Charge Distributions and Chemical Effects. IV. On a Criterion for Selecting a Theoretical Method for the Study of Molecular Properties Involving Charges

S. Fliszár,* G. Kean, and R. Macaulay

Contribution from the Department of Chemistry, University of Montreal, Montreal, Quebec, Canada. Received December 17, 1973

Abstract: The charge distributions in alkanes, derived from the inductive Taft-like equations $q_H = a\sigma^* + b$ (for R–H compounds) and $q(CH_3) = a\sigma^*$ (for R–CH$_3$), can be expressed in terms of a parameter n ($= -10/3a$), in a scale of relative units where $q_C(C_2H_6) = 1$. Multiplication of relative charges by a unity factor β yields electron densities in charge units. A comparison of these "inductive" charge distributions with quantum mechanical results from semiempirical and partially or nonoptimized *ab initio* methods reveals an overall agreement which is satisfactory. With fully optimized *ab initio* (STO-3G) results the agreement is perfect. The set of "inductive" charge distributions represents, hence, a summary of sets of theoretical charge distributions obtained from the various methods, each of which can be generated by inserting the appropriate n and β values in the "inductive" equations. This illustrates the flexibility of the inductive approach. From the conclusions that (i) any choice of a particular theoretical method reduces to choosing a particular n value and (ii) that the relative scaling of the charge densities depends on n, then it follows that in any theoretical study of a molecular property involving charges attention should be given to the adequacy of the particular n value corresponding to the choice of the theoretical method.

One of the most popular (though ill defined) concepts, that of charge distributions in organic molecules, has received considerable attention, in part because of the interest in charges for the discussion of chemical problems. Present theoretical results moderate this interest, however, because quantum chemical results are not free from difficulties, some of which are quite disturbing.

Firstly, there is an uncertainty about the C–H bond polarity in alkanes. Most of the MO calculations indicate (*e.g.*, in methane) a C$^-$–H$^+$ polarity, but INDO[1] results indicate the reverse, *i.e.*, a C$^+$–H$^-$ polarity.[2] The latter result is also supported by Bader's calculation on methane.[3] In favor of the C$^-$–H$^+$ polarity, one may argue that *ab initio* calculations consistently support this result. STO-3G calculations, for example, indicate that the C atom in methane carries a net charge of -72 melectrons;[4] "better" *ab initio* results, *i.e.*, those

derived using larger Gaussian basis sets, indicate, however, much larger charge separations than those derived using a minimal basis set. A consideration of André's 7s3p/3s results[5] (q_C in CH$_4$ = -0.79 electron) or Allen's results[6] (q_C = -1.072 electron in CH$_4$) then raises the questions: (i) why do alkanes with such positive H atoms never form hydrogen bonds, or (ii) why are *ab initio* results for charge distributions less credible for (good) large Gaussian basis sets than those obtained from the less sophisticated 3G calculations?

Another difficulty which has been noticed is that MO calculations indicate the phenyl ring in toluene to be electron poorer than the phenyl group in benzene, which is contrary to chemical expectation. For example, Hoffmann's extended Hückel calculations[7] indicate a net charge for the C$_6$H$_5$ group of -101 melectrons in benzene and of -21 melectrons in toluene. Pople's *ab initio* results[8] indicate the same trend for the charge

(1) J. A. Pople, D. L. Beveridge, and P. A. Dobosh, *J. Chem. Phys.*, **47**, 2026 (1967); R. N. Dixon, *Mol. Phys.*, **12**, 1397 (1967).

(2) D. R. Salahub and C. Sándorfy, *Theor. Chim. Acta*, **20**, 227 (1971).

(3) R. F. W. Bader and H. J. T. Preston, *Theor. Chim. Acta*, **17**, 384 (1970).

(4) W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.*, **51**, 2657 (1969).

(5) J. M. André, P. Degand, and G. Leroy, *Bull. Soc. Chim. Belg.*, **80**, 585 (1971).

(6) J. E. Williams, V. Buss, and L. C. Allen, *J. Amer. Chem. Soc.*, **93**, 6867 (1971).

(7) R. Hoffmann, *J. Chem. Phys.*, **39**, 1397 (1963).

(8) W. J. Hehre and J. A. Pople, *J. Amer. Chem. Soc.*, **92**, 2191 (1970).

in the C_6H_5 group: -37 melectrons in toluene, and -51 melectrons in benzene.

In order to study the origin of these difficulties we proceed in two steps. (I) In this paper we summarize and rediscuss some aspects of previous results[9-11] and lend them additional support. Primarily, this is an analysis of the similarities in the charge distributions calculated by various methods; once these similarities are clearly recognized, it becomes easy to point at a specific reason which causes the divergence in the theoretical charge distributions. (2) The accompanying paper discusses this specific reason and suggests an answer to the difficulties outlined above.

Inductive Effects

In the interpretation of the chemical behavior of organic molecules, important arguments are related to changes in electron densities brought about by substitution. The study[9-11] of the distribution of electrons in a series of simple organic compounds has revealed several interesting features. Particular attention was given to the fact that in simple alkanes R–H and R'–CH_3 the net charges q_H and $q(CH_3)$ are related in a simple manner to Taft's polar σ^* constants[12] by the equations

$$q_H = a\sigma^*_R + b \qquad (1)$$

$$q(CH_3) = a\sigma^*_{R'} \qquad (2)$$

which enable the back-calculation of charges from the knowledge of the polar σ^* constants. These results indicated that Taft's polar constants adequately reflect the inductive effects in terms of electron release (or withdrawal), as calculated by various quantum-chemical approaches,[13] thus providing an important and useful link between quantum– and physical–organic chemistry, two very rich bodies of information.

Charge Distributions Derived from Equations 1 and 2

The back-calculation of charge densities using eq 1 and 2 can be made as follows. Let β = net charge on C in ethane, i.e.

$$\beta = q_C(\text{ethane}) \qquad (3)$$

be chosen as reference for the calculation of charge densities. If the charges are to be expressed in *relative units*, one may well choose $q_C(\text{ethane}) = +1$ as reference for this relative scale.[17]

The quantity a can be expressed as follows, for convenience,[18] in terms of a new variable n, i.e.

$$a = -10/3n \text{ (relative units)} \qquad (4)$$

$$a = -\beta 10/3n \text{ (charge units)} \qquad (5)$$

From eq 1 is is clear that a must be positive in order to satisfy the Taft order of electron releasing abilities tert-C_4H_9 > ... > methyl, since $\sigma^*(\text{tert-butyl})$ = -0.300 and $\sigma^*(CH_3)$ = 0, because only then will a hydrogen atom be electron richer when attached to a tert-butyl group (H_{tert} in isobutane) than in methane. This means that in eq 5, β and n are opposite in sign. Since β has the same sign as q_C in ethane, *the Taft order requires that n is positive if the polarity is C^--H^+ and negative for C^+-H^-.* Most of the results presented in the next section correspond to $n > 0$ with C^--H^+. The INDO results[2] are an exception among the theoretical charge distributions since they correspond to $n = -2$, with a positive C atom in ethane.

With $q_C(\text{ethane}) = \beta$, $q_H(\text{ethane})$ is necessarily $-\beta/3$. By means of eq 1, using $\sigma^*(C_2H_5) = -0.100$, it follows that

$$b = -\beta(n + 1)/3n \text{ (charge unit)} \qquad (6)$$

Insertion of the appropriate σ^* values[19] (Table I) in

Table I. Polar σ^* Constants

R	σ^*	R	σ^*
CH_3	0	i-C_3H_7	-0.190
C_2H_5	-0.100	sec-C_4H_9	-0.210
n-C_3H_7	-0.115	$(C_2H_5)_2CH$	-0.225
n-C_4H_9	-0.124	$tert$-C_4H_9	-0.300
i-C_4H_9	-0.129	$CH_3CH=CHCH_2$	0.130
neo-C_5H_{11}	-0.151	$CH_3CH=CH$	0.360

eq 1 and 2, where a and b are given by eq 5 and 6, generates the charge distributions indicated in Table II, in relative units. A multiplication of the relative charges by the appropriate β yields the results in charge units which can be used for a direct comparison with the quantum mechanical results. This terminates the "chemical" population analysis. We shall now focus attention on the MO charge distributions.

Results

A comparison of the charge distribution for alkanes[9] calculated by Del Re's method[15] with charges deduced from charge alternation has proved very satisfactory with $n = 34$ and $\beta = -117$ melectrons.[10] A similar agreement has also been found[10] for the charges cal-

(9) S. Fliszár, *J. Amer. Chem. Soc.*, **94**, 1068 (1972).

(10) S. Fliszár, *J. Amer. Chem. Soc.*, **94**, 7386 (1972).

(11) S. Fliszár and J. Sygusch, *Can. J. Chem.*, **51**, 991 (1973).

(12) (a) R. W. Taft in "Steric Effects in Organic Chemistry," M. S. Newman, Ed., Wiley, New York, N. Y., 1956; (b) R. W. Taft, *J. Amer. Chem. Soc.*, **75**, 4231 (1953).

(13) These comparisons were made using Hoyland's BO results,[14] Del Re's method,[15] and the PCILO method.[16]

(14) J. R. Hoyland, *J. Chem. Phys.*, **50**, 473 (1969).

(15) G. Del Re, *J. Chem. Soc., London*, 4031 (1958).

(16) S. Diner, J. P. Malrieu, and P. Claverie, *Theor. Chim. Acta*, **13**, 1 (1969); J. P. Malrieu, P. Claverie, and S. Diner, *ibid.*, **13**, 18 (1969); S. Diner, J. P. Malrieu, F. Jordan, and M. Gilbert, *ibid.*, **15**, 100 (1969).

(17) This choice appears to be somewhat more practical than q_C-(ethane) = -1, which was used in previous work.[10,11]

(18) The meaning of n can be deduced from eq 5, where a, in a way, measures the sensitivity of charge transfers to substituent effects (eq 2

and 3). Small n values indicate strong substituent effects; if, on the contrary, inductive effects were nonexistent, the charge distributions would be those corresponding to $n = \infty$, i.e., $a = 0$. Previous work[10] indicates that the coefficients (a) are, at least to a good approximation, the same in eq 2 and 3. This point has not received rigorous proof. Although the calculation of charge densities is no more difficult using different a values in eq 2 and 3, it is felt that nothing would be gained by doing so (except occasional better fits with MO charges) and moreover as fully optimized *ab initio* results can be very adequately reproduced with one single a value.

(19) From a consideration of all the trends exhibited by the alkyl groups, the slightly modified σ^* values for the n-butyl (-0.124) and isobutyl (-0.129) groups appear to be more consistent than the usual ones[15] (-0.130 and -0.125, respectively). In the present ordering, any R–CH_2 group is a better electron donor when R is a better donor, whereas the usual σ^* values for n-C_4H_9 and i-C_4H_9 would indicate an inversion in the ordering for these groups. At any rate, the σ^* values used here are well within their error limits. This applies also to the σ^* value for the neopentyl group which is used here (lit.[15a] -0.165; lit.[15b] -0.14, as determined from the kinetics of sulfation of alcohols). All other σ^* values are those of ref 15. Finally, the prime importance of the σ^* values does not rest here in their exact numerical value (which may, in the future, be subject to minor revisions) but in the fact that the same set of σ^* values is used in all the comparisons with the different theoretical methods.

Table II. Net Charges in Relative Units, with $q_C(C_2H_6) = +1.000$

Molecule	Atom	Net charge
Methane	C	$4(n + 1)/3n$
Ethane	C	1.000
Propane	C_{prim}	$(3n + 0.55)/3n$
	C_{sec}	$(2n - 3.8)/3n$
	H_{prim}	$(0.15 - n)/3n$
	H_{sec}	$(0.9 - n)/3n$
Butane	C_{prim}	$(3n + 0.43)/3n$
	C_{sec}	$(2n - 3.35)/3n$
	H_{prim}	$(0.24 - n)/3n$
	H_{sec}	$(1.1 - n)/3n$
Pentane	C_{centr}	$(2n - 2.8)/3n$
	H_{centr}	$(1.25 - n)/3n$
Isobutane	C_{prim}	$(3n + 1.03)/3n$
	C_{tert}	$(n - 7.7)/3n$
	H_{prim}	$(0.29 - n)/3n$
	H_{tert}	$(2 - n)/3n$
Neopentane	C_{prim}	$(n + 0.49)/n$
	C_{quat}	$-4/n$
	H	$(0.51 - n)/3n$

berg) atomic orbitals in the basis, was presented by Salahub and Sándorfy.[2] In this RCNDO method, H 2s, H 2p, C 3s, and C 3p Slater atomic orbitals were added to the basis. In these calculations the energy of the ground state was first minimized in the usual way then configuration interaction was applied including the lowest 30 singly excited configurations for both singlets and triplets. The agreement of the RCNDO charges with "chemical intuition" (Table III) is superior to what can be obtained from the original CNDO/2 method.

The INDO[1] charge distributions calculated by Sándorfy, et al.,[2] are compared in Table III with those deduced from the "inductive" method, with $n = -2$ and $\beta = +71$ melectrons. Contrary to all the other quantum mechanical results discussed in this paper, INDO predicts a C^+-H^- polarity. It is noted that the inductive approach has the flexibility of describing such a situation, with $n < 0$. Except for neopentane, where the INDO charge for the quaternary C atom appears to

Table III. A Comparison of Semiempirical Theoretical Results with the Charges Deduced from Table II (10^{-3} eu)

Molecule	Atom	RCNDO	Inductive $n = 1.6$	INDO	Inductive $n = -2$	EHMO	Inductive $n = 9.14$
Methane	C	-84	-82	$+43$	$+47$	-503	-500
Ethane	C	-34	-38	$+77$	$+71$	-326	-338
Propane	C_{prim}	-44	-42	$+67$	$+65$	-343	-345
	C_{sec}	$+12$	$+5$	$+94$	$+92$	-164	-178
	H_{prim}			-25	-25	$+110$	$+111$
	H_{sec}			-37	-34	$+95$	$+102$
Butane	C_{prim}	-47	-42	$+65$	$+66$	-342	-343
	C_{sec}	$+2$	$+1$	$+85$	$+87$	-181	-184
	H_{prim}			-26	-26	$+110$	$+110$
	H_{sec}			-36	-37	$+96$	$+99$
Pentane	C_{centr}			$+76$	$+80$		
	H_{centr}			-36	-38		
Isobutane	C_{prim}	-52	-47	$+62$	$+59$	-358	-351
	C_{tert}	$+53$	$+48$	$+103$	$+115$	-8	-18
	H_{prim}			-27	-27	$+111$	$+109$
	H_{tert}			-48	-47	$+82$	$+88$
Neopentane	C_{prim}	-57	-50	$+58$	$+54$	-373	-356
	C_{quat}	$+92$	$+95$	$+97$	$+142$	$+140$	$+148$
	H			-27	-30	$+113$	$+106$

culated by Hoffmann,[7] with $n = 9.5$ and $\beta = -356$ melectrons. Similarly, PCILO charge distributions are well reproduced by the inductive formulas, with $n = 0.525$ and $\beta = -11$ melectrons.[11]

With CNDO/2[20] results the situation is, however, somewhat different.[11] While the CNDO/2 results are clearly in qualitative agreement with the "inductive" results in that they reproduce correctly all the major trends predicted by the latter method, the actual numerical agreement between the charges calculated by both methods appears to be erratic. It must be noted that no attempt has been made in that case to improve the CNDO/2 results by an adequate choice of parameters. An "agreement" of similar sort between "inductive" and theoretical calculations is also observed for the results obtained by Sándorfy, et al.,[2] using a modified CNDO approach. The latter is a variant of that of Del Bene and Jaffé,[21] which had been modified to interpret correctly σ-electronic spectra.

A calculation of CNDO type, including higher (Ryd-

be less positive than that of the tertiary C atom in isobutane (which is contrary to the results given by all other methods), it appears that all the major trends are reflected in the same manner by both methods.

A basically different theoretical approach is that suggested by Hoffmann.[7] A variant of this extended Hückel (EHMO) method has been discussed by Whitehead, et al.[22] Their results are found to be in satisfactory agreement with "chemical expectation" (Table III) with $n = 9.14$ and $\beta = -338$ melectrons. Similarly, the SCF-type results calculated by Lipscomb, et al.,[23] also appear to follow the general trends predicted by the "inductive" approach (Table IV).

We shall now proceed with the comparison between our "inductive" formulas and ab initio results, obtained at different degrees of sophistication (Table V).

The "standard" STO-3G calculation was performed[24] using Pople's recipe.[4] Standard geometries were as-

(20) J. A. Pople and G. A. Segal, *J. Chem. Phys.*, **44**, 3289 (1966).
(21) J. Del Bene and H. H. Jaffé, *J. Chem. Phys.*, **48**, 1807 (1968).

(22) J. M. Sichel and M. A. Whitehead, *Theor. Chim. Acta*, **5**, 35 (1966).
(23) M. D. Newton, F. P. Boer, and W. N. Lipscomb, *J. Amer. Chem. Soc.*, **88**, 2367 (1966).
(24) G. Kean and S. Fliszár, to be submitted for publication.

Table IV. A Comparison of SCF-Type Calculations with the "Inductive" Charges Deduced from Table II (eu)

Molecule	Atom	SCF	Inductive $n = 1.6964$
Methane	C	−0.49	−0.49
Ethane	C	−0.19	−0.23
Propane	C_{prim}	−0.25	−0.25
	C_{sec}	−0.05	+0.02
Butane	C_{prim}	−0.26	−0.25
	C_{sec}	+0.01	0.00
Pentane	C_{cetnr}	−0.03	−0.03
Isobutane	C_{prim}	−0.28	−0.28
	C_{tert}	+0.30	+0.27
Neopentane	C_{prim}	−0.29	−0.29
	C_{quat}	+0.54	+0.54

similar to that (+62.92 melectrons) deduced from a fully optimized minimal basis set *ab initio* calculation. As with Andre's results, the charge separations calculated by Hoyland are much larger than those deduced from a minimal basis set. Again, the extended basis set calculation corresponds to a higher n value ($n = 25$) than the n value corresponding to a minimal basis set calculation. Nonetheless, the quaternary C atom in neopentane (as calculated from Hoyland's data) is similar (+81 melectrons) to that deduced from the other *ab initio* calculations reported here.

Discussion

The comparisons presented in Tables III–V indicate

Table V. A Comparison of *Ab Initio* Results with Charges Deduced from Table II

Molecule	Atom	STO-3G[b] "standard"	Inductive[b] $n = 1$	STO-3G[b] "optimized"	Inductive[b] $n = 1.3325$	7s3p/3s[c]	Inductive[c] $n = 42.316$	BO[b]	Inductive[b] $n = 25$
Methane	C	−72	−69	−48.92	−48.92	−0.79	−0.78	−699	−699
Ethane	C	−26	−26	−20.96	−20.96	−0.57	−0.57	−504	−504
Propane	C_{prim}	−31	−31	−23.81	−23.84	−0.58	−0.58	−508	−508
	C_{sec}	+19	+16	+5.94	+5.95	−0.38	−0.37	−303	−310
	H_{prim}	+7	+7.4	+6.23	+6.20			+167	+167
	H_{sec}	+0.3	+0.9	+2.20	+2.27			+160	+162
Butane	C_{prim}					−0.56	−0.57	−504	−507
	C_{sec}					−0.37	−0.37	−309	−313
	H_{prim}							+165	+166
	H_{sec}							+159	+160
Isobutane	C_{prim}	−35	−35	−26.39	−26.36	−0.55	−0.58		
	C_{tert}	+57	+58	+33.36	+33.39	−0.18	−0.16		
	H_{prim}	+6	+6.2	+5.50	+5.47				
	H_{tert}	−6.2	−8.7	−3.53	−3.50				
Neopentane	C_{prim}	−39	−39	−28.66	−28.67				
	C_{quat}	+93	+104	+62.92	+62.92				
trans-2-Butene	$C_{prim}{}^{a}$					−0.60	−0.59		

[a] From eq 1 and 2: $qc = (n + 1.1)\beta/n$. [b] Millielectrons. [c] Electrons.

sumed, with $r(CH) = 1.091$ Å, $r(CC) = 1.540$ Å, $\angle HCH = 109.47°$, and $\angle CCC = 109.47°$. Pople's optimized scale factors[4] were used. The results indicate a satisfactory agreement between the two methods of obtaining charge distributions, with $n = 1$ and $\beta = -26$ melectrons. The greatest discrepancy is observed for the quaternary C atom in neopentane. In order to examine whether this discrepancy should be regarded as an indication for a possible departure of "inductive" results from their theoretical counterpart, a more sophisticated minimal basis set calculation has been performed.[24] First all geometry parameters have been optimized. Second all exponents have been optimized individually for each molecule, including those of the carbon 1s electrons. Of course, different C and H atoms in the molecule have been optimized separately. The charges obtained from this "best" calculation using a minimal basis set are virtually identical with their "inductive" counterpart. The departure observed for neopentane in the case of the "standard" STO-3G calculation should therefore not be considered as indicative for a disagreement for which the inductive approach can be made responsible.

The *ab initio* 7s3p/3s results by André, *et al.*,[5] are also satisfactorily reproduced by the "inductive" method. These results correspond to a very high n value, as compared to a minimal basis set calculation. Interestingly, the charge for the neopentane quaternary C atom (+54 melectrons) which can be deduced from these data is

that the usual interpretation of inductive effects in terms of electron release appears to be adequate, as its description (eq 1 and 2) is in general satisfactorily reproduced by various quantum-mechanical approaches. This observation is important in that it provides a clear link between quantum-mechanical information and linear free energy relationships of the type

$$\Delta G^{\ddagger} = \Delta G_0^{\ddagger} - 2.3RT\rho^*\sigma^*$$

which represent the basic tenet for the discussion of structure-reactivity problems.[12] These comparisons illustrate the flexibility which is contained in the set of equations of Table II. These equations represent any possible scheme of charge distributions, ranging from C^+-H^- to C^--H^+ situations, depending upon the choice of n, whereas any individual theoretical method represents only one case, corresponding to a definite n value.

This situation now raises the following questions. Suppose we were to study a correlation between a certain property (*e.g.*, ^{13}C or H nmr spectra) and charge densities. The first question would then concern the choice of the theoretical method to be used for this comparison. Present results indicate that *any choice of any particular theoretical method in fact reduces to choosing a particular n value* (*i.e.*, that corresponding to the method selected). While all sorts of arguments may be invoked for justifying the preference for one method over the others, it appears difficult to justify *a priori*

Table VI. Summary of the "Inductive" Parameters Which Generate the Results of Various Theoretical Methods

Method	n	β, melectrons
INDO	−2	+71
CNDO/2	0.35	−7.6
PCILO	0.525	−11
RCNDO	1.6	−38
PPP	9.4	−160
EHMO[a]	9.14	−338
EHMO[b]	9.5	−356
Del Re	34	−117
STO-3G standard	1.0	−26
STO-3G optimized	1.3325	−20.96
SCF-type	1.7	−230
BO	25	−504
7s3p/3s	42.3	−573

[a] Reference 22. [b] Reference 7.

why any particular n value should be preferred for describing the charge densities to which the property under observation should be correlated. This is best illustrated by Figure 1, which represents a comparison between C net charges calculated from the "inductive" set of equations (Table II) using different n values.

Figure 1 indicates clearly that the scaling, and even the relative ordering, of the various C charges depend strongly on n. Consequently, *if a property is expected to correlate with C net charges, it is imperative that the latter correspond to the "proper" n value: the use of any theoretical set of charge distributions corresponding to an n value other than the "proper" one is bound to fail in showing the expected correlation.*

The final step consists in understanding the physical meaning of n. Anticipating from the forthcoming study, we note that n can be related very simply to the
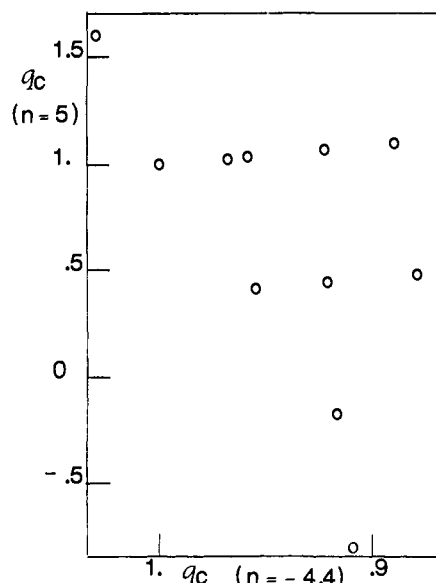

Figure 1. A comparison of carbon net charges, in relative units q_C(ethane) = 1, calculated for $n = 5$ with charges calculated for $n = -4.4$.

mode of partitioning overlap populations. In this sense, the flexibility brought by the set of equations (Table II) comes from the fact that the "inductive" equations are not restricted to any particular artificial mode of partitioning overlap populations.[25]

(25) S. Fliszár, A. Goursot, and H. Dugas, *J. Amer. Chem. Soc.*, **96**, 4358 (1974).