INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again – beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

75-25,364

COHEN, Carol Frishberg, 1947-
A METHOD FOR OBTAINING N-REPRESENTABLE ELECTRON DENSITIES FROM X-RAY DIFFRACTION DATA.

The City University of New York, Ph.D., 1975
Chemistry, physical

Xerox University Microfilms, Ann Arbor, Michigan 48106

A Method for Obtaining N-Representable Electron Densities from X-ray Diffraction Data

by

Carol Frishberg

A dissertation submitted to the Graduate Faculty in Chemistry in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1975

This manuscript has been read and accepted for the Graduate Committee in Chemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

June 10, 1975
date

Chairman of Examining Committee

June 10, 1975
date

Leonard H. Schwartz
Executive Officer

David L. Beveridge
Dr. David L. Beveridge

C. R. Fischer
Dr. C. Rutherford Fischer

Angelo Santoro
Dr. Angelo Santoro

The City University of New York

Abstract

A Method for Obtaining N-Representable Electron Densities from X-ray Diffraction Data

by

Carol Frishberg

Adviser: Professor Louis Massa

An iterative matrix equation for generating idempotent one-body reduced density matrices which give a least absolute value fit to a set of x-ray scattering data is presented and its properties studied. The data may be in either intensity or structure factor form. There is no restriction upon the type of basis functions employed; the density need not be described by a sum of spherical atoms. Other constraints upon ρ_1 can easily be enforced. The density matrices found are automatically N-representable by a single Slater determinant, which appears to be a suitable approximation for the highest quality x-ray data presently being achieved. The algorithm is shown to converge to satisfactory solutions for simple atomic and molecular (model) systems. The method is employed in simple cases to determine structural parameters. The algorithm as formulated is not in general fully constrained and there exists the possibility that the solution found will be influenced by the initial guess matrix used. Some suggestions for alleviating this problem are given.

Conventional least squares analyses are carried out in some cases for comparison to the idempotent results. It is found that the unconstrained solutions always fit better then the idempotent solutions but are never N-representable. The not-idempotent least squares solutions are frequently found to contain regions of negative density. The idempotent matrix has fewer parameters than the least squares determined matrix and is unaffected by any correlation among the parameters.

A steepest descent method in which the sums of the squares of the residuals of the intensity data is minimized subject to the constraint that the one-body reduced density matrix be idempotent, is formulated and investigated.

Acknowledgment

I would like to express my gratititude to Professor Louis Massa for his guidance and assistance during the course of this work.

Table of Contents

I. Introduction 1

II. Formalism 14

III. Application of Algorithm 29

IV. The Phase Problem 68

V. Correlation of Parameters 72

VI. Method of Steepest Descent 77

VIII. Conclusions and Suggestions for Future Work 81

Appendix I - Factorization of an Idempotent Matrix Referred to a Non-Orthogonal Basis 85

Appendix II - Method of Least Squares 87

References 90

List of Figures

III - 1. Hydrogen atom, $\xi = 0.90$ 48
III - 2. Hydrogen atom, $\xi = 1.00$ 49
III - 3. Hydrogen atom, $\xi = 1.12$ 50
III - 4. Hydrogen atom, $\xi = 1.2$ 51
III - 5a. Hydrogen atom, $\xi = 5.0$, not-idempotent solution 52
III - 5b. Hydrogen atom, $\xi = 5.0$, idempotent solution 53
III - 6a. Hydrogen atom, $\xi = 7.0$, not-idempotent solution 54
III - 6b. Hydrogen atom, $\xi = 7.0$, idempotent solution 55
III - 7. Model molecular density, exact basis 59
III - 8. Model molecular density, $\xi = 1.05\xi$exact 60
III - 9. Model molecular density, $\xi = 0.95\xi$exact 61
III - 10. Model molecular density, $\xi = 0.85\xi$exact and 1.5ξexact 62
III - 11. Hydrogen molecule 64
III - 12. Hydrogen molecule 65
III - 13. Hydrogen molecule 66
III - 14. Hydrogen molecule 67

IV - 1. Determination of bond length for a model density 70
IV - 2. Determination of bond length for the hydrogen molecule 71

List of Tables

II - 1. Comparison of the number of independent parameters in several $\underset{\sim}{P}$-matrices 28

III - 1. Hydrogen atom analysis using two basis functions, comparison of intensity, structure factor and steepest descent solutions 45

III - 2. Hydrogen atom analysis using four basis functions, illustration of initial guess dependency 46

III - 3. Hydrogen atom solutions, comparison of idempotent and not-idempotent densities for $\xi = 0.90, 1.00, 1.12, 1.20, 5.0, 7.0$ 47

III - 4. Model molecular density data 58

I. Introduction

The traditional task of the crystallographer has been to deduce from a set of intensities, i.e., amplitudes, the phases of the waves and determine the positions of all atoms in the unit cell. The expression for a wave scattered coherently by an electron density distribution, $\rho(\vec{r})$, relative to one electron at the origin is given by

$$F(\vec{S}) = \int \rho(\vec{r})\, e^{i\vec{r}\cdot\vec{S}}\, d\vec{r} \quad . \tag{1}$$

$\vec{S}$ is the vector difference of the emerging and impinging waves; its magnitude is sin θ/λ where 2θ is the scattering angle and λ is the wavelength of the x-rays. For the periodic density distribution of a crystal the integral (1) is taken over the volume of a unit cell. Only the intensities,

$$I(\vec{S}) = F(\vec{S})F^{*}(\vec{S}) \tag{2}$$

of the scattered waves are observable. In an x-ray diffraction experiment, where the wavelength of the radiation is comparable to the dimensions of the unit cell, the intensities have non-zero values only at those points in reciprocal space where the vector $\vec{S}$ satisfies the Laue conditions. The intensities are real quantities which are origin independent. They lack the phase information contained in the structure factors, equation (1), which if known, would allow us to obtain the electron density directly from its diffraction pattern. The crystal structure follows easily from the electron density.

A typical structure determination might proceed in the following way (see Stout & Jensen 1968 for more details). First, the raw data is put in more usable form by correcting for extinction and polarization effects, systematic errors arising from geometric factors of the particular experiment, and the absorbtion of radiation. Next a starting approximation, called a phasing model, is determined, most likely by "direct" methods or by construction and examination of a Patterson map. Once an approximate structure has been found Fourier synthesis techniques are generally employed to find the atomic positions. The electron density of the unit cell is assumed to be composed of the linear sum of the spherical electron densities of the atoms it contains.

$$\rho(\vec{r}) = \sum_{\substack{n \\ \text{atoms}}} \rho_n(\vec{r}) \tag{3}$$

To calculate the crystal structure factors based on this spherical atom model, one need only know the Fourier transforms of the electron densities of the various isolated atoms.

$$F(\vec{S}) = \sum_{\substack{n \\ \text{atoms}}} f_n(|\vec{S}|)\, e^{i\vec{r}_n \cdot \vec{S}}\, T_n(\vec{S}) \tag{4}$$

where $$f_n(|\vec{S}|) = \int \rho_n(\vec{r})\, e^{i\vec{r}\cdot\vec{S}}\, d\vec{r} \tag{5}$$

and $\vec{r}_n$ defines the location of the nth atom in the unit cell. Tn is a thermal smearing function that accounts for the fact that the atoms are vibrating about their equilibrium positions. The atomic $f(|\vec{S}|)$'s are calculated from isolated atom Hartree-Fock wave functions and tabulated in the literature for routine use by crystallographers. In a Fourier difference synthesis the observed structure amplitudes are subtracted from the calculated structure factors, assuming their phases are equal to the calculated phases. The resulting "difference structure factor" is Fourier inverted to find a difference density.

$$\Delta\rho(\vec{r}) = \frac{1}{V}\sum_{\vec{S}} |F_{calc} - F_{obs}|\, e^{-i\vec{r}\cdot\vec{S}} \tag{6}$$

The sum is taken over all observed reflections although theoretically an infinite number of terms should contribute. This difference density function reflects the ways in which the true density differs from the assumed model. The crystallographer seeks clues from $\Delta\rho$ that will help him improve his model so that the process may be repeated until satisfactory agreement is obtained. The final step is a least squares refinement in which all atomic position and thermal parameters are modified slightly to improve the fit to the experimental data.

The forgoing outline of a structure determination is greatly oversimplified. Since 1912 when von Laue did the first x-ray experiment a vast number of techniques have evolved to unravel diffraction patterns. Each data set is unique and the crystallographer must make a decision at each step as to how to proceed. Until recently all alternatives were founded on the assumption that a crystal was composed of isolated atoms arranged on a periodic lattice.

The spherical atom approximation has been extraordinarily successful in structure determination. It is apparent from equation (1) however that the x-ray diffraction experiment can potentially provide much more detailed information about the crystalline electron density. The appearance of the "forbidden" 222-reflection in diamond indicates that non-spherical bonding electron density and anisotropic thermal vibrations do indeed modify the diffraction pattern. One reason that partially accounts for the good fits to experimental data is that unrealistically large anisotropic temperature factors can simulate bonding density when spherical atomic form factors are used. It has also been found (Coppens 1968) that the positions of first row atoms will be altered by about 0.1 A° to 0.2 A° to more adequately represent bonding or lone pair electrons. In an analysis of very accurate data for cyanuric acid, Vershoor and Keulen (1971) found that the structural parameters determined will be affected by the choice of atomic form factors.

Attempts have been made to achieve more accurate structures but still maintain spherical form factors. One approach is to take advantage of the fact that since bonding electrons are more diffuse than core electrons they contribute practically nothing to the scattering at high angles. Low angle reflections are either eliminated or weighted less than high angle reflections in the refinement process. Alternatively, Little, Pautler and Coppens (1971) have used a double atom refinement in which valence electrons are allowed to float away from the core positions.

Although these methods have eliminated some inconsistencies they still do not always agree with atomic position and thermal parameters derived from neutron scattering experiments. Since neutrons are scattered by nuclei rather than electrons, for non-magnetic systems, the parameters obtained by this method will be unaffected by aspherical electron density. It is possible to use this fact to see the effect of molecule formation on charge distribution. A difference density is calculated from equation (6) using position and thermal parameters determined by the neutron scattering experiment, with the isolated atom form factors generally used in x-ray work. These difference functions can then be compared to theoretical studies in which atomic densities are subtracted from molecular densities of comparable quality. The resulting density is then convoluted with a thermal smearing function derived from the unbiased neutron experiment. Studies of this sort on cyanuric acid (Jones, Pautler and Coppens 1972) show that satisfactory agreement is only obtained by extended basis Hartree-Fock calculations. Semi- empirical INDO (McIver, Coppens and Nowack 1972) and minimal basis ab-initio calculations fail to reproduce the quantitative and sometimes qualitative aspects of the difference density map.

The above results seem to indicate that present day accuracy of x-ray data contains information about the details of electron density in crystals. Matthews, Stucky and Coppens (1972) have a data set for tetracyanoethylene oxide with an internal consistency of 0.8%. Aldred and Hart (1973) have reported intensities for silicon which have been measured to an internal consistency of 0.1%. Data of this quality cry out for more accurate theoretical analysis.

The first step is to get away from the restrictions imposed by using spherical atomic form factors. Dawson (1967) included Kubic Harmonics in calculating form factors for carbon in order to analyze diamond data and reveal information about bonding. McWeeny (1951) suggested the use of "effective" scattering factors calculated by partitioning the molecular charge to the various atoms such that it is completely reproduced. He used a Heitler-London wave function to calculate an effective scattering factor for a bonded hydrogen atom. Following this idea, Stewart, Davidson, and Simpson (1965) used the more sophisticated Kolos-Roothaan wave function for H_2 to obtain a non-spherical but localized H-atom form factor. The success of this approach depends upon the transferability of an atom from molecule to molecule. A more flexible model should yield better results.

Jones and Lipscomb (1970) calculated molecular structure factors (equation (1)) for four different geometries of diborane from SCF minimum basis set LCAO MO wave functions. These molecular structure factors were fitted to the experimental results using the technique of least squares by varying the thermal motion parameters. A plot of the agreement factor vs. bond lengths and bond angles were used to obtain the best structural parameters. They found the B-H bond length to be 0.05A° longer than determined by the spherical atom analysis; a value which accounts for about two-thirds of the discrepancy between x-ray and electron scattering results. This approach involves solving SCF equations for several geometries, fourier inverting the electron densities obtained and finally fitting all sets of structure factors to the observed data. It is considerably more expensive than a method which fits structure factors directly.

Following Stewart's (1969) work on generalized scattering factors, Coppens and coworkers (Coppens, Willoughby and Csonka 1971; Coppens, Pautler and Griffin 1971; Matthews, Stuckey and Coppens 1972; Jones, Pautler and Coppens 1972) have developed and used a formalism which allows for great flexibilty in the charge density. This formalism includes bond form factors as well as atomic form factors and allows variable occupancies. The electron density is expressed as the sum of the electron densities of each molecular orbital

$$\rho(\vec{r}) = \sum_{i}^{occ} \phi_i(\vec{r})\, \phi_i^*(\vec{r}) \tag{7}$$

Each molecular orbital is further expanded in an atomic basis

$$\phi_i(\vec{r}) = \sum_{\mu} c_{i\mu} \Psi_{\mu}(\vec{r}) \tag{8}$$

The total electron density can then be written in terms of a charge and bond order matrix $\underset{\sim}{P}$,

$$\rho(\vec{r}) = \sum_{\mu} \sum_{\nu} P_{\mu\nu} \Psi_{\nu}(\vec{r}) \Psi_{\mu}^*(\vec{r}) \tag{9}$$

whose elements are defined by

$$P_{\mu\nu} \equiv \sum_{i} c_{\mu i}^* c_{\nu i} \tag{10}$$

The elements of $\underset{\sim}{P}$ are parameters to be determined by a least squares analysis of accurate x-ray data. To insure that correlation between vibrational motion and population coefficients does not affect the convergence of the least squares equations, the temperature factors and atomic positions are fixed by analysis of neutron scattering data. Although this doubles the experimental effort involved, Coppens (1967) feels it is a necessary step to arrive at meaningful population coefficients. The charge densities of cyanuric acid (Jones, Pautler and Coppens 1972; Coppens, Pautler and Griffin 1971; Coppens and Vos 1971), tetracyanoethylene oxide (Matthews, Stucky and Coppens 1972; Matthews and Stucky 1972), 1,3,5- triacetylbenzene (O'Connor and Maslen 1974), oxalic acid dihydrate (Coppens, Pautler and Griffin 1971) and fumaramic acid (Hirshfeld 1971) have been investigated with this procedure.

One of the major drawbacks of this formalism is the large number of parameters to be determined. For a least squares fit to make any sense there must be many more experimental points then parameters. Estimates for the optimum ratio for this type of analysis vary from 8:1 to 10:1 (personal communications to L. Massa). The number of observed reflections is limited by the wavelength of the radiation used and the symmetry of the crystal. To deal with this problem all analyses to date have taken $P_{\mu\nu}$ equal to zero unless basis functions μ and ν are on the same atom or adjacent atoms. Refinements (Coppens, Pautler and Griffin 1971; Matthews, Stucky and Coppens 1972) in which all two- center terms were neglected failed to produce satisfactory results. Although the contribution to the charge density from overlap terms on non-adjacent atoms may be small, going from a two-center to an all-center calculation on diborane reduced the agreement index by a factor of six (Jones, Pautler and Coppens 1972) The number of parameters is sometimes reduced further by imposing symmetry constraints that are only approximately adhered to by the molecule (Hirshfield 1971). All core electrons will be frozen at their Hartree-Fock values; only the populations of the valence electrons will be allowed to change. A recent study by Bentley and Stewart, (1974) using atomic and molecular wave functions near the restricted Hartree-Fock limit, calculated that the error in scattering factors for first row diatomics due to deformation of the core electrons upon binding is less than 1.%. This is within the experimental error of most x-ray work done today, but clearly not all if one refers to Aldred & Hart (1973). The authors point out, however, that certain molecular properties may be grossly in error if this effect is neglected, in particular local electric fields and field gradients.

Strictly speaking the results of this method do not adhere to its formalism. To write the charge density in the form of equation (7) implies the orthonormality of the molecular orbitals, ϕ_i. However, no constraints are placed upon the $\underset{\sim}{P}$- matrix to require that this is so. Even if one is to obtain a charge density which is a good fit to the experimental data, there is no assurance that the density so obtained describes any physically realizable system.

Determining whether a charge density is quantum mechanically acceptable is part of a larger problem, named N-representability by A.J. Coleman in 1963, but previously recognized by others (see Löwdin 1955 and McWeeny 1960). This well known dilemma can be stated in general terms as follows. Given a normalized, antisymmetric m-body function, where $m<N$, what are the necessary and sufficient conditions to insure that it is related to an antisymmetric N-body wave function by:

$$\rho_m(1,2,\ldots m;\, 1',2',\ldots m') = \int \Psi^*(1',2'\ldots m', m+1\ldots N)\,\Psi(1,2\ldots m, m+1\ldots N)\, d\tau_{m+1}\cdots d\tau_N\,, \tag{11}$$

or to a weighted sum of such wave functions. To find the expectation value of any m- body operator in this formalism one need only know the m-body reduced density matrix.

$$\langle\Theta\rangle = \int \hat{\Theta}\,\rho_m\,(1,2\ldots m;\, 1',2'\ldots m)\, d\tau_1\cdots d\tau_m \tag{12}$$

where the primed coordinates are immune to operation but are unprimed before the integration is carried out. For $m=1$ and $1'=1$, expression (10) reduces to the charge density. For most problems of chemical interest all necessary information is contained in ρ_2 . If one takes the Hamiltonian for the operator in equation (12), the variational principle could be applied directly to ρ_2 instead of the full N-body wave function to find an approximation to the ground state. The danger in doing this without imposing sufficient N-representability restraints on ρ_2, however, is that one can obtain an energy lower than the ground state! Mayer (1955) tried to do this for the case of a gas of electrons immersed in a uniform positive charge . His results were in marked contrast to similar systems with fewer particles which can be solved exactly. Tredgold (1957) points out that this is due to a lack of N-representability in Mayer's trial function, and then goes on to show what physical contradictions may be obtained if sufficient N-representability conditions are not enforced.

Although the appropriate conditions on the 1-matrix are known (its eigenvalues must be between 0 and 1), there is to date no practicable solution to the N-representability problem for the 2-matrix. In fact, it appears (Coleman 1973) that there may never be N-representability conditions which will make it any easier to deal with the 2-matrix than with Ψ itself. In contrast to the 1-matrix case where conditions on the eigenvalues alone are necessary and sufficient, N-representability forces complex relations among the 2-matrix eigenfunctions.

This dilemma has made the technique of working directly with the reduced density matrices rarely employed. The exception to this is the case where Ψ is taken to be a single Slater determinant. The wave function is then completely determined by the 1-matrix which must be idempotent. In 1969 a series of papers by Clinton and co-workers exploited this property of idempotency and developed a simple iterative equation to find such a $\underset{\sim}{p}$-matrix constrained to yield any number of expectation values one wishes to impose. This algorithm was applied to first row diatomic molecules using the virial and electrostatic theorems as constraints. Following Clinton, we have developed an algorithm to find an idempotent P-matrix as defined by equation (9) which is fit in a "least squares" manner to a set of x-ray diffraction intensities. It is automatically N-representable by a single Slater determinant and can be used to determine any physical properties of the crystal.

This is the first time a "least squares" - type constraint (to be defined later) has has been used in the P-equations, although in 1972, Clinton and Massa published a calculation which found an idempotent 2 x 2 matrix which was fit to a set of theoretical crystallographic constraints. This matrix had only one parameter which was found by a least squares analysis of the data. The analogous non-idempotent matrix was determined in the same manner; this matrix contained two adjustable parameters. The authors found that while the non-idempotent matrix always gave a better fit to the data, it was never a quantum mechanically acceptable density. Also, when systematic errors were introduced into the data the non-idempotent matrix was more seriously affected than the idempotent solution.

II Formalism

If we take as our wave function one normalized, antisymmetrized product of ortho- normal molecular orbitals,

$$\Psi(1,2\ldots N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) \cdots & \phi_1(N) \\ \phi_2(1) & \phi_2(2) \cdots & \phi_2(N) \\ \vdots & & \vdots \\ \phi_N(1) & \cdots \quad \cdots & \phi_N(N) \end{vmatrix} \tag{1}$$

it is completely characterized by the 1-particle reduced density matrix (Lowdin 1955) defined by equation (I.10).

$$\rho_1(1,1') = \sum_N \phi_N(1)\,\phi_N^*(1) \tag{2}$$

All cross products in expression (2) are zero due to the ortho-normality of the molecular orbitals. Writing (2) in matrix notation and expanding the ϕ's in the discrete basis $\{\psi\}$, whose metric is the matrix $\underset{\sim}{S}$,

$$\underset{\sim}{\phi} = \underset{\sim}{c}\,\underset{\sim}{\psi} \quad , \tag{3}$$

we get,

$$\rho_1(1,1') = \mathrm{tr}\,\underset{\sim}{\phi}(1)\,\underset{\sim}{\phi}^{\dagger}(1') = \mathrm{tr}\,\underset{\sim}{c}\,\underset{\sim}{\psi}(1)\,\underset{\sim}{\psi}^{\dagger}(1')\,\underset{\sim}{c}^{\dagger} \tag{4}$$

$$\rho_1(1,1') = \mathrm{tr}\,\underset{\sim}{\rho}\,\underset{\sim}{\psi}(1)\,\underset{\sim}{\psi}^{\dagger}(1') \; ; \quad \underset{\sim}{\rho} \equiv \underset{\sim}{c}^{\dagger}\underset{\sim}{c} \tag{5}$$

where $\underset{\sim}{\phi}$ and $\underset{\sim}{\psi}$ are column matrices. It can be shown (Appendix I) that the conditions

$$\underset{\sim}{P} = \underset{\sim}{P}^{+} \tag{6a}$$

$$\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{S} = N \tag{6b}$$

$$\underset{\sim}{P}\underset{\sim}{S}\underset{\sim}{P} = \underset{\sim}{P} \tag{6c}$$

are necessary and sufficient for ρ_1 to be N-representable by a single Slater determinant. One-body expectation values of the operator $\hat{\theta}$ are given by

$$\langle\hat{\theta}\rangle = \mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{\hat{\theta}} \tag{7}$$

where $\underset{\sim}{\hat{\theta}}$ is the matrix representative of $\hat{\theta}$ in the $\{\psi\}$ basis:

$$\underset{\sim}{\hat{\theta}}_{ij} \equiv \int \psi_i^{*}\,\hat{\theta}\,\psi_j\,d\tau \tag{8}$$

Expectation values of operators involving two or more electrons at a time can be calculated from $\underset{\sim}{P}$, but the process is slightly more complicated than expression (7).

We seek a matrix, $\underset{\sim}{P}$, which satisfies equations (6) and gives the best possible fit to a set of x-ray diffraction data. By reducing the quantity $\mathrm{tr}(PSP-P)^2$ to zero, we derive an iterative equation for an idempotent (in the sense of (6c)) matrix. The Lagrangian multiplier technique is used to assure that our matrix is normalized (equation (6b)) and fit in a "least absolute value" manner to a set of empirical data.

$$\sum_{\vec{s}} w(\vec{s}) \left| \mathrm{tr}\, \underset{\sim}{P}\underset{\sim}{\Theta}(\vec{s}) - \langle \Theta(\vec{s}) \rangle \right| = \epsilon \tag{9}$$

where $w(\vec{s})$ is a weighting factor and ϵ is a small number. This type of constraint is applicable when there are more data points than parameters, and no point must be satisfied exactly. In equation (9) the absolute value rather than the more conventional square is taken to make the problem more tractable. The merits of this will be discussed later.

$$\begin{aligned}\delta\Big\{ & \mathrm{tr}(\underset{\sim}{P}\underset{\sim}{S}\underset{\sim}{P}-\underset{\sim}{P})^2 - 2\lambda_1 \mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{S} \\ & -2\lambda_2\Big[\sum_{S_+} w(\vec{S}_+)\,\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{\Theta}(S_+) - \sum_{S_-} w(\vec{S}_-)\,\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{\Theta}(S_-)\Big]\Big\} = 0\end{aligned} \tag{10}$$

where a data point belongs in the group $\{S+\}$ if the quantity, $\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{\Theta}(\vec{s}) - \langle \Theta(\vec{s}) \rangle$, is greater than zero, otherwise it is in the $\{S-\}$ group. To first order

$$\begin{aligned}2\mathrm{tr}\Big\{ & \underset{\sim}{P}\underset{\sim}{S}\underset{\sim}{P}\underset{\sim}{P}\underset{\sim}{S} + \underset{\sim}{S}\underset{\sim}{P}\underset{\sim}{P}\underset{\sim}{S}\underset{\sim}{P} - \underset{\sim}{P}\underset{\sim}{S}\underset{\sim}{P} - \underset{\sim}{P}\underset{\sim}{P}\underset{\sim}{S} - \underset{\sim}{S}\underset{\sim}{P}\underset{\sim}{P} + \underset{\sim}{P} \\ & -\lambda_1 \underset{\sim}{S} - \lambda_2\Big[\sum_{S_+} w(\vec{S}_+)\,\underset{\sim}{\Theta}(S_+) - \sum_{S_-} w(\vec{S}_-)\,\underset{\sim}{\Theta}(S_-)\Big]\Big\}\,\delta\underset{\sim}{P} = 0 \, .\end{aligned} \tag{11}$$

The arbitrariness in $\delta\underset{\sim}{P}$ implies that the matrix in brackets must be the null matrix and we obtain the following iterative equation:

$$\utilde{P}n+1 = \utilde{PSP} + \utilde{PPS} + \utilde{SPP} - \utilde{PSPPS} - \utilde{SPPSP} + \lambda_1 \utilde{S} + \lambda_2 \utilde{G} \tag{12}$$

where the matrix $\utilde{G}$ is defined by

$$\utilde{G} \equiv \sum_{S_+} w(\vec{S}_+)\,\Theta(\vec{S}_+) - \sum_{S_-} w(\vec{S}_-)\,\Theta(\vec{S}_-) \;. \tag{13}$$

For ease of writing we also define the matrix $\utilde{\mathcal{P}}$.

$$\utilde{\mathcal{P}} \equiv \utilde{PSP} + \utilde{PPS} + \utilde{SPP} - \utilde{PSPPS} - \utilde{SPPSP} \tag{14}$$

The Lagrangian multipliers are found by substituting $\utilde{P}n+1$ in the constraint equations (6a) and (9).

$$\mathrm{tr}\{\utilde{\mathcal{P}} + \lambda_1 \utilde{S} + \lambda_2 \utilde{G}\}\utilde{S} = N \tag{15a}$$

$$\mathrm{tr}\{\utilde{\mathcal{P}} + \lambda_1 \utilde{S} + \lambda_2 \utilde{G}\}\utilde{G} = \epsilon + \sum_{S_+}\langle\Theta(\vec{S}_+)\rangle - \sum_{S_-}\langle\Theta(\vec{S}_-)\rangle \;. \tag{15b}$$

The above simultaneous equations are simply written in matrix notation

$$\utilde{T}\,\utilde{\lambda} = \utilde{I} \tag{16}$$

where $\utilde{\lambda}$ is the column matrix of Lagrangian multipliers, $\utilde{T}$ is the symmetric matrix of traces

$$\underset{\sim}{T} = \begin{pmatrix} \mathrm{tr}\,\underset{\sim}{S}\underset{\sim}{S} & \mathrm{tr}\,\underset{\sim}{S}\underset{\sim}{G} \\ \mathrm{tr}\,\underset{\sim}{S}\underset{\sim}{G} & \mathrm{tr}\,\underset{\sim}{G}\underset{\sim}{G} \end{pmatrix} \tag{17}$$

and

$$\underset{\sim}{I} = \begin{pmatrix} N - \mathrm{tr}\,\underset{\sim}{\not{P}}\underset{\sim}{S} \\ \epsilon + \sum_{S_+} \langle \Theta(\vec{S}_+) \rangle - \sum_{S_-} \langle \Theta(\vec{S}_-) \rangle - \mathrm{tr}\,\underset{\sim}{\not{P}}\underset{\sim}{G} \end{pmatrix} \tag{18}$$

The solution to equation (16) requires the inversion of the 2 x 2 matrix of traces. Alternatively one could define the matrix $\underset{\sim}{R} = \underset{\sim}{S}^{\frac{1}{2}}\underset{\sim}{P}\underset{\sim}{S}^{\frac{1}{2}}$ which is idempotent, normalized and hermitian:

$$\begin{aligned} &\underset{\sim}{R}^2 = \underset{\sim}{S}^{1/2}\underset{\sim}{P}\underset{\sim}{S}^{1/2}\underset{\sim}{S}^{1/2}\underset{\sim}{P}\underset{\sim}{S}^{1/2} = \underset{\sim}{R} \\ &\mathrm{tr}\,\underset{\sim}{R} = N \\ &\underset{\sim}{R} = \underset{\sim}{R}^{\dagger} \end{aligned} \tag{19}$$

The algorithm to find a constrained $\underset{\sim}{R}$-matrix is then

$$\underset{\sim}{R}_{n+1} = 3\underset{\sim}{R}^2 - 2\underset{\sim}{R}^3 + \sum_k \lambda_k \underset{\sim}{S}^{-1/2}\underset{\sim}{\Theta}_k\underset{\sim}{S}^{-1/2} \tag{20}$$

where $\underset{\sim}{\Theta}_k$ is defined in the ψ-basis. The Lagrangian multipliers are found by substitution in the following expectation value expressions

$$\mathrm{tr}\,\underset{\sim}{R}\underset{\sim}{S}^{-1/2}\underset{\sim}{\Theta}_k\underset{\sim}{S}^{-1/2} = \langle \Theta_k \rangle \tag{21}$$

$\underset{\sim}{P}$ can be recovered by

$$\underset{\sim}{P} = \underset{\sim}{S}^{-1/2}\underset{\sim}{R}\underset{\sim}{S}^{-1/2} \tag{22}$$

This is simpler than equation (II.12) and involves fewer matrix multiplications. The inverse square root matrix, $S^{-\frac{1}{2}}$, need only be found once for a given basis set.

Solution is achieved when $PSP = P$ and $\lambda_1 = \lambda_2 = 0$. The normalization condition will be satisfied on every iteration although the "least absolute value" conditon may not be. This is due to the fact that individual data points may change from the S+group to the S-group with each iteration; at solution, of course, there will be no changes and equation (9) must hold.

The best possible fit is obtained by minimizing ϵ in a super- iterative procedure. An ϵ is guessed, an idempotent solution is found, ϵ is decreased and the procedure is repeated until the minimum ϵ compatible with equations (6) is determined. Taking the absolute value of the differences in condition (9) rather than the square leads to linear equations for the Lagrangian multipliers. The ease of solution justifies its use and the procedure is not without theoretical basis. The conventional Method of Least Squares, developed in the early Nineteenth century in different ways by Legendre, Laplace and Gauss, minimizes the sum of the squares of the residuals:

$$\mathcal{F} = \sum_{\vec{S}} w(\vec{S}) \left\{ \mathrm{tr}\, \underset{\sim}{P}\,\underset{\sim}{\Theta}(\vec{S}) - \langle \Theta(\vec{S}) \rangle \right\}^2 \qquad (19)$$

This leads to the same result as the Principle of Maximum Likelihood if all variables have a gaussian error distribution. If the error distribution were described by $e^{-h|x-\bar{x}|}$, the principle of maximum likelihood would lead to the method of least absolute value. The central limit theorem makes the first alternative more attractive but does not negate the value of the second. In fact, the minimization of ϵ generally takes us in the same direction as the minimum of the functional defined in equation (19). Should there be a discrepancy one could always choose the $\underset{\sim}{P}$-matrix corresponding to minimum $\mathcal{F}$. The Method of Least Absolute Value was proposed by Edgeworth in 1888 to replace the Method of Least Squares. He argues that its solution is less cumbrous (computers were human then), and that for certain probability curves will yield superior results.

Crystallographers have similar difficulty in choosing the best fit to experimental data (see comment by Lipscomb, Trans Am Cryst 1972 p. 108). The most frequently used criterion is the agreement index.

$$R = \frac{\sum\limits_{\substack{\text{all}\\ \text{observations}}} \left| \, |F_{obs}| - |F_{calc}| \, \right|}{\sum |F_{obs}|} \qquad (20)$$

These three criteria, ϵ , $\mathfrak{F}$, R, have not lead to any serious contradictory decisions on a best solution for any calculations we have performed to date.

This derivation closely follows Clinton, Galli and Massa (1969) in which the matrix $\underset{\sim}{R} = \underset{\sim}{P}\underset{\sim}{S}$ is defined and the quantity minimized is subject to a set of constraints of the form, $\langle O \rangle = \mathrm{tr}\underset{\sim}{P}\underset{\sim}{S}^{-1}\underset{\sim}{O}$. The matrix $\underset{\sim}{R}$ is equal to its square but is not hermitian. Thus the $\mathrm{tr}(\underset{\sim}{R}^2 - \underset{\sim}{R})^2 = 0$ does not imply that $\underset{\sim}{R}^2 = \underset{\sim}{R}$ and minimization of the former quantity does not necessarily lead to solution. Equation (12) has the advantage of preserving the hermiticity of $\underset{\sim}{P}$. For an orthogonal basis, $\underset{\sim}{S} = \underset{\sim}{1}$, both algorithms collapse to

$$\underset{\sim}{P}_{n+1} = 3\underset{\sim}{P}^2 - 2\underset{\sim}{P}^3 + \sum_k \lambda_k \underset{\sim}{O}_k \tag{21}$$

We wish to use x-ray diffraction data for the constraints to be inserted in equation (9). Substituting equation (5) in equation (I.1) we obtain

$$F(\vec{S}) = \mathrm{tr}\underset{\sim}{P}\underset{\sim}{f}(\vec{S}) \tag{22}$$

where $$f_{ij}(\vec{S}) = \int \psi_i^* \psi_j \, e^{i\vec{r}\cdot\vec{S}} \, d\vec{r} \tag{23}$$

If the structure has been solved by conventional methods, the atomic positions and phases of the scattered waves are known, and the experimental structure factors can be used.

In this case we take $\underset{\sim}{\hat{O}}(\vec{S}) = \underset{\sim}{f}(\vec{S})$. The use of structure factors in this formalism is applicable only when there is a center of symmetry present. Otherwise the matrix $\underset{\sim}{f}$ is not hermitian and cannot be used in equation (12).

Using equation (12) and definition (23) we can write x-ray diffraction intensities as

$$I(\vec{S}) = \mathrm{tr}\,Pf(\vec{S})\,\mathrm{tr}\,Pf^{\dagger}(\vec{S}) \tag{24}$$

We chose for $\underset{\sim}{\hat{O}}$ this time

$$\underset{\sim}{\hat{O}}(\vec{S}) = \frac{1}{2}\left\{\underset{\sim}{f}^{\dagger}(\vec{S})\,\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{f}(\vec{S}) + \underset{\sim}{f}(\vec{S})\,\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{f}^{\dagger}(\vec{S})\right\} \tag{25}$$

so that $I(\vec{S}) = \mathrm{tr}\underset{\sim}{P}\,\underset{\sim}{O}(\vec{S})$. It is convenient to use expression (25) because the intensities appear as linear functions of $\underset{\sim}{P}$, which makes the iterative solution of equation (12) quite simple. Note that $\underset{\sim}{O}(\vec{S})$ is a function of $\underset{\sim}{P}$; it must be recomputed on every iteration. At solution $\underset{\sim}{O}n+1(S)$ must equal $\underset{\sim}{O}n(S)$.

Implicit in every single-determinant wave function is the assumption that the electrons do not interact (although the antisymmetry does introduce some correlation between particles of like spin). The self consistent field procedure will account for the effect of electrons upon each other in an average way only. This can never be exactly true for systems containing more than one electron but it is nevertheless the most widely used approximation in quantum chemistry today. How applicable is it to the analysis of x-ray data? Bats and Feil (1974) found that difference densities for CN^- computed from extended basis Hartree-Fock wave functions compared favorably with experimental difference maps, while limited basis set ab-initio and semi-empirical calculations did not.

Benesch and Smith (1970) compared scattering factors calculated from a highly accurate Hylleraas-type wave function for lithium (which accounted for 99.90% of the correlation energy) with scattering factors calculated from various independent particle model (IPM) wave functions. These were restricted Hartree-Fock (RHF), spin-polarized Hartree-Fock (SPHF), best density (BD) and symmetry-adapted best density (SABD). The difference of the IPM and "exact" scattering factors are a function of the scattering angle, but for the BD and SABD wave functions the maximum difference was about 0.1%, with virtually no difference for $\sin \theta/\lambda$ less than 0.2 or greater than 1.2 A° . The RHF and SPHF scattering curves did not fit quite so well. The authors extrapolate these results to assume that BD and SABD will yield superior expectation values for other one-electron properties.

From the forgoing evidence we conclude that the single- determinant approximation is adequate for interpretation of x-ray data of the quality presently being achieved.

The major contribution of this work is a formalism that produces densities which are N-representable. It is not yet clear how significant this is with respect to the interpretation of diffraction data, but there are indications that a good fit at the expense of violating N-representability may have profound and undesirable effects. For example, it is possible to obtain densities which are negative in parts of the unit cell. We expect that a lack of N- representability may have severe effects on other calculated physical properties.

The result of the algorithm presented here is an analytical representation of the first-order reduced density matrix. It includes the so-called "off diagonal" terms which allow us to calculate any one-body property. The diagonal terms (the charge density) obtained in least squares fitting procedures will be sufficient to determine one-body properties of multiplicative operators only. Due to the special nature of the density matrix derived by our algorithm, i.e., its idempotency, any quantum mechanical property can be calculated from it, without regard to the number of particles involved in the interaction.

In addition to fitting the density to diffraction data, inclusion of other constraints can easily be accomodated. There is no change in the basic formalism; a Lagrangian multiplier is needed for each constraint and the size of the matrix to be inverted is equal to the number of constraints used. In particular, it may be desirable to incorporate a nuclear cusp condition. In an analysis of the charge density of diamond, Stewart (1973) found that cusp constrained densities gave better temperature factors, and a more appropriate representation of the charge near the carbon nucleus. Clinton and Massa (1972) applied such a constraint along with the Hellman-Feynman and virial theorems in a similar P-algorithm to determine the density matrix for the nitrogen molecule. They found satisfactory convergence, as well as reasonable molecular properties.

One of the greatest impediments to a conventional least squares determination of population coefficients is that the number of parameters increases rapidly with the size of the system, while the number of data points is limited by the x-ray experiment. This problem confines the application of this method to crystals with unit cells containing few electrons and low crystallographic symmetry. The number of parameters to be determined, K_{NI}, when only normalization is imposed is

$$K_{NI} = \frac{1}{2}(m^2+m) - 1 \tag{26}$$

where m is the number of basis functions. An idempotent matrix representative with the same basis has

$$K_I = mN - N^2 \tag{27}$$

degrees of freedom, where N is the number of molecular orbitals. The subtraction of N^2 parameters is due to the imposition of normalization and orthogonality of the molecular orbitals. In the non-idempotent case the number of parameters increases as the square of the basis while in the idempotent case the number of parameters is a linear function of the size of the basis. Table I lists comparisons of the degrees of freedom for an idempotent vs. non- idempotent matrix for several molecules which have been analzed by crystallographers with the Method of Least Squares. Use of our algorithm will enable larger systems to be studied, as well as allow larger basis sets for a given system.

Further advantage over a typical least squares analysis is gained by the fact that linear dependencies in the basis have no effect upon the convergence of our algorithm. Even 100% correlation between scattering factors, which would prevent inversion of the least squares matrix, would not stop solution of the P-equations. A very small determinant, due to less than 100% correlation would decrease the accuracy of least squares determined parameters. This will be discussed in more detail in another section.

The major disadvantage in implementing our formalism as presented, is the possible initial guess dependence. For any system in which $K_I > 1$, the problem is underconstrained. Past experience (Lou Massa, personal communication) with non-"least squares" type constraints has shown an infinity of solutions for any under- constrained problem. We have found no initial guess problem for K =2 when the minimum is reached. This shows that satisfaction of one "least squares" type constraints is not equivalent to one ordinary constraint. For K>2, not all initial guesses reach the same solution, nor do all "solutions" found yield equally good fits to the data. It is important to eliminate this initial guess dependence, but we do not feel it is an insurmountalbe problem.

A minor disadvantage of our method is that no variance- covariance matrix is computed in the course of solving equation (12), so that no estimate of the standard deviations of each parameter is immediately available.

Table I. Comparison of the number of independent parameters in several P-matrices.

Crystal	Total number of parameters in non-idempotent P*	Total number of parameters in idempotent P	number of observed reflections	References
tetracyanoethylene oxide	1539	684	3021	Matthews and Stucky (1971, 1972) Matthews, Swanson, Meuller and Stucky (1971); Matthews, Stucky and Coppens (1971).
fumaramic acid	1034	450	1178	Hirschfeld (1971).
1,3,5-triacetylbenzene	3827	1782	2058	O'Connor and Maslen (1974).
diborane	83	32	273	Jones, Pautler and Coppens (1972)

* The entry in this column represents the total number of parameters in a non-idempotent P-matrix for a limited basis set. The actual least squares analyses done approximate many of the elements to be zero.

III Application of the Algorithm

1. Error Analysis

For the matrix equation (II.6c) to be satisfied to four decimal places for every element, the quantity $\mathrm{tr}(PSP-P)^2$ must be zero to eight decimal places. Assuming that the accuracy of the operators, $\underset{\sim}{\hat{\theta}}(\xi)$, can be as good as we wish, the elements of $\underset{\sim}{P}$ cannot be more significant than the constraints, $\langle\theta(\xi)\rangle$, and possibly less if the scattering factors are less than one. At solution, the Lagrangian multipliers should be zero to at least as many decimal places as are significant in the observational data; certainly the product of $\lambda_2\underset{\sim}{G}$ can have no element larger than the required number of decimal places. The measure for goodness of fit, ϵ , can have no more meaningful decimal places than the $\{\langle\theta(\xi)\rangle\}$. To decide whether the ϵ's for two different P-matrices represent significantly different fits to the data requires the use of a hypothesis testing table, such as Student's t-test or Hamilton's R-ratio test (Hamilton, 1964).

2. Solution of the iterative equations

The iterative solution of equation(II.12) requires the computation of the fourier transforms of all basis orbital products. For m basis functions, $\frac{1}{2}(m^2 + m)$ fourier transforms for each experimental point are needed. These are calculated only once for a given basis set. For a multi-atomic system this is facilitated by the use of gaussian or linear combinations of gaussian functions for basis orbitals. The analytic formulas involved for the various 1s, 2s, and 2p products have been given by McWeeny (1953) and Stewart (1969a). Linear combinations of one to ten gaussians to give the best fit to Slater-type-orbitals and Hartree-Fock atomic orbitals have been worked out by Stewart (1969b), Huzinaga (1965) and Huzinaga and Sakai (1969) for first row atoms. An expansion length of four is generally found to be sufficient for most work.

Each experimental point must be tested on every iteration to determine whether it belongs in the S_+ or S_- group; this means that the matrix $\underset{\sim}{G}$ (equation (II.13)) and the constraint quantity, $\epsilon + \sum_{S_+} I(\vec{S}_+) - \sum_{S_-} I(\vec{S}_-)$, must be reconstructed on every cycle. For the use of intensity constraints where the operator is a function of $\underset{\sim}{P}$, the $\underset{\sim}{G}$-matrix is simultaneously altered to reflect the current $\underset{\sim}{P}$-matrix. Since a reflection may change groups (S_+ and S_-) in the course of one iteration, the ϵ constraint may not be satisfied. The normalization constraint will always be satisfied because the Lagrangian multipliers are chosen to ensure this. At solution both constraints must hold exactly.

To obtain the Lagrangian multipliers a simple 2-dimensional matrix must be inverted. The only other operations involved are matrix multiplications and additions.

3. Computer Requirements

The implementation of this algorithm requires the storage of N + 4 symmetric matrices of the dimension of the basis, where N is the number of data points used in the fitting. The N matrices are the fourier transforms of all orbital products. The efficient calculation of these fourier transforms for atomic orbitals has been investigated by others (Stewart 1969; Coppens, Willoughby, Csonka 1971). At no time in the course of this project has any effort been made to optimize the computing time or the storage requirement. For the case of three basis functions and eighty reflections, each iteration takes approximately 0.07 seconds of CPU time on the IBM 370/168. Use of the alternate algorithm mentioned in chapter II should reduce the time per iteration by a factor of one-quarter. Experience indicates that less than twenty-five iterations are needed when ϵ is far from the minimum ϵ, even if the initial guess is far from the solution. As ϵ_{min} is closely approached the number of iterations required to reach an idempotent solution increases to one or two thousand.

4. Experimental Data

To test our algorithm we chose for "experimental" data a set of scattering factors published by Stewart, Davidson and Simpson (1965) for the hydrogen atom. These scattering factors were calculated by taking the Kolos-Roothaan "exact" wave function for the H -molecule and fitting it as closely as possible by a function of the form $\rho_{molecule} = \rho_{atom}^{(a)} + \rho_{atom}^{(b)}$, with the requirement that ρ_{atom} is a function only of $|\vec{r}|$. Eleven s- type Slater orbitals were used to determine ρ_{atom} which was then fourier inverted to obtain a set of forty-six real structure factors in the region $\sin\theta/\lambda$ equal 0.to 1.7176 Å^{-1}. The structure factors ranged in value from 1.000 to 0.0010. Intensity data for this system was obtained by squaring the structure factors. We attempted to fit this data with a small basis of spherical hydrogenic functions.

$$\psi_{1s} = (\xi^3/\pi)^{1/2} e^{-\xi r}$$

$$\psi_{2s} = (\xi^3/8\pi)^{1/2}(1 - \xi r/2) e^{-\xi r/2}$$

$$\psi_{3s} = (\xi^3/27\pi)^{1/2}(1 - 2\xi r/3 + 2\xi^2 r^2/27) e^{-\xi r/3}$$

$$\psi_{4s} = (\xi^3/64\pi)^{1/2}(1 - 3\xi r/4 + \xi^2 r^2/8 - \xi^3 r^3/192) e^{-\xi r/4} \qquad (1)$$

The fourier transforms for the products of all the above functions on one center can be found exactly and are given by the formula

$$I_n(\xi,\vec{S}) = \frac{2\pi(n+1)!}{i|\vec{S}|}\left\{\frac{(\xi+i|\vec{S}|)^{n+2}-(\xi-i|\vec{S}|)^{n+2}}{(\xi^2+|\vec{S}|^2)^{n+2}}\right\} \tag{2}$$

for In defined by

$$I_n(\xi,\vec{S}) \equiv \int \vec{r}^{\,n} e^{-\xi\vec{r}} e^{i\vec{r}\cdot\vec{S}} d\vec{r} \tag{3}$$

5. Thermal Motion

To account for atomic vibrations, the true crystalline electron density can be approximated by the convolution of a static charge density and a thermal smearing function. The structure factors then become the product of the fourier transforms of the two functions, both of which are dependent upon $\vec{S}$. The traditional approach in crystallography is to apply the temperature factor to the atomic form factor and describe it by

$$T_i(|\vec{S}|) = e^{-B_i \sin^2\theta/\lambda^2} \tag{4}$$

where B_i is a thermal parameter equal to $8\pi^2\overline{\mu_i^2}$, $\overline{\mu_i^2}$ being the mean-square amplitude of the vibration of atom i. If one wished to allow for the fact that the atom may not vibrate equally in all directions, six thermal parameters are needed for each atom to describe the ellipsoid of vibration.

Temperature factors are generally determined in the refinement stages of a crystal structure determination and they are partially responsible for the good fits obtained by spherical atomic form factors. An overestimation of anisotropic temperature factors can simulate bonding density (Coppens, 1968). To avoid this problem, crystallographers employing a least squares analysis of the charge density will take the thermal parameters to be those determined by a neutron diffraction experiment on the same crystal. In our formalism we can apply these independently determined temperature factors to basis functions using $B_{ij} = \frac{1}{2}(B_i + B_j)$ for the overlap contributions. Jones, Pautler and Coppens (1972) have found the results of their least squares analyses to be insensitive to the particular form of the overlap thermal parameter.

In matrix notation we can write the structure factors for the thermally averaged density as

$$F_T(\vec{S}) = \mathrm{tr}\, \underset{\sim}{P}\, \underset{\sim}{t}(\vec{S})\, \underset{\sim}{f}(\vec{S})\, \underset{\sim}{t}(\vec{S}) \qquad (5)$$

where $\underset{\sim}{t}(\vec{s})$ is a diagonal matrix whose elements, t_{ii} is the square root of the thermal function of the atom upon which the basis function is centered. We have not tested this formalism with the systems studied here. All the following calculations were performed on systems at rest.

6. Multiple Solutions

Because the conditions on the $\underset{\sim}{P}$-matrix are of the form $P^2 = P$, it is possible to have multiple solutions satisfying even a fully constrained problem. This can be seen more clearly if we consider the 2-dimensional case where $\underset{\sim}{P}$ is normalized to 1.0 and write it in the form

$$\underset{\sim}{P} = \begin{pmatrix} p & \pm p^{1/2}(1-p)^{1/2} \\ \pm p^{1/2}(1-p)^{1/2} & 1-p \end{pmatrix} \qquad (6)$$

The above matrix is idempotent and normalized with only one parameter, p, to be determined by applying the condition $\mathrm{tr}\underset{\sim}{P}\underset{\sim}{\Theta} = \langle\Theta\rangle$. Whatever value is found for p, two matrices corresponding to the ± signs in the P_{12} term are solutions.

Using our "experimental" H-atom density and a basis consisting of a 1s and 2s hydrogenic function with $\xi = 2.0$, idempotent $\underset{\sim}{P}$-matrices were found in three ways:

(a) directly by solving the quadratic equation $\mathrm{tr}Pf(\vec{s}) = F(\vec{s})$, when $\underset{\sim}{P}$ is written in the form described above, for each F(s),

(b) with the iterative equations and one structure factor constraint to be satisfied exactly, and

(c) with the iterative equations and one intensity constraint to be solved exactly.

For method (a) two solutions for each reflection were found as expected. For methods (b) and (c) the same two solutions were found for each reflection as were determined by method (a). Which of the two possible solutions methods (b) and (c) converged to was dependent upon the initial guess used. Four initial guesses were tired.

When all forty-six reflections are used in one least absolute value constraint, table (I), with all weighting factors unity, only one result was found with intensity algorithm and six initial guesses tried. A steepest descent procedure (Chap. VI) which minimizes the sum of the squares of the residuals of the intensity data yields a slightly different P-matrix but an almost imperceptibly different density (0.02% at the maximum). The structure factor algorithm found two solutions, which we call F1 and F2. They are of equal quality as judged by the sum of the absolute values of the residuals of the structure factor data, ϵ_F.

$$\epsilon_F = \sum_{\vec{S}} w(\vec{S}) \left| \mathrm{tr}\, \underset{\sim}{P} \underset{\sim}{f}(\vec{S}) - F(\vec{S}) \right| \tag{7}$$

Neither solution, F1 nor F2 is identical to the one obtained by using the intensities. In fact, while F1 and F2 are equally well fit to the structure factor data, they do not fit equally well to the intensity data as measured by

$$\epsilon_I = \sum_{\vec{S}} w(\vec{S}) \left| \mathrm{tr}\, \underset{\sim}{P} \underset{\sim}{f}(\vec{S})\, \mathrm{tr}\, \underset{\sim}{P} \underset{\sim}{f}^{\dagger}(\vec{S}) - I(\vec{S}) \right| \tag{8}$$

Although (7) and (8) are not required to have their minima at exactly the same point, great differences would not be expected. It might be of interest to note that the intensity solution, I , was passed by in the structure factor algorithm heading for F1 . In least squares refinements of the structural parameters of three different crystalline forms of D,L-homocysteine thiolactone hydrochloride, Freer and Kraut(1965) compared the results of refinement on the structure factors to results using the intensity data. As expected, the two methods produced slightly different values for the structural parameters. The authors feel that an analysis based on the intensity data directly is preferable to one based on the square root of the observed data.

7. Initial guess dependence

In addition to the theoretical possibility of the existence of multiple solutions, since our $\underset{\sim}{P}$-matrix generally has more parameters than conditons imposed, there may be some initial guess dependence. For any idempotent matrix representative of a wave function with N molecular orbitals expanded in m basis functions there are N x m coefficients to be determined. The requirement that all molecular orbitals are normalized and orthogonal to one another imposes N^2 conditions and we arrive at equation (II.27) for the number of free parameters in an idempotent matrix (Clinton, Galli, Massa 1969). As formulated, equation (II.12) imposes only one constraint in addition to normalization, and therefore any system for which K(m,N) is greater than one is underconstrained. This implies that the solution found may depend upon the initial guess used. As discussed previously this is a highly undersirable situation. In principle there should be an infinite number of $\underset{\sim}{P}$-matrices that can satisfy all imposed constraints when the number of constraints is less than K(m,N). This point is illustrated in table II. When four basis functions were used with one reflection as a constraint, every initial guess tried lead to a different solution. This is in accord with previous experience with underconstrained problems in similar algorithms. However, when we use a least absolute value constraint and minimize ϵ, the $\underset{\sim}{P}$-matrix is in effect being made to fit at many points in reciprocal space. Since we have frequently found unique solutions for such underconstrained systems when widely different initial guesses have been used, it appears that one least absolute value condition is a much stronger requirement than one ordinary constraint.

Referring to table II, we see that the use of 46 structure factor constraints finds no initial guess dependence. Use of the least absolute value algorithm with intensity constraints produces an intermediate situation. Three of the initial guesses have solutions very close to one another at the same value of ϵ, a fourth initial guess has a very different solution at a much smaller value of ϵ. If we are counting on the "strength" of one least absolute value constraint to force unique solutions, its ability to do so will depend upon the number and magnitudes of the reflections included. Structure factors in this system are larger than the corresponding intensities. The size of our problem (one electron) makes the hydrogen atom the most unfavorable system to study from the point of view of initial guess dependence.

A weighting scheme whose effect is to increase the magnitude of the constraints should have some influence upon the initial guess dependency. When the weight of each reflection is taken to be $(\sin \theta/\lambda)^2$, the reflections for large scattering angles (hence smaller intensities) will be weighted more. Such a weighting scheme brings all initial guesses to same solution (table II).

One way to eliminate the initial guess dependence is to add enough additional constraints to fully determine $\underset{\sim}{P}$. In light of our experience with underconstrained problems including least absolute value constraints, it may not be necessary to impose exactly K(m,N) constraints, fewer may do. One constraint to be considered is the cusp condition since it is not only easily applied (Clinton and Massa, 1972), but it is highly desirable for our wave functions to have this quality. As mentioned previously, it has been found (Stewart 1973) that densities obeying the cusp condition give more realistic temperature factors. However, any theoretical or experimental constraint may be used.

Another approach to alleviate this problem is to reformulate equation (II.12). Instead of using all data points in a one-package constraint, data points could be divided into groups and the ϵ of each group minimized independently. A method similar to this idea was widely used before the acceptance of the Method of Least Squares. It consisted of arranging the equations of condition into sets, taking the sum of all equations in each set and using the new equations to solve for the unknowns. In 1920 Campbell proposed that this method, which he called the Method of Zero Sum, replace the Method of Least Squares on the grounds that it is easier to implement and on a firmer theoretical basis.

Another possibility is to solve the iterative P-equations exactly for K(m,N) data points at a time, and then take some sort of "averaged" guess. How to extract the best initial guess from these results needs more study. The least squares solution or an idempotent solution from a smaller basis not subject to initial guess dependency might be used.

A last point of caution to keep in mind is the practical application of this algorithm is that wherever initial guess dependence may be a problem the starting value of ϵ should be chosen with care. Since ϵ is the sum of the absolute values of the residuals, solving the P-equations with a given ϵ forces the solution to be exactly that far away from the data. Starting with an ϵ

much greater than the minimum ϵ may push the initial guess in the wrong direction and allow the solution to get stuck in a local minimum.

8. Hydrogen Atom

A 1s, 2s, 3s hydrogenic basis (equation (III.4 (1)) with variable exponent, ξ, was used to fit to the hydrogen atom data. Conventional least squares analyses on the structure factors, in which the only imposed constraint is normalization, were performed for comparison. The results for several values of ξ are shown in figures I through VI. Sensible idempotent solutions are found for all systems in which a reasonable basis ($.9 < \xi < 2.0$) is used. This indicates that solution of the iterative equations is not very basis dependent. In every case the greater flexibility in the not idempotent density allows it to fit better than the idempotent density. Although the idempotent and not idempotent curves reported are in general very similar, (the exception is for $\xi = 7.0$) their eigenvalues are quite different. In no case is the least squares not-idempotent density N-representable as determined by its eigenvalues, whose only allowable values are zeros and one for a one electron system. As the fit to the density gets better (R decreases), the not-idempotent eigenvalues approach the acceptable values. The rigorous requirement of zeros and one eigenvalues is never acheieved for any basis tested, and the eigenvalues diverge from the acceptable values much more rapidly than the agreement index increases.

For ξ = 1.12 (figure III) the idempotent and not-idempotent curves are almost identical. This is due to the fact that the exact density for our hydrogen atom is idempotent. If the exact basis were used in the analysis the Method of Least Squares and the idempotency algorithm should yield identical results.

As judged by any of the criteria discussed in chapter II, the best idempotent fit to the data occurred at ξ =1.12. While the unrestricted least squares fit also has a local minimum R at ξ = 1.12, the solution at ξ = 5.00 is an order of magnitude better. For this highly contracted basis the 2s function makes the largest contribution. For a hydrogen atom in a ground state hydrogen molecule the solution at ξ =1.12 is more appealing. We would expect it to be more contracted than the isolated atom (ξ =1.00) but still predominantly of 1s character. Inspection of figures III and V leads to the observation that when fitting to fourier transform of the density, the solution is not as greatly influenced by the high density region as would be when fitting was made directly to the density. Although ξ =5.00 gives a better fit to the structure factors (this was also found to be the case with the idempotent intensity algorithm) we would not expect it to furnish better expectation values for other properties. Benesch and Smith (1972) and Benesch, Singh and Smith (1971) have shown how it is possible to obtain a good fit to the structure factors of a system, while the momentum distribution can be in error by a much greater amount. By enforcing N-representability when fitting to X-ray data we hope the restriction upon the class of density matrices considered to physically acceptable systems will improve results when other properties are subsequently calculated. This hypothesis is yet to be tested.

There are four examples (ξ = 0.9,1.0, 1.12, 7.0) in Table III where negative diagonal elements appear for $\underset{\sim}{P}$not-idempotent. This makes Pii impossible to interpret as the square of a coefficient (c.f. equation I.10 for one molecular orbital). This is an obvious expression of the fact that such a $\underset{\sim}{P}$not-idempotent cannot describe any physical system.

As an extreme example of what the violation of N-representability may allow we refer to figure VI. For this case ξ = 7.0, a rather poor basis for describing the hydrogen atom. The fit obtained by the not-idempotent density is nevertheless quite good (R = .009) although the electron density takes on dramatically negative values at and near the nucleus. The agreement factor for the idempotent density is worse by a factor of almost 10 (R = .085), yet the absence of negative density makes it far more reasonable than the not-idempotent case. Further examples of negative densities have been found with better fitting bases, but these regions are generally far from the nucleus and only slightly, though significantly negative. It would be interesting to see how the failure to impose N-representability affects other physical properties besides the charge density.

Table I. Hydrogen Atom $\xi = 1.00$

	I^a	F_1^b	F_2^c	$S.D.^d$
$\underset{\sim}{P}_{solution}$	$\begin{bmatrix} .9591 & .1980 \\ & .0409 \end{bmatrix}$	$\begin{bmatrix} .9525 & .2128 \\ & .0475 \end{bmatrix}$	$\begin{bmatrix} .9425 & .2327 \\ & .0575 \end{bmatrix}$	$\begin{bmatrix} .9594 & .1974 \\ & .0406 \end{bmatrix}$
ϵ_I	.5194	.5248	.5403	.5194
ϵ_F	.6404	.6357	.6357	
$\lvert\vec{r}\rvert$ (a.u.)	Densities (e/a.u.^{3})			
0.	.3515	.3530	.3547	.3514
0.1	.2877	.2889	.2903	.2877
0.2	.2354	.2364	.2376	.2354
0.3	.1926	.1934	.1943	.1925
0.4	.1574	.1581	.1588	.1574
0.5	.1287	.1292	.1297	.1287
0.6	.1051	.1055	.1059	.1051
0.7	.0858	.0861	.0864	.0858
0.8	.0700	.0702	.0704	.0700
0.9	.0570	.0572	.0573	.0570
1.0	.0464	.0465	.0467	.0464

a) Solution for intensity data

b) Solution for structure factor data when $\underset{\sim}{P}$initial $= \begin{bmatrix} 1.0 & 0. \\ 0. & 0. \end{bmatrix}$

c) Solution for structure factor data when $\underset{\sim}{P}$initial $= \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix}; \begin{bmatrix} .40 & .49 \\ .49 & .60 \end{bmatrix}; \begin{bmatrix} .60 & .49 \\ .49 & .40 \end{bmatrix}$

d) Solution obtained by steepest descent method using intensity data.

Table I. Hydrogen Atom $\xi = 1.00$

	I^a	F_1^b	F_2^c	$S.D.^d$
$\underset{\sim}{P}_{\text{solution}}$	$\begin{bmatrix} .9591 & .1980 \\ & .0409 \end{bmatrix}$	$\begin{bmatrix} .9525 & .2128 \\ & .0475 \end{bmatrix}$	$\begin{bmatrix} .9425 & .2327 \\ & .0575 \end{bmatrix}$	$\begin{bmatrix} .9594 & .1974 \\ & .0406 \end{bmatrix}$
ϵ_I	.5194	.5248	.5403	.5194
ϵ_F	.6404	.6357	.6357	
$\lvert\vec{r}\rvert$ (a.u.)	Densities $(e/a.u.^3)$			
0.	.3515	.3530	.3547	.3514
0.1	.2877	.2889	.2903	.2877
0.2	.2354	.2364	.2376	.2354
0.3	.1926	.1934	.1943	.1925
0.4	.1574	.1581	.1588	.1574
0.5	.1287	.1292	.1297	.1287
0.6	.1051	.1055	.1059	.1051
0.7	.0858	.0861	.0864	.0858
0.8	.0700	.0702	.0704	.0700
0.9	.0570	.0572	.0573	.0570
1.0	.0464	.0465	.0467	.0464

a) Solution for intensity data

b) Solution for structure factor data when $\underset{\sim}{P}_{\text{initial}} = \begin{bmatrix} 1.0 & 0. \\ 0. & 0. \end{bmatrix}$

c) Solution for structure factor data when $\underset{\sim}{P}_{\text{initial}} = \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix}; \begin{bmatrix} .40 & .49 \\ .49 & .60 \end{bmatrix}; \begin{bmatrix} .60 & .49 \\ .49 & .40 \end{bmatrix}$

d) Solution obtained by steepest descent method using intensity data.

TABLE II

initial guess	.2500				1.0000				.5000				.0000			
	.2500	.2500			.0000	.0000			.5000	.5000			.0000	1.0000		
	.2500	.2500	.2500		.0000	.0000	.0000		.0000	.0000	.0000		.0000	.0000	.0000	
	.2500	.2500	.2500	.2500	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.00
underconstrained solutions,K(m,N)= 3 one reflection, F=.8350 used	.6616				.8575				.6714				.6326			
	.3807	.2191			-.3068	.1098			.4644	.3213			.4471	.3160		
	.2110	.1214	.0673		-.1472	.0527	.0253		.0612	.0423	.0056		.1551	.1096	.0380	
	.1855	.1068	.0592	.0520	-.0799	.0286	.0137	.0074	.0349	.0241	.0032	.0018	.0918	.0649	.0225	.01
underconstrained solutions,K(m,N)= 3 one reflection, F=.0010 used	.0067				.3946				.0008				.0050			
	.0373	.2082			-.4092	.4244			.0265	.8821			-.0705	.9947		
	-.0487	.2721	.3555		-.2240	.2323	.1272		-.0081	-.2693	.0822		.0010	-.0141	.0002	
	-.0535	.2991	.3908	.4296	-.1458	.1512	.0828	.0539	-.0053	-.1753	.0535	.0348	.0006	-.0092	.0001	.00
least absolute value structure factor solution, 46 reflections					.7688											
					-.3670	.1752										
					-.1791	.0855	.0417		All initial guesses result in same solution.							
					-.1048	.0500	.0244	.0143								
least absolute value intensity solution, 46 reflections W(I) = 1.0	.5071				.8261				.5091				.5073			
	.4721	.4395			-.3438	.1431			.4750	.4433			.4726	.4402		
	.1493	.1390	.0439		-.1394	.0580	.0235		.1414	.1320	.0393		.1480	.1379	.0432	
	.0694	.0646	.0204	.0095	-.0774	.0322	.0131	.0072	.0652	.0608	.0181	.0083	.0688	.0641	.0201	.0
least absolute value intensity solution, 46 reflections					.7123											
					-.3645	.1865										
					-.2232	.1142	.0700		All initial guesses result in same solution.							
					-.1492	.0764	.0468	.0313								

TABLE III-HYDROGEN ATOM

ξ	$\underset{\sim}{P}$ idempotent	eigen-values	$\underset{\sim}{P}$ not-idempotent	eigen-values
.90	$\begin{bmatrix} .9483 & & \\ .1983 & .0415 & \\ .0985 & .0206 & .0102 \end{bmatrix}$ R = .118	1.0000 0.0000 0.0000	$\begin{bmatrix} 3.1923\ (.31)^{*} & & \\ -28.3055\ (3.7) & -2.0229\ (.27) & \\ 48.6904\ (6.2) & -.7343\ (.15) & -.1694\ (.05) \end{bmatrix}$ R = .027	57.9407 - 2.1925 -54.7482
1.00	$\begin{bmatrix} .9460 & & \\ .2088 & .0460 & \\ .0872 & .0192 & .0080 \end{bmatrix}$ R = .049	1.0000 0.0000 0.0000	$\begin{bmatrix} 1.9893\ (.11) & & \\ -12.9117\ (1.3) & -.9163\ (.10) & \\ 22.3742\ (2.2) & -.3214\ (.06) & .0730\ (.02) \end{bmatrix}$ R = .010	26.8444 - .9837 -24.8607
1.12	$\begin{bmatrix} .9936 & & \\ -.0395 & .0016 & \\ -.0692 & .0028 & .0048 \end{bmatrix}$ R = .006	1.0000 0.0000 0.0000	$\begin{bmatrix} 1.0448\ (.04) & & \\ -.7032\ (.44) & -.0445\ (.03) & \\ 1.3489\ (.74) & -.0065\ (.02) & -.0002\ (.007) \end{bmatrix}$ R = .004	2.1293 - .0403 - 1.0890
1.20	$\begin{bmatrix} .9977 & & \\ -.0333 & .0011 & \\ -.0345 & .0012 & .0012 \end{bmatrix}$ R = .015	1.0000 0.0000 0.0000	$\begin{bmatrix} .6566\ (.08) & & \\ -4.3608\ (.97) & .3152\ (.07) & \\ -7.4685\ (1.6) & .1187\ (.04) & .0282\ (.01) \end{bmatrix}$ R = .009	8.9821 .3457 - 8.3277
5.00	$\begin{bmatrix} .2267 & & \\ -.4055 & .7254 & \\ .1042 & -.1864 & .0479 \end{bmatrix}$ R = .028	1.0000 0.0000 0.0000	$\begin{bmatrix} .1340\ (.005) & & \\ -.9715\ (.06) & .6351\ (.004) & \\ .8811\ (.09) & -.2250\ (.002) & .0510\ (.0007) \end{bmatrix}$ R = .0009	1.7974 .1002 - .8976
7.00	$\begin{bmatrix} .0000 & & \\ -.0055 & .8208 & \\ .0026 & -.3835 & .1791 \end{bmatrix}$ R = .085	1.0000 0.0000 0.0000	$\begin{bmatrix} -.1522\ (.06) & & \\ 3.0254\ (.78) & .8543\ (.06) & \\ -5.4171\ (1.3) & -.2024\ (.03) & .2979\ (.008) \end{bmatrix}$ R = .009	6.4519 .5476 - 5.9995

* The numbers in parentheses are the standard deviations.

Figure I. Hydrogen atom $\zeta = 0.90$

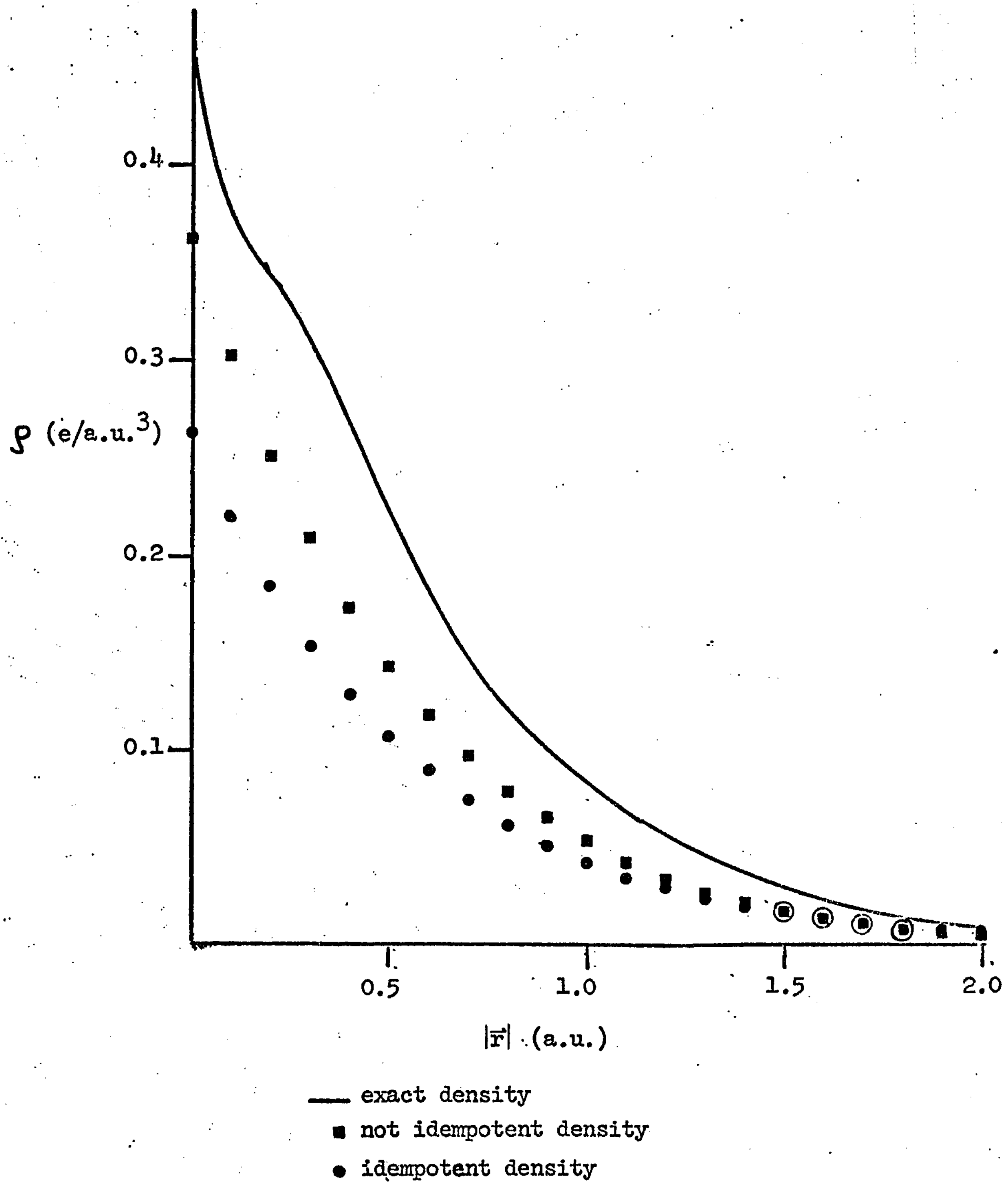


Figure II. Hydrogen atom $\xi = 1.00$

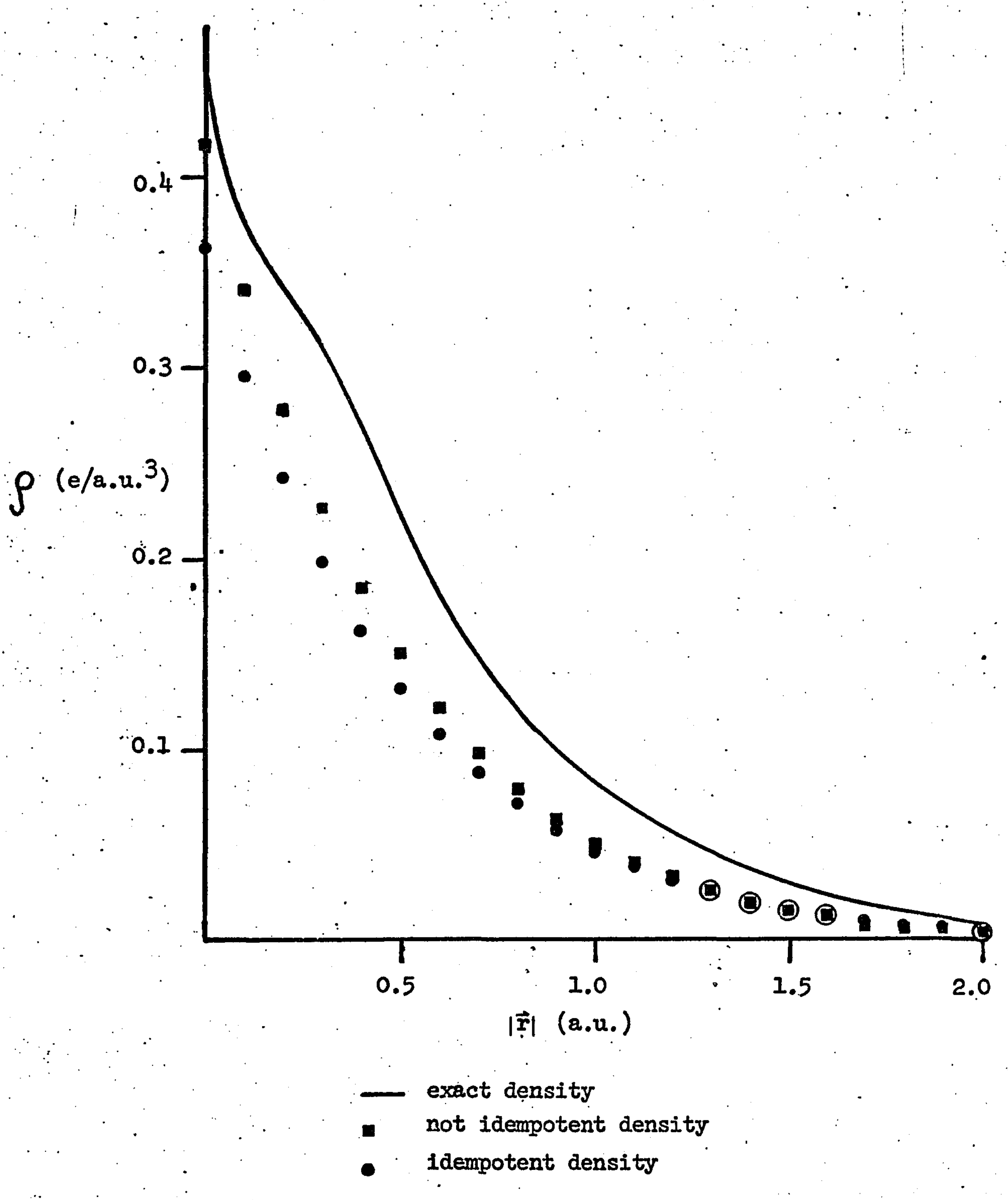


Figure III. Hydrogen atom $\xi = 1.12$

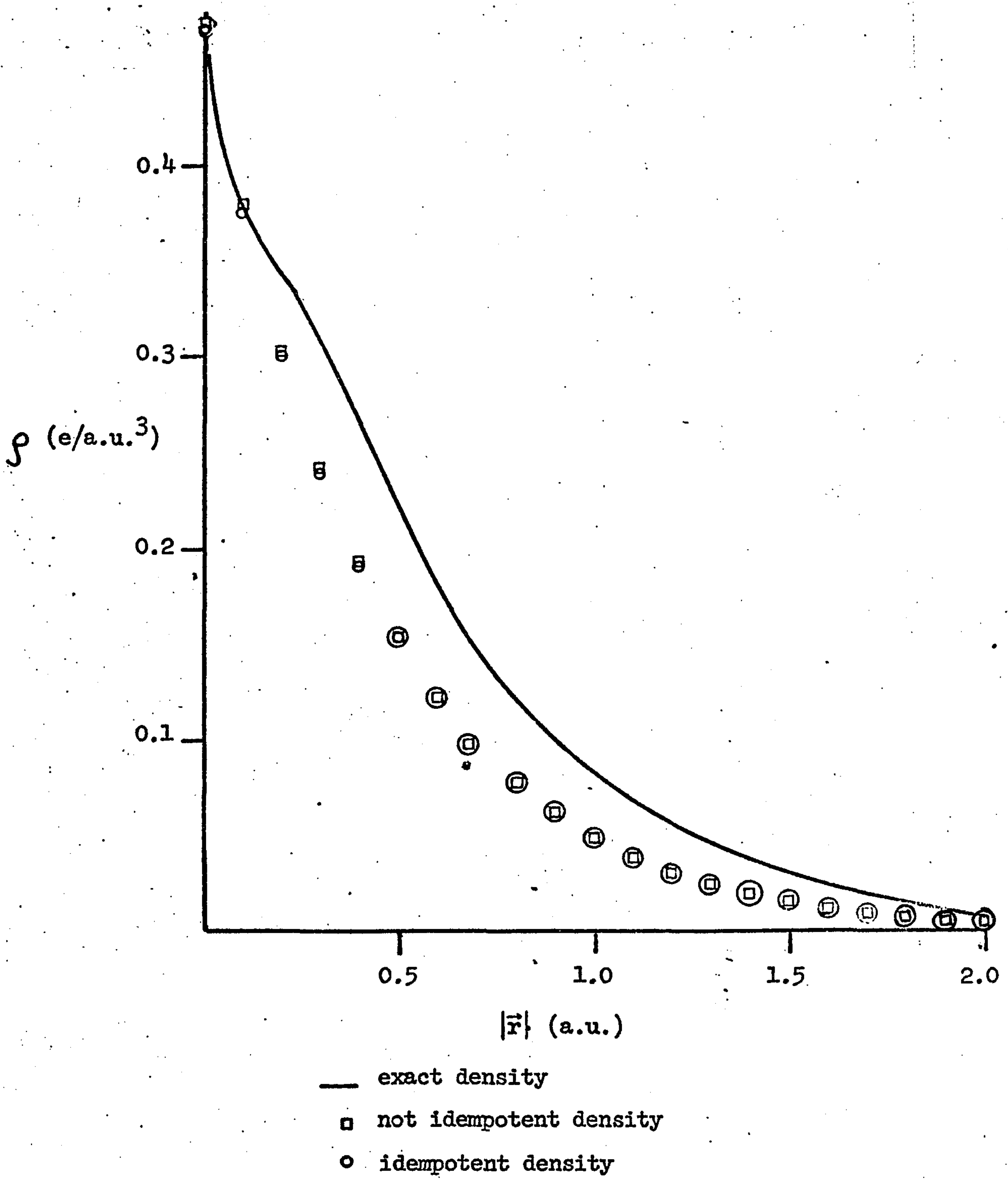


Figure IV. Hydrogen atom $\xi = 1.20$

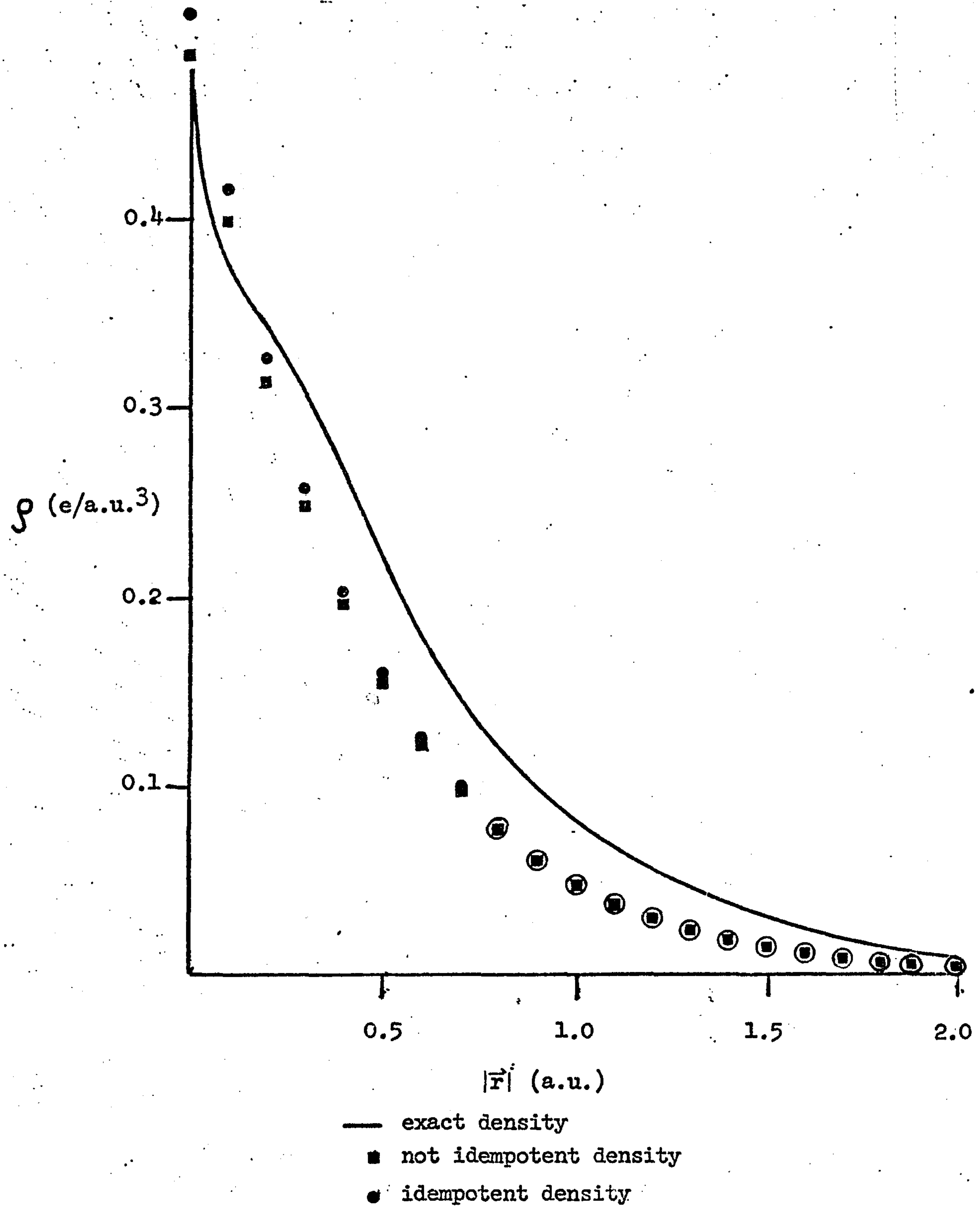


Figure Va. Hydrogen atom $\xi = 5.00$

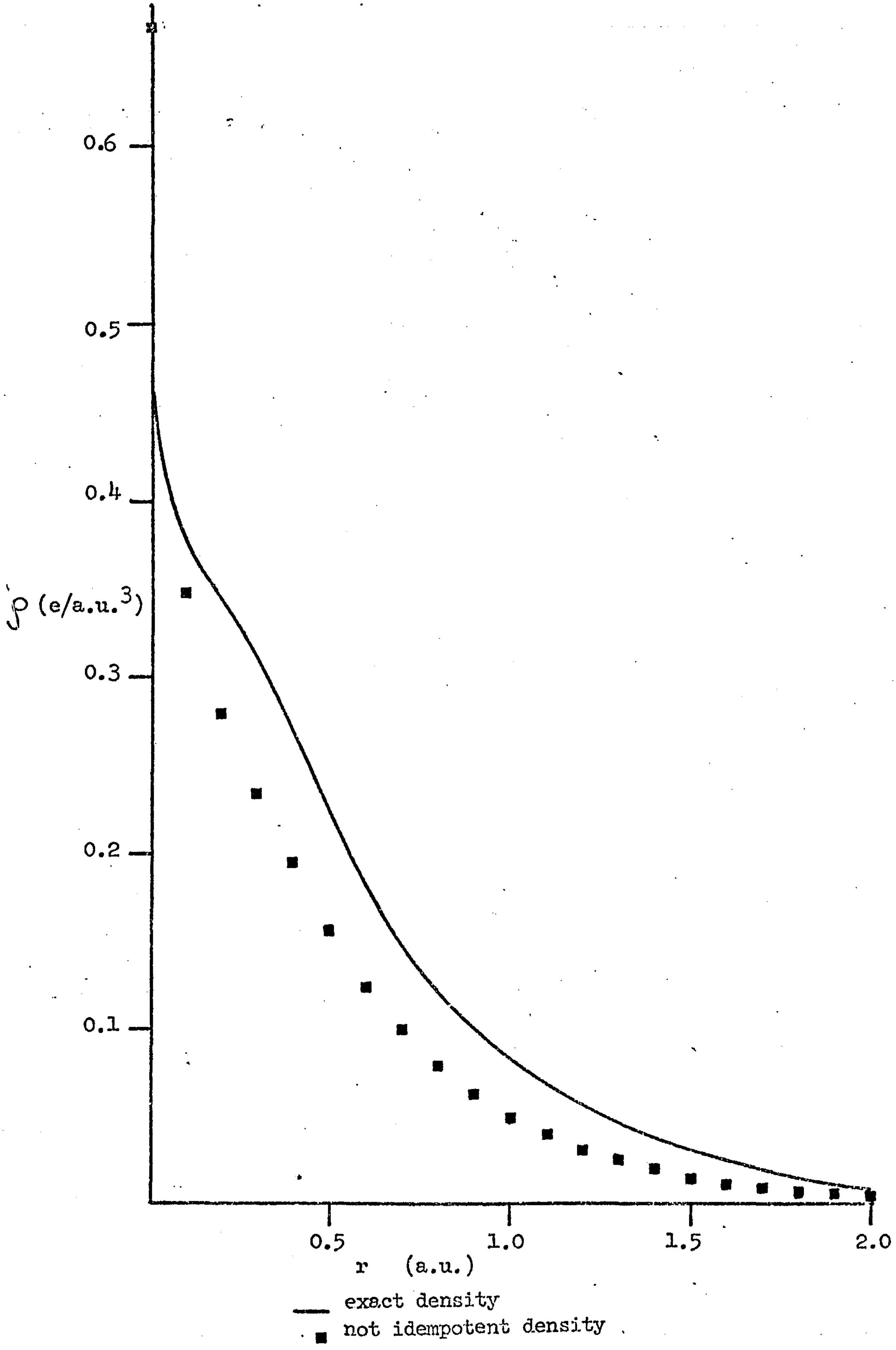
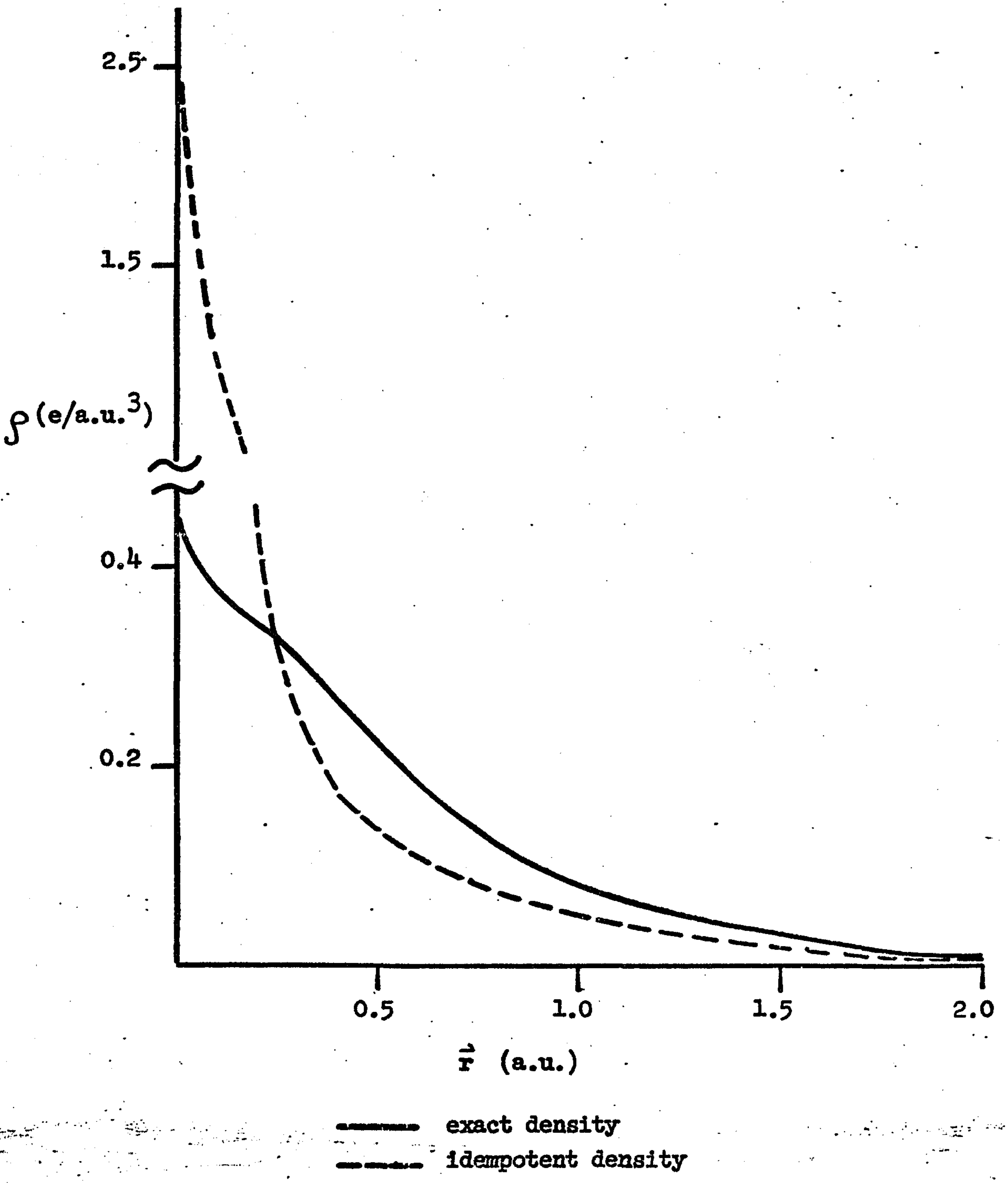


Figure Vb. Hydrogen Atom $\xi = 5.00$



[illegible]

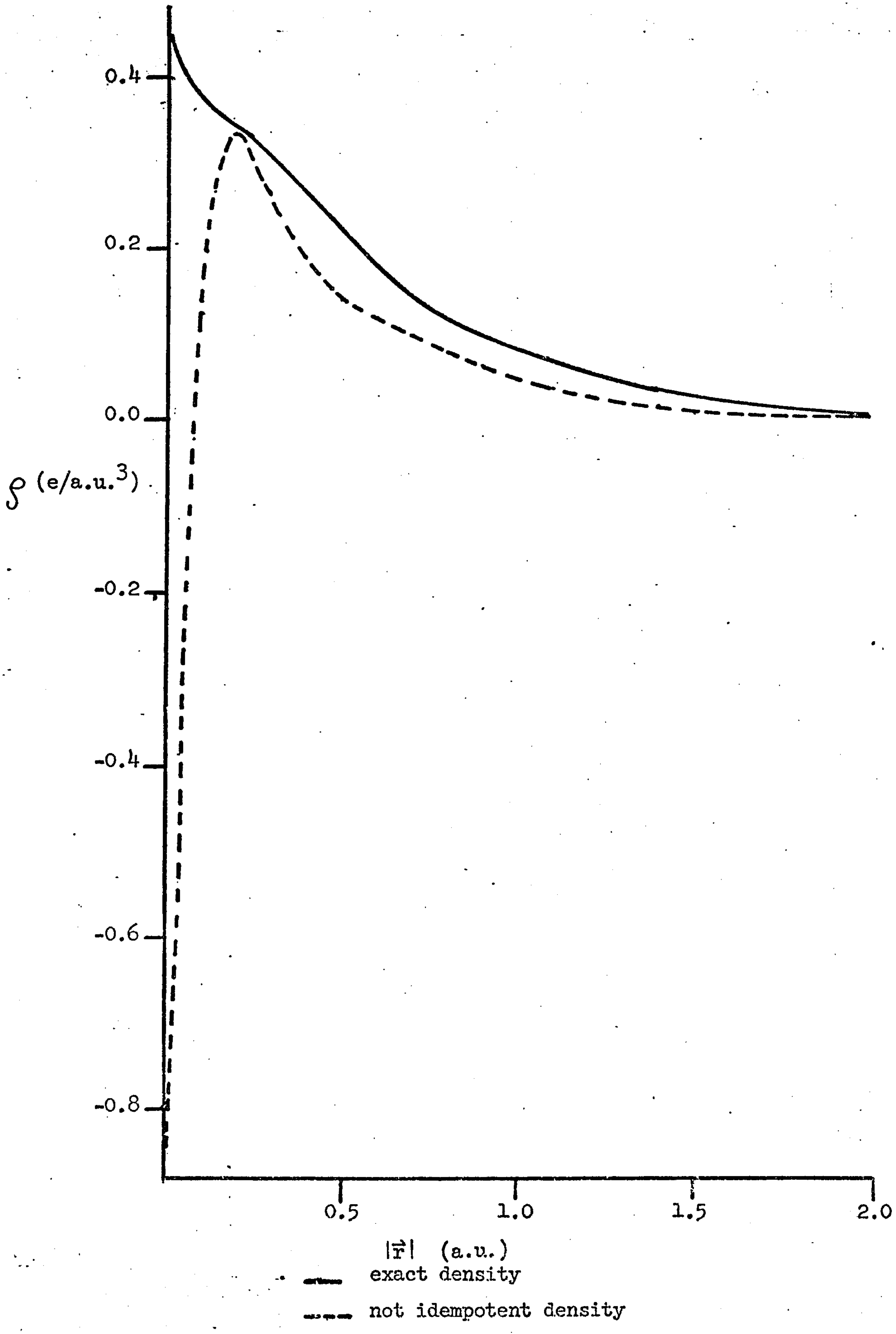
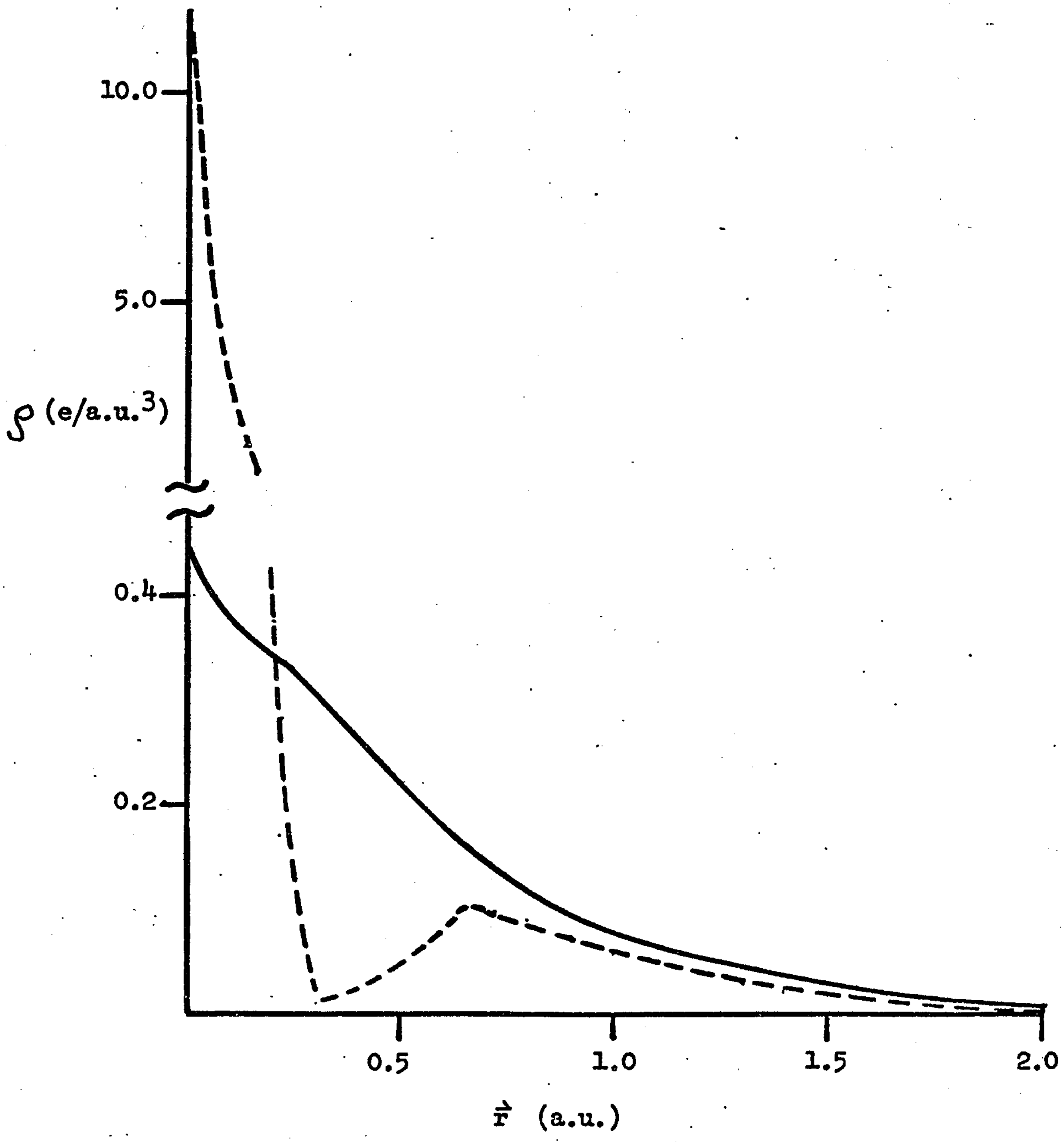

0.4
0.2
0.0
ρ (e/a.u.^{3})
-0.2
-0.4
-0.6
-0.8
0.5
1.0
1.5
2.0
|r⃗| (a.u.)
exact density
not idempotent density

Figure VIb. Hydrogen Atom $\xi = 7.00$



━━━ exact density

╍╍╍ idempotent density

9. A Model Density

To test the convergence and properties of our algorithm for a two center system we constructed an arbitrary normalized idempotent density containing two electrons and two protons at a distance of 1.28 a.u. This density is defined by three symmetry basis functions which are sums of eight gaussians each. From this density we calculate eighty intensities where the magnitude of $\vec{S}$ ranges from .2702 to 18.3456 $\mathring{A}^{-1}$ and the angles it makes with the inter- nuclear bond are 0°, 90°, 50°, 20°.All intensities are tabulated in table IV. Using the exact basis but starting with an initial guess far from solution, the algorithm converged to essentially the exact solution within a reasonable period of time. The minimum ϵ iterated to was .003 rather than zero, but the differences in densities are so small (.0001 e/au^3 at the greatest) that they are indistinguishable in figure VII. The four greatly different initial guesses converged to the same result (no initial guess dependence found) but a fifth (the farthest initial guess density tried) did not converge at all.

Altering the basis so that the density could not be reproduced exactly, provides encouraging results. All exponential factors were multiplied by 1.05 (figure VIII) or

0.95 (figure IX). Three initial guesses

$$P_A = \begin{bmatrix} 1.0 & 0. & 0. \\ & 0. & 0. \\ & & 0. \end{bmatrix} ; \quad P_B = \begin{bmatrix} 0.5 & 0. & 0. \\ & 0.5 & 0. \\ & & 0. \end{bmatrix} ; \quad P_C = \begin{bmatrix} .3333 & 0. & 0. \\ & .3333 & 0. \\ & & .3333 \end{bmatrix}$$

were tried in each case. For $\xi = 1.05 \times \xi$exact all calculations produce the identical result which is very close to the true density (figure VIII). When the $\xi = .95 \times \xi$exact basis is employed, initials guesses (B) and (C) result in the same solution density but initial guess (A) yields a slightly different density.

Grosser changes in the basis functions, $\xi = .85 \times \xi$exact and $\xi = 1.5 \times \xi$exact find solutions very close to the true density with agreement factors of .0046 and .0048 respectively (figure X).

Table IV. Model Molecular Density Data

$$\underset{\sim}{P}_{exact} = \begin{bmatrix} 1.2321 & .1665 & -.4456 \\ .1665 & .0225 & -.0602 \\ -.4456 & -.0602 & .1612 \end{bmatrix}$$

basis:

$\psi_n = \Phi_n(a) + \Phi_n(b)$ where $|\vec{r}_{ab}| = 1.4009$ a.u.

$\Phi_n = \sum_i d_{in} e^{-\alpha_{in} r^2}$

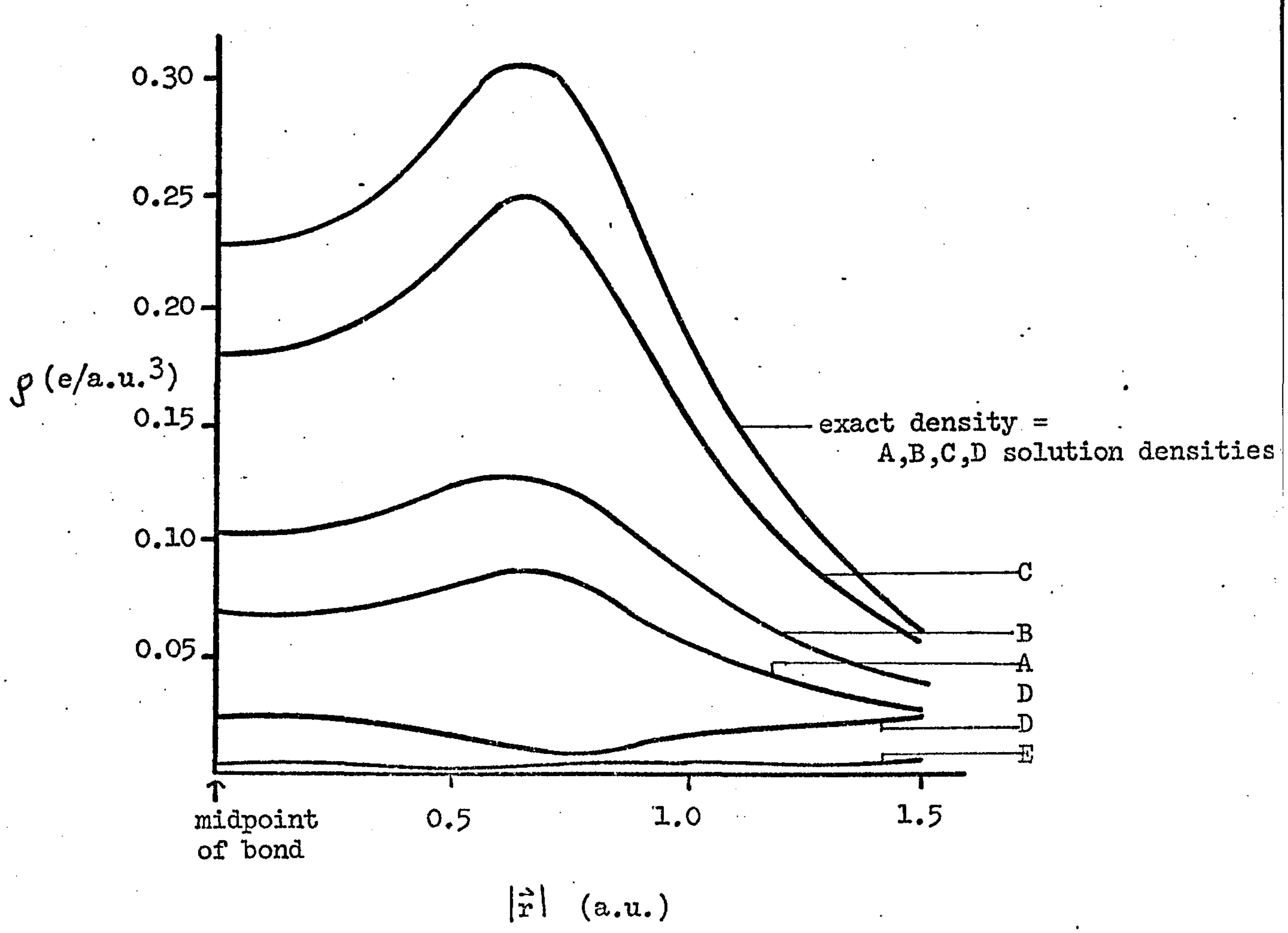
	d_{in}	α_{in}
H1S		
	0.271919D 00	0.1233169
	0.253448D 00	0.4537570
	0.717041D-01	2.0132999
	0.101809D-01	13.3614998
H2S		
	0.280042D 00	0.0660000
	0.260930D 00	0.1709040
	-0.292617D-01	2.3309598
	-0.432158D-02	18.7739983
H3S		
	0.239513D 00	0.0426870
	0.314183D 00	0.0778104
	-0.851380D-01	0.5144040
	-0.962054D-02	2.0043896

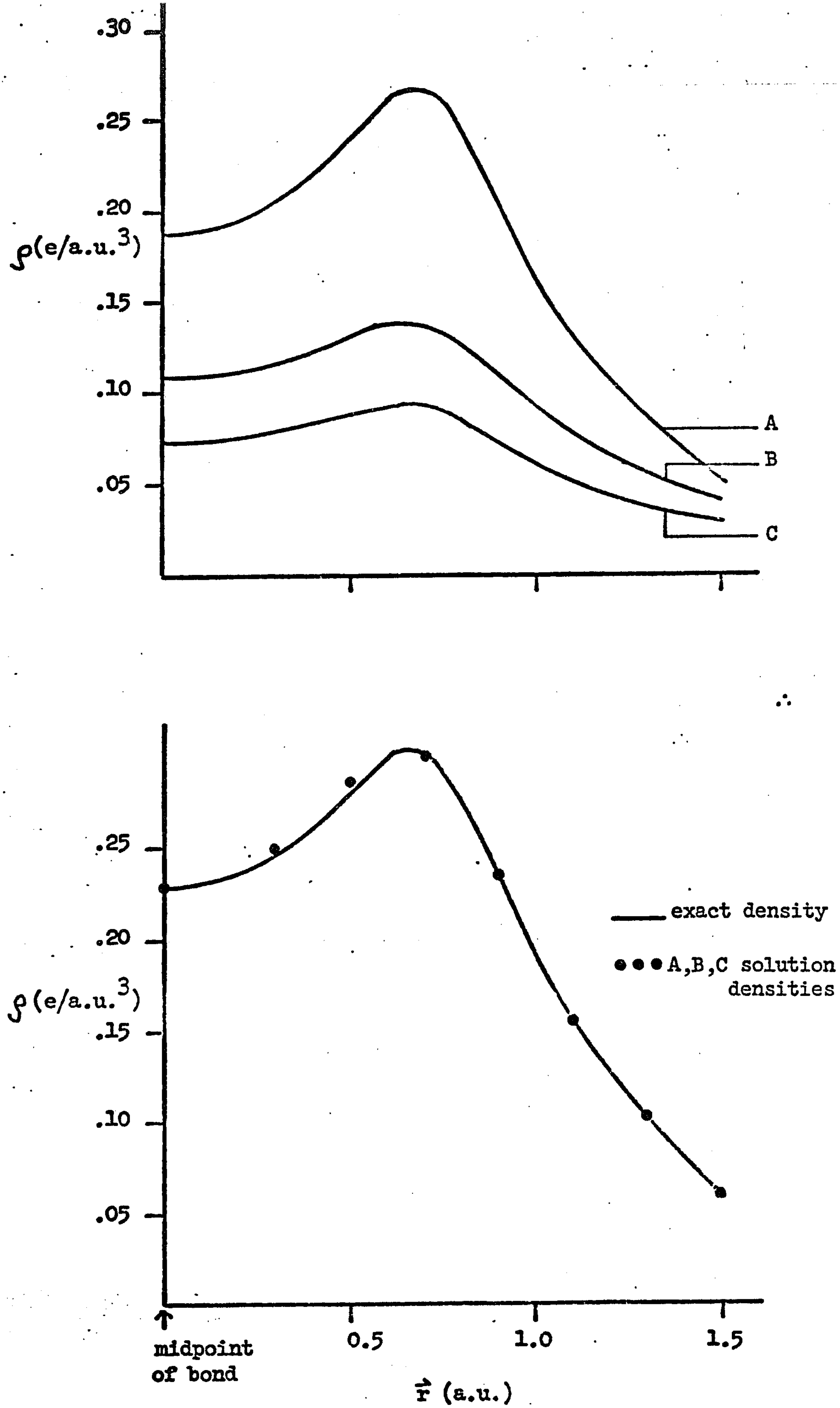
Intensities = $4 \mathrm{tr} P f(S) \mathrm{tr} P f^{\dagger}(S)$

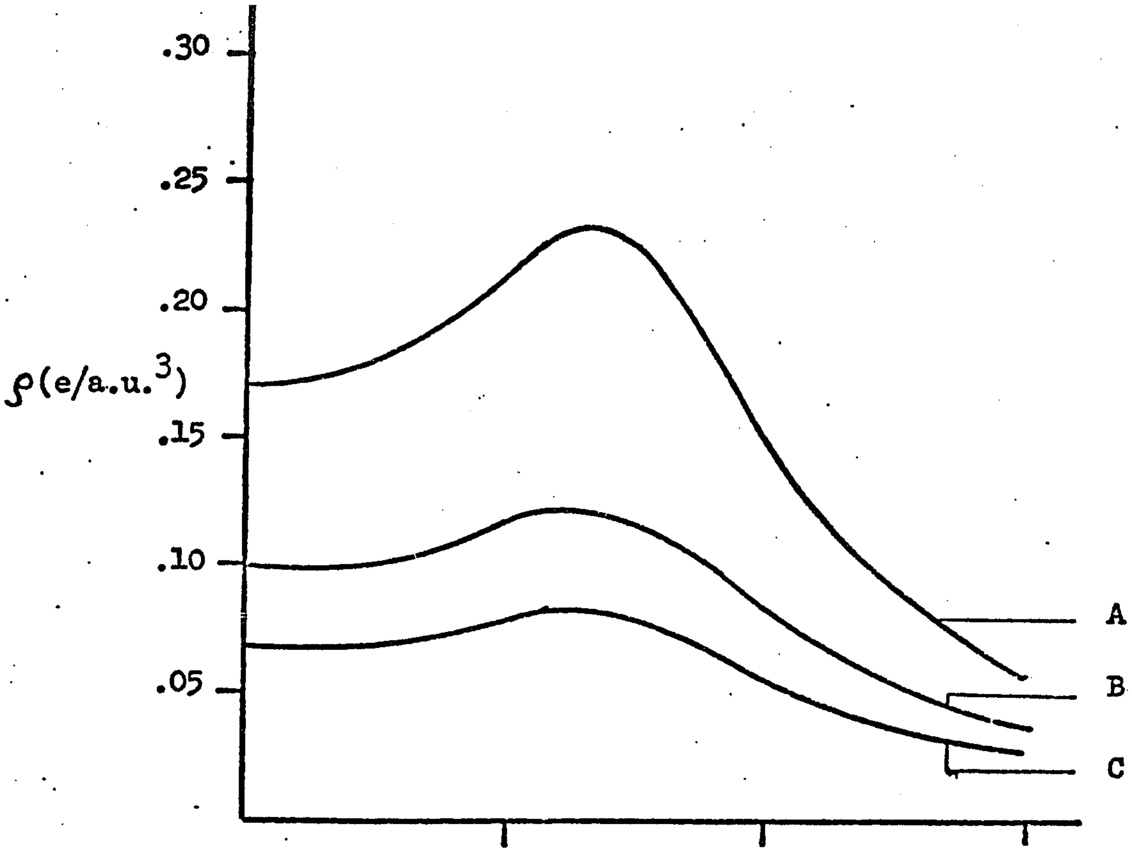
Angle between $\vec{R}$ and $\vec{S}$	90°	50°	20°	0°
$\sin\theta/\lambda$				
0.0	0.37624E 01	0.37238E 01	0.36802E 01	0.36694E 01
0.0429	0.31648E 01	0.30342E 01	0.28906E 01	0.28557E 01
0.0859	0.24205E 01	0.21964E 01	0.19606E 01	0.19050E 01
0.1288	0.16889E 01	0.14131E 01	0.11412E 01	0.10798E 01
0.1718	0.10827E 01	0.81128E 00	0.56626E 00	0.51457E 00
0.2147	0.37084E 00	0.19924E 00	0.81704E-01	0.62324E-01
0.3006	0.11451E 00	0.35368E-01	0.29093E-02	0.67573E-03
0.3864	0.35694E-01	0.40688E-02	0.11172E-02	0.26356E-02
0.4723	0.12075E-01	0.10240E-03	0.29221E-02	0.40542E-02
0.5582	0.44735E-02	0.11304E-03	0.23839E-02	0.27559E-02
0.6441	0.17652E-02	0.27609E-03	0.12770E-02	0.12689E-02
0.7300	0.73062E-03	0.26584E-03	0.51113E-03	0.41958E-03
0.8158	0.31983E-03	0.19009E-03	0.15186E-03	0.89287E-04
0.9017	0.15177E-03	0.11920E-03	0.27727E-04	0.57791E-05
0.9876	0.79310E-04	0.69748E-04	0.63548E-06	0.27233E-05
1.0735	0.45445E-04	0.38437E-04	0.33317E-05	0.13445E-04
1.1593	0.27900E-04	0.19377E-04	0.99452E-05	0.18645E-04
1.2452	0.17873E-04	0.84312E-05	0.12603E-04	0.16458E-04
1.3311	0.11673E-04	0.28255E-05	0.10862E-04	0.10507E-04
1.4170	0.76470E-05	0.53000E-06	0.69801E-05	0.47546E-05

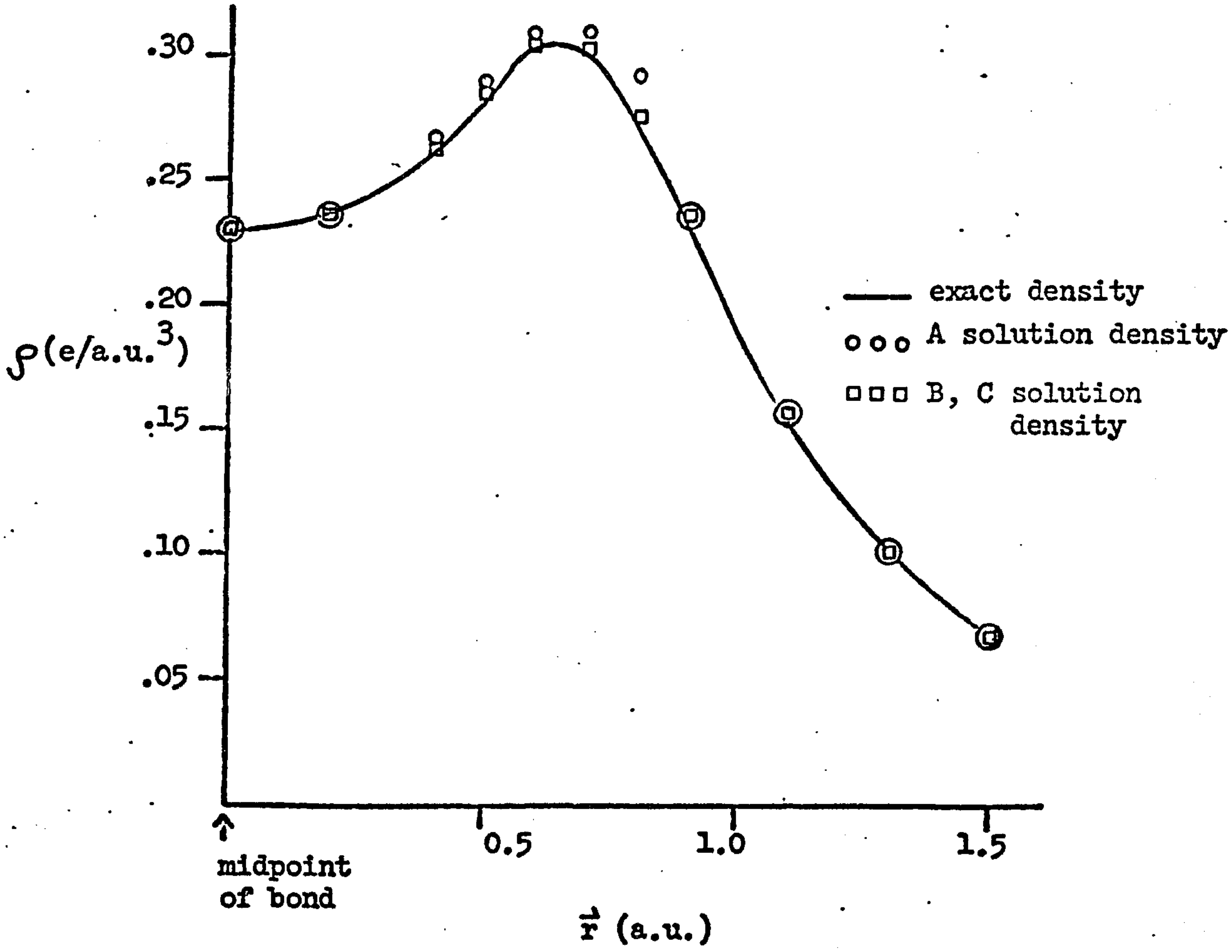
Figure VII. Model molecular density, exact basis

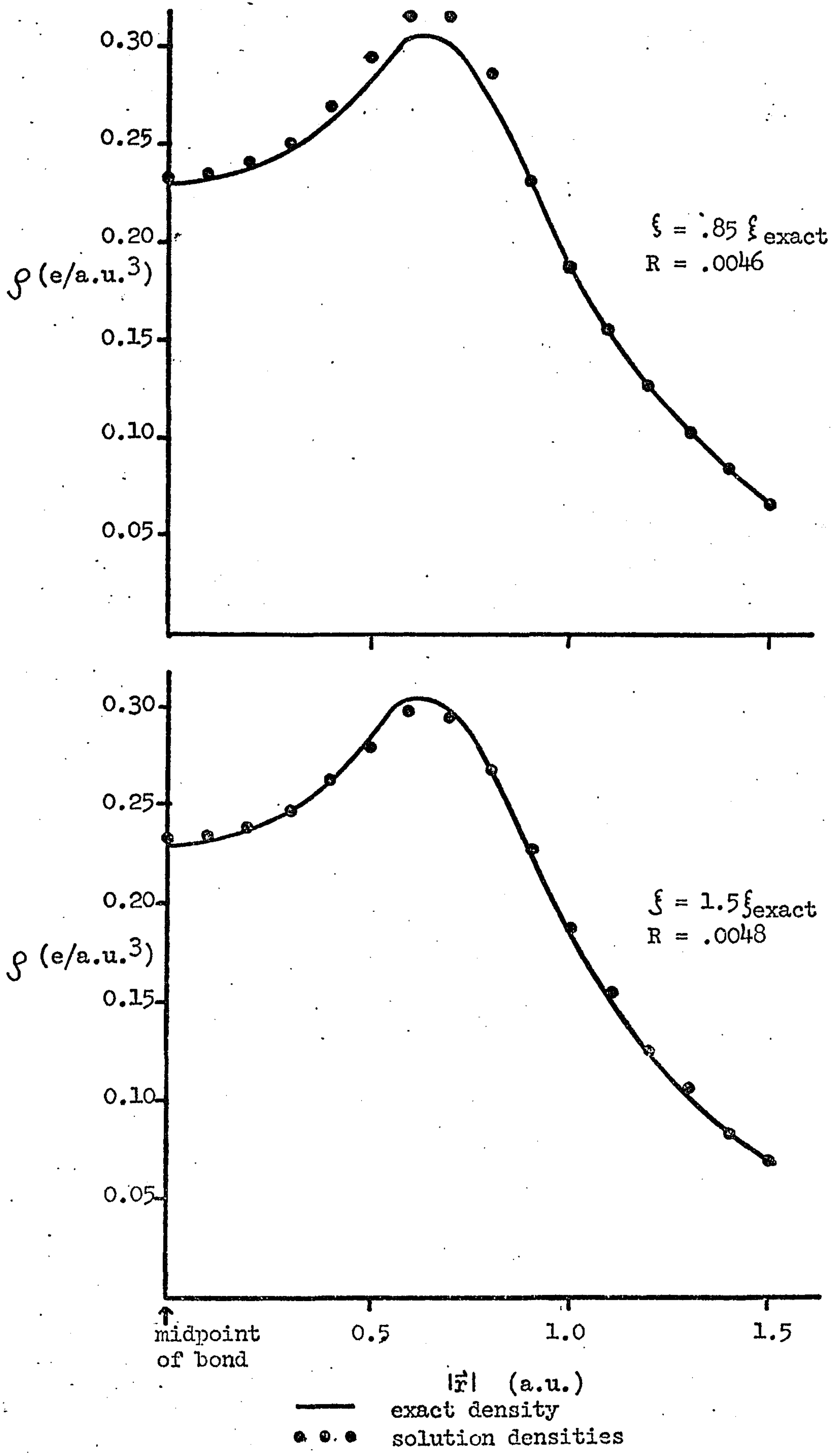
five initial guesses




.30
.25
.20
ρ(e/a.u.^{3})
.15
.10
.05
A
B
C
.25
.20
ρ(e/a.u.^{3})
.15
.10
.05
exact density
● ● ● A,B,C solution densities
midpoint of bond
0.5
1.0
1.5
$\vec{r}$ (a.u.)


.30
.25
.20
ρ(e/a.u.³)
.15
.10
.05
A
B
C


.30
.25
.20
ρ(e/a.u.³)
.15
.10
.05
exact density
○○○ A solution density
□□□ B, C solution density
midpoint of bond
0.5
1.0
1.5
r⃗ (a.u.)


0.30
0.25
0.20
0.15
0.10
0.05
ρ (e/a.u.³)
ξ = .85 ξexact
R = .0046
0.30
0.25
0.20
0.15
0.10
0.05
ρ (e/a.u.³)
ξ = 1.5 ξexact
R = .0048
midpoint of bond
0.5
1.0
1.5
|r⃗| (a.u.)
exact density
solution densities

10. Hydrogen Molecule

To investigate a somewhat more realistic molecular problem, the spherical hydrogen atomic scattering factors reported by Stewart, Davidson and Simpson (1965) were used to construct a set of molecular hydrogen scattering factors according to

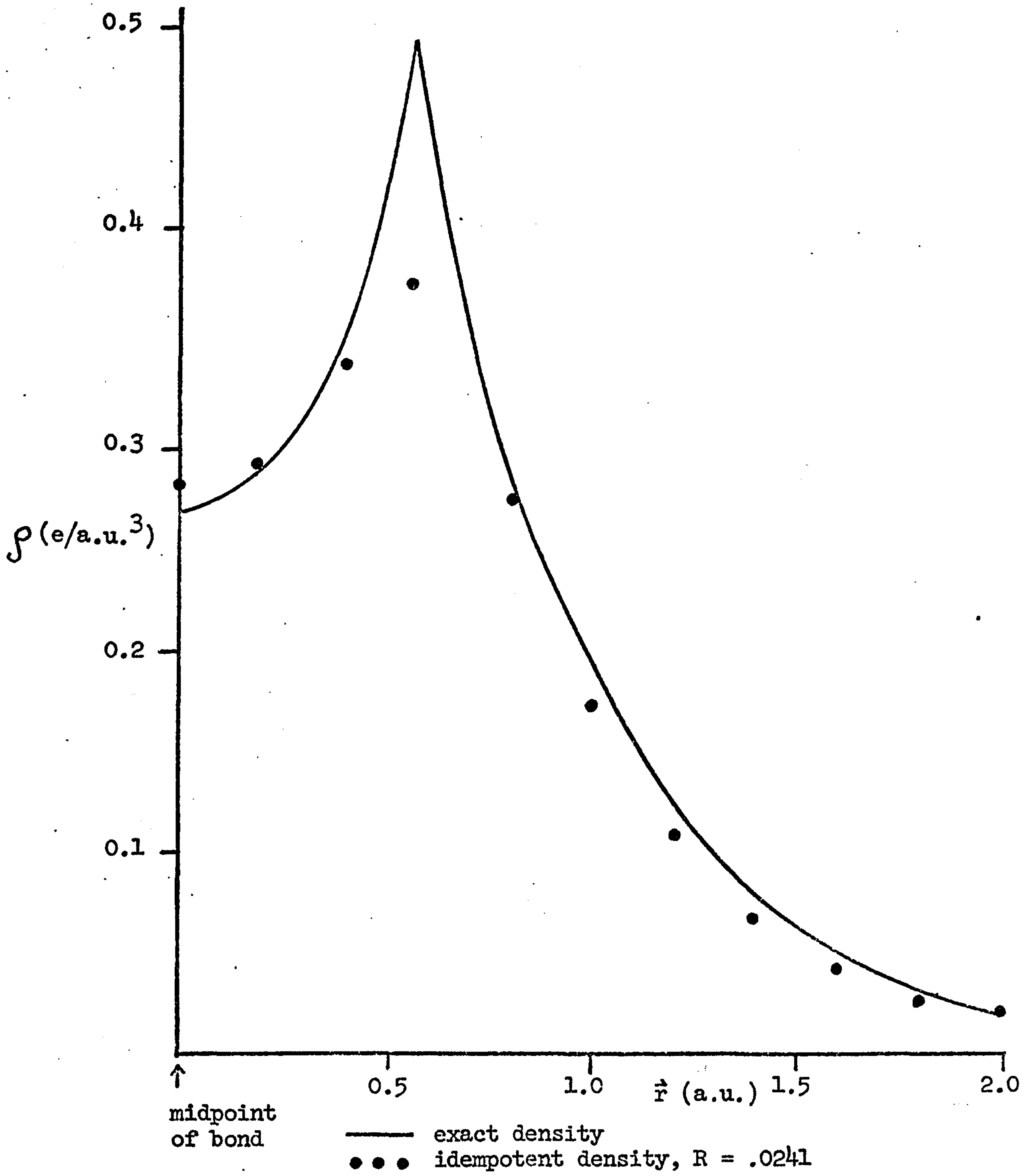
$$F_{mol}(\vec{S}) = F_{atom}(\vec{S})\left[e^{i\frac{\vec{R}}{2}\cdot\vec{S}} + e^{-i\frac{\vec{R}}{2}\cdot\vec{S}}\right] \qquad (9)$$

and

$$I_{mol}(\vec{S}) = F_{mol}(\vec{S}) * F^{*}_{mol}(\vec{S}) \qquad (10)$$

The interatomic distance $\vec{R}$, is .81 x 1.4009 a.u. which Stewart, Davidson and Simpson (1965) found to be the best distance to fit the Kolos-Roothaan H wave function by the sum of two spherical atoms. The largest nineteen scattering factors, from $\sin\theta/\lambda = .0215$ to $\sin\theta/\lambda = .6011$ $\mathring{A}^{-1}$ were used for each of four planes whose angle with the internuclear bond was 0°, 20°, 50° and 90°. This makes a total of seventy-six scattering factors. Huzinaga's (1965) gaussian expansions of Slater-type-orbitals with variable exponential factors were used to expand the density. The results of various basis sets are shown in figures XI through XIV. The similarity of all the results indicates that the convergence of and solutions to the iterative equations are not very sensitive to the particular form of the basis.

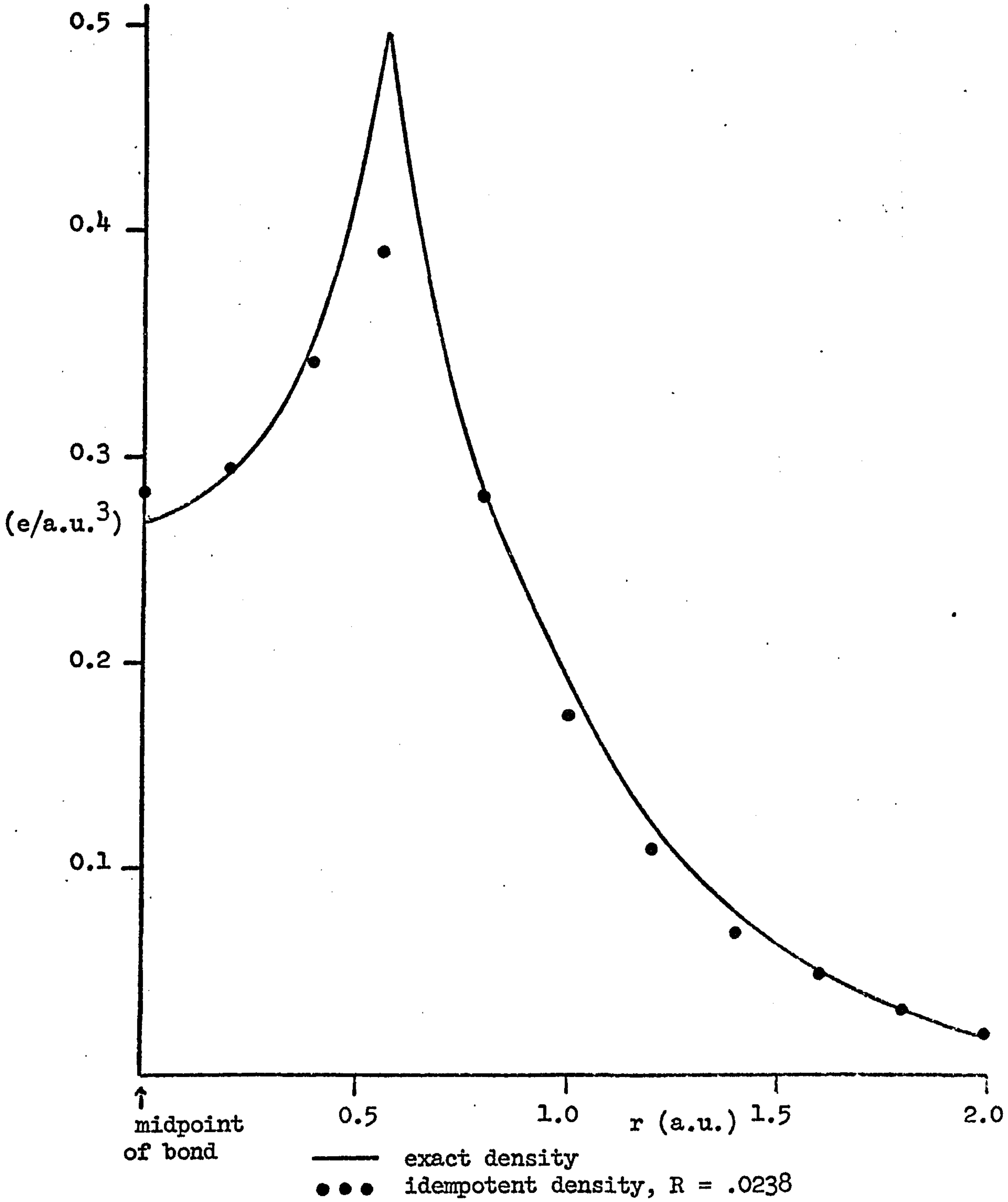
Figure XI. Hydrogen Molecule



basis:

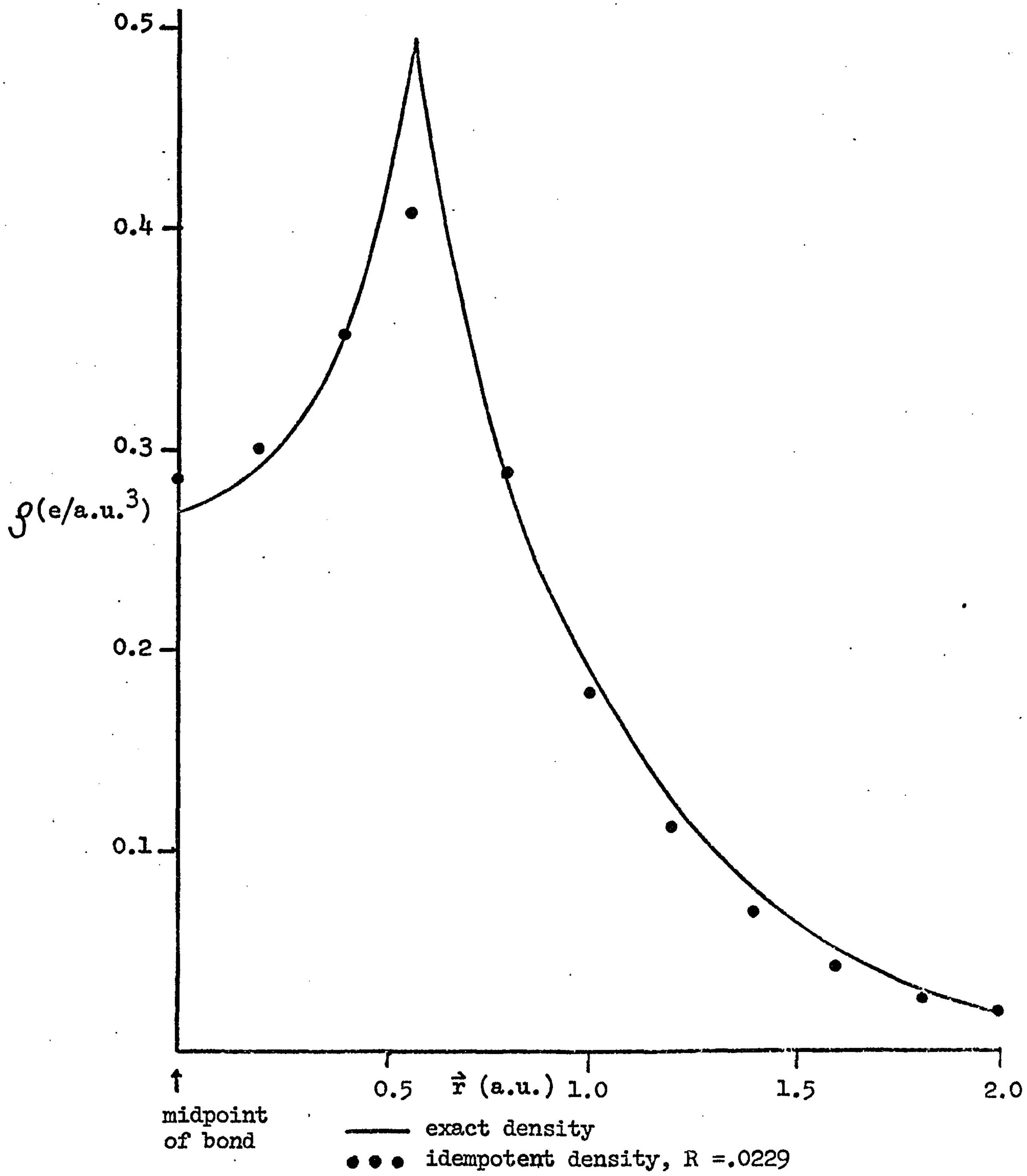
$H1s\ \xi = 1.2$
$H2s\ \xi = 1.2$
$H3s\ \xi = 1.2$

Figure XII. Hydrogen Molecule



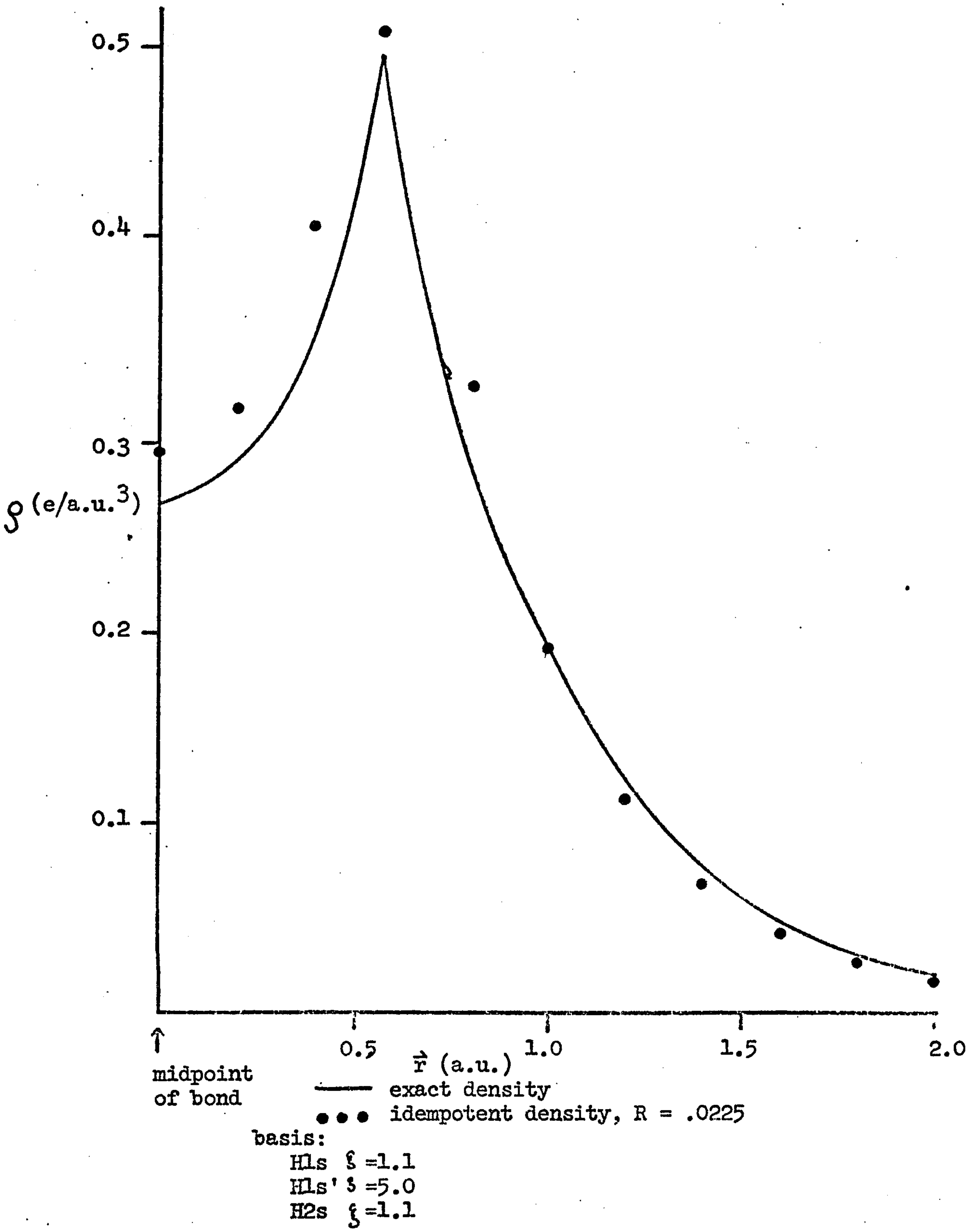
basis:
H1s ξ =1.2
H1s' ξ =1.3
H2s ξ =1.2

Figure XIII. Hydrogen Molecule



basis:
H1s ξ =1.2
H1s' ξ =0.95
H2s ξ =1.2

Figure XIV. Hydrogen Molecule



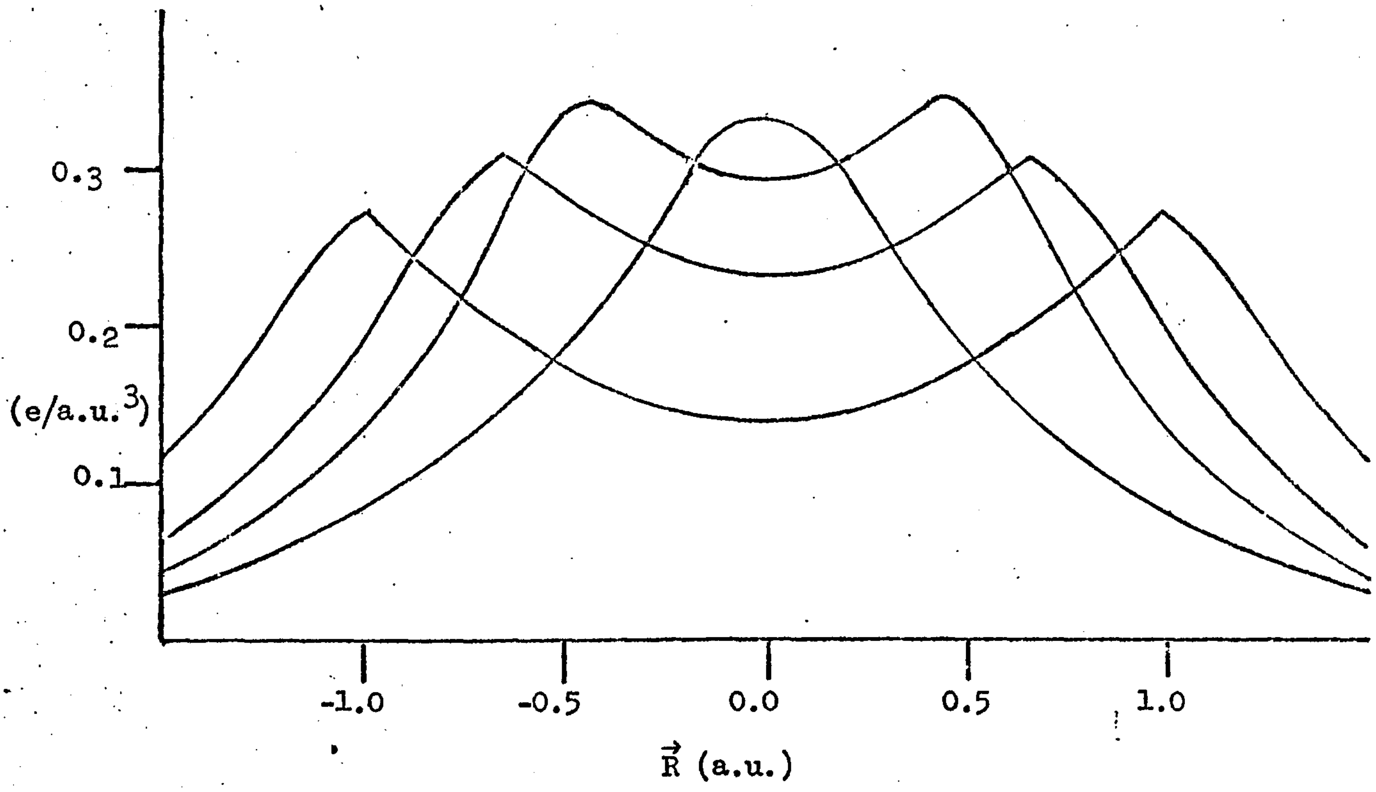
IV. The Phase Problem

The major problem of crystallography has always been the determination of the phases from knowledge of only the magnitudes of the structure factors. One of the most successful techniques was introduced by A.L. Patterson in 1935 when he showed that performing a fourier synthesis, (equation I.6) using the phaseless intensities, $|F^2|$, as coefficients will produce a map where peaks correspond to inter- atomic vectors in the unit cell. The peak heights are proportional to the product of the atomic numbers of the two atoms to which the peak corresponds. Although it is not possible to obtain a structure directly from a Patterson map, it may be possible to deduce the positions of a few atoms and thus obtain a model from which to start. For a unit cell containing N atoms, a Patterson map will occupy the same volume as the unit cell and contain N^2 peaks. (N of these peaks will be at the origin corresponding to the inter- atomic vector of each atom with itself.) The crowding and broadening of peaks may make Patterson maps difficult to unravel unless the unit cell contains one or two heavy atoms. In this case the peaks will be of greater intensity than the others and easy to pick out. While the presence of a heavy atom makes a phasing model easier to find, it dominates the scattering and makes the final fourier syntheses relatively insensitive to the location of the lighter atoms.

Since our formalism requires only a knowledge of the intensities it seems worthwhile to examine its ability to determine structural parameters. Using our defined "H_2 -molecular" density described in Chapter III, Sec.9 , the same basis functions used to define the density were moved to 0., 0.5, 1.0, 1.5 atomic units along the internuclear axis in separate calculations. The solution densities all peaked close to where the basis functions were placed (figure I), but a plot of ϵ vs. R has a minimum at the correct internuclear distance. Similar calculations on the molecular density published by Stewart, et al (1965), (figure II), found the correct internuclear distance for the hydrogen molecule composed of two non-spherical atoms. The shapes of the ϵ vs. R curves are quite similar; they increase steeply once the internuclear distance has been exceeded.

Although the above results are encouraging this method may not lead to the correct structural parameters unless an excellent basis is used. If a less than an adequate basis is employed, the best fit to the density may not peak at the nuclei. Using a more generalized basis set than atomic functions, e.g., plane waves, with the iterative equations may be sufficient to find a phasing model without recourse to the trial and error approach. If an investigation with generalized, space filling functions were to prove successful, this will be a major contribution to the phase problem.

Figure I. Determination of Bond Length for a Model Density



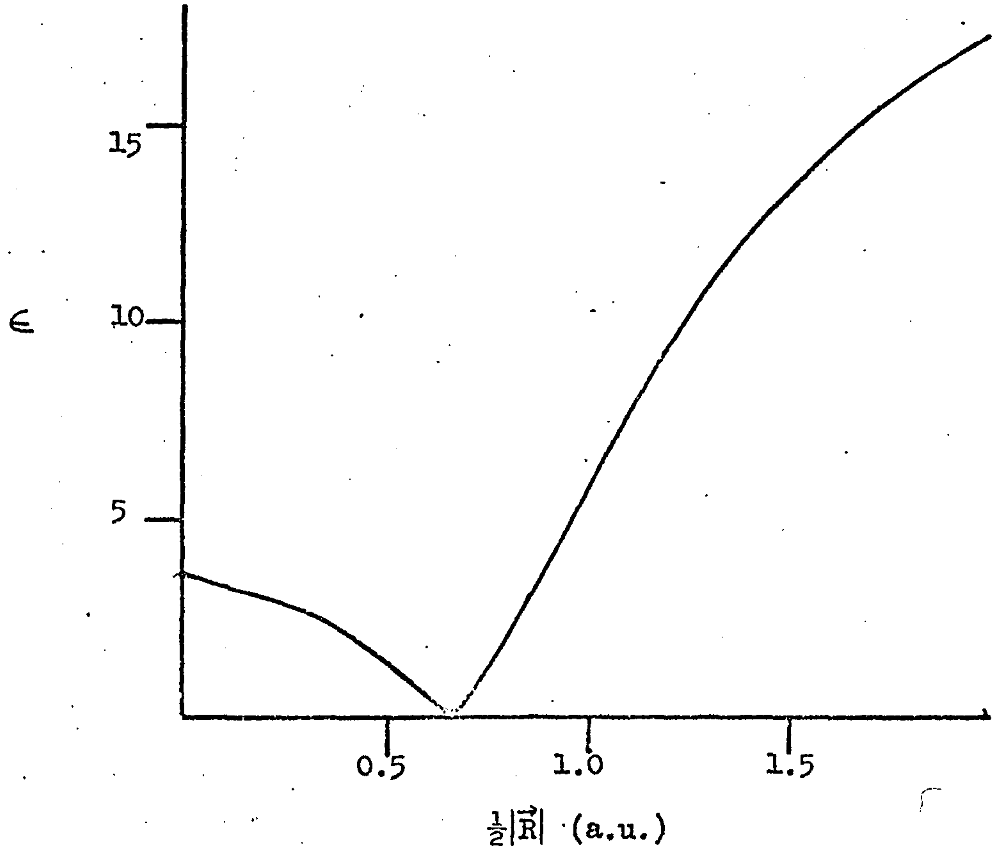
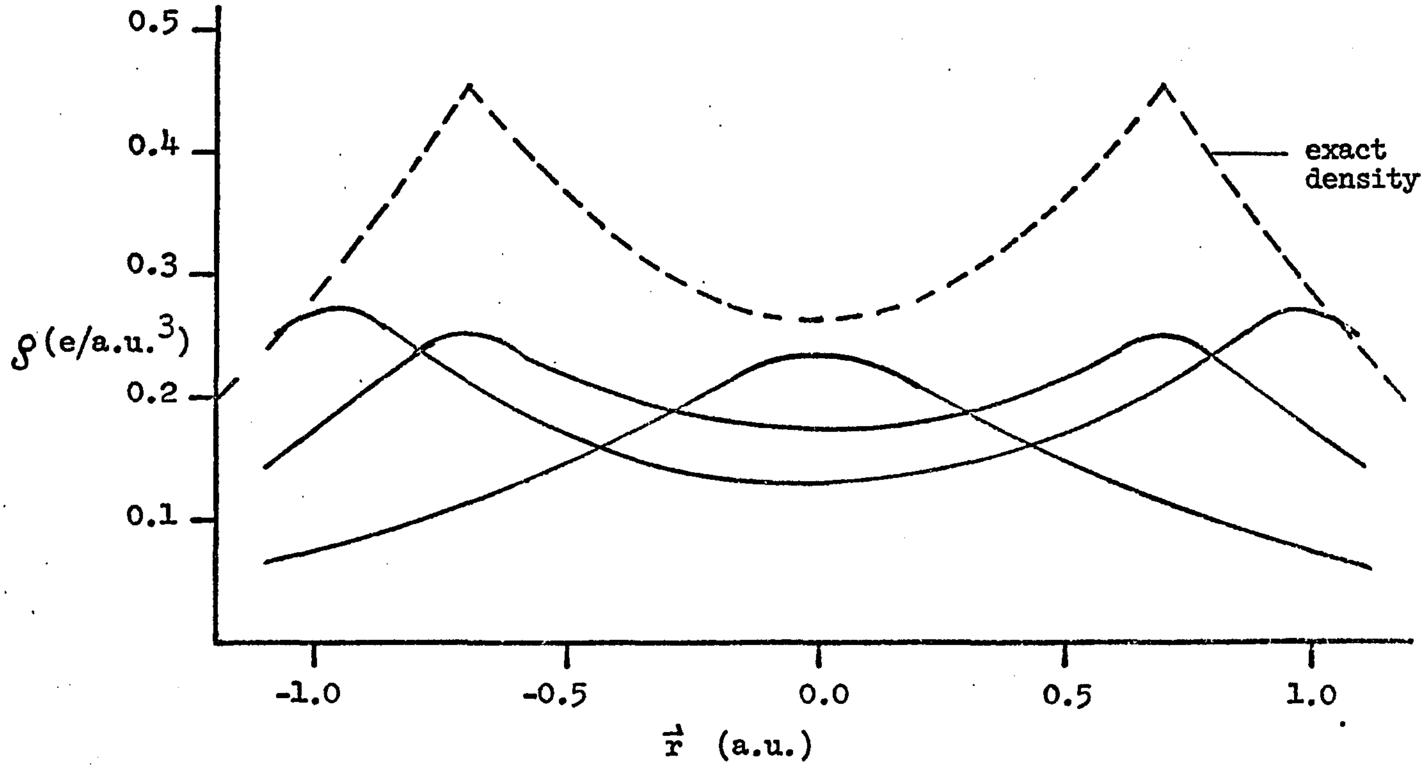
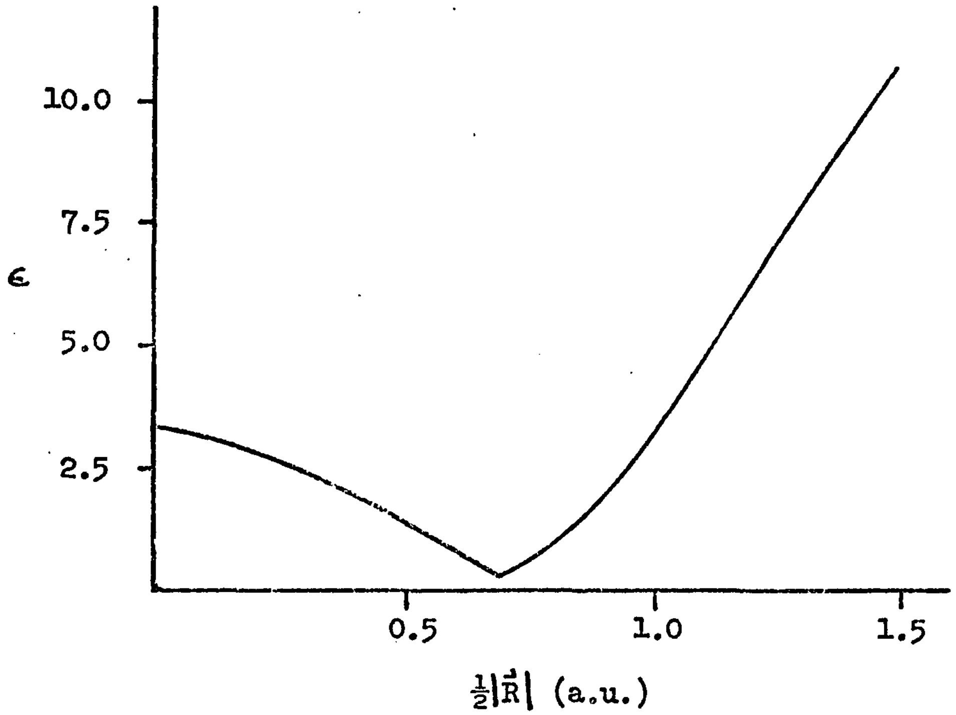


Figure II. Determination of Bond Length for the Hydrogen Molecule





V. Correlation of Parameters

The direct minimization of the functional, equation II.19, leads to the least squares normal equations for the "best" unrestricted parameters. If all weights are taken to be unity we can write the normal equations in matrix form as

$$\underset{\sim}{p} = (\underset{\sim}{f}^{+}\underset{\sim}{f})^{-1}\underset{\sim}{f}^{+}\underset{\sim}{F} \tag{1}$$

where $\underset{\sim}{p}$ is a column matrix of the parameters (in this case all independent elements of the $\underset{\sim}{P}$-matrix), $\underset{\sim}{f}$ is the rectangular matrix of the coefficients of these parameters in the observational equations, and $\underset{\sim}{F}$ is the column matrix of observations. The matrix $\underset{\sim}{f}^{+}\underset{\sim}{f}$ contains the correlation coefficients of all the parameters (see Appendix II). It is clear that no solution can be found if the matrix $\underset{\sim}{f}^{+}\underset{\sim}{f}$ is singular. A major disadvantage of a least squares analysis of the charge density of a molecular crystal is that any linear dependencies in the basis set will not allow a solution. This is not an uncommon situation among atomic basis functions. For example, the 2s S.T.O. and the spherical average of the three 2p S.T.O.'s have identical charge distributions. A least squares analyses cannot use these functions simultaneously but must reduce the basis by taking a linear combination of them. Since the parameters, $\underset{\sim}{p}$, include overlap terms between basis functions some of these products may also be completely correlated. An example is the $p_x(a)\ p_y(b)$ and $p_y(a)\ p_x(b)$ terms where the z-axis is along the bond of atoms a and b (Coppens, Willoughby and Csonka 1971). Another complication in the analysis of diffraction data is that thermal vibrations obscure nodes in density functions and linear dependencies may appear where they would not be expected.

The iterative equations are not affected in the same manner. The only matrix to be inverted in the course of solution is the 2-dimensional matrix of traces, equation (II.17). The determinant of this matrix is

$$\Delta = \mathrm{tr}\,\underset{\sim}{S}\underset{\sim}{S}\,\mathrm{tr}\,\underset{\sim}{G}\underset{\sim}{G} - (\mathrm{tr}\,\underset{\sim}{S}\underset{\sim}{G})^2 \tag{2}$$

where $\underset{\sim}{S}$ is the overlap matrix and $\underset{\sim}{G}$ has been defined previously. If $\mathrm{tr}\underset{\sim}{A}\underset{\sim}{B}$ is thought of as the scalar product of two n -dimensional vectors, A and B, Schwarz' Inequality requires expression (2) to always be greater than or equal to zero. Furthermore, it can only take on the value zero when $\underset{\sim}{S} = k\underset{\sim}{G}$ (when the only reflection considered is scattering in the forward direction); or if either $\underset{\sim}{S}$ or $\underset{\sim}{G}$ is the null vector (an impossible situation for $\underset{\sim}{S}$ and highly unlikely for $\underset{\sim}{G}$). Clearly linear dependencies in the basis will not force expression (2) to be zero. The iterative equations will and have converged even when two or more of the basis functions are chosen to be identical, i.e., 100% correlation. The $\underset{\sim}{P}$-matrices found will of necessity not be unique because the same density can be obtained by mixing the occupations among the identical orbitals. Of course, these equations may not converge in any given case but not due to linear dependencies in the basis set.

To illustrate this point consider the following simple example. We choose 2 "f"-matrices which were constructed so that 100% correlation will occur. They are

$$f_1 = \begin{bmatrix} .9 & .2 \\ .2 & .7 \end{bmatrix} \quad ; \quad f_2 = \begin{bmatrix} .5 & .1 \\ .1 & .4 \end{bmatrix}$$

We define the solution by $\underset{\sim}{P}_{\text{solution}} = \begin{pmatrix} .64 & .48 \\ .48 & .36 \end{pmatrix}$ so that the constraints are F1=1.02, F2=0.56. Taking

$$P_{NI} = \begin{bmatrix} p & q \\ q & 1-p \end{bmatrix}$$

to be the normalized form of the not idempotent matrix we construct the least squares matrix, $\underset{\sim}{a}^{+}\underset{\sim}{a}$ according to the equations of condition, $\text{tr}\underset{\sim}{P}_{NI}\,\underset{\sim}{f}_n = F_n$.

$$a^{+}a = \begin{bmatrix} .2 & .1 \\ .1 & .05 \end{bmatrix}$$

from which it is clearly seen to be singular. Although a solution exists it is impossible for the least squares method to reach it without first eliminating the linear dependencies. On the other hand the $\underset{\sim}{P}$-equations converge in very few iterations starting with the initial guess

$$P_0 = \begin{bmatrix} 1.0 & 0. \\ 0. & 0. \end{bmatrix}$$

to the exact solution (above).

When a correlation is large but less than 100%, the standard deviations of the individual parameters are large, and the determinant of $\underset{\sim}{f}^{+}\underset{\sim}{f}$ is small. There exists a unique solution to the least squares normal equations but numerical problems might make this best solution difficult to find. Many solutions will satisfy the normal equations to a given number of decimal places. These solutions will not increase the value of the functional significantly because a deviation in one parameter is compensated by a deviation in a correlated parameter. The scattering experiment is not sufficient to make a decision between these two parameters. Although the individual population coefficients may not be well determined, it is still possible to obtain an accurate description of the charge density.

Crystallographers engaged in a least squares refinement on population parameters are faced with the task of deciding which parameters should be eliminated from the set based upon the magnitudes of the correlation matrix. This procedure has two motivations, one of which is to reduce the number of parameters in order to make the problem tractable. The other is to aid in the comparison of their results with independent semi-empirical or ab-initio calculations. Some effort has been spent in converting the basis into an eigenparameter basis (i.e., one in which the correlations are identically zero) by a unitary transformation on the correlation matrix. This runs into complications when a scale factor is included in the parameter set because it gets mixed along with the population parameters and the interpretation of these mixed eigenparameters is ambiguous. It would be wiser to compare charge densities directly rather than do a population analysis based on eigenparameters or otherwise, because comparison of densities is always valid, regardless of correlations. A way to reduce correlations would be to include information from other experiments which does not depend solely on the charge density. Inclusion of other constraints are easily accomodated with the iterative equations, as discussed previously.

VI. Method of Steepest Descent

In order to explore another way of obtaining idempotent solutions and as a check on the iterative equations, the steepest descent algorithm published in 1973 by Clinton, Frishberg, Massa and Oldfield, was implemented. The algorithm was modified to use intensity constraints rather than structure factors. (The algorithm as published uses a matrix which is only hermitian when all basis functions share the same center of symmetry.) It is similar to the steepest descent algorithm published by Clinton, Galli and Massa, 1969, which was shown there to be equivalent to the iterative P- equations in the same paper, but differs in one important aspect. In that case the functional minimized was $\mathrm{tr}(P^2 - P)^2$, for this problem we minimized

$$\mathcal{J}(\underset{\sim}{P}) = \sum_{\vec{S}} w(\vec{S})\left[I(\vec{S}) - 4\,\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{f}(\vec{S})\,\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{f}^{\dagger}(\vec{S})\right]^2 \qquad (1)$$

subject to the conditions $\mathrm{tr}P = N$ and $P^2 = P$ for an ortho-normal basis. The procedure follows McWeeny (1960). To first order, the variation with respect to $\underset{\sim}{P}$ of functional (1) is

$$\begin{aligned}\delta\mathcal{J} = -8\,\mathrm{tr}\Big\{&\sum_{\vec{S}} w(\vec{S})\left[I(\vec{S}) - 4\,\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{f}(\vec{S})\,\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{f}^{\dagger}(\vec{S})\right]\\ &\times\left[(\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{f}(\vec{S}))\underset{\sim}{f}^{\dagger}(\vec{S}) + (\mathrm{tr}\,\underset{\sim}{P}\underset{\sim}{f}^{\dagger}(\vec{S}))\underset{\sim}{f}(\vec{S})\right]\Big\}\delta\underset{\sim}{P}\end{aligned} \qquad (2)$$

which can be written as

$$\delta \mathcal{F}(\underset{\sim}{P}) = \mathrm{tr}\, \delta \underset{\sim}{P}\, \underset{\sim}{A} \tag{2'}$$

where $\underset{\sim}{A}$ is the hermitian matrix defined by

$$\begin{aligned} \underset{\sim}{A} \equiv -8 \sum_{\vec{S}} w(\vec{S}) \left\{ I(\vec{S}) - 4\, \mathrm{tr}\, \underset{\sim}{P} \underset{\sim}{f}(\vec{S})\, \mathrm{tr}\, \underset{\sim}{P} \underset{\sim}{f}^{+}(\vec{S}) \right\} \\ \times \left\{ \left[\mathrm{tr}\, \underset{\sim}{P} \underset{\sim}{f}(\vec{S}) \right] \underset{\sim}{f}^{+}(\vec{S}) + \left[\mathrm{tr}\, \underset{\sim}{P} \underset{\sim}{f}^{+}(\vec{S}) \right] \underset{\sim}{f}(\vec{S}) \right\} \end{aligned} \quad . \tag{3}$$

$\underset{\sim}{A}$ will be the null matrix if all data points fit exactly. Starting with a normalized, idempotent initial guess, $\underset{\sim}{P}_0$, we expand $\delta \underset{\sim}{P}$ as

$$\begin{aligned} \delta \underset{\sim}{P} = \underset{\sim}{P}_0 \underset{\sim}{\Delta} (\underset{\sim}{1} - \underset{\sim}{P}_0) + (\underset{\sim}{1} - \underset{\sim}{P}_0) \underset{\sim}{\Delta} \underset{\sim}{P}_0 - \underset{\sim}{P}_0 \underset{\sim}{\Delta} (\underset{\sim}{1} - \underset{\sim}{P}) \underset{\sim}{\Delta} \underset{\sim}{P}_0 \\ + (\underset{\sim}{1} - \underset{\sim}{P}_0) \underset{\sim}{\Delta} \underset{\sim}{P}_0 \underset{\sim}{\Delta} (\underset{\sim}{1} - \underset{\sim}{P}_0) + - - - \end{aligned} \tag{4}$$

where Δ is an arbitrary hermitian matrix. Substituting only the first two terms of the expansion of $\delta \underset{\sim}{P}$ in equation (2') we get

$$\delta \mathcal{F}(\underset{\sim}{P}) = \mathrm{tr}\, \underset{\sim}{\Delta} \left\{ (\underset{\sim}{1} - \underset{\sim}{P}_0) \underset{\sim}{A} \underset{\sim}{P}_0 + \underset{\sim}{P}_0 \underset{\sim}{A} (\underset{\sim}{1} - \underset{\sim}{P}_0) \right\} \tag{5}$$

after a cyclic permutation of terms under the trace. The descent of $\mathcal{F}$ towards zero is optimized if $\underset{\sim}{\Delta}$ is chosen anti-parallel to the matrix in curly brackets. We therefore take

$$\underset{\sim}{\Delta} = -\lambda\left\{(\underset{\sim}{1}-\underset{\sim}{P}_0)\underset{\sim}{A}\underset{\sim}{P}_0 + \underset{\sim}{P}_0\underset{\sim}{A}(\underset{\sim}{1}-\underset{\sim}{P}_0)\right\} \tag{6}$$

where λ is a real positive number. To find the optimum magnitude of the step we insert equation (6) into equation (5) and determine λ such that $d(\delta\mathcal{F}(P))/d\lambda=0)$. This leads to

$$\lambda = \frac{\mathrm{tr}\,\underset{\sim}{P}_0\underset{\sim}{A}(\underset{\sim}{1}-\underset{\sim}{P}_0)\underset{\sim}{A}}{\mathrm{tr}\,(\underset{\sim}{1}-\underset{\sim}{P}_0)\underset{\sim}{A}\underset{\sim}{P}_0\underset{\sim}{A}(\underset{\sim}{1}-\underset{\sim}{P}_0)\underset{\sim}{A} - \mathrm{tr}\,\underset{\sim}{P}_0\underset{\sim}{A}(\underset{\sim}{1}-\underset{\sim}{P}_0)\underset{\sim}{A}\underset{\sim}{P}_0\underset{\sim}{A}} \tag{7}$$

and

$$\underset{\sim}{P}' = \underset{\sim}{P}_0 - \lambda\left\{\underset{\sim}{P}_0\underset{\sim}{A}(\underset{\sim}{1}-\underset{\sim}{P}_0) + (\underset{\sim}{1}-\underset{\sim}{P}_0)\underset{\sim}{A}\underset{\sim}{P}_0\right\} \tag{8}$$

Due to the truncation of expansion (4), $\underset{\sim}{P}'$ will not be idempotent unless $\underset{\sim}{P}o$ is very close to solution. It is therefore necessary to use as $\underset{\sim}{P}_{n+1}$ the nearest idempotent matrix to $\underset{\sim}{P}'$ which can usually be found by applying the iteration

$$\underset{\sim}{P}_{n+1} = 3\underset{\sim}{P}_n^2 - 2\underset{\sim}{P}_n^3 \tag{9}$$

until $\underset{\sim}{P}_{n+1}^2 = \underset{\sim}{P}_{n+1}$ (McWeeny 1956).

The only successful applications of this steepest descent procedure has been to the spherical hydrogen atom data described in Section III.4 using a hydrogenic basis. The convergence of this algorithm is highly initial guess dependent, and when it does converge it may not lead to the best solution. For those initial guesses which do lead to the lowest possible ϵ, convergence is quite rapid generally requiring about fifteen to twenty-five cycles. This is much faster than the algorithm previously discussed as there is no successive lowering of ϵ, ϵ is decreased directly. The usefulness of this approach is limited because of the frequently encountered situation where the steepest descent procedure gets stuck at a solution far from the minimum functional, or diverges upon application of equation (9). The latter problem may (or may not) be overcome by carrying out the variation in (5) to higher order or by using an iteration to idempotency with convergence properties superior to (9).

VII. Conclusions and Suggestions for Future Work

We have presented in this dissertation a formalism for obtaining idempotent density matrices for the unit cell of a molecular crystal whose x-ray diffraction intensities have been measured. The convergence of this algorithm to reasonable results for simple atomic and molecular systems has been demonstrated. The idempotency insures that the 1-matrix solutions are derived from a wave function which takes the form of a single Slater determinant. Where conventional unrestricted (except for normalization) least squares analyses have been applied to the same data sets, the 1-matrices found were never N-representable as shown by their eigenvalues which always exceeded the boundaries of zero and one.

The algorithm is sufficiently developed to proceed to more complicated structures. It is recommended that the next data set to be tested is the diborane crystal. It is a simple polyatomic molecule but still represents a significant departure from the diatomic systems investigated thus far. Theoretical structure factors and intensities can be calculated from Jone's and Lipscomb's (1970) work. This is an intermediate step between the theoretical hydrogen molecule and real data on molecular crystals. The results can be compared to the least squares analyses done by Jones, Pautler and Coppens (1972). Using this theoretical data the temperature factor formalism (Chapter III.5) can be tested and the effect of imposing nuclear cusp conditions can be investigated. The ideas discussed in Chapter III.7 for eliminating possible initial guess dependency can be implemented and their effectiveness evaluated.

Since it is difficult to measure absolute intensities experimentally, most refinements on diffraction data include a scale factor, k, defined by

$$I(\vec{S}) = kI_{obs}(\vec{S}) \qquad (1)$$

as a parameter to be determined. Least squares refinements on charge density coefficients which include a scale factor are hampered by correlations between these two very different types of parameters. Preliminary results (not reported here) indicate that our algorithm will find the correct scale factor when k is varied and a plot of ϵ vs. k is made. Further study into the effects of different basis sets in determining k is required. The information gained about the properties of the algorithm can then be used in analyzing the original data.

One of the most serious impediments to a charge density analysis of x-ray data using atomic orbital products is the large number of parameters to be determined. With the exception of diborane, no system has been analyzed by coventional least squares methods which includes contributions to the density from overlap terms on non-adjacent atoms. Going from an unrestricted $\underset{\sim}{P}$-matrix to one which is idempotent cuts down the degrees of freedom drastically but, as table II.1 indicates, for many molecules of interest this reduction is insufficient. The number of parameters can be further reduced by imposing upon our idempotent $\underset{\sim}{P}$ constraints of the form

$$\mathrm{tr}\, P\Theta(i,j) = 0. \tag{2}$$

when the elements of $\underset{\sim}{\Theta}(i,j)$ are defined by

$$\Theta_{kl}(i,j) \equiv \delta_{ki}\delta_{lj} \tag{3}$$

Because idempotency severely limits the number of parameters in a matrix, many of the non-adjacent atom overlap terms with substantial magnitudes (not an uncommon situation for the more diffuse functions) can still be retained. In particular those terms which have significant effects on other properties of interest (Bentley and Stewart 1974) need not be eliminated.

As discussed in Chapter II, the single determinant wave function appears to be an adequate approximation for the analysis of diffraction data of the highest quality presently being obtained. Should future advances in intensity measurements and data reduction exceed the independent particle model, it would be desirable to develop an algorithm which finds a non-idempotent but N-representable density matrix. In this case one could never go back to the wave function as in the idempotent case, but all one- body expectation values would be calculable. One alternative is to look for a sum of idempotent matrices. Another approach might take advantage of the fact that every N-body density matrix for a pure state is idempotent and apply the algorithm as it stands with an appropriate N-body basis. Knowledge of the N-matrix allows all expectation values to be determined.

Appendix I. Factorization of an Idempotent Matrix Referred to a Non Orthogonal Basis

Starting with a $\mathbf{P}$ matrix (referred to a basis whose overlap matrix is $\mathbf{S}$) satisfying the conditions

$$\mathbf{P}\mathbf{S}\mathbf{P} = \mathbf{P}$$

$$\mathrm{tr}\mathbf{P}\mathbf{S} = N \tag{1}$$

$$\mathbf{P} = \mathbf{P}^{\dagger}$$

we wish to show that $\mathbf{P}$ can be written as

$$\mathbf{P} = \mathbf{C}^{\dagger}\,\mathbf{C}\ . \tag{2}$$

We define the matrix

$$\mathbf{R} \equiv \mathbf{S}^{1/2}\mathbf{P}\mathbf{S}^{1/2}, \tag{3}$$

which is idempotent, normalized and hermitian:

$$\mathbf{R}^{2} = \mathbf{S}^{1/2}\mathbf{P}\mathbf{S}^{1/2}\mathbf{S}^{1/2}\mathbf{P}\mathbf{S}^{1/2} = \mathbf{R}$$

$$\mathrm{tr}\mathbf{R} = \mathrm{tr}\mathbf{S}^{1/2}\mathbf{P}\mathbf{S}^{1/2} = \mathrm{tr}\mathbf{P}\mathbf{S} = N \tag{4}$$

$$\mathbf{R}^{\dagger} = \mathbf{S}^{1/2\dagger}\mathbf{P}^{\dagger}\mathbf{S}^{1/2\dagger} = \mathbf{R}$$

We proceed to diagonalized R with a unitary transformation

$$\underset{\sim}{U}\underset{\sim}{R}\underset{\sim}{U}^{\dagger} = \underset{\sim}{d} \tag{5}$$

Since $\underset{\sim}{d}^{2} = \underset{\sim}{d}$, its only elements can be zeros and ones. The normalization condition further requires that there be exactly N ones along the diagonal. Performing the reverse transformation on $\underset{\sim}{d}^{2}$ to recover $\underset{\sim}{R}$

$$\underset{\sim}{R} = \underset{\sim}{U}^{\dagger}\underset{\sim}{d}\underset{\sim}{d}\underset{\sim}{U} \tag{6}$$

or

$$\underset{\sim}{R} = \underset{\sim}{T}^{\dagger}\underset{\sim}{T} \tag{7}$$

where $\underset{\sim}{T}$ is a rectangular matrix of the N columns picked out of $\underset{\sim}{U}$ by $\underset{\sim}{d}$. Thus

$$\underset{\sim}{P} = \underset{\sim}{S}^{-1/2}\underset{\sim}{R}\underset{\sim}{S}^{-1/2} = \underset{\sim}{S}^{-1/2}\underset{\sim}{T}^{\dagger}\underset{\sim}{T}\underset{\sim}{S}^{-1/2} \tag{8}$$

for which (2) follows when the identification

$$\underset{\sim}{C} = \underset{\sim}{T}\underset{\sim}{S}^{-1/2} \tag{9}$$

is made.

Appendix II. Method of Least Squares

This presentation closely follows that given by W.C. Hamilton (1964). We start with a set of observational equations written in matrix form

$$\underset{\sim}{A}\underset{\sim}{X}_t = \underset{\sim}{F}_t \tag{1}$$

where $\underset{\sim}{X}_t$ is a n-dimensional column matrix of the true parameters and $\underset{\sim}{F}_t$ is an m-dimensional column matrix of the true values of the functions defined by the matrix $\underset{\sim}{A}$. Since $\underset{\sim}{F}_t$ is not obtainable by experiment so that (1) could be solved exactly we take many more observations than parameters, $m > n$, and seek a "best" fit by minimizing the sum of the squares of the residuals defined by

$$\underset{\sim}{V} \equiv \underset{\sim}{F}_{obs} - \underset{\sim}{A}\underset{\sim}{X}_e \tag{2}$$

where $\underset{\sim}{X}_e$ are the least squares estimates to the parameters, $\underset{\sim}{X}_t$, and $\underset{\sim}{F}_{obs}$ are the experimentally observed values of the functions (1). Assuming that all observations are uncorrelated and of equal weight (see Hamilton 1964 for the procedure when this assumption is invalid) we minimize

$$S = \underset{\sim}{V}^{+}\underset{\sim}{V} \tag{3}$$

or

$$\delta[(\underset{\sim}{F}_{obs} - \underset{\sim}{A}\underset{\sim}{X}_e)^{+}(\underset{\sim}{F}_{obs} - \underset{\sim}{A}\underset{\sim}{X}_e)] = 0 \tag{4}$$

which leads to

$$2(\delta\underset{\sim}{X}_e)^{+}(-\underset{\sim}{A}^{+}\underset{\sim}{F}_{obs} + \underset{\sim}{A}^{+}\underset{\sim}{A}\underset{\sim}{X}_e) = 0 \;. \tag{5}$$

For arbitrary variation $\delta\underset{\sim}{X}e$ this implies

$$\underset{\sim}{A}^{+}\underset{\sim}{A}\underset{\sim}{X}_e - \underset{\sim}{A}^{+}\underset{\sim}{F}_{obs} = 0 \;. \tag{6}$$

Solving for $\underset{\sim}{X}e$ we obtain the least squares normal equations

$$\underset{\sim}{X}_e = (\underset{\sim}{A}^{+}\underset{\sim}{A})^{-1}\underset{\sim}{A}^{+}\underset{\sim}{F}_{obs} \tag{7}$$

The variances and covariances of all determined parameters will be given by the matrix

$$\underset{\sim}{M} = \mathcal{E}\{(\underset{\sim}{X}_e - \underset{\sim}{X}_t)^{+}(\underset{\sim}{X}_e - \underset{\sim}{X}_t)\} \tag{8}$$

where $\mathcal{E}\{y\}$ means the expectation value of y. Substitution of the matrix equations (2) and (7) gives

$$\underset{\sim}{M} = \mathcal{E}\{[(\underset{\sim}{A}^{+}\underset{\sim}{A})^{-1}\underset{\sim}{A}^{+}(\underset{\sim}{F}_{obs} - \underset{\sim}{F}_t)][(\underset{\sim}{F}_{obs} - \underset{\sim}{F}_t)^{+}\underset{\sim}{A}(\underset{\sim}{A}^{+}\underset{\sim}{A})^{-1}]\} \tag{9}$$

$$\underset{\sim}{M} = (\underset{\sim}{A}^{+}\underset{\sim}{A})^{-1}\underset{\sim}{A}^{+}\,\mathcal{E}\{(\underset{\sim}{F}_{obs} - \underset{\sim}{F}_t)(\underset{\sim}{F}_{obs} - \underset{\sim}{F}_t)^{+}\}\underset{\sim}{A}(\underset{\sim}{A}^{+}\underset{\sim}{A})^{-1} \tag{10}$$

But $\mathcal{E}\left\{(Fobs - Ft)(Fobs - Ft)^{+}\right\}$ is equal to σ^2 x unit matrix under the assumption of uncorrelated observations of equal weight with variance σ^2. Thus

$$\underset{\sim}{M} = \sigma^2 (\underset{\sim}{A}^+\underset{\sim}{A})^{-1}\underset{\sim}{A}^+\underset{\sim}{A}(\underset{\sim}{A}^+\underset{\sim}{A})^{-1} = \sigma^2(\underset{\sim}{A}^+\underset{\sim}{A})^{-1} \tag{11}$$

An unbiased estimate of σ^2 is given by

$$\sigma_e^2 = \underset{\sim}{V}^+\underset{\sim}{V}\left(\frac{1}{m-n}\right) \tag{12}$$

Dividing each element of $(A^+A)^{-1}_{ij}$ by $\sqrt{(A^+A)^{-1}_{ii}\,(A^+A)^{-1}_{jj}}$ will result in a matrix of correlation coefficients.

If a set of constraints,

$$\underset{\sim}{Q}\,\underset{\sim}{X}_e' = \underset{\sim}{Z} \tag{13}$$

are to be imposed upon the solution, use of the Lagrangian multiplier technique yields the result

$$\underset{\sim}{X}_e' = \underset{\sim}{X}_e + (\underset{\sim}{Z}^+ - \underset{\sim}{X}_e^+\underset{\sim}{Q}^+)(\underset{\sim}{Q}(\underset{\sim}{A}^+\underset{\sim}{A})^{-1}\underset{\sim}{Q}^+)^{-1}\underset{\sim}{Q}(\underset{\sim}{A}^+\underset{\sim}{A})^{-1} \tag{14}$$

for the constrained least squares parameters, $\underset{\sim}{Xe}'$.

References

Aldred, P.J.E. and M. Hart (1973), Proc. R. Soc. Lond. A333, 223.

Bats, W. and D. Feil (1974), preprint.

Benesch, R. and V.H. Smith, Jr. (1970), J. Chem. Phys., 53 4, 1466.

Benesch, R. and V.H. Smith (1972), Phys. Rev. A, 5, 1, 114.

Benesch, R., S.R. Singh and V.H. Smith (1971), Chem. Phys. Lett., 10, 2, 151.

Bentley, J. and R.F. Stewart (1974), Acta Cryst,A30, 60.

Campbell, N., (1920), Phil. Mag. S.6., 39, 230, 177.

Clinton, W.L., C.A. Frishberg, L.J. Massa and P.A. Oldfield, (1973), I.J.Q.C., Symp. 7, 505.

Clinton, W.L., A.J. Galli and L.J. Massa (1969), Phys. Rev., 177, 1, 7.

Clinton, W.L. and L.J. Massa (1972a), Phys. Rev. Lett., 20, 1363.

Clinton, W.L. and L.J. Massa (1972b), I.J.Q.C., VI, 519.

Coleman, A.J. (1963), Rev. Mod. Phys., 35, 3, 668.

Coleman, A.J. (1972), J. Math. Phys., 13, 2, 214.

Coppens, P. (1967), Science, 158, 1577.

Coppens, P. (1968), Acta Cryst., B24, 1272.

Coppens, P. and A. Vos (1971), Acta Cryst., B27, 146.

Coppens, P., L. Csonka and T.V. Willoughby (1970),

Science, 167, 1126.

Coppens, P., D. Pautler and J.F. Griffin (1971), J.A.C.S., 93, 1051.

Coppens, P., T.V. Willoughby and L.N. Csonka (1971), Acta Cryst., A27,248.

Dawson, B. (1966), Proc. Roy. Soc. A, 298, 264.

Edgeworth, F.Y. (1888), Phil. Mag., 25, 184.

Freer, S.T. and J. Kraut (1965), Acta Cryst., 19, 992.

Hamilton, W.C. (1964), Statistics in Physical Science, Ronald Press, New York.

Hamilton, W.C. (1965), Acta Cryst., 18, 502.

Hirshfeld, F.L. (1971), Acta Cryst., B27, 769.

Huzinaga, S. (1965), J. Chem. Phys., 42, 4, 1293.

Huzinaga, S. and Y. Sakai (1969), J. Chem. Phys., 50, 3, 1371.

Jones, D.S. and W.N. Lipscomb (1970), Acta Cryst., A26, 196.

Jones, D.S., D. Pautler and P. Coppens (1972), Acta Cryst., A28, 635.

Little, R.G., D. Pautler and P. Coppens (1971), Acta Cryst., B27, 1931.

Lowdin, P.O. (1955), Phys. Rev, 97, 1474.

Matthews, D.A. and G.D. Stucky (1972), Trans. Am. Cryst. Assoc., 113.

Matthews, D.A. and G.D. Stucky (1971), J.A.C.S., 93, 23, 5954.

Matthews, D.A., G.D. Stucky and P. Coppens (1972), J.A.C.S., 9423, 8001.

Mayer, J.E. (1955), Phys. Rev., 100, 1579.

McIver, J.W., P.Coppens and D. Nowack (1971), Chem. Phys. Lett., 11, 1 82.

McWeeny, R. (1951), Acta Cryst., 4, 513.

McWeeny, R. (1952), Acta Cryst., 5, 463.

McWeeny, R. (1953), Acta Cryst., 6, 631.

McWeeny, R. (1959), Phys. Rev., 114, 6, 1528.

McWeeny, R. (1960), Rev. Mod. Phys., 32, 2, 335.

O'Connor, B.H. and E.N. Maslen (1974), Acta Cryst., B30, 383.

Stewart, R.F. (1969a), J. Chem. Phys., 51, 10, 4569.

Stewart, R.F. (1969b), J. Chem. Phys., 50, 6, 2485.

Stewart, R.F. (1973a), J. Chem. Phys., 58, 10, 4430.

Stewart, R.F. (1973b), J. Chem. Phys., 58, 4, 1668.

Stewart, R.F., E.R. Davidson and W.T. Simpson (1965), J. Chem. Phys., 42, 9, 3175.

Stout, G.H. and L.H. Jensen (1968), X-Ray Structure Determination, MacMillan Company, London.

Tredgold, R.H. (1957), Phys. Rev., 105, 5, 1421.

Verschoor, G.C. and E. Keulen (1971), Acta Cryst., B27, 134.