

# 1

## Introduction and theoretical background

### 1.1 The Schrödinger equation and models of chemistry

#### *The Schrödinger equation and its elements*

As early as 1929, the noted physicist P. A. M. Dirac wrote<sup>1</sup>

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.

A similar view was echoed in a 1944 textbook of quantum chemistry:<sup>2</sup>

In so far as quantum mechanics is correct, chemical questions are problems in applied mathematics.

The implication is that chemical phenomena are determined by the laws of quantum mechanics, as expressed in the fundamental time-independent Schrödinger equation

$$\hat{H}\psi = E\psi \quad (1.1)$$

Exact answers to practically *all* chemical questions are, in principle, obtainable from solutions of this equation.<sup>3</sup> Thus, Eq. (1.1) is the ultimate oracle of chemical knowledge.

Equation (1.1) contains three mathematical entities: (i) the *Hamiltonian operator*  $\hat{H}$ , determined by the choice of chemical system; (ii) the *wavefunction*  $\psi$ , describing the allowed spatial distribution of electrons and nuclei of the system; and (iii) the *energy level*  $E$  associated with  $\psi$ . The Hamiltonian  $\hat{H}$  contains terms representing kinetic- and potential-energy contributions, depending only on fixed properties (e.g., mass, charge) of the electrons and nuclei that compose the chosen system of interest. Hence,  $\hat{H}$  is the “known” and  $\psi$  and  $E$  are the “unknowns” of Eq. (1.1). Mathematically,  $\hat{H}$  is an operator that modifies the wavefunction  $\psi(\vec{r}, \vec{R})$  appearing on its right, where we write  $\vec{r} = (\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$  to denote the collective coordinates of  $N$  electrons and  $\vec{R} = (\vec{R}_1, \vec{R}_2, \dots, \vec{R}_\nu)$  those of  $\nu$  nuclei. Only for



Erwin Schrödinger

exceptional choices of  $\Psi$  (“eigenstates”) does this operation give a scalar multiple of the original wavefunction, with characteristic multiplier (“eigenvalue”)  $E$ . The possible solutions ( $\Psi_n, E_n$ ) of Eq. (1.1) for a specific  $\hat{H}$  are commonly labeled with a quantum-number index  $n$ , ordered according to the energy of the ground state and successive excited states of the system. However, in the present work we shall generally restrict attention to the ground solution, so this index can be temporarily omitted.

#### *Perturbation theory of “model chemistry” systems*

In practice, the chemist seldom requires numerically exact answers to chemical questions. Answers that are sufficiently accurate *in the context of the chemical investigation* will therefore be considered satisfactory for practical applications. More specifically, this means that energy differences  $\Delta E$  should be reliable to within a few kcal mol<sup>-1</sup> (i.e., a small percentage of a chemical bond energy), but the necessary accuracy may be higher or lower according to context. To achieve this goal, we introduce an approximate *model* Hamiltonian  $\hat{H}^{(0)}$  that is somehow simplified (for example, by neglecting some of the potential-energy terms in  $\hat{H}$ ), but is expected to retain the most important features of the true Hamiltonian for describing chemical phenomena. The Schrödinger-type equation associated with the model  $\hat{H}^{(0)}$

$$\hat{H}^{(0)}\psi^{(0)} = E^{(0)}\psi^{(0)} \quad (1.2)$$

may then be said to describe a “model chemistry” (in the terminology introduced by J. A. Pople), just as Eq. (1.1) describes<sup>4</sup> the true chemistry of  $\hat{H}$ . The adequacy or inadequacy of this model chemistry to describe the actual phenomena of chemistry can of course be tested through direct comparisons with experimental results. However, it is also feasible to formulate  $\hat{H}^{(0)}$  in a systematically improvable manner, so

## 1.1 The Schrödinger equation

3

that the model chemistry of Eq. (1.2) can be made to approach the exact solutions of Eq. (1.1) as closely as desired (or as far as computational resources permit). From the rate of convergence of these successive corrections, useful estimates of theoretical accuracy that are independent of experimental data can be given. Thus, a central goal of modern computational quantum chemistry is to formulate and improve systematically model  $\hat{H}^{(0)}$ 's that can predict chemical phenomena at a useful level of accuracy.

For our purposes, the most general way to perform systematic correction of a specified model  $\hat{H}^{(0)}$  is by means of *perturbation theory*, as first developed for such problems by Schrödinger himself.<sup>5</sup> The difference between the true  $\hat{H}$  and the model  $\hat{H}^{(0)}$  is defined as the perturbation operator  $\hat{H}^{(\text{pert})}$ ,

$$\hat{H}^{(\text{pert})} = \hat{H} - \hat{H}^{(0)} \quad (1.3)$$

The perturbed Schrödinger equation (1.1) is rewritten in terms of the model  $\hat{H}^{(0)}$  and  $\hat{H}^{(\text{pert})}$  as

$$(\hat{H}^{(0)} + \hat{H}^{(\text{pert})})\psi = E\psi \quad (1.4)$$

By systematic perturbation-theoretic procedures,<sup>6</sup> the exact  $E$  can be obtained in terms of successive orders of correction to  $E^{(0)}$ ,

$$E = E^{(0)} + E^{(1)} + E^{(2)} + \dots \quad (1.5a)$$

where  $E^{(1)}$  is the first-order correction,  $E^{(2)}$  the second-order correction, and so forth;  $\psi$  is similarly expanded as

$$\psi = \psi^{(0)} + \psi^{(1)} + \psi^{(2)} + \dots \quad (1.5b)$$

For example, the first-order energy correction in Eq. (1.5a) is given by

$$E^{(1)} = \langle \psi^{(0)} | \hat{H}^{(\text{pert})} | \psi^{(0)} \rangle \quad (1.5c)$$

where  $\langle \rangle$  is the Dirac “bra-ket” symbol<sup>7</sup>

$$\langle \psi^{(0)} | \hat{H}^{(\text{pert})} | \psi^{(0)} \rangle = \int \psi^{(0)*} \hat{H}^{(\text{pert})} \psi^{(0)} d\tau$$

representing integration over all coordinates, with symbolic differential  $d\tau$ . (Because Eq. (1.2) remains true when  $\psi^{(0)}$  is multiplied by any number, we assume without loss of generality that  $\psi^{(0)}$  is normalized,  $\langle \psi^{(0)} | \psi^{(0)} \rangle = 1$ .) The second-order correction  $E^{(2)}$  can be evaluated from the variational inequality<sup>8</sup>

$$E^{(2)} \leq - \frac{\langle \psi^{(0)} | \hat{H}^{(\text{pert})} | \tilde{\psi}^{(1)} \rangle^2}{\langle \tilde{\psi}^{(1)} | \hat{H}^{(0)} - E^{(0)} | \tilde{\psi}^{(1)} \rangle} \quad (1.5d)$$

where  $\tilde{\Psi}^{(1)}$  is any variational trial function orthogonal to  $\Psi^{(0)}$  ( $\langle \Psi^{(0)} | \tilde{\Psi}^{(1)} \rangle = 0$ ) and the inequality becomes *exact* when  $\tilde{\Psi}^{(1)} = \Psi^{(1)}$ . Note that (1.5d) suggests a numerical method for determining both  $\Psi^{(1)}$  (as the  $\tilde{\Psi}^{(1)}$  that makes the right-hand side as negative as possible) and  $E^{(2)}$  (as the extremal possible value of the right-hand side). The leading corrections  $E^{(1)}$  and  $E^{(2)}$  will suffice for the applications of this book.

### Example 1.1

*Exercise:* Use the perturbation equations (1.5) to estimate the lowest orbital energy of  $\alpha$  spin for a Li atom in a basis of orthogonalized 1s and 2s orbitals, for which the matrix elements of the effective one-electron Hamiltonian operator are<sup>9</sup>  $\langle 1s | \hat{H} | 1s \rangle = -2.3200$ ,  $\langle 1s | \hat{H} | 2s \rangle = -0.3240$ , and  $\langle 2s | \hat{H} | 2s \rangle = -0.2291$ .

*Solution:* The desired orbital energy is an eigenvalue of a  $2 \times 2$  matrix, which can be identified as the “ $\hat{H}$ ” for the application of Eqs. (1.1)–(1.5):

$$\hat{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{12} & H_{22} \end{pmatrix} = \begin{pmatrix} -2.3200 & -0.3240 \\ -0.3240 & -0.2291 \end{pmatrix}$$

To apply the perturbation-theory formalism we can first separate  $\hat{H}$  into diagonal (unperturbed) and off-diagonal (perturbation) matrices,

$$\hat{H}^{(0)} = \begin{pmatrix} H_{11} & 0 \\ 0 & H_{22} \end{pmatrix} = \begin{pmatrix} -2.3200 & 0 \\ 0 & -0.2291 \end{pmatrix}$$

$$\hat{H}^{(\text{pert})} = \begin{pmatrix} 0 & H_{12} \\ H_{12} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -0.3240 \\ -0.3240 & 0 \end{pmatrix}$$

The solutions of the eigenvalue equation for  $\hat{H}^{(0)}$  are evidently

$$E^{(0)} = H_{11}, \quad \Psi^{(0)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and the first-order correction is

$$E^{(1)} = \langle \Psi^{(0)} | \hat{H}^{(\text{pert})} | \Psi^{(0)} \rangle = (1 \ 0) \begin{pmatrix} 0 & H_{12} \\ H_{12} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$$

For the second-order correction, we can recognize that the only possible normalized trial function  $\tilde{\Psi}^{(1)}$  orthogonal to  $\Psi^{(0)}$  in this  $2 \times 2$  case is

$$\tilde{\Psi}^{(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

so that

$$\begin{aligned}\langle \psi^{(0)} | \hat{H}^{(\text{pert})} | \hat{\psi}^{(1)} \rangle &= (1 \quad 0) \begin{pmatrix} 0 & H_{12} \\ H_{12} & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = H_{12} \\ \langle \hat{\psi}^{(1)} | \hat{H}^{(0)} - E^{(0)} | \tilde{\psi}^{(1)} \rangle &= (0 \quad 1) \begin{pmatrix} 0 & 0 \\ 0 & H_{22} - H_{11} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = H_{22} - H_{11} \\ E^{(2)} &= -\frac{H_{12}^2}{H_{22} - H_{11}}\end{aligned}$$

The lowest eigenvalue (1s orbital energy) is therefore estimated as

$$E = E^{(0)} + E^{(1)} + E^{(2)} = H_{11} - \frac{H_{12}^2}{H_{22} - H_{11}} = -2.3702 \text{ a.u.}$$

The corresponding estimate for the second eigenvalue (2s orbital energy) is  $-0.1789$ . These results are in good agreement with the actual HF/STO-3G (“Hartree–Fock method with a variational basis set of three-term Gaussians for each Slater-type orbital”<sup>10</sup>) eigenvalues:  $\epsilon_{1s} = -2.3692$  and  $\epsilon_{2s} = -0.1801$ .

Among various model  $\hat{H}^{(0)}$ 's that could be considered, the best such model is evidently that for which the perturbative corrections are most rapidly convergent, i.e., for which  $\hat{H}^{(\text{pert})}$  is in some sense smallest and the model  $E^{(0)}$  and  $\psi^{(0)}$  are closest to the true  $E$  and  $\psi$ . Perturbation theory can therefore be used to *guide* selection of the best possible  $\hat{H}^{(0)}$  within a class of competing models, as well as to evaluate systematic corrections to this model.

### Conceptual constructs in model systems

Perturbation theory also provides the natural mathematical framework for developing chemical concepts and “explanations.” Because the model  $\hat{H}^{(0)}$  corresponds to a simpler physical system that is presumably well understood, we can determine how the properties of the more complex system  $\hat{H}$  evolve term by term from the perturbative corrections in Eq. (1.5a), and thereby elucidate how these properties originate from the terms contained in  $\hat{H}^{(\text{pert})}$ . For example, Eq. (1.5c) shows that the first-order correction  $E^{(1)}$  is merely the average (quantum-mechanical expectation value) of the perturbation  $\hat{H}^{(\text{pert})}$  in the unperturbed eigenstate  $\psi^{(0)}$ , a highly intuitive result. Most physical explanations in quantum mechanics can be traced back to this kind of perturbative reasoning, wherein the connection is drawn from what is “well understood” to the specific phenomenon of interest.

Perturbative reasoning can be used to justify conceptual models of chemistry that are far from evident in Eq. (1.1) itself. An important example is the concept of *molecular structure* – the notion that nuclei assume a definite equilibrium configuration  $\vec{R}_0$ , which determines the spatial shape and symmetry of the molecule. At first glance, this concept appears to have no intrinsic meaning in Eq. (1.1),

because the true molecular Hamiltonian  $\hat{H}$  has *identical* dependence on all identical nuclei,<sup>11</sup> and thus could not assign a distinct structural role to a particular nucleus. Indeed,  $\hat{H}$  is totally symmetric with respect to all rotation axes, mirror planes, or inversion centers that might be chosen to classify the “structure” of the molecule, and thus could never distinguish between, e.g., *dextro* and *levo* optical isomers, in-plane and out-of-plane H’s of hydrocarbons, carbonyl-type and hydroxyl-type O’s of carboxylic acids, and so forth. This means, for example, that no true eigenstate of Eq. (1.1) can correspond to a chiral molecule of definite handedness,<sup>12</sup> even though the experimental existence of distinct enantiomeric species is well established.

The resolution of this paradox lies in the *Born–Oppenheimer* approximation,<sup>13</sup> which is based on the fact that nuclei are thousands of times more massive than electrons. The nuclear motions are therefore so sluggish that electrons can be considered to rearrange virtually instantaneously around each static nuclear configuration  $\vec{R}_0$ . In this limit,  $\hat{H}$  can be replaced by a model  $\hat{H}^{(0)} = \hat{H}(\vec{r}; \vec{R}_0)$  that depends only parametrically on nuclear positions, which are considered fixed at  $\vec{R}_0$ . Solving Eq. (1.2) for the electronic motions alone, while holding  $\vec{R}_0$  fixed, then leads to solutions in which the energy  $E^{(0)} = E(\vec{R}_0)$  varies with nuclear configuration, the “potential-energy surface” for subsequent treatment of nuclear motion. The high accuracy of the Born–Oppenheimer model, i.e., the generally negligible values of its higher-order perturbative corrections, amply justifies the chemist’s faith in the existence of well-defined molecular structures. More generally, such considerations validate the direct formulation of Eq. (1.1) as an *electronic* Schrödinger equation in the Born–Oppenheimer framework, as we do throughout this book.

It is also routine to assume the *non-relativistic* approximation<sup>14</sup> in writing Eq. (1.1). This is based on the fact that molecular electronic velocities are generally far less than the speed of light, and the magnetic forces arising from electronic motions can therefore be neglected compared with the dominant electrical forces between charged particles. In this limit, the model Hamiltonian contains only potential-energy terms corresponding to Coulomb’s law of classical electrostatics. However, in reducing  $\hat{H}$  from relativistic (Dirac-like) to non-relativistic form, one must recognize the two possible orientations of the intrinsic “spin” angular momentum of each electron (a relativistic effect) and include suitable spin labels in the wavefunction  $\Psi$ . Thus, we should generalize the spatial coordinate  $\vec{r}$  to include the orientation of each electron in “spin space.” Furthermore, we must insure that the total electronic  $\Psi(\vec{r})$  is antisymmetric with respect to exchange of space–spin coordinates of any two electrons  $i$  and  $j$ , as required by the *Pauli exclusion principle*<sup>15</sup> (Section 1.7),

$$\Psi(\vec{r}_i, \vec{r}_j) = -\Psi(\vec{r}_j, \vec{r}_i) \quad (1.6)$$

## 1.1 The Schrödinger equation

7

However, in other respects spin plays no direct role in construction of the non-relativistic  $\hat{H}$ . If higher accuracy is desired, perturbative expressions such as (1.5) can be used to evaluate spin-dependent corrections to the non-relativistic model.

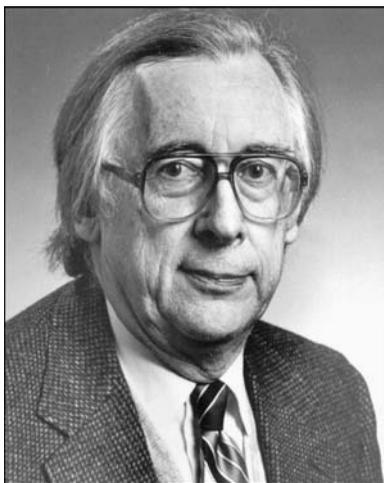
*Variational models*

Still another useful approximation is introduced by reformulating Eq. (1.1) as a *variational principle*,<sup>16</sup>

$$E \leq E^{(0)} = \frac{\langle \psi^{(0)} | \hat{H} | \psi^{(0)} \rangle}{\langle \psi^{(0)} | \psi^{(0)} \rangle} \quad (1.7)$$

The inequality (1.7) is true for any possible variational trial function  $\psi^{(0)}$ , subject only to the usual antisymmetry and boundary conditions for square-integrable functions of proper symmetry, and the *best* such  $\psi^{(0)}$  is that leading to the lowest possible value of  $E^{(0)}$ , closest to the true  $E$ . However, it can be shown<sup>17</sup> that any such trial function  $\psi^{(0)}$  and variational energy  $E^{(0)}$  are also solutions of a Schrödinger-type equation (1.2) for a suitably defined model Hamiltonian  $\hat{H}^{(0)}$ . Thus, *any* variational approximation (1.7) can be formulated in terms of a model  $\hat{H}^{(0)}$ , and the errors of this model can be systematically corrected with perturbative expressions such as Eqs. (1.5).<sup>18</sup> This view of variational calculations, although somewhat unconventional, allows us to treat both variational and perturbative approximation methods in a common “model chemistry” language, along the lines enunciated by Pople.<sup>19</sup>

Summing up, we may say that approximation methods in quantum chemistry generally involve (either explicitly or implicitly) a model  $\hat{H}^{(0)}$  and associated model chemistry that more or less mimics the true behavior of Eq. (1.1). Such models might be closely patterned after the well-known conceptual models of empirical



John A. Pople

chemistry, such as valence bond or ionic models, or chosen for purely mathematical reasons. Perturbation theory always allows us to choose the best among a family of such proposed models and to calculate systematically the corrections that bring the model into improved agreement with the exact Schrödinger equation (1.1).

The primary goal of a theory of valency and bonding is to find the model  $\hat{H}^{(0)}$  that most simply describes the broad panorama of chemical bonding phenomena, or, as Gibbs said,<sup>20</sup> “to find the point of view from which the subject appears in its greatest simplicity.” In the past, conceptual models were often cobbled together from diverse empirical patterns, guided only weakly by theory. This resulted in a patchwork of specialized “effects,” with incommensurate seams, indeterminate limits of applicability, and little overall theoretical coherence. However, remarkable advances in computational technology<sup>21</sup> now make it possible to construct improved conceptual models *directly* from accurate *ab initio* (“first-principles”) wavefunctions.

The fundamental starting point for a rational electronic theory of valency and bonding is the Lewis-structure representation of the shared and unshared electrons in each atomic valence configuration, as formulated by G. N. Lewis. In the present work, we shall focus on a *natural Lewis-structure* model, based on associating the electron pairs of the familiar Lewis-structure diagram with a set of optimal, intrinsic, “natural” bond orbitals (NBOs),<sup>22</sup> as outlined in Section 1.5. This viewpoint is deeply tied to traditional chemical bonding concepts – including hybridization, polarization, and bond transferability – and takes advantage of a model  $\hat{H}^{(0)}$  that describes localized electron pairs and their interactions in rapidly convergent fashion. By employing a non-empirical theoretical methodology to construct quantitative bonding concepts, we can better achieve the goal of unifying and harmonizing conceptual models of valency and bonding with the deepest principles of chemistry, as expressed by Eq. (1.1).

## 1.2 Hydrogen-atom orbitals

### *Orbitals and electron-density distributions*

In his first communication<sup>23</sup> on the new wave mechanics, Schrödinger presented and solved his famous Eq. (1.1) for the one-electron hydrogen atom. To this day the H atom is the only atomic or molecular species for which exact solutions of Schrödinger’s equation are known. Hence, these hydrogenic solutions strongly guide the search for accurate solutions of many-electron systems.

The essence of Schrödinger’s treatment was to replace the classical orbit of Bohr’s semi-classical (particle) model of the H-atom by a corresponding wave-like *orbital* (single-electron wavefunction)  $\psi$ . Instead of specifying the electron’s

## 1.2 Hydrogen-atom orbitals

9

position at a particular point  $\vec{r}$  of its orbit (as we should expect classically), the orbital  $\Psi(\vec{r})$  determines only the electron *density*  $\rho(\vec{r})$  at each spatial point  $\vec{r}$ ,

$$\rho(\vec{r}) = |\Psi(\vec{r})|^2 \quad (1.8)$$

The density  $\rho(\vec{r})$  might also be described as the fractional *probability* of finding the (entire) electron at point  $\vec{r}$ . However, chemical experiments generally do not probe the system in this manner, so it is preferable to picture  $\rho(\vec{r})$  as a *continuous* distribution of *fractional* electric charge. This change from a “countable” to a “continuous” picture of electron distribution is one of the most paradoxical (but necessary) conceptual steps to take in visualizing chemical phenomena in orbital terms. Bohr’s “orbits” and the associated “particulate” picture of the electron can serve as a temporary conceptual crutch, but they are ultimately impediments to proper wave-mechanical visualization of chemical phenomena.

Equally paradoxical is the fact that  $\rho(\vec{r})$  depends only on the absolute square of the orbital, and is everywhere sensibly non-negative, whereas  $\Psi(\vec{r})$  oscillates in “wavy” fashion between positive and negative values.<sup>24</sup> The phase patterns corresponding to such sign changes are of utmost importance in chemistry. Solutions of Schrödinger’s equation are generally governed by the *superposition principle*, such that two interacting orbitals may interfere with one another in wave-like constructive (in-phase) or destructive (out-of-phase) patterns<sup>25</sup> that strongly alter the form of  $\rho(\vec{r})$ . Visualizing and understanding the subtle chemical consequences of orbital phase patterns and superposition is a central goal of this book.

*Quantum numbers and shapes of atomic orbitals*

Let us denote the one-electron hydrogenic Hamiltonian operator by  $\hat{h}$ , to distinguish it from the many-electron  $\hat{H}$  used elsewhere in this book. This operator contains terms to represent the electronic kinetic energy ( $\hat{t}_e$ ) and potential energy of attraction to the nucleus ( $\hat{v}_{ne}$ ),

$$\hat{h} = \hat{t}_e + \hat{v}_{ne} \quad (1.9)$$

The associated Schrödinger equation for the H atom can then be written as

$$\hat{h}\Psi_{nlm} = \epsilon_n\Psi_{nlm} \quad (1.10)$$

Each orbital eigenstate  $\Psi_{nlm} = \Psi_{nlm}(\vec{r})$  is labeled by three quantum numbers:

$$\text{principal: } n = 1, 2, 3, \dots, \infty \quad (1.11a)$$

$$\text{azimuthal: } l = 0, 1, 2, \dots, n - 1 \quad (1.11b)$$

$$\text{magnetic: } m = 0, \pm 1, \pm 2, \dots, \pm l \quad (1.11c)$$

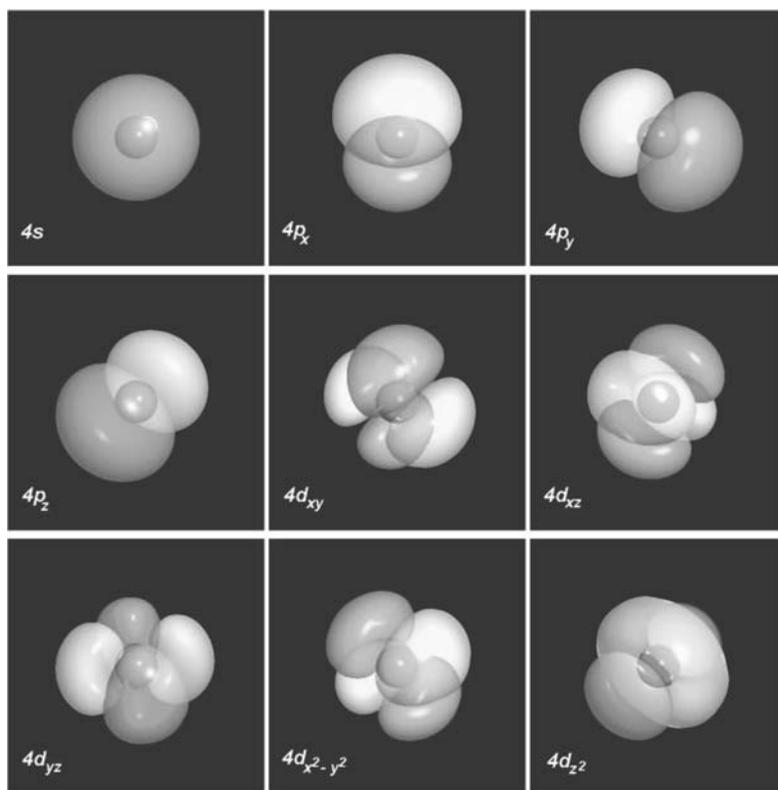


Figure 1.1 Surface plots of representative s, p, and d atomic orbitals (from the Kr valence shell).

The energy eigenvalue  $\epsilon_n$  depends only on the principal quantum number  $n$ ; its value is given (in atomic units; see Appendix C) by

$$\epsilon_n = -\frac{Z^2}{2n^2} \text{ a.u.} \quad (1.12)$$

for atomic number  $Z$  ( $Z = 1$  for H).

The three quantum numbers may be said to control the size ( $n$ ), shape ( $l$ ), and orientation ( $m$ ) of the orbital  $\Psi_{nlm}$ . Most important for orbital visualization are the angular shapes labeled by the azimuthal quantum number  $l$ : s-type (spherical,  $l = 0$ ), p-type (“dumbbell,”  $l = 1$ ), d-type (“cloverleaf,”  $l = 2$ ), and so forth. The shapes and orientations of basic s-type, p-type, and d-type hydrogenic orbitals are conventionally visualized as shown in Figs. 1.1 and 1.2. Figure 1.1 depicts a surface of each orbital, corresponding to a chosen electron density near the outer fringes of the orbital. However, a wave-like object intrinsically lacks any definite boundary, and surface plots obviously cannot depict the interesting variations of orbital amplitude *under* the surface. Such variations are better represented by *radial* or *contour*