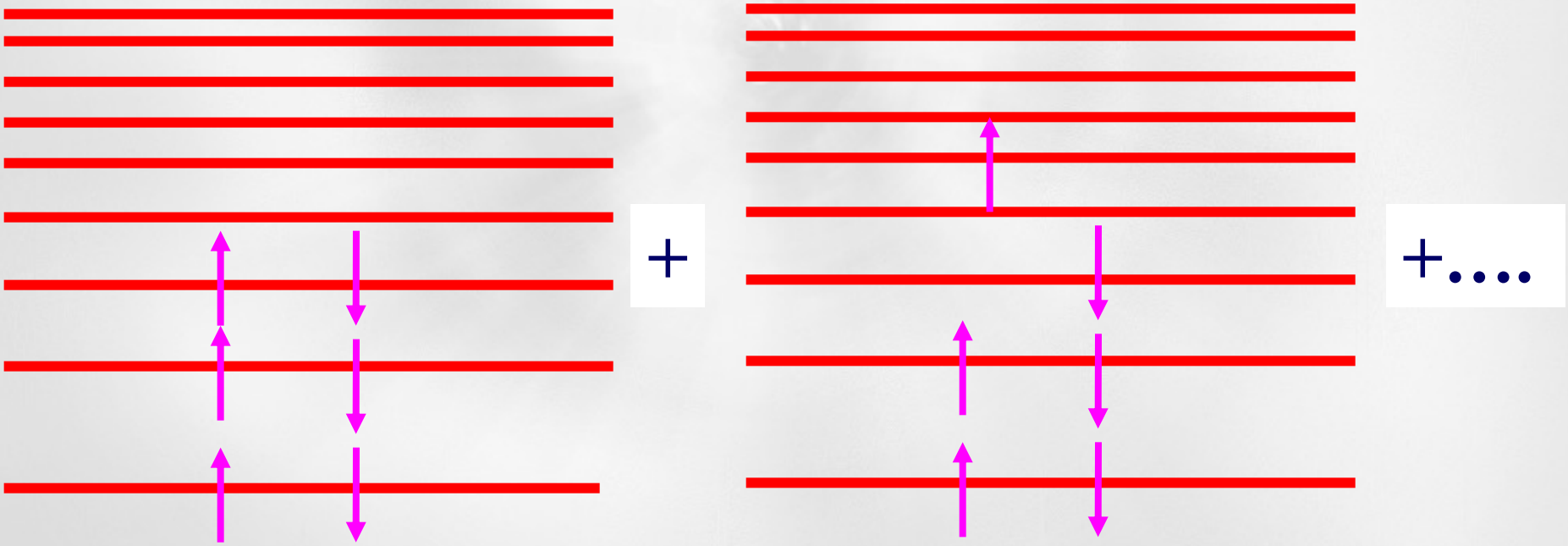


Человеческая отталкивающая корреляция

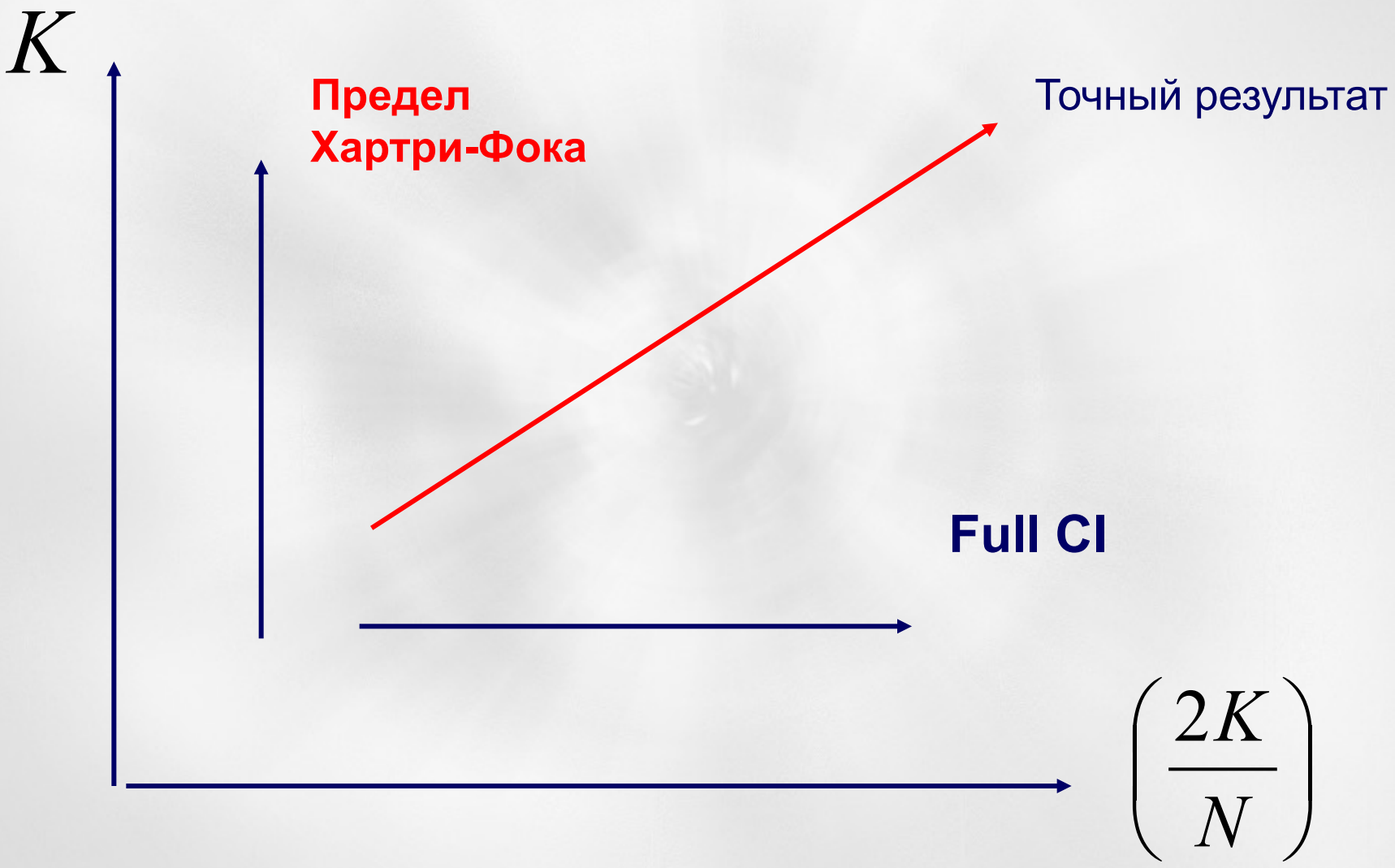
КОНФИГУРАЦИОННОЕ ВЗАИМОДЕЙСТВИЕ (CI) – общий метод учета электронных корреляций

В данном методе **возбужденные состояния** включаются в расчет электронного состояния системы. Тогда волновая функция всей системы является линейной комбинацией волновых функций основного и возбужденного состояния

$$\Psi = c_0\Psi_0 + c_1\Psi_1 + c_2\Psi_2 + \cdots$$



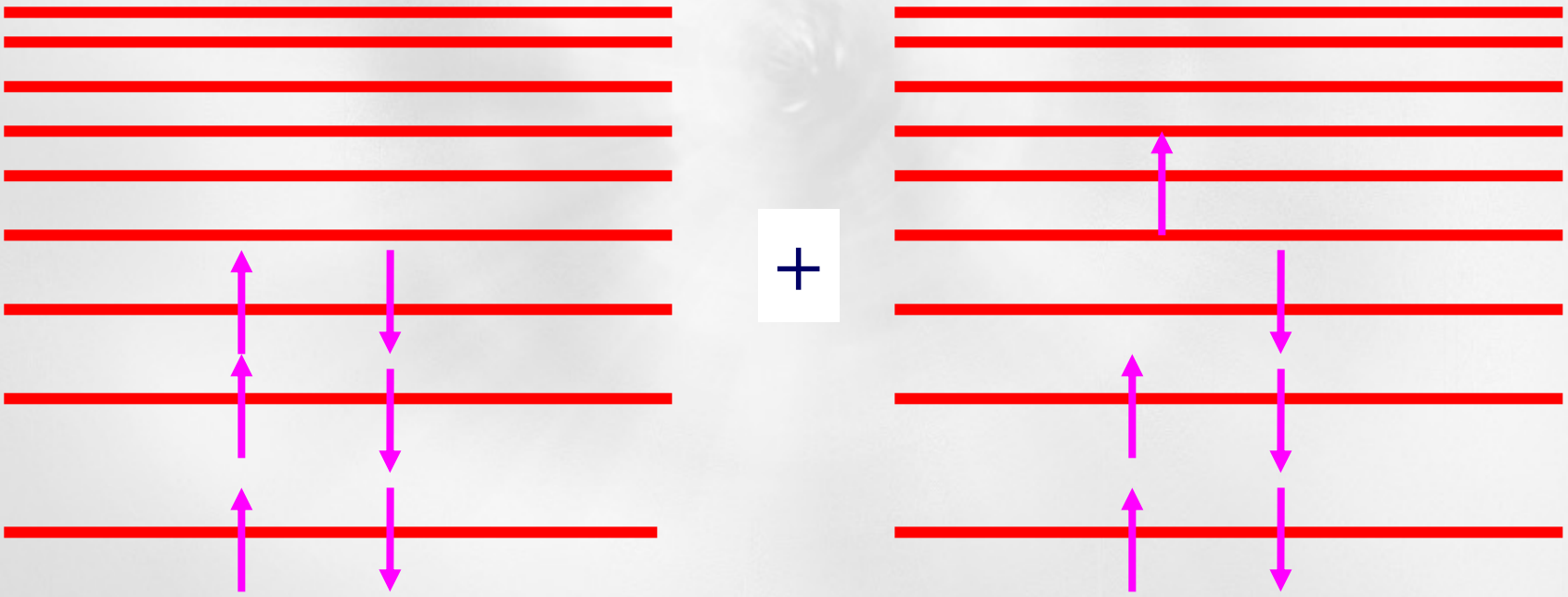
Конфигурационное взаимодействие (CI)



КОНФИГУРАЦИОННОЕ ВЗАИМОДЕЙСТВИЕ (CI)

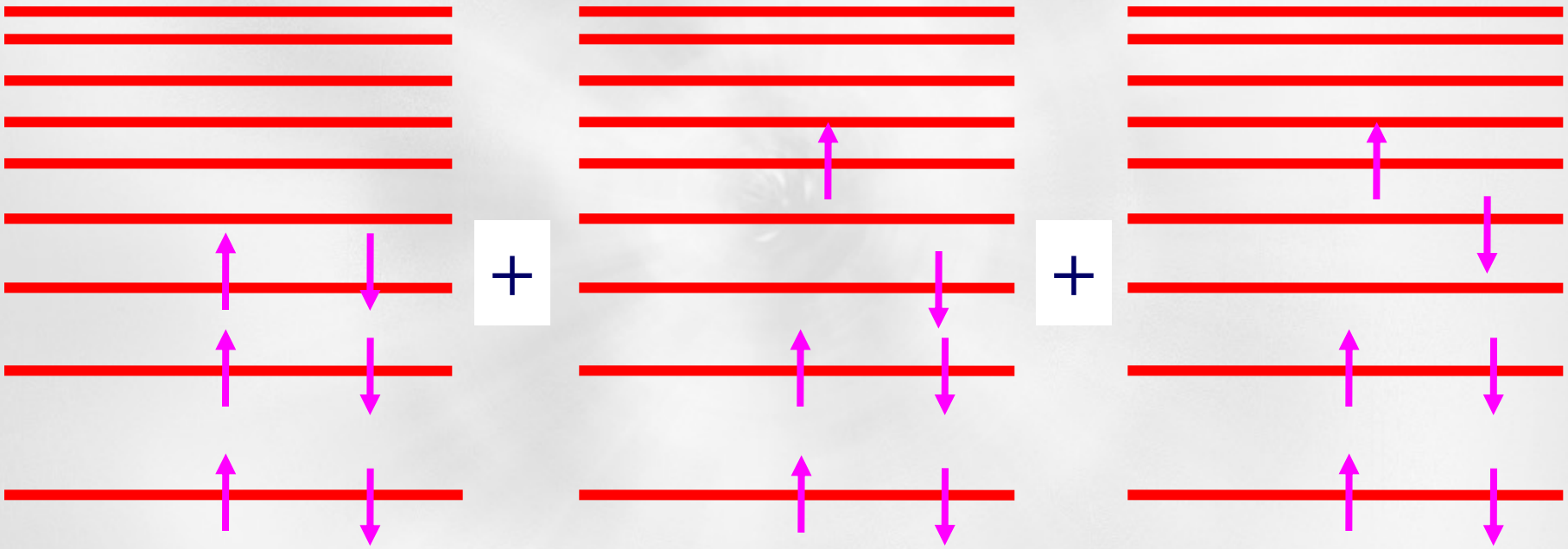
Всеобщей практикой является ограничение рассматриваемых возбужденных состояний

Так, например, в методе **CIS (Configuration Interaction Singles)** рассматривается возбуждение только **одного электрона:**



КОНФИГУРАЦИОННОЕ ВЗАИМОДЕЙСТВИЕ (CI)

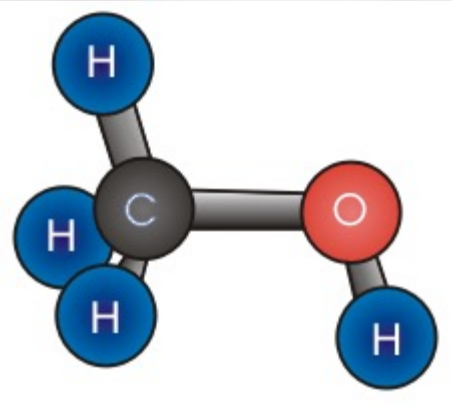
Следующие уровни теории включают **CID (CI Doubles)** или **CISD (CI Singles and Doubles)**



Основное состояние **CIS** **CID**

ПОЛНОЕ КОНФИГУРАЦИОННОЕ ВЗАИМОДЕЙСТВИЕ (Full CI)

Полное конфигурационное взаимодействие возможно только для очень маленьких молекул, поскольку перевод электрона на виртуальную орбиталь может порождать **огромное** число состояний. Например, для метанола базис 6-31G+ дает число базисных функций - 38. Для того, чтобы посчитать полное конфигурационное взаимодействие, надо оптимизировать **2.4 *10^{13}** коэффициентов.



post-HF МЕТОДЫ
МНОГОКОНФИГУРАЦИОННЫЙ МЕТОД ССП

Орбитали, найденные из решения уравнения Хартри-Фока не являются лучшим выбором для вычислений CI.

Многоконфигурационный метод ССП **(Multiconfiguration Self-Consistent Field method MCSCF)**:

$$\Psi_{MCSCF} = \sum_I c_I \left| \Psi_I \right\rangle$$

Чтобы достичь минимума энергии варьируются и коэффициенты c_I и орбитали, входящие в детерминантные функции $\left| \Psi_I \right\rangle$

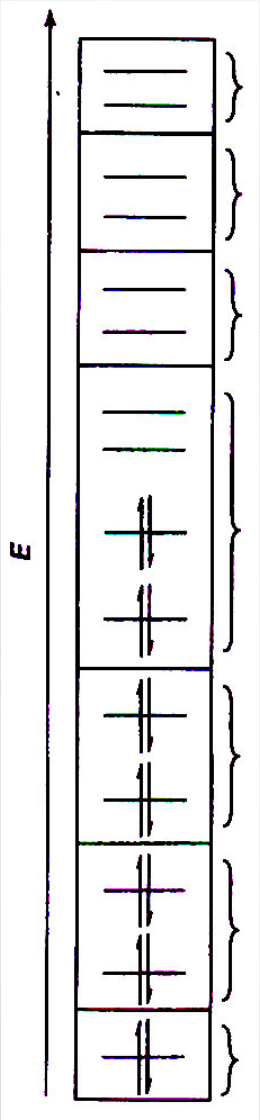
В этом методе обязательно разложение CI обрезают – иначе не решить.

ССП МЕТОД ПОЛНОГО АКТИВНОГО ПРОСТРАНСТВА

ССП метод полного активного пространства (complete active space SCF method - CASSCF) - был предложен Roos and co-workers в 1980.

В нем учитывается большое число конфигураций путем деления всех молекулярных орбиталей на три типа: **полностью занятые** во всех конфигурациях, **полностью не занятые** во всех конфигурациях и оставшиеся **«активные»** орбитали, которые могут быть заняты либо нет.

ССП МЕТОД ПОЛНОГО АКТИВНОГО ПРОСТРАНСТВА



Замороженные HF валентные орбитали

Принудительно не занятые

Принудительно занятые частично

Полное активное пространство

Все состояния могут быть заняты или пусты

Принудительно занятые частично

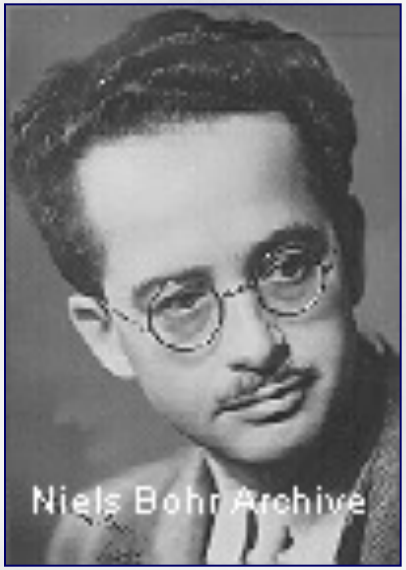
Ограниченное активное пространство

Принудительно дважды занятые

Замороженные HF занятые орбитали

post-HF МЕТОДЫ
ТЕОРИЯ ВОЗМУЩЕНИЙ

Использование **теории возмущений** для решения проблемы электронной корреляции в 1934 году было предложено C.Møller и M.S.Plesset. В квантовохимических программах она носит название **MP theory**.



Christiane Møller

Milton Spinoza Plesset

ТЕОРИЯ ВОЗМУЩЕНИЙ

MP метод основан на теории возмущений Rayleigh-Schrödinger, в которой «настоящий» Гамильтониан ***H*** является суммой Гамильтониана «нулевого порядка» $\boldsymbol{H_0}$ и возмущения ***V***

$$\boldsymbol{H = H_0 + \lambda V}$$

где $\boldsymbol{\lambda}$ **-** параметр, который может меняться от 0 до 1

Тогда собственные волновые функции $\mathbf{\Psi_i}$ и собственные энергии $\mathbf{E_i}$

$$\Psi_i = \Psi_i^{(0)} + \lambda\Psi_i^{(1)} + \lambda^2\Psi_i^{(2)} + \cdots = \sum_{n=0} \lambda^n \Psi_i^{(n)}$$

$$E_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \cdots = \sum_{n=0} \lambda^n E_i^{(n)}$$

ТЕОРИЯ ВОЗМУЩЕНИЙ

E_1 - поправка первого порядка

E_2 - поправка второго порядка и т.д.

В теории МР показано, что сумма энергии **«нулевого порядка»** и возмущающей поправки **первого порядка** соответствует энергии Хартри-Фока, поэтому для учета корреляции необходимо рассматривать поправку по крайней мере **второго порядка (E_2)**. Этот уровень теории обозначается в квантовохимических программах как **MP2**

$$E_0^{(2)} = \sum_{i}^{\text{occupied}} \sum_{j>1} \sum_{a}^{\text{virtual}} \sum_{b>a} \frac{\iint d\tau_1\, d\tau_2\, \chi_i(1)\chi_j(2)\left(\frac{1}{r_{12}}\right)[\chi_a(1)\chi_b(2) - \chi_b(1)\chi_a(2)]}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}$$

$$\mathbf{E = E_{HF} + E_{MP2}}$$

post-HF МЕТОДЫ
ТЕОРИЯ СВЯЗАННЫХ КЛАСТЕРОВ

Теория связанных кластеров **(Coupled-cluster CC)** - была предложена для оценки электронной корреляции Jiři Čížek и Josef Paldus в 1966.

Josef Paldus

Волновая функция системы с учетом корреляции может быть записана в виде

$$\psi = e^{\hat{T}}\Phi_0$$

где **Ψ** - нерелятивистская электронная волновая функция, **Φ_0** - нормализованная волновая функция Хартри-Фока, $e^{\hat{T}}$ - оператор и $\hat{T}$ - **кластерный оператор** или оператор возбуждения электронов

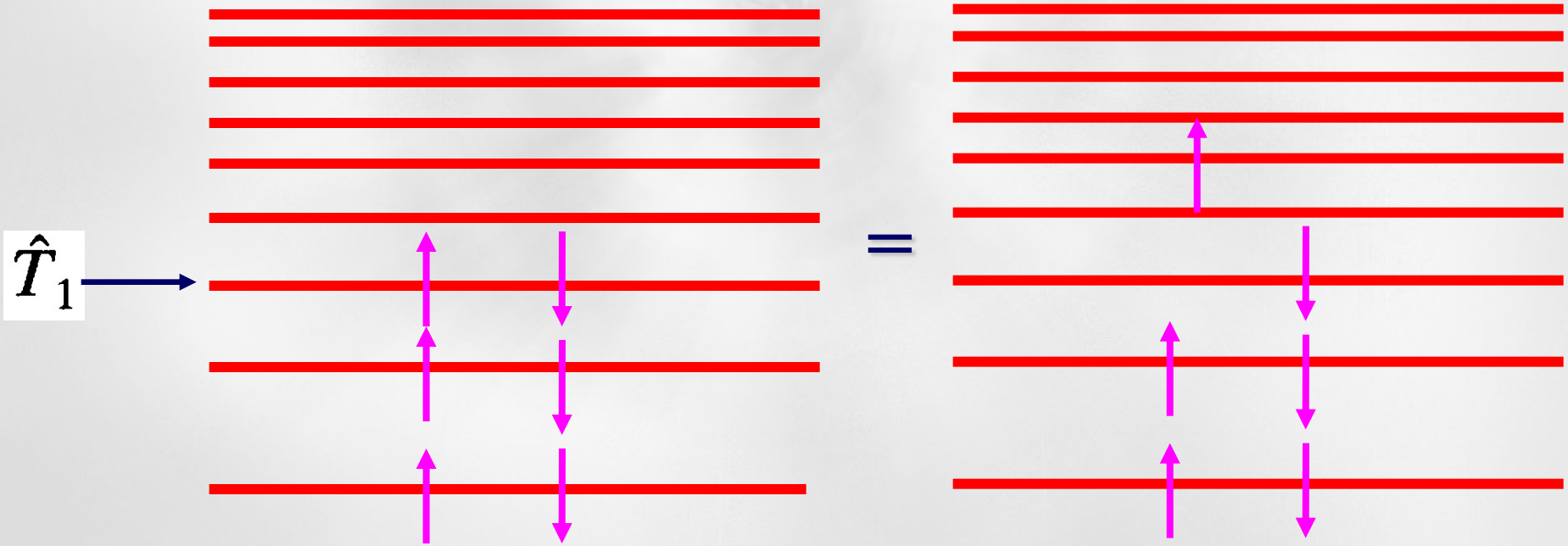
ТЕОРИЯ СВЯЗАННЫХ КЛАСТЕРОВ

Кластерный оператор $\hat{T}$ определяется как

$$\hat{T} \equiv \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_n$$

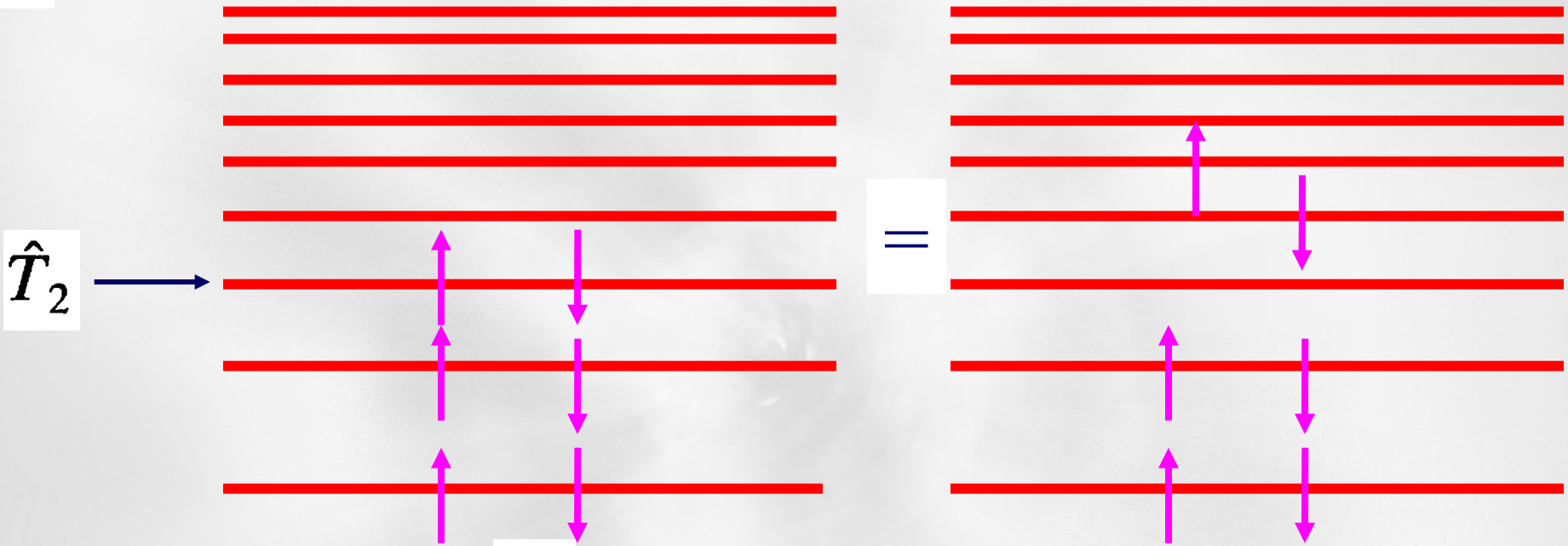
n - число электронов в молекуле

$\hat{T}_1$ - одночастичный оператор возбуждения **(CCS метод)**



ТЕОРИЯ СВЯЗАННЫХ КЛАСТЕРОВ

$\hat{T}_2$ - двухчастичный оператор возуждения **(CCD метод)**



Действие оператора $e^{\hat{T}}$ заключается в определении **ψ** как линейной комбинации Слэтеровских детерминантов, которые включают **все возможные возбуждения** электронов с заполненных на свободные спин-орбитали. Это соответствует **полному конфигурационному взаимодействию (Full CI)**

ТЕОРИЯ СВЯЗАННЫХ КЛАСТЕРОВ

В реальных квантовохимических пакетах используют следующие приближения связанных кластеров:

CCSD метод (CC singles and doubles)

$$\hat{T} = \hat{T}_1 + \hat{T}_2$$

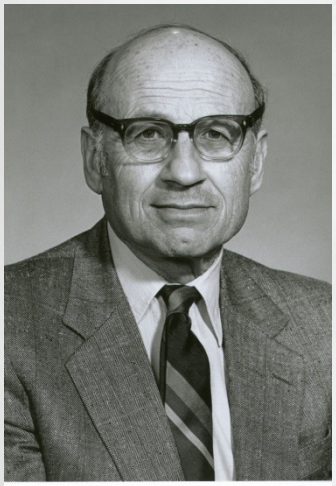
CCSDT метод (CC singles, doubles and triples)

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$$

Однако компьютерные затраты масштабируются как N^6 для метода **CCSD** и N^8 для метода **CCSDT**

МЕТОДЫ ФУНКЦИОНАЛА ПЛОТНОСТИ

В 1964 году P.Hohenberg и W.Kohn доказали, что для молекул в основном состоянии энергия, волновая функция и молекулярные свойства **однозначно** определяются **распределением электронной плотности** в молекуле.



Walter Kohn

Pierre Hohenberg

Иными словами энергия $\mathbf{E_0}$ - это **<u>функционал</u>** электронной плотности $\boldsymbol{\rho_0}$

$$\boxed{E_0 = E_0[\rho_0]}$$

Inhomogeneous Electron Gas, Phys. Rev. 136, B864, Published 09.11.1964.
DOI: http://dx.doi.org/10.1103/PhysRev.136.B864

МЕТОДЫ ФУНКЦИОНАЛА ПЛОТНОСТИ

Теория функционала плотности (Density Functional Theory - DFT) вычисляет энергию и молекулярные свойства системы в основном состоянии по электронной плотности в основном состоянии. Данная теория заменяет вычисление **многоэлектронной волновой функции** в методе Хартри-Фока и последующих приближениях на расчет всего лишь **одной характеристики** системы, зависящей от трех переменных. В 1965 W.Kohn и L.J.Sham предложили метод для нахождения электронной плотности без первоначального определения волновой функции.

МЕТОД Kohn-Sham

Данный метод сводится к решению задачи о **не взаимодействующих электронах**, движущихся в **эффективном потенциале**. Эффективный потенциал состоит из внешнего потенциала, определяемого положениями атомных ядер в системе, и функционала **F** – суммы кинетической энергии электронов и вклада от межэлектронного взаимодействия.

$$E[\rho(\mathbf{r})] = \int V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})\,d\mathbf{r} + F[\rho(\mathbf{r})]$$

ПРИБЛИЖЕНИЕ ЛОКАЛЬНОЙ ПЛОТНОСТИ

Для определения формы функционала **F** используется **приближение локальной плотности (local density approximation - LDA)**. В данном приближении предполагается, что плотность невзаимодействующего гомогенного электронного газа соответствует плотности электронов в рассматриваемой системе и таким образом, выражение для определения обменно-корреляционной энергии газа справедливо и для рассматриваемой системы. **Локальность** подразумевает, что обменно-корреляционная энергия электрона в любой точке пространства - это функция электронной плотности **только в этой точке**.

ФОРМА ФУНКЦИОНАЛА В LDA

Обменно-корреляционная энергия имеет следующий вид:

$$E_{xc} = E_x + E_c$$

В приближении LDA – обменная часть:

$$E_x^{\mathrm{LDA}} \equiv \int \rho \varepsilon_x \, d\mathbf{r} = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \int [\rho(\mathbf{r})]^{4/3} \, d\mathbf{r}$$

В то же время для описания корреляционной части существует несколько приближений:

- Vosko-Wilk-Nusair (VWN)
- Perdew-Wang (PW91)

МЕТОДЫ ГРАДИЕНТНОЙ КОРРЕКЦИИ

Дальнейшее улучшение приближения LDA – это моделирование негомогенного электронного газа. В этом случае обменная и корреляционная энергии зависят не только от электронной плотности, но и от производных этой плотности. Эти методы называются **градиентной коррекцией или обобщенным градиентным приближением (Gradient correction or Generalized gradient approximation – GGA)**

Примерами таких функционалов являются:

- **PW92**
- **LYP**
- **B95**

ГИБРИДНЫЕ МЕТОДЫ

Гибридные функционалы включают в обменную часть точный расчет по методу Хартри-Фока, взятый с определенным коэффициентом, а оставшиеся обменные и корреляционные члены рассчитываются по выбранному функционалу.

Так наиболее популярный функционал **B3LYP** имеет форму:

$$E_{xc}^{\mathrm{B3LYP}} = (1 - a_0 - a_x)E_x^{\mathrm{LSDA}} + a_0 E_x^{\mathrm{exact}} + a_x E_x^{\mathrm{B88}} + (1 - a_c)E_c^{\mathrm{VWN}} + a_c E_c^{\mathrm{LYP}}$$

Другие гибридные функционалы

- **B3PW91**
- **B1B96**

ПРОГРАММЫ, ВКЛЮЧАЮЩИЕ DFT

- **SIESTA** - www.uam.es/departamentos/ciencias/fismateriac/siesta (Universidad Autonoma de Madrid) - свободная
- **VASP – Разработана в Университете Вены - платная**
- **deMon2k** - http://demon-software.com/public_html/program.html (University of Montreal)
- **ADF** - www.scm.com/Welcome.html (Scientific computing & modelling)

Кроме того, приближения функционала плотности включены в пакеты

GAUSSIAN-09

GAMESS

MOLPRO

Dalton

Постулаты полуэмпирических методов

- Они базируются на теории молекулярных орбиталей и используют **минимальные базисные наборы** для описания **валентных** электронов
- Традиционно используется три уровня приближения для описания интегралов:

  CNDO – complete neglect of differential overlap (1965)

  INDO – intermediate neglect of differential overlap (1967)

  NDDO – neglect of diatomic differential overlap (1965)
- Эти методы **параметризуются** таким образом, чтобы **воспроизводить экспериментальные данные**.

МЕТОД ПРЕНЕБРЕЖЕНИЯ ДВУХЦЕНТРОВЫМ ДИФФЕРЕНЦИАЛЬНЫМ ПЕРЕКРЫВАНИЕМ - NDDO

Впервые это приближение было введено: **Pople and co-workers 1960-х**

В этом приближении матрица перекрываний **S** заменяется единичной матрицей, что позволяет заменить вековое уравнение Хартри-Фока

$$|\mathbf{H} - \mathbf{ES}| = 0 \quad \text{на}$$

$$|\mathbf{H} - \mathbf{E}| = 0$$

т.е. нулю приравниваются интегралы, описывающие перекрывание атомных орбиталей, центрированные на разных атомах

Это приближение является основой для современных методов **MNDO, AM1, PM3**

СОВРЕМЕННЫЕ ПОЛУЭМПИРИЧЕСИКИЕ МЕТОДЫ

MNDO - опубликован Dewar and co-workers в 1977. Данная модель базируется на приближении NDDO, но включает **только одноцентровые и двухцентровые члены.**

AM1 (Austin Model 1) – предложена Dewar and co-workers в 1985. В данной модели член, описывающий **отталкивание внутренних оболочек**, модифицировался с помощью гауссовых функций.

PM3 (MNDO parametric method 3) – предложена Stewart 1989. Форма Гамильтониана похожа на AM1, но параметры для этой модели определялись с использованием специальной процедуры, разработанной J.J.P. Stewart .

СРАВНЕНИЕ МЕТОДОВ

Усредненные абсолютные ошибки
для молекул (C, H, N, O)

Property[a]	N[b]	MNDO	AM1	PM3
ΔH_f (kcal/mol)	133	6.3	5.5	4.2
R(Å)	228	0.015	0.017	0.011
θ (deg)	92	2.69	2.01	2.22
IP (eV)	51	0.47	0.36	0.43
μ (D)	57	0.32	0.25	0.27

N^b – число сравнений
ΔH_f – теплота образования
R – длина связи
Θ – валентный угол
IP – потенциал ионизации
μ - дипольный момент

ПОЛУЭМПИРИЧЕСКИЕ ПРОГРАММЫ

Существует ряд специальных полуэмпирических программ:

- **MOPAC2007** - http://openmopac.net (Stewart Computational Chemistry)
- **AMPAC** - http://www.semichem.com (Semichem Inc.)
- **SIBIQ** – http://www.chem.ac.ru/Chemistry/Soft/SIBIQ.en.html (Serge Pachkovsky)

Также возможность проводить полуэмпирические расчеты содержатся в таких распространенных пакетах, как: **Gaussian, GAMESS, HyperChem**

ПРОГРАММА MOPAC2007

<u>Пример задания входного файла</u>

```
AM1 1SCF
Formic acid
Example of a single-point calculation
O
C   1.20   1
O   1.32   1  116.8  1   0.0     0    2 1
H   0.98   1  123.9  1   0.0     0    3 2 1
H   1.11   1  127.3  1   180.0 0    2 1 3
0   0.00    0 0.0     0    0.0    0    0 0 0
```

M.Dewar

J.Stewart

ПРОГРАММА MOPAC2007

```
*****************************************************************************
 ** Site#:  1321          For non-commercial use only        Version 7.060W **
 *****************************************************************************
 ** Cite this work as: MOPAC2007, James J. P. Stewart, Stewart Computational  **
 ** Chemistry, Version 7.060W web: HTTP://OpenMOPAC.net   Days remaining: 366 **
 *****************************************************************************
 **                                                  **
 **                    MOPAC2007                          **
 **                                                  **
 *****************************************************************************

                    AM1 CALCULATION RESULTS


GEOMETRY OPTIMISED USING EIGENVECTOR FOLLOWING (EF).
   SCF FIELD WAS ACHIEVED


                     AM1    CALCULATION
                                   MOPAC2007 (Version: 7.060W)
                                   Thu Mar  1 16:17:07 2007
                                   No. of days left = 366
FINAL HEAT OF FORMATION =       -97.41479 KCAL =    -407.58348 KJ


       TOTAL ENERGY            =       -797.12666 EV
       ELECTRONIC ENERGY       =     -1697.04783 EV  POINT GROUP:     Cs
       CORE-CORE REPULSION     =       899.92117 EV
       COSMO AREA              =        71.42 SQUARE ANGSTROMS
       COSMO VOLUME            =        52.65 CUBIC ANGSTROMS

       IONIZATION POTENTIAL    =        11.81903
       NO. OF FILLED LEVELS    =         9
       MOLECULAR WEIGHT        =        46.026

       MOLECULAR DIMENSIONS (Angstroms)

        Atom      Atom      Distance
        H    5    H    4    2.86513
        O    3    O    1    2.10425
        C    2    H    4    0.00000
```

MOPAC 2012-2016 – J. Stewart

- Optimize systems of up to 15,000 atoms, e.g. proteins
- PM6, PM7 parameterization from experimental & *ab initio*
- *PM6 – базовый метод: поправлены ошибки и сделаны улучшения по сравнению с PM3*
- *PM6-D3H4X – метод PM6 с поправками на дисперсию, водородные и галогенные связи*
- *PM7 – сразу был параметризован с учетом поправок на дисперсию, водородные и галогенные связи*
- More accurate (than AM1, PM3 and some ab initio methods) heats of formation and geometries
- ALL main group elements & transition metals
- Serious errors from PM3 and AM1 corrected
- Crystals, surfaces & polymers with periodic boundaries
- Сейчас MOPAC – свободно распросраняется

MOPAC 2012

Year	Method	# compounds used for parameterization
1977	MNDO	39 experimental
1985	AM1	~200 experimental
1989	PM3	~500 experimental
2007	**PM6**	**> 9,000 experimental & *ab initio* + *corr: dispersion, hydrogen bonds, halogen bonds***
2013	***PM7***	**> 9,000 experimental & *ab initio: + SCF corr: dispersion, hydrogen bonds, halogen bonds***

•More accurate heats of formation from PM6 and PM7

Method	Average unsigned error	Largest error (kcal/mol)
PM6	**4.79**	**-42.2**
B3LYP 6-31G(d)	5.19	35.8
PM3	6.26	135.6
HF 6-31G(d)	7.37	72.5
AM1	10.01	111.9

Comparison of errors in heats of formation for a set of 1,373 compounds containing only C, H, O, N, F, Cl, S, P and Br.

Intermolecular: PM7 vs PM6

Average unsigned error, ккал/моль			
	PM7	**PM6-D3H4X**	**Количество комплексов**
Set S22	**0.74**	**3.38**	**22**
Set L7	**4.67**	**13.73**	**7**
Set X40	**1.70**	**1.67**	**40**
Set S66	**0.78**	**2.68**	**66**
Set S12L	**21.37**	**13.37**	**12**

Преимущество PM7/PM6 по сравнению с более ранними методами, например с PM3:

1. Гораздо больший объем данных для параметризации
2. Введены поправки на описание межмолекулярных взаимодействий
3. Введены поправки на описание водородных и галогенных связей

BSSE – basis set superposition error

Энергию вазимодействия между двумя молекулами или атомами можно выразить как разницу между энергией комплекса и отдельных компонентов:

$$\Delta E = E(AB) - E(A) - E(B)$$

Для систем, связанных **дисперсионными взаимодействиями** или **водородными связями**, энергия, рассчитанная таким образом сильно искажена

Это следствие не полноты одноэлектронного базисного набора.

BSSE

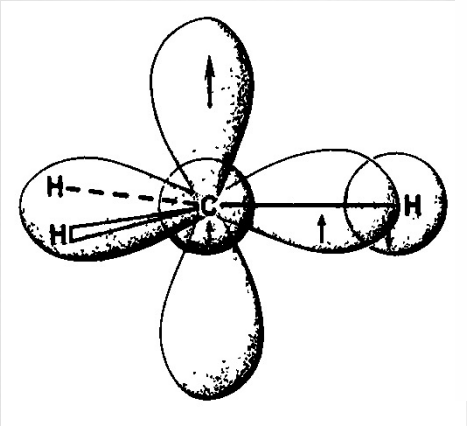
Пример: расчет димера He_2

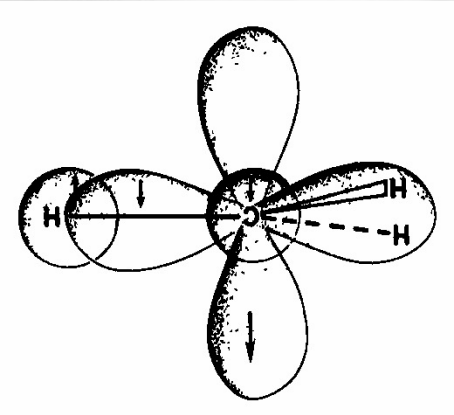
Метод	Базисные функции	Длина связи (пм)	ΔE, кДж/моль
RHF/6-31G	2	323.0	-0.0035
RHF/cc-pVDZ	5	321.1	-0.0038
RHF/cc-pVTZ	14	366.2	-0.0023
эксперимент		**297**	**-0.091**
MP2/6-31G	2	321.0	-0.0042
MP2/cc-pVDZ	5	309.4	-0.0159
MP2/cc-pVTZ	14	331.8	-0.0211

УРАВНОВЕШИВАЮЩАЯ ПОПРАВКА ДЛЯ BSSE

Для исправления BSSE применяется метод **уравновешивающей поправки**:

рассчитывается **энергия фрагментов** в **полном** базисном наборе всего комплекса и вычитается из энергии всей системы





Спасибо за внимание

- ***…Surely every medicine is an innovation; and he that will not apply new remedies, must expect new evils…***
- ***…Каждый медицинский метод есть инновация; а кто не хочет применять новые средства, должен ждать новых бед…***

Sir Francis Bacon (1561-1626)

OF INNOVATIONS