Theoretical Chemistry Accounts (2019) 138:9
https://doi.org/10.1007/s00214-018-2386-x

REGULAR ARTICLE

Atomic orbitals revisited: generalized hydrogen-like basis sets for 2nd-row elements

Ilya V. Popov[1,2] · **Andrei L. Tchougréeff**[1,2,3]

Received: 16 June 2018 / Accepted: 15 November 2018 / Published online: 4 December 2018

Abstract

In the present work, we revisit the problem of atomic orbitals from the positions mostly dictated by semiempirical approaches in quantum chemistry. To construct basis set, having proper nodal structure and simple functional form of orbitals and representing atomic properties with reasonable accuracy, authors propose an *Ansatz* based on gradual improvement of hydrogen atomic orbitals. According to it, several basis sets with different numbers of variable parameters are considered and forms of orbitals are obtained for the 2nd-row elements either by minimization of their ground state energy (direct problem) or by extracting from atomic spectra (inverse problem). It is shown that so-derived three- and four-parametric basis sets provide accurate description of atomic properties, being, however, substantially provident for computational requirements and, what is more important, simple to handle in analytic models of quantum chemistry. Since the discussed *Ansatz* allows a generalization for heavier atoms, our results may be considered not only as a solution for light elements, but also as a proof of concept with possible further extension to a wider range of elements.

Keywords Atomic orbitals · Atoms · Analytic models · Semiempirical methods

1 Introduction

Quantum chemical description of molecules involves one-electron states, expanded against finite sets of basis functions. Quality and efficiency of electronic structure calculations and f eatures of their computational implementation substantially rely on the properties of the underlying basis set. This fact is emphasized in almost each handbook on quantum chemistry (see, e.g., [1]) and reflected in the huge number of basis sets available in the literature for description of various objects and properties [2, 3]. Two major classes of basis functions (coming from two main types of objects in theoretical chemistry) are local (atomic) basis functions and plane waves. Although using the latter is extremely efficient numerically, the incurred loss of local chemical information generated demand for *a posteriori* analysis tools projecting the results obtained in the plane wave basis onto local basis sets as successfully implemented, e.g., in the LOBSTER package [4–6]. In this work, we focus only on local atomic orbitals, and thus our further discussion will be restricted to them.

Atomic functions appear in either numerical (tabular) or analytic form [1]. Numerical atomic orbitals (for example, Ref. [7]) come from accurate *ab initio* calculations on many-electron atoms, but their actual application is restricted to very simple and highly symmetric (usually linear) molecules. By contrast, analytic atomic orbitals (AOs) are the main tool of quantum chemistry. Analytic AOs are in their turn linear combinations of primitives, the latter being either Slater-type (STO) [8] or Gaussian-type functions [9]. The numbers of primitives and variable parameters are determined by two target characteristics of a basis set: flexibility (growing with number of primitives and parameters) and efficiency in computational and analytic applications

---

Published as part of the special collection of articles In Memoriam of János Ángyán.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00214-018-2386-x) contains supplementary material, which is available to authorized users.

✉ Andrei L. Tchougréeff
andrei.tchougreeff@ac.rwth-aachen.de

[1] A.N. Frumkin Institute of Physical Chemistry and Electrochemistry of RAS, Moscow, Russia

[2] Independent University of Moscow, Moscow, Russia

[3] Chair of Solid State and Quantum Chemistry, RWTH - Aachen University, Aachen, Germany

(obviously decreasing with the numbers of primitives and parameters). Balance between these two characteristics depends on specific class of problems to be addressed and requirements of the method used.

Two complementary approaches may be pointed out in quantum chemistry: numerical one, motivated by accurate reproduction of different properties with heavy computational modeling, and analytic one, trying to find conceptual *explanation* of chemical trends within semiempirical analytic models (in the spirit of works [10–12]; earlier analytic models of polyenes which profited from the analogy with the translationally invariant models of solids had been studied by Lennard-Jones [13] and Coulson [14] and later complemented by electron–electron interaction terms by Misurkin and Ovchinnikov [15]). Principal differences in these two approaches cause differences in their requirements to the basis sets to be used. Numerical treatment can partially sacrifice physical clarity to make basis orbitals more parametrically flexible and convenient in computational implementation, and analytic approach, in its turn, can partially sacrifice numerical precision to obtain functions of simple form, containing parameters with clear physical meaning. The latter sets the target for basis sets development from the point of analytic quantum chemistry: to employ simple and physically reasonable functional forms of the AOs, yielding adequate estimates of the parameters (ultimately, matrix elements) of the semiempirical Hamiltonians from these orbitals. This problem, however, is not that trivial, since all known simplest semiempirical basis sets (single STOs) usually fail badly in reproducing of one-center core attraction parameters (see related discussion in Section S6[1]), while more involved basis sets present in the literature cannot be used in analytic treatment. Because of that, systematic studies on the basis sets with a simple functional form, suitable for analytic work, are required.

Due to present dominance of the numeric approach in quantum chemistry, there are not that many studies of such kind in the literature, and a few examples of the basis sets potentially usable for that purpose are quite old. For example, yet in 1935 Morse et al. [16] proposed basis set with simple functional form of orbitals for light elements (see below for details), which was later studied by several authors (Duncanson and Coulson [17]; Pujol and Simon [18]). Approximately the same time Gombás and Szondy [19, 20] developed atomic basis sets of STOs with variable parameters, which was actually one of the first prototypes of later (more involved) basis sets using STO primitives (e.g., Ref. [21, 22]). In Ref. [1], one can find further examples of the basis sets with simple forms of AOs used to reproduce exact Hartree–Fock (HF) solution for the carbon atom, which are actually already quite involved for any analytic applications.

Trying to change this situation, we revisit the problem of the AOs having in mind deriving a basis set useful for analytic semiempirical approaches similar to those cited above. This work should be considered as the first step of such revision, where we propose a series of hydrogen-like basis sets and test them on 2nd-row elements. We demonstrate the principal possibility of using proposed hydrogen-like basis sets by comparing calculated atomic properties with both experimental data and the results obtained with the Bunge [22] basis sets.

2 Specification of basis sets and methods of calculations

The hydrogen-like functions, being the well-known solutions for the bound states of one-electron atom (for example Ref. [1]), may serve as a starting point in a search for basis sets suitable for analytic models of quantum chemistry. Obviously, they cannot be applied for the many-electron atom directly, since only one variable parameter (effective core charge Z_{eff}) makes them too rigid to reflect different screening patterns (arising from electrostatic interactions) of electrons located in different orbital subshells well known in the pseudopotential theory (see [23] and references therein). However, hydrogen-like functions have correct radial nodal structure assuring the orthogonality of the radial parts of the AOs with the different principal quantum numbers n belonging to the same azimuthal quantum number l may be prototypes for constructing more advanced basis sets.

The hydrogen-like basis functions can be modified in several directions to acquire more flexibility, but retaining their simple analytic form and satisfying the orthonormalization conditions. The simplest way is to use the orthogonality of the spherical harmonics with the different azimuthal quantum numbers l. Because of it, effective nuclear charges for different l's may be unequal, yielding the orbital exponents, e.g., $\xi_{1s} = Z^{(s)}_{\text{eff}}; \xi_{2s} = Z^{(s)}_{\text{eff}}/2 \neq \xi_{2p} = Z^{(p)}_{\text{eff}}$ for the 2nd-row elements), keeping the form of the hydrogen-like functions and taking $Z^{(s)}_{\text{eff}}$ and $Z^{(p)}_{\text{eff}}$ for variational parameters. It leads to a two-parametric hydrogen-like basis set (H-2 set). For the 2nd-row atoms, such basis has the form:[2]

$$\begin{aligned} R_{1s}(r) &= 2\xi_{1s}^{3/2} \exp\left(-\xi_{1s} r\right) \\ R_{2s}(r) &= 2\xi_{2s}^{3/2} \left\{1 - \xi_{2s} r\right\} \exp\left(-\xi_{2s} r\right) \\ R_{2p}(r) &= \frac{2\xi_{2p}^{5/2}}{\sqrt{3}} r \exp\left(-\xi_{2p} r\right) \end{aligned} \tag{1}$$

---

[1] Sections numbered SX refer to Section X of the Supplementary material.

[2] Generalization of eq. (1) for further periodic rows can be obtained by inserting in $R_{nl}(r)$ the orbital exponents $\xi_{nl} = Z^{(l)}_{\text{eff}}/n$ with $Z^{(l)}_{\text{eff}}$'s as variational parameters.

Although the H-2 basis Eq. (1) is more flexible than purely hydrogen-like [1], there is still an obvious physical inconsistency left. Clearly, the effect of the occupied $1s$-shell upon $2s$-electrons does not reduce to the bare orthogonality condition since the $2s$- and $2p$-shells feel the Coulomb repulsions of $1s$-electrons differently which translates into the deviation of the orbital exponents from the simple $\xi_{nl} = Z_{\mathrm{eff}}^{(l)}/n$ rule. It is impossible to reach with use of the hydrogen-like functions [1], since the non-equivalence of the effective nuclear charges for the $1s$- and $2s$-shells will break their orthogonality. Thus, we modify the functional form of the AOs to restore the orthogonality. It can be done by introducing polynomials $P_{n-l-1}^{2l+1}\left(x, \mu_1^{(l)}, \dots, \mu_{n-l-1}^{(l)}\right)$ for each value of l, which depend on $n-l-1$ parameters $\mu_1^{(l)}, \dots, \mu_{n-l-1}^{(l)}$: the ratios of the orbital exponents with the same l and different n's (the ratios of $\xi_{n_{\min}l}$ to ξ_{nl}). Such polynomials are of the same power as the Laguerre polynomials entering the hydrogen-like AOs [1] with the same indices thus having the required number of nodes (details are given in Section S1). For the 2nd-row elements, such an approach yields the following three-parametric generalized hydrogen-like basis set (H-3):

$$\begin{aligned} R_{1s}(r) &= 2\xi_{1s}^{3/2} \exp\left(-\xi_{1s} r\right) \\ R_{2s}(r) &= 2N_\mu \xi_{2s}^{3/2} \left\{ 1 - \frac{1+\mu}{3}\left(\xi_{2s} r\right) \right\} \exp\left(-\xi_{2s} r\right) \\ R_{2p}(r) &= \frac{2\xi_{2p}^{5/2}}{\sqrt{3}} r \exp\left(-\xi_{2p} r\right) \end{aligned} \tag{2}$$

where $\mu = \xi_{1s}/\xi_{2s}$ and

$$N_\mu = \left(\frac{3}{1 - \mu + \mu^2} \right)^{\frac{1}{2}}.$$

The H-3 basis set turns into the H-2 set, when the ratio of the $1s$- to $2s$-exponents is equal to that in the hydrogen-like setting: $\mu = 2$.

It should be noted that the idea, leading to the H-3 orbitals, is quite similar (but not same) to that used by Gombas and Szondy [19, 20]. They as well started from the simple STO functional form for the AOs and parameterized them. Further, they applied the Schmidt orthogonalization algorithm to restore the orthonormality of the orbitals with different effective core charges which assured the correct radial nodal structure of so-derived AOs. The mixed orbitals coming from the Schmidt procedure cannot be characterized by any specific effective core charge and become very cumbersome for analytic work, when their principal quantum number increases. Thus, to reach the same result we change the functional form of polynomial coefficients. Our approach keeps the general functional form of the hydrogen-like orbitals (a polynomial guaranteeing the required number of radial nodes times an exponent) and retains the physical meaning of the orbital exponents as of measures of the subshell-specific screening by lower-lying electrons.

Keeping the number of the involved STO primitives constant, the basis can be made even more flexible. A four-parametric set (H-4), originally proposed by Morse et al. [16], is given by:

$$R_{2s}(r) = N_{\mu\nu} \left(\xi_{2s}^{(2)}\right)^{3/2} \left[A \exp\left(-\xi_{2s}^{(1)} r\right) - \xi_{2s}^{(2)} r \exp\left(-\xi_{2s}^{(2)} r\right) \right] \tag{3}$$

where $\mu = \xi_{1s}/\xi_{2s}^{(2)}$, $\nu = \xi_{2s}^{(1)}/\xi_{2s}^{(2)}$ and with the single STOs for the $1s$- and $2p$-subshells. Orthonormality of R_{2s} and R_{1s} uniquely determines the constants in Eq. (3):

$$A = 3\frac{(\mu+\nu)^3}{(1+\mu)^4}$$

$$N_{\mu\nu} = \left(\frac{1}{1 - \frac{16A}{(1+\nu)^4} + \frac{A^2}{3\nu^3}} \right)^{\frac{1}{2}}$$

and the $1s$- and $2p$-orbitals have the same form as in the H-3 setting Eq. (2). Thus, the H-4 basis turns into the H-3 one by making the exponents $\xi_{2s}^{(1)} = \xi_{2s}^{(2)}$ ($\nu = 1$) in Eq. (3) equal.

It is quite surprising that to the best of our knowledge there were no studies on the H-3-like basis sets (for heavier atoms bases, constructed in the same manner as H-3, must contain more parameters, so technically they can not be called 'H-3') in the literature, while the H-4 like bases were studied [16–18, 24]. As it is shown below, the H-3 basis set allows us to reproduce the atomic properties with an accuracy comparable to that of the H-4 one. At the same time, the H-3 basis contains parameters of clear physical meaning (effective core charges for each subshell rendering the respective effective screening by the core electrons) and also is easier in analytic treatment (especially, when it comes to generalization to heavier atoms). The explicit presence of the effective core charges in H-3-like basis sets makes them a promising tool, which will hopefully permit to get insight into observable patterns of changing in atomic properties. For instance, it may be possible to provide a more detailed study of the influence of core electrons upon valence state of an atom, believed to be the reason of secondary periodicity, firstly noticed by Biron [25] and later discussed in more advanced theoretical terms by Pyykkö [26].

In each of the proposed settings Eqs. (1)–(3), the values of the orbital exponents may be obtained either in the frame of the *direct* problem: by minimizing the Hartree–Fock (HF) energy of atomic ground state (Φ) or by solving certain *inverse* problems: fitting the parameters (exponents) of the AOs against some external data. We shall investigate both approaches for all “hydrogen-like” basis sets introduced above.

2.1 Direct problem

Taking Φ as a single Slater determinant (with the amendments necessary to assure the correct spin and symmetry in the sense of the *old* MC SCF theory [27]), one comes to the following expression for the energy:

$$E = \langle \Phi | \hat{H} | \Phi \rangle = \sum_{R} \langle \phi_R | \hat{h} | \phi_R \rangle + \frac{1}{2} \sum_{R} \sum_{R \neq S} \left(\langle \phi_R \phi_S | \hat{g} | \phi_R \phi_S \rangle - \langle \phi_R \phi_S | \hat{g} | \phi_S \phi_R \rangle \right) \tag{4}$$

where $\phi_{R,S}$ are one-electron states occupied in the determinant Φ, $\hat{h}$ and $\hat{g}$ are one- and two-electron energy operators:

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i}$$

$$\hat{g}_{ij} = \frac{1}{\left| r_i - r_j \right|}$$

The average energy of the $1s^2 s^m 2p^p$ states (vis. *configuration*) easily expresses through the core attraction parameters (U) and linear combinations (K_1–K_6) of the Coulomb (F^k) and exchange (G^k) Slater–Condon parameters [28] as:

$$E(n, m, p) = nU(1s) + mU(2s) + pU(2p) + \frac{n(n-1)}{2}K_1 + \frac{m(m-1)}{2}K_2 + \frac{p(p-1)}{2}K_3 + nmK_4 + npK_5 + mpK_6 \tag{5}$$

Exact expressions of the combinations K_1–K_6 are given in Section S2. All parameters entering Eq. (5) depend only on the orbital exponents of whatever basis we use (see Section S2 for exact formulae), which allows us to write analytic expressions for the energy through the variation parameters (exponents) of any basis set. An important observation coming from Section S2 is that K_1–K_6 are all homogeneous functions of the 1st degree with respect to the orbital exponents, whereas U’s are combinations of homogeneous functions of the 1st and 2nd degrees. It allows us to use gradient method to solve the direct problem: find the minimal value of the energy and the corresponding values of the orbital exponents.

2.2 Inverse problem

As we mentioned above, an inverse problem can be used to determine the orbital exponents for various proposed types of the basis sets by fitting them against different types of external data. In the present work we consider two inverse problems. The first one is based on fitting the orbital exponents against available experimental data on atomic spectra. The second inverse problem is to fit our approximate radial parts of the AOs against the more involved basis set derived from the HF calculation and constructed from the extended set of primitives.

2.2.1 Experimental spectra fit

The first investigated inverse problem extracts the exponents from the experimental relative energies of atomic configurations. To solve this problem, only the configurations, which do not bring any new unknown orbital exponents, should be considered. For the 2nd-row element, we consider the configurations, having electrons on the $1s$-, $2s$- or $2p$-orbitals only (for example, the $1s^2 2s^2 2p^1 3s^1$ configuration for carbon does not satisfy this condition—it brings a new unknown exponent for the $3s$-subshell). It also needs to be assumed that the orbital exponents (and, consequently, the atomic parameters) are equal for different configurations.

Simple accounting shows that the total number of K_m parameters (six) is smaller than the number of the Slater–Condon parameters required for description of the three involved subshells, which means that these latter *cannot* be all determined by a linear fit of the expressions like Eq. (5) against experimental data on the configuration energies: ill-defined inverse problem in terms of Ref. [29]. Expressing the Hamiltonian parameters involved in the energy expression Eq. (5) through the orbital exponents of either of AOs Eqs. (1)–(3) brings in additional information about the relation between these parameters by this *regularizing* [29] otherwise ill-defined *inverse* problem.

Even the proposed regularization of the inverse problem of determining one- and two-electron parameters with use of the AOs model does not allow to uniquely fix the answer since the reference “zero” energy of the ground state is missing. This lack of data is cured by extending the set of spectral data on account of the ionization potentials (IP’s), calculated as differences: $I_p = E\left(A^+\right) - E(A)$. For the sake of the consistency control, we minimized the energies of the cations $E\left(A^+\right)$ with respect to the orbital exponents (that is, solved the direct problem for the cations as well). The values of $E\left(A^+\right)$ calculated with the exponents optimized for the neutral atom are higher than the minimized $E\left(A^+\right)$ by only 3–5% (for the H-2, H-3 and H-4 basis sets)—imprecision, not affecting the trends in the chemical series. Thus, for the standard semiempirical setting: the calculation (or analytic treatment) performed in a fixed basis, the minimization of the cation energy with respect to the exponents is superfluous.

Two types of the ionization potentials (see Table 1) are considered in the present work: $s^2p^n \rightarrow s^2p^{n-1}$ ($\mathrm{IP}^{(1)}$) and $s^2p^n \rightarrow s^1p^n$ ($\mathrm{IP}^{(2)}$). Experimental values of $\mathrm{IP}^{(1)}$’s may be directly found in the NIST database [30, 31]. Experimental values of $\mathrm{IP}^{(2)}$’s are

Table 1 Terms of 2nd-row ions used in our calculations

Element	Ground state	State of ion IP$^{(1)}$	State of ion IP$^{(2)}$
Li	$1s^2 2s^1\left(^2S\right)$	–	$1s^2\left(^1S\right)$
Be	$2s^2\left(^1S\right)$	–	$2s^1\left(^2S\right)$
B	$2s^2 2p^1\left(^2P\right)$	$2s^2\left(^1S\right)$	$2s2p^1\left(^1P\right)$
C	$2s^2 2p^2\left(^3P\right)$	$2s^2 2p^1\left(^2P\right)$	$2s2p^2\left(^2P\right)$
N	$2s^2 2p^3\left(^4S\right)$	$2s^2 2p^2\left(^3P\right)$	$2s2p^3\left(^3S\right)$
O	$2s^2 2p^4\left(^3P\right)$	$2s^2 2p^3\left(^4S\right)$	$2s2p^4\left(^4P\right)$
F	$2s^2 2p^5\left(^2P\right)$	$2s^2 2p^4\left(^3P\right)$	$2s2p^5\left(^3P\right)$

obtained as a sum of IP$^{(1)}$'s and the excitation energies of the s^2p^{n-1} ions available from the atomic spectra [32].

2.2.2 Fitting against external basis set

Another option for the inverse problem is to seek for the orbital exponents yielding the maximum overlap of our model AOs Eqs. (1)–(3) with those of some external basis set, obtained by highly accurate calculations (*ab initio* basis sets, for example). This problem is easiest to solve and it provides fast results on the orbital exponents, allowing to reproduce the atomic properties with an accuracy pretty much close to that of the reference basis set for each given analytic form of the model AOs. In the present work, we use the Bunge basis set [22] as a reference for such a fitting also having in mind to establish a relation with the procedures of the bonding analysis of solids [4–6], where the latter is used for projecting the results numerically obtained in the plane wave setting.

Parameters of the Bunge basis [22], which we use for comparing, are known only for neutral atoms and their values are not established for cations. Thus we use parameters, obtained for the neutral atoms, for the cations as well. Within such setting, the IP's obtained with the Bunge orbitals are expected to be slightly overstated as compared to the experimental values. It indeed happens, and the IP's so calculated are all higher than the experimental values by about 0.3–1 eV (see Sect. 3.1.3).

3 Results and discussion

3.1 Direct problem

3.1.1 Orbital exponents and shapes of orbitals

The orbital exponents, obtained from the energy minimization (*direct* problem) for H-2, H-3 and H-4 basis sets, are collected in Tables 1, 2 and 3 of Section S4. In Section S7, the distance dependencies of the radial parts (compared to the Bunge orbitals) are shown for all elements of the 2nd period. In addition, the overlap integrals with the Bunge AOs are listed for all basis sets in Table 4 of Section S4.

As it may be seen from the plots of the radial parts shown in Figs. 1–7 (Section S7), the shapes of all orbitals obtained within the H-3 and H-4 settings are very similar to the shapes of the Bunge AOs taken as a reference, whereas the 2s-orbital in the H-2 setting differs substantially from the corresponding orbitals in other sets. This is also reflected in the overlap integrals: 2s-functions of the H-2 basis overlap with the Bunge 2s-AOs by less than 93%, while typical overlaps of the 2s-AOs for the H-3 and H-4 sets arrive to more than 99%. The reason of such deviation was already mentioned previously—the inconsistency of assuming the equal effective core charges for the 1s- and 2s-electrons in the H-2 basis, making it impossible to correctly represent the shapes of the 1s- and 2s-orbitals simultaneously. It leads to much less diffuse 2s-orbitals in the H-2 basis (overestimated core charge) and, as we shall see later, to a poor reproduction of the atomic parameters calculated with use of this basis.

As for the H-3 and H-4 basis sets, they, having minimal set of primitives (two STO primitives for an 2s-orbital and one for each 1s- and 2p-orbitals), allows us to represent the shapes of the Bunge AOs (seven STO primitives for s- and five for p-orbitals) very closely. Only exceptions are 2p-orbitals of oxygen and fluorine, for the reasons discussed below.

Analyzing the screening constants (S_{nl}) derived from our calculations on the direct problem for the H-3 (Table 2) and H-4 sets (for H-4 sets 1s- and 2p-shielding constants are very close to corresponding values for the H-3 set, because the 1s- and 2p-orbital exponents are similar in these two basis sets), we may see that the Slater rules [8] work pretty good for 1s-orbitals. All screening constants for 1s-orbitals have values close to the Slater value of 0.30. The simplest estimate of $S_{1s} = 5/16 = 0.3125$, coming from minimizing the energy of the $1s^2$ configuration in the field of the nucleus with the charge Z, fairly sets a "theoretical" reference point. For the 2p-orbitals of the H-3 and H-4 basis sets, the Slater's rules in their original form work not so well, but still we have quite close screening constants (especially for B, C and N) to those, obtained by Slater. For the 2s-orbitals of the H-3 shape, the original Slater's rules do not work at all, which is anticipated, since that latter has functional form, differing from one Slater used. However, an analogous additive rule fairly applies for the 2s-screening constants:

$$S_{2s} = 0.733 n_{1s} + 0.194\left(n_{2s} + n_{2p} - 1\right) \tag{6}$$

Table 2 Screening constants for the orbitals of the H-3 basis set

	Li	Be	B	C	N	O	F
S_{1s}							
	0.306	0.293	0.289	0.289	0.288	0.288	0.289
[8]	0.30						
[33]	0.307	0.314	0.322	0.329	0.336	0.343	0.350
S_{2s}							
	1.467	1.681	1.842	2.037	2.246	2.439	2.642
Equation (6)	1.467	1.662	1.856	2.051	2.245	2.440	2.634
[33]	1.721	2.081	2.441	2.801	3.161	3.522	3.882
S_{2p}							
	–	–	2.627	2.908	3.208	3.799	4.141
Equation (7)			2.553	2.945	3.337	3.729	4.120
[33]	–	–	2.579	2.911	3.244	3.577	3.909
$= S_{2s}$ [8]	1.70	2.05	2.40	2.75	3.10	3.45	3.80

and for the 2*p*-ones:

$$S_{2p} = 0.885n_{1s} + 0.392\left(n_{2s} + n_{2p} - 1\right), \tag{7}$$

where n_i is the number of electrons in the *i*th shell.

The rules for the 2*s*- and 2*p*-screening constant correspond to simple relations for the exponents themselves:

$$\xi_{2s}^{H-3} = 0.403(Z - 1.097)$$
$$\xi_{2p}^{H-3} = 0.304(Z - 0.975)$$

valid for $Z > 2$. For the 2*p*-exponent of the H-4 basis, we have nearly the same relation as for the H-3 basis:

$$\xi_{2p}^{H-4} = 0.304(Z - 0.972)$$

These expressions together with the calculated values of the exponents are plotted in Fig. 1. The coefficients of determination, characterizing the quality of the fits, are: $R_{2s}^2 = 0.9999$; $R_{2p}^2(H-3) = 0.9958$; $R_{2p}^2(H-4) = 0.9916$.

3.1.2 Position of the node of the 2s orbital

A remarkable qualitative feature of the basis sets considered in the present paper is the position r_0 of the radial node of the 2*s*-orbitals for all considered forms of the AOs, which is almost independent of the basis set (differences occur only on the 4th decimal place for all elements except Li—for Li on the 2nd decimal place). For any given element, it depends on the nuclear charge (r_0 decreases with increasing *Z*). Thus, this point may be considered as a characteristic of an element. For the H-2 and H-3 basis sets, its explicit expressions are simply:

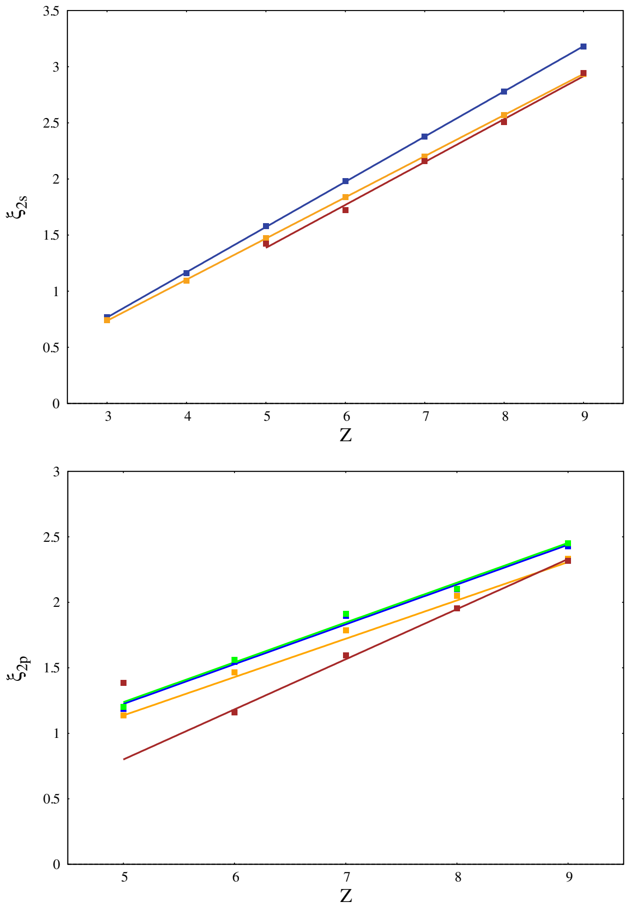

Fig. 1 Orbital exponents as functions of nuclear charge. Left: the 2*s*-exponent; right: the 2*p*-exponent. Color code: blue corresponds to the H-3 basis set, green: H-4, orange: H-3, fitted from the Bunge AOs, brown: H-3, fitted against atomic spectra

Table 3 The $2s$ node position (a.u.) calculated within the direct problem in the H-3 setting and estimated from exact numerical and approximate [Eq. (12)] WKB values

	Li	Be	B	C	N	O	F
Calcd.	0.867	0.616	0.477	0.390	0.330	0.286	0.252
Numerical WKB	1.082	0.671	0.495	0.394	0.328	0.281	0.246
r_0	0.995	0.663	0.497	0.398	0.332	0.284	0.249

$$r_0^{H-2} = \frac{1}{\xi_{2s}};\ r_0^{H-3} = \frac{3}{\xi_{1s} + \xi_{2s}}. \tag{8}$$

In the case of the H-3 basis, $2s$-orbital exponent may be expressed through the $1s$-orbital exponent and the position of the node as:

$$\xi_{2s} = \frac{3}{r_0} - \xi_{1s}. \tag{9}$$

With use of the analytic estimate for ξ_{1s} (coming from the energy minimum of the $1s^2$ configuration), the $2s$-orbital exponent turns a function of the nuclear charge and the node position:

$$\xi_{2s} = \frac{3}{r_0} - Z + \frac{5}{16}. \tag{10}$$

Since r_0 is a function of the nuclear charge as well, it is of interest to find a theoretical estimate for it, making it possible to derive a theoretical estimate for the $2s$-orbital exponent through the nuclear charge only. In order to do so, let us consider $2s$-electron moving in the field of a nucleus and two $1s$-electrons. One-electron effective Hamiltonian for this is:

$$\hat{H} = -\frac{1}{2}\nabla^2 - \left(\frac{Z}{r} - 2v_{1s}(r)\right), \tag{11}$$

where $v_{1s}(r)$ is the electrostatic potential of a $1s$-electron, coming from the averaging of the electron–electron repulsion operator over the $1s$-states [Section S3, Eq. (27)]. The exact analytic solution for the eigenstates of the Hamiltonian Eq. (11) is not feasible, and thus with use of the Wentzel–Kramers–Brillouin (WKB) method [34, 35] we obtained an approximation as described in Section S3. This move leads to the system of equations (Eqs. (30)–(32) of Section S3), yielding the position of the node. It may be solved numerically, which is done, and results are presented in Table 3. Moreover, in Section S3 an analytic estimate of the node position is obtained :

$$r_0 = \frac{x_0}{Z - 1} \tag{12}$$

($x_0 = 1.9893$ is the solution of a transcendental equation derived in Section S3). As one can see, the numerical solution of the WKB problem yields a very good estimate (except for Li) of the results obtained within the direct problem. This expression works very well and may be useful in some applications.

3.1.3 Spectra and ionization potentials as derived from direct problem

The core attraction and Slater–Condon parameters calculated with use of the introduced basis sets with the exponents derived from the solution of the direct problem compared to their "experimental values" (those extracted from the available spectral data) or those calculated previously in the literature are presented in Tables 10–16 of Section S5, where we also list the values of the first ionization potentials $s^2p^n \rightarrow s^2p^{n-1}$ ($\mathrm{IP}^{(1)}$) and $s^2p^n \rightarrow s^1p^n$ ($\mathrm{IP}^{(2)}$). In Fig. 2 we plot the relative errors of the calculated ionization potentials.

Regarding the ionization potentials $\mathrm{IP}^{(1)}$, one can see that the H-2 basis gives totally inappropriate results, while the H-3 and H-4 results look not bad in general, but for oxygen and fluorine the relative error is still close to 20%. This is caused by electronic correlations not taken into account, throughout minimization of the HF energy with respect to the exponents. Values of the ionization potential $\mathrm{IP}^{(2)}$ for all basis sets (H-3, H-4 and Bunge) deviate quite substantially from the experimental data, especially for B (the relative error of 15–20%) and C (the relative error of 10–15%).

In Table 4, the relative energies of the terms, calculated with the H-3, H-4 and Bunge sets for B, C, N, O and F, are compared with the corresponding experimental values. These results are characterized by the quantity

$$\phi = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{E_i - E_i^{(\mathrm{exp})}}{E_i^{(\mathrm{exp})}}\right)^2} \cdot 100\%$$

calculated for B, C, N and O (for F there are no experimental data for the single line, which can be calculated with only $1s$, $2s$ and $2p$ exponents) for the H-3, H-4 and Bunge sets and plotted in Fig. 3 together with the mean absolute deviation (in the formula N stands for the number of considered atomic lines, E_i is the calculated energy of ith transition and

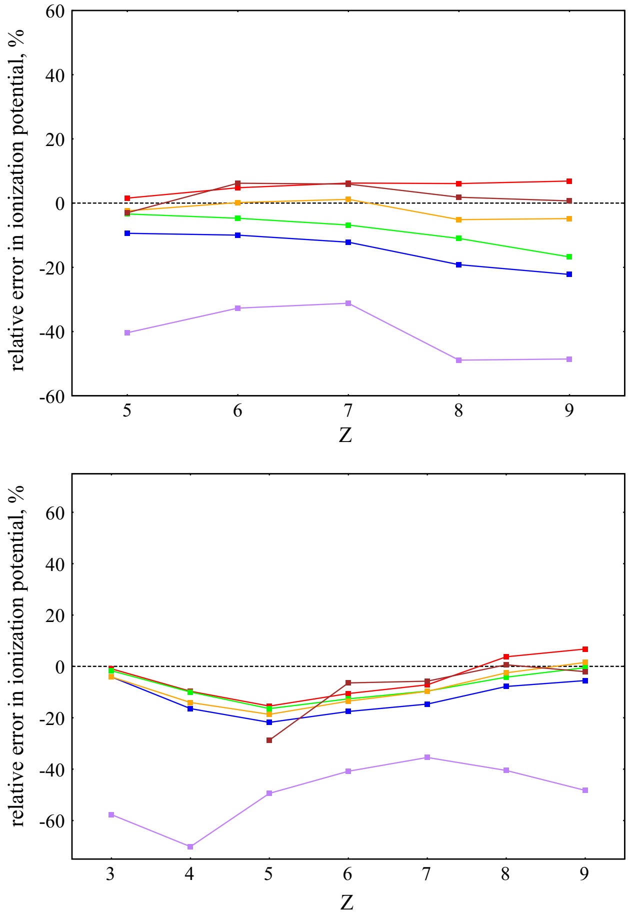

Fig. 2 The relative errors of the ionization potentials for different basis sets. Left: $IP^{(1)}$; right: $IP^{(2)}$. Color code: red corresponds to the results obtained in the present work with the Bunge orbitals, purple: the same with H-2, blue: with H-3, green: with H-4, yellow: with H-3, fitted from the Bunge AOs, brown: with H-3, extracted from atomic spectra

$E_i^{(exp)}$—its experimental value). Dimensionless quantity ϕ is thus the average relative error for the term energies. Being compared to the standard mean square deviation, it has that advantage that its values do not depend on the range of energies, and thus it may be considered as a uniform parameter, allowing to compare quality of calculations for different elements.

As it may be concluded from Fig. 3 for the H-3 basis set, the average relative errors lay in the range 20–40%, for H-4 in the range 15–40% and for the Bunge basis set—15–25%. At the same time, mean absolute deviations obtained with the H-3 set are actually not much higher (by 0.25–0.40 eV) than those obtained with the Bunge set. It shows that the term energies calculated with the H-3 set are not that bad as it may be concluded from quite high relative errors, coming mostly from errors in low-energy transitions.

However, errors in the atomic spectra and IP's are quite high for the H-3 and H-4 basis sets (interestingly, the H-3 setting permits to reproduce the spectral transitions even better than the H-4 setting for C, N and O), and thus results should be improved. Two possible ways may be chosen to do so. First, we can modify our way of extracting exponents from the theoretical model (change the setting of the direct problem explicitly accounting the correlations). This will be done elsewhere. In the present work, we use another possible way to improve our results, based on solution of two inverse problems as described below.

3.2 Inverse problems

3.2.1 Spectral data fit

The exponents of the H-3 basis extracted from the spectral data for the B, C, N, O and F atoms are plotted in Fig. 1,

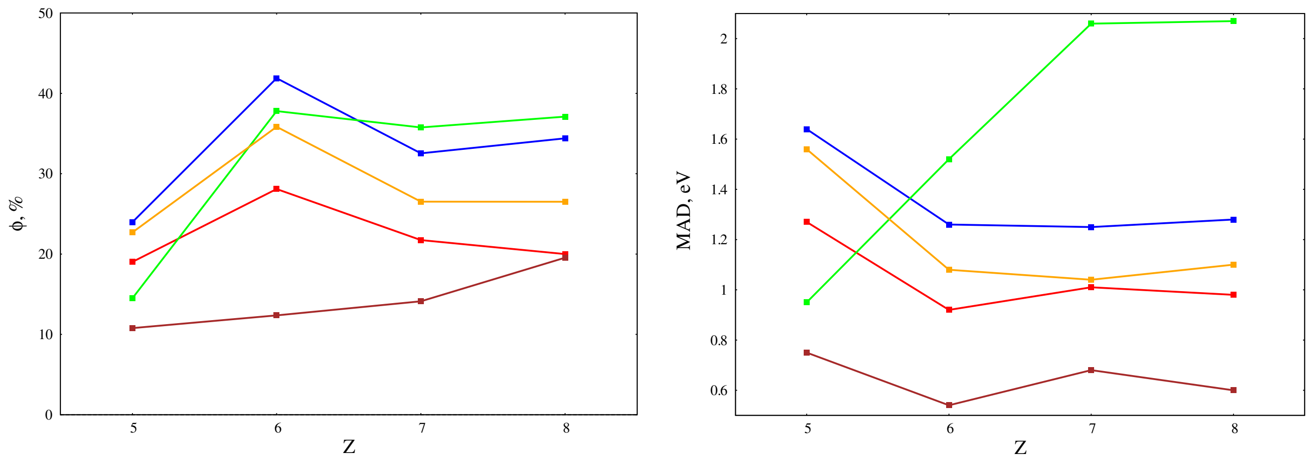

Fig. 3 Errors in the calculated term energies for different basis sets. Left: average relative errors (ϕ); right: mean absolute deviations. Color code red: calculated with the Bunge orbitals; blue: with H-3; green: with H-4; yellow: with H-3, fit from the Bunge AOs: brown: with H-3, extracted from experimental spectra

Table 4 The calculated and observed relative term energies

Element	Atomic state	H-3 direct	H-3 orbital fit	H-3 spectral fit	H-4	Bunge	Exp [30–32]
B: $2s^22p\,(^2P)$	$2s2p^2\,(^4P)$	2.92	3.08	3.98	3.73	3.46	3.55
	$2s2p^2\,(^2P)$	4.95	5.03	5.98	5.71	5.32	5.93
	$2s2p^2\,(^2D)$	5.63	5.68	6.78	6.40	5.94	7.88
	$2s2p^2\,(^2S)$	6.31	6.33	7.57	7.09	6.56	8.99
C: $2s^22p^2\,(^3P)$	$2s^22p^2\,(^1D)$	1.77	1.68	1.33	1.80	1.59	1.26
	$2s^22p^2\,(^1S)$	4.44	4.20	3.33	4.50	3.97	2.68[a]
	$2s2p^3\,(^5S)$	1.45	1.84	3.63	2.72	2.54	4.16[a]
	$2s2p^3\,(^3D)$	7.57	7.66	8.32	8.81	8.04	7.94
	$2s2p^3\,(^3P)$	9.35	9.34	9.66	10.61	9.63	9.33
	$2s2p^3\,(^3S)$	15.33	15.03	14.40	16.28	15.01	13.12
N: $2s^22p^3\,(^4S)$	$2s^22p^3\,(^3D)$	3.26	3.07	2.74	3.29	2.89	2.39
	$2s^22p^3\,(^2P)$	5.44	5.12	4.57	5.48	4.82	3.58
	$2s2p^4\,(^4P)$	10.76	10.81	10.64	12.57	11.30	10.93
	$2s2p^4\,(^2D)$	17.13	16.83	16.10	18.84	16.97	15.03
O: $2s^22p^4\,(^3P)$	$2s^22p^4\,(^1D)$	2.53	2.35	2.24	2.55	2.19	1.96
	$2s^22p^4\,(^1S)$	6.33	5.88	5.60	6.37	5.49	4.18
	$2s2p^5\,(^3P)$	16.77	16.87	15.76	19.08	17.06	15.66
F: $2s^22p^5\,(^2P)$	$2s2p^6\,(^2S)$	22.45	32.96	22.69	36.17	23.93	–

[a] As it may be seen, the theoretical H-3, H-4 and Bunge basis sets give inverse order of the terms $s^2p^2\,(^1S)$ and $sp^3\,(^5S)$ for carbon. Unlike the situation with boron, noted in the text, in this case one cannot rigorously establish any "correct" order of terms within the *LS*-scheme, since negative core attraction parameters enter the expression of terms energy difference (so we have negative and positive contributions and the sign of the expression is not definite as it is in the case of boron). However, the H-3 spectral parameterization allowed us to reproduce the experimental order of terms. It is, however, highly unlikely that the correlated setting assures the experimental order in case of theoretical basis sets since it is largely predetermined by the relative positions of the 2*s*- and 2*p*-subshells on the energy scale

and their numerical values may be found in Table 5 of Section S4. Linear fit for these "spectral" exponents yields:

$$\xi_{2s} = 0.382\,(Z - 1.373)$$

$$\xi_{2p} = 0.383\,(Z - 2.910)$$

($R^2_{2s} = 0.9965$; $R^2_{2p} = 0.9978$). As one can see, spectral exponents are quite smaller than the theoretical values. It may correspond to the implicit influence of the correlation effects, covered by the spectral parameterization (especially fitted to reproduce experimental atomic properties). It is qualitatively clear that correlation effects should decrease effective core charges, but explicit quantitative estimates of these effects lay beyond the scope of the present paper and are postponed for the future work.

Several issues related to this inverse problem should be mentioned. First, for the Li and Be atoms there are not enough spectral data to extract the orbital exponents, and thus we do not consider these elements here. For the F atom, only one line suitable for the spectral fit is expected to be present in the atomic spectra, but there is no experimental value for it in the literature. Thus, in this case the orbital exponents were extracted from two ionization potentials $\mathrm{IP}^{(1)}$ and $\mathrm{IP}^{(2)}$ and electronic affinity. Another issue relates to the experimental data for the boron atom. In the NIST database, one finds the relative energies (in eV): 3.55 $(^4P;sp^2)$, 5.93 $(^2D;sp^2)$, 7.88 $(^2S;sp^2)$, 8.99 $(^2P;sp^2)$. This consequently comes from all original sources cited by the NIST. However, our calculations with the H-3, H-4 (obtained from the direct problem) and the Bunge orbitals show the order of terms to be ^4P, ^2P, ^2D, ^2S. Moreover, it can be easily shown, for example, that in the uncorrelated setting:

$$E\left(^2S\right) - E\left(^2P\right) = \frac{6}{25}F^2_{pp} > 0$$

in contradiction with the NIST ordering. Although one cannot exclude some extraordinarily strong correlation manifestations changing the order of the excited states on the energy scale, in the present work we have chosen to arrange the experimental energies according to the non-correlated ("old") MC SCF order, leaving the investigation of the other options along the correlated setting in a line with Ref. [36] for the future.

Atomic parameters, obtained with the "spectral" exponents, are listed in Tables 10–16 of Section S5 and in Table 4; relative term energies, calculated with the spectroscopic H-3 and the Bunge basis sets, are compared with the experimental values in Table 4.

As it can be seen from Fig. 2, the ionization potentials $\mathrm{IP}^{(1)}$, obtained with the spectroscopically determined exponents of the H-3 basis sets are very close to experimental values (maximal deviation of 5%). The spectroscopic H-3 basis set also gives the best values of $\mathrm{IP}^{(2)}$ (better than the Bunge basis set), which are very close to experiment (maximal deviation of 6%). The term energies obtained from the spectroscopic parameterization are better then those, calculated with the Bunge orbitals. It is quite surprising that such a simple basis set as H-3 can be minimally adjusted to assure even better agreement with experiment than the Bunge basis, having seven STO primitives for *s*- and five for *p*-orbitals.

3.2.2 Fit against Bunge orbitals

The exponents of the H-3 basis fitted from the Bunge AOs for all elements are presented in Table 7 of Section S4 and atomic parameters, obtained with these exponents are listed in Tables 10–16 (Section S5) and in Table 4. The linear fit of exponents *vs* nuclear charge yields:

$$\xi_{2s} = 0.367\,(Z - 0.990)$$
$$\xi_{2p} = 0.297\,(Z - 1.101)$$

($R^2_{2s} = 0.9999$; $R^2_{2p} = 0.9980$). One can see that the $2s$ and $2p$ exponents, obtained here, are smaller than the corresponding theoretical exponents and their values increase somewhat slower with Z than the theoretical ones. It corresponds to the tendency, already noted above, that the theoretical effective core charges, coming form the SCF setting, are overestimated, and decreasing their values (theoretically it should come from some correlation effects) allows us to reproduce the atomic properties in a better agreement with the experimental data.

The ionization potentials and the average relative errors in atomic spectra become better than those calculated with the H-3 set (relative error in the IP's is lower than 5%, and the relative error is in the range 20–35% for the atomic spectra), derived from the direct problem. Their values predictably converge to the IP's calculated with the Bunge AOs. All these results show that the H-3 basis can be parameterized to provide very accurate representation of the atomic quantities. It is anticipated that more flexible H-4 basis will also be able to reproduce atomic properties after similar fitting. However, further exploring the H-4 basis goes beyond our task of developing basis sets suitable for analytic treatments.

4 Conclusions

A series of the hydrogen-like basis sets for the 2nd-row elements is proposed, and their parameters (orbital exponents) are obtained by solving both *direct* and *inverse* problems. The orbitals, obtained by solving the *direct* problem within the H-3 and H-4 settings, are shown to reproduce atomic properties of 2nd-row elements with quite satisfying accuracy. The exponents, obtained by solving two *inverse* problems, show that the considered model basis sets may be parameterized up to the level of quantitative reproducing of the atomic spectra, sometimes even better, than the much more involved Bunge basis sets. At the same time, the proposed basis sets (especially the H-3 one) have simple functional form, making them easily useful for analytic models. The linear dependencies of the orbital exponents on nuclear charges (extensions of the Slater rules) were established, and their accuracy was checked for all types of the generalized H-like basis sets and for the direct and inverse methods of determining the values of the exponents. The node position of the $2s$ orbital for the 2nd-row elements is characteristic for each given atom. Based on the WKB approximation, we proposed theoretical estimates of the node position from the first principles and found simple formulas, respectively, bracketing and expressing the position of the node as a function of the nuclear charge. It was shown that the knowledge of the node position allows one to express the exponents of all *s*-orbitals as analytic functions of nuclear charge, justifying the empirical fits. Together with the fact that all matrix elements of the Hamiltonian may be expressed through the orbital exponents of the proposed basis sets in the easiest possible way, it makes possible a construction of analytic quantum chemical models, depending only on nuclear charges of atoms and by this enables analytic studies of the former.

Acknowledgements This work has been supported by the Volkswagenstiftung (Grant No 151110). The authors acknowledge valuable discussion with Dr. Nikolai S. Mosyagin (St. Petersbourg, Russia) and thank the Referees for their valuable comments and references to the previous important studies, allowing to improve the present work significantly.

References

1. Helgaker T, Jørgensen P, Olsen J (2014) Molecular electronic-structure theory. Wiley, Hoboken
2. Jensen F (2012) WIREs Comput Mol Sci 3:273–295

3. Davidson E, Feller D (1986) Chem Rev 86:681–696
4. Deringer VL, Tchougréeff AL, Dronskowski R (2011) J Phys Chem A 115:5461–5466
5. Maintz S, Deringer VL, Tchougréeff AL, Dronskowski R (2013) J Comput Chem 34:2557–2567
6. Maintz S, Deringer VL, Tchougréeff AL, Dronskowski R (2016) J Comput Chem 37:1030–1035
7. Junquera J, Paz O, Sanchez-Portal D, Artacho E (2001) Phys Rev B 64:235111
8. Slater JC (1930) Phys Rev 16:57–64
9. Boys SF (1950) Proc R Soc London A Math Phys Sci 200:542–554
10. Tchougréeff AL, Dronskowski R (2016) Mol Phys 114:1423–1444
11. Tchougréeff AL (2017) AIP Conf Proc 1906:030004
12. Tchougréeff AL, Theor Chem Acc (submitted)
13. Lennard-Jones JE (1937) Proc R Soc A158:280
14. Coulson CA (1938) Proc R Soc A 164:383
15. Misurkin IA, Ovchinnikov AA (1977) Russ Chem Rev 46:967–987
16. Morse PM, Young LA, Haurwitz ES (1935) Phys Rev 48:948–954
17. Duncanson WE, Coulson CA (1944) Proc R Soc Edinb 62:37–40
18. Pujol L, Simon J-C (1968) Theor Chim Acta (Berl) 11:59–74
19. Gombás P, Szondy T (1970) Solutions of the simplified self-consistent field for all atoms of the periodic system of elements from Z=2 to Z=92. Adam Hilger, London
20. Gombás P, Szondy T (1968) Acta Phys Acad Sci Hung 25:345–359
21. Clementi E, Roetti C (1974) At Data Nucl Data Tables 14:177
22. Bunge CF, Barrientos JA, Bunge AV (1993) At Data Nucl Data Tables 53:113–162
23. Szasz L (1985) Pseudopotential theory of atoms and molecules. Wiley, Hoboken
24. Moffit W (1950) Proc R Soc Lond A 202:534–547
25. Бирон ЕВ (1915) ЖРФХО, ч. хим., 47, вып. 4, 964–968; Biron EV (1915) J Russ Chem Part 47(4):964–968 (in Russian)
26. Pyykkö P (1979) J Chem Res (S) 11:380–381
27. Roothaan CCJ (1960) Rev Mod Phys 32:179–185
28. Condon EU, Shortley GH (1935) The theory of atomic spectra. Cambridge University Press, London
29. Тихонов АН, Гончарский АВ, Степанов ВВ, Ягола АГ, (1983) Регуляризующие алгоритмы и априорная информация. Наука, Москва [in Russian]; Tikhonov AN, Goncharsky AV, Stepanov VV, Yagola AG (1995) Numerical methods for the solution of Ill-posed problems. Kluwer, Dordrecht (in English)
30. NIST Ionization Energies Databse. https://physics.nist.gov/PhysRefData/ASD/ionEnergy.html. Accessed May 2018
31. NIST Atomic Spectra Database. https://physics.nist.gov/PhysRefData/ASD/lines_form.html. Accessed May 2018
32. Verner DA, Verner EM, Ferland GJ (1996) At Data Nucl Data Tables 64:1–180
33. Clementi E, Raimondi DJ (1963) Chem Phys 38:2686
34. Kramers HA (1926) Z Phys 39:828–840
35. Wentzel G (1926) Z Phys 38:518–529
36. Tchougréeff AL, Ángyán J (2010) Int J Quant Chem 110:454–475