Density matrix refinement for molecular crystals

S. T. Howard
Department of Chemistry, University of Wales College of Cardiff, Cardiff CF1 3TB, United Kingdom

J. P. Huke
Defence Research Agency (Malvern), St. Andrews Road, Great Malvern, Worcestershire WR14 3PS, United Kingdom

P. R. Mallinson and C. S. Frampton
Chemistry Department, University of Glasgow, Glasgow G12 8QQ
(Received 25 May 1993)

The fitting of an idempotent density matrix to x-ray structure factors is discussed. If the density matrix is identified with a single Slater determinant of orbitals, then a model wave function can be recovered, enabling the computation of two-electron properties as well as the usual one-electron properties derived from the electron density. Hence this approach is potentially superior to the usual multipole expansion used in charge-density analysis, although the current formalisms are only useful for molecular crystals of very small molecules. A procedure is described, and applications to one theoretical example (methylamine) and one experimental example (formamide) are presented.

I. INTRODUCTION

The determination of charge distributions from scattering experiments on crystals is a growing area of research.[1–3] Progress has been made in elastic x-ray scattering, which yields total electron densities, and in polarized neutron scattering for spin densities. Developments in inelastic and Compton scattering have led to determinations of momentum densities. All of these techniques can in principle be used to find the electron density $\rho(\mathbf{r})$. Most methods for analyzing such experiments involve fitting a parametrized model density; if this model is derived from a model wave function such as a Slater determinant of linear combination of atomic orbitals (LCAO) molecular orbitals (MO's), the process also gives an estimate of the system's wave function, making possible the computation of many-electron properties. The work described below concentrates on determining approximate (molecular) wave functions using elastic x-ray scattering experiments on molecular crystals, although the technique presented may be more widely applicable.

X-ray scattering is commonly analyzed in terms of the Bragg structure factor $F(\mathbf{s})$, the Fourier transform of $\rho(\mathbf{r})$. A parametrized model of $\rho(\mathbf{r})$ is chosen, and the parameters are fitted to the experimental data by minimizing differences between the observed structure factors and those calculated from the model. The so-called "multipole model"[4,5] of a charge distribution has come to be widely adopted. Two types of multipole model are in common use, which we might call "deformation models" and "valence-shell models." Both share the philosophy of expanding the density as a sum of "pseudoatomic," nuclear-centered contributions: $\rho(\mathbf{r}) = \sum_j \rho_j(\mathbf{r})$. In the first of these, only the deformation density is described by a multipole expansion:

$$\rho_j(\mathbf{r}_j) = \rho_{j,\mathrm{HF}}(\mathbf{r}_j) + \sum_{l,m} P_{jlm} R_{jl}(r_j) Y_{lm}(\mathbf{r}_j). \quad (1)$$

Here $\rho_{j,\mathrm{HF}}(\mathbf{r}_j)$ is the spherically averaged density from a Hartree-Fock (HF) calculation on the jth atom, R_{jl} is a radial function, Y_{lm} is a spherical harmonic, and the multipole populations P_{jlm} are adjustable parameters. In the second scheme, HF core and valence shell populations are also adjustable:

$$\rho_j(\mathbf{r}) = P_{jc}\rho_{jc}(r_j) + \kappa_j'^3 P_{jv}\rho_{jv}(\kappa_j' r_j) + \sum_{l,m} \kappa_j''^3 P_{jlm} R_{jl}(\kappa_j'' r_j) Y_{lm}(\mathbf{r}_j). \quad (2)$$

In either model, the basis function exponents may be adjusted by varying the κ_j', κ_j'' parameters in addition to the multipole populations.

Equations (1) and (2) provide compact descriptions of $\rho(\mathbf{r})$, in which the numbers of adjustable parameters are sufficiently small that they can be fitted to the results of an elastic x-ray scattering experiment. Experimental charge densities determined by multipole refinement can be in very good agreement with theory, where such comparisons are possible.[6–8] The results of such studies have been applied in modeling intermolecular interactions[9–11] and have helped to elucidate the nature of hydrogen bonding.[12]

Although an associated wave function is not obtained in the course of fitting the multipole model, it is always possible in principle to find a wave function Ψ that gives rise to the fitted density through

$$\rho(\mathbf{r}) = N \int \Psi^* \Psi \, d\mathbf{r}_2 d\mathbf{r}_3 \ldots d\mathbf{r}_N. \quad (3)$$

That is to say, any density is *N representable.*[13,14] However, in general many wave functions can be constructed for a given $\rho(\mathbf{r})$, so the choice of a suitable wave function would require further work, over and above the fitting of the density parameters. One approach to doing this would be to try to find a wave function which (say) minimized the total energy. A more direct approach than trying to recover a wave function from a multipole model density such as (1) or (2) is to use a density based upon a model wave function and to determine the wave function parameters by a fit to the x-ray data. Of course, since many different functions Ψ may give rise to the same ρ via (3), we must choose our model wave function carefully: if the set of models contains more than one wave function corresponding to a given density, we may not be able to determine the parameters uniquely by a fit to the data. In practice, we restrict the wave function to be a Slater determinant of MO's which are linear combinations of a fixed basis set. This will certainly remove any ambiguity if the products of basis functions are linearly independent,[22] but even if they are not, imposition of an idempotency constraint on the density matrix should lead to a unique fit, as mentioned below.

There are other reasons for preferring model densities derived from wave functions. It is well known that the population parameters of multipole models can correlate strongly with other parameters in the total scattering model, such as nuclear positions, thermal motion parameters, and basis function exponents.[4] In particular, a correlation between anisotropic thermal parameters β and multipoles on the same nucleus probably ultimately limits the accuracy with which the charge density and thermal motion can be deconvoluted. It seems reasonable to expect that this problem will be reduced by the use of two-center products to describe the density.

Here we consider only molecular crystals, and we assume that an expansion of the wave function in Bloch-type states is unnecessary, i.e., the (molecular) states are genuinely localized. However, thermal averaging causes electron distributions from (solid-state) x-ray diffraction to differ fundamentally from the Born-Oppenheimer (BO) $\rho(\mathbf{r})$ of a free molecule. The elastic scattering intensity is given by[16]

$$I_{\text{elastic}}(\mathbf{s},T) = \sum_m W_m(T) \left| \int X_m^*(\mathbf{Q}) F(\mathbf{s},\mathbf{Q}) X_m(\mathbf{Q}) d\mathbf{Q} \right|^2, \quad (4)$$

where $W_m(T)$ is a Boltzmann-type weight for the crystal normal mode $X_m(\mathbf{Q})$ at temperature T, $\mathbf{Q}$ is the nuclear configuration, and $F(\mathbf{s},\mathbf{Q})$ is the structure factor. The charge distribution obtained from an analysis of (4) thus refers to an ensemble of vibronic states and $\rho = \rho(\mathbf{r},T)$. As the temperature is lowered, successive $\rho(\mathbf{r},T)$ converge towards the ground vibronic state average. Although this is still not identical to the BO $\rho(\mathbf{r})$, it is widely assumed that this ensemble density contains the same pertinent chemical information.

If we model the wave function of the system as an LCAO Slater determinant, then the density is given by

$$\rho(\mathbf{r}) = \sum_{i,j}^{K} P_{ij} \chi_i(\mathbf{r})^* \chi_j(\mathbf{r}), \quad (5)$$

where the χ's are the atomic orbital basis functions and P is known as the density matrix. Furthermore, a necessary and sufficient condition for a density in the form of (5) to be derivable from such a Slater determinant is that $P^2 = P$: the matrix P is then said to be *idempotent.*[17] (If the basis functions are not orthogonal, this condition is modified to include the overlap matrix.) The eigenvectors of P, which have eigenvalues of 1 or 0, are the MO coefficients of sets of occupied and virtual orbitals, respectively.

So by fitting a model of the form (5) and imposing idempotency as a constraint, we not only obtain a density, but also an approximate wave function, with which we can calculate any desired property. Only a limited number of interesting properties can be computed from $\rho(\mathbf{r})$ alone — one-electron properties, such as molecular multipole moments, the electric field, electrostatic potential, and $\nabla^2\rho$.

In the absence of the idempotency constraint on P, model (5) becomes a general expression for $\rho(\mathbf{r})$ expanded in one and two-center products of basis functions. Various early attempts to determine charge distributions from x-ray structure factors employed this type of parametrization.[18–20] These studies encountered problems with linear dependence of the basis function products. However, as pointed out by Clinton *et al.*,[21] the imposition of idempotency as a constraint considerably reduces the number of parameters to be determined and should thus reduce or even eliminate this problem. According to Levy and Goldstein,[22] even in the presence of linear dependence in the basis set products $\{\chi_i\chi_j\}$ a unique Slater determinant can be found if the constraint is imposed, provided that the number of linearly independent products is sufficiently large.

The fitting procedure involves minimizing the average squared error between the N_{obs} observed and calculated structure factors

$$\chi^2 = \frac{1}{N_{\text{obs}}} \sum_{\mathbf{s}} w_{\mathbf{s}} (|F_{\text{obs}}(\mathbf{s})| - k|F_{\text{calc}}(\mathbf{s})|)^2 \quad (6)$$

subject to the idempotency constraint. The $w_{\mathbf{s}}$ are weighting factors and k is a scale factor (although it is the $\{F_{\text{obs}}\}$ which are not on an absolute scale, crystallographers conventionally scale the $\{F_{\text{calc}}\}$). Two approaches can be distinguished. We might try to impose the constraint analytically, for example, by using Lagrange multipliers; or we might attempt to construct idempotent density matrices in the vicinity of some current guess for P and iteratively choose those with lower values of χ^2 until a minimum is found. The first approach leads to a pseudoeigenvalue problem of the HF type, except that χ^2 is minimized in place of the energy;[23] as in the HF method, the solution of this problem will be an iterative procedure. The second approach is the one we adopt here.

One of the simplest schemes for fitting an idempotent P to experimental data has been reported by Pecora.[24]

This involves making changes in the density-matrix elements based on the gradient of χ^2 and then restoring the matrix to idempotency by an iterative technique due to McWeeny.[17] Successful applications were made to some model problems, and to positron-annihilation data for a Cu-Ge alloy.[24] Clinton, Frishberg, Massa, and Oldfield[21] have developed a procedure for fitting an idempotent density matrix to elastic x-ray scattering data, with a subsequent experimental application to beryllium metal.[25] Aleksandrov, Tsirelson, Reznik, and Ozerov[26] have recently presented a variation on this scheme and applied it to diamond and silicon. To date, we are unaware of any application of these techniques to x-ray scattering from a molecular crystal.

In this paper we describe a scheme related to that of Pecora, and apply it to fit idempotent density matrices to x-ray structure factors. Two examples are presented: in the first, the structure factors are computed from an *ab initio* wave function for methylamine; the second uses experimental low temperature data for the molecular crystal formamide.

II. METHOD

Our model wave function is a single LCAO MO Slater determinant; hence the molecular density is given by (5), where P is idempotent. We consider a crystal with one closed-shell molecule in an asymmetric unit and $(Z-1)$ symmetry-related molecules in the remainder of the unit cell. Such a crystal scatters x rays elastically according to the structure factor

$$F_{\text{calc}}(\mathbf{s}) = \sum_{i,j}^{K} P_{ij} \sum_{k}^{Z} T_{ijk}(\mathbf{s}) e^{2\pi i \mathbf{s}\cdot\mathbf{R}_{ijk}} \int \chi_{ik}\chi_{jk} e^{2\pi i \mathbf{s}\cdot\mathbf{r}} d\mathbf{r}$$
$$= \sum_{i,j}^{K} P_{ij} \sum_{k}^{Z} T_{ijk}(\mathbf{s}) O_{ijk}(\mathbf{s}) e^{2\pi i \mathbf{s}\cdot\mathbf{R}_{ijk}}, \tag{7}$$

where T_{ijk} is the temperature factor for the ijth basis function product in the kth molecule $\chi_i\chi_j$ and $\mathbf{R}_{ijk}$ locates this product in the unit cell. [If the integral in (7) refers to an origin at the midpoint of the two basis function centers, then $\mathbf{R}_{ijk}$ is the position of this point in an appropriate Cartesian system referred to the unit cell.] Formulas for the evaluation of the two-center integrals O_{ijk} over Gaussian-type basis functions are reported by Stewart.[27]. For one-center products, the temperature factor is that of the nucleus on which the basis functions are centered. The treatment of temperature factors for two-center products is more arbitrary; in this work, we use the mean of the factors for the two nuclei, as justified for rigid-body molecular motion by Stewart.[27] An alternative procedure based on the translation-libration description of rigid molecular motion has been described by Stevens, Rys, and Coppens.[28]

The squared error (6) is to be minimized subject to the constraint of idempotency, which takes the form

$$PSP = 2P,$$

where S is the basis function overlap matrix.[17] (The factor of 2 arises from assuming doubly occupied MO's.) The number of independent density-matrix elements is $p = N(K-N)$, where N is the number of (doubly) occupied MO's and K is the number of basis functions. For a unique fit, p must be less than or equal to N_{obs}.[22]

We take as a starting point the density matrix from an *ab initio* self-consistent field calculation, using nuclear coordinates which are assumed known, e.g., from a neutron diffraction experiment. Ideally, the temperature factors will also be predetermined in a neutron diffraction experiment and fixed during the density-matrix refinement.

The optimization strategy proceeds as follows: (i) Make a small, random perturbation to P, to get P'. (ii) Make P' idempotent. (iii) Compute $\chi^2(P')$ and decide whether to adopt P' as current best fit density matrix. (iv) Iterate (back to step 1) until χ^2 is minimized (P is converged).

In step (i), random numbers uniformly distributed between $-\delta$ and $+\delta$ are added to each element P_{ij} of P, where δ is a parameter gradually reduced in size during refinement. Typical initial and final values are 0.05 and 0.0005, respectively. Step (ii) uses McWeeny's procedure.[17] Beginning with a non-idempotent matrix P' $(= P_0)$, a sequence of matrices is generated according to

$$P_{n+1} = \frac{3}{2}P_nSP_n - \frac{1}{2}P_nSP_nSP_n. \tag{8}$$

In the Appendix, we show that this sequence converges to the idempotent matrix nearest to P', if δ is small enough. In practice, Eq. (8) is iterated until the elements of the matrix

$$P_nSP_n - 2P_n \tag{9}$$

are all less than some specified cutoff value P_{tol} (0.001 was used).

In step (iii), $\chi^2(P')$ is computed from Eqs. (6) and (7). This is the rate-determining step, since it involves a sum over hundreds or thousands of observations. If $\chi^2(P') < \chi^2(P)$, then P' is adopted as the new P. If $\chi^2(P') > \chi^2(P)$, then P' is either rejected ("direct descent optimization") or adopted with some temperature-dependent probability ("simulated annealing") given by

$$\text{Prob} = \exp\{-[\chi^2(P') - \chi^2(P)]/T_a\}. \tag{10}$$

At the beginning of refinement, T_a is chosen so that in-

creases in χ^2 are accepted with high probability. T_a is lowered during refinement, so that finally only decreases in χ^2 are accepted, leading to minimization. For either method, the criterion for convergence was that none of the last Q density matrices generated which gave lower χ^2 values should differ from the current matrix in any element by more than 0.001 (values of Q varied between 10 and 50).

After choosing a basis set, the obvious starting guess for P is the ground state density matrix obtained from a molecular HF calculation. In fact, only such starting guesses were successful in the two examples presented below. Cruder alternatives were tried, including block-diagonal matrices constructed from the density matrices of the free atoms, but these proved to be so far from idempotency that the McWeeny procedure diverged.

The fitted density can be constrained to have the appropriate symmetry by choosing symmetry adapted basis functions. The basis functions should be symmetry adapted to reflect the site symmetry of the molecule in its crystalline environment. Neither system considered below has nontrivial symmetry elements.

Since our optimization procedure uses no gradient information it is rather slow—typically tens of thousands of idempotent P must be sampled. Providing that this is computationally feasible for a given problem, this approach has certain advantages. Because of its simplicity it is easily extended to optimize model parameters other than the P_{ij}, for example, the temperature factors and basis function exponents.

III. APPLICATIONS

A. Methylamine

To test the procedure, a theoretical study was first carried out on the molecular crystal methylamine (CH_3NH_2; Z=8; space group $D_{2h}^{15} - Pcab$). A hypothetical crystalline density was used, in which each molecule was assigned a density taken from a 6-311G SDCI (singles and doubles configuration interaction) *ab initio* calculation with GAUSSIAN90.[29] The fractional coordinates and unit cell were taken from the experimental x-ray study of Atoji and Lipscomb.[30] Structure factors were computed with Eq. (7), ignoring thermal motion, i.e., the T_{ijk}'s were all set to unity. These are then to be treated as the $\{F_{\rm obs}\}$ in (6). Two sets of data were used in the refinements: (i) 221 structure factors, applying $\sin(\theta)/\lambda < 0.5$, $|F_{\rm obs}| > 1$ and omitting $F(000)$, and (ii) 845 structure factors, applying $\sin(\theta)/\lambda < 0.75$, $|F_{\rm obs}| > 0.1$ and omitting $F(000)$.

Data set (i) might be typical of a routine experimental data collection, with a relatively low $\sin(\theta)/\lambda$ cutoff, and with weak reflections likely to be omitted in the refinement. Data set (ii) is more typical of a charge-density quality data set, although the $\sin(\theta)/\lambda$ cutoff should be around 1.0 or better for accurate work.

The basis set used for the fit was STO-3G. With 15 basis functions and 9 doubly occupied MO's, this gives 54 independent density-matrix elements. So even with data set (i) the ratio observations:parameters is greater than 4, and a unique fitted density matrix should be obtained. The starting guess density matrix was taken from an HF STO-3G calculation.

We first consider the results of direct descent refinement. Approximately 15 000 new idempotent density matrices were considered during the optimization, which used δ values ranging from 0.025 to 0.003; more than two-thirds were generated with this smallest δ. With the largest value of δ, typically five iterations of Eq. (8) were needed to restore idempotency to within the specified tolerance ($P_{\rm tol} = 0.001$), whereas one iteration was sufficient with the smallest δ value.

Table I summarizes the values of χ^2 and the agreement factor, defined as

$$R = \frac{\sum_{\mathbf{s}} |F_{\rm obs}(\mathbf{s})| - k|F_{\rm calc}(\mathbf{s})|}{\sum_{\mathbf{s}} |F_{\rm obs}(\mathbf{s})|}. \tag{11}$$

The scale factor k is unity in this example, since both $\{F_{\rm obs}\}$ and $\{F_{\rm calc}\}$ are on an absolute scale. Compared to structure factors computed from the relativistic HF (RHF) Slater-type orbital (STO)-3G density matrix, a fivefold improvement in χ^2 is obtained when a density matrix is fitted to data set (i). The agreement factor also shows a significant improvement. Figures 1(a)–1(c) are molecular density difference maps in a plane containing $N1$, $C1$ and $H1$. Figure 1(a) shows the difference between the RHF STO-3G density and the exact (6-311G SDCI) density. Figures 1(b) and 1(c) show how this difference becomes smaller as the STO-3G density is fitted to the 6-311G SDCI density. The more extensive data set (ii) gives the smallest values of χ^2, R, and the flattest difference map. The inclusion of higher-angle scattering in data set (ii) means that a more realistic representation of the core electron density can be extracted from the data. This corresponds more closely to the representation built into the basis set, so a better fit is obtained.

The Laplacian $\nabla^2\rho(\mathbf{r})$ contains information about where charge is concentrated and depleted.[31] Figures 2(a)–2(c) are Laplacian maps for the exact, RHF STO-3G and fitted [data set (ii)] densities, in the same plane as Fig. 1. It is evident that the RHF STO-3G and exact Laplacians differ considerably and not clear that the fitted STO-3G Laplacian represents an improvement. This is not surprising, given the well-known limitations of the minimal basis set.

A critical point (CP) analysis of $\rho(\mathbf{r})$ (Ref. 31) is a quantitative tool for comparing the fitted and exact densities. Table II compares several properties evaluated

TABLE I. Direct descent refinement for CH_3NH_2.

	Hartree-Fock STO-3G	Density-matrix refinement, STO-3G		
		Data set (i)	Data set (ii)	Data set (ii)[a]
χ^2	0.215	0.044	0.0202	0.0026
R	0.033	0.0151	0.0146	0.0104

[a]No idempotency constraint.

at the bond critical points (BCP's): the points where $\nabla\rho(\mathbf{r}) = 0$. The last column of Table II gives the distance (in bohrs) from the BCP positions in the 6-311G SDCI density, to the BCP positions in the various STO-3G densities. The fitting procedure moves all the STO-3G BCP positions towards the exact ones, with data set (ii) giving the best agreement. A similar improvement is seen in the CP density values ρ_c, with one exception in the bond C1-H5. All of the CP Laplacian values $\nabla^2\rho_c$, without exception, are taken further from the exact values by the fitting procedure. This may be because the RHF STO-3G density optimizes the total energy, which depends on the second derivative of the wave function; the fitted density minimizes χ^2, which depends only on the wave function itself.

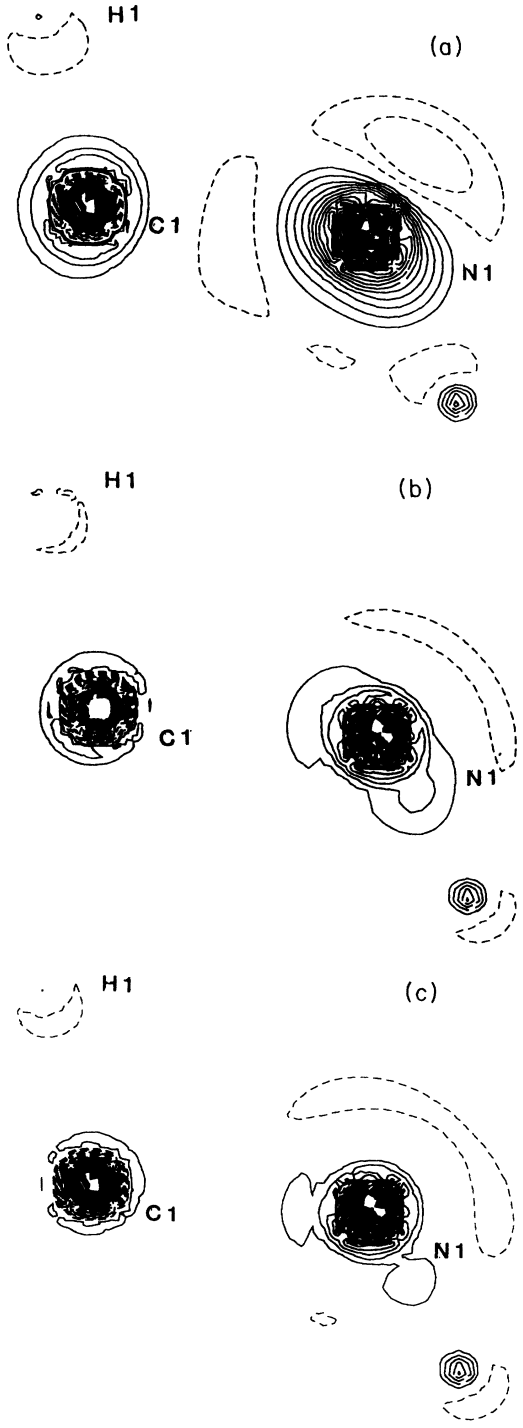


FIG. 1. Molecular difference densities for CH_3NH_2. (a) HF/6-311G SDCI $\rho(\mathbf{r})$, HF/STO-3G $\rho(\mathbf{r})$. (b) HF/6-311G SDCI $\rho(\mathbf{r})$, fit (i)/STO-3G $\rho(\mathbf{r})$. (c) HF/6-311G SDCI $\rho(\mathbf{r})$, fit (ii)/STO-3G $\rho(\mathbf{r})$.

A direct descent refinement was carried out with P_{tol} set to a large positive number (effectively, no idempotency or normalization constraint). This of course produced a better fit (in the sense of lower R and χ^2; see Table I), but a population analysis of the fitted density matrix gave only 17.36 electrons instead of the correct 18. Since such fits are neither normalized nor idempotent they are of little interest, except perhaps to indicate the best possible fit for a given basis set.

Refinements were also carried out with the simulated annealing method of optimization. Some ten runs all gave different fitted density matrices, with final values of R and χ^2 which were higher than the direct descent optimization. So although the technique finds additional local minima, the lowest minimum is obtained by descending directly from the RHF starting guess, in this particular example.

B. Formamide

The experimental study differs from the first in that the true $\rho(\mathbf{r})$ is unknown. However, it is often assumed that the electron distribution from a large basis set RHF calculation on a single molecule (with the crystalline geometry) gives an adequate estimate of the true $\rho(\mathbf{r})$. Thus we might use such a density as a benchmark for different refinement models. An experimental charge-density study of formamide has been carried out by Stevens,[32] although he used only monopole functions on each nucleus to describe the density. Here we will compare the results of multipole refinement and density matrix refinement in some detail.

The statistics R and χ^2 alone are insufficient for comparing models with different numbers of variables N_{var}, so we will also use the goodness of fit:

$$S = \left[\frac{\sum_{\mathbf{s}} w_{\mathbf{s}} (|F_{\text{obs}}(\mathbf{s})| - k|F_{\text{calc}}(\mathbf{s})|)^2}{N_{\text{obs}} - N_{\text{var}}} \right]^{1/2} . \quad (12)$$

For the density-matrix refinements, the contribution of the density matrix elements to N_{var} is taken as $K(N-K)$.

The collection and multipole analysis of the low-temperature experimental data for formamide will be described elsewhere.[33] A multipole refinement of the data with LSMOL,[34] using a model with 88 variables (no variation of κ' and κ'' variables) resulted in R=0.0304, χ^2=3.061, and $S = 1.791$.

To investigate the effect of basis set size on the quality of *ab initio* structure factors, three sets (STO-3G, 6-31G, and 6-31G**) of structure factors were computed

TABLE II. CH_3NH_2 bond critical point analysis.

	ρ_c (e bohr^{-3})				$\nabla^2\rho_c$ (e bohr^{-5})				RCP(HF/CI)-RCP (bohr)		
Bond	CI	HF	Fit 1	Fit 2	CI	HF	Fit 1	Fit 2	HF	Fit 1	Fit 2
C1-N1	0.229	0.242	0.227	0.229	-0.41	-0.44	-0.29	-0.30	0.08	0.06	0.04
N1-H1	0.312	0.329	0.309	0.305	-1.07	-1.19	-0.88	-0.84	0.04	0.03	0.02
N1-H2	0.321	0.338	0.305	0.307	-1.11	-1.21	-0.60	-0.65	0.04	0.02	0.02
C1-H3	0.245	0.254	0.238	0.244	-0.65	-0.62	-0.42	-0.48	0.05	0.03	0.03
C1-H4	0.267	0.271	0.261	0.266	-0.78	-0.71	-0.56	-0.61	0.06	0.05	0.04
C1-H5	0.278	0.279	0.269	0.275	-0.86	-0.75	-0.58	-0.65	0.06	0.05	0.05

for the same 1926 reflections used in the experimental refinement. These were thermally "smeared" with the temperature factors from the multipole refinement, so a direct comparison could be made between the experiment and theory. The scale factor for the $\{F_{\text{calc}}\}$ was chosen to minimize χ^2, i.e., it was found from

$$k = \frac{\sum_{\mathbf{s}} w(\mathbf{s})|F_{\text{obs}}(\mathbf{s})||F_{\text{calc}}(\mathbf{s})|}{\sum_{\mathbf{s}} w(\mathbf{s})|F_{\text{obs}}(\mathbf{s})|^2}. \quad (13)$$

The agreement factors obtained in order of increasing basis size were STO-3G 0.0566, 6-31G 0.0379, and 6-31G** 0.0366. Evidently the HF structure factors are relatively insensitive to the addition of d-type polarization functions for such (isolated) organic molecules. In crystals, however, the intermolecular interactions present should lead to larger contributions to $\rho(\mathbf{r})$ from polarization functions.

A direct descent density-matrix refinement was carried out with the STO-3G basis set, taking fixed anisotropic temperature factors from the multipole refinement and using the same weighting scheme. This resulted in R=0.0424, $\chi^2 = 6.116$, and $S = 2.521$. The scale factor k from density-matrix refinement was 6.06, which is rather different from the 5.90 obtained from multipole refinement. This is probably due to the limited basis set, especially in the core region (only three Cartesian Gaussians per contracted function). Therefore we tried a refinement in the basis STO-6G (still 18 basis functions, but now six Cartesian Gaussians per contracted function). This gave an improved fit: R=0.0406, $\chi^2 = 5.752$, and $S = 2.445$, and the scale factor ($k = 6.03$) was slightly closer to the multipole-refined value. A population analysis of the wave function gave a total of 23.995 electrons.

The description of the scattering used in the density-matrix refinements incorporated no corrections for anomalous dispersion, whereas the multipole refinement had incorporated a correction for core electrons in the usual way.[35] To check that this correction could reasonably be neglected for these first-row atoms, the multipole refinement was repeated with the correction terms f' and f'' set to zero. Refinements with and without these correction terms did not differ significantly.

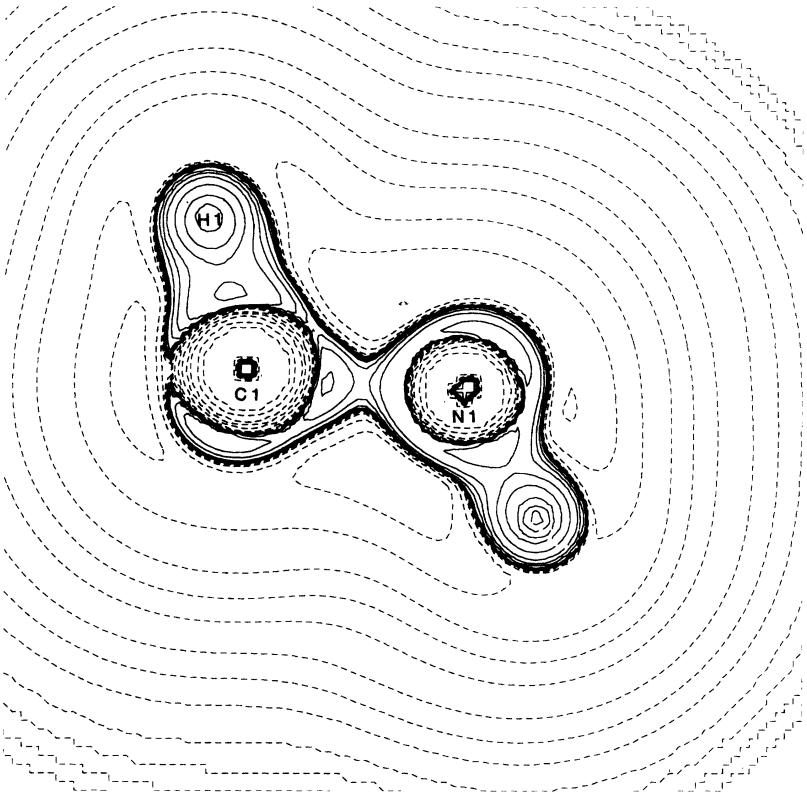
Table III compares CP analyses obtained from RHF 6-31G**, RHF STO-6G, and fitted multipole and STO-6G densities. The STO-6G ρ_c values show no overall better agreement with the 6-31G** results after fitting. Multipole-fitted values of ρ_c are not in conspicuously better agreement with the RHF 6-31G** values than are the density-matrix-fitted ones. The carbonyl and C-N CP's lie further from the 6-31G** positions after refinement, but the CP's in bonds to H atoms lie nearer. The multipole and density matrix-fitted BCP positions are closer to each other than to either of the (free-molecule) *ab initio* densities.

The Laplacian values $\nabla^2\rho_c$ in the two heavy atom bonds show a considerable "improvement" for the fitted density matrix, with a positive value (characteristic of ionic interactions between atoms[31]) in the C-N bond becoming negative (characteristic of covalent interactions), as in the 6-31G** density. The multipole $\nabla^2\rho_c$ values, as with ρ_c, do not agree better on average with the 6-31G** values. More information about the shape of the electron distribution in bonds can be derived from the Hessian of $\rho(\mathbf{r}_c)$. About half of the the fifteen STO-6G Hessian eigenvalues move closer to the corresponding 6-31G** values when the density matrix is refined. The multipole fit Hessian eigenvalues are, on average, no closer to the 6-31G** values than the density matrix fit results.

Figures 3(a)–3(d) are maps of $-\nabla^2\rho$ in the plane containing the C, N and O nuclei. The HF and fitted STO-6G densities [Figs. 3(b) and 3(d)] differ considerably, with the latter showing a much thinner valence shell of charge concentration[31] (VSCC) in the heavy atom bonds—in fact, the oxygen VSCC no longer connects with the carbon atom. The multipole density [Fig. 3(c)] also shows this effect, to a lesser extent. It is apparent that both the experimental (fitted) Laplacians are quite different to the HF 6-31G** result [Fig. 3(a)].

On the whole, it appears that the STO-6G fitted density is not more like the HF 6-31G** density than its HF STO-6G counterpart, but that the two models (multipole and STO-6G density matrix) fitted to the experimental data give similar results. This is encouraging, since one would hope that the fitted density would not be too model dependent. It also suggests that the 6-31G** (free molecule) density is an inadequate benchmark for the various experimental fits. Stevens[32] has pointed out that there are rather large differences between experimental gas phase and in-crystal geometries of formamide, so hydrogen bonding effects in $\rho(\mathbf{r})$ may be considerable. The carbonyl and C-N bond lengths in angstroms are 1.219, 1.352 (from microwave studies[36]),

(a)



(b)

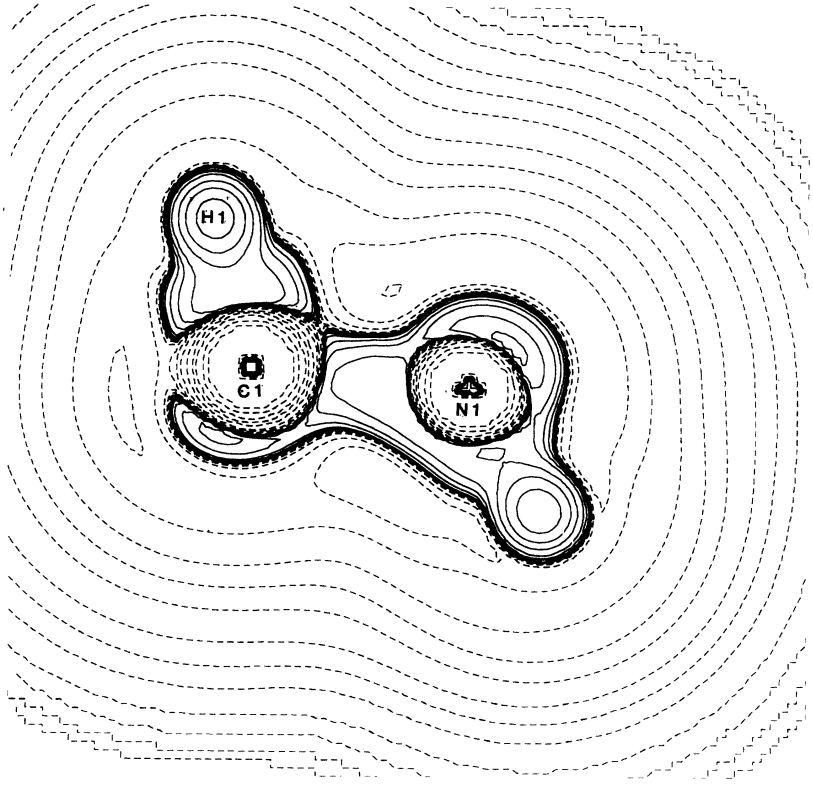


FIG. 2. Laplacian maps $\nabla^2\rho(\mathbf{r})$ for CH_3NH_2. (a) HF/6-311G SDCI. (b) HF/STO-3G. (c) Fit (ii)/STO-3G.

(c)

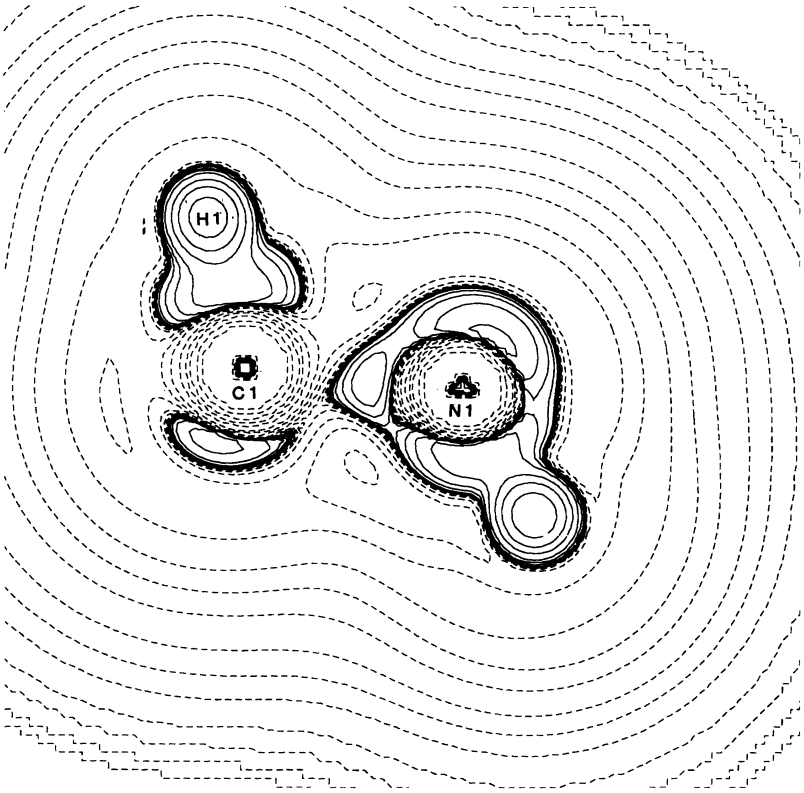
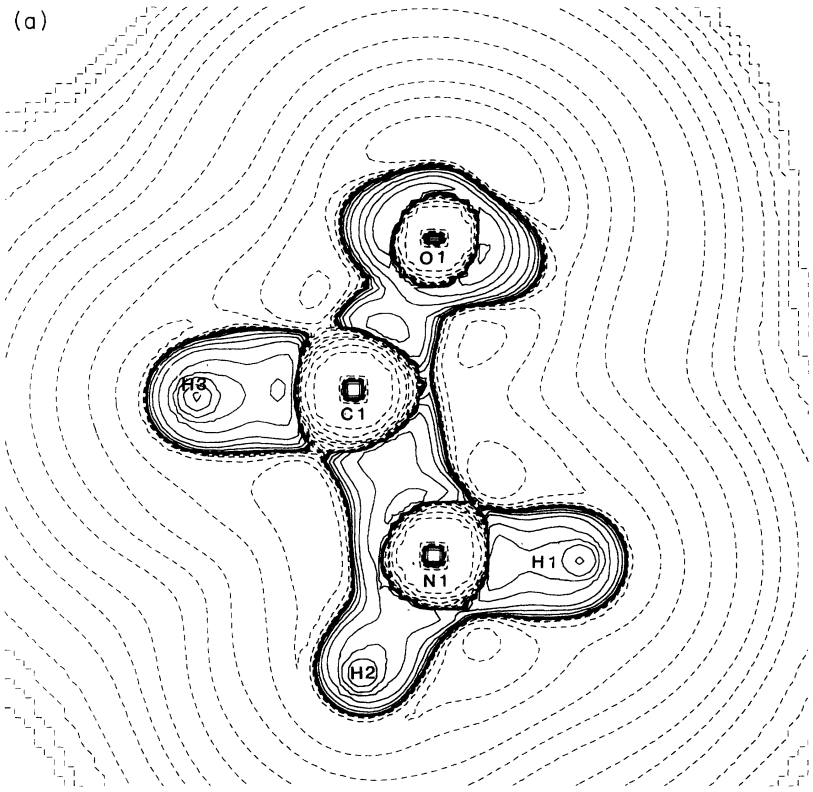


FIG. 2 *(Continued).*

TABLE III. (a) Formamide BCP analysis for heavy atoms. d is the distance of the first-named atomic nucleus to the BCP associated with that bond. (b) Formamide BCP analysis for bonds to hydrogen atoms.

Bond	Method	Hessian eigenvalues (e bohr^{-5}) λ_1	λ_2	λ_3	ρ_c (e bohr^{-3})	$\nabla^2\rho_c$ (e bohr^{-5})	d (bohr)
				(a)			
C=O	6-31G**(HF)	-1.087	-1.015	1.967	0.397	-1.346	0.77
	STO-6G (HF)	-0.705	-0.512	1.999	0.340	+0.782	0.76
	Multipole(exp)	-1.116	-1.022	0.615	0.410	-1.523	0.87
	STO-6G (exp)	-0.688	-0.492	1.342	0.339	+0.163	0.87
C-N	6-31G**(HF)	-0.873	-0.855	1.066	0.350	-0.661	0.81
	STO-6G (HF)	-0.491	-0.408	1.302	0.290	+0.403	0.79
	Multipole(exp)	-0.954	-0.845	0.600	0.380	-1.199	1.10
	STO-6G (exp)	-0.766	-0.492	1.342	0.335	-0.409	1.09
				(b)			
N-H1	6-31G**(HF)	-1.359	-1.285	0.727	0.347	-1.917	1.45
	STO-6G (HF)	-1.096	-1.020	0.836	0.317	-1.279	1.39
	Multipole(exp)	-1.137	-1.042	1.209	0.309	-0.970	1.41
	STO-6G (exp)	-1.133	-1.115	0.966	0.306	-1.283	1.45
N-H2	6-31G**(HF)	-1.377	-1.307	0.760	0.346	-1.924	1.46
	STO-6G (HF)	-1.104	-1.029	0.838	0.317	-1.295	1.39
	Multipole(exp)	-1.132	-1.076	1.283	0.307	-0.925	1.42
	STO-6G (exp)	-1.122	-1.098	1.203	0.309	-1.016	1.43
C-H3	6-31G**(HF)	-0.870	-0.855	0.448	0.308	-1.277	1.28
	STO-6G (HF)	-0.648	-0.628	0.441	0.269	-0.835	1.25
	Multipole(exp)	-0.767	-0.690	0.827	0.268	-1.629	1.37
	STO-6G (exp)	-0.659	-0.618	0.585	0.265	-0.692	1.31



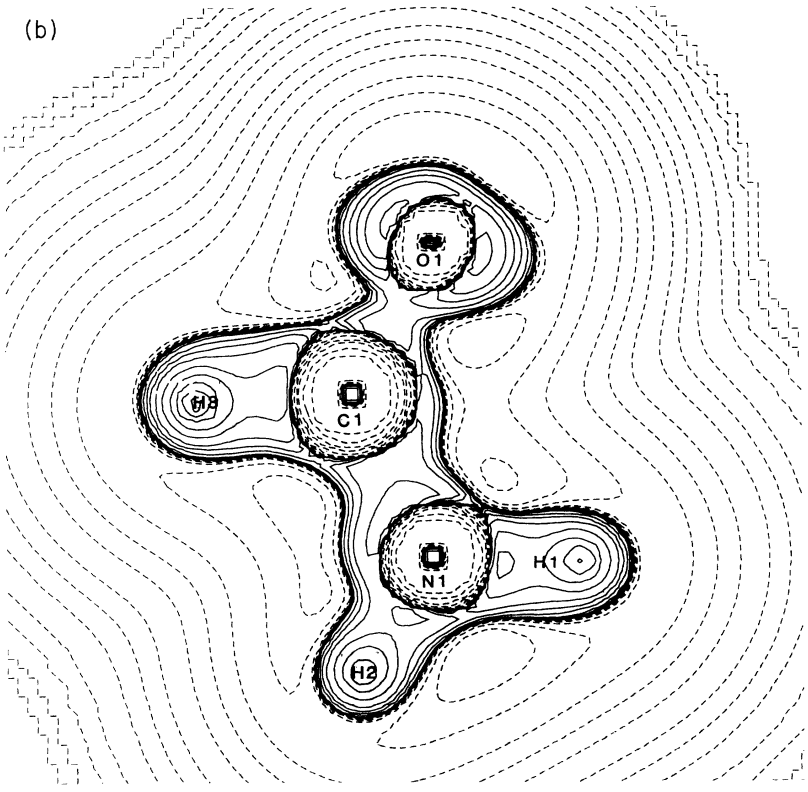


FIG. 3. Laplacian maps $\nabla^2\rho(\mathbf{r})$ for $HCONH_2$. (a) HF/6-31G**. (b) HF/STO-6G. (c) Multipole refined. (d) Density-matrix fit STO-6G.

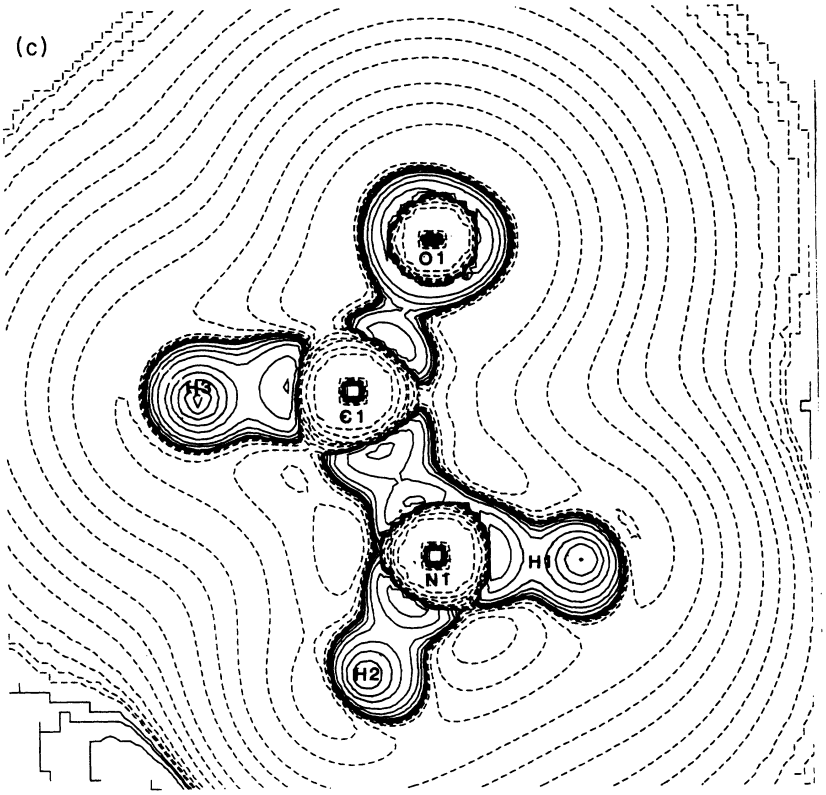


FIG. 3 *(Continued).*

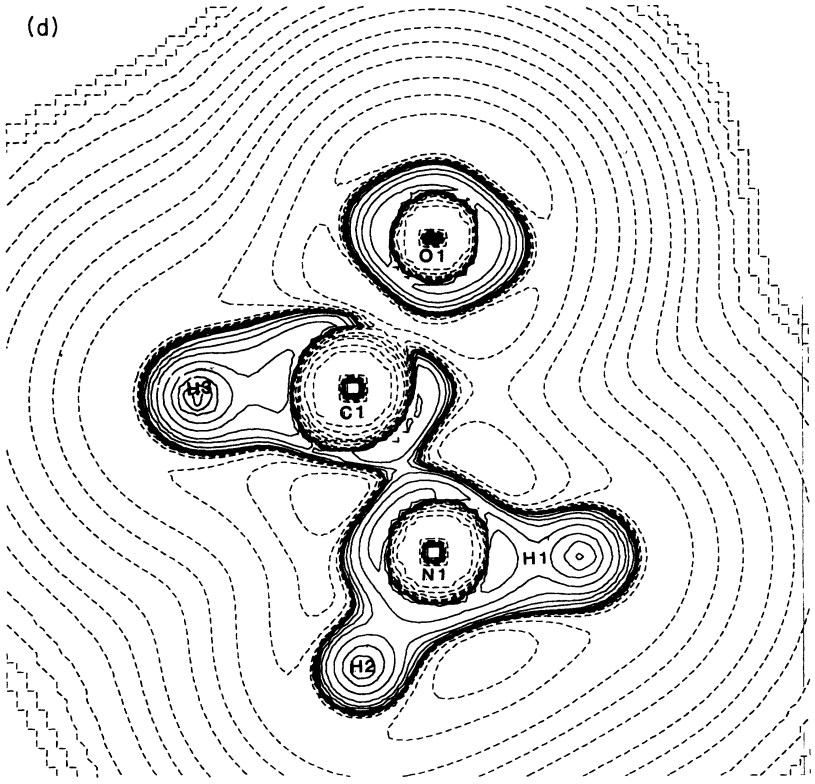


TABLE IV. Formamide population analysis charges.

Atom	HF/6-31G**	HF/STO-6G	Fit/STO-6G
O	-0.66	-0.29	-0.29
N	-0.87	-0.45	-0.52
C	0.71	0.25	0.39
H1	0.33	0.22	0.20
H2	0.34	0.22	0.21
H3	0.12	0.05	0.01

1.242, 1.319 (from this study and x-ray multipole refinement), 1.223, and 1.360 (from a free molecule *ab initio* MP2/6-31G** optimization[33]).

Table IV lists the atomic charges obtained from a Mulliken population analysis[38] of the fitted and HF wave functions. Since such population analysis charges are highly basis-set dependent, the absolute values of these charges mean little. However, as the same basis set and coordinates are used for the HF calculation and the density-matrix fit, the differences may reflect the presence of intermolecular interactions in the latter case. The data suggest that, in the crystal, the N-C bond is more polarized by such interactions than the C-O bond, although the Laplacian maps indicated large differences between (isolated molecule) HF/STO-6G and fitted STO-6G densities in the carbonyl bond. Hydrogen charges are barely altered by refinement.

IV. CONCLUSIONS

This is a viable method for fitting idempotent density matrices, although it is only likely to be useful for small molecules (say ten or fewer atoms). In the study with theoretical data, we were able to show a measurable improvement in the fitted density as compared with the HF density in the same (minimal) basis. In the experimental study, the evidence that the fitted density was superior to the HF density (in the same basis) was less compelling, perhaps largely because the true density is unknown in this case. A multipole model with 88 variables (including thermal parameters for heavy atoms) fitted the experimental data rather better in the sense of R and χ^2 than the 72-variable STO-6G (idempotent) density matrix. There are at least three reasons for this; in probable order of importance: (i) the density matrix fit is tightly constrained by idempotency; (ii) temperature factors were not optimized for the fitted STO-6G density; and (iii) a minimal basis set was used. Temperature factor refinement should appreciably improve the quality of the fit in the density-matrix case.

Multipole models frequently include κ' and κ'' variables on heavy atoms, whereas the LCAO model allows no such radial variation (or basis function optimization). It would be a straightforward task to optimize the basis function exponents simultaneously with the P_{ij} using the random perturbation technique. Here we have chosen to illustrate the results with the simplest, nonoptimized type of LCAO model. The multipole model of the density has the advantage of very accurate HF atomic core functions utilizing STO's, whereas most molecular basis sets necessarily use Gaussian functions; so basis sets with the largest possible number of Gaussians per contracted core function should be used. Since it is only required that one-electron (overlap) integrals be computed in the current fitting procedure, it is feasible to use STO basis sets rather than the usual contractions of Gaussians. These should give much-improved wave functions and may be especially advantageous for modelling x-ray scattering, which is dominated by core-electron contributions. However, the usual problems with computing many-center two-electron matrix elements over Slater-type functions would remain if any two-electron properties were desired. But these are problems only of computational practice: there is no problem in principle.

Based on our limited experience with simulated annealing, it seems that the HF starting guess may be sufficiently close to the global minimum to make it unnecessary. This may not be the case with larger density matrices (larger molecules or basis sets). In cases $N(K-N) > N_{\text{obs}}$ a nonunique fit may be obtained, possibly an exact fit. In such cases the usefulness of any nonunique density matrix is questionable, unless perhaps some other criteria can be applied to choose between density matrices with the same χ^2. For example, this might be the total energy, virial ratio, or Hellmann-Feynman forces on the nuclei.

The computational effort involved in fitting an idempotent density matrix with this technique is not negligible. Without refinement of the parameters describing thermal motion, approximately 12 h on a VAX 4060 workstation are required for formamide in a minimal (STO-6G) basis. If the thermal motion is also refined this increases considerably, since the O_{ijk} in Eq. (7) must be recomputed every cycle. In comparison to multipole refinement, this is then a factor of 100–1000 times slower. However, we anticipate that fitting larger basis sets, coupled with optimization of exponents and (perhaps) more realistic core basis functions (STO's), will give a fit competitive with (or better than) the multipole one. Since the fitted wave function obtained from this study is of rather low quality, we have not attempted to evaluate any two-electron properties from it, but this remains an important aim of future studies.

ACKNOWLEDGMENTS

We are grateful to Professor R. Nalewajski of the Jagiellonian University in Kraków for stimulating discussions which contributed to this work. S.T.H. and C.S.F. acknowledge the SERC for support while this work was being undertaken.

APPENDIX

We want to show that McWeeny's iterative procedure finds the closest $n \times n$ symmetric idempotent matrix P to a given symmetric matrix A, provided that A and P are not too far apart (in the sense of Frobenius norm),[37]

$$\|A-P\|^2=\sum_{i=1}^{n}\sum_{j=1}^{n}(A_{ij}-P_{ij})^2.$$

We begin by showing that if P_* is a matrix such that $\|A-P\|^2$ is a minimum at $P=P_*$, then A commutes with P_*. To show this we use Lagrange multipliers to minimize $\|A-P\|^2$ subject to the constraint that P be idempotent, that is, $\sum_k P_{ik}P_{kj}=P_{ij}$ for each i,j. (We assume for simplicity that the basis set is orthonormal.) Say λ_{ij} is a Lagrange multiplier for the i,jth constraint, then we minimize

$$E=\sum_{i=1}^{n}\sum_{j=1}^{n}\left[(a_{ij}-P_{ij})^2-\lambda_{ij}\left(\sum_k P_{ik}P_{kj}-P_{ij}\right)\right] \tag{A1}$$

with respect to the elements of P.

Evidently

$$\frac{\partial E}{\partial P_{mn}}=-2(a_{mn}-P_{mn})-\sum_{i=1}^{n}\lambda_{in}P_{im}-\sum_{j=1}^{n}\lambda_{mj}P_{nj}+\lambda_{mn}=\left[-2(A-P)-P^\dagger\Lambda-\Lambda P^\dagger+\Lambda\right]_{mn},$$

where Λ is the matrix of Lagrange multipliers. Hence

$$2(A-P_*)+P_*^\dagger\Lambda+\Lambda P_*^\dagger-\Lambda=0.$$

Multiplying on the left by P_*, and using the facts that $P=P^\dagger$ and $P_*^2=P_*$ this reduces to $2P_*A-2P_*+P_*\Lambda P=0$. Similarly, multiplying on the right by P_* gives $2AP_*-2P_*+P_*\Lambda P_*=0$, and so $P_*A=AP_*$.

It follows that A and P_* can be diagonalized simultaneously. Since the Frobenius norm is invariant under orthogonal transformations

$$\|A-P\|^2=\sum_i(a_i-p_i)^2,$$

where a_i are the eigenvalues of A (ordered most positive first) and the p_i are the eigenvalues of P_*. If P_* is of rank k, then k of the p_i equal 1 and the others equal zero. It follows that

$$\|A-P\|^2\geq\sum_{i=1}^{k}(a_i-1)^2+\sum_{i=k+1}^{n}a_i^2$$

Furthermore, we can find an idempotent matrix which has precisely this distance from A, by performing the eigenvalue decomposition of A and replacing its eigenvalues with 1 or 0 appropriately.

All that remains to be shown is that the McWeeny procedure yields the P_* described in the preceding sentence. The iterative procedure is defined by $P_{n+1}=3P_n^2-2P_n^3$, with $P_0=A$. Now suppose v is an eigenvector of P_n with eigenvalue λ; then

$$P_{n+1}v=(3P_n^2-2P_n^3)v=(3\lambda^2-2\lambda^3)v,$$

so v is an eigenvector of P_{n+1} with eigenvalue $3\lambda^2-2\lambda^3$. It follows that if A has the eigenvalue decomposition $C_A\Lambda_AC_A^\dagger$, then applying the McWeeny procedure will generate a sequence of matrices $C_A\Lambda_nC_A^\dagger$ with $\Lambda_{n+1}=3\Lambda_n^2-2\Lambda_n^3$ and $\Lambda_0=\Lambda_A$. The procedure will converge if the sequence $\lambda_{n+1}=3\lambda_n^2-2\lambda_n^3$ converges for $\lambda_0=a_i, i=1,\ldots,n$.

It is easy to see that the system $\lambda\rightarrow 3\lambda^2-2\lambda^3$ has two stable fixed points 0 and 1 (and also one unstable one at $\lambda=1/2$). Thus if the McWeeny procedure converges, it will converge to an idempotent matrix. Let us suppose that A is the result of perturbing some idempotent matrix $\hat{P}$, where $\hat{P}$ has k unit eigenvalues and $n-k$ zero eigenvalues. We suppose the perturbation is so small that A has k eigenvalues in the range $1/2$ to $(1+\sqrt{3})/2$ and $n-k$ in the range $(1-\sqrt{3})/2$ to $1/2$. Then the procedure will clearly converge to produce k unit eigenvalues and its result will be to replace the k largest eigenvalues of A by 1 and the others by 0. As we saw above, the resulting idempotent matrix minimizes $\|A-P\|^2$.

[1] P. Coppens, J. Phys. Chem. **93**, 7979 (1989), and references therein.

[2] *The Application of Charge Density Research to Chemistry and Drug Design*, Vol. 250 of *NATO Advanced Study Institute, Series B: Physics*, edited by G. A. Jeffrey and J. F. Piniella (Plenum, New York, 1991).

[3] *Sagamore X, Conference on Charge, Spin and Momentum Densities: Collected Abstracts, Konstanz, Germany*, 1991, edited by M. Springborg, A. Saenz, and W. Weyrich (University of Konstanz, Konstanz, Germany, 1991).

[4] R.F. Stewart, J. Chem. Phys. **58**, 1668 (1973).

[5] N. K. Hansen and P. Coppens, Acta Cryst. A **34**, 909 (1978).

[6] S. Swaminathan, B. M. Craven, M. A. Spackman, and R. F. Stewart, Acta Cryst. B **40**, 398 (1984).

[7] M. P. C. M. Krijn, H. Graafsma, and D. Feil, Acta Cryst. B **44**, 609 (1988).

[8] S. T. Howard, M. B. Hursthouse, C. W. Lehmann, P. R. Mallinson, and C. S. Frampton, J. Chem. Phys. **97**, 5616 (1992).

[9] G. Moss and D. Feil, Acta Cryst. **37**, 414 (1981).

[10] D. Feil, and G. Moss, Acta Cryst. **39**, 14 (1983).

[11] M. A. Spackman, J. Chem. Phys. **85**, 6579 (1986).

[12] M. P. C. M. Krijn and D. Feil, J. Chem. Phys. **89**, 4199 (1988).

[13] T. L. Gilbert, Phys. Rev. B **12**, 2111 (1975).

[14] J. E. Harriman, Phys. Rev. A **24**, 680 (1981).

[15] J. E. Harriman, Phys. Rev. A **34**, 29 (1986).

[16] R. F. Stewart and D. Feil, Acta Cryst. A **36**, 503 (1980).
[17] R. McWeeny, Rev. Mod. Phys. **32**, 335 (1960).
[18] D. S. Jones, D. Pautler, and P. Coppens, Acta Cryst. A **28**, 635 (1972).
[19] D. Feil *et al.* (unpublished).
[20] P. Coppens, T. V. Willoughby, and L. N. Csonka, Acta Cryst. A **27**, 248 (1971).
[21] W. L. Clinton, C. A. Frishberg, L. J. Massa, and P. A. Oldfield, Int. J. Quantum Chem. Symp. **7**, 505 (1973).
[22] M. Levy and J. A. Goldstein, Phys. Rev. B **35**, 7887 (1987).
[23] J. P. Huke (unpublished).
[24] L. M. Pecora, Phys. Rev. B **33**, 5987 (1986).
[25] L. Massa, M. Goldberg, C. Frishberg, R. F. Boehme, and S. J. La Placa, Phys. Rev. Lett. **55**, 622 (1985).
[26] Y. V. Aleksandrov, V. G. Tsirelson, I. M. Reznik, and R. P. Ozerov, Phys. Status Solidi. **155**, 201 (1989).
[27] R. F. Stewart, J. Chem. Phys. **51**, 4569 (1969).
[28] E. D. Stevens, J. Rys, and P. Coppens, Acta Cryst. A **33**, 333 (1977).
[29] M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. Robb, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol, and J. A. Pople, Gaussian 90, Revision F, Gaussian Inc., Pittsburgh, PA, 1990.
[30] M. Atoji and W. N. Lipscomb, Acta Cryst. **6**, 770 (1953).
[31] R. F. Bader, *Atoms in Molecules—A Quantum Theory* (Oxford University Press, Oxford, 1990).
[32] E. D. Stevens, Acta Cryst. B **34**, 544 (1978).
[33] S. T. Howard and J. P. Huke (unpublished).
[34] T. Koritzansky, *LSMOL Multipole Refinement Package* (State University of New York, Buffalo, 1987).
[35] *International Tables For Crystallography* (Kynoch Press, Birmingham, 1974), Vol. IV.
[36] E. Hirota, R. Sugisaki, C. J. Nielsen, and G. O. Sorensen, J. Mol. Spectrosc. **49**, 251 (1974).
[37] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985).
[38] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).