6 | Basis Sets

In the previous chapter we introduced the Hartree–Fock method. The goal of this chapter is to discuss some practical aspects of how Hartree–Fock computations actually run on a computer and the way the are approximated. We will encounter the idea of using a linear combination of basis functions to approximate the orbitals $\phi(\mathbf{r})$ in systematically improvable way. Lastly, we will look at some applications of these ideas.

6.1 Exact solutions to the Schrödinger equation for the hydrogen atom

As it was mentioned earlier, we know exact solutions of the Schrödinger only for a handful of models. An example that should be familiar to you is the hydrogen atom. For this system, the solutions are known to depend on three quantum numbers n, l, and m_l, and that the energy levels (eigenvalues) are given by (in atomic units)

$$E_n = -\frac{1}{2n^2} E_h, \quad n = 1, 2, \ldots \tag{6.1}$$

The wave function ψ_{nlm_l} is usually expressed in **spherical coordinates** (r, θ, φ)[1] as a product of a function that depends on the electron distance $R_{nl}(r)$ and a term that depends on the angles θ and φ, which we write as $Y_l^{m_l}(\theta, \varphi)$

$$\psi_{nlm_l}(r, \theta, \varphi) = R_{nl}(r)Y_l^{m_l}(\theta, \varphi) \tag{6.2}$$

For example, the wave function for the ground state of hydrogen, the 1s orbital, is given by

$$1s = \psi_{1,0,0}(r, \theta, \varphi) = R_{1,0}(r)Y_0^0(\theta, \varphi) = 2e^{-r}\frac{1}{\sqrt{4\pi}} = \frac{1}{\sqrt{\pi}}e^{-r} \tag{6.3}$$

while the 2p$_z$ orbital is given by

$$2p_z = \psi_{2,1,0}(r, \theta, \varphi) = R_{2,1}(r)Y_1^0(\theta, \varphi) = \frac{1}{4\sqrt{2\pi}}re^{-r/2}\cos\theta. \tag{6.4}$$

When we turn to more complicated systems, like atoms with more than one electron and molecules, the Schrödinger equation becomes intractable and we have to turn to approximate numerical methods to solve it (recall Dirac's quote on this point). This is what we will look into in the next section.

[1] If you are curious, spherical coordinates are connected to Cartesian coordinates via the following equations

$$x = r\sin\theta\cos\varphi,$$
$$y = r\sin\theta\sin\varphi,$$
$$z = r\cos\theta$$
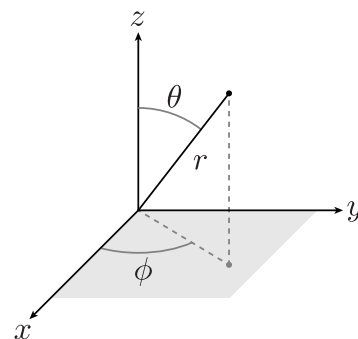
Fig. 6.1: Definition of the spherical coordinates r, θ, and φ.
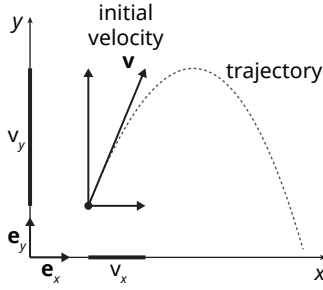
Fig. 6.2: The trajectory of a particle with initial velocity vector $\mathbf{v} = (v_x, v_y)$. This vector may be decomposed in the basis of orthogonal unit vectors $\mathbf{e}_x$ and $\mathbf{e}_y$ as $\mathbf{v} = v_x\mathbf{e}_x + v_y\mathbf{e}_y$.

[2] A linear combination of vectors/functions/etc. (f_i) is a sum

$$f_1 c_1 + f_2 c_2 + \ldots + f_K c_K$$

where the quantities $c_1, c_2, \ldots$ are numbers (they can be real or complex).

[3] A basis can have orthogonal or nonorthogonal vectors. However, bases that are orthogonal are typically preferred.

6.2   Bases, basis vectors, and basis functions

An important concept when constructing numerical approximations is that of a basis. To build some intuition about the concept of a basis you can think back to your introductory physics courses and Newton's equations. Consider the problem of describing the trajectory of a particle under the effect of gravity. To solve this problem you need to specify the position and initial velocity of the particle. The position can be specified by the x and y coordinates, and the initial velocity by the velocity in x and y directions (v_x, v_y). These quantities are essentially vectors, which we indicate with the symbols $\mathbf{r} = (x, y)$ and $\mathbf{v} = (v_x, v_y)$. Figure 6.2 shows the position and velocity vectors.

Every vector can be written as a **linear combination**[2] of vectors that are linearly independent and span the full space, which we call a **basis**. For example, the velocity vector $\mathbf{v}$ can be expressed as a sum of a component pointing in the x direction ($\mathbf{e}_x$) and one pointing in the y direction ($\mathbf{e}_y$)

$$\mathbf{v} = v_x\mathbf{e}_x + v_y\mathbf{e}_y, \text{ with } \mathbf{e}_x = (1, 0) \text{ and } \mathbf{e}_y = (0, 1) \tag{6.5}$$

The vectors $\mathbf{e}_x$ and $\mathbf{e}_y$ are called a basis, because any vector in two dimensions $\mathbf{a}$ can be written as a sum of coefficients multiplied by the basis vectors

$$\mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y \tag{6.6}$$

The vectors $\mathbf{e}_x$ and $\mathbf{e}_y$ have another special property: they are orthogonal, which means that their dot product is zero[3]

$$\mathbf{e}_x \cdot \mathbf{e}_y = 0. \tag{6.7}$$

In the more general case, the velocity is a three dimensional vector, and must be described as a sum of three terms

$$\mathbf{v} = v_x\mathbf{e}_x + v_y\mathbf{e}_y + v_z\mathbf{e}_z \tag{6.8}$$

Interestingly, this idea can also be extended to functions. For example, suppose we want to approximate some function $g(x)$. We could start from choosing two simple functions $f_1(x)$ and $f_2(x)$ and approximate $g(x)$ as

$$g(x) = c_1 f_1(x) + c_2 f_2(x) \tag{6.9}$$

One way to think about this is to discretize the integral as a sum. The integral of a function $f(x)$ is the limit of $\Delta x \to 0$ of the sum

$$\int_a^b dx\, f(x) = \Delta x \sum_{k=1}^{n} f(x_k)$$

where the points x_k form a grid with spacing Δx. The integral of a product of two functions $u^*(x)v(x)$ is then also given by the limit of the following expression

$$\int_a^b dx\, u^*(x)v(x) = \Delta x \sum_{k=1}^{n} u^*(x_k)v(x_k)$$

If we identify the value of the functions at the points x_k as the components of a vector, for example, $v(x_1) = v_1, v(x_2) = v_2, \ldots$, then the above expression can be written as

$$\int_a^b dx\, u^*(x)v(x) = \Delta x [u_1 v_1 + u_2 v_2 + \ldots]$$

This looks very similar to the definition of the dot product of two vectors of arbitrary dimension ($\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \ldots$).

We can think of this equation as being analogous to taking a combination of basis vectors, with the difference being that we are combining the functions $f_1(x)$ and $f_2(x)$. This is not just an analogy. We use the same language because you can think of a function as a vector with infinite dimensions and most of the concepts that apply to vectors (like length, being orthogonal, etc.) can be generalized to functions.

When "fitting" the function $g(x)$ in this way, it is convenient to work with functions that are **orthogonal**. But how does one define the idea of orthogonality for functions? A natural generalization is the overlap integral ($\langle f_1 | f_2 \rangle$) [4]

$$"f_1 \cdot f_2" = \langle f_1 | f_2 \rangle = \int_{-\infty}^{\infty} dx\, f_1^*(x)f_2(x) \tag{6.10}$$

From this definition, we define two functions to be orthogonal if their dot product (inner product) is null, $\langle f_1 | f_2 \rangle = 0$.

6.3 Approximate solutions via expansion in a basis

We can now look into how the idea discussed in the previous section can be applied to approximate solutions of the Schrödinger equation. The main point is to use a **linear combination of basis functions** to approximate the wave function. This is a very common strategy for studying differential equations that are too difficult to solve in analytical form. In Eq. (6.9) we considered the case of a basis composed by two functions. However, in the more general case, we approximate $g(x)$ as a sum of a finite number (K) of fixed functions $f_i(x)$, called basis functions. Each basis function is multiplied by the coefficients c_i and the approximation is given by

$$g(x) \approx c_1 f_1(x) + c_2 f_2(x) + \cdots + c_K f_K(x) = \sum_{i=1}^{K} c_i f_i(x) \tag{6.11}$$

You can think of this approximation as fitting the function $g(x)$ with the basis functions, where the unknowns are the fitting coefficients c_i. An important point is that, if the basis $\{f_i(x)\}$ is chosen judiciously, the approximate solution should converge rapidly to the exact $g(x)$ when the number of basis functions (K) is increased.

This process is illustrated in Fig. 6.3, where the function $\exp(x)\sin(x)$ is approximated with a basis of polynomials of x. In this example, we stop at the polynomial of order K (here we also include the constant term $x^0 = 1$)

$$g(x) \approx c_0 + c_1 x + \cdots + c_K x^K = \sum_{i=0}^{K} c_i x^i \tag{6.12}$$

As you can see, the first two approximations are not very good, and by the time we use a forth-order polynomial we cannot easily distinguish $\exp(x)\sin(x)$ from its approximation.

6.4 Gaussian-type orbitals

To understand how we use bases in quantum chemistry we will start by learning how atomic orbitals are approximated with a Gaussian basis. Since the early days of computational chemistry it has become common practice to approximate atomic orbitals using a basis of **Gaussian functions**. For example, if we want to approximate the 1s orbital of the hydrogen atom, $1s = \frac{1}{\sqrt{\pi}} e^{-r}$, an expansion in Gaussian functions consists of the following approximation

$$\frac{1}{\sqrt{\pi}} e^{-r} \approx \sum_{\mu} C_{\mu} \exp(-a_{\mu} r^2) \tag{6.13}$$

At this point we should ask: **Why would we want to approximate a function that we already know how to write in closed form?** The answer is that all quantum chemistry methods need to know integrals of atomic orbitals like

$$\int d\mathbf{r}_1 \, d\mathbf{r}_2 \, \frac{\phi_i^*(\mathbf{r}_1)\phi_j^*(\mathbf{r}_2)\phi_k(\mathbf{r}_1)\phi_l(\mathbf{r}_2)}{r_{12}} \tag{6.14}$$
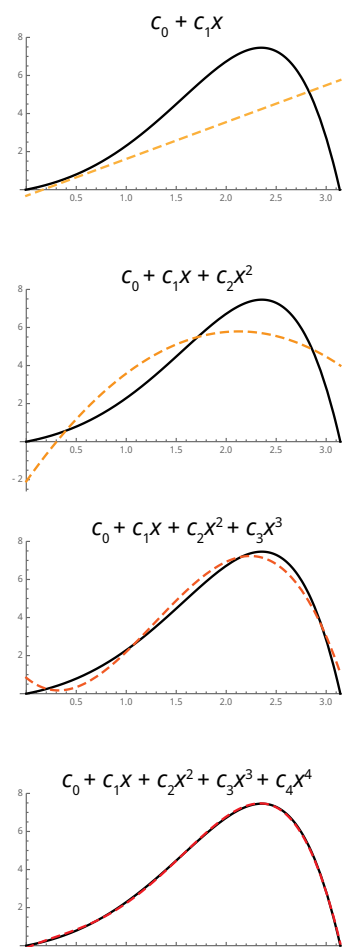

Fig. 6.3: Example of expansion of the function $\exp(x)\sin(x)$ in a basis of polynomials x^i. By the time we include fourth powers of x this function is approximated well in the entire range $0 \le x \le \pi$.

[5]This is due to the Gaussian product theorem, which states that the product of two multidimensional Gaussians is still a Gaussian function.

These integrals are difficult to compute with function of the form e^{-ar}, but they turn out to be doable if one instead uses functions of the form $\exp(-ar^2)$.[5]

For each type of atomic orbital (s, p, d, etc.) one can write a corresponding **Gaussian-type orbital** obtained by multiplying a 3D spherical Gaussian function $\exp(-ar^2)$ times a polynomial in the coordinates x, y, and z

$$x^l y^m z^n e^{-ar^2} \tag{6.15}$$

As we have seen, s orbitals are obtained from a combination of Gaussians

$$e^{-ar^2} \tag{6.16}$$

while p orbitals can be represented with a product of x, y, or z and a Gaussian

$$xe^{-ar^2}, ye^{-ar^2}, ze^{-ar^2} \tag{6.17}$$

Atomic orbitals of d type can be formed by quadratic polynomials times a Gaussian. In this case there are six combinations

$$xye^{-ar^2}, yze^{-ar^2}, xze^{-ar^2}, x^2 e^{-ar^2}, y^2 e^{-ar^2}, z^2 e^{-ar^2} \tag{6.18}$$

We call these 6 d orbitals **Cartesian** d functions, to distinguish them from the 5 **spherical** or **pure angular momentum** d functions that are obtained when we solve the Schrödinger equation

$$xye^{-ar^2}, yze^{-ar^2}, xze^{-ar^2}, (x^2 - y^2)e^{-ar^2}, (2z^2 - x^2 - y^2)e^{-ar^2} \tag{6.19}$$

Similarly, there are more Cartesian f functions (10) than spherical f function (7), etc. In a quantum chemistry code, integrals are first computed using Cartesian and then these are combined to form spherical functions. Some codes might allow you to use Cartesian or spherical functions, but you should always use the default.
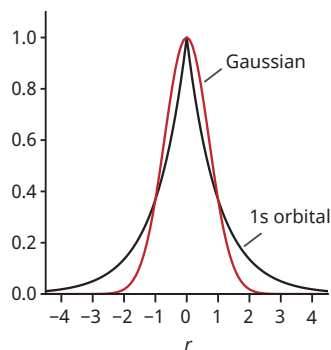
6.5   Contracted Gaussian-Type Orbitals

The Gaussians functions defined in Eq. (6.15) are also called **primitives**. Figure 6.4 reports a comparison of the shape of the hydrogen 1s orbital and a 1s Gaussian primitive. As you notice, the Gaussian and the 1s orbital are both peaked at the origin. However, the Gaussian has a rounded top at $r = 0$ and decays too quickly when $r > 1$. To represent the correct shape of atomic orbitals several primitives are combined together using a fixed set of coefficients to form numerically accurate approximations to atomic orbitals. We call these function **contracted Gaussians** and to avoid confusing them for orbitals, they will be denoted with the Greek letter chi (χ). For example, a p_z contracted Gaussian centered at the origin is approximated as

$$\chi_{p_z}(\mathbf{r}) = \sum_\mu b_\mu z \exp(-a_\mu r^2) \tag{6.20}$$

The quantities a_μ are called the **exponents** of the basis, while the coefficients b_μ are called **contraction coefficients**.

Fig. 6.4: Plot of the STO-6G Gaussian basis set and the six Gaussian primitives for the hydrogen atom.

A **Gaussian basis set** is a collection of contracted Gaussian functions with optimized exponents and contraction coefficients. Making a basis sets is not trivial and there are many articles in the literature describing how to build basis set. The most commonly used basis sets are collected in the **basis set exchange** on-line repository (https://www.basissetexchange.org).

It is instructive to take a look at an example of a Gaussian basis used by a computer program. The following shows the **STO-6G** basis set for the hydrogen atom. This basis approximates one atomic orbital of hydrogen using six Gaussians (hence the "6G"). This basis includes only one atomic orbital for each hydrogen. The first line lists the element name, while the second one is a tag that indicates that the next few lines show the exponent and contraction coefficients for a s-type ("S") orbital approximated with 6 Gaussians. The left column shows the exponent of the Gaussians, while the second column the corresponding contraction coefficients

```
H  0
S    6    1.00
        0.3552322122E+02        0.9163596281E-02
        0.6513143725E+01        0.4936149294E-01
        0.1822142904E+01        0.1685383049E+00
        0.6259552659E+00        0.3705627997E+00
        0.2430767471E+00        0.4164915298E+00
        0.1001124280E+00        0.1303340841E+00
```

This basis corresponds to the function

$$\chi(\mathbf{r}) = 0.009163596281 \, \exp(-35.52322122 \, r^2) + \dots \tag{6.21}$$

which is plotted in Fig. 6.5. This figure also shows the six Gaussian primitives that are combined together to obtain the STO-6G basis function. The STO-6G function looks like a Gaussian but it is more pointy at the origin and decays more slowly for large values of *r*. **Even though the STO-6G basis uses six Gaussians, this basis consists of only one function.**

The STO-6G basis is not a very flexible basis, since it does not allow the orbital of the hydrogen atom to shrink, grow, or polarize. For example, if we consider the H⁻ anion, the STO-6G basis does not allow the occupied orbital to expand due to the increased electron-electron repulsion experienced due to the extra negative charge. In practice, computation basis include more functions than the strict valence atomic orbitals. For example, the def2-SVP basis for hydrogen is defined as

```
H      0
S    3    1.00
       13.0107010          0.19682158E-01
        1.9622572          0.13796524
        0.44453796         0.47831935
S    1    1.00
        0.12194962         1.0000000
P    1    1.00
        0.8000000          1.0000000
```
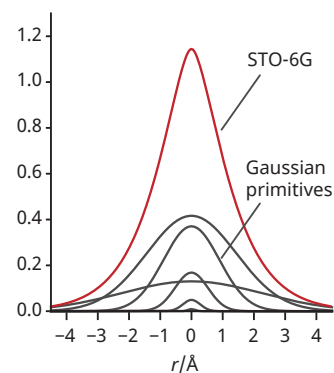

Fig. 6.5: Plot of the STO-6G Gaussian basis set and the six Gaussian primitives for the hydrogen atom.
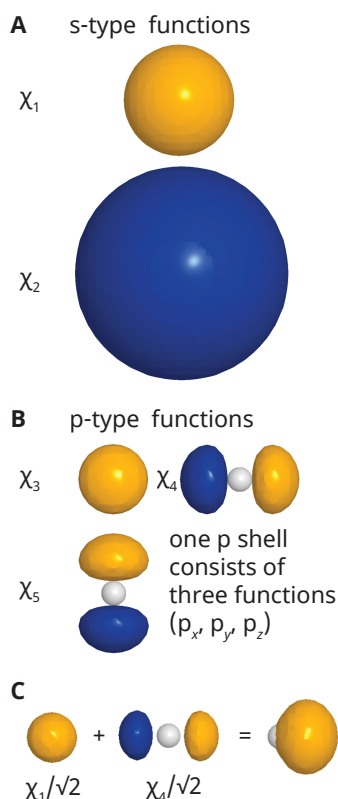
A   s-type functions

χ_1

χ_2

B   p-type functions

χ_3 χ_4

χ_5 one p shell
consists of
three functions
(p_x, p_y, p_z)

C

$\chi_1/\sqrt{2}$ $\chi_4/\sqrt{2}$ =

Fig. 6.6: Plot of the def2-SVP Gaussian basis set for the hydrogen atom. (**A**) split-valence s-type functions and (**B**) p-type polarization functions. The combination of a s-type and p-type function is a s-like function polarized along the y direction (**C**)

[6]Other examples include using plane waves $[\exp(i\mathbf{k} \cdot \mathbf{r})]$ and numerical grids to approximate the orbitals.

This basis contains two s-type contracted Gaussians (SV = split valence, meaning two sets of valance orbitals) plus a set of three p orbitals (p_x, p_y, p_z), which are called **polarization functions** (P = polarization). Figure 6.6 shows these five basis functions. By combining orbitals χ_1 and χ_2 it is possible to form spherical orbitals of different size. The p orbitals instead allow to polarize the wave function along any direction in space as shown in part **C** of this figure.

Gaussian basis sets can be built in such a way that the accuracy of the orbitals can be systematically improved. The following table shows the energy of the hydrogen atom computed with a family of basis sets (cc–pVXZ, with $X =$D,T,Q,5,6) that contains an increasing number of Gaussian functions (this basis contains more than just s functions). As you can see, the cc-pVDZ basis is already very accurate (the absolute energy error is less than 0.001 mE_h) and the error can be reduced by almost one order of magnitude as we go from one basis to next one in this family.

Table 6.1: Convergence of the energy of the hydrogen atom as a function of the computational basis. The exact solution corresponds to the analytic solution of the Schrödinger equation (E = $-0.5\ E_h$).

Basis	Basis Functions	Energy (E_h)	Error (E_h)
cc-pVDZ	5	-0.499278403	0.000721597
cc-pVTZ	14	-0.499809811	0.000190189
cc-pVQZ	30	-0.499945569	0.000054431
cc-pV5Z	55	-0.499994535	0.000005465
cc-pV6Z	91	-0.499999245	0.000000755
Exact	∞	-0.500000000	0.000000000

6.6   The linear combination of atomic orbitals approximation

The Gaussian basis functions introduced in the previous section are the building blocks used to construct approximate orbitals (atomic and molecular). A sensible choice for approximating the shape of molecular orbitals is to use a basis that is as close as possible to the atomic orbitals of isolated atoms. The intuition here is that a lot of the details of the structure of atoms is preserved in molecules. So it makes sense to use atomic orbitals as building blocks for approximating molecular orbitals.

This idea is called the **linear combination of atomic orbitals** (LCAO). You have already encountered LCAO when you applied molecular orbital theory to simple molecules like H_2, where the bonding orbital is the in-phase combination (superposition) of two 1s atomic orbitals, or when you were introduced to the concept of hybrid orbitals like the sp^3 hybrids of carbon in methane. This is not the only way quantum chemistry computations are done,[6] but it is by far the most common approach used in most quantum chemistry programs.

Mathematically, the LCAO approximation expresses a generic orbital ϕ_i as a linear combination of Gaussian basis functions χ_μ, so that

$$\phi_i(\mathbf{r}) \approx \sum_\mu C_\mu^{(i)} \chi_\mu(\mathbf{r}) \tag{6.22}$$

The quantity $C_\mu^{(i)}$ is a vector of coefficients for orbital ϕ_i. **When you run a Hartree–Fock computation, this is the quantity that is optimized by the SCF procedure.**

For very simple molecules, the coefficients C_μ^i are simply determined by symmetry. As an example, we will look at the orbitals of the H_2 molecule using the STO-6G basis. For convenience, we will label the two hydrogen atoms 1 and 2. Each hydrogen atom carries one basis function, which we label as χ_1 and χ_2. These functions are shown in Fig. 6.7.

From these basis functions we can form two atomic orbitals. The first one is a in-phase combination (g, from the German word *gerade*) of χ_1 and χ_2, times a normalization factor N_g

$$\phi_g = N_g(\chi_1 + \chi_2), \text{ which corresponds to } C_1^{(1)} = N_g, C_2^{(1)} = N_g \qquad (6.23)$$

The second orbital is the out-of-phase combination (u, *ungerade*)

$$\phi_u = N_u(\chi_1 - \chi_2), \text{ which corresponds to } C_1^{(2)} = N_u, C_2^{(2)} = -N_u \qquad (6.24)$$

These are the orbitals that you will get if you do a Hartree–Fock computation on H_2. The orbital ϕ_g is a **bonding orbital** while ϕ_u is an **antibonding orbital**. Since ϕ_g has no nodes, its energy is lower than that of ϕ_u.

6.7   A survey of common basis sets

The basis set exchange website counts a total of 589 basis sets! So how do you go about choosing the right basis set for your computation? In choosing the right basis set we always have to balance two competing factors: the accuracy of the result and the cost of a computation. When one is willing to tolerate a given average error, it is possible to choose a basis using previous results in the literature and use it in a study. However, if your goal is to get the best possible result, the it might be necessary to use a range of basis sets. By repeating computations using basis sets of increasing size, we can gauge the so-called **basis set incompleteness error**, the error that arises from approximating the orbitals using a finite basis set.

For computations on large molecules with Hartree–Fock or DFT it is customary to employ a fixed basis set. Basis sets like Pople's 6-31G, 6-311G, 6-31G*,[7] 6-31G**, etc. are ofter used in DFT computations on large molecules. However, I would discourage you from using this family of basis sets because it is now obsolete and some serious problems have been discovered (e.g., when used with the MP2 method this basis gives a nonplanar structure for benzene!).

A more modern choice is the family of Karlsruhe basis sets (def2-SVP, def2-TZVP, def2-QZVP, def2-TZVPD, ...). For routine DFT computations, a def2-TZVP basis should provide nearly-converged results. The "TZV" and "QZV" stand for triple zeta and quadruple zeta valence (for the Greek letter zeta, ζ, often used to indicate the exponent of a Gaussian instead of the letter α used in these notes), that is three or four sets of valence orbitals. "P" and "D" indicate polarization and **diffuse functions**, respectively. Diffuse functions are extra Gaussian basis functions with a small exponent, which are added to a basis to give it more flexibility in the description of
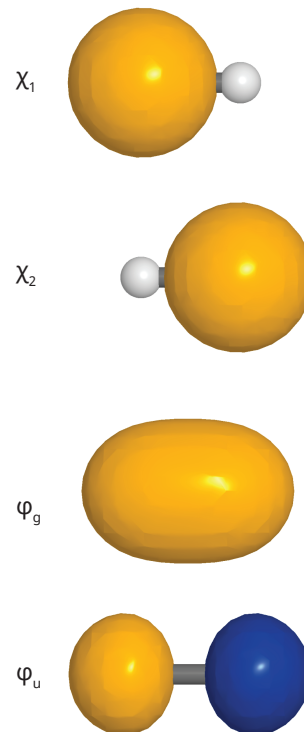

Fig. 6.7: Hydrogen molecule. Plot of the STO-6G basis functions (χ_1 and χ_2) and bonding (gerade, ϕ_g,) and antibonding (ungerade, ϕ_u) molecular orbitals.

[7]The 6-31G* basis set is similar to the def2-SVP basis. It is defined for the atoms H through Zn. It uses two sets of valence orbitals plus polarization functions (d-type polarization functions on atoms Li–Ca and f-type polarization functions on atoms Sc–Zn.

orbitals that are diffuse. These functions increase the accuracy of computations on anions, long-range interactions, and excited states.
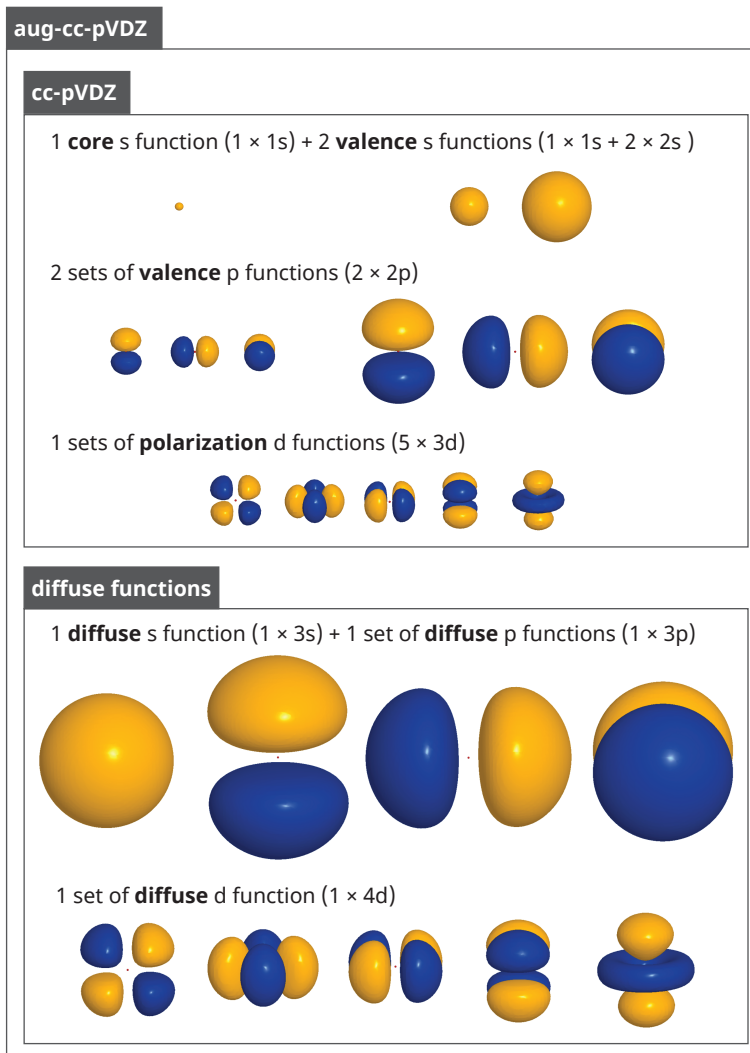

Fig. 6.8: Oxygen atom. Plot of the cc-pVDZ and aug-cc-pVDZ contracted Gaussian-type basis functions. The aug-cc-pVDZ basis is built from the cc-pVDZ by adding an extra set of diffuse functions.

For wave function methods like MP2, CCSD, CCSD(T), etc., Dunning's correlation-consistent basis sets (cc-pVXZ, with X = D, T, Q, 5, 6) are recommended and represent the state-of-the-art. These are families of basis sets that systematically converge to the full basis set limit. Figure 6.8 shows the basis functions included in the cc-pVDZ basis for oxygen. This basis includes a double set of valence orbitals (2s and 2p) and adds one set of d polarization functions. The cc-pVDZ basis is likely to provide reasonable, but potentially inaccurate results. A cc-pVTZ basis is generally recommended

for routine applications. For highly-accurate benchmark results it is necessary to use the cc-pVQZ or cc-pV5Z basis sets. There are many variants of the Dunning's correlation-consistent basis sets.

- The basis sets aug-cc-pV*X*Z include **diffuse functions** and are useful to describe Rydberg states, long range interactions, and anions. Figure 6.8 shows the extra diffuse functions included in the aug-cc-pVDZ basis.

- Dunning's basis sets (cc-pV*X*Z) are meant to correlate only valence electrons, so core electrons should be frozen in computations that use them. To account for core electron correlation it is necessary to use a **core-valence basis set**, usually indicated with the symbol cc-pCV*X*Z, or the newer weighted core-valence sets (cc-pwCV*X*Z).

- Other variants of Dunning's basis sets include **basis extended for period 3 atoms** [cc-pV(*X*+d)Z], **pseudopotential basis sets** for heavy elements [cc-pV*X*Z-PP], and **basis sets for relativistic computations** [cc-pV*X*Z-DK].

This is by no means a comprehensive overview of basis sets. A good practice is to use the literature as a starting point for getting a feeling for what basis sets are used by other researchers. Ideally, before working on a project you should benchmark the convergence of the energy, the molecular geometry, or other properties that you are interested in computing using different basis sets.

Note, that **it is common practice in Quantum Chemistry to cite the original reference to the basis sets employed in a computation**. Crafting a basis set for a large number of elements in the periodic table takes a lot of work and the people that invested their time doing so deserve credit for it.

6.8 The complete basis set (CBS) limit and the basis set error

To analyze the error introduced by using a finite basis we define the **complete basis set (CBS) limit** as the limit of the energy (or any other property) for the basis size that goes to infinity.

$$E_{CBS} = \lim_{K \to \infty} E(\text{basis of size } K) \tag{6.25}$$

The underlying assumption is that with a **complete basis** we can approximate any well-behaved function. It is convenient to introduce the idea of the CBS limit and not just call this energy exact, because even if we use a complete basis, we are still introducing other sources of error that will make our results deviate from experiment.

The error due to the finite basis set is called the **basis set incompleteness error** (BSIE). For any method, the BSIE can be formally defined as the difference between the energy computed with a finite basis $E_{\text{finite basis}}$ and the CBS limit (E_{CBS})

$$\Delta E_{BSIE} = E_{\text{finite basis}} - E_{CBS} \tag{6.26}$$

We will now take a look at the convergence of the Hartree–Fock energy as a function of the basis set for the H_2^+ molecule (see Table 6.2). This systems has only one electron and, therefore, a Hartree–Fock computation is only affected by the error

that arises from using a finite basis set. In this case the energy converges rapidly, like in the case of the hydrogen atom. However, the rate at which the energy converges is slower. For example, the change in energy when going from the cc-pVDZ to the cc-pVTZ basis is only 0.53 mE_h for H, but this number is 1.98 mE_h for H_2^+. Comparing the energy to the best result for H_2^+ (cc-pV6Z basis), we conclude that cc-pVTZ is already sufficiently accurate to make quantitative predictions, and the cc-pVQZ energy is only 0.1 mE_h (ca. 0.15 kcal/mol) off from the best result.

Basis	Basis Functions	Energy (E_h)
cc-pVDZ	10	-0.600264667
cc-pVTZ	28	-0.602244426
cc-pVQZ	60	-0.602520583
cc-pV5Z	110	-0.602619758
cc-pV6Z	182	-0.602632085

Table 6.2: Convergence of the energy of the H_2^+ molecule at a bond distance $r_{HH} = 2$ bohr as a function of the computational basis.