

A General Population Analysis Preserving the Dipole Moment

Bernard T. Thole and Petrus Th. van Duijnen

Theoretical Chemistry Group, Laboratory of Chemical Physics, University of Groningen, Nijenborgh 16, 9747 AG Groningen, The Netherlands

A new and efficient method is proposed for obtaining atomic charges from molecular wave functions, preserving both total charge and dipole moment. The method is independent of the type of wave function (SCF, CI) and does not refer explicitly to the basis set used, nor to integral approximations (e.g. CNDO) applied. The method takes very little time and is better than Mulliken's analysis as a generator of electric potentials.

Key words: Population analysis – Atomic charges.

1. Introduction

In many fields of theoretical chemistry and molecular physics, the concept of atomic charge plays an important role. Atomic charges are used for analysing molecular wave functions, as index for the reactivity of atoms in molecules, and as sources for molecular electrostatic potentials [1]. In contrast to the properties one is trying to represent (charge distribution, electric potential), atomic charge is not a well defined quantum mechanical observable.

The most widely used method for obtaining atomic charges has been Mulliken's population analysis [2]. Although its results reflect trends which are chemically realistic and compatible with concepts like electronegativity, they have hardly any physical meaning. This is caused by the fact that a physically not well defined property – the atomic gross charge – is obtained from an even less physical concept: the assignment of orbital charge to a particular atom, only because the corresponding orbitals are centred on that atom, and the arbitrary reallocation

of orbital overlap populations. Löwdin [3] and Jug [4] suggested the reallocation of overlap populations in a way preserving the bond dipole moment, or at least its component along the axis between the atoms involved. Although this involves little more effort than a standard Mulliken analysis, this method has never become popular. Presumably, because one faces the problem of dealing with the other (one-centre) contributions to the dipole moment. Claverie [5] discussed this problem, and suggested to assign charges to the nearest neighbours of any atom carrying a local dipole, such that this dipole is preserved as well as possible.

The preservation of the molecular dipole and higher moments is, obviously, of paramount importance if one wants to compute, in a simple way, molecular electric potentials. Other, more physical (or less basis set dependent) methods for defining atomic charges [6, 7] still only try to preserve the total charge, and are therefore inadequate. In recent years, a score of methods was reported [1, 8–15] for generating these potentials from reduced representations of molecular charge distributions. They range from using simple Mulliken charges [8], via localized bond contributions [1, 9–12] to fitting charges to the expectation value of a molecular electrostatic potential [13–15]. For use in large molecules, or in calculations which have to be repeated many times (Monte Carlo, Molecular Dynamics), point charge models are preferable.

In this paper we report a method which first uses a generalization of Mulliken's approach, after which a modification of Claverie's method is applied. This modification is chosen in order to make the method "automatic", i.e. no external specification of bonded atoms is necessary. The general idea of this modification leaves several possibilities to construct a population analysis. From experience on a large number of molecules we ultimately chose the form presented in the next section. As criterion for the choice, we used the accuracy of the molecular potential generated by the point charges. Further much attention was paid to obtain a stable method in the sense that small changes in the geometry of the molecule produce no discontinuous jumps in the point charges generated.

2. Method

The expectation value of any one-electron property in an (atomic) basis $\{\chi\}$ can be written as

$$\langle A \rangle = \sum_{ij} D_{ij} A'_{ji} \equiv \sum_{ij} A_{ij} \quad (1)$$

with D_{ij} an element of the density matrix [16] in the basis used, and

$$A'_{ji} = \langle \chi_j | A | \chi_i \rangle. \quad (2)$$

Since the density matrix is used, Eq. (1) applies to any type of wave function expressed in a finite basis set. The summations over i and j may be contracted such that we are left with smaller sums over orbital (usually atomic) centres:

$$\langle A \rangle = \sum_{m,n} A_{mn} \quad (3a)$$

with

$$A_{mn} = \sum_{i \text{ on } n} \sum_{j \text{ on } m} A_{ij}. \quad (3b)$$

This sum over “one-centre” and “overlap” contributions can be contracted further to “net atomic contributions” by the method which may be called “Mulliken contraction”.

$$\langle A \rangle = \sum_n A_n \quad (3c)$$

$$A_n = \sum_m A_{mn}. \quad (3d)$$

Note that $A_{mn} = A_{nm}$ and that A_{mn} is added to A_n and A_{nm} to A_m and thus the overlap value $A_{mn} + A_{nm}$ is divided equally between the centres involved.

The first step in our method is the construction of the net atomic contributions, Q_n and \mathbf{P}_n , to the total electronic charge and dipole moment, from the matrix elements:

$$Q'_{ij} = -\langle \chi_i | \chi_j \rangle \quad \mathbf{P}'_{ij} = -\langle \chi_i | \mathbf{r} | \chi_j \rangle. \quad (4)$$

Next the nuclear charge Z_n is added to Q_n and $Z_n \mathbf{r}_n$ to \mathbf{P}_n to obtain gross atomic charges and dipole moments. This procedure gave the most satisfying results but for the following treatment it is immaterial whether the Q and \mathbf{P} elements are on orbital basis, or contracted either to one-centre and overlap or to gross atomic contributions, or whether the nuclear contributions are added before or after the next step. But the results do depend on the method of contraction. In the following we will assume the method given above which is the strongest contraction of all.

The first essential element in our analysis is that, apart from Q , also \mathbf{P} is contracted. The second is that the local, one-centre dipoles are each redistributed over neighbouring centres such that the dipoles are represented by point charges only. Our procedure is reminiscent of Claverie's but it is more general. Claverie represents his one-centre dipoles on an atom by point charges on that atom and the ones bound to it. We propose to distribute such a dipole over all atoms in the molecule, thus omitting the determination of which atoms are bound. By suitable choice of a weighting function the charge will practically only be distributed over the nearest neighbours. The method does not rest on very deep theoretical foundations. It is an attempt to find a general mathematical expression for the intuitions we have about the various special cases occurring in population analysis.

Here we will first derive the most general method of representing any charge distribution by point charges, preserving both monopole and dipole, because the method is instructive.

Consider an arbitrary one-centre or overlap charge distribution ρ_{mn} , with monopole Q_{mn} and dipole moment \mathbf{P}_{mn} . This distribution may be represented

by point charges q_k on the atomic centres k of the molecule, such that its charge and dipole moment are preserved. Suppressing the indices m and n we require

$$\sum_k q_k = Q \quad \sum_k \mathbf{r}_k q_k = \mathbf{P} \quad (5)$$

where q_k is the part of the charge assigned to atom k . In general, Eq. (5) can be satisfied in many ways. We use this freedom to take the smallest possible charges, placed on the nearest atoms, compatible with Eq. (5). This is achieved by minimizing

$$\sum_k q_k^2 / 2w_k. \quad (6)$$

Here, w_k is a weighting function which decreases rapidly with the distance between atom k and the charge distribution. The form of Eq. (6) ensures that all charges are minimized, but especially for atoms with low w_k , i.e. for distant atoms. Combining Eqs. (5) and (6) and differentiating with respect to the q_k , we have

$$q_k / w_k - \alpha - \mathbf{r}_k^\dagger \boldsymbol{\beta} = 0$$

or

$$q_k = w_k (\alpha + \mathbf{r}_k^\dagger \boldsymbol{\beta}) \quad (7)$$

with α and $\boldsymbol{\beta}$ Langrangian multipliers. We define

$$W \equiv \sum_k w_k; \quad \langle a \rangle \equiv \sum_k w_k a_k / W$$

and solve for α and $\boldsymbol{\beta}$.

$$\begin{aligned} Q &= \sum_k q_k = \sum_k w_k (\alpha + \mathbf{r}_k^\dagger \boldsymbol{\beta}) \\ &= W\alpha + \langle \mathbf{r} \rangle^\dagger W\boldsymbol{\beta} \end{aligned} \quad (8a)$$

$$\begin{aligned} \mathbf{P} &= \sum_k \mathbf{r}_k q_k = \sum_k w_k \mathbf{r}_k (\alpha + \mathbf{r}_k^\dagger \boldsymbol{\beta}) \\ &= \langle \mathbf{r} \rangle W\alpha + \langle \mathbf{r} \mathbf{r}^\dagger \rangle W\boldsymbol{\beta}. \end{aligned} \quad (8b)$$

Hence,

$$W\alpha = Q - \langle \mathbf{r} \rangle^\dagger W\boldsymbol{\beta} \quad (8c)$$

$$W\boldsymbol{\beta} = [\langle \mathbf{r} \mathbf{r}^\dagger \rangle - \langle \mathbf{r} \rangle \langle \mathbf{r} \rangle^\dagger]^{-1} (\mathbf{P} - Q \langle \mathbf{r} \rangle) \quad (8d)$$

from which follows

$$q_k = (w_k / W) \{ Q + (\mathbf{r}_k - \langle \mathbf{r} \rangle)^\dagger \boldsymbol{\gamma}^{-1} (\mathbf{P} - Q \langle \mathbf{r} \rangle) \} \quad (9)$$

where $\boldsymbol{\gamma}$ is the positive semi-definite matrix in square brackets in Eq. (8d). In practice $\boldsymbol{\gamma}$ may be (nearly) singular. This problem is attacked by transforming to coordinates which diagonalize the tensor $\boldsymbol{\gamma}$ and which, for simplicity, have

$\langle \mathbf{r} \rangle$ as the origin. Then

$$q_k = \frac{w_k}{W} [Q - \mathbf{r}_k^{\dagger} (a^{-1} \ b^{-1} \ c^{-1}) \mathbf{P}'] \quad (10)$$

with a, b, c the eigenvalues of $\boldsymbol{\gamma}$ and $\mathbf{P}' = \mathbf{P} - Q\langle \mathbf{r} \rangle$.

According to Eq. (10), Q is simply redistributed over all atoms with distribution function w_k . The tensor $\boldsymbol{\gamma}$ "looks" in which direction \mathbf{P}' can be most easily constructed. If there are many atoms with high weights in direction "a", the eigenvalue, a , will be large and a^{-1} small. This means that P'_a can be constructed by putting small charges on these atoms. If, in contrast, there are only few atoms of sufficient weight in some direction, "b" say, the corresponding eigenvalue, b , will be small, and P'_b will be constructed with relatively large charges.

Hence, $\boldsymbol{\gamma}$ is (nearly) singular when it is impossible (or difficult) to construct a component of \mathbf{P}' in a particular direction. When a molecule is planar, this will occur in the direction perpendicular to the plane. It also occurs when in some direction there are only nuclei with small weights, indicating that they are far away from the distribution at hand. Then the molecule is locally planar or linear. In either case, most likely the corresponding component of P' will be small, and one would rather discard its contribution than reproduce it by using excessively large charges – in the case of a (nearly) planar molecule – or by putting charges on distant nuclei. One could neglect such components by simply putting a^{-1} to zero when a is small compared to the largest eigenvalue. However, in order to avoid discontinuities in the charges, for example when a molecule is deformed from planar to just non-planar, it is better to use the *smooth* function

$$a^{-1} \rightarrow (a + \Delta(c + \Delta))^{-1} \quad \{a = a, b, c\} \quad (11)$$

where c is the largest eigenvalue, and Δ is a small positive number. This means that we deliberately introduce a small error in order to keep the method numerically stable. In practice, we always use $\Delta = 10^{-5}$.

3. The Weighting Function

The weighting function, w_k , measures the relative importance of atom k for charge distribution (m, n) . A natural choice for w_k follows from Eq. (9). According to this equation, Q is distributed over the nuclei using w_k as the distribution function. Also the charges representing the dipole moment \mathbf{P}' are distributed according to w_k . In order to make the charge distribution q_k resemble the distribution ρ_{mn} , w_k should have about the same shape and width as ρ . However, there is no sense in trying to make a really "true" representation of a continuous distribution by the discrete point charges. Therefore we took a reasonable guess for the shape and width of w_k :

$$w_k = \exp[-|\mathbf{r}_k - \mathbf{r}_{mn}|^2/d_{mn}^2] \quad (12)$$

i.e. a gaussian function centred in \mathbf{r}_{mn} and with width d_{mn} . For one-centre distributions (n, n) we might take $\mathbf{r}_{nn} = \mathbf{r}_n$ and d_{nn} proportional to the nearest

neighbour distance. For overlap distributions (m, n) we might take $r_{mn} = (r_m + r_n)/2$ and d_{nm} proportional to $|r_m - r_n|$. The proportionality constant should be of the order of unity. Then one-centre distributions will practically be distributed over the nearest neighbours only and two-centre distributions over the two nuclei involved. This distribution seems reasonable, at least when the basis set does not contain extremely diffuse functions extending over many centres. Thus the method applied to overlap distributions gives essentially the same results as other dipole preserving methods [3, 4] and the one-centre distributions are treated in about the same way as in Claverie's method [5].

We chose the exact value of the proportionality factor for d_{mn} and d_{nn} by inspection of the charges obtained and of the error in the electrostatic potential of the molecule generated by these charges. The potential was chosen because generating potentials was one of our aims, and also because the potential is the most physical, and therefore least arbitrary, criterion. We also compared our charges to potential derived charges (PD). These charges minimize the sum, over a set of grid points of the squares of the differences between the expectation value of the potential and the potential calculated from the point charges. The method used was that of Cox and Williams [15] which constrains the total charge to be equal to the molecular charge. A strong argument in favour of PD is that it would be the exact method if the charge density could be written as the sum of spherical distributions centred at the nuclei. In general this is not possible and the PD method is somewhat unstable because it may try to describe small non spherical distributions by relatively large atomic charges. Therefore PD charges must not be taken too absolutely.

4. Calculations

Using the general method for representing any charge distribution ρ by point charges q_k , in the neighbourhood of ρ , and preserving charge and dipole moment, we have studied the results obtained when the method was applied after various stages of contraction of charge distributions and various values of the proportionality factor for d_{mn} and d_{nn} . As a measure for the relative errors in the potential (V^{model}) generated by Mulliken charges (M), by the monopole and dipole preserving charges (MD), and by the potential derived charges (PD), we used

$$S = \left[\frac{\sum_{\text{grid}} (V_i - V_i^{\text{model}})^2}{\sum_{\text{grid}} V_i^2} \right]^{1/2} \quad (13)$$

where V is the expectation value for the potential. We used a rectangular grid spaced 3 bohr, from which we took the points with a distance between 4 and 7 bohr to the nearest atom in the molecule. The PD charges minimize S . Typical values for S are 2–10% for the PD charges, and 20–50% for M charges (Table 1). The MD charges were generally better than M by a factor of ≈ 2 except for the small molecules without dipole moment (CH_4 , C_2H_2 , C_2H_4 , B_2H_6) which had S values of 100–200%. These molecules often had the wrong sign for their charges. The H atoms were negatively charged. The same effect occurred in

CH_3 and CH_2 groups in the larger dipolar molecules, but there it did not show up clearly in the potential because there the potential was governed largely by the dipole moment (for very large molecules by the local dipole moments of their parts) which was reproduced within 1 per mille. The fact that in small molecules without dipole moment MD is not automatically better than M can be understood if one considers CH_4 . In this molecule the restriction to conserve the dipole moment does not determine the charges. One may take any amount of charge from the central C atom and divide it equally among the surrounding H atoms without changing the dipole (or even the quadrupole). Only octupole conservation will determine the charges uniquely. The same argument can be applied to any tetrahedral (or triangular) part of a large molecule. If the charges are not uniquely determined by dipole conservation (as they are in e.g. HF, H_2O and NH_3), they are determined, somewhat arbitrarily, by the method of contraction and by the choice for the weighting function. This arbitrariness is not fatal however. In Mulliken's analysis only the monopole is conserved exactly, but higher poles can be reproduced roughly if the constituting charge distributions are skillfully assigned to "appropriate" atoms. Going one order higher, our method additionally preserves the total dipole but we have to put in some skill and intuition to obtain realistic atomic charges which roughly preserve higher poles.

While experimenting with the MD method it appeared that it was hard to improve on Mulliken's method if one looks at the signs and the rough magnitudes of the charges of individual atoms. Thus, even though the fifty-fifty distribution of overlap populations looks arbitrary, it has its virtues. The good properties of Mulliken's analysis suggested to us to follow it as closely as possible and to use the MD method only at the latest moment as a correction to the Mulliken analysis. If it is true that the M charges are reasonable, then this correction is expected to be small and the MD charges will be very similar to M charges.

Indeed this procedure proved to be the best of the various versions we have tried. It is defined by the following steps:

1. use Mulliken contraction for overlap and dipole moment matrix elements, cf. Eq. (3c). This gives one-centre monopoles and dipoles Q_k and P_k .
2. add nuclear terms Z_k to Q_k and $Z_k r_k$ to P_k . Now the Q_k are the Mulliken gross charges. The P_k may be called gross dipoles.
3. represent P_n by point charges on the atoms k and add these to Q_k . Here the method of the previous section is used with Q in Eq. (5) equal to zero and $P = P_n$, and with d_{nn} in Eq. (12) equal to the distance to the nearest neighbour of nucleus n , i.e. the proportionality factor is unity. This factor gave about the best fit to the potential.

The MD charges thus obtained for the small molecules are displayed in Table 1. Table 2 gives a comparison of the errors in the potential for the three methods for a number of molecules we happened to do research on. For MD the typical error is 5–20% and thus it is about three times better than M. The largest error occurs in B_2H_6 . This molecule has an interesting potential because it cannot be

Table 1. Atomic net charges ($\times 1000$) from a Roos-Siegbahn basis (RS) contracted to double zeta [17]. For hydrogen 3s basis of Jonkman et al. [18] contracted to (21). M = Mulliken charges, MD = monopole and dipole preserving charges (this work), PD = potential derived charges. *S* = relative error in potential ($\times 100$) cf. Eq. (13). Geometries from Ref. [21]

Molecule	Atom	M	MD	PD
CH ₄	C	-705	-707	-567
	H	176	177	142
	<i>S</i>	26	26	10
C ₂ H ₂	C	-294	-445	-287
	H	294	445	287
	<i>S</i>	3.2	55	2.2
C ₂ H ₄	C	-380	-365	-348
	H	190	183	174
	<i>S</i>	28	27	26
CO ₂	C	887	917	1110
	O	-444	-458	-555
	<i>S</i>	20	18	2.5
B ₂ H ₆	B	-82	-15	-879
	H	58	2	231
	H(bridge)	-34	11	416
	<i>S</i>	127	97	48
CH ₃ OH	C	-175	-92	361
	O	-683	-694	-767
	H (eclipsed)	191	153	25
	H (staggered)	157	81	-34
	H (hydroxyl)	352	471	449
<i>S</i>	44	10	7.6	
CH ₃ CN	C (methyl)	-378	-303	-708
	C (nitril)	-78	143	534
	N	-290	-327	-514
	H	249	162	229
	<i>S</i>	55	19	2.9
CHOOH	C	564	619	854
	O (hydroxyl)	-519	-526	-619
	O	-648	-659	-755
	H	218	89	18
	H (hydroxyl)	385	477	502
<i>S</i>	60	13	8.0	
HCONH ₂	C	547	613	749
	N	-819	-732	-935
	O	-570	-590	-627
	H (carbon)	176	27	5
	H (trans)	331	329	384
	H (cis)	336	354	424
<i>S</i>	31	5.9	4.8	
H ₂ CO	C	111	134	352
	O	-464	-379	-440
	H	177	122	44
	<i>S</i>	30	9.6	7.1

Table 2. Error in potential (S) for some large molecules. the Heptacain H^+ is protonated at N15 (Fig. 1). $(Ala)_n$ means n alanines in the helical conformation, cf. [22, 23], RS basis (cf. Table 1), except $(Ala)_n$ which is in a minimal basis [20]

	M	MD	PD
Heptacain	191	37	6.5
Heptacain H^+	19	3.0	0.6
$PO_4(CH_3)_2^-$	5.8	1.2	0.7
$PO_4(CH_3)_2H$	40	5.6	3.3
$HN(CH_3)_3^+$	7.4	0.6	0.3
$N(CH_3)_3$	66	19	13
$PO_4(CH_3)_2HN(CH_3)_3$	36	5.0	2.7
$(Ala)_2$	26	9.7	3.9
$(Ala)_3$	24	8.5	3.0
$(Ala)_4$	22	8.8	2.2
$(Ala)_5$	19	8.3	1.6
$(Ala)_6$	18	7.8	1.3
$(Ala)_7$	17	7.6	1.2

Table 3. Charges ($\times 10^3$) in 1[2-(2 methoxy-phenylcarbamoyl)ethyl]piperidine (heptacain, cf. Fig. 1). RS-basis (cf. Table 1)

Atom	M	MD	PD	Atom	M	MD	PD
C ₁	-180	-154	29	H ₁₃	219	115	135
C ₂	-223	-190	-401	H ₁₃	209	44	91
C ₃	-260	-254	-547	C ₁₄	-85	41	94
C ₄	364	459	659	H ₁₄	169	26	30
C ₅	-185	-186	-40	H ₁₄	170	31	33
C ₆	142	233	-86	N ₁₅	-691	-340	-561
H ₁	216	107	132	C ₁₆	-140	31	3
H ₂	217	125	197	H ₁₆	206	44	92
H ₃	220	146	217	H ₁₆	161	62	94
H ₅	327	172	199	C ₁₇	-388	-191	-244
O ₇	-759	-529	-501	H ₁₇	189	30	71
C ₈	-117	-24	25	H ₁₇	206	121	122
H ₈	196	100	106	C ₁₈	-398	-163	-24
H ₈	175	95	74	H ₁₈	185	118	51
H ₈	175	97	76	H ₁₈	195	53	31
N ₉	-826	-778	-655	C ₁₉	-386	-189	-154
H ₉	308	368	333	H ₁₉	205	121	90
C ₁₀	1067	1131	1138	H ₁₉	189	29	40
O ₁₁	-621	-691	-716	C ₂₀	-145	30	73
O ₁₂	-656	-477	-489	H ₂₀	163	65	50
C ₁₃	-8	139	73	H ₂₀	193	35	63
error: S					191	37	6.5

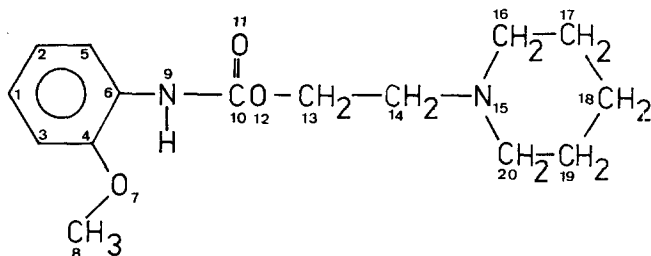


Fig. 1. Structure of 1[2-(2-methoxy-phenylcarbamoyl) ethyl] piperidine (heptacain)

fitted accurately by point charges on the atoms, so the failure of M and MD is not surprising, nor is the great difference between the PD and the M and MD charges. The same argument can be applied to C_2H_4 which also has a potential that is hard to fit. In C_2H_4 however, the M and MD charges are nearly optimal, which may be fortuitous. The MD method is least convincing in the case of C_2H_2 where no definite reason can be found for the bad performance, whereas Mulliken is very good. Maybe C_2H_2 is just an exception or maybe the triple bond causes the trouble.

The next worse case is heptacain [24] with an error of 37%, but here M is still much worse. It is surprising that the individual M charges do not compare badly to the PD charges (Table 3), but that the potential they generate on our grid does not look in the least like the true potential. Large areas, with large absolute potential, have the wrong sign of the potential and in other large areas the potential is overestimated by a factor of two. The improvement given by MD is striking. Almost everywhere the sign is correct and in large areas the magnitude is accurate within a few per cent. But in a few regions, which all have a negative potential, the potential is underestimated by a factor of up to two. These are the regions below and above the aromatic ring and between O_{12} and N_{15} (Fig. 1). The chain between O_{12} and N_{15} is bent such that these atoms are rather close. It is hard to attribute the errors in the potential to particular atoms, but inspection of the errors seems to support the chemical intuition that the π -electron system of the aromatic ring and the lone pairs of O and N atoms give excessive negative potentials which are not represented by MD.

Table 4 contains some information on basisset effects. We calculated formamide in three basissets. In making comparisons we must keep in mind that part of the basisset effects are real, because the charge densities and the potentials are different in different basissets. We assume that the changes in the PD charges are the best measure for the real effects. In Table 4 the changes in the atomic charges with the basisset are smallest for PD and largest for M. MD is about halfway between them. Also the trends in MD and PD charges are more alike than those in M and PD charges. These facts suggest that MD shows less artificial basisset dependency than M.

As an illustration of the behaviour of the three methods in large molecules the charges for the 3- and 4-peptide α -helices of polyalanine are given in Table 5.

Table 4. Basis set effects on charges in formamide

Basis	MB ^a			RS ^b			RS* ^c		
	M	MD	PD	M	MD	PD	M	MD	PD
H	197	-2	-31	176	27	5	25	-25	-10
C	315	499	651	547	613	749	473	436	534
O	-426	-475	-500	-570	-590	-627	-457	-494	-513
N	-795	-632	-906	-819	-732	-935	-385	-401	-736
H(cis)	378	326	410	336	354	424	167	215	354
H(trans)	367	283	377	331	329	384	177	269	371
μ^d	4.12	3.67	3.66	4.21	4.41	4.41	3.18	4.25	4.25
S^e	33	10	8.3	31	5.9	4.8	30	8.2	6.0

^a Minimal basis according to Mehler and Paul [20].

^b Roos-Siegbahn basis (cf. Table 1).

^c RS plus one polarization function per atom [19].

^d Dipole moment in Debye.

^e Relative error in potential: Eq. (13).

The charges are interesting because they may be used for the calculation of Coulomb interactions in proteins. Again a comparison of individual M- and MD-charges with PD charges is not particularly unfavourable for M. It is the better distribution of charge on a larger scale which gives the better MD potential. Table 5 also gives an impression of the relative instability of PD charges. More or less equivalent atoms in one molecule show more variations in their PD charges than in M and MD charges.

5. Conclusions

We present a new, general and efficient method for obtaining charges from molecular wave functions. The method preserves the dipole moment and, in larger molecules tends to preserve higher moments because of the representation of local dipoles. The general method can lead to a variety of charge analyses but the best results are obtained by a version which is as close as possible to the Mulliken analysis. The potentials obtained are about three times as accurate as those from Mulliken charges and also the basisset dependency is smaller. New elements of the method are the Mulliken contraction of the dipole matrix elements and the automated method of representing dipole moments by atomic charges.

The method is superior to Mulliken in the sense that the potentials generated by the charges and hence the Coulomb interactions between molecules are better. However, as a tool to analyse the numerical results of calculations the Mulliken charges have the advantage that they are more simply related to the overlap and density matrix elements, which makes interpretation easier. In many cases the two methods will give complementary information, and because they are both computationally very fast they may be generated as standard output of

Table 5. Charges ($\times 10^3$) in (Ala)₃ and (Ala)₄; cf. Table 2. HC' and underlined charges designate H atoms replacing C_α at the terminal units. Minimal basis [20]

HC'	C	O	N	H	C _α	H _α	C _β	H _{β1}	H _{β2}	H _{β3}
201	311	-444	-630	374	-156	274	-648	249	206	227
	530	-427	-615	357	-153	268	-647	246	207	222
	530	-450	-775	354	<u>387</u>					(Ala) ₃ M
216	331	-477	-627	383	-158	275	-649	254	209	229
	525	-425	-615	357	-151	271	-647	250	206	222
	533	-430	-610	356	-150	260	-653	237	233	210
	527	-465	-788	390	<u>372</u>					(Ala) ₄ M
18	434	-442	-694	371	683	-62	-932	236	185	224
	347	-460	-344	300	251	-38	-767	200	171	190
	968	-588	-1087	392	<u>442</u>					(Ala) ₃ PD
	259	-257	-612	374	608	-80	-819	215	162	196
63	575	-511	-793	379	814	-148	-810	192	153	176
	513	-500	-566	363	212	13	-697	201	126	168
	738	-554	-714	213	<u>347</u>					(Ala) ₄ PD
	441	-447	-336	241	215	-14	-532	149	123	105
-46	702	-555	-408	244	214	-18	-530	146	124	101
	770	-585	-675	288	<u>282</u>					(Ala) ₃ MD
	436	-417	-341	251	213	-14	-534	154	126	107
-41	697	-551	-410	245	200	-20	-532	150	120	100
	702	-559	-405	247	219	-26	-536	136	145	90
	768	-600	-693	305	<u>268</u>					(Ala) ₄ MD

LCAO calculations. The present method is simpler and much faster than the generation of potential derived charges, which also suffers from some instability against small changes in geometry.

References

1. Scrocco, E., Tomasi, J.: *Adv. Quantum Chem.* **11**, 115 (1978)
2. Mulliken, R. S.: *J. Chem. Phys.* **23**, 1833, 1841 (1955)
3. Löwdin, P. O.: *Adv. Phys.* **5**, 2 (1956)
4. Jug, K.: *Theor. Chim. Acta* **39**, 301 (1975)
5. Claverie, P.: *Intermolecular interactions. From diatomics to biopolymers* (B. Pullman, ed.) p. 283 New York: Wiley, 1978
6. Heinzmann, R., Ahlrichs, R.: *Theoret. Chim. Acta (Berl.)* **42**, 33 (1976)
7. Polák, R.: *Theoret. Chim. Acta (Berl.)* **50**, 21 (1978)
8. Hayes, D. M., Kollman, P. A.: *J. Am. Chem. Soc.* **98**, 3335 (1976)
9. Agresti, A., Bonnacorsi, R., Tomasi, J.: *Theoret. Chim. Acta (Berl.)* **53**, 215 (1979)
10. Naráy-Szabó, G.: *Int. J. Quantum Chem.* **16**, 265 (1979)
11. Shipman, L. L.: *Chem. Phys. Letters* **31**, 361 (1975)
12. Martin, D., Hall, G. G.: *Theoret. Chim. Acta (Berl.)* **59**, 281 (1981)
13. Smit, P. H., Derissen, J. L., van Duijneveldt, F. B.: *Mol. Phys.* **37**, 521 (1979)
14. Momany, F.: *J. Phys. Chem.* **82**, 529 (1978)
15. Cox S. R., Williams, D. A.: *J. Comp. Chem.* **2**, 304 (1981)
16. McWeeny, R., Sutcliffe, B. T.: *Methods of molecular quantum mechanics*, ch. 4. New York: Academic Press, 1969
17. Roos, B., Siegbahn, P.: *Theoret. Chim. Acta (Berl.)* **17**, 209 (1970)
18. Jonkman, H., van der Velde, G. A., Nieuwpoort, W. C.: *Chem. Phys. Letters* **25**, 62 (1974)
19. Roos, B., Siegbahn, P.: *Theoret. Chim. Acta (Berl.)* **17**, 199 (1970)
20. Mehler, E. L., Paul, C.: *Chem. Phys. Letters* **63**, 145 (1979)
21. Sutton, L. E.: *Tables of interatomic distances*, Special publication No 18. London: The Chemical Society, 1965
22. van Duijnen, P. Th., Thole, B. T.: *Chem. Phys. Letters* **83**, 129 (1981)
23. van Duijnen, P. Th., Thole, B. T.: *Biopolymers* **21**, 1749, (1982)
24. Remko, M., van Duijnen, P. Th.: to be published

Received October 28, 1982