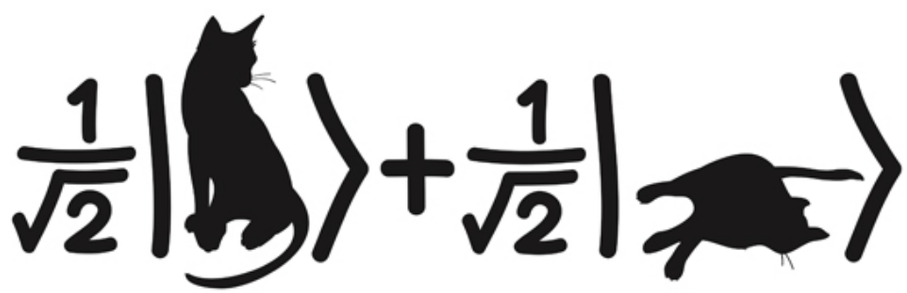
CHEM 7520:

Advanced Quantum Mechanics and Spectroscopy



Atom-Centered Basis Sets

THE OHIO STATE UNIVERSITY

Alexander Yu. Sokolov

Atom-Centered Basis Sets

We have introduced basis sets in the Hartree-Fock theory

Schrödinger equation

$$\hat{H}\,|\Psi(1,\ldots,N)\rangle = E\,|\Psi(1,\ldots,N)\rangle$$

Many-electron wavefunction

$$|\Psi(1,\ldots,N)\rangle \approx \mathcal{A}(1,\ldots,N)\prod_{i}^{N}|\psi_i(i)\rangle$$

Molecular orbitals

$$|\phi_i\rangle \;=\; \sum_{\mu}^{n} C_{\mu}^{i}\,|\chi_\mu\rangle$$

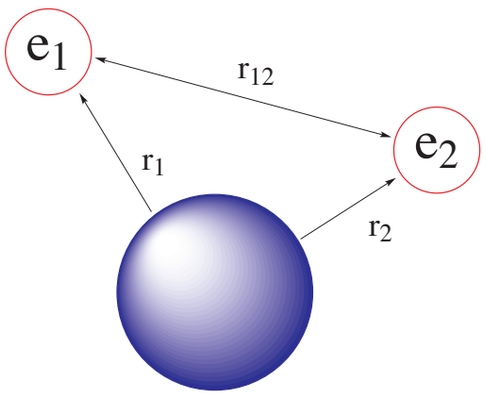
Atomic orbitals

$$|\chi_\mu^{\mathrm{GTO}}(r)\rangle = N_c\,x^k y^l z^m e^{-\alpha_\mu r^2}$$

Other electronic structure methods also use basis sets as "building blocks" to construct many-electron wavefunctions

Exact Wavefunction

Basis sets must be flexible enough to represent all essential features of the many-electron wavefunctions



Example: two-electron system

Hamiltonian of the 2e⁻ system:

$$\hat{H} = -\frac{1}{2}\frac{\partial^2}{\partial r_1^2} - \frac{1}{r_1}\frac{\partial}{\partial r_1} - \frac{Z}{r_1} - \frac{1}{2}\frac{\partial^2}{\partial r_2^2} - \frac{1}{r_2}\frac{\partial}{\partial r_2} - \frac{Z}{r_2}$$
$$- \frac{r_1^2 - r_2^2 + r_{12}^2}{2r_1 r_{12}}\frac{\partial^2}{\partial r_1 \partial r_{12}} - \frac{r_2^2 - r_1^2 + r_{12}^2}{2r_2 r_{12}}\frac{\partial^2}{\partial r_2 \partial r_{12}}$$
$$- \frac{\partial^2}{\partial r_{12}^2} - \frac{2}{r_{12}}\frac{\partial}{\partial r_{12}} + \frac{1}{r_{12}}$$

For a two-electron system, it is possible to compute wavefunction numerically, without using basis sets

Electron-Nuclear Cusp

Let's have a closer look at the Hamiltonian of the 2e⁻ system:

$$\hat{H} = -\frac{1}{2}\frac{\partial^2}{\partial r_1^2} \boxed{- \frac{1}{r_1}\frac{\partial}{\partial r_1} - \frac{Z}{r_1}} - \frac{1}{2}\frac{\partial^2}{\partial r_2^2} - \frac{1}{r_2}\frac{\partial}{\partial r_2} - \frac{Z}{r_2}$$
$$- \frac{r_1^2 - r_2^2 + r_{12}^2}{2r_1 r_{12}}\frac{\partial^2}{\partial r_1 \partial r_{12}} - \frac{r_2^2 - r_1^2 + r_{12}^2}{2r_2 r_{12}}\frac{\partial^2}{\partial r_2 \partial r_{12}}$$
$$- \frac{\partial^2}{\partial r_{12}^2} - \frac{2}{r_{12}}\frac{\partial}{\partial r_{12}} + \frac{1}{r_{12}}$$

What happens when $r_1 = 0$ (or $r_2 = 0$)?

The Schrödinger equation must hold over the entire space, what is the condition to avoid a singularity at $r_1 = 0$?

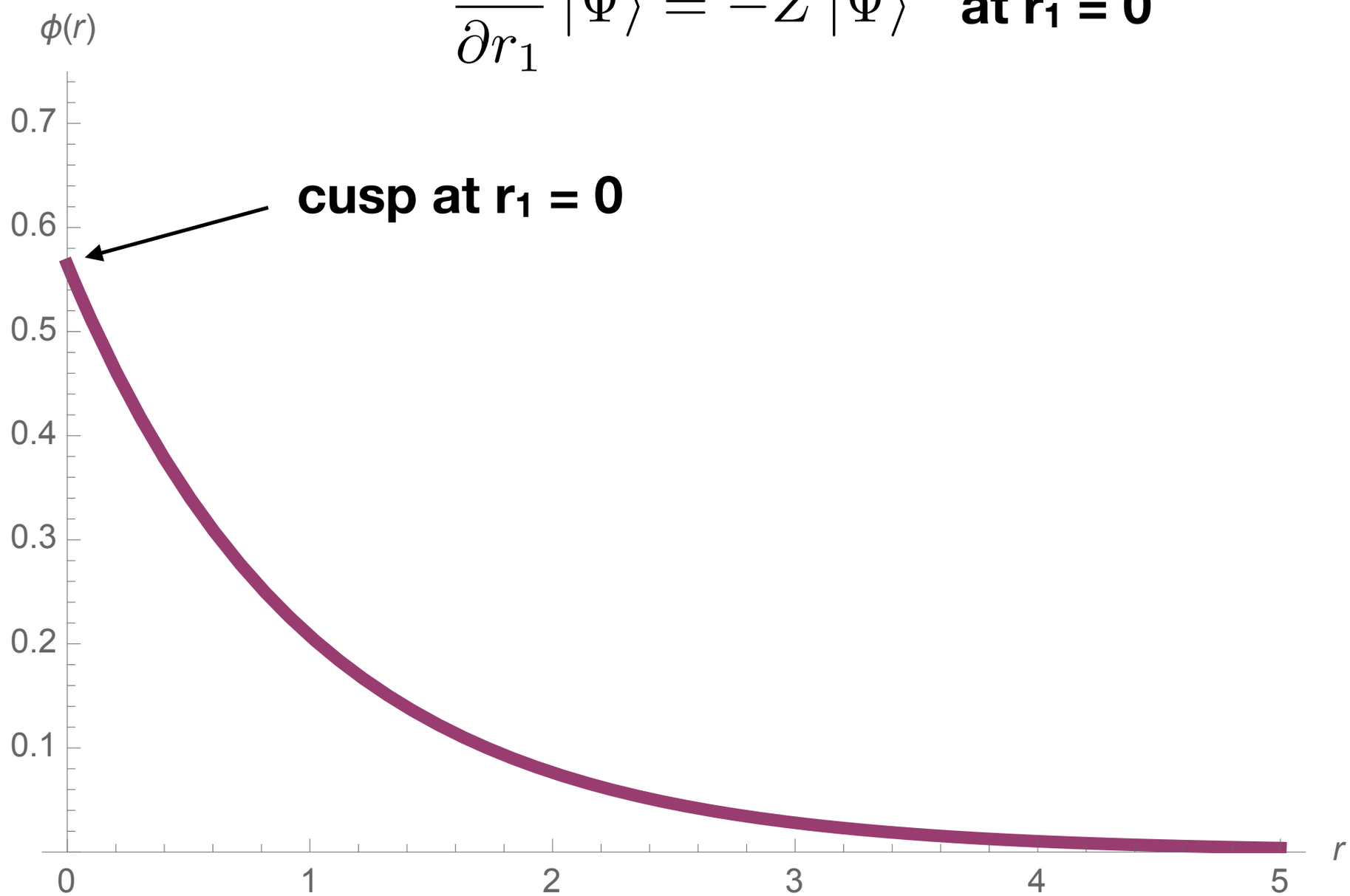
$$\left(-\frac{1}{r_1}\frac{\partial}{\partial r_1} - \frac{Z}{r_1}\right)|\Psi\rangle = 0 \qquad \Rightarrow \qquad \frac{\partial}{\partial r_1}|\Psi\rangle = -Z\,|\Psi\rangle$$

Exact wavefunction has a cusp at $r_1 = 0$!

Electron-Nuclear Cusp

Electron-nuclear cusp condition (Kato theorem):

$$\frac{\partial}{\partial r_1}\left|\Psi\right\rangle = -Z\left|\Psi\right\rangle \quad \textbf{at } r_1 = 0$$



Electron-Electron Cusp

Let's have a closer look at the Hamiltonian of the 2e⁻ system:

$$\hat{H} = -\frac{1}{2}\frac{\partial^2}{\partial r_1^2} - \frac{1}{r_1}\frac{\partial}{\partial r_1} - \frac{Z}{r_1} - \frac{1}{2}\frac{\partial^2}{\partial r_2^2} - \frac{1}{r_2}\frac{\partial}{\partial r_2} - \frac{Z}{r_2}$$
$$- \frac{r_1^2 - r_2^2 + r_{12}^2}{2r_1 r_{12}}\frac{\partial^2}{\partial r_1 \partial r_{12}} - \frac{r_2^2 - r_1^2 + r_{12}^2}{2r_2 r_{12}}\frac{\partial^2}{\partial r_2 \partial r_{12}}$$
$$- \frac{\partial^2}{\partial r_{12}^2} \boxed{- \frac{2}{r_{12}}\frac{\partial}{\partial r_{12}} + \frac{1}{r_{12}}}$$

What happens when $r_{12} = 0$?

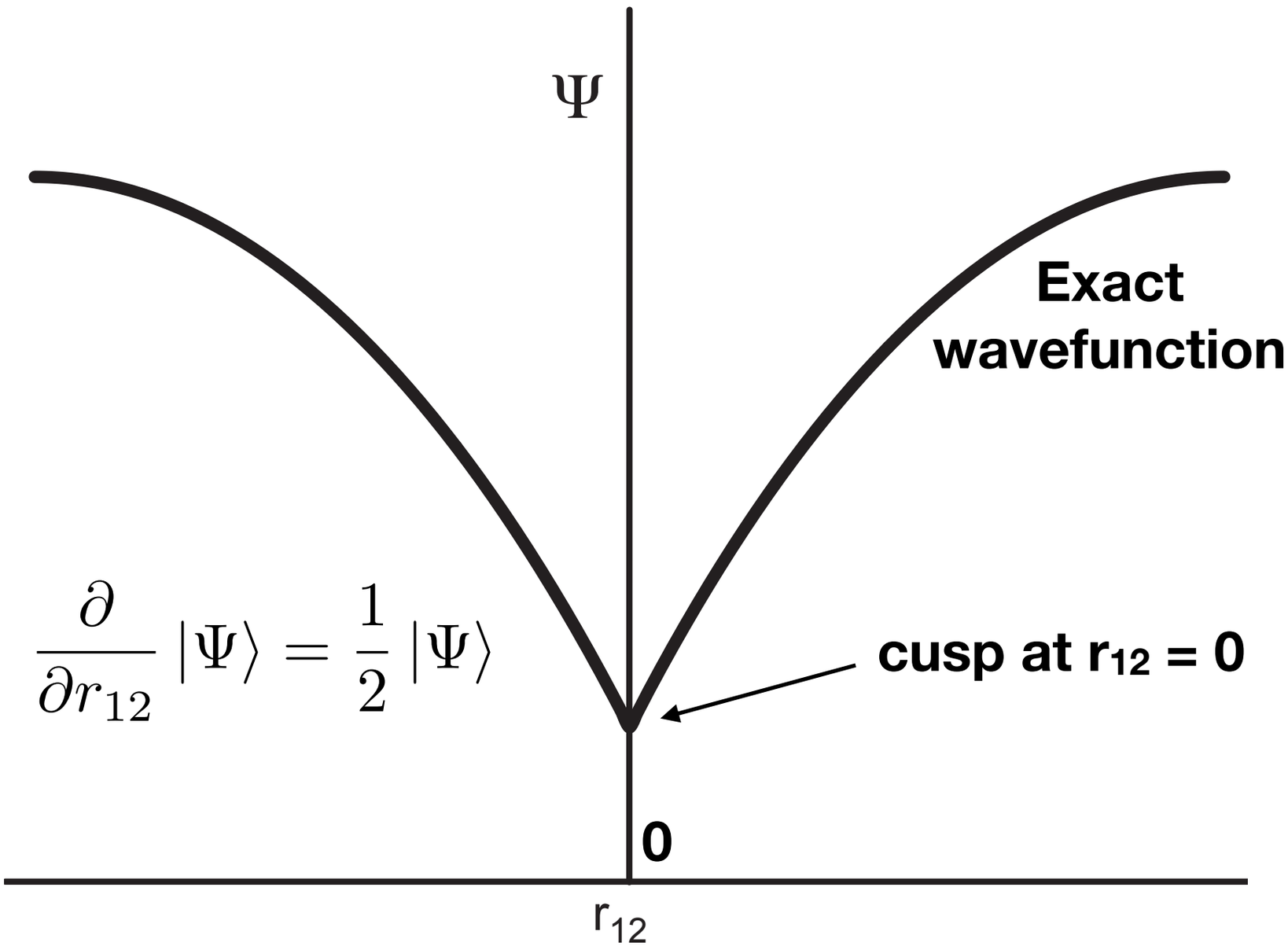
We must avoid a singularity at $r_{12} = 0$:

$$\left(-\frac{2}{r_{12}}\frac{\partial}{\partial r_{12}} + \frac{1}{r_{12}}\right)|\Psi\rangle = 0 \quad \Rightarrow \quad \frac{\partial}{\partial r_{12}}|\Psi\rangle = \frac{1}{2}|\Psi\rangle$$

Exact wavefunction has a cusp at $r_{12} = 0$!

Electron-Electron Cusp

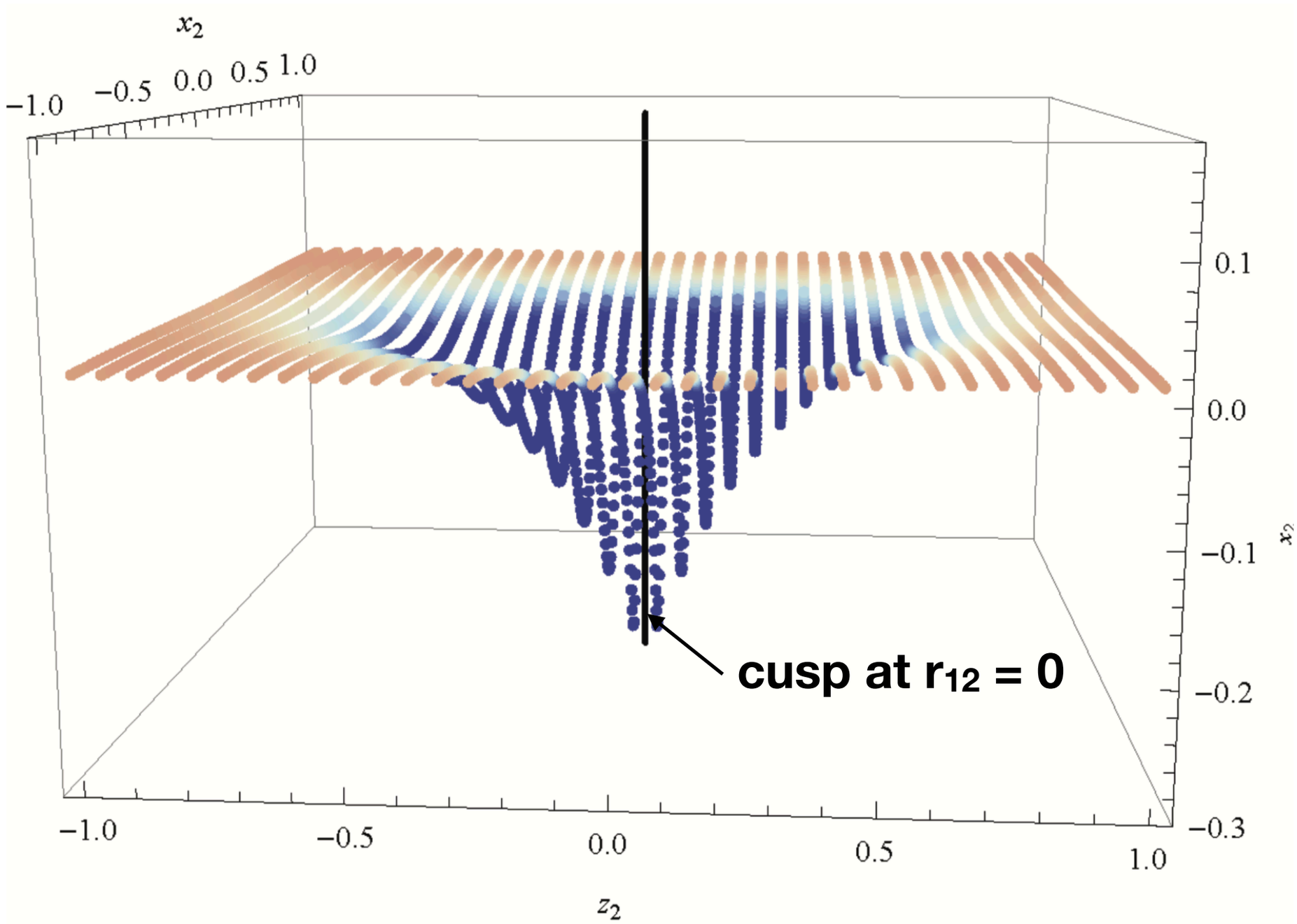
Electron-electron cusp condition (Kato theorem):



$$\frac{\partial}{\partial r_{12}} |\Psi\rangle = \frac{1}{2} |\Psi\rangle$$

Basis sets must be able to accurately describe electron-nuclear and electron-electron cusps

Electron-Electron Cusp in 3D



L. Kong, F.A. Bischoff, and E.F. Valeev, Chem. Rev. 112, 75 (2012)

Types of Basis Sets

1. Slater-type orbitals (STO's)

$$|\chi_\mu^{\mathrm{STO}}(r)\rangle = N_c\, x^k y^l z^m e^{-\alpha_\mu r}$$

- ⋆ **Defined in Cartesian space (x, y, z)**
- ⋆ **Normalization constant: N_c**
- ⋆ **Angular momentum is controlled by positive integers k, l, m:**

$$L = k + l + m$$

- ⋆ **Radial width of the orbital is controlled by the constant α: large α gives tight orbital, small α gives diffuse orbital**
- ⋆ **"Hydrogen-atom-like": correct short-range and long-range behavior**
- ⋆ **Problem: one- and two-electron integrals that involve products of STO's are difficult to compute**

Types of Basis Sets

2. Gaussian-type orbitals (GTO's)

$$|\chi_\mu^{\mathrm{GTO}}(r)\rangle = N_c\, x^k y^l z^m e^{-\alpha_\mu r^2}$$

- ⋆ **Defined in Cartesian space (x, y, z)**
- ⋆ **Normalization constant: N_c**
- ⋆ **Angular momentum is controlled by positive integers k, l, m:**

$$L = k + l + m$$

- ⋆ **Radial width of the orbital is controlled by the constant α: large α gives tight orbital, small α gives diffuse orbital**
- ⋆ **Problem: no longer "Hydrogen-atom-like", even for 1s**
- ⋆ **One- and two-electron integrals that involve products of GTO's are much easier to compute**

Types of Basis Sets

3. Contracted gaussian-type orbitals (CGTO's)

- ⋆ **Problem: STO's are more accurate, but GTO's are more computationally efficient**
- ⋆ **Solution: use a linear combination of several GTO's to accurately represent an STO**

$$|\chi_\mu^{\mathrm{CGTO}}(r)\rangle = N_c\, x^k y^l z^m \sum_\nu c_\nu e^{-\alpha_\nu r^2}$$

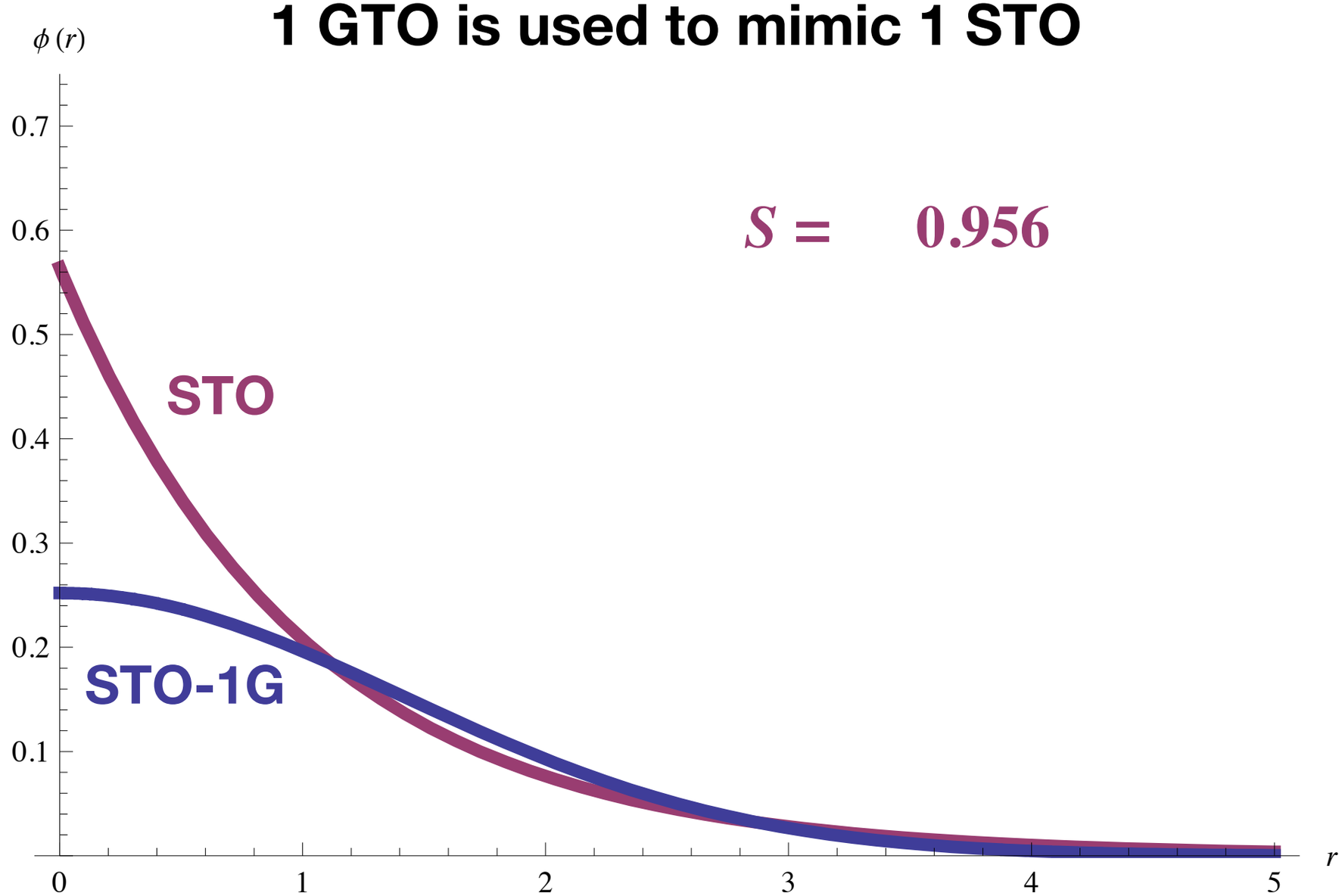
- ⋆ **CGTO's are used in almost all quantum-chemical computations**
- ⋆ **Coefficients c and α are kept fixed during a computation**
- ⋆ **Optimal values of c and α are obtained from computations of isolated atoms. They are tabulated for each basis set and can be found in papers, quantum chemistry programs, or online**

Types of Basis Sets

3. Contracted gaussian-type orbitals (CGTO's)

Example: STO-1G basis set

1 GTO is used to mimic 1 STO

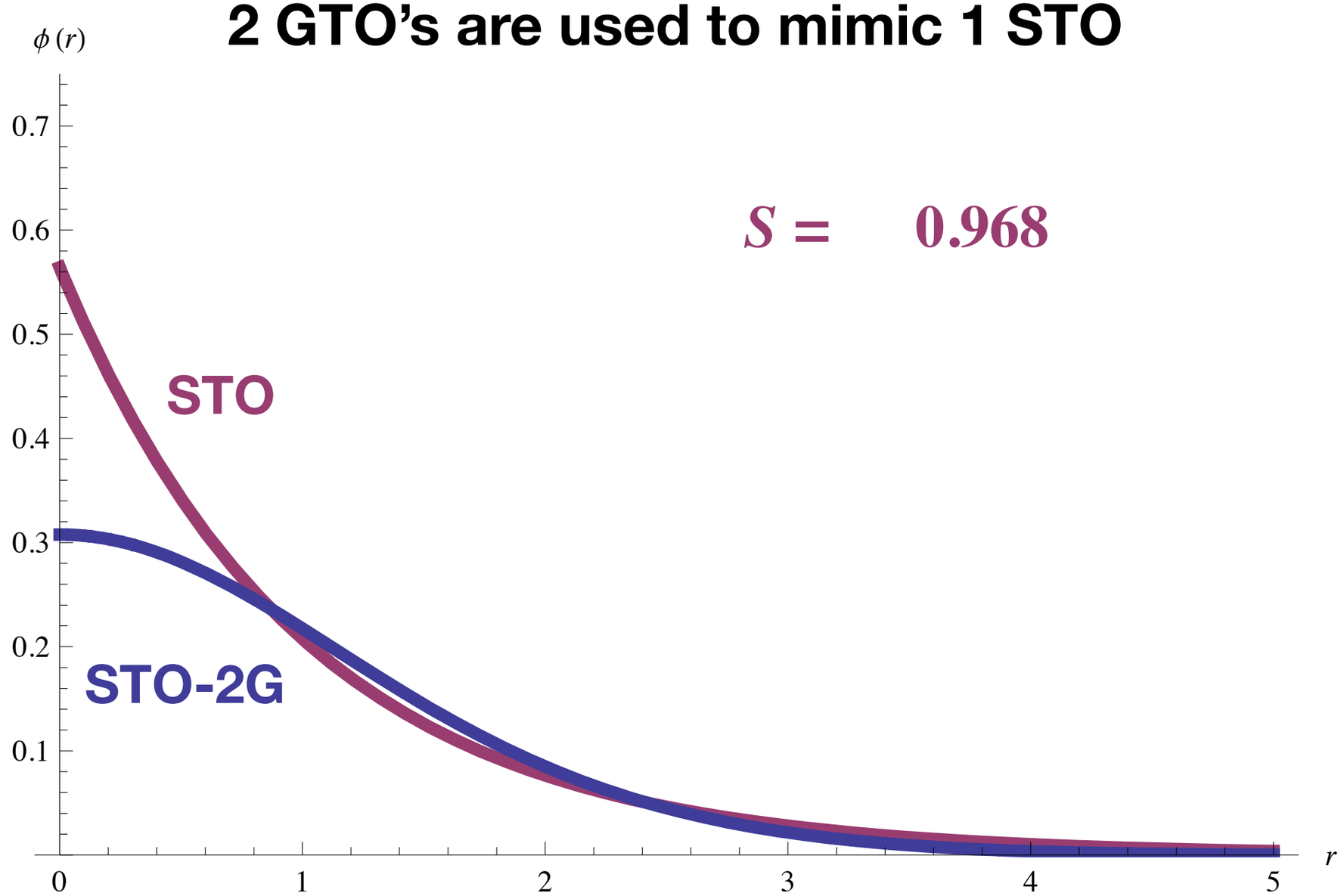


Types of Basis Sets

3. Contracted gaussian-type orbitals (CGTO's)

Example: STO-2G basis set

2 GTO's are used to mimic 1 STO

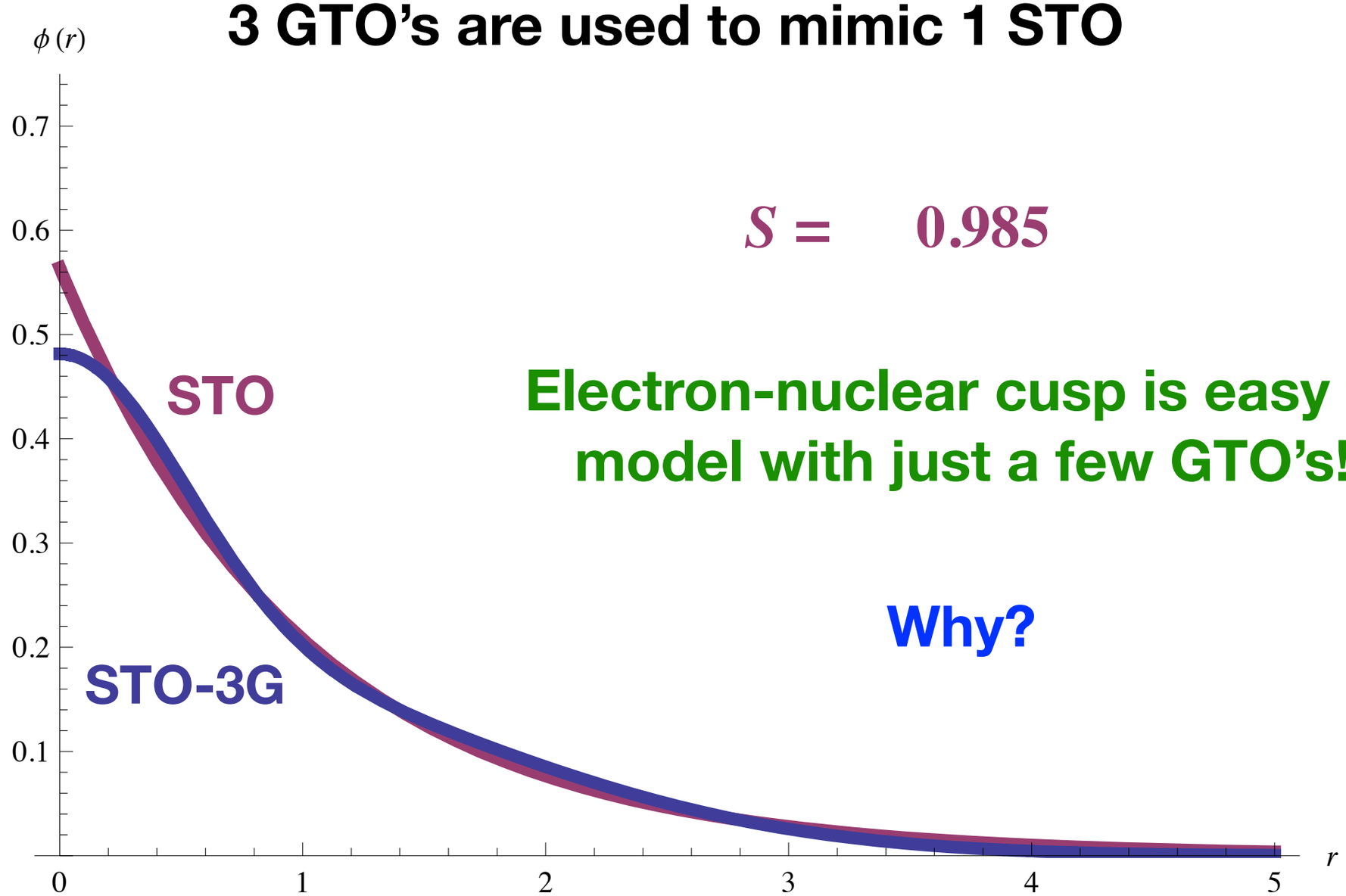


Types of Basis Sets

3. Contracted gaussian-type orbitals (CGTO's)

Example: STO-3G basis set

3 GTO's are used to mimic 1 STO



$S = \quad 0.985$

Electron-nuclear cusp is easy to model with just a few GTO's!

Why?

Classes of Basis Sets

CGTO's can be divided into several classes:

1. **Minimal:**
   - ⋆ **One basis function for each AO in an atom**
   - ⋆ **Example (N atom): 1s, 2s, $2p_x$, $2p_y$, $2p_z$ (5 basis functions)**
2. **Double-zeta (triple-zeta, quadruple-zeta, etc):**
   - ⋆ **Two (3, 4, etc) basis functions for each AO in an atom**
   - ⋆ **Example (N atom, double-zeta):**
     2 x (1s, 2s, $2p_x$, $2p_y$, $2p_z$) (10 basis functions)
3. **Split-valence double-zeta (triple-zeta, quadruple-zeta, etc):**
   - ⋆ **One basis function for each core AO, two (3, 4, etc) basis functions for the valence AO's**
   - ⋆ **Example (N atom, split-valence double-zeta):**
     1s, 2 x (2s, $2p_x$, $2p_y$, $2p_z$) (9 basis functions)

Classes of Basis Sets

CGTO's can be divided into several classes:

4. Polarized double-zeta (triple-zeta, quadruple-zeta, etc):

- ⋆ **As atoms approach each other, their orbitals may want to shift to one side or the other (polarize). An s orbital can polarize if it mixes with a p orbital**
- ⋆ **In general, to polarize a basis function with angular momentum L (usually, valence), mix it with basis functions of angular moment L + 1**
- ⋆ **This results in "polarized double-zeta" basis sets, etc**
- ⋆ **There are also polarized split-valence double-zeta basis sets (triple-zeta, etc)**
- ⋆ **Example (N atom, polarized double-zeta):**
  2 x (1s, 2s, $2p_x$, $2p_y$, $2p_z$), $3d_{z2}$, $3d_{x2-y2}$, $3d_{xy}$, $3d_{xz}$, $3d_{yz}$
  (15 basis functions)

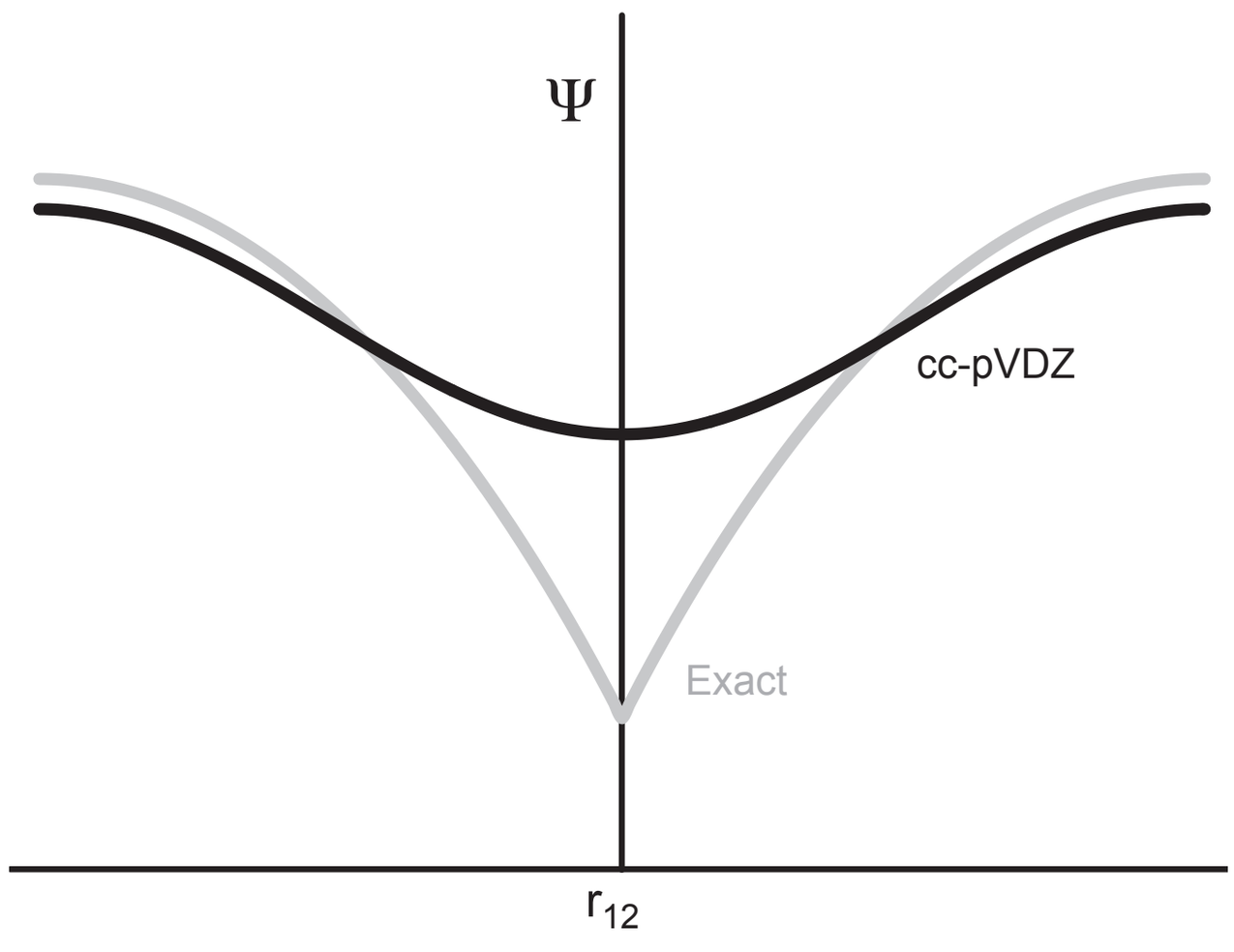
Classes of Basis Sets

CGTO's can be divided into several classes:

4. Polarized double-zeta (triple-zeta, quadruple-zeta, etc):

- ⋆ **Some polarized basis sets contain polarization functions with very high angular momentum (L = 4, 5, 6, …). These functions are used to model electron-electron cusp**

cc-pVDZ (N atom):
1s
2 x 2s
2 x 2p
3d



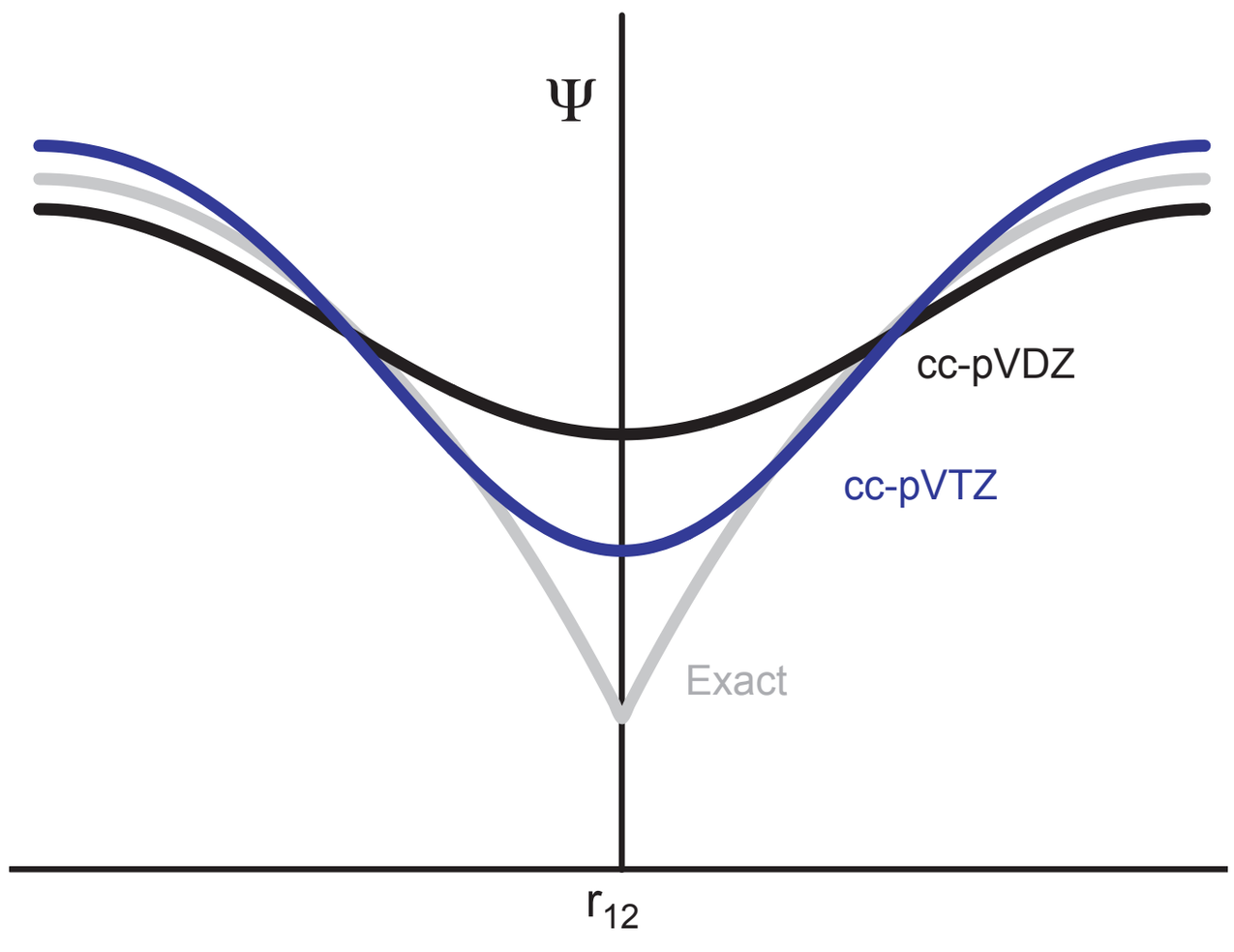
Classes of Basis Sets

CGTO's can be divided into several classes:

4. Polarized double-zeta (triple-zeta, quadruple-zeta, etc):

- ⋆ **Some polarized basis sets contain polarization functions with very high angular momentum (L = 4, 5, 6, …). These functions are used to model electron-electron cusp**

cc-pVTZ (N atom):
1s
3 x 2s
3 x 2p
2 x 3d
4f



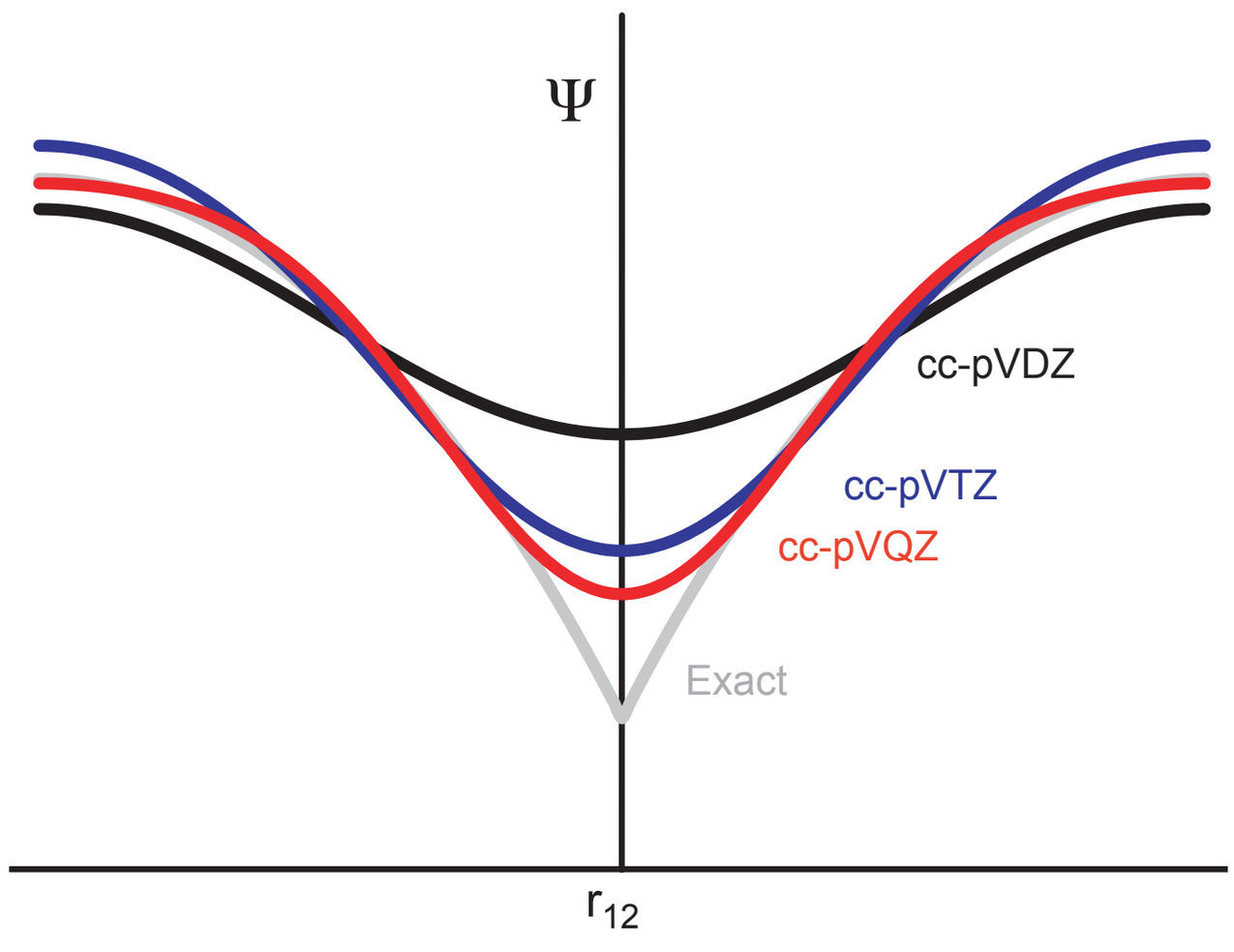
Classes of Basis Sets

CGTO's can be divided into several classes:

4. Polarized double-zeta (triple-zeta, quadruple-zeta, etc):

- ⋆ **Some polarized basis sets contain polarization functions with very high angular momentum (L = 4, 5, 6, …). These functions are used to model electron-electron cusp**

cc-pVQZ (N atom):
1s
4 x 2s
4 x 2p
3 x 3d
2 x 4f
5g



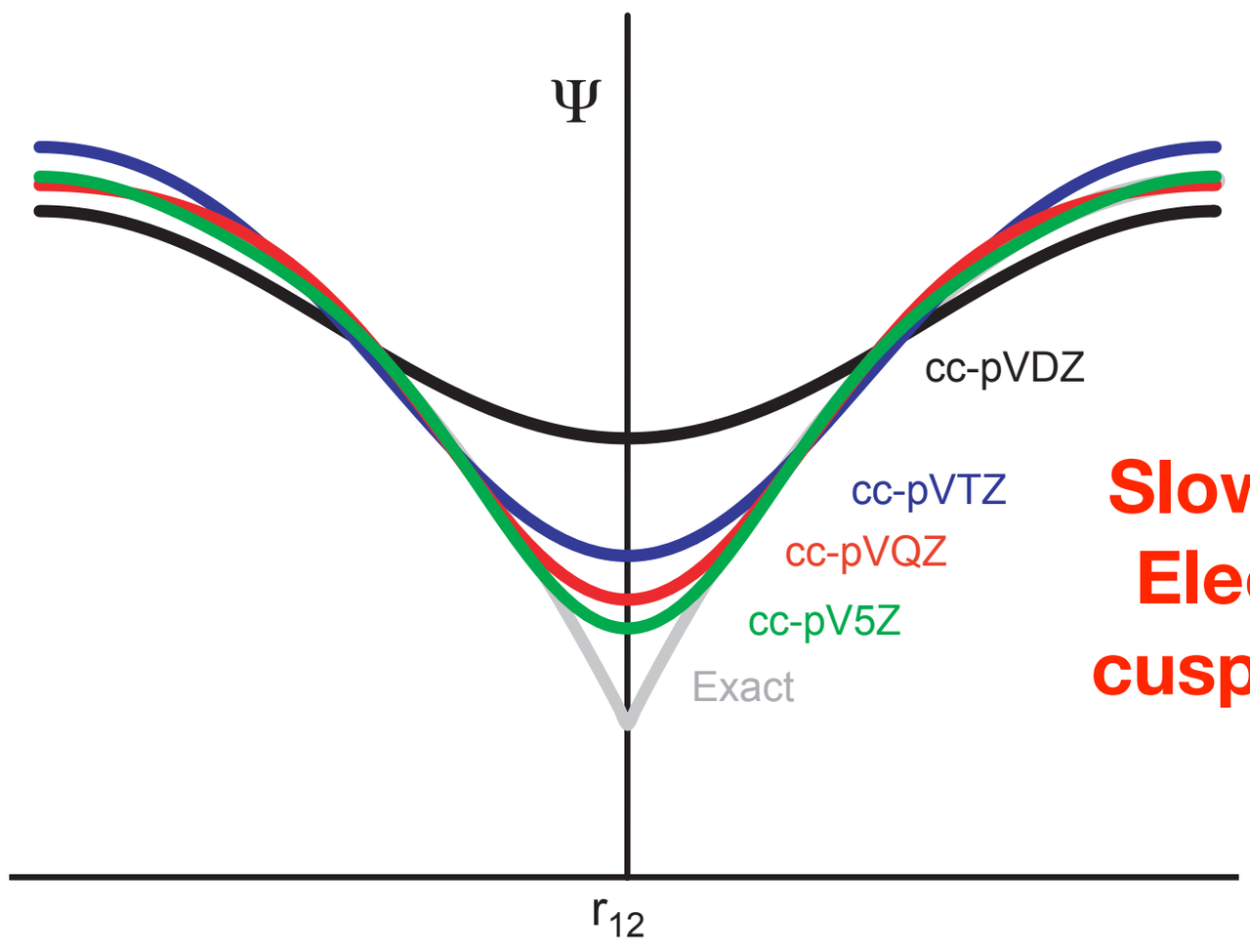
Classes of Basis Sets

CGTO's can be divided into several classes:

4. Polarized double-zeta (triple-zeta, quadruple-zeta, etc):

- ⋆ **Some polarized basis sets contain polarization functions with very high angular momentum (L = 4, 5, 6, …). These functions are used to model electron-electron cusp**

cc-pV5Z (N atom):
1s
5 x 2s
5 x 2p
4 x 3d
3 x 4f
2 x 5g
6h



Slow convergence! Electron-electron cusp is very difficult to model!

Classes of Basis Sets

CGTO's can be divided into several classes:

5. Diffuse basis sets:

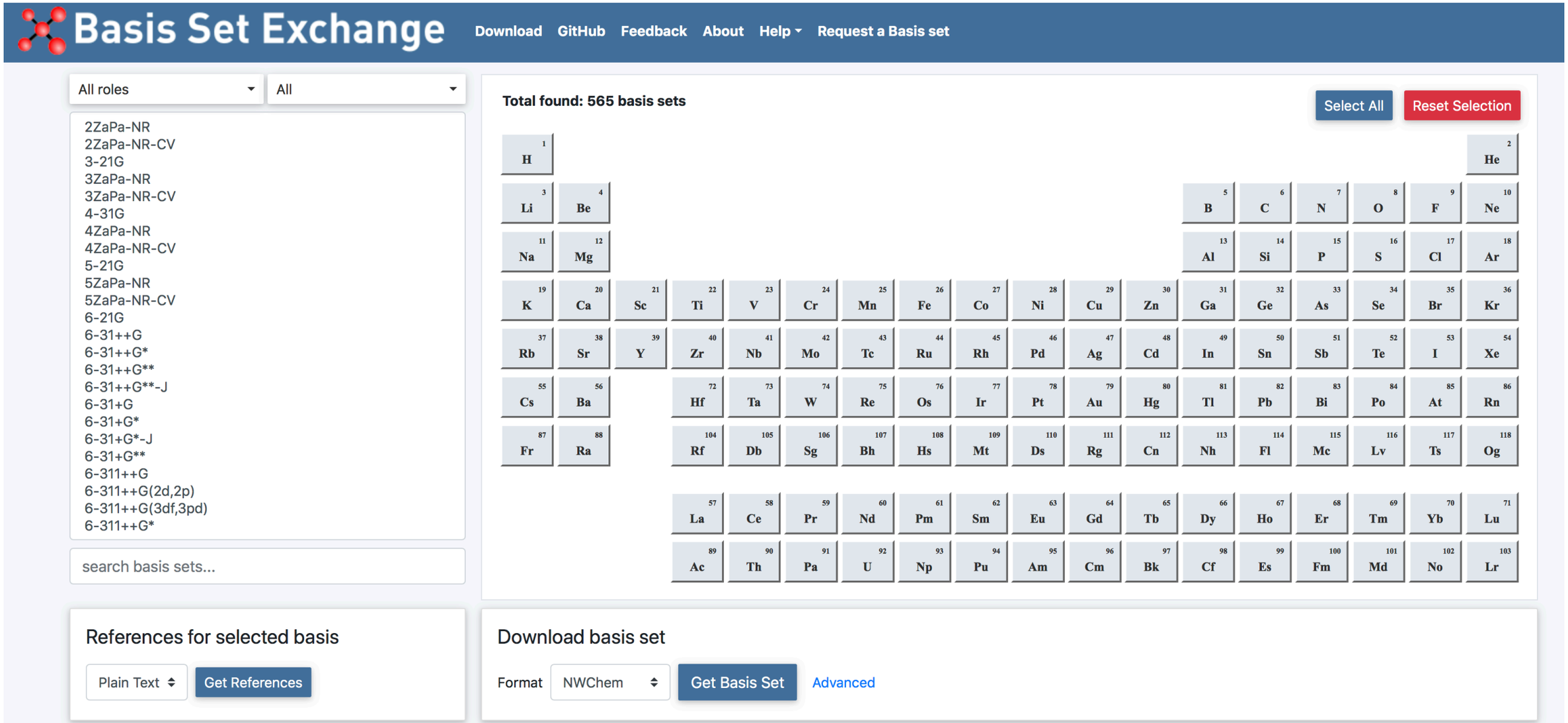
- ⋆ **Diffuse basis sets contain one or several basis functions with small α exponents**
- ⋆ **This allows electron to be far away from the nucleus if it needs to**
- ⋆ **Diffuse basis functions are often added in addition to polarization functions (diffuse polarized double-zeta ...)**
- ⋆ **Necessary for: anions, excited states, very electronegative atoms, non-bonded interactions (van der Waals complexes)**
- ⋆ **Lack of diffuse functions for anions can lead to qualitatively incorrect results!**

Popular Basis Sets

1. **STO-nG:**
   - ⋆ **Minimal basis sets: STO-3G, STO-6G**
2. **XZ (X = D, T, Q):**
   - ⋆ **DZ (double-zeta), TZ (triple-zeta), QZ (quadruple-zeta)**
3. **XZP:**
   - ⋆ **DZP (polarized double-zeta), TZP (polarized triple-zeta), …**
4. **Pople basis sets (split-valence):**
   - ⋆ **Double-zeta: 3-21G, 6-31G; triple-zeta: 6-311G**
   - ⋆ **Polarized: 6-31G*, 6-31G**, 6-311G*, 6-311G****
   - ⋆ **Diffuse: 6-31+G, 6-31++G, 6-311+G, 6-311++G**
   - ⋆ **Diffuse polarized: 6-31+G*, 6-31++G*, 6-31++G****
5. **Correlation-consistent basis sets (split-valence):**
   - ⋆ **Polarized: cc-pVXZ (X = D, T, Q, 5, …)**
   - ⋆ **Diffuse polarized: aug-cc-pVXZ (X = D, T, …)**

Basis Sets: Examples

Basis sets can be found in the Basis Set Exchange:
https://www.basissetexchange.org



Basis Sets: Examples

$$|\chi_\mu^{\mathrm{CGTO}}(r)\rangle = N_c\, x^k y^l z^m \sum_\nu c_\nu e^{-\alpha_\nu r^2}$$

atom type

H atom, STO-3G:

```
H      0
S   3   1.00
        3.42525091              0.15432897
        0.62391373              0.53532814
        0.16885540              0.44463454
```

angular momentum (L)

exponents α_ν

contraction coefficients c_ν

Minimal basis:
one s basis function composed of 3 primitive GTO's

Basis Sets: Examples

$$|\chi_\mu^{\mathrm{CGTO}}(r)\rangle = N_c\, x^k y^l z^m \sum_\nu c_\nu e^{-\alpha_\nu r^2}$$

H atom, DZ:

```
H      0
S   3   1.00
      19.2406000                0.0328280
       2.8992000                0.2312080
       0.6534000                0.8172380
S   1   1.00
       0.1776000                1.0000000
```

Double-zeta basis:
one s basis function composed of 3 primitive GTO's
plus one s basis function composed of 1 primitive GTO

Basis Sets: Examples

N atom, DZ:

```
N     0
S   6   1.00
   5909.4400000              0.0020040
    887.4510000              0.0153100
    204.7490000              0.0742930
     59.8376000              0.2533640
     19.9981000              0.6005760
      2.6860000              0.2451110
S   1   1.00
      7.1927000              1.0000000
S   1   1.00
      0.7000000              1.0000000
S   1   1.00
      0.2133000              1.0000000
P   4   1.00
     26.7860000              0.0182570
      5.9564000              0.1164070
      1.7074000              0.3901110
      0.5314000              0.6372210
P   1   1.00
      0.1654000              1.0000000
```

Double-zeta basis

What is the number of basis functions per N atom?

Basis Sets: Examples

N atom, cc-pVDZ:

```
N     0
S   9   1.00
   9046.0000000              0.0007000
   1357.0000000              0.0053890
    309.3000000              0.0274060
     87.7300000              0.1032070
     28.5600000              0.2787230
     10.2100000              0.4485400
      3.8380000              0.2782380
      0.7466000              0.0154400
      0.2248000             -0.0028640
S   9   1.00
   9046.0000000             -0.0001530
   1357.0000000             -0.0012080
    309.3000000             -0.0059920
     87.7300000             -0.0245440
     28.5600000             -0.0674590
     10.2100000             -0.1580780
      3.8380000             -0.1218310
      0.7466000              0.5490030
      0.2248000              0.5788150
S   1   1.00
      0.2248000              1.0000000
P   4   1.00
     13.5500000              0.0399190
      2.9170000              0.2171690
      0.7973000              0.5103190
      0.2185000              0.4622140
P   1   1.00
      0.2185000              1.0000000
D   1   1.00
      0.8170000              1.0000000
```

What basis set class is this?

What is the number of basis functions per N atom?

Cartesian vs Spherical Functions

$$|\chi_\mu^{\mathrm{CGTO}}(r)\rangle = N_c\, x^k y^l z^m \sum_\nu c_\nu e^{-\alpha_\nu r^2}$$

In Cartesian space, there are 6 functions with L = 2:

$$k + l + m = 2 \quad \Rightarrow \quad d_{x^2} \quad d_{y^2} \quad d_{z^2} \quad d_{xz} \quad d_{xy} \quad d_{yz}$$

But there are only five d-orbitals! Why?

Orbitals are spherically-symmetric!

Six functions with L = 2 form a basis for five d-orbitals and one s-orbital!

$d_{x^2} + d_{y^2} + d_{z^2}$ **behaves like an s orbital**

For L = 3, there are 10 Cartesian and 7 spherical functions

Some programs work with Cartesian functions (6d, 10f), while some use spherical (5d, 7f)

Basis Set Convergence

Convergence of bond distances with the basis set

cc-pVXZ (X = D, T, Q, 5, 6) basis set

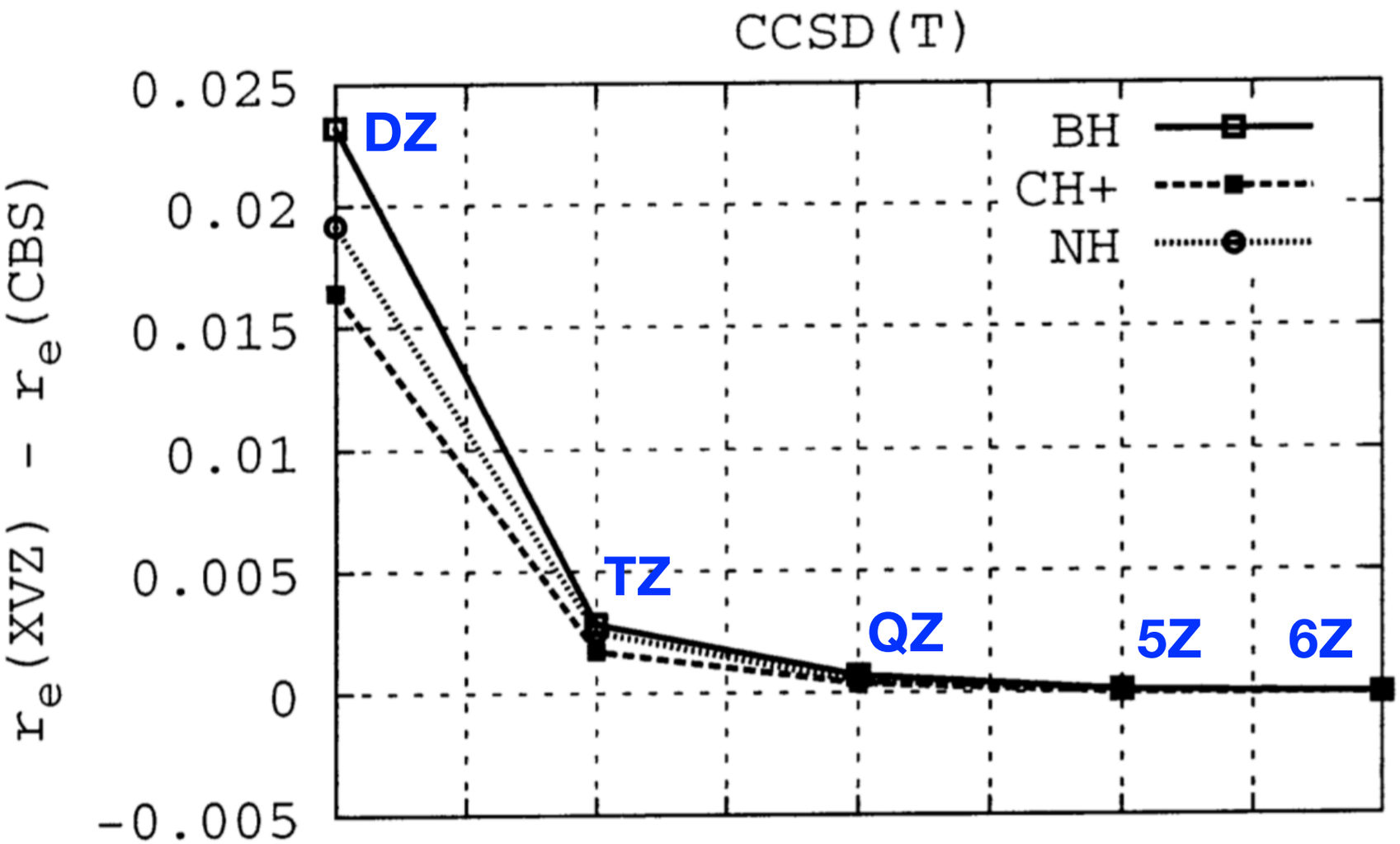


Figure 1. Convergence of CCSD(T) r_e and ω_e toward the CBS limit derived for valence-only (cc-pVNZ) basis sets.

B. Temelso, E.F. Valeev, and C.D. Sherrill, J. Phys. Chem. A 108, 3068 (2004)

Basis Set Convergence

Convergence of vibrational frequencies with the basis set

cc-pVXZ (X = D, T, Q, 5, 6) basis set

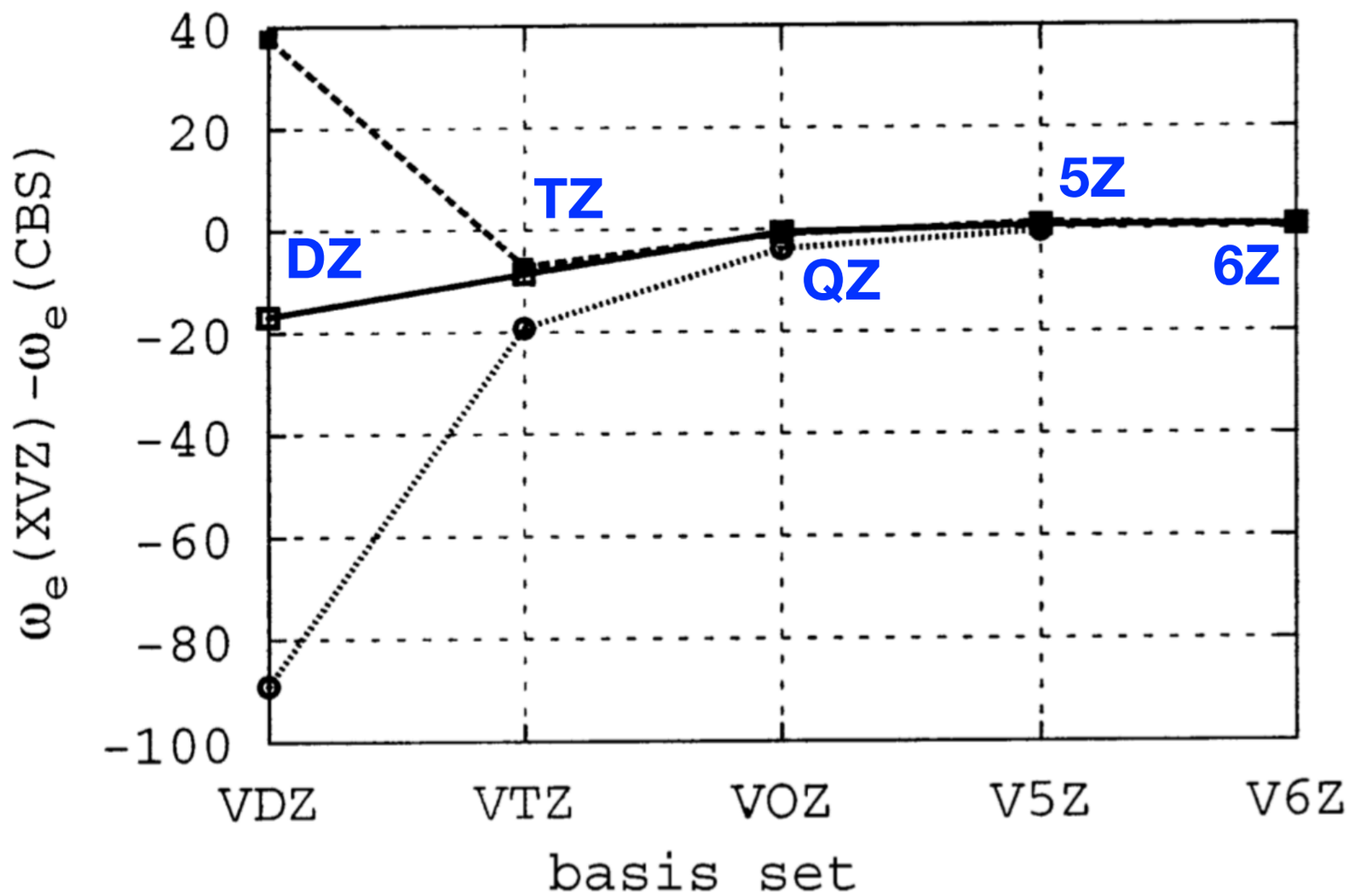


Figure 1. Convergence of CCSD(T) r_e and ω_e toward the CBS limit derived for valence-only (cc-pVNZ) basis sets.

B. Temelso, E.F. Valeev, and C.D. Sherrill, J. Phys. Chem. A 108, 3068 (2004)

Basis Set Convergence: Methods

Different methods show different convergence with basis set

cc-pCVXZ (X = D, T, Q) basis set

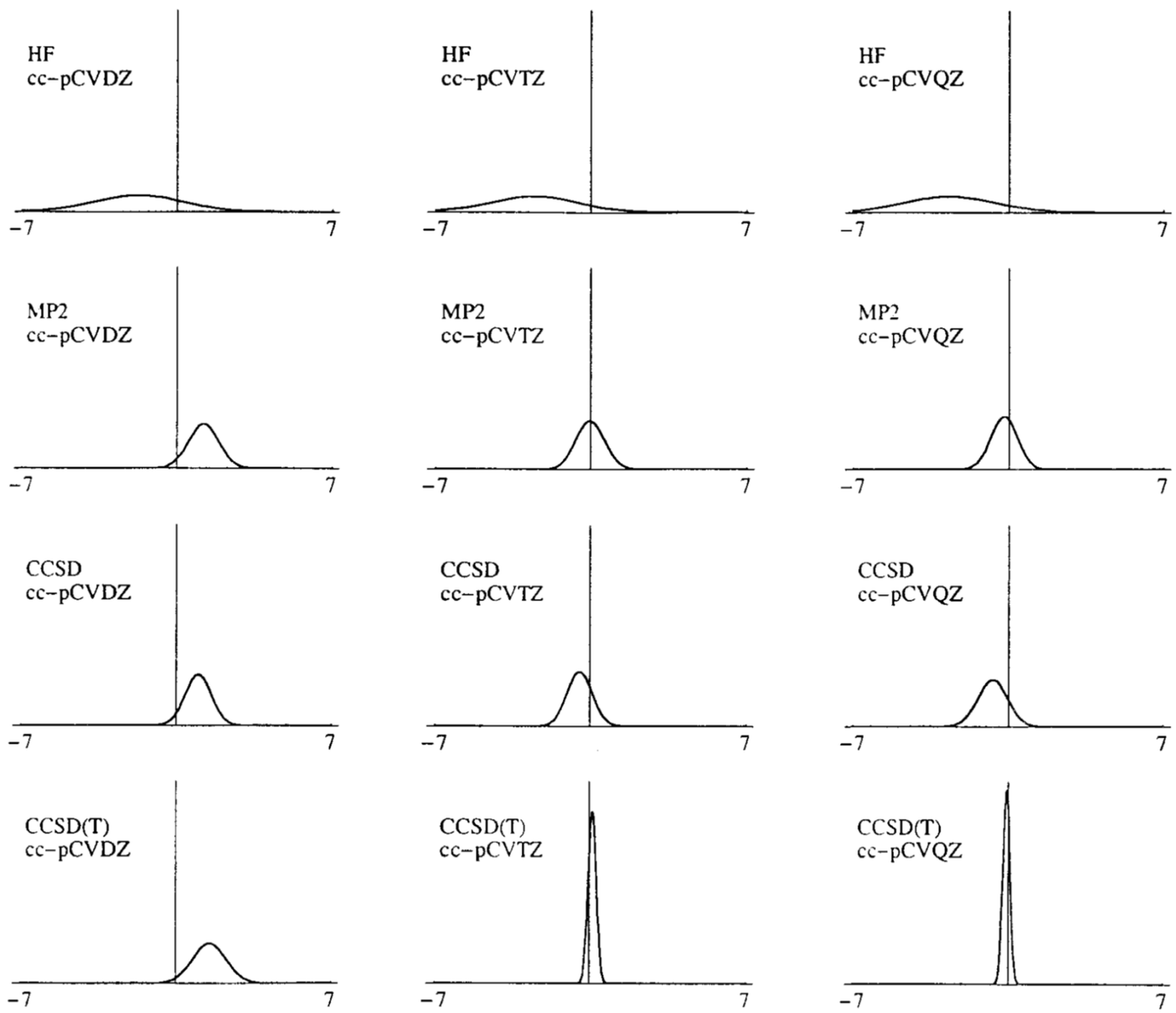


FIG. 2. Normal distributions of errors in the calculated bond distances (pm). For ease of comparison, all distributions have been normalized to one and plotted against the same horizontal and vertical scales.

J. Chem. Phys. 114, 6548 (2001)

Basis Sets: General Comments

- ⋆ **The bigger the basis, the better? Usually yes, but for some methods increasing the basis can make the results worse (loss of error cancellation)**
- ⋆ **For some methods (e.g., HF or DFT), basis set convergence is faster than for others (e.g., coupled cluster)**
- ⋆ **It is usually better to start with a modest basis set and work your way up, monitoring the basis set convergence**
- ⋆ **Minimal basis sets should only be used for test purposes**
- ⋆ **For accurate results, basis set should be at least double-zeta (ideally, triple-zeta) quality with polarization functions**
- ⋆ **Diffuse functions must be used for anions, excited states**
- ⋆ **There are many basis sets out there, some may be better than others. Make sure you test them before you use them**