PCCP

Emergence of Linnett's "Double Quartets" from a Model of "Lewis Dots"

SCHOLARONE™
Manuscripts

Emergence of Linnett's "Double Quartets" from a Model of "Lewis Dots"

Judith Herzfeld*

Department of Chemistry, Brandeis University, Waltham, Massachusetts, USA

ABSTRACT

Chemists routinely explicate molecular structures and chemical reactions in terms of the propensities of semi-classical valence electrons (aka "Lewis dots"). Typically, the electrons are viewed as forming spin pairs and recent efforts to translate this concise and intuitive qualitative picture into an efficient and relatable quantitative model have made good progress. But electrons are not always paired and advanced quantum calculations have shown that this is so even in small diamagnetic species such as dicarbon and benzene. Here we show that the latest semi-classical model for paired electrons can clarify the limitations on pairing simply by dissecting the elements of the interparticle potentials. Although not trained to do so, these elements produce a Linnett-like benzene, with three valence electrons in each CC bond, when the electrons are free to move singly. At the same time, sustaining higher order bonds with independently mobile electrons requires adjustments in the details of the model potentials at short distances. This is addressed with new training data and new forms for the contributions from Coulomb integrals. Although trained on hydrogen and carbon species separately, the combination applied to ethyne

predicts the pairing of spins in the CH bonds and the dispersion of spins in the CC bond that is found in *ab initio* calculations. This adjusted force field is named for Linnett, in appreciation of his insightful double quartet interpretation of the Lewis octet.

INTRODUCTION

Electron pairing has been a compelling and central concept in chemistry for over a century.[1] Before the discovery of electron spin, it rationalized the tendency of atoms to form bonds with even numbers of electrons and, subsequent to the discovery of electron spin, it comported with the diamagnetic nature of the ground states of most molecules. With the advent of computational chemistry, electron pairing justified the convenience of assigning electrons to orbitals in pairwise fashion. Only relatively recently has computational power allowed broadening the consideration of unpaired electrons beyond free radicals, including in such diamagnetic molecules as C_2 and C_6H_6.[2-5] This has led to an appreciative revisiting of Linnett's concept that, not only do electrons of like spin avoid each other, but electrons of unlike spin also tend to separate unless the attractive forces of nuclei cause them to co-localize. With his "double-quartet theory" of octets, Linnett rationalized the paramagnetic ground state of O_2 and the symmetry of benzene without invoking molecular orbitals.[6, 7]

Recent years have seen efforts aimed at a quantitative rendering of the user-friendly semi-classical picture of electrons. Greater progress has been made for paired valence electrons

because they entail fewer particles and spin need not be explicitly considered.[8] In the recent

LEWIS-B sub-atomistic force field,[9] pairwise interactions of valence electron pairs with each

other and with kernels, are able to describe single, double, bridge and bent bonds in linear,

branched and cyclic hydrocarbons, including anionic and cationic states. Beyond predicting

structures and energies, this force field efficiently simulates carbocation addition to a double

bond and cation migration to a neighboring carbon. A critical feature of this force field is

allowing a variable spread for the electron pairs, amounting to a fourth degree of freedom

beyond the three classical translational coordinates. The spread of each valence pair is a telling

feature: since more diffuse electron distributions decrease both the quantum kinetic energy and

the strengths of interactions, the trade-off is such that electrons are least diffuse when their

interactions are most favorable. The lack of this freedom in the older LEWIS force field[10]

restricts its applicability: while it provides an excellent description of the acid-base behavior of

water in the bulk[11, 12] and at surfaces,[13] it does not adequately describe either H_2 or H_2O_2. In

effect, the electron spread that is suitable in the vicinity of a single oxygen kernel (in H_2O in all

its protonation and cluster states), is not suitable in the absence of an oxygen kernel (as in H_2) or

between two oxygen kernels (as in HOOH).

Although LEWIS-B models electron pairs, it contains all the elements relevant to modeling

unpaired electrons with their spin dependent interactions. In particular, development of the

LEWIS-B potentials attended to energy contributions corresponding specifically to individual

Coulomb and exchange integrals for the electrons comprising the electron pairs. Since the

Coulomb integrals apply equally to electrons of like and unlike spin while the exchange integrals apply only to electrons of like spin,[14] the LEWIS-B potentials can, in principle, be deconstructed to provide spin-dependent potentials for single electrons. In the present work, we take this as a starting point, identifying some strengths and limitations of the deconstructed LEWIS-B potential. Then we proceed to address the primary limitation in developing a more adequate LINNETT force field.

MODEL

Under suitable interaction potentials, independently mobile kernels (X, Y, etc.) and valence electrons (e) form molecules that are flexible, anisotropically polarizable, and capable of breaking and making bonds, all without any need for the atom types employed in conventional molecular mechanics where there is no explicit account of the electrons. Here kernels differ from each other only in charge (q_X, q_Y, etc) according to the corresponding element and electrons differ only in having one of two spins (α, β). We model the compact kernels classically, with three translational degrees of freedom (R) and pairwise interactions between them given by (q_X q_Y /$4\pi\varepsilon_o$r). We model the valence electrons semi-classically, with a variable spread (σ) in addition to three translational degrees of freedom (r), and with interactions that deviate from classical at short distances in ways that reflect wave properties and depend on spin. Energy terms involving the electrons arise from four types of integrals:

C_k = contribution from a coulomb integral over the kinetic operator

C_q = contribution from a coulomb integral over the electrostatic operator

X_k = contribution from an exchange integral over the kinetic operator

X_q = contribution from an exchange integral over the electrostatic operator

(It should be noted that, since the electrons are not assigned to orthogonal orbitals, X_k cannot be assumed to be zero.) In developing the potentials corresponding to these integrals, we have been guided by the analytical integrals obtained for electrons occupying floating spherical Gaussian orbitals (FSGOs),[14] as described in the ESI.†

DECONSTRUCTING LEWIS-B

Given pairwise interactions,

$$U_{\text{LINNETT}} = \Sigma_i U_e (\sigma_i) + \Sigma_{i<j} U_{ee} (\sigma_i, r_i ; \sigma_j, r_j) + \Sigma_{i,I} U_{eX} (\sigma_i, r_i; R_I) + \Sigma_{I<J} U_{XY}(R_I; R_J) \quad (1)$$

where the single particle term $U_e = C_k$. On the other hand,

$$U_{\text{LEWIS-B}} = \Sigma_i U_V (\sigma_i) + \Sigma_{i<j} U_{VV} (\sigma_i, r_i ; \sigma_j, r_j) + \Sigma_{i,I} U_{VX} (\sigma_i, r_i; R_I) + \Sigma_{I<J} U_{XY}(R_I; R_J) \quad (2)$$

where V represents a valence pair.[9] The classical U_{XY} is identical in the two constructs and the

attraction to a kernel for a single electron is just half of that for an electron pair, such that $U_{eX} = U_{VX}/2$. Relating the electron-electron interactions is more complicated. In LEWIS-B,

$$U_{VV} = 4C_{\text{æ}} + 2X_{k} + 2X_{q} \tag{3}$$

taking into account that, of the four electron-electron interactions across two valence pairs, only

the two between electrons of like spin have non-zero exchange integrals. Pulling these apart for

LINNETT, yields

$$U_{\text{ee-unlike}} = C_{\text{æ}} \tag{4}$$

$$U_{\text{ee-like}} = C_{\text{æ}} + X_{k} + X_{q} \tag{5}$$

for the interactions between electrons of unlike and like spins respectively.

Figure 1 illustrates the behavior of the deconstructed LEWIS-B. As expected, the tetrahedral

quartets of α and β electrons roughly coincide in CH_4: the attraction of the electrons to the

internuclear regions overcomes the relatively weak tendency of unlike-spin electrons to separate.

On the other hand, in benzene where there are fewer internuclear regions to occupy, the

tetrahedral quartets do not coincide. At a given carbon, each quartet comprises an electron in the

CH bond, an electron in one of the two CC bonds and the remaining two in the other CC bond.

This corresponds to Linnett's description of fractional bond order without resonance[7] and to the electron configuration for benzene obtained by dynamic Voronoi Metropolis sampling.[5] It is noteworthy that the alternation of spins from one CC bond to the next provides favorable quartet arrangements for the even number of carbons in benzene and would be unfavorable for rings with odd numbers of carbons. Furthermore, the arrangements of electrons around each carbon would be increasingly distorted from tetrahedral in rings of fewer than five or more than six carbons.
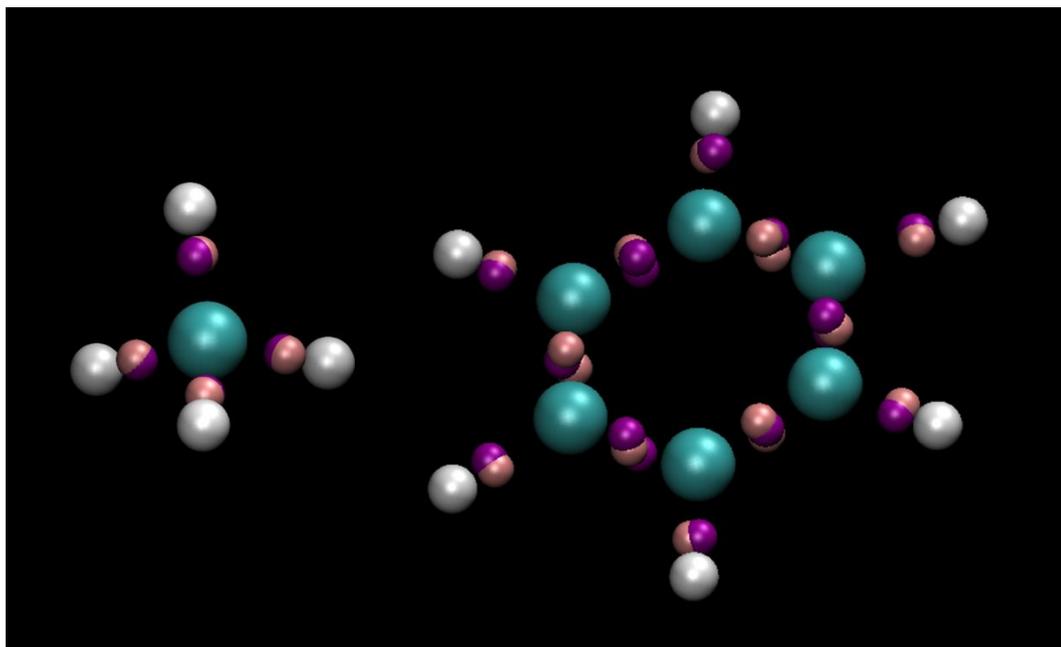

Figure 1. Energy minimized structures for CH_4 and C_6H_6 under the deconstructed LEWIS-B potential. Rendered with VMD, protons are white, carbons are teal, and electrons are pink or magenta depending on their spin.

It is important to appreciate that this electron distribution in benzene is predicted by a force field

that was developed for and trained on paired electrons: the LEWIS-B model comprised only

kernels and valence electron pairs, and the force field training set included structural and

thermodynamic data exclusively for molecules with paired electrons. Thus, the key for the

separation of electrons of unlike-spin is seen to be in a correct understanding of the physical

origins of different terms in the force field.

While these results for CH_4 and C_6H_6 are encouraging, the deconstructed LEWIS-B potential

fails for higher order bonds: with this set of potentials, four unpaired electrons are unable to

coexist in the CC region of ethene, not to mention six in the CC region of ethyne. This is not

surprising given that the training set for the LEWIS-B potentials did not cover short distances

between electrons because, with their –2 charges, valence pairs never get very close to one

another. A new training set is clearly required to obtain a more adequate LINNETT force field.

LINNETT TRAINING SET

Despite the relatively few adjustable parameters involved for any given choice of potential

forms, it is impractical to optimize all the potentials simultaneously. In addition, we want to be

able to test different forms for the Coulomb contributions to different interactions. Thus, our

strategy, as for LEWIS-B, is to optimize in stages

    — first for U_{eH} using the training set in Table S1 of the ESI,†

— then for $U_{ee\text{-unlike}} = C_{\mathcal{e}}$ (given U_{eH}) using the training set in Table S2 of the ESI,†

— then for $U_{ee\text{-like}} = C_{\mathcal{e}} + X_k + X_q$ (given U_{eH} and $C_{\mathcal{e}}$) using the training set in Table S3 of the ESI,†

— and finally for U_{eC} (given $U_{ee\text{-unlike}}$ and $U_{ee\text{-like}}$) using the training set in Table S4 of the ESI.†

Given the small species involved, the training data comprise mostly ionization, spin excitation and bond energies. For structural data there are just a few bond lengths and two bond orders, supplemented by an expectation of zero dipole moment for all species other than H_2^+ (because it separates into H and H^+ when stretched). Significantly omitted is any information about the pairing or unpairing of electrons. For higher order bonding, the expectation is only that three electrons of each spin will be located in the CC bond of C_2 (Table S4 of the ESI†), whether they are paired or not.

CURRENT BEST SET OF POTENTIALS

The process of evaluating potential forms, including optimizing parameters for each potential form, was as described previously.[9] The best final fits to the training data (shown in the last columns of Tables S1 – S4 of the ESI†) were obtained with the parameter values shown in Table 1 for the potential forms

$$U_e = (3/2)\,(\hbar^2/m_e)\,\sigma^{-2} \tag{6}$$

$$U_{eH} = -\ (e^2/4\pi\varepsilon_o)\ \sigma_{eff}^{-1}\ [(\tau_H + \rho_H r_{eff}\ )^{vH} + (1 - \rho_H^{vH})\ r_{eff}^{\ vH}\ ]^{-(1/vH)} \tag{7}$$

$$\text{with } \sigma_{eff} = \sigma + \sigma_H,$$

$$U_{eC} = -\ 4\ (e^2/4\pi\varepsilon_o)\ \sigma_{eff}^{-1}\ [[(\tau_C^{\ 2})^{(vC1+vC2)} + \chi_C\ (r_{eff}^{\ 2})^{\ vC2} + (r_{eff}^{\ 2})^{\ (vC1+vC2)}\ ]^{-(1/2(vC1+vC2))}$$

$$\tag{8}$$

$$\text{with } \sigma_{eff} = \sigma + \sigma_C,$$

$$X_k = \ (\kappa_k/\ \sigma_{eff}^{\ 2})\ (\hbar^2/2m_e)\ [3(\gamma^2-1) + (r_{eff}^{\ 2}/\lambda_k^{\ 2})]/\ [\gamma^3 \exp(r_{eff}^{\ 2}/\lambda_k^{\ 2}) - 1] \tag{9}$$

$$\text{with } \ \gamma = [(\sigma_i/\sigma_j) + (\sigma_j/\sigma_i)]/2$$

$$X_q = -\ (\kappa_q/\ \sigma_{eff})\ (e^2/6\pi^{3/2}\varepsilon_o)\ [1 - (r_{eff}^{\ 2}/\lambda_a^{\ 2})\ ]\ \exp\ [-(r_{eff}^{\ 2}/\lambda_q^{\ 2})]\ . \tag{10}$$

and

$$C_{ee} = (e^2/4\pi\varepsilon_o)\ \sigma_{eff}^{-1}\ [(\tau_e^{\ 2})^{\ (ve1+ve2)}\ \text{erfc}\ (r_{eff}/\lambda_e) + \chi_e\ (r_{eff}^{\ 2})^{\ ve2} + (r_{eff}^{\ 2})^{\ (ve1+ve2)}\ ]^{-(1/2(ve1+ve2))}$$

$$\tag{11}$$

$$\text{with } \sigma_{eff} = (\sigma_i^2 + \sigma_j^2)^{1/2}.$$

where, in each case, the scaled distance $r_{eff} = (r/\sigma_{eff})$.

Table 1. Values of optimized potential parameters (dimensionless except as otherwise noted)

	value
U_{eH}	
τ_H	0.57751
ρ_H	0.25111
v_H	15.1087
σ_H (Å)	(0.00000)
U_{eC}	
τ_C	1.02804
χ_C	1.68769
v_{C1}	3.35670
v_{C2}	0.69819
σ_C (Å)	0.02464
C_{ee}	
τ_e	0.74804
χ_e	107.616
λ_e	0.04614
v_{e1}	3.57550
v_{e2}	8.92264
X_k	
κ_k	5.77993

λ_k	0.63039
X_q	
κ_q	2.70606
λ_a	1.87274
λ_q	1.06718

The quality of the fit to the training data is illustrated in Figure 2. The correlations between

model predictions and reference data are generally better for hydrogen than for carbon. This may

be due to the sequence of training. The weakest energy correlations are for the neutral C atom

where the electron affinity is overestimated by more than a factor of two and the energy of spin

excitation is underestimated by about a factor of two. The weakest bond length correlation is for

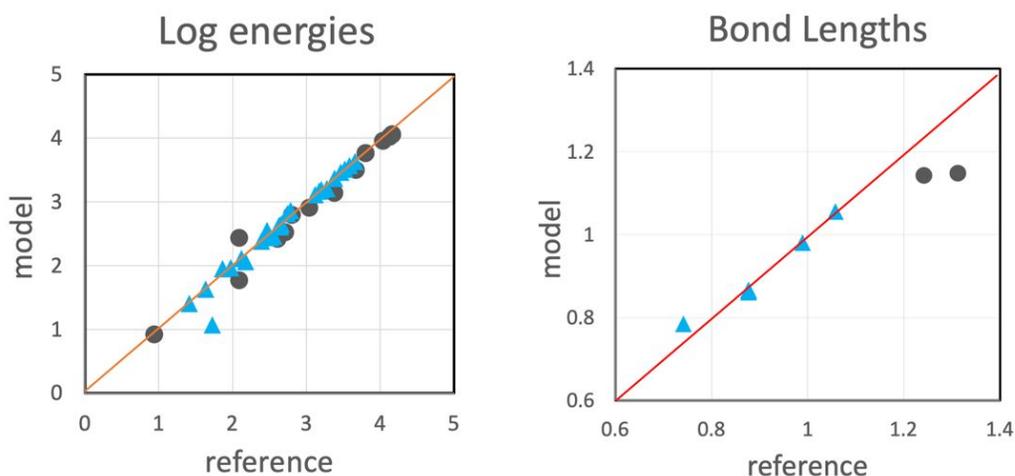the dicarbons; the bond lengths are not only too short, but are hardly sensitive to spin state.

Figure 2. Training results for carbon species (black circles) and hydrogen (blue triangles). Left: the

logarithms of the energies in the training set (absolute values in kJ/mol). Right: the bond lengths in the

training set (in Å).

Several features of these potentials are noteworthy.

— In general, the optimized parameter values conform to expectations. The scaling parameters,

τ, κ and λ are all reasonably close to 1.0, suggesting that the analytical results for FSGO's

provided a reasonable starting point. The λ values are quite close to those for LEWIS-B indicating similar decay distances for the exchange energies. However, the κ values are larger, indicating a stronger contribution from exchange, especially the electrostatic portion. As for our LEWIS• potential (an early potential for single semi-classical valence electrons),[15-17] the parameter σ_H was initially assumed to be zero because the H kernel has no core electrons. In subsequent testing, no improvement was obtained by allowing non-zero values for σ_H. On the other hand, good fits required $\sigma_C \neq 0$. However, the magnitude is small, as befits a representation of the core electrons of the carbon kernel.

— The contributions of different Coulomb integrals are more effectively described by different potential forms. For U_{eH}, the simplest form was required to obtain well-determined parameter values. On the other hand, for C_{ee} to fit the range of data in its training set (Table S2 of the ESI†) required a non-monotonic polynomial in r_{eff} (corresponding to the small value of λ_e).

— The small value of λ_e in C_{ee} is intriguing. For a fixed σ_{eff}, it gives C_{ee} a non-monotonic dependence on r, rising sharply at very small r before declining more gradually with increasing r. On the other hand, the dependence of C_{ee} on σ_{eff} for a given r remains monotonic, although accentuated for large σ_{eff}, especially for small r. These trends suggest that the form of the potential reflects changes in the shapes of two electron clouds at close approach, such that the meaning of a given value of σ_{eff} changes at small r.

— The FSGO-based exchange forms used here (and in LEWIS-B) are very different from those used in the more heursitic LEWIS•.[15-17] In particular, the curvature of X_k at short distances is opposite in the two cases. For the present potentials, the positive curvature at short r_{eff} favors

equal spacing of three electrons of like spin and similar diameter.[18] On the other hand, the

negative curvature at all distances in LEWIS• favors clumping. This difference probably

explains why the very rough potential energy surface that afflicts LEWIS• is not observed with

the present potential.

VIRIAL THEOREM

A test of the validity of the distinction made between electrostatic contributions to the energy (C_q

and X_q) and kinetic contributions (C_k and X_k) is adherence to the virial theorem for electrostatic

potentials: $T = -V/2$, where T and V represent the total kinetic and potential energies,

respectively. Figure 3 shows the values of $-V/T$ for equilibrated molecules in the training set.

The equilibrated hydrogen species give values very close to the expected 2, while the stretched

and compressed diatomics (not shown) deviate markedly as expected. These results for hydrogen

species indicate that our U_e, U_{eH} and U_{ee} potentials (in Eqn. 1) distinguish appropriately between

kinetic and electrostatic contributions. For all the carbon species, the values of $-V/T$ deviate

somewhat from 2. The worst two cases among the monatomics are for the highest spin species,

^5C and $^4C^{+1}$. We found that the carbon deviations can be cured by setting σ_C equal to zero, like

σ_H, but that results in an otherwise less satisfactory description of the molecules. The question is

whether there is some other reasonably simple way of accounting for core electrons that is more

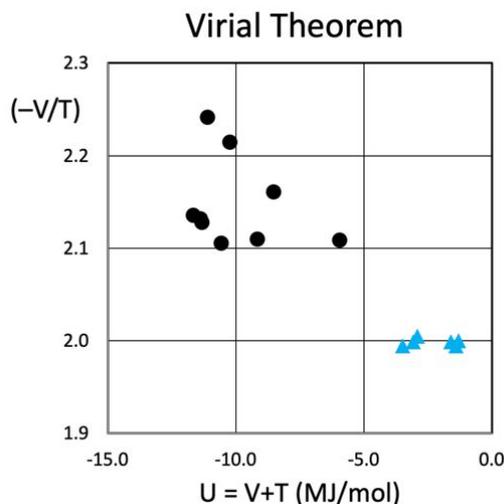compatible with the virial theorem.

Figure 3. Adherence to virial theorem by the total potential energies (V) and total kinetic energies (T) of the individual equilibrated species in the training set. For both of the dicarbons (not shown in order to avoid compressing the horizontal scale) equilibrated $-V/T = 2.23$.

STRUCTURES

Figure 4 shows minimum energy structures for some of the species in the training set. The hydrogen species in the top panel show that the electrons are held most tightly in the positively charged H_3^+ and most loosely in the negatively charged H^-. This is because tighter clouds enhance electrostatic interactions and the tradeoff against higher kinetic energy is more favorable in H_3^+ than in H^-. Although the electrons are relatively tightly held in H_3^+, the advantage of the in-plane position is not sufficient to cause the two opposite spin electrons to co-localize. In H^-, where the two electrons are co-localized, one electron cloud is much more diffuse than the other because its domination of σ_{eff} (see Eqn. 11) allows it to mitigate the electron-electron repulsion

with minimum overall sacrifice of the electron-kernel attractions. From an *ab initio* point of

view, this consistent with some mixing of the 2s state for one of the electrons.

The middle panel of Figure 4 shows the three spin states of the neutral carbon atom. Both the

sizes and positions of the electron clouds are telling.

— The force field correctly describes the non-monotonic spin sequence triplet → singlet →

quintet with increasing energy. Evidence for this energy sequence can be seen in the electron

cloud sizes. In the ground state, one electron is compact and co-localized with the kernel (such

that it is visible here only as a thin halo around the kernel). In the first excited state, one electron

remains compact and co-localized with the kernel, however another is diffuse. Finally, in the

highest energy spin state, two electrons are relatively diffuse while none are compact or co-

localized with the kernel.

— In the triplet state, three electrons of one spin are situated roughly according to the major

lobes of sp^2 orbitals while the other electron is centered on the kernel as in an s orbital. In the

singlet state, the two electrons of each spin are situated with symmetries corresponding to the

major lobes of sp orbitals, but displaced to optimize cloud sizes. In the quintet we effectively

have an isolated Linnett quartet: all four electrons are of the same spin and roughly tetrahedrally

disposed, as for the major lobes of sp^3 orbitals.

The bottom panel of Figure 4 shows the singlet and triplet states of dicarbon. In both cases,

electrons in the CC bond are unpaired as found by ab initio methods.[2, 3] In the singlet, a rough

tetrahedron of α electrons surrounds one carbon and a rough tetrahedron of β electrons surrounds

the other. In the triplet, a rough tetrahedron of α electrons surrounds each carbon and a rough

triangle of β electrons is situated in the midplane. An interesting feature of these molecules is

that the triangles of α and β electrons at the centers of these molecules are not only rotated

relative to one another around the C-C axis, but also translated relative to one another

perpendicular to the C-C axis, with the electron of each spin that is closer to the C-C axis more

compact. Another way of looking at this central configuration is as two chevrons, each with two

electrons of one spin and one of the other. In this view, the triple bond looks like a doubling of

the order 1.5 bond in benzene (see Figure 1). We see this chevron pattern with every potential
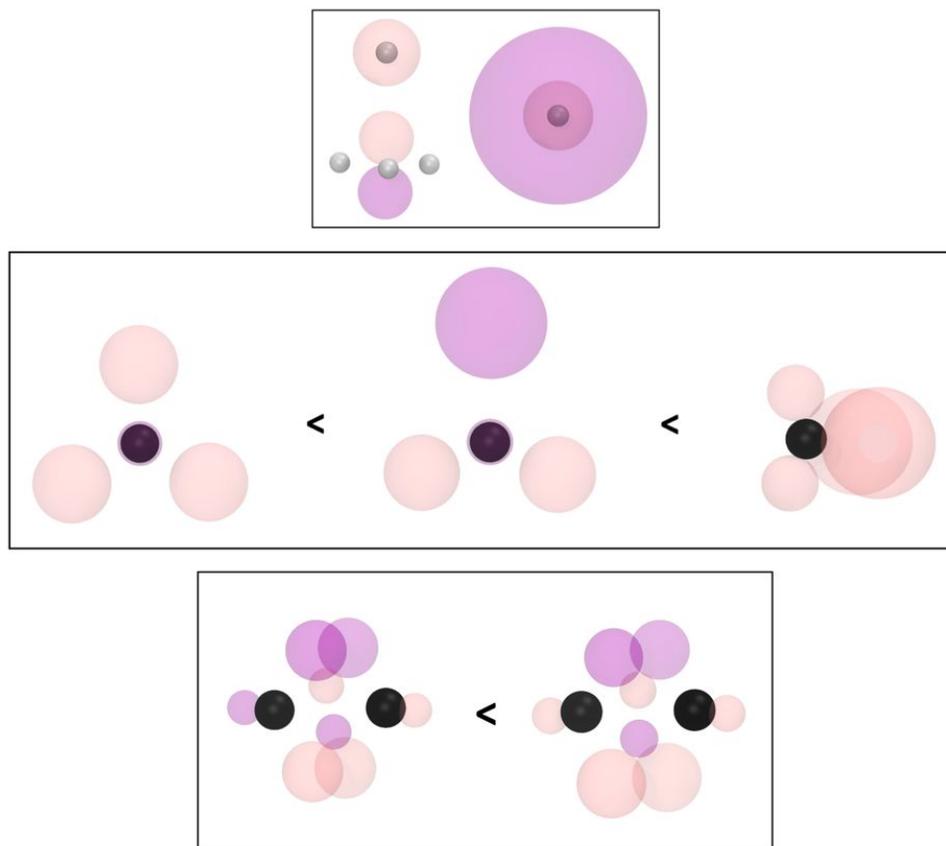
that sustains triple bonds.

Figure 4.   Energy minimized structures of some species in the training sets. Top: H, H_3^+ and H^-. Middle: ^3C, ^1C and ^5C. Bottom: $^1C_2$ and $^3C_2$. Rendered with VMD, protons are white with radius 0.1 Å, carbon kernels are black with radius 0.2 Å, and electrons are pink or magenta according to spin, with radius σ/3. The electron scaling is chosen to clarify spatial relationships even though it erroneously suggests, at first sight, that there are gaps between the densities for different electrons and that the electron density does not fully enclose the kernels. It should be noted that the electrons are rendered with a constant transparency that cannot reflect the lower density that is associated with greater spread.

Figure 5 shows structures for two species outside the training set.

   — Symmetric H_3 is the transition state in the displacement reaction $H + H_2 \rightarrow H_2 + H$. The force field correctly predicts a linear conformation. Bent conformations seen with other

potentials indicate that the present force field does a better job of balancing electron-electron

repulsions vs. electron-kernel attractions over the two bonds. However, the predicted H_3 is too

stable and its bonds are too short vs. ab initio CI results.[19] This suggests that it would be good to

include these features of symmetric H_3 in future training sets.

   — The structure of C_2H_2 shows the expected coincidence of α and β electrons in the CH bonds

and the expected, C_2-like splaying of the α and β electrons in the middle of the molecule. Since

the potentials were trained on hydrogen and carbon species separately, the successful joint

application to a hydrocarbon supports the underlying premises of the model.
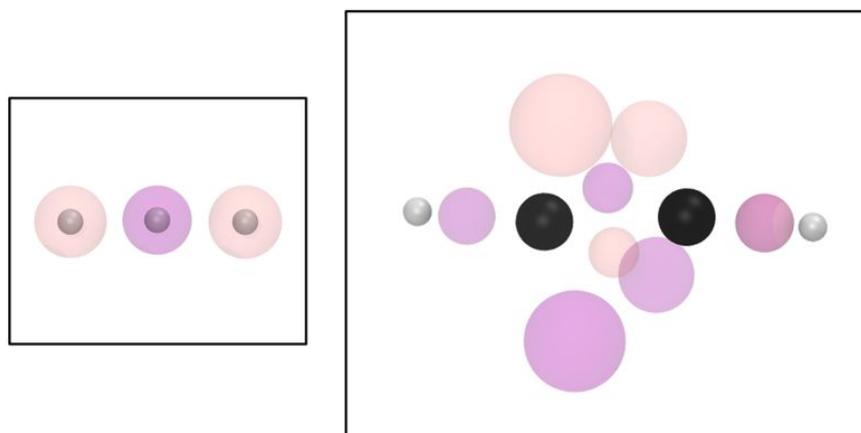

Figure 5. Energy minimized structures for symmetric H_3 and C_2H_2, rendered as in Figure 4. In C_2H_2, the

overlap between the two electrons in each of the CH bonds is so complete that they are impossible to

distinguish in this rendering. Two types of electrons are found in the CC bond: one type occupies the two

relatively tight clouds located relatively close to the C-C axis and the other type occupies the four more

diffuse clouds located further from the C-C axis. The appearance of differences among the latter is due to

perspective.

Larger validation species would obviously be of great interest. However, MD simulations are required, as MC is not effective with more particles at the particle densities involved in these models, especially when many have spread as a fourth degree of freedom. We have used MC to explore different potential forms because this does not require deriving forces for each potential. But, with settled potentials, a switch to MD would allow broader exploration of chemical space.

CONCLUSIONS

Six decades ago, Linnett argued that Lewis' octet should be regarded as a "double quartet" with electrons of like spin tending to form tetrahedra that will only coincide to form four electron pairs under special conditions.[6, 7] The result for noble gas atoms is electrons of each spin on alternating vertices of a cube, just what Lewis' cubical atom would have looked like if it had been informed by the subsequent discovery of electron spin. Attention to the vertices, edges and faces of the constituent tetrahedra not only allows for the full range of single, double and triple bonds, but also paramagnetic ground states (e.g., in dioxygen) and fractional bonds without resonance (e.g., in benzene).

Here we have shown that this qualitative semi-classical picture emerges unprompted from a quantitative model that attends closely to the physical origins of different contributions to electron-electron interaction energies. This is intuitively satisfying as electrons should repel each

other, even when they have unlike spin. In fact, it is a bit concerning that most students of introductory chemistry readily accept the pairwise association of electrons without explanation of when and how electrostatic repulsions are overcome. And without understanding how pairing comes about, students are not prepared to reason semi-classically (i.e., without orbitals) about the exceptions. An example relevant to organic chemistry instruction is the separation and counter propagation of α and β spins in the Diels-Alder reaction.[20]

As we have seen previously,[9, 15-17] in addition to non-classical interaction potentials, semiclassical electrons require a degree of freedom for spread, beyond the classical three for translation. In addition, considerable care is required as to the form of the potentials for interactions at the very short electron-electron distances that obtain in higher order bonds. This called for new training data and exploration of potential forms not considered previously. Persistence with closed-form potentials (vs. machine learned potentials[21]) is rewarded by their connection to physical insight and chemical intuition. Closed-form potentials also involve optimization of relatively few parameters.

Of course, force fields can always be improved with new potential forms and new training sets. For example, training could include properties of symmetric H_3 and compliance with the virial theorem. For C_2, one can remove the assumption that the CC bond in the triplet ground state is analogous to that found by advanced quantum theory for the singlet (i.e., involves three electrons of each spin). In particular, a CC bond comprising three electrons of one spin and just one of the

other (leaving an electron pair beyond each C) might be more consistent with spectroscopic findings.

Assumptions can also be reconsidered for the potential forms. Here we have assumed that kernels have frozen cores and interact with each other in a strictly classical manner. We have also assumed a Hamiltonian that includes only kinetic and electrostatic operators, ignoring any spin interactions. Furthermore, within that construct, we have required our pair potentials to compensate for the omission of potentially non-negligible three-body terms that arise from some electrostatic exchange integrals.

It would be of particular interest to extend the work in several other directions. An obvious one is to develop potentials for other elements. For example, with oxygen one can model water and broaden the range of organic molecules. Finally, with potentials proven on small molecules, it becomes worthwhile to derive forces for MD simulations of larger molecules and reactions.

† ELECTRONIC SUPPLEMENTARY INFORMATION

Trial potential functions

Tables of training data with corresponding model predictions.

ACKNOWLEDGMENTS

REFERENCES

1. Lewis, G. N., The Atom and the Molecule. *J. Am. Chem. Soc.* **1916**, *38* (4), 762-785.
2. Shaik, S.; Rzepa, H. S.; Hoffmann, R., One Molecule, Two Atoms, Three Views, Four Bonds? *Angew. Chem. Int. Ed.* **2013**, 3020-3033.
3. Liu, Y.; Frankcombe, T. J.; Schmidt, T. W., Chemical bonding motifs from a tiling of the many-electron wavefunction. *Phys.Chem.Chem.Phys.* **2016**, *18*, 13385-13394.
4. Liu, Y.; Frankcombe, T. J.; Schmidt, T. W., Electronic Wavefunction Tiles. *Aust. J. Chem.* **2020**, *73*, 757-766.
5. Liu, Y.; Kilby, P.; Frankcombe, T. J.; Schmidt, T. W., The electronic structure of benzene from a tiling of the correlated 126-dimensional wavefunction. *Nature Communications* **2020**, *11*, 1210.
6. Linnett, J. W., Valence-Bond Structures - New Proposal. *Nature* **1960**, *187*, 859-861.
7. Linnett, J. W., A modification of the Lewis-Langmuir octet rule. *J. Am. Chem. Soc.* **1961**, *83* (12), 2643-2653.

8.   Bai, C.; Kale, S.; Herzfeld, J., Chemistry with semi-classical electrons: reaction trajectories auto-generated by sub-atomistic force fields. *Chem. Sci.* **2017,** *8* (6), 4203-4210.

9.   Li, J.; Song, X.; Li, P.; Herzfeld, J., A Carbon is a Carbon is a Carbon: Orbital-free Simulations of Hydrocarbon Chemistry without Resort to Atom Types. *J. Phys.Chem. A* **2022,** *in press.*

10.   Kale, S.; Herzfeld, J., Natural polarizability and flexibility via explicit valency: The case of water. *J. Chem. Phys.* **2012,** *136,* 084109.

11.   Bai, C.; Herzfeld, J., Special Pairs Are Decisive in the Autoionization and Recombination of Water. *J. Phys. Chem. B* **2017,** *121* (16), 4213-4219.

12.   Kale, S.; Herzfeld, J., Proton defect solvation and dynamics in aqueous acid and base. *Angew. Chem. Int. Ed.* **2012,** *51* (44), 11029-11032.

13.   Bai, C.; Herzfeld, J., Surface Propensities of the Self-Ions of Water. *ACS Cent. Sci.* **2016,** *2* (4), 225-231.

14.   Herzfeld, J.; Ekesan, S., Exchange potentials for semi-classical electrons. *Phys. Chem. Chem. Phys.* **2016,** *18,* 30748-30753.

15.   Ekesan, S.; Herzfeld, J., Pointillist rendering of electron charge and spin density suffices to replicate trends in atomic properties. *Proc. R. Soc. A* **2015,** *471* (2181), 20150370.

16.   Ekesan, S.; Kale, S.; Herzfeld, J., Transferable pseudoclassical electrons for aufbau of atomic ions. *J. Comput. Chem.* **2014,** *35* (15), 1159-1164.

17.   Ekesan, S.; Lin, D. Y.; Herzfeld, J., Magnetism and Bond Order in Diatomic Molecules Described by Semi-Classical Electrons. *J. Phys. Chem. B* **2016,** *120,* 6264-6269.

18.   Herzfeld, J.; Song, X.; Li, J.; Li, P., Bending the Curve: Molecular Manifestations of Electron Antisymmetry. *ChemistryOpen* **2021,** *10,* 1-6.

19.   Liu, B., Ab initio potential energy surface for linear H_3. *J. Chem. Phys.* **1973,** *58,* 1925-1937.

20.   Liu, Y.; Kilby, P.; Frankcombe, T. J.; Schmidt, T. W., Calculating curly arrows from ab initio wavefunctions. *Nature Communications* **2018,** *9,* 1436.

21.   Cools-Ceuppens, M.; Dambre, J.; Verstraelen, T., Modeling Electronic Response Properties with an Explicit-Electron Machine Learning Potential. *J. Chem. Theory Comput.* **2022,** *18,* 1672-1691.