

# Базисные наборы

Иван Федянин

ВХК РАН

Например:  $(0, 1), (1, 0)$  — базисный набор 2D декартового пространства

В вычислительной химии «базис», «базисный набор» — линейно-независимые функции, с помощью которых представляется волновая функция системы

- Атом-центрированные базисные функции
  - Универсальные (подходят для изолированных и периодических систем)
  - Теоретически, требуется меньше функций/расчётов на электрон
  - Подходят только для борн-оппенгеймеровской МД
  - Всегда приводят к ошибке суперпозиции базисного набора (Basis Set Superposition Error, BSSE)
- Плоские волны
  - Требуют трёхмерной периодичности системы
  - Теоретически, намного медленнее, из-за необходимости очень большого числа функций для корректного описания поведения волновой функции вблизи ядра
  - На практике используются с псевдопотенциалами и присоединёнными плоскими волнами, что приводит к большей скорости для кристаллов
  - Подходят как для МД в приближении Борна-Оппенгеймера, так и Кара-Паренелло (более быстрой!)
  - Проблема BSSE обычно не возникает

$$\varphi_i(\mathbf{r}) = \sum_{\mu}^K c_{\mu i} \tilde{\varphi}_{\mu}$$

$\{\tilde{\varphi}_\mu(\mathbf{r})\}$  — атомные орбитали: базисный набор

Все молекулярные орбитали  $\{\varphi_i(\mathbf{r})\}$  строятся из одних и тех же базисных функций (но с разными коэффициентами)!

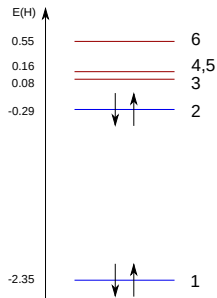
Минимальный базисный набор — включает все занятые орбитали каждого из атомов

Для закрытых оболочек общее число (пространственных) орбиталей  $L$  равно числу базисных функций  $K$ .

При  $N$  электронах занятыми являются  $N/2$  орбиталей, а незаполненными  $L^* = L - N/2$ .

LiH (d=1.595Å; HF / STO-3G)

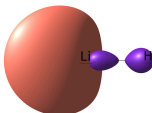
	1	2	3	4	5	6
1s(Li)	0.99	-0.17	-0.21	0.00	0.00	0.09
2s(Li)	0.03	0.45	0.80	0.00	0.00	-0.70
2p <sub>x</sub> (Li)	0.00	0.00	0.00	1.00	0.00	0.00
2p <sub>y</sub> (Li)	0.00	0.00	0.00	0.00	1.00	0.00
2p <sub>z</sub> (Li)	0.01	-0.35	0.61	0.00	0.00	0.98
1s(H)	0.00	0.55	-0.14	0.00	0.00	1.19



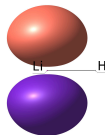
1



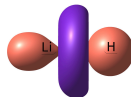
2



3



4 (5)

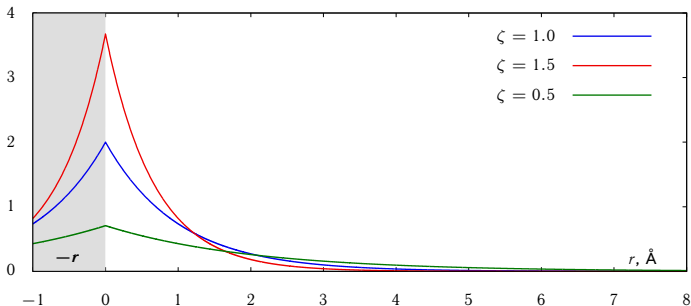


6

### Орбитали Слейтеровского типа ("Slater-Type Orbitals", STO)

$$\begin{aligned}\phi_{abc}^{\text{STO}} &= N_c x^a y^b z^c e^{-\zeta r} \\ \phi_{n\ell m}^{\text{STO}} &= N Y_{\ell}^m(\theta, \phi) r^{n-1} e^{-\zeta r}\end{aligned}\quad (1)$$

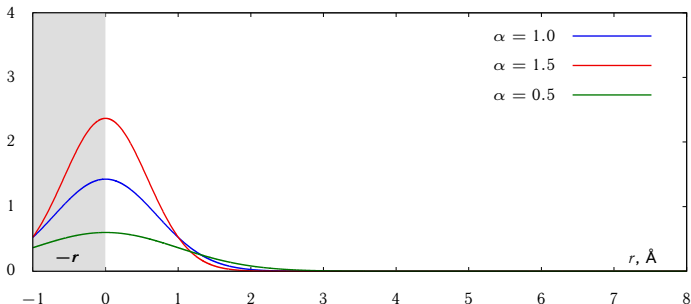
- $N_c, N$  — константы для нормировки,  $N = (2\zeta)^n \sqrt{2\zeta/(2n)!}$
- $a + b + c = L$ ,  $0 \leq a, b, c \leq L$  — определяют момент импульса (угловой момент)
- $\zeta$  — определяет скорость спадания функции



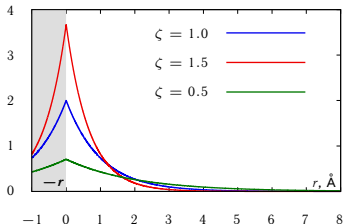
### Орбитали гауссова типа (“Gaussian-Type Orbitals”, GTO)

$$\begin{aligned}\phi_{abc}^{\text{GTO}} &= N_c x^a y^b z^c e^{-\alpha r^2} \\ \phi_{n\ell m}^{\text{GTO}} &= N Y_\ell^m(\theta, \phi) r^{2n-2-\ell} e^{-\alpha r^2}\end{aligned}\quad (2)$$

- $N_c, N$  — константы для нормировки
- $a + b + c = L, 0 \leq a, b, c \leq L$  — определяют момент импульса (угловой момент)
- $\alpha$  — определяет скорость спадания функции

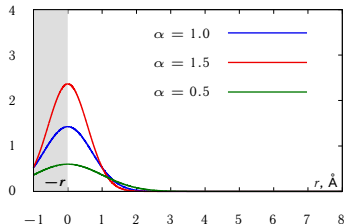


## STO vs. GTO



STO

- Правильное поведение вблизи 0
- Совпадает с  $1s$  для атома H
- Считать более затратно

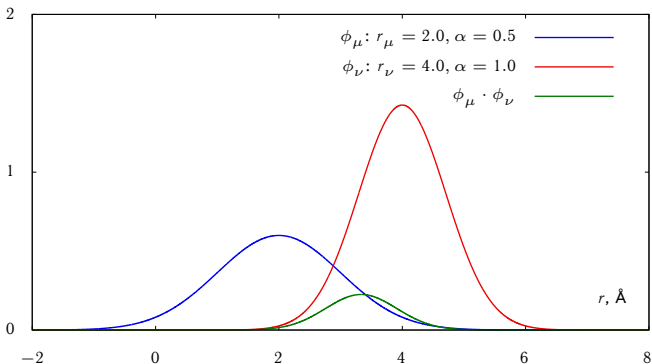


GTO

- Неправильное поведение вблизи 0
- Спадает быстрее, чем STO
- Считать менее затратно

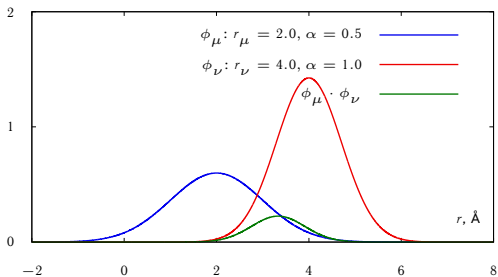


Почему расчёты с GTO менее затратные?



Произведение двух GTO — также GTO с центром между ними

$$\int [\varphi_\mu(\mathbf{r}_1) \varphi_\nu(\mathbf{r}_1) \frac{1}{r_{12}} \varphi_\lambda(\mathbf{r}_2) \varphi_\sigma(\mathbf{r}_2)] d\mathbf{r}_1 d\mathbf{r}_2$$



$$\varphi_{\mu}(r) = e^{-\alpha(r-A)^2} \quad \varphi_{\nu}(r) = e^{-\beta(r-B)^2}$$

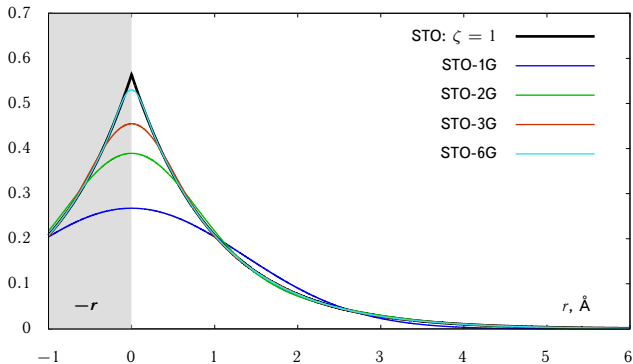
$$\varphi_{\mu}(r)\varphi_{\nu}(r) = e^{-\frac{\alpha\beta}{\alpha+\beta}(A-B)^2} e^{-(\alpha+\beta)(r-\frac{\alpha A+\beta B}{\alpha+\beta})^2}$$

Можно приблизить STO с помощью нескольких GTO!

$$\phi^{\text{STO}}(\zeta, r) = \sum_{i=1}^P d_i \phi^{\text{GTO}}(\alpha_i, r)$$

Семейство базисных наборов STO-nG

(W.J. Hehre, R.F. Stewart, J.A. Pople, J. Chem. Phys. 1969, 51(6), 2657–2664, 10.1063/1.1672392)



В общем случае, базисные орбитали строятся из GTO

$$\phi(r) = \sum_{i=1}^P d_i \phi^{\text{GTO}}(\alpha_i, r)$$

$\{d_i\}$  — коэффициенты контракции. В процессе расчёта *не меняются!*

$$\tilde{\varphi}_{\mu}(\mathbf{r}) = \sum_{k=1}^{P_{\mu}} d_{k\mu} \phi_{\mu}^{\text{GTO}}(\alpha_{k\mu}, r) \Rightarrow \varphi_i(r) = \sum_{\mu=1}^K c_{\mu i} \tilde{\varphi}_{\mu}(\mathbf{r}) \Rightarrow \{\chi_i(\mathbf{x})\} \Rightarrow$$

$$\Psi = |\chi_1 \dots \chi_N\rangle$$

$$\phi(r) = \sum_{i=1}^P d_i \phi^{\text{GTO}}(\zeta_i, r)$$

n GTO  $\rightarrow$  STO-nG  $\Rightarrow$  контракция; коэффициенты  $d_i$  фиксированы.

STO-3G для H (в формате программы Gaussian, с сайта [basissetexchange.org](http://basissetexchange.org))

```

H      0
S      3      1.00
        0.3425250914D+01      0.1543289673D+00
        0.6239137298D+00      0.5353281423D+00
        0.1688554040D+00      0.4446345422D+00
****

```

STO-3G для Na [ $1s^2 2s^2 2p^6 3s^1 3p^0$ ]

```

Na      0
S      3      1.00
        0.2507724300D+03      0.1543289673D+00
        0.4567851117D+02      0.5353281423D+00
        0.1236238776D+02      0.4446345422D+00
SP      3      1.00
        0.1204019274D+02      -0.9996722919D-01      0.1559162750D+00
        0.2797881859D+01      0.3995128261D+00      0.6076837186D+00
        0.9099580170D+00      0.7001154689D+00      0.3919573931D+00
SP      3      1.00
        0.1478740622D+01      -0.2196203690D+00      0.1058760429D-01
        0.4125648801D+00      0.2255954336D+00      0.5951670053D+00
        0.1614750979D+00      0.9003984260D+00      0.4620010120D+00
****

```

STO-nG  $\rightarrow$  STO-(n-1)G+GTO: деконтракция

«расфиксируем» один из коэффициентов  $d_i$  для большей гибкости базисного набора!

**3-21G** для H

```

H      0
S      2      1.00
        0.5447178000D+01      0.1562849787D+00
        0.8245472400D+00      0.9046908767D+00
S      1      1.00
        0.1831915800D+00      1.00000000
****

```

**3-21G** для Na [ $1s^2 2s^2 2p^6 3s^1 3p^0$ ]

```

Na      0
S      3      1.00
        0.5476130000D+03      0.6749111801D-01
        0.8206780000D+02      0.3935051050D+00
        0.1769170000D+02      0.6656051776D+00
SP     3      1.00
        0.1754070000D+02      -0.1119370290D+00      0.1282330093D+00
        0.3793980000D+01      0.2546540659D+00      0.4715330341D+00
        0.9064410000D+00      0.8444172187D+00      0.6042730437D+00
SP     2      1.00
        0.5018240000D+00      -0.2196604975D+00      0.9066487958D-02
        0.6094580000D-01      0.1089122467D+01      0.9972017754D+00
SP     1      1.00
        0.2443490000D-01      0.1000000000D+01      0.1000000000D+01
****

```

Эффект деконтракции для молекулы  $H_2$ 

## STO-3G

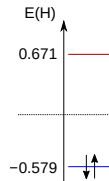
H	0		
S	3	1.00	
		0.3425250914D+01	0.1543289673D+00
		0.6239137298D+00	0.5353281423D+00
		0.1688554040D+00	0.4446345422D+00

## Gaussian output:

```

2 basis functions,      6 primitive gaussians,      2 cartesian basis functions
1 alpha electrons      1 beta electrons
.....
SCF Done:  E(RHF) =  -1.11675930734

```



## 3-21G

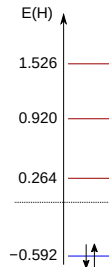
H	0		
S	2	1.00	
		0.5447178000D+01	0.1562849787D+00
		0.8245472400D+00	0.9046908767D+00
S	1	1.00	
		0.1831915800D+00	1.00000000

## Gaussian output:

```

4 basis functions,      6 primitive gaussians,      4 cartesian basis functions
1 alpha electrons      1 beta electrons
.....
SCF Done:  E(RHF) =  -1.12294025881

```



**3-21G, 6-31G** — двухрасщеплённые

**6-311G** — трёхрасщеплённый (для 1 и 2 периода(!)) и т.д.

Функции для s и p орбиталей в пределах периода имеют одни и те же  $\zeta$ .

6-311G для C [ $1s^2 2s^1 2p^3$ ]

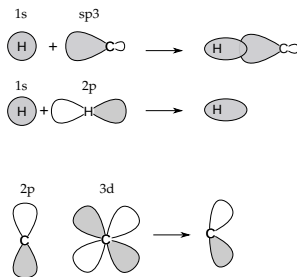
C	0		
S	6	1.00	
	4563.240	0.00196665	
	682.0240	0.0152306	
	154.9730	0.0761269	
	44.45530	0.2608010	
	13.02900	0.6164620	
	1.827730	0.2210060	
SP	3	1.00	
	20.96420	0.114660	0.0402487
	4.803310	0.919999	0.237594
	1.459330	-0.00303068	0.815854
SP	1	1.00	
	0.4834560	1.000000	1.000000
SP	1	1.00	
	0.1455850	1.000000	1.000000



6-311G для 3 (и 4) периодов (уже не поплровский базис): а хотелось, чтобы было просто. . .

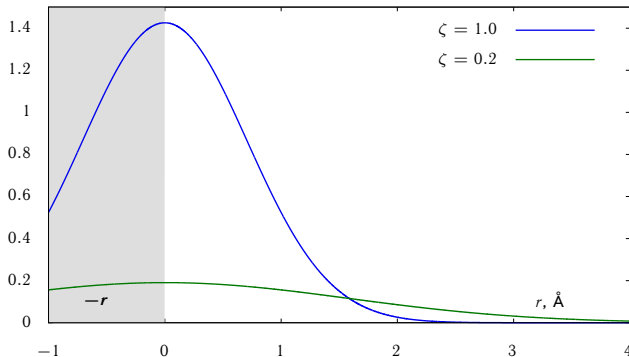
Na	0	
S	6	1.00
	36166.4	0.001032
	5372.58	0.008071
	1213.21	0.042129
	339.623	0.169789
	109.553	0.514621
	38.7773	0.379817
S	3	1.00
	38.7773	0.374762
	14.5759	0.575769
	5.26993	0.112933
S	1	1.00
	1.82777	1.000000
S	1	1.00
	0.619948	1.000000
S	1	1.00
	0.057240	1.000000
S	1	1.00
	0.024048	1.000000
P	4	1.00
	144.645	0.011485
	33.9074	0.082383
	10.6285	0.319658
	3.82389	0.701295
P	2	1.00
	1.44429	0.638506
	0.552621	0.425365
P	1	1.00
	0.188720	1.000000
P	1	1.00
	0.046501	1.000000
P	1	1.00
	0.016285	1.000000

При образовании связей электронная плотность исходно сферически симметричных атомов изменяется: «поляризуется». Орбитали с исходной симметрией оказываются неоптимальными для описания нового состояния.



Можно или взять сильно больше функций, или использовать *поляризационные функции*: функции АО более высокого порядка, чем заполненные оболочки (H, He:  $\sim 2p$ ; 2-й период:  $\sim 3d$ ; 3-й период:  $\sim 4f$  и т. д.)

В анионах и при образовании невалентных межмолекулярных взаимодействий важны «хвосты» волновой функции на большом расстоянии от ядер.



Можно или взять сильно больше функций, или добавить *диффузные функции*: функции с маленькими  $\zeta$ .

6-31+G\*\*

6 Число контрактированных GTO для внутренних оболочек

31 Число контрактированных GTO для валентных оболочек

+ Дополнительная диффузная функция с малым  $\zeta$  для атомов 2-го периода

\*\* d-функции для атомов 2-го периода + p-функции для H

H	0	6-31G		
S	3	1.00		
		0.1873113696D+02	0.3349460434D-01	
		0.2825394365D+01	0.2347269535D+00	
		0.6401216923D+00	0.8137573261D+00	
S	1	1.00		
		0.1612777588D+00	1.00000000	

H	0	6-31+G**		
S	3	1.00		
		0.1873113696D+02	0.3349460434D-01	
		0.2825394365D+01	0.2347269535D+00	
		0.6401216923D+00	0.8137573261D+00	
S	1	1.00		
		0.1612777588D+00	1.00000000	
P	1	1.00		
		0.1100000000D+01	1.00000000	

C	0			
S	6	1.00		
		0.3047524880D+04	0.1834737132D-02	
		0.4573695180D+03	0.1403732281D-01	
		0.1039486850D+03	0.6884262226D-01	
		0.2921015530D+02	0.2321844432D+00	
		0.9286662960D+01	0.4679413484D+00	
		0.3163926960D+01	0.3623119853D+00	
SP	3	1.00		
		0.7868272350D+01	-0.1193324198D+00	0.68999
		0.1881288540D+01	-0.1608541517D+00	0.31642
		0.5442492580D+00	0.1143456438D+01	0.74430
SP	1	1.00		
		0.1687144782D+00	0.1000000000D+01	0.10000

C	0			
S	6	1.00		
		0.3047524880D+04	0.1834737132D-02	
		0.4573695180D+03	0.1403732281D-01	
		0.1039486850D+03	0.6884262226D-01	
		0.2921015530D+02	0.2321844432D+00	
		0.9286662960D+01	0.4679413484D+00	
		0.3163926960D+01	0.3623119853D+00	
SP	3	1.00		
		0.7868272350D+01	-0.1193324198D+00	0.6899906659D-01
		0.1881288540D+01	-0.1608541517D+00	0.3164239610D+00
		0.5442492580D+00	0.1143456438D+01	0.7443082909D+00
SP	1	1.00		
		0.1687144782D+00	0.1000000000D+01	0.1000000000D+01
SP	1	1.00		
		0.4380000000D-01	0.1000000000D+01	0.1000000000D+01
D	1	1.00		
		0.8000000000D+00	1.00000000	

$$6-31G^* \equiv 6-31G(d)$$

$$6-31+G^{**} \equiv 6-31+G(d,p)$$

$$6-311++G(2df, 2pd):$$

на атом H добавлены две p и одна d функция

на атомы 2-го периода две d и одна f функция

$$6-311++G(3df, 3pd) \text{ — максимальный из размещённых на } \text{basisetexchange.org}$$

Добавление поляризационных и диффузных функций не изменяет основу валентно-расщеплённого базисного набора

Можно ли расщепить все функции для всех АО, а не только для валентных? Да, но не очень практично.

Базисы Даннинга и Хэя:

- DZ — Double Zeta
- TZ — Tripe Zeta

C	0	
S	6	1.00
	4232.61	0.002029
	634.882	0.015535
	146.097	0.075411
	42.4974	0.257121
	14.1892	0.596555
	1.96660	0.242517
S	1	1.00
	5.14770	1.000000
S	1	1.00
	0.4962	1.000000
S	1	1.00
	0.1533	1.000000
P	4	1.00
	18.1557	0.018534
	3.98640	0.115442
	1.14290	0.386206
	0.35940	0.640089
P	1	1.00
	0.11460	1.000000

- cc-pVDZ: correlation-consistent polarized valence double-zeta  
Для H  $\sim$  41G(p), для C  $\sim$  9-91/41(d), для Si  $\sim$  12-12 1/81(d)
- cc-pVTZ  
Для H  $\sim$  511G(2pd), для C  $\sim$  10-10 11/511(2df), для Si  $\sim$  15-15 11/911(2df)
- cc-pVQZ  
Для H  $\sim$  6111G(3p2df), для C  $\sim$  12-12 111/6111(3d2fg), для Si  $\sim$  16-16 11/11 111(3d2fg)
- cc-pV5Z
- cc-pV6Z

Поляризационные функции добавляются автоматически в зависимости от степени расщепления.

Версии с диффузными функциями:

aug-cc-pVDZ, aug-cc-pVTZ, aug-cc-pVQZ и т. д.

Для каждого  $n$  свой набор  $\{\zeta_i\}$ : “segemted contraction” (сегментированная? раздельная?)

Для каждого  $n$  одинаковый набор  $\{\zeta_i\}$ , но с разными коэффициентами: “general contraction”

## cc-pVDZ

C	0		
S	9	1.00	
		6.665000D+03	6.920000D-04
		1.000000D+03	5.329000D-03
		2.280000D+02	2.707700D-02
		6.471000D+01	1.017180D-01
		2.106000D+01	2.747400D-01
		7.495000D+00	4.485640D-01
		2.797000D+00	2.850740D-01
		5.215000D-01	1.520400D-02
		1.596000D-01	-3.191000D-03
S	9	1.00	
		6.665000D+03	-1.460000D-04
		1.000000D+03	-1.154000D-03
		2.280000D+02	-5.725000D-03
		6.471000D+01	-2.331200D-02
		2.106000D+01	-6.395500D-02
		7.495000D+00	-1.499810D-01
		2.797000D+00	-1.272620D-01
		5.215000D-01	5.445290D-01
		1.596000D-01	5.804960D-01
S	1	1.00	
		1.596000D-01	1.000000D+00
P	4	1.00	
		9.439000D+00	3.810900D-02
		2.002000D+00	2.094800D-01
		5.456000D-01	5.085570D-01
		1.517000D-01	4.688420D-01
P	1	1.00	
		1.517000D-01	1.000000D+00
D	1	1.00	
		5.500000D-01	1.000000

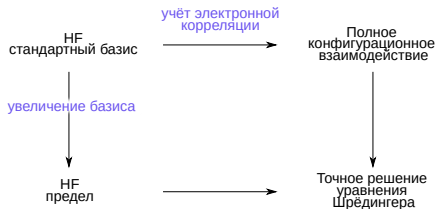


- **soutdef2-DZVP** :  
def2-SV(**P**) (H: 31, B-Ne: 511/31/1  $\approx$  "5-31G\*"),  
def2-SV**P** (H: 31/1, B-Ne 511/31/1  $\approx$  "5-31G\*\*")
- def2-TZV**P**, def2-TZV**PP**, def2-TZV**PD**, def2-TZV**PPD**
- def2-QZV**P**, def2-QZV**P**, def2-QZV**PD**

Q: Какой базисный набор выбрать?

A: Чем больше, тем лучше!

- Чем больше базис, тем ниже энергия
- Сложность HF  $\sim N^4$  функций, и это минимум...



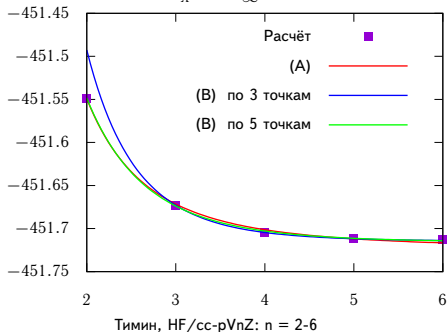
- Нужна сравнительно редко (для бенчмарков, термохимии)
- Несколько формул:

Требующая минимум две точки:

$$E_X^{\text{HF}} = E_\infty^{\text{HF}} + Ax^{-3} \quad (\text{A})$$

Требующая минимум три точки, или фиксированный коэффициент ( $\alpha$ ):

$$E_X^{\text{HF}} = E_\infty^{\text{HF}} + Ae^{-\alpha\sqrt{X}} \quad (\text{B})$$



- (A,2-6): -451.723 H
- (B,2-3): -451.715 H,  $\alpha = 5.250$
- (B,2-6): -451.716 H
- (B,3/4): -451.714 H, фикс.  $\alpha = 5.460$

- Кристалл отличается от изолированной молекулы отсутствием диффузных «хвостов» волновой функции.
- Диффузные функции (с малыми  $\zeta$ ) могут привести к линейной зависимости базисного набора.
- Могут быть использованы обычные базисы без диффузных функций, но они не оптимальны
- Специальные базисные наборы без функций с малыми  $\zeta$ : POB-DZVP, POB-TZVP
- Увеличение базисного набора также ограничено проблемой малых  $\zeta$
- Проблема «сжатия» кристаллических ячеек при оптимизации: BSSE?

- Для атомов 4+ периода число электронов возрастает
- Требуется большое число GTO для описания внутренних оболочек (9-25)
- Релятивистские эффекты! Корректный учёт требует особых методов (DKH, ZORA)

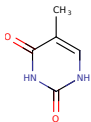
Внутренние электроны обычно мало интересны! Можно заменить их влияние на валентную (или пред-валентную) оболочки псевдопотенциалом, включая релятивистские эффекты.

[“ECP” - “effective core potentials”]

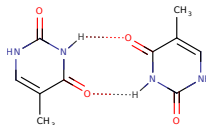
CO	0	
CO-ECP	2	10
d potential		
1		
1	19.4455900	-3.9474600
s-d potential		
3		
0	2.2593500	3.0770500
2	6.6666500	215.4838200
2	5.7681800	-174.1181600
p-d potential		
2		
0	50.2814600	3.9532100
2	12.5261000	90.9004300

Проблема всех расчётов с атом-центрированными базисными наборами —  
ошибка суперпозиции базисного набора [“Basis Set Superposition Error”, BSSE]

Тимин и димер: def2-TZVPP



363 базисные функции



726 базисных функций

Базисный набор для димера в два раза больше, чем для мономера!

- Более общая проблема: димер и мономер считаются в *разных базисах*
- Следствие: даже две конформации молекулы считаются в разных базисах (*внутримолекулярная BSSE*)
- Приводит к нефизично большей стабильности димеров, комплексов и т. д.
- Для плоского димера бензола в 6-31G(d) BSSE сравнима с величиной энергии связывания
- Чем больше базисный набор, тем меньше BSSE

## Методы учёта BSSE:

- Метод «противовеса» [“counterpoise”]

$$E_{\text{bond}} = E_{\text{dim}}^{\text{B:dim,G:dim}} - E_{\text{mon}}^{\text{B:dim,G:dim}}$$

$$\text{BSSE} = E_{\text{dim}}^{\text{B:dim,G:dim}} - E_{\text{mon}}^{\text{B:mon,G:dim}}$$

$$E_{\text{bond}} = E_{\text{dim}}^{\text{B:dim,G:dim}} - E_{\text{mon}}^{\text{B:mon,G:mon}} + \text{BSSE}$$

$E_{\text{bond}}$  — энергия связывания в димере

$E_{\text{form}}$  — энергия образования димера

B : dim, B : mon — базисы димера и мономера

G : dim, G : mon — геометрия димера и мономера

Применимы только к молекулярным ассоциатам / кристаллам.

- (Полу)эмпирические методы (“gCP” — “Geometric counterpoise”)

$$E_{\text{gCP}} = \sigma \sum_a \sum_{b \neq a} e_a^{\text{miss}} f_{\text{dec}}(R_{ab})$$

$a, b$  — атомы системы

$e_a^{\text{miss}}$  — величина «малости базисного набора»

$f_{\text{dec}}(r_{ab})$  — функция, спадающая с увеличением  $r_{ab}$

$\sigma$  — подбирается для сочетания базис/метод (напр., HF/def2-DVP)

Найдите любой базис на  $bse.pnl.gov \rightarrow basisetexchange.org$