Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals

W. J. Hehre, R. F. Stewart and J. A. Pople

View Online   Export Citation

ARTICLES YOU MAY BE INTERESTED IN

AIP Publishing

Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals

W. J. Hehre, R. F. Stewart, and J. A. Pople

Department of Chemistry, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

(Received 17 March 1969)

Least-squares representations of Slater-type atomic orbitals as a sum of Gaussian-type orbitals are presented. These have the special feature that common Gaussian exponents are shared between Slater-type $2s$ and $2p$ functions. Use of these atomic orbitals in self-consistent molecular-orbital calculations is shown to lead to values of atomization energies, atomic populations, and electric dipole moments which converge rapidly (with increasing size of Gaussian expansion) to the values appropriate for pure Slater-type orbitals. The ζ exponents (or scale factors) for the atomic orbitals which are optimized for a number of molecules are also shown to be nearly independent of the number of Gaussian functions. A standard set of ζ values for use in molecular calculations is suggested on the basis of this study and is shown to be adequate for the calculation of total and atomization energies, but less appropriate for studies of charge distribution.

I. INTRODUCTION

There is a need for well-defined *ab initio* molecular-orbital calculations which are simple enough to be applied systematically to moderately large organic molecules. One such method, which has already been used quite extensively, is the linear combination of atomic-orbitals self-consistent-field (LCAO SCF) method using a minimal basis set of Slater-type atomic orbitals (STO's). For molecules with no atom heavier than fluorine, only $1s$, $2s$, and $2p$ atomic orbitals are involved and these have the form

$$\phi_{1s}(\zeta_1, \mathbf{r}) = (\zeta_1^3/\pi)^{1/2} \exp(-\zeta_1 r),$$

$$\phi_{2s}(\zeta_2, \mathbf{r}) = (\zeta_2^5/3\pi)^{1/2} r \exp(-\zeta_2 r),$$

$$\phi_{2p}(\zeta_2, \mathbf{r}) = (\zeta_2^5/\pi)^{1/2} r \exp(-\zeta_2 r) \cos\theta, \quad (1.1)$$

and similar expressions for other $2p$ functions. In general, different values of the exponents ζ_2 may be used for $2s$ and $2p$ functions, but in the simplest form of the theory, with which we shall be concerned here, they are constrained to be identical.

Variational calculations of LCAO SCF molecular orbitals

$$\psi_i = \sum_\mu c_{\mu i} \phi_\mu \quad (1.2)$$

using the basis set (1.1) with fixed ζ exponents have been published for a number of molecules.[1-7] These involve the use of somewhat arbitrarily chosen standard sets of ζ values (usually Slater atomic values[8]) but it is preferable to treat the exponents as further variational parameters leading to final optimized values.

Such calculations have been published for first-row diatomics,[1] methane,[9] and water.[10]

Full Slater-type calculations are time consuming largely because of the evaluaton of two-electron integrals, even though efficient programs involving Gaussian transforms and numerical integration are being developed.[11,12] Another possibility is to replace each Slater-type orbital by a linear combination of a small number of Gaussian-type orbitals, since integrals involving Gaussian functions can be evaluated analytically. A technique of this sort was first used by Foster and Boys[2] and it was discussed in some detail by Reeves, Fletcher, O-ohata, Takata, and Huzinaga,[13] who proposed obtaining the Gaussian representations by least-squares fitting. In calculations of this sort, the Slater-type orbitals (1.1) are replaced by a finite Gaussian expansion at all points in the calculation. The results will approach those of the full Slater-type calculation as the size of the Gaussian set increases. We have obtained representations of Slater-type orbitals using 2–6 Gaussian functions by least-squares methods, and the main purpose of this paper is to use these basis functions (referred to as STO-2G, STO-3G, STO-4G, STO-5G, and STO-6G) to study the approach of LCAO SCF calculations towards the full Slater-type results. Since the integrals can be evaluated rapidly, a study of such a series of calculations may be the most efficient way of obtaining the information implicit in a full Slater-type wavefunction, either with fixed or with optimized ζ exponents. In this preliminary work, emphasis will be laid on the mathematical convergence and the choice of exponents, rather than comparison of any calculated results with experimental data.

[1] B. J. Ransil, Rev. Mod. Phys. **32**, 239, 245 (1960).

[2] J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 303 (1960).

[3] W. E. Palke and W. N. Lipscomb, J. Am. Chem. Soc. **88**, 2384 (1966).

[4] M. D. Newton and W. E. Palke, J. Chem. Phys. **45**, 2329 (1966).

[5] W. E. Palke and R. M. Pitzer, J. Chem. Phys. **46**, 3948 (1967).

[6] U. Kaldor and I. Shavitt, J. Chem. Phys. **45**, 888 (1966); **48**, 191 (1968).

[7] R. M. Pitzer, J. Chem. Phys. **47**, 965 (1967).

[8] J. C. Slater, Phys. Rev. **36**, 57 (1930).

[9] R. M. Pitzer, J. Chem. Phys. **46**, 4871 (1967).

[10] S. Aung, R. M. Pitzer, and S. I. Chan, J. Chem. Phys. **49**, 2071 (1968).

[11] I. Shavitt and M. Karplus, J. Chem. Phys. **36**, 550 (1962); **43**, 398 (1965).

[12] R. M. Stevens (private communication).

[13] C. M. Reeves and R. Fletcher, J. Chem. Phys. **42**, 4073 (1965); K. O-ohata, H. Taketa, and S. Huzinaga, J. Phys. Soc. (Japan) **21**, 2306 (1966).

TABLE I. Coefficients and exponents for Gaussian expansion of Slater orbitals.[a]

K	α_{1s}	d_{1s}	ϵ_{1s}	α_2	d_{2s}	ϵ_{2s}	d_{2p}	ϵ_{2p}
2	1.51623(−1)	6.78914(−1)	3.16(−3)	9.74545(−2)	9.63782(−1)	5.60(−3)	6.12820(−1)	3.48(−3)
	8.51819(−1)	4.30129(−1)		3.84244(−1)	4.94718(−2)		5.11541(−1)	
3	1.09818(−1)	4.44635(−1)	3.31(−4)	7.51386(−2)	7.00115(−1)	6.42(−4)	3.91957(−1)	3.60(−4)
	4.05771(−1)	5.35328(−1)		2.31031(−1)	3.99513(−1)		6.07684(−1)	
	2.22766	1.54329(−1)		9.94203(−1)	−9.99672(−2)		1.55916(−1)	
4	8.80187(−2)	2.91626(−1)	4.38(−5)	6.28104(−2)	4.97767(−1)	7.95(−5)	2.46313(−1)	4.82(−5)
	2.65204(−1)	5.32846(−1)		1.63541(−1)	5.58855(−1)		5.83575(−1)	
	9.54620(−1)	2.60141(−1)		5.02989(−1)	2.97680(−5)		2.86379(−1)	
	5.21686	5.67523(−2)		2.32350	−6.22071(−2)		4.36843(−2)	
5	7.44527(−2)	1.93572(−1)	6.88(−6)	5.44949(−2)	3.46121(−1)	1.17(−5)	1.56828(−1)	7.28(−6)
	1.97572(−1)	4.82570(−1)		1.27920(−1)	6.12290(−1)		5.10240(−1)	
	5.78648(−1)	3.31816(−1)		3.29060(−1)	1.28997(−1)		3.73598(−1)	
	2.07173	1.13541(−1)		1.03250	−6.53275(−2)		1.07558(−1)	
	1.13056(1)	2.21406(−2)		5.03629	−2.94086(−2)		1.25561(−2)	
6	6.51095(−2)	1.30334(−1)	1.24(−6)	4.85690(−2)	2.40706(−1)	2.01(−6)	1.01708(−1)	1.22(−6)
	1.58088(−1)	4.16492(−1)		1.05960(−1)	5.95117(−1)		4.25860(−1)	
	4.07099(−1)	3.70563(−1)		2.43977(−1)	2.50242(−1)		4.18036(−1)	
	1.18506	1.68538(−1)		6.34142(−1)	−3.37854(−2)		1.73897(−1)	
	4.23592	4.93615(−2)		2.04036	−4.69917(−2)		3.76794(−2)	
	2.31030(1)	9.16360(−3)		1.03087(1)	−1.32528(−2)		3.75970(−3)	

[a] Six significant figures are tabulated for each Gaussian parameter. The appropriate power of 10 is enclosed in parentheses to the right of each entry

II. METHOD

As indicated in the introduction, each Slater-type atomic orbital is replaced by an atomic orbital ϕ_μ' which is a sum of K Gaussian-type orbitals $(K=2–6)$. These combinations are obtained for Slater-type orbitals with $\zeta=1$ and then uniformly scaled. Thus

$$\phi_\mu'(\zeta, \mathbf{r}) = \zeta^{3/2}\phi_\mu'(1, \zeta\mathbf{r}), \qquad (2.1)$$

where

$$\phi_{1s}'(1, \mathbf{r}) = \sum_k^K d_{1s,k}g_{1s}(\alpha_{1k}, \mathbf{r}),$$

$$\phi_{2s}'(1, \mathbf{r}) = \sum_k^K d_{2s,k}g_{1s}(\alpha_{2k}, \mathbf{r}),$$

$$\phi_{2p}'(1, \mathbf{r}) = \sum_k^K d_{2p,k}g_{2p}(\alpha_{2k}, \mathbf{r}). \qquad (2.2)$$

Here g_{1s} and g_{2p} are the Gaussian-type orbitals

$$g_{1s}(\alpha, \mathbf{r}) = (2\alpha/\pi)^{3/4} \exp(-\alpha r^2),$$

$$g_{2p}(\alpha, \mathbf{r}) = (128\alpha^5/\pi^3)^{1/4} r \exp(-\alpha r^2) \cos\theta. \qquad (2.3)$$

Two points should be noted about the expansions (2.2).

In the first place, the 2s exponential function is replaced by a linear combination of 1s Gaussian functions. This is because integrals involving 2s-type Gaussian functions cannot be handled by such elementary methods. Second, the combinations for ϕ_{2s}' and ϕ_{2p}' share the same set of Gaussian exponents α_{2k}. In previous studies of this sort,[13,14] two different sets of α values have been taken for 2s and 2p. The restriction of a common set reduces the flexibility of the representation somewhat, but we shall see that this is more than offset by the possibility of a more efficient computer program for the two-electron integrals.

The constants d and α in (2.2) are chosen to minimize the integrals

$$\epsilon_{1s} = \int (\phi_{1s}-\phi_{1s}')^2 d\tau,$$

$$\epsilon_{2s}+\epsilon_{2p} = \int (\phi_{2s}-\phi_{2s}')^2 d\tau + \int (\phi_{2p}-\phi_{2p}')^2 d\tau, \qquad (2.4)$$

subject to normalization constraints. Actual values of α and d along with the corresponding ϵ values are listed in Table I.

[14] W. J. Hehre, R. F. Stewart, and J. A. Pople, Symp. Faraday Soc. 2, 15 (1968).

TABLE II. Total energy with fixed exponents.

Molecule	STO-2G	STO-3G	STO-4G	STO-5G	STO-6G	STO[a]
HF	−95.5572	−98.5274	−99.2204	−99.3969	−99.4501	−99.4785
CO	−107.9282	−111.2297	−112.0337	−112.2443	−112.3086	−112.3436
CH₄	−38.5719	−39.7153	−39.9985	−40.0771	−40.1011	−40.1141
C₂H₂	−73.5950	−75.8415	−76.3952	−76.5468	−76.5935	−76.6185
C₂H₄	−74.7950	−77.0506	−77.6089	−77.7618	−77.8090	−77.8343
C₂H₆	−76.0208	−78.2795	−78.8410	−78.9958	−79.0434	−79.0689
BH₃	−25.2960	−26.0647	−26.2583	−26.3121	−26.3288	−26.3377
B₂H₆	−50.6064	−52.1351	−52.5201	−52.6268	−52.6603	−52.6782
NH₃	−53.8350	−55.4536	−55.8490	−55.9553	−55.9874	−56.0050
HCN	−88.9456	−91.6622	−92.3275	−92.5058	−92.5608	−92.5903
H₂CO	−109.0037	−112.3295	−113.1362	−113.3497	−113.4146	−113.4496

[a] HF, CO, Ref. 1; polyatomics except H_2CO, Ref. 3; H_2CO, Ref. 4; for C_2H_2 and NH_3 our calculated nuclear repulsion energies do not agree exactly with those of Ref. 3. Consequently, the total energies quoted here differ slightly from those in the original paper.

The method for solving the least-squares equations is reported elsewhere.[15] The two-, three-, and four- (1s) set expansions were solved in single precision on a CDC 1604A. All derivative terms were less than 10^{-9} in absolute value and the eigenvalues of the least-squares matrix were all positive. The four- $(2s, 2p)$, five-, and six-set expansions were solved in double precision. For these cases all derivatives had absolute values less than 10^{-21} and again the eigenvalues of the matrix of second partial derivative terms were positive. The smallest eigenvalue was 4×10^{-9} for the six-set $1s$ expansion. Thus, at worst, the absolute error in a Gaussian parameter is 10^{-12} for the double-precision work. In the single-precision results the largest absolute error is 5×10^{-8} which is for the largest α_{1s} term in the four-set expansion.

Given the basis functions (2.2), the molecular-orbital coefficients $c_{\mu i}$ can be found by minimizing the total electronic energy calculated from a single-determinant wavefunction. For closed-shell molecules, this leads to the Roothaan equations[16]

$$\sum_{\nu} F_{\mu\nu} c_{\nu i} = \epsilon_i \sum_{\nu} S_{\mu\nu} c_{\nu i}. \qquad (2.5)$$

Here $S_{\mu\nu}$ is the overlap matrix and $F_{\mu\nu}$ is the Fock matrix with elements

$$F_{\mu\nu} = H_{\mu\nu}^{\text{core}} + \sum_{\lambda\sigma} P_{\lambda\sigma} [(\mu\nu \mid \lambda\sigma) - \tfrac{1}{2}(\mu\lambda \mid \nu\sigma)], \qquad (2.6)$$

where $P_{\lambda\sigma}$ is the density matrix

$$P_{\lambda\sigma} = 2 \sum_{i}^{\text{occ}} c_{\lambda i} c_{\sigma i}. \qquad (2.7)$$

$H_{\mu\nu}^{\text{core}}$ is the matrix of the core Hamiltonian (kinetic energy and potential in the field of fixed nuclei), and

$(\mu\nu \mid \lambda\sigma)$ is the two-electron integral

$$(\mu\nu \mid \lambda\sigma) = \iint \phi_{\mu}'(1)\phi_{\nu}'(1) \frac{1}{r_{12}} \phi_{\lambda}'(2)\phi_{\sigma}'(2) d\tau_1 d\tau_2.$$

Using the Gaussian expansions (2.2), all the integrals can be evaluated by the methods originally due to Boys.[17] Equation (2.5) for the LCAO coefficients $c_{\mu i}$ can then be solved in an iterative manner, leading to calculated energies, density matrices, gross orbital and atom populations,[18] and electric dipole moments.

For open-shell molecules, we shall use unrestricted single-determinant wavefunctions, with different molecular orbitals for α and β electrons. The LCAO SCF equations (2.5) then have to be generalized.[19] Such wavefunctions are appropriate for all first-row atoms in their ground states.

All calculations have been carried out by a single-precision FORTRAN-63 program using a CDC-1604A computer (48-bit word). In the generation of two-electron integrals, the basis orbitals ϕ_{μ}' are grouped into K shells (a single $1s$ function) and L shells (one $2s$ and three $2p$ functions). All integrals arising from a particular set of four shells are evaluated together to make maximum use of common information involving the Gaussian α exponents. All integrals are believed to be accurate to 2×10^{-6} hartree. A number of internal tests on the program (recalculation with arbitrarily rotated Cartesian axes and with increased integral accuracy) suggest that the computational error in the calculated total energy is around 10^{-5} hartree for molecules as large as benzene. An external check was attempted by repeating the pyrolle calculation of

[15] R. F. Stewart, J. Chem. Phys. 50, 2485 (1969).
[16] C. C. J. Roothaan, Rev. Mod. Phys. 23, 69 (1951).
[17] S. F. Boys, Proc. Roy. Soc. (London) A200, 542 (1950).
[18] R. S. Mulliken, J. Chem. Phys. 23, 1833, 1841, 2338, 2343 (1955).
[19] J. A. Pople and R. K. Nesbet, J. Chem. Phys. 22, 571 (1954).

TABLE III. Atomization energy with fixed exponents.

Molecule	STO-2G	STO-3G	STO-4G	STO-5G	STO-6G	STO
HF	0.0683	0.0492	0.0476	0.0469	0.0469	0.0469
CO	0.2627	0.1999	0.1930	0.1924	0.1921	0.1921
CH$_4$	0.6250	0.5862	0.5757	0.5761	0.5758	0.5755
C$_2$H$_2$	0.4772	0.4350	0.4209	0.4220	0.4220	0.4214
C$_2$H$_4$	0.7518	0.6935	0.6775	0.6780	0.6778	0.6772
C$_2$H$_6$	1.0523	0.9719	0.9525	0.9529	0.9525	0.9518
BH$_3$	0.4254	0.4079	0.4030	0.4032	0.4031	0.4027
B$_2$H$_6$	0.8652	0.8214	0.8095	0.8092	0.8090	0.8083
NH$_3$	0.3482	0.3085	0.2999	0.3003	0.3000	0.2999
HCN	0.2879	0.2396	0.2270	0.2270	0.2271	0.2266
H$_2$CO	0.4129	0.3491	0.3383	0.3387	0.3384	0.3381

Clementi, Clementi, and Davis.[20] The basis functions listed in this reference appear not to be fully normalized, but when they were multiplied by appropriate normalization factors, a total energy of -207.93065 hartree was found which differs from the value of Ref. 20 by 7×10^{-4}. The origin of this small discrepancy is not clear, but it is unlikely to modify any conclusion significantly.

The average time per two-electron integral varies widely from one molecule to another, depending on the number of hydrogen atoms present. Integrals involving four L shells are evaluated most rapidly, since one set of values of error and exponential functions suffices for $4^4 = 256$ distinct integral values (four functions per shell). Typical times per integral (involving four ϕ functions) on the CDC-1604A are 6, 20, 50, 100, and 180 msec for $K = 2$, 3, 4, 5, and 6, respectively. If different Gaussian exponents α_k are used for the $2s$ and $2p$ atomic orbitals (so that not more than $3^4 = 81$ integrals can be obtained from one set of error and exponential functions) these times are approximately doubled.

III. CONVERGENCE STUDIES WITH FIXED ζ EXPONENTS

In this section we shall investigate the convergence (with increasing size of Gaussian set) of calculated energies, energy differences, and other features of the wavefunctions using fixed ζ exponents. To study this, we have taken a number of molecular-orbital calculations from the literature with a minimal STO basis and have repeated them *using the same ζ exponents* but with the STO basis replaced by STO-2G, STO-3G, STO-4G, STO-5G, and STO-6G. The results are presented in Tables II–V together with the literature STO values.

[20] E. Clementi, H. Clementi, and D. R. Davis, J. Chem. Phys. 46, 4725 (1967).

Table II contains calculated total energies and Table III energies of atomization obtained by subtracting the molecular energy from the energy of a corresponding set of isolated neutral atoms in their ground states, using the same exponents. (The atomic energies are all calculated using an open-shell unrestricted program permitting different orbitals for different spins.)

The principal point to be made about Tables II and III is that the atomization energies converge towards the full STO result much more rapidly than the total energies. Changes in total energy with increasing size of Gaussian representation are of the order of 100 times are great as changes in the atomization energy. This indicates that the error introduced when the Slater-type atomic orbitals are replaced by small Gaussian expansions is very similar in the molecule and in the separate atoms. These errors are probably largely associated with the inadequate behavior of Gaussian functions near the nuclei (e.g., lack of cusps), a region which will not be greatly modified by bond formation. As a result, it is possible to calculate atomization energies close to those for a full STO basis set with only a small Gaussian expansion. For the molecules listed the STO-2G, atomization energies are of the order of 0.05–0.1 hartree greater than the STO values. The STO-3G values are of the order of 0.01 greater. This difference is quite systematic, there being positive contribution of about 0.002–0.003 hartree for each bond in the molecule. Going to the STO-4G level leads to atomization energies much closer to the STO values, the maximum difference now being 0.0013 hartree. For the 5G and 6G bases, convergence continues less rapidly, but it is clearly possible to reproduce the STO result to about 0.001 hartree for all these molecules.

Tables IV and V give the gross atom populations and the calculated electric dipole moments for the same set of molecules, geometries, and exponents. The greatest difference in the populations (between the

TABLE IV. Atomic populations with fixed exponents.

Molecule	Atom	STO-2G	STO-3G	STO-4G	STO-5G	STO-6G	STO[a]
HF	F	9.046	9.143	9.153	9.154	9.154	9.154
	H	0.954	0.857	0.847	0.846	0.846	0.846
CO	C	5.976	5.904	5.889	5.888	5.888	5.889
	O	8.024	8.096	8.111	8.112	8.112	8.111
CH₄	C	6.416	6.524	6.540	6.535	6.534	6.534
	H	0.896	0.869	0.865	0.866	0.867	0.867
C₂H₂	C	6.142	6.187	6.190	6.189	6.188	6.188
	H	0.858	0.813	0.810	0.811	0.812	0.812
C₂H₄	C	6.211	6.276	6.283	6.280	6.280	6.278
	H	0.894	0.862	0.858	0.860	0.860	0.860
C₂H₆	C	6.288	6.367	6.377	6.373	6.373	6.372
	H	0.904	0.878	0.874	0.876	0.876	0.876
BH₃	B	5.115	5.182	5.179	5.174	5.172	5.172
	H	0.962	0.939	0.940	0.942	0.943	0.943
B₂H₆	B	5.240	5.296	5.291	5.286	5.284	5.283
	H	0.923	0.904	0.906	0.907	0.908	0.908
	H(br)	0.913	0.896	0.898	0.900	0.900	0.901
NH₃	N	7.335	7.450	7.471	7.467	7.467	7.467
	H	0.888	0.850	0.843	0.844	0.844	0.845
HCN	C	6.131	6.145	6.141	6.140	6.139	6.139
	N	7.034	7.069	7.077	7.076	7.076	7.075
	H	0.835	0.786	0.782	0.784	0.784	0.784
H₂CO	C	6.137	6.168	6.168	6.164	6.164	6.163
	O	8.007	8.070	8.078	8.079	8.079	8.080
	H	0.928	0.881	0.877	0.878	0.878	0.879

[a] HF, CO, S. I. Fraga and B. J. Ransil, J. Chem. Phys. **34**, 727 (1961); polyatomics except H₂CO, Ref. 3; H₂CO, M. D. Newton (private communication).

Gaussian value and the corresponding STO value) is 0.017 for STO-3G, 0.008 for STO-4G, 0.003 for STO-5G, and 0.002 for STO-6G. The dipole moments are fairly sensitive to the size of the Gaussian expansions, the errors approaching 0.1 D for the 3G set. For both populations and dipole moments, the 2G results show much larger errors.

IV. OPTIMIZED ζ EXPONENTS

It is already well known that the best ζ exponents for atoms in molecules differ significantly from the free atom values.[9] It is therefore desirable to find whether variation of the ζ values in the Gaussian calculations

will lead to the same values as the full STO method. A study of the convergence of this property has been carried out by optimizing the ζ exponents for the K shell of hydrogen and the L shells of heavy atoms for a series of atoms and molecules. These were varied one at a time in steps of 0.01 until a minimum in the energy was found. Usually three or four cycles on the complete set of exponents were necessary. The inner (K) shells for heavy atoms were given fixed ζ values corresponding to free atom values[21] rounded to two decimal places (5.67, 6.67, 7.66, 8.65 for C, N, O, F,

[21] E. Clementi and D. L. Raimondi, J. Chem. Phys. **38**, 2686 (1963).

TABLE V. Dipole moments with fixed exponents.

Molecule	STO-2G	STO-3G	STO-4G	STO-5G	STO-6G	STO[a]
HF	0.496	0.850	0.874	0.880	0.879	0.878
CO	1.104	0.805	0.738	0.731	0.731	0.730
NH_3	1.617	1.742	1.769	1.766	1.765	1.76
HCN	1.858	2.099	2.122	2.122	2.121	2.11
H_2CO	0.545	0.941	1.003	1.008	1.008	1.006

[a] HF, CO, Ref. 1; NH_3, H_2CO, Refs. 3 and 4; correction given in M. D. Newton and W. N. Lipscomb, J. Am. Chem. Soc. **89**, 4261 (1967); HCN, Ref. 3.

respectively). The optimum valence-shell ζ values for free ground-state atoms are given in Table VI and the corresponding total energies in Table VII. For the molecules considered, we have used the standard geometry defined elsewhere.[22] The optimum exponents found are listed in Table VIII. It is clear from these values and those in Table VI that the convergence is good and that exponents optimized with small Gaussian representations are close to those with the exponential functions. The maximum difference between STO-3G and STO-5G results is 0.02.

It is evident from previous studies[9,14,23] and from Table VIII that optimum ζ values for the valence shells of atoms in molecules differ substantially from the values appropriate to free atoms. For hydrogen and carbon in particular, optimized ζ values are uniformly greater than atomic values. It is therefore useful to introduce a new set of *standard molecular ζ values* appropriate for an average molecular environment. The standard we propose is listed in Table IX. These are based on the results in Table VIII and other similar studies. Values for Li, Be, and B are preliminary and are close to optimum values for LiH, BeH_2 (linear), and BH_3. However, they may need modification when more molecules are considered.

Since the standard exponents represent an average of the optimized values of Table VIII, it is useful to compare corresponding energies, populations, and dipole moments so that the error introduced by using standard exponents can be assessed. Such a comparison is displayed for STO-3G in Table X (atomization energies being measured relative to the optimized atomic energies of Table VII). Very similar tables can be constructed at the 4G and 5G levels. These results suggest that exponent optimization is relatively unimportant for the calculation of total or atomization energy, but that the electron distribution and related properties such as the dipole moment are sensitive to the exponents.

[22] J. A. Pople and M. Gordon, J. Am. Chem. Soc. **89**, 4253 (1967).

[23] W. J. Hehre and J. A. Pople, Chem. Phys. Letters **2**, 379 (1968).

V. CONCLUSIONS

On the basis of the results reported in this paper, we may draw the following general conclusions:

(1) The least-squares representation of Slater-type atomic orbitals (STO's) by a small sum of Gaussian functions provides a rapidly convergent method for finding self-consistent molecular orbitals simulating those directly based on an STO set. Although convergence of total energy (with increasing size of Gaussian representation) is limited, that of atomization energy is much more effective.

(2) The smaller Gaussian representations seem suitable as universal atomic-orbital basis functions which may be used in molecular calculations instead of STO sets. The STO-3G set, in particular, is very economical to use and as a minimal ($1s$, $2s$, $2p$) basis, can be applied to quite large organic molecules. Any further economy achieved by using an STO-2G set, however, is hardly worthwhile, since the calculation time is then dominated by the solution of the Roothaan equations (2.5), rather than integral evaluation. Of the three larger Gaussian representations, STO-4G appears

TABLE VI. Optimum exponents for isolated atoms.

Atom	STO-3G	STO-4G	STO-5G	STO-6G	STO
H(^2S)	1.00	1.00	1.00	1.00	1.00
Li(^2S)	0.65	0.64	0.64	0.64	0.64
Be(^1S)	0.97	0.96	0.96	0.96	0.96
B(^2P)	1.28	1.27	1.27	1.27	1.27
C(^3P)	1.60	1.59	1.59	1.59	1.59
N(^4S)	1.93	1.92	1.92	1.92	1.92
O(^3P)	2.24	2.24	2.24	2.23	2.23
F(^2P)	2.56	2.56	2.56	2.56	2.56

TABLE VII. Atomic energies with optimized exponents.

Atom	STO-3G	STO-4G	STO-5G	STO-6G	STO
H(^2S)	−0.49491	−0.49848	−0.49951	−0.49983	−0.50000
Li(^2S)	−7.32823	−7.39185	−7.40971	−7.41536	−7.41848
Be(^1S)	−14.39180	−14.50884	−14.54080	−14.55098	−14.55670
B(^2P)	−24.23160	−24.42216	−24.47226	−24.48829	−24.49717
C(^3P)	−37.22866	−37.51069	−37.58578	−37.60906	−37.62182
N(^4S)	−53.72010	−54.11585	−54.21972	−54.25155	−54.26886
O(^3P)	−73.80425	−74.33740	−74.47555	−74.51749	−74.54011
F(^2P)	−97.98709	−98.68185	−98.85976	−98.91327	−98.94193

TABLE VIII. Optimum exponents for molecules.

Molecule	Atom	STO-3G	STO-4G	STO-5G
HF	F	2.55	2.55	2.54
	H	1.33	1.31	1.32
H_2O	O	2.24	2.23	2.23
	H	1.28	1.26	1.26
NH_3	N	1.95	1.95	1.95
	H	1.25	1.23	1.23
CH_4	C	1.76	1.76	1.76
	H	1.18	1.17	1.16
C_2H_2	C	1.68	1.67	1.67
	H	1.31	1.31	1.31
C_2H_4	C	1.70	1.70	1.70
	H	1.23	1.22	1.21
C_2H_6	C	1.76	1.75	1.75
	H	1.18	1.17	1.17
C_6H_6	C	1.72	1.71	1.71
	H	1.21	1.21	1.21
CH_3F	C	1.79	1.79	1.79
	F	2.56	2.55	2.55
	H	1.19	1.18	1.18
CF_4	C	1.84	1.83	1.84
	F	2.56	2.56	2.55
CH_2O	C	1.76	1.75	1.75
	O	2.25	2.24	2.24
	H	1.21	1.21	1.20

to be suitable for resolving uncertainties which might arise in the application of the smaller sets. Atomization energies, for example, converge quite rapidly in going from STO-2G to STO-4G, gaining almost an order of magnitude for each additional Gaussian. After the STO-4G level, however, the rate of convergence tapers off markedly.

TABLE IX. Standard molecular exponents for Slater-type atomic orbitals.

Atom	K shell	L shell	L shell (Slater atomic value)
H	1.24	•••	•••
Li	2.69	(0.75)	0.65
Be	3.68	(1.10)	0.975
B	4.68	(1.45)	1.3
C	5.67	1.72	1.625
N	6.67	1.95	1.95
O	7.66	2.25	2.275
F	8.65	2.55	2.6

(3) The study of ζ exponents indicates that the optimized values (for one type of atom) very considerably from one molecule to another but are largely independent of the size of the Gaussian representation. A standard set of ζ exponents (Table IX) appears to be sufficient for the calculation of atomization energies, but the charge distribution is significantly modified by ζ optimization.

ACKNOWLEDGMENTS

TABLE X. Comparison of calculations with standard and optimized exponents at STO-3G.

Molecule	Atomization energy		Atomic population			Dipole moment	
	Standard	Optimized	Atom	Standard	Optimized	Standard	Optimized
HF	0.0891	0.0916	F	9.209	9.228	1.285	1.406
			H	0.791	0.772		
H_2O	0.1667	0.1676	O	8.372	8.406	1.689	1.782
			H	0.814	0.797		
NH_3	0.2477	0.2478	N	7.481	7.486	1.648	1.657
			H	0.840	0.838		
CH_4	0.5184	0.5222	C	6.255	6.073		
			H	0.936	0.982		
C_2H_2	0.4062	0.4123	C	6.109	6.182		
			H	0.891	0.818		
C_2H_4	0.6342	0.6356	C	6.127	6.156		
			H	0.937	0.922		
C_2H_6	0.8787	0.8848	C	6.171	6.026		
			H	0.943	0.991		
C_6H_6	1.5486	1.5497	C	6.063	6.051		
			H	0.937	0.949		
CH_3F	0.4674	0.4740	C	6.042	5.831	1.128	0.960
			F	9.153	9.157		
			H	0.935	1.004		
CF_4	0.4038	0.4271	C	5.427	5.326		
			F	9.143	9.169		
CH_2O	0.3298	0.3316	C	5.941	5.833	1.520	1.527
			O	8.186	8.208		
			H	0.937	0.979		

APPENDIX: GENERAL NOTATION FOR GAUSSIAN REPRESENTATIONS OF ATOMIC BASIS FUNCTIONS

Since the replacement of an atomic-orbital basis set by least-squares-fitted combinations of Gaussian functions is likely to have wide utility, it seems worthwhile to propose a general notation for such representations. This should clearly specify the way in which Gaussian exponents are shared between different atomic orbitals. We suggest that the basis sets used in this paper should be described as

$$STO[1s \rightarrow N(1s_g) \mid 2s \rightarrow N(1s_g), 2p \rightarrow N(2p_g)]. \quad (A1)$$

The symbols $1s \rightarrow N(1s_g)$ mean that the $1s$ Slater-type orbital is replaced by a least-squares-fitted set of N $1s$ Gaussian orbitals. Vertical bars (or square end brackets) are used to separate shells or groups of functions which share common Gaussian exponents. This applies to the $2s$ and $2p$ Slater functions in (A1).

If the Slater $2s$ and $2p$ functions are fitted to N Gaussian orbitals independently, as in Refs. 13 and 14, the basis set would be described as

$$STO[1s \rightarrow N(1s_g) \mid 2s \rightarrow N(1s_g) \mid 2p \rightarrow N(2p_g)]. \quad (A2)$$

Mixed representations can be written in a similar way. Thus an expansion of a $3s$ function as a sum of one $1s$ Gaussian and two $3s$ Gaussians would be

$$STO[3s \rightarrow (1s_g, 2(3s_g))]. \quad (A3)$$

Finally, we may note that the same notation may be used for Gaussian representations of other types of atomic orbital. Thus, the expansions of Hartree–Fock atomic orbitals (HFAO) given previously [15] are

$$HFAO[1s \rightarrow N(1s_g) \mid 2s \rightarrow N(1s_g) \mid 2p \rightarrow N(2p_g)],$$

$$HFAO[1s \rightarrow N(1s_g), 2s \rightarrow N(1s_g) \mid 2p \rightarrow N(2p_g)],$$

$$HFAO[1s \rightarrow N(1s_g) \mid 2s \rightarrow N(1s_g), 2p \rightarrow N(2p_g)]. \quad (A4)$$