c.1

THEORY AND APPLICATION OF EIGENVALUE INDEPENDENT PARTITIONING IN THEORETICAL CHEMISTRY

by

DAVID WARREN SABO
B. Sc. (Hons.) University of Alberta, 1972

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
in the Department of
CHEMISTRY

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

July, 1977

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Chemistry

The University of British Columbia
Vancouver, British Columbia, Canada V6T 1W5

Date August 26, 1977

ABSTRACT

This work concerns the description of eigenvalue independent partitioning theory, and its application to quantum mechanical calculations of interest in chemistry. The basic theory for an m-fold partitioning of a hermitian matrix H, ($2 \leq m \leq n$, the dimension of the matrix), is developed in detail, with particular emphasis on the 2 x 2 partitioning, which is the most useful. It consists of the partitioning of the basis space into two subspaces -- an n_A-dimensional subspace ($n_A \geq 1$), and the complementary $n-n_A = n_B$-dimensional subspace. Various n_A- (or n_B-) dimensional effective operators, and projections onto n_A- (or n_B-) dimensional eigenspaces of H, are defined in terms of a mapping, f, relating the parts of eigenvectors lying in each of the partitioned subspaces. This mapping is shown to be determined by a simple nonlinear operator equation, which can be solved by iterative methods exactly, or by using a perturbation expansion. Properties of approximate solutions, and various alternative formulas for effective operators, are examined. The theory is developed for use with both orthonormal and non-orthonormal bases.

Being a generalization of well known one-dimensional partitioning formalisms, this eigenvalue independent partitioning theory has a number of important areas of application. New and efficient methods are developed for the simultaneous determination of several eigenvalues and eigenvectors of a large hermitian matrix, which are based on the construction and

diagonalization of an appropriate effective operator. Perturbation formulas are developed both for effective operators defined in terms of f, and for projections onto whole eigenspaces of H. The usefulness of these formulas, especially when the zero order states of interest are degenerate, is illustrated by a number of examples, including a formal uncoupling of the four component Dirac hamiltonian to obtain a two component hamiltonian for electrons only, the construction of an effective nuclear spin hamiltonian in esr theory, and the derivation of perturbation series for the one-particle density matrix in molecular orbital theory (in both Huckel-type and closed shell self-consistent field contexts).

A procedure is developed for the direct minimization of the total electronic energy in closed shell self-consistent field theory in terms of the elements of f, which are unconstrained and contain no redundancies. This formalism is extended straightforwardly to the general multi-shell single determinant case. The resulting formulas, along with refinements of the basic conjugate gradient minimization algorithm, which involve the use of scaled variables and frequent basis modification, lead to efficient, rapidly convergent methods for the determination of stationary values of the electronic energy. This is illustrated by some numerical calculations in the closed shell and unrestricted Hartree-Fock cases.

TABLE OF CONTENTS

	Page
Abstract	iii
List of Tables	xii
List of Figures	xvi
Acknowledgements	xviii
Chapter 1 Eigenvalue Independent Partitioning, An Introduction	1
Chapter 2 2 x 2 Partitioning Theory	9
2.1 Basic Theory	10
2.1.a The f-operator	10
2.1.b The Defining Condition for f	14
2.1.c Rederivation From A Projection Point of View	16
2.1.d The Relationship Between $\hat{T}$ and the Eigenprojections--Covariant and Contravariant Representations	21
2.1.e Variational Formulation of D(f)=0	23
2.1.f Relation Between σ^2 and D(f) --Eigenvalue Dispersion	25
2.1.g Transformation of f Under a Change of Basis	27
2.2 Effective Operators	30
2.2.a Basic Definitions	30
2.2.b Eigenvectors and Eigenvalues of the Effective Operators	34

Page

2.2.c Relationships With Other Formulations 37

2.3 Generalization to a Nonorthonormal Basis 42

Chapter 3 The Effective Hamiltonians--Practical Considerations 46

3.1 Alternative Formulas 47

3.2 Implications of Inexact Solutions 51

3.3 Perturbation Theory for $\hat{H}_A$, G_A, and $\tilde{H}_A$ . . 55

3.3.a The $\hat{H}_A$ Scheme 55

3.3.b The G_A Scheme 59

3.3.c The $\tilde{H}_A$ Scheme 62

Chapter 4 Multiple Partitioning Theory 64

4.1 The m x m Partitioning Formalism 66

4.1.a Basic Theory 66

4.1.b The Defining Condition on the f_{JI} 72

4.1.c Variational Formulation of the Equations for the f_{JI} 74

4.1.d Transformation of the f_{JI} Under a Change of Basis 77

4.2 Effective Operators 79

4.2.a Basic Definitions 79

4.2.b Eigenvalues and Eigenvectors of the Effective Operators 82

4.3 Generalization to a Non-orthonormal Basis 85

4.4 Practical Considerations 91

4.4.a Alternative Formulas 91

		Page
	4.4.b Implications of Inexact Solutions	95
Chapter 5	Exact Determination of $\hat{T}$	98
5.1	The Calculation of a Few Eigenvalues of a Large Hermitian Matrix	100
5.2	2 x 2 Partitioning--Orthonormal Basis . .	104
	5.2.a General Considerations	104
	5.2.b Methods Based on $D^{(1)}(f)$	109
	5.2.c Methods Based on $D^{(2)}(f)$	113
	5.2.d Solution of the Newton-Raphson Equations by Descent Methods . .	117
	5.2.e Extremizing the Trace	120
	5.2.f Minimization of the Norm of D . .	122
	5.2.g Test Calculations	123
5.3	Generalization to a Non-orthonormal Basis -- 2 x 2 Case	131
	5.3.a General Considerations	131
	5.3.b Methods Based on G_{BA} and g_{BA} -- A Generalization of Algorithm SDNR	132
	5.3.c Methods Based on $D^{(2)}(f)$ -- Generalized Nesbet Procedures. .	135
	5.3.d Other Methods	137
	5.3.e Choice of an Initial Estimate, and Improvement of Convergence Rates	140
	5.3.f Test Calculations With Overlap . .	144
5.4	Multiple Partitioning	159
Chapter 6	Perturbation Theory	161
6.1	Introduction	162

Page

6.2 2 x 2 Partitioning -- Orthonormal Basis 164

6.2.a General Discussion 164

6.2.b A-states Degenerate 173

6.2.c A-states Non-degenerate 178

6.3 Examples 183

6.3.a The Dirac Equation 183

6.3.b Derivation of a Spin Hamiltonian -- Strong Field Case 191

6.4 Non-orthonormal Basis -- 2 x 2 Partitioning 195

Chapter 7 Eigenvalue Independent Partitioning And Molecular Orbital Theory 209

7.1 Introduction 210

7.2 Perturbation of the Density Matrix -- Orthonormal Basis 212

7.2.a General Theory 212

7.2.b Huckel Molecular Orbital Theory . 221

7.2.c Numerical Example -- Huckel Theory 224

7.3 Perturbation of the Density Matrix -- Non-orthonormal Basis 235

7.3.a General Theory 235

7.3.b Extended Huckel Molecular Orbital Theory 240

7.4 Self-Consistent Perturbation Theory . . . 243

7.4.a General Theory 243

7.4.b Coupled Hartree-Fock Perturbation Theory 246

Page

7.4.c Uncoupled Hartree-Fock Perturbation Theory 249

Chapter 8 Direct Minimization Self-consistent Field Theory 251

8.1 Introduction 252

8.2 Closed Shell Systems 257

8.2.a Orthonormal Basis 257

8.2.b Non-orthonormal Basis 263

8.2.c Results of Test Calculations . . . 266

8.3 Unrestricted Hartree-Fock Theory 271

8.3.a Energy Derivatives 271

8.3.b Test Calculations and Computational Refinements 275

8.3.c Use of Scaled Variables 282

8.4 Theory for the General Single Determinant Case 290

8.4.a The Basic Variables of the Calculation 290

8.4.b The Energy Variation and First Derivatives 293

8.4.c The Second Derivatives 297

8.4.d Incorporation of the Intershell Orthogonality Constraints . . . 302

8.4.e Example -- The Two Shell System . 305

Bibliography 313

		Page
Appendices		316
Appendix 1	Proofs of Alternative Formulas -- 2 x 2 Partitioning	317
Appendix 2	The 3 x 3 and 4 x 4 Case -- Orthonormal Basis	323
Appendix 3	Proofs of Alternative Formulas -- Multiple Partitioning	327
Appendix 4	Description of Algorithms -- 2 x 2 Case	333
Appendix 5	Rates of Convergence and Asymptotic Error Constants	341
Appendix 6	Algorithms for the Determination of $\hat{T}$ -- Multiple Partitioning Case	349
A6.1	Methods Based on $D_{JI}^{(1)}(\hat{T}) = 0$	349
A6.2	Methods Based on $D_{JI}^{(2)}(\hat{T}) = 0$	353
A6.3	Methods Based on the Simultaneous Solution of $G_{JI}(\hat{T})=0$ and $g_{JI}(\hat{T})=0$	356
A6.4	Methods Based on $D_{JI}^{(3)}(\hat{T}) = 0$	358
Appendix 7	Additional Perturbation Series -- Orthonormal Basis	359
Appendix 8	Non-relativistic Approximation of the Dirac Hamiltonian	370
Appendix 9	Additional Perturbation Series -- Non-orthonormal Basis	377
Appendix 10	Self-consistent Perturbation Theory When $F^{(0)}$ is not Block Diagonal	382

Page

Appendix 11 Minimization Algorithms 385

A11.1 Method of Conjugate Gradients 385

A11.2 The Newton-Raphson Method 387

Appendix 12 Derivatives With Respect to Real and Imaginary Parts of f. 390

Appendix 13 Covariant and Contravariant Representations -- the General Case 391

List of Tables

Page

Table 5.1 Linear Convergence Rates of the Algorithms in Selected Calculations 125

Table 5.2 Rates of Convergence 146

Table 5.3 Rates of Convergence 149

Table 5.4 Rates of Convergence 155

Table 6.1 $D^{(n)}(f)$ 169

Table 6.2 $\hat{H}_A^{(n)}(f)$ 170

Table 6.3 $G_A^{(n)}$ 171

Table 6.4 $g_A^{(n)}$ 171

Table 6.5 $\tilde{H}_A^{(n)}$ 172

Table 6.6 $f^{(n)}$ (A-states degenerate) 174

Table 6.7 $f^{(n)}$ (A-states degenerate) 174

Table 6.8 $\hat{H}_A^{(n)}$ -- Reduced Formulas (A-states degen.) 175

Table 6.9 $\hat{H}_A^{(n)}$ (A-states degenerate) 176

Table 6.10 $G_A^{(n)}$ (A-states degenerate) 176

Table 6.11 $\tilde{H}_A^{(n)}$ (A-states degenerate) 177

Table 6.12 $f^{(n)}$ 180

Table 6.13 $\hat{H}_A^{(n)}$ 180

Table 6.14 $G_A^{(n)}$ 181

Table 6.15 $\tilde{H}_A^{(n)}$ 182

Table 6.16 $D^{(n)}(f)$ -- Non-orthonormal Basis 198

Table 6.17 $\hat{H}_A^{(n)}(f)$ -- Non-orthonormal Basis 199

Table 6.18 $\tilde{H}_A^{(n)}$ -- Non-orthonormal Basis 199

Page

Table 6.19 $G_{BA}^{(n)}$ -- Non-orthonormal Basis 200

Table 6.20 $g_{BA}^{(n)}$ -- Non-orthonormal Basis 200

Table 6.21 $f^{(n)}$ -- Non-orthonormal Basis (A-states degenerate) 201

Table 6.22 $\hat{H}_A^{(n)}$ -- Non-orthonormal Basis (A-states degenerate) 202

Table 6.23 $G_A^{(n)}$ -- Non-orthonormal Basis (A-states degenerate) 203

Table 6.24 $\tilde{H}_A^{(n)}$ -- Non-orthonormal Basis (A-states degenerate) 203

Table 6.25 $f^{(n)}$ -- Non-orthonormal Basis 207

Table 6.26 $\hat{H}_A^{(n)}$ -- Non-orthonormal Basis 207

Table 6.27 $G_A^{(n)}$ -- Non-orthonormal Basis 208

Table 6.28 $\tilde{H}_A^{(n)}$ -- Non-orthonormal Basis 208

Table 7.1 $(P_A')_{AA}^{(n)}$ -- Molecular Orbital Basis . . . 216

Table 7.2 $(P_A')_{BA}^{(n)}$ -- Molecular Orbital Basis . . . 217

Table 7.3 $(P_A')_{BB}^{(n)}$ -- Molecular Orbital Basis . . . 218

Table 7.4 $E^{(n)}$ -- Molecular Orbital Basis 219

Table 7.5 $E^{(n)}$ -- Molecular Orbital Basis 220

Table 7.6 $(P_{AO})^{(i)}$ for A_6 System ($H_{11}^{(0)} = 0$) 227

Table 7.7 $(P_{AO})^{(i)}$ for an A_4B_4 System ($H_{11}^{(0)} = -1$) . 228

Table 7.8 $(P_{AO})^{(i)}$ for an A_4B_4 System ($H_{11}^{(0)} = 0$) . . 228

Table 7.9 $(P_A')_{AA}^{(n)}$ -- Non-orthonormal Basis 237

Table 7.10 $(P_A')_{BA}^{(n)}$ -- Non-orthonormal Basis 237

	Page
Table 7.11 $(P_A')_{BB}^{(n)}$ -- Non-orthonormal Basis	238
Table 7.12 $E^{(n)}$ -- Non-orthonormal Basis	238
Table 7.13 $E^{(n)}$ -- Non-orthonormal Basis	239
Table 8.1 Closed Shell Case -- Test Calculations .	270
Table 8.2 Details of Direct Minimization Calculations, CN Molecule (r = 2.0 a.u.) . .	278
Table 8.3 Details of Direct Minimization Calculations, CN Molecule (r = 2.2 a.u.) . .	283
Table A7.1 $g_A^{\frac{1}{2}(n)}$	363
Table A7.2 $g_A^{-\frac{1}{2}(n)}$	364
Table A7.3 $g_A^{-1(n)}$	365
Table A7.4 $g_A^{\frac{1}{2}(n)}$	365
Table A7.5 $g_A^{-\frac{1}{2}(n)}$	366
Table A7.6 $g_A^{-1(n)}$	366
Table A7.7 $\tilde{H}_A^{(n)}$ in Terms of the $g_A^{(n)}$ and $\hat{H}_A^{(n)}$ . . .	367
Table A7.8 $\tilde{H}_A^{(n)}$ in Terms of the $g_A^{(n)}$ and $G_A^{(n)}$ . . .	368
Table A7.9 $G_A^{(n)}$	369
Table A8.1 Pauli Hamiltonian (adapted from DeVries (1970))	374
Table A8.2 g_A -- Non-relativistic Approximation .	374
Table A8.3 $g_A^{\frac{1}{2}}$ -- Non-relativistic Approximation .	375
Table A8.4 $g_A^{-\frac{1}{2}}$ -- Non-relativistic Approximation .	375
Table A8.5 Eriksen Hamiltonian (adapted from DeVries (1970))	376

Page

Table A8.6 Transformation Connecting H_{Pauli} and H_{Er} (adapted from DeVries (1970)) ... 376

Table A9.1 $g_A^{(n)}$ -- Non-orthonormal Basis 378

Table A9.2 $g_A^{\frac{1}{2}(n)}$ -- Non-orthonormal Basis 379

Table A9.3 $g_A^{-\frac{1}{2}(n)}$ -- Non-orthonormal Basis 379

Table A9.4 $g_A^{-1(n)}$ -- Non-orthonormal Basis 380

Table A9.5 $\tilde{H}_A^{(n)}$ -- Non-orthonormal Basis 380

Table A9.6 $\tilde{H}_A^{(n)}$ -- Non-orthonormal Basis 381

List Of Figures

		Page
Figure 5.1	Algorithm SDNRS	153
Figure 5.2	Algorithm SDNRS	154
Figure 7.1	P_{11} vs. H_{ii} for the A_6 System	229
Figure 7.2	P_{12} vs. H_{ii} for the A_6 System	230
Figure 7.3	Variation of P_{11} with H_{11} for the A_4B_4 System, $H_{11}^{(0)} = -1.0$	231
Figure 7.4	P_{21} vs H_{11} for the A_4B_4 System, $H_{11}^{(0)} = -1.0$	232
Figure 7.5	Variation of P_{11} with H_{11} for the A_4B_4 System, $H_{11}^{(0)} = 0.0$	233
Figure 7.6	Variation of P_{21} with H_{11} for the A_4B_4 System, $H_{11}^{(0)} = 0.0$	234
Figure 8.1	Total electronic energy as a function of iteration number for the CN molecule, bond length = 2.0 a.u.	279
Figure 8.2	Total electronic energy as a function of iteration number for the CN molecule, bond length = 2.0 a.u.	280
Figure 8.3	Total electronic energy as a function of iteration number for the CN molecule, bond length = 2.0 a.u.	281
Figure 8.4	Total electronic energy as a function of iteration number for the CN molecule, bond length = 2.2 a.u.	284

Page

Figure 8.5 Total electronic energy as a function of iteration number for the CN molecule, bond length = 2.2 a.u. 285

Figure 8.6 Total electronic energy as a function of iteration number for the CN molecule, bond length = 2.2 a.u. 286

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratitude and sincere thanks to Dr. John A. R. Coope for his guidance and many helpful suggestions during my time as a graduate student at U. B. C., and especially during the preparation of this thesis. His accessibility and willingness to become involved in my research, and to demonstrate how to communicate the results of that work, have made this a rewarding and productive time.

I would like to thank my wife, Marlene, not only for transcribing the figures and hand drawn tables in this thesis, but also for her support, and her patience, tolerance, and understanding of a husband often obsessed with writing a Ph. D. thesis. My parents also deserve much credit for their support and wise counsel over the many years of my education.

I would like to thank the National Research Council of Canada, the H. R. MacMillan Foundation, and the University of British Columbia, for financial support, without which these studies would not have been possible.

I would also like to acknowledge the contribution of the U. B. C. Computing Center to the more practical aspects of this work. Their extensive program library, and their extensive and powerful hardware facilities have made this part of the work far less painful than could be expected at many other institutions.

CHAPTER 1

EIGENVALUE INDEPENDENT PARTITIONING,

AN INTRODUCTION

"The average Ph. D. thesis is nothing but a transference of bones from one graveyard to another."
(J. Frank Dobie, A Texan in England, 1945)

Matrix partitioning is a well established technique in linear algebra, and such techniques have been found to be very useful in quantum chemistry. In a series of papers, Löwdin (1968, and references cited therein) has demonstrated the power and generality of a one-dimensional partitioning formalism, which contains, as special cases, many conventional methods used in quantum mechanical calculations. Through the partitioning of the basis space into two subspaces -- a one-dimensional space spanned by a chosen reference function, and the complementary n-1 dimensional space -- he obtains an expression for the eigenvalues, ϵ_a, of the matrix H as

$$\begin{aligned} \epsilon_a &= \tilde{H}_{aa}(\epsilon_a) \\ &= H_{aa} + H_{ab}(\epsilon_a 1_b - H_{bb})^{-1} H_{ba}, \end{aligned} \tag{1.1}$$

where $\tilde{H}_{aa}$ is a function not only of the elements of H, but also, of ϵ_a itself. Further development of the formalism leads to a variety of perturbation formulas (including, among others, the Rayleigh-Schrodinger and Brillouin-Wigner types), iterative methods for determining a single eigenvalue, formulas for upper and lower bounds to eigenvalues, and many other useful results.

The function $\tilde{H}_{aa}(\epsilon_a)$ in eq. (1.1) can be regarded as a one-dimensional effective operator which depends implicitly on the eigenvalues ϵ_a of H. A number of attempts have been made to construct effective operators without implicit eigenvalues

(see Klein (1974) and references cited therein), one of which is the eigenvalue independent partitioning of Coope (1970), which has some similarities to a non-canonical approach to the construction of effective operators, in elementary particle theory, first formulated by Okubo (1954). This thesis is primarily concerned with the development of this partitioning formalism, and its application in quantum mechanical calculations. The basic theory is described in considerable detail in chapters 2 - 4.

In the simplest (2 x 2) case, the basis space is partitioned into two subspaces -- an n_A-dimensional subpace and the complementary $n-n_A = n_B$ dimensional subspace, where $1 \leqslant n_A \leqslant n-1$ -- but now, the fundamental quantity is taken to be a mapping, f, relating the parts of the eigenvectors lying in these two subspaces. It is possible to define a variety of n_A-dimensional (and also, n_B-dimensional) effective operators in terms of this mapping. The set of eigenvalues of these effective operators form a subset of the eigenvalues of the matrix H, but the effective operators themselves no longer depend explicitly or implicitly on these eigenvalues. Also, the corresponding eigenvectors of the full matrix H are obtained straightforwardly from those of the effective operators using the mapping f.

Löwdin and Goscinski (1971) are quite correct in pointing out that implicitness of some sort is unavoidable in a partitioning formalism, and that this eigenvalue independent partitioning formalism could be described, in a particular sense,

as an eigenvector implicit partitioning. This implicitness is basically a result of the fact that the eigenvalues (and through them, the eigenvectors) of a matrix are nonlinear functions of the elements of the matrix. As indicated by Coope (1970), the one-dimensional partitioning formalism of Löwdin can be obtained as a special case of this eigenvalue independent partitioning formalism when $n_A = 1$ (as is, in fact, also demonstrated, but not emphasized, by Löwdin and Goscinski (1971)).

The adoption of this more general point of view, in which the partitioning theory is formulated in terms of a mapping between the partitioned spaces rather than in terms of the eigenvalues and eigenvectors of the matrix, leads to new and important areas of application. In particular, it is especially suitable when groups of eigenvalues or eigenvectors are to be treated simultaneously. In chapter 2, it is shown that the mapping f can be used to define projections onto whole eigenspaces of H. The condition defining f can be formulated variationally, and is also seen to be related to measures of errors in such eigenprojections. It is also shown that f transforms nonlinearly under a linear transformation of the basis vectors, and that this has important practical implications.

The simplest (2 x 2) case of the eigenvalue independent partitioning described above is straightforwardly generalized to partitioning of the basis space (and eigenvector space) into m, $(2 \leqslant m \leqslant n)$, subspaces, as is demonstrated in chapter 4.

There are two main areas of application of this partitioning formalism. One of them is in the construction of effective operators in n_A-dimensional spaces, with $n_A \geq 1$. For eigenvalues which are well separated from all others, one-dimensional partitioning formalisms, as in eq. (1.1), are useful, but when degeneracy or near degeneracy occurs, these formulas become ill-conditioned. Traditionally, multi-dimensional effective operators have been constructed using a canonical procedure,

$$\tilde{H} = U^{\dagger} H U, \tag{1.2}$$

requiring the calculation of a unitary transformation, U, which uncouples the desired eigenspace of the operator from the rest of the eigenvector space (for example, Van Vleck perturbation theory, (Van Vleck, 1929), also, see Tani (1954) and Klein (1974)). The unitarity of U is commonly ensured by writing it as

$$U = e^{iS}, \tag{1.3}$$

where S is a hermitian operator. Thus, in obtaining the desired uncoupled operator $\tilde{H}$, one must determine the exponential operator, e^{iS}. This can be done straightforwardly using a perturbation formalism when that is appropriate, but it is very difficult, in general, to calculate S exactly otherwise. On the other hand, the mapping f, in this partitioning formalism, is defined by a much simpler, though still nonlinear, equation, which can not only be solved using a perturbation expansion, when appropriate, but can also be solved iteratively in a very

straightforward manner to obtain f to any desired level of accuracy.

Methods for the iterative determination of f, and its generalization in a multi-partitioning formalism, are given in chapter 5, and accompanying appendices. The particular application to the calculation of a small number of eigenvalues and eigenvectors of a large hermitian matrix is considered in detail, and test calculations demonstrate the usefulness of this new approach to the problem.

Because of the simple algebraic form of the condition defining f, compared to those defining the operator S in eq. (1.3), perturbation formulas for f and for effective operators defined in terms of f, are obtained straightforwardly for arbitrary order, unlike the involved step by step procedure required in the canonical approach. Certain of the more useful series are developed in chapter 6. Two examples are included to demonstrate the scope and ease of use of these formulas. It is shown that a formal uncoupling of the four-component Dirac equation, to obtain a two-component relativistic wave equation for electrons, is obtained by a particularly simple application of the basic formulas derived in the early part of chapter 6. Also, a nuclear spin hamiltonian for the strong field case is derived to second order. In all cases, the presence of degeneracy in zero order is of no concern as long as all degenerate or nearly degenerate levels are treated at the same time.

Another major application of this eigenvalue independent partitioning formalism is in the use of the mapping operators to describe projections onto particular eigenspaces. As shown in chapter 2, projections onto eigenspaces can be written in terms of f in a form which is automatically idempotent and self-adjoint for any value of f. Because the elements of f are required to satisfy only a single simple defining condition, perturbation formulas to arbitrarily high order are again obtained straightforwardly. In chapter 7, perturbation formulas for such projections are developed with reference to molecular orbital theory. In particular, perturbation formulas for the density matrix in Huckel, extended Huckel, and closed shell self-consistent field theory are produced.

The density matrix (the projection onto the occupied orbitals) in closed shell self-consistent field theory can be written solely in terms of the operator f corresponding to a partitioning of the eigenvectors of the Fock operator, F, into two sets, consisting, respectively, of the occupied and the unoccupied orbitals, and thus, the total electronic energy is completely specified by f. The application of this partitioning formalism in self-consistent field theory represents a generalization of the simple matrix partitioning described above, in that the operator, F, to be brought to block diagonal form, itself depends on the partitioning operator f through its dependence on the density matrix R. Since the matrix elements of f are not constrained in any way and do not contain any

redundancy (see section 8.2), they are a particularly suitable set of variables in terms of which to determine the stationary values of the energy directly. The derivatives of the Hartree-Fock energy with respect to these variables are given very compactly using the columns of the density matrix and its complement. This formalism is extended straightforwardly to the general multi-shell single determinant case using the multi-partitioning formalism described in chapter 4. Some numerical calculations in the closed shell and unrestricted Hartree-Fock cases are described in chapter 8, and they indicate that refinements involving the use of scaled variables and the adoption of bases which nearly diagonalize the Fock matrices, result in practical procedures which are superior to the Roothaan procedure and to other currently available direct minimization self-consistent field procedures.

CHAPTER 2

2 x 2 PARTITIONING THEORY

"The White Rabbit put on his spectacles. 'Where shall I begin, please your Majesty?' he asked.
'Begin at the beginning', the King said gravely, 'and go on till you come to the end: then stop.' "
(Alice's Adventures in Wonderland, Lewis Carroll)

2.1 Basic Theory

2.1.a The f-operator

Consider the matrix eigenvalue equation,

$$H X = X \mathcal{E} , \tag{2.1a}$$

$$X^{\dagger} X = 1, \tag{2.1b}$$

where H is an n x n hermitian matrix, X is the n x n unitary matrix whose columns are the orthonormal eigenvectors of H, and $\mathcal{E}$ is the n x n diagonal matrix whose elements are the corresponding real eigenvalues of H. If the n-dimensional basis set being used is partitioned into two subsets spanning spaces S_A and S_B of dimensions n_A and n_B, respectively, and the eigenvectors of H are similarly partitioned into two sets $X^{(A)}$ and $X^{(B)}$, spanning spaces S_A' and S_B', also of dimensions n_A and n_B, respectively, then, the matrix, X, above can be written in the block form,

$$X = [X^{(A)} \; X^{(B)}] = \begin{bmatrix} X_{AA} & X_{AB} \\ X_{BA} & X_{BB} \end{bmatrix} = \begin{bmatrix} 1_A & h \\ f & 1_B \end{bmatrix} \begin{bmatrix} X_{AA} & 0 \\ 0 & X_{BB} \end{bmatrix}$$

$$= \hat{T} \hat{X}. \tag{2.2}$$

Formally, one has,

$$f = X_{BA} X_{AA}^{-1} , \tag{2.3a}$$

and,

$$h = X_{AB} X_{BB}^{-1} . \tag{2.3b}$$

The operator f maps the part of an eigenvector $x_r^{(A)}$ lying in S_A into the part lying in S_B. It can be considered as a

generalization of the operator f(E), defined by Löwdin (1962), in connection with a partitioning formalism with $n_A = 1$ (that is, with the space S_A one-dimensional). The function of the space S_A here is analogous to that of the so-called reference function in one-dimensional partitioning formalisms. Similarly, the operator h maps the part of an eigenvector $x_\sigma^{(B)}$ lying in S_B into the part lying in S_A. From eqs. (2.3), it is seen that f exists if the matrix block X_{AA} is non-singular, while h exists if the matrix block X_{BB} is non-singular. Since the eigenvectors of a hermitian matrix are orthogonal if the basis functions are linearly independent, the above conditions on X_{AA} and X_{BB} are satisfied simultaneously for at least one partitioning of the basis functions.

The orthonormality condition, (2.1b), on X can be used to show that

$$h = -f^\dagger . \tag{2.4}$$

Thus, in the simple 2 x 2 case,

$$\hat{T} = \begin{bmatrix} 1_A & -f^\dagger \\ f & 1_B \end{bmatrix} . \tag{2.5}$$

The operator f is the fundamental quantity in this 2 x 2 partitioning formalism. Because of (2.4), it completely determines projection operators, P_A' and P_B', onto the two eigenspaces S_A' and S_B'. One has

$$P_A' = X^{(A)} X^{(A)\dagger} = \begin{bmatrix} 1_A \\ f \end{bmatrix} (X_{AA}\, X_{AA}^\dagger) \begin{bmatrix} 1_A & f^\dagger \end{bmatrix} .$$

However, from the orthonormality condition, (2.1b), on X, one can write,

$$X^{\dagger}X = \hat{X}^{\dagger}\, g\, \hat{X} = 1_n \,, \tag{2.6}$$

where,

$$g = \hat{T}^{\dagger}\hat{T} = \begin{bmatrix} 1_A + f^{\dagger}f & 0 \\ 0 & 1_B + f\, f^{\dagger} \end{bmatrix} = \begin{bmatrix} g_A & 0 \\ 0 & g_B \end{bmatrix}. \tag{2.7}$$

The matrices g_A and g_B define metrics, with respect to which the truncated eigenvectors in X_{AA} and X_{BB}, are orthonormal. That is,

$$X_{AA}^{\dagger}\, g_A\, X_{AA} = 1_A \,, \tag{2.8a}$$

and,

$$X_{BB}^{\dagger}\, g_B\, X_{BB} = 1_B \,. \tag{2.8b}$$

These truncated eigenvectors are not orthonormal with respect to unity unless f = 0. Since $\hat{X}$ is invertible, from (2.6) or (2.8), one has,

$$g_A = (X_{AA}\, X_{AA}^{\dagger})^{-1} = (1_A + f^{\dagger}f), \tag{2.9a}$$

and,

$$g_B = (X_{BB}\, X_{BB}^{\dagger})^{-1} = (1_B + ff^{\dagger}). \tag{2.9b}$$

Using (2.9a), the projection $P_A^{\cdot}$ can now be written,

$$P_A^{\cdot} = \begin{bmatrix} 1_A \\ f \end{bmatrix} g_A^{-1} \begin{bmatrix} 1_A & f^{\dagger} \end{bmatrix} = \begin{bmatrix} g_A^{-1} & g_A^{-1}\, f^{\dagger} \\ f\, g_A^{-1} & f\, g_A^{-1}\, f^{\dagger} \end{bmatrix} \tag{2.10}$$

In a similar manner, the projection $P_B^{\cdot}$ onto the eigenspace $S_B^{\cdot}$, can be written solely in terms of f as,

$$P_B' = \begin{bmatrix} -f^\dagger \\ 1_B \end{bmatrix} g_B^{-1} \begin{bmatrix} -f & 1_B \end{bmatrix} = \begin{bmatrix} f^\dagger g_B^{-1} f & -f^\dagger g_B^{-1} \\ -g_B^{-1} f & g_B^{-1} \end{bmatrix}. \qquad (2.11)$$

It is easily verified that $P_A' + P_B' = 1$. The operators P_A' and P_B' above are manifestly self-adjoint. Furthermore, using the definitions of g_A and g_B in terms of f, given in eqs. (2.9), these matrices can be shown to be idempotent by direct matrix multiplication. Finally,

$$\begin{aligned} \text{tr } P_A' &= \text{tr } g_A^{-1} + \text{tr } f g_A^{-1} f^\dagger = \text{tr } g_A^{-1}(1_A + f^\dagger f) \\ &= \text{tr } 1_A = n_A, \end{aligned} \qquad (2.12a)$$

and similarly,

$$\text{tr } P_B' = \text{tr } 1_B = n_B, \qquad (2.12b)$$

where the cyclic property of the trace has been used. Thus, for arbitrary f, the operators P_A' and P_B' satisfy all the criteria necessary to be orthogonal projection operators. The usefulness of the formulation in terms of the operator f is essentially that, while the operators P_A' and P_B' must satisfy a complicated set of general constraints in order to be projections onto (any) spaces S_A' and S_B', the partitioning operator f is not constrained in any way.

The eigenprojection P_A' is completely specified by the $n_A n$ complex components of the vectors $x_r^{(A)}$, $(r = 1, \ldots, n_A)$, spanning S_A'. However, the space S_A' is also spanned by any other set of n_A vectors $y_r^{(A)}$ related to the $x_r^{(A)}$ by a non-singular $n_A \times n_A$ linear transformation. This transformation

corresponds merely to a change of basis in S_A'. Therefore, there are n_A^2 arbitrary, or redundant, complex parameters present in the specification of S_A' using $X^{(A)}$. Thus only $n_A(n - n_A) = n_A n_B$ complex parameters are necessary to specify the eigenspace S_A'. But this is exactly the number of degrees of freedom (or matrix elements) in f. Thus the operator f represents the minimum amount of information necessary to specify a projection onto the eigenspace S_A' (and therefore, also onto S_B', which is the complement of S_A') of H. This partitioning formalism is therefore particularly useful in situations in which only eigenspaces have significance, rather than specific eigenvectors.

2.1.b The Defining Condition For f

The matrices f and h, defined in the previous subsection, can be obtained by diagonalizing H to get its eigenvectors, X, and then applying the formulas (2.3) directly. However, it is possible to formulate a system of equations for f and h, which do not require knowledge of the eigenvectors of H.

The eigenvalue equation (2.1a) is rewritten as

$$H \hat{T} = \hat{T} \hat{H}, \tag{2.13}$$

where

$$\hat{H} = \hat{X} \not{f} \hat{X}^{-1} = \begin{bmatrix} \hat{H}_A & 0 \\ 0 & \hat{H}_B \end{bmatrix}, \tag{2.14}$$

is to be block diagonal. The diagonal blocks of eq. (2.13)

give expressions for $\hat{H}_A$ and $\hat{H}_B$ in terms of f, h, and H,

$$\hat{H}_A = H_{AA} + H_{AB}f, \tag{2.15a}$$

and,

$$\hat{H}_B = H_{BA}h + H_{BB}. \tag{2.15b}$$

If these expressions are substituted back into eq. (2.13), the two off-diagonal blocks become nonlinear matrix equations,

$$\begin{aligned} D(f) &= H_{BA} + H_{BB}f - f\,\hat{H}_A(f) \\ &= H_{BA} + H_{BB}f - f\,H_{AA} - f\,H_{AB}f = 0, \end{aligned} \tag{2.16}$$

and,

$$\begin{aligned} D'(h) &= H_{AA}h + H_{AB} - h\,\hat{H}_B(h) \\ &= H_{AB} + H_{AA}h - h\,H_{BB} - h\,H_{BA}h = 0. \end{aligned} \tag{2.17}$$

Equations (2.16) and (2.17) are both systems of $n_A n_B$ simultaneous nonlinear equations, the first for the matrix elements of f, and the second for the matrix elements of h. It is noteworthy that the two systems are not coupled, and thus can be solved independently.

Of course, in this case, it is not necessary to solve both (2.16) and (2.17), because if one or the other has been solved, the solution of the remaining system is given by eq. (2.4). In fact, it can easily be seen that $D'^{\dagger}$ is of the same form in $-h^{\dagger}$, as D(f) is in f, implying eq. (2.4) without explicitly making use of orthogonality (the hermiticity of H is used, and this, of course, implies the orthogonality condition anyway).

In the 2 x 2 partitioning formalism, eq. (2.16) is the fundamental equation determining the operator f, if an orthonormal basis is used. A number of efficient iterative techniques for the exact solution of (2.16) will be detailed later. The quantity D(f) is closely related to other more commonly used quantities in the determination of eigenprojections. In particular, D(f) will be seen to be related in several ways to the error in an eigenprojection.

2.1.c Rederivation From a Projection Point of View

An alternative approach to this partitioning formalism can be made via the projection operators themselves. The objective is to determine the eigenprojection P_A' onto a space S_A' spanned by n_A eigenvectors of H, in terms of some minimal set of variables which number $n_A n_B$, as shown previously. It is useful to examine this approach in some detail, not only because it provides a different point of view, but also, because the projections themselves are manifestly basis independent.

The conditions that P_A' be an eigenprojection of H are that P_A' commute with H,

$$[\, H \,,\, P_A' \,] = 0, \tag{2.18}$$

and that P_A' be a projection operator, that is,

$$P_A'^2 = P_A', \qquad P_A'^\dagger = P_A', \qquad \mathrm{tr}\, P_A' = n_A. \tag{2.19}$$

It is convenient to define a partitioning of the basis

functions into two sets, spanning spaces S_A and S_B of dimensions n_A and n_B, respectively. Projections onto the spaces S_A and S_B are given by,

$$P_A = \begin{bmatrix} 1_A & 0 \\ 0 & 0 \end{bmatrix}, \qquad P_B = \begin{bmatrix} 0 & 0 \\ 0 & 1_B \end{bmatrix}. \tag{2.20}$$

This partitioning of the basis functions implies that the projection P_A' can be written in block form,

$$P_A' = \begin{bmatrix} P_{AA}' & P_{AB}' \\ P_{BA}' & P_{BB}' \end{bmatrix}, \tag{2.21a}$$

where,

$$\begin{aligned} P_{AA}' &= P_A P_A' P_A , \\ P_{BA}' &= P_B P_A' P_A , \\ P_{AB}' &= P_A P_A' P_B , \end{aligned} \tag{2.21b}$$

and

$$P_{BB}' = P_B P_A' P_B .$$

In terms of the partitioned matrix, (2.21), the idempotency condition, $P_A'^2 = P_A'$, is equivalent to the three block equations,

$$\begin{aligned} P_{AA}' - P_{AA}'^2 - P_{AB}' P_{BA}' &= 0, \\ P_{BA}' - P_{BA}' P_{AA}' - P_{BB}' P_{BA}' &= 0, \end{aligned} \tag{2.22}$$

and

$$P_{BB}' - P_{BB}'^2 - P_{BA}' P_{AB}' = 0,$$

the remaining block equation being just the adjoint of the second one. Since there are only $n_A n_B$ independent variables in P_A', it is possible, in principle, to express P_{AA}' and P_{BB}'

in terms of P'_{BA}. However, the equations are nonlinear, and while formal general solutions can be written down,

$$P'_{AA} = \tfrac{1}{2}[1 \pm (1 - 4P'^{\dagger}_{BA} P'_{BA})^{\frac{1}{2}}] , \qquad (2.23a)$$

and

$$P'_{BB} = \tfrac{1}{2}[1 \pm (1 - 4P'_{BA} P'^{\dagger}_{BA})^{\frac{1}{2}}] , \qquad (2.23b)$$

they are seen to contain the ambiguity in the square root, and are generally difficult to evaluate.

A more useful result is reached by a different route. The matrix P'_A is of rank n_A, because the projection operator onto an n_A-dimensional space has precisely n_A non-zero eigenvalues, corresponding to the n_A eigenvectors of P'_A which span the image space S'_A. This means that there is at least one $n_A \times n_A$ submatrix of P'_A which is non-singular. It will be assumed that the partitioning of the basis set, (2.20), is carried out so that P'_{AA} is such a submatrix, that is, $\det(P'_{AA}) \neq 0$. With this assumption, the first equation of (2.22) can be rewritten as

$$P'_{AA} = P'_{AA} (1_A + P'^{-1}_{AA} P'^{\dagger}_{BA} P'_{BA} P'^{-1}_{AA}) P'_{AA} . \qquad (2.24)$$

The quantity inside the brackets in (2.24) will be greatly simplified if P'_{BA} is written as some factor times P'_{AA}, that is, if

$$P'_{BA} = f P'_{AA} , \qquad P'_{AB} = P'_{AA} f^{\dagger} , \qquad (2.25a)$$

where f is an $n_B \times n_A$ matrix, and thus represents a suitable quantity, in terms of which the matrix P'_A could be expressed. The existence of f is assured by the invertibility of P'_{AA}.

$$f = P'_{BA}\, P'^{-1}_{AA} \,. \tag{2.25b}$$

Now, (2.24) yields

$$P'_{AA} = (1_A + f^\dagger f)^{-1}. \tag{2.26}$$

From (2.25),

$$P'_{BA} = f\,(1_A + f^\dagger f)^{-1}. \tag{2.27}$$

Finally, substituting (2.25) into the second of eqs. (2.22), and multiplying from the right by $f^\dagger(ff^\dagger)^{-1}$, yields,

$$P'_{BB} = f\,(1_A + f^\dagger f)^{-1} f^\dagger. \tag{2.28}$$

Equation (2.18) can now be used to derive an equation defining the operator f. Expansion of the commutator again yields three unique block equations,

$$\begin{aligned}
(EQ)_{AA} &= (1_A + f^\dagger f)^{-1}(H_{AA} + f^\dagger H_{AB}) \\
&\quad - (H_{AA} + H_{AB}f)(1_A + f^\dagger f)^{-1} = 0, \\
(EQ)_{BA} &= (1_B + ff^\dagger)^{-1}(fH_{AA} + ff^\dagger H_{BA}) \\
&\quad -(H_{BA} + H_{BB}f)(1_A + f^\dagger f)^{-1} = 0,
\end{aligned} \tag{2.29}$$

and,

$$\begin{aligned}
(EQ)_{BB} &= (1_B + ff^\dagger)^{-1}(fH_{AB} + ff^\dagger H_{BB}) \\
&\quad - (H_{BA}f^\dagger + H_{BB}ff^\dagger)(1_B + ff^\dagger)^{-1} = 0.
\end{aligned}$$

Here, use has been made of the relations,

$$\begin{aligned}
f(1_A + f^\dagger f)^{-1} &= (1_B + ff^\dagger)^{-1} f, \\
\text{and} \qquad & \qquad (2.30) \\
(1_A + f^\dagger f)^{-1} f^\dagger &= f^\dagger (1_B + ff^\dagger)^{-1},
\end{aligned}$$

to move all of the inverse operators to the outside of each

term. It is then seen that

$$-f(1_A + f^\dagger f)\,(EQ)_{AA}\,(1_A + f^\dagger f) + (1_B + ff^\dagger)\,(EQ)_{BA}(1_A + f^\dagger f)$$
$$= H_{BA} + H_{BB}f - fH_{AA} - fH_{AB}f = D(f) = 0, \tag{2.31}$$

and also,

$$(1_B + ff^\dagger)\,(EQ)_{BA}(1_A + f^\dagger f) + (1_B + ff^\dagger)\,(EQ)_{BB}(1_B + ff^\dagger)f$$
$$= D(f) = 0,$$

where the quantity D(f) has been defined in eq. (2.16) That is, the operator f defined in eq. (2.16) is of the same size and satisfies the same defining equation as the partitioning operator f described in the previous two subsections. This result re-emphasizes the fact that this partitioning formalism is based on the idea of defining an eigenspace of a hermitian operator, rather than individual eigenvectors.

The 'pull-through' relations, (2.30), are used extensively in the 2 x 2 partitioning formalism. They are most usefully written as

$$f\,g_A^{-1} = g_B^{-1}\,f, \tag{2.32a}$$

and

$$g_A^{-1}\,f^\dagger = f^\dagger g_B^{-1}\,, \tag{2.32b}$$

in the notation established in the previous subsections.

2.1.d The Relationship Between $\hat{T}$ and the Eigenprojections-- Covariant and Contravariant Representations

The columns of the partitioning matrix $\hat{T}$, of eq. (2.2) or (2.5), can be regarded as a set of non-orthonormal basis vectors spanning the original n-dimensional basis space. These vectors will be denoted here by e_r, (r = 1, ..., n), that is

$$\hat{T} = [e_1, e_2, \ldots, e_n] = \begin{bmatrix} 1_A & -f^\dagger \\ f & 1_B \end{bmatrix}. \tag{2.33}$$

The metric defined by the scalar products of these vectors, $g_{rs} = e_r \cdot e_s$, is given by

$$\mathcal{g} = \hat{T}^\dagger \hat{T} = \begin{bmatrix} g_A & 0 \\ 0 & g_B \end{bmatrix}, \tag{2.34}$$

using the notation developed in eq. (2.7). Using the inverse metric, $\tilde{g} = g^{-1}$, a set of contragredient basis vectors e^r, (r = 1, ..., n), can be defined by

$$e^r = \sum_{s=1}^{n} g^{rs} e_s, \tag{2.35}$$

or

$$[e^1, e^2, \ldots, e^n] = \begin{bmatrix} g_A^{-1} & 0 \\ 0 & g_B^{-1} \end{bmatrix} \begin{bmatrix} 1_A & -f^\dagger \\ f & 1_B \end{bmatrix} = \begin{bmatrix} g_A^{-1} & -g_A^{-1} f^\dagger \\ g_B^{-1} f & g_B^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} g_A^{-1} & -g_A^{-1} f^\dagger \\ f g_A^{-1} & 1_B - f g_A^{-1} f^\dagger \end{bmatrix}. \tag{2.36}$$

On comparison with eq. (2.10), the first n_A of these vectors e^r can be identified as the first n_A columns of the projection P_A' onto S_A'. Similarly, the last n_B of the e^r are the last

n_B columns of $P'_B = (1 - P'_A)$, the projection onto the complementary subspace S'_B. Thus the two sets of n_A vectors

$$\underset{\sim}{e}_A = \begin{bmatrix} 1_A \\ f \end{bmatrix}, \quad \text{and} \quad \tilde{e}_A = \begin{bmatrix} (P'_A)_{AA} \\ (P'_A)_{BA} \end{bmatrix}, \tag{2.37}$$

are dual (contragredient), both spanning the eigenspace S'_A, while the two sets of n_B vectors,

$$\underset{\sim}{e}_B = \begin{bmatrix} -f^\dagger \\ 1_B \end{bmatrix}$$

and

$$\tilde{e}_B = \begin{bmatrix} -(P'_A)_{AB} \\ 1_B - (P'_A)_{BB} \end{bmatrix} = \begin{bmatrix} (P'_B)_{AB} \\ (P'_B)_{BB} \end{bmatrix}, \tag{2.38}$$

are also dual, both sets spanning the eigenspace S'_B.

From a different point of view, a metric $\tilde{\Delta}$ can be defined, with respect to which the e_r are orthonormal, namely,

$$e_r^\dagger \cdot \tilde{\Delta} \cdot e_s = \delta_{rs}, \quad (r,s = 1, \ldots, n), \tag{2.39}$$

That is,

$$\tilde{\Delta} = \tilde{e}\tilde{e}^\dagger.$$

Here

$$\tilde{\Delta} = \tilde{e}\tilde{e}^\dagger = \begin{bmatrix} g_A^{-1} & 0 \\ 0 & g_B^{-1} \end{bmatrix}, \tag{2.40}$$

is the same as $\underset{\sim}{g}^{-1}$ above. Similarly, the e^r are orthonormal with respect to $\tilde{\Delta}^{-1}$, which is the same numerically as $\underset{\sim}{g}$ above. It should be noted that $\tilde{\Delta}^{-1}$ and $\underset{\sim}{g}$, and $\tilde{\Delta}$ and $\underset{\sim}{g}^{-1}$ as denoted here are, in principle, quite different quantities. They happen to be numerically identical here because $\tilde{e}\tilde{e}^\dagger = \tilde{e}^\dagger\tilde{e}$

(and $\underline{e}\underline{e}^{\dagger} = \underline{e}^{\dagger}\underline{e}$). Such is not the case, however, if the original basis is non-orthonormal. These sets of contragredient vectors are very useful for writing a number of important relations, to be developed later, in a very compact manner.

2.1.e Variational Formulation of D(f) = 0

The expectation value of an operator with respect to one of its eigenvectors is stationary with respect to arbitrary small variations in that eigenvector. As a result, if P_A' is a projection onto the eigenspace S_A' of the operator H, then the expectation value of H over S_A', given by

$$E = \text{tr } P_A' H \,, \tag{2.41}$$

will be stationary with respect to arbitrary small variations in P_A'. That is

$$\begin{aligned} E(P_A' + \delta P_A') - E(P_A') &= \text{tr}[P_A'(f+\delta f) - P_A'(f)]H \\ &= \text{tr } \delta P_A' H + O(\delta^2), \end{aligned} \tag{2.42}$$

must vanish to first order in the infinitesimals. It is assumed here that H is independent of P_A' or f. From eqs. (2.10), to first order,

$$\begin{aligned} (\delta P_A')_{AA} &= -g_A^{-1}\delta g_A\, g_A^{-1} \,, \\ (\delta P_A')_{BA} &= -f g_A^{-1}\delta g_A\, g_A^{-1} + \delta f g_A^{-1} \,, \\ (\delta P_A')_{AB} &= -g_A^{-1}\delta g_A\, g_A^{-1}\, \underline{f}^{\dagger} + g_A^{-1}\delta f^{\dagger} \,, \end{aligned} \tag{2.43a}$$

and

$$(\delta P_A')_{BB} = -f g_A^{-1}\delta g_A\, g_A^{-1} f^{\dagger} + \delta f g_A^{-1} f^{\dagger} + f g_A^{-1}\delta f^{\dagger} \,,$$

where to first order,

$$\delta g_A = \delta f^\dagger f + f^\dagger \delta f. \tag{2.43b}$$

Substitution of (2.43a,b) into eq. (2.41), followed by use of the cyclic property of the trace, and the 'pull-through' relations (2.32) for f and $f^\dagger$, results in an expression of the form

$$\delta E = \mathrm{tr}\delta f^\dagger \bar{D} + \mathrm{tr}\delta f\, \bar{D}^\dagger + O(\delta^2), \tag{2.44}$$

where

$$\bar{D} = g_B^{-1}\, D(f)\, g_A^{-1}. \tag{2.45}$$

Because δf and $\delta f^\dagger$ are arbitrary variations in f and $f^\dagger$, the condition that E vanish in first order is that the matrix $\bar{D}$ vanish. The matrices g_A^{-1} and g_B^{-1} are positive definite, however, and thus $\bar{D}$ can vanish only if D(f) itself vanishes. Thus, the condition that the expectation value of H over the image space of the projection P_A' be stationary is equivalent to the condition D(f) = 0, eq. (2.16).

It is also interesting to note here that the quantity $\bar{D}$ in eq. (2.45) is the BA block of the hamiltonian H, in the basis of contragredient, non-orthonormal vectors $\tilde{e}$ of eq. (2.36). Thus one can write

$$\frac{\partial E}{\partial f^*_{\sigma r}} = \left([(1 - P_A')\, H\, P_A']_{BA}\right)_{\sigma r} \tag{2.46a}$$

$$= H_{e_B \sigma e_A r} \tag{2.46b}$$

and the rate of change of the expectation value E with an element $f_{\sigma r}$ of f is seen to be proportional to the corresponding element of the off-diagonal block of the hamiltonian, in this particular basis.

2.1.f Relation Between σ^2 and D(f) -- Eigenvalue Dispersion

In the study of matrix eigenvalue problems, it is useful to define the variance σ^2, which is a measure of the error in an approximate eigenvector, x, of a matrix H, given by

$$\sigma^2 = \frac{(Hx - \lambda x)^\dagger (Hx - \lambda x)}{x^\dagger x} . \qquad (2.47)$$

If the approximate eigenvalue, λ, is calculated as the Rayleigh quotient of H with respect to x,

$$\lambda = \frac{x^\dagger H x}{x^\dagger x} , \qquad (2.48)$$

then eq. (2.47) becomes

$$\sigma^2 = \frac{x^\dagger H^2 x}{x^\dagger x} - \left(\frac{x^\dagger H x}{x^\dagger x}\right)^2 = \langle H^2 \rangle_x - \langle H \rangle_x^2 , \qquad (2.49)$$

which is in the form of the usual definition of variance. In terms of the projection

$$P_x = x\, x^\dagger ,$$

onto the one-dimensional space spanned by the normalized vector x, eq. (2.49) can be written as

$$\sigma^2 = \mathrm{tr}\, H(1 - P_x)\, H\, P_x . \qquad (2.50)$$

Equation (2.50) suggests a generalization of the concept of the variance σ^2 to apply to projections P_A' onto a multi-dimensional space spanned by several approximate eigenvectors of H. Substitution of eq. (2.10) for P_A' into eq. (2.50), and use of (2.16), gives

$$\sigma^2 = \mathrm{tr}\, H(1-P_A')HP_A' = \mathrm{tr}\, g_B^{-1} D(f) g_A^{-1} D(f)^\dagger$$
$$= \| g_B^{-\frac{1}{2}} D(f) g_A^{-\frac{1}{2}} \|^2 , \qquad (2.51)$$

where $\|A\| = \left(\sum_{r,s} |A_{rs}|^2 \right)^{\frac{1}{2}}$ denotes the Hilbert-Schmidt norm of the matrix A. This may also be written in the form

$$\sigma^2 = -\tfrac{1}{2}\mathrm{tr}([\, H \,,\, P_A' \,]_-^2). \tag{2.52}$$

If P_A' is an exact eigenprojection of H, the variance σ^2 in (2.51) must vanish, because then $[\, H \,,\, P_A' \,]_- = 0$. Since g_A^{-1} and g_B^{-1} are positive definite matrices, σ^2 can vanish only if $D(f) = 0$. In this case, D(f) is seen to give a quantitative measure of the error in P_A', rather than merely a criterion for the presence or absence of error.

In terms of matrix elements, one has

$$\sigma^2 = \sum_{\rho,t} |\langle \phi_\rho | g_B^{-\frac{1}{2}} D(f) g_A^{-\frac{1}{2}} | \phi_t \rangle|^2 = \sum_{\rho,t} |\langle \phi_\rho^o | H | \phi_t^o \rangle|^2, \tag{2.53}$$

where ϕ_ρ, $(\rho = 1, \ldots, n_B)$, and ϕ_t, $(t = 1, \ldots, n_A)$, are basis elements in the subspaces S_B and S_A, respectively. The ϕ_ρ^o, $(\rho = 1, \ldots, n_B)$, and the ϕ_t^o, $(t = 1, \ldots, n_A)$, are the orthogonalized transformed basis vectors,

$$\{\phi_t^o\} = \begin{bmatrix} 1_A \\ f \end{bmatrix} g_A^{-\frac{1}{2}} \quad , \qquad \{\phi_\rho^o\} = \begin{bmatrix} -f \\ 1_B \end{bmatrix} g_B^{-\frac{1}{2}} \quad , \tag{2.54}$$

in the basis of the ϕ_ρ and ϕ_t, above. Thus σ^2 is seen to be a measure of the smallness of the elements of the off-diagonal block of H in this basis. Using the closure relation $\sum_\rho |\rho\rangle\langle\rho| = 1 - \sum_t |t\rangle\langle t|$, eq. (2.53) can be rewritten as

$$\sigma^2 = \sum_{t \in S_A} \langle \phi_t^o | H^2 | \phi_t^o \rangle - \sum_{t,s \in S_A} |\langle \phi_t^o | H | \phi_s^o \rangle|^2, \tag{2.55}$$

where ϕ_t^o, ϕ_s^o, in these summations run over eigenvectors in the

space S_A' only. On transforming these vectors to a new set ψ_r, ($r = 1, \ldots, n_A$), which diagonalizes H in S_A', eq. (2.55) becomes

$$\sigma^2 = \sum_{n=1}^{n_A} [\langle \psi_n | H^2 | \psi_n \rangle - \langle \psi_n | H | \psi_n \rangle^2]$$
$$= \sum_{n=1}^{n_A} \sigma_n^2 \quad . \tag{2.56}$$

If f is an exact solution of (2.16), uncoupling the parts of the ϕ_t^o, ($t = 1, \ldots, n_A$), in S_A, and the ϕ_ρ^o, ($\rho = 1, \ldots, n_B$), in S_B, exactly, then the ψ_n in (2.56) are exact eigenvectors of H, and each σ_n is identically zero. If f is not exact, then the ψ_n will be only approximate eigenvectors of H, and σ_n^2 is the variance of H with respect to the single approximate eigenvector ψ_n. Thus σ^2 is the sum over these individual variances, and is useful not only as a quantitative measure of the accuracy of f, but also as an upper bound to the individual σ_n^2.

2.1.g Transformation of f Under a Change of Basis

The quantity f defined by eq. (2.2) is clearly dependent upon the basis set being used. Because of eq. (2.3), it does not transform linearly under a linear transformation of the basis vectors.

Consider the linear transformation,

$$\phi_\nu' = \sum_{\mu=1}^{n} \phi_\mu V_{\mu\nu}^\dagger \quad , \tag{2.57}$$

of the basis vectors $\{\phi_\mu\}$, so that the eigenvectors of H, referred to the new basis $\{\phi'_\nu\}$ have coefficients,

$$X' = V\,X. \tag{2.58}$$

In the new basis, partitionings of eigenvectors and basis vectors similar to those described in section 2.1.a can be carried out, yielding,

$$X' = \begin{bmatrix} X'_{AA} & X'_{AB} \\ X'_{BA} & X'_{BB} \end{bmatrix} = \begin{bmatrix} 1_A & -f'^{\dagger} \\ f' & 1_B \end{bmatrix} \begin{bmatrix} X'_{AA} & 0 \\ 0 & X'_{BB} \end{bmatrix}, \tag{2.59a}$$

where

$$f' = X'_{BA}\,X'^{-1}_{AA}\,, \tag{2.59b}$$

analogous to eqs. (2.2) and (2.3).

To obtain the relationship between f and f', we proceed as follows. From (2.2),

$$X' = V\,X = V\,\hat{T}\,\hat{X}$$
$$= \hat{T}'\,\hat{X}'\,,$$

or

$$\hat{T}' = V\,\hat{T}\,\hat{X}\,\hat{X}'^{-1}, \tag{2.60}$$

where the right hand side of eq. (2.60) is independent of f', but does depend on the truncated new eigenvectors $\hat{X}'$. However, from the AA block equation of (2.60), it follows that,

$$X'_{AA} = (V_{AA} + V_{AB}\,f)\,X_{AA}. \tag{2.61}$$

Substitution of this equation into the BA block equation of (2.60) gives f' in terms of f and V only,

$$f' = (V_{BA} + V_{BB}f)(V_{AA} + V_{AB}f)^{-1}. \tag{2.62}$$

While such a complicated transformation can be very

inconvenient in some cases, it is also a feature which can be usefully exploited. In calculations in which the elements of f are acting as coordinates, the metric character of the object function can be radically altered by a simple basis change, because of the nonlinear dependence of f' on V. For quantities transforming linearly in V, such a basis change merely results in a rotation of the object function. This point is discussed in greater detail in chapters 5 and 8.

If f is small, the inverse matrix in (2.62) can be expanded as

$$\begin{aligned}(V_{AA} + V_{AB}f)^{-1} &= V_{AA}^{-1}\,(1_A + V_{AB}fV_{AA}^{-1})^{-1} \\ &= V_{AA}^{-1} - V_{AA}^{-1}V_{AB}fV_{AA}^{-1} + \ldots\,, \qquad (2.63)\end{aligned}$$

and thus, to first order in f,

$$f' = V_{BA}V_{AA}^{-1} + (V_{BB} - V_{BA}V_{AA}^{-1}V_{AB})fV_{AA}^{-1} + O(f^2). \qquad (2.64)$$

Thus, if f is small, the transformation (2.62) is nearly linear, although not homogeneous.

2.2 Effective Operators

2.2.a Basic Definitions

The primary application of the partitioning formalism just described is in the construction of effective operators. In this context, such operators are defined in either of the subspaces S_A or S_B of the full basis space, but their eigenvalues form a subset of the eigenvalues of the original operator in the full basis space, and the corresponding eigenvectors are related in some way to those of the original operator. There are two ways of regarding the matrices of such effective operators. They can be regarded as the matrix of a transformed operator in the old basis (active sense), or, alternatively, as the matrix of the old operator in a transformed basis (passive sense). Both points of view are equivalent, but in what follows, the former will be emphasized.

The simplest set of such effective operators for the matrix H has already been defined in equations (2.14) and (2.15). In S_A, we have the operator

$$\hat{H}_A = H_{AA} + H_{AB} f, \tag{2.65a}$$

with the eigenvalue equation

$$\hat{H}_A X_{AA} = X_{AA} \varepsilon^{(A)}, \tag{2.65b}$$

and in S_B,

$$\hat{H}_B = H_{BB} - H_{BA} f^{\dagger}, \tag{2.66a}$$

with the eigenvalue equation,

$$\hat{H}_B X_{BB} = X_{BB} \varepsilon^{(B)}. \tag{2.66b}$$

Both $\hat{H}_A$ and $\hat{H}_B$ are non-hermitian in general, although their eigenvalues $\xi^{(A)}$ and $\xi^{(B)}$ are real, since they are subsets of the eigenvalues of the hermitian operator H. The eigenvectors X_{AA} and X_{BB} are not orthonormal in general, because they are truncations of the orthonormal eigenvectors X of the full hamiltonian H.

It is possible to derive a pair of self-adjoint effective operators directly from the eigenvalue equation (2.1a). Premultiplication by $\hat{T}^\dagger$, and use of eq. (2.17), yields,

$$G_A X_{AA} = g_A X_{AA} \xi^{(A)}, \tag{2.67a}$$

where

$$G_A = (\hat{T}^\dagger H \hat{T})_{AA} = H_{AA} + H_{AB}f + f^\dagger H_{BA} + f^\dagger H_{BB} f, \tag{2.67b}$$

and,

$$G_B X_{BB} = g_B X_{BB} \xi^{(B)}, \tag{2.68a}$$

where

$$G_B = (\hat{T}^\dagger H \hat{T})_{BB} = H_{BB} - H_{BA}f^\dagger - fH_{AB} + fH_{AA}f^\dagger. \tag{2.68b}$$

The off-diagonal block of $\hat{T}^\dagger H \hat{T}$ is given by,

$$G_{BA} = H_{BA} + H_{BB}f - fH_{AA} - fH_{AB}f, \tag{2.69}$$

which is just the quantity D(f), defined in eq. (2.16). When $G_{BA} = 0$, it can be shown that,

$$G_A = g_A \hat{H}_A, \qquad G_B = g_B \hat{H}_B, \tag{2.70}$$

using eqs. (2.9), and the definitions of the effective operators presented above. Thus, when f is known exactly, the self-adjoint effective operators G_A and G_B could be considered to be obtained from the non-selfadjoint effective operators $\hat{H}_A$

and $\hat{H}_B$ by orthonormalizing the eigenvectors of the latter.

It is also possible to obtain self-adjoint effective operators in S_A and S_B by orthonormalizing the truncated eigenvectors. The effective operators $\hat{H}_A$, $\hat{H}_B$, and G_A and G_B above, are uniquely determined once particular partitionings of the basis and eigenvector spaces are defined. The self-adjoint effective operators obtained by orthonormalization are not unique, however, in that they depend on the particular orthonormalization procedure employed.

The symmetrical orthogonalization procedure of Lowdin (1970) and others, has the feature that the new orthonormalized vectors resemble the initial vectors as closely as possible, in a particular sense.[1] Applied to the present case, the new orthonormal eigenvectors are given by

$$C_{AA} = g_A^{\frac{1}{2}} X_{AA} , \tag{2.71a}$$

in S_A, and,

$$C_{BB} = g_B^{\frac{1}{2}} X_{BB} , \tag{2.71b}$$

in S_B. Thus one has,

$$\begin{aligned} C_{AA}^{\dagger} C_{AA} &= X_{AA}^{\dagger} g_A X_{AA} = 1_A , \\ C_{BB}^{\dagger} C_{BB} &= X_{BB}^{\dagger} g_B X_{BB} = 1_B , \end{aligned} \tag{2.72}$$

by eq. (2.8). The eigenvalue equation in C_{AA} is obtained either by premultiplying (2.65b) by $g_A^{\frac{1}{2}}$ or (2.67a) by $g_A^{-\frac{1}{2}}$ to get,

[1] In the notation used above, the difference between the two sets is measured by

$$\sum_{i,j} |C_{ij} - X_{ij}|^2$$

which is minimized if C is given by eq. (2.71) (Lowdin, 1970).

$$\tilde{H}_A C_{AA} = C_{AA} \xi^{(A)}, \tag{2.73}$$

where

$$\tilde{H}_A = g_A^{\frac{1}{2}} \hat{H}_A g_A^{-\frac{1}{2}} \tag{2.74a}$$

$$= g_A^{-\frac{1}{2}} G_A g_A^{-\frac{1}{2}}. \tag{2.74b}$$

Similarly, premultiplication of (2.66b) by $g_B^{\frac{1}{2}}$ or (2.68a) by $g_B^{-\frac{1}{2}}$, gives the equation

$$\tilde{H}_B C_{BB} = C_{BB} \xi^{(B)}, \tag{2.75}$$

where

$$\tilde{H}_B = g_B^{\frac{1}{2}} \hat{H}_B g_B^{-\frac{1}{2}} \tag{2.76a}$$

$$= g_B^{-\frac{1}{2}} G_B g_B^{-\frac{1}{2}}. \tag{2.76b}$$

It is also possible to define effective operators in either S_A or S_B for any other operator defined in the unpartitioned space. For some operator M,

$$X^{(A)\dagger} M X^{(A)} = X_{AA}^{\dagger} \tilde{M}_A X_{AA}, \tag{2.77}$$

where

$$\tilde{M}_A = \begin{bmatrix} 1_A & f^{\dagger} \end{bmatrix} \begin{bmatrix} M_{AA} & M_{AB} \\ M_{BA} & M_{BB} \end{bmatrix} \begin{bmatrix} 1_A \\ f \end{bmatrix}$$

$$= M_{AA} + M_{AB} f + f^{\dagger} M_{BA} + f^{\dagger} M_{BB} f. \tag{2.78}$$

$\tilde{M}_A$ has the same form in M as G_A defined in (2.67b) has in H. Here $\tilde{M}_A$ has the same expectation values for the truncated eigenvectors X_{AA} as the operator M has for the full eigenvectors $X^{(A)}$. An effective operator with the same properties with respect to the orthonormalized eigenvectors C_{AA} is clearly given by,

$$\tilde{\tilde{M}}_A = g_A^{-\frac{1}{2}} \tilde{M}_A g_A^{-\frac{1}{2}}, \tag{2.79}$$

which is analogous to $\tilde{H}_A$ defined in eq. (2.74). The analogue of the effective operator $\hat{H}_A$ of eq. (2.65a) can be obtained by premultiplying $\tilde{M}_A$ by g_A^{-1}, following eq. (2.67b). Effective operators for M restricted to S_B, analogous to (2.77) - (2.79), can be obtained in a similar manner.

2.2.b Eigenvectors and Eigenvalues of the Effective Operators

In order to amplify the material in the immediately preceding subsection, the connection between the eigenvalues and eigenvectors of the operator H and those of the effective operators $\hat{H}$, G, and $\tilde{H}$, will be illustrated here from a different point of view. The full operator H has the eigenvalue equation

$$\begin{aligned} &H\psi_i = \xi_i \psi_i, \qquad (i = 1, \ldots, n), \\ &\langle \psi_i | \psi_j \rangle = \delta_{ij} . \end{aligned} \tag{2.80}$$

Once the two basis spaces, S_A and S_B, are defined, each eigenvector ψ_i can be written as a sum of two parts,

$$\psi_i = \psi_{iA} + \psi_{iB} , \tag{2.81}$$

one part in S_A and one in S_B. The eigenvectors are themselves divided into two sets, $\psi_i^{(A)}$, $(i = 1, \ldots, n_A)$, and $\psi_i^{(B)}$, $(i = 1, \ldots, n_B)$, where $n_A + n_B = n$, according as they lie in S_A' or S_B'. The basic property of f is to map the part of $\psi_i^{(A)}$ in S_A into the part in S_B, according to

$$\psi_{iB}^{(A)} = f\,\psi_{iA}^{(A)}. \tag{2.82a}$$

Similarly, one has,

$$\psi_{iA}^{(B)} = (-f^{\dagger})\,\psi_{iB}^{(B)}. \tag{2.82b}$$

Combination with eq. (2.81) yields,

$$\begin{aligned}\psi_{i}^{(A)} &= \psi_{iA}^{(A)} + \psi_{iB}^{(A)} = \psi_{iA}^{(A)} + f\psi_{iA}^{(A)} \\ &= (1_A + f)\psi_{iA}^{(A)},\end{aligned} \tag{2.83a}$$

and,

$$\begin{aligned}\psi_{i}^{(B)} &= \psi_{iA}^{(B)} + \psi_{iB}^{(B)} = -f^{\dagger}\psi_{iB}^{(B)} + \psi_{iB}^{(B)} \\ &= (1_B - f^{\dagger})\,\psi_{iB}^{(B)}.\end{aligned} \tag{2.83b}$$

In the notation used here and throughout this subsection, the operator f is to be regarded, when necessary, as embedded in the n-dimensional basis space, but will be denoted by the same symbol as before.

The eigenvalue equation for $\hat{H}_A$ is

$$\hat{H}_A\,\psi_{iA}^{(A)} = \xi_i^{(A)}\,\psi_{iA}^{(A)}, \qquad (i = 1, \ldots, n_A), \tag{2.84a}$$

where the eigenvectors satisfy

$$\langle \psi_{iA}^{(A)} | g_A | \psi_{jA}^{(A)} \rangle = \delta_{ij}, \qquad (i,j = 1, \ldots, n_A). \tag{2.84b}$$

For the effective operator G_A, it is

$$G_A\,\psi_{iA}^{(A)} = \xi_i^{(A)}\,g_A\,\psi_{iA}^{(A)}, \qquad (i = 1, \ldots, n_A), \tag{2.85}$$

with the same orthonormality conditions (2.84b). Finally, for the effective operator $\tilde{H}_A$, the eigenvalue equation is

$$\tilde{H}_A\,\chi_{iA}^{(A)} = \xi_i^{(A)}\,\chi_{iA}^{(A)}, \qquad (i = 1, \ldots, n_A), \tag{2.86a}$$

where,

$$\chi_{iA}^{(A)} = g_A^{\frac{1}{2}} \psi_{iA}^{(A)}, \qquad (i = 1, \ldots, n_A), \qquad (2.86b)$$

and,

$$\langle \chi_{iA}^{(A)} | \chi_{jA}^{(A)} \rangle = \delta_{ij}, \qquad (i,j = 1, \ldots, n_A). \qquad (2.86c)$$

In terms of the eigenvectors of $\tilde{H}_A$, the eigenvectors of the original operator H are

$$\psi_i^{(A)} = (1_A + f) g_A^{-\frac{1}{2}} \chi_{iA}^{(A)}. \qquad (2.87)$$

In all of these equations, the eigenvalues $\xi_i^{(A)}$, $(i = 1, \ldots, n_A)$, are exactly the n_A eigenvalues of the original operator H corresponding to the eigenvectors $\psi_i^{(A)}$, $(i = 1, \ldots, n_A)$. The eigenvalue equations for the effective operators $\hat{H}_B$, G_B, and H_B, defined in S_B, are of the same form as those given above for the corresponding effective operators in S_A.

Finally, consider the projections P_A' and P_B' onto the eigenspaces S_A', and S_B', respectively. For P_A',

$$P_A' = \sum_{i=1}^{n_A} |\psi_i^{(A)}\rangle\langle\psi_i^{(A)}|$$

$$= \sum_{i=1}^{n_A} (1_A + f) |\psi_{iA}^{(A)}\rangle\langle\psi_{iA}^{(A)}| (1_A + f^\dagger) \qquad (2.88a)$$

Here,

$$\sum_{i=1}^{n_A} |\psi_{iA}^{(A)}\rangle\langle\psi_{iA}^{(A)}| = g^{(A)-1}, \qquad (2.88b)$$

defines an embedding of the inverse of the metric g_A in the n-dimensional basis space. Similarly,

$$P_B' = \sum_{i=1}^{n_B} |\psi_i^{(B)}\rangle\langle\psi_i^{(B)}|$$

$$= \sum_{i=1}^{n_B} (1_B - f^\dagger) |\psi_{iB}^{(B)}\rangle\langle\psi_{iB}^{(B)}| (1_B - f), \qquad (2.89a)$$

where,

$$\sum_{i=1}^{n_B} |\psi_{iB}^{(B)}\rangle\langle\psi_{iB}^{(B)}| = g^{(B)-1}, \qquad (2.89b)$$

is an embedding of the inverse of the metric g_B in the full n-dimensional basis space.

2.2.c Relationships With Other Formulations

Many of the quantities defined or derived above have appeared in one form or another in the literature, usually in connection with the calculation of effective operators in a perturbation formalism. The treatment by Friedrichs (1965) of an isolated part of the spectrum of an operator H, is particularly interesting in this regard. Several interrelations between the current non-canonical formulation and the more commonly used unitary methods are illustrated by rewriting some of the quantities introduced in that treatment, using the block notation employed here.

Following Friedrichs, the aim here is to obtain an expression for a projection operator P_ϵ onto a space spanned by a set of eigenvectors which correspond to an isolated part of the spectrum of some perturbed operator H. Rather than requiring that the projection P_ϵ be orthogonal (that is, that the operator P_ϵ be hermitian), or explicitly idempotent, it is

required only that

$$P_\epsilon P_o = P_\epsilon, \qquad P_o P_\epsilon = P_o, \tag{2.90}$$

where P_o is a projection onto the corresponding eigenspace of the unperturbed operator H_o. These linear conditions, (2.90), imply idempotency,

$$P_\epsilon^2 = P_\epsilon P_o P_\epsilon = P_\epsilon P_o = P_\epsilon,$$

and

$$P_o^2 = P_o P_\epsilon P_o = P_o P_\epsilon = P_o,$$

thus verifying that P_o and P_ϵ are projections. However, by themselves, they do not imply that $P_\epsilon^\dagger = P_\epsilon$, or that $P_o^\dagger = P_o$. Equations (2.90) represent the minimal conditions for P_ϵ to be a projection, without prescribing any information about the internal structure of its image space.

In a basis adapted to the solution of the zero order problem, that is, with the matrix representation,

$$P_o = \begin{bmatrix} 1_A & 0 \\ 0 & 0 \end{bmatrix}, \tag{2.91}$$

where the subscript A denotes the space spanned by the zero order eigenvectors of interest, the form of the matrix representation of P_ϵ is restricted by (2.90) to

$$P_\epsilon = \begin{bmatrix} 1_A & 0 \\ f_\epsilon & 0 \end{bmatrix}, \tag{2.92}$$

where f_ϵ is a matrix undetermined by (2.90).

It is now possible to define mappings, $U_\epsilon^-(S_o \to S_\epsilon)$ and $U_\epsilon^+(S_\epsilon \to S_o)$, between the spaces S_ϵ spanned by the eigenfunctions

of the perturbed operator, and the space S_o spanned by the corresponding eigenfunctions of the unperturbed operator. In terms of the projections P_o and P_ϵ, these mappings are

$$U_\epsilon^{\pm} = 1 \mp (P_\epsilon - P_o), \tag{2.93}$$

as given by Friedrichs. It then follows that the operator

$$H_\epsilon^F = U_\epsilon^{+} H_\epsilon U_\epsilon^{-}, \tag{2.94}$$

is from S_o to S_o, but has the same spectrum as H_ϵ, the perturbed operator. That is, H_ϵ^F is an effective operator in the space S_o.

In the matrix notation introduced above, the mappings $U_\epsilon^{\pm}$ are

$$U_\epsilon^{\pm} = \begin{bmatrix} 1_A & 0 \\ \mp f_\epsilon & 1_B \end{bmatrix}, \tag{2.95}$$

where the subscript B denotes the space of all eigenvectors of H_o except those of interest. Thus, in the notation developed in the previous sections,

$$\tilde{H}_\epsilon^F = \begin{bmatrix} \hat{H}_A(f_\epsilon) & H_{AB} \\ D(f_\epsilon) & \hat{H}_B^{\dagger}(f_\epsilon) \end{bmatrix}. \tag{2.96}$$

It is possible to define a new set of unitary mappings, $U_\epsilon^{\pm\prime}$, which map between S_ϵ and S_o, and vice versa, as

$$U_\epsilon^{\pm\prime} = \begin{bmatrix} 1_A & \pm f_\epsilon^{\dagger} \\ \mp f_\epsilon & 1_B \end{bmatrix}. \tag{2.97}$$

Using (2.97) instead of (2.95) in eq. (2.94), a new transformed perturbed operator is obtained,

$$\tilde{H}_\epsilon^{F'} = U_\epsilon^{+'} H_\epsilon U_\epsilon^{-'} = \begin{bmatrix} G_A(f_\epsilon) & D(f_\epsilon)^\dagger \\ D(f_\epsilon) & G_B(f_\epsilon) \end{bmatrix} . \qquad (2.98)$$

which is self-adjoint. The operators G_A and G_B are given by eqs. (2.67b) and (2.68b). These results are in accord with the fact that the non-selfadjointness of the operators $\hat{H}_A$ and $\hat{H}_B$, introduced in the previous section, is associated with the fact that the mappings between the two spaces S_A and S_B are not unitary (that is, they do not leave the inner product unchanged).

We point out that $\tilde{H}_\epsilon^{F'}$ is block diagonal when the matrix block f_ϵ of $U_\epsilon^{\pm'}$ satisfies $D(f_\epsilon) = 0$, eq. (2.16). It is interesting to note that choosing the matrix block f_ϵ in $U_\epsilon^{\pm}$ of eq. (2.95) to satisfy (2.16), is equivalent to a partial reduction of H_ϵ toward the upper Hessenberg form, the result of a non-unitary procedure used in numerical matrix diagonalization. However, $\tilde{H}_\epsilon^{F}$ is not exactly upper Hessenberg even if $D(f_\epsilon)$ vanishes, because the diagonal blocks of $\hat{H}_A$ and $\hat{H}_B$ are not upper triangular in general.

Finally, note that Friedrichs introduces an operator $(P_\epsilon^\dagger P_\epsilon)^{-1}$, which is defined only in the image space S_o of P_o. In the matrix notation used above,

$$P_\epsilon^\dagger P_\epsilon = 1_A + f_\epsilon^\dagger f_\epsilon = (g_\epsilon)_A . \qquad (2.99)$$

Thus, the orthogonal projection onto S_ϵ, given by Friedrichs as

$$\hat{P}_\epsilon = P_\epsilon (P_\epsilon^\dagger P_\epsilon)^{-1} P_\epsilon ,$$

is written in matrix notation here as

$$\hat{P}_\epsilon = \lim_{\alpha \to 0} \begin{bmatrix} 1_A & 0 \\ f_\epsilon & 0 \end{bmatrix} \begin{bmatrix} (\hat{g}_\epsilon)_A^{-1} & 0 \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} 1_A & f_\epsilon^\dagger \\ 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} (g_\epsilon)_A^{-1} & (\hat{g}_\epsilon)_A^{-1} f_\epsilon^\dagger \\ f_\epsilon (g_\epsilon)_A^{-1} & f_\epsilon (g_\epsilon)_A^{-1} f_\epsilon^\dagger \end{bmatrix}, \qquad (2.100)$$

which is identical to the projection P_A' of eq. (2.10).

Finally, we also point out that the operator $\tilde{H}$, defined by symmetrical orthogonalization in eqs. (2.74) and (2.76), coincides with operators of Sz.-Nagy (1946/47; see also Riesz and Sz.-Nagy, 1955, §136), Primas (1961, 1963), and also Kato (1966, Remark 4.4 of chapter 2).

2.3 Generalization to a Nonorthonormal Basis Set

The formalism presented in the first part of this chapter can easily be generalized to the situation in which the basis functions ϕ_i, $(i = 1, \ldots, n)$, being used, are not orthonormal. In this case, the eigenvalue equation has the form

$$H X = S X E, \tag{2.101a}$$

with normalization

$$X^\dagger S X = 1_n , \tag{2.101b}$$

where the elements of the matrix S are the inner products of the basis functions,

$$S_{ij} = \langle \phi_i | \phi_j \rangle .$$

The partitioning of the basis set, and of the eigenvectors of H into two sets of dimensions n_A and n_B, respectively, is carried out exactly as before, leading to eq. (2.2),

$$X = \begin{bmatrix} 1_A & h \\ f & 1_B \end{bmatrix} \begin{bmatrix} X_{AA} & 0 \\ 0 & X_{BB} \end{bmatrix} = \hat{T}\,\hat{X} ,$$

where f and h are again formally given by (2.3). However, as a result of the more complicated normalization condition (2.101b), the simple relation (2.4) is now replaced by

$$h = -(S_{AA} + f^\dagger S_{BA})^{-1}(S_{AB} + f^\dagger S_{BB}). \tag{2.102}$$

Because of the complexity of (2.102), it is convenient here to retain the notation h and f throughout, rather than eliminate h entirely, as was done for the orthonormal case.

The metric matrices for the truncated eigenvectors, as in eq. (2.8), are given by the diagonal blocks of the product

$\hat{T}^{\dagger} S \hat{T}$,

$$g_A = S_{AA} + S_{AB}f + f^{\dagger}S_{BA} + f^{\dagger}S_{BB}f, \tag{2.103a}$$

and

$$g_B = S_{BB} + S_{BA}h + h^{\dagger}S_{AB} + h^{\dagger}S_{AA}h. \tag{2.103b}$$

The projection P_A' still has the form (2.10), but the projection P_B' must here be written,

$$P_B' = \begin{bmatrix} h \\ 1_B \end{bmatrix} g_B^{-1} \begin{bmatrix} h^{\dagger} & 1_B \end{bmatrix} = \begin{bmatrix} hg_B^{-1}h^{\dagger} & hg_B^{-1} \\ g_B^{-1}h^{\dagger} & g_B^{-1} \end{bmatrix}. \tag{2.104}$$

These projections are self-adjoint, but now the idempotency conditions become $(P_A'S)^2 = P_A'S$, and $(P_B'S)^2 = P_B'S$, as can be verified by direct matrix multiplication. Also, it can easily be shown that tr $P_A'S = n_A$, and tr $P_B'S = n_B$.

The defining conditions on f and h can be obtained from the analogue of eq. (2.13), namely,

$$H \hat{T} = S \hat{T} \hat{H}, \tag{2.105a}$$

where

$$\hat{H} = \hat{X} \mathcal{E} \hat{X}^{-1}, \tag{2.105b}$$

is to be block diagonal. The non-selfadjoint effective operators $\hat{H}_A$ and $\hat{H}_B$ are given by the diagonal blocks of (2.105a) as

$$\hat{H}_A = (S_{AA} + S_{AB}f)^{-1}(H_{AA} + H_{AB}f), \tag{2.106}$$

and

$$\hat{H}_B = (S_{BA}h + S_{BB})^{-1}(H_{BA}h + H_{BB}). \tag{2.107}$$

With these definitions, the eigenvalue equations for these effective operators have exactly the same form as in the orthonormal case.

Alternatively, the inverse matrices in (2.106) and (2.107) could be transferred to the right hand sides of the eigenvalue equations for $\hat{H}_A$ and $\hat{H}_B$, respectively, and be regarded as effective overlap matrices, giving eigenvalue equations of the form

$$\hat{H}'_A X_{AA} = \hat{S}_A X_{AA} \xi^{(A)}, \tag{2.108}$$

and

$$\hat{H}'_B X_{BB} = \hat{S}_B X_{BB} \xi^{(B)}, \tag{2.109}$$

where now, $\hat{H}'_A$ and $\hat{H}'_B$ are given formally by eq. (2.15), as

$$\hat{H}'_A = H_{AA} + H_{AB}f, \tag{2.110}$$

and

$$\hat{H}'_B = H_{BB} + H_{BA}h. \tag{2.111}$$

The operators $\hat{S}_A$ and $\hat{S}_B$ are of the same form in S,

$$\hat{S}_A = S_{AA} + S_{AB}f, \qquad \hat{S}_B = S_{BA}h + S_{BB}. \tag{2.112}$$

Equations (2.108) and (2.109) are generalized eigenvalue equations for a non-selfadjoint operator.

Using (2.106) and (2.107) in the off-diagonal blocks of eq. (2.105a), the defining equations for f and h, analogous to (2.16) and (2.17), are now found to be

$$\begin{aligned} D(f) &= H_{BA}+H_{BB}f-(S_{BA}+S_{BB}f)\hat{H}_A \\ &= H_{BA}+H_{BB}f-(S_{BA}+S_{BB}f)(S_{AA}+S_{AB}f)^{-1}(H_{AA}+H_{AB}f) = 0, \end{aligned} \tag{2.113}$$

and

$$\begin{aligned} D'(h) &= H_{AB}+H_{AA}h-(S_{AA}h+S_{AB})\hat{H}_B \\ &= H_{AB}+H_{AA}h-(S_{AA}h+S_{AB})(S_{BA}h+S_{BB})^{-1}(H_{BA}h+H_{BB}) = 0. \end{aligned} \tag{2.114}$$

As for an orthonormal basis, the equations for f here are not coupled to those for h. Equations (2.113) and (2.114) are not the only useful equations defining f and h. An alternative approach is used in some detail in the next chapter.

Self-adjoint effective operators can again be obtained by premultiplying the eigenvalue equation, (2.101a), by $\hat{T}^{\dagger}$. The resulting operator, G_A, in S_A is given by eq. (2.67b), but the corresponding effective operator in S_B must now be written

$$G_B = H_{BB} + H_{BA}h + h^{\dagger}H_{AB} + h^{\dagger}H_{AA}h \qquad (2.115a)$$

$$= g_B H_B, \qquad (2.115b)$$

with (2.115b) holding only if eqs. (2.113) and (2.114) are satisfied. The eigenvalue equations for these effective operators are as in eqs. (2.67a) and (2.68a), applicable also in an orthonormal basis. The BA block of $\hat{T}^{\dagger}H\hat{T}$ is

$$G_{BA} = H_{BA} + H_{BB}f + h^{\dagger}(H_{AA} + H_{AB}f), \qquad (2.116)$$

which becomes identical to D(f) if h is given by (2.102).

The effective operators $\tilde{H}_A$ and $\tilde{H}_B$ are given by eqs. (2.74) and (2.76), respectively, in this case. Their eigenvalue equations are given by (2.73) and (2.75).

Sets of contragredient vectors can be defined here in terms of the columns of P_A' and $(1 - P_A'S)$, and their reciprocal vectors. These are useful in writing various quantities in a compact manner when a nonorthonormal basis is used. These vectors are considerably more complicated in this case than those given in section 2.1.d. Their detailed examination will be deferred until some motivation has been provided for defining them.

CHAPTER 3

THE EFFECTIVE HAMILTONIANS--

PRACTICAL CONSIDERATIONS

"So I prophesied as I was commanded: and as I prophesied, there was a noise, and behold a shaking, and the bones came together, bone to his bone. And when I beheld, lo, the sinews and the flesh came up upon them, and the skin covered them above: but there was no breath in them."

Ezekiel 37: 7,8 (KJV)

3.1 Alternative Formulas

The purpose of this section is to examine some of the inter-relationships between the effective operators described in sections 2.2 and 2.3, in somewhat greater detail, especially when f is known only approximately. As has been pointed out before, the two alternative expressions for the operators G_A and G_B given in eqs. (2.67b), (2.68b), and (2.70), are equivalent only if f satisfies D(f)=0. If f satisfies D(f)=0 only approximately, it is possible to distinguish two types of approximate effective operators, $\hat{H}_A$ and $\hat{H}_B$, namely,

$$\hat{H}_A^{(1)} = H_{AA} + H_{AB}f, \tag{3.1a}$$

$$\hat{H}_B^{(1)} = H_{BB} - H_{BA}f^\dagger, \tag{3.1b}$$

and

$$\hat{H}_A^{(2)} = g_A^{-1} G_A , \tag{3.2a}$$

$$\hat{H}_B^{(2)} = g_B^{-1} G_B , \tag{3.2b}$$

These two types of operators are related by

$$\hat{H}_A^{(2)} = \hat{H}_A^{(1)} + g_A^{-1}f^\dagger D^{(1)}(f), \tag{3.3a}$$

and

$$\hat{H}_B^{(2)} = \hat{H}_B^{(1)} + g_B^{-1}fD^{(1)}(f)^\dagger, \tag{3.3b}$$

where the notation $D^{(1)}(f)$ is defined below in eq. (3.4). Thus the two sets of formulas, (3.1) and (3.2), are equivalent only if $D^{(1)}(f) = 0$. In effect, $\hat{H}_A^{(2)}$ and $\hat{H}_B^{(2)}$, here are generalizations of the Rayleigh quotient $\langle\psi|H|\psi\rangle/\langle\psi|\psi\rangle$ for a single eigenfunction. The operators $\hat{H}_A^{(1)}$ and $\hat{H}_B^{(1)}$ correspond

to the use of an intermediate normalization, involving writing the expectation value of H as $\langle\phi|H|\psi\rangle/\langle\phi|\psi\rangle$, where ϕ is some arbitrary reference function. The Rayleigh quotient is second order in ψ, while this intermediate normalization is only first order in ψ.

In terms of the operators $\hat{H}_A^{(1)}$ and $\hat{H}_A^{(2)}$, eq. (2.16) can be written in one of two forms,

$$D^{(1)}(f) = H_{BA} + H_{BB}f - f\hat{H}_A^{(1)} = 0, \tag{3.4}$$

and

$$D^{(2)}(f) = H_{BA} + H_{BB}f - f\hat{H}_A^{(2)} = 0. \tag{3.5}$$

These two equations are equivalent in that they both have the same solutions. However, their detailed forms are quite different away from this solution. Equation (3.5) can be obtained directly by requiring that $\hat{T}^{-1}H\hat{T}$, rather than $\hat{T}^{\dagger}H\hat{T}$, be block diagonal, the latter being implicit in the derivation of (2.16). It can be shown that the relationship between these two quantities is given by

$$D^{(2)}(f) = g_B^{-1}D^{(1)}(f) , \tag{3.6}$$

in the case of an orthonormal basis.

It is also possible to distinguish between three different formulas for calculating operators of the type designated $\tilde{H}_A$, depending on which form of eq. (2.74) and also which form of $\hat{H}_A$ is used. Only one such form is useful, and for practical purposes, is given by either eq. (2.74b) or by

$$\tilde{H}_A = g_A^{+\frac{1}{2}}\,\hat{H}_A^{(2)}\,g_A^{-\frac{1}{2}} . \tag{3.7}$$

In the case of a nonorthonormal basis, the situation is considerably more complicated. Because the orthogonality condition, (2.101b), is no longer simple, it is necessary to allow for the possibility that if f and h do not exactly satisfy eqs. (2.113) and (2.114), they may also fail to satisfy eq. (2.102). As a result, the off-diagonal blocks of both the matrices,

$$G = \hat{T}^{\dagger}H\hat{T} = \begin{bmatrix} G_A & G_{AB} \\ G_{BA} & G_B \end{bmatrix} , \tag{3.8a}$$

and

$$g = \hat{T}^{\dagger}S\hat{T} = \begin{bmatrix} g_A & g_{AB} \\ g_{BA} & g_B \end{bmatrix} , \tag{3.8b}$$

must be considered to be potentially nonzero in what follows.

Two pairs of operators $\hat{H}_A$ and $\hat{H}_B$ are again defined in this case,

$$\hat{H}_A^{(1)} = (S_{AA} + S_{AB}f)^{-1}(H_{AA} + H_{AB}f), \tag{3.9a}$$

and

$$\hat{H}_B^{(1)} = (S_{BA}h + S_{BB})^{-1}(H_{BB} + H_{BA}h), \tag{3.9b}$$

identical to eqs. (2.106) and (2.107), and

$$\hat{H}_A^{(2)} = g_A^{-1}G_A , \tag{3.10a}$$

and

$$\hat{H}_B^{(2)} = g_B^{-1}G_B . \tag{3.10b}$$

These two sets of operators are shown in Appendix 1 to be related by the equations,

$$\hat{H}_A^{(2)} = \hat{H}_A^{(1)} + g_A^{-1} f^\dagger \, D^{(1)}(f), \tag{3.11}$$

and

$$\hat{H}_B^{(2)} = \hat{H}_B^{(1)} + g_B^{-1} f \, D^{(1)}(f)^\dagger, \tag{3.12}$$

where $D^{(1)}(f)$, given formally by eq. (3.4), is the quantity in eq. (2.113) defining f. Thus, these two pairs of operators are identical only when both $D^{(1)}(f) = 0$.

The two operators $\hat{H}_A^{(1)}$ and $\hat{H}_A^{(2)}$ give rise to two different defining conditions on f, given by

$$D^{(1)}(f) = H_{BA} + H_{BB} f - (S_{BA} + S_{BB} f)\hat{H}_A^{(1)} , \tag{3.13}$$

and

$$D^{(2)}(f) = H_{BA} + H_{BB} f - (S_{BA} + S_{BB} f)\hat{H}_A^{(2)} . \tag{3.14}$$

In this case, the relationship between the two quantities $D^{(1)}(f)$ and $D^{(2)}(f)$ is

$$D^{(2)}(f) = (S_{BB}+S_{BA}h)(g_B+h^\dagger g_{AB})^{-1}(1_B - g_{BA}g_A^{-1}f^\dagger)D^{(1)}(f). \tag{3.15a}$$

When $g_{BA} = 0$, this reduces to

$$D^{(2)}(f) = (S_{BB} + S_{BA}h)g_B^{-1}D^{(1)}(f), \tag{3.15b}$$

or

$$D^{(2)}(f) = [1_B - (S_{BA} + S_{BB}f)g_A^{-1}f^\dagger]D^{(1)}(f). \tag{3.15c}$$

The derivations of eqs. (3.12) - (3.15) are quite long, and have been outlined in Appendix 1.

3.2 Implications of Inexact Solutions

Consider an approximate solution, f^{approx}, to eq. (2.16), given by

$$f^{approx} = f + \delta f, \tag{3.16}$$

where f is an exact solution of (2.16). If the effective operators $\hat{H}_A$, G_A, and $\tilde{H}_A$ are calculated using f^{approx}, the error δf will result in errors in the effective operators at some order in δf.

Starting with the operator G_A, and writing

$$G_A^{approx} = G_A + \delta G_A , \tag{3.17a}$$

where G_A is exact, it is easily verified, using (2.16), that

$$\delta G_A = (\delta f^\dagger f)H_A + H_A(f^\dagger \delta f) + O(\delta^2), \tag{3.17b}$$

to first order in the errors. Here the operator $\hat{H}_A$ is exact. Thus the error in G_A^{approx} is first order in δf.

Similarly, from eq. (2.70),

$$G_A + \delta G_A = (g_A + \delta g_A)(\hat{H}_A^{(2)} + \delta\hat{H}_A^{(2)}),$$

or

$$\delta\hat{H}_A^{(2)} = g_A^{-1}[\delta G_A - \delta g_A \hat{H}_A] + O(\delta^2). \tag{3.18}$$

Since

$$\delta g_A = \delta f^\dagger f + f^\dagger \delta f + O(\delta^2), \tag{3.19}$$

eq. (3.18) then yields

$$\delta\hat{H}_A^{(2)} = g_A^{-1}[\hat{H}_A^\dagger(f^\dagger\delta f) - (f^\dagger\delta f)\hat{H}_A] + O(\delta^2). \tag{3.20}$$

On the other hand, from (3.1a), one has

$$\delta\hat{H}_A^{(1)} = H_{AB}\delta f, \tag{3.21}$$

exactly. Except for $n_A = 1$, both these errors, $\delta\hat{H}_A^{(1)}$ and $\delta\hat{H}_A^{(2)}$, are first order in δf. However, (3.20) consists of the difference of two very similar terms, which actually vanishes for $n_A = 1$. This corresponds to the familiar property that the error in an eigenvalue calculated as the Rayleigh quotient of an approximate eigenvector is second order in the error in the eigenvector. For $n_A > 1$, the first order error, (3.20), does not vanish in general, but, as will be shown presently, the first order correction to the eigenvalues does vanish.

Using eq. (3.20), and the result

$$\delta g_A^{-\frac{1}{2}} = -g_A^{-\frac{1}{2}}\delta g_A\, g_A^{\frac{1}{2}} g_A^{-\frac{1}{2}} + O(\delta^2), \tag{3.22}$$

it is easy to show from (3.7) that

$$\delta\tilde{H}_A = [\tilde{H}_A,\ g_A^{-\frac{1}{2}} f^{\dagger}\delta f g_A^{-\frac{1}{2}} - \delta g_A\, g_A^{\frac{1}{2}} g_A^{-\frac{1}{2}}]_- + O(\delta^2). \tag{3.23}$$

This also vanishes in first order when $n_A = 1$, but is in general non-vanishing when $n_A > 1$.

For a non-orthonormal basis, eqs. (3.17b) and (3.23) remain the same because the formula for the operators G_A and $\tilde{H}_A$ used in deriving these results does not contain the overlap matrix explicitly. However, for the two operators $\hat{H}_A^{(1)}$ and $\hat{H}_A^{(2)}$, the form of the errors caused by an error in f does differ from (3.20) and (3.21). From eq. (3.9a),

$$\delta\hat{H}_A^{(1)} = (S_{AA}+S_{AB}f)^{-1}(H_{AB}\delta f - S_{AB}\delta f\hat{H}_A^{(1)}) + O(\delta^2). \tag{3.24}$$

From (3.10a), and using the same procedure as was used to obtain eq. (3.20), one obtains,

$$\delta\hat{H}_A^{(2)} = g_A^{-1}[\hat{H}_A^{\dagger}(S_{AB} + f^{\dagger}S_{BB})\delta f - (S_{AB} + f^{\dagger}S_{BB})\delta f\hat{H}_A] + O(\delta^2). \tag{3.25}$$

Except for (3.25) vanishing when $n_A = 1$, both (3.24) and (3.25) are first order in δf. Although the first order term in (3.24) now involves a difference between two terms, the two terms are not very similar, as is the case in (3.25), and therefore $\hat{H}_A^{(2)}$ is still expected to be inherently more accurate for a non-orthonormal basis than $\hat{H}_A^{(1)}$.

The first order variation in the eigenvalues of G_A due to some variations δG_A, δg_A, in the operators G_A and g_A, respectively, is given by

$$\delta\xi_i = \langle\psi_{iA}^{(A)}|\delta G_A - \xi_i\delta g_A|\psi_{iA}^{(A)}\rangle + O(\delta^2). \tag{3.26}$$

The functions $\psi_{iA}^{(A)}$ are the eigenfunctions of the exact operator G_A, and eq. (3.26) follows directly from the eigenvalue equation, (2.85), for G_A. But upon substitution of (3.17a) and (3.20) into (3.26), and using the eigenvalue equation (2.84a) for $\hat{H}_A$, it is seen that $\delta\xi_i$ of (3.26) vanishes in the first order in δf.

In the case of the operator $\hat{H}_A^{(1)}$, the first order error in the eigenvalues is given by

$$\delta\xi_i^{(1)} = \langle\psi_{iA}^{(A)}|g_A\delta\hat{H}_A^{(1)}|\psi_{iA}^{(A)}\rangle = \langle\psi_{iA}^{(A)}|g_AH_{AB}\delta f|\psi_{iA}^{(A)}\rangle, \tag{3.27}$$

which clearly does not vanish in general. On the other hand, one has,

$$\delta\xi_i^{(2)} = \langle\psi_{iA}^{(A)}|g_A\delta\hat{H}_A^{(2)}|\psi_{iA}^{(A)}\rangle$$

$$= \langle \psi_{iA}^{(A)} | \hat{H}_A^{\dagger}(f^{\dagger}\delta f) - (f^{\dagger}\delta f)\hat{H}_A | \psi_{iA}^{(A)} \rangle + O(\delta^2)$$

$$= O(\delta^2). \qquad (3.28)$$

The operator $\hat{H}_A^{(2)}$ is thus inherently more accurate than the operator $\hat{H}_A^{(1)}$, when evaluated using an inexact f.

The first order errors in the eigenvalues of the operators $\tilde{H}_A$ are just given as the expectation values of the first order error operator, (3.23), with respect to the eigenfunctions $\chi_{iA}^{(A)}$ of the exact operator $\tilde{H}_A$, defined in eqs. (2.86b). Since $\delta\tilde{H}_A$ can be written as a commutator to first order, its expectation value will be zero, and therefore,

$$\delta \tilde{\varepsilon}_i = \langle \chi_{iA}^{(A)} | \delta\tilde{H}_A | \chi_{iA}^{(A)} \rangle = O(\delta^2). \qquad (3.29)$$

For a monorthonormal basis, only the expectation values of $\delta\hat{H}_A^{(1)}$ and $\delta\hat{H}_A^{(2)}$ are different in nature from those given above for an orthonormal basis. From (3.24), one obtains,

$$\delta \varepsilon_i^{(1)} = \langle \psi_{iA}^{(A)} | g_A \delta\hat{H}_A^{(1)} | \psi_{iA}^{(A)} \rangle + O(\delta^2)$$

$$= \langle \psi_{iA}^{(A)} | g_A (S_{AA}+S_{AB}f)^{-1}(H_{AB}\delta f - S_{AB}\delta f \hat{H}_A^{(1)}) | \psi_{iA}^{(A)} \rangle + O(\delta^2), \qquad (3.30)$$

which does not vanish in first order in δf under any obvious general conditions. However, from (3.25),

$$\delta \varepsilon_i^{(2)} = \langle \psi_{iA}^{(A)} | g_A \delta\hat{H}_A^{(2)} | \psi_{iA}^{(A)} \rangle + O(\delta^2)$$

$$= \langle \psi_{iA}^{(A)} | \hat{H}_A^{\dagger}(S_{AB}+f^{\dagger}S_{BB})\delta f - (S_{AB}+f^{\dagger}S_{BB})\hat{H}_A | \psi_{iA}^{(A)} \rangle + O(\delta^2)$$

$$= O(\delta^2). \qquad (3.31)$$

Therefore the eigenvalues of $\hat{H}_A^{(2)}$ in a non-orthonormal basis are affected only in second order by errors in f.

3.3 Perturbation Theory For $\hat{H}_A$, G_A, and H_A

The purpose of this section is to outline perturbation formulas for the eigenvalues and eigenvectors of the effective operators $\hat{H}_A$, G_A, and $\tilde{H}_A$, defined in the space S_A. For the most part, the formulas presented below are not new, however, those for $\hat{H}_A$ and G_A are not well known. These formulas are necessary if the eigenvalues and eigenvectors of the full operator are to be calculated via a perturbation procedure based on this partitioning formalism.

The formulas for $\hat{H}_A$ will be derived in some detail, because they are unusual in that $\hat{H}_A$ is a non-selfadjoint operator (but with real eigenvalues). Those for G_A and $\tilde{H}_A$ will then just be summarized.

3.3a The $\hat{H}_A$ Scheme

The eigenvalue equation in this case is written,

$$\hat{H}_A \psi_i = \xi_i \psi_i \ , \qquad \langle \psi_i | g_A | \psi_j \rangle = \delta_{ij} , \tag{3.32}$$

where the subscripts and superscripts 'A' on the ξ_i and the ψ_i have been suppressed, and will be throughout this section. The metric matrix g_A is selfadjoint. We have

$$\begin{aligned} \hat{H}_A &= \sum_{n=0}^{\infty} \hat{H}_A^{(n)} , & \psi_i &= \sum_{n=0}^{\infty} \psi_i^{(n)} , \\ \xi_i &= \sum_{n=0}^{\infty} \xi_i^{(n)} , & g_A &= \sum_{n=0}^{\infty} g_A^{(n)} , \end{aligned} \tag{3.33}$$

where the superscript is to indicate the order of the term

in the perturbation parameter (or parameters), and the solution of the zero order eigenvalue equation

$$\hat{H}_A^{(0)}\psi_i^{(0)} = \mathcal{E}_i^{(0)}\psi_i^{(0)} \quad , \quad \langle\psi_i^{(0)}|g_A^{(0)}|\psi_i^{(0)}\rangle = \delta_{ij} \tag{3.34}$$

is known. The terms in the series for $\hat{H}_A$ and g_A are given, and the terms in the series for $\mathcal{E}_i$ and ψ_i are to be calculated.

Consider the first order term of the eigenvalue equation and normalization condition (3.32), given by

$$(\hat{H}_A^{(1)} - \mathcal{E}_i^{(1)})\psi_i^{(0)} + (\hat{H}_A^{(0)} - \mathcal{E}_i^{(0)})\psi_i^{(1)} = 0, \tag{3.35a}$$

and

$$2\langle\psi_i^{(0)}|g_A^{(0)}|\psi_i^{(1)}\rangle + \langle\psi_i^{(0)}|g_A^{(1)}|\psi_i^{(0)}\rangle = 0, \tag{3.35b}$$

when all quantities are real. The first order eigenvalue $\mathcal{E}_i^{(1)}$ is obtained by premultiplying (3.35a) by $(g_A^{(0)}\psi_i^{(0)})^\dagger$, and integrating, to give

$$\mathcal{E}_i^{(1)} = \langle\psi_i^{(0)}|g_A^{(0)}\hat{H}_A^{(1)}|\psi_i^{(0)}\rangle \, . \tag{3.36}$$

No contribution is obtained from the second term of (3.35a), since from (3.34),

$$\langle\psi_i^{(0)}|g_A^{(0)}\hat{H}_A^{(0)}|\psi_i^{(1)}\rangle = \mathcal{E}_i^{(0)}\langle\psi_i^{(0)}|g_A^{(0)}|\psi_i^{(1)}\rangle,$$

cancelling the rest of the term. The first order wavefunction is obtained in a similar manner. Premultiplying (3.35a) by $(g_A^{(0)}\psi_k^{(0)})^\dagger$ and integrating, gives

$$(\mathcal{E}_k^{(0)} - \mathcal{E}_i^{(0)})\langle\psi_k^{(0)}|g_A^{(0)}|\psi_i^{(1)}\rangle = -\langle\psi_k^{(0)}|g_A^{(0)}(\hat{H}_A^{(1)} - \mathcal{E}_i^{(1)})|\psi_i^{(0)}\rangle.$$

Writing $\psi_i^{(1)}$ here as,

$$\psi_i^{(1)} = \sum_j \psi_j^{(0)} a_{ji}^{(1)}, \tag{3.37}$$

then gives

$$a_{ki}^{(1)} = \frac{\langle \psi_k^{(0)} | g_A^{(0)} \hat{H}_A^{(1)} | \psi_i^{(0)} \rangle}{\mathcal{E}_i^{(0)} - \mathcal{E}_k^{(0)}}, \quad k \neq i. \tag{3.38a}$$

The coefficient $a_{ii}^{(1)}$ is obtained from eq. (3.35b) as

$$a_{ii}^{(1)} = -\tfrac{1}{2} \langle \psi_i^{(0)} | g_A^{(1)} | \psi_i^{(0)} \rangle , \tag{3.38b}$$

and thus

$$\psi_i^{(1)} = \sum_{j \neq i} \frac{\langle \psi_j^{(0)} | g_A^{(0)} \hat{H}_A^{(1)} | \psi_i^{(0)} \rangle}{\mathcal{E}_i^{(0)} - \mathcal{E}_j^{(0)}} \psi_j^{(0)} - \tfrac{1}{2} \langle \psi_i^{(0)} | g_A^{(1)} | \psi_i^{(0)} \rangle \psi_i^{(0)}. \tag{3.39}$$

The second order terms of eqs. (3.32) are

$$(\hat{H}_A^{(2)} - \mathcal{E}_i^{(2)}) \psi_i^{(0)} + (\hat{H}_A^{(1)} - \mathcal{E}_i^{(1)}) \psi_i^{(1)} + (\hat{H}_A^{(0)} - \mathcal{E}_i^{(0)}) \psi_i^{(2)} = 0, \tag{3.40a}$$

and

$$2\langle \psi_i^{(2)} | g_A^{(0)} | \psi_i^{(0)} \rangle + 2\langle \psi_i^{(1)} | g_A^{(1)} | \psi_i^{(0)} \rangle + \langle \psi_i^{(1)} | g_A^{(0)} | \psi_i^{(1)} \rangle + \langle \psi_i^{(0)} | g_A^{(2)} | \psi_i^{(0)} \rangle = 0. \tag{3.40b}$$

The approach here is the same as that in the first order case. Premultiplication of (3.40a) by $(g_A^{(0)} \psi_i^{(0)})^\dagger$ and integration, leads to,

$$\mathcal{E}_i^{(2)} = \langle \psi_i^{(0)} | g_A^{(0)} \hat{H}_A^{(2)} | \psi_i^{(0)} \rangle + \langle \psi_i^{(0)} | g_A^{(0)} (\hat{H}_A^{(1)} - \mathcal{E}_i^{(1)}) | \psi_i^{(1)} \rangle . \tag{3.41a}$$

Substitution of eqs. (3.36) and (3.39) into (3.41) to eliminate $\mathcal{E}_i^{(1)}$ and $\psi_i^{(1)}$ from the latter, results in a formula somewhat

reminiscent of the usual Rayleigh-Schrodinger second order energy formula,

$$\mathcal{E}_i^{(2)} = \sum_{j\neq i} \frac{\langle \psi_i^{(0)} | g_A^{(0)} \hat{H}_A^{(1)} | \psi_j^{(0)} \rangle \langle \psi_j^{(0)} | g_A^{(0)} \hat{H}_A^{(1)} | \psi_i^{(0)} \rangle}{\mathcal{E}_i^{(0)} - \mathcal{E}_j^{(0)}} + \langle \psi_i^{(0)} | g_A^{(0)} \hat{H}_A^{(2)} | \psi_i^{(0)} \rangle . \tag{3.41b}$$

The second order wavefunction $\psi_i^{(2)}$ is expanded in terms of the zero order wavefunctions,

$$\psi_i^{(2)} = \sum_j \psi_j^{(0)} a_{ji}^{(2)} , \tag{3.42}$$

and the coefficients $a_{ji}^{(2)}$ determined from eqs. (3.40a,b) in the same manner as was used in the first order case. The final result is

$$\psi_i^{(2)} = \sum_{k\neq i} \left[\frac{\langle \psi_k^{(0)} | g_A^{(0)} \hat{H}_A^{(2)} | \psi_i^{(0)} \rangle + \langle \psi_k^{(0)} | g_A^{(0)} (\hat{H}_A^{(1)} - \mathcal{E}_i^{(1)}) | \psi_i^{(1)} \rangle}{\mathcal{E}_i^{(0)} - \mathcal{E}_k^{(0)}} \right] \psi_k^{(0)} \tag{3.43}$$

$$-\tfrac{1}{2}[2\langle \psi_i^{(1)} | g_A^{(1)} | \psi_i^{(0)} \rangle + \langle \psi_i^{(0)} | g_A^{(2)} | \psi_i^{(0)} \rangle + \langle \psi_i^{(1)} | g_A^{(0)} | \psi_i^{(1)} \rangle] \psi_i^{(0)} .$$

The pattern is now clear. The n^{th} order terms of eq. (3.32) can be written as

$$\sum_{j=0}^{n} (\hat{H}_A^{(j)} - \mathcal{E}_i^{(j)}) \psi_i^{(n-j)} = 0, \tag{3.44a}$$

and

$$\sum_{j=0}^{n} \sum_{k=0}^{n-j} \langle \psi_i^{(j)} | g_A^{(n-j-k)} | \psi_i^{(k)} \rangle = 0. \tag{3.44b}$$

Premultiplying (3.44a) by $(g_A^{(0)} \psi_i^{(0)})^\dagger$, and integrating gives,

$$\mathcal{E}_i^{(n)} = \langle \psi_i^{(0)} | g_A^{(0)} \hat{H}_A^{(n)} | \psi_i^{(0)} \rangle + \sum_{j=1}^{n-1} \langle \psi_i^{(0)} | g_A^{(0)} (\hat{H}_A^{(j)} - \mathcal{E}_i^{(j)}) | \psi_i^{(n-j)} \rangle . \tag{3.45}$$

The n^{th} order wavefunction, $\psi_i^{(n)}$, is expanded as a linear combination of the zero order wavefunctions, and the expansion coefficients are deduced from eqs. (3.44a,b). The result is

$$\psi_i^{(n)} =$$

$$\sum_{k\neq i} \frac{<\psi_k^{(0)}|g_A^{(0)}\hat{H}_A^{(n)}|\psi_i^{(0)}> + \sum_{j=1}^{n-1} <\psi_k^{(0)}|g_A^{(0)}(\hat{H}_A^{(j)} - \mathcal{E}_i^{(j)})|\psi_i^{(n-j)}>}{\mathcal{E}_i^{(0)} - \mathcal{E}_k^{(0)}} \psi_k^{(0)} \tag{3.46}$$

$$-\tfrac{1}{2}\sum_{j=0}^{n-1}\sum_{\substack{k=0\\k\neq n}}^{n-j} <\psi_i^{(j)}|g_A^{(n-j-k)}|\psi_i^{(k)}>\psi_i^{(0)}$$

No attempt has been made in any of these formulas to eliminate higher order quantities in terms of lower order ones, because that leads to computationally less efficient formulas. All formulas above are given in terms of the eigenfunctions of $\hat{H}_A$. To obtain formulas applicable in a matrix notation, the functions $\psi_i^{(j)}$ are replaced by column vectors $x_i^{(j)}$, and all operators by their matrix representations.

3.3.b The G_A Scheme

The eigenvalue equation in this case is written,

$$G_A\psi_i = \mathcal{E}_i g_A \psi_i. \qquad <\psi_i|g_A|\psi_j> = \delta_{ij}. \tag{3.47}$$

Both G_A and g_A are selfadjoint. We have

$$G_A = \sum_{n=0}^{\infty} G_A^{(n)}, \qquad \psi_i = \sum_{n=0}^{\infty} \psi_i^{(n)}$$
$$\mathcal{E}_i = \sum_{n=0}^{\infty} \mathcal{E}_i^{(n)}, \qquad g_A = \sum_{n=0}^{\infty} g_A^{(n)}, \tag{3.48}$$

where the $G_A^{(n)}$ and the $g_A^{(n)}$ are given, and the $\psi_i^{(n)}$ and the $\mathcal{E}_i^{(n)}$ are to be calculated. It is assumed that the zero order eigenvalue equation,

$$G_A^{(0)}\,\psi_i^{(0)} = \mathcal{E}_i^{(0)} g_A^{(0)}\,\psi_i^{(0)},$$
$$\langle \psi_i^{(0)} | g_A^{(0)} | \psi_j^{(0)} \rangle = \delta_{ij}, \qquad (3.49)$$

has been solved. The n^{th} order term of (3.47) is then

$$\sum_{j=0}^{n} [G_A^{(j)} - (\sum_{k=0}^{j} \mathcal{E}_i^{(j)} g_A^{(j-k)})]\psi_i^{(n-j)} = 0, \qquad (3.50a)$$

and

$$\sum_{j=0}^{m}\sum_{k=0}^{n-j} \langle \psi_i^{(j)} | g_A^{(n-j-k)} | \psi_i^{(k)} \rangle = 0. \qquad (3.50b)$$

The n^{th} order eigenvalue can be obtained by pre-multiplying (3.50a) by $\psi_i^{(0)\dagger}$ and integrating to give

$$\mathcal{E}_i^{(n)} = \sum_{j=1}^{n-1} \langle \psi_i^{(0)} | G_A^{(j)} - (\sum_{k=0}^{j} \mathcal{E}_i^{(k)} g_A^{(j-k)}) | \psi_i^{(n-j)} \rangle$$
$$+ \langle \psi_i^{(0)} | G_A^{(n)} - \sum_{k=0}^{n-1} \mathcal{E}_i^{(k)} g_A^{(j-k)} | \psi_i^{(0)} \rangle. \qquad (3.51)$$

The n^{th} order wavefunction is expanded in terms of the $\psi_j^{(0)}$, and the expansion coefficients deduced from eqs. (3.50a,b). The final result is,

$$\psi_i^{(n)} = \sum_{k\neq i}^{m}\sum_{j=1}^{n} \frac{\langle \psi_k^{(0)} | G_A^{(j)} - \sum_{i=0}^{j} \mathcal{E}_i^{(1)} g_A^{(j-1)} | \psi_i^{(n-j)} \rangle}{\mathcal{E}_i^{(0)} - \mathcal{E}_k^{(0)}} \psi_k^{(0)}$$
$$-\tfrac{1}{2}\sum_{\substack{j=0\\ k\neq n}}^{m-1}\sum_{k=0}^{n-j} \langle \psi_i^{(j)} | g_A^{(n-k-j)} | \psi_i^{(k)} \rangle \psi_i^{(0)} \qquad (3.52)$$

These formulas can be shown to be equivalent to those derived

in the $\hat{H}_A$ scheme, by expressing the $G_A^{(j)}$ in terms of the $g_A^{(j)}$ and $\hat{H}_A^{(j)}$, according to eq. (2.70). These results also agree with those given by A. Imamura (1968), in a different notation, to second order.

The first order formulas here are

$$\mathcal{E}_i^{(1)} = \langle \psi_i^{(0)} | G_A^{(1)} - \mathcal{E}_i^{(0)} g_A^{(1)} | \psi_i^{(0)} \rangle , \tag{3.53a}$$

and

$$\psi_i^{(1)} = \sum_{j \neq i} \frac{\langle \psi_j^{(0)} | G_A^{(1)} - \mathcal{E}_i^{(0)} g_A^{(1)} | \psi_i^{(0)} \rangle}{\mathcal{E}_i^{(0)} - \mathcal{E}_j^{(0)}} \psi_j^{(0)} - \tfrac{1}{2} \langle \psi_i^{(0)} | g_A^{(1)} | \psi_i^{(0)} \rangle \psi_i^{(0)} . \tag{3.53b}$$

The second order formulas are

$$\mathcal{E}_i^{(2)} = \sum_{j \neq i} \frac{| \langle \psi_i^{(0)} | G_A^{(1)} - \mathcal{E}_i^{(0)} g_A^{(1)} | \psi_j^{(0)} \rangle |^2}{\mathcal{E}_i^{(0)} - \mathcal{E}_j^{(0)}} + \langle \psi_i^{(0)} | G_A^{(2)} - \mathcal{E}_i^{(0)} g_A^{(2)} - \mathcal{E}_i^{(1)} g_A^{(1)} | \psi_i^{(0)} \rangle , \tag{3.54a}$$

and

$$\begin{aligned} \psi_i^{(2)} = \sum_{j \neq i} (\mathcal{E}_i^{(0)} - \mathcal{E}_j^{(0)})^{-1} & [\langle \psi_j^{(0)} | G_A^{(1)} - \mathcal{E}_i^{(1)} g_A^{(0)} - \mathcal{E}_i^{(0)} g_A^{(1)} | \psi_i^{(1)} \rangle \\ & + \langle \psi_j^{(0)} | G_A^{(2)} - \mathcal{E}_i^{(0)} g_A^{(2)} - \mathcal{E}_i^{(1)} g_A^{(1)} | \psi_i^{(0)} \rangle] \psi_j^{(0)} \\ & - \tfrac{1}{2} [\langle \psi_i^{(0)} | g_A^{(2)} | \psi_i^{(0)} \rangle + 2 \langle \psi_i^{(1)} | g_A^{(1)} | \psi_i^{(0)} \rangle \\ & + \langle \psi_i^{(1)} | g_A^{(0)} | \psi_i^{(1)} \rangle] \psi_i^{(0)} \end{aligned} \tag{3.54b}$$

Here, eq. (3.53b) was used to eliminate $\psi_i^{(1)}$ from the expression for $\mathcal{E}_i^{(2)}$. The resulting formula is longer, but the first term is now of the more familiar form for a second

order energy formula. The extra terms here compared to the usual Rayleigh-Schrodinger formula are due to the presence of $g_A^{(1)}$ and $g_A^{(2)}$.

3.3.c The $\tilde{H}_A$ Scheme

The eigenvalue equation in this case is

$$\tilde{H}_A \chi_i = \xi_i \chi_i , \qquad \langle \chi_i | \chi_j \rangle = \delta_{ij}, \tag{3.55}$$

where

$$\chi_i = g_A^{-\frac{1}{2}} \psi_i ,$$

and the ψ_i are the eigenfunctions considered in the $\hat{H}_A$ and G_A schemes. We have,

$$\tilde{H}_A = \sum_{n=0}^{\infty} \tilde{H}_A^{(n)}, \qquad \chi_i = \sum_{n=0}^{\infty} \chi_i^{(n)}, \qquad \xi_i = \sum_{n=0}^{\infty} \xi_i^{(n)}. \tag{3.56}$$

The solution of the zero order eigenvalue equation,

$$\tilde{H}_A^{(0)} \chi_i^{(0)} = \xi_i^{(0)} \chi_i^{(0)}, \quad \langle \chi_i^{(0)} | \chi_j^{(0)} \rangle = \delta_{ij}, \tag{3.57}$$

is assumed known. The $\tilde{H}_A^{(n)}$ are given, and the $\chi_i^{(n)}$ and $\xi_i^{(n)}$ are to be calculated. Since $\tilde{H}_A$ is selfadjoint, this is just the usual Rayleigh-Schrodinger perturbation theory.

The m^{th} order term of (3.55) is

$$\sum_{j=0}^{n} (\tilde{H}_A^{(j)} - \xi_i^{(j)}) \chi_i^{(n-j)} = 0, \tag{3.58a}$$

and

$$\sum_{j=0}^{n} \langle \chi_i^{(j)} | \chi_i^{(n-j)} \rangle = 0, \quad (n \neq 0). \tag{3.58b}$$

Pre-multiplying (3.58a) by $\chi_i^{(0)\dagger}$ and integrating gives the n^{th}

order eigenvalues,

$$\mathcal{E}_i^{(n)} = \langle \chi_i^{(0)} | \tilde{H}_A^{(n)} | \chi_i^{(0)} \rangle + \sum_{j=1}^{n-1} \langle \chi_i^{(0)} | \tilde{H}_A^{(j)} - \mathcal{E}_i^{(j)} | \chi_i^{(n-j)} \rangle . \tag{3.59}$$

The n^{th} order eigenfunction is expanded in terms of the zero order eigenfunctions,

$$\chi_i^{(n)} = \sum_j \chi_j^{(0)} a_{ji}^{(n)} , \tag{3.60}$$

where

$$a_{ki}^{(m)} = \sum_{j=1}^{n} \frac{\langle \chi_k^{(0)} | \tilde{H}_A^{(j)} - \mathcal{E}_i^{(j)} | \chi_i^{(n-j)} \rangle}{\mathcal{E}_i^{(0)} - \mathcal{E}_j^{(0)}} , \quad (k \neq i), \tag{3.61a}$$

and

$$a_{ii}^{(n)} = -\tfrac{1}{2} \sum_{j=1}^{n-1} \langle \chi_i^{(j)} | \chi_i^{(n-j)} \rangle . \tag{3.61b}$$

The coefficient $a_{ii}^{(1)}$ vanishes, but, in general, the $a_{ii}^{(n)}$ for $n > 1$, are not zero. Equations (3.59) and (3.61) can be written in terms of the eigenfunctions $\psi_i^{(j)}$ used in the $\hat{H}_A$ and G_A schemes using

$$\chi_i^{(j)} = \sum_k g_A^{\frac{1}{2}}(k) \psi_i^{(j-k)} . \tag{3.62}$$

The terms in the series

$$g_A^{\frac{1}{2}} = \sum_{j=0}^{\infty} g_A^{\frac{1}{2}}(j) , \tag{3.63}$$

are given below, in chapter 6.

CHAPTER 4

MULTIPLE PARTITIONING THEORY

"Great fleas have little fleas upon their back to bite 'em,
and little fleas have lesser fleas, and so ad infinitum.
The great fleas themselves in turn have greater fleas to go on,
while these again have greater still, and greater still, and so on."
(quoted in C. F. Froberg, Introduction to Numerical Analysis, (1969))

In this chapter, the possibilities of generalizing the formulas derived in the preceding two chapters to a more extensive partitioning will be examined. Such m x m partitioning formalisms, for $m > 2$, have a number of applications in the construction of effective operators and in the derivation of perturbation formulas in eigenvalue problems in which it is convenient or necessary to divide the eigenvalues and their eigenvectors into several distinct sets. The limiting partitioning formalism is that in which $m = n$, that is, the n-dimensional space spanned by the basis functions, and by the eigenvectors, is partitioned into n one-dimensional spaces. This is the ordinary eigenvalue problem.

4.1 The m x m Partitioning Formalism

4.1.a Basic Theory

For the present, it is assumed that the basis set used consists of orthonormal functions, so that the eigenvalue equation to be examined is

$$H X = X \mathcal{E} , \qquad X^{\dagger}X = 1_n , \tag{4.1}$$

where H is hermitian, X, the matrix of the eigenvectors of H, is unitary, and $\mathcal{E}$ is the real diagonal matrix of the eigenvalues of H. The set of basis functions is now divided into m subsets, each spanning one of m subspaces, S_1, S_2, ..., S_m, of dimensions, n_1, n_2, ..., n_m, respectively. Here, $\sum_{I=1}^{m} n_I$ is equal to n, the dimension of the full space. Similarly, the set of n eigenvectors of H, which are represented as the columns of X above, are divided into m subsets, $X^{(1)}$, $X^{(2)}$, ..., $X^{(m)}$, each spanning one of m subspaces S_1', S_2', ..., S_m', of the same respective dimensions n_1, n_2, ..., n_m. Because of this double partitioning, the matrices H and X can be written in an m x m block form,

$$H = \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1M} \\ H_{21} & H_{22} & \cdots & H_{2M} \\ \vdots & \vdots & & \vdots \\ H_{M1} & H_{M2} & \cdots & H_{MM} \end{bmatrix} , \quad X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1M} \\ X_{21} & X_{22} & \cdots & X_{2M} \\ \vdots & \vdots & & \vdots \\ X_{M1} & X_{M2} & \cdots & X_{MM} \end{bmatrix} , \tag{4.2}$$

where the symbols H_{IJ} and X_{IJ} represent n_I x n_J dimensional matrix blocks. Let the diagonal part of the eigenvector

matrix be denoted by

$$\hat{X} = \begin{bmatrix} X_{11} & & & 0 \\ & X_{22} & & \\ & & \ddots & \\ 0 & & & X_{MM} \end{bmatrix} . \qquad (4.3)$$

The basic quantities in this m x m partitioning formalism are now obtained as the off-diagonal blocks of the operator $\hat{T}$, defined, as for a 2 x 2 partitioning, by the equation,

$$X = \hat{T}\,\hat{X} . \qquad (4.4)$$

In the notation to be adopted, one has

$$\hat{T}_{II} = 1_I ,$$

and $\qquad (I,J = 1, \ldots, m), \qquad (4.5)$

$$\hat{T}_{IJ} = f_{IJ} , \quad J \neq I ,$$

where 1_I is the identity matrix in the space S_I, and f_{IJ} is an n_I x n_J matrix given by

$$X_{JI} = f_{JI} X_{II} , \qquad (4.6a)$$

or

$$f_{JI} = X_{JI}\, X_{II}^{-1} . \qquad (4.6b)$$

The operators f_{IJ} are straightforward generalizations of the two operators f and h defined for the 2 x 2 partitioning (where $f \equiv f_{21}$, and $h \equiv f_{12}$). The specific operator f_{KL} maps the part of an eigenvector, $x_r^{(L)}$, lying in the space S_L, into the part lying in the space S_K, where the eigenvector $x_r^{(L)}$ is

in the space $S_L^{\bullet}$. It is seen from eq. (4.6) that all the f_{IJ}, $(I,J = 1, \ldots, m, I \neq J)$, exist only if the matrix, $\hat{X}$, of eq. (4.3) is non-singular, or alternatively, only if the diagonal blocks, X_{II}, $(I = 1, \ldots, m)$, are nonsingular. However, if the full eigenvector matrix, X, is itself nonsingular (as it must be if H is hermitian, since then X is an orthogonal matrix, with inverse given by $X^{\dagger}$), then there is at least one partitioning of the basis functions for which X is nonsingular. A particular block X_{II} will be singular only if at least one of the eigenvectors $x_r^{(I)}$, $(r = 1, \ldots, n_I)$, is orthogonal to the basis subspace S_I.

The blocks f_{IJ} of the partitioning operator $\hat{T}$ in eq. (4.4) are not entirely independent. From the orthonormality condition (4.1), one has

$$X^{\dagger}X = \hat{X}^{\dagger}\hat{T}^{\dagger}\hat{T}\,\hat{X} = 1_n,$$

or

$$\hat{T}^{\dagger}\hat{T} = (\hat{X}\,\hat{X}^{\dagger})^{-1} = g, \qquad (4.7)$$

which is to be block diagonal. Thus the blocks of $\hat{T}$ are related by the equations,

$$g_{JK} = f_{KJ}^{\dagger} + f_{JK} + \sum_{\substack{L=1 \\ L\neq J,K}}^{m} f_{LJ}^{\dagger} f_{LK} = 0, \quad (J,K=1,\ldots,m, \; J\neq K). \qquad (4.8)$$

Since g is symmetric, eqs. (4.8) represent $\frac{1}{2}m(m-1)$ unique matrix block equations, involving the $m(m-1)$ different off-diagonal blocks of $\hat{T}$. Equations (4.8) could be used to eliminate half of the elements of the off-diagonal blocks of $\hat{T}$ in favour of the remaining half. While this procedure leads

to the very simple result $f_{12} = -f_{21}^{\dagger}$ in the 2 x 2 case, when m is larger than 2, the increase in complexity of eqs. (4.8) makes it impossible in practice to incorporate these equations explicitly into the general formalism. As a result, in what follows, a notation involving all m(m-1) off-diagonal blocks of $\hat{T}$ will be used (though eqs. (4.8) are implicit, at least when the f_{IJ} are exact). For the cases m = 3 and 4, eqs. (4.8) are examined in somewhat greater detail in Appendix 2.

Equations (4.8) express the orthogonality of eigenvectors of H belonging to different sets $X^{(J)}$ and $X^{(K)}$. Thus it is not necessary to impose them explicitly, since if H is hermitian, this orthogonality is automatic. As a result, in many applications, the increasing complexity of eqs. (4.8) with increasing m is of no practical concern.

The diagonal blocks g_I of the matrix g of (4.7) are metrics, with respect to which the corresponding truncated eigenvectors, X_{II}, are orthonormal. That is,

$$X_{II}^{\dagger}\, g_I\, X_{II} = 1_I \,, \tag{4.9}$$

where

$$g_I = (\hat{T}^{\dagger}\hat{T})_{II} = 1_I + \sum_{\substack{J=1 \\ J\neq I}}^{m} f_{JI}^{\dagger}\, f_{JI} \,. \tag{4.10}$$

The projections P_I' onto the eigenspaces S_I' can be written solely in terms of the f_{JI}, (J = 1, ..., m, J ≠ I), for each I. Using (4.7) and $P_I' = X^{(I)}X^{(I)\dagger} = \hat{T}^{(I)} g_I^{-1} \hat{T}^{(I)\dagger}$, it is seen that

$$(P_I')_{KL} = f_{KI}\, g_I^{-1}\, f_{LI}^{\dagger} \,. \tag{4.11}$$

The definition, (4.10), of g_I can be used to establish the idempotency of P_I', and if eqs. (4.8) are satisfied, it is easy to show that $P_I P_J = 0$. Equation (4.10) alone is sufficient to establish that tr $P_I' = n_I$. Thus, in expressing the projection operators, P_I', (I = 1, ..., m), in terms of the f_{JI} as in (4.11), it is necessary to constrain the m(m-1) blocks f_{JI} only if the P_I' are to be mutually orthogonal. Furthermore, this formalism provides an apparatus via eqs. (4.8), however tedious it may be, to express the P_I' in terms of a minimum number of unconstrained variables.

The minimum number of variables required to describe the eigenspaces S_1', ..., S_m', is of considerable importance here, as in the 2 x 2 case. The projection P_I' onto the eigenspace S_I' is completely specified by the n n_I complex components of the eigenvectors $X^{(I)}$, which span S_I'. However, the space S_I' is equally well spanned by any set of n_I vectors obtained from the $x_r^{(I)}$, (r = 1, ..., n_I), by a nonsingular linear transformation. Thus, there are n_I^2 complex variables in $X^{(I)}$, which serve only to specify a particular basis in S_I'. Furthermore, the orthogonality constraints in (4.1), written,

$$X^{(I)\dagger} X^{(J)} = 0, \qquad I < J,$$

can be used to eliminate a further $\sum_{I=2}^{m} n_I \sum_{J=1}^{I-1} n_J$ complex variables from all of the $X^{(I)}$. The remaining number of non-redundant and unconstrained variables required to simply specify the eigenspaces S_1', ..., S_m', is thus

$$\sum_{I,J=1}^{m} n_I n_J - \sum_{I=1}^{m} n_I^2 - \sum_{I=2}^{m} n_I \sum_{J=1}^{I-1} n_J = \sum_{I=2}^{m} \sum_{J=1}^{I-1} n_I n_J . \tag{4.12}$$

This is just the number of elements in the upper (or lower) block triangle of $\hat{T}$, and is also the number of independent variables left in $\hat{T}$ when eqs. (4.8) are explicitly incorporated into the formalism.

This multiple partitioning formalism can be defined completely from the point of view of the determination of the eigenprojections P'_I, $(I = 1, \ldots, m)$, in a manner analogous to that used in section 2.1.c. From eq. (4.11), it is easily seen that

$$f_{KL} = (P''_L)_{KL} \, (P'_L)^{-1}_{LL} . \tag{4.13}$$

One of the difficulties in manipulating quantities in this multipartitioning formalism arises from the fact that there is no counterpart here to the "pull-through" relations, (2.32), which were used extensively in the 2 x 2 case to simplify the various expressions arising. In fact, in this case, the analogue of eqs. (2.33) is

$$(X X^\dagger)_{JK} = \delta_{JK} = \sum_{I=1}^{m} f_{JI} g_I^{-1} f^\dagger_{KI} ,$$

which, for $J \neq K$, gives

$$f_{JK} g_K^{-1} = -g_J^{-1} f^\dagger_{KJ} - \sum_{I \neq J,K} f_{JI} g_I^{-1} f^\dagger_{KI} , \qquad (J,K = 1, \ldots, m, \; J \neq K). \tag{4.14}$$

Equations (4.14) are not of great use in general because of the summation term on the right hand side. In the 2 x 2 case, this term does not occur, leaving eqs. (2.32).

4.1.b The Defining Conditions on the f_{JI}.

As in the 2 x 2 case, the off-diagonal blocks of the partitioning matrix $\hat{T}$ can be determined by diagonalizing the matrix H, to obtain its eigenvectors, X, and then using eqs. (4.6) directly. However, it is again possible to formulate systems of nonlinear equations which can be solved to obtain the f_{JI} directly, thus making it unnecessary to fully diagonalize H.

Consider first the eigenvalue equation (4.1), written, using (4.4), as

$$H \hat{T} = \hat{T} \hat{X} \not{E} \hat{X}^{-1} = \hat{T} \hat{H}, \tag{4.15}$$

where the matrix $\hat{H}$, given as

$$\hat{H} = \hat{X} \not{E} \hat{X}^{-1} = \begin{bmatrix} \hat{H}_1 & & & 0 \\ & \hat{H}_2 & & \\ & & \ddots & \\ 0 & & & \hat{H}_M \end{bmatrix}, \tag{4.16}$$

is to be block diagonal. Equation (4.15) is valid only if the diagonal block part, $\hat{X}$, of X is nonsingular, which is exactly the condition that must be satisfied if the f_{JI} are to exist. The diagonal block parts of (4.15) define the operators $\hat{H}_I$ and the off-diagonal block parts provide equations for the f_{JI}. Thus, one has

$$\hat{H}_I = H_{II} + \sum_{\substack{J=1 \\ J \neq I}}^{m} H_{IJ} f_{JI} . \tag{4.17}$$

The equations determining the f_{JI} are then,

$$D_{JI}(\hat{T}) = 0 = H_{JI} + \sum_{\substack{K=1 \\ K\neq I}}^{m} H_{JK}f_{KI} - f_{JI}\hat{H}_I \tag{4.18}$$

$$= H_{JI} + \sum_{\substack{K=1 \\ K\neq I}}^{m} H_{JK}f_{KI} - f_{JI}H_{II} - f_{JI}\sum_{\substack{K=1 \\ K\neq I}}^{m} H_{IK}f_{KI} ,$$

$$(I,J = 1, \ldots, m, I\neq J).$$

Equations (4.18) consist of $\sum_{\substack{I,J=1 \\ I\neq J}}^{m} n_I n_J$ coupled nonlinear equations in the matrix elements of the f_{JI}. The solutions of these equations will automatically also satisfy eqs. (4.8) because of the hermiticity of H. Explicit incorporation of eqs. (4.8) into (4.18) could be used to reduce the total number of equations and variables by a factor of two, but at the expense of greatly increasing the complexity of the equations to be solved. In eqs. (4.18), coupling occurs only between f_{JI} in the same block column of $\hat{T}$. Thus, the $(n - n_I)n_I$ equations $D_{JI}(\hat{T}) = 0$, $(J = 1, \ldots, m, J\neq I)$, can be solved for the elements of the I^{th} block column of $\hat{T}$, namely, the f_{JI}, $(J = 1, \ldots, m, J\neq I)$, without having to determine any of the f_{KL} for $L=I$.

A somewhat different set of equations for the off-diagonal blocks of $\hat{T}$ result if the eigenvalue equation is rewritten as

$$G = \hat{T}^{\dagger}H\hat{T} = \hat{T}^{\dagger}\hat{T}\ \hat{X}\ \beta\ \hat{X}^{-1}, \tag{4.19a}$$

and

$$g = \hat{T}^{\dagger}\hat{T} = (\hat{X}\ \hat{X}^{\dagger})^{-1} , \tag{4.19b}$$

where the second equality in (4.19a) is obtained using (4.15). Both G and g are to be block diagonal, and the condition that their off-diagonal blocks vanish provides equations for the f_{JI}. Since both G and g are hermitian, the vanishing of their off-diagonal blocks can each provide only $\frac{1}{2}\sum_{I\neq J} n_I n_J$ unique equations, and thus both of (4.19a) and (4.19b) must be used together to determine all the f_{JI}. This results in a set of coupled nonlinear equations of the form

$$G_{JI}(\hat{T}) = 0 = H_{JI} + \sum_{L\neq I} H_{JL}f_{LI} + \sum_{K\neq I} f^{\dagger}_{KJ}H_{KI} + \sum_{\substack{K,L\\ \neq I,J}} f^{\dagger}_{KJ}H_{KL}f_{LI}\,, \tag{4.20}$$

and

$$g_{JI} = 0 = f^{\dagger}_{IJ} + f_{JI} + \sum_{\substack{L=1\\ L\neq J,K}}^{m} f^{\dagger}_{LJ}f_{LI}\,, \tag{4.21}$$

$$(I,J = 1, \ldots, m,\ J < I),$$

where eqs. (4.21) have appeared before in (4.8). These equations effectively couple all of the off-diagonal blocks of $\hat{T}$, and therefore, the entire matrix $\hat{T}$ must be determined at once if (4.20) and (4.21) are used. As a result, while the system of equations (4.20)-(4.21) has the same solutions as the system (4.18), the two systems must be treated quite differently from a computational point of view.

4.1.c Variational Formulation of the Equations for the f_{JI}

In this multiple partitioning procedure, it is also possible to show that eqs. (4.18), determining the f_{JI}, are equivalent

to a variational criterion, in that the vanishing of the quantities $D_{KI}(\hat{T})$, (K = 1, ..., m, K≠I), implies that the trace of the operator H over the image space of the projection operator P_I' is stationary. This stationarity implies that P_I' is an eigenprojection of H.

The algebra required to demonstrate this is considerably more tedious here than in the case of a 2 x 2 partitioning. The objective is to obtain an expression for the first order variation of the quantity,

$$E_I = \mathrm{tr}\, P_I' H = \sum_{J,K=1}^{m} \mathrm{tr}(P_I')_{KJ} H_{JK} , \tag{4.22}$$

with respect to small variations in the f_{KI}, (K=1, ..., m, K≠I). From (4.11), one obtains,

$$\delta(P_I')_{KJ} = \delta f_{KI} g_I^{-1} f_{JI}^{\dagger} + f_{KI} \delta g_I^{-1} f_{JI}^{\dagger} + f_{KI} g_I^{-1} \delta f_{JI}^{\dagger} + O(\delta^2), \tag{4.23a}$$

where,

$$\begin{aligned} \delta g_I^{-1} &= -g_I^{-1} \delta g_I g_I^{-1} + O(\delta^2) \\ &= -g_I^{-1} \sum_{L\neq I} (\delta f_{LI}^{\dagger} f_{LI} + f_{LI}^{\dagger} \delta f_{LI}) g_I^{-1} + O(\delta^2). \end{aligned} \tag{4.23b}$$

Substitution of eqs. (4.23a,b) into the equation,

$$\delta E_I = \mathrm{tr} \sum_{J,K=1}^{m} (\delta P_I')_{KJ} H_{JK} , \tag{4.24}$$

when H is independent of the f_{JI}, leads to the rather complicated expression,

$$\begin{aligned} \delta E_I = \mathrm{tr} \sum_{\substack{P=1 \\ P\neq I}}^{m} & \delta f_{PI} g_I^{-1} [(\hat{T}^{\dagger} H)_{IP} - (\hat{T}^{\dagger} H \hat{T})_{II} g_I^{-1} f_{PI}^{\dagger}] \\ & + \mathrm{tr} \sum_{\substack{P=1 \\ P\neq I}}^{m} \delta f_{PI}^{\dagger} [(H\hat{T})_{PI} - f_{PI} g_I^{-1} (\hat{T}^{\dagger} H \hat{T})_{II}] g_I^{-1} + O(\delta^2). \end{aligned} \tag{4.25}$$

The passage from (4.24) to (4.25) uses the cyclic property of the trace.

Consider the coefficient of δf_{PI} in (4.25). From (4.18), one has,

$$\begin{aligned}
&(H\hat{T})_{PI} - f_{PI}g_I^{-1}(\hat{T}^\dagger H\hat{T})_{II} \\
&\quad = D_{PI}(\hat{T}) + f_{PI}g_I^{-1}[g_I\hat{H}_I - (\hat{T}^\dagger H\hat{T})_{II}] \\
&\quad = D_{PI}(\hat{T}) + f_{PI}g_I^{-1}[(\hat{T}^\dagger\hat{T})_{II}\hat{H}_I - (\hat{T}^\dagger H\hat{T})_{II}] \\
&\quad = D_{PI}(\hat{T}) + f_{PI}g_I^{-1}\sum_{K\neq I}^{m} f_{KI}^\dagger[f_{KI}\hat{H}_I - (H\hat{T})_{KI}] \\
&\quad = D_{PI}(\hat{T}) - f_{PI}g_I^{-1}\sum_{K\neq I}^{m} f_{KI}^\dagger D_{KI}(\hat{T}) \\
&\quad = \sum_{K\neq I}(1 - P_I')_{PK}D_{KI}(\hat{T}). \qquad (4.26)
\end{aligned}$$

Consequently, the vanishing of all the $D_{KI}(\hat{T})$, (K=1, ..., m, K≠I), implies the vanishing of the coefficients of $\delta f_{PI}^\dagger$ in eq. (4.25). Since the coefficients of δf_{PI} in (4.25) are just the adjoints of those of $\delta f_{PI}^\dagger$, they vanish also, causing δE_I to vanish to first order in the infinitesimals. The vanishing of δE_I to first order implies the converse, namely, that all D_{KI}, (K=1, ..., m, K≠I), vanish. This follows from the fact that the rank of the matrix $(1 - P_I')$ must be $n-n_I$ if it is to project onto the complement of the eigenspace S_I' of H of dimension $n-n_I$. Because of this, the set of linear systems (one for each column $D_{,i}$ of $D_{,I}$), written compositely as

$$\sum_{\substack{K=1\\K\neq I}}^{m}(1 - P_I')_{PK}D_{KI}(\hat{T}) = 0\ ,$$

has only the trivial solution $D_{KI}(\hat{T}) = 0$, (K=1, ..., m, K≠I).

4.1.d Transformation of the f_{JI} Under a Change of Basis

Under a linear transformation of the basis functions $\{\phi_\mu\}$, (see section 2.1.g), the eigenvectors of H become (eq. (2.50)),

$$X' = V X$$
$$= \hat{T}' \hat{X}', \qquad (4.27)$$

where

$$\hat{T}'_{II} = 1_I ,$$

and $\qquad (I,J = 1, \ldots, m). \qquad (4.28)$

$$\hat{T}'_{JI} = f'_{JI} , \; J \neq I ,$$

The off-diagonal blocks of $\hat{T}'$ in the new basis and those of $\hat{T}$ in the old basis are related by

$$\hat{T}' = V X \hat{X}'^{-1} = X \hat{T} \hat{X} \hat{X}'^{-1}$$

from which,

$$f'_{JI} = (V\hat{T})_{JI} X_{II} X'^{-1}_{II} . \qquad (4.29)$$

But, from (4.27), one has,

$$X'_{II} = (V\hat{T})_{II} X_{II}$$

and thus,

$$f'_{JI} = (V\hat{T})_{JI} [(V\hat{T})_{II}]^{-1}$$
$$= (V_{JI} + \sum_{K \neq I} V_{JK} f_{KI})(V_{II} + \sum_{K \neq I} V_{IK} f_{KI})^{-1}, \qquad (4.30)$$

which gives the f'_{JI} solely in terms of the transformation coefficients V_{IJ}, and the f_{JI} in the old basis.

Again, the transformation for the f_{JI} under a linear basis change is complicated and nonlinear in both the coefficients of the transformation and the old variables. While

such a complicated transformation is doubtless disadvantageous under some circumstances, it can be also usefully exploited, as pointed out in section 2.1.g.

Note also, that, writing,

$$(V_{II} + \sum_{K\neq I} V_{IK}f_{KI})^{-1} = (1_I + \sum_{K\neq I} V_{II}^{-1}V_{IK}f_{KI})^{-1}V_{II}^{-1}$$

$$= (1_I - \sum_{K\neq I} V_{II}^{-1}V_{IK}f_{KI} + O(f^2))V_{II}^{-1} ,$$

it is seen that

$$f'_{JI} = V_{JI}V_{II}^{-1} + \sum_{K\neq I} (V_{JK} - V_{JI}V_{II}^{-1}V_{IK})f_{KI}V_{II}^{-1} + O(f^2), \tag{4.31}$$

which, for small f, is nearly linear, but not homogeneous.

4.2 Effective Operators

4.2.a Basic Definitions

Like the 2 x 2 partitioning formalism, one of the primary applications of this multiple partitioning formalism is in the construction of effective operators. Such operators would be defined in one of the subspaces, S_I, of the full basis space, but would have as eigenvalues, a particular subset of the eigenvalues of the original operator H in the full space. Those eigenvalues correspond to the eigenvectors of H spanning the space S_I'. Since they are restricted to the subspaces in which the effective operators are defined, the corresponding eigenvectors of these operators are simple, or orthonormalized, truncations of eigenvectors of the operator H in the full space. However, given the matrix $\hat{T}$ and the eigenvectors of the effective operators, those of the original operator H in the full space can be obtained straightforwardly.

The types of effective operators arising here are analogous to those defined previously in the 2 x 2 case. The simplest set of effective operators has been defined already in eq. (4.14). These operators,

$$\hat{H}_I = H_{II} + \sum_{\substack{J=1 \\ J\neq I}}^{m} H_{IJ} f_{JI} , \qquad (I = 1, \ldots, m), \qquad (4.32)$$

are defined in the corresponding subspaces S_I, and have the eigenvalue equations,

$$\hat{H}_I X_{II} = \xi^{(I)} X_{II} , \qquad (I = 1, \ldots, m), \qquad (4.33)$$

as seen from eq. (4.13). Here, $\mathcal{E}^{(I)}$ is the I^{th} diagonal block of the matrix of eigenvalues, $\mathcal{E}$, in eq. (4.1). The operators $\hat{H}_I$ are non-selfadjoint, in general. However, their eigenvectors, X_{II}, are orthonormal with respect to the non-unit metrics g_I, according to eqs. (4.9).

The corresponding basic set of self-adjoint effective operators are those defined by the diagonal blocks of eq. (4.19a), namely,

$$G_I = (\hat{T}^\dagger H \hat{T})_{II} , \tag{4.34}$$

with the eigenvalue equation,

$$G_I X_{II} = g_I X_{II} \mathcal{E}^{(I)} , \quad (I = 1, \ldots, m), \tag{4.35}$$

where

$$g_I = (\hat{T}^\dagger \hat{T})_{II} . \tag{4.36}$$

In detail, one has,

$$G_I = H_{II} + \sum_{J \neq I} (f^\dagger_{IJ} H_{JI} + H_{IJ} f_{JI}) + \sum_{\substack{J \neq I \\ K \neq I}} f^\dagger_{JI} H_{JK} f_{KI} . \tag{4.37}$$

If eqs. (4.16) are satisfied, this can also be written as,

$$\begin{aligned} G_I &= \sum_J (\hat{T}^\dagger)_{IJ} (H\hat{T})_{JI} = \sum_J (\hat{T}^\dagger)_{IJ} (\hat{T}\hat{H})_{JI} \\ &= (\hat{T}^\dagger \hat{T})_{II} \hat{H}_I = g_I \hat{H}_I . \end{aligned} \tag{4.38}$$

As in the 2 x 2 case, other sets of self-adjoint effective operators can be obtained by orthogonalizing the truncated eigenvectors by other procedures. Lowdin's (1970) symmetrical orthogonalization (see section 2.2.a) leads to the set of orthonormal eigenvectors,

$$C_{II} = g_I^{\frac{1}{2}} X_{II} , \qquad (I = 1, \ldots, m), \tag{4.39}$$

where $C_{II}^{\dagger} C_{II} = X_{II}^{\dagger} g_I X_{II} = 1_I$, by eq. (4.9). These new vectors satisfy the eigenvalue equation,

$$\tilde{H}_I \, C_{II} = C_{II} \, \varepsilon^{(I)}, \tag{4.40}$$

where

$$\tilde{H}_I = g_I^{\frac{1}{2}} \hat{H}_I \, g_I^{-\frac{1}{2}} , \tag{4.41a}$$

$$= g_I^{-\frac{1}{2}} G_I \, g_I^{-\frac{1}{2}} , \tag{4.41b}$$

the equivalences here being based implicitly on the assumption that the partitioning operator $\hat{T}$ used is known exactly.

Effective operators in the spaces $S_1, \ldots, S_m$, can be defined for any other operator in the full space, using the definitions (4.4) and (4.39). In particular, matrix elements of some operator O can be written,

$$X^{(I)\dagger} O \, X^{(I)} = X_{II}^{\dagger} \, \tilde{O}_I \, X_{II} . \tag{4.42a}$$

Here

$$\tilde{O}_I = (\hat{T}^{\dagger} O \hat{T})_{II} , \tag{4.42b}$$

is an operator confined to the subspace S_I, but possessing the same expectation values with respect to the X_{II} as the original operator does with respect to the full eigenvectors $X^{(I)}$. A second type of effective operator,

$$\tilde{\tilde{O}}_I = g_I^{-\frac{1}{2}} \tilde{O}_I \, g_I^{-\frac{1}{2}} , \tag{4.43}$$

will give the same matrix elements with respect to the orthonormalized vectors C_{II}, of (4.39), as the original operator O does with respect to the $X^{(I)}$. Here $\tilde{O}_I$ has the same form

in O as the operator G_I, eq. (4.34), does in H, and $\tilde{\tilde{O}}_I$ is of the same form as $\tilde{H}_I$ given by (4.41b).

4.2.b Eigenvalues and Eigenvectors of the Effective Operators

Up to this point, the m x m partitioning formalism has been presented almost totally in matrix notation. It is instructive, however, to re-examine some of the relationships quoted previously, from the point of view of the actual eigenfunctions of the operator H, and derived effective operators. The eigenvalue equation, (4.1), for H is written as

$$\begin{aligned} H\psi_i &= \xi_i \psi_i \,, \\ \langle \psi_i \mid \psi_j \rangle &= \delta_{ij} \,, \end{aligned} \qquad (i,j = 1, \ldots, n). \qquad (4.44)$$

If a partitioning of the basis space into m subspaces S_1, ..., S_m, is carried out, these eigenfuctions of H can be written as a sum of parts,

$$\psi_i = \psi_{i1} + \psi_{i2} + \ldots + \psi_{im} = \sum_{J=1}^{m} \psi_{iJ}, \qquad (4.45)$$

with ψ_{iJ} being the part of ψ_i lying in the subspace S_J. The partitioning of the eigenvectors of H into m sets, spanning eigenspaces S_1', ... S_m', merely divides the ψ_i into m sets -- the notation $\psi_i^{(J)}$, (J=1, ...,m; i=1, ..., n_I), now denoting the i^{th} member of the j^{th} such set. The basic equations, (4.6), of the partitioning formalism then are,

$$\psi_{iK}^{(J)} = f_{KJ}\, \psi_{iJ}^{(J)}. \qquad (4.46)$$

This means that the eigenfunctions of H in the full space can be written as

$$\psi_i^{(J)} = \sum_{K=1}^{m} \psi_{iK}^{(J)} = \psi_{iJ}^{(J)} + \sum_{K\neq J}^{m} f_{KJ}\, \psi_{iJ}^{(J)}$$

$$= [1_J + \sum_{K\neq J}^{m} f_{KJ}]\, \psi_{iJ}^{(J)}. \qquad (4.47)$$

In the notation used in this section, the symbol f_{KJ} represents an embedding of the mapping f_{KJ}, $(S_J \rightarrow S_K)$, in the whole n-dimensional basis space.

It is a simple matter to write down the eigenvalue equations for the effective operators in this notation. The counterpart of eq. (4.33) for the operators $\hat{H}_I$ is,

$$\hat{H}_I\, \psi_{iI}^{(I)} = \varepsilon_i^{(I)}\, \psi_{iI}^{(I)}, \qquad (i,j=1,\ldots,n_I),$$
$$\langle \psi_{iI}^{(I)} | g_I | \psi_{jI}^{(I)} \rangle = \delta_{ij}, \qquad (I = 1,\ldots,m). \qquad (4.48)$$

The eigenvalue equation, (4.35), for the operators G_I becomes,

$$G_I\, \psi_{iI}^{(I)} = \varepsilon_i^{(I)} g_I\, \psi_{iI}^{(I)}, \quad (i=1,\ldots,n_I;\ I=1,\ldots,m), \qquad (4.49)$$

with the same orthonormality condition as in (4.48). The eigenfunctions obtained from the $\psi_{iI}^{(I)}$ by the symmetric orthogonalization procedure are given by,

$$\chi_{iI}^{(I)} = g_I^{\frac{1}{2}}\, \psi_{iI}^{(I)}. \qquad (4.50)$$

Thus, the eigenvalue equation, (4.40), for the operators $\tilde{H}_I$ is

$$\tilde{H}_I\, \chi_{iI}^{(I)} = \varepsilon_i^{(I)} \chi_{iI}^{(I)}, \qquad (i,j=1,\ldots,n_I;\ I=1,\ldots,m).$$
$$\langle \chi_{iI}^{(I)} | \chi_{jI}^{(I)} \rangle = \delta_{ij}, \qquad (4.51)$$

The full eigenfunctions of H are given in terms of the eigenfunctions, (4.50), of the operators $\tilde{H}_I$, by

$$\psi_i^{(I)} = (1 + \sum_{\substack{K=1 \\ K\neq I}}^{m} f_{KI}) g_I^{-\frac{1}{2}} \chi_{iI}^{(I)}. \tag{4.52}$$

Finally, for the eigenprojections, P_I', it is seen that,

$$\begin{aligned} P_I' &= \sum_{i=1}^{n_I} |\psi_i^{(I)}\rangle\langle\psi_i^{(I)}| \\ &= \sum_{i=1}^{n_I} (1 + \sum_{\substack{K=1 \\ K\neq I}}^{m} f_{KI}) |\psi_{iI}^{(I)}\rangle\langle\psi_{iI}^{(I)}| (1 + \sum_{\substack{K'=1 \\ K'\neq I}}^{m} f_{K'I}^{\dagger}) \\ &= (1 + \sum_{\substack{K=1 \\ K\neq I}}^{m} f_{KI}) g^{(I)} (1 + \sum_{\substack{K'=1 \\ K'\neq I}}^{m} f_{K'I}^{\dagger}), \end{aligned} \tag{4.53}$$

where

$$g^{(I)} = \sum_{i=1}^{n_I} |\psi_{iI}^{(I)}\rangle\langle\psi_{iI}^{(I)}|, \tag{4.54}$$

defines an embedding of the metric g_I in the full n-dimensional basis space.

4.3 Generalization to a Non-orthonormal Basis

The generalization of the multiple partitioning formalism to the case of a non-orthonormal basis is straightforward. The eigenvalue equation is now

$$H X = X S E, \tag{4.55a}$$

with

$$X^{\dagger} S X = 1_n, \tag{4.55b}$$

as in eq. (2.90), where S is the matrix of overlap integrals of the basis functions. The set of basis functions, and the eigenvectors, X, of H, are each partitioned into m subsets, exactly as described in section 4.1.a, making it possible to write the eigenvector matrix X in the partitioned form (4.2). The full matrix X can then be written in terms of some matrix $\hat{T}$, and the diagonal block part, $\hat{X}$, of X, as given in eq. (4.4),

$$X = \hat{T} \hat{X}. \tag{4.56}$$

The matrix elements of $\hat{T}$ are given here also by eqs. (4.5). The conditions under which eqs. (4.56) will be valid are identical to those under which (4.4) are valid, namely, that the partitioning of the basis functions must be so defined that $\hat{X}$ is invertible. While X is no longer a unitary matrix, the hermiticity of H implies that the columns of X are linearly independent (except possibly if S is singular) and thus there will be at least one partitioning of the basis functions for which $\hat{X}$ is invertible.

The m(m-1) off-diagonal blocks of the matrix $\hat{T}$ are not

all independent, as can be demonstrated using the orthogonality condition, (4.55b). The analogue of eq. (4.7) is

$$\hat{T}^{\dagger}S\hat{T} = (\hat{X}\,\hat{X}^{\dagger})^{-1} = g, \tag{4.57}$$

which must be block diagonal if the orthogonality condition is to be satisfied. This implies the equations,

$$g_{IJ} = S_{IJ} + \sum_{\substack{L=1 \\ L\neq J}}^{m} S_{IL}f_{LJ} + \sum_{\substack{L=1 \\ L\neq I}}^{m} f^{\dagger}_{LI}S_{LJ} + \sum_{\substack{K,L=1 \\ K\neq I, L\neq J}}^{m} f^{\dagger}_{KI}S_{KL}f_{LJ}$$

$$= 0, \qquad (I,J = 1, \ldots, m;\ I\neq J). \tag{4.58}$$

These equations could be used in specific cases to eliminate half of the elements of the off-diagonal blocks of $\hat{T}$ from the formalism. However, they are considerably more complicated than the corresponding equations, (4.8), for an orthonormal basis. Therefore, the remarks following eqs. (4.8) apply here with even greater emphasis. From a practical point of view, such an elimination procedure is not to be recommended.

The diagonal blocks of g, given by

$$g_I = S_{II} + \sum_{L\neq I}^{m} S_{IL}f_{LI} + \sum_{L\neq I}^{m} f^{\dagger}_{LI}S_{LI} + \sum_{L,K\neq I}^{m} f^{\dagger}_{KI}S_{KL}f_{LI}, \tag{4.59}$$

serve as metrics for the truncated eigenvectors X_{II}, as indicated in eq. (4.9). Because of the explicit presence of the overlap matrix, S, the leading term of g_I here is S_{II}, rather than a unit matrix of the same dimension, as occurs with an orthonormal basis, eq. (4.10)

The defining conditions on the off-diagonal blocks of $\hat{T}$ are obtained in a manner similar to that employed with an

orthonormal basis. Direct substitution of (4.56) into (4.55), and use of the fact that $\hat{X}$ is invertible leads to

$$H\hat{T} = S\hat{T}(\hat{X}\,\xi\,\hat{X}^{-1}) = S\hat{T}\hat{H}, \tag{4.60}$$

where $\hat{H}$ is to be block diagonal, as in eq. (4.14). The diagonal blocks of (4.60) give $\hat{H}$ in terms of H, S, and $\hat{T}$, as

$$\begin{aligned}\hat{H}_I &= [(S\hat{T})_{II}]^{-1}(H\hat{T})_{II} \\ &= [S_{II} + \sum_{\substack{J=1 \\ J\neq I}}^{m} S_{IJ}f_{JI}]^{-1}[H_{II} + \sum_{\substack{J=1 \\ J\neq I}}^{m} H_{IJ}f_{JI}].\end{aligned} \tag{4.61}$$

If the overlap matrix is a unit matrix, the inverse matrix in (4.61) reduces to an identity matrix, and eq. (4.17) for an orthonormal basis is recovered. The expression, (4.61), for the effective operators $\hat{H}_I$, (I=1, ..., m), are of the same form as eqs. (2.95) and (2.96), given for the operators $\hat{H}_A$ and $\hat{H}_B$ in the 2 x 2 partitioning formalism. From (4.60), it is seen that the eigenvalue equations for these $\hat{H}_I$ are given by

$$\hat{H}_I X_{II} = X_{II}\,\xi^{(I)}, \quad (I=1, \ldots, m), \tag{4.62}$$

exactly as in (4.33) for an orthonormal basis. As pointed out in chapter 2, however, a new set of effective operators $\hat{H}'_I$ could be defined by

$$\hat{H}'_I = H_{II} + \sum_{\substack{J=1 \\ J\neq I}}^{m} H_{IJ}f_{JI}, \quad (I=1, \ldots, m), \tag{4.63}$$

which is identical to (4.17), but leads to an effective eigenvalue equation of the form

$$\hat{H}_I X_{II} = \hat{S}_I X_{II}\,\xi^{(I)}, \tag{4.64}$$

where

$$\hat{S}_I = S_{II} + \sum_{\substack{J=1 \\ J\neq I}}^{m} S_{IJ} f_{JI} \,, \tag{4.65}$$

can be regarded as an effective overlap matrix. Equations (4.64) and (4.62) are simply restatements of the same eigenvalue equation, and one typical way of actually solving (4.64) is by using (4.62) as an intermediate.

Defining conditions on the f_{JI} are now obtained from the off-diagonal blocks of eq. (4.60), after substitution of (4.61). The result is

$$\begin{aligned} D_{IJ}(\hat{T}) &= (H\hat{T})_{IJ} - (S\hat{T})_{IJ}\hat{H}_J \\ &= H_{IJ} + \sum_{\substack{K=1 \\ K\neq I}}^{m} H_{IK} f_{KJ} - \left(S_{IJ} + \sum_{\substack{K=1 \\ K\neq I}}^{m} S_{IK} f_{KJ}\right)\hat{H}_J \\ &= H_{IJ} + \sum_{\substack{K=1 \\ K\neq I}}^{m} H_{IK} f_{KJ} - \left(S_{IJ} + \sum_{\substack{K=1 \\ K\neq J}}^{m} S_{IK} f_{KJ}\right)\left(S_{JJ} + \sum_{\substack{K=1 \\ K\neq J}}^{m} S_{JK} f_{KJ}\right)^{-1} \\ &\quad \times \left(H_{JJ} + \sum_{\substack{K=1 \\ K\neq J}}^{m} H_{JK} f_{KI}\right) = 0. \end{aligned} \tag{4.66}$$

Clearly, the presence of the overlap matrix severely complicates the determination of the f_{JI}. It is seen that these equations still retain the property of being separately soluble for individual block columns of $\hat{T}$. Despite the complexity of eqs. (4.66), it is still possible to devise efficient iterative schemes for their solution.

An alternative set of defining conditions, analogous to eqs. (4.19), are obtained by premultiplying the eigenvalue

equation, (4.60), by $\hat{T}^\dagger$, to give

$$G = \hat{T}^\dagger H \hat{T} = (\hat{T}^\dagger S \hat{T})\, \hat{X}\, \mathcal{E}\, \hat{X}^{-1}, \tag{4.67a}$$

where

$$\hat{T}^\dagger S \hat{T} = (\hat{X}\, \hat{X}^\dagger)^{-1} = g, \tag{4.67b}$$

must be block diagonal. Thus G itself must be block diagonal and its diagonal blocks form a second set of effective operators, when this is so, with the eigenvalue equations

$$G_I\, X_{II} = g_I\, X_{II}\, \mathcal{E}^{(I)}, \quad (I=1, \ldots, m), \tag{4.68}$$

This is identical in form to eqs. (4.35), the effects of the presence of the overlap matrix being buried in the detailed form of g_I. The operator G_I here is identical in form with the corresponding quantity for an orthonormal basis. When eqs. (4.66) are satisfied, implying that the second equality in (4.67a) is satisfied, it is seen that

$$G_I = g_I \hat{H}_I, \qquad (I = 1, \ldots, m). \tag{4.69}$$

The matrices f_{JI} can therefore also be determined by the condition that the off-diagonal blocks of G and g vanish, eqs. (4.20) and (4.58). Since both G and g are hermitian, both eqs. (4.67a) and (4.67b) are required to determine all the f_{JI}. These equations effectively couple all of the off-diagonal blocks of $\hat{T}$, which must therefore be completely determined simultaneously, rather than block column-wise, as is possible using (4.66). This drawback in using eqs. (4.67) is probably more than compensated for by the much simpler form of these equations.

Two of the three types of effective operators which were defined for an orthonormal basis have been introduced above in eqs. (4.62) and (4.68) for the present case. The third type of effective operator, namely, the $\tilde{H}_I$, (I = 1, ..., m), given by eqs. (4.41), are identical in form here because the overlap matrix does not appear explicitly in their definitions. The corresponding eigenvalue equations are given by (4.40) with the eigenfunctions of $\tilde{H}_I$ being related to those of $\hat{H}_I$ and G_I by (4.39).

4.4 Practical Considerations

4.4.a Alternative Formulas

In the development of iterative procedures for the determination of the f_{JI} either from eqs. (4.18) or (4.20), (4.21), or their counterparts in the case of a non-orthonormal basis, it is necessary to take into account the manner in which a given f-dependent quantity is evaluated. This point has been explored in detail in section 3.1 for a 2 x 2 partitioning. The purpose of this section is to outline the corresponding (more complicated) results for a multiple partitioning formalism.

Consider first the case of an orthonormal basis. The operators G_I are given by eq. (4.37) as

$$G_I = (\hat{T}^\dagger H \hat{T})_{II} \,. \tag{4.70}$$

When the $D_{IJ}(\hat{T})$ are not all zero, it is possible to distinguish two distinct forms for the operators $\hat{H}_I$, namely,

$$\hat{H}_I^{(1)} = H_{II} + \sum_{\substack{J=1 \\ J\neq I}}^{m} H_{IJ} f_{JI}, \qquad (I=1, \ldots, m), \tag{4.71}$$

and in eq. (4.17), and,

$$\hat{H}_I^{(2)} = g_I^{-1} G_I \,, \qquad (I = 1, \ldots, m), \tag{4.72}$$

from eq. (4.69). It is a relatively simple matter to demonstrate that (see Appendix 3),

$$\hat{H}_I^{(2)} = \hat{H}_I^{(1)} + g_I^{-1} \sum_{\substack{J=1 \\ J\neq I}}^{m} f_{JI}^\dagger D_{JI}^{(1)} \,, \quad (I = 1, \ldots, m), \tag{4.73}$$

where the $D_{JI}^{(1)}$, defined below, are essentially the conditions

(4.18), defining the f_{JI}. Thus, if all $D_{JI}^{(1)}$, (J=1, ..., m; J≠I), vanish for a particular value of I, the two operators $\hat{H}_I^{(1)}$ and $\hat{H}_I^{(2)}$ (for that particular value of I) are identical.

Given the two forms of the operator $\hat{H}_I$, it is possible to write the first form of the defining conditions, (4.18), on the f_{JI} in one of two alternative ways, namely,

$$D_{JI}^{(1)}(\hat{T}) = H_{JI} + \sum_{K\neq I} H_{JK}f_{KI} - f_{JI}\hat{H}_I^{(1)}, \qquad (4.74)$$

and

$$D_{JI}^{(2)}(\hat{T}) = H_{JI} + \sum_{K\neq I} H_{JK}f_{KI} - f_{JI}\hat{H}_I^{(2)}. \qquad (4.75)$$

These two forms are equivalent in the sense that they both have the same zeros, by virtue of (4.73). In detailed form, they are quite different, however, away from a zero. Substitution of (4.73) into (4.75) gives

$$D_{JI}^{(2)}(\hat{T}) = D_{JI}^{(1)}(\hat{T}) - f_{JI}g_I^{-1}\sum_{L\neq I} f_{LI}^{\dagger}D_{LI}^{(1)}(\hat{T}), \qquad (4.76)$$

verifying that (4.74) and (4.75) are only equal where they vanish, and that they do have all their zeros in common. Equation (4.76) is the generalization of eq. (3.6) in the 2 x 2 partitioning formalism.

In the 2 x 2 case, it was shown that the conditions $D_{JI}^{(2)} = 0$ also arose out of the requirement that $\hat{T}^{-1}H\hat{T}$ be block diagonal. In the present multiple partitioning case, this is no longer true, because of the increased complexity of the orthogonality conditions, (4.8). Since $\hat{T}^{\dagger}\hat{T} = g$, one has

$$\hat{T}^{-1} = g^{-1}\hat{T}^{\dagger}, \qquad (4.77)$$

and therefore,

$$\hat{T}^{-1}H\hat{T} = g^{-1}\hat{T}^{\dagger}H\hat{T} = g^{-1}G. \tag{4.78}$$

In the 2 x 2 case, g could easily be made block diagonal, and in so doing, G_{BA} became identical to $D^{(1)}(f)$, leading to eq. (3.6). In the present case, the general block diagonalization of g is not possible, and therefore, a result similar to that of the former case cannot be obtained.

Here, as in the 2 x 2 case, it is possible to calculate the operators $\tilde{H}_I$ using one of three different formulas, in terms of $\hat{H}_I^{(1)}$, $\hat{H}_I^{(2)}$, and G_I, respectively. The first form, in terms of $\hat{H}_I^{(1)}$, as indicated in (4.41a), is not of practical interest, because it represents only a partial re-normalization of the truncated eigenvectors. The latter two formulas,

$$\tilde{H}_I = g_I^{\frac{1}{2}}\hat{H}_I^{(2)}g_I^{-\frac{1}{2}} \tag{4.79a}$$

$$= g_I^{-\frac{1}{2}}\, G_I\, g_I^{-\frac{1}{2}}, \tag{4.79b}$$

are effectively identical from a practical point of view.

Consider now the case of a non-orthonormal basis. Many of the results presented earlier, for the simple 2 x 2 partitioning, have analogues in the present m x m partitioning formalism which are too complicated to be likely to be useful. Again, it is useful to distinguish two sets of effective operators of the $\hat{H}_I$-type, namely,

$$\hat{H}_I^{(1)} = \left[S_{II} + \sum_{J\neq I} S_{IJ}f_{JI}\right]^{-1}\left[H_{II} + \sum_{J\neq I} H_{IJ}f_{JI}\right], \tag{4.80}$$

and,

$$\hat{H}_I^{(2)} = g_I^{-1} G_I$$
$$= g_I^{-1}[H_{II} + \sum_{J\neq I}(f_{JI}^{\dagger} H_{JI} + H_{IJ} f_{JI}) + \sum_{\substack{J\neq I \\ K\neq I}} f_{JI}^{\dagger} H_{JK} f_{KI}]. \qquad (4.81)$$

$I = 1, \ldots, m$, in both cases. The relationship between these two sets of operators is found to be (see Appendix 3),

$$\hat{H}_I^{(2)} = \hat{H}_I^{(1)} + g_I^{-1} \sum_{J\neq I} f_{JI}^{\dagger} D_{JI}^{(1)}, \qquad (4.82)$$

exactly as for an orthonormal basis.

The two types of effective operators, $\hat{H}_I^{(1)}$ and $\hat{H}_I^{(2)}$, lead to two different sets of defining conditions for the f_{JI}, of the type (4.18). They are written

$$D_{JI}^{(1)}(\hat{T}) = H_{JI} + \sum_{K\neq I} H_{JK} f_{KI} - (S_{JI} + \sum_{K\neq I} S_{JK} f_{KI}) \hat{H}_I^{(1)}, \qquad (4.83)$$

and,

$$D_{JI}^{(2)}(\hat{T}) = H_{JI} + \sum_{K\neq I} H_{JK} f_{KI} - (S_{JI} + \sum_{K\neq I} S_{JK} f_{KI}) \hat{H}_I^{(2)}, \qquad (4.84)$$

where, in both (4.83) and (4.84), $I, J = 1, \ldots, m$, $I \neq J$. Direct application of eq. (4.82) to eq. (4.84) gives the relationship between these two types of quantities,

$$D_{JI}^{(2)} = D_{JI}^{(1)} - (S_{JI} + \sum_{K\neq I} S_{JK} f_{KI}) g_I^{-1} \sum_{K\neq I} f_{KI} D_{KI}^{(1)}. \qquad (4.85)$$

Equation (4.85) is the generalization of eq. (3.15c) to the multiple partitioning case. A generalization of (3.15b) can also be obtained here, but only at the expense of a great deal of tedious algebra. The final result contains many additional terms not appearing in (3.15b), and thus is not likely to be useful.

4.4.b Implications of Inexact Solutions

The purpose of this section is to examine the errors in the effective operators $\hat{H}_I$, G_I, and $\tilde{H}_I$, arising from the use of inexact f_{JI}, (I,J = 1, ..., m, I≠J). These results closely correspond to those given in section 3.2, and thus, only a brief summary is required here.

Consider an approximate solution to eqs. (4.18) or (4.20), (4.21), written as

$$f_{IJ}^{approx} = f_{IJ} + \delta f_{IJ} , \quad (I,J = 1, \ldots, m, I \neq J), \tag{4.86}$$

where the f_{IJ}, here, are to represent an exact solution to those equations. The error δf_{IJ} in f_{IJ} gives rise to errors in the effective operators $\hat{H}_I$, G_I, and $\tilde{H}_I$. The only complicating factor here, compared to a 2 x 2 partitioning, is that the errors in several f_{IJ} will contribute to the overall error in a given effective operator.

From eq. (4.37), it is seen that

$$\delta G_I = \sum_{J \neq I} (\delta f_{JI}^{\dagger} f_{JI} \hat{H}_I + \hat{H}_I^{\dagger} f_{JI}^{\dagger} \delta f_{JI}) + O(\delta^2). \tag{4.87}$$

Similarly, from (4.71),

$$\delta \hat{H}_I^{(1)} = \sum_{J \neq I} H_{IJ} \delta f_{JI} + O(\delta^2). \tag{4.88}$$

Using the equation,

$$(g_I + \delta g_I)(\hat{H}_I^{(2)} + \delta \hat{H}_I^{(2)}) = G_I + \delta G_I ,$$

obtained from the definition (4.72) of $\hat{H}_I^{(2)}$, the error in $\hat{H}_I^{(2)}$ induced by the above errors in the f_{IJ} is given by

$$\delta\hat{H}_I^{(2)} = g_I^{-1}[\delta G_I - \delta g_I \hat{H}_I^{(2)}] + O(\delta^2)$$
$$= g_I^{-1} \sum_{J\neq I} [\hat{H}_I^{\dagger}(f_{JI}^{\dagger}\delta f_{JI}) - (f_{JI}^{\dagger}\delta f_{JI})\hat{H}_I] + O(\delta^2). \quad (4.89)$$

In obtaining eq. (4.89), eq. (4.10) has been used to write

$$\delta g_I = \sum_{J\neq I} (\delta f_{JI}^{\dagger} f_{JI} + f_{JI}^{\dagger}\delta f_{JI}) + O(\delta^2).$$

Finally, using $\delta g_I^{-\frac{1}{2}} = -g_I^{-\frac{1}{2}}\,\delta g_I^{\frac{1}{2}}\,g_I^{-\frac{1}{2}} + O(\delta^2)$, it can be shown that

$$\delta\tilde{H}_I = \delta(g_I^{\frac{1}{2}}\,\hat{H}_I^{(2)} g_I^{-\frac{1}{2}})$$
$$= [\tilde{H}_I,\ g_I^{-\frac{1}{2}} \sum_{J\neq I} f_{JI}^{\dagger}\delta f_{JI} g_I^{-\frac{1}{2}} - \delta g_I^{\frac{1}{2}} g_I^{-\frac{1}{2}}] + O(\delta^2), \quad (4.90)$$

using eq. (4.79). A similar, though not identical, form for $\delta\tilde{H}_I$ is obtained if the formula (4.79) for $\tilde{H}_I$ in terms of G_I is used.

In eqs. (4.87)-(4.90), all quantities on the right hand sides which are not incremental, are exact. The formulas (4.87)-(4.90) exhibit substantial similarities with the corresponding formulas for a 2 x 2 partitioning. In fact, in most cases, it is seen that terms involving the single block f in the 2 x 2 case here contain sums over similar terms for each block f_{JI} in the I^{th} block column of $\hat{T}$.

As for a 2 x 2 partitioning, the error expressions (4.87)-(4.90) are all first order in the errors δf_{JI}. However, the errors δG_I of (4.87), $\delta\hat{H}_I^{(2)}$ of (4.89), and $\delta\tilde{H}_I$ of (4.90), have a vanishing expectation value in first order with respect to the exact eigenvectors of these operators. The error $\delta\hat{H}_I^{(1)}$

of (4.88) does not have such an expectation value which vanishes in first order in the errors in the f_{JI}. For this reason, the $\hat{H}_I^{(1)}$ can be considered as inherently less accurate than the former three effective operators, when inexact values for the elements of $\hat{T}$ are used.

CHAPTER 5

EXACT DETERMINATION OF $\hat{T}$

"'Why,' said the Dodo, 'the best way to explain it is to do it'. (And, as you might like to try the thing yourself some winter day, I will tell you how the Dodo managed it.)"
(Alice's Adventures in Wonderland, Lewis Carroll)

Several sets of simultaneous non-linear equations, defining the off-diagonal blocks of the partitioning operator $\hat{T}$, have been derived. In general, these equations can only be solved numerically. Some numerical iterative techniques are described in this chapter, and some assessment of their efficiency and reliability is made. A number of additional ways of defining $\hat{T}$ are also discussed, together with the numerical procedures they suggest.

The methods described can be applied in a wide range of quantum mechanical calculations. They are particularly useful when only a small number of the eigenvalues and eigenvectors, or only a projection onto a whole eigenspace (rather than the individual eigenvectors) of a hermitian operator are desired. The techniques described below represent new and practical approaches to such calculations.

5.1 The Calculation of a Few Eigenvalues of a Large Hermitian Matrix

The choice of algorithms to determine f depends to some extent on the nature of the applications which are anticipated. One important application of the methods of this chapter is the calculation of a small number of the lowest (or highest) eigenvalues, and corresponding eigenvectors, of a large hermitian matrix. Such applications arise in the determination of electronic wavefunctions for the lower lying energy levels of atoms and molecules in large scale configuration interaction calculations, and in a variety of calculations in applied mathematics and physics. The matrices arising may have dimensions up to tens of thousands (Roos, 1975).

Algorithms for the partial diagonalization of large matrices must satisfy a number of conditions to be practical. With a matrix so large that it must be stored on some auxiliary device, rather than in the central computer memory, only small sections are available to random access at one time. Techniques which involve many successive modifications to the original matrix thus become very inefficient, and their vulnerability to significant cumulative round-off error increases with the dimension of the matrix. Further, in techniques in which the entire matrix must be brought to some standard form before the calculation of a single eigenvalue and eigenvector, the calculation of a small number of eigenvalues and eigenvectors may require nearly as much work as the calculation of all of them.

In iterative techniques, on the other hand, these difficulties can be minimized. With proper organization, small sections of the matrix can be used sequentially, and the work per iteration can be made proportional to the actual number of eigenvalues being calculated. For large matrices, this work should then also be roughly proportional to the square of the dimension of the matrix, rather than the third power.

Most iterative techniques now available[1] for the partial diagonalization of large matrices are based on the calculation of successive corrections to some starting vector, to obtain a sequence of vectors converging to a single eigenvector. Since these techniques typically use the maximization or minimization of the Rayleigh quotient with respect to the approximate eigenvector as the criterion for the calculation of the appropriate corrections, the single eigenvector obtained usually corresponds to the largest or smallest eigenvalue of the matrix. To find other eigenvalues and eigenvectors of the matrix, the same procedure is repeated, but convergence onto previously calculated eigenvectors is prevented using one of several techniques (Shavitt, 1973).

A different approach to the partial diagonalization of a large hermitian matrix by iterative methods, is provided by this eigenvalue independent partitioning formalism. If a

[1]See Shavitt et. al., (1973); Shavitt, (1970); Nesbet, (1965); and Feler, (1974).

matrix f, corresponding to the uncoupling of an n_A-dimensional subspace spanned by the desired eigenvectors, can be determined, then the calculation of these n_A eigenvalues and eigenvectors reduces to the construction and solution of an n_A-dimensional eigenvalue equation, to get truncated eigenvectors X_{AA}, only, followed by the matrix multiplication,

$$X^{(A)} = \begin{bmatrix} 1_A \\ f \end{bmatrix} X_{AA},$$

(see eq. (2.3)). The n_A eigenvalues and eigenvectors are determined simultaneously, and thus, no error prone and time consuming deflation or eigenvalue shifting procedures need be employed to obtain eigenvalues greater than the smallest one. If the accuracy of the elements of f is uniform, the accuracy of the n_A eigenvalues and eigenvectors calculated should be uniform, rather than slowly deteriorating in the order in which they are calculated. These methods are especially useful when the desired eigenvalues are nearly, or exactly, equal, but well separated from the remaining eigenvalues of the matrix. Existing procedures which consist of successive calculation of the desired eigenvalues, one at a time, may perform very poorly in such a situation.

The major part of the procedures described here involves the calculation of f. In developing suitable algorithms for the iterative determination of f, two criteria were satisfied whenever possible, namely, that the amount of computation per iteration be proportional to $n_A n_B^2$, and that the columns or

rows of H_{BB} be required only sequentially. With $n_B >> n_A$, manipulations of $n_B \times n_B$ matrices (such as inversion, or the evaluation of the product of two of them) require of the order of n_B^3 computational operations, which is of the same order as the amount of work required to completely diagonalize the entire matrix by traditional methods.

To maximize their accuracy, given f to some accuracy, the eigenvalues and eigenvectors should be computed from one of the effective operators $\hat{H}_A^{(2)}$, G_A, or $\tilde{H}_A$, rather than from $\hat{H}_A^{(1)}$, even though the latter is easier to calculate. The computed eigenvalues will then be accurate to second order in the error in f (see section 3.2). For $n_B >> n_A$, the calculation of G_A requires of the order of $n_B^2 n_A$ computational operations. The remainder of the calculation, including the calculation of $\hat{H}_A^{(2)}$ or $\tilde{H}_A$, if desired, the diagonalization of the $n_A \times n_A$ effective operator, and the determination of $X^{(A)}$ all represent negligible additional computation.

5.2 2 x 2 Partitioning -- Orthonormal Basis

5.2.a General Considerations

This section is concerned with the determination of f by solution of eq. (2.16),

$$D(f) = H_{BA} + H_{BB}f - f\hat{H}_A = 0. \tag{5.1}$$

This matrix equation represents a system of $n_A n_B$ simultaneous nonlinear equations for the individual matrix elements $f_{\sigma r}$. A general solution can be written down in only two special cases. If the hamiltonian is already block diagonal, then, clearly, f = 0. If the diagonal blocks of H vanish, so that H is block off-diagonal or "alternant", then (5.1) reduces to

$$H_{BA} - fH_{AB}f = 0, \tag{5.2}$$

which has the solution,

$$f = (H_{BA}H_{AB})^{-\frac{1}{2}}H_{BA} = H_{BA}(H_{AB}H_{BA})^{-\frac{1}{2}}, \tag{5.3}$$

as can be verified by direct substitution.

When H does not have one of the special forms mentioned above, some iterative procedure or perturbation method must be used to solve (5.1). Iterative methods to successively correct the approximation to a solution are considered here. Perturbation methods are discussed in the following chapter.

Among the simplest iterative techniques to apply are those in which eq. (5.1) is rewritten as a fixed point problem,

$$f = \mathcal{F}(f) = \mathcal{A}^{-1}[D(f) + \mathcal{A}f], \tag{5.4}$$

where $\mathcal{A}$ is some non-singular, possibly f-dependent super-

operator. Successive substitutions, $f_{m+1} = \mathcal{F}(f_m)$, starting from some initial guess f_o, give the scheme,

$$\delta f_{m+1} = \mathcal{A}^{-1} D(f_m), \tag{5.5a}$$

$$f_{m+1} = f_m + \delta f_{m+1}, \quad m = 0,1,2,\ldots, \tag{5.5b}$$

hopefully convergent to a solution of (5.1). If the sequence $\{f_m\}$ converges, the rate of convergence will be linear in general if $\mathcal{A}$ is independent of f.[2]

Iterative procedures with better than linear convergence invariably involve the use of an f-dependent operator $\mathcal{A}$. The Newton-Raphson procedure is the simplest of this type. The generalized Newton-Raphson equations,

$$-J(f_m)\delta f_{m+1} = D(f_m), \tag{5.6}$$

are a special case of (5.5a), in which $\mathcal{A}$ is the negative of the Jacobian matrix, J(f), which consists of the first derivatives of the elements of D(f) with respect to the elements of f. Iteration on eqs. (5.6) and (5.5b) results in a second order convergent sequence $\{f_m\}$. That is, the error in the estimate, f_m, of f, after the m^{th} iteration is given as a linear combination of second order products of the errors in f_{m-1}, the result of the previous iteration (in the sense described in Appendix 5), so that convergence becomes very rapid as the solution is approached. For eq. (5.6), as for any iteration formula of order greater than one, convergence

[2] See Rall (1969), especially section 12; Traub, (1964); and also, Appendix 5 of this thesis.

will always occur if a sufficiently accurate initial approximation, f_o, can be obtained. For linear iteration functions, there need not be any initial estimate of f which will lead to convergence.

A set of related iterative procedures with high order convergence properties can be generated according to the scheme,

$$\begin{aligned}
\delta f_m^{(1)} &= -J(f_{m-1})^{-1} D(f_{m-1}), \\
\delta f_m^{(2)} &= -J(f_{m-1})^{-1} D(f_{m-1} + \delta f_m^{(1)}) \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
\delta f_m^{(j+1)} &= -J(f_{m-1})^{-1} D(f_{m-1} + \sum_{k=1}^{j} \delta f_m^{(k)}).
\end{aligned} \tag{5.7}$$

It can be shown that the error in $f_m = f_{m-1} + \sum_{k=1}^{j} \delta f_m^{(k)}$ is a linear combination of $(j+1)^{th}$ order products of errors in f_{m-1} (Traub, 1964). The advantage of using an iteration formula of the type (5.7) is that the Jacobian matrix, which is typically of large dimension ($n_A n_B$ x $n_A n_B$ here), need be constructed and inverted only once for each cycle of the type (5.7).

Iteration schemes with second order convergence require the evaluation and manipulation of the $(n_A n_B)^2$ first derivatives of D(f). Similarly, third order convergent iteration schemes generally require the evaluation and manipulation of the $\frac{1}{2}(n_A n_B)^3$ second derivatives of D(f). Algebraic expressions for these sets of derivatives are easily obtained. Third and higher order derivatives of D(f), eq. (2.16), with respect to f are zero.

For the particular application to large matrices, these

iteration schemes with better than linear convergence involve the manipulation of unacceptably large amounts of information. The solution of eq. (5.6) for δf_{m+1} involves of the order of $(n_A n_B)^3$ computational operations, and for $n_B \approx m >> n_A$, this is comparable to the amount of work required to diagonalize H completely. For $n_B >> n_A$, a third order formula involves of the order of m^4 operations per iteration -- equivalent to the complete diagonalization of the matrix m times over. For large matrices, it is therefore necessary to concentrate on computationally efficient, linearly convergent iteration procedures.

When H is diagonally dominant, with the diagonal elements of H_{AA} closely grouped about the value λ_A^o, the simple choice

$$\mathcal{A} = (\lambda_A^o 1_B - H_{BB}^{(d)}) \otimes 1_A \qquad (5.8)$$

(direct product notation) suggests itself. Here $H_{BB}^{(d)}$ is the diagonal part of H_{BB}. This gives an iteration scheme based on the correction

$$\delta f_{\sigma r} = \frac{D_{\sigma r}(f)}{\lambda_A^o - H_{\sigma\sigma}} \quad , \qquad (5.9)$$

closely related to degenerate perturbation theory. In eq. (5.9), and throughout the treatment of the 2 x 2 case, Greek letters refer to basis elements in S_B, and Roman letters to basis elements in S_A. The iteration index m will be dropped wherever the context does not require it.

More generally, for diagonally dominant matrices, the simple choice,

$$\mathcal{A} = 1_B \otimes H_{AA}^{(d)} - H_{BB}^{(d)} \otimes 1_A \ , \qquad (5.10)$$

leads to the corrections,

$$\delta f_{\sigma r} = \frac{D_{\sigma r}(f)}{H_{rr} - H_{\sigma\sigma}} \ , \qquad (5.11)$$

also closely related to perturbation theory. The procedure based on (5.11) will be designated as the "Simple Perturbation" (SP) algorithm. Numerical calculations indicate that it converges well only when the diagonal elements of H are ordered monotonically, and when the diagonal elements of H_{AA} are well separated from those of H_{BB}. Details of test calculations, using this and other algorithms, are given in section 5.2.g.

A better approach is to base the choice of $\mathcal{A}$ on approximations to the appropriate Newton-Raphson equations. As demonstrated in Appendix 5, these methods are still linearly convergent, but hopefully exhibit some of the stability of the Newton-Raphson equations, over a range of problems. Different approximations to (5.6) lead to algorithms exhibiting different rates of linear convergence. In assessing the computational efficiency of such algorithms, however, it is necessary to consider both the amount of computation per iteration and the number of iterations required to obtain desired accuracy.

During the iterative solution of (5.1), the required f-dependent quantities must be evaluated using the current approximation to f. Thus, the considerations in sections 3.1

and 3.2 are relevant here, and it is useful to classify the algorithms developed below according to the way in which the f-dependent quantities involved are evaluated.

5.2.b Methods Based on $D^{(1)}(f)$

If f_o is an approximation to the solution of eq. (5.1), and δf is the exact correction, so that $f = f_o + \delta f$ is the exact solution of $D^{(1)}(f) = 0$, then, it follows from the definition, (3.4), of $D^{(1)}(f)$, that

$$\hat{H}_B^{(1)}(f_o)^\dagger \delta f - \delta f \hat{H}_A(f) = -D^{(1)}(f_o). \qquad (5.12)$$

This is an exact equation for δf. The Newton-Raphson equations for the system $D^{(1)}(f) = 0$ are

$$\hat{H}_B^{(1)}(f_o)^\dagger \delta f - \delta f \hat{H}_A^{(1)}(f_o) = -D^{(1)}(f_o), \qquad (5.13)$$

the matrix elements of the Jacobian in this case being,

$$J^{(1)}_{\rho t,\sigma r} = \frac{\partial D^{(1)}_{\rho t}}{\partial f_{\sigma r}} = (\hat{H}_B^{(1)\dagger})_{\rho\sigma}\delta_{rt} - \delta_{\rho\sigma}(\hat{H}_A^{(1)})_{tr}. \qquad (5.14)$$

Equation (5.13) differs from the exact equation (5.12) only in that the exact operator $\hat{H}_A(f)$ appearing in (5.12) is replaced by the current approximation $\hat{H}_A^{(1)}(f_o)$ in (5.13).

Despite the sparseness of the Jacobian matrix here, the Newton-Raphson method is still computationally inefficient. A non-iterative method, such as Gaussian elimination, for solution of (5.13), does not easily allow proper exploitation

of the blocked structure of the Jacobian. Straightforward application of the Gauss-Seidel method and its refinements, to the determination of δf from (5.13) with $J^{(1)}(f)$ and $D^{(1)}(f)$ fixed, does allow the sparseness of the Jacobian matrix to be exploited. However, such a procedure is inefficient in that it does not make use of all the information available about f at all times if $J^{(1)}$ and $D^{(1)}$ are held fixed during the iteration to determine δf. Thus, a modified Gauss-Seidel procedure applied to (5.13) is required.

The simplest linear iteration formula based on the Newton-Raphson equations, (5.13), is one in which the operator $\mathcal{A}$ in (5.4) is taken as the negative of the diagonal part of the Jacobian matrix. The successive corrections to $f_{\sigma r}$ are then given by

$$\delta f_{\sigma r} = \frac{D^{(1)}_{\sigma r}}{(\hat{H}_A^{(1)})_{rr} - (\hat{H}_B^{(1)\dagger})_{\sigma\sigma}} \tag{5.15}$$

In view of the simplicity of the matrices involved, the most efficient computational procedure is to change only one element of f at a time, calculating $D^{(1)}_{\sigma r}$ at that time, and updating $\hat{H}_A^{(1)}$ and the diagonal part of $\hat{H}_B^{(1)\dagger}$ continually. After a change in a single $f_{\sigma r}$, these quantities are easily updated because they are linear in f,

$$(\delta\hat{H}_A^{(1)})_{sr} = H_{s\sigma}\delta f_{\sigma r}, \qquad (s = 1, \ldots, n_A), \tag{5.16a}$$

and

$$(\delta\hat{H}^{(1)\dagger})_{\sigma\sigma} = -\delta f_{\sigma r}H_{r\sigma} = -(\delta\hat{H}_A^{(1)})_{rr}. \tag{5.16b}$$

Calculation of $D_{\sigma r}^{(1)}$ as required involves the same number of computational operations per sweep through δf as the continual updating of $D^{(1)}$ (for which $D^{(1)}$ must be stored), but there is a likelihood of significant accumulation of round-off error as the solution is approached if such an updating procedure is used for $D^{(1)}$. Where the diagonal elements of H_{AA} are fairly well separated from those of H_{BB}, the usual starting approximation is $f_o = 0$. In this case, the starting approximations to $\hat{H}_A^{(1)}$ and $\hat{H}_B^{(1)\dagger}$ are simply H_{AA} and H_{BB}. The iterative scheme based on (5.15) and (5.16) will be referred to here as the "Simple Diagonal Newton-Raphson" (SDNR) algorithm. A precise statement of computational details is given in Appendix 4.

The idea of the correction $\delta f_{\sigma r}$, calculated in (5.15), is that it should reduce the corresponding $D_{\sigma r}^{(1)}$ approximately to zero. This may be far from true early in the calculation if $\delta f_{\sigma r}$ is large. The change $\delta f_{\sigma r}$ required to reduce $D_{\sigma r}^{(1)}$ exactly to zero can be determined from (5.12). The result is a quadratic equation in $\delta f_{\sigma r}$, namely,

$$H_{r\sigma}\delta f_{\sigma r}^2 + [(\hat{H}_A^{(1)})_{rr}-(\hat{H}_B^{(1)\dagger})_{\sigma\sigma}]\delta f_{\sigma r}-D_{\sigma r}^{(1)} = 0. \quad (5.17)$$

The iterative scheme based on this equation will be referred to as the "Quadratic Diagonal Newton-Raphson" (QDNR) algorithm. Precise details are given in Appendix 4. If (5.17) has two real roots, the desired correction is the one of smallest magnitude numerically. When $(\hat{H}_A^{(1)})_{rr}-(\hat{H}_B^{(1)\dagger})_{\sigma\sigma}$ is much greater than either or both of $D_{\sigma r}^{(1)}$ or $H_{r\sigma}$, this correction

differs negligibly from that given by (5.15). As the solution of (5.1) is approached, and the magnitude of $D_{\sigma r}^{(1)}$ becomes progressively smaller compared to the other coefficients in (5.17) (which are constant, or effectively constant once a reasonable approximation to f is achieved), it is necessary to use the formula for the root of a quadratic equation with a rationalized numerator to avoid serious round-off error, that is,

$$\delta f_{\sigma r} = \frac{2\chi D_{\sigma r}^{(1)}}{|(\hat{H}_A^{(1)})_{rr}-(\hat{H}_B^{(1)\dagger})_{\sigma\sigma}|+\{[(\hat{H}_A^{(1)})_{rr}-(\hat{H}_B^{(1)\dagger})_{\sigma\sigma}]^2-4H_{r\sigma}D_{\sigma r}^{(1)}\}^{\frac{1}{2}}} \qquad (5.18)$$

where

$$\chi = \mathrm{sgn}[(\hat{H}_A^{(1)})_{rr}-(\hat{H}_B^{(1)\dagger})_{\sigma\sigma}], \qquad (5.19)$$

when all coefficients are real.

This equation can be used instead of eq. (5.15) in cases where difficulty is experienced in establishing convergence. If diagonal elements of $\hat{H}_A^{(1)}$ and $\hat{H}_B^{(1)}$ become very nearly equal at some stage of the iterative calculation, eq. (5.15) may lead to divergence. Such diverging tendencies may be damped if (5.18) is used. On the other hand, situations occur in which eqs. (5.18) accelerate the divergent process. The results of some numerical calculations using both of these algorithms are included in section 5.2.g and Table 5.1.

If diagonal elements of H_{AA} and H_{BB} are equal, it is necessary either to use a non-zero starting approximation for f, or to use algorithm QDNR initially, since application of the

SDNR algorithm may lead to a division by zero early in the calculation. It is unlikely that either of these stratagems will lead to a rapidly converging calculation, however, unless a reasonable separation is soon established between the diagonal elements of $\hat{H}_A$ and $\hat{H}_B$.

Im the limit $n_B >> n_A$, the quadratic algorithm, QDNR, requires effectively the same amount of computation per sweep through δf as the linear algorithm SDNR. In both cases, the time consuming part of the calculation is the evaluation of $D^{(1)}_{\sigma r}$, and possibly, the updating of $\hat{H}^{(1)}_A$ and $\hat{H}^{(1)\dagger}_B$, rather than the calculation of $\delta f_{\sigma r}$ from either (5.15) or (5.18). Iteration on (5.15) for $\delta f_{\sigma r}$, while updating $(\hat{H}^{(1)}_A)_{rr}$, but keeping $(\hat{H}^{(1)\dagger}_B)_{\sigma\sigma}$ and $D^{(1)}_{\sigma r}$ fixed, is equivalent, if convergent, to using (5.18). This is not necessarily an efficient procedure, however.

5.2.c Methods Based on $D^{(2)}(f)$

The operator $\hat{H}^{(1)}(f)$ appearing in $D^{(1)}(f)$ must be considered to have errors of the same order as those in f itself. As shown in chapter 3, however, the error in $\hat{H}^{(2)}_A$ is smaller, in some sense, the eigenvalues being unaffected in first order by a first order error in f. This insensitivity can be exploited in iterative procedures for solving the equation,

$$D^{(2)}(f) = H_{BA} + H_{BB}f - f\hat{H}^{(2)}_A = 0, \qquad (5.19)$$

which has the same solutions as does eq. (5.1).

Because of the inverse operator g_A^{-1} in $\hat{H}_A^{(2)}$, the exact Jacobian matrix of $D^{(2)}(f)$ is no longer simple,

$$J^{(2)}_{\sigma r,\rho t} = \frac{\partial D^{(2)}_{\sigma r}}{\partial f_{\rho t}} = H_{\sigma\rho}\delta_{rt} - (\hat{H}_A^{(2)})_{rt}\delta_{\rho\sigma} - \sum_{s=1}^{n_A} f_{\sigma s}\frac{\partial(\hat{H}_A^{(2)})_{sr}}{\partial f_{\rho t}} . \tag{5.20a}$$

Here, one has,

$$\frac{\partial(\hat{H}_A^{(2)})_{sr}}{\partial f_{\rho t}} = (g_A^{-1}\hat{H}_A f^{\dagger})_{s\rho}\delta_{rt} - (g_A^{-1}f^{\dagger})_{s\rho}(\hat{H}_A)_{tr}. \tag{5.20b}$$

Because of the term involving the derivatives of the elements of $\hat{H}_A^{(2)}$, the exact Jacobian matrix is not at all sparse, in general, unlike $J^{(1)}$ of eq. (5.14). However, since $\hat{H}_A^{(2)}$ varies slowly with f near the solution of (5.19), it is expected that those elements of $J^{(2)}$ arising solely from the third term of (5.20a) will be relatively smaller than the remaining non-zero ones. On neglecting this term in (5.20a), the approximation

$$J^{(2)}_{\sigma r,\rho t} \approx \tilde{J}^{(2)}_{\sigma r,\rho t} = H_{\sigma\rho}\delta_{rt} - (\hat{H}_A^{(2)})_{rt}\delta_{\rho\sigma} , \tag{5.21}$$

is obtained. This gives the simple equation,

$$H_{BB}\delta f - \delta f\hat{H}_A^{(2)} = -D^{(2)}(f), \tag{5.22}$$

as an approximation to the Newton-Raphson equations for the system (5.19). In contrast to (5.13), this equation involves the original H_{BB} only, and not some modified $n_B \times n_B$ matrix. On the other hand, it is more complicated to update $\hat{H}_A^{(2)}$ than $\hat{H}_A^{(1)}$. For any change δf in f, the change in $\hat{H}_A^{(2)}$ is given exactly, by,

$$\delta\hat{H}_A^{(2)} = g_A^{-1(\text{new})} G_A^{(\text{new})} - g_A^{-1(\text{old})} G_A^{(\text{old})}$$

$$= g_A^{-1(\text{new})} [\delta G_A - \delta g_A \hat{H}_A^{(2)}] \qquad (5.23)$$

$$= g_A^{-1(\text{new})} [\delta f^\dagger (D^{(2)} + H_{BB}\delta f - \delta f \hat{H}_A^{(2)}) + W_{BA}^\dagger \delta f - f^\dagger \delta f \hat{H}_A^{(2)}],$$

where

$$W_{BA} = H_{BA} + H_{BB} f. \qquad (5.24)$$

All quantities on the right side of (5.23) are before updating, except where explicitly indicated.

Since an $n_A \times n_A$ matrix inversion is required for each updating of $\hat{H}_A^{(2)}$, the use of (5.22) is efficient only if groups of elements of f are changed simultaneously before updating $\hat{H}_A^{(2)}$. In application to large matrices, it is most efficient to change entire n_A-dimensional rows of f at one time. For $n_B >> n_A$, this leads to an algorithm requiring comparable work, per iterative sweep through δf, to algorithm SDNR (that is, of the order of $n_A n_B^2$ computational operations per sweep). As in SDNR, only single columns of the block H_{BB} are required at one time.

Two iterative methods based on eq. (5.22) appear useful. The first is the simplest diagonal approximation, which corresponds to taking $\mathcal{A}$ of eq. (5.4) as the negative of the diagonal part of $J^{(2)}$. This leads to the iteration formula,

$$\delta f_{\sigma r} = \frac{D_{\sigma r}^{(2)}}{(\hat{H}_A^{(2)})_{rr} - H_{\sigma\sigma}}, \qquad (r = 1, \ldots, n_A). \qquad (5.25)$$

When $\delta f_{\sigma r}$ is given by this equation, the expression, (5.23),

for $\delta\hat{H}_A^{(2)}$ simplifies somewhat to

$$\delta\hat{H}_A^{(2)} = g_A^{-1(new)}[(f^{(new)\dagger})_{A\sigma}\delta f_{\sigma A}\hat{H}_A^{(2)} + (W_{BA}^{\dagger})_{A\sigma}\delta f_{\sigma A} + (\delta f^{\dagger})_{A\sigma}\delta f_{\sigma A}\hat{H}_A^{(2)d}] , \qquad (5.26)$$

where $\hat{H}_A^{(2)d}$ is the diagonal part of $\hat{H}_A^{(2)}$, and where $(W_{BA}^{\dagger})_{A\sigma}$, $(\delta f^{\dagger})_{A\sigma}$, $(f^{(new)\dagger})_{A\sigma}$, and $(\delta f)_{\sigma A}$, refer respectively to the σ^{th} rows of $W_{BA}^{\dagger}$, $f^{(new)\dagger}$, and $\delta f^{\dagger}$, and the σ^{th} column of δf.

The second method is to treat the n_A equations in (5.22) for each fixed σ as a system of simultaneous linear equations for the $\delta f_{\sigma r}$, $(r = 1, \ldots, n_A)$. This corresponds to taking $\mathcal{A}$ to be block diagonal, each diagonal block being the negative of the diagonal block of $\tilde{J}^{(2)}$ referring to a row of δf. The resulting iteration formula can be written,

$$\delta f_{\sigma A} = -D_{\sigma A}^{(2)}[H_{\sigma\sigma}1_A - \hat{H}_A^{(2)}]^{-1} , \qquad (5.27)$$

which, in practice, involves the solution of a system of n_A simultaneous linear equations in n_A unknowns. For this change δf, the first term of (5.23) vanishes, so that the updating formula for $\hat{H}_A^{(2)}$ reduces to

$$\delta\hat{H}_A^{(2)} = g_A^{-1(new)}[(W_{BA}^{\dagger})_{A\sigma}\delta f_{\sigma A} - (f^{\dagger})_{A\sigma}\delta f_{\sigma A}\hat{H}_A^{(2)}]. \qquad (5.28)$$

This method involves somewhat more computation per sweep through δf than the preceding one, but may be expected to converge in fewer overall iterations in certain cases where the off-diagonal elements of H_{AA} are large.

The two procedures described above will be referred to as the "Diagonal Generalized Nesbet" (DGN), and the "Full

Generalized Nesbet" (FGN) algorithms, respectively. A precise statement of computational details is given in Appendix 4. In the case $n_A = 1$, they both reduce to an algorithm of Nesbet (1965). There are also certain similarities to that of Davidson (1975). Test calculations using them are described in section 5.2.g.

5.2.d Solution of the Newton-Raphson Equations by Descent Methods

The approximation of the full Newton-Raphson equations by much simpler equations, to avoid prohibitively costly calculations, reduces both the rate of convergence and the range of calculations for which convergence occurs. A major factor in non-convergence of any of the algorithms SDNR, QDNR, DGN, or FGN, must be the neglect of some or all of the coupling between elements of δf in the Newton-Raphson equations. The successive correction of individual (or at most a few) elements of f can lead to very slow convergence ("spiraling"), and also divergence, in the case of systematic over-estimation of the elements of δf. It is desirable to vary all of the elements of f simultaneously, but using methods which are less costly than solving the Newton-Raphson equations exactly. The Gauss-Seidel method applied to the full Newton-Raphson equations, with updating of J and D only after one or more sweeps through δf have been completed, is one possible way to approximate the coupling between the elements of δf. However, in this sub-

section, alternative methods based on the minimization of the residual, $J\delta f + D$, of the full Newton-Raphson equations, will be examined.

When δf is real, the solution of the Newton-Raphson equations is equivalent to determination of the stationary points of

$$Q(\delta f) = \tfrac{1}{2}\delta f^{\dagger} J \delta f + \delta f^{\dagger} D, \qquad (5.29)$$

considered as a function of δf. Further, if J is positive-definite, the solutions of the Newton-Raphson equations are equivalent to local minima of (5.29), so that δf can be determined using a gradient minimization technique. The eigenvalues of $J^{(1)}$, of eq. (5.14), are evidently the differences between the eigenvalues of $\hat{H}_B^{(1)}$ and $\hat{H}_A^{(1)}$. Thus, as long as all the eigenvalues of $\hat{H}_B^{(1)}$ are larger than all the eigenvalues of $\hat{H}_A^{(1)}$, the Jacobian $J^{(1)}$ will be positive definite. The eigenvalues of the Jacobian matrix $J^{(2)}$ are less easy to deduce because of the terms involving the derivatives of $\hat{H}_A^{(2)}$ in eq. (5.20a). If these derivatives are sufficiently small, then $J^{(2)}$ will be positive definite as long as some minimum separation is maintained between the largest eigenvalue of $\hat{H}_A^{(2)}$ and the smallest eigenvalue of H_{BB}. The condition that the Jacobian be positive definite implies that it is the lowest n_A eigenvalues which are sought.

If the Jacobian matrix is not positive definite, the solution of the Newton-Raphson equations is equivalent to minimization of the functional,

$$Q(\delta f) = (J\,\delta f + D)^{\dagger}(J\,\delta f + D)$$

$$= D^{\dagger}D + D^{\dagger}J\delta f + \delta f^{\dagger}J\,D + \delta f^{\dagger}J^{2}\delta f, \qquad (5.30)$$

which is more difficult to handle than (5.29) because of the generally large dimension of J.

When J is positive definite, the use of an iterative gradient minimization technique (such as the method of steepest descents or the method of conjugate gradients) to calculate δf by minimizing (5.29) while holding J and D fixed, involves modifying δf as a whole by successive amounts $\alpha_i v_i$, where α_i is a scalar step length chosen to minimize Q along the search direction v_i. The search directions v_i are chosen equal to, or related to, the directions along which Q changes most rapidly. Computational details of the application of these minimization techniques to quadratic forms like (5.29) are given by Ralston (1965, pp. 439-445).

If the steepest descent method is used with the Newton-Raphson equations based on $D^{(1)}(f)$, the most costly part of the minimization iteration is the determination of the step lengths α_i, which involves evaluation of the scalar product

$$v_i J^{(1)} v_i = \sum_{\sigma,\rho,t} (v_i)_{\rho t}(\hat{H}_B^{(1)\dagger})_{\rho\sigma}(v_i)_{\sigma t} - \sum_{\rho,t,r} (v_i)_{\rho t}(\hat{H}_A^{(1)})_{tr}(v_i)_{\rho r}. \qquad (5.31)$$

For $n_B >> n_A$, this requires of the order of $n_A n_B^2$ computational operations. In the conjugate gradient method, an additional $n_A n_B^2$ operations are required to evaluate the product $v_i J \cdot D$,

necessary in determining v_{i+1}. Thus, if m minimization iterations are carried out, the calculation of δf, including the initial evaluation of $J^{(1)}$ and $D^{(1)}$ requires of the order of $(m+2)n_A n_B^2$ operations using steepest descents, and a minimum of $(2m+2)n_A n_B^2$ operations for conjugate gradients. This is roughly equivalent to m+2 and 2m+2 iterations, respectively, of the algorithms discussed in the previous three subsections. The advantage here is that a very good estimate of δf may be obtained for m small, because the first few iterations in such minimization techniques frequently result in the greatest movement towards the minimum. Coupling between the elements of δf is taken into account here, while the computation per iteration is still proportional to $n_A n_B^2$ for $n_B >> n_A$, as is desired.

While such descent methods are not expected to be of much use in application to large matrices, they are very useful in the slightly more complicated self-consistent field problem in molecular orbital theory (see chapter 8), where it is considerably more costly to update D(f) and J(f), because of the complicated dependence on f of the matrix being block diagonalized.

5.2.e Extremizing the Trace

Another alternative to eq. (5.1) is to determine f such that the trace of the matrix H over the eigenspace $S_A^{\bullet}$ is

stationary (see section 2.1.e), that is, such that

$$E(f+\delta f) - E(f) = \delta E = \mathrm{tr}[P_A'(f+\delta f)-P_A'(f)]H$$
$$= \delta \mathrm{tr}\, P_A' H, \qquad (5.32)$$

vanishes to first order in δf. From eq. (2.39), this is equivalent to the vanishing of the quantity

$$\bar{D} = \nabla_f E = g_B^{-1}\, D^{(1)}(f)\, g_A^{-1}\ , \qquad (5.33)$$

$\bar{D}_{\sigma r}$ being the derivative $\partial E/\partial f^*_{\sigma r}$. The derivatives of $\bar{D}$ with respect to the elements of f and $f^\dagger$ are given by

$$\frac{\partial \bar{D}_{\rho t}}{\partial f^*_{\sigma r}} = -(fg_A^{-1})_{\sigma t}\bar{D}_{\rho r} - (fg_A^{-1})_{\rho r}\bar{D}_{\sigma t}, \qquad (5.34a)$$

and

$$\frac{\partial \bar{D}_{\rho t}}{\partial f_{\sigma r}} = -(g_B^{-1})_{\sigma\rho}(g_A^{-1}G_A g_A^{-1})_{tr} + (g_A^{-1})_{tr}(g_B^{-1}G_B g_B^{-1})_{\sigma\rho}. \qquad (5.34b)$$

Thus, the Newton-Raphson equations for the system $\bar{D} = 0$ can be written

$$(g_B^{-1}G_B g_B^{-1})\delta f g_A^{-1} - g_B^{-1}\delta f(g_A^{-1}G_A g_A^{-1}) - \bar{D}\delta f^\dagger f g_A^{-1} - f g_A^{-1}\delta f^\dagger \bar{D}$$
$$= -\bar{D}. \qquad (5.35)$$

On multiplying from the right by g_A and from the left by g_B, these become

$$\hat{H}_B^{(2)\dagger}\delta f - \delta f \hat{H}_A^{(2)} = -D^{(1)} + [D^{(1)} g_A^{-1}\delta f^\dagger f + f\delta f^\dagger D^{(2)}]\ . \qquad (5.36)$$

If $D^{(1)}$ and $D^{(2)}$ are considered to be of the same order as δf as the solution is approached, then the last term of (5.36) is of higher order than the remaining three terms. If this term

is neglected, the resulting equation is of the type (5.13) with the operators $\hat{H}_A^{(1)}$ and $\hat{H}_B^{(1)}$, replaced, respectively, by $\hat{H}_A^{(2)}$ and $\hat{H}_B^{(2)}$. Because the difference between $\hat{H}_A^{(1)}$, $\hat{H}_B^{(1)}$, and $\hat{H}_A^{(2)}$, $\hat{H}_B^{(2)}$, is of the same order as the term neglected, the resulting approximate equation is not necessarily an improvement over (5.13), despite the presence of the more accurate effective operators.

The exact equations, (5.35) and (5.36), could be significant for gradient minimization techniques, which can be set up so that divergence cannot occur. However, the evaluation of $\hat{H}_B^{(2)}$ involves the inversion of the n_B x n_B matrix, g_B, as well as the formation of the product $g_B^{-1}G_B$. For $n_B >> n_A$, these two computations could be prohibitive.

5.2.f Minimization of the Norm of D

A further alternative to determining f by solution of (5.1) is to minimize the square of the Hilbert-Schmidt norm of D,

$$\|D\|^2 = \sum_{\sigma,r} |D_{\sigma r}|^2, \tag{5.37}$$

with respect to f. The required gradients are

$$\frac{\partial \|D\|^2}{\partial f_{\tau s}} = 2 \sum_{\sigma,r} D_{\sigma r} \frac{\partial D_{\sigma r}}{\partial f_{\tau s}}$$

$$= 2 (\hat{H}_B D + D \hat{H}_A)_{\tau s}. \tag{5.38}$$

This approach is attractive because it involves a suitable

convergence criterion directly. If a gradient minimization technique is used, it is easy to ensure that a maximal rate of convergence is maintained. By minimizing $\|D\|^2$ itself along some search direction in each iteration, problems of overshoot and undershoot can largely be avoided.

5.2.g Test Calculations

The algorithms described in sections 5.2.a - 5.2.c have been applied to a series of matrices based on that considered by Nesbet (1965). The off-diagonal elements of these matrices are all unity, and the diagonal elements are some combination of the first n odd integers, 1,3,5, ... Matrices with dimensions up to 250 x 250 were considered, this being sufficient for testing. The calculations were carried out on an IBM 370/168 computer using double precision arithmetic. The convergence criterion was based on the Hilbert-Schmidt norm, $\|D\|=(\mathrm{tr}\ D^{\dagger}D)^{\frac{1}{2}}$, of the particular form of D(f) used in each method. A criterion based on the maximum change δf_{cr} in the elements of f during an iterative sweep can also be useful. In all examples, the basis space, S_A, is defined by the first n_A basis functions in order, so that re-ordering the diagonal elements of H is equivalent to varying S_A.

For convergent calculations, it was found, except for the first few iterations in some cases, that $\log\|D\|$ is usually well approximated as a linear function of the iteration number.

That is, convergence was linear once the calculation stabilized, with the value of $\|D\|$ decreasing on the average by some constant factor for each iteration. This factor can be regarded as an average asymptotic error constant. Table 5.1 gives these error constants (or convergence rates) for a number of examples, to illustrate the effects of varying the size of the matrix, varying the differences between the diagonal elements of H_{AA} and H_{BB}, and varying the ordering of the diagonal elements of the full matrix to change S_A. For comparison, Nesbet's algorithm (Nesbet, 1965) was used to obtain a single eigenvalue of each of the matrices considered. The square root, σ, of the variance for the approximate eigenvector, as defined in eq. (2.49), was used as the convergence criterion in this case, and log σ was also found to be a linear function of the iteration number. Note that the smallest numbers in Table 5.1 represent the fastest convergence.

For the basic Nesbet matrix, with S_A the space corresponding to the n_A smallest (or largest) diagonal elements of H, all methods converge to give the eigenspace of the n_A smallest (or largest) eigenvalues. The rate of convergence varies little with n_A, either increasing or decreasing slightly as n_A increases. When the largest eigenvalues are sought (or equivalently, when the off-diagonal elements are -1), convergence is considerably poorer for $n_A = 1$ than for $n_A = 5$, except for algorithm DGN. The five algorithms tested have rates of convergence generally comparable to the Nesbet method

TABLE 5.1

Linear[a] Convergence Rates of the Algorithms in Selected Calculations.

n_A	n ($n_A + n_B$)	DIAGONAL MATRIX ELEMENTS IN ORDER	METHOD[b]					Nesbet $n_A = 1$
			SP	SDNR	QDNR	DGN	FGN	
1	10	1,3,5...17,19[c]	0.27	0.24	0.24	0.23	0.23	0.23
1	20	1,3,5...37,39	0.33	0.31	0.31	0.30	0.30	0.30
1	250	1,3,5...497,499	0.52	0.51	-	0.50	0.50	0.50
5	10	1,3,5...17,19	0.26	0.22	0.23	0.44	0.19 (4)	0.23
5	20	1,3,5...37,39	0.38	0.29	0.30	0.37	0.29 (4)	0.30
5	250	1,3,5...497,499	0.55	0.51	-	0.49	0.50 (4)	-
5	10	19,17,15...3,1	0.24[h]	0.23[h]	0.23[h]	0.74(3)[h]	0.24 (2)[h]	0.54
5	10	1,3,5...11,9...17,19	K	0.33 (5)	0.994	0.40[d]	0.30 (2)[d]	0.23
5	10	1,3,...11,9,7,13...17,19	0.61[e]	0.55[e]	0.55[e]	0.42[e]	0.31 (4)[e]	0.23
5	10	1,3,...11,13,7,9...17,19	0.58[f]	0.31[f]	0.31[f]	0.42[f]	0.999(4)[g]	0.23
5	10	1,3...13,11,9,7...17,19	0.57[f]	0.31[f]	0.31[f]	0.43[f]	0.999(4)[g]	0.23
5	10	19,17...9,11,...3,1	0.59[i]	j	div.	0.64[i]	0.23[i]	0.54
5	20	1,1.1,1.2,1.8,1.4,11,13,...37,39	0.20	0.18	0.18	0.21	0.16	m
5	20	1,1.1,1.2,1.3,1.4,1.5,13,15...31,39	K	0.20	0.19	0.98	0.90	m
5	20	1,1.1,1.2,1.3,1.4,3,5...29,31	0.56	0.41	0.40	0.57	l	m
5	20	1,1.1,1.2,1.3,1.4,1.5,3,5,...27,29	K	0.45	0.44	K	K	m

TABLE 5.1 (continued)

[a]The tabulated numbers represent the average factor by which the norm $\|D\|$ is decreased per iteration, once a linear convergence rate is established. S_A is spanned by the first n_A basis functions. All off-diagonal elements are unity. The numbers are obtained by a least squares calculation of the slope of $\log\|D\|$ as a function of iteration number.

[b]the number of iterations before linear convergence is established is indicate in brackets to the right of the convergence factor when not zero.

[c]the eigenvalues of this matrix are: 0.386, 2.461, 4.519, 6.753, 8.629, 10.691, 12.766, 14.868, 17.037, 22.072.

[d]converges to the eigenvalues 0.386, 2.461, 4.519, 6.573, 10.691.

[e]converges to the eigenvalues 0.386, 2.461, 4.519, 8.629, 10.691.

[f]converges to the eigenvalues 0.386, 2.461, 4.519, 10.691, 12.766.

[g]apparently converges to the eigenvalues 0.386, 2.461, 4.519, 10.691, 14.868.

[h]converges to the eigenvalues 10.691, 12.766, 14.868,17.037, 22.072.

[i] converges to the eigenvalues 8.629, 12.766, 14.868, 17.037, 22.072.

[j]$\|D\|$ is oscillatory.

[k]$\|D\|$ is apparently divergent.

[l]$\|D\|$ becomes constant (= 4.34) after 25 iterations.

[m] σ becomes constant or increases very slowly after about 50 iterations.

in cases where convergence is straightforward, as it is when the diagonal elements of H are ordered monotonically and are well separated. Frequently, large matrices arising in various applications, for which only a few of the lowest eigenvalues and their eigenvectors are required, have diagonal elements arranged in roughly increasing order, with variation in the diagonal elements large compared to individual off-diagonal elements. As seen from the results in the first part of the table, these algorithms are well suited for such calculations. The simple perturbation (SP) algorithm generally exhibits the poorest convergence rates in these examples, as may be expected, since it represents the crudest approximation to the exact Newton-Raphson equations. The algorithm DGN works relatively poorly in two cases for $n_A = 5$. Presumably, one of the diagonal elements of $\hat{H}_A^{(2)}$ approaches one of H_{BB} too closely during the calculation. The algorithm QDNR has convergence rates identical to SDNR, because these two calculations differ significantly only at initial stages before linear convergence is established.

The effect of varying the spaces S_A and S_B, as defined by the associated diagonal elements of H, is illustrated by the third part of Table 5.1. Rates of convergence deteriorate markedly when one or more diagonal elements of H_{AA} exceed at least one diagonal element of H_{BB}. The algorithms DGN and FGN sometimes converge to different eigenspaces S_A' than do SDNR and QDNR. It is noteworthy that QDNR gives no improvement

over SDNR in these examples, and is actually non-convergent in one case where SDNR converges well.

The uncertain convergence is presumably due to one of the differences $(\hat{H}_B^{(1)\dagger})_{\sigma\sigma} - (\hat{H}_A^{(1)})_{rr}$, or $H_{\sigma\sigma} - (\hat{H}_A^{(2)})_{rr}$, appearing in the denominators of the iteration formulas, becoming small or changing sign. The iteration formulas become ill-conditioned or even singular under such circumstances. The presence of an "induction period" before linear convergence is established is presumably associated with an initial uncertainty in the selection of a space S_A', when the diagonal elements of the approximate A-space and B-space effective operators are not well separated.

In principle, the space S_A', specified by the calculated f, may correspond to any group of n_A eigenvalues of the matrix H. Thus, in principle, any subset of n_A eigenvalues, none of whose eigenvectors are orthogonal to the subspace S_A of the full basis space, can be calculated without previous determination of any of the other eigenvalues. However, the first three sections of Table 5.1 show that the iterative calculation converges best when S_A' corresponds to the n_A lowest (or highest) eigenvalues of the matrix, and S_A to the smallest (or largest) diagonal elements of the matrix. Deviations from this arrangement entail considerable risk of poor convergence or no convergence at all. If n_A of the lowest m eigenvalues ($m > n_A$) are desired, these convergence problems can be avoided by prediagonalizing a block of H containing the m smallest diagonal

elements, as described in section 5.3.e.

The last section of Table 5.1 shows the superiority of the algorithms developed here over Nesbet's algorithm when the lowest few diagonal elements of the matrix are nearly equal, but well separated from the remaining ones. Generally, the f-operator calculated must correspond to the space S_A' spanned by the eigenvectors belonging to all of the nearly equal eigenvalues if good convergence rates are to be obtained. However, a surprising feature of the results is that the algorithm SDNR performs very well, even when a diagonal element of H_{BB} is relatively close to one of H_{AA}.

These computations do not indicate any clear cut superiority of one algorithm in all cases. When convergence is straightforward, all converge effectively equally rapidly. When convergence is not straightforward, any one of the methods may be more stable or rapidly convergent than the others. However, the algorithm DGN appears to be less successful than SDNR and FGN, generally. The simple diagonal Newton-Raphson procedure, based on $D^{(1)}(f) = 0$, is somewhat easier to program efficiently for $n_A > 1$, than the methods based on $D^{(2)}(f)$, and from this standpoint is particularly attractive. In fact, in most cases, the rates of convergence for this method compare very favourably with those of the other, more complex, methods. The extra computation involved in using QDNR rather than SDNR appears to be of little value, in general, even though this represents a negligible amount of additional work as n_B

becomes very large. While SDNR yields only the approximation $\hat{H}_A^{(1)}$ directly, a calculation of $\hat{H}_A^{(2)}$ at the end of the iterative sequence takes only of the order of the time of one iteration, and can be carried out if the eigenvalues and eigenvectors of H corresponding to $S_A^{\bullet}$ are desired. For n_A =1, SDNR offers an alternative to Nesbet's method, of comparable efficiency.

5.3 Generalization to a Non-orthonormal Basis -- 2 x 2 Case

5.3.a General Considerations

When the basis is not orthonormal, the relation of the two off-diagonal blocks of the partitioning operator is more complicated than before (see section 2.3). The off-diagonal blocks, f and h, of $\hat{T}$, eq. (2.2), can be defined by the pair of simultaneous equations representing the vanishing of the off-diagonal blocks of $G = \hat{T}^{\dagger}H\hat{T}$ and $g = \hat{T}^{\dagger}S\hat{T}$, namely,

$$G_{BA} = H_{BA} + H_{BB}f + h^{\dagger}\hat{H}'_{A} = 0, \tag{5.39a}$$

and

$$g_{BA} = S_{BA} + S_{BB}f + h^{\dagger}\hat{S}_{A} = 0, \tag{5.39b}$$

where $\hat{H}'_{A} = H_{AA} + H_{AB}f$, and $\hat{S}_{A} = S_{AA} + S_{AB}f$. Alternatively, these equations may be combined to give separate equations for f and h,

$$D_{BA}(f) = H_{BA} + H_{BB}f - (S_{BA} + S_{BB}f)\hat{H}_{A} = 0, \tag{5.40a}$$

and

$$D_{AB}(h) = H_{AB} + H_{AA}h - (S_{AA}h + S_{AB})\hat{H}_{B} = 0, \tag{5.40b}$$

as in eqs. (2.113) and (2.114). Here, one has,

$$\hat{H}_{A} = \hat{S}_{A}^{-1}\hat{H}'_{A}, \tag{5.41a}$$

and

$$\hat{H}_{B} = \hat{S}_{B}^{-1}\hat{H}'_{B}, \tag{5.41b}$$

with $\hat{H}'_{B} = H_{BB} + H_{BA}h$, and $\hat{S}_{B} = S_{BB} + S_{BA}h$.

Algorithms to calculate f and h, or either separately, have again been founded on approximations of the full Newton-

Raphson equations by less costly linearly convergent iteration scheme.

5.3.b Methods Based on G_{BA} and g_{BA} -- A Generalization of Algorithm SDNR

A direct generalization of the SDNR algorithm based on the equation $D_{BA}^{(1)}(f) = 0$, does not lead to an efficient computational scheme in a non-orthonormal basis. Because of the inverse matrix $\hat{S}_A^{-1}$, a change in even a single element of f changes all of the elements of $\hat{H}_A^{(1)}$, making updating costly. However, a very simple procedure can be based on the simultaneous solution of eqs. (5.39), this procedure reducing to algorithm SDNR when the overlap matrix S is replaced by 1_n, and h by $-f^\dagger$.

The Newton-Raphson equations corresponding to the system (5.39) can be written as the pair

$$\hat{H}_B^{\bullet\dagger}\delta f + \delta h^\dagger \hat{H}_A^{\bullet} = -G_{BA} \,, \tag{5.42a}$$

and

$$\hat{S}_B^\dagger \delta f + \delta h^\dagger \hat{S}_A = -g_{BA} \,. \tag{5.42b}$$

These represent $2n_An_B$ equations for the n_An_B elements of f and the n_An_B elements of h. The diagonal parts of these equations are of the form

$$\begin{bmatrix} (\hat{H}_B^{\bullet\dagger})_{\sigma\sigma} & (\hat{H}_A^{\bullet})_{rr} \\ (\hat{S}_B^\dagger)_{\sigma\sigma} & (\hat{S}_A)_{rr} \end{bmatrix} \begin{bmatrix} \delta f_{\sigma r} \\ \delta h^*_{r\sigma} \end{bmatrix} = - \begin{bmatrix} G_{\sigma r} \\ g_{\sigma r} \end{bmatrix}, \tag{5.43}$$

with the solution,

$$\delta f_{\sigma r} = \frac{-G_{\sigma r}(\hat{S}_A)_{rr} - g_{\sigma r}(\hat{H}'_A)_{rr}}{\triangle_{\sigma r}} , \qquad (5.44a)$$

and

$$\delta h^*_{\sigma r} = \frac{-(\hat{H}'^{\dagger}_B)_{\sigma\sigma} g_{\sigma r} + (\hat{S}^{\dagger}_B)_{\sigma\sigma} G_{\sigma r}}{\triangle_{\sigma r}} , \qquad (5.44b)$$

where

$$\triangle_{\sigma r} = (\hat{H}'^{\dagger}_B)_{\sigma\sigma}(\hat{S}_A)_{rr} - (\hat{S}^{\dagger}_B)_{\sigma\sigma}(\hat{H}'_A)_{rr} . \qquad (5.45)$$

A computational procedure based on these equations involves roughly double the work per iteration as algorithm SDNR. The quantities $G_{\sigma r}$ and $g_{\sigma r}$ are calculated as required, while $\hat{H}'_A$, $\hat{S}_A$, and the diagonal elements of $\hat{S}^{\dagger}_B$ and $\hat{H}'^{\dagger}_B$ are stored. These latter matrices are easily updated, because they are linear in f and h. For a change in $f_{\sigma r}$ and $h_{r\sigma}$, one has

$$\left.\begin{aligned} (\delta\hat{H}'_A)_{sr} &= H_{s\sigma}\delta f_{\sigma r} , \\ (\delta\hat{S}_A)_{sr} &= S_{s\sigma}\delta f_{\sigma r} , \end{aligned}\right\} \quad (s = 1, \ldots, n_A), \qquad (5.46a)$$

and

$$\left.\begin{aligned} (\delta\hat{H}'^{\dagger}_B)_{\sigma\sigma} &= \delta h^*_{r\sigma} H_{r\sigma} , \\ (\delta\hat{S}^{\dagger}_B)_{\sigma\sigma} &= \delta h^*_{r\sigma} S_{r\sigma} . \end{aligned}\right. \qquad (5.46b)$$

Precise computational details of this procedure, designated as the "Simple Diagonal Newton-Raphson with Overlap" (SDNRS) algorithm, are given in Appendix 4.

Again, a quadratic generalization (here QDNRS) can be obtained. Equations for the exact corrections δf and $\delta h^{\dagger}$,

required to reduce G_{BA} and g_{BA} exactly to zero, are obtained from (5.39) in the form,

$$\hat{H}_B'^{\dagger}\delta f + \delta h^{\dagger}\hat{H}_A' + \delta h^{\dagger}H_{AB}\delta f = -G_{BA} , \qquad (5.47a)$$

and

$$\hat{S}_B^{\dagger}\delta f + \delta h^{\dagger}\hat{S}_A + \delta h^{\dagger}S_{AB}\delta f = -g_{BA} . \qquad (5.47b)$$

The corresponding "diagonal equations",

$$(\hat{H}_B'^{\dagger})_{\sigma\sigma}\delta f_{\sigma r} + \delta h^{*}_{r\sigma}(\hat{H}_A')_{rr} + \delta h^{*}_{r\sigma}H_{r\sigma}\delta f_{\sigma r} = -G_{\sigma r}, \qquad (5.48a)$$

and

$$(\hat{S}_B^{\dagger})_{\sigma\sigma}\delta f_{\sigma r} + \delta h^{*}_{r\sigma}(\hat{S}_A)_{rr} + \delta h^{*}_{r\sigma}S_{r\sigma}\delta f_{\sigma r} = -g_{\sigma r} , \qquad (5.48b)$$

can be combined to give $\delta f_{\sigma r}$ and $\delta h^{*}_{r\sigma}$ as the roots of quadratic equations. The correction to $f_{\sigma r}$ is the smallest root of

$$A\delta f_{\sigma r}^2 + B\delta f_{\sigma r} + C = 0, \qquad (5.49)$$

where

$$\begin{aligned} A &= (\hat{S}_B^{\dagger})_{\sigma\sigma}H_{r\sigma} - S_{r\sigma}(\hat{H}_B'^{\dagger})_{\sigma\sigma} , \\ B &= -\Delta_{\sigma r} - S_{r\sigma}G_{\sigma r} + H_{r\sigma}g_{\sigma r} , \qquad (5.50) \end{aligned}$$

and

$$C = g_{\sigma r}(\hat{H}_A')_{rr} - (\hat{S}_A)_{rr}G_{\sigma r} .$$

The correction to $h^{\dagger}$ is then

$$\delta h^{*}_{r\sigma} = \frac{-G_{\sigma r} - (\hat{H}_B'^{\dagger})_{\sigma\sigma}\delta f_{\sigma r}}{(\hat{H}_A')_{rr} + H_{r\sigma}\delta f_{\sigma r}} . \qquad (5.51)$$

Precise computational details again appear in Appendix 4.

5.3.c Methods Based on $D^{(2)}(f)$ -- Generalized Nesbet Procedures

As explained in section 3.2, equation (5.40a) can be understood either as $D^{(1)}(f) = 0$, or $D^{(2)}(f) = 0$, according as $\hat{H}_A$ is taken as the approximation $\hat{H}_A^{(1)}$ or $\hat{H}_A^{(2)}$. In either case, the Jacobian matrix has the elements

$$J^{(i)}_{\sigma r,\rho t} = \frac{\partial D^{(i)}_{\sigma r}}{\partial f_{\rho t}} = H_{\sigma\rho}\delta_{rt} - S_{\sigma\rho}(\hat{H}_A^{(i)})_{tr} - \left[(S_{BA}+S_{BB}f)\frac{\partial \hat{H}_A^{(i)}}{\partial f_{\rho t}}\right]_{\sigma r} . \tag{5.52}$$

Even without the last term involving the derivatives of $\hat{H}_A^{(i)}$, this matrix is no longer sparse, and the convergence of iterative methods based on "diagonal" approximations may be adversely affected.

For $D^{(1)}(f)$, this Jacobian matrix can be written,

$$J^{(1)}_{\sigma r,\rho t} = [H_{\sigma\rho} - Y_{\sigma A}\hat{S}_A^{-1}H_{A\rho}]\delta_{rt} - [S_{\sigma\rho} - Y_{\sigma A}\hat{S}_A^{-1}S_{A\rho}](\hat{H}_A)_{tr} . \tag{5.53}$$

This is considerably more complicated than before, and must be handled in a similar way to the $D^{(2)}$ methods.

The $D^{(2)}$ (generalized Nesbet) methods are extended to a non-orthonormal basis straightforwardly. As before, it is reasonable to neglect the derivatives of $\hat{H}_A^{(2)}$, giving,

$$\frac{\partial D^{(2)}_{\sigma r}}{\partial f_{\rho t}} \approx \tilde{J}^{(2)}_{\sigma r,\rho t} = H_{\sigma\rho}\delta_{rt} - S_{\sigma\rho}(\hat{H}_A^{(2)})_{tr} . \tag{5.54}$$

For some change δf in f, the operator $H_A^{(2)}$ is updated according to

$$\delta\hat{H}_A^{(2)} = g_A^{-1(new)}[\delta f^\dagger(D^{(2)} + H_{BB}\delta f - S_{BB}\delta f\hat{H}_A^{(2)}) + W_{BA}^\dagger\delta f - Y_{BA}^\dagger\delta f\hat{H}_A^{(2)}], \quad (5.55)$$

where $W_{BA} = H_{BA} + H_{BB}f$ and $Y_{BA} = S_{BA} + S_{BB}f$. The "Diagonal Generalized Nesbet with Overlap" (DGNS) iteration formula is

$$\delta f_{\sigma r} = \frac{D_{\sigma r}^{(2)}}{S_{\sigma\sigma}(\hat{H}_A^{(2)})_{rr} - H_{\sigma\sigma}}, \quad (r = 1, \ldots, n_A), \quad (5.56)$$

for which the updating formula for $\hat{H}_A^{(2)}$, eq. (5.55), becomes

$$\delta\hat{H}_A^{(2)} = g_A^{-1(new)}[(W_{BA}^\dagger)_{A\sigma}\delta f_{\sigma A} + (Y_{BA}^{(new)\dagger})_{A\sigma}\delta f_{\sigma A}\hat{H}_A^{(2)} + S_{\sigma\sigma}\delta f_{\sigma A}^\dagger\delta f_{\sigma A}\hat{H}_A^{(2)d}]. \quad (5.57)$$

The "Full Generalized Nesbet with Overlap" (FGNS) iteration formula is

$$\delta f_{\sigma A} = D_{\sigma A}^{(2)}[S_{\sigma\sigma}\hat{H}_A^{(2)} - H_{\sigma\sigma}]^{-1}, \quad (5.58)$$

which again, in practice, is treated as a system of n_A simultaneous linear equations. The first term of eq. (5.55) now vanishes, so that the updating formula for $\hat{H}_A^{(2)}$ becomes

$$\delta\hat{H}_A^{(2)} = g_A^{-1(new)}[(W_{BA}^\dagger)_{A\sigma}\delta f_{\sigma A} - (Y_{BA}^\dagger)_{A\sigma}\delta f_{\sigma A}\hat{H}_A^{(2)}]. \quad (5.59)$$

For both algorithms DGNS and FGNS, approximately twice the computation is required per sweep through δf, as for their counterparts in an orthonormal basis. A precise statement of computational details is given in Appendix 4.

5.3.d Other Methods

For a non-orthonormal basis, the gradient of the trace of the matrix H over the image space, S_A', of the projection P_A' is given by

$$\frac{\partial E}{\partial f^*_{\sigma r}} = \frac{\partial}{\partial f^*_{\sigma r}} \operatorname{tr} (P_A' H) = \bar{D}_{\sigma r}, \tag{5.60}$$

where

$$\bar{D} = -Y_{BA} g_A^{-1} G_A g_A^{-1} + W_{BA} g_A^{-1}. \tag{5.61}$$

As this trace is stationary if and only if S_A' is an eigenspace of H (section 2.1.e), one way to determine f is to solve the equation $\bar{D} = 0$. Using eqs. (3.11) - (3.15), eqs. (5.61) can be transformed to

$$\begin{aligned} \bar{D} &= D^{(2)}(f)\, g_A^{-1} \\ &= [1_B - Y_{BA} g_A^{-1} f^\dagger] D^{(1)}(f) g_A^{-1}, \end{aligned} \tag{5.62}$$

which vanishes, as it should, when $D^{(1)}(f)$ and $D^{(2)}(f)$ vanish. This equation reduces to eq. (2.45) in an orthonormal basis.

Algebraic expressions for the derivatives of $\bar{D}$ with respect to the elements of f can be obtained without difficulty, and give the Newton-Raphson equations for the system $\bar{D}=0$, as

$$\begin{aligned} -(S_{BB} - Y_{BA} g_A^{-1} Y_{BA}^\dagger)\delta f (g_A^{-1} G_A g_A^{-1}) + (Y_{BA} g_A^{-1} W_{BA}^\dagger + H_{BB} - \bar{D} Y_{BA}^\dagger)\delta f g_A^{-1} \\ - Y_{BA} g_A^{-1} \delta f^\dagger \bar{D} - \bar{D} \delta f^\dagger Y_{BA} g_A^{-1} = -\bar{D}. \end{aligned} \tag{5.63}$$

These are somewhat more complicated than the previous eq. (5.35), but not hopelessly so.

The Newton-Raphson equations, (5.42), or those arising from $D^{(1)}(f) = 0$, or $D^{(2)}(f) = 0$, can be solved for δf using descent methods as described in section 5.2.d. As before, the costly part of the minimization iteration, the evaluation of products like $v_i \cdot J \cdot v_i$, generally require of the order of $n_A^2 n_B^2$ computational steps, in addition to the work required in calculating the Jacobian, and the other vectors entering the product. For the Newton-Raphson equations, (5.42), reduction in computation by a factor of $\frac{1}{2}n_A$ results if the blocked structure of the Jacobian is explicitly taken into account, yielding,

$$v \cdot J \cdot v = \sum_{\sigma,\rho,r} [(v_f)_{\sigma r}(\hat{H}_B'^{\dagger})_{\sigma\rho}(v_f)_{\rho r} + (v_h)_{r\sigma}(\hat{S}_B^{\dagger})_{\sigma\rho}(v_f)_{\rho r}] + \sum_{\sigma,r,t} [(v_f)_{\sigma r}(\hat{H}_A')_{tr}(v_h)_{t\sigma} + (v_h)_{r\sigma}(\hat{S}_A)_{tr}(v_h)_{t\sigma}] \tag{5.64a}$$

where the search vector, v, has been divided into an f part and an h part,

$$v = \begin{bmatrix} v_f \\ v_h \end{bmatrix} . \tag{5.64b}$$

Equation (5.64a) represents of the order of $2n_A n_B^2$ computational operations for $n_B >> n_A$. As for an orthonormal basis, then, the approximate calculation of δf (or δf and δh) from the Newton-Raphson equations using a gradient minimization procedure is as costly as several iterations in algorithms SDNRS, QDNRS, DGNS, or FGNS.

A third alternative is to determine f by minimizing the

Hilbert-Schmidt norm of G_{BA} and g_{BA}, or of $D^{(1)}(f)$ or $D^{(2)}(f)$ directly. Only the scheme based on G_{BA} and g_{BA} is considered here because the derivatives of G_{BA} and g_{BA} with respect to f and h are particularly simple, and because the form of the quantity to be minimized is not as simple as (5.37). Since G_{BA} has dimensions of energy, whereas g_{BA} is dimensionless, the quantity to be minimized should be of the form

$$N = \|G_{BA}\|^2 + \alpha^2\|g_{BA}\| = \sum_{\sigma,r} (|G_{\sigma r}| + \alpha^2|g_{\sigma r}|), \quad (5.65)$$

where α is a constant scale factor with dimensions of energy. The first derivatives of N are

$$\frac{\partial N}{\partial f_{\tau s}} = 2(\hat{H}'_B G_{BA} + \alpha^2 \hat{S}_B g_{BA})_{\tau s}, \quad (5.66a)$$

and

$$\frac{\partial N}{\partial h^*_{s\tau}} = 2(G_{BA}\hat{H}'^{\dagger}_A + \alpha^2 g_{BA}\hat{S}^{\dagger}_A)_{\tau s}. \quad (5.66b)$$

Actual test calculations are required to develop criteria for the choosing of α. It is desirable to choose α in some way which maximizes the rate of convergence, but such a criterion is not easily translated into an algebraic condition on α. The interpretation of α as an average energy scaling factor suggests $\alpha \approx n_A^{-1}$ tr G_A.

5.3.e Choice of an Initial Estimate, and Improvement of Convergence Rates

When the off-diagonal elements of H are small compared to differences between diagonal elements in H_{AA} and H_{BB}, the matrix elements of f are small, and a reasonable (and practical) starting approximation in an iterative calculation is $f_o = 0$.

An improved starting approximation may be provided by the solution of a similar problem when available, or more easily calculated. For example, a possible starting estimate of f for a non-orthonormal basis is an approximate solution of the corresponding problem with S replaced by a unit matrix (orthonormal basis). Similarly, the operator,

$$f' = \begin{bmatrix} f'_{\bar{m}} \\ 0 \end{bmatrix}, \tag{5.67a}$$

for the matrix H', related to H by

$$H' = \left[\begin{array}{c|c} H_{mm} & 0 \\ \hline 0 & H_{\bar{B}\bar{B}} \end{array}\right] = \left[\begin{array}{ccc} H_{AA} & H_{A\bar{m}} & 0 \\ H_{\bar{m}A} & H_{\bar{m}\bar{m}} & 0 \\ \hline 0 & 0 & H_{\bar{B}\bar{B}} \end{array}\right], \tag{5.67b}$$

with $m > n_A$, $\bar{m} = m - n_A$, and $n_{\bar{B}} = n_B - m$, will also be an improved initial estimate for f, especially when $H_{\bar{m}A}$ contains the most significant elements of H_{BA}. If m is not too large, the $(m - n_A) \times n_A$ block $f'_{\bar{m}}$ is easily calculated from the eigenvectors of the block H_{mm}, eq. (5.67b), using eq. (2.3). The idea here is to improve the initial estimate of the larger

elements of f.

The consideration of asymptotic error constants and rates of convergence (Appendix 5) implies that an improved initial estimate of f may make the difference between convergence and divergence, but, in general, will have little effect on the rate of convergence eventually established. This has been born out in test calculations.

Generally, the rate of linear convergence in these algorithms is inversely related to the ratio between the off-diagonal elements of H and the denominators occurring in the iteration formulas. Thus, the rate of convergence will be increased if these ratios are decreased by carrying out a linear transformation to reduce the size of the off-diagonal elements of H, and perhaps increase the size of the denominators in the iteration formulas. Therefore, a partial diagonalization of H to reduce to zero those off-diagonal elements which are coefficients of the potentially largest errors in the error formulas given in Appendix 5, followed by the iterative calculation of f (with $f_o = 0$) in this new basis, will result in improved rates of convergence. The desired mapping, f, in the original basis is obtained using the transformation equations given in section 2.1.f. Typically, this prediagonalization would involve an m x m block of H ($m > n_A$) containing H_{AA} in addition to that part of the remainder of H with the strongest coupling to H_{AA}.

This prediagonalization is especially useful when some

denominators in the iteration formulas are small (implying that H_{AA} and $H_{\bar{B}\bar{B}}$ have some nearly equal or equal diagonal elements), since in the new basis, these denominators may be much larger, and rates of convergence correspondingly become significantly improved. If the diagonal elements of H_{AA} and $H_{\bar{B}\bar{B}}$ are initially well separated, the effect of prediagonalizing a small block of a relatively larger matrix may not be as noticeable.

It must be emphasized that this procedure is not the same thing as the prediagonalization procedure described earlier to obtain the starting approximation f' of eq. (5.67a). The linear basis transformation corresponds to a nonlinear transformation on the elements of f, and the metric properties of the iteration formula are changed, thereby changing the entire character of the iterative calculation.

It is easily seen that for $n_B >> m > n_A$, the transformation of H to the new basis, given by the columns of the matrix V relative to the old basis, and the subsequent back-transformation of f requires at most of the order of $n_A n_B^2$ operations, because the greater part of the forward transformation matrix is a unit matrix, that is

$$V = \begin{bmatrix} V_{m\bar{m}} & 0 \\ 0 & 1_{\bar{B}} \end{bmatrix} = \begin{bmatrix} V_{AA} & V_{A\bar{m}} & 0 \\ V_{\bar{m}A} & V_{\bar{m}\bar{m}} & 0 \\ 0 & 0 & 1_{\bar{B}} \end{bmatrix} . \qquad (5.68)$$

Here $V_{m\bar{m}}$ is the m x m matrix of the eigenvectors of the m x m

block of H, $1_{\bar{B}}$ is an (n-m) x (n-m) unit matrix, and $\bar{m} = m-n_A$. The transformed matrix H' is then

$$H' = V^\dagger H V = \begin{bmatrix} V_{mm}^\dagger H_{mm} V_{mm} & (H_{\bar{B}m} V_{mm})^\dagger \\ H_{\bar{B}m} V_{mm} & H_{\bar{B}\bar{B}} \end{bmatrix} \qquad (5.69a)$$

$$= \begin{bmatrix} H_{AA}^{(d)} & 0 & H_{\bar{B}A}'^\dagger \\ 0 & H_{\bar{m}\bar{m}}^{(d)} & H_{\bar{B}m}'^\dagger \\ H_{\bar{B}A}' & H_{\bar{B}\bar{m}}' & H_{\bar{B}\bar{B}} \end{bmatrix}, \qquad (5.69b)$$

The reverse transformation for f is

$$f = [(V^{\dagger -1})_{BA} + (V^{\dagger -1})_{BB} f'][(V^{\dagger -1})_{AA} + (V^{\dagger -1})_{AB} f']^{-1}. \qquad (5.70)$$

The eigenvectors in V_{mm} are normalized with respect to the corresponding m x m block of S, that is, $V_{mm}^\dagger S_{mm} V_{mm} = 1_m$, and therefore, the inverse of the transformation $V^\dagger$ is

$$V^{\dagger -1} = \begin{bmatrix} S_{mm} V_{mm} & 0 \\ 0 & 1_{\bar{B}} \end{bmatrix}. \qquad (5.71)$$

Using this, the transformation (5.70) becomes,

$$f = \begin{bmatrix} (SV)_{\bar{m}A} + (SV)_{\bar{m}\bar{m}} f_{\bar{m}} \\ f_{\bar{B}}' \end{bmatrix} [(SV)_{AA} + (SV)_{A\bar{m}} f_{\bar{m}}]^{-1}, \qquad (5.72)$$

where the operator f' in the partially diagonalizing basis has been written

$$f' = \begin{bmatrix} f_{\bar{m}}' \\ f_{\bar{B}}' \end{bmatrix}.$$

The evaluation of the right hand side of eq. (5.72) requires of the order of $n_A n_B^2$ operations when $n_B >> n_A$. No direct handling or manipulation of an n x n matrix is required.

5.3.f Test Calculations With Overlap

A series of test calculations have been carried out using algorithms SDNRS, QDNRS, DGNS, and FGNS. In the model problems examined, the basic matrix H was the same as that used previously in the calculations without overlap, namely, with diagonal elements equal to the first n odd integers, and the off-diagonal elements all unity. The overlap matrices were of the form

$$S = \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-1} \\ \alpha & 1 & \alpha & \cdots & \alpha^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \alpha^{n-1} & \alpha^{n-2} & & \cdots & 1 \end{bmatrix}, \qquad (5.73)$$

This matrix is positive definite for all $\alpha < 1$. It resembles the overlap matrix for a linear chain of atoms, with overlap falling off with distance ($S_{ij} = \alpha^{|i-j|}$), but it also serves to model a configuration interaction calculation in which overlaps fall off with energy differences. For $\alpha = 0$, the orthonormal case is recovered, while as α approaches the maximum value unity, the eigenvalue equation, (2.101) becomes ill-conditioned. At $\alpha = 1$, all but one of the eigenvalues of S vanish, and the eigenvalue equation is singular.[3]

All other computational details are the same as for the

[3]For large n, the eigenvalues of (5.73) will not differ significantly from those for the corresponding "circulant" matrix (Rutherford, 1949) of the same dimension. For such a matrix, it can be shown generally that the eigenvalues range between $(1-\alpha)/(1+\alpha)$ and $(1+\alpha)/(1-\alpha)$, with the greatest concentration of eigenvalues near the lower end as α approaches unity.

calculations of Table 5.1. As before, S_A is the space of the basis functions corresponding to the first n_A diagonal elements of H. The results of three series of calculations are given in Tables 5.2 - 5.4, which include information on the effect of varying the initial approximations to f and h, and of prediagonalization. As before, the rates of convergence decreased only slowly with increasing dimension of the eigenvalue problem.

Table 5.2 shows how convergence rates vary as the overlap integral α increases from zero to 0.9, with n_A and n held constant. It is seen that all the calculations diverge between $\alpha = 0.4$ and $\alpha = 0.6$, except those with prediagonalization, for which the upper limit overlap is between 0.8 and 0.9. In this case, the rate of convergence of the algorithms SDNRS and QDNRS at first deteriorates only slowly, but changes abruptly to divergence between $\alpha = 0.8$ and $\alpha = 0.9$. For DGNS and FGNS, the deterioration of convergence rates and onset of divergence is more gradual.

Initializations in this series of calculations tend to favour convergence to eigenspaces corresponding to the n_A lowest eigenvalues. Except where noted, convergence in all cases was to the space S_A' corresponding to the n_A lowest eigenvalues. The only other combination of eigenvalues obtained from a convergent calculation consisted of the $n_A - 1$ lowest eigenvalues together with the largest eigenvalue. A possible explanation of this is that the iterative corrections are

TABLE 5.2 RATES of CONVERGENCE [e]

				SDNRS				QDNRS				DGNS				FGNS			
n_A	n	Hii	α	①[a]	②[b]	③[c]	④[d]	①	②	③	④	①	②	③	④	①	②	③	④
5	20	1, 3, 5 . . . , 39	0.0	0.29				0.30				0.37				0.29 (4)			
"	"	"	0.1	0.29	0.29	0.18	0.30	0.29	0.29	0.18	0.30	0.22	0.22	0.13	0.21	0.28	0.28	0.17	0.28
"	"	"	0.2	0.46	0.46	0.22	0.46	0.46	0.46	0.22	0.47	0.57	0.57	0.27	0.57	0.33	0.33	0.22	0.30
"	"	"	0.4	0.71	0.70	0.33	0.71	0.71	0.70	0.33	0.71	0.90	0.90	0.37	1.00[h]	0.98[f] (4)	0.87 (3)	0.32	0.98[f] (4)
"	"	"	0.6	div.[m]	div.	0.41	div.	div.	div.	0.41	div.	0.90[f] (3)	0.99[g] (8)	0.60	0.96[f]	0.96[i]	0.98[g] (4)	0.59	div.
"	"	"	0.7			0.52				0.52				0.89[j]				0.88	
"	"	"	0.8			0.51 (4)				0.51 (4)				0.99				0.99[K]	
"	"	"	0.9			div.				div.				div.				1.06[L]	

TABLE 5.2 (continued)

[a] $f_o = 0$, no basis change.

[b] f_o calculated from eigenvectors of upper diagonal 10 x 10 block.

[c] $f_o = 0$, iterative calculation carried out in basis diagonalizing upper diagonal 10 x 10 block.

[d] $f_o = f$ calculated for $\alpha=0$ (non-overlap case).

[e] Blank spaces indicate no calculation was carried out; bracketed numbers indicate number of iterations before linear convergence is established (error constants were determined neglecting these points).

[f] Convergence to lowest four eigenvalues, and largest eigenvalue, $\|D\|$ slowly decreasing.

[g] $\|D\| = 0.34$ after 50 iterations -- but convergence apparently is to lowest 5 eigenvalues.

[h] $\|D\|$ slowly increasing -- calculation may be divergent.

[i] $\|D\| \sim 10^8$ after 50 iterations and is oscillatory.

[j] Calculation divergent if iterative scheme restarted in original basis.

[k] $\|D\|$ diverges slowly in partially diagonalizing basis, but begins to converge in original basis.

[l] Calculation is possibly converging slowly after 28 iterations.

[m] Indicates that iterative calculation is divergent, $\|D\|$ increases over most of first 50 iterations.

reflecting the situation which would occur if α actually were greater than one. As α increases through unity, the largest eigenvalue increases to $+\infty$, re-emerges from $-\infty$, and becomes the new lowest eigenvalue, while the corresponding eigenvector direction presumably changes little.

Table 5.2 shows that the rate of convergence, once linear convergence is established, is effectively independent of the starting f. An improvement of the initial f may slightly reduce the overall number of iterations, but does not increase the rate of convergence.

These results also illustrate the substantial improvement in convergence rates (as well as the substantially wider range of values of α over which convergence is obtained) resulting from prediagonalization of a small block of H. This improvement is not due to the improved starting approximation, but to the change of basis.

Table 5.3 gives rates of convergence for a set of calculations in which the basis space S_A does not correspond to the n_A smallest diagonal elements of H. It is seen that convergence rates are very poor indeed, and that a large proportion of the calculations did not converge at all. When convergent, DGNS and FGNS did not give the lowest n_A eigenvalues in these calculations. However, except with prediagonalization, SDNRS and QDNRS still give the lowest n_A eigenvalues; the rates of convergence being far superior to those of DGNS and FGNS in these cases.

TABLE 5.3 RATES of CONVERGENCE[a]

				SDNRS				GDNRS				DGNS				FGNS				$n_A=1$
n_A	n	Hii	α	①[b]	②[c]	③[d]	④[e]	①	②	③	④	①	②	③	④	①	②	③	④	nesbet
5	20	1,3,...11,9... 37,39	0.2	0.34 (4)	0.34 (7)	0.26[f]	0.31 (8)	0.34 (7)	0.34 (14)	0.26[f]	0.34 (15)	0.72[f]	0.73[f]	0.27[f]	0.71[f]	0.63[f]	0.62[f]	0.27[f]	0.60[f]	0.30
5	20	1,3,...11,9,7,13...37, 39	0.2	div.	div.	0.70[g] (14)	div.	div.	div.	0.70 (14)	1.00[g]	1.00 (4)	0.993 (5)	0.95[g]	1.03[h] (6)	0.97[g] (2)	0.97[g]	0.95[g]	0.97[g] (8)	0.30
5	20	1,3,...11,13,7,9...37, 39	0.2	0.45 (31)	div.	0.34[i]	div.	0.40 (35)	div.	0.34[i]	div.	0.97[i] (3)	0.97[i]	0.80[i] (4)	0.97[i]	0.97[i] (3)	0.96[i]	0.73[i]	0.96[i]	0.30
5	20	1,3,...13,11,9,7,...37, 39	0.2	div.	div.	0.28[i]	div.	div.	div.	0.28[i]	div.	0.95[i] (3)	0.95[i]	0.63[j] (5)	0.96[i] (3)	0.95[i] (5)	0.93[i]	0.57[j] (7)	0.89[i]	0.30

TABLE 5.3 (continued)

[a]Bracketed numbers indicate the number of iterations before linear convergence is established. These points are ignored in calculating the rates of convergence. Convergence is to the lowest n_A eigenvalues in all calculations, unless otherwise noted.

[b]$f_o = 0$, no basis change.

[c]f_o calculated from eigenvectors of upper diagonal 10 x 10 block.

[d]$f_o = 0$, iterative calculation carried out in basis diagonalizing upper diagonal 10 x 10 block.

[e]$f_o = f$ calculated for $\alpha = 0$ (non-overlap case).

[f]converges to eigenvalues #1, 2, 3, 4, and 6.

[g]slow convergence (or possibly slow divergence), eigenvalues after 50 iterations apparently #1, 2, 3, 5, and 6.

[h]slow oscillation in $\|D\|$, eigenvalues after 50 iterations are #1, 2, 3, 5, and 6.

[i]convergence apparently to eigenvalues #1, 2, 3, 6, and 7 -- modification (3) unstable after transformation back to original basis.

[j]convergence to eigenvalues #1, 2, 3, 6, and 7 -- convergence continues after back transformation.

On the other hand, comparatively good rates of convergence to eigenspaces S_A' not corresponding to the lowest n_A eigenvalues of H were obtained if the iterative calculation was carried out after prediagonalization ($m = 2n_A$ in Table 5.3). In fact, it is clear that prediagonalization is necessary if higher eigenvalues are to be obtained reliably and efficiently.

If the resulting back-transformed f-operator was used as an initial approximation for a calculation in the original basis, the calculation diverged rapidly in a number of cases, even though this initial approximation to f yielded values for $\|D^{(2)}\|$ or $[\sum_{\sigma,r} (G_{\sigma r}^2 + g_{\sigma r}^2)]^{\frac{1}{2}}$ which were less than 10^{-12}. This indicates that in certain cases, no improvement in the starting approximation for f (without also changing the basis) will lead to convergence -- the asymptotic error constants defined in Appendix 5 must be predominantly greater than one, leading to an increase in the errors $e_{\sigma r}$ in $f_{\sigma r}$, regardless of how small the $e_{\sigma r}$ are initially. By transforming to the partially diagonalizing basis, the most important of these error constants are reduced to zero, and convergence occurs.

The calculations using the generalized Nesbet algorithms, DGNS and FGNS, frequently consisted of a few initial iterations during which $\|D^{(2)}\|$ changed relatively rapidly, either increasing, or decreasing, or both, followed by a region of apparent convergence in which $\|D^{(2)}\|$ decreased extremely slowly. In such cases, it was not unusual for $\|D^{(2)}\|$ to decrease by only one part in 10^4 - 10^5 per iteration. In

many of these calculations where convergence was very slow, certain of the n_A eigenvalues of the effective operator G_A were surprisingly accurate in view of the large value of $\|D^{(2)}\|$. In several cases, with final $\|D^{(2)}\|$ in the range 0.1 - 0.2, those eigenvalues of G_A belonging to the lowest n_A of H were obtained accurate to eight or more figures, whereas the remaining eigenvalues of G_A were much less accurate. The poor convergence is thus apparently associated with determining that part of S_A' corresponding to eigenvalues not among the lowest n_A.

For convergent calculations, also, the plots of $\log(\|G_{BA}\|^2 + \|g_{BA}\|^2)^{\frac{1}{2}}$, or $\log \|D^{(2)}\|$ as a function of iteration number often exhibited "induction periods" before linear convergence was established. Figures 5.1 and 5.2 show such plots for two groups of calculations. The shape and length of these induction periods depends strongly on the initial f. Typically, only 5 - 10 iterations are involved -- the example in Fig. 5.2 is an extreme case in which over 30 iterations are required before convergence finally occurs. As indicated in Table 5.3, the two converging calculations in Fig. 5.2 are to different eigenspaces S_A'. Figure 5.1 illustrates clearly the independence of the rate of convergence on the starting approximation of f.

Table 5.4 gives rates of convergence for a series of calculations in which the first n_A or n_A+1 diagonal elements of H are nearly equal. When the first n_A diagonal elements of H are well separated from the rest, convergence is rapid.

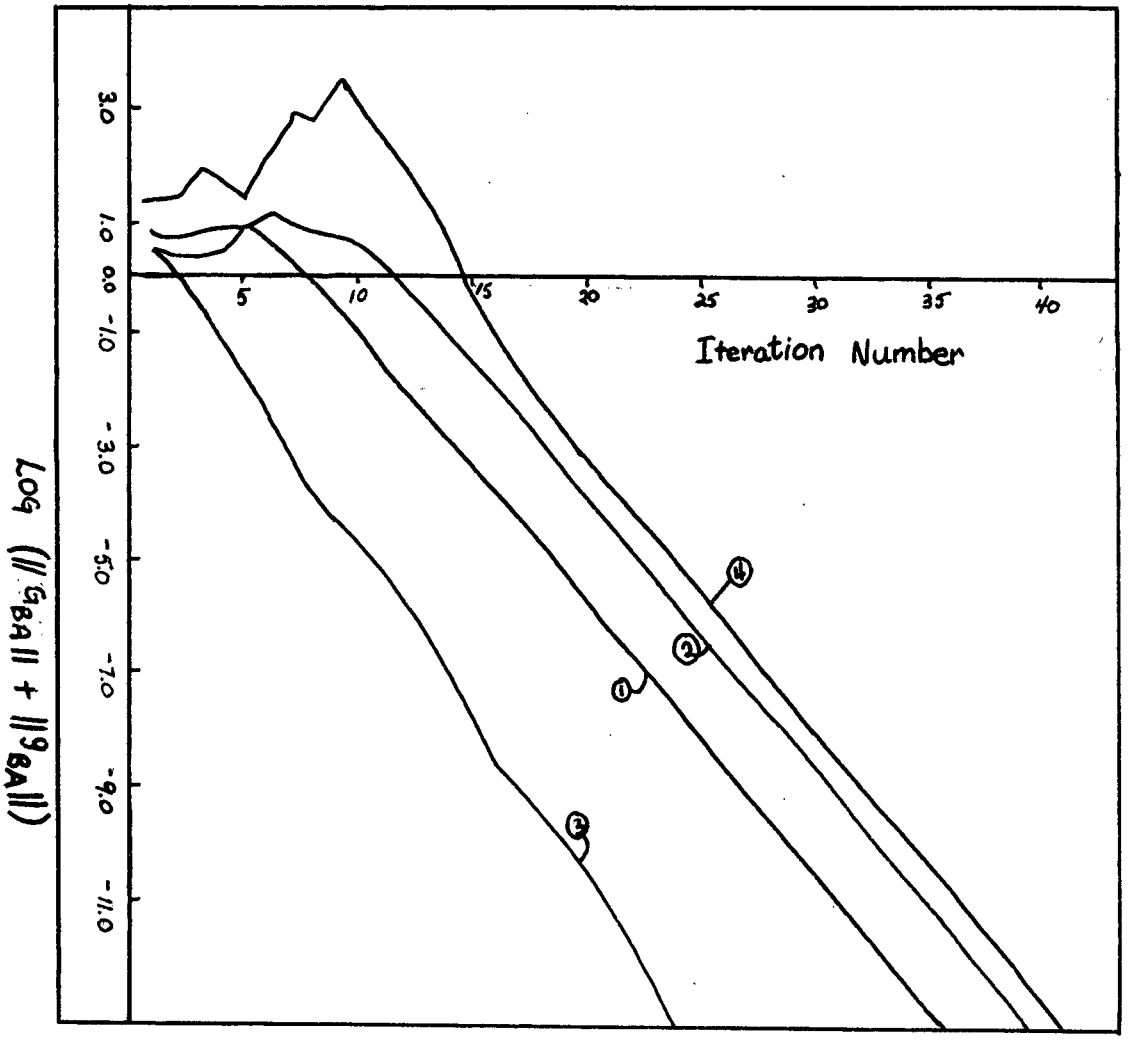

FIGURE 5.1 Algorithm SDNRS

H_{ii} = 1, 3, 5, 7, 11, 9, 13, 15, ..., 39.

n_A = 5, n = 20, α = 0.2

1. f_o = 0, iterative calculation in original basis.
2. f_o calculated from eigenvectors of upper 10 x 10 block of H.
3. f_o = 0, iterative calculation carried out in basis diagonalizing upper diagonal 10 x 10 block.
4. f_o = f calculated for α = 0 (non-overlap case).

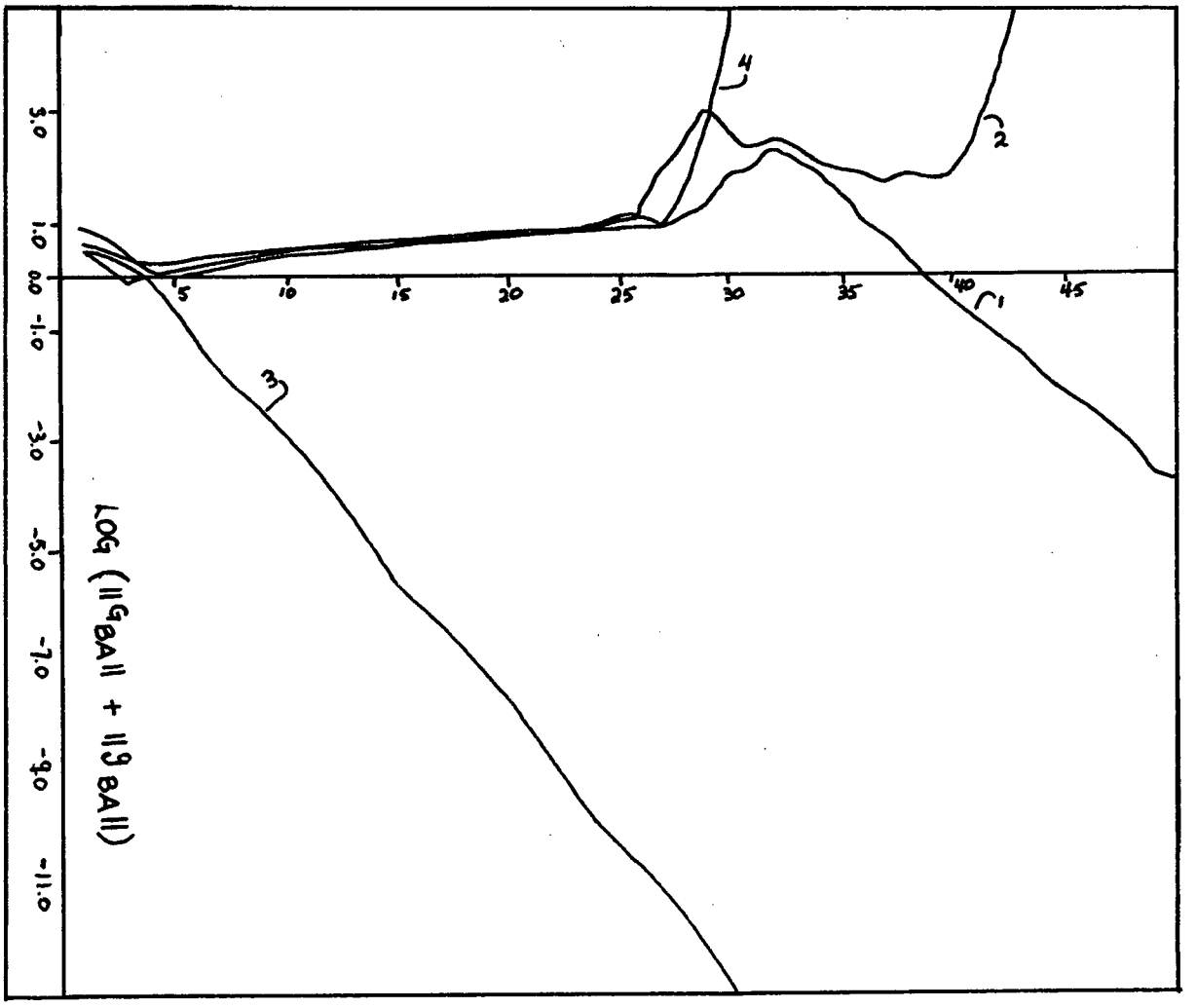

FIGURE 5.2 Algorithm SDNRS

H_{ii} = 1, 3, 5, 11, 13, 7, 9, 15, 17, ..., 39; n_A=5, n=20, α=0.2.

1. f_o = 0, iterative calculation in original basis.
2. f_o calculated from eigenvectors of upper 10 x 10 block of H.
3. f_o = 0, iterative calculation carried out in basis diagonalizing upper diagonal 10 x 10 block.
4. f_o = f calculated for α = 0 (non-overlap case).

TABLE 5.4 RATES of CONVERGENCE[a]

n_A	n	Hii	α	SDRS				GDNRS				DGNS				FGNS				$n_A=1$ nested
				①[b]	②[f]	③[d]	④[e]	①	②	③	④	①	②	③	④	①	②	③	④	
5	20	1, 1.1, 1.2, 1.3, 1.4, 11, 13, ...	0.2	0.18	0.18	0.12	0.18	0.18	0.18	0.12	0.18	0.14	0.13	0.11	0.14	0.14	0.145	0.11	0.14	0.94[f]
5	20	1, 1.1, 1.2, 1.3, 1.4, 1.5, 13, 15...	0.2	0.19	0.17	0.16	0.19	0.18	0.17	0.16	0.19	1.02 (5)	0.40	0.16	0.41	0.93 (5)	0.38	0.16	0.35	0.91[f] (26)
5	20	1, 1.1, 1.2, 1.3, 1.4, 3, 5, 7...	0.2	0.40	0.40	0.15	0.40	0.40	0.40	0.15	0.40	0.43	0.43	0.20	0.44	0.33	0.33	0.20	0.33	0.89[f] (21)
5	20	1, 1.1, 1.2, 1.3, 1.4, 1.5, 3, 5...	0.2	0.48	0.49	0.25	0.47 (4)	0.48	0.49	0.25	0.47 (4)	0.95[g] (8)	0.49	0.25	0.50	1.00[h] (7)	0.44	0.25	0.45	0.92[f] (27)
5	20	1, 1.1, 1.2, 1.3, 1.4, 11, 13...	0.4	0.24	0.24	0.11	0.25	0.24	0.24	0.11	0.25	0.13	0.13	0.10	0.13	0.13	0.13	0.10	0.13	0.88[f] (13)
5	20	1, 1.1, 1.2, 1.3, 1.4, 1.5, 13...	0.4	0.34	0.34	0.25	0.35	0.34	0.34	0.25	0.34	0.33	0.35	0.25	0.35	0.94[i] (7)	0.31	0.25	0.30	0.92[f] (35)
5	20	1, 1.1, 1.2, 1.3, 1.4, 3, 5, 7...	0.4	0.36	0.37	0.18	0.37	0.37	0.37	0.18	0.37	0.28	0.28	0.16	0.28	0.31 (9)	0.31	0.16	0.31	0.89[f] (16)
5	20	1, 1.1, 1.2, 1.3, 1.4, 1.5, 3, 5...	0.4	0.57	0.58 (5)	0.38	0.57	0.57	0.58 (4)	0.38	0.57	0.50	0.49	0.38	0.49	1.00[i] (11)	0.52	0.38	0.52	0.92[f] (23)

TABLE 5.4 (continued)

[a] All calculations converge to the lowest 5 eigenvalues unless otherwise noted--bracketed numbers indicate the number of iterations before linear convergence rates are established.

[b] $f_o = 0$, no basis change.

[c] f_o calculated from eigenvectors of upper diagonal 10 x 10 block.

[d] $f_o = 0$, iterative calculation carried out in basis diagonalizing upper diagonal 10 x 10 block.

[e] $f_o = f$ calculated for $\alpha = 0$ (non-overlap case).

[f] converges to second lowest eigenvalue--in cases with long induction period, there is a shallow minimum in $\|D\|$ after between 10 and 20 iterations, after which it increases to a maximum before decreasing again.

[g] apparently converging to eigenvalues #1, 2, 3, 4, 9; after 50 iterations, $\|D\| = 14.8$.

[h] apparently converging to eigenvalues #1, 2, 3, 4, and 8; after 50 iterations, $\|D\| = 5.9$.

[i] apparently converging to eigenvalues # 1, 2, 3, 4, and 7; after 50 iterations, $\|D\| = 5.8$.

The denominators in the iteration formulas are large, so that the asymptotic error constants are small, and the iterative calculations well-conditioned. When the first diagonal element of H_{BB} is close to diagonal elements of H_{AA} (and $\alpha = 0.2$), the rates of convergence of the algorithms SDNRS and QDNRS are virtually unaffected, whereas, those of DGNS and FGNS deteriorate to a much greater extent. The greater the number of diagonal elements of H_{BB} near those of H_{AA}, the slower the rate of convergence, as evidenced by the poor convergence here when $n_A = 1$ (Nesbet algorithm). The Nesbet algorithm is apparently converging here to the second lowest eigenvalue of H at the 50^{th} iteration, but the convergence is very slow. In the calculations reported in Table 5.4, convergence is normally to the space S_A' corresponding to the lowest n_A eigenvalues of H. The only exceptions are the most poorly converging generalized Nesbet calculations, in which the space S_A' corresponds to the lowest $n_A - 1$ plus the $(n_A+1)^{st}$, or $(n_A+2)^{nd}$ eigenvalues of H. Generally, the rates of convergence shown in Table 5.4 decrease as the overlap α increases.

From Tables 5.2 - 5.4, it is seen that the two algorithms, SDNRS and QDNRS are generally more reliable than the generalized Nesbet algorithms. However, when convergence occurs, the rates of all algorithms are similar. Again, algorithm SDNRS is easier to program efficiently than the others, and, unless the size of the problem makes the extra storage to hold h a critical factor, this is probably the algorithm of choice. While the

eigenvalue problem is more difficult with overlap, these methods are useful, especially with prediagonalization. It appears that attempts to obtain improved starting approximations for f are not of great value, and in particular, the solution of the corresponding problem for an orthonormal basis was frequently the worst starting approximation tried in these calculations.

5.4 Multiple Partitioning

Finally, the possibilities for efficient, practical iteration procedures for solving the m x m partitioning equations of chapter 4 are considered. The basic strategy is again to obtain efficient linearly convergent iterative schemes by approximating the second order convergent Newton-Raphson equations corresponding to the nonlinear system to be solved.

Three sets of equations were introduced in chapter 4 for the determination of the off-diagonal blocks of $\hat{T}$ in an orthonormal basis. They are:

$$(1)\quad D_{JI}^{(1)}(\hat{T}) = H_{JI} + \sum_{\substack{K=1 \\ K\neq I}}^{m} H_{JK}f_{KI} + f_{JI}\hat{H}_I^{(1)} = 0, \qquad (5.74)$$

$$(I,J = 1,\ldots, m,\ I\neq J),$$

$$(2)\quad D_{JI}^{(2)}(\hat{T}) = H_{JI} + \sum_{\substack{K=1 \\ K\neq I}}^{m} H_{JK}f_{KI} + f_{JI}\hat{H}_I^{(2)} = 0, \qquad (5.75)$$

$$(I,J = 1,\ldots, m,\ I\neq J),$$

and the pair of systems,

$$(3)\quad G_{JI}(\hat{T}) = H_{JI} + \sum_{L\neq I} H_{JL}f_{LI} + \sum_{L\neq J} f_{LJ}^{\dagger}H_{LI} + \sum_{\substack{L\neq J \\ K\neq I}} f_{LJ}^{\dagger}H_{LK}f_{KI} = 0,$$

$$g_{JI}(\hat{T}) = f_{IJ}^{\dagger} + f_{JI} + \sum_{\substack{L\neq I \\ L\neq J}} f_{LJ}^{\dagger}f_{JI} = 0, \qquad (5.76b)$$

$$(I < J = 1,\ \ldots,\ m).$$

A fourth set of equations, intermediate between (5.74), (5.75) and eqs. (5.76) are

$$D^{(3)}_{JI}(\hat{T}) = (g^{-1}G)_{JI} = 0, \quad (I,J = 1, \ldots, m, \; I \neq J), \qquad (5.77)$$

which arise from the condition that $(S\hat{T})^{-1}H\hat{T}$ be block diagonal. These four sets of equations have the same solutions, however, they lead to algorithms which are quite different.

While the iteration schemes derived from eqs. (5.74) - (5.76) are very similar to those developed for a 2 x 2 partitioning, a major difference is that the complexity of the orthogonality conditions (4.8) makes it a practical impossibility to explicitly eliminate half the elements of the off-diagonal blocks of $\hat{T}$ before solving one of these sets of equations for the remaining elements. In all algorithms, therefore, it is assumed that the elements of all m(m-1) off-diagonal blocks of $\hat{T}$ are to be determined. A description of possible algorithms for solving the four sets of equations above is given in Appendix 6. No numerical testing of these algorithms has been carried out.

The determination of the matrix elements of the off-diagonal blocks of $\hat{T}$ for an m x m partitioning in a non-orthonormal basis involves complications only in detail due to the presence of the overlap matrix. Equations for this case have been given in chapter 4, and they may be handled in essentially the same way as those defining equations given above.

CHAPTER 6

PERTURBATION THEORY

O polish'd perturbation! golden care!
That keep'st ports of slumber open wide,
To many a watchful night!
(Shakespeare, King Henry IV, Part II)

"I wouldn't lose any sleep over it"
(wise old saying)

6.1 Introduction

Only the simplest problems in quantum mechanics can be solved exactly. As a result, perturbation expansions of some sort are involved in many quantum mechanical calculations.

Perturbation series for effective operators are useful in treating a set of degenerate, or nearly degenerate levels with one or more perturbation parameters, especially when the degeneracy is not split in first order. Effective operator perturbation series are also useful in developing physical pictures, as, for example, in uncoupling the Dirac equation to obtain equations for electrons only.

In this chapter, perturbation series are developed for the effective operators $\hat{H}_A$, G_A, and $\tilde{H}_A$, defined for a 2 x 2 partitioning in terms of the operator f. These series can be derived straightforwardly because of the relatively simple algebraic form of the relations defining the operators. The absence of constraints or auxiliary conditions on f makes possible efficient computational schemes for automatic sequential calculation of the terms in the perturbation series to arbitrary high order. The perturbation formulas are not complicated by degeneracies at any order, so long as all eigenfunctions in a given degenerate set in zero order are partitioned into the same space. In fact, as will be seen below, the presence of degeneracy tends to simplify the use of these perturbation series.

Two examples are presented to illustrate the perturbation

formulas derived. These are the uncoupling of the Dirac equation for a spin-½ particle, and the construction of a nuclear spin hamiltonian in esr theory. Perturbation of the projection $P_A^{\bullet}$ which, in molecular orbital theory, becomes the one-particle density operator, is considered in the following chapter, and the formalism is extended there to the related, but more complicated, self-consistent field molecular orbital problem.

6.2 2 x 2 Partitioning -- Orthonormal Basis

6.2.a General Discussion

A perturbation formalism based on the material presented in chapter 2, and in particular, on the eigenvalue equation (2.1), will be considered first. It is assumed that the hamiltonian H can be written as an infinite series

$$H = \sum_{n=0}^{\infty} H^{(n)}, \tag{6.1}$$

where the perturbation parameter or parameters are to be considered to be included implicitly in symbols like $H^{(n)}$, which is of order n in the perturbation.

The operator f is written as an infinite series,

$$f = \sum_{n=0}^{\infty} f^{(n)}. \tag{6.2}$$

Substitution of these two series into the condition D(f) = 0, eq. (2.16), defining f, yields the series

$$D(f) = \sum_{n=0}^{\infty} D^{(n)}(f) = 0, \tag{6.3a}$$

where

$$D^{(n)}(f) = H_{BA}^{(n)} + \sum_{j=0}^{n} \left(H_{BB}^{(n-j)} f^{(j)} - f^{(j)} H_{AA}^{(n-j)} - f^{(j)} \sum_{i=0}^{n-j} H_{AB}^{(n-i-j)} f^{(i)} \right) \tag{6.3b}$$

$$= H_{BA}^{(n)} + \sum_{j=0}^{n} \left(H_{BB}^{(n-j)} f^{(j)} - f^{(j)} \hat{H}_{A}^{(n-j)} \right). \tag{6.3c}$$

The series for $\hat{H}_A$ is given below. Since (6.3a) is implicitly a power series in one or more arbitrary perturbation parameters, D(f) will vanish as a whole only if each term vanishes. Thus,

a hierarchy of perturbation equations is obtained,

$$D^{(n)}(f) = 0, \qquad n = 0,1,2,\ldots , \tag{6.4}$$

from which the $f^{(n)}$ can be determined.

The zero order condition is formally

$$D^{(0)}(f)=H_{BA}^{(0)}+H_{BB}^{(0)}f^{(0)}-f^{(0)}H_{AA}^{(0)}-f^{(0)}H_{AB}^{(0)}f^{(0)}= 0, \tag{6.5}$$

which is just the original condition defining f for the zero order operator $H^{(0)}$. Unless $H^{(0)}$ is block diagonal, $f^{(0)}$ will not vanish, and consequently, the $D^{(n)}$ will depend on $f^{(0)}$, taking the form

$$\begin{aligned} D^{(n)}(f)&=(H_{BB}^{(0)}-f^{(0)}H_{AB}^{(0)})f^{(n)}-f^{(n)}(H_{AA}^{(0)}+H_{AB}^{(0)}f^{(0)}) \\ &\quad + A^{(n)}(f^{(n-1)}, \ldots, f^{(0)}) \\ &= \hat{H}_B^{(0)\dagger}f^{(n)}-f^{(n)}\hat{H}_A^{(0)}+A^{(n)} = 0, \end{aligned} \tag{6.6}$$

where $A^{(n)}$ is a quantity depending on terms in the series for f of order n-1 or lower. General solutions for the n_An_B-dimensional system of simultaneous linear equations, (6.6), cannot usually be written down, and the $f^{(n)}$ must therefore be determined by numerical methods.

If $H^{(0)}$ is block diagonal, then $f^{(0)} = 0$, and eqs. (6.6) become

$$H_{BB}^{(0)}f^{(n)}-f^{(n)}H_{AA}^{(0)}=-A^{(n)}(f^{(n-1)}, \ldots, f^{(1)}), \tag{6.7}$$

which is again a system of n_An_B simultaneous linear equations, which, in general, must also be solved numerically. However, these equations are considerably simpler than eqs. (6.6).

Finally, if $H^{(0)}$ is diagonal, eqs. (6.7) reduce to

$$(H^{(0)}_{\sigma\sigma} - H^{(0)}_{rr})f^{(n)}_{\sigma r} = -A^{(n)}_{\sigma r}, \tag{6.8a}$$

and in this case, the solution can be given explicitly,

$$f^{(n)}_{\sigma r} = \frac{A^{(n)}_{\sigma r}}{H^{(0)}_{rr} - H^{(0)}_{\sigma\sigma}} \quad . \tag{6.8b}$$

Here, again, Greek letters refer to basis elements in the subspace S_B, and Roman letters to basis elements in S_A.

In general, for $f^{(0)} = 0$, the $A^{(n)}$ are given by

$$A^{(n)} = H^{(n)}_{BA} + \sum_{j=1}^{n-1}(H^{(n-j)}_{BB}f^{(j)} - f^{(j)}H^{(n-j)}_{AA}) - \sum_{j=1}^{n-2}\sum_{i=1}^{n-j-1} f^{(j)}H^{(n-i-j)}_{AB}f^{(i)}, \tag{6.9}$$

which is obtained by deleting terms depending on $f^{(n)}$ and $f^{(0)}$ from eq. (6.3b). When the series for H contains only a few terms -- for example, when $H = H^{(0)} + H^{(1)}$ only -- it is more useful to group terms in the $D^{(n)}$ and $A^{(n)}$ according to the order of the hamiltonian, H, rather than f, in the term. For $A^{(n)}$, this gives

$$A^{(n)} = H^{(n)}_{BA} + \sum_{k=1}^{n-1}(H^{(k)}_{BB}f^{(n-k)} - f^{(n-k)}H^{(k)}_{AA} - \sum_{j=1}^{n-k-1} f^{(j)}H^{(k)}_{AB}f^{(n-j-k)}) \tag{6.10a}$$

$$= H^{(n)}_{BA} + \sum_{k=1}^{n-1}(H^{(k)}_{BB}f^{(n-k)} - f^{(n-k)}\hat{H}^{(k)}_{A}). \tag{6.10b}$$

Table 6.1 lists the first few members of the perturbation hierarchy, $D^{(n)}(f)$, for the case $f^{(0)} = 0$. Explicit formulas, in the format of eq. (6.10a), for low order $A^{(n)}$ are obtainable from Table 6.1 by deleting the $f^{(n)}$ dependent terms in the $D^{(n)}$.

Perturbation formulas, in terms of the $f^{(n)}$ and $H^{(n)}$, for all of the other quantities defined in sections 2.1 and 2.2 follow directly from their definitions. The formulas for them presented in the remainder of this section apply when $f^{(0)} = 0$.

Using eq. (2.65a), the series for the effective operator $\hat{H}_A$ is found to be

$$\hat{H}_A = \sum_{n=0}^{\infty} \hat{H}_A^{(n)}, \tag{6.11a}$$

where

$$\hat{H}_A^{(0)} = H_{AA}^{(0)},$$

$$\hat{H}_A^{(1)} = H_{AA}^{(1)}, \tag{6.11b}$$

and

$$\hat{H}_A^{(n)} = H_{BA}^{(n)} + \sum_{j=1}^{n-1} H_{AB}^{(n-j)} f^{(j)}, \quad (n > 1).$$

For G_A, given in eq. (2.67b), one obtains

$$G_A = \sum_{n=0}^{\infty} G_A^{(n)}, \tag{6.12a}$$

where,

$$G_A^{(0)} = H_{AA}^{(0)},$$

$$G_A^{(1)} = H_{AA}^{(1)}, \tag{6.12b}$$

and,

$$G_A^{(n)} = H_{AA}^{(n)} + \sum_{j=1}^{n-1} (H_{AB}^{(j)} f^{(n-j)} + f^{(n-j)\dagger} H_{BA}^{(j)}) + \sum_{j=1}^{n-2} \sum_{i=1}^{n-j-1} f^{(n-j-i)\dagger} H_{BB}^{(j)} f^{(i)}, \quad (n > 1).$$

The metric g_A has the very simple series

$$g_A = \sum_{n=0}^{\infty} g_A^{(n)}, \tag{6.13a}$$

where

$$g_A^{(0)} = 1_A,$$
$$g_A^{(1)} = 0, \tag{6.13b}$$

and

$$g_A^{(n)} = \sum_{j=1}^{n-1} f^{(n-j)\dagger} f^{(j)}.$$

If the hierarchy, (6.4), is used to explicitly eliminate the terms in $H_{BB}^{(j)}$ from eqs. (6.12b), the resulting series for G_A will be identical to that obtained by expansion of the relation $G_A = g_A \hat{H}_A$, eq. (2.70), that is

$$G_A^{(n)} = \sum_{j=0}^{n} g_A^{(j)} \hat{H}_A^{(n-j)}. \tag{6.14}$$

That the two expressions, (6.12) and (6.14), for G_A are equivalent if and only if the equations of the hierarchy (6.4) are satisfied, is in accord with the fact (section 3.1) that the two definitions, $G_A = (\hat{T}^\dagger H \hat{T})_{AA}$ and $G_A = g_A \hat{H}_A$, are equivalent if and only if $D(f) = 0$. An advantage of eqs. (6.12) is that they are the same whether or not the basis is orthonormal, whereas, any formulas incorporating the relations $D^{(n)}(f) = 0$ explicitly must be different in a non-orthonormal basis, since the condition $D(f)$ depends explicitly on the overlap matrix in that case.

Perturbation series for the powers, $g_A^{\pm\frac{1}{2}}$, of g_A, can be obtained in several ways outlined in Appendix 7. Given these,

TABLE 6.1 $D^{(n)}(f)$

$$D^{(1)}(f) = H_{BB}^{(0)} f^{(1)} - f^{(1)} H_{AA}^{(0)} + H_{BA}^{(1)}$$

$$D^{(2)}(f) = H_{BB}^{(0)} f^{(2)} - f^{(2)} H_{AA}^{(0)} + H_{BB}^{(1)} f^{(1)} - f^{(1)} H_{AA}^{(1)} + H_{BA}^{(2)}$$

$$D^{(3)}(f) = H_{BB}^{(0)} f^{(3)} - f^{(3)} H_{AA}^{(0)} + H_{BB}^{(1)} f^{(2)} - f^{(2)} H_{AA}^{(1)} - f^{(1)} H_{AB}^{(1)} f^{(1)} + H_{BB}^{(2)} f^{(1)} - f^{(1)} H_{AA}^{(2)} + H_{BA}^{(3)}$$

$$D^{(4)}(f) = H_{BB}^{(0)} f^{(4)} - f^{(4)} H_{AA}^{(0)} + H_{BB}^{(1)} f^{(3)} - f^{(3)} H_{AA}^{(1)} - f^{(2)} H_{AB}^{(1)} f^{(1)} - f^{(1)} H_{AB}^{(1)} f^{(2)} + H_{BB}^{(2)} f^{(2)} - f^{(2)} H_{AA}^{(2)} - f^{(1)} H_{AB}^{(2)} f^{(1)} + H_{BB}^{(3)} f^{(1)} - f^{(1)} H_{AA}^{(3)} + H_{BA}^{(4)}$$

$$D^{(5)}(f) = H_{BB}^{(0)} f^{(5)} - f^{(5)} H_{AA}^{(0)} + H_{BB}^{(1)} f^{(4)} - f^{(4)} H_{AA}^{(1)} - f^{(1)} H_{AB}^{(1)} f^{(3)} - f^{(2)} H_{AB}^{(1)} f^{(2)} - f^{(3)} H_{AB}^{(1)} f^{(1)} + H_{BB}^{(2)} f^{(3)} - f^{(3)} H_{AA}^{(2)} - f^{(1)} H_{AB}^{(2)} f^{(2)} - f^{(2)} H_{AB}^{(1)} f^{(1)} + H_{BB}^{(3)} f^{(2)} - f^{(2)} H_{AA}^{(3)} - f^{(1)} H_{AB}^{(3)} f^{(1)} + H_{BB}^{(4)} f^{(1)} - f^{(1)} H_{AA}^{(4)} + H_{BA}^{(5)}$$

TABLE 6.2 $\hat{H}_A^{(n)}$

$$\hat{H}_A^{(0)} = H_{AA}^{(0)}$$

$$\hat{H}_A^{(1)} = H_{AA}^{(1)}$$

$$\hat{H}_A^{(2)} = H_{AA}^{(2)}+H_{AB}^{(1)}f^{(1)}$$

$$\hat{H}_A^{(3)} = H_{AA}^{(3)}+H_{AB}^{(2)}f^{(1)}+H_{AB}^{(1)}f^{(2)}$$

$$\hat{H}_A^{(4)} = H_{AA}^{(4)}+ H_{AB}^{(3)}f^{(1)}+H_{AB}^{(2)}f^{(2)}+H_{AB}^{(1)}f^{(3)}$$

$$\hat{H}_A^{(5)} = H_{AA}^{(5)}+H_{AB}^{(4)}f^{(1)}+H_{AB}^{(3)}f^{(2)}+H_{AB}^{(2)}f^{(3)}+H_{AB}^{(1)}f^{(4)}$$

$$\hat{H}_A^{(6)} = H_{AA}^{(6)}+H_{AB}^{(5)}f^{(1)}+H_{AB}^{(4)}f^{(2)}+H_{AB}^{(3)}f^{(3)}+H_{AB}^{(2)}f^{(4)}+H_{AB}^{(1)}f^{(5)}$$

---

the series for the effective operator $\tilde{H}_A$ can be written down from eqs. (2.74) as

$$\tilde{H}_A = \sum_{n=0}^{\infty} \tilde{H}_A^{(n)}, \tag{6.15a}$$

where,

$$\tilde{H}_A^{(n)} = \sum_{i=0}^{n} \sum_{j=0}^{n} g_A^{\frac{1}{2}(i)} \hat{H}_A^{(n-i-j)} g_A^{-\frac{1}{2}(j)}, \tag{6.15b}$$

or

$$\tilde{H}_A^{(n)} = \sum_{i=0}^{n} \sum_{j=0}^{n} g_A^{-\frac{1}{2}(i)} G_A^{(n-i-j)} g_A^{-\frac{1}{2}(j)}. \tag{6.15c}$$

Explicit expressions for the lower order $\hat{H}_A^{(n)}$, $G_A^{(n)}$, and $g_A^{(n)}$, according to eqs. (6.11), (6.12), and (6.13), and for the $\tilde{H}_A^{(n)}$, are given in Tables 6.2 - 6.5. All three effective operators, $\hat{H}_A$, G_A, and $\tilde{H}_A$, are identical in zero and first order. In general, they differ in second and higher order.

TABLE 6.3 $G_A^{(n)}$

$$G_A^{(0)} = H_{AA}^{(0)}$$

$$G_A^{(1)} = H_{AA}^{(1)}$$

$$G_A^{(2)} = H_{AA}^{(2)} + H_{AB}^{(1)} f^{(1)} + f^{(1)\dagger} H_{BA}^{(1)} + f^{(1)\dagger} H_{BB}^{(0)} f^{(1)}$$

$$\begin{aligned} G_A^{(3)} = {} & H_{AA}^{(3)} + H_{AB}^{(2)} f^{(1)} + f^{(1)\dagger} H_{BA}^{(2)} \\ & + H_{AB}^{(1)} f^{(2)} + f^{(2)\dagger} H_{BA}^{(1)} + f^{(1)\dagger} H_{BB}^{(1)} f^{(1)} \\ & + f^{(2)\dagger} H_{BB}^{(0)} f^{(1)} + f^{(1)\dagger} H_{BB}^{(0)} f^{(2)} \end{aligned}$$

$$\begin{aligned} G_A^{(4)} = {} & H_{AA}^{(4)} + H_{AB}^{(3)} f^{(1)} + f^{(1)\dagger} H_{BA}^{(3)} \\ & + H_{AB}^{(2)} f^{(2)} + f^{(2)\dagger} H_{BA}^{(2)} + f^{(1)\dagger} H_{BB}^{(2)} f^{(1)} \\ & + H_{AB}^{(1)} f^{(3)} + f^{(3)\dagger} H_{BA}^{(1)} + f^{(1)\dagger} H_{BB}^{(1)} f^{(2)} + f^{(2)\dagger} H_{BB}^{(1)} f^{(1)} \\ & + f^{(1)\dagger} H_{BB}^{(0)} f^{(3)} + f^{(2)\dagger} H_{BB}^{(0)} f^{(2)} + f^{(3)\dagger} H_{BB}^{(0)} f^{(1)} \end{aligned}$$

TABLE 6.4 $g_A^{(n)}$

$$g_A^{(0)} = 1_A$$

$$g_A^{(1)} = 0$$

$$g_A^{(2)} = f^{(1)\dagger} f^{(1)}$$

$$g_A^{(3)} = f^{(2)\dagger} f^{(1)} + f^{(1)\dagger} f^{(2)}$$

$$g_A^{(4)} = f^{(3)\dagger} f^{(1)} + f^{(2)\dagger} f^{(2)} + f^{(1)\dagger} f^{(3)}$$

$$g_A^{(5)} = f^{(4)\dagger} f^{(1)} + f^{(3)\dagger} f^{(2)} + f^{(2)\dagger} f^{(3)} + f^{(1)\dagger} f^{(4)}$$

$$g_A^{(6)} = f^{(5)\dagger} f^{(1)} + f^{(4)\dagger} f^{(2)} + f^{(3)\dagger} f^{(3)} + f^{(2)\dagger} f^{(4)} + f^{(1)\dagger} f^{(5)}$$

TABLE 6.5 $\tilde{H}_A^{(n)}$

$$\tilde{H}_A^{(0)} = H_{AA}^{(0)}$$

$$\tilde{H}_A^{(1)} = H_{AA}^{(1)}$$

$$\tilde{H}_A^{(2)} = H_{AA}^{(2)} + H_{AB}^{(1)}f^{(1)} + \tfrac{1}{2}[f^{(1)\dagger}f^{(1)}, H_{AA}^{(0)}]_-$$

$$\tilde{H}_A^{(3)} = H_{AA}^{(3)} + H_{AB}^{(2)}f^{(1)} + H_{AB}^{(1)}f^{(2)} + \tfrac{1}{2}[f^{(1)\dagger}f^{(1)}, H_{AA}^{(1)}]_-$$
$$+ \tfrac{1}{2}[f^{(1)\dagger}f^{(2)} + f^{(2)\dagger}f^{(1)}, H_{AA}^{(0)}]_-$$

$$\tilde{H}_A^{(4)} = H_{AA}^{(4)} + H_{AB}^{(3)}f^{(1)} + H_{AB}^{(2)}f^{(2)} + H_{AB}^{(1)}f^{(3)}$$
$$+ \tfrac{1}{2}[f^{(1)\dagger}f^{(3)} + f^{(2)\dagger}f^{(2)} + f^{(3)\dagger}f^{(1)}, H_{AA}^{(0)}]_-$$
$$+ \tfrac{1}{2}[f^{(1)\dagger}f^{(2)} + f^{(2)\dagger}f^{(1)}, H_{AA}^{(1)}]_- + \tfrac{1}{2}[f^{(1)\dagger}f^{(1)}, H_{AA}^{(2)} + H_{AB}^{(1)}f^{(1)}]_-$$
$$- \tfrac{1}{8}(f^{(1)\dagger}f^{(1)})^2 H_{AA}^{(0)} + \tfrac{3}{8}f^{(1)\dagger}f^{(1)}H_{AA}^{(0)}f^{(1)\dagger}f^{(1)} - \tfrac{1}{4}H_{AA}^{(0)}(f^{(1)\dagger}f^{(1)})^2$$

Several alternative formulas for the terms of the perturbation series of these effective operators can be given, the usefulness of a given set depending on the situation. The formulas given above are not particularly well suited in some cases for the calculation of high order terms. Procedures for deriving alternative series are given in Appendix 7, along with a tabulation of some alternative formulas.

The perturbation series (6.9) through (6.15) are rather general in that they give the various effective operators ultimately in terms of the $f^{(n)}$ and $H^{(n)}$. However, as indicated in eq. (6.8), if $H^{(0)}$ is diagonal, the $f^{(n)}$ can be written explicitly in terms of the matrix elements of the $H^{(n)}$. Expansions corresponding to each of eqs. (6.8) through (6.15) in terms of only the perturbed operator H will be given in the next two sections.

6.2.b A-states Degenerate

Explicit perturbation formulas are especially simple when the eigenvalues of $H^{(0)}$ corresponding to the subspace S_A are all equal, say, to ϵ_A^o. In its eigenbasis, $H_{AA}^{(0)}$ is just a multiple of the unit matrix, and the $f^{(n)}$ are defined by

$$(H_{BB}^{(0)} - \epsilon_A^o 1_B) f^{(n)} = -A^{(n)}, \tag{6.16a}$$

with the solution

$$f^{(n)} = LA^{(n)}, \tag{6.16b}$$

where

$$L = (\epsilon_A^o 1_B - H_{BB}^{(0)})^{-1}, \tag{6.17}$$

is the reduced resolvent matrix evaluated at ϵ_A^o and restricted to S_B. The useful point here is that L is a matrix, not a supermatrix. If $H_{BB}^{(0)}$ is diagonal, the $f^{(n)}$ are given simply as the products of the matrices $A^{(n)}$ with the $n_B \times n_B$ diagonal matrix L, and thus, relatively simple matrix expressions can be given in terms of H only, for the various perturbation series derived in the previous subsection. Perturbation formulas of this type are given in Tables 6.6 - 6.11. When $H_{BB}^{(0)}$ is not diagonal, the $f^{(n)}$ must be determined by solving a system of $n_A n_B$ simultaneous equations.

Substitution of eq. (6.16b) into eqs. (6.10) yields

$$H_{AB}^{(j)} = A^{(j)\dagger} - \sum_{k=1}^{n-1} (H_{BB}^{(k)} f^{(n-k)} - f^{(n-k)} \hat{H}_A^{(k)})$$

$$= f^{(j)\dagger} L^{-1} - \sum_{k=1}^{n-1} (H_{BB}^{(k)} f^{(n-k)} - f^{(n-k)} \hat{H}_A^{(k)}), \tag{6.18}$$

which can be used to eliminate high order $f^{(j)}$ from perturbation

TABLE 6.6 $f^{(n)}$

(A-states degenerate)

$$L^{-1}f^{(1)} = H_{BA}^{(1)}$$

$$L^{-1}f^{(2)} = H_{BA}^{(2)} + H_{BB}^{(2)}f^{(1)} - f^{(1)}\hat{H}_A^{(1)}$$

$$L^{-1}f^{(3)} = H_{BA}^{(3)} + H_{BB}^{(2)}f^{(1)} - f^{(1)}\hat{H}_A^{(2)} + H_{BB}^{(1)}f^{(2)} - f^{(2)}\hat{H}_A^{(1)}$$

$$L^{-1}f^{(4)} = H_{BA}^{(4)} + H_{BB}^{(3)}f^{(1)} - f^{(1)}\hat{H}_A^{(3)} + H_{BB}^{(2)}f^{(2)} - f^{(2)}\hat{H}_A^{(2)} + H_{BB}^{(1)}f^{(3)} - f^{(3)}\hat{H}_A^{(1)}$$

TABLE 6.7 $f^{(n)}$

(A-states degenerate)

$$L^{-1}f^{(1)} = H_{BA}^{(1)}$$

$$L^{-1}f^{(2)} = H_{BA}^{(2)} + H_{BB}^{(1)}LH_{BA}^{(1)} - LH_{BA}^{(1)}H_{AA}^{(1)}$$

$$L^{-1}f^{(3)} = H_{BA}^{(3)} + H_{BB}^{(2)}LH_{BA}^{(1)} - LH_{BA}^{(1)}H_{AA}^{(2)} + H_{BB}^{(1)}LH_{BA}^{(2)} - LH_{BA}^{(2)}H_{AA}^{(1)} + (H_{BB}^{(1)}LH_{BB}^{(1)} - LH_{BA}^{(1)}H_{AB}^{(1)})LH_{BA}^{(1)} + \{L, H_{BB}^{(1)}\}_+ LH_{BA}^{(1)}H_{AA}^{(1)} + L^2H_{BA}^{(1)}H_{AA}^{(1)^2}$$

TABLE 6.8 $\hat{H}_A^{(n)}$ -- REDUCED FORMULAS

(A-states degenerate)

$$\hat{H}_A^{(0)} = H_{AA}^{(0)}$$

$$\hat{H}_A^{(1)} = H_{AA}^{(1)}$$

$$\hat{H}_A^{(2)} = H_{AA}^{(2)} + H_{AB}^{(1)} f^{(1)}$$

$$\hat{H}_A^{(3)} = H_{AA}^{(3)} + H_{AB}^{(2)} f^{(1)} + f^{(1)\dagger} H_{BA}^{(2)} + f^{(1)\dagger} H_{BB}^{(1)} f^{(1)} - f^{(1)\dagger} f^{(1)} H_{AA}^{(1)}$$

$$\begin{aligned}\hat{H}_A^{(4)} = {} & H_{AA}^{(4)} + H_{AB}^{(3)} f^{(1)} + f^{(1)\dagger} H_{BA}^{(3)} \\ & + H_{AB}^{(2)} f^{(2)} + f^{(1)\dagger} H_{BB}^{(2)} f^{(1)} - f^{(1)\dagger} f^{(1)} \hat{H}_A^{(2)} \\ & + f^{(1)\dagger} H_{BB}^{(1)} f^{(2)} - f^{(1)\dagger} f^{(2)} \hat{H}_A^{(1)}\end{aligned}$$

$$\begin{aligned}\hat{H}_A^{(5)} = {} & H_{AA}^{(5)} + H_{AB}^{(4)} f^{(1)} + f^{(1)\dagger} H_{BA}^{(4)} \\ & + H_{AB}^{(3)} f^{(2)} + f^{(2)\dagger} H_{BA}^{(3)} + f^{(1)\dagger} H_{BB}^{(3)} f^{(1)} - f^{(1)\dagger} f^{(1)} \hat{H}_A^{(3)} \\ & + f^{(1)\dagger} H_{BB}^{(2)} f^{(2)} + f^{(2)\dagger} H_{BB}^{(2)} f^{(1)} - f^{(1)\dagger} f^{(2)} \hat{H}_A^{(2)} \\ & - f^{(2)\dagger} f^{(1)} \hat{H}_A^{(2)} + f^{(2)\dagger} H_{BB}^{(1)} f^{(2)} - f^{(2)\dagger} f^{(2)} \hat{H}_A^{(1)} \\ & - \left[f^{(1)\dagger} f^{(3)}, \hat{H}_A^{(1)} \right]_-\end{aligned}$$

$$\begin{aligned}\hat{H}_A^{(6)} = {} & H_{AA}^{(6)} + H_{AB}^{(5)} f^{(1)} + f^{(1)\dagger} H_{BA}^{(5)} \\ & + H_{AB}^{(4)} f^{(2)} + f^{(2)\dagger} H_{BA}^{(4)} + f^{(1)\dagger} H_{BB}^{(4)} f^{(1)} - f^{(1)\dagger} f^{(1)} \hat{H}_A^{(4)} \\ & + H_{AB}^{(3)} f^{(3)} + f^{(2)\dagger} H_{BB}^{(3)} f^{(1)} - f^{(2)\dagger} f^{(1)} \hat{H}_A^{(3)} \\ & + f^{(1)\dagger} H_{BB}^{(3)} f^{(2)} - f^{(1)\dagger} f^{(2)} \hat{H}_A^{(3)} + f^{(2)\dagger} H_{BB}^{(2)} f^{(2)} \\ & + f^{(2)\dagger} f^{(2)} \hat{H}_A^{(2)} + f^{(1)\dagger} H_{BB}^{(2)} f^{(3)} - f^{(1)\dagger} f^{(2)} \hat{H}_A^{(3)} \\ & + f^{(2)\dagger} H_{BB}^{(1)} f^{(3)} - f^{(2)\dagger} f^{(3)} \hat{H}_A^{(1)} + \left[\hat{H}_A^{(1)}, f^{(1)\dagger} f^{(4)} \right]_-\end{aligned}$$

TABLE 6.9 $\hat{H}_A^{(n)}$

(A-states degenerate)

$$\hat{H}_A^{(0)} = H_{AA}^{(0)}$$

$$\hat{H}_A^{(1)} = H_{AA}^{(1)}$$

$$\hat{H}_A^{(2)} = H_{AA}^{(2)} + H_{AB}^{(1)}LH_{BA}^{(1)}$$

$$\hat{H}_A^{(3)} = H_{AA}^{(3)} + H_{AB}^{(2)}LH_{BA}^{(1)} + H_{AB}^{(1)}LH_{BA}^{(2)} + H_{AB}^{(1)}LH_{BB}^{(1)}LH_{BA}^{(1)} - H_{AB}^{(1)}L^2H_{BA}^{(1)}H_{AA}^{(1)}$$

$$\begin{aligned}\hat{H}_A^{(4)} = H_{AA}^{(4)} &+ H_{AB}^{(3)}LH_{BA}^{(1)} + H_{AB}^{(1)}LH_{BA}^{(3)} + H_{AB}^{(2)}LH_{BA}^{(2)} - H_{AB}^{(1)}L^2H_{BA}^{(1)}\hat{H}_A^{(2)} \\ &- (H_{AB}^{(2)}L^2H_{BA}^{(1)} + H_{AB}^{(1)}L^2H_{BA}^{(2)})H_{AA}^{(1)} + H_{AB}^{(2)}LH_{BB}^{(1)}LH_{BA}^{(1)} \\ &+ H_{AB}^{(1)}LH_{BB}^{(2)}LH_{BA}^{(1)} + H_{AB}^{(1)}LH_{BB}^{(1)}LH_{BA}^{(2)} - H_{AB}^{(1)}L^2H_{BA}^{(1)}H_{AB}^{(1)}LH_{BA}^{(1)} \\ &+ H_{AB}^{(1)}LH_{BB}^{(1)}LH_{BB}^{(1)}LH_{BA}^{(1)} - H_{AB}^{(1)}LH_{BB}^{(1)}L^2H_{BA}^{(1)}H_{AA}^{(1)} \\ &- H_{AB}^{(1)}L^2H_{BB}^{(1)}LH_{BA}^{(1)}H_{AA}^{(1)} + H_{AB}^{(1)}L^3H_{BA}^{(1)}H_{AA}^{(1)^2}\end{aligned}$$

TABLE 6.10 $G_A^{(n)}$

(A-states degenerate)

$$G_A^{(0)} = H_{AA}^{(0)}$$

$$G_A^{(1)} = H_{AA}^{(1)}$$

$$G_A^{(2)} = H_{AA}^{(2)} + 2H_{AB}^{(1)}LH_{BA}^{(1)} + H_{AB}^{(1)}LH_{BB}^{(0)}LH_{BA}^{(1)}$$

$$\begin{aligned}G_A^{(3)} = H_{AA}^{(3)} &+ 2H_{AB}^{(2)}LH_{BA}^{(1)} + 2H_{AB}^{(1)}LH_{BA}^{(2)} + 3H_{AB}^{(1)}LH_{BB}^{(1)}LH_{BA}^{(1)} \\ &- H_{AB}^{(1)}L^2H_{BA}^{(1)}H_{AA}^{(1)} - H_{AA}^{(1)}H_{AB}^{(1)}L^2H_{BA}^{(1)} \\ &+ H_{AB}^{(1)}LH_{BB}^{(0)}L(H_{BA}^{(2)} + H_{BB}^{(1)}LH_{BA}^{(1)} - LH_{BA}^{(1)}H_{AA}^{(1)}) \\ &+ (H_{AB}^{(2)} + H_{AB}^{(1)}LH_{BB}^{(1)} - H_{AA}^{(1)}H_{AB}^{(1)}L)LH_{BB}^{(0)}LH_{BA}^{(1)}\end{aligned}$$

TABLE 6.11 $\tilde{H}_A^{(n)}$

(A-states degenerate)

$$\tilde{H}_A^{(0)} = H_{AA}^{(0)}$$

$$\tilde{H}_A^{(1)} = H_{AA}^{(1)}$$

$$\tilde{H}_A^{(2)} = H_{AA}^{(2)} + H_{AB}^{(1)} L H_{BA}^{(1)}$$

$$\tilde{H}_A^{(3)} = H_{AA}^{(3)} + H_{AB}^{(2)} L H_{BA}^{(1)} + H_{AB}^{(1)} L H_{BA}^{(2)} + H_{AB}^{(1)} L H_{BB}^{(1)} L H_{BA}^{(1)} - \tfrac{1}{2}\{H_{AB}^{(1)} L^2 H_{BA}^{(1)}, H_{AA}^{(1)}\}_+$$

For $H = H^{(0)} + H^{(1)}$ only:

$$\tilde{H}_A^{(4)} = H_{AB}^{(1)} L H_{BB}^{(1)} L H_{BB}^{(1)} L H_{BA}^{(1)} - \tfrac{1}{2}\{\tilde{H}_A^{(1)}, (H_{AB}^{(1)} L\, L, H_{BB}^{(1)} + L H_{BA}^{(1)})\}_+ + \tfrac{1}{2}\{\tilde{H}_A^{(1)2}, H_{AB}^{(1)} L^3 H_{BA}^{(1)}\}_+ - \tfrac{1}{2}\{\tilde{H}_A^{(2)}, H_{AB}^{(1)} L\, H_{BA}^{(1)}\}_+$$

$$\tilde{H}_A^{(5)} = H_{AB}^{(1)} L H_{BB}^{(1)} L H_{BB}^{(1)} L H_{BB}^{(1)} L H_{BA}^{(1)} - \tfrac{1}{2}\{\tilde{H}_A^{(1)}, H_{AB}^{(1)} L(L H_{BB}^{(1)} L H_{BB}^{(1)} + H_{BB}^{(1)} L^2 H_{BB}^{(1)} + H_{BB}^{(1)} L H_{BB}^{(1)} L) L H_{BA}^{(1)}\}_+ + \tfrac{1}{2}\{\tilde{H}_A^{(1)2}, H_{AB}^{(1)} L(L^2 H_{BB}^{(1)} + L H_{BB}^{(1)} L + H_{BB}^{(1)} L^2) L H_{BA}^{(1)}\}_+ - \tfrac{1}{2}\{\tilde{H}_A^{(1)3}, H_{AB}^{(1)} L^4 H_{BA}^{(1)}\}_+ - \tfrac{1}{2}\{\tilde{H}_A^{(2)}, H_{AB}^{(1)} L\, L, H_{BB}^{(1)} + L H_{BA}^{(1)}\}_+ + \tfrac{1}{2}\{\{\tilde{H}_A^{(1)}, \tilde{H}_A^{(2)}\}_+, H_{AB}^{(1)} L^3 H_{BA}^{(1)}\}_+ - \tfrac{1}{2}\{\tilde{H}_A^{(3)}, H_{AB}^{(1)} L^2 H_{BA}^{(1)}\}_+ + \tfrac{1}{8}[H_{AB}^{(1)} L\, H_{BA}^{(1)}[H_{AB}^{(1)} L^2 H_{BA}^{(1)}, \tilde{H}_A^{(1)}]_-]_-$$

expressions for the effective operators. Substitution of eqs. (6.18) for $H_{AB}^{(j)}$, (j = 1, ..., n-1), in the formulas for $\hat{H}_A^{(2n)}$ and $\hat{H}_A^{(2n+1)}$, and simplification of the resulting expressions using the formulas for $L^{-1}f^{(n)}$ given in Table 6.6, results in the so-called "reduced" formulas for the $\hat{H}_A^{(n)}$ given in Table 6.8. It is seen that, in the formula for $\hat{H}_A^{(2n)}$ and $\hat{H}_A^{(2n+1)}$, all terms containing $f^{(n+1)}$ through $f^{(2n)}$ are in the form of commutators with a lower order term in the expansion of $\hat{H}_A$. When $n_A = 1$, these commutators vanish, and one obtains a 2n+1 rule in the sense that f through order n is sufficient to determine $\hat{H}_A$ correct through order 2n+1. Similar explicit results have not been obtained for the operators G_A and $\tilde{H}_A$, but the discussion of section 3.2 implies that errors in the eigenvalues of these two operators, when calculated from f correct through order n, should be of order 2n+2. For $n_A \neq 1$, none of these effective operators can be given correct to order n+2 or higher solely in terms of $f^{(1)}$ through $f^{(n)}$.

In this case, $\hat{H}_A^{(2)}$ and $\tilde{H}_A^{(2)}$ are identical. In general, all three effective operators are different in third and higher order here.

6.2.c A-states Non-degenerate

If the eigenvalues of $H^{(0)}$ corresponding to the subspace S_A are not all equal, then the factorization (6.16a) is not possible, and it is necessary to calculate the matrix elements

of the $f^{(n)}$ via eq. (6.8), which can be written formally as

$$f^{(n)} = \mathcal{L}(A^{(n)}). \tag{6.19}$$

Here, $\mathcal{L}$ is a superoperator, which, when acting on the operator $A^{(n)}$ produces the operator $f^{(n)}$. The superoperator $\mathcal{L}$ can be represented as a four index matrix, so that (6.19) becomes

$$f^{(n)}_{\sigma r} = \sum_{\rho,t} \mathcal{L}_{\sigma r,\rho t} A^{(n)}_{\rho t}. \tag{6.20a}$$

If $H^{(0)}$ is diagonal, with eigenvalues ϵ^o_i, then

$$\mathcal{L}_{\sigma r,\rho t} = \frac{1}{\epsilon^o_r - \epsilon^o_\sigma} \delta_{\sigma\rho}\delta_{rt}. \tag{6.20b}$$

In this case, the perturbation formulas are for single matrix elements. Tables 6.12 - 6.15 give some low order formulas of this type. One application of these formulas is in molecular orbital theory. The application to the derivation of Coulson-Longuet Higgins type Hückel theory is outlined in the next chapter.

TABLE 6.12 $f^{(n)}$

$$f_{\sigma r}^{(1)} = \frac{H_{\sigma r}^{(1)}}{\epsilon_r^o - \epsilon_\sigma^o}$$

$$f_{\sigma r}^{(2)} = \left[H_{\sigma r}^{(2)} + \sum_{\mu=1}^{n_B} \frac{H_{\sigma\mu}^{(1)} H_{\mu r}^{(1)}}{\epsilon_r^o - \epsilon_\mu^o} - \sum_{t=1}^{n_A} \frac{H_{\sigma t}^{(1)} H_{tr}^{(1)}}{\epsilon_t^o - \epsilon_\sigma^o} \right] \cdot \frac{1}{\epsilon_r^o - \epsilon_\sigma^o}$$

$$f_{\sigma r}^{(3)} = \left[H_{\sigma r}^{(3)} + \sum_{\mu=1}^{n_B} \frac{H_{\sigma\mu}^{(2)} H_{\mu r}^{(1)}}{\epsilon_r^o - \epsilon_\mu^o} - \sum_{t=1}^{n_A} \frac{H_{\sigma t}^{(1)} H_{tr}^{(2)}}{\epsilon_t^o - \epsilon_\sigma^o} - \sum_{\mu=1}^{n_B} \sum_{t=1}^{n_A} \frac{H_{\sigma t}^{(1)} H_{t\mu}^{(1)} H_{\mu r}^{(1)}}{(\epsilon_t^o - \epsilon_\sigma^o)(\epsilon_r^o - \epsilon_\mu^o)} \right.$$

$$+ \sum_{\mu=1}^{n_B} \frac{H_{\sigma\mu}^{(1)}}{\epsilon_r^o - \epsilon_\mu^o} \left(H_{\mu r}^{(2)} + \sum_{\gamma=1}^{n_B} \frac{H_{\mu\gamma}^{(1)} H_{\gamma r}^{(1)}}{\epsilon_r^o - \epsilon_\gamma^o} - \sum_{t=1}^{n_A} \frac{H_{\mu t}^{(1)} H_{tr}^{(1)}}{\epsilon_t^o - \epsilon_\mu^o} \right)$$

$$\left. - \sum_{t=1}^{n_A} \left(H_{\sigma t}^{(2)} + \sum_{\mu=1}^{n_B} \frac{H_{\sigma\mu}^{(1)} H_{\mu t}^{(1)}}{\epsilon_t^o - \epsilon_\mu^o} - \sum_{s=1}^{n_A} \frac{H_{\sigma s}^{(1)} H_{st}^{(1)}}{\epsilon_s^o - \epsilon_\sigma^o} \right) \frac{H_{tr}^{(1)}}{\epsilon_t^o - \epsilon_\sigma^o} \right] \cdot \frac{1}{\epsilon_r^o - \epsilon_\sigma^o}$$

TABLE 6.13 $\hat{H}_A^{(n)}$

$$\hat{H}_{rs}^{(0)} = H_{rs}^{(0)}$$

$$\hat{H}_{rs}^{(1)} = H_{rs}^{(1)}$$

$$\hat{H}_{rs}^{(2)} = H_{rs}^{(2)} + \sum_{\mu=1}^{n_A} \frac{H_{r\mu}^{(1)} H_{\mu s}^{(1)}}{\epsilon_s^o - \epsilon_\mu^o}$$

$$\hat{H}_{rs}^{(3)} = H_{rs}^{(3)} + \sum_{\mu=1}^{n_B} \frac{H_{r\mu}^{(2)} H_{\mu s}^{(1)} + H_{r\mu}^{(1)} H_{\mu s}^{(2)}}{\epsilon_s^o - \epsilon_\mu^o}$$

$$+ \sum_{\mu=1}^{n_B} \frac{H_{r\mu}^{(1)}}{\epsilon_s^o - \epsilon_\mu^o} \left[\sum_{\gamma=1}^{n_B} \frac{H_{\mu\gamma}^{(1)} H_{\gamma s}^{(1)}}{\epsilon_s^o - \epsilon_\gamma^o} - \sum_{t=1}^{n_A} \frac{H_{\mu t}^{(1)} H_{ts}^{(1)}}{\epsilon_t^o - \epsilon_\mu^o} \right]$$

TABLE 6.14 $G_A^{(n)}$

$$G_{rs}^{(0)} = H_{rs}^{(0)}$$

$$G_{rs}^{(1)} = H_{rs}^{(1)}$$

$$G_{rs}^{(2)} = H_{rs}^{(2)} + \sum_{\mu=1}^{n_B} H_{r\mu}^{(1)} H_{\mu s}^{(1)} \frac{(\epsilon_r^o + \epsilon_s^o - \epsilon_\mu^o)}{(\epsilon_r^o - \epsilon_\mu^o)(\epsilon_s^o - \epsilon_\mu^o)}$$

$$G_{rs}^{(3)} = H_{rs}^{(3)} + \sum_{\sigma=1}^{n_B} \left[\frac{H_{r\sigma}^{(2)} H_{\sigma s}^{(1)}}{\epsilon_s^o - \epsilon_\sigma^o} + \frac{H_{r\sigma}^{(1)} H_{\sigma s}^{(2)}}{\epsilon_r^o - \epsilon_\sigma^o} \right] + \sum_{\sigma=1}^{n_B} \sum_{\mu=1}^{n_B} \frac{H_{r\sigma}^{(1)} H_{\sigma\mu}^{(1)} H_{\mu s}^{(1)}}{(\epsilon_r^o - \epsilon_\sigma^o)(\epsilon_s^o - \epsilon_\mu^o)}$$

$$+ \sum_{\sigma=1}^{n_B} \frac{\epsilon_r^o H_{r\sigma}^{(1)}}{\epsilon_r^o - \epsilon_\sigma^o} \left[H_{\sigma s}^{(2)} + \sum_{\mu=1}^{n_B} \frac{H_{\sigma\mu}^{(1)} H_{\mu s}^{(1)}}{\epsilon_s^o - \epsilon_\mu^o} - \sum_{t=1}^{n_A} \frac{H_{\sigma t}^{(1)} H_{ts}^{(1)}}{\epsilon_t^o - \epsilon_\sigma^o} \right] \frac{1}{\epsilon_s^o - \epsilon_\sigma^o}$$

$$+ \sum_{\sigma=1}^{n_B} \left[H_{r\sigma}^{(2)} + \sum_{\mu=1}^{n_B} \frac{H_{r\mu}^{(1)} H_{\mu\sigma}^{(1)}}{\epsilon_r^o - \epsilon_\mu^o} - \sum_{t=1}^{n_A} \frac{H_{rt}^{(1)} H_{t\sigma}^{(1)}}{\epsilon_t^o - \epsilon_\sigma^o} \right] \frac{1}{\epsilon_r^o - \epsilon_\sigma^o} \cdot \frac{\epsilon_s^o H_{\sigma s}^{(1)}}{\epsilon_s^o - \epsilon_\sigma^o}$$

TABLE 6.15 $\tilde{H}_A^{(n)}$

$$\tilde{H}_{rs}^{(0)} = H_{rs}^{(0)}$$

$$\tilde{H}_{rs}^{(1)} = H_{rs}^{(1)}$$

$$\tilde{H}_{rs}^{(2)} = H_{rs}^{(2)} + \sum_{\sigma=1}^{n_B} \frac{\left[\frac{1}{2}(\epsilon_r^o+\epsilon_s^o)-\epsilon_\sigma^o\right]}{(\epsilon_r^o-\epsilon_\sigma^o)(\epsilon_s^o-\epsilon_\sigma^o)} H_{r\sigma}^{(1)}H_{\sigma s}^{(1)}$$

$$\tilde{H}_{rs}^{(3)} = H_{rs}^{(3)} + \sum_{\sigma=1}^{n_B} \frac{H_{r\sigma}^{(1)}H_{\sigma s}^{(1)}}{\epsilon_s^o - \epsilon_\sigma^o}$$

$$+ \sum_{\sigma=1}^{n_B}\sum_{t=1}^{n_A} \left[\frac{H_{r\sigma}^{(1)}H_{\sigma t}^{(1)}H_{ts}^{(1)}}{(\epsilon_r^o-\epsilon_\sigma^o)(\epsilon_t^o-\epsilon_\sigma^o)} - \frac{H_{rt}^{(1)}H_{t\sigma}^{(1)}H_{\sigma s}^{(1)}}{(\epsilon_t^o-\epsilon_\sigma^o)(\epsilon_s^o-\epsilon_\sigma^o)} \right]$$

$$+ \sum_{\sigma=1}^{n_B} \frac{H_{r\sigma}^{(1)}}{\epsilon_s^o-\epsilon_\sigma^o} \left[1+\frac{1}{2}\frac{\epsilon_s^o-\epsilon_r^o}{\epsilon_r^o-\epsilon_\sigma^o} \right] \left[H_{\sigma s}^{(2)} + \sum_{\mu=1}^{n_B} \frac{H_{\sigma\mu}^{(1)}H_{\mu s}^{(1)}}{\epsilon_s^o - \epsilon^o} - \sum_{t=1}^{n_A} \frac{H_{\sigma t}^{(1)}H_{ts}^{(1)}}{\epsilon_t^o - \epsilon_\sigma^o} \right]$$

$$+ \frac{1}{2}\sum_{\sigma=1}^{n_B} \left[H_{r\sigma}^{(2)} + \sum_{\mu=1}^{n_B} \frac{H_{r\mu}^{(1)}H_{\mu s}^{(1)}}{\epsilon_r^o - \epsilon_\mu^o} - \sum_{t=1}^{n_A} \frac{H_{rt}^{(1)}H_{t\sigma}^{(1)}}{\epsilon_t^o - \epsilon_\sigma^o} \right] \frac{H_{\sigma s}^{(1)}(\epsilon_s^o - \epsilon_r^o)}{(\epsilon_s^o-\epsilon_\sigma^o)(\epsilon_r^o-\epsilon_\sigma^o)}$$

6.3 Examples

6.3.a The Dirac Equation

A particularly simple application of the perturbation formalism just developed is to the formal uncoupling of the Dirac equation for a spin-½ particle in an electric and magnetic field. Not only are the A-states ($E^{(0)} = mc^2$) degenerate in zero order here, but the B-states ($E^{(0)} = -mc^2$) are degenerate as well. Historically, much effort has been expended on the problem of obtaining a two component effective operator describing the behaviour of a spin-½ particle in an electromagnetic field, from the four component Dirac Hamiltonian. In several cases, special algebraic properties of the Dirac hamiltonian were used to construct the desired effective operators, so that it appeared that such operators were unique, in some sense, to the Dirac equation, and not necessarily analogous to effective operators constructed in other contexts. In this section and the accompanying Appendix 8, it is shown that the perturbation formulas tabulated in section 6.2.b yield the desired effective operators immediately.

The Dirac equation is special in that if only a magnetic field is present, the condition $D(f) = 0$ can be solved exactly. Other methods for the exact uncoupling of the Dirac equation in the absence of an electric field have, of course, been known for many years (for example, Foldy and Wouthuysen, 1950).

The Dirac hamiltonian, including electromagnetic interactions, will be written here as

$$H = H^{(0)} + H^{(1)} , \tag{6.21}$$

where,

$$H^{(0)} = \begin{bmatrix} mc^2 & 0 \\ 0 & -mc^2 \end{bmatrix} , \tag{6.22a}$$

and

$$H^{(1)} = \begin{bmatrix} e\phi & c\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi} \\ c\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi} & e\phi \end{bmatrix} , \tag{6.22b}$$

Here ϕ is the electric potential, $(\sigma_x,\sigma_y,\sigma_z)$ are the Pauli spin matrices, $\underset{\sim}{\pi} = \underset{\sim}{p} - \frac{e}{c}\underset{\sim}{A}$ is the mechanical momentum of the system, and m, e, c, are the mass and charge of the electron, and the velocity of light. With the perturbation defined in (6.22b), the implicit perturbation parameter is 1/m. This is strictly not the usual non-relativistic approximation, in which the terms of the series are ordered by powers of v/c, and which will be dealt with later. Both ordering schemes eventually give the same terms in the perturbation series, but the order in which specific terms occur may be different.

The hamiltonian (6.22a,b) is blocked according to the partitioning of the basis space into the subspaces S_A $(E^{(0)}=mc^2)$, and S_B $(E^{(0)}=-mc^2)$. The reduced resolvent (6.17) is just a multiple of the unit matrix because both the A-states and the B-states are degenerate, that is,

$$L = \frac{1}{2mc^2} \cdot 1 . \tag{6.23}$$

Referring now to Table 6.8, it is seen that

$$\hat{H}_A^{(0)} = mc^2 ,$$

$$\hat{H}_A^{(1)} = e\phi ,$$

$$\hat{H}_A^{(2)} = \frac{1}{2m} (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2 , \tag{6.24}$$

and

$$\hat{H}_A^{(3)} = \frac{e}{4m^2c^2} (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})\phi(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi}) - (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2\phi .$$

Standard methods[1] can be used to transform these expressions into the more explicit forms,

$$\hat{H}_A^{(2)} = \frac{-he}{2mc} \underset{\sim}{\sigma}\cdot\underset{\sim}{\mathcal{H}} + \frac{1}{2m} \underset{\sim}{\pi}\cdot\underset{\sim}{\pi},$$

and (6.25)

$$\hat{H}_A^{(3)} = \frac{-he}{4m^2c^2} \underset{\sim}{\sigma}\cdot(\underset{\sim}{E} \times \underset{\sim}{\pi}) - \frac{ihe}{4m^2c^2} \underset{\sim}{E}\cdot\underset{\sim}{\pi} + \frac{eh^2}{4m^2c^2} \underset{\sim}{\nabla}\cdot\underset{\sim}{E}.$$

Here $\underset{\sim}{\mathcal{H}} = \underset{\sim}{\nabla} \times \underset{\sim}{A}$ is the magnetic field, and $\underset{\sim}{E} = -\underset{\sim}{\nabla}\phi$ is the electric field. The various terms in (6.24) are readily identifiable. $\hat{H}_A^{(0)}$ is the rest energy of the system in the absence of any fields, and $\hat{H}_A^{(1)}$ is the electrostatic energy. $\hat{H}_A^{(2)}$ includes both the kinetic energy of the system, and the magnetic dipole interaction. The non-hermiticity of the

---

[1] Using the well known commutation properties of both the Pauli matrices and of differential operators, one obtains,

$$[\phi, \underset{\sim}{\pi}]_- = ih\underset{\sim}{\nabla}\phi = -ih\underset{\sim}{E},$$

$$(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})\phi(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi}) = -h\underset{\sim}{\sigma}\cdot(\underset{\sim}{E} \times \underset{\sim}{\pi}) - \frac{he}{c}\phi\underset{\sim}{\sigma}\cdot\underset{\sim}{\mathcal{H}} + ih\underset{\sim}{E}\cdot\underset{\sim}{\pi} + \phi\underset{\sim}{\pi}\cdot\underset{\sim}{\pi},$$

$$(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2 = 2i\underset{\sim}{\sigma}\cdot(\underset{\sim}{E} \times \underset{\sim}{\pi}) + ih\underset{\sim}{\nabla}\cdot\underset{\sim}{E},$$

$$(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2\phi = -\frac{he}{c}\phi\ \underset{\sim}{\sigma}\cdot\underset{\sim}{\mathcal{H}} + \underset{\sim}{\pi}\cdot\underset{\sim}{\pi}\phi,$$

$$[\phi, \underset{\sim}{\pi}\cdot\underset{\sim}{\pi}]_- = h^2\underset{\sim}{\nabla}^2\phi - 2ih\underset{\sim}{E}\cdot\underset{\sim}{\pi}.$$

operator $\hat{H}_A$ is seen to appear first in third order. The first term in $\hat{H}_A^{(3)}$, eq. (6.25), is the spin-orbit interaction, and the second term is the so-called Darwin term. Equations (6.24) are identical to the results obtained using the Pauli elimination method to uncouple the Dirac hamiltonian. The correction to the kinetic energy due to the relativistic variation of mass arises out of the term $\frac{1}{8m^3c^2}(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^4$, appearing in $\hat{H}_A^{(4)}$.

Similarly, using Table 6.11, one obtains,

$$\tilde{H}_A^{(0)} = mc^2,$$

$$\tilde{H}_A^{(1)} = e\phi$$

$$\tilde{H}_A^{(2)} = \frac{1}{2m}(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2 \qquad (6.26)$$

and

$$\tilde{H}_A^{(3)} = \frac{e}{8m^2c^2}[\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi}, [\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi}, \phi]_-]_-$$

$$= \frac{e}{4m^2c^2}(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})\phi(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi}) - \frac{e}{8m^2c^2}[\phi(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2 + (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2\phi].$$

The second and third order terms can be rewritten

$$\tilde{H}_A^{(2)} = \frac{-he}{2mc}\underset{\sim}{\sigma}\cdot\underset{\sim}{\mathcal{H}} + \frac{1}{2m}\underset{\sim}{\pi}\cdot\underset{\sim}{\pi},$$

and (6.27)

$$\tilde{H}_A^{(3)} = \frac{-he}{4m^2c^2}\underset{\sim}{\sigma}\cdot(\underset{\sim}{E}\times\underset{\sim}{\pi}) - \frac{eh^2}{8m^2c^2}\underset{\sim}{\nabla}\cdot\underset{\sim}{E}.$$

Except for the fourth order relativistic correction to the kinetic energy, which will appear here in $\tilde{H}_A^{(4)}$, this is effectively the result quoted by DeVries (1970), which was obtained from a perturbation series to fourth order in v/c calculated

via the Foldy-Wouthuysen procedure. The point of eqs. (6.24)-(6.27) is that, except for some algebraic manipulations necessary to obtain the effective hamiltonians in a more familiar form, these expressions could be written down without any other calculation, using the tabulated formulas in section 6.2.

The first and second order terms in the expansion of f are particularly simple here also, being given by

$$f^{(1)} = \frac{1}{2mc}\, \underset{\sim}{\sigma}\cdot\underset{\sim}{\pi},$$

and (6.28)

$$f^{(2)} = \frac{e}{4m^2c^3}\,[\phi,\ \underset{\sim}{\sigma}\cdot\underset{\sim}{\pi}]_- = -\frac{ihe}{4m^2c^3}\,\underset{\sim}{\sigma}\cdot\underset{\sim}{E}.$$

That is, each is made up of only one term. These two terms are sufficient to determine $\hat{H}_A$ and $\tilde{H}_A$ to fourth order. Equations (6.28) are also useful for calculating effective operators for other properties of the system.

For an expansion in powers of v/c (the non-relativistic approximation), the Dirac Hamiltonian is usually rewritten as (DeVries, 1970),

$$H = H^{(0)} + H^{(1)} + H^{(2)}, \tag{6.29a}$$

where $H^{(0)}$ is as in (6.22), but now

$$H^{(1)} = \begin{bmatrix} 0 & c\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi} \\ c\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi} & 0 \end{bmatrix}, \tag{6.29b}$$

and

$$H^{(2)} = \begin{bmatrix} e\phi & 0 \\ 0 & e\phi \end{bmatrix}, \tag{6.29c}$$

Actual expressions for $\hat{H}_A$, G_A, and $\tilde{H}_A$ to sixth order, based on eqs. (6.29), are given in Appendix 8 in a somewhat more abstract notation. Because of the particular form of the first and second order perturbations here (the first order perturbation couples two states only if they are in different subspaces S_A and S_B, whereas, the second order perturbation couples states in the same subspace only), the perturbation series for these effective operators have only even order terms nonzero, while the series for f has only odd order terms nonzero. To sixth order, the operator $\hat{H}_A$ is exactly equal to the result obtained to sixth order in v/c using the Pauli elimination method to obtain a non-relativistic approximation. Similarly, to sixth order, $\tilde{H}_A$ is identical to the results of a canonical uncoupling of the Dirac hamiltonian, such as that carried out by Eriksen, (1958). DeVries (1970) demonstrates that the Pauli hamiltonian is related to the Eriksen hamiltonian to sixth order by a fourth order similarity transformation defined in the space of positive energy states only. Such a relationship is evident from the definition of $\tilde{H}_A$, eq. (2.74a), in terms of $\hat{H}_A$, namely that

$$\tilde{H}_A = g_A^{\frac{1}{2}} \hat{H}_A \, g_A^{-\frac{1}{2}}. \tag{6.30}$$

Thus, the required similarity transformation matrix is just $g_A^{-\frac{1}{2}}$. Because $H_{AA}^{(0)}$ here is a multiple of the unit matrix, the terms in $g_A^{(6)}$ in eq. (6.30) exactly cancel if $\tilde{H}_A$ is desired to sixth order only. Since $g_A^{(5)} = 0$, as seen from the tabulations in Appendix 8, the similarity transformation $g_A^{-\frac{1}{2}}$ need be known only

to fourth order to determine $\tilde{H}_A$ to sixth order.

A treatment of the Dirac equation with some similarity to the above application of the partitioning formalism has been given by Morpurgo (1960). In the course of a rather complicated derivation of a unitary transformation to bring the Dirac hamiltonian to an uncoupled form, it becomes convenient for Morpurgo to define an operator of the form

$$G = U_{BB}\, U_{AB}^{-1}, \tag{6.31}$$

where the quantities U_{BB} and U_{AB} are blocks of the unitary transformation matrix when partitioned in the same manner as H in eq. (6.22). It is not difficult to show from the defining condition given by Morpurgo that, for the Dirac hamiltonian only,

$$G = (-f^{\dagger})^{-1}. \tag{6.32}$$

It is not known if the quantity G has useful generalizations in other contexts, as does the operator f. Certainly, the relationship (6.32) can possibly hold only in cases where S_A and S_B have identical dimensions (so that U_{AB} has an inverse).

If an electric field is not present, these effective operators can be calculated exactly, since the perturbation is nonzero only in off-diagonal blocks. This was, in fact, the basis of Foldy and Wouthuysen's free particle calculation (1950). The equation D(f) = 0, defining the operator f, becomes,

$$\underset{\sim}{\sigma}\cdot\underset{\sim}{\Pi} - 2mcf - f\underset{\sim}{\sigma}\cdot\underset{\sim}{\Pi}f = 0. \tag{6.33}$$

Multiplication by $\underset{\sim}{\sigma}\cdot\underset{\sim}{\Pi}$ from the right yields a quadratic equation for $(\underset{\sim}{\sigma}\cdot\underset{\sim}{\Pi})f$. The desired solution is

$$f = \frac{\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi}}{mc + [m^2c^2 + (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2]^{\frac{1}{2}}} , \qquad (6.34)$$

since the root with the plus sign in the denominator leads to the expansion (6.28). Given this exact expression for f, all other quantities defined in chapter 2 can be written, exactly, in terms of mc and $\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi}$. The operator $\hat{H}_A$ is particularly simple,

$$\hat{H}_A = H_{AA} + H_{AB}f$$

$$= mc^2 + \frac{c(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2}{mc + [m^2c^2 + (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2]^{\frac{1}{2}}} . \qquad (6.35)$$

The operators G_A and $\tilde{H}_A$ are obtained in the same way, but the expressions are much more complicated. It is also possible to write down an exact expression for the projection onto the space S_A', spanned by the eigenvectors of the perturbed hamiltonian which have zero order energy $E^{(0)} = mc^2$. Since

$$g_A = 1_A + f^\dagger f = 1_A + f^2$$

$$= \frac{[mc + \sqrt{(mc)^2 + (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2}]^2 + (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2}{mc + [(mc)^2 + (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2]^{\frac{1}{2}}} ,$$

by eq. (2.9a), one obtains, from eq. (2.10),

$$P_A' = \{[mc + [(mc)^2 + (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2]^{\frac{1}{2}}]^2 + (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2\}^{-1}$$

$$\times \begin{bmatrix} mc+[(mc)^2+(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2]^{\frac{1}{2}} & [mc+[(mc)^2+(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2]^{\frac{1}{2}}](\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi}) \\ (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})[mc+[(mc)^2+(\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2]^{\frac{1}{2}}] & (\underset{\sim}{\sigma}\cdot\underset{\sim}{\pi})^2 \end{bmatrix}$$

(6.36)

6.3.b Derivation of a Spin Hamiltonian -- Strong Field Case

Consider the hamiltonian operator

$$H = \underline{h}\cdot\underline{S} + \underline{\mathcal{I}}\cdot\underline{S} + \underline{S}\cdot\underline{\underline{D}}\cdot\underline{S}, \tag{6.37}$$

where $\underline{h}$ is the effective magnetic field, $-\beta\underline{\underline{g}}\cdot\underline{H}$, and where

$$\underline{\mathcal{I}} = \sum_j \underline{I}^{(j)}\cdot\underline{\underline{A}}^{(j)}, \tag{6.38}$$

the $\underline{\underline{A}}^{(j)}$ being hyperfine tensors. This hamiltonian describes the interaction of a system of nuclear spins with an electronic spin. In the strong field case, the term $\underline{h}\cdot\underline{S}$ (the electronic Zeeman interaction) is large enough that the energy separation between levels of different electronic spin is greater than the energy separations between nuclear spin levels. Therefore, a perturbation expansion with respect to $H^{(0)} = \underline{h}\cdot\underline{S}$ is appropriate in examining the characteristics of the nuclear spin system. In this subsection, a nuclear spin hamiltonian is constructed from (6.37), in which the electron spin quantum numbers are present only as parameters.

In the strong field case, the electronic spin is quantized in the field direction, taken to be the z-axis in the notation adopted here. Thus, the zero order hamiltonian is

$$H^{(0)} = hS_z, \tag{6.39}$$

where $h = |\underline{h}|$. It is convenient to expand the perturbation $H^{(1)} = \underline{\mathcal{I}}\cdot\underline{S}+\underline{S}\cdot\underline{\underline{D}}\cdot\underline{S}$ in the form,

$$\begin{aligned} H^{(1)} = {} & \mathcal{I}_z S_z + \tfrac{1}{2}\mathcal{I}_- S_+ + \tfrac{1}{2}\mathcal{I}_+ S_- \\ & + D_{+2}S_-^2 + D_{+1}(S_-S_z+S_zS_-) + D_0(S_z^2-\tfrac{1}{3}S^2) + D_{-1}(S_+S_z+S_zS_+) + D_{-2}S_+^2 \end{aligned} \tag{6.40}$$

Here $S_{\pm} = S_x \pm S_y$ are the usual shift operators for the electronic spin, and $\mathcal{I}_{\pm} = \mathcal{I}_x \pm \mathcal{I}_y$ are of the same form in the components of $\mathcal{I}$. The coefficients D_q are

$$D_o = \frac{3}{2} D_{zz},$$

$$D_{\pm 1} = \tfrac{1}{2}(D_{xz} \pm iD_{yz}), \tag{6.41}$$

and

$$D_{\pm 2} = \tfrac{1}{2}[\tfrac{1}{2}(D_{xx} - D_{yy}) \pm iD_{xy}].$$

The zero order levels having $S_z = m$ define the space S_A for the effective nuclear spin hamiltonian $\mathcal{H}_m$. Because the perturbation $H^{(1)}$ has matrix elements with $\Delta m = 0, \pm 1, \pm 2$ only, there are only a small number of nonzero matrix elements in $f^{(1)}$, which can be written down directly using Table 6.6 and eq. (6.40). These are

$$f^{(1)}_{m+1,m} = -\frac{1}{h}[D_{-1}(2m+1) + \tfrac{1}{2}\mathcal{I}_-][\underset{\sim}{S}^2 - m(m+1)]^{\frac{1}{2}},$$

$$f^{(1)}_{m-1,m} = \frac{1}{h}[D_{+1}(2m-1) + \tfrac{1}{2}\mathcal{I}_+][\underset{\sim}{S}^2 - m(m-1)]^{\frac{1}{2}},$$

$$f^{(1)}_{m+2,m} = \frac{-1}{2h} D_{-2}[\underset{\sim}{S}^2 - m(m+1)]^{\frac{1}{2}}[\underset{\sim}{S}^2 - (m+1)(m+2)]^{\frac{1}{2}}, \tag{6.42}$$

and

$$f^{(1)}_{m-2,m} = \frac{1}{2h} D_{+2}[\underset{\sim}{S}^2 - m(m-1)]^{\frac{1}{2}}[\underset{\sim}{S}^2 - (m-1)(m-2)]^{\frac{1}{2}}.$$

The A-states are degenerate, and so $\hat{H}_A$ is identical to $\tilde{H}_A$ to second order. Therefore, Tables 6.8, 6.9, or 6.11 yield (to second order),

$$\mathcal{H}^{(0)}_m = H^{(0)}_{AA} = hm,$$

$$\mathcal{H}_m^{(1)} = H_{AA}^{(1)} = D_o(m^2 - \tfrac{1}{3}\underset{\sim}{S}^2) + m\,\mathcal{J}_z,$$

and

$$\mathcal{H}_m^{(2)} = \frac{2m}{h}D_{+2}D_{-2}[2\underset{\sim}{S}^2-2m^2-1]-\frac{2m}{h}D_{+1}D_{-1}[4\underset{\sim}{S}^2-8m-1]$$
$$-\frac{[\underset{\sim}{S}^2-3m^2]}{h}[D_{+1}\,\mathcal{J}_- + D_{-1}\,\mathcal{J}_+] \qquad (6.43)$$
$$+\frac{1}{4h}\left\{[\mathcal{J}_-,\mathcal{J}_+]_-(-\underset{\sim}{S}^2+m^2) + m\,\{\mathcal{J}_+,\mathcal{J}_-\}_+\right\}.$$

This derivation appears to assume a special, and inconvenient, coordinate system. However, all reference to special coordinate directions (x and y) perpendicular to the field direction disappears on developing the terms in (6.43). If $\hat{\underset{\sim}{h}}$ specifies a unit vector in the direction of $\underset{\sim}{h}$, then the effective hamiltonian to second order can be written

$$\mathcal{H}_m = C_m + \hat{\underset{\sim}{h}}\cdot\sum_j \tilde{\underset{\approx}{A}}^{(j)}\cdot\underset{\sim}{I}^{(j)} + \sum_{i,j}\underset{\sim}{I}^{(i)}\cdot\tilde{\underset{\approx}{Q}}^{(ij)}\cdot\underset{\sim}{I}^{(j)}, \qquad (6.44)$$

Here C_m is

$$C_m = hm + D_o(m^2-\tfrac{1}{3}\underset{\sim}{S}^2) + \frac{m}{h}\left[\frac{(\hat{\underset{\sim}{h}}\cdot\underset{\approx}{D}\cdot\hat{\underset{\sim}{h}})^2}{8} - \frac{\hat{\underset{\sim}{h}}\cdot\underset{\approx}{D}\cdot\hat{\underset{\sim}{h}}}{2}\right][2\underset{\sim}{S}^2-2m^2-1]$$
$$-\frac{2m}{h}[\hat{\underset{\sim}{h}}\cdot\underset{\approx}{D}^2\cdot\hat{\underset{\sim}{h}} - (\hat{\underset{\sim}{h}}\cdot\underset{\approx}{\hat{D}}\cdot\hat{\underset{\sim}{h}})^2][4\underset{\sim}{S}^2-8m-1], \qquad (6.45a)$$

the latter two terms giving the overall second order shift of the endor levels. The other effective parameters are

$$\tilde{\underset{\approx}{A}}^{(j)} = m\underset{\approx}{A}^{(j)}\cdot\hat{\underset{\sim}{h}} + \frac{2}{h}(3m^2-\underset{\sim}{S}^2)\underset{\approx}{D}\cdot(1-\hat{\underset{\sim}{h}}\hat{\underset{\sim}{h}})\cdot\underset{\approx}{A}^{(j)} + \frac{1}{2h}(\underset{\sim}{S}^2-m^2)\underset{\approx}{A}^{(j)}, \qquad (6.45b)$$

and

$$\tilde{\underset{\approx}{Q}}^{(ij)} = \underset{\approx}{A}^{(j)}\cdot(1 - \hat{\underset{\sim}{h}}\hat{\underset{\sim}{h}})\cdot\underset{\approx}{A}^{(j)}. \qquad (6.45c)$$

Also, $\hat{\underset{\approx}{D}}$ and $\hat{\underset{\approx}{A}}$ are the cofactor matrices,

$$\hat{\underset{\approx}{D}} = \underset{\approx}{D}^{-1}\det(\underset{\approx}{D}), \qquad \hat{\underset{\approx}{A}} = \underset{\approx}{A}^{-1}\det(\underset{\approx}{A}). \tag{6.46}$$

The third order $\mathcal{H}_m^{(3)}$ has been obtained in a similar way.

6.4 Non-orthonormal Basis -- 2 x 2 Partitioning

Many quantum mechanical calculations are carried out using a non-orthonormal basis. In such situations, it may be inconvenient or undesirable, (or even impossible for certain kinds of perturbation) to transform to an orthonormal basis in order to carry out a perturbation calculation. This section outlines perturbation expansions based on the formalism in section 2.3, applicable in a non-orthonormal basis.

It is assumed that the given perturbed hamiltonian and overlap matrices have the expansions

$$H = \sum_{n=0}^{\infty} H^{(n)}, \qquad S = \sum_{n=0}^{\infty} S^{(n)}, \tag{6.47}$$

where $H^{(n)}$ and $S^{(n)}$ are implicitly of order n in the perturbation parameter or parameters.

As observed in section 2.3, there are two alternative types of conditions defining the off-diagonal blocks of the partitioning operator $\hat{T}$, in this case, the first being eqs. (2.113) and (2.114). If perturbation expansions are desired only for effective operators in the A-space, then only eq. (2.113) need be considered,

$$D(f) = H_{BA} + H_{BB}f - (S_{BA} + S_{BB}f)\hat{S}_A^{-1}\hat{H}_A, \tag{6.48a}$$

$$= H_{BA} + H_{BB}f - (S_{BA} + S_{BB}f)(S_{AA} + S_{AB}f)^{-1}(H_{AA} + H_{AB}f) = 0. \tag{6.48b}$$

Substitution of the series, (6.47), for H and S, and expansion of f in a similar series then leads, as before, to a hierarchy

of equations,

$$D^{(n)}(f) = 0, \qquad n = 0,1,2,\ldots . \tag{6.49}$$

It will be assumed in the remainder of this section that $H^{(0)}$ and $S^{(0)}$ are at least block diagonal, so that $f^{(0)}=0$. Formulas for the $D^{(n)}(f)$ in terms of H, S, and f are then obtained in stages. Writing

$$\hat{S}_A = S_{AA}^{(0)}+S_{AA}^{(1)}+\sum_{n=2}^{\infty}(S_{AA}^{(n)}+\sum_{j=1}^{n-1}S_{AB}^{(n-j)}f^{(j)})$$

$$=S_{AA}^{(0)}[1_A+S_{AA}^{(0)-1}S_{AA}^{(1)}+S_{AA}^{(0)-1}\sum_{n=2}^{\infty}(S_{AA}^{(n)}+\sum_{j=1}^{n-1}S_{AB}^{(n-j)}f^{(j)})], \tag{6.51}$$

one obtains

$$\hat{S}_A^{-1}=S_{AA}^{(0)-1}+\sum_{k=1}^{\infty}(-1)^k[S_{AA}^{(0)-1}S_{AA}^{(1)}+ \tag{6.52}$$

$$+S_{AA}^{(0)-1}\sum_{n=2}^{\infty}(S_{AA}^{(n)}+\sum_{j=1}^{n-1}S_{AB}^{(n-j)}f^{(j)})]^k S_{AA}^{(0)-1},$$

from which low order terms in the perturbation series for $\hat{S}_A^{-1}$ can be obtained. Then for the operator $\hat{H}_A = \hat{S}_A^{-1}\hat{H}_A'$, one has

$$\hat{H}_A = \sum_{n=0}^{\infty}\hat{H}_A^{(n)}, \tag{6.53a}$$

where

$$\hat{H}_A^{(n)} = \sum_{j=0}^{n}\hat{S}_A^{-1(j)}\hat{H}_A'^{(n-j)}$$

$$= \sum_{j=0}^{n}\hat{S}_A^{-1(j)}(H_{AA}^{(n-j)}+\sum_{k=1}^{n-j-1}H_{AB}^{(n-j-k)}f^{(k)}). \tag{6.53b}$$

Given (6.52) and (6.53), eq. (6.48) can be expanded straightforwardly, and the hierarchy (6.49) can finally be written

$$H^{(0)}_{BB} f^{(n)} - S^{(0)}_{BB} f^{(n)} S^{(0)-1}_{AA} H^{(0)}_{AA} = -A^{(n)}, \quad n=1,2,\ldots, \tag{6.54}$$

where

$$A^{(n)} = H^{(n)}_{BA} + \sum_{j=1}^{n-1} H^{(n-j)}_{BB} f^{(j)} - \sum_{j=1}^{n-1} (S^{(n-j)}_{BA} + \sum_{k=1}^{n-j} S^{(n-j-k)}_{BB} f^{(k)}) \hat{H}^{(j)}_{A} - (S^{(n)}_{BA} + \sum_{k=1}^{n-1} S^{(n-k)}_{BB} f^{(k)}) \hat{H}^{(0)}_{A}. \tag{6.55}$$

The equations (6.54) for the $f^{(n)}$ are more complicated than those for an orthonormal basis because of the blocks $S^{(0)}_{BB}$ and $S^{(0)-1}_{AA}$ appearing on the left hand side. This linear system can be solved for $f^{(n)}$ using numerical methods, but a general solution in terms of S and H cannot be given.

The expansions (6.51) - (6.55) become much simpler when H and S are diagonal in zero order. Then $S^{(0)} = 1_n$ (so that $S^{(0)}_{BB} = 1_B$, $S^{(0)-1}_{AA} = 1_A$), and the equations defining the $f^{(n)}$ become identical in form to eqs. (6.8),

$$f^{(n)}_{\sigma r} = \frac{A^{(n)}_{\sigma r}}{H^{(0)}_{rr} - H^{(0)}_{\sigma\sigma}}, \tag{6.56}$$

where the $A^{(n)}$ are now given by (6.55).

Low order terms in the expansions (6.48) and (6.53) are given in Tables 6.16 and 6.17 for the case $S^{(0)} = 1_n$. In terms of f and H, the effective operator G_A is independent of S, and therefore the expansion (6.12) still holds (Table 6.3). Low order terms for the metric g_A have been listed in Table A9.1 of Appendix 9. The formal definition of $\tilde{H}_A$, eq. (2.74), in terms of $g_A^{\pm\frac{1}{2}}$ is also unaffected. However, the series for g_A

TABLE 6.16 $D^{(n)}(f)$ -- Non-orthonormal Basis

$$D^{(1)} = H_{BB}^{(0)}f^{(1)} - f^{(1)}H_{AA}^{(0)} + H_{BA}^{(1)} - S_{BA}^{(1)}H_{AA}^{(0)}$$

$$D^{(2)} = H_{BB}^{(0)}f^{(2)} - f^{(2)}H_{AA}^{(0)}$$
$$+H_{BA}^{(2)} + H_{BB}^{(1)}f^{(1)} - (S_{BA}^{(1)} + f^{(1)})(H_{AA}^{(1)} - S_{AA}^{(1)}H_{AA}^{(0)}) - (S_{BA}^{(2)} + S_{BB}^{(1)}f^{(1)})H_{AA}^{(0)}$$

$$D^{(3)} = H_{BB}^{(0)}f^{(3)} - f^{(3)}H_{AA}^{(0)}$$
$$+H_{BA}^{(3)} + H_{BB}^{(2)}f^{(1)} + H_{BB}^{(1)}f^{(2)}$$
$$-(S_{BA}^{(1)} + f^{(1)})[H_{AA}^{(2)} + H_{AB}^{(1)}f^{(1)} - S_{AA}^{(1)}H_{AA}^{(1)} - (S_{AA}^{(2)} + S_{AB}^{(1)}f^{(1)})H_{AA}^{(0)}$$
$$+S_{AA}^{(1)2}H_{AA}^{(0)}]$$
$$-(S_{BA}^{(2)} + S_{BB}^{(1)}f^{(1)} + f^{(2)})(H_{AA}^{(1)} - H_{AA}^{(1)}H_{AA}^{(0)})$$
$$-(S_{BA}^{(3)} + S_{BB}^{(2)}f^{(1)} + S_{BB}^{(1)}f^{(2)})H_{AA}^{(0)}$$

$$D^{(4)} = H_{BB}^{(0)}f^{(4)} - f^{(4)}H_{AA}^{(0)}$$
$$+H_{BA}^{(4)} + H_{BB}^{(3)}f^{(1)} + H_{BB}^{(2)}f^{(2)} + H_{BB}^{(1)}f^{(3)}$$
$$-(S_{BA}^{(1)} + f^{(1)})\hat{H}_{A}^{(3)} - (S_{BA}^{(2)} + S_{BB}^{(1)}f^{(1)} + f^{(2)})\hat{H}_{A}^{(2)}$$
$$-(S_{BA}^{(3)} + S_{BB}^{(2)}f^{(1)} + S_{BB}^{(1)}f^{(2)} + f^{(3)})(H_{AA}^{(1)} - S_{AA}^{(1)}H_{AA}^{(0)})$$
$$-(S_{BA}^{(4)} + S_{BB}^{(3)}f^{(1)} + S_{BB}^{(2)}f^{(2)} + S_{BB}^{(1)}f^{(3)})H_{AA}^{(0)}$$

TABLE 6.17 $\hat{H}_A^{(n)}(f)$ -- Non-orthonormal Basis

$$\hat{H}_A^{(0)} = H_{AA}^{(0)}$$

$$\hat{H}_A^{(1)} = H_{AA}^{(1)} - S_{AA}^{(1)} H_{AA}^{(0)}$$

$$\hat{H}_A^{(2)} = H_{AA}^{(2)} + H_{AB}^{(1)} f^{(1)} - S_{AA}^{(1)} H_{AA}^{(1)} - (S_{AA}^{(2)} + S_{AB}^{(1)} f^{(1)}) H_{AA}^{(0)} + S_{AA}^{(1)2} H_{AA}^{(0)}$$

$$\begin{aligned}\hat{H}_A^{(3)} = {} & H_{AA}^{(3)} + H_{AB}^{(2)} f^{(1)} + H_{AB}^{(1)} f^{(2)} - S_{AA}^{(1)} (H_{AA}^{(2)} + H_{AB}^{(1)} f^{(1)}) \\ & - (S_{AA}^{(2)} + S_{AB}^{(1)} f^{(1)}) H_{AA}^{(1)} + S_{AA}^{(1)2} H_{AA}^{(1)} \\ & - (S_{AA}^{(3)} + S_{AB}^{(2)} f^{(1)} + S_{AB}^{(1)} f^{(2)}) H_{AA}^{(0)} + \{S_{AA}^{(1)}, S_{AA}^{(2)} + S_{AB}^{(1)} f^{(1)}\}_+ H_{AA}^{(0)} \\ & - S_{AA}^{(1)3} H_{AA}^{(0)}\end{aligned}$$

TABLE 6.18 $\tilde{H}_A^{(n)}$ -- Non-orthonormal Basis

$$\tilde{H}_A^{(0)} = H_{AA}^{(0)}$$

$$\tilde{H}_A^{(1)} = H_{AA}^{(1)} - \tfrac{1}{2}\{S_{AA}^{(1)}, H_{AA}^{(0)}\}_+$$

$$\begin{aligned}\tilde{H}_A^{(2)} = {} & H_{AA}^{(2)} + H_{AB}^{(1)} f^{(1)} + \tfrac{1}{2}[f^{(1)\dagger} (S_{BA}^{(1)} + f^{(1)}), H_{AA}^{(0)}]_- \\ & - \tfrac{1}{2}\{S_{AA}^{(2)} + S_{AB}^{(1)} f^{(1)}, H_{AA}^{(0)}\}_+ + \tfrac{3}{8}\{S_{AA}^{(1)2}, H_{AA}^{(0)}\}_+ \\ & - \tfrac{1}{2}\{S_{AA}^{(1)}, H_{AA}^{(1)}\}_+ + \tfrac{1}{4} S_{AA}^{(1)} H_{AA}^{(1)} S_{AA}^{(1)}\end{aligned}$$

TABLE 6.19 $G_{BA}^{(n)}$ -- Non-orthonormal Basis

$$G_{BA}^{(1)} = H_{BB}^{(0)} f^{(1)} + h^{(1)\dagger} H_{AA}^{(0)} + H_{BA}^{(1)}$$

$$G_{BA}^{(2)} = H_{BB}^{(0)} f^{(2)} + h^{(2)\dagger} H_{AA}^{(0)} + H_{BA}^{(2)} + H_{BB}^{(1)} f^{(1)} + h^{(1)\dagger} H_{AA}^{(1)}$$

$$G_{BA}^{(3)} = H_{BB}^{(0)} f^{(3)} + h^{(3)\dagger} H_{AA}^{(0)} + H_{BA}^{(3)} + H_{BB}^{(2)} f^{(1)} + H_{BB}^{(1)} f^{(2)} + h^{(1)\dagger} (H_{AA}^{(2)} + H_{AB}^{(1)} f^{(1)}) + h^{(2)\dagger} H_{AA}^{(1)}$$

$$G_{BA}^{(4)} = H_{BB}^{(0)} f^{(4)} + h^{(4)\dagger} H_{AA}^{(0)} + H_{BA}^{(4)} + H_{BB}^{(3)} f^{(1)} + H_{BB}^{(2)} f^{(2)} + H_{BB}^{(1)} f^{(3)} + h^{(1)\dagger} (H_{AA}^{(3)} + H_{AB}^{(2)} f^{(1)} + H_{AB}^{(1)} f^{(2)}) + h^{(2)\dagger} (H_{AA}^{(2)} + H_{AB}^{(1)} f^{(1)}) + h^{(3)\dagger} H_{AA}^{(1)}$$

TABLE 6.20 $g_{BA}^{(n)}$ -- Non-orthonormal Basis

$$g_{BA}^{(1)} = f^{(1)} + h^{(1)\dagger} + S_{BA}^{(1)}$$

$$g_{BA}^{(2)} = f^{(2)} + h^{(2)\dagger} + S_{BA}^{(2)} + S_{BB}^{(1)} f^{(1)} + h^{(1)\dagger} S_{AA}^{(1)}$$

$$g_{BA}^{(3)} = f^{(3)} + h^{(3)\dagger} + S_{BA}^{(3)} + S_{BB}^{(2)} f^{(1)} + S_{BB}^{(1)} f^{(2)} + h^{(1)\dagger} (S_{AA}^{(2)} + S_{AB}^{(1)} f^{(1)}) + h^{(2)\dagger} S_{AA}^{(1)}$$

$$g_{BA}^{(4)} = f^{(4)} + h^{(4)\dagger} + S_{BA}^{(4)} + S_{BB}^{(3)} f^{(1)} + S_{BB}^{(2)} f^{(2)} + S_{BB}^{(1)} f^{(3)} + h^{(1)\dagger} (S_{AA}^{(3)} + S_{AB}^{(2)} f^{(1)} + S_{AB}^{(1)} f^{(2)}) + h^{(2)\dagger} (S_{AA}^{(2)} + S_{AB}^{(1)} f^{(1)}) + h^{(3)\dagger} S_{AA}^{(1)}$$

TABLE 6.21 $f^{(n)}$ -- Non-orthonormal Basis

(A-states degenerate)

$$\tilde{\tilde{H}} = H - \epsilon_A^o S$$

$$f^{(1)} = L(H_{BA}^{(1)} - \epsilon_A^o S_{BA}^{(1)})$$

$$\begin{aligned} f^{(2)} = {} & LH_{BA}^{(2)} + LH_{BB}^{(1)} L(H_{BA}^{(1)} - \epsilon_A^o S_{BA}^{(1)}) \\ & - L[S_{BA}^{(1)} + L(H_{BA}^{(1)} - \epsilon_A^o S_{BA}^{(1)})](H_{AA}^{(1)} - \epsilon_A^o S_{AA}^{(1)}) \\ & + \epsilon_A^o L[S_{BA}^{(2)} + S_{BB}^{(1)} L(H_{BA}^{(1)} - \epsilon_A^o S_{BA}^{(1)})] \end{aligned}$$

$$\begin{aligned} f^{(3)} = {} & LH_{BA}^{(3)} + LH_{BB}^{(2)} L\tilde{\tilde{H}}_{BA}^{(1)} + LH_{BB}^{(1)} [LH_{BA}^{(2)} + LH_{BB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)} - L(S_{BA}^{(1)} + L\tilde{\tilde{H}}_{BA}^{(1)})\tilde{\tilde{H}}_{AA}^{(1)} \\ & \qquad + \epsilon_A^o L(S_{BA}^{(2)} + S_{BB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)})] \\ & - L(S_{BA}^{(1)} + \tilde{\tilde{H}}_{BA}^{(1)})[H_{AA}^{(2)} + H_{AB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)} - S_{AA}^{(1)} H_{AA}^{(1)} \\ & \qquad - \epsilon_A^o (S_{AA}^{(2)} + S_{AB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)}) + \epsilon_A^o S_{AA}^{(1)2}] \\ & - L[S_{BA}^{(2)} + S_{BB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)} + LH_{BA}^{(2)} + LH_{BB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)} - L(S_{BA}^{(1)} + L\tilde{\tilde{H}}_{BA}^{(1)})\tilde{\tilde{H}}_{AA}^{(1)} \\ & \qquad + \epsilon_A^o L(S_{BA}^{(2)} + S_{BB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)})]\tilde{\tilde{H}}_{AA}^{(1)} \\ & - \epsilon_A^o [S_{BA}^{(3)} + S_{BB}^{(2)} L\tilde{\tilde{H}}_{BA}^{(1)} + S_{BB}^{(1)} L\{H_{BA}^{(2)} + H_{BB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)} - (S_{BA}^{(1)} + L\tilde{\tilde{H}}_{BA}^{(1)})\tilde{\tilde{H}}_{AA}^{(1)} \\ & \qquad + \epsilon_A^o (S_{BA}^{(2)} + S_{BB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)})\}] \end{aligned}$$

---

now depends on S, and has a first order term, so that, while eqs. (6.15) hold here also, the formulas in Tables 6.5, A7.7, and A7.8 are no longer valid. Explicit expressions for low order terms in the series for $\tilde{\tilde{H}}_A$ in terms of H, S, and f are given in Table 6.18.

As in an orthonormal basis, the perturbation series for the effective operators $\hat{H}_A$, G_A, and $\tilde{\tilde{H}}_A$, are much more compact when the A-states are degenerate in zero order. Equations

TABLE 6.22 $\hat{H}_A^{(n)}$ -- Non-orthonormal Basis
(A-states degenerate)

$$\hat{H}_A^{(0)} = H_{AA}^{(0)} = \epsilon_A^o 1_A$$

$$\hat{H}_A^{(1)} = H_{AA}^{(1)} - \epsilon_A^o S_{AA}^{(1)} = \tilde{\tilde{H}}_{AA}^{(1)}$$

$$\hat{H}_A^{(2)} = H_{AA}^{(2)} + H_{AB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)} - S_{AA}^{(1)} H_{AA}^{(1)} - \epsilon_A^o (S_{AA}^{(2)} + S_{AB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)}) + \epsilon_A^o S_{AA}^{(1)2}$$

$$\begin{aligned}\hat{H}_A^{(3)} = {} & H_{AA}^{(3)} + H_{AB}^{(2)} L\tilde{\tilde{H}}_{BA}^{(1)} \\ & + H_{AB}^{(1)} L[H_{BA}^{(2)} + H_{BB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)} - (S_{BA}^{(1)} + L\tilde{\tilde{H}}_{BA}^{(1)})\tilde{\tilde{H}}_{AA}^{(1)} + \epsilon_A^o (S_{BA}^{(2)} + S_{BB}^{(1)} \tilde{\tilde{H}}_{BA}^{(1)})] \\ & - S_{AA}^{(1)} (H_{AA}^{(2)} + H_{AB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)}) - (S_{AA}^{(2)} + S_{AB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)}) H_{AA}^{(1)} \\ & - \epsilon_A^o \{S_{AA}^{(3)} + S_{AB}^{(2)} L\tilde{\tilde{H}}_{BA}^{(1)} + S_{AB}^{(1)} L[H_{BA}^{(2)} + H_{BB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)} - (S_{BA}^{(1)} + L\tilde{\tilde{H}}_{BA}^{(1)})\tilde{\tilde{H}}_{AA}^{(1)} \\ & \qquad + \epsilon_A^o (S_{BA}^{(2)} + S_{BB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)})]\} \\ & + S_{AA}^{(1)2} H_{AA}^{(1)} - \epsilon_A^o S_{AA}^{(1)3} + \epsilon_A^o \{S_{AA}^{(1)}, S_{AA}^{(2)} + S_{AB}^{(1)} L\tilde{\tilde{H}}_{BA}^{(1)}\}_+\end{aligned}$$

(6.16) and (6.17) apply here, with the modified $A^{(n)}$ of eq. (6.55), and can be used to obtain explicit matrix formulas for f, $\hat{H}_A$, G_A, and $\tilde{H}_A$, solely in terms of H and S. The lower order terms for these expansions are given in Tables 6.21 - 6.24.

If the A-states are not degenerate in zero order, use of eqs. (6.56) yields formulas for the individual matrix elements of the operators f, $\hat{H}_A$, G_A, and $\tilde{H}_A$. Low order formulas of this type are given in Tables 6.25 - 6.28.

The second set of conditions defining f and h arise from the requirements

TABLE 6.23 $G_A^{(n)}$ -- Non-orthonormal Basis

(A-states degenerate)

$$G_A^{(0)} = H_{AA}^{(0)} = \epsilon_A^o 1_A$$

$$G_A^{(1)} = H_{AA}^{(1)}$$

$$G_A^{(2)} = H_{AA}^{(2)} + H_{AB}^{(1)} L\tilde{H}_{BA}^{(1)} + \tilde{H}_{AB}^{(1)} LH_{BA}^{(1)} + \tilde{H}_{AB}^{(1)} LH_{BB}^{(0)} L\tilde{H}_{BA}^{(1)}$$

$$\begin{aligned} G_A^{(3)} = {} & H_{AA}^{(3)} + H_{AB}^{(2)} L\tilde{H}_{BA}^{(1)} + \tilde{H}_{AB}^{(1)} LH_{BA}^{(2)} \\ & + H_{AB}^{(1)} L[H_{BA}^{(2)} + H_{BB}^{(1)} L\tilde{H}_{BA}^{(1)} - L(S_{BA}^{(1)} + L\tilde{H}_{BA}^{(1)})\tilde{H}_{AA}^{(1)} + \epsilon_A^o L(S_{BA}^{(2)} + S_{BB}^{(1)} L\tilde{H}_{BA}^{(1)})] \\ & + [H_{AB}^{(2)} + \tilde{H}_{AB}^{(1)} LH_{BB}^{(1)} - \tilde{H}_{AA}^{(1)}(S_{AB}^{(1)} + \tilde{H}_{AB}^{(1)} L)L + \epsilon_A^o (S_{AB}^{(2)} + \tilde{H}_{AB}^{(1)} LS_{BB}^{(1)})] LH_{BA}^{(1)} \\ & + [H_{AB}^{(2)} + \tilde{H}_{AB}^{(1)} LH_{BB}^{(1)} - \tilde{H}_{AA}^{(1)}(S_{AB}^{(1)} + \tilde{H}_{AB}^{(1)} L)L + \epsilon_A^o (S_{AB}^{(2)} + \tilde{H}_{AB}^{(1)} LS_{BB}^{(1)}] \\ & \qquad \times LH_{BB}^{(0)} L\tilde{H}_{BA}^{(1)} \\ & + \tilde{H}_{AB}^{(1)} LH_{BB}^{(0)} L[H_{BA}^{(2)} + H_{BB}^{(1)} L\tilde{H}_{BA}^{(1)} - L(S_{BA}^{(1)} + L\tilde{H}_{BA}^{(1)})\tilde{H}_{AA}^{(1)} \\ & \qquad + \epsilon_A^o L(S_{BA}^{(2)} + S_{BB}^{(1)} L\tilde{H}_{BA}^{(1)})] \\ & + \tilde{H}_{AB}^{(1)} LH_{BB}^{(1)} L\tilde{H}_{BA}^{(1)} \end{aligned}$$

TABLE 6.24 $\tilde{H}_A^{(n)}$ -- Non-orthonormal Basis

(A-states degenerate)

$$\tilde{H}_A^{(0)} = H_{AA}^{(0)} = \epsilon_A^o 1_A$$

$$\tilde{H}_A^{(1)} = H_{AA}^{(1)} - \epsilon_A^o S_{AA}^{(1)} = \tilde{H}_{AA}^{(1)}$$

$$\tilde{H}_A^{(2)} = H_{AA}^{(2)} + H_{AB}^{(1)} L\tilde{H}_{BA}^{(1)} - \epsilon_A^o (S_{BA}^{(2)} + S_{AB}^{(1)} L\tilde{H}_{BA}^{(1)}) - \tfrac{1}{2}\{S_{AA}^{(1)}, H_{AA}^{(1)}\} + \epsilon_A^o S_{AA}^{(1)2}$$

$$G_{BA} = (\hat{T}^{\dagger} H \hat{T})_{BA} = H_{BA} + H_{BB} f + h^{\dagger}(H_{AA} + H_{AB} f) = 0, \qquad (6.57a)$$

and

$$g_{BA} = (\hat{T}^{\dagger} S \hat{T})_{BA} = S_{BA} + S_{BB} f + h^{\dagger}(S_{AA} + S_{AB} f) = 0, \qquad (6.57b)$$

When $H^{(0)}$ and $S^{(0)}$ are diagonal, expansion of these equations yields

$$H^{(0)}_{BB} f^{(n)} + h^{(n)\dagger} H^{(0)}_{AA} = -B_1^{(n)}, \qquad (6.58a)$$

and

$$f^{(n)} + h^{(n)\dagger} = -B_2^{(n)}, \qquad (6.58b)$$

where explicit use has been made of the condition $S^{(0)}=1$. Here, one has,

$$B_1^{(n)} = H_{BA}^{(n)} + \sum_{j=1}^{n-1} H_{BB}^{(n-j)} f^{(j)} + \sum_{j=1}^{n-1} h^{(j)\dagger} H_{AA}^{(n-j)} + \sum_{i=1}^{n-2} \sum_{j=1}^{n-i-1} h^{(i)\dagger} H^{(n-i-j)} f^{(j)}, \qquad (6.59a)$$

and

$$B_2^{(n)} = S_{BA}^{(n)} + \sum_{j=1}^{n-1} (S_{BB}^{(n-j)} f^{(j)} + h^{(j)\dagger} S_{AA}^{(n-j)}) + \sum_{i=1}^{n-2} \sum_{j=1}^{n-i-1} h^{(i)\dagger} S_{AB}^{(n-i-j)} f^{(j)}. \qquad (6.59b)$$

Equations (6.58) can be solved simultaneously for the matrix elements of $f^{(n)}$ and $h^{(n)}$, which are given by

$$f_{\sigma r}^{(n)} = \frac{(B_1^{(n)})_{\sigma r} - \epsilon_r^o (B_2^{(n)})_{\sigma r}}{\epsilon_r^o - \epsilon_\sigma^o}, \qquad (6.60a)$$

and

$$h_{\sigma r}^{(n)*} = \frac{-(B_1^{(n)})_{\sigma r} + \epsilon_\sigma^o (B_2^{(2)})_{\sigma r}}{\epsilon_r^o - \epsilon_\sigma^o}, \qquad (6.60b)$$

where the ϵ_i^o are the eigenvalues of $H^{(0)}$. If the overlap matrix is not perturbed, the quantities $B_2^{(n)}$ all vanish, and eq. (6.60a) reduces to eq. (6.9), and (6.60b) implies eq. (2.4).

If eqs. (6.57) are to be used as the basis of a perturbation formalism here, the series for h and f must both be considered simultaneously. This complication is offset by the simpler for of the expansions (6.59). Several low order terms in the series for G_{BA} and g_{BA} are given in Tables 6.19 and 6.20. Explicit expressions for the corresponding quantities $B_1^{(n)}$ and $B_2^{(n)}$ can be obtained from these tables by deleting the terms in $h^{(n)}$ and $f^{(n)}$.

If the A-states are degenerate, eqs. (6.58) can be written as

$$H_{BB}^{(0)} f^{(n)} - \epsilon_A^o h^{(n)\dagger} = -B_1^{(n)},$$

and (6.61)

$$f^{(n)} + h^{(n)\dagger} = -B_2^{(n)},$$

which can be solved as a system of two matrix equations in two unknown matrices. The solution is

$$f^{(n)} = L[B_1^{(n)} - \epsilon_A^o B_2^{(n)}],$$

and (6.62)

$$h^{(n)\dagger} = (\epsilon_A^o L - 1)B_2^{(n)} - LB_1^{(n)},$$

where L is given in eq. (6.17). If these equations are used to obtain expressions for the $f^{(n)}$ solely in terms of H and S, the expressions obtained are naturally identical to those from (6.16) and (6.17) with (6.55). However, eqs. (6.62) provide a more

efficient computational scheme for the calculation of high order terms in the series for f.

A collection of alternative formulas for the terms in the series for $\tilde{H}_A$ along with formulas for the metric g_A and related quantities have been given in Appendix 9. This type of perturbation theory is useful, for example, in extended Huckel molecular orbital theory.

TABLE 6.25 $f^{(n)}$ -- Non-orthonormal Basis

$$f_{\sigma r}^{(1)} = \frac{H_{\sigma r}^{(1)} - \epsilon_r^o S_{\sigma r}^{(1)}}{\epsilon_r^o - \epsilon_\sigma^o}$$

$$f_{\sigma r}^{(2)} = \Bigg[H_{\sigma r}^{(2)} + \sum_\rho (H_{\sigma\rho}^{(1)} - \epsilon_r^o S_{\sigma\rho}^{(1)}) \frac{(H_{\rho r}^{(1)} - \epsilon_r^o S_{\rho r}^{(1)})}{\epsilon_r^o - \epsilon_\rho^o} + \epsilon_r^o S_{\sigma r}^{(2)}$$

$$- \sum_t \left(S_{\sigma t}^{(1)} + \frac{H_{\sigma t}^{(1)} - \epsilon_t^o S_{\sigma t}^{(1)}}{\epsilon_t^o - \epsilon_\sigma^o} \right) (H_{tr}^{(1)} - \epsilon_r^o S_{tr}^{(1)}) \Bigg] \frac{1}{\epsilon_r^o - \epsilon_\sigma^o}$$

TABLE 6.26 $\hat{H}_A^{(n)}$ -- Non-orthonormal Basis

$$\hat{H}_{rs}^{(0)} = H_{rs}^{(0)}$$

$$\hat{H}_{rs}^{(1)} = H_{rs}^{(1)} - \epsilon_s^o S_{rs}^{(1)}$$

$$\hat{H}_{rs}^{(2)} = H_{rs}^{(2)} + \sum_\rho (H_{r\rho}^{(1)} - \epsilon_s^o S_{r\rho}^{(1)}) \frac{(H_{\rho s}^{(1)} - \epsilon_s^o S_{\rho s}^{(1)})}{\epsilon_s^o - \epsilon_\rho^o} - (S_{AA}^{(1)} H_{AA}^{(1)})_{rs}$$

$$- \epsilon_s^o S_{rs}^{(2)} + \epsilon_r^o (S_{AA}^{(1)2})_{rs}$$

TABLE 6.27 $G_A^{(n)}$ -- Non-orthonormal Basis

$$G_{rs}^{(0)} = H_{rs}^{(0)}$$

$$G_{rs}^{(1)} = H_{rs}^{(1)}$$

$$G_{rs}^{(2)} = H_{rs}^{(2)} + \sum_{\rho} H_{r\rho}^{(1)} \frac{(H_{\rho s}^{(1)} - \epsilon_s^o S_{\rho s}^{(1)})}{\epsilon_s^o - \epsilon_\rho^o} + \sum_{\rho} \frac{(H_{r\rho}^{(1)} - \epsilon_r^o S_{r\rho}^{(1)})}{\epsilon_r^o - \epsilon_\rho^o} H_{\rho s}^{(1)}$$

$$+ \sum_{\rho} \epsilon_\rho^o \frac{(H_{r\rho}^{(1)} - \epsilon_r^o S_{r\rho}^{(1)})(H_{\rho s}^{(1)} - \epsilon_s^o S_{\rho s}^{(1)})}{(\epsilon_r^o - \epsilon_\rho^o)(\epsilon_s^o - \epsilon_\rho^o)}$$

TABLE 6.28 $\tilde{H}_A^{(n)}$ -- Non-orthonormal Basis

$$\tilde{H}_{rs}^{(0)} = H_{rs}^{(0)}$$

$$\tilde{H}_{rs}^{(1)} = H_{rs}^{(1)} - \tfrac{1}{2}(\epsilon_r^o + \epsilon_s^o) S_{rs}^{(1)}$$

$$\tilde{H}_{rs}^{(2)} = H_{rs}^{(2)} + \sum_{\rho} H_{r\rho}^{(1)} \frac{(H_{\rho s}^{(1)} - \epsilon_s^o S_{\rho s}^{(1)})}{\epsilon_s^o - \epsilon_\rho^o} - \tfrac{1}{2} \sum_{t} (H_{rt}^{(1)} S_{ts}^{(1)} + S_{rt}^{(1)} H_{ts}^{(1)})$$

$$+\tfrac{1}{2}(\epsilon_r^o + \epsilon_s^o) \sum_{\rho} \frac{(H_{r\rho}^{(1)} - \epsilon_r^o S_{r\rho}^{(1)})}{\epsilon_r^o - \epsilon_\rho^o} \left[S_{\rho s}^{(1)} + \frac{H_{\rho s}^{(1)} - \epsilon_s^o S_{\rho s}^{(1)}}{\epsilon_s^o - \epsilon_\rho^o} \right]$$

$$+\tfrac{1}{2}(\epsilon_r^o + \epsilon_s^o) \left[S_{rs}^{(2)} + \sum_{\rho} S_{r\rho}^{(1)} \frac{(H_{\rho s}^{(1)} - \epsilon_s^o S_{\rho s}^{(1)})}{\epsilon_s^o - \epsilon_\rho^o} \right]$$

$$+\tfrac{3}{8}(\epsilon_r^o + \epsilon_s^o)(S_{AA}^{(1)^2})_{rs} + \tfrac{1}{4} \sum_{t} \epsilon_t^o S_{rt}^{(1)} S_{ts}^{(1)}$$

CHAPTER 7

EIGENVALUE INDEPENDENT PARTITIONING AND MOLECULAR ORBITAL THEORY

"'We have applied the same process,' Mein Herr continued, not noticing Bruno's question, 'to many other purposes. We have gone on selecting walking-sticks--always keeping those that walk best--till we have obtained some, that can walk by themselves! ...'"
(Sylvie and Bruno Concluded, Lewis Carroll

7.1 Introduction

The eigenvalue independent partitioning formalism developed in chapter 2 is particularly suited to situations in which only the whole space spanned by a subset of eigenvectors of some operator has significance, rather than the individual eigenvectors themselves. The mapping f is sufficient to determine the projection P_A', eq. (2.10), onto the subspace of interest, so that, in principle, all relevant properties can be determined once f has been calculated. One of the more important areas of quantum chemistry in which these aspects of the partitioning formalism can be exploited is in molecular orbital theory.

In molecular orbital theory, a closed shell system containing $2n_A$ electrons is represented by a Slater determinant made up from n_A doubly occupied orbitals. Since this determinantal wavefunction changes by at most a complex scalar factor under an arbitrary linear transformation of these occupied orbitals, the individual orbitals have no direct significance. In an n-dimensional basis space, the n_A occupied molecular orbitals are specified by $n_A n = n_A(n_A + n_B)$ complex numbers, the lcao (linear combination of atomic orbitals) coefficients. Since these n_A molecular orbitals are arbitrary up to an $n_A \times n_A$ linear transformation, so that n_A^2 of these complex numbers must be redundant, there are only $n_A n_B$ independent complex variables in the problem, which is exactly equal to the number of Brillouin conditions that must be satisfied. This is also the number of (complex) matrix elements in the mapping f defined in eq. (2.2), arising out of a partitioning of the eigenvector space of the hamiltonian into

an n_A-dimensional subspace spanned by the occupied molecular orbitals, and an n_B-dimensional subspace spanned by the unoccupied orbitals.[1] Thus, not only is the mapping f sufficient to determine the projection onto the space of the occupied orbitals, but it also represents the minimum amount of information required to specify that projection. The matrix elements of f contain no redundancies, and are subject to no constraints. These two properties of f are of considerable practical importance.

This chapter is primarily concerned with the derivation of perturbation formulas for the projection P_A onto the space of the occupied molecular orbitals. This projection is also frequently referred to as the one-particle density matrix in molecular orbital theory, and is equal to the charge-bond order matrix except for a factor of two. Both the simple matrix (Huckel theory), and the self-consistent field, cases are considered. The latter is more general than the matrix uncoupling considered hitherto in that the operator to be block diagonalized by f, itself depends on f. This chapter is restricted to consideration of closed shell systems only.

[1] The detailed nature of the partitioning of the basis space is not of central importance here, as long as f exists. Nevertheless, particular partitionings may be of special interest in certain cases because the elements of f then have a particular physical significance. One example is a basis made up of localized bond, lone pair, and antibond orbitals. When this space is partitioned into an n_A-dimensional subspace, S_A, spanned by the bond and lone pair orbitals, and an n_B-dimensional subspace, S_B, spanned by the antibonding orbitals, then the elements of f measure the delocalization of the bond and lone pair orbitals through the mixing in of antibond orbitals. In the same way, in a self-consistent field calculation, carried out in a Huckel basis, which diagonalizes the hamiltonian in the absence of explicit electron-electron interaction, and partitioned into occupied and unoccupied orbitals, the elements of f represent the magnitude of the mixing of these initially occupied and unoccupied orbitals because of the electron repulsion terms.

7.2 Perturbation of the Density Matrix --Orthonormal Basis

7.2.a General Theory

Consider a partitioning of the basis space into two subspaces, S_A and S_B, spanned by the orbitals occupied and unoccupied, respectively, in zero order. The projection P_A' onto a subspace S_A', spanned by the occupied perturbed orbitals, can be written (eq. (2.10)), as,

$$P_A' = \begin{bmatrix} g_A^{-1} & g_A^{-1}f^\dagger \\ fg_A^{-1} & fg_A^{-1}f^\dagger \end{bmatrix}, \qquad (7.1)$$

where $g_A = 1_A + f^\dagger f$. The perturbation series for f therefore determines series for each of the blocks of P_A', given by

$$\begin{aligned} (P_A')_{AA}^{(n)} &= g_A^{-1(n)}, \\ (P_A')_{BA}^{(n)} &= \sum_{j=1}^{n} f^{(j)} g_A^{-1(n-j)}, \end{aligned} \qquad (7.2)$$

and

$$(P_A')_{BB}^{(n)} = \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} f^{(i)} g_A^{-1(n-i-j)} f^{(j)\dagger},$$

when the zero order hamiltonian is at least block diagonal (that is, when $f^{(0)} = 0$). In the simple matrix case, the terms in the perturbation series for f are determined from the hierarchy of conditions $D^{(n)}(f) = 0$, eqs. (6.4). In a self-consistent field formalism, the equations defining the $f^{(n)}$ may be more complicated. They are considered in some detail in section 7.4.

From eqs. (7.2), it is seen that the first two terms in the series for P_A' are given by,

$$P_A'^{(0)} = P_A = \begin{bmatrix} 1_A & 0 \\ 0 & 0 \end{bmatrix}, \tag{7.3}$$

and

$$P_A'^{(1)} = \begin{bmatrix} 0 & f^{(1)\dagger} \\ f^{(1)} & 0 \end{bmatrix}, \tag{7.4}$$

where P_A is the projection onto the basis space S_A. The blocks of the second and higher order terms of P_A' are all non-vanishing in general. The special form of $P_A'^{(1)}$, in which the only non-vanishing matrix elements are those between the zero order occupied and unoccupied spaces, is a consequence of the absence of a first order contribution to the metric g_A.

Using the formulas given in section 6.2, the perturbation series for P_A' can be expressed solely in terms of the perturbed hamiltonian H. Tables 7.1 - 7.3 give these formulas for the elements of $(P_A')^{(n)}_{AA}$, $(P_A')^{(n)}_{BA}$, and $(P_A')^{(n)}_{BB}$, for n = 0,1,2, and 3. The case in which the A-states are all degenerate is not of great importance in molecular orbital theory, and no formulas for $(P_A')^{(n)}$ applicable to that case are included here.

The formulas in Tables 7.1 - 7.3 give the matrix elements of the perturbed density matrix in the basis of the zero order orbitals. These, in turn may be known in terms of some more primitive basis functions, for example, as a linear combination of atomic orbitals. The coefficients with respect to such a basis (eg. the lcao coefficients) will be denoted here as the columns of a unitary matrix C. The perturbed density matrix in the original basis will be denoted by R. The terms in the

perturbation series for R are given by

$$R^{(n)} = CP_A'^{(n)}C^\dagger,$$

or

$$R_{ij}^{(n)} = \sum_{r,s}^{occ} C_{ir}[(P_A')_{AA}^{(n)}]_{rs}C_{js}^* + \sum_{\sigma,\tau}^{unocc} C_{i\sigma}[(P')_{BB}^{(n)}]_{\sigma} C_{j\tau}^* \qquad (7.5)$$

$$+ \sum_{r}^{occ}\sum_{\sigma}^{unocc} C_{ir}[(P_A')_{AB}^{(n)}]_{r\sigma}C_{j\sigma'}^* + C_{i\sigma'}[(P_A')_{BA}^{(n)}]_{\sigma r}C_{jr}^*,$$

where the primes on the summation indices indicate that they are referring to basis elements in S_B (that is, numerically, $\sigma' = \sigma + n_A$ in $C_{r\sigma'}$).

In the simple matrix case (Huckel theory), the energy of the system described by the determinantal wavefunction made up of orbitals which span the image space of P_A' is given by

$$E = \nu \operatorname{tr} P_A'H, \qquad (7.6a)$$

ν being an occupation number for the orbitals. Using eqs. (7.2), a perturbation series for E in terms of f and H, and ultimately, in terms of H only, can be derived. The general formula is

$$E^{(n)} = \nu \sum_{j=0}^{n} \operatorname{tr}[(P_A')_{AA}^{(j)}H_{AA}^{(n-j)} + (P_A')_{AB}^{(j)}H_{BA}^{(n-j)} + (P_A')_{BA}^{(j)}H_{AB}^{(n-j)}$$

$$+ (P_A')_{BB}^{(j)}H_{BB}^{(n-j)}]. \qquad (7.6b)$$

Formulas for $E^{(n)}$ in terms of $f^{(n)}$ and $H^{(n)}$ are given in Table 7.4 through fifth order. By using the conditions $D^{(n)}(f) = 0$, eqs. (6.4), it is possible to obtain a 2n+1 rule here, in that

$E^{(2n)}$ and $E^{(2n+1)}$ can be written in terms of $f^{(1)}$ through $f^{(n)}$ only, as is done in the formulas in Table 7.4. Table 7.5 gives formulas for the $E^{(n)}$ in terms of the perturbed hamiltonian only, through third order.

The formalism presented above corresponds to some extent to that developed by McWeeny (1962), for the perturbation of the density matrix in the context of self-consistent field theory. Since self-consistency terms are not indicated explicitly in much of that derivation, the resulting formulas correspond closely to those derived above. The procedure used by McWeeny to derive a perturbation series for P_A' was to expand the equations[2]

$$[H, P_A'] = 0, \tag{7.7a}$$

and

$$P_A'^2 = P_A', \qquad P_A'^\dagger = P_A', \tag{7.7b}$$

in perturbation series, and then successively solve the hierarchy of simultaneous equations effectively for the blocks of P_A', as it is partitioned in eq. (7.1). The series obtained by McWeeny for P_A' is identical to that obtained here -- only the derivation is different. Here P_A' has been written in terms of a matrix f in such a way that eq. (7.7b) is automatically satisfied. The hierarchy of equations, $D^{(n)}(f) = 0$, defining the series for f, is equivalent to the hierarchy resulting from eq. (7.7a), as shown in section 2.3. In his derivation,

[2] McWeeny refers to the one-particle density matrix, denoted as P_A' here, by the symbol ρ in his 1962 paper.

McWeeny effectively takes the elements of $(P_A')^{(n)}_{BA}$ as the independent variables to be determined in the calculation, and calculates the elements of the other blocks of $(P_A')^{(n)}$ from them.

TABLE 7.1 $(P_A')^{(n)}_{AA}$ -- Molecular Orbital Basis

$$(P_A')^{(0)}_{AA} = 1_A$$

$$(P_A')^{(1)}_{AA} = 0$$

$$[(P_A')^{(2)}_{AA}]_{rs} = - \sum_{\sigma=1}^{n_B} \frac{H^{(1)}_{\sigma r} H^{(1)}_{\sigma s}}{(\epsilon^o_s - \epsilon^o_\sigma)(\epsilon^o_r - \epsilon^o_\sigma)}$$

$$[(P_A')^{(3)}_{AA}]_{rs} = - \sum_{\sigma=1}^{n_B} \Bigg[-H^{(1)}_{\sigma r} \sum_{\rho=1}^{n_B} \frac{H^{(1)}_{\sigma\rho} H^{(1)}_{\rho s}}{(\epsilon^o_s - \epsilon^o_\rho)} - \left(\sum_{\rho=1}^{n_B} \frac{H^{(1)}_{\sigma\rho} H^{(1)}_{\rho r}}{(\epsilon^o_r - \epsilon^o_\rho)} \right) H^{(1)}_{\sigma s}$$

$$-H^{(1)}_{\sigma r} \sum_{t=1}^{n_A} \frac{H^{(1)}_{\sigma t} H^{(1)}_{ts}}{(\epsilon^o_t - \epsilon^o_\sigma)} - \left(\sum_{t=1}^{n_A} \frac{H^{(1)}_{\sigma t} H^{(1)}_{tr}}{(\epsilon^o_t - \epsilon^o_\sigma)} \right) H^{(1)}_{\sigma s}$$

$$+ H^{(1)}_{\sigma r} H^{(2)}_{\sigma s} + H^{(2)}_{\sigma r} H^{(1)}_{\sigma s} \Bigg] \frac{1}{(\epsilon^o_s - \epsilon^o_\sigma)(\epsilon^o_r - \epsilon^o_\sigma)}$$

TABLE 7.2 $(P_A')_{BA}^{(n)}$ -- Molecular Orbital Basis

$$(P_A')_{BA}^{(0)} = 0$$

$$[(P_A')_{BA}^{(1)}]_{\sigma r} = \frac{H_{\sigma r}^{(1)}}{(\epsilon_r^o - \epsilon_\sigma^o)}$$

$$[(P_A')_{BA}^{(2)}]_{\sigma r} = H_{\sigma r}^{(2)} + \sum_{\rho=1}^{n_B} \frac{H_{\sigma\rho}^{(1)} H_{\rho r}^{(1)}}{(\epsilon_r^o - \epsilon_\rho^o)} - \sum_{t=1}^{n_A} \frac{H_{\sigma t}^{(1)} H_{tr}^{(1)}}{(\epsilon_t^o - \epsilon_\sigma^o)} \quad \frac{1}{(\epsilon_r^o - \epsilon_\sigma^o)}$$

$$[(P_A')_{BA}^{(3)}]_{\sigma r} = \Bigg[H_{\sigma r}^{(3)} + \sum_{\rho=1}^{n_B} \frac{H_{\sigma\rho}^{(2)} H_{\rho r}^{(1)}}{(\epsilon_r^o - \epsilon_\rho^o)} - \sum_{t=1}^{n_A} \frac{H_{\sigma t}^{(1)} H_{tr}^{(2)}}{(\epsilon_t^o - \epsilon_\sigma^o)}$$

$$+ \sum_{\rho=1}^{n_B} \frac{H_{\sigma\rho}^{(1)}}{(\epsilon_r^o - \epsilon_\rho^o)} \left(H_{\rho r}^{(2)} + \sum_{\gamma=1}^{n_B} \frac{H_{\rho\gamma}^{(1)} H_{\gamma r}^{(1)}}{(\epsilon_r^o - \epsilon_\gamma^o)} - \sum_{t=1}^{n_A} \frac{H_{\rho t}^{(1)} H_{tr}^{(1)}}{(\epsilon_t^o - \epsilon_\rho^o)} \right)$$

$$- \sum_{t=1}^{n_A} \left(H_{\sigma t}^{(2)} + \sum_{\rho=1}^{n_B} \frac{H_{\sigma\rho}^{(1)} H_{\rho t}^{(1)}}{(\epsilon_t^o - \epsilon_\rho^o)} - \sum_{s=1}^{n_A} \frac{H_{\sigma s}^{(1)} H_{sr}^{(1)}}{(\epsilon_s^o - \epsilon_\sigma^o)} \right) \frac{H_{tr}^{(1)}}{(\epsilon_t^o - \epsilon_\sigma^o)}$$

$$+ \sum_{\rho=1}^{n_B} \sum_{t=1}^{n_A} \frac{H_{\sigma t}^{(1)} H_{t\rho}^{(1)} H_{\rho r}^{(1)}}{(\epsilon_t^o - \epsilon_\sigma^o)(\epsilon_r^o - \epsilon_\rho^o)} \Bigg] \frac{1}{(\epsilon_r^o - \epsilon_\sigma^o)}$$

$$- \sum_{s=1}^{n_A} \sum_{\rho=1}^{n_B} \frac{H_{\sigma s}^{(1)} H_{s\rho}^{(1)} H_{\rho r}^{(1)}}{(\epsilon_s^o - \epsilon_\sigma^o)(\epsilon_s^o - \epsilon_\rho^o)(\epsilon_r^o - \epsilon_\rho^o)}$$

TABLE 7.3 $(P_A')_{BB}^{(n)}$ -- Molecular Orbital Basis

$$(P_A')_{BB}^{(0)} = 0$$

$$(P_A')_{BB}^{(1)} = 0$$

$$\left[(P_A')_{BB}^{(2)}\right]_{\sigma\rho} = \sum_{r=1}^{n_A} \frac{H_{\sigma r}^{(1)} H_{\rho r}^{(1)}}{(\epsilon_r^o - \epsilon_\sigma^o)(\epsilon_r^o - \epsilon_\rho^o)}$$

$$\left[(P_A')_{BB}^{(3)}\right]_{\sigma\rho} = \sum_{r=1}^{n_A} \Bigg[H_{\sigma r}^{(1)} H_{\rho r}^{(2)} + H_{\sigma r}^{(2)} H_{\rho r}^{(1)} + H_{\sigma r}^{(1)} \sum_{\gamma=1}^{n_B} \frac{H_{\rho\gamma}^{(1)} H_{\gamma r}^{(1)}}{(\epsilon_r^o - \epsilon_\gamma^o)}$$

$$+ \left(\sum_{\gamma=1}^{n_B} \frac{H_{\sigma\gamma}^{(1)} H_{\gamma r}^{(1)}}{(\epsilon_r^o - \epsilon_\gamma^o)} \right) H_{\rho r}^{(1)} - H_{\sigma r}^{(1)} \sum_{t=1}^{n_A} \frac{H_{\rho t}^{(1)} H_{t r}^{(1)}}{(\epsilon_t^o - \epsilon_\rho^o)}$$

$$+ \left(\sum_{t=1}^{n_A} \frac{H_{\sigma t}^{(1)} H_{t r}^{(1)}}{(\epsilon_t^o - \epsilon_\sigma^o)} \right) H_{\rho r}^{(1)} \Bigg] \frac{1}{(\epsilon_r^o - \epsilon_\sigma^o)(\epsilon_r^o - \epsilon_\rho^o)}$$

TABLE 7.4 $E^{(n)}$ -- Molecular Orbital Basis

$$E^{(0)} = \nu \, \mathrm{tr} \, H_{AA}^{(0)}$$

$$E^{(1)} = \nu \, \mathrm{tr} \, H_{AA}^{(1)}$$

$$E^{(2)} = \nu \, \mathrm{tr} \left[H_{AA}^{(2)} + H_{AB}^{(1)} f^{(1)} \right]$$

$$E^{(3)} = \nu \, \mathrm{tr} \left[H_{AA}^{(3)} - f^{(1)\dagger} f^{(1)} H_{AA}^{(1)} + f^{(1)} f^{(1)\dagger} H_{BB}^{(1)} + f^{(1)\dagger} H_{BA}^{(2)} + f^{(1)} H_{AB}^{(2)} \right]$$

$$E^{(4)} = \nu \, \mathrm{tr} \left[H_{AA}^{(4)} + f^{(1)\dagger} H_{BA}^{(3)} + f^{(1)} H_{AB}^{(3)} - f^{(1)\dagger} f^{(1)} H_{AA}^{(2)} + f^{(1)} f^{(1)\dagger} H_{BB}^{(2)} - f^{(1)} f^{(1)\dagger} f^{(1)} H_{AB}^{(1)} + H_{AA}^{(0)} f^{(2)\dagger} f^{(2)} - f^{(2)} f^{(2)\dagger} H_{BB}^{(0)} \right]$$

$$\begin{aligned} E^{(5)} = \nu \, \mathrm{tr} \big[& H_{AA}^{(5)} + f^{(1)} H_{AB}^{(4)} + f^{(1)\dagger} H_{BA}^{(4)} + f^{(1)} f^{(1)\dagger} H_{BB}^{(3)} - f^{(1)\dagger} f^{(1)} H_{AA}^{(3)} \\ & + f^{(2)\dagger} H_{BA}^{(3)} + f^{(2)} H_{AB}^{(3)} + (f^{(2)} f^{(1)\dagger} + f^{(1)} f^{(2)\dagger}) H_{BB}^{(2)} \\ & - (f^{(2)\dagger} f^{(1)} + f^{(1)\dagger} f^{(2)}) H_{AA}^{(2)} - f^{(2)\dagger} f^{(2)} H_{AA}^{(1)} + f^{(2)} f^{(2)\dagger} H_{AA}^{(1)} \\ & - f^{(1)\dagger} f^{(2)} f^{(1)\dagger} H_{BA}^{(1)} - f^{(1)} f^{(2)\dagger} f^{(1)} H_{AB}^{(1)} \\ & + (f^{(1)} f^{(1)\dagger})^2 H_{BB}^{(1)} - (f^{(1)\dagger} f^{(1)})^2 H_{AA}^{(1)} + f^{(1)\dagger} f^{(1)} f^{(2)\dagger} H_{BB}^{(0)} f^{(1)} \\ & + f^{(1)\dagger} f^{(1)} f^{(1)\dagger} H_{BB}^{(0)} f^{(2)} - f^{(1)} f^{(1)\dagger} f^{(2)} H_{AA}^{(0)} f^{(1)} \\ & - f^{(1)} f^{(1)\dagger} f^{(1)} H^{(0)} f^{(2)\dagger} \big] \end{aligned}$$

TABLE 7.5 $E^{(n)}$ -- Molecular Orbital Basis

$$E^{(0)} = \nu \sum_{r=1}^{n_A} \epsilon_r^o$$

$$E^{(1)} = \nu \sum_{r=1}^{n_A} H_{rr}^{(1)}$$

$$E^{(2)} = \nu \sum_{r=1}^{n_A} \left[H_{rr}^{(2)} \quad \sum_{\sigma=1}^{n_B} \frac{H_{r\sigma}^{(1)} H_{\sigma r}^{(1)}}{(\epsilon_r^o - \epsilon_\sigma^o)} \right]$$

$$E^{(3)} = \nu \sum_{r=1}^{n_A} \left[H_{rr}^{(3)} - \sum_{s=1}^{n_A} \sum_{\sigma=1}^{n_B} \frac{H_{r\sigma}^{(1)} H_{\sigma s}^{(1)} H_{sr}^{(1)}}{(\epsilon_r^o - \epsilon_\sigma^o)(\epsilon_s^o - \epsilon_\sigma^o)} + \sum_{\sigma=1}^{n_B} \frac{H_{r\sigma}^{(1)} H_{\sigma r}^{(2)}}{(\epsilon_r^o - \epsilon_\sigma^o)} \right]$$

$$+ \nu \sum_{\sigma=1}^{n_B} \left[\sum_{\rho=1}^{n_B} \sum_{s=1}^{n_A} \frac{H_{\sigma s}^{(1)} H_{s\rho}^{(1)} H_{\rho\sigma}^{(1)}}{(\epsilon_s^o - \epsilon_\sigma^o)(\epsilon_s^o - \epsilon_\rho^o)} + \sum_{s=1}^{n_A} \frac{H_{\sigma s}^{(1)} H_{s\sigma}^{(1)}}{(\epsilon_s^o - \epsilon_\sigma^o)} \right]$$

7.2.b Huckel Molecular Orbital Theory

As an illustration of the straightforward way in which the tabulated perturbation formulas for $P_A^{\bullet}$ can be used, expressions for the bond-bond, bond-atom, atom-bond, and atom-atom polarizabilities, as defined by Coulson and Longuett-Higgins (1947) in Huckel molecular orbital theory, will be derived. These quantities are proportional to the first order response of diagonal and off-diagonal elements of the charge-bond order matrix, $P = 2R$, where R denotes the density matrix in the atomic orbital basis, (or the second order response of the energy, (7.6)), to a perturbation of diagonal or off-diagonal elements of the hamiltonian, $H^{(0)}$, in the atomic orbital basis. Thus, the results below are determined by combining the formulas in Tables 7.1 - 7.3 with eq. (7.5).

First, consider the single center perturbation given by

$$(H_{AO}^{(1)})_{pq} = \delta\alpha_t \, \delta_{pt}\delta_{qt} \, , \qquad (7.8a)$$

representing a change in $H_{tt}^{(0)}$ by an amount $\delta\alpha_t$. On transforming to the zero order molecular orbital basis, this becomes

$$(H_{MO}^{(1)})_{ij} = \delta\alpha_t C_{ti} C_{tj} \, , \quad (i,j = 1, \ldots, n), \qquad (7.8b)$$

assuming the C_{ij} are all real. Substitution of (7.8b) into the first order formulas for $P_A^{\bullet}$ in Tables 7.1 - 7.3, and back-transformation to the atomic orbital basis via (7.5) then gives the results,

$$\frac{\partial R_{ij}}{\partial H_{tt}} = \frac{1}{\delta\alpha_t} (R^{(1)})_{ij} \, , \qquad (7.9a)$$

where,

$$\frac{1}{\delta\alpha_t} R_{ij}^{(1)} = \sum_{r}^{occ} \sum_{\sigma}^{unocc} \frac{C_{tr}C_{t\sigma'}}{\epsilon_r^o - \epsilon_{\sigma'}^o} [C_{i\sigma'}C_{jr} + C_{ir}C_{j\sigma'}], \quad (7.9b)$$

and, in particular, for the diagonal elements,

$$\frac{1}{\delta\alpha_t} R_{ii}^{(1)} = 2 \sum_{r}^{occ} \sum_{\sigma}^{unocc} \frac{C_{t\sigma'}C_{tr}C_{i\sigma'}C_{ir}}{\epsilon_r^o - \epsilon_{\sigma'}^o} \quad (7.9c)$$

These quantities are respectively the atom-bond and atom-atom polarizabilities ($\pi_{ij,t}$ and $\pi_{i,t}$) to within a factor of two, as defined by Coulson and Longuett-Higgins (1947).

One may obtain second derivatives of the elements of R with respect to one or more diagonal elements of H in an analogous manner. A summation over all $\delta\alpha_t$ is incorporated into eq. (7.8), allowing for a simultaneous perturbation of all diagonal elements of $H_{AO}^{(0)}$. The second derivatives of elements of R with respect to two diagonal elements, H_{pp} , and H_{qq}, of H_{AO} are obtained by isolating the coefficient of $\delta\alpha_p\delta\alpha_q$ in $R^{(2)}$,

$$\frac{\partial^2 R_{ij}}{\frac{1}{2}\partial H_{pp}\partial H_{qq}} = \sum_{r}^{occ} \sum_{\sigma}^{unocc} \frac{C_{p\sigma'}C_{pr}}{\epsilon_r^o - \epsilon_{\sigma'}^o} \left[- \sum_{s}^{occ} \frac{C_{qs}C_{q\sigma'}C_{ir}C_{js}}{\epsilon_s^o - \epsilon_{\sigma'}^o} \right.$$

$$\left. + \sum_{\rho}^{unocc} \frac{C_{q\rho'}C_{qr}C_{i\sigma'}C_{j\rho'}}{\epsilon_r^o - \epsilon_{\rho'}^o} \right]$$

$$+ \sum_{r}^{occ} \sum_{\sigma}^{unocc} \frac{C_{p\sigma'}C_{qr}}{\epsilon_r^o - \epsilon_{\sigma'}^o} [C_{ir}C_{j\sigma'} + C_{i\sigma'}C_{jr}] \quad (7.10)$$

$$\times \left[\sum_{\rho}^{unocc} \frac{C_{p\rho'}C_{q\rho'}}{\epsilon_r^o - \epsilon_{\rho'}^o} - \sum_{s}^{occ} \frac{C_{ps}C_{qs}}{\epsilon_s^o - \epsilon_{\sigma'}^o} \right].$$

While this is considerably more complicated than the first order formula, it is nevertheless obtained quite straightforwardly from the formulas tabulated for P_A'. This procedure can be continued to arbitrary order, but the explicit formulas rapidly become more complex and less useful. The computation of high order terms can be done more efficiently by successively calculating the $f^{(i)}$ and $g_A^{-1(i)}$, evaluating the $P_A'^{(i)}$ in terms of these quantities numerically using eqs. (7.2), and then transforming to the atomic orbital basis. The attractive feature of the derivation of (7.9) and (7.10) above is that the various summations in the formulas for the $R_{ij}^{(n)}$ appear automatically as being either over occupied orbitals or over unoccupied orbitals. This is not so when conventional perturbation formulas, based on the perturbation series for the occupied orbitals, are used, for which the derivation and simplification of formulas for $R^{(n)}$ for $n > 1$ becomes very laborious.

For a two-center perturbation given by $(H_{AO}^{(1)})_{pq} = (H_{AO}^{(1)})_{qp} = \delta\beta_{pq}$, the matrix of the perturbation in the molecular orbital basis is

$$(H_{MO}^{(1)})_{rs} = \delta\beta_{pq}[C_{pr}C_{qs} + C_{qr}C_{ps}]. \qquad (7.11)$$

The formulas in Tables 7.1 -7.3 and eq. (7.5) yield immediately the bond-bond and bond-atom polarizabilities, $\pi_{ij,pq}$ and $\pi_{i,pq}$, (to within a factor of 2), respectively,

$$\frac{\partial R_{ij}}{\partial H_{pq}} = \frac{1}{\delta\beta_{pq}} R_{ij}^{(1)}, \qquad (7.12a)$$

where

$$\frac{1}{\delta\beta_{pq}} R^{(1)}_{ij} = \sum_{r}^{occ} \sum_{\sigma}^{unocc} \frac{[C_{p\sigma}, C_{qr}+C_{q\sigma}, C_{pr}][C_{i\sigma}, C_{jr}+C_{ir}C_{i\sigma},]}{\epsilon^{o}_{r} - \epsilon^{o}_{\sigma},} \qquad (7.12b)$$

and

$$\frac{1}{\delta\beta_{pq}} R^{(1)}_{ii} = 2 \sum_{r}^{occ} \sum_{\sigma}^{unocc} \frac{C_{i\sigma}, C_{ir}}{\epsilon^{o}_{r} - \epsilon^{o}_{\sigma},} [C_{p\sigma}, C_{qr}+C_{q\sigma}, C_{pr}]. \qquad (7.12c)$$

Higher order formulas are obtained straightforwardly here also, but even in second order, they are lengthy and none will be given here.

7.2.c Numerical Example -- Huckel Theory

To obtain some information on the nature and usefulness of high order perturbation series for the charge-bond order matrix, P = 2R, a number of numerical calculations were carried out, based on three Huckel-like hamiltonians. The first two of them were

$$H^{(0)}_{A_6} = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & -1 \\ -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & -1 \\ -1 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}, \qquad (7.13)$$

which could represent a six-membered ring of identical atoms in Huckel theory (eg. benzene), and,

$$H^{(0)}_{A_4B_4} = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 & 0 & 0 & -1 \\ -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 \\ -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}, \quad (7.14)$$

representing an eight-membered ring around which two kinds of atoms alternate (eg. P_4N_4). The third example, denoted as $H^{(0)}_{A_5B_3}$, is obtained by setting the (1,1) element of $H^{(0)}_{A_4B_4}$ to zero. Series for P_{11} and P_{12} for a single-center perturbation (H_{11} varying) only will be described. They can be written as

$$P_{ij} = \sum_{n=0} P^{(n)}_{ij} (\delta H_{11})^n, \quad (7.15)$$

and coefficients in each case for n = 0,1,2,3, and 4, are given in Tables 7.6 - 7.8.

Because of the symmetry of $H^{(0)}_{A_6}$, the series for P_{11} contains only odd order terms, while that for P_{12} contains only even order terms. For the A_4B_4 and A_5B_3 systems, none of the coefficients in the series for P_{11} and P_{12} through fourth order are zero. The coefficients in the series obtained decrease in magnitude quite rapidly, with the fourth order coefficient being smaller than the zero order term by a factor of up to several hundred. The coefficients given in Tables 7.6 - 7.8 are both positive and negative, but no pattern in sign is recognizable to fourth order. Plots of exact values of P_{11} and P_{12} as a

function of H_{11}, along with plots of the first through fourth order series approximating these quantities are given in Figures 7.1 - 7.6.

The two matrices, $H^{(0)}_{A_4B_4}$ and $H^{(0)}_{A_5B_3}$, can be considered as alternative zero order terms when the (1,1) element only is to be perturbed. Thus, the exact quantities P_{11} and P_{12}, considered as functions of H_{11}, are identical in both cases. The alternative series given in Tables 7.7 and 7.8 are then seen to be power series expansions of the functions $P_{11}(H_{11})$ and $P_{12}(H_{11})$ around two different values of H_{11}.

Two potential pitfalls in the use of high order perturbation series, which warrant some emphasis, are illustrated by the results here. These are rather obvious dangers which apply quite generally to the use of any truncated power series expansion, which is what such finite perturbation approximations actually are. Firstly, as the size of the perturbation increases, the error in a higher (but finite) order partial sum eventually becomes larger than the errors in the lower order truncations of the series, although, by the time this occurs, none of the partial sums of lower order may be sufficiently accurate to be useful. Thus, while the inclusion of the next higher order term in a given series will generally increase the accuracy of the approximation when the perturbation is small, it may substantially decrease the accuracy if the perturbation is large. Secondly, the range of acceptable accuracy of an approximation to a given order depends signifi-

cantly on the zero order approximation used. As seen from Figures 7.4 and 7.6, the first order approximation for $P_{12}(H_{11})$ has a considerably wider range of usefulness around $H_{11}^{(0)} = -1$ than around $H_{11}^{(0)} = 0$. If an approximation for P_{12} as a function of H_{11} is desired for $-1 < H_{11} < 0$, it is clear that the zero order hamiltonian $H^{(0)}_{A_4B_4}$ is superior to $H^{(0)}_{A_5B_3}$.

TABLE 7.6 $(P_{A0})^{(i)}$ for A_6 System ($H_{11}^{(0)} = 0$)

$P_{11}^{(0)} = 1.0$	$P_{12}^{(0)} = 0.666667$ (2/3)
$P_{11}^{(1)} = -0.398148$	$P_{12}^{(1)} = 0.0$
$P_{11}^{(2)} = 0.0$	$P_{12}^{(2)} = -0.053626$
$P_{11}^{(3)} = 0.031875$	$P_{12}^{(3)} = 0.0$
$P_{11}^{(4)} = 0.0$	$P_{12}^{(4)} = 0.006489$

TABLE 7.7 $(P_{AO})^{(i)}$ for an A_4B_4 System ($H_{11}^{(0)} = -1$)

$P_{11}^{(0)} = 1.477301$	$P_{12}^{(0)} = 0.575869$
$P_{11}^{(1)} = -0.273157$	$P_{12}^{(1)} = 0.104447$
$P_{11}^{(2)} = -0.068263$	$P_{12}^{(2)} = -0.009937$
$P_{11}^{(3)} = -0.001732$	$P_{12}^{(3)} = -0.011540$
$P_{11}^{(4)} = 0.005201$	$P_{12}^{(4)} = -0.002410$

TABLE 7.8 $(P_{AO})^{(i)}$ for an A_5B_3 System ($H_{11}^{(0)} = 0$)

$P_{11}^{(0)} = 1.140825$	$P_{12}^{(0)} = 0.657296$
$P_{11}^{(1)} = -0.387179$	$P_{12}^{(1)} = 0.045056$
$P_{11}^{(2)} = -0.030147$	$P_{12}^{(2)} = -0.048380$
$P_{11}^{(3)} = 0.027841$	$P_{12}^{(3)} = -0.008863$
$P_{11}^{(4)} = 0.005568$	$P_{12}^{(4)} = 0.004953$

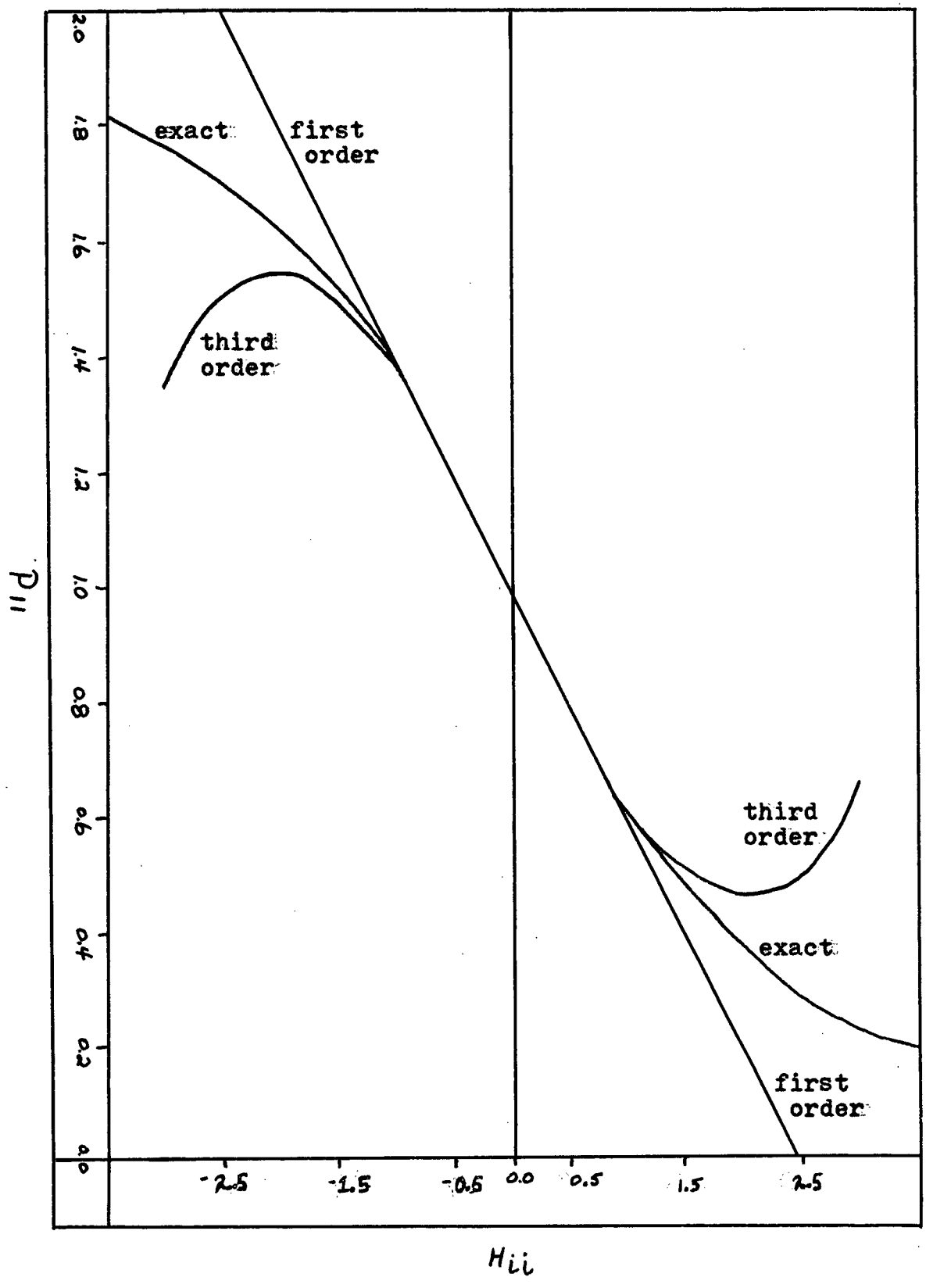

FIGURE 7.1 P_{11} vs. H_{ii} for the A_6 System (units of β).

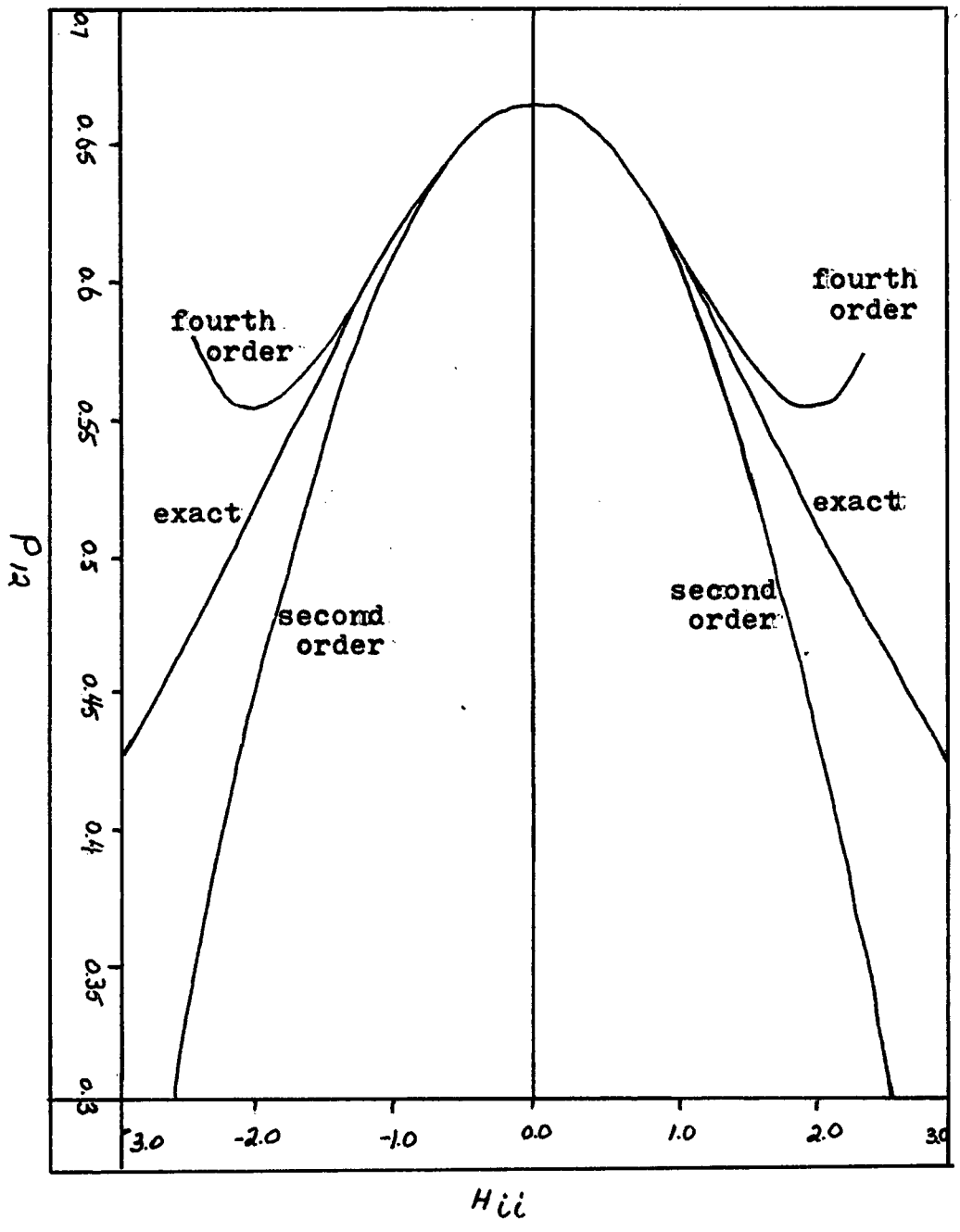

FIGURE 7.2 P_{12} vs H_{ii} for the A_6 System (units of β).

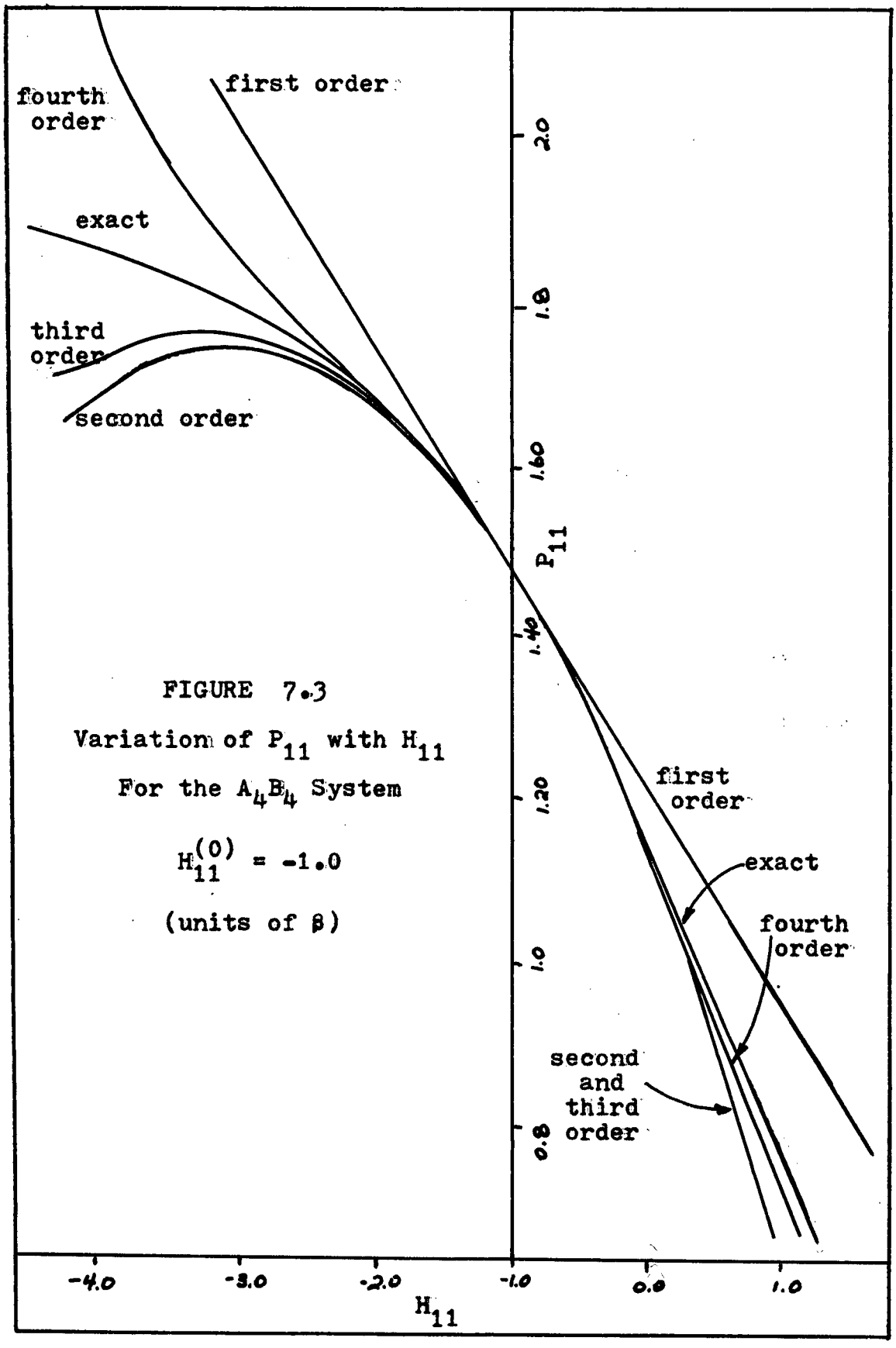

FIGURE 7.3

Variation of P_{11} with H_{11}

For the A_4B_4 System

$H_{11}^{(0)} = -1.0$

(units of β)

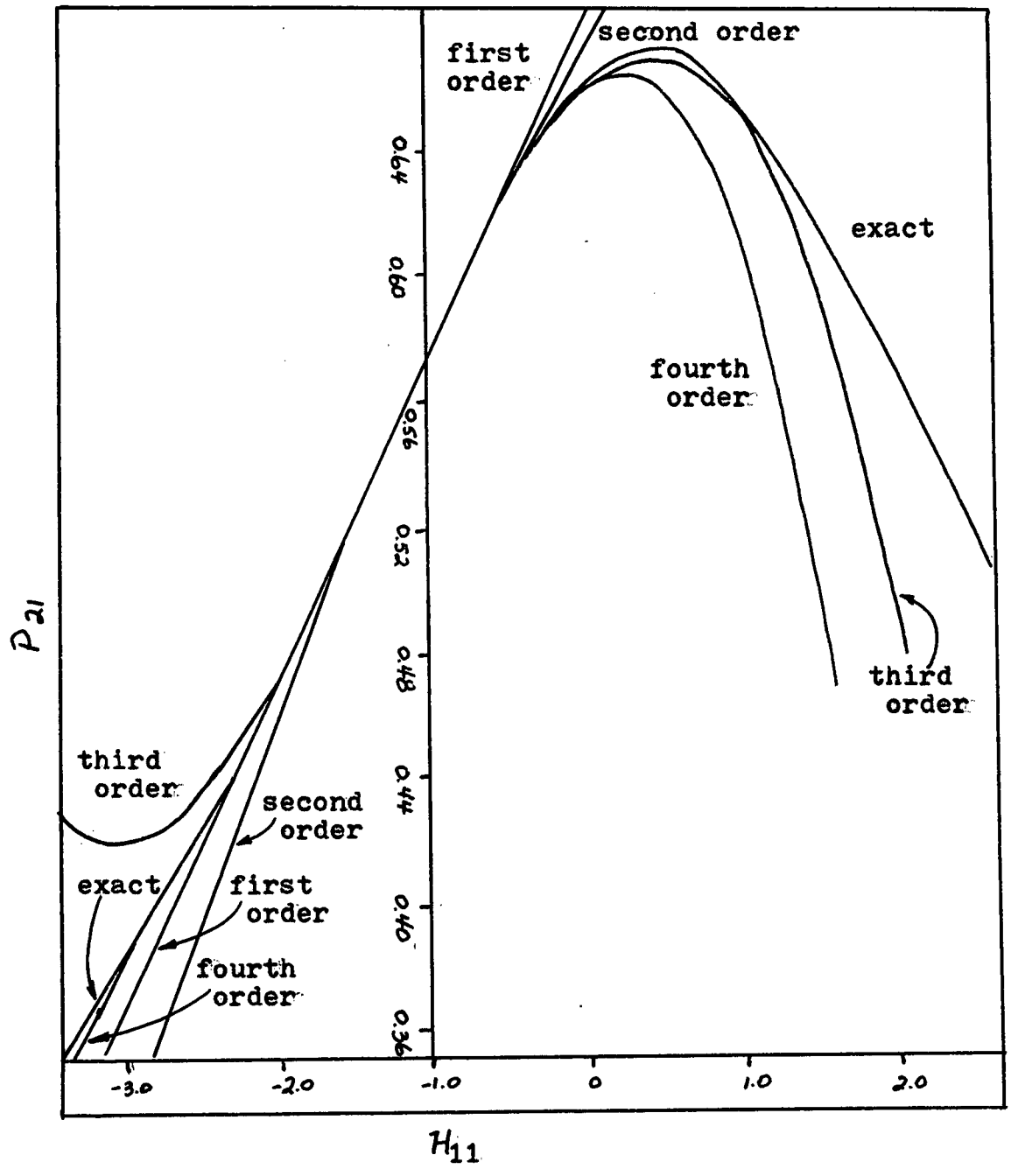

FIGURE 7.4 P_{21} vs. H_{11} for the A_4B_4 System (units of β).

$$H_{11}^{(0)} = -1.0$$

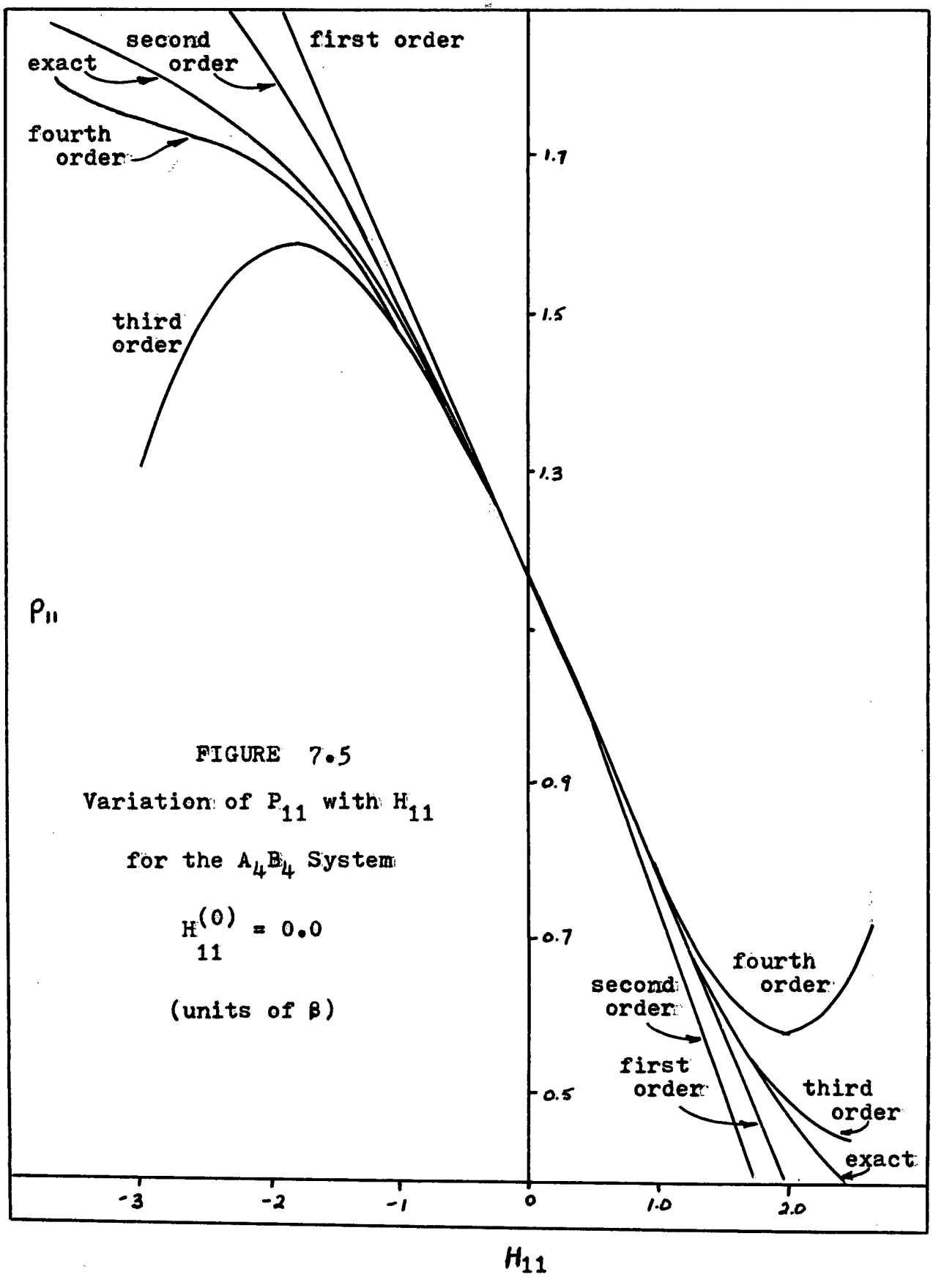

FIGURE 7.5

Variation of P_{11} with H_{11}

for the A_4B_4 System

$H_{11}^{(0)} = 0.0$

(units of β)

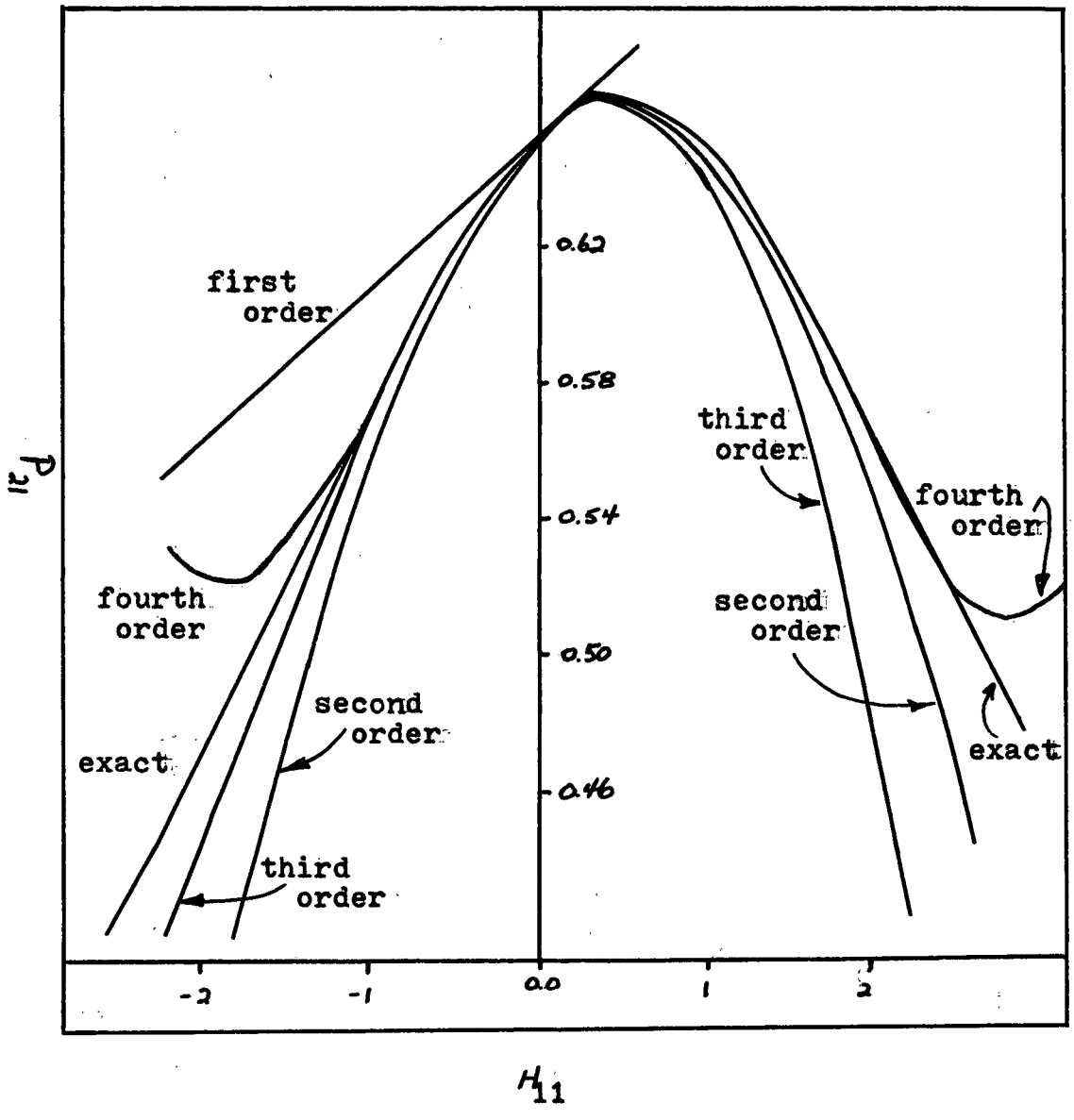

FIGURE 7.6 Variation of P_{21} with H_{11} for the A_4B_4 System.

$H_{11}^{(0)} = 0.0$ (units of β)

7.3 Perturbation of the Density Matrix -- Non-orthonormal Basis

7.3.a General Theory

In this section, perturbation formulas for the density matrix are developed for the case in which the primitive (atomic orbital) basis is initially non-orthonormal, and in which the overlap between these orbitals may itself be perturbed. Such a situation would arise, for example, for a perturbation involving a bond length change in a Huckel-type molecular orbital formalism. The major complications here are the explicit presence of the perturbed overlap matrix, and the fact that the transformation between the zero order molecular orbitals and the atomic orbital basis is now non-unitary. The projection, P_A', onto the space of occupied orbitals, is still given by eq. (7.2), and therefore, the formal expansions, (7.3), still hold. However, now the formulas for the $f^{(j)}$ and $g_A^{-1(j)}$ must be obtained from section 6.4.

The initial perturbation series are calculated in a basis of zero order molecular orbitals, with coefficients relative to the atomic orbital basis denoted here as the columns of the (generally non-unitary) matrix C. That is, in the calculation of the $f^{(j)}$ and $g_A^{-1(j)}$, the perturbations are

$$H_{MO}^{(n)} = C^\dagger H_{AO}^{(n)} C \ , \qquad S_{MO}^{(n)} = C^\dagger S_{AO}^{(n)} C, \quad (n=0,1,2,\ldots), \tag{7.16}$$

where $H_{MO}^{(0)}$ and $S_{MO}^{(0)}$ are to be at least block diagonal (so that $f^{(0)}=0$). When $C^\dagger S_{AO}^{(0)} C = S_{MO}^{(0)} = 1$, the transformation of the density matrix, P_A', from the molecular orbital basis to

the original atomic orbital basis can be written as

$$R = C P_A' C^\dagger \tag{7.17}$$

When $H_{MO}^{(0)}$ is diagonal, explicit formulas for the elements of P_A' (and R), in terms of those of H, S, and C, only, can be written down. It will be assumed that $H_{MO}^{(0)}$ is diagonal and that $S_{MO}^{(0)} = 1$ in what follows.

The zero order term of P_A' is still given by eq. (7.3). However, the first order term is now

$$P_A'^{(1)} = \begin{bmatrix} -(S_{MO}^{(1)})_{AA} & f^{(1)\dagger} \\ f^{(1)} & 0 \end{bmatrix} . \tag{7.18}$$

The matrix elements between zero order occupied orbitals appear as a result of the perturbation of the overlap matrix. Explicit formulas for the matrix elements of the blocks of $P_A'^{(n)}$ in terms of H_{MO} and S_{MO} only are given in Tables 7.9, 7.10, and 7.11 for n = 0, 1, and 2.

A perturbation series for the Huckel energy E (eq. (7.6)) can again be obtained using eq. (7.7). Expressions for the $E^{(n)}$ in terms of f, H, and S only are given in Table 7.12 for n = 0, 1, 2, and 3. No difficulty is encountered in eliminating $f^{(2)}$ from the expression obtained via (7.7) for $E^{(3)}$. However, no attempt was made to verify that $E^{(4)}$ and $E^{(5)}$ can be written down solely in terms of $f^{(1)}$ and $f^{(2)}$ by using the conditions defining the $f^{(n)}$, as was done for the case of an orthonormal basis. Formulas for the $E^{(n)}$ in terms of the elements of H and S only are given for n = 0, 1, and 2 in Table 7.12.

TABLE 7.9 $(P_A')^{(n)}_{AA}$ -- Non-orthonormal Basis

$$(P_A')^{(0)}_{AA} = 1_A$$

$$(P_A')^{(1)}_{AA} = -S^{(1)}_{AA}$$

$$[(P_A')^{(2)}_{AA}]_{rs} = -S^{(2)}_{rs} + \sum_{\rho=1}^{n_B} S^{(1)}_{r\rho}\frac{(H^{(1)}_{\rho s}-S^{(1)}_{\rho s}\epsilon^o_s)}{(\epsilon^o_s-\epsilon^o_\rho)} + \sum_{t=1}^{n_A} S^{(1)}_{rt}S^{(1)}_{ts}$$

$$+ \sum_{\rho=1}^{n_B} \frac{(H^{(1)}_{r\rho}-\epsilon^o_r S^{(1)}_{r\rho})(H^{(1)}_{\rho s}-S^{(1)}_{s\rho}\epsilon^o_\rho)}{(\epsilon^o_r-\epsilon^o_\rho)(\epsilon^o_s-\epsilon^o_\rho)}$$

TABLE 7.10 $(P_A')^{(n)}_{BA}$ -- Non-orthonormal Basis

$$(P_A')^{(0)}_{BA} = 0$$

$$[(P_A')^{(1)}_{BA}]_{\sigma r} = \frac{H^{(1)}_{\sigma r} - \epsilon^o_r S^{(1)}_{\sigma r}}{\epsilon^o_r - \epsilon^o_\sigma}$$

$$[(P_A')^{(2)}_{BA}]_{\sigma r} = \left[H^{(2)}_{\sigma r} - \epsilon^o_r S^{(2)}_{\sigma r} + \sum_{\rho=1}^{n_B} \frac{(H^{(1)}_{\sigma\rho}-\epsilon^o_r S^{(1)}_{\sigma\rho})(H^{(1)}_{\rho r}-\epsilon^o_r S^{(1)}_{\rho r})}{(\epsilon^o_r - \epsilon^o_\rho)} \right.$$

$$\left. - \sum_{t=1}^{n_A} \frac{(H^{(1)}_{\sigma t}-S^{(1)}_{\sigma t}\epsilon^o_\sigma)}{(\epsilon^o_t-\epsilon^o_\sigma)} H^{(1)}_{tr} \right] \frac{1}{\epsilon^o_r - \epsilon^o_\sigma}$$

$$- \sum_{t=1}^{n_A} \frac{(H^{(1)}_{\sigma t}-\epsilon^o_t S^{(1)}_{\sigma t})}{(\epsilon^o_t-\epsilon^o_\sigma)} S^{(1)}_{tr}$$

TABLE 7.11 $(P_A')^{(n)}_{BB}$ -- Non-orthonormal Basis

$$(P_A')^{(0)}_{BB} = 0$$

$$(P_A')^{(1)}_{BB} = 0$$

$$[(P_A')^{(2)}_{BB}]_{\sigma\rho} = \sum_{r=1}^{n_A} \frac{(H^{(1)}_{\sigma r} - S^{(1)}_{\sigma r}\epsilon^o_r)(H^{(1)}_{r\rho} - S^{(1)}_{r\rho}\epsilon^o_r)}{(\epsilon^o_r - \epsilon^o_\sigma)(\epsilon^o_r - \epsilon^o_\rho)}$$

TABLE 7.12 $E^{(n)}$ -- Non-orthonormal Basis

$$E^{(0)} = \nu \operatorname{tr} H^{(0)}_{AA}$$

$$E^{(1)} = \nu \operatorname{tr}(H^{(1)}_{AA} - S^{(1)}_{AA}H^{(0)}_{AA})$$

$$E^{(2)} = \nu \operatorname{tr}[H^{(2)}_{AA} - S^{(1)}_{AA}H^{(1)}_{AA} + f^{(1)}H^{(1)}_{AB} - (S^{(2)}_{AA} + S^{(1)}_{AB}f^{(1)} - S^{(1)2}_{AA})H^{(0)}_{AA}]$$

$$\begin{aligned} E^{(3)} = \nu \operatorname{tr}\; & H^{(3)}_{AA} + f^{(1)}H^{(2)}_{AB} + H^{(2)}_{BA}f^{(1)\dagger} \\ & - [S^{(2)}_{AA} + S^{(1)}_{AB}f^{(1)} + f^{(1)\dagger}(S^{(1)}_{BA} + f^{(1)}) - S^{(1)2}_{AA}]H^{(1)}_{AA} \\ & - f^{(1)}S^{(1)}_{AA}H^{(1)}_{AB} - S^{(1)}_{AB}f^{(1)\dagger}H^{(1)}_{BA} + f^{(1)}f^{(1)\dagger}H^{(1)}_{BB} \\ & + [-(S^{(3)}_{AA} + S^{(2)}_{AB}f^{(1)} + f^{(1)\dagger}(S^{(2)}_{BA} + S^{(1)}_{BB}f^{(1)})) \\ & + (S^{(2)}_{AA} + S^{(1)}_{AB}f^{(1)} + f^{(1)\dagger}(S^{(1)}_{BA} + f^{(1)}))S^{(1)}_{AA} \\ & + S^{(1)}_{AA}(S^{(2)}_{AA} + S^{(1)}_{AB}f^{(1)} + f^{(1)\dagger}(S^{(1)}_{BA} + f^{(1)})) - S^{(1)3}_{AA}]H^{(0)}_{AA} \\ & - f^{(1)}S^{(1)}_{AA}f^{(1)\dagger}H^{(0)}_{BB} \end{aligned}$$

TABLE 7.13 $E^{(n)}$ -- Non-orthonormal Basis

$$E^{(0)} = \nu \sum_{r=1}^{n_A} \epsilon_r^o$$

$$E^{(1)} = \nu \sum_{r=1}^{n_A} (H_{rr}^{(1)} - \epsilon_r^o S_{rr}^{(1)})$$

$$E^{(2)} = \nu \sum_{r=1}^{n_A} \left[H_{rr}^{(2)} - \sum_{t=1}^{n_A} S_{rt}^{(1)} H_{tr}^{(1)} \right.$$

$$\left. -\epsilon_r^o \left(S_{rr}^{(2)} + \sum_{t=1}^{n_A} S_{rt}^{(1)} S_{tr}^{(1)} + \sum_{\sigma=1}^{n_B} S_{r\sigma}^{(1)} \frac{(H_{\sigma r}^{(1)} - \epsilon_r^o S_{\sigma r}^{(1)})}{(\epsilon_r^o - \epsilon_\sigma^o)} \right) \right]$$

$$+ \nu \sum_{\sigma=1}^{n_B} \frac{(H_{\sigma r}^{(1)} - \epsilon_r^o S_{\sigma r}^{(1)})}{(\epsilon_r^o - \epsilon_\sigma^o)} H_{r\sigma}^{(1)}$$

7.3.b Extended Huckel Molecular Orbital Theory

The formulas just developed for use with a non-orthonormal basis will now be used to derive explicit expressions for the first order response of the elements of the density matrix R (in the atomic orbital basis) -- equal to the charge-bond order matrix divided by 2 -- under a perturbation of both the hamiltonian and overlap matrices. These formulas are analogous to those of section 7.2.b for ordinary Huckel theory, and would be applicable, for example, in an extended Huckel formalism. The increased complexity of the formulas due to the presence of the overlap matrix probably accounts in part for the lack of a detailed treatment of this problem in the literature, although a number of low order formulas have appeared in connection with particular applications (Fujimoto et al, 1974; Coope, 1956; Libit and Hoffmann, 1974).

For single center perturbations, we have

$$(H_{AO}^{(1)})_{pq} = \sum_{t=1}^{n} \delta\alpha_t \, \delta_{pt}\delta_{qt},$$

and

$$(S_{AO}^{(1)})_{pq} = \sum_{t=1}^{n} \delta S_{tt} \, \delta_{pt}\delta_{qt}, \qquad (7.19a)$$

so that

$$(H_{MO}^{(1)})_{ij} = \sum_{t=1}^{n} \delta\alpha_t \, C_{ti}C_{tj},$$

and

$$(S_{MO}^{(1)})_{ij} = \sum_{t=1}^{n} \delta S_{tt} C_{ti}C_{tj}. \qquad (7.19b)$$

The derivatives $\partial R_{ij}/\partial(H_{AO})_{pp}$ and $\partial R_{ij}/\partial(S_{AO})_{pp}$ are given by

the coefficients of $\delta\alpha_p$ and δS_{pp}, respectively, in the expression obtained for $R_{ij}^{(1)}$ by inserting the first order formulas in Tables 7.9 - 7.11 with eq. (7.19b), in eq. (7.17). In detail

$$R_{ij}^{(1)} = -\sum_{t=1}^{n} \delta\alpha_t \sum_{r}^{occ} \sum_{s}^{occ} C_{ir}C_{tr}^{*}C_{ts}C_{js}^{*}$$

$$+ \sum_{t=1}^{n} \sum_{r}^{occ} \sum_{\sigma}^{unocc} \frac{[\delta\alpha_t - \epsilon_r^o \delta S_{tt}]}{\epsilon_r^o - \epsilon_\sigma^o} C_{tr}C_{t\sigma}^{*} \tag{7.20}$$

$$\times [C_{ir}C_{\sigma j}^{*} + C_{i\sigma}C_{rj}^{*}],$$

where the ϵ_i^o are the eigenvalues of $H^{(0)}$, and it has been assumed that $H_{MO}^{(0)}$ is diagonal and $S_{MO}^{(0)} = 1$.

In the same way, for two-center perturbations, $(H_{AO}^{(1)})_{pq} = (H_{AO}^{(1)})_{qp} = \delta\beta_{pq}$, and $(S_{AO}^{(1)})_{pq} = (S_{AO}^{(1)})_{qp} = \delta S_{pq}$, for all p,q, $(p \neq q)$, which implies that

$$(H_{MO}^{(1)})_{ij} = \sum_{p,q=1}^{n} \delta\beta_{pq}[C_{pi}C_{qj} + C_{pj}C_{qi}],$$

and (7.21)

$$(S_{AO}^{(1)})_{ij} = \sum_{p,q=1}^{n} \delta S_{pq}[C_{pi}C_{qj} + C_{pj}C_{qi}],$$

one obtains

$$R_{ij}^{(1)} = -\sum_{p,q=1}^{n} \delta S_{pq} \sum_{r}^{occ} \sum_{s}^{occ} C_{ir}[C_{pr}C_{qs} + C_{ps}C_{qr}]C_{js}^{*}$$

$$+ \sum_{p,q=1}^{n} \sum_{r}^{occ} \sum_{\sigma}^{unocc} \frac{(\delta\beta_{pq} - \epsilon_r^o \delta S_{pq})}{\epsilon_r^o - \epsilon_\sigma^o}[C_{pr}C_{q\sigma}^{*} + C_{p\sigma}C_{qr}^{*}] \tag{7.22}$$

$$\times [C_{ir}C_{\sigma j}^{*} + C_{j\sigma}C_{rj}^{*}],$$

from which the first derivatives of the elements of R with respect to off-diagonal elements of H_{A0} and S_{A0} may be obtained.

Formulas for higher order terms in the series for R can be obtained here in the same way. However, they are long and tedious, and not very informative by themselves. Nevertheless, using formulas developed in section 6.4, it is possible to compute these higher order terms numerically for specific applications, in a relatively efficient manner.

7.4 Self-consistent Perturbation Theory

The object of this section is to develop a perturbation formalism for the one-particle density matrix in closed shell Hartree-Fock theory. The formulas developed here allow a more rigourous calculation of various properties of atoms and molecules than those given in previous sections of this chapter, because electron repulsion terms are included explicitly. The entire effect of the self-consistency terms is buried in the detailed calculation of the perturbation series for f, and therefore, formulas for the density matrix and related quantities in terms of f, which were derived for the simple matrix case, will apply here also.

7.4.a General Theory

In this case, the perturbation series for the operator f is obtained by expansion of the equation

$$D(f) = F_{BA} + F_{BB}f - f(F_{AA} + F_{AB}f) = 0, \qquad (7.23)$$

leading to a hierarchy of equations determining the $f^{(n)}$. This hierarchy is formally identical to that obtained in the simple matrix case, except that now, the matrix F (the closed shell Fock matrix) itself depends on f through its dependence on the density matrix, P_A',

$$F(P_A') = H + G(P_A'). \qquad (7.24)$$

Here H is the core hamiltonian, and the two-electron part,

$G(P_A')$, representing electron repulsion, is given by

$$G(P_A')_{rs} = 2 \sum_{t,u=1}^{m} (P_A')_{tu}([rs\|ut] - \tfrac{1}{2}[rt\|us]), \qquad (7.25a)$$

with

$$[rs\|ut] = \iint \phi_r^*(1)\phi_s(1)r_{12}^{-1}\phi_u^*(2)\phi_t(2)d\tau_1 d\tau_2 , \qquad (7.25b)$$

the ϕ_r being elements of the zero order molecular orbital basis used for the calculation. Direct iterative solution of eq. (7.23) for f, without making a perturbation expansion, is equivalent to solving the Hartree-Fock equations exactly.

We will consider here only those perturbations which can be introduced as perturbations of the core hamiltonian,

$$H = \sum_{n=0}^{\infty} H^{(n)} . \qquad (7.26)$$

This perturbation will induce changes in the electron distribution described by P_A', and through that, the two-electron part of the Fock matrix is perturbed. Thus, the n^{th} order term in the perturbation series for the Fock matrix consists not only of the n^{th} order term, $H^{(n)}$, in (7.26), but also includes an n^{th} order two-electron term. The exact form of this n^{th} order two-electron term depends on the manner in which eq. (7.23) is expanded into a hierarchy of equations determining the terms of the series for f, as is explained below.

It is convenient, but not strictly necessary, to require that $F^{(0)}(f)$ be at least block diagonal, so that f itself is at least a first order quantity, $f = \sum_{n=1}^{\infty} f^{(n)}$. In some applications, it may be desirable to relax this requirement,

and the modifications which must be made in such a case to the formalism given below are indicated in Appendix 10. However, in this section, and sections 7.4.b and 7.4.c following, it will be assumed that we are working in a basis in which $F^{(0)}$ is at least block diagonal.

Formal substitution of the series for f and F into (7.23) gives

$$\begin{aligned} D^{(n)} &= F_{BA}^{(n)} + \sum_{j=1}^{n} \left(F_{BB}^{(n-j)} f^{(j)} - f^{(j)} F_{AA}^{(n-j)}\right) \\ &\quad + \sum_{i=1}^{n-2} \sum_{j=1}^{n-i-1} f^{(j)} F_{AB}^{(n-i-j)} f^{(j)} \\ &= F_{BA}^{(n)} + F_{BB}^{(0)} f^{(n)} - f^{(n)} F_{AA}^{(0)} + A_{BA}^{(n)} = 0, \\ &\qquad n = 0, 1, 2, \ldots . \end{aligned} \tag{7.27}$$

Here, one has

$$\begin{aligned} A_{BA}^{(n)} &= \sum_{j=1}^{n-1} \left(F_{BB}^{(n-j)} f^{(j)} - f^{(j)} F_{AA}^{(n-j)}\right) \\ &\quad + \sum_{i=1}^{n-2} \sum_{j=1}^{n-i-1} f^{(i)} F_{AB}^{(n-i-j)} f^{(j)}, \end{aligned} \tag{7.28}$$

which does not depend explicitly or implicitly (through $F^{(n)}$) on $f^{(n)}$. The quantity (7.28) is not the same as the analogous quantity in eq. (6.9) in the simple matrix case, because $F_{BA}^{(n)}$ may now depend on $f^{(n)}$, and therefore, for the purposes of solving (7.27), it must appear explicitly. The extension of these equations to a non-orthonormal basis will lead to equations similar to (6.54) and (6.55) in place of (7.27) and (7.28).

Despite the similarity of the basic equations, the determination of the perturbation series for f is more complicated here than in the simple matrix case when the term $F_{BA}^{(n)}$ in (7.27) is considered to depend on $f^{(n)}$, that is, when the n^{th} order Fock matrix is considered to be

$$F^{(n)} = H^{(n)} + G(P_A'^{(n)}). \qquad (7.29)$$

When (7.29) is incorporated into (7.27), the so-called "coupled Hartree-Fock" perturbation scheme results, and the equations for the $f^{(n)}$ in this case are derived in the following subsection. Formalisms in which the dependence of $F^{(n)}$ on $P_A'^{(n)}$ is partially or completely neglected, leading to schemes referred to as "uncoupled Hartree-Fock" perturbation theory, are discussed in section 7.4.c.

7.4.b Coupled Hartree-Fock Perturbation Theory

In the coupled Hartree-Fock perturbation scheme, $F^{(n)}$ is considered to be dependent on $f^{(n)}$ as indicated in eq. (7.29). That is, the two-electron integrals are considered to be order neutral. It is convenient to write the n^{th} order density matrix, $P_A'^{(n)}$, in the form

$$P_A'^{(n)} = \tilde{P}_A'^{(n)} + \begin{bmatrix} 0 & f^{(n)\dagger} \\ f^{(n)} & 0 \end{bmatrix} = \tilde{P}_A'^{(n)} + f_1^{(n)}, \qquad (7.30)$$

where $\tilde{P}_A'^{(n)}$ depends on $f^{(j)}$ for $j \leq n-1$ only. Thus, eqs. (7.27)

become

$$D^{(n)}(f) = G_{BA}(f^{(n)}) + F^{(0)}_{BB} f^{(n)} - f^{(n)} F^{(0)}_{AA} + F^{(n)}_{BA}(\tilde{P}'^{(n)}_{A}) + A^{(n)}_{BA} = 0, \quad n = 0, 1, 2, \ldots , \tag{7.31}$$

where

$$F^{(n)}_{BA}(\tilde{P}'^{(n)}_{A}) = H^{(n)}_{BA} + G_{BA}(\tilde{P}'^{(n)}_{A}). \tag{7.32}$$

The last two terms of (7.31) are now independent of $f^{(n)}$. The important feature of eqs. (7.31), however, is that the first three terms -- those which are dependent on $f^{(n)}$ -- are of a particularly simple form, which is the same for all values of n.

From eq. (7.25a),

$$G_{BA}(f_1^{(n)})_{\tau s} = \sum_{\sigma}^{unocc} \sum_{r}^{occ} f^{(n)}_{\sigma r} \, 2[\tau' s \| r\sigma'] - [\tau' \sigma' \| rs] + \sum_{\sigma}^{unocc} \sum_{r}^{occ} f^{(n)*}_{\sigma r} \, 2[\tau' s \| \sigma' r] - [\tau' r \| \sigma' s] . \tag{7.33}$$

If all quantities are real, this reduces to

$$G_{BA}(f_1^{(n)})_{\tau s} = \sum_{\sigma}^{unocc} \sum_{r}^{occ} f^{(n)}_{\sigma r} \, 4[\tau' s \| \sigma' r] - [\tau' \sigma' \| rs] - [\tau' r \| \sigma' s]$$

$$= \sum_{\sigma}^{unocc} \sum_{r}^{occ} f^{(n)}_{\sigma r} A_{\tau' s r \sigma'} . \tag{7.34}$$

The four index quantity, $A_{\tau' s r \sigma'}$, has sometimes been referred to as the Nesbet supermatrix. Thus, if the zero order Fock matrix is block diagonal, the n^{th} order equation (7.31) can

be written

$$D^{(n)}_{\tau s} = \sum_{r}^{occ} \sum_{\sigma}^{unocc} B_{\tau s r \sigma} f^{(n)}_{\sigma r} + [F^{(n)}_{BA}(\tilde{P}'^{(n)}_{A}) + A^{(n)}_{BA}]_{\tau s} = 0,$$

$$(\tau = 1, \ldots, n_B, \; s = 1, \ldots, n_A), \tag{7.35a}$$

where

$$B_{\tau s r \sigma} = (F^{(0)}_{BB})_{\tau\sigma}\delta_{rs} - (F^{(0)}_{AA})_{rs}\delta_{\tau\sigma} + A_{\tau s r \sigma}. \tag{7.35b}$$

When the zero order Fock matrix is diagonal, eq. (7.35a) becomes

$$D^{(n)}_{\tau s} = (\epsilon^{o}_{\tau} - \epsilon^{o}_{s}) f^{(n)}_{\tau s} + \sum_{\sigma}^{unocc} \sum_{r}^{occ} A_{\tau s r \sigma} f^{(n)}_{\sigma r} + [F^{(n)}_{BA}(\tilde{P}'^{(n)}_{A}) + A^{(n)}_{BA}]_{\tau s} = 0, \tag{7.36}$$

$$(\tau = 1, \ldots, n_B; \; s = 1, \ldots, n_A),$$

where the ϵ^{o}_{i} are the eigenvalues of the zero order Fock matrix. In either case, the calculation of $f^{(n)}$ reduces to the solution of a system of $n_A n_B$ simultaneous linear equations. Even when the zero order Fock matrix is diagonal, it is no longer possible to obtain a closed formula for the elements of $f^{(n)}$, because of the self-consistency term. However, only the terms $F^{(n)}_{BA}(\tilde{P}'^{(n)}_{A})$ and $A^{(n)}_{BA}$ need be calculated for each value of n, since the coefficients of the $f^{(n)}_{\sigma r}$ in $D^{(n)}(f)$ are the same for every value of n. The calculation of these two quantities is easily done automatically, and therefore, the formalism above can be used to calculate high order perturbation series for P'_A without having to derive and use explicit perturbation equations.

Because of the potentially large dimension of the matrix B in eq. (7.35a), non-iterative methods of solution (such as Gaussian elimination) may not be practical in application to that linear system, especially if B is a sparse matrix. Perhaps the simplest iterative technique is the Gauss-Seidel procedure, with the iteration formula

$$f_{\tau r}^{(n)} = \frac{[F_{BA}^{(n)}(\tilde{P}_A'^{(n)}) + A_{BA}^{(n)}]_{\tau s} - \sum_{\sigma \neq \tau}^{\text{unocc}} \sum_{r \neq s}^{\text{occ}} B_{\tau s r \sigma} f_{\sigma r}^{(n)}}{B_{\tau s s \tau}} . \quad (7.37)$$

This procedure has been found to be satisfactory in the small number of calculations we have done, although no data has been obtained on actual rates of convergence. Many other efficient techniques are available for the iterative solution of large linear systems. We shall not explore this aspect further here, however.

7.4.c Uncoupled Hartree-Fock Perturbation Theory

The term "uncoupled Hartree-Fock perturbation theory" has been applied to a number of related approximations, proposed over a period of years, in order to simplify the solution of (7.27) (for example, Langhoff, Karplus, and Hurst, 1965; Musher, 1967), usually only in first order. The complicated coupling term in eqs. (7.35a) or (7.36) arises directly from the requirement that self-consistency be maintained in all orders in the perturbation. However, in a situation in which the perturbation is expected to distort the electronic distri-

bution only slightly, it may be possible to obtain acceptable results by relaxing this self-consistency requirement somewhat.

In first order, this amounts to ignoring the dependence of $F_{BA}^{(1)}$ on $f^{(1)}$, leading to the same result as in the simple matrix case,

$$f_{\sigma r}^{(1)\mathrm{unc}} = \frac{H_{\sigma r}^{(1)}}{\epsilon_r^o - \epsilon_\sigma^o} \quad , \tag{7.38}$$

when $F^{(0)}$ is diagonal. A degree of ambiguity enters if this formalism is to be extended to higher order, however. It is not clear whether one should ignore just the $f^{(n)}$ dependent part of $P_A^{\prime(n)}$, or all of $P_A^{\prime(n)}$ in the n^{th} order equation, (7.27). There may be an accumulation of non-self-consistency as one proceeds to higher orders, depending on the exact form of the approximations employed, and this may cast doubt on the validity of these higher order terms.

An internally consistent and unambiguous perturbation formalism does result if the two-electron integrals are considered to be first order quantities except where they enter implicitly in $F^{(0)}$. Then

$$F^{(n)} = H^{(n)} + G(P_A^{\prime(n-1)}), \tag{7.39}$$

and no self-consistency term in $f^{(n)}$ will occur in the n^{th} order equation of (7.27). In fact, except for the implicit dependence of the $F^{(j)}$ in $A_{BA}^{(n)}$ on lower order $f^{(j)}$, the resulting hierarchy of equations determining the $f^{(n)}$ will be identical to that in the simple matrix case. Only by actual calculations can the validity of the assumption (7.39) be assessed, however.

CHAPTER 8

DIRECT MINIMIZATION SELF-CONSISTENT FIELD THEORY

"'In that case', said the Dodo solemnly, rising to its feet, 'I move that the meeting adjourn, for the immediate adoption of more energetic remedies--'"

"'Would you tell me, please, which way I ought to go from here?'
'That depends a good deal on where you want to get to', said the Cat.
'I don't much care where--', said Alice.
'Then it doesn't matter which way you go', said the Cat.
'--so long as I get somewhere', Alice added as an explanation.
'Oh, you're sure to do that', said the Cat, 'if you only walk long enough'.
Alice felt that this could not be denied, s she tried another question. 'What sort of people live about here?'
'In that direction,' said the Cat, waving its right paw around, 'lives a Hatter: and in that direction', waving the other paw, 'lives a March Hare. Visit either you lik they're both mad.'" (Alice's Adventures in Wonderland, Lewis Carroll)

8.1 Introduction

In the Hartree-Fock approximation, the total electronic energy of an atomic or molecular system described by a single determinant wavefunction, Ψ, can be written in a given basis as

$$E = \sum_{i=1}^{m} \nu_i \sum_{r,s=1}^{n} R^{(i)}_{sr} h_{rs} + \tfrac{1}{2} \sum_{i,j=1}^{m} \nu_i \nu_j \sum_{\substack{r,s \\ t,u=1}}^{n} R^{(i)}_{sr} R^{(j)}_{tu} \times ([rs||ut] - a_{ij}[rt||us])$$

$$= \sum_{i=1}^{m} \nu_i \sum_{r,s=1}^{n} \sum_{\mu}^{occ} X^{(i)}_{s\mu} X^{(i)*}_{r\mu} h_{rs} \qquad (8.1)$$

$$+ \tfrac{1}{2} \sum_{i,j=1}^{m} \nu_i \nu_j \sum_{\substack{r,s \\ t,u=1}}^{n} \sum_{\alpha,\beta=1}^{occ} X^{(i)}_{s\alpha} X^{(i)*}_{r\alpha} X^{(j)}_{t\beta} X^{(j)*}_{u\beta} \times ([rs||ut] - a_{ij}[rt||us]).$$

Here h is the core hamiltonian for the system, and the $[rs||ut]$ are two-electron integrals defined in eq. (7.25b). The summation indices, i,j, refer to electronic shells. The $X^{(i)}_{\mu\nu}$ are expansion (lcao) coefficients, expressing the occupied normalized orbitals as linear combinations of the given basis functions. The ν_i are occupation numbers for these orbitals, and the a_{ij} are constants determined according to the values of ν_i and ν_j. The operator $R^{(i)}$ (the one-particle density matrix for the i^{th} shell) is a projection onto the space of the i^{th} shell occupied orbitals,

$$R^{(i)} = X^{(i)} X^{(i)\dagger}. \qquad (8.2)$$

The electronic energies of the stationary states of the system are approximated by the stationary values of E, eq. (8.1), considered as a function of a suitable set of variables, such as the $X_{\mu\nu}^{(i)}$ or $R^{(i)}$. The traditional and still, at present, the most commonly used procedure for determining the stationary values of E has been by solving the corresponding Hartree-Fock equations (Roothaan, 1951),

$$F^{(i)}X^{(i)} = S\,X^{(i)}\,\varepsilon^{(i)}, \qquad (i = 1, \ldots, m). \qquad (8.3)$$

The matrices $F^{(i)}$ depend on all of the occupied orbitals $X^{(j)}$, $(j = 1, \ldots, m)$. An initial estimate of the $X^{(i)}$ is used to construct approximations to the $F^{(i)}$, which are then diagonalized to yield, it is hoped, an improved estimate of the $X^{(i)}$, which can be used to obtain a further improved approximation for the $F^{(i)}$. This iterative procedure is continued until self-consistency is achieved. It is conceptually very simple, and in applications to the simplest (single shell) systems, rates of convergence relative to the work required in each iteration are quite good. Difficulties in obtaining convergence do arise, however, especially in calculations involving more complicated multi-shell systems.

An alternative to the use of the Hartree-Fock equations is to minimize the energy, E, directly with respect to a chosen set of variables. One problem in using the elements of the density matrices, $R^{(i)}$, or the lcao coefficients, $X_{\mu\nu}^{(i)}$, for this, is that a relatively large number of constraints must be imposed if the simple functional form, (8.1), of the energy

is to be preserved.

When using the lcao coefficients, the presence of redundant variables also causes difficulties for procedures, such as the Newton-Raphson method (Appendix 11), which require that the matrix of second derivatives of E (the Hessian matrix) be non-singular near stationary points. Redundancy among the expansion coefficients is associated with the invariance (to within a complex scalar factor) of the determinantal wavefunction to non-singular linear transformations of occupied orbitals in the same shell. Under orthonormality constraints, the redundancy associated with unitary transformations still remains. The density matrices contain no redundancy, but must satisfy more complicated constraints. The presence of q redundant variables implies the existence of a q-dimensional constant energy surface through each point in the coordinate space of the unconstrained variables. A serious consequence for some gradient minimization techniques is that the Hessian matrix is then singular at stationary points of the energy (Sutcliffe, 1974, 1975; Coope, unpubl.).

An efficient technique for eliminating the orthogonality constraints on the lcao coefficients for closed shell systems has been developed by Fletcher (1970), and extended by Kari and Sutcliffe (1970, 1973) to more general multi-shell and multi-determinant cases. However, calculations in which Fletcher's method is used in conjunction with the conjugate gradient minimization technique, are frequently poorly

convergent near the energy minimum. As a result, such direct energy minimization procedures have been used more to provide improved starting estimates for the solution of the Hartree-Fock equations than as an alternative to the Hartree-Fock equations (see, for example, Claxton and Smith, 1971).

Sutcliffe (1974, 1975) has explicitly exhibited the singularity of the Hessian matrix at the energy minimum in formalisms based on Fletcher's method, and he has suggested that this singularity may contribute to the slowness of convergence near the energy minimum. This suggestion is questioned below, both on theoretical grounds, and by examination of rates of convergence for calculations involving minimization of the energy with respect to a set of unconstrained variables containing no redundancies. It is our contention that the observed poor convergence rates arise rather out of deficiencies in the straightforward implementation of the conjugate gradient minimization algorithm.

Sutcliffe (1974, 1975) has proposed several solutions to the redundancy problem, but clearly, the simplest would be to write the total electronic energy, from the beginning, in terms of a set of unconstrained variables not possessing such redundancies. The eigenvalue independent partitioning formalism developed in chapters 2 and 4 provides such sets of variables, namely the matrix elements of the off-diagonal blocks of the matrix $\hat{T}$. In the following sections, the application of the partitioning formulas to the minimization of the energy of a

system represented by a single determinant wavefunction is described. One of the major advantages here is the fact that the derivatives of E with respect to these variables can be expressed very simply in terms of the columns of the projections, (8.2), onto the occupied orbitals, and their complements. A scaled descent method, based on partitioning with respect to current occupied and unoccupied molecular orbitals, is proposed, (section 8.3.c), which appears to be very successful in practice.

8.2 Closed Shell Systems

8.2.a Orthonormal Basis

The square matrix, X, of the eigenvectors of the closed shell Fock matrix is partitioned into the n_A occupied and the n_B unoccupied molecular orbitals. The orthonormal basis functions, in terms of which these orbitals are expressed, are partitioned into two sets of the same dimensions, n_A and n_B, defining spaces S_A and S_B. In this way, the coefficient matrix X can be written in the blocked form (2.2). The projection, R, onto the space of occupied orbitals, is given by eq. (2.10) as

$$R = \begin{bmatrix} g_A^{-1} & g_A^{-1}f^\dagger \\ fg_A^{-1} & fg_A^{-1}f^\dagger \end{bmatrix}, \qquad (8.4)$$

where $g_A = 1_A + f^\dagger f$ is the metric for the eigenvectors $X^{(A)}$, truncated to the space S_A.

In the closed shell case, the energy functional is particularly simple,

$$E = 2 \text{ tr } Rh + \text{tr } RG(R)$$

$$= 2 \sum_{r,s}^{n} R_{sr} h_{rs} + \sum_{t,u}^{n} R_{tu}([rs\|ut] - \tfrac{1}{2}[rt\|us]). \qquad (8.5)$$

Substitution of (8.4) into (8.5) gives the energy in terms of the matrix elements of f only. Since the degrees of freedom available[1], $n_A n_B$, exactly equal the number of matrix elements

[1]The argument involving numbers of variables is of central importance here, and is as follows for the closed shell case. (cont'd)

of f, there can be no redundancy. Also, the matrix f is completely unconstrained because R, eq. (8.4), automatically satisfies the criteria necessary to be a projection (section 2.1.a). In short, the matrix elements of f represent a set of unconstrained variables possessing no redundancies, with respect to which the energy can be minimized.

In principle, the elements of the block R_{BA} also provide a set of non-redundant and unconstrained variables, but they are not very suitable for specifying the energy, because of the complicated relationship between R_{BA} and R_{AA} or R_{BB}.[2]

If no constraints are imposed on the occupied molecular orbitals, $X^{(A)}$, specified by $n_A n$ complex parameters, then these orbitals are arbitrary up to an $n_A \times n_A$ linear transformation. Therefore, there must be n_A^2 complex redundant variables among the lcao coefficients in the single determinant wavefunction, leaving $n_A n_B$ complex variables which are independent. If the molecular orbitals are constrained to be orthonormal, then n_A^2 real parameters are eliminated by the constraints, and n_A^2 of the real parameters remaining are redundant (equal to the independent parameters in a unitary transformation -- this includes the n_A arbitrary phase factors), again leaving $2n_A n_B$ real parameters or $n_A n_B$ complex parameters which are independent.

[2]For the density matrix, the requirement of idempotency leads to the $n_A^2 + n_B^2$ (complex) constraints: $R_{AA}^2 - R_{AA} + R_{AB}R_{BA} = 0$, and $R_{BB} = R_{BA}R_{AA}^{-1}R_{AB}$, which give the blocks R_{AA} and R_{BB} in terms of R_{BA}, specified by $n_A n_B$ complex parameters. Given R_{AA} and R_{BA}, it is easy to calculate R_{BB}, but the first equation here is not easily solved for R_{AA} (see section 2.1.c, and in particular, eq. (2.23)).

The first and second derivatives of E with respect to the elements of f are most easily obtained by the incremental approach used in section 2.1.e, but now retaining terms to second order in the variation. Writing $g_A^{-1}(f + \delta f) = g_A^{-1}(f) + \delta g_A^{-1}$, to second order one has,

$$\delta g_A^{-1} = -g_A^{-1}\delta g_A\, g_A^{-1} + g_A^{-1}\delta g_A\, g_A^{-1}\delta g_A\, g_A^{-1} + O(\delta^3), \quad (8.6)$$

where

$$\delta g_A = \delta f^\dagger f + f^\dagger \delta f + \delta f^\dagger \delta f. \quad (8.7)$$

For the density matrix, the variation $R(f + \delta f) = R(f) + \delta R$, is given, exactly, by

$$\begin{aligned}
\delta R_{AA} &= \delta g_A^{-1}, \\
\delta R_{AB} &= \delta g_A^{-1} f^\dagger + g_A^{-1}\delta f^\dagger + \delta g_A^{-1}\delta f^\dagger, \\
\delta R_{BA} &= f\delta g_A^{-1} + \delta f g_A^{-1} + \delta f \delta g_A^{-1}, \\
\delta R_{BB} &= f\delta g_A^{-1} f^\dagger + \delta f g_A^{-1} f^\dagger + f g_A^{-1}\delta f^\dagger + \delta f \delta g_A^{-1} f^\dagger + f\delta g_A^{-1}\delta f^\dagger \\
&\quad + \delta f g_A^{-1}\delta f^\dagger + \delta f \delta g_A^{-1}\delta f^\dagger.
\end{aligned} \quad (8.8)$$

The first order term in the expansion of $E(f + \delta f)$ can be simplified to give

$$\delta^{(1)}E = 2\ \mathrm{tr}\ \delta R\, F \quad (8.9a)$$

$$= 2\ \mathrm{tr}\ \delta f^\dagger \bar{D} + 2\ \mathrm{tr}\ \delta f \bar{D}^\dagger, \quad (8.9b)$$

where

$$\bar{D} = g_B^{-1} D(f) g_A^{-1}. \quad (8.9c)$$

This is identical to the result obtained for a simple matrix

(section 2.1.e), except that here, the quantity D(f), given by

$$D(f) = F_{BA} + F_{BB}f - fF_{AA} - fF_{AB}f, \tag{8.10}$$

is defined in terms of the Fock operator,

$$F = h + G(R), \tag{8.11}$$

which itself is a function of f. As before, $g_B = 1_B + ff^\dagger$, is the metric for the eigenvectors $X^{(B)}$ of F truncated to the space S_B. From (8.9b), the first derivatives of the energy are seen to be

$$\frac{\partial E}{\partial f^*_{\sigma r}} = 2\bar{D}_{\sigma r} \tag{8.12a}$$

$$= 2[(1 - R)FR]_{\sigma r} \tag{8.12b}$$

$$= 2\, F_{e_B^\sigma e_A^r}, \tag{8.12c}$$

using the notation developed in section 2.1.d. Here and below, Greek letters denote basis elements in S_B, and Roman letters denote basis elements in S_A. The first derivatives of the energy with respect to the variables $f_{\sigma r}$ are therefore given by elements of the off-diagonal blocks of the current Fock matrix between contragredient non-orthonormal vectors given by the first n_A columns of R and the last n_B columns of (1-R).

Because the metric matrices g_A and g_B are positive definite (as, therefore, are their inverses), it is seen that the first derivatives of the energy with respect to the elements of f can vanish only if D(f) = 0. In fact, this condition, or the more

general one, eq. (2.13),

$$F\hat{T} = \hat{T}\hat{F},$$

with $\hat{F}$ defined as in (2.15a), can be regarded as an expression for the Hartree-Fock equations in the present formalism. As indicated in section 2.2.a, the condition $D(f) = 0$ (and therefore, $\nabla_{f^\dagger} E = 0$) is also equivalent to $X^\dagger FX$ being block diagonal.

Isolation of the coefficients of the second order terms in $E(f + \delta f)$ yields the second derivatives of the energy with respect to the elements of f. After considerable algebraic manipulation, one obtains

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial f^*_{\sigma r}\partial f_{\tau s}} = F_{e_B^\sigma e_B^\tau} R_{rs} - (1 - R)_{\sigma\tau} F_{e_A^r e_A^s} - [e_B^\sigma e_B^\tau \| e_A^s e_A^r] + [e_B^\sigma e_A^r \| e_A^s e_B^\tau], \tag{8.13a}$$

and

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial f_{\sigma r}\partial f_{\tau s}} = -[e_B^\sigma e_A^s \| e_B^\tau e_A^r] + 2[e_B^\sigma e_A^r \| e_B^\tau e_A^s] - \frac{\partial E}{\partial f^*_{\tau s}} R_{\sigma s} + \frac{\partial E}{\partial f^*_{\sigma s}} R_{\tau r}. \tag{8.13b}$$

In the particular case that the partitioning chosen is defined by the current projection R, so that $f = 0$ and $R = 1_A$, and further, when particular bases adapted to R and $(1 - R)$ are chosen which diagonalize $F_{AA}(R)$ and $F_{BB}(R)$, respectively, then the dominant terms in eqs. (8.13) are the derivatives $\partial^2 E/\partial f^*_{\sigma r}\partial f_{\sigma r}$ with the value,

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial f^*_{\sigma r}\partial f_{\sigma r}} = \epsilon_\sigma - \epsilon_r - J_{r\sigma} + 2K_{r\sigma} = {}^1\Delta E_{r\to\sigma}, \qquad (8.14)$$

(the ϵ_i being the eigenvalues of F) equal to the singlet single excitation energies. At the energy minimum, the remaining second derivatives all reduce to combinations of two-electron integrals. To the extent that these combinations are small, the excitation energies, (8.14), approximate the eigenvalues of the Hessian matrix, which are thus positive, as they should be for an energy minimum.

In the case that all quantities are real, the above derivative formulas become (see Appendix 12),

$$\frac{\partial E}{\partial f_{\sigma r}} = 4F_{e_B^\sigma e_A^r}, \qquad (8.15)$$

and

$$\tfrac{1}{4}\frac{\partial^2 E}{\partial f_{\sigma r}\partial f_{\tau s}} = 4[e_A^r e_B^\sigma \| e_A^s e_B^\tau] - [e_A^r e_B^\tau \| e_B^\sigma e_A^s] - [e_B^\tau e_B^\sigma \| e_A^r e_A^s]$$
$$- \frac{\partial E}{\partial f_{\tau r}} R_{\sigma s} + \frac{\partial E}{\partial f_{\sigma s}} R_{\tau s} + F_{e_B^\sigma e_B^\tau} R_{rs} - (1-R)_{\sigma\tau} F_{e_A^r e_A^s}, \qquad (8.16a)$$

with

$$\tfrac{1}{4}\frac{\partial^2 E}{\partial f^2_{\sigma r}} = -[e_B^\sigma e_B^\sigma \| e_A^r e_A^r] + 3[e_B^\sigma e_A^r \| e_B^\sigma e_A^r] - 2\frac{\partial E}{\partial f_{\sigma r}} R_{r\sigma}$$
$$+ F_{e_B^\sigma e_B^\sigma} R_{rr} - (1-R)_{\sigma\sigma} F_{e_A^r e_A^r}. \qquad (8.16b)$$

8.2.b Non-orthonormal Basis

In this case, the density matrix, R, and the electronic energy, E, are still given by eqs. (8.4) and (8.5), respectively. However, now both E and R depend, through the metric g_A, on the overlap matrix.

According to eq. (2.103a), the metric g_A is given here by

$$g_A = S_{AA} + S_{AB}f + f^\dagger S_{BA} + f^\dagger S_{BB}f, \tag{8.17}$$

so that now eq. (8.7) must be replaced by

$$\begin{aligned}\delta g_A &= (S_{AB} + f^\dagger S_{BB})\delta f + \delta f^\dagger (S_{BA} + S_{BB}f) + \delta f^\dagger S_{BB}\delta f \\ &= Y_{AB}\delta f + \delta f^\dagger Y_{BA} + \delta f^\dagger S_{BB}\delta f,\end{aligned} \tag{8.18}$$

in the energy variation. The quantity $Y_{BA} = S_{BA} + S_{BB}f$ has been defined previously in section 5.3.c, and reduces to f for an orthonormal basis.

Isolation of the first order part of $E(f + \delta f)$ gives the first derivatives of the energy with respect to the elements of f as

$$\frac{\partial E}{\partial f^*_{\sigma r}} = 2F e^\sigma_B e^r_A , \tag{8.19}$$

which is identical to eq. (8.12c). The orbitals e^r_A, $(r=1,\ldots,n_A)$, are the same as before, but now the e^σ_B, $(\sigma=1,\ldots,n_B)$, are given as the columns of

$$e_B = \begin{bmatrix} -g_A^{-1}Y_{AB} \\ 1_B - fg_A^{-1}Y_{AB} \end{bmatrix} = [1 - RS]^{(B)}. \tag{8.20}$$

That is, the first derivatives of the energy with respect to the elements of f are given as matrix elements of the current Fock matrix between two sets of contragredient non-orthonormal vectors consisting, respectively, of the first n_A columns of the density matrix, R, and the last n_B columns of the complementary matrix (1 - RS). [3]

As before, the second derivatives are obtained by isolating the coefficients of second order terms in E(f + δf). The calculation of these coefficients is considerably more lengthy and tedious than for an orthonormal basis, but the final results are given simply by

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial f^*_{\sigma r}\partial f^*_{\tau s}} = 2[e^{\sigma}_{B}e^{r}_{A}\|e^{\tau}_{B}e^{s}_{A}] - [e^{\sigma}_{B}e^{s}_{A}\|e^{\tau}_{B}e^{r}_{A}] - (SR)_{\sigma s}\frac{\partial E}{\partial f^*_{\tau r}} + (SR)_{\tau r}\frac{\partial E}{\partial f^*_{\sigma s}} \tag{8.21a}$$

and

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial f^*_{\sigma r}\partial f^*_{\tau s}} = 2[e^{\sigma}_{B}e^{r}_{A}\|e^{s}_{A}e^{\tau}_{B}] - [e^{\sigma}_{B}e^{\tau}_{B}\|e^{s}_{A}e^{r}_{A}] + R_{sr}F_{e^{\sigma}_{B}e^{\tau}_{B}} - (1 - RS)_{\sigma\tau}F_{e^{s}_{A}e^{r}_{A}} . \tag{8.21b}$$

These formulas are identical in form to those obtained in an orthonormal basis, eqs. (8.13), except for the factors R and (1-R) being replaced by SR and (1-RS), respectively, in certain places.

---

[3] This is not the complement of R in the usual sense of the word. Since $(RS)^2 = RS$, one has (1-RS)R = 0, however, the reverse product R(1-RS) = R - R^2S is not zero in general.

An analysis of the metric properties of the e_A^r and e_B^σ similar to that of section 2.1.d for an orthonormal basis can be carried out. The algebra is tedious and only the major results are listed here.

Writing these vectors as columns of a matrix $\tilde{e}$,

$$\tilde{e} = [\ \tilde{e}_A \quad \tilde{e}_B\ ] = \begin{bmatrix} g_A^{-1} & -g_A^{-1}Y_{AB} \\ fg_A^{-1} & 1_B - fg_A^{-1}Y_{AB} \end{bmatrix}, \tag{8.22}$$

one can show that

$$\tilde{e}^\dagger\tilde{e} = \tilde{g} = \begin{bmatrix} g_A^{-1}(1_A+f^\dagger f)g_A^{-1} & g_A^{-1}[f^\dagger-(1_A+f^\dagger f)g_A^{-1}Y_{AB}] \\ [f-Y_{BA}g_A^{-1}(1_A+f^\dagger f)]g_A^{-1} & 1_B+Y_{BA}g_A^{-1}(1_A+f^\dagger f)g_A^{-1}Y_{AB} \\ & -fg_A^{-1}Y_{AB}-Y_{BA}g_A^{-1}f^\dagger \end{bmatrix} \tag{8.23}$$

verifying the non-orthonormality of the columns of $\tilde{e}$. A set of vectors dual to the e^i (that is, such that $\tilde{e}^\dagger\underset{\sim}{e} = 1 = \underset{\sim}{e}^\dagger\tilde{e}$) are given by

$$\underset{\sim}{e} = \begin{bmatrix} \hat{S}_A & -f^\dagger \\ Y_{BA} & 1_B \end{bmatrix}, \tag{8.24}$$

where $\hat{S}_A = S_{AA} + S_{AB}f$. These vectors are also non-orthonormal, as is seen from

$$\underset{\sim}{e}^\dagger\underset{\sim}{e} = \underset{\sim}{g} = \begin{bmatrix} \hat{S}_A^\dagger\hat{S}_A + Y_{AB}Y_{BA} & Y_{AB} - \hat{S}_A^\dagger f^\dagger \\ -f\hat{S}_A + Y_{BA} & 1_B + ff^\dagger \end{bmatrix}, \tag{8.25}$$

It is seen from eq. (2.33) that the last n_B of the e_i are

formally the same here as in an orthonormal basis.

Metric matrices, with respect to which the e^i and e_i are orthonormal, can be constructed explicitly. One obtains,

$$\tilde{\Delta} = \tilde{e}\tilde{e}^\dagger = \begin{bmatrix} \hat{S}_A\hat{S}_A^\dagger + f^\dagger f & \hat{S}_A Y_{AB} - f^\dagger \\ Y_{BA}\hat{S}_A - f & 1_B + Y_{BA}Y_{AB} \end{bmatrix}, \qquad (8.26)$$

and

$$\underset{\sim}{\Delta} = \underset{\sim}{e}\underset{\sim}{e}^\dagger = \begin{bmatrix} -g_A^{-1}(1_A + Y_{AB}Y_{BA})g_A^{-1} & \\ -Y_{BA}g_A^{-1} + fg_A^{-1}(1_A + Y_{AB}Y_{BA})g_A^{-1} & \end{bmatrix}$$

$$\begin{bmatrix} -g_A^{-1}Y_{AB} + g_A^{-1}(1_A + Y_{AB}Y_{BA})g_A^{-1}f^\dagger \\ 1_B + fg_A^{-1}(1_A + Y_{AB}Y_{BA})g_A^{-1}f^\dagger - Y_{BA}g_A^{-1}f - fg_A^{-1}Y_{AB} \end{bmatrix}, \qquad (8.27)$$

for which it is easy to verify that

$$\underset{\sim}{e}^\dagger\tilde{\Delta}\,\underset{\sim}{e} = 1 , \qquad \tilde{e}^\dagger\underset{\sim}{\Delta}\,\tilde{e} = 1 .$$

Not only are these results more complicated than for an orthonormal basis, but now $\tilde{g} \neq \tilde{\Delta}$ and $\underset{\sim}{g} \neq \underset{\sim}{\Delta}$, in contrast to the previous case. The matrices $\underset{\sim}{e}$ and $\tilde{e}$ are no longer normal.

8.2.c Results of Test Calculations -- Closed Shell Case

A set of CNDO/2 calculations were carried out to obtain information on the convergence properties of direct energy minimization procedures based on the formalism presented in sections 8.2.a and 8.2.b. The calculations were carried out

on an IBM 370/168 computer using double precision arithmetic.[4] In all calculations, the convergence criteria imposed were $|\delta E| < 10^{-12}$ a.u. and $|\delta R_{ii}| < 10^{-6}$ per iteration. The number of iterations required to satisfy both these criteria are given in Table 8.1 for selected calculations. In practice, a single iteration in addition to those indicated in the table is required in each case to verify that the convergence criteria have been satisfied.

The seven molecules chosen are ones for which the Roothaan iteration method can be used with varying degrees of success. Four of them, CH_4, HF, LiF, and H_2O, present no problems at all. For two of them, BeO and BN, Roothaan's method is only slowly convergent, and the last one, PN, leads to oscillations between definite charge distributions after about thirty Roothaan iterations. For each of these last three difficult cases, convergence of Roothaan's method will occur or can be accelerated if a suitable inter-iteration density matrix averaging procedure is employed.

The variables $f_{\sigma r}$ were defined by a partitioning between 'occupied' bond and lone pair orbitals, and 'unoccupied' antibond and atomic orbitals. The bond orbitals were non-polar combinations of hybrid atomic orbitals, the hybrid AOs used being far from optimal in some cases (for example, sp^3 hybrids

[4] The parts of the programs involved in calculating the CNDO/2 integrals and core hamiltonian were adapted from the CNDO/2 program of Pople and Beveridge (1970).

on the F atom). In this bond/antibond/lone pair orbital basis, the starting approximation was f = 0. For calculations done directly in the atomic orbital basis, the starting value of f was calculated using eq. (2.3a), where X defines the starting orbitals in the AO basis.

It is seen that in all but a small number of calculations, substantially fewer iterations were required to satisfy the convergence criteria when using the direct minimization methods, than when using Roothaan's method. Even in the cases causing difficulty for the Roothaan method, convergence appears striaghtforward for the direct methods. When variables, f_{or}, defined by an arbitrary partitioning of the AO basis are used, the number of iterations required increases somewhat. Rates of convergence for Fletcher's method and the partitioning method are generally comparable, indicated that the presence of redundant variables in the former has no observable effect on convergence rates. Generally, it was found that the overall rate of convergence depends very little on the accuracy of the step length as long as some minimal accuracy is maintained.

Assuming that the construction of the Fock matrix is by far the most costly single step in an SCF calculation, direct energy minimization procedures based on the conjugate gradient algorithm are at least twice as costly per iteration as the Roothaan method. Therefore, even a rather substantial decrease in the number of iterations required for a direct method may not represent a more efficient overall calculation. However,

the direct methods do have an advantage of reliability -- they can never diverge, if set up appropriately.

With the partitioning defined in the bond/antibond/lone pair basis, the full Newton-Raphson equations converge very rapidly -- in none of the seven examples studied are more than five iterations required to satisfy the stringent convergence criteria. In the case of CH_4, this rate of convergence can be duplicated using the conjugate gradient technique if the step lengths during the linear search are calculated sufficiently accurately (correct to four figures), but not for H_2O. If an arbitrary partitioning is defined in the atomic orbital basis, initial convergence of the Newton-Raphson method is generally very much poorer. For two of the molecules, the calculation actually diverges, while for a third, it converges to a stationary point above the minimum value of the energy.

Because of the expense involved in using the full Newton-Raphson equations, both a diagonal block and diagonal approximation were tested, these being analogous, respectively to algorithm FGN, and to algorithms DGN and SDNR, as described in chapter 5. While these approximations represent a very significant reduction in computation required, the methods are seen to be generally unreliable. Convergence is not only much poorer, but some calculations actually diverge in cases where Roothaan's method converges. The Newton-Raphson equations can, nevertheless, be usefully exploited in other ways, one of which is described and illustrated in section 8.3.c.

TABLE 8.1 Closed Shell Case -- Test Calculations[a]

Method[c]		Molecule						
		CH_4	HF	LiF	H_2O	BeO	BN	PN
Roothaan		10	16	17	20	49	>80[b]	osc.
Fletcher	d_1	7	5	14	7	18	19	16
	d_3	2	4	14	4	18	19	16
	c	7	5	14	9	18	18	16
Partitioning								
bond orbital basis	d_1	7	5	15	7	19	19	18
	d_3	2	4	14	5	19	20	15
	c	7	5	14	9	19	19	16
atomic orbital basis	d_1	8	7	20	10	18	19	44
	d_3	4	5	18	10	18	20	43
	c	5	5	18	10	18	20	43
Partitioning (steepest descents)		5						>40
Newton-Raphson								
Full (B/A basis)		2	3	4	2	5	5	5
Full (AO basis)		3	4	13[d]	4	div.	5	div.
Block Diagonal (B/A basis)		4	3	27	7	div.	div.	div.
Diagonal (B/A basis)		17	11	div.	24	div.	div.	div.

[a] Number of iterations required to satisfy $|\delta E| < 10^{-12}$, $|\delta R_{ii}| < 10^{-6}$ per iteration.

[b] convergent

[c] interpolation schemes: d_i -- secant formula, i times; c -- cubic formula

[d] converged to an excited state.

8.3 Unrestricted Hartree-Fock Theory

8.3.a Energy Derivatives

The formalism developed in the previous section for the closed shell case can be carried over with minor modifications to unrestricted Hartree-Fock calculations. In fact, it is possible, in some sense, to view the resulting formalism as that of two coupled closed shell systems, one for the α-spin electrons, and one for the β-spin electrons.

The energy functional is now

$$E = \tfrac{1}{2}\sum_{r,s} R^{\alpha}_{sr}(2h_{rs} + \sum_{t,u} R^{\alpha}_{tu}([rs\|ut]-[rt\|us]) + \sum_{t,u} R^{\beta}_{tu}[rs\|ut])$$

$$+\tfrac{1}{2}\sum_{r,s} R^{\beta}_{sr}(2h_{rs} + \sum_{t,u} R^{\beta}_{tu}([rs\|ut]-[rt\|us]) + \sum_{t,u} R^{\alpha}_{tu}[rs\|ut])$$

$$= \tfrac{1}{2}\sum_{r,s} R^{\alpha}_{sr}(2h_{rs}+G^{\alpha}_{rs}) + \tfrac{1}{2}\sum_{r,s} R^{\beta}_{sr}(2h_{rs}+G^{\beta}_{rs}). \qquad (8.28)$$

The matrices R^{α} and R^{β} are the one-particle density matrices referring to the α-spin and β-spin occupied orbitals, respectively.

A set of unconstrained, non-redundant variables completely specifying E can be introduced as follows. In the chosen basis, the n^{α}_{A} occupied α-spin orbitals are written as columns of a matrix $X^{\alpha(A)}$, and similarly, the n^{β}_{A} occupied β-spin orbitals as $X^{\beta(A)}$. These orbitals will be eigenvectors of the appropriate Fock operators. Now, two different partitionings of the basis set are carried out. In the first case, the basis functions are partitioned into two sets of dimensions n^{α}_{A} and n^{α}_{B} spanning

spaces S_A^α and S_B^α. A second partitioning is defined in which the basis functions are partitioned into two sets of dimensions n_A^β and n_B^β spanning spaces S_A^β and S_B^β, respectively. As a result, the occupied α-spin and β-spin orbitals can be written in the block form

$$X^{\alpha(A)} = \begin{bmatrix} X_{AA}^{\alpha} \\ X_{BA}^{\alpha} \end{bmatrix}, \qquad X^{\beta(A)} = \begin{bmatrix} X_{AA}^{\beta} \\ X_{BA}^{\beta} \end{bmatrix}. \tag{8.29}$$

It is now possible to define two f-operators, namely,

$$f^{\alpha} = X_{BA}^{\alpha} X_{AA}^{\alpha-1}, \qquad f^{\beta} = X_{BA}^{\beta} X_{AA}^{\beta-1}, \tag{8.30}$$

in terms of the two sets of occupied orbitals. Then one has

$$R^i = X^{i(A)} X^{i(A)\dagger} = \begin{bmatrix} g_A^{i-1} & g_A^{i-1} f^{i\dagger} \\ f^i g_A^{i-1} & f^i g_A^{i-1} f^{i\dagger} \end{bmatrix}, \tag{8.31}$$

with

$$g_A^i = 1_A + f^{i\dagger} f^i, \tag{8.32}$$

giving the two density matrices, and thus the electronic energy solely in terms of the $n_A^\alpha n_B^\alpha + n_A^\beta n_B^\beta$ elements of f^α and f^β.

That the elements of f^α and f^β are the minimum number of variables necessary to specify the energy, but not subject to any constraints nor possessing any redundancies, can be established in the same way as for the closed shell case. The requirement that the α-spin occupied orbitals be orthonormal, and the redundancy associated with the invariance of the energy, (8.28), to an $n_A^\alpha \times n_A^\alpha$ unitary transformation of these orbitals

together eliminates $n_A^{\alpha 2}$ of the $n_A^\alpha(n_A^\alpha + n_B^\alpha)$ lcao coefficients in $X^{\alpha(A)}$, leaving $n_A^\alpha n_B^\alpha$ independent variables -- equal to the number of elements in f^α. Similarly, orthonormality constraints and redundancy leave only $n_A^\beta n_B^\beta$ independent variables in the β-spin occupied orbitals, which is equal to the number of elements in f^β. Orthogonality between α-spin orbitals and the β-spin orbitals is automatic, due to the orthogonality of the spin parts. The so-called "pairing conditions" sometimes used in the derivation of this different-orbitals-different-spin (DODS) formalism (Rosenberg and Martino, 1975), merely represent a particular choice of some of the redundant variables in the orbital coefficients of the two sets of spin-orbitals, and thus need not be considered in the above arguments concerning the number of degrees of freedom in the problem.

For a variation δR^i in the R^i ($i=\alpha,\beta$), the corresponding change δE in E, eq. (8.28), is given exactly as

$$\delta E = \mathrm{tr}[\delta R^\alpha F^\alpha + \delta R^\beta F^\beta] + \tfrac{1}{2}\mathrm{tr}[\delta R^\alpha \delta G^\alpha + \delta R^\beta \delta G^\beta]. \quad (8.33)$$

Here F^α and F^β are the α-spin and β-spin orbital Fock matrices respectively,

$$F^\alpha = h + G^\alpha = h + J(R^\alpha) - K(R^\alpha) + J(R^\beta) , \quad (8.34a)$$

and

$$F^\beta = h + G^\beta = h + J(R^\beta) - K(R^\beta) + J(R^\alpha) . \quad (8.34b)$$

The first order part of (8.33) is the sum of two terms of the same form as (8.9a) for the closed shell case. Therefore, one has immediately that

$$\frac{\partial E}{\partial f^{i*}_{\sigma r}} = [(1 - R^i S)F^i R^i]_{\sigma r} \qquad (8.35a)$$

$$= F^i_{(e^i_B)^\sigma (e^i_A)^r}, \quad (i = \alpha, \beta). \qquad (8.35b)$$

Thus, the first derivatives of the energy with respect to the elements of the f^i are again just matrix elements of the corresponding current Fock operator between two sets of contragredient (non-orthonormal) molecular orbitals, which are, respectively, the first n^i_A columns of the density matrix R^i and the last n^i_B columns of the matrix $(1 - R^i S)$, $(i = \alpha, \beta)$. It is seen that the first derivatives of the energy with respect to elements of f^α depend on f^β only implicitly through the dependence of F^α on R^β, and vice versa.

The second derivatives of the energy are given by,

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial f^{i*}_{\sigma r}\partial f^{i*}_{\tau s}} = -\tfrac{1}{2}\Big\{(SR^i)_{\sigma s}\frac{\partial E}{\partial f^{i*}_{\tau r}} + (SR^i)_{\tau r}\frac{\partial E}{\partial f^{i*}_{\sigma s}} + [(e^i_B)^\sigma (e^i_A)^\tau \| (e^i_B)^\tau (e^i_A)^s] - [(e^i_B)^\sigma (e^i_A)^s \| (e^i_B)^\tau (e^i_A)^r]\Big\} \qquad (8.36)$$

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial f^{i*}_{\sigma r}\partial f^{i}_{\tau s}} = \tfrac{1}{2}\Big\{R^i_{sr}F^i_{(e^i_B)^\sigma (e^i_B)^\tau} - (1 - R^i S)_{\sigma\tau}F^i_{(e^i_A)^s (e^i_A)^r} + [(e^i_B)^\sigma (e^i_A)^r \| (e^i_A)^s (e^i_B)^\tau] - [(e^i_B)^\sigma (e^i_B)^\tau \| (e^i_A)^s (e^i_A)^r]\Big\},$$

and

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial f^{i*}_{\sigma r}\partial f^{j*}_{\tau s}} = \tfrac{1}{2}[(e^i_B)^\sigma (e^i_A)^r \| (e^j_B)^\tau (e^j_A)^s], \quad i \neq j, \qquad (8.37)$$

$$\frac{1}{2}\frac{\partial^2 E}{\partial f^{i*}_{\sigma r}\partial f^{j}_{\tau s}} = \tfrac{1}{2}[(e_B^i)^\sigma (e_A^i)^r \| (e_A^j)^s (e_B^j)^\tau], \quad i \neq j \ .$$

These formulas are different from those in the closed shell case because the coupling between α-spin and β-spin orbitals is explicit in the second order variation of E with f^α and f^β.

8.3.b Test Calculations and Computational Refinements

A series of minimal basis set (STO) ab initio calculations were carried out on the molecule CN in order to obtain information on the practical implementation of the UHF-SCF formalism just described. Claxton and Smith (1971) have reported convergence problems in similar calculations. A Roothaan iteration procedure converges very slowly when the interatomic distance is 2.0 a.u., and exhibits oscillatory behaviour, failing to converge, when this distance is increased to 2.2 a.u. (see Figures 8.1 and 8.4). When a direct minimization procedure based on Fletcher's method was used, it was found that convergence was rapid at first, but became very slow as the minimum was approached. They concluded that the most efficient procedure was to use the direct method initially, until a good estimate of the energy minimum was obtained, and then complete the calculation using a Roothaan iteration procedure, which converges well when provided with a good starting approximation.

The calculations here were carried out on an IBM 370/168 computer using double precision arithmetic. The integrals in

the Slater orbital basis were obtained from a version of the POLYCAL program. Orbital exponents were taken from Clementi (1963). The linear search step in the conjugate gradient algorithm was required to reduce $\partial E/\partial\lambda$ by a factor ϵ compared to its value at $\lambda = 0$, and ϵ was usually chosen as 0.1. The starting approximation in all but one case was equivalent to the eigenvectors of the core hamiltonian.

It was found that the convergence of the direct minimization calculations based on the partitioning formalism was very poor if the $f^i_{\sigma r}$ were defined by an arbitrary partitioning of the atomic orbital basis. Convergence improves greatly if they are defined by the partitioning of a set of molecular orbitals, X_o, which more nearly block diagonalize the Fock operator. In practice, this involves evaluating the energy gradient and f-operator in the new basis, that is,

$$\nabla_{fi}E^{MO} = (e^{i\dagger}_B X^\dagger_o)F^{iAO}(X_o e^i_A), \tag{8.38}$$

the calculation requiring less computation if the quantities in the brackets are evaluated first, and then the back-transformation of the density matrix as calculated from the MO basis f-operator using (8.31),

$$(R^i)^{AO} = X_o(R^i)^{MO}X^\dagger_o, \tag{8.39}$$

if $X^\dagger_o S X_o = 1$. No transformation of the two-electron integrals is necessary.[5] The Fletcher and Roothaan calculations were done

[5]The transformation to the MO basis has an additional advantage when working in a non-orthonormal AO basis, because if the new (cont'd)

in the original atomic orbital basis. Tables 8.2 and 8.3 summarize the results of sixteen different calculations done here. The relative rates of convergence of some of the methods and refinements are also illustrated in Figures 8.1 - 8.3 for the CN molecule with a bond length of 2.0 a.u., and in Figures 8.4 - 8.6 for a bond length of 2.2 a.u. The energy range is Figures 8.1 and 8.4 is larger by a ratio of 400:15 than that in the other four figures.

On comparing the results of the calculations involving Fletcher's method (2, 4, 5) to those based on the use of the $f^i_{\sigma r}$, (6, 7), it is seen that not only do both methods converge poorly near the energy minimum, but that Fletcher's method actually slightly outperforms the method based on the partitioning formalism. This is also seen in Figures 8.1 and 8.4.

A number of modifications of the basic method based on the use of the $f^i_{\sigma r}$ were examined. Slow rates of convergence near the minimum imply significant linear dependence between successive search directions in the conjugate gradient calculation. Simply restarting the calculation with a steepest descent direction more frequently resulted in no improvement (Figures 8.2 and 8.5). However, a major increase in the rate of convergence was obtained when the basis, X_o, in which the partitioning was defined was replaced by the eigenbasis of the current Fock

---

basis vectors satisfy $X_o^\dagger S X_o = 1$, then the energy gradient formulas applying in an orthonormal basis can be used since $S^{MO} = 1$. This partially, if not completely, offsets the additional cost of the transformations in (8.38) and (8.39).

TABLE 8.2 Details of Direct Minimization Calculations

CN Molecule (r = 2.0 a.u.)

	Type[a]	min. alg.[b]	ϵ	modification			Final Energy[f] a.u.	rank[m]
				c	d	e		
1	R						-112.691216[g]	16
2	Fl	c.g.					-112.835540[h]	10
3	R						-112.845722[i]	3
4	Fl	c.g.	0.1				-112.841230	8
5	Fl	c.g.	0.01				-112.840916	9
6	P	c.g.	0.1				-112.818824	14
7	P	c.g.	0.01				-112.822130	13
8	P	c.g.	0.1	3			-112.805275	15
9	P	c.g.	0.1	3	x		-112.845708	4
10	P	c.g.	0.1			x[k]	-112.845365	6
11	P	c.g.	0.1	3		x	-112.845418	5
12	P	c.g.	0.1	3	x	x	-112.845722[j]	1
13	P	s.d.	0.1				-112.824592	12
14	P	s.d.	0.1		3[l]		-112.845121	7
15	P	s.d.	0.1			x[k]	-112.832094	11
16	P	s.d.	0.1		3[l]	x	-112.845722	2

[a] R=Roothaan, Fl=Fletcher, P=Partitioning

[b] c.g. = conjugate gradient, s.d. = steepest descent.

[c] steepest descent restart frequency.

[d] basis update at steepest descent restart.

[e] gradient scaling in effect

[f] after 30 iterations unless otherwise noted, exact energy is -112.845722 a.u.

[g] 28 iterations

[h] 29 iterations

[i] uses final result from calculation #2 as starting approximation.

[j] convergence criteria $|\delta E|<10^{-12}$, $|\delta R_{ii}|<10^{-6}$ satisfied in 25 its.

[k] using eigenvalues of core hamiltonian.

[l] indicates the frequency of basis modification.

[m] indicates the order of the final energies, from lowest to highest.

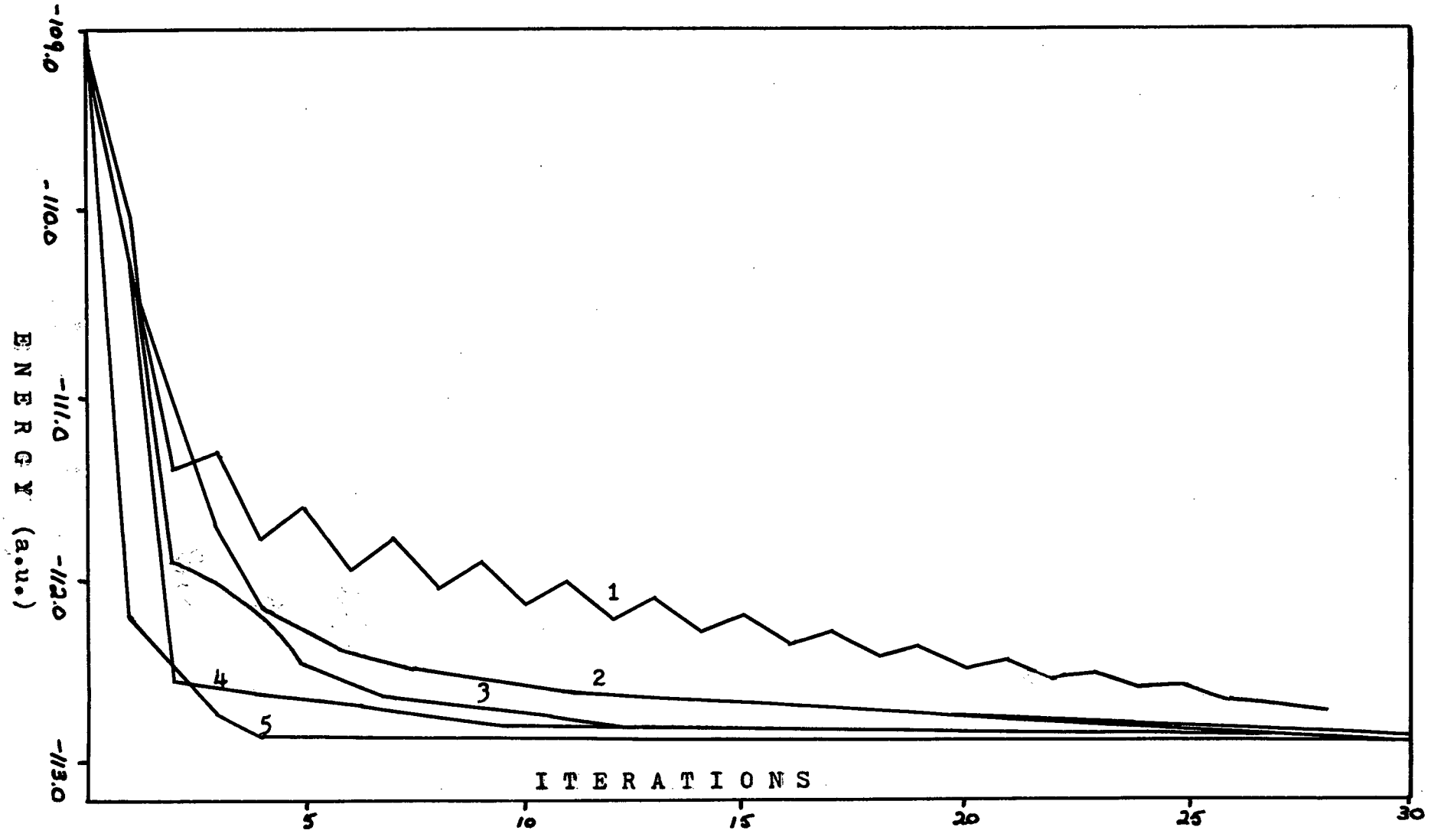

FIGURE 8.1 Total electronic energy as a function of iteration number for the CN molecule, (bond length = 2.0 a.u.). (1) Roothaan; (2) partitioning, steepest descent search directions only; (3) partitioning, conjugate gradients; (4) Fletcher, conjugate gradients; (5) partitioning conjugate gradients with gradient scaling and basis update with steepest descent restart every 3 iterations. In all direct minimization calculations, $\epsilon = 0.1$ (see Table 8.2).

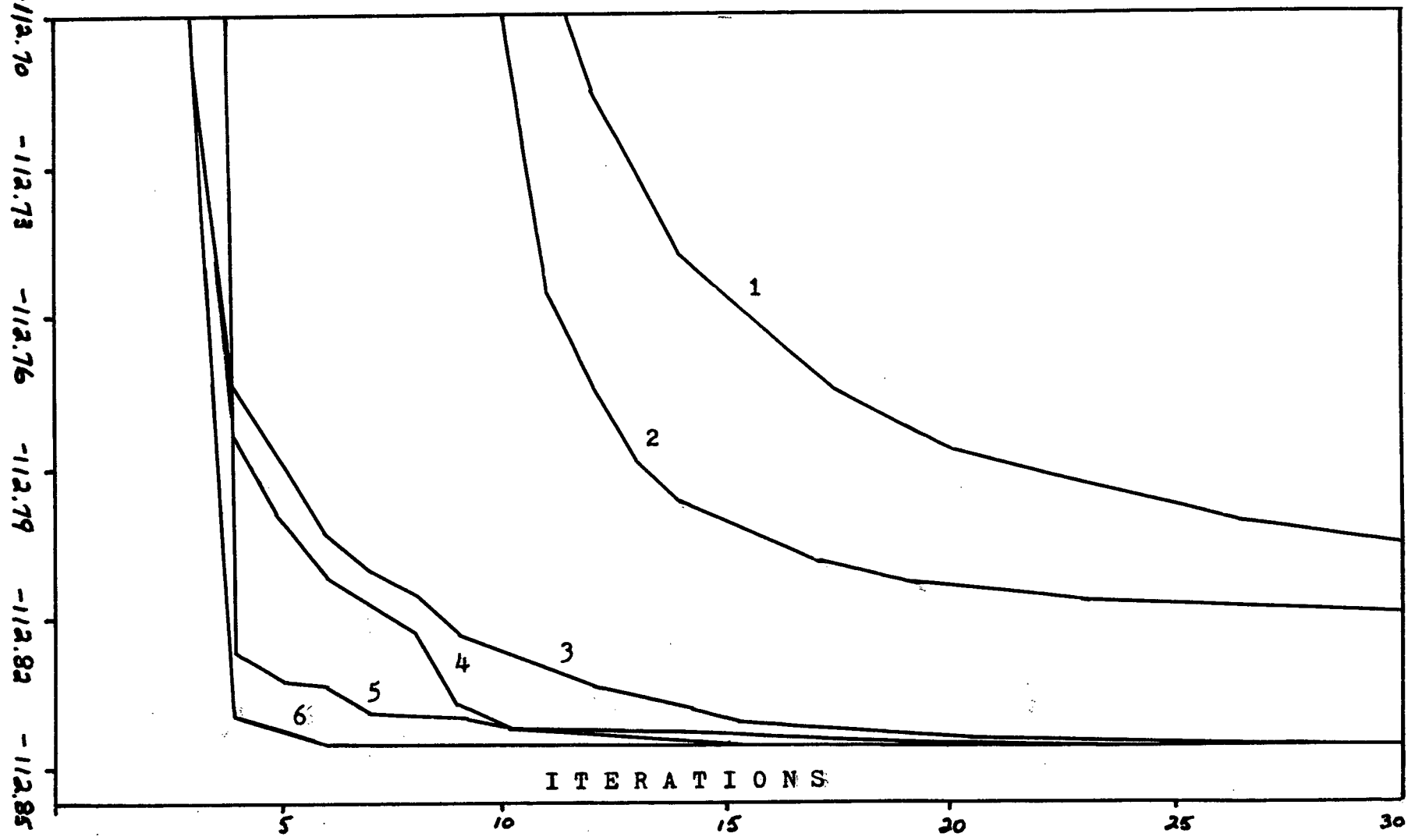

FIGURE 8.2 Total electronic energy as a function of iteration number for the CN molecule, (bond length = 2.0 a.u.). Comparison of the effect of various modifications on the conjugate gradient algorithm--partitioning approach only: (1) steepest descent restart every 3 iterations only; (2) basic conjugate gradient algorithm; (3) gradient scaling only; (4) gradient scaling and steepest descent restart every 3 iterations; (5) steepest descent restart every 3 iterations with basis update at restart; (6) gradient scaling, steepest descent restart every 3 iterations with basis update at restart (see Table 8.2).

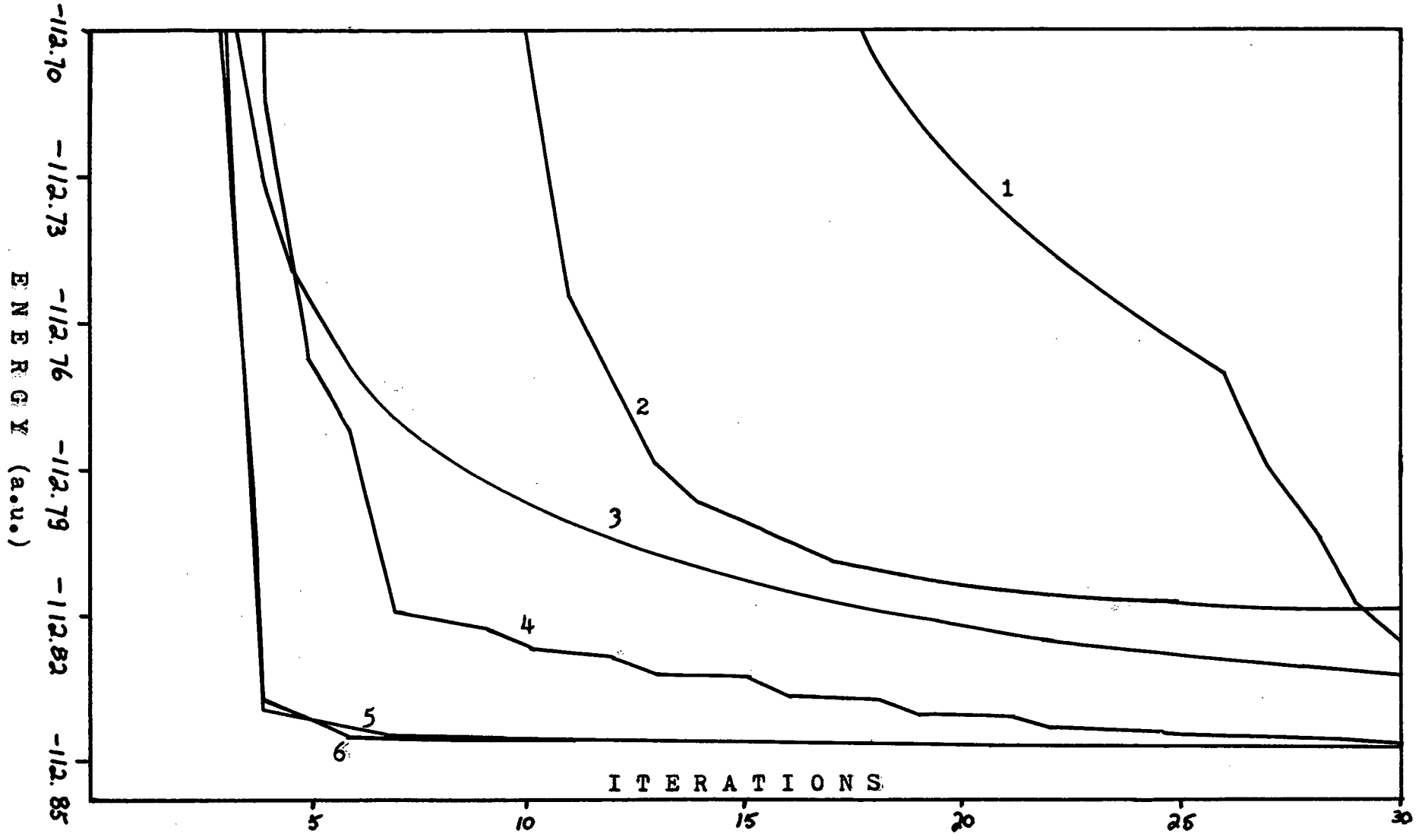

FIGURE 8.3 Total electronic energy as a function of iteration number for the CN molecule, (bond length = 2.0 a.u.). Comparison of the effect of various modifications on the steepest descent algorithm--partitioning approach only: (1) steepest descent algorithm only; (2) basic conjugate gradient algorithm; (3) steepest descents with gradient scaling only; (4) steepest descents with basis update every 3 iterations; (5) steepest descents with gradient scaling and basis update every 3 iterations; (6) conjugate gradients with gradient scaling, steepest descent restart every 3 iterations with basis update at restart (see Table 8.2).

operators at the time of the steepest descent restart, as was done for the calculations numbered 9 in the tables. The partitioning operators f^i are set to zero when the new basis is incorporated into the calculation, and therefore, this basis modification is equivalent to a single Roothaan iteration.

8.3.c Use of Scaled Variables

A second modification of the basic algorithm which results in a major improvement in convergence, is suggested by the Newton-Raphson equations for determining the zeros of the energy gradient (see Appendix 11). Upon neglecting the two-electron integrals in eqs. (8.36) and (8.37), and in a basis diagonalizing the current Fock operators, it is seen that

$$\frac{\partial^2 E}{\partial f^{i\,2}_{\sigma r}} \approx \epsilon^i_\sigma - \epsilon^i_r . \tag{8.40}$$

Thus, the diagonal approximation of the Newton-Raphson equations can be written,

$$\delta f^i_{\sigma r} \approx -\lambda(\epsilon^i_\sigma - \epsilon^i_r)^{-1} \frac{\partial E}{\partial f^i_{\sigma r}} , \tag{8.41}$$

where λ is some constant independent of σ and r. If the minimization problem is rewritten in terms of a new set of variables,

$$\tilde{f}^i_{\sigma r} = (\epsilon^i_\sigma - \epsilon^i_r)^{\frac{1}{2}} f^i_{\sigma r} , \tag{8.42}$$

TABLE 8.3 Details of Direct Minimization Calculations

CN Molecule (r = 2.2 a.u.)

	Type[a]	min. alg.[b]	ϵ	modification			Final Energy[f] a.u.	rank[l]
				c	d	e		
1	R						osc.[g]	16
2	Fl	c.g.					-110.991333[h]	8
3	R						-111.012914[i]	2
4	Fl	c.g.	0.1				-110.995595	5
5	Fl	c.g.	0.01				-110.994936	6
6	P	c.g.	0.1				-110.976114	13
7	P	c.g.	0.01				-110.977213	12
8	P	c.g.	0.1	3			-110.975463	14
9	P	c.g.	0.1	3	x		-111.011872	3
10	P	c.g.	0.1			x[j]	-110.981282	11
11	P	c.g.	0.1	3		x	-110.984485	9
12	P	c.g.	0.1	3	x	x	-111.012980	1
13	P	s.d.	0.1				-110.951620	15
14	P	s.d.	0.1		3[k]		-110.992565	7
15	P	s.d.	0.1			x[j]	-110.981113	10
16	P	s.d.	0.1		3[k]	x	-111.010863	4

[a] R=Roothaan, Fl=Fletcher, P=Partitioning

[b] c.g. = conjugate gradient, s.d. = steepest descent.

[c] steepest descent restart frequency.

[d] basis update at steepest descent restart.

[e] gradient scaling in effect.

[f] after 30 iterations unless otherwise noted, exact energy is -111.012980 a.u.

[g] 28 iterations

[h] 29 iterations

[i] uses final result from calculation #2 as starting approximation.

[j] using eigenvalues of core hamiltonian.

[k] indicates frequency of basis modification.

[l] indicates the order of the final energies from lowest to highest.

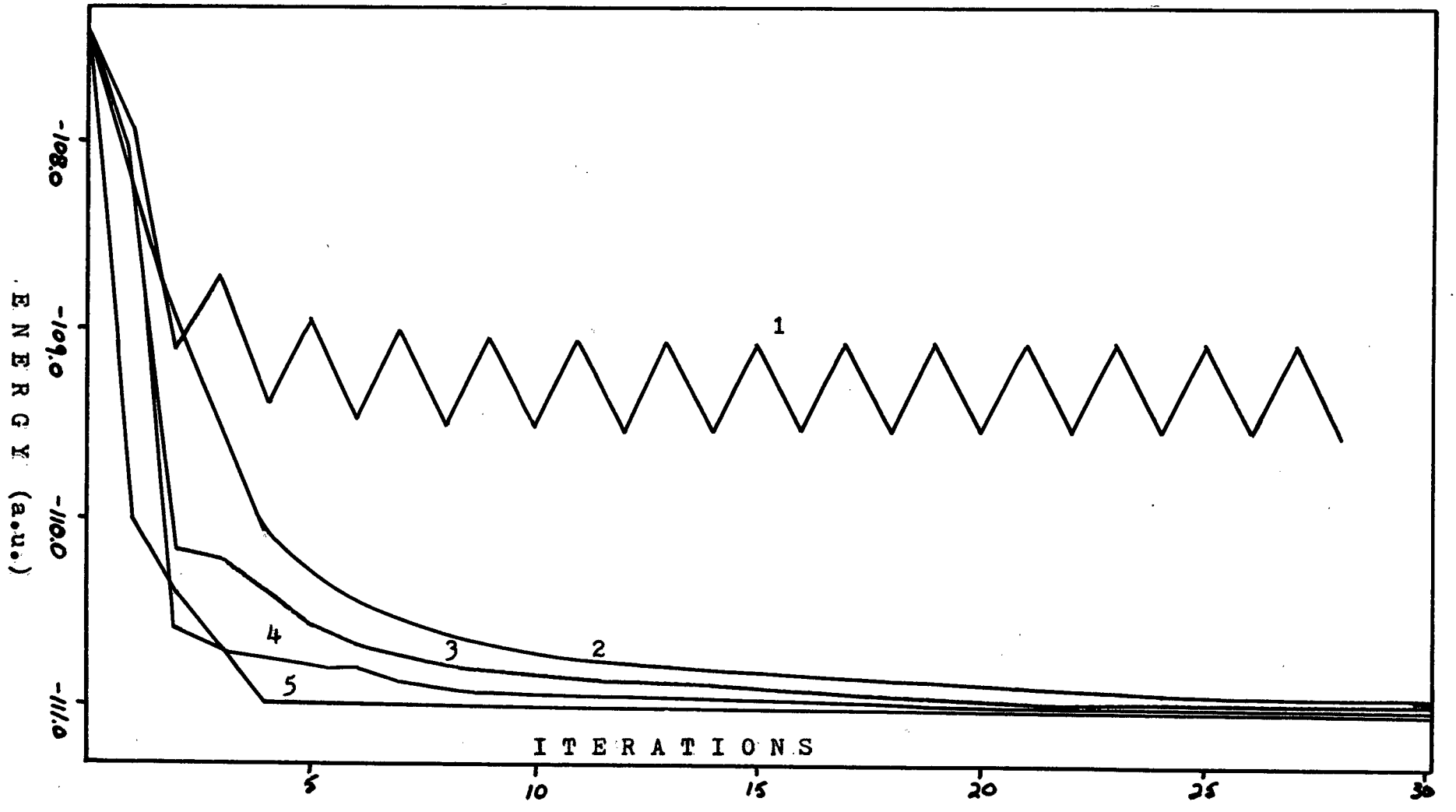

FIGURE 8.4 Total electronic energy as a function of iteration number for the CN molecule, (bond length = 2.2 a.u.). (1) Roothaan; (2) partitioning, steepest descent search directions only; (3) partitioning, conjugate gradients; (4) Fletcher, conjugate gradients; (5) partitioning, conjugate gradients with gradient scaling and basis update with steepest descent restart every 3 iterations. In all direct minimization calculations, $\epsilon = 0.1$ (see Table 8.3).

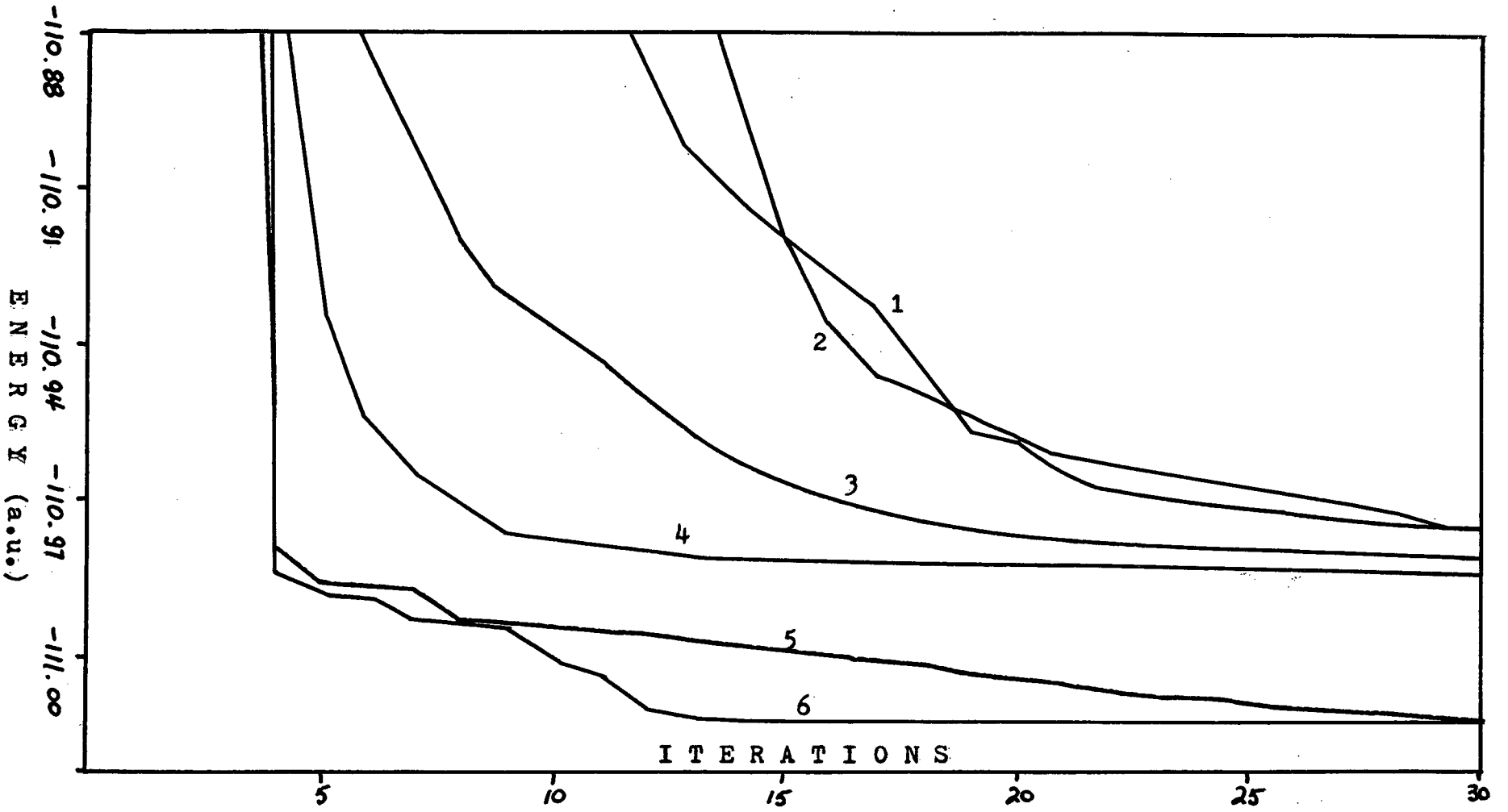

FIGURE 8.5 Total electronic energy as a function of iteration number for the CN molecule, (bond length = 2.2 a.u.). Comparison of the effect of various modifications on the conjugate gradient algorithm--partitioning approach only: (1) steepest descent restart every 3 iterations only; (2) basic conjugate gradient algorithm; (3) gradient scaling only; (4) gradient scaling and steepest descent restart every 3 iterations; (5) steepest descent restart every 3 iterations with basis update at restart; (6) gradient scaling, steepest descent restart every 3 iterations with basis update at restart (see Table 8.3).

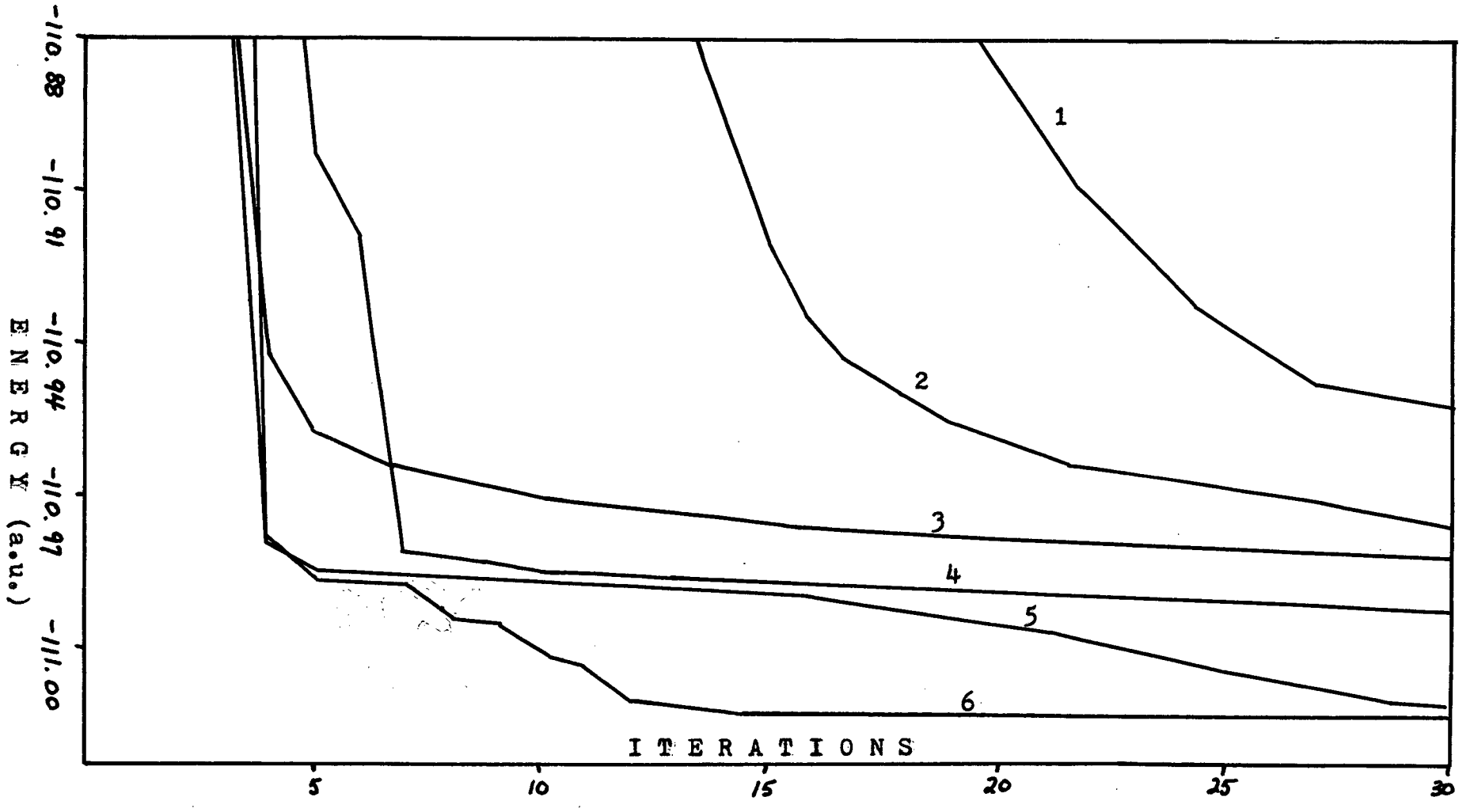

FIGURE 8.6 Total electronic energy as a function of iteration number for the CN molecule, (bond length = 2.2 a.u.). Comparison of the effect of various modifications on the steepest descent algorithm--partitioning approach only: (1) basic steepest descent algorithm; (2) basic conjugate gradient algorithm; (3) steepest descents with gradient scaling only; (4) steepest descents with basis update every 3 iterations; (5) steepest descents with gradient scaling and basis update every 3 iterations; (6) conjugate gradients with gradient scaling, steepest descent restart every 3 iterations with basis update at restart (see Table 8.3).

then the diagonal Newton-Raphson equations, which are approximations of a second order convergent method, give the correction $\delta\tilde{f}^{i}_{\sigma r}$ as a simple step (the same for all σ and r) along the steepest direction. This modification can easily be incorporated into the ordinary conjugate gradient formalism by scaling the energy gradient,

$$\frac{\partial E}{\partial \tilde{f}^{i*}_{\sigma r}} = (\epsilon^{i}_{\sigma} - \epsilon^{i}_{r})^{-\frac{1}{2}} \frac{\partial E}{\partial f^{i*}_{\sigma r}} \quad . \tag{8.43}$$

Then the appropriate correction for the unscaled variables is

$$\delta f^{i}_{\sigma r} = (\epsilon^{i}_{\sigma} - \epsilon^{i}_{r})^{-\frac{1}{2}} \delta\tilde{f}^{i}_{\sigma r} = \lambda(\epsilon^{i}_{\sigma} - \epsilon^{i}_{r})^{-\frac{1}{2}} v_{\sigma r} \quad , \tag{8.44}$$

where v is the conjugate search direction, computed from the scaled gradients, and λ is the step length computed by interpolation in the usual manner. This gradient scaling (or the implicit use of the scaled variables $\tilde{f}^{i}_{\sigma r}$) is aimed at correcting the problems in descent methods caused by anisotropy in the curvature of the energy surface. In practice, the numbers ϵ^{i}_{μ} used are the best available estimates of the eigenvalues of the Fock operator at any stage of the calculation. Initially, any suitable estimate may be used (for example, orbital energies from a semi-empirical calculation of some sort, or even the eigenvalues of the core hamiltonian, as was done for the calculations described in Tables 8.2 and 8.3). This scaling procedure has no simple counterpart for Fletcher's method.

The calculations numbered 10 were done by incorporating only this scaling procedure into the basic conjugate gradient

algorithm, resulting in a substantial improvement in the rate of convergence. Increasing the steepest descent restart frequency to every three iterations (compared with 45 as recommended by Fletcher and Reeves, (1964)), resulted in a further small improvement. However, when the molecular orbital basis defining the partitioning is replaced by the eigenbasis of the current Fock matrix at the steepest descent restart, the most rapidly convergent algorithm resulted. For the 2.0 a.u. interatomic distance, the energy became correct to fifteen figures (effectively the limit of the machine precision), and the diagonal elements of the density matrices to seven figures, in only 25 iterations. Nearly the same results were obtained for the 2.2 a.u. bond length.

The test calculations described above support the assertion that the singularity of the Hessian matrix at the energy minimum has no observable effect on the rate of convergence of the conjugate gradient algorithm.[6] Rather, they indicate that much of the poor convergence is due to the fact that the energy curvature is highly anisotropic in general, and the usual conjugate gradient algorithm does not take proper account of this. A single average

[6] For a quadratic form, it is easily demonstrated that a singular Hessian matrix has no effect on the convergence properties of the conjugate gradient algorithm, except that the minimum will be located in fewer iterations, since no linear search is required in directions corresponding to those along which the form has zero curvature. In a converging energy minimization calculation, the part of the coordinate space corresponding to the redundant variables should effectively act as a null space as far as the choice of search directions is concerned. This is especially true near a minimum, where the energy is most like a quadratic form.

step length along the descent direction generated tends to overestimate the necessary correction for some variables and underestimate it for the others. This is a well known shortcoming of steepest descent procedures also. In fact, the steepest descent algorithm employing a cubic interpolation linear search, does not converge much more slowly than the conjugate gradient algorithm. It appears that in application to direct minimization self-consistent field theory, the finite termination property of the conjugate gradient method is of little advantage, since, even for the smallest systems, this finite termination in principle requires considerably more iterations than are acceptable if efficient calculations are to result.

No attempt was made in this series of calculations to determine an optimal restart frequency. The rate of convergence is usually greatest either in an iteration involving a restart, or in the one immediately following, and therefore, it is unlikely that an interval between restarts of much more than three iterations will result in faster overall convergence. Despite such frequent steepest descent restarts, there still appears to be some advantage to using the conjugate gradient search directions, as can be seen on comparison of calculations 13 - 16, respectively, with calculations 6, 9, 10, and 12. The calculation of these search directions from the steepest directions is a small part of the whole calculation. Even so, the steepest directions by themselves give remarkably good results here (see Figures 8.3 and 8.6).

8.4 Theory for the General Single Determinant Case

8.4.a The Basic Variables of the Calculation

We now consider the case of an N-electron system represented by a single determinant wavefunction constructed from occupied orbitals for which there is a natural grouping into m sets, called shells, which are relatively weakly coupled by the hamiltonian operator. The n_I occupied orbitals, $X_I^{(i)}$, associated with the i^{th} shell, are chosen from a set of n orbitals $X^{(i)}$, which are eigenfunctions of the Fock operator, $F^{(i)}$, referring to the i^{th} shell. The total energy of the system, eq. (8.1), is then completely determined by the projections (one-particle density matrices in molecular orbital theory),

$$R^{(i)} = X_I^{(i)} X_I^{(i)\dagger}, \qquad (i = 1, \ldots, m), \tag{8.45}$$

onto the individual n_I-dimensional subspaces of the full n-dimensional basis space, each subspace spanned by one of these sets of occupied orbitals. It will be shown that the columns of these projections and their complements again provide non-orthonormal basis vectors, in terms of which the first and second derivatives of the energy, (8.1), with respect to a set of variables provided by the multi-partitioning formalism of chapter 4, can be written effectively as compactly as in the closed shell case.

If the simple form of the energy, (8.1), is to be preserved, these projections must satisfy the constraints

$$R^{(i)\dagger} = R^{(i)}, \tag{8.46a}$$

and

$$R^{(i)}SR^{(j)} = R^{(i)}\delta_{ij}, \tag{8.46b}$$

or, equivalently, the occupied orbitals must satisfy the orthonormality conditions,

$$X_I^{(i)\dagger}SX_J^{(j)} = 1_I\delta_{ij}. \tag{8.47}$$

Here S is the matrix of overlap integrals of the fixed basis functions in terms of which the $X^{(i)}$ and $R^{(i)}$ are defined.

The number of independent parameters necessary to specify the energy exactly is determined as follows. The total number of parameters in the lcao coefficients, $X_I^{(i)}$, $(i = 1, \ldots, m)$, is $n\sum_{I=1}^{m} n_I$, where $n = \sum_{I=1}^{m+1} n_I$, is the dimension of the full basis space. Within each set $X_I^{(i)}$, the orthonormality constraint, and the redundancy due to the invariance of the energy, (8.1), to an arbitrary unitary transformation of orbitals in the same shell, together account for n_I^2 parameters. The orthogonality of orbitals in different shells is expressed by $\sum_{I=2}^{m} n_I \sum_{J=1}^{I-1} n_J$ unique conditions, making it possible, in principle, to eliminate an equal number of parameters. Thus, the total number of unconstrained and non-redundant parameters required to specify the energy in the form (8.1) is $\sum_{I=1}^{m} n_I(n - \sum_{J=1}^{I} n_J)$.

The intra-shell constraints and redundancy can be eliminated by rewriting the energy in terms of a new set of parameters chosen as follows. For each i, $(i = 1, \ldots, m)$, the n eigenvectors $X^{(i)}$, of the Fock operator $F^{(i)}$, are divided into m+1

subsets, $X_J^{(i)}$, of dimension n_J, respectively, such that the subset $X_I^{(i)}$ is the set of occupied i^{th} shell orbitals. Similarly, the n-dimensional basis space is partitioned into m+1 subspaces S_J, (J = 1, ..., m+1), of the same dimensions, n_J, respectively. The matrices $X^{(i)}$, (i = 1, ..., m), can now be written in an (m+1) x (m+1) block form, similar to that in eq. (4.2). A set of uncoupling operators, $\hat{T}^{(i)}$, is defined, such that (eq. (4.4)),

$$X^{(i)} = \hat{T}^{(i)}\hat{X}^{(i)}, \qquad (i = 1, \ldots, m), \tag{8.48}$$

where $\hat{X}^{(i)}$ is the diagonal block part of $X^{(i)}$, making it possible to write the block columns of interest as

$$X_I^{(i)} = \begin{bmatrix} X_{1I}^{(i)} \\ X_{2I}^{(i)} \\ \vdots \\ X_{m+1,I}^{(i)} \end{bmatrix} = \hat{T}_I^{(i)} X_{II}^{(i)} = \begin{bmatrix} f_{1I}^{(i)} \\ f_{2I}^{(i)} \\ \vdots \\ 1_I \\ \vdots \\ f_{m+1,I}^{(i)} \end{bmatrix} X_{II}^{(i)}, \tag{8.49a}$$

(i = 1, ..., m), where

$$f_{JI}^{(i)} = X_{JI}^{(i)} X_{II}^{(i)-1}, \qquad (J = 1, \ldots, m+1). \tag{8.49b}$$

That is, we have implicitly set up m (m+1)-fold partitionings, one for each of the $X^{(i)}$. The parts of each shown in (8.49a) are the only ones which enter the energy expression, (8.1). The $f_{JI}^{(i)}$, (J = 1, ..., m+1, J≠I), are specified by $n_I(n-n_I)$

complex parameters, which is exactly the number of parameters in the $X_I^{(i)}$ after intra-shell orthonormality and redundancy have been accounted for.

The intershell orthogonality constraints imply $\sum_{I=2}^{m} \sum_{J=1}^{I-1} n_I n_J$ relations between the elements of the $f_{JI}^{(i)}$, (i = 1, ..., m). The explicit incorporation of these relations into the theory will be considered in section 8.4.d.

8.4.b The Energy Variation and First Derivatives

The energy functional will be written here in terms of the $R^{(i)}$ as

$$E = \sum_{i=1}^{m} \text{tr } v_i R^{(i)} h + \tfrac{1}{2} \sum_{i=1}^{m} \text{tr } v_i R^{(i)} G_i \,, \qquad (8.50)$$

where h is the core hamiltonian matrix, representing the electronic kinetic energy, and the interaction between the electrons and the nuclei, and the G_i represent the inter-electronic repulsion terms of the hamiltonian operator. In detail, one has

$$G_i = \sum_{j=1}^{m} G_{i,j}(v_j R^{(j)}) = \sum_{j=1}^{m} v_j G_{i,j}(R^{(j)}), \qquad (8.51)$$

where

$$G_{i,j}(R) = J(R) - a_{ij} K(R), \qquad (8.52a)$$

and

$$a_{ij} = 1 \quad \text{if } v_i = v_j = 1,$$
$$= \tfrac{1}{2} \quad \text{otherwise.} \qquad (8.52b)$$

The matrices J(R) and K(R) are the usual Coulomb and exchange matrices with elements given by

$$J(R)_{rs} = \sum_{t,u} R_{tu}[rs\|ut], \quad K(R)_{rs} = \sum_{t,u} R_{tu}[rt\|us], \tag{8.52c}$$

where the symbol $[rs\|ut]$ is defined in eq. (7.25b). The occupation numbers, ν_i, may have values of 1 or 2 only.

An incremental approach is employed to obtain the derivatives of the energy. A change $\delta R^{(i)}$ in $R^{(i)}$, $(i = 1, \ldots, m)$, produces a change in the energy given exactly by

$$\delta E = \sum_{i=1}^{m} \nu_i \mathrm{tr}\, \delta R^{(i)} F^{(i)} + \tfrac{1}{2} \sum_{i=1}^{m} \nu_i \mathrm{tr}\, \delta R^{(i)} \delta G_i, \tag{8.53}$$

where $F^{(i)}$ is the Fock matrix associated with the i^{th} shell,

$$F^{(i)} = h + G_i. \tag{8.54}$$

In the notation established in the previous subsection, the blocks of the projection $R^{(i)}$ are given by

$$R^{(i)}_{JK} = f^{(i)}_{JI} g^{(i)-1}_{I} f^{(i)\dagger}_{KI}, \tag{8.55}$$

where

$$g^{(i)}_{I} = (X^{(i)}_{II} X^{(i)\dagger}_{II})^{-1}$$

$$= S_{II} + \sum_{L\neq I} S_{IL} f^{(i)}_{LI} + \sum_{K\neq I} f^{(i)\dagger}_{KI} S_{KI} + \sum_{\substack{K\neq I \\ L\neq I}} f^{(i)\dagger}_{KI} S_{KL} f^{(i)}_{LI} \tag{8.56}$$

Then, one has,

$$\delta R^{(i)}_{JK} = \delta f^{(i)}_{JI} g^{(i)-1}_{I} f^{(i)\dagger}_{KI} + f^{(i)}_{JI} \delta g^{(i)-1}_{I} f^{(i)\dagger}_{KI} + f^{(i)}_{JI} g^{(i)-1}_{I} \delta f^{(i)\dagger}_{KI}$$
$$+ \delta f^{(i)}_{JI} \delta g^{(i)-1}_{I} f^{(i)\dagger}_{KI} + \delta f^{(i)}_{JI} g^{(i)-1}_{I} \delta f^{(i)\dagger}_{KI} + f^{(i)}_{JI} \delta g^{(i)-1}_{I} \delta f^{(i)\dagger}_{KI}$$
$$+ \delta f^{(i)}_{JI} \delta g^{(i)-1}_{I} \delta f^{(i)\dagger}_{KI}. \tag{8.57}$$

To second order, one has

$$\delta g_I^{(i)-1} = -g_I^{(i)-1}\delta g_I^{(i)} g_I^{(i)-1} + g_I^{(i)-1}\delta g_I^{(i)} g_I^{(i)-1}\delta g_I^{(i)} g_I^{(i)-1} + O(\delta^3), \tag{8.58}$$

where eq. (8.56) yields

$$\delta g_I^{(i)} = \sum_{L\neq I} S_{IL}\delta f_{LI}^{(i)} + \sum_{K\neq I} \delta f_{KI}^{(i)\dagger} S_{KI} + \sum_{\substack{L\neq I \\ K\neq I}} [\delta f_{KI}^{(i)\dagger} S_{KL} f_{LI}^{(i)} + f_{KI}^{(i)\dagger} S_{KL}\delta f_{LI}^{(i)} + \delta f_{KI}^{(i)\dagger} S_{KL}\delta f_{LI}^{(i)}]. \tag{8.59}$$

The calculation of the first derivatives is simplified by noting that terms in δE linear in $\delta f_{JI}^{(i)}$ or $\delta f_{JI}^{(i)\dagger}$, for a specific value of i, can only enter via a single term of the first summation (over shells) in eq. (8.53). Substituting (8.57) - (8.59) into (8.53), and retaining terms only to first order in the $\delta f_{JI}^{(i)}$ and their adjoints, one obtains

$$\delta^{(1)}E = \sum_{i=1}^{m} \nu_i \mathrm{tr} \sum_{J=1}^{m+1}\sum_{K=1}^{m+1} \Big\{ \delta f_{JI}^{(i)} g_I^{(i)-1} f_{KI}^{(i)\dagger} + f_{JI}^{(i)} g_I^{(i)-1}\delta f_{KI}^{(i)\dagger} - f_{JI}^{(i)} g_I^{(i)-1} \Big[\sum_{L\neq I} S_{IL}\delta f_{LI}^{(i)} + \sum_{M\neq I} \delta f_{MI}^{(i)\dagger} S_{MI} + \sum_{\substack{M\neq I \\ L\neq I}} (\delta f_{MI}^{(i)\dagger} S_{ML} f_{LI}^{(i)} + f_{MI}^{(i)\dagger} S_{ML}\delta f_{LI}^{(i)}) \Big] \times g_I^{(i)-1} f_{KI}^{(i)\dagger} F_{KJ}^{(i)} \Big\}$$

$$= \sum_{i=1}^{m} \nu_i \mathrm{tr} \sum_{\substack{P=1 \\ \neq I}}^{m+1} \delta f_{PI}^{(i)\dagger} \Big[(F^{(i)}\hat{T}^{(i)})_{PI} - (S\hat{T}^{(i)})_{PI} g_I^{(i)-1} \times (\hat{T}^{(i)\dagger} F^{(i)} \hat{T}^{(i)})_{II} \Big] g_I^{(i)-1} +$$

$$+\sum_{\substack{P=1\\ \neq I}}^{m+1} \delta f_{PI}^{(i)} g_I^{(i)-1}\left[(\hat{T}^{(i)\dagger}F^{(i)})_{IP}-(\hat{T}^{(i)\dagger}F^{(i)}\hat{T}^{(i)})_{II}\right.$$

$$\left. \times g_I^{(i)-1}(\hat{T}^{(i)\dagger}S)_{IP}\right]$$

$$= \sum_{i=1}^{m} \nu_i \mathrm{tr}\Big[\sum_{\substack{P=1\\ \neq I}}^{m+1} \delta f_{PI}^{(i)\dagger} \sum_{J,K} (1-SR^{(i)})_{PK} F_{KJ}^{(i)} R_{JI}^{(i)}$$

$$+\sum_{\substack{P=1\\ \neq I}}^{m+1} \delta f_{PI}^{(i)} \sum_{J,K} R_{IK}^{(i)} F_{KJ}^{(i)} (1-R^{(i)}S)_{JP}\Big]. \qquad (8.60a)$$

Thus, on defining a new set of non-orthonormal basis vectors,

$$\left.\begin{aligned} e_{JK}^{(i)} &= (1 - R^{(i)}S)_{JK}, \quad K\neq I, \\ e_{JI}^{(i)} &= R_{JI}^{(i)}, \end{aligned}\right\} \begin{aligned} &(J,K = 1, \ldots, m+1), \\ &(i = 1, \ldots, m), \end{aligned} \qquad (8.61)$$

one can write,

$$\delta^{(1)}E = \sum_{i=1}^{m} \nu_i \mathrm{tr} \sum_{\substack{P=1\\ \neq I}}^{m+1} \left[\delta f^{(i)\dagger} F^{(i)}_{e_P^{(i)\dagger} e_I^{(i)}} + \delta f^{(i)} F^{(i)}_{e_I^{(i)\dagger} e_P^{(i)}}\right]. \qquad (8.60b)$$

From this, one obtains,

$$\frac{1}{\nu_i} \frac{\partial E}{\partial (f_{PI}^{(i)*})_{\mu\nu}} = F^{(i)}_{(e_P^{(i)})^\mu (e_I^{(i)})^\nu}, \qquad (8.62)$$

as the formal first derivatives of the energy with respect to the elements of the $f_{JI}^{(i)}$ and their hermitian conjugates. As in the simpler closed shell case, the first derivatives of the energy are matrix elements of the appropriate current Fock operator in a basis of non-orthonormal molecular orbitals

which are columns of the corresponding projection $R^{(i)}$ and the complementary matrix $(1-R^{(i)}S)$. The metric properties of the basis vectors, (8.61), are examined in Appendix 13.

8.4.c The Second Derivatives

The second derivatives of the energy are obtained in a straightforward, but somewhat tedious, manner, by isolating the second order terms in (8.53). These terms consist of two types. The first arises from the trace over the product of second order variations in the projections $R^{(i)}$ and the corresponding Fock operator $F^{(i)}$, while the second arises from the terms of the form $\mathrm{tr}\ \delta R^{(i)}\delta G_i$, which contain products of first order variations of the density matrices.

Consider first the simple term

$$\mathrm{tr}\ \delta R^{(i)}\delta G_i = \sum_{j=1}^{m} v_j \sum_{\substack{r,s\\t,u}}^{n} \delta R^{(i)}_{sr}\delta R^{(i)}_{tu}\{[rs\|ut]-a_{ij}[rt\|us]\}. \tag{8.64}$$

This equation can be viewed as representing linear transformations on the basis functions in terms of which the two-electron integrals are evaluated. The summations over r,s can be treated independently of those over t and u, above. To first order, one has

$$\sum_{r,s} \delta R^{(i)}_{sr}\langle r\|s\rangle = \sum_{J,K=1}^{m+1} \sum_{\substack{r\in K\\ s\in J}} (\delta R^{(i)}_{JK})_{sr}\langle r\|s\rangle$$

$$= \sum_{\substack{P=1\\ \neq I}}^{m+1} \sum_{\substack{\mu\in P\\ \nu\in I}} (\delta f_{PI}^{(i)})_{\mu\nu} \sum_{J,K=1}^{m+1} \sum_{\substack{r\in K\\ s\in J}} [\delta_{JP} - \sum_{i=1}^{m+1} f_{JI}^{(i)} g_I^{(i)-1} f_{LI}^{(i)\dagger} S_{LP}]_{s\mu}$$

$$\times [g_I^{(i)-1} f_{KI}^{(i)\dagger}]_{\nu r} \langle r \| s \rangle$$

$$+ \sum_{\substack{P=1\\ \neq I}}^{m+1} \sum_{\substack{\mu\in I\\ \nu\in P}} (f_{PI}^{(i)\dagger})_{\mu\nu} \sum_{J,K=1}^{m+1} \sum_{\substack{r\in K\\ s\in J}} [f_{JI}^{(i)} g_I^{(i)-1}]_{s\mu}$$

$$\times [\delta_{KP} - \sum_{i=1}^{m+1} S_{KP} f_{LI}^{(i)} g_I^{(i)-1} f_{KI}^{(i)\dagger}]_{\nu r} \langle r \| s \rangle$$

$$= \sum_{\substack{P=1\\ \neq I}}^{m+1} \sum_{\substack{\mu\in P\\ \nu\in I}} (\delta f_{PI}^{(i)})_{\mu\nu} \langle (e_I^{(i)})^\nu \| (e_P^{(i)})_\mu \rangle \qquad (8.65)$$

$$+ \sum_{\substack{P=1\\ \neq I}}^{m+1} \sum_{\substack{\mu\in I\\ \nu\in P}} (\delta f_{PI}^{(i)\dagger})_{\mu\nu} \langle (e_P^{(i)})^\nu \| (e_I^{(i)})_\mu \rangle .$$

Here, the notation $\langle r \| s \rangle$ is to indicate symbolically only that the basis function ϕ_r enters the expression antilinearly, while ϕ_s enters it linearly. It is not meant to imply that matrix elements of the type $[rs\|ut]$, given by (7.25b), can be written as the product of two simpler matrix elements. Combining (8.65) with the corresponding result for the sum over indices t,u, in the original expression (8.64), then leads to the result

$$\mathrm{tr}\, \delta R^{(i)} \delta G_i = \sum_{j=1}^{m+1} \sum_{\substack{P=1\\ \neq I}}^{m+1} \sum_{\substack{Q=1\\ \neq J}}^{m+1} \sum_{\mu\in P,\ \nu\in I} \sum_{\alpha\in J,\ \beta\in Q}$$

$$\Big((\delta f_{PI}^{(i)\dagger})_{\nu\mu} (\delta f_{QJ}^{(j)\dagger})_{\alpha\beta} \Big\{ [(e_P^{(i)})^\mu (e_I^{(i)})^\nu \| (e_Q^{(j)})^\beta (e_J^{(i)})^\alpha]$$

$$- a_{ij} [(e_P^{(i)})^\mu (e_J^{(j)})^\alpha \| (e_Q^{(j)})^\beta (e_I^{(i)})^\nu] \Big\}$$

$$+(\delta f_{PI}^{(i)\dagger})_{\nu\mu}(\delta f_{QJ}^{(j)})_{\beta\alpha}\Big\{\big[(e_P^{(i)})^\mu(e_I^{(i)})^\nu\big\|(e_J^{(j)})^\alpha(e_Q^{(j)})^\beta\big]$$

$$-a_{ij}\big[(e_P^{(i)})^\mu(e_Q^{(j)})^\beta\big\|(e_J^{(j)})^\alpha(e_I^{(i)})^\nu\big]\Big\}$$

$$+(\delta f_{PI}^{(i)})_{\mu\nu}(\delta f_{QJ}^{(j)\dagger})_{\alpha\beta}\Big\{\big[(e_I^{(i)})^\nu(e_P^{(i)})^\mu\big\|(e_Q^{(j)})^\beta(e_J^{(j)})^\alpha\big]$$

$$-a_{ij}\big[(e_I^{(i)})^\nu(e_J^{(j)})^\alpha\big\|(e_Q^{(j)})^\beta(e_P^{(i)})^\mu\big]\Big\}$$

$$+(\delta f_{PI}^{(i)})_{\mu\nu}(\delta f_{QJ}^{(i)})_{\beta\alpha}\Big\{\big[(e_I^{(i)})^\nu(e_P^{(i)})^\mu\big\|(e_J^{(j)})^\alpha(e_Q^{(j)})^\beta\big]$$

$$-a_{ij}\big[(e_I^{(i)})^\nu(e_Q^{(j)})^\beta\big\|(e_J^{(j)})^\alpha(e_P^{(i)})^\mu\big]\Big\}\Big)\,. \tag{8.66}$$

Each of the four terms here consists of a Coulomb and exchange integral combination, evaluated in the particular non-orthonormal molecular orbital basis given by (8.61). The contributions of the second term in (8.53) to the second derivatives of E are easily obtained from (8.66).

Consider now the first term of eq. (8.53). A considerable amount of algebraic manipulation is required to obtain the second order terms in compact form. The final result is

$$\mathrm{tr}\,\delta^{(2)}R^{(i)}F^{(i)} = -\mathrm{tr}\sum_{\substack{P,Q=1\\ \neq I}}^{m+1}\delta f_{PI}^{(i)\dagger}(SR^{(i)})_{PI}\delta f_{QI}^{(i)\dagger}F^{(i)}_{(\tilde{e}_Q^{(i)})(\tilde{e}_I^{(i)})}$$

$$-\mathrm{tr}\sum_{\substack{P,Q=1\\ \neq I}}^{m+1}\Big\{\delta f_{PI}^{(i)\dagger}[S(1-R^{(i)}S]_{PQ}\delta f_{QI}^{(i)}F^{(i)}_{(\tilde{e}_I^{(i)})(\tilde{e}_I^{(i)})}$$

$$+\,\delta f_{PI}^{(i)}R_{II}^{(i)}\delta f_{QI}^{(i)\dagger}F^{(i)}_{(\tilde{e}_Q^{(i)})(\tilde{e}_P^{(i)})}\Big\}$$

$$- \mathrm{tr} \sum_{\substack{P,Q=1 \\ \neq I}}^{m+1} \delta f_{PI}^{(i)} (R^{(i)}S)_{IQ} \delta f_{QI}^{(i)} F^{(i)} (e_I^{(i)})(e_P^{(i)})^{\bullet} \tag{8.67}$$

This equation contains no terms involving products of matrices $\delta f_{PI}^{(i)}$ with different values of i. Thus, the first term of (8.53) gives contributions only to second derivatives of the energy with respect to variables referring to the same shell.

The complete second derivatives of the energy can now be written down by combining eqs. (8.66) and (8.67), and incorporating constant factors and occupation numbers where indicated by (8.53). In all, there are only six different formulas (of which two pairs are complex conjugates of each other):

$$\frac{1}{2}\frac{\partial^2 E}{\partial (f_{PI}^{(i)*})_{\mu\nu} \partial (f_{QI}^{(i)*})_{\alpha\beta}} = \frac{-\nu_i}{2} \left[(SR^{(i)})_{\mu\beta} \frac{\partial E}{\partial (f_{QI}^{(i)*})_{\alpha\nu}} + (SR^{(i)})_{\alpha\nu} \frac{\partial E}{\partial (f_{PI}^{(i)*})_{\mu\beta}} \right]$$

$$+ \frac{\nu_i^2}{2} \left\{ [(e_P^{(i)})^{\mu} (e_I^{(i)})^{\nu} \| (e_Q^{(i)})^{\alpha} (e_I^{(i)})^{\beta}] - a_{ii} [(e_P^{(i)})^{\mu} (e_I^{(i)})^{\beta} \| (e_Q^{(i)})^{\alpha} (e_I^{(i)})^{\nu}] \right\},$$

$$\frac{1}{2}\frac{\partial^2 E}{\partial (f_{PI}^{(i)})_{\mu\nu} \partial (f_{QI}^{(i)})_{\alpha\beta}} = - \frac{\nu_i}{2} \left[(R^{(i)}S)_{\nu\alpha} \frac{\partial E}{\partial (f_{PI}^{(i)})_{\mu\beta}} + (R^{(i)}S)_{\beta\mu} \frac{\partial E}{\partial (f_{QI}^{(i)})_{\alpha\nu}} \right]$$

$$+\frac{\nu_i^2}{2}\Big\{[(e_I^{(i)})^\nu(e_P^{(i)})^\mu\,\|\,(e_I^{(i)})^\beta(e_Q^{(i)})^\alpha]$$

$$-a_{ii}[(e_I^{(i)})^\nu(e_Q^{(i)})^\alpha\,\|\,(e_I^{(i)})^\beta(e_P^{(i)})^\mu]\Big\},$$

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial(f_{PI}^{(i)*})_{\mu\nu}\partial(f_{QI}^{(i)})_{\alpha\beta}} = \frac{\nu_i}{2}\Big[-[S(1-R^{(i)}S]_{\mu\alpha}F^{(i)}_{(e_I^{(i)})^\beta(e_I^{(i)})^\nu}$$

$$+R^{(i)}_{\beta\nu}F^{(i)}_{(e_P^{(i)})^\mu(e_Q^{(i)})^\alpha}\Big]$$

$$+\frac{\nu_i^2}{2}\Big\{[(e_P^{(i)})^\mu(e_I^{(i)})^\nu\,\|\,(e_I^{(i)})^\beta(e_Q^{(i)})^\alpha]$$

$$-a_{ij}[(e_P^{(i)})^\mu(e_I^{(i)})^\alpha\,\|\,(e_I^{(i)})^\beta(e_Q^{(i)})^\nu]\Big\},$$

(8.68)

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial(f_{PI}^{(i)*})_{\mu\nu}\partial(f_{QJ}^{(j)*})_{\alpha\gamma}} = -\frac{\nu_i\nu_j}{2}\Big\{[(e_P^{(i)})^\mu(e_I^{(i)})^\nu\,\|\,(e_Q^{(j)})^\alpha(e_J^{(j)})^$$

$$-a_{ij}[(e_P^{(i)})^\mu(e_J^{(j)})^\gamma\,\|\,(e_Q^{(j)})^\alpha(e_I^{(i)})^\nu]\Big\},$$

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial(f_{PI}^{(i)})_{\mu\nu}\partial(f_{QJ}^{(j)})_{\alpha\gamma}} = -\frac{\nu_i\nu_j}{2}\Big\{[(e_I^{(i)})^\nu(e_P^{(i)})^\mu\,\|\,(e_J^{(j)})^\gamma(e_Q^{(j)})^\alpha]$$

$$-a_{ij}[(e_I^{(i)})^\nu(e_Q^{(j)})^\alpha\,\|\,(e_J^{(j)})^\gamma(e_P^{(i)})^\mu]\Big\},$$

$$\tfrac{1}{2}\frac{\partial^2 E}{\partial(f_{PI}^{(i)*})_{\mu\nu}\partial(f_{QJ}^{(j)})_{\alpha\gamma}} = \frac{\nu_i\nu_j}{2}\Big\{[(e_P^{(i)})^\mu(e_I^{(i)})^\nu\,\|\,(e_J^{(j)})^\gamma(e_Q^{(j)})^\alpha]$$

$$-a_{ij}[(e_P^{(i)})^\mu(e_Q^{(j)})^\alpha\,\|\,(e_J^{(j)})^\gamma(e_I^{(i)})^\nu]\Big\}.$$

In all these formulas, the convention $\mu\in P$, $\nu,\beta\in I$, $\alpha\in Q$, and $\gamma\in J$, is implied.

8.4.d Incorporation of the Intershell Orthogonality Constraints

The intershell orthogonality constraints on the lcao coefficients are given by eq. (8.47). The equivalent expressions in terms of the $f_{PI}^{(i)}$ are

$$g_{IJ}^{(ij)} = \sum_{K,L=1}^{m+1} f_{KI}^{(i)\dagger} S_{KL} f_{LJ}^{(j)} = 0, \quad i \neq j, \tag{8.69}$$

analogous to eqs. (4.58). Only half of these equations are unique, the other half being their adjoints.

There are two ways to incorporate these constraints into the theory. The constraint equations, (8.69), can be used to explicitly eliminate an appropriate number of elements of the $f_{PI}^{(i)}$ occurring in the energy functional. The derivatives of the energy with respect to the remaining unconstrained variables are then obtained from eqs. (8.62) by a simple application of the chain rule. The advantage of using this method to handle the intershell constraints is that the energy and its derivatives are then expressed in terms of a <u>minimum</u> number of unconstrained and non-redundant variables. The resulting formalism is suitable to use with true minimization techniques, such as the conjugate gradient method with variations discussed previously (section 8.3.b). Such procedures would be reliable, since divergence could not occur, and efficient, as long as the number of shells is small. As the number of shells increases, the intershell constraint equations become considerably more complicated (see Appendix 2) and the additional cost of calculating

the energy derivatives with respect to the independent variables may soon offset the other advantages. Another problem here is that the elimination procedure is not easily automated for use with an arbitrary number of shells. This approach is illustrated in detail in section 8.4.e for a two shell system.

A second approach to incorporating the intershell constraints into the theory is to consider the $\sum_{I=2}^{m} n_I \sum_{J=1}^{I-1} n_J$ unique constraint equations, (8.69), and the $\sum_{I=1}^{m} n_I \sum_{J=I+1}^{m+1} n_J$ independent equations expressing the vanishing of an appropriate set of the same number of first derivatives of the energy, as a system of $\sum_{I=1}^{m} n_I(n - n_I)$ simultaneous nonlinear equations for the elements of the $f_{PI}^{(i)}$, (P = 1, ..., m+1, P≠I; i = 1, ..., m). The derivatives of the $g_{IJ}^{(ij)}$, eq. (8.69), are

$$\frac{\partial(g_{IJ}^{(ij)})_{rs}}{\partial(f_{ML}^{(l)*})_{pq}} = \delta_{il}\delta_{qr}\delta_{LI}\left(\sum_{K=1}^{m+1} S_{MK}f_{KJ}^{(j)}\right)_{ps}, \qquad (8.70a)$$

and

$$\frac{\partial(g_{IJ}^{(ij)})_{rs}}{\partial(f_{ML}^{(l)})_{pq}} = \delta_{jl}\delta_{LJ}\delta_{sq}\left(\sum_{K=1}^{m+1} f_{KI}^{(i)\dagger}S_{KM}\right)_{rp}. \qquad (8.70b)$$

With these formulas, and eqs. (8.62), the Jacobian matrix for the complete system can be constructed, and the $f_{PI}^{(i)}$ can be determined iteratively, using one of several methods (for example, the Newton-Raphson equations).

The advantage of using this approach is that the energy

derivatives, (8.62), and the derivatives, (8.70), of the constraints, can be used without further modification, and can be calculated automatically for systems involving an arbitrary number of shells. The calculation now involves twice as many variables as in the first approach. The large number of variables may preclude the use of the full Newton-Raphson equations, making it necessary to develop linearly convergent approximations to them which are more efficient overall, much as was done in chapter 5, in a different context.[7] These methods are not descent methods, and therefore, will not necessarily yield an energy minimum at all times. Nevertheless, for systems involving a large number of shells, this would appear to be the approach of choice.

---

[7]The situation is admittedly greatly complicated here by the presence of the large number of two-electron integrals (which must be transformed to a new molecular orbital basis in each iteration) entering the second derivatives of the energy. They would have to be partially or totally neglected, or else approximated in some manner if a computationally efficient algorithm based on the Newton-Raphson equations is to result. An alternative procedure would be to use a method not requiring these second derivatives (for example, a generalization of the secant method for the solution of a single nonlinear equation). No information on the performance of such methods has yet been obtained, however.

8.4.e Example -- The Two Shell System

As an illustration of the general formalism just described, formulas applicable to a two shell system are given here explicitly. A 3 x 3 partitioning must be used in this case.

The variables entering the calculation are

$$\hat{T}_1^{(1)} = \begin{bmatrix} 1_1 \\ f_{21}^{(1)} \\ f_{31}^{(1)} \end{bmatrix}, \qquad \hat{T}_2^{(2)} = \begin{bmatrix} f_{12}^{(2)} \\ 1_2 \\ f_{32}^{(2)} \end{bmatrix}, \tag{8.71}$$

where the occupied orbitals for the two shells are written,

$$X_1^{(1)} = \hat{T}_1^{(1)} X_{11}^{(1)}, \qquad X_2^{(2)} = \hat{T}_2^{(2)} X_{22}^{(2)}. \tag{8.72}$$

The projection operators onto these two occupied spaces are given explicitly by

$$R^{(1)} = \begin{bmatrix} g_1^{(1)-1} & g_1^{(1)-1} f_{21}^{(1)\dagger} & g_1^{(1)-1} f_{31}^{(1)\dagger} \\ f_{21}^{(1)} g_1^{(1)-1} & f_{21}^{(1)} g_1^{(1)-1} f_{21}^{(1)\dagger} & f_{21}^{(1)} g_1^{(1)-1} f_{31}^{(1)\dagger} \\ f_{31}^{(1)} g_1^{(1)-1} & f_{31}^{(1)} g_1^{(1)-1} f_{21}^{(1)\dagger} & f_{31}^{(1)} g_1^{(1)-1} f_{31}^{(1)\dagger} \end{bmatrix}, \tag{8.73}$$

and

$$R^{(2)} = \begin{bmatrix} f_{12}^{(2)} g_2^{(2)-1} f_{12}^{(2)\dagger} & f_{12}^{(2)} g_2^{(2)-1} & f_{12}^{(2)} g_2^{(2)-1} f_{32}^{(2)\dagger} \\ g_2^{(2)-1} f_{12}^{(2)\dagger} & g_2^{(2)-1} & g_2^{(2)-1} f_{32}^{(2)\dagger} \\ f_{32}^{(2)} g_2^{(2)-1} f_{12}^{(2)\dagger} & f_{32}^{(2)} g_2^{(2)-1} & f_{32}^{(2)} g_2^{(2)-1} f_{32}^{(2)\dagger} \end{bmatrix}. \tag{8.74}$$

In an orthonormal basis, one has,

$$g_1^{(1)} = 1_1 + f_{21}^{(1)\dagger} f_{21}^{(1)} + f_{31}^{(1)\dagger} f_{31}^{(1)}, \tag{8.75a}$$

and

$$g_2^{(2)} = 1_2 + f_{12}^{(2)\dagger} f_{12}^{(2)} + f_{32}^{(2)\dagger} f_{32}^{(2)}. \tag{8.75b}$$

If the basis is non-orthonormal, explicit formulas for the $g_I^{(i)}$ are considerably lengthier, for example,

$$g_1^{(1)} = S_{11} + S_{12} f_{21}^{(1)} + S_{13} f_{31}^{(1)} + f_{21}^{(1)\dagger} S_{21} + f_{31}^{(1)\dagger} S_{31} + f_{21}^{(1)\dagger} S_{22} f_{21}^{(1)} + f_{21}^{(1)\dagger} S_{23} f_{31}^{(1)} + f_{31}^{(1)\dagger} S_{32} f_{21}^{(1)} + f_{31}^{(1)\dagger} S_{33} f_{31}^{(1)}, \tag{8.76}$$

and similarly for $g_2^{(2)}$.

For an orthonormal fixed basis, the non-orthonormal contragredient molecular orbital basis, (8.61), in terms of which the energy derivatives can be written very compactly, are given by

$$\tilde{e}_1^{(1)} = R_{,1}^{(1)} = \begin{bmatrix} 1_1 \\ f_{21}^{(1)} \\ f_{31}^{(1)} \end{bmatrix} g_1^{(1)^{-1}},$$

$$\tilde{e}_2^{(1)} = (1 - R^{(1)})_{,2} = \begin{bmatrix} -g_1^{(1)^{-1}} f^{(1)\dagger} \\ 1_2 - f_{21}^{(1)} g_1^{(1)^{-1}} f_{21}^{(1)\dagger} \\ -f_{31}^{(1)} g_1^{(1)^{-1}} f_{21}^{(1)\dagger} \end{bmatrix}, \tag{8.77a}$$

and

$$\tilde{e}_3^{(1)} = (1 - R^{(1)})_{,3} = \begin{bmatrix} -g_1^{(1)-1} f_{31}^{(1)\dagger} \\ -f_{21}^{(1)} g_1^{(1)-1} f_{31}^{(1)\dagger} \\ 1_3 - f_{31}^{(1)} g_1^{(1)-1} f_{31}^{(1)\dagger} \end{bmatrix} .$$

The expressions for the $\tilde{e}_K^{(2)}$ are analogous,

$$\tilde{e}_1^{(2)} = (1-R^{(2)})_{,1}, \quad \tilde{e}_2^{(2)} = R_{,2}^{(2)}, \quad \tilde{e}_3^{(2)} = (1-R^{(2)})_{,3}. \tag{8.77b}$$

For a non-orthonormal fixed basis, explicit expressions for the $\tilde{e}_K^{(i)}$ in terms of S and the $f_{PI}^{(i)}$ and $g_I^{(i)-1}$, (i = 1,2), are considerably lengthier. In the course of a calculation, the current projections $R^{(1)}$ and $R^{(2)}$ would always be known, so that the $e_K^{(i)}$ would be obtained directly from formulas like (8.77b), rather than being evaluated using formulas like (8.77a).

The vectors dual to the $\tilde{e}^{(1)}$ are given by

$$\underset{\sim}{e}^{(1)} = \begin{bmatrix} 1_1 & -f_{21}^{(1)\dagger} & -f_{31}^{(1)\dagger} \\ f_{21}^{(1)} & 1_2 & 0 \\ f_{31}^{(1)} & 0 & 1_3 \end{bmatrix} . \tag{8.78}$$

The scalar products of these vectors and the metric matrices with respect to which they are orthonormal are given by

$$\tilde{g}^{(1)} = \tilde{e}^{(1)\dagger}\tilde{e}^{(1)} = \tilde{e}^{(1)}\tilde{e}^{(1)\dagger} = \tilde{\Delta}^{(1)}$$

$$= \begin{bmatrix} g_1^{(1)-1} & 0 & 0 \\ 0 & 1_2 - f_{21}^{(1)} g_1^{(1)-1} f_{21}^{(1)\dagger} & -f_{21}^{(1)} g_1^{(1)-1} f_{31}^{(1)\dagger} \\ 0 & -f_{31}^{(1)} g_1^{(1)-1} f_{21}^{(1)\dagger} & 1_3 - f_{31}^{(1)} g_1^{(1)-1} f_{31}^{(1)\dagger} \end{bmatrix} , \quad (8.79)$$

and

$$\underset{\sim}{g}^{(1)} = \underset{\sim}{e}^{(1)\dagger} \underset{\sim}{e}^{(1)} = \underset{\sim}{e}^{(1)} \underset{\sim}{e}^{(1)\dagger} = \underset{\sim}{\Delta}^{(1)}$$

$$= \begin{bmatrix} g_1^{(1)} & 0 & 0 \\ 0 & 1_2 + f_{21}^{(1)} f_{21}^{(1)\dagger} & f_{21}^{(1)} f_{31}^{(1)\dagger} \\ 0 & f_{31}^{(1)} f_{21}^{(1)\dagger} & 1_3 + f_{31}^{(1)} f_{31}^{(1)\dagger} \end{bmatrix} , \quad (8.80)$$

with similar results for $\underset{\sim}{e}^{(2)}$ and $\tilde{e}^{(2)}$. Similar, but lengthier, results are obtained in a non-orthonormal fixed basis, but, in that case, $\tilde{\underset{\sim}{g}}^{(i)} \neq \tilde{\underset{\sim}{\Delta}}^{(i)}$, and $\underset{\sim}{g}^{(i)} \neq \underset{\sim}{\Delta}^{(i)}$, (i = 1,2), and thus, the number of formulas doubles. The similarities between eqs. (8.77) - (8.80), and the results given in section 2.1.d, applicable to a single shell system, are easily seen.

The formal first derivatives of the energy are

$$\begin{aligned} \frac{\partial E}{\partial (f_{21}^{(1)*})_{\sigma r}} &= \nu_1 F^{(1)}_{(e_2^{(1)})^\sigma (e_1^{(1)})^r} , \\ \frac{\partial E}{\partial (f_{31}^{(1)*})_{\alpha r}} &= \nu_1 F^{(1)}_{(e_3^{(1)})^\alpha (e_1^{(1)})^r} , \\ \frac{\partial E}{\partial (f_{12}^{(2)*})_{r\sigma}} &= \nu_2 F^{(2)}_{(e_1^{(2)})^r (e_2^{(2)})^\sigma} , \end{aligned} \quad (8.81)$$

and

$$\frac{\partial E}{\partial (f_{32}^{(2)*})_{\alpha\sigma}} = \nu_2 F^{(2)} (e_3^{(2)})^{\alpha} (e_2^{(2)})^{\sigma} \quad .$$

Formal second derivatives can be written down explicitly from eqs. (8.68). Even in this simple case, there are thirty-two different explicit second derivative formulas (E depends on $f_{21}^{(1)}$, $f_{31}^{(1)}$, $f_{12}^{(2)}$, and $f_{32}^{(2)}$, and their adjoints) neglecting those which are complex conjugates.

There is only one intershell constraint equation in this case. In an orthonormal basis, it is

$$g_{12}^{(12)} = f_{12}^{(2)} + f_{21}^{(2)\dagger} + f_{31}^{(1)\dagger} f_{32}^{(2)} = 0. \qquad (8.82)$$

This equation is easily used to obtain

$$f_{12}^{(2)} = -f_{21}^{(1)\dagger} - f_{31}^{(1)\dagger} f_{32}^{(2)} \quad , \qquad (8.83)$$

giving $f_{12}^{(2)}$ in terms of $f_{21}^{(1)}$, $f_{31}^{(1)}$, and $f_{32}^{(2)}$, whose elements can be used as a set of unconstrained and non-redundant variables, in terms of which the energy may be minimized. Equation (8.83) is unusually simple. For the next simplest case, a three shell system, there are three intershell constraint equations, which, while similar to (8.82), cannot be used to obtain three "dependent" blocks, $f_{PI}^{(i)}$, in terms of the remaining six "independent" blocks without introduction of an inverse matrix (see Appendix 2). In fact, for a two shell system, when the fixed basis is non-orthonormal, the intershell constraint becomes

$$g_{12}^{(12)} = (\hat{T}_1^{(1)\dagger} S \hat{T}_2^{(2)})_{12}$$

$$= S_{11}f_{12}^{(2)} + S_{12} + S_{13}f_{32}^{(2)} + f_{21}^{(1)\dagger}(S_{21}f_{12}^{(2)}+S_{22}+S_{23}f_{32}^{(2)})$$

$$+ f_{31}^{(1)\dagger}(S_{31}f_{12}^{(2)}+S_{32}+S_{33}f_{32}^{(2)}) = 0, \quad (8.84)$$

from which one obtains

$$f_{12}^{(2)} = -[S_{11} + f_{21}^{(1)\dagger}S_{21} + f_{31}^{(1)\dagger}S_{31}]^{-1} \quad (8.85a)$$

$$\times [S_{12}+f_{21}^{(1)\dagger}S_{22}+f_{31}^{(1)\dagger}S_{32}+(S_{13}+f_{31}^{(1)\dagger}S_{23}+f_{31}^{(1)\dagger}S_{33})f_{32}^{(2)}]$$

$$= -A^{-1}B. \quad (8.85b)$$

Not only is this expression considerably lengthier than (8.83), but the presence of the inverse matrix complicates the application of the chain rule, and leads to more complicated formulas for the energy derivatives with respect to the remaining independent variables. From (8.83), one obtains

$$\frac{\partial(f_{12}^{(2)})_{\sigma\rho}}{\partial(f_{21}^{(1)})_{\mu\nu}} = -\delta_{\mu\rho}\delta_{\sigma\nu}, \qquad \frac{\partial(f_{12}^{(2)})_{\sigma\rho}}{\partial(f_{31}^{(1)*})_{\mu\nu}} = -\delta_{\sigma\nu}(f_{32}^{(2)})_{\mu\rho},$$

$$\frac{\partial(f_{12}^{(2)})_{\sigma\rho}}{\partial(f_{31}^{(2)})_{\mu\nu}} = -\delta_{\nu\rho}(f_{31}^{(1)*})_{\mu\sigma}. \quad (8.86)$$

Combining these with eqs. (8.81) then yields,

$$\frac{\partial E}{\partial(f_{21}^{(1)*})_{\sigma r}} = \nu_1 F^{(1)}_{(e_2^{(1)})\sigma(e_1^{(1)})r} - \nu_2 F^{(2)}_{(e_2^{(2)})\sigma(e_1^{(2)})r},$$

$$\frac{\partial E}{\partial(f_{31}^{(1)*})_{\alpha r}} = \nu_1 F^{(1)}_{(e_3^{(1)})\alpha(e_1^{(1)})r} - \nu_2\left[f_{32}^{(2)}F^{(2)}_{e_2^{(2)\dagger}e_1^{(2)}}\right]_{\alpha r} \quad (8.87)$$

and

$$\frac{\partial E}{\partial (f_{32}^{(2)*})_{\alpha\sigma}} = \nu_2 F^{(2)}_{(e_3^{(2)})^\alpha (e_2^{(2)})^\sigma} = \nu_2 \left[f_{31}^{(1)} F^{(2)}_{e_1^{(2)\dagger} e_2^{(2)}} \right]_{\alpha\sigma} ,$$

which require little additional work once the derivatives in (8.81) are known. For a non-orthonormal basis, eqs. (8.85) can be used to obtain,

$$\frac{\partial (f_{12}^{(2)})_{r\sigma}}{\partial (f_{21}^{(1)*})_{\tau s}} = A^{-1}_{rs} [S_{21} A^{-1} B - S_{22} - S_{23} f_{32}^{(2)}]_{\tau\sigma},$$

$$\frac{\partial (f_{12}^{(2)})_{r\sigma}}{\partial (f_{31}^{(1)*})_{\alpha s}} = A^{-1}_{rs} [S_{31} A^{-1} B - S_{32} - S_{33} f_{32}^{(2)}]_{\alpha\sigma} , \quad (8.88)$$

and

$$\frac{\partial (f_{12}^{(2)})_{r\sigma}}{\partial (f_{32}^{(2)})_{\alpha\tau}} = -\delta_{\sigma\tau} [A^{-1} (f_{21}^{(1)\dagger} S_{23} + f_{31}^{(1)\dagger} S_{33})]_{r\alpha},$$

which lead to

$$\frac{\partial E}{\partial (f_{21}^{(1)*})_{\tau s}} = \nu_1 F^{(1)}_{(e_2^{(1)})^\tau (e_1^{(1)})^s}$$

$$+ \nu_2 [(S_{21} A^{-1} B - S_{22} - S_{23} f_{32}^{(2)}) F^{(2)}_{e_2^{(2)\dagger} e_1^{(2)}} A^{-1}]_{\tau s},$$

$$\frac{\partial E}{\partial (f_{31}^{(1)*})_{\alpha s}} = \nu_1 F^{(1)}_{(e_3^{(1)})^\alpha (e_1^{(1)})^s}$$

$$+ \nu_2 [(S_{31} A^{-1} B - S_{32} - S_{33} f_{32}^{(2)}) F^{(2)}_{e_2^{(2)\dagger} e_1^{(2)}} A^{-1}]_{\alpha s}$$

and

$$\frac{\partial E}{\partial (f_{32}^{(2)*})_{\alpha\tau}} = \nu_2 F^{(2)}_{(e_3^{(2)})^\alpha (e_2^{(2)})^\tau}$$

$$-\nu_2 [(S_{33} f_{31}^{(1)} + S_{32} f_{21}^{(1)})(A^\dagger)^{-1} F^{(2)}_{e_1^{(2)} \dagger e_2^{(2)}}]_{\alpha\tau}. \tag{8.89}$$

Both sets of equations, (8.87) and (8.89), are suitable for use with gradient minimization algorithms. Second derivatives of E with respect to elements of $f_{21}^{(1)}$, $f_{31}^{(1)}$, and $f_{32}^{(2)}$, and their adjoints are obtained in a similar way. The greater complexity of the formulas (8.89) compared to those in (8.87) is not of much concern here, since in an actual calculation, one would expect to carry out the energy minimization in a molecular orbital basis in which S = 1 (see section 8.3.b).

In this case, the stationary points of the energy can also be determined by solving the system of $2n_1n_2 + n_1n_3 + n_2n_3$ (in general complex) simultaneous nonlinear equations given by

$$F^{(1)}_{\tilde{e}_2^{(1)} \dagger \tilde{e}_1^{(1)}} = 0,$$

$$F^{(1)}_{\tilde{e}_3^{(1)} \dagger \tilde{e}_1^{(1)}} = 0, \tag{8.89}$$

$$F^{(2)}_{\tilde{e}_3^{(2)} \dagger \tilde{e}_2^{(2)}} = 0,$$

and

$$g_{12}^{(12)} = 0.$$

BIBLIOGRAPHY

Claxton, T. A., and Smith, N. A., Theoret. Chim. Acta, 22, (1971), 399.

Clementi, E., J. Chem. Phys., 38, (1963), 2686.

Coope, J. A. R., Molecular Physics, 18, (1970), 571.

-- Ph. D. Thesis, Oxford University, (1956).

Coulson, C., and Longuet-Higgins, H. C., Proc. Roy. Soc., A191, (1947), 39.

Daniel, J. W., Numerische Mathematik, 10, (1965), 125.

Davidson, E. R., J. Comp. Phys., 17, (1975), 87.

DeVries, E., Fortschritte der Physik, 18, (1970), 149.

Eriksen, E., Phys. Rev., 111, (1958), 1011.

Feler, M. G., J. Comp. Phys., 14, (1974), 341.

Fletcher, R., Molecular Physics, 19, (1970), 55.

Fletcher, R., and Reeves, C.M., Computer Journal, 7, (1964), 149.

Foldy, L. L., and Wouthuysen, S. A., Phys. Rev., 78, (1950), 29.

Friedrichs, K. O., Perturbation of Spectra in Hilbert Space, (1965), American Mathematical Society, Prov., R. I.

Fujimoto, H. et. al., J. Phys. Chem., 78, (1974), 1167.

Garton, D., and Sutcliffe, B. T., Theoretical Chemistry, Vol. 1. Quantum Chemistry, Specialist Periodical Reports, Chemical Society, London, (1974).

Imamura, A., Molecular Physics, 15, (1968), 225.

Kari, R., and Sutcliffe, B. T., Chem. Phys. Lett., 7, (1970), 149.

-- International Journal of Quantum Chemistry, 7, (1973), 459.

Kato, T., Perturbation Theory for Linear Operators, Springer-Verlag, New York, (1966).

Klein, D., J. Chem. Phys., 61, (1974), 786.

Langhoff, P. W., Karplus, M., and Hurst, R. P., J. Chem. Phys., 44, (1966), 505.

Libit, L., and Hoffmann, R., Journal of the American Chemical Society, 96, (1974), 1370.

Lowdin, P. O., J. Math. Phys., 3, (1962), 969.

-- International Journal of Quantum Chemistry, 2, (1968), 867.

-- Advances in Quantum Chemistry, 5, (1970), 185.

Lowdin, P. O., and Goscinski, O., International Journal of Quantum Chemistry, 5, (1971), 685.

McWeeny, R., Phys. Rev., 126, (1962), 1028.

Morpurgo, G., Nuovo Cimento, 25, (1960), 624.

Musher, J., J. Chem. Phys., 46, (1967), 369.

Nesbet, R. K., J. Chem. Phys., 43, (1965), 311.

Okubo, S., Prog. Theor. Phys., 12, (1954), 102.

Pople, J. A., and Beveridge, D. L., Approximate Molecular Orbital Theory, McGraw-Hill, Toronto, (1970).

Primas, H., Helv. Phys. Acta, 34, (1961), 331.

-- Rev. Mod. Phys., 35, (1963), 710.

Rall, L. B., Computational Solution of Nonlinear Operator Equations, Wiley, New York, (1969).

Ralston, A., A First Course in Numerical Analysis, McGraw-Hill, Toronto, (1965).

Riesz, F., and Sz.-Nagy, B., Functional Analysis, Frederick Ungar, (1955).

Roos, B., "The Configuration Interaction Method" in Computational Techniques in Quantum Chemistry and Molecular Physics, (G. H. F. Diercksen, B. T. Sutcliffe, and A. Veillard, eds.), Reidel, Boston, (1975).

Roothaan, C. C. J., Rev. Mod. Phys., 23, (1951), 69.

Rosenberg, M., and Martino, F., J. Chem. Phys., 63, (1975), 5354.

Rutherford, D. E., Roy. Soc. of Edinburgh, Proc. A, 63, (1949/52), 232.

Shavitt, I., J. Comp. Phys., 6, (1970), 124.

Shavitt, I., et. al., J. Comp. Phys., 11, (1973), 90.

Sutcliffe, B. T., Theoret. Chim. Acta, 33, (1974), 201.

-- Theoret. Chim. Acta, 39, (1975), 93.

Sz.-Nagy, B., Comment. Math. Helvet., 19, (1946/47), 347.

Tami, S., Prog. Theor. Phys., 12, (1954), 104.

Traub, J. F., Iterative Methods for the Solution of Equations, Prentice Hall, Englewood Cliffs, N. J., (1964).

Van Vleck, J. H., Phys. Rev., 33, (1929), 467.

APPENDICES

"Humpty Dumpty looked doubtful. 'I'd rather see that done on paper,' he said."
(Through the Looking Glass, Lewis Carroll)

APPENDIX 1

Proofs of Alternative Formulas -- 2 x 2 Partitioning

This appendix outlines some of the manipulations necessary to establish a number of inter-relations which have been quoted in section 3.1.

Consider first the orthonormal case. The relationship between the two sets of effective operators (3.1) and (3.2) is easily established. From the definition

$$\hat{H}_A^{(2)} = g_A^{-1} G_A,$$

one obtains,

$$\begin{aligned}\hat{H}_A^{(2)} &= g_A^{-1}[\, H_{AA} + H_{AB}f + f^\dagger(H_{BA} + H_{BB}f)] \\ &= g_A^{-1}[\hat{H}_A^{(1)} + f^\dagger(D^{(1)}(f) + f\hat{H}_A^{(1)})] \\ &= \hat{H}_A^{(1)} + g_A^{-1}f^\dagger D^{(1)}(f).\end{aligned}$$

A similar procedure establishes the relation (3.3b) between $\hat{H}_B^{(2)}$ and $\hat{H}_B^{(1)}$. To establish eq. (3.6), the result (3.3a) is substituted into (3.5), and the "pull-through" relations, (2.32), used. This yields

$$\begin{aligned}D^{(2)}(f) &= H_{BA} + H_{BB}f - f\hat{H}_A^{(2)} \\ &= (H_{BA} + H_{BB}f - f\hat{H}_A^{(1)}) - fg_A^{-1}f^\dagger D^{(1)}(f) \\ &= (1_B - fg_A^{-1}f^\dagger)D^{(1)}(f) \\ &= g_B^{-1}D^{(1)}(f).\end{aligned}$$

The condition that $\hat{T}^{-1}H\hat{T}$ be block diagonal is easily determined in a direct manner. The inverse of $\hat{T}$ is

$$\hat{T}^{-1} = \begin{bmatrix} g_A^{-1} & g_A^{-1}f^\dagger \\ -fg_A^{-1} & 1_B - fg_A^{-1}f^\dagger \end{bmatrix} .$$

Matrix multiplcation, followed by use of the "pull-through" relations, (2.32), then establishes that the off-diagonal blocks of $\hat{T}^{-1}H\hat{T}$ are given by $D^{(2)}(f) = g_B^{-1}D^{(1)}(f)$.

Before deriving eqs. (3.11) - (3.15), applying in the case of a nonorthonormal basis, it is necessary to examine the orthonormality condition, (2.101b), in more detail. The blocks of the matrix g are

$$g_A = S_{AA} + S_{AB}f + f^\dagger S_{BA} + f^\dagger S_{BB}f,$$

$$g_B = S_{BB} + S_{BA}h + h^\dagger S_{AB} + h^\dagger S_{AA}h,$$

and

$$g_{BA} = h^\dagger S_{AA} + h^\dagger S_{AB}f + S_{BA} + S_{BB}f = g_{AB}^\dagger . \qquad (A1.1)$$

Thus, one has,

$$g_A = (1_A - f^\dagger h^\dagger)\hat{S}_A + f^\dagger g_{BA} , \qquad (A1.2)$$

and

$$g_B = (1_B - h^\dagger f^\dagger)\hat{S}_B + h^\dagger g_{AB} . \qquad (A1.3)$$

Here, and throughout this appendix, the notation

$$\hat{S}_A = S_{AA} + S_{AB}f , \qquad \hat{S}_B = S_{BB} + S_{BA}h,$$

established in eq. (2.112), is used to simplify the equations. From (A1.2) and (A1.3), one obtains,

$$g_A\hat{S}_A^{-1} = (1_A - f^\dagger h^\dagger) + f^\dagger g_{BA}\hat{S}_A^{-1} , \qquad (A1.4)$$

and

$$g_B\hat{S}_B^{-1} = (1_B - h^\dagger f^\dagger) + h^\dagger g_{AB}\hat{S}_B^{-1} . \qquad (A1.5)$$

From these last relations, two generalizations of the "pull-through" relations in the orthonormal case, can be derived. They are

$$h^\dagger g_A \hat{S}_A^{-1} = g_A \hat{S}_B^{-1} h^\dagger + h^\dagger f^\dagger g_{BA} \hat{S}_A^{-1} - h^\dagger g_{AB} \hat{S}_B^{-1} h^\dagger , \quad \text{(A1.6)}$$

and

$$g_A \hat{S}_A^{-1} f^\dagger = f^\dagger g_B \hat{S}_B^{-1} + f^\dagger h^\dagger g_{AB} \hat{S}_B^{-1} - f^\dagger g_{BA} \hat{S}_A^{-1} f^\dagger . \quad \text{(A1.7)}$$

The last two terms vanish in each if $g_{BA} = 0$, leaving the simpler expressions

$$h^\dagger g_A \hat{S}_A^{-1} = g_B \hat{S}_B^{-1} h^\dagger , \quad \text{(A1.8)}$$

and

$$g_A \hat{S}_A^{-1} f^\dagger = f^\dagger g_B \hat{S}_B^{-1} . \quad \text{(A1.9)}$$

Two other relations will be useful below in deriving (3.15). They are,

$$h^\dagger g_A = -(1_B - h^\dagger f^\dagger)(S_{BA} + S_{BB} f) + g_{BA} , \quad \text{(A1.10)}$$

and

$$f^\dagger g_B = -(1_A - f^\dagger h^\dagger)(S_{AA} h + S_{AB}) + g_{AB} . \quad \text{(A1.11)}$$

The first, (A1.10), is obtained as follows,

$$\begin{aligned} h^\dagger g_A^{-1} &= (1_B - h^\dagger f^\dagger) h^\dagger \hat{S}_A + h^\dagger f^\dagger g_A \\ &= -(1_B - h^\dagger f^\dagger)(S_{BA} + S_{BB} f - g_{BA}) + h^\dagger f^\dagger g_{BA} \\ &= -(1_B - h^\dagger f^\dagger)(S_{BA} + S_{BB} f) + g_{BA} . \end{aligned}$$

The first line here is obtained by premultiplying (A1.2) by $h^\dagger$. The second line then follows directly from the definition, (A1.1), of g_{BA}. The relation (A1.11) is derived analogously.

by first premultiplying (A1.3) by $f^\dagger$, and then using (A1.1) for $g_{AB} = g_{BA}^\dagger$. A number of other relations similar to these could be derived here also, but these are sufficient for what follows.

To establish the relationship between the operators $\hat{H}_A^{(1)}$ and $\hat{H}_A^{(2)}$, we proceed as follows. First, from (A1.1),

$$h^\dagger = [g_{BA} - (S_{BA} + S_{BB}f)]\hat{S}_A^{-1} ,$$

and thus, eq. (3.13) yields

$$H_{BA} + H_{BB}f = -[h^\dagger - g_{BA}\hat{S}_A^{-1}](H_{AA} + H_{AB}f) + D^{(1)}(f). \quad (A1.12)$$

Then, using (3.8), one obtains

$$\begin{aligned}
\hat{H}_A^{(2)} &= g_A^{-1}G_A \\
&= g_A^{-1}[H_{AA} + H_{AB}f + f^\dagger(H_{BA} + H_{BB}f)] \\
&= g_A^{-1}[H_{AA} + H_{AB}f - f^\dagger[(h^\dagger - g_{BA}\hat{S}_A^{-1})(H_{AA} + H_{AB}f) \\
&\qquad\qquad - D^{(1)}(f)]] \\
&= g_A^{-1}(1_A - f^\dagger h^\dagger)(H_{AA} + H_{AB}f) + g_A^{-1}f^\dagger D^{(1)}(f) \\
&\qquad + g_A^{-1}f^\dagger g_{BA}\hat{H}_A^{(1)}.
\end{aligned}$$

But, from eq. (A1.2),

$$(1_A - f^\dagger h^\dagger) = (g_A - f^\dagger g_{BA})\hat{S}_A^{-1},$$

and thus,

$$\begin{aligned}
\hat{H}_A^{(2)} &= \hat{S}_A^{-1}(H_{AA} + H_{AB}f) + g_A^{-1}f^\dagger[D^{(1)} - g_{BA}\hat{S}_A^{-1}(H_{AA} + H_{AB}f) \\
&\qquad\qquad + g_{BA}\hat{H}_A^{(1)}] \\
&= \hat{H}_A^{(1)} + g_A^{-1}f^\dagger D^{(1)}(f),
\end{aligned}$$

establishing (3.11). Equation (3.12) is obtained in an analogous manner.

A number of approaches can be used to obtain eqs. (3.15), one of which is as follows:

$$\begin{aligned} D^{(2)} &= H_{BA} + H_{BB}f - (S_{BA} + S_{BB}f)\, g_A^{-1} G_A \\ &= H_{BA} + H_{BB}f + (1_B - h^\dagger f^\dagger)^{-1}(h^\dagger - g_{BA}g_A^{-1})G_A \\ &= (1_B - h^\dagger f^\dagger)^{-1}[(1_B - h^\dagger f^\dagger)(H_{BA}+H_{BB}f) + h^\dagger G_A - g_{BA}g_A^{-1}G_A] \\ &= (1_B - h^\dagger f^\dagger)^{-1}[D^{(1)} - g_{BA}(\hat{H}_A^{(2)} - \hat{H}_A^{(1)})] \\ &= (1_B - h^\dagger f^\dagger)^{-1}(1_B - g_{BA}g_A^{-1}f^\dagger)D^{(1)}(f). \end{aligned} \qquad \text{(A1.13)}$$

The transition from the first line to the second is effected using the relation (A1.10), and the remainder by use of previous definitions, including (A1.12) above. From (A1.3), one has

$$(1_B - h^\dagger f^\dagger)^{-1} = (S_{BB} + S_{BA}h)(g_B + h^\dagger g_{AB})^{-1} ,$$

which, upon substitution into (A1.12), gives eq. (3.15a). Equation (3.15b) follows directly from (3.15a) by simply dropping the terms in g_{AB} and g_{BA}. The easiest way to obtain (3.15c) is to begin again from the definition, (3.14), of $D^{(2)}(f)$,

$$\begin{aligned} D^{(2)}(f) &= H_{BA} + H_{BB}f - (S_{BA} + S_{BB}f)\hat{H}_A^{(2)} \\ &= H_{BA} + H_{BB}f - (S_{BA} + S_{BB}f)[\hat{H}_A^{(1)} + g_A^{-1}f^\dagger D^{(1)}(f)] \\ &= [1_B - (S_{BA} + S_{BB}f)g_A^{-1}f^\dagger]D^{(1)}(f). \end{aligned}$$

This derivation is analogous to that establishing (3.6) in the case of an orthonormal basis.

For an orthonormal basis, it was found that the condition $D^{(2)}(f) = 0$ could be obtained by requiring that the product $\hat{T}^{-1}H\hat{T}$ be block diagonal. For a nonorthonormal basis, the

corresponding condition is to require the product $(S\hat{T})^{-1}H\hat{T}$ to be block diagonal. But writing $\hat{T}^{\dagger}S\hat{T} = g$, one has,

$$(S\hat{T})^{-1} = g^{-1}\hat{T}^{\dagger},$$

and

$$(S\hat{T})^{-1}H\hat{T} = g^{-1}\hat{T}^{\dagger}H\hat{T} = g^{-1}G.$$

Thus, when g is block diagonal, $(S\hat{T})^{-1}H\hat{T}$ has diagonal blocks $\hat{H}_A^{(2)}$ and $\hat{H}_B^{(2)}$, and off-diagonal blocks,

$$D_{BA}^{(2)'} = [(S\hat{T})^{-1}H\hat{T}]_{BA} = g_B^{-1}D^{(1)},$$

and

$$D_{AB}^{(2)'} = [(S\hat{T})^{-1}H\hat{T}]_{AB} = g_A^{-1}D^{(1)\dagger}.$$

While this is of the form of eq. (3.6), $D_{BA}^{(2)'}$ can not be written in terms of $\hat{H}_A^{(2)}$ as in eq. (3.14), unlike the analogous result in an orthonormal basis.

APPENDIX 2

The 3 x 3 and 4 x 4 Case -- Orthonormal Basis

To illustrate some of the complications which arise in a multiple partitioning formalism, a number of explicit formulas are given here for quantities arising out of a 3 x 3 and a 4 x 4 partitioning formalism.

In the 3 x 3 partitioning formalism, the three non-self-adjoint effective operators given by (4.17) are

$$\hat{H}_1 = H_{11} + H_{12}f_{21} + H_{13}f_{31} ,$$

$$\hat{H}_2 = H_{22} + H_{21}f_{12} + H_{23}f_{32} , \qquad \text{(A2.1)}$$

and

$$\hat{H}_3 = H_{33} + H_{31}f_{13} + H_{32}f_{23} ,$$

each containing only one extra term compared to the 2 x 2 case. A considerably greater increase in complexity occurs when m=3 in the defining conditions on the f_{JI}, given by (4.18). There are now six matrix block equations with six terms each, in place of the two block equations with four terms each, as in the 2 x 2 case. They are,

$$D_{21} = H_{21} + H_{22}f_{21} + H_{23}f_{31} - f_{21}(H_{11} + H_{12}f_{21} + H_{13}f_{31}) = 0,$$

$$D_{31} = H_{31} + H_{32}f_{21} + H_{33}f_{31} - f_{31}(H_{11} + H_{12}f_{21} + H_{13}f_{31}) = 0,$$

$$D_{12} = H_{12} + H_{11}f_{12} + H_{13}f_{32} - f_{12}(H_{22} + H_{21}f_{12} + H_{23}f_{32}) = 0,$$

$$D_{32} = H_{32} + H_{31}f_{12} + H_{33}f_{32} - f_{32}(H_{22} + H_{21}f_{12} + H_{23}f_{32}) = 0,$$

$$D_{13} = H_{13} + H_{11}f_{13} + H_{12}f_{23} - f_{13}(H_{33} + H_{31}f_{13} + H_{32}f_{23}) = 0,$$

and,

$$D_{23} = H_{23} + H_{23}f_{13} + H_{22}f_{23} - f_{23}(H_{33} + H_{31}f_{13} + H_{32}f_{23}) = 0. \tag{A2.2}$$

The orthogonality condition in (4.1) gives rise to three matrix block equations here,

$$g_{12} = f_{12} + f_{21}^{\dagger} + f_{31}^{\dagger}f_{32} = 0,$$
$$g_{13} = f_{13} + f_{21}^{\dagger}f_{23} + f_{31}^{\dagger} = 0, \tag{A2.3}$$

and,

$$g_{23} = f_{12}^{\dagger}f_{13} + f_{23} + f_{32}^{\dagger} = 0.$$

These equations can be used to eliminate f_{12}, f_{13} and f_{23} from the remainder of the formalism, in favour of f_{21}, f_{31} and f_{32}. In fact, it is not difficult to show that

$$f_{12} = -f_{21}^{\dagger} - f_{31}^{\dagger}f_{32},$$
$$f_{23} = (1_1 - f_{12}^{\dagger}f_{21}^{\dagger})^{-1}(-f_{32}^{\dagger} + f_{12}^{\dagger}f_{31}^{\dagger})$$
$$= -[1_1 + (f_{21}^{\dagger} + f_{31}^{\dagger}f_{32})f_{21}]^{-1}[f_{32} + (f_{21}^{\dagger} + f_{31}^{\dagger}f_{32})f_{31}^{\dagger}], \tag{A2.4}$$

and

$$f_{13} = -f_{31}^{\dagger} - f_{21}^{\dagger}f_{23}$$
$$= -f_{31}^{\dagger} + f_{21}^{\dagger}[1_1 + (f_{21}^{\dagger} + f_{31}^{\dagger}f_{32})f_{21}^{\dagger}]^{-1}[f_{32} + (f_{21}^{\dagger} + f_{31}^{\dagger}f_{32})f_{31}^{\dagger}].$$

The problems involved in eliminating f_{12}, f_{23} and f_{13} from eqs. (A2.2), and thereby reducing the number of block equations which must be considered from six to three, are clear from these equations. It would be quite difficult to derive efficient procedures to solve such a system, because of the generally complex dependence of the remaining three block equations on the elements of f_{21}, f_{31} and f_{32}.

For a 4 x 4 partitioning, the orthogonality conditions, $g_{IJ} = 0$, give the following six unique matrix block equations,

$$g_{12} = f_{12} + f_{21}^{\dagger} + f_{31}^{\dagger} f_{32} + f_{41}^{\dagger} f_{42} = 0, \tag{A2.5}$$

$$\begin{aligned} g_{13} &= f_{13} + f_{21}^{\dagger} f_{23} + f_{31}^{\dagger} + f_{41}^{\dagger} f_{43} = 0, \\ g_{23} &= f_{12}^{\dagger} f_{13} + f_{23} + f_{32}^{\dagger} + f_{42}^{\dagger} f_{43} = 0, \end{aligned} \tag{A2.6}$$

$$\begin{aligned} g_{14} &= f_{14} + f_{21}^{\dagger} f_{24} + f_{31}^{\dagger} f_{34} + f_{41}^{\dagger} = 0, \\ g_{24} &= f_{12}^{\dagger} f_{14} + f_{24} + f_{32}^{\dagger} f_{34} + f_{42}^{\dagger} = 0, \\ g_{34} &= f_{13}^{\dagger} f_{14} + f_{23}^{\dagger} f_{24} + f_{34} + f_{43}^{\dagger} = 0. \end{aligned} \tag{A2.7}$$

These six equations can be used to write the f_{IJ}, $(J > I)$, solely in terms of the f_{IJ}, $(J < I)$, as follows. Equation (A2.5) gives f_{12} directly as

$$f_{12} = -(f_{21}^{\dagger} + f_{31}^{\dagger} f_{32} + f_{41}^{\dagger} f_{42}). \tag{A2.8}$$

Then the two equations, (A2.6), are solved simultaneously for f_{13} and f_{23}, yielding,

$$f_{23} = (f_{12}^{\dagger} f_{21}^{\dagger} - 1)^{-1} [f_{32}^{\dagger} + f_{42}^{\dagger} f_{43} - f_{12}^{\dagger} (f_{31}^{\dagger} + f_{41}^{\dagger} f_{43})], \tag{A2.9a}$$

and

$$f_{13} = -(f_{31}^{\dagger} + f_{41}^{\dagger} f_{43}) - f_{21}^{\dagger} f_{23}. \tag{A2.9b}$$

Substitution of the adjoint of (A2.8) into (A2.9a) then gives f_{23} in terms only of the f_{IJ}, $(J < I)$, and substitution of that result into (A2.9b) does the same for f_{13}. The three equations (A2.7) can be solved simultaneously for f_{14}, f_{24}, and f_{34}, yielding,

$$f_{34} = -[1-f_{13}^{\dagger}f_{31}^{\dagger}-(f_{23}^{\dagger}-f_{13}^{\dagger}f_{21}^{\dagger})(1-f_{12}^{\dagger}f_{21}^{\dagger})^{-1}(f_{32}^{\dagger}-f_{12}^{\dagger}f_{31}^{\dagger})]^{-1}$$
$$\times [f_{43}^{\dagger}+f_{13}^{\dagger}f_{41}^{\dagger}+(f_{23}^{\dagger}-f_{13}^{\dagger}f_{21}^{\dagger})(1-f_{12}^{\dagger}f_{21}^{\dagger})^{-1}(-f_{42}^{\dagger}+f_{12}^{\dagger}f_{41}^{\dagger})], \quad \text{(A2.10a)}$$

$$f_{24} = -(1-f_{12}^{\dagger}f_{21}^{\dagger})^{-1}[f_{42}^{\dagger}+f_{12}^{\dagger}f_{41}^{\dagger}+(f_{32}^{\dagger}-f_{12}^{\dagger}f_{31}^{\dagger})f_{34}], \quad \text{(A2.10b)}$$

and

$$f_{14} = -(f_{41}^{\dagger} + f_{31}^{\dagger}f_{34} + f_{21}^{\dagger}f_{24}). \quad \text{(A2.10c)}$$

Substitution of eqs. (A2.8) and (A2.9) into (A2.10a) gives f_{34} in terms of the f_{IJ}, $(J < I)$, only. Similarly, (A2.8), (A2.9), and (A2.10a) can then be used to write f_{24} in terms of the same set of variables. Equations (A2.10a,b) then can be used to eliminate f_{24} and f_{34} from (A2.10c). The resulting expressions will clearly be very lengthy.

It should be noted that the elements of the f_{IJ}, $(J > I)$, can be calculated numerically much more easily from those of the f_{IJ}, $(J < I)$, than eqs. (A2.3) or (A2.8) - (A2.10) indicate. Such a calculation involves the solution of $\sum_{\substack{I,J \\ J<I}} n_I n_J$ simultaneous linear equations in the same number of scalar variables. The complicated formulas above arise only when analytic formulas are desired relating these different matrix blocks f_{IJ}.

APPENDIX 3

Proofs of Alternative Formulas -- Multiple Partitioning

This appendix outlines some of the manipulations necessary to establish a number of inter-relations which have been quoted in section 4.4.

Equations (4.73) for an orthonormal basis are obtained as follows. From eq. (4.37) and eq. (4.72), one has,

$$\hat{H}_I^{(2)} = g_I^{-1} G_I$$

$$= g_I^{-1}[\hat{H}_I^{(1)} + \sum_{J\neq I} f_{JI}^{\dagger}[H_{JI} + \sum_{K\neq I} H_{JK} f_{KI}]] \,. \qquad \text{(A3.1)}$$

Elimination of the quantity in the inner brackets in this equation using

$$D_{JI}^{(1)} = (H\,\hat{T})_{JI} - (\hat{T}\,\hat{H}^{(1)})_{JI}$$

$$= H_{JI} + \sum_{K\neq I} H_{JK} f_{KI} - f_{JI}\hat{H}_I^{(1)}, \qquad \text{(A3.2)}$$

leads to the desired expression,

$$\hat{H}_I^{(2)} = g_I^{-1}[\hat{H}_I^{(1)} + \sum_{J\neq I} f_{JI}^{\dagger}(D_{JI}^{(1)} + f_{JI}\hat{H}_I^{(1)})]$$

$$= \hat{H}_I^{(1)} + g_I^{-1} \sum_{J\neq I} f_{JI}^{\dagger} D_{JI}^{(1)}.$$

Equation (4.10) has also been used in the last step.

The relation (4.76) between $D^{(1)}$ and $D^{(2)}$, in an orthonormal basis, is established immediately by substituting eq. (4.73) into the definition, (4.75), of $D^{(2)}$.

The non-orthonormal case presents many more complications here. From eq. (4.58), one has,

$$S_{KI} + \sum_{L\neq I} S_{KL} f_{LI} = g_{KI} = \sum_{L\neq K} f^{\dagger}_{LK}(S_{LI} + \sum_{M\neq I} S_{LM} f_{MI}). \qquad (A3.3)$$

Upon substitution of this equation into eq. (4.59), a series of alternative formulas for the metric g_I can be obtained, among them,

$$g_I = \sum_{K\neq I} f^{\dagger}_{KI}[g_{KI} - \sum_{L\neq K} f^{\dagger}_{LK}(S_{LI} + \sum_{M\neq I} S_{LM} f_{MI})] + S_{II} + \sum_{K\neq I} S_{IK} f_{KI}$$

$$= -\sum_{\substack{K\neq I \\ L\neq K}} f^{\dagger}_{KI} f^{\dagger}_{LK}(S_{LI} + \sum_{M\neq I} S_{LM} f_{MI}) + \sum_{K\neq I} f^{\dagger}_{KI} g_{KI} + S_{II} + \sum_{K\neq I} S_{IK} f_{KI}$$

$$= [1 - \sum_{K\neq I} f^{\dagger}_{KI} f^{\dagger}_{IK}](S_{II} + \sum_{M\neq I} S_{IM} f_{MI}) - \sum_{K\neq I} f^{\dagger}_{KI}[\sum_{\substack{L\neq K \\ L\neq I}} f^{\dagger}_{LK}(S_{LI} + \sum_{M\neq I} S_{LM} f_{MI})]$$

$$+ \sum_{K\neq I} f^{\dagger}_{KI} g_{KI}, \quad (I = 1, \ldots, m). \qquad (A3.5)$$

This last form is the generalization to the m x m partitioning of eqs. (A1.1) and (A1.2) of Appendix 1. In the case of a 2 x 2 partitioning, the second term of (A3.5) does not occur at all because of the restrictions on the range of the inner summation. Also, the summation symbols in the first and third terms of (A3.5) can be deleted in that case, since the summation is over only one term.

Using the notation $\hat{S}_I$ of eq. (4.65), the generalizations of eqs. (A1.3) and (A1.4) of Appendix 1 are obtained from (A3.5) as

$$g_I \hat{S}_I^{-1} = (1 - \sum_{K\neq I} f^{\dagger}_{KI} f^{\dagger}_{IK}) + \Big\{ \sum_{K\neq I} f^{\dagger}_{KI}[\sum_{\substack{L\neq K \\ L\neq I}} f^{\dagger}_{LK}(S_{LI} + \sum_{M\neq I} S_{LM} f_{MI})] - \sum_{K\neq I} f^{\dagger}_{KI} g_{KI} \Big\} \hat{S}_I^{-1}, \quad (I=1,\ldots,m). \qquad (A3.6)$$

The analogues of eqs. (A1.5) and (A1.6) of Appendix 1 are now obtained by right multiplying (A3.6) by $f^{\dagger}_{IP}$, $(P \neq I)$, and left multiplying the equation for $g_P \hat{S}_P^{-1}$ by the same factor, and combining the two equations to get,

$$f^{\dagger}_{IP} g_I \hat{S}_I^{-1} = g_P \hat{S}_P^{-1} f^{\dagger}_{IP} - \mathcal{D} f^{\dagger}_{IP} + f^{\dagger}_{IP} \mathcal{B} ,$$

and (A3.7)

$$g_I \hat{S}_I^{-1} f^{\dagger}_{PI} = f^{\dagger}_{PI} g_P \hat{S}_P^{-1} - f^{\dagger}_{PI} \mathcal{D} + \mathcal{B} f^{\dagger}_{PI},$$

where

$$\mathcal{B} = \sum_{\substack{K \neq I \\ K \neq P}} f^{\dagger}_{KI} f^{\dagger}_{IK} + \sum_{K \neq I} f^{\dagger}_{KI} \Big[\sum_{\substack{L \neq K \\ L \neq I}} f^{\dagger}_{LK} \Big(S_{LI} + \sum_{M \neq I} S_{LM} f_{MI} \Big) \Big] - \sum_{K \neq I} f^{\dagger}_{KI} g_{KI} \hat{S}_I^{-1},$$

and (A3.8)

$$\mathcal{D} = \sum_{\substack{K \neq P \\ K \neq I}} f^{\dagger}_{KP} f^{\dagger}_{PK} + \sum_{K \neq P} f^{\dagger}_{KP} \Big[\sum_{\substack{L \neq K \\ L \neq P}} f^{\dagger}_{LK} \Big(S_{LP} + \sum_{M \neq I} S_{LM} f_{MP} \Big) \Big] - \sum_{K \neq P} f^{\dagger}_{KP} g_{KP} \hat{S}_P^{-1}.$$

The two equations, (A3.7), as well as the two quantities, $\mathcal{B}$ and $\mathcal{D}$, can be obtained from each other by interchanging the indices P and I. The generalization of eqs. (A1.7) and (A1.8 of Appendix 1 is then obtained by dropping those terms in (A3.7) above involving g_{KI}, $K \neq I$. Here, however, this amounts to dropping only the last term inside the curly brackets of $\mathcal{B}$ and $\mathcal{D}$. Thus the usefulness of the resulting equations as generalized multiple partitioning 'pull-through' relations is severely hampered, because of the complexity and size of the last two terms of (A3.7), even when the orthogonality condition is satisfied.

Finally, the generalization of eqs. (A1.9) and (A1.10) of

Appendix 1, which were used to obtain one of the relations between $D^{(2)}$ and $D^{(1)}$ for a 2 x 2 partitioning, can be obtained as follows. From (4.58), one has

$$f^{\dagger}_{IJ}\hat{S}_I = g_{JI} - \sum_{\substack{K\neq J\\ K\neq I}} f^{\dagger}_{KJ}(S_{KI} + \sum_{L\neq I} S_{KL}f_{LI}) - (S_{JI} + \sum_{K\neq I} S_{JK}f_{KI}), \tag{A3.9}$$

so that, from eq. (4.59) and (A3.4),

$$\begin{aligned} f^{\dagger}_{IJ}g_I &= (1-f^{\dagger}_{IJ}f^{\dagger}_{JI})\Big[g_{JI} - \sum_{\substack{K\neq J\\ K\neq I}} f^{\dagger}_{KJ}(S_{KI} + \sum_{L\neq I} S_{KL}f_{LI}) - (S_{JI} + \sum_{K\neq I} S_{JK}f_{KI})\Big] \\ &\quad + f^{\dagger}_{IJ}\sum_{\substack{K\neq I\\ K\neq J}} f^{\dagger}_{KI}f^{\dagger}_{IK}\hat{S}_I - f^{\dagger}_{IJ}\sum_{K\neq I} f^{\dagger}_{KI}\Big[\sum_{\substack{L\neq K\\ L\neq I}} f^{\dagger}_{LK}(S_{LI} + \sum_{M\neq I} S_{LM}f_{MI})\Big] \\ &\quad + f^{\dagger}_{IJ}\sum_{K\neq I} f^{\dagger}_{KI}g_{KI} \\ &= -(1-f^{\dagger}_{IJ}f^{\dagger}_{IJ})(S_{JI} + \sum_{K\neq I} S_{JK}f_{KI}) + g_{JI} + f^{\dagger}_{IJ}\sum_{\substack{K\neq I\\ K\neq J}} f^{\dagger}_{KI}g_{KI} \\ &\quad + f^{\dagger}_{IJ}\sum_{\substack{K\neq I\\ K\neq J}} f^{\dagger}_{KI}f^{\dagger}_{IK}\hat{S}_I - f^{\dagger}_{IJ}\sum_{\substack{K\neq J\\ K\neq I}} f^{\dagger}_{KI}\Big[\sum_{\substack{L\neq K\\ L\neq I}} f^{\dagger}_{LK}(S_{LI} + \sum_{M\neq I} S_{LM}f_{MI})\Big] \\ &\quad - f^{\dagger}_{KJ}(S_{KI} + \sum_{\substack{K\neq J\\ K\neq I}} S_{KL}f_{LI}). \end{aligned} \tag{A3.10}$$

The last four complicated summation terms in (A3.10) make the result effectively useless, and therefore, no generalization of eqs. (3.15a,b) is given here.

The proof of eqs. (4.82) is as follows. From eq. (A3.9), one has,

$$-(S_{JI} + \sum_{K\neq I} S_{JK}f_{KI})\hat{S}_I^{-1} = f^{\dagger}_{IJ} - \Big[g_{JI} - \sum_{\substack{K\neq J\\ K\neq I}} f^{\dagger}_{KJ}(S_{KI} + \sum_{J\neq I} S_{KL}f_{LI})\Big]\hat{S}_I^{-1}\ .$$

Then, eq. (4.83) becomes,

$$D_{JI}^{(1)} = H_{JI} + \sum_{K\neq I} H_{JK}f_{KI} - (S_{JI} + \sum_{K\neq I} S_{JK}f_{KI})\hat{S}_I^{-1}(H_{II} + \sum_{J\neq I} H_{IJ}f_{JI})$$

$$= H_{JI} + \sum_{K\neq I} H_{JK}f_{KI} + [f_{IJ}^{\dagger} - g_{JI} - \sum_{\substack{K\neq J \\ K\neq I}} f_{KJ}^{\dagger}(S_{KI} + \sum_{L\neq I} S_{KL}f_{LI})\hat{S}_I^{-1}]$$

$$\times (H_{II} + \sum_{J\neq I} H_{IJ}f_{JI}), \qquad (A3.11)$$

from which, an expression for the combination $H_{JI} + \sum_{K\neq I} H_{JK}f_{KI}$ can be obtained. By definition, one has,

$$\hat{H}_I^{(2)} = g_I^{-1}G_I$$

$$= g_I^{-1}[H_{II} + \sum_{J\neq I} H_{IJ}f_{JI} + \sum_{J\neq I} f_{JI}^{\dagger}(H_{JI} + \sum_{K\neq I} H_{JK}f_{KI})],$$

which becomes, after using (A3.11),

$$\hat{H}_I^{(2)} = g_I^{-1}\, H_{II} + \sum_{J\neq I} H_{IJ}f_{JI} - \sum_{J\neq I} f_{JI}^{\dagger}[f_{IJ}^{\dagger} - g_{IJ}$$

$$- \sum_{\substack{K\neq I \\ K\neq J}} f_{KJ}^{\dagger}(S_{KI} + \sum_{L\neq I} S_{KL}f_{LI})\hat{S}_I^{-1}](H_{II} + \sum_{J\neq I} H_{IJ}f_{JI}) - D_{JI}^{(1)}$$

$$= g_I^{-1}\,(1 - \sum_{J\neq I} f_{JI}^{\dagger}f_{IJ}^{\dagger})(H_{II} + \sum_{J\neq I} H_{IJ}f_{JI})$$

$$+ \sum_{J\neq I} f_{JI}^{\dagger}[g_{JI} - \sum_{\substack{K\neq J \\ K\neq I}} f_{KJ}^{\dagger}(S_{KI} + \sum_{L\neq I} S_{KL}f_{LI})]\hat{H}_I^{(1)} + \sum_{J\neq I} f_{JI}^{\dagger}D_{JI}^{(1)}$$

$$= g_I^{-1}\,[g_I + \sum_{K\neq I} f_{KI}^{\dagger}[\sum_{\substack{L\neq K \\ L\neq I}} f_{LK}^{\dagger}(S_{LI} + \sum_{M\neq I} S_{LM}f_{MI})] + \sum_{K\neq I} f_{KI}^{\dagger}g_{KI}]$$

$$\times \hat{S}_I^{-1}(H_{II} + \sum_{J\neq I} H_{IJ}f_{JI})$$

$$+ \sum_{J\neq I} f_{JI}^{\dagger}[g_{JI} - \sum_{\substack{K\neq J \\ K\neq I}} f_{KJ}^{\dagger}(S_{KI} + \sum_{L\neq I} S_{KL}f_{LI})]\hat{H}_I^{(1)} + \sum_{J\neq I} f_{JI}^{\dagger}D_{JI}^{(1)},$$

using eq. (A3.5) to obtain this last form. A large amount of cancellation now occurs among the coefficients of

$$\hat{H}_I^{(1)} = \hat{S}_I^{-1}(H_{II} + \sum_{J\neq I} H_{IJ} f_{JI}),$$

with the final result being eq. (4.82),

$$\hat{H}_I^{(2)} = \hat{H}_I^{(1)} + g_I^{-1} \sum_{J\neq I} f_{JI}^{\dagger} D_{JI}^{(1)}.$$

APPENDIX 4

Description of Algorithms -- 2 x 2 Case

This appendix gives detailed descriptions of the implementation of the algorithms discussed in chapter 5. In various instances below, especially in the updating cycles, the order in which the computations are done is important. Greek indices refer to basis elements in S_B, Roman indices to basis elements in S_A.

1. Simple Diagonal Newton-Raphson (SDNR).

initialization:

$$f = 0$$

$$\hat{H}_A^{(1)} = H_{AA}$$

$$\hat{H}_B^{(1)\mathrm{diag}} = H_{BB}^{\mathrm{diag}}$$

then:

$$D_{\sigma r}^{(1)} = H_{\sigma r} + \sum_{\rho=1}^{n_B} H_{\sigma\rho} f_{\rho r} - \sum_{t=1}^{n_A} f_{\sigma t} (\hat{H}^{(1)})_{tr} ,$$

$$\delta f_{\sigma r} = D_{\sigma r}^{(1)} / [(\hat{H}_A^{(1)})_{rr} - (\hat{H}_B^{(1)\dagger})_{\sigma\sigma}] \quad .$$

$r = 1, \ldots, n_A$

$\sigma = 1, \ldots, n_B$

update:

$$(\hat{H}_A^{(1)})_{sr} \rightarrow (\hat{H}_A^{(1)})_{sr} + H_{s\sigma} \delta f_{\sigma r}, \quad (s = 1, \ldots, n_A),$$

$$(\hat{H}_B^{(1)\dagger})_{\sigma\sigma} \rightarrow (\hat{H}_B^{(1)\dagger})_{\sigma\sigma} - \delta f_{\sigma r} H_{r\sigma} ,$$

$$f_{\sigma r} \rightarrow f_{\sigma r} + \delta f_{\sigma r} \quad .$$

2. Quadratic Diagonal Newton-Raphson (QDNR).

initialization: $f = 0$,

$\hat{H}_A^{(1)} = H_{AA}$,

$\hat{H}_B^{(1)\mathrm{diag}} = H_{BB}^{\mathrm{diag}}$,

then

$$D_{\sigma r}^{(1)} = H_{\sigma r} + \sum_{\rho=1}^{n_B} H_{\sigma\rho} f_{\rho r} - \sum_{t=1}^{n_A} f_{\sigma t} (\hat{H}_A^{(1)})_{tr},$$

$$\Delta_{\sigma r} = (\hat{H}_A^{(1)})_{rr} - (\hat{H}_B^{(1)\dagger})_{\sigma\sigma},$$

if $H_{\sigma r} = 0$, then $\delta f_{\sigma r} = D_{\sigma r}^{(1)} / \Delta_{\sigma r}$,

if $\Delta_{\sigma r} = 0$, then $\delta f_{\sigma r} = +\sqrt{D_{\sigma r}^{(1)} / H_{r\sigma}}$ if real,

$\delta f_{\sigma r} = 0$, otherwise,

if both $H_{\sigma r} = 0, \Delta_{\sigma r} = 0$, then $\delta f_{\sigma r} = 0$,

otherwise,

$$\delta f_{\sigma r} = \frac{-2\chi D_{\sigma r}^{(1)}}{\Delta_{\sigma r} + [\Delta_{\sigma r}^2 + 4H_{\sigma r} D_{\sigma r}^{(1)}]^{\frac{1}{2}}},$$

$$\chi = \mathrm{sgn}(\Delta_{\sigma r}),$$

$r = \ , \ldots, n_A$,

$\sigma = 1, \ldots, n_B$.

update:

$$(\hat{H}_A^{(1)})_{sr} \rightarrow (\hat{H}_A^{(1)})_{sr} + H_{s\sigma} \delta f_{\sigma r}, \quad (s = 1, \ldots, n_A),$$

$$(\hat{H}^{(1)\dagger})_{\sigma\sigma} \rightarrow (\hat{H}^{(1)\dagger})_{\sigma\sigma} - \delta f_{\sigma r} H_{r\sigma},$$

$$f_{\sigma r} \rightarrow f_{\sigma r} + \delta f_{\sigma r}.$$

3. Diagonal Generalized Nesbet (DGN).

initialization: $f = 0$,

$\hat{H}_A^{(2)} = H_{AA}$,

$g_A = 1_A$,

then:

$$W_{\sigma r} = H_{\sigma r} + \sum_{\rho=1}^{n_B} H_{\sigma\rho} f_{\rho r} ,$$

$$D_{\sigma r}^{(2)} = W_{\sigma r} - \sum_{t=1}^{n_A} f_{\sigma t} (\hat{H}_A^{(2)})_{tr} ,$$

$$\delta f_{\sigma r} = D_{\sigma r}^{(2)} / [(\hat{H}_A^{(2)})_{rr} - H_{\sigma\sigma}] .$$

$r = 1, \ldots, n_A$,

update:

$$(\delta g_A)_{ts} = \delta f_{t\sigma}^{\dagger} f_{\sigma s} + f_{t\sigma}^{\dagger} \delta f_{\sigma s} + \delta f_{t\sigma}^{\dagger} \delta f_{\sigma s} ,$$

$$(s,t = 1, \ldots, n_A) ,$$

$$g_A \rightarrow g_A + \delta g_A ,$$

$$f_{\sigma r} \rightarrow f_{\sigma r} + \delta f_{\sigma r} , \quad (r = 1, \ldots, n_A),$$

$$A_{rs} = f_{r\sigma}^{\dagger} \sum_{t=1}^{n_A} \delta f_{\sigma t} (\hat{H}_A^{(2)})_{ts} + \delta f_{r\sigma}^{\dagger} \delta f_{\sigma s} (\hat{H}_A^{(2)})_{ss} ,$$

$$+ W_{r\sigma}^{\dagger} \delta f_{\sigma s} , \quad (r,s = 1, \ldots, n_A),$$

$$\hat{H}_A^{(2)} \rightarrow \hat{H}_A^{(2)} + g^{-1} A .$$

$\sigma = 1, \ldots, n_B$

4. Full Generalized Nesbet (FGN).

initialization: $f = 0$,

$$\hat{H}_A^{(2)} = H_{AA} ,$$

$$g_A = 1_A ,$$

then:

$$\left.\begin{aligned} W_{\sigma r} &= H_{\sigma r} + \sum_{\rho=1}^{n_B} H_{\sigma\rho} f_{\rho r} , \\ D_{\sigma r}^{(2)} &= W_{\sigma r} - \sum_{t=1}^{n_A} f_{\sigma t} (\hat{H}_A^{(2)})_{tr} , \end{aligned}\right\} r=1,\ldots,n_A ,$$

solve $\quad \delta f_{\sigma A}[\hat{H}_A^{(2)} - H_{\sigma\sigma} 1_A] = D_{\sigma A}^{(2)} .$

update:

$$(\delta g_A)_{ts} = \delta f_{t\sigma}^{\dagger} f_{\sigma s} + f_{t\sigma}^{\dagger} \delta f_{\sigma s} + \delta f_{t\sigma}^{\dagger} \delta f_{\sigma s} ,$$
$$(s,t = 1, \ldots, n_A) ,$$

$$g_A = g_A + \delta g_A ,$$

$$A_{rs} = W_{r\sigma}^{\dagger} \delta f_{\sigma s} - f_{r\sigma}^{\dagger} \sum_{t=1}^{n_A} \delta f_{\sigma t} (\hat{H}_A^{(2)})_{ts} ,$$
$$(r,s=1,\ldots,n_A),$$

$$\hat{H}_A^{(2)} \rightarrow \hat{H}_A^{(2)} + g_A^{-1} A ,$$

$$f_{\sigma r} \rightarrow f_{\sigma r} + \delta f_{\sigma r} , \quad (r = 1, \ldots, n_A),$$

$\sigma=1, \ldots, n_B.$

5. Simple Diagonal Newton-Raphson With Overlap (SDNRS).

initialization: $f = 0$, $h = 0$,

$$\hat{H}'_A = H_{AA}, \qquad \hat{S}_A = S_{AA},$$

$$\hat{H}'^{\dagger diag}_B = H^{\dagger diag}_{BB}, \qquad \hat{S}^{\dagger diag}_B = S^{\dagger diag}_{BB},$$

then:

$$G_{\sigma r} = H_{\sigma r} + \sum_{\rho=1}^{n_B} H_{\sigma\rho} f_{\rho r} + \sum_{s=1}^{n_A} h^*_{s\sigma} (\hat{H}'_A)_{sr},$$

$$g_{\sigma r} = S_{\sigma r} + \sum_{\rho=1}^{n_B} S_{\sigma\rho} f_{\rho r} + \sum_{s=1}^{n_A} h^*_{s\sigma} (\hat{S}_A)_{sr},$$

$$\Delta_{\sigma r} = (\hat{H}'^{\dagger}_B)_{\sigma\sigma} (\hat{S}_A)_{rr} - (\hat{H}'_A)_{rr} (\hat{S}^{\dagger}_B)_{\sigma\sigma}$$

$$\delta f_{\sigma r} = [g_{\sigma r} (\hat{H}'_A)_{rr} - G_{\sigma r} (\hat{S}_A)_{rr}] / \Delta_{\sigma r},$$

$$\delta h_{r\sigma} = [(\hat{S}^{\dagger}_B)_{\sigma\sigma} G_{\sigma r} - (\hat{H}'^{\dagger}_B)_{\sigma\sigma} g_{\sigma r}] / \Delta_{\sigma r},$$

update:

$$(\hat{H}'_A)_{sr} \to (\hat{H}'_A)_{sr} + H_{s\sigma} \delta f_{\sigma r},$$

$$(\hat{S}_A)_{sr} \to (\hat{S}_A)_{sr} + S_{s\sigma} \delta f_{\sigma r},$$

$$s = 1, \ldots, n_A, \quad r = 1, \ldots, n_A, \quad \sigma = 1, \ldots, n_B.$$

$$(\hat{H}'^{\dagger}_B)_{\sigma\sigma} \to (\hat{H}'^{\dagger}_B)_{\sigma\sigma} + \delta h_{r\sigma} H_{r\sigma},$$

$$(\hat{S}^{\dagger}_B)_{\sigma\sigma} \to (\hat{S}^{\dagger}_B)_{\sigma\sigma} + \delta h_{r\sigma} S_{r\sigma},$$

$$f_{\sigma r} \to f_{\sigma r} + \delta f_{\sigma r},$$

$$h_{r\sigma} \to h_{r\sigma} + \delta h_{r\sigma},$$

6. Quadratic Diagonal Newton-Raphson with Overlap (QDNRS).

initialization: $f = 0$, $h = 0$,

$\hat{H}'_A = H_{AA}$, $\hat{S}_A = S_{AA}$,

$\hat{H}'^{\dagger diag}_B = H^{\dagger diag}_{BB}$, $\hat{S}^{\dagger diag}_B = S^{\dagger diag}_{BB}$,

then:

$$G_{\sigma r} = H_{\sigma r} + \sum_{\rho=1}^{n_B} H_{\sigma\rho} f_{\rho r} + \sum_{s=1}^{n_A} h_{s\sigma} (\hat{H}'_A)_{sr} ,$$

$$g_{\sigma r} = S_{\sigma r} + \sum_{\rho=1}^{n_B} S_{\sigma\rho} f_{\rho r} + \sum_{s=1}^{n_A} h_{s\sigma} (\hat{S}_A)_{sr}$$

$$A = (\hat{S}_B)_{\sigma\sigma} H_{r\sigma} - S_{r\sigma} (\hat{H}'_B)_{\sigma\sigma} ,$$

$$B = (\hat{S}_B)_{\sigma\sigma} (\hat{H}'_A)_{rr} - (\hat{S}_A)_{rr} (\hat{H}'_B)_{\sigma\sigma} - S_{r\sigma} G_{\sigma r} + H_{r\sigma} g_{\sigma r},$$

$$C = -(\hat{S}_A)_{rr} G_{\sigma r} + g_{\sigma r} (\hat{H}'_A)_{rr},$$

if $A=0$, $B=0$, $\delta f_{\sigma r} = 0$,

if $A\neq 0$, $B=0$, $\delta f_{\sigma r} = (-C/A)^{\frac{1}{2}}$ if $C/A < 0$,

$\delta f_{\sigma r} = 0$ if $C/A > 0$,

if $A=0$, $B\neq 0$, $\delta f_{\sigma r} = -C/B$,

if $A\neq 0$, $B\neq 0$, $\delta f_{\sigma r} = \frac{-B}{|B|} \left(\frac{2C}{|B| + \sqrt{B^2 - 4AC}} \right)$

$r = 1, \ldots, n_A$, $\sigma = 1, \ldots, n_B$.

$$\delta h_{r\sigma} = \frac{G_{\sigma r} + (H'^{\dagger}_B)_{\sigma\sigma} \delta f_{\sigma r}}{-H_{r\sigma} \delta f_{\sigma r} - (H'_A)_{rr}} .$$

update:

$$(\hat{H}'_A)_{sr} \rightarrow (\hat{H}'_A)_{sr} + H_{s\sigma} \delta f_{\sigma r},$$

$$(\hat{S}_A)_{sr} \rightarrow (\hat{S}_A)_{sr} + S_{s\sigma} \delta f_{\sigma r},$$

$(s = 1, \ldots, n_A)$,

$$(\hat{H}'^{\dagger}_B)_{\sigma\sigma} \rightarrow (\hat{H}'^{\dagger}_B)_{\sigma\sigma} + \delta h_{r\sigma} H_{r\sigma},$$

$$(\hat{S}^{\dagger}_B)_{\sigma\sigma} \rightarrow (\hat{S}^{\dagger}_B)_{\sigma\sigma} + \delta h_{r\sigma} S_{r\sigma},$$

$$f_{\sigma r} \rightarrow f_{\sigma r} + \delta f_{\sigma r} ,$$

$$h_{r\sigma} \rightarrow h_{r\sigma} + \delta h_{r\sigma} .$$

7. Non-Orthogonal Diagonal Generalized Nesbet (DGNS).

initialization: $f = 0$,

$$\hat{H}_A^{(2)} = S_{AA}^{-1} H_{AA} ,$$

$$g_A = S_{AA} ,$$

then:

$$Y_{\sigma r} = S_{\sigma r} + \sum_{\rho=1}^{n_B} S_{\sigma\rho} f_{\rho r}, \quad (r=1,\ldots,n_A),$$

$$W_{\sigma r} = H_{\sigma r} + \sum_{\rho=1}^{n_B} H_{\sigma\rho} f_{\rho r},$$

$$D_{\sigma r}^{(2)} = W_{\sigma r} - \sum_{t=1}^{n_A} Y_{\sigma t} (\hat{H}_A^{(2)})_{tr},$$

$$\delta f_{\sigma r} = D_{\sigma r}^{(2)} / [(\hat{H}_A^{(2)})_{rr} S_{\sigma\sigma} - H_{\sigma\sigma}],$$

$r=1,\ldots,n_A$,

update:

$$(\delta g_A)_{ts} = \delta f_{t\sigma}^{\dagger} Y_{\sigma s} + Y_{t\sigma}^{\dagger} \delta f_{\sigma s} + \delta f_{t\sigma}^{\dagger} S_{\sigma\sigma} \delta f_{\sigma s} ,$$

$$(s,t = 1,\ldots, n_A),$$

$$g_A \rightarrow g_A + \delta g_A ,$$

$$Y_{\sigma r} \rightarrow Y_{\sigma r} + S_{\sigma\sigma} \delta f_{\sigma r}, \quad (r=1,\ldots,n_A),$$

$$A_{rs} = W_{r\sigma}^{\dagger} \delta f_{\sigma s} + Y_{r\sigma}^{\dagger} \sum_{t=1}^{n_A} \delta f_{\sigma t} (\hat{H}_A^{(2)})_{ts}$$

$$+ \delta f_{r\sigma}^{\dagger} S_{\sigma\sigma} \delta f_{\sigma s} (\hat{H}_A^{(2)})_{ss} ,$$

$$(r,s=1,\ldots,n_A),$$

$$\hat{H}_A^{(2)} \rightarrow \hat{H}_A^{(2)} + g_A^{-1} A ,$$

$$f_{\sigma r} \rightarrow f_{\sigma r} + \delta f_{\sigma r} , \quad (r=1,\ldots,n_A).$$

$\sigma=1,\ldots,n_B$.

8. Non-Orthogonal Full Generalized Nesbet (FGNS).

initialization: $f = 0$,

$$\hat{H}_A^{(2)} = S_{AA}^{-1} H_{AA} ,$$

$$g_A = S_{AA} ,$$

then:

$$Y_{\sigma r} = S_{\sigma r} + \sum_{\rho=1}^{n_B} S_{\sigma\rho} f_{\rho r} , \quad (r=1,\ldots,n_A),$$

$$W_{\sigma r} = H_{\sigma r} + \sum_{\rho=1}^{n_B} H_{\sigma\rho} f_{\rho r} ,$$

$$D_{\sigma r}^{(2)} = W_{\sigma r} - \sum_{t=1}^{n_A} Y_{\sigma t} (\hat{H}_A^{(2)})_{tr} ,$$

$r=1,\ldots,n_A$,

solve: $\delta f_{\sigma A} [S_{\sigma\sigma} \hat{H}_A^{(2)} - H_{\sigma\sigma} 1_A] = D_{\sigma A}^{(2)}$.

update:

$$(\delta g_A)_{ts} = \delta f_{t\sigma}^{\dagger} Y_{\sigma s} + Y_{t\sigma}^{\dagger} \delta f_{\sigma s} + \delta f_{t\sigma}^{\dagger} S_{\sigma\sigma} \delta f_{\sigma s},$$

$$(s,t=1,\ldots,n_A),$$

$$g_A \rightarrow g_A + \delta g_A ,$$

$$A_{ts} = W_{t\sigma}^{\dagger} \delta f_{\sigma s} - Y_{t\sigma}^{\dagger} \sum_{r=1}^{n_A} \delta f_{\sigma r} (\hat{H}_A^{(2)})_{rs} ,$$

$$(s,t=1,\ldots,n_A),$$

$$\hat{H}_A^{(2)} \rightarrow \hat{H}_A^{(2)} + g_A^{-1} A ,$$

$$f_{\sigma r} \rightarrow f_{\sigma r} + \delta f_{\sigma r}, \quad (r=1,\ldots,n_A),$$

$\sigma=1,\ldots,n_B$.

APPENDIX 5.

Rates of Convergence and Asymptotic Error Constants

This appendix analyses the rates of convergence of some of the algorithms, for the determination of f, described in chapter 5. These considerations are based to some degree on the work of Traub (1964).

To avoid confusion between subscripts denoting iteration number, and those denoting matrix elements, the fixed point iteration formula, eq. (5.4), will be rewritten here as

$$\phi(f) = f - \mathcal{K} D(f), \qquad (A5.1)$$

where $\mathcal{K} = -\mathcal{A}^{-1}$. Thus, $\phi(f^{exact}) = f^{exact}$, because $D(f^{exact})=0$. The basic iteration formula, eqs. (5.5), can be written in this notation as

$$f_{m+1} = f_m - \mathcal{K} D(f_m) = \phi(f_m). \qquad (A5.2)$$

If the necessary derivatives of $\phi(f)$ exist at f^{exact}, then $\phi(f)$ can be expanded in a Taylor series centered on f^{exact}, which allows one to write an expression for the error in the current estimate of f, given by $\phi(f)$, in terms of the error in the result of the previous iteration. One has

$$e_{\rho t}^{(m+1)} = \sum_{\sigma,r} \frac{\partial \phi_{\rho t}}{\partial f_{\sigma r}}\bigg|_{f^{exact}} e_{\sigma r}^{(m)} + \sum_{\substack{\sigma,r \\ \sigma',r'}} \frac{\partial^2 \phi_{\rho t}}{\partial f_{\sigma r} \partial f_{\sigma' r'}}\bigg|_{f^{exact}} e_{\sigma r}^{(m)} e_{\sigma' r'}^{(m)} + \ldots\ldots, \qquad (A5.3)$$

where

$$e_{\sigma r}^{(m)} = (f_m)_{\sigma r} - f_{\sigma r}^{exact}. \qquad (A5.4)$$

The iteration function $\phi(f)$ is then said to be of order p if all derivatives of the elements of ϕ, with respect to elements of f, of order less than p vanish at $f = f^{exact}$, while at least one derivative of order p does not vanish. Near the solution, the dominant term in the error will then be a sum over terms containing the product of p errors from the previous iteration. The asymptotic error constants for this iteration functions are taken here as the p^{th} order coefficients in (A5.3).

For the exact Newton-Raphson equations, (5.6), the iteration function is

$$\phi^{NR}(f) = f - J^{-1}D. \qquad (A5.5)$$

Thus, one has

$$\frac{\partial \phi^{NR}_{\rho t}}{\partial f_{\sigma r}} = - \sum_{\tau,s} \frac{\partial (J^{-1})_{\rho t,\tau s}}{\partial f_{\sigma r}} D_{\tau s} , \qquad (A5.6)$$

which vanishes at $f = f^{exact}$, because $D(f^{exact}) = 0$. The second derivatives are

$$\frac{\partial^2 \phi^{NR}_{\rho t}}{\partial f_{\sigma r}\partial f_{\sigma' r'}} = - \sum_{\tau,s} \frac{\partial^2 (J^{-1})_{\rho t,\tau s}}{\partial f_{\sigma r}\partial f_{\sigma' r'}} D_{\tau s} + \sum_{\tau,s} (J^{-1})_{\rho t,\tau s} \frac{\partial^2 D_{\tau s}}{\partial f_{\sigma r}\partial f_{\sigma' r'}} , \qquad (A5.7)$$

where the identity, $\partial(J^{-1}J)/\partial f_{\sigma r} = 0$, has been used to obtain the last term. At $f = f^{exact}$, the first summation in (A5.7) vanishes, but not the second one, in general. The Newton-Raphson equations are thus second order convergent, as is well known, and one can write

$$e^{(m+1)}_{\rho t} = -\tfrac{1}{2} \sum_{\substack{\sigma,r\,\sigma',r' \\ \tau,s}} (J^{-1})_{\rho t,\tau s} \left.\frac{\partial^2 D_{\tau s}}{\partial f_{\sigma r}\partial f_{\sigma' r'}}\right|_{f^{exact}} e^{(m)}_{\sigma r} e_{\sigma' r'} + O(e^3). \qquad (A5.8)$$

indicating explicitly the second order nature of $\phi^{NR}(f)$.

The algorithm SDNR is based on the equation $D^{(1)}(f)=0$. As seen from eq. (5.15), the operator $\mathcal{K}$ in the iteration formula $\phi^{SDNR}(f)$ is just the inverse of the diagonal part of the Jacobian matrix,

$$\mathcal{K}^{SDNR}_{\rho t,\sigma r} = \frac{1}{J^{(1)}_{\rho t,\rho t}} \delta_{\rho\sigma}\delta_{rt} , \qquad (A5.9)$$

so that,

$$\phi^{SDNR}_{\rho t}(f) = f_{\rho t} - \frac{D^{(1)}_{\rho t}(f)}{J^{(1)}_{\rho t,\rho t}} , \qquad (A5.10)$$

and

$$\left.\frac{\partial \phi^{SDNR}_{\rho t}}{\partial f_{\sigma r}}\right|_{f^{exact}} = \delta_{\rho\sigma}\delta_{rt} - \left.\frac{J^{(1)}_{\rho t,\sigma r}}{J^{(1)}_{\rho t,\rho t}}\right|_{f^{exact}} , \qquad (A5.11)$$

which vanishes only for $\rho=\sigma$, and $t=r$, in general. Thus, for ϕ^{SDNR}, one can write

$$e^{(m+1)}_{\rho t} = \frac{\sum_{\sigma\neq\rho} (\hat{H}^{(1)\dagger}_B)_{\rho\sigma} e^{(m)}_{\sigma t} - \sum_{r\neq t} (\hat{H}^{(1)}_A)_{rt} e^{(m)}_{\rho r}}{(\hat{H}^{(1)}_A)_{tt} - (\hat{H}^{(1)\dagger}_B)_{\rho\rho}} + O(e^2), \qquad (A5.12)$$

which verifies that ϕ^{SDNR} is indeed linearly convergent, and gives an expression for the dominant error term near the solution. A sufficient condition for convergence to occur is

$$\left|e^{(m+1)}_{\rho t}\right| < \left|e^{(m)}_{\rho t}\right|, \quad (\rho=1,\ldots,n_B;\ t=1,\ldots,n_A). \qquad (A5.13)$$

Assuming that

$$\left|\frac{e^{(m)}_{\sigma r}}{e^{(m)}_{\rho t}}\right| \leq 1, \quad (\sigma=1,\ldots,n_B;\ r=1,\ldots,n_A), \qquad (A5.14)$$

is true when $\phi_{\rho t}$ is to be evaluated, it can be seen from (A5.12)

that convergence will definitely occur if

$$\sum_{\sigma\neq} \left|(\hat{H}_B^{(1)\dagger})_{\sigma}\right| + \sum_{r\neq t} \left|(\hat{H}_A^{(1)})_{rt}\right| < \left|(\hat{H}_A^{(1)})_{tt} - (\hat{H}_B^{(1)\dagger})\right| \quad , \qquad (A5.15)$$

which is obtained by replacing all the ratios of the type (A5.14) occurring in (A5.13) by unity. When cycling systematically through the elements of f, all $f_{\sigma r}$, ($\sigma\neq\rho$, $r\neq t$), will have been updated more recently than $f_{\rho t}$ at this point, and thus, in an appropriate basis, the condition (A5.14), is not unreasonable, as long as the calculation is converging and the errors thus decreasing. While the condition (A5.15) is too crude to be of any practical use, it does indicate that the rate of convergence is related to the relative magnitudes of the differences between diagonal elements in $\hat{H}_A$ and $\hat{H}_B$, and of their off-diagonal elements. Convergence requires only that the errors in the elements of f decrease over a number of iterations, rather than that the errors in each element of f decrease in every iteration. The results of test calculations in Table 5.1 show that good rates of convergence occur when (A5.15) is violated substantially for some elements of f (as for the example with n = 250).

A more detailed error analysis indicates that a crucial factor for convergence is the calculation of δf one element at a time, with continual updating of $\hat{H}_A^{(1)}$ and $\hat{H}_B^{(1)}$ (and implicitly, of $D^{(1)}$). If these quantities are updated only after a complete sweep through δf, convergence occurs only for small n_A, and n_B, and is very slow, at best.

For the generalized Nesbet algorithms, based on the equation $D^{(2)}(f)=0$, the same sort of result is obtained. From eq. (5.25), one has

$$\left.\frac{\partial \phi^{DGN}_{\rho t}}{\partial f_{\sigma r}}\right|_{f^{exact}} = \delta_{\rho\sigma}\delta_{tr} - \left.\frac{J^{(2)}_{\rho t,\sigma r}}{H_{\rho\rho} - (\hat{H}_A^{(2)})_{tt}}\right|_{f^{exact}}, \quad (A5.16)$$

which does not vanish in general, and thus ϕ^{DGN} is linearly convergent in general, with

$$e^{(m+1)}_{\rho t} = \frac{\sum_{\sigma\neq\rho} H_{\rho\sigma} e^{(m)}_{\sigma t} - \sum_{r\neq t} (\hat{H}_A^{(2)})_{rt} e^{(m)}_{\rho r}}{(\hat{H}_A^{(2)})_{tt} - H_{\rho\rho}} + O(e^2). \quad (A5.17)$$

For $n_A=1$, one has

$$e^{(m+1)}_{\rho} = \frac{\sum_{\sigma\neq\rho} H_{\rho\sigma} e^{(m)}_{\sigma}}{(\hat{H}_A^{(2)})_{11} - H_{\rho\rho}} + O(e^2). \quad (A5.18)$$

Therefore, algorithm DGN is second order convergent when $n_A=1$ only if H_{BB} is diagonal.

For the algorithm FGN, the operator $\mathcal{K}$ is the inverse of the diagonal block part of $J^{(2)}$, defined in eq. (5.21), each such diagonal block of $\mathcal{K}$ corresponding to a row of $D^{(2)}$. The algorithm is linearly convergent, since,

$$\frac{\partial \phi^{FGN}_{\rho t}}{\partial f_{\sigma r}} = \delta_{\rho\sigma}\delta_{rt} + \sum_{\gamma=1}^{n_B}\sum_{s=1}^{n_A} \mathcal{K}^{FGN}_{\rho t,\gamma s} J^{(2)}_{\gamma s,\sigma r}, \quad (A5.19a)$$

with

$$\mathcal{K}^{FGN}_{\rho t,\gamma s} = -\delta_{\rho\gamma}[H_{\rho\rho}\,1_A - \hat{H}_A^{(2)}]^{-1}_{st}, \quad (A5.19b)$$

does not vanish in general at f^{exact}.

The generalized Nesbet algorithms for use with a non-orthonormal basis give similar results. From section 5.3.c, it is seen that

$$\mathcal{K}^{DGNS}_{\rho t,\sigma r} = \frac{-1}{S_{\sigma\sigma}(\hat{H}_A^{(2)})_{rr} - H_{\sigma\sigma}}\,\delta_{\rho\sigma}\delta_{rt}, \tag{A5.20}$$

and

$$\mathcal{K}^{FGNS}_{\rho t,\sigma r} = \delta_{\rho\sigma}[H_{\sigma\sigma} - S_{\sigma\sigma}\hat{H}_A^{(2)}]^{-1}. \tag{A5.21}$$

Therefore, one has,

$$\left.\frac{\partial \phi^{DGNS}_{\rho t}}{\partial f_{\sigma r}}\right|_{f^{exact}} = \delta_{\rho\sigma}\delta_{rt} + \left.\frac{J^{(2)}_{\rho t,\sigma r}}{S_{\rho\rho}(\hat{H}_A^{(2)})_{tt} - H_{\rho\rho}}\right|_{f^{exact}}, \tag{A5.22}$$

which does not vanish in general for any values of ρt and σr. Thus, for algorithm DGNS, one has,

$$e^{(m+1)}_{\rho t} = \sum_{\sigma,r}\left[\delta_{\rho\sigma}\delta_{rt} - \left.\frac{J^{(2)}_{\rho t,\sigma r}}{H_{\rho\rho} - S_{\rho\rho}(\hat{H}_A^{(2)})_{tt}}\right|_{f^{exact}}\right] e^{(m)}_{\sigma r} + O(e^2). \tag{A5.23}$$

It is seen from eq. (5.52), defining $J^{(2)}$, that unless $n_A=1$, the coefficient of $e^{(m)}_{\rho t}$ on the right hand side here is not zero, although it is likely very small. The expression for the error in ϕ^{FGNS} is of the same form as (A5.19a) with (A5.21) substituted for (A5.19b).

The error analysis for algorithms SDNRS and QDNRS requires an extension of the procedures used above. The iteration formula must now be written as the pair of equations

$$\begin{bmatrix}\phi_f(f,h)\\ \phi_h(f,h)\end{bmatrix} = \begin{bmatrix}f\\ h\end{bmatrix} - \begin{bmatrix}\mathcal{K}^{(11)}(f,h) & \mathcal{K}^{(12)}(f,h)\\ \mathcal{K}^{(21)}(f,h) & \mathcal{K}^{(22)}(f,h)\end{bmatrix}\begin{bmatrix}G_{BA}(f,h)\\ g_{BA}(f,h)\end{bmatrix}. \tag{A5.24}$$

Comparison to equation (5.44) yields the result,

$$\mathcal{K}^{(11)}_{t,\sigma r} = \frac{(\hat{S}_A)_{tt}}{\Delta_{\rho t}} \delta_{\rho\sigma}\delta_{tr}, \qquad \mathcal{K}^{(12)}_{t,\sigma r} = \frac{-(\hat{H}'_A)_{tt}}{\Delta_{\rho t}} \delta_{\rho\sigma}\delta_{rt}$$

$$\mathcal{K}^{(21)}_{t,\sigma r} = \frac{-(\hat{S}^{\dagger}_B)}{\Delta_{\rho t}} \delta_{\rho\sigma}\delta_{tr}, \qquad \mathcal{K}^{(22)}_{t,\sigma r} = \frac{(\hat{H}'^{\dagger}_B)}{\Delta_{\rho t}} \delta_{\rho\sigma}\delta_{tr}, \tag{A5.25}$$

where $\Delta_{\rho t}$ is defined in eq. (5.45). An expression for the errors in the iteration formula (A5.24) must now be obtained from a Taylor series in the elements of both f and h, which yields the result,

$$(e_f^{(m+1)})_{\rho t} = \sum_{\sigma,r}\left[\frac{\partial(\phi_f)_{\rho t}}{\partial f_{\sigma r}}(e_f^{(m)})_{\sigma r} + \frac{\partial(\phi_f)_{\rho t}}{\partial h_{r\sigma}}(e_h^{(m)})_{r\sigma}\right]$$

$$+\tfrac{1}{2}\sum_{\substack{\sigma,r\\ \sigma',r'}}\left[\frac{\partial^2(\phi_f)_{\rho t}}{\partial f_{\sigma r}\partial f_{\sigma' r'}}(e_f^{(m)})_{\sigma r}(e_f^{(m)})_{\sigma' r'}\right. \tag{A5.26}$$

$$\left.+2\frac{\partial^2(\phi_f)_{\rho t}}{\partial f_{\sigma r}\partial h_{r'\sigma'}}(e_f^{(m)})_{\sigma r}(e_h^{(m)})_{r'\sigma'} + \frac{\partial^2(\phi_f)_{\rho t}}{\partial h_{r\sigma}\partial h_{r'\sigma'}}(e_h^{(m)})_{r\sigma}(e_h^{(m)})_{r'\sigma'}\right]$$

$$+ O(e^3),$$

for ϕ_f, with all derivatives evaluated at f^{exact} and h^{exact}. A similar expansion can be written for $(e_h^{(m+1)})$. Substitution of (A5.25) and (A5.24), then gives

$$(e_f^{(m+1)})_{\rho t} = \sum_{\sigma\neq\rho}\frac{[(\hat{H}'_A)_{tt}(\hat{S}^{\dagger}_B)_{\rho\sigma} - (\hat{S}_A)_{tt}(\hat{H}'^{\dagger}_B)_{\rho\sigma}]}{\Delta_{\rho t}}(e_f^{(m)})_{\sigma t} \tag{A5.27}$$

$$+\sum_{r\neq t}\frac{[(\hat{H}'_A)_{tt}(\hat{S}_A)_{rt} - (\hat{S}_A)_{tt}(\hat{H}'_A)_{rt}]}{\Delta_{\rho t}}(e_h^{(m)})_{r\rho} + O(e^2),$$

and

$$(e_h^{(m+1)})_{t\rho} = \sum_{\sigma\neq\rho} \frac{[(\hat{S}_B^{\dagger})_{\rho\rho}(\hat{H}_B'^{\dagger})_{\rho\sigma} - (\hat{H}_B'^{\dagger})_{\rho\rho}(\hat{S}_B^{\dagger})_{\rho\sigma}]}{\Delta_{\rho t}} (e_f^{(m)})_{\sigma t}$$

$$+ \sum_{r\neq t} \frac{[(\hat{S}_B^{\dagger})_{\rho\rho}(\hat{H}_A')_{rt} - (\hat{H}_B'^{\dagger})_{\rho\rho}(\hat{S}_A)_{rt}]}{\Delta_{\rho t}} (e_h^{(m)})_{r\rho} + O(e^2). \quad \text{(A5.28)}$$

Neither $(e_f^{(m)})_{\rho t}$ or $(e_h^{(m)})_{t\rho}$ occur in the first order term of either of these equations.

In all the first order error estimates derived in this Appendix, the denominator of the error estimate is seen to be identical to the denominator in the iteration formula. Thus, if this denominator becomes small, not only does δf (or δf and δh) become large, but so do the errors in f (or f and h). Also, it is seen that these error estimates all involve off-diagonal elements of H (or H and S), and therefore, improved convergence is expected in all algorithms if these matrices are made more diagonal.

APPENDIX 6

Algorithms for the Determination of $\hat{T}$ -- Multiple Partitioning Case

The purpose of this appendix is to outline, in some detail, algorithms for solving eqs. (5.74) - (5.77) for the matrix elements of the off-diagonal blocks of the uncoupling operator $\hat{T}$ in an m x m partitioning.

A6.1 Methods Based on $D_{JI}^{(1)}(\hat{T}) = 0$.

If the f_{JI}^{o}, (I,J=1, ..., m, I≠J), are approximate solutions to any of the defining conditions (5.74) - (5.77), and the exact solutions are given by $f_{JI} = f_{JI}^{o} + \delta f_{JI}$, then, from the equations $D_{JI}^{(1)}(\hat{T}) = 0$, it is seen that the exact corrections δf_{JI} to the f_{JI}^{o} are given by

$$\sum_{K\neq I} \hat{\bar{H}}_{JK}^{I}(\hat{T}^{o})\delta f_{KI} - \delta f_{JI}\hat{H}_{I}(\hat{T}) = -D_{JI}^{(1)}(\hat{T}^{o}), \quad (I,J=1,\ldots,m,\ I\neq J), \tag{A6.1}$$

where

$$\hat{\bar{H}}_{JK}^{I}(\hat{T}^{o}) = H_{JK} - f_{JI}^{o}H_{IK} \,. \tag{A6.2}$$

If the exact effective operators $\hat{H}_{I}(\hat{T})$, (I=1,...,m), were known, the linear system (A6.1) could be solved directly for the δf_{JI}.

The Newton-Raphson equations corresponding to the nonlinear system $D_{JI}^{(1)} = 0$, eqs. (5.74), are

$$\sum_{K\neq I} \hat{\bar{H}}_{JK}^{I}(\hat{T}^{o})\delta f_{KI} - \delta f_{JI}\hat{H}_{I}^{(1)}(\hat{T}^{o}) = -D_{JI}^{(1)}(\hat{T}^{o}), \quad (I,J=1,\ldots,m,\ I\neq J), \tag{A6.3}$$

which differ from (A6.1) only in that the approximate effective operators $\hat{H}_I^{(1)}(\hat{T}^o)$ appear in (A6.3) in place of the exact operators $\hat{H}_I(\hat{T})$ in (A6.1). Equations (A6.3) are obtained by substituting the Jacobian matrix with elements

$$J^{(1)}_{\sigma_J r_I, \rho_K t_L} = \frac{\partial (D^{(1)}_{JI})_{\sigma r}}{\partial (f_{KL})_{\rho t}} = -(\hat{H}_I^{(1)})_{tr}\delta_{\sigma\rho}\delta_{KJ}\delta_{LI} + (\hat{\hat{H}}^I_{JK})_{\sigma\rho}\delta_{rt}\delta_{LI}\,, \tag{A6.4}$$

into eq. (5.6), and isolating the JI block. The similarity of eqs. (A6.1) - (A6.4) to the corresponding equations for a 2 x 2 partitioning is seen if it is noted that $\hat{\hat{H}}^A_{BB} = \hat{H}_B^{(1)\dagger}$ in that case.

If solved exactly, eqs. (A6.3) would lead to a second order convergent algorithm. In fact, if the $\hat{H}_I^{(1)}(\hat{T}^o)$ are replaced by the $\hat{H}_I^{(2)}(\hat{T}^o)$, the resulting iteration formula is nearly third order convergent, just as in the 2 x 2 case. However, the linear system (A6.3) is of dimension $2\sum_{I<J} n_I n_J$, which can be unacceptably large even when $n = \sum_I n_I$ is itself not unusually large.

There are at least two levels of diagonal approximations possible here. In the diagonal block approximation, only terms involving δf_{JI} itself in eq. (A6.3) are retained, leaving

$$\hat{\hat{H}}^I_{JJ}\delta f_{JI} - \delta f_{JI}\hat{H}_I^{(1)} = -D^{(1)}_{JI}, \quad (I,J=1,\ldots,m,\ I\neq j). \tag{A6.5}$$

This involves the solution of $m(m-1)$ smaller systems of linear equations in each iteration, of dimensions, respectively, $n_I n_J$, $(I,J=1,\ldots,m,\ I\neq J)$, a considerable reduction in computation per iterative sweep. These equations become

especially useful if each of the n_I is small, that is, if the partitioning divides up the full space into a large number of subspaces of small dimension. It might also be necessary to use (A6.5) if the off-diagonal elements of the $\hat{\hat{H}}^I_{JJ}$ and the $\hat{H}^{(1)}_I$ are large. In the 2 x 2 case, eqs. (A6.5) are still the full Newton-Raphson equations, however.

The lowest level of diagonal approximation of (A6.3) gives an iteration formula which reduces to that of algorithm SDNR in the 2 x 2 case. It consists of retention of only the individual diagonal elements of the Jacobian matrix (A6.4), which leads to the iteration formula,

$$\delta f_{\sigma_J r_I} = \frac{D^{(1)}_{\sigma_J r_I}}{[(\hat{H}^{(1)}_I)_{rr} - (\hat{\hat{H}}^I_{JJ})_{\sigma\sigma}]} \qquad \text{(A6.6)}$$

Like SDNR, an efficient iterative scheme based on eq. (A6.6) would consist of cycling through the δf_{JI} one element at a time, calculating the $D^{(1)}_{\sigma_J r_I}$ as required, and storing the $\hat{H}^{(1)}_I$ and diagonal elements of the $\hat{\hat{H}}^I_{JJ}$ continuously. Because the $\hat{H}^{(1)}_I$ and $\hat{\hat{H}}^I_{JJ}$ are linear in the f_{JI}, they are easily updated, according to

$$(\delta\hat{H}^{(1)}_I)_{sr} = (H_{IJ})_{s\sigma}\delta f_{\sigma_J r_I}, \quad (s=1,\ldots,n_I), \qquad \text{(A6.7a)}$$

and

$$(\delta\hat{\hat{H}}^I_{JJ})_{\sigma\sigma} = (H_{IJ})_{r\sigma}\delta f_{\sigma_J r_I} = (\delta\hat{H}^{(1)}_I)_{rr} . \qquad \text{(A6.7b)}$$

A change in f_{KL} affects only $\hat{H}^{(1)}_L$ and $\hat{\hat{H}}^L_{KK}$. If the diagonal elements in different diagonal blocks of H are well separated, and the off-diagonal elements are small compared to these

separations, then a reasonable starting approximation is $\hat{T}=1_n$, or $f_{JI} = 0$, (I,J=1,...,m, I≠J). The block columns of $\hat{T}$ can be determined individually here, in any order, because the effective operators $\hat{H}_I^{(1)}$, and $\hat{\hat{H}}_{JJ}^{I}$, as well as the quantities $D_{JI}^{(1)}$, (J=1, ...,m, J≠I), depend only on the f_{LI}, (L=1,...,m, L≠I).

Substitution of $f_{JI} = f_{JI}^{o} + \delta f_{JI}$ into eq. (A6.1) yields the exact equation for the δf_{JI},

$$D_{JI}^{(1)}(\hat{T}^o) = -\sum_{K\neq I} \hat{\hat{H}}_{JK}^{I}(\hat{T}^o)\delta f_{KI} - \delta f_{JI}\hat{H}_I^{(1)}(\hat{T}^o) + \delta f_{JI}\sum_{K\neq I} H_{IK}\delta f_{KI},$$

$$(I,J=1, \ldots, m, I\neq J). \qquad (A6.8)$$

The diagonal block approximation,

$$D_{JI}^{(1)}(\hat{T}^o) = -\hat{\hat{H}}_{JJ}^{I}\delta f_{JI} + \delta f_{JI}\hat{H}_I^{(1)} + \delta f_{JI}H_{IJ}\delta f_{JI}, \qquad (A6.9)$$

has a form in δf_{JI} like the form of the basic defining condition, eq. (2.16), for a 2 x 2 partitioning. The diagonal elements of this equation give a quadratic iteration formula,

$$(H_{IJ})_{r\sigma}\delta f_{\sigma_J r_I}^2 + [(\hat{H}_I^{(1)})_{rr} - (\hat{\hat{H}}_{JJ}^{I})_{\sigma\sigma}]\delta f_{\sigma_J r_I} - (D_{JI}^{(1)})_{\sigma r} = 0. \qquad (A6.10)$$

When $\delta f_{\sigma_J r_I}$ is large, this formula may give improved convergence. The relative increase in cost accompanying the use of (A6.10) in place of (A6.6) depends on the dimension of the problem, but becomes negligible as H becomes large.

A6.2 Methods Based on $D^{(2)}_{JI}(\hat{T}) = 0$.

The elements of the Jacobian matrix in this case are given by,

$$J^{(2)}_{\sigma_J r_I, \rho_L t_K} = \frac{\partial (D^{(2)}_{JI})_{\sigma r}}{\partial (f_{LK})_{\rho t}}$$

$$= [(H_{JJ})_{\sigma\rho}\delta_{rt} - (\hat{H}^{(2)}_I)_{tr}\delta_{\sigma\rho}]\delta_{LJ}\delta_{KI} + (H_{JL})_{\sigma\rho}\delta_{rt}\delta_{KI}$$

$$- \sum_{s=1}^{n_I} (f_{JI})_{\sigma s} \frac{\partial (\hat{H}^{(2)}_I)_{sr}}{\partial (f_{LI})_{\rho t}} \delta_{KI}, \qquad \text{(A6.11)}$$

where, from (4.89), one obtains,

$$\frac{\partial (\hat{H}^{(2)}_I)_{sr}}{\partial (f_{LI})_{\rho t}} = \sum_{M \neq I} [(g_I^{-1}\hat{H}^{\dagger}_I f^{\dagger}_{MI})_{s\rho}\delta_{rt} - (g_I^{-1} f^{\dagger}_{MI})_{s\rho}(\hat{H}_I)_{tr}]\delta_{LM}. \qquad \text{(A6.12)}$$

For any n_I which are unity, the first derivatives of the corresponding $\hat{H}^{(2)}_I$ are zero. The relative insensitivity of the effective operators $\hat{H}^{(2)}_I(\hat{T}^o)$, as approximations to the exact operators $\hat{H}_I(\hat{T})$, to errors in the off-diagonal blocks of $\hat{T}$, can again be exploited by neglecting the derivatives of $\hat{H}^{(2)}_I$ in (A6.11). This truncation leaves the simplified Jacobian matrix, $\tilde{J}^{(2)}$, with the only nonzero elements given by,

$$\tilde{J}^{(2)}_{\sigma_J r_I, \rho_L t_I} = [(H_{JJ})_{\sigma\rho}\delta_{rt} - (\hat{H}^{(2)}_I)_{tr}\delta_{\sigma\rho}]\delta_{LJ} + (H_{JL})_{\sigma\rho}\delta_{rt}. \qquad \text{(A6.13)}$$

This yields the Newton-Raphson equations

$$H_{JJ}\delta f_{JI} - \delta f_{JI}\hat{H}^{(2)}_I + \sum_{\substack{K \neq J \\ K \neq I}} H_{JK}\delta f_{KI} = -D^{(2)}_{JI}, \quad (I,J=1,\ldots,m,\ I \neq J). \qquad \text{(A6.14)}$$

In these equations, only the coefficient matrices $\hat{H}_I^{(2)}$ depend on $\hat{T}$, and these only on the I^{th} block column of $\hat{T}$, respectively. Therefore, in principle, eqs. (A6.14) can be used to solve for the f_{LI}, (L=1,...,m, L≠I) for each value of I individually, that is, for a single block column of $\hat{T}$ at a time.

If the separations between diagonal elements in different diagonal blocks are not small compared to the off-diagonal elements of H, then diagonal approximations of (A6.14) are useful. The diagonal block approximation is

$$H_{JJ}\delta f_{JI}-\delta f_{JI}\hat{H}_I^{(2)}=-D_{JI}^{(2)}, \quad (J=1,\ldots,m,\ J\neq I), \qquad (A6.15)$$

where the $n_I(n-n_I)$ dimensional system of linear equations (A6.14) for a given value of I is replaced by (m-1) linear systems of the smaller respective dimensions $n_I n_J$, (J=1,...,m, J≠I). Equations (A6.15) can still lead to relatively costly overall iterations unless the products $n_I n_J$ are all small. However, they can be approximated further to give generalizations of the algorithms DGN and FGN.

The exact change in $\hat{H}_I^{(2)}$ due to a change in a set of f_{KI}, (K=1,...,m, K≠I), is

$$\delta\hat{H}_I^{(2)}=g_I^{-1(new)}\Big[\sum_{J\neq I}\delta f_{JI}^{\dagger}\Big(D_{JI}^{(2)}+\sum_{K\neq I}H_{JK}\delta f_{KI}-\delta f_{JI}\hat{H}_I^{(2)}\Big) + \sum_{J\neq I}W_{JI}^{\dagger}\delta f_{JI}-\sum_{J\neq I}f_{JI}^{\dagger}\delta f_{JI}\hat{H}_I^{(2)}\Big], \qquad (A6.16)$$

where

$$W_{JI} = H_{JI} + \sum_{L\neq I} H_{JL}f_{LI} \cdot \qquad (A6.17)$$

If only a single f_{JI} is changed, eq. (A6.16) reduces to the

same form as eq. (5.23),

$$\delta\hat{H}_I^{(2)}=g_I^{-1(new)}[\delta f_{JI}^{\dagger}(D_{JI}^{(2)}+H_{JJ}\delta f_{JI}-\delta f_{JI}\hat{H}_I^{(2)}) + W_{JI}^{\dagger}\delta f_{JI}-f_{JI}^{\dagger}\delta f_{JI}\hat{H}_I^{(2)}]. \tag{A6.18}$$

For any change in f_{JI}, the entire operator $\hat{H}_I^{(2)}$ is changed, and thus, as for a 2 x 2 partitioning, it is most efficient to change groups of elements within the matrix f_{JI} before updating $\hat{H}_I^{(2)}$.

The generalization of the algorithm DGN is

$$\delta f_{\sigma_J r_I} = \frac{(D_{JI}^{(2)})_{\sigma r}}{(\hat{H}_I^{(2)})_{rr} - (H_{JJ})_{\sigma\sigma}}, \quad (r=1,\ldots,n_I), \tag{A6.19}$$

and when all n_I elements of the σ^{th} row of f_{JI} are changed in this way, the change in $\hat{H}_I^{(2)}$ is

$$\delta\hat{H}_I^{(2)}=g_I^{-1(new)}[(f_{JI}^{(new)\dagger})_{I\sigma}(\delta f_{JI})_{\sigma I}\hat{H}_I^{(2)}+(W_{JI}^{\dagger})_{I\sigma}(\delta f_{JI})_{\sigma I} +(\delta f_{JI}^{\dagger})_{I\sigma}(\delta f_{JI})_{\sigma I}\hat{H}_I^{(2)d}]. \tag{A6.20}$$

The generalization of algorithm FGN is

$$(\delta f_{JI})_{\sigma I}=(D_{JI}^{(2)})_{\sigma I}[(H_{JJ})_{\sigma\sigma}1_I-\hat{H}_I^{(2)}]^{-1}, \tag{A6.21}$$

and the corresponding change in $\hat{H}_I^{(2)}$ is

$$\delta\hat{H}_I^{(2)}=g_I^{-1(new)}[(W_{JI}^{\dagger})_{I\sigma}\delta f_{\sigma I}-(f_{JI}^{\dagger})_{I\sigma}\delta f_{\sigma I}\hat{H}_I^{(2)}]. \tag{A6.22}$$

In eqs. (A6.20) - (A6.22), the symbols $I\sigma$ (σI) indicate the σ^{th} columns (rows) of the I^{th} block row (column).

A6.3 Methods Based on the Simultaneous Solution of $G_{JI}(\hat{T})=0$ and $g_{JI}(\hat{T})=0$.

The Newton-Raphson equations arising from eqs. (5.76) are the pairs

$$\sum_{L\neq J} (\delta f_{LJ})^{\dagger} W_{LI} + \sum_{L\neq I} (W_{LJ})^{\dagger} \delta f_{LI} = -G_{JI}, \tag{A6.23a}$$

and

$$(\delta f_{JI})^{\dagger} + \delta f_{IJ} + \sum_{\substack{L\neq I \\ L\neq J}} [(\delta f_{LJ})^{\dagger} f_{LI} + (f_{LJ})^{\dagger} \delta f_{LI}] = -g_{JI}, \tag{A6.23b}$$

$$(I < J = 1, \ldots, m),$$

where the quantities W_{LI} are defined in eq. (A6.17). All of these equations must be solved simultaneously. That is, they cannot be separated easily into a number of subsets without common variables.

The diagonal block approximation to (A6.23) is the pair

$$(\delta f_{IJ})^{\dagger} W_{II} + W_{JJ}^{\dagger} \delta f_{JI} = -G_{JI}, \tag{A6.24a}$$

and

$$\delta f_{JI} + (\delta f_{IJ})^{\dagger} = -g_{JI}. \tag{A6.24b}$$

Solving the diagonal parts of these equations simultaneously for corresponding elements of δf_{JI} and $(\delta f_{IJ})^{\dagger}$, gives the iteration formulas

$$(\delta f_{JI})_{\sigma r} = \frac{-(g_{JI})_{\sigma r}(W_{II})_{rr} + (G_{JI})_{\sigma r}}{(W_{II})_{rr} - (W_{JJ}^{\dagger})_{\sigma\sigma}}, \tag{A6.25a}$$

and

$$(\delta f_{IJ}^{*})_{r\sigma} = \frac{-(G_{JI})_{\sigma r} + (g_{JI})_{\sigma r}(W_{JJ}^{\dagger})_{\sigma\sigma}}{(W_{II})_{rr} - (W_{JJ}^{\dagger})_{\sigma\sigma}}, \tag{A6.25b}$$

which are similar to the iteration formulas for algorithm SDNRS.

The simplicity of eq. (A6.24b) makes it possible to eliminate either δf_{JI} or $(\delta f_{JI})^\dagger$ from eq. (A6.24a), leaving an equation in terms of only one of these. Substitution of

$$(\delta f_{IJ})^\dagger = -(g_{JI} + \delta f_{JI}) \tag{A6.26}$$

into (A6.24a), gives the equation

$$-\delta f_{JI} W_{II} + W_{JJ}^\dagger \delta f_{JI} = -G_{JI} + g_{JI} W_{II} , \tag{A6.27}$$

for δf_{JI}. The diagonal part of this equation yields

$$(\delta f_{JI})_{\sigma r} = \frac{(G_{JI})_{\sigma r} - (g_{JI} W_{II})_{\sigma r}}{(W_{II})_{rr} - (W_{JJ}^\dagger)_{\sigma\sigma}} , \tag{A6.28}$$

and, using this in eq. (A6.26) gives

$$(\delta f_{IJ}^*)_{r\sigma} = \frac{-(G_{JI})_{\sigma r} + (g_{JI} W_{II})_{\sigma r} - (g_{JI})_{\sigma r}[(W_{II})_{rr} - (W_{JJ}^\dagger)_{\sigma\sigma}]}{(W_{II})_{rr} - (W_{JJ}^\dagger)_{\sigma\sigma}} . \tag{A6.29}$$

These formulas amount to addition of the quantity $\sum_{t \neq r} (g_{JI})_{\sigma t} (W_{II})_{tr}$ to the numerator of (A6.25b) and its subtraction from the numerator of (A6.25a). If δf_{JI} is first eliminated from (A6.24a), then the same sort of result is obtained, but now, the quantity $\sum_{\rho \neq \sigma} (W_{JJ}^\dagger)_{\sigma\rho} (g_{JI})_{\rho r}$ is added to the numerator of (A6.25b) and subtracted from the numerator of (A6.25a). This ambiguity in iteration formula is undesirable. Only actual numerical studies can determine whether the inclusion of these potentially costly additional terms in the iteration formulas (A6.28) and (A6.29) is justified, in comparison to (A6.25).

The implementation of these iteration formulas is similar

to the SDNRS algorithm. The quantities W_{JI}, (I,J=1,...,m) are stored, and updated continuously as the elements of f_{KL} are changed. The elements of G_{JI} are calculated from

$$G_{JI} = W_{JI} + \sum_{L\neq J} (f_{LJ})^{\dagger} W_{LI}, \qquad \text{(A6.30)}$$

and those of g_{JI} from (5.76b), as required.

A6.4 Methods Based on $D_{JI}^{(3)}(\hat{T}) = 0$.

The Newton-Raphson equations corresponding to eqs. (5.77) are

$$-\sum_{M}[(\delta f_{MK})^{\dagger} + \delta f_{KM} + \sum_{\substack{N\neq K \\ N\neq M}} ((\delta f_{NK})^{\dagger} f_{NM} + (f_{NK})^{\dagger} \delta f_{NM})] D_{MI}^{(3)} \qquad \text{(A6.31)}$$

$$+ \sum_{L\neq K} (\delta f_{LK})^{\dagger} W_{LI} + \sum_{L\neq I} (W_{LK})^{\dagger} \delta f_{LI} = -G_{KI}, \quad (K,I=1,...,m,\ K\neq I).$$

All of the off-diagonal blocks of $\hat{T}$ occur in each of these equations in a complicated manner. If only those terms involving δf_{KI} on the left hand side of (A6.31) are retained, a much simpler approximation results,

$$W_{KK}^{\dagger} \delta f_{KI} - \delta f_{KI} D_{II}^{(3)} = -G_{KI}, \quad (K,I=1,...,m,\ K\neq I). \qquad \text{(A6.32)}$$

Note that $D_{II}^{(3)} = (g^{-1}G)_{II} \neq \hat{H}_I^{(2)}$ unless eqs. (5.77) are satisfied. It has not been determined whether or not an efficient iterative procedure can be based on eqs. (A6.32). The quantities $D_{II}^{(3)}$ are comparatively costly to calculate, and the iterative scheme would apparently require the maintainence of some estimate of g^{-1} throughout the calculation.

APPENDIX 7

Additional Perturbation Series -- Orthonormal Basis

In section 6.2, several perturbation formulas for the mapping f, and the effective operators $\hat{H}_A$, G_A, and $\tilde{H}_A$, were given. The purpose there was to give the low order terms of the perturbation series solely in terms of the perturbed operator H. Such perturbation formulas for the effective operators would then have a general significance, in that they are not necessarily obtainable only from the partitioning formalism presented in chapters 2 and 3. For example, the formulas for $\tilde{H}_A$ can also be obtained using a canonical transformation formalism.

The purpose of this Appendix is to supplement the material in section 6.2. Additional information on the efficient calculation, especially of high order terms, of the perturbation series for $g_A^{\pm 1}$, $g_A^{\pm \frac{1}{2}}$, and $\tilde{H}_A$ is presented. The formulas tabulated in section 6.2 become too lengthy with increasing order to be of practical use much beyond third order.

Perturbation series for those powers of the metric g_A, namely $g_A^{+\frac{1}{2}}$, $g_A^{-\frac{1}{2}}$, and g_A^{-1}, can be obtained in a number of ways. In terms of the perturbation series for f, the series for these quantities is obtained by generalization of the familiar power series expansions,

$$g_A^{\frac{1}{2}} = (1_A + f^\dagger f)^{\frac{1}{2}} = \sum_{n=0}^{\infty} g_A^{\frac{1}{2}(n)}$$

$$= 1_A + \tfrac{1}{2}f^\dagger f - \tfrac{1}{8}(f^\dagger f)^2 + \tfrac{1}{16}(f^\dagger f)^3 - \tfrac{5}{128}(f^\dagger f)^4 + \tfrac{7}{256}(f^\dagger f)^5 - \ldots, \tag{A7.1}$$

$$g_A^{-\frac{1}{2}} = (1_A + f^\dagger f)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} g_A^{-\frac{1}{2}}(n)$$

$$= 1_A - \tfrac{1}{2}f^\dagger f + \tfrac{3}{8}(f^\dagger f)^2 - \tfrac{5}{16}(f^\dagger f)^3 + \tfrac{35}{128}(f^\dagger f)^4 - \tfrac{63}{256}(f^\dagger f)^5 + \ldots, \tag{A7.2}$$

and

$$g_A^{-1} = (1_A + f^\dagger f)^{-1} = \sum_{n=0}^{\infty} g_A^{-1}(n)$$

$$= 1_A - f^\dagger f + (f^\dagger f)^2 - (f^\dagger f)^3 + (f^\dagger f)^4 - \ldots \ . \tag{A7.3}$$

In each of (A7.1), (A7.2), and (A7.3), actual expressions for the $g_A^{\frac{1}{2}}(n)$, $g_A^{-\frac{1}{2}}(n)$, and $g_A^{-1}(n)$ are obtained by substituting eq. (6.8) into each product, and isolating all terms of order n. Tables A7.1, A7.2, and A7.3 contain some low order formulas of this type. Perturbation formulas based on eqs. (A7.1) - (A7.3) can be generated to high order automatically without great difficulty, but with increasing order, they rapidly become very lengthy, and thus costly to use.

For automatic computation of high order terms in each of these series, a more efficient procedure is available. It is possible to obtain a series for any of $g_A^{\frac{1}{2}}$, $g_A^{-\frac{1}{2}}$, and g_A^{-1}, in terms of that of g_A and other powers of g_A by expanding identities of the form

$$g_A g_A^{-1} = 1_A \ , \tag{A7.4a}$$

$$(g_A^{-\frac{1}{2}})^2 = g_A^{-1} , \tag{A7.4b}$$

$$(g_A^{\frac{1}{2}})^2 = g_A \,, \tag{A7.4c}$$

$$g_A^{\frac{1}{2}} g_A^{-\frac{1}{2}} = 1_A \,, \tag{A7.4d}$$

$$g_A^{\frac{1}{2}} g_A^{-1} g_A^{\frac{1}{2}} = 1_A \,, \tag{A7.4e}$$

$$g_A^{-\frac{1}{2}} g_A \, g_A^{-\frac{1}{2}} = 1_A \,, \tag{A7.4f}$$

and so on. From (A7.4a), one obtains

$$g_A^{-1(n)} = 1_A \delta_{n0} - \sum_{j=0}^{n-1} g_A^{(n-j)} g_A^{-1(j)}, \quad n \geqslant 0, \tag{A7.5}$$

giving $g_A^{-1(n)}$ in terms of $g_A^{(j)}$, $(j = 0, \ldots, n)$, and lower order $g_A^{-1(j)}$, $(j = 0, \ldots, n-1)$. Here, the second term is assumed to make no contribution when $(n-1) < 0$. Similarly, eq. (A7.4c) yields the expansion

$$g_A^{\frac{1}{2}(n)} = \tfrac{1}{2} g_A^{(n)} - \tfrac{1}{2} \sum_{j=1}^{n-1} g_A^{\frac{1}{2}(n-j)} g_A^{\frac{1}{2}(j)}, \tag{A7.6}$$

giving $g_A^{\frac{1}{2}(n)}$ in terms of $g_A^{(n)}$ and lower order terms in the expansion of $g_A^{\frac{1}{2}}$. Equation (A7.4b) yields a similar result for $g_A^{-\frac{1}{2}(n)}$ in terms of $g_A^{-1(n)}$ and lower order terms in the series for $g_A^{-\frac{1}{2}}$. From eq. (A7.4d), one obtains

$$g_A^{\frac{1}{2}(n)} = - \sum_{j=1}^{n} g_A^{\frac{1}{2}(n-j)} g_A^{-\frac{1}{2}(j)}, \quad n \geqslant 2, \tag{A7.7a}$$

and

$$g_A^{-\frac{1}{2}(n)} = - \sum_{j=0}^{n-1} g_A^{\frac{1}{2}(n-j)} g_A^{-\frac{1}{2}(j)}, \quad n \geqslant 2, \tag{A7.7b}$$

which inter-relate the perturbation series for $g_A^{\frac{1}{2}}$ and $g_A^{-\frac{1}{2}}$.

Somewhat more complex expressions are obtained from eqs. (A7.4e) and (A7.4f). The important feature of eqs. (A7.5) - (A7.7), and other similar ones, is that the evaluation of a k^{th} order quantity generally requires the evaluation of no more than k products of two $n_A \times n_A$ matrices. As k increases, this represents a rapidly decreasing fraction of the computation required if formulas like (A7.1) - (A7.3) are used explicitly. This substantial computational advantage is a result of not having to repeatedly re-evaluate certain often-occurring combinations in (A7.1) - (A7.3). The need to store all lower order terms in these series for use in the calculation of higher order terms may be regarded as a disadvantage, but it is of no consequence if $n_A \ll n_B$. An appropriate combination of eqs. (A7.5) - (A7.7) with eqs. (6.15), together with eqs. (6.12) or (6.13), is certainly the most practical procedure for the calculation of high order terms in the series for $\tilde{H}_A$.

Equations (A7.5) - (A7.7) can also be used to obtain the terms in the series for $g_A^{\pm\frac{1}{2}}$ and g_A^{-1} solely in terms of the $g_A^{(j)}$. Such expressions are particularly useful when moderately high order calculations are being done by hand and algebraically, rather than numerically by machine. Tables A7.4, A7.5 and A7.6 contain several low order formulas of this type. Note the simplicitly of these formulas relative to those in Tables A7.1 - A7.3. Tables A7.7 and A7.8 contain expressions for low order terms in the series for $\tilde{H}_A$ in terms of only g_A and $\hat{H}_A$ or G_A. Finally, Table A7.9 contains low order formulas for G_A

in which the equations $D^{(n)}(f) = 0$ have been used to eliminate all terms involving H_{BB}. These formulas are the same as those derived using eqs. (6.14).

TABLE A7.1 $g_A^{\frac{1}{2}(n)}$

$$g_A^{\frac{1}{2}(0)} = 1_A$$

$$g_A^{\frac{1}{2}(1)} = 0$$

$$g_A^{\frac{1}{2}(2)} = \tfrac{1}{2} f^{(1)} \dagger_f(1)$$

$$g_A^{\frac{1}{2}(3)} = \tfrac{1}{2}\left(f^{(1)}\dagger_f(2) + f^{(1)}\dagger_f(2)\right)$$

$$g_A^{\frac{1}{2}(4)} = \tfrac{1}{2}\left(f^{(1)}\dagger_f(3) + f^{(2)}\dagger_f(2) + f^{(3)}\dagger_f(1)\right) - \tfrac{1}{8}\left(f^{(1)}\dagger_f(1)\right)^2$$

$$g_A^{\frac{1}{2}(5)} = \tfrac{1}{2}\left(f^{(1)}\dagger_f(4) + f^{(2)}\dagger_f(3) + f^{(3)}\dagger_f(2) + f^{(4)}\dagger_f(1)\right)$$
$$-\tfrac{1}{8}\left\{f^{(1)}\dagger_f(2) + f^{(2)}\dagger_f(1),\ f^{(1)}\dagger_f(1)\right\}_+$$

$$g_A^{\frac{1}{2}(6)} = \tfrac{1}{2}\left(f^{(1)}\dagger_f(5) + f^{(2)}\dagger_f(4) + f^{(3)}\dagger_f(3) + f^{(4)}\dagger_f(2) + f^{(5)}\dagger_f(1)\right)$$
$$-\tfrac{1}{8}\left\{f^{(1)}\dagger_f(3) + f^{(2)}\dagger_f(2) + f^{(3)}\dagger_f(1),\ f^{(1)}\dagger_f(1)\right\}_+$$
$$-\tfrac{1}{8}\left(f^{(1)}\dagger_f(2) + f^{(2)}\dagger_f(1)\right)^2 + \tfrac{1}{16}\left(f^{(1)}\dagger_f(1)\right)^3$$

TABLE A7.2 $g_A^{-\frac{1}{2}(n)}$

$$g_A^{-\frac{1}{2}(0)} = 1_A$$

$$g_A^{-\frac{1}{2}(1)} = 0$$

$$g_A^{-\frac{1}{2}(2)} = -\tfrac{1}{2}(f^{(1)\dagger}f^{(1)})$$

$$g_A^{-\frac{1}{2}(3)} = -\tfrac{1}{2}(f^{(1)\dagger}f^{(2)}+f^{(2)\dagger}f^{(1)})$$

$$g_A^{-\frac{1}{2}(4)} = -\tfrac{1}{2}(f^{(1)\dagger}f^{(3)}+f^{(2)\dagger}f^{(2)}+f^{(3)\dagger}f^{(1)}) + \tfrac{3}{8}(f^{(1)\dagger}f^{(1)})^2$$

$$g_A^{-\frac{1}{2}(5)} = -\tfrac{1}{2}(f^{(1)\dagger}f^{(4)}+f^{(2)\dagger}f^{(3)}+f^{(3)\dagger}f^{(2)}+f^{(4)\dagger}f^{(1)}) + \tfrac{3}{8}\{f^{(1)\dagger}f^{(2)}+f^{(2)\dagger}f^{(1)},\ f^{(1)\dagger}f^{(1)}\}_+$$

$$g_A^{-\frac{1}{2}(6)} = -\tfrac{1}{2}(f^{(1)\dagger}f^{(5)}+f^{(2)\dagger}f^{(4)}+f^{(3)\dagger}f^{(3)}+f^{(4)\dagger}f^{(2)}+f^{(5)\dagger}f^{(1)}) + \tfrac{3}{8}\{f^{(1)\dagger}f^{(3)}+f^{(2)\dagger}f^{(2)}+f^{(3)\dagger}f^{(1)},\ f^{(1)\dagger}f^{(1)}\}_+ + \tfrac{3}{8}(f^{(2)\dagger}f^{(1)}+f^{(1)\dagger}f^{(2)})^2 - \tfrac{5}{16}(f^{(1)\dagger}f^{(1)})^3$$

TABLE A7.3 $g_A^{-1(n)}$

$$g_A^{-1(0)} = 1_A$$

$$g_A^{-1(1)} = 0$$

$$g_A^{-1(2)} = -f^{(1)}\dagger_f(1)$$

$$g_A^{-1(3)} = -(f^{(1)}\dagger_f(2)+f^{(2)}\dagger_f(1))$$

$$g_A^{-1(4)} = -(f^{(1)}\dagger_f(3)+f^{(2)}\dagger_f(2)+f^{(3)}\dagger_f(1)) + (f^{(1)}\dagger_f(1))^2$$

$$g_A^{-1(5)} = -(f^{(1)}\dagger_f(4)+f^{(2)}\dagger_f(3)+f^{(3)}\dagger_f(2)+f^{(4)}\dagger_f(1)) + \{f^{(2)}\dagger_f(1)+f^{(1)}\dagger_f(2),\ f^{(1)}\dagger_f(1)\}_+$$

$$g_A^{-1(6)} = -(f^{(1)}\dagger_f(5)+f^{(2)}\dagger_f(4)+f^{(3)}\dagger_f(3)+f^{(4)}\dagger_f(2)+f^{(5)}\dagger_f(1)) + \{f^{(1)}\dagger_f(3)+f^{(2)}\dagger_f(2)+f^{(3)}\dagger_f(1),\ f^{(1)}\dagger_f(1)\}_+ + (f^{(1)}\dagger_f(2)+f^{(2)}\dagger_f(1))^2 - (f^{(1)}\dagger_f(1))^3$$

TABLE A7.4 $g_A^{\frac{1}{2}(n)}$

$$g_A^{\frac{1}{2}(0)} = 1_A$$

$$g_A^{\frac{1}{2}(1)} = 0$$

$$g_A^{\frac{1}{2}(2)} = \tfrac{1}{2}g_A^{(2)}$$

$$g_A^{\frac{1}{2}(3)} = \tfrac{1}{2}g_A^{(3)}$$

$$g_A^{\frac{1}{2}(4)} = \tfrac{1}{2}g_A^{(4)} - \tfrac{1}{8}g_A^{(2)^2}$$

$$g_A^{\frac{1}{2}(5)} = \tfrac{1}{2}g_A^{(5)} - \tfrac{1}{8}\{g_A^{(2)},\ g_A^{(3)}\}_+$$

$$g_A^{\frac{1}{2}(6)} = \tfrac{1}{2}g_A^{(6)} - \tfrac{1}{8}\{g_A^{(4)},\ g_A^{(2)}\}_+ - \tfrac{1}{8}g_A^{(3)^2} + \tfrac{1}{16}g_A^{(2)^3}$$

TABLE A7.5 $g_A^{-\frac{1}{2}(n)}$

$$g_A^{-\frac{1}{2}(0)} = 1_A$$

$$g_A^{-\frac{1}{2}(1)} = 0$$

$$g_A^{-\frac{1}{2}(2)} = -\tfrac{1}{2}g_A^{(2)}$$

$$g_A^{-\frac{1}{2}(3)} = -\tfrac{1}{2}g_A^{(3)}$$

$$g_A^{-\frac{1}{2}(4)} = -\tfrac{1}{2}g_A^{(4)} + \tfrac{3}{8}g_A^{(2)^2}$$

$$g_A^{-\frac{1}{2}(5)} = -\tfrac{1}{2}g_A^{(5)} + \tfrac{3}{8}\{g_A^{(2)}, g_A^{(3)}\}_+$$

$$g_A^{-\frac{1}{2}(6)} = -\tfrac{1}{2}g_A^{(6)} + \tfrac{3}{8}\{g_A^{(2)}, g_A^{(4)}\}_+ + \tfrac{3}{8}g_A^{(3)^2} - \tfrac{5}{16}g_A^{(2)^3}$$

TABLE A7.6 $g_A^{-1(n)}$

$$g_A^{-1(0)} = 1_A$$

$$g_A^{-1(1)} = 0$$

$$g_A^{-1(2)} = -g_A^{(2)}$$

$$g_A^{-1(3)} = -g_A^{(3)}$$

$$g_A^{-1(4)} = -g_A^{(4)} + g_A^{(2)^2}$$

$$g_A^{-1(5)} = -g_A^{(5)} + \{g_A^{(2)}, g_A^{(3)}\}_+$$

$$g_A^{-1(6)} = -g_A^{(6)} + \{g_A^{(2)}, g_A^{(3)}\}_+ + g_A^{(3)^2} - g_A^{(2)^3}$$

TABLE A7.7 $\tilde{H}_A^{(n)}$ in Terms of the $g_A^{(n)}$ and $\hat{H}_A^{(n)}$

$$\tilde{H}_A^{(0)} = \hat{H}_A^{(0)}$$

$$\tilde{H}_A^{(1)} = \hat{H}_A^{(1)}$$

$$\tilde{H}_A^{(2)} = \hat{H}_A^{(2)} + \tfrac{1}{2}[g_A^{(2)}, \hat{H}_A^{(0)}]_-$$

$$\tilde{H}_A^{(3)} = \hat{H}_A^{(3)} + \tfrac{1}{2}[g_A^{(3)}, \hat{H}_A^{(0)}]_- + \tfrac{1}{2}[g_A^{(2)}, \hat{H}_A^{(1)}]_-$$

$$\tilde{H}_A^{(4)} = \hat{H}_A^{(4)} + \tfrac{1}{2}[g_A^{(4)}, \hat{H}_A^{(0)}]_- + \tfrac{1}{2}[g_A^{(3)}, \hat{H}_A^{(1)}]_- + \tfrac{1}{2}[g_A^{(2)}, \hat{H}_A^{(2)}]_- - \tfrac{1}{8}g_A^{(2)2}\hat{H}_A^{(0)} + \tfrac{3}{8}\hat{H}_A^{(0)}g_A^{(2)2} - \tfrac{1}{4}g_A^{(2)}\hat{H}_A^{(0)}g_A^{(2)}$$

$$\tilde{H}_A^{(5)} = \hat{H}_A^{(5)} + \tfrac{1}{2}[g_A^{(5)}, \hat{H}_A^{(0)}]_- + \tfrac{1}{2}[g_A^{(4)}, \hat{H}_A^{(1)}]_- + \tfrac{1}{2}[g_A^{(3)}, \hat{H}_A^{(2)}]_- + \tfrac{1}{2}[g_A^{(2)}, \hat{H}_A^{(3)}]_- - \tfrac{1}{8}\{g_A^{(2)}, g_A^{(3)}\}_+\hat{H}_A^{(0)} + \tfrac{3}{8}\hat{H}_A^{(0)}\{g_A^{(2)}, g_A^{(3)}\}_+ - \tfrac{1}{8}g_A^{(2)2}\hat{H}_A^{(1)} + \tfrac{3}{8}\hat{H}_A^{(1)}g_A^{(2)2} - \tfrac{1}{4}(g_A^{(2)}\hat{H}_A^{(0)}g_A^{(3)} + g_A^{(3)}\hat{H}_A^{(0)}g_A^{(2)}) - \tfrac{1}{4}g_A^{(2)}\hat{H}_A^{(1)}g_A^{(2)}$$

$$\tilde{H}_A^{(6)} = \hat{H}_A^{(6)} + \tfrac{1}{2}[g_A^{(6)}, \hat{H}_A^{(0)}]_- + \tfrac{1}{2}[g_A^{(5)}, \hat{H}_A^{(1)}]_- + \tfrac{1}{2}[g_A^{(4)}, \hat{H}_A^{(2)}]_- + \tfrac{1}{2}[g_A^{(3)}, \hat{H}_A^{(3)}]_- + \tfrac{1}{2}[g_A^{(2)}, \hat{H}_A^{(4)}]_- - \tfrac{1}{8}\{g_A^{(4)}, g_A^{(2)}\}_+\hat{H}_A^{(0)} + \tfrac{3}{8}\hat{H}_A^{(0)}\{g_A^{(4)}, g_A^{(2)}\}_+ - \tfrac{1}{8}g_A^{(3)2}\hat{H}_A^{(0)} + \tfrac{3}{8}\hat{H}_A^{(0)}g_A^{(3)2} + \tfrac{1}{16}g_A^{(2)3}\hat{H}_A^{(0)} - \tfrac{5}{16}\hat{H}_A^{(0)}g_A^{(2)3} - \tfrac{1}{8}\{g_A^{(2)}, g_A^{(3)}\}_+\hat{H}_A^{(1)} + \tfrac{3}{8}\hat{H}_A^{(1)}\{g_A^{(2)}, g_A^{(3)}\}_+ - \tfrac{1}{8}g_A^{(2)2}\hat{H}_A^{(2)} + \tfrac{3}{8}\hat{H}_A^{(2)}g_A^{(2)2} - \tfrac{1}{4}g_A^{(4)}\hat{H}_A^{(0)}g_A^{(2)} - \tfrac{1}{4}g_A^{(2)}\hat{H}_A^{(0)}g_A^{(4)} + \tfrac{1}{16}g_A^{(2)2}\hat{H}_A^{(0)}g_A^{(2)} + \tfrac{3}{16}g_A^{(2)}\hat{H}_A^{(0)}g_A^{(2)2} - \tfrac{1}{4}g_A^{(3)}\hat{H}_A^{(0)}g_A^{(3)} - \tfrac{1}{4}g_A^{(3)}\hat{H}_A^{(1)}g_A^{(2)} - \tfrac{1}{4}g_A^{(2)}\hat{H}_A^{(1)}g_A^{(3)} - \tfrac{1}{4}g_A^{(2)}\hat{H}_A^{(2)}g_A^{(2)}$$

TABLE A7.8 $\tilde{H}_A^{(n)}$ in Terms of the $g_A^{(n)}$ and $G_A^{(n)}$

$$\tilde{H}_A^{(0)} = G_A^{(0)}$$

$$\tilde{H}_A^{(1)} = G_A^{(1)}$$

$$\tilde{H}_A^{(2)} = G_A^{(2)} - \tfrac{1}{2}\{g_A^{(2)}, G_A^{(0)}\}_+$$

$$\tilde{H}_A^{(3)} = G_A^{(3)} - \tfrac{1}{2}\{g_A^{(3)}, G_A^{(0)}\}_+ - \tfrac{1}{2}\{g_A^{(2)}, G_A^{(1)}\}_+$$

$$\tilde{H}_A^{(4)} = G_A^{(4)} + \{G_A^{(0)}, -\tfrac{1}{2}g_A^{(4)} + \tfrac{3}{8}g_A^{(2)2}\}_+ - \tfrac{1}{2}\{g_A^{(3)}, G_A^{(1)}\}_+$$
$$- \tfrac{1}{2}\{g_A^{(2)}, G_A^{(2)}\}_+ + \tfrac{1}{4}g_A^{(2)}G_A^{(0)}g_A^{(2)}$$

$$\tilde{H}_A^{(5)} = G_A^{(5)} + \{G_A^{(0)}, -\tfrac{1}{2}g_A^{(5)} + \tfrac{3}{8}\{g_A^{(2)}, g_A^{(3)}\}_+\}_+ + \{G_A^{(1)}, -\tfrac{1}{2}g_A^{(4)} + \tfrac{3}{8}g_A^{(2)2}\}_+$$
$$- \tfrac{1}{2}\{G_A^{(2)}, g_A^{(3)}\}_+ - \tfrac{1}{2}\{G_A^{(3)}, g_A^{(2)}\}_+$$
$$+ \tfrac{1}{4}(g_A^{(3)}G_A^{(0)}g_A^{(2)} + g_A^{(2)}G_A^{(0)}g_A^{(3)}) + \tfrac{1}{4}g_A^{(2)}G_A^{(1)}g_A^{(2)}$$

$$\tilde{H}_A^{(6)} = G_A^{(6)} - \tfrac{1}{2}\{g_A^{(6)}, G_A^{(0)}\}_+ - \tfrac{1}{2}\{g_A^{(5)}, G_A^{(1)}\}_+ - \tfrac{1}{2}\{g_A^{(4)}, G_A^{(2)}\}_+$$
$$- \tfrac{1}{2}\{g_A^{(3)}, G_A^{(3)}\}_+ - \tfrac{1}{2}\{g_A^{(2)}, G_A^{(4)}\}_+ + \tfrac{3}{8}\{G_A^{(0)}, \{g_A^{(2)}, g_A^{(4)}\}_+\}_+$$
$$+ \tfrac{3}{8}\{G_A^{(0)}, g_A^{(3)2}\}_+ - \tfrac{5}{16}\{G_A^{(0)}, g_A^{(2)3}\}_+ + \tfrac{3}{8}\{G_A^{(1)}, \{g_A^{(2)}, g_A^{(3)}\}_+\}_+$$
$$+ \tfrac{3}{8}\{G_A^{(2)}, g_A^{(2)2}\}_+ + \tfrac{1}{4}g_A^{(4)}G_A^{(0)}g_A^{(2)} + \tfrac{1}{4}g_A^{(2)}G_A^{(0)}g_A^{(4)} + \tfrac{1}{4}g_A^{(3)}G_A^{(0)}g_A^{(3)}$$
$$+ \tfrac{1}{4}g_A^{(3)}G_A^{(1)}g_A^{(2)} + \tfrac{1}{4}g_A^{(2)}G_A^{(1)}g_A^{(3)} - \tfrac{3}{16}g_A^{(2)2}G_A^{(0)}g_A^{(2)}$$
$$- \tfrac{3}{16}g_A^{(2)}G_A^{(0)}g_A^{(2)2} + \tfrac{1}{4}g_A^{(2)}G_A^{(2)}g_A^{(2)}$$

TABLE A7.9 $G_A^{(n)}$

$$G_A^{(0)} = H_{AA}^{(0)}$$

$$G_A^{(1)} = H_{AA}^{(1)}$$

$$G_A^{(2)} = H_{AA}^{(2)}+H_{AB}^{(1)}f^{(1)}+f^{(1)\dagger}f^{(1)}H_{AA}^{(0)}$$

$$G_A^{(3)} = H_{AA}^{(3)}+H_{AB}^{(2)}f^{(1)}+H_{AB}^{(1)}f^{(2)}+f^{(1)\dagger}f^{(1)}H_{AA}^{(1)}$$
$$+(f^{(1)\dagger}f^{(2)}+f^{(2)\dagger}f^{(1)})H_{AA}^{(0)}$$

$$G_A^{(4)} = H_{AA}^{(4)}+H_{AB}^{(3)}f^{(1)}+H_{AB}^{(2)}f^{(2)}+f^{(1)\dagger}f^{(1)}H_{AA}^{(2)}$$
$$+H_{AB}^{(1)}f^{(3)}+f^{(1)\dagger}f^{(1)}H_{AB}^{(1)}f^{(1)}+(f^{(1)\dagger}f^{(2)}+f^{(2)\dagger}f^{(1)})H_{AA}^{(1)}$$
$$+(f^{(1)\dagger}f^{(3)}+f^{(2)\dagger}f^{(2)}+f^{(3)\dagger}f^{(1)})H_{AA}^{(0)}$$

APPENDIX 8

Non-relativistic Approximation of the Dirac Hamiltonian

The purpose of this appendix is to list expressions for the various effective operators dealt with in section 6.3a, to order $(v/c)^6$, for comparison with expressions reported by DeVries (1970). In order to facilitate this, the terms in the Dirac hamiltonian will be written in the symbolic form

$$H^{(0)} = \begin{bmatrix} m & 0 \\ 0 & -m \end{bmatrix}, \quad H^{(1)} = \begin{bmatrix} 0 & \alpha^\dagger \\ \alpha & 0 \end{bmatrix}, \quad H^{(2)} = \begin{bmatrix} \phi & 0 \\ 0 & \phi \end{bmatrix}. \tag{A8.1}$$

Here all the natural constants have been dropped except for the mass m, which is useful in comparing the formulas below to those in section 6.3a. In this notation, the reduced resolvent is

$$L = \frac{1}{2m} 1_B. \tag{A8.2}$$

Formulas of the type given in Table 6.6 or 6.7 give the following first six terms in the series for f,

$$\begin{aligned} f^{(1)} &= \frac{1}{2m}\alpha, \\ f^{(2)} &= 0, \\ f^{(3)} &= \frac{1}{4m^2}[\phi\alpha - \alpha\phi] - \frac{1}{8m^3}\alpha\alpha^\dagger\alpha, \\ f^{(4)} &= 0, \\ f^{(5)} &= \frac{1}{8m^3}[\phi^2\alpha+\alpha\phi^2-2\phi\alpha\phi] + \frac{1}{16m^4}[-2\phi\alpha\alpha^\dagger\alpha+2\alpha\alpha^\dagger\alpha\phi+\alpha\phi\alpha^\dagger\alpha-\alpha\alpha^\dagger\phi\alpha] \\ &\quad + \frac{1}{16m^5}\alpha(\alpha^\dagger\alpha)^2, \end{aligned} \tag{A8.3}$$

and

$$f^{(6)} = 0.$$

Only odd order terms are nonvanishing in the expansion for f. Formulas of the type given in Table 6.8 can now be used to obtain,

$$\begin{aligned}
\hat{H}_A^{(0)} &= m, \\
\hat{H}_A^{(1)} &= 0, \\
\hat{H}_A^{(2)} &= \phi + \frac{1}{2m}\alpha^\dagger\alpha, \\
\hat{H}_A^{(3)} &= 0, \\
\hat{H}_A^{(4)} &= \frac{1}{4m^2}(\alpha^\dagger\phi\alpha - \alpha^\dagger\alpha\phi) - \frac{1}{8m^2}(\alpha^\dagger\alpha)^2, \\
\hat{H}_A^{(5)} &= 0, \\
\hat{H}_A^{(6)} &= \frac{1}{8m^3}(\alpha^\dagger\phi^2\alpha+\alpha^\dagger\alpha\phi^2-2\alpha^\dagger\phi\alpha\phi) \\
&\quad + \frac{1}{16m^4}(-2\alpha^\dagger\phi\alpha\alpha^\dagger\alpha-\alpha^\dagger\alpha\alpha^\dagger\phi\alpha+2(\alpha^\dagger\alpha)^2\phi+\alpha^\dagger\alpha\phi\alpha^\dagger\alpha)+\frac{1}{16m^5}(\alpha^\dagger\alpha)^3.
\end{aligned} \tag{A8.4}$$

As expected from the structure of the perturbation, only even order terms are non-vanishing in this expansion. It is seen that the fourth and sixth order terms here are explicitly non-hermitian. Comaprison of eqs. (A8.4) with the formulas in Table A8.1, obtained using the Pauli elimination method, indicates that both sets are identical.

For calculations of G_A and $\tilde{H}_A$, up to sixth order, formulas of the type given in Tables 6.10 and 6.11 are cumbersome. It

is preferable to calculate the series for g_A, $g_A^{\frac{1}{2}}$, and $g_A^{-\frac{1}{2}}$, and to use eqs. (6.14) and (6.15b) or (6.15c), respectively. The perturbation series for these metric quantities are given in Tables A8.2 through A8.4. Again, it is seen that they contain only even order terms since they are defined in only a single subspace, S_A. Equation (6.14) yields,

$$
\begin{aligned}
G_A^{(0)} &= m, \\
G_A^{(1)} &= 0, \\
G_A^{(2)} &= \phi + \frac{3}{4m}\alpha^\dagger\alpha, \\
G_A^{(3)} &= 0, \\
G_A^{(4)} &= \frac{1}{8m^2}(4\alpha^\dagger\phi\alpha - \phi\alpha^\dagger\alpha - \alpha^\dagger\alpha\phi) - \frac{1}{8m^3}(\alpha^\dagger\alpha)^2, \\
G_A^{(5)} &= 0,
\end{aligned}
\tag{A8.5}
$$

and

$$
\begin{aligned}
G_A^{(6)} = {} & \frac{1}{16m^3}(5\alpha^\dagger\phi^2\alpha - 3\alpha^\dagger\phi\alpha\phi - 3\phi\alpha^\dagger\phi\alpha + \alpha^\dagger\alpha\phi^2 + \phi^2\alpha^\dagger\alpha - \phi\alpha^\dagger\alpha\phi) \\
& + \frac{1}{32m^4}(-4\alpha^\dagger\phi\alpha\alpha^\dagger\alpha - 4\alpha^\dagger\alpha\alpha^\dagger\phi\alpha + (\alpha^\dagger\alpha)^2\phi + \phi(\alpha^\dagger\alpha)^2 + 2\alpha^\dagger\alpha\phi\alpha^\dagger\alpha) \\
& + \frac{3}{64m^5}(\alpha^\dagger\alpha)^3.
\end{aligned}
$$

Equation (6.15) yields,

$$
\begin{aligned}
\tilde{H}_A^{(0)} &= m, \\
\tilde{H}_A^{(1)} &= 0, \\
\tilde{H}_A^{(2)} &= \phi + \frac{1}{2m}\alpha^\dagger\alpha,
\end{aligned}
$$

$$\tilde{H}_A^{(3)} = 0,$$

$$\tilde{H}_A^{(4)} = \frac{1}{8m^2}(2\alpha^\dagger\phi\alpha-\alpha^\dagger\alpha\phi-\phi\alpha^\dagger\alpha) - \frac{1}{8m^3}(\alpha^\dagger\alpha)^2, \qquad \text{(A8.6)}$$

$$\tilde{H}_A^{(5)} = 0,$$

$$\tilde{H}_A^{(6)} = \frac{1}{16m^3}(2\alpha^\dagger\phi^2\alpha-2\alpha^\dagger\phi\alpha\phi-2\phi\alpha^\dagger\phi\alpha+\alpha\alpha^\dagger\phi^2+\phi^2\alpha^\dagger\alpha)$$

$$+\frac{1}{128m^4}(12\alpha^\dagger\phi\alpha\alpha^\dagger\alpha-12\alpha^\dagger\alpha\alpha^\dagger\phi\alpha+7(\alpha^\dagger\alpha)^2\phi+7\phi(\alpha^\dagger\alpha)^2+10\alpha^\dagger\alpha\phi\alpha^\dagger\alpha)$$

$$+\frac{1}{16m^5}(\alpha^\dagger\alpha)^3.$$

Both of these operators are manifestly self-adjoint. Comparison of eqs. (A8.6) with the formulas in Table A8.5 obtained using Eriksen's method (Eriksen, 1958) indicates that the Eriksen hamiltonian is identical to $\tilde{H}_A$, at least to sixth order. The transformation, V, used by DeVries (1970) to transform the Pauli hamiltonian into the Eriksen hamiltonian,

$$H_{Er} = V\, H_{Pauli}\, V^{-1}, \qquad \text{(A8.7)}$$

is given in Table A8.6. On comparison of Tables A8.4 and A8.6, the similarity transformation, V^{-1}, implied by eq. (A8.7) is seen to be identical, to fourth order, to $g_A^{-\frac{1}{2}}$.

TABLE A8.1 Pauli Hamiltonian (adapted from DeVries (1970))

$$H^{(0)}_{Pauli} = m$$

$$H^{(1)}_{Pauli} = 0$$

$$H^{(2)}_{Pauli} = \phi + \frac{1}{2m}\alpha^\dagger\alpha$$

$$H^{(3)}_{Pauli} = 0$$

$$H^{(4)}_{Pauli} = \frac{1}{4m^2}(-\alpha^\dagger\alpha\phi+\alpha^\dagger\phi\alpha) - \frac{1}{8m^3}(\alpha^\dagger\alpha)^2$$

$$H^{(5)}_{Pauli} = 0$$

$$H^{(6)}_{Pauli} = \frac{1}{8m^3}(\alpha^\dagger\alpha\phi^2-2\alpha^\dagger\phi\alpha\phi+\alpha^\dagger\phi^2\alpha) + \frac{1}{16m^4}(2(\alpha^\dagger\alpha)^2\phi-\alpha^\dagger\alpha\alpha^\dagger\phi\alpha+\alpha^\dagger\alpha\phi\alpha^\dagger\alpha-2\alpha^\dagger\phi\alpha\alpha^\dagger\alpha) + \frac{1}{16m^5}(\alpha^\dagger\alpha)^3$$

TABLE A8.2 g_A -- Non-relativistic Approximation

$$g_A^{(0)} = 1_A$$

$$g_A^{(1)} = 0$$

$$g_A^{(2)} = \frac{1}{4m^2}\alpha^\dagger\alpha$$

$$g_A^{(3)} = 0$$

$$g_A^{(4)} = \frac{1}{8m^3}(2\alpha^\dagger\phi\alpha-\phi\alpha^\dagger\alpha-\alpha^\dagger\alpha\phi) - \frac{1}{8m^4}(\alpha^\dagger\alpha)^2$$

$$g_A^{(5)} = 0$$

$$g_A^{(6)} = \frac{1}{16m^4}(3\alpha^\dagger\phi^2\alpha-3\phi\alpha^\dagger\phi\alpha-3\alpha^\dagger\phi\alpha\phi+\phi^2\alpha^\dagger\alpha+\alpha^\dagger\alpha\phi^2+\phi\alpha^\dagger\alpha\phi) + \frac{1}{32m^5}(-4\alpha^\dagger\alpha\alpha^\dagger\phi\alpha-\alpha^\dagger\phi\alpha\alpha^\dagger\alpha+3\phi(\alpha^\dagger\alpha)^2+3(\alpha^\dagger\alpha)^2\phi+2\alpha^\dagger\alpha\phi\alpha^\dagger\alpha) - \frac{5}{64m^6}(\alpha^\dagger\alpha)^3$$

TABLE A8.3 $g_A^{\frac{1}{2}}$ -- Non-relativistic Approximation

$$g_A^{\frac{1}{2}}(0) = 1_A$$

$$g_A^{\frac{1}{2}}(1) = 0$$

$$g_A^{\frac{1}{2}}(2) = \frac{1}{8m^2}\alpha^\dagger\alpha$$

$$g_A^{\frac{1}{2}}(3) = 0$$

$$g_A^{\frac{1}{2}}(4) = \frac{1}{16m^3}(2\alpha^\dagger\phi\alpha-\phi\alpha^\dagger\alpha-\alpha^\dagger\alpha\phi) - \frac{9}{128m^4}(\alpha^\dagger\alpha)^2$$

$$g_A^{\frac{1}{2}}(5) = 0$$

$$g_A^{\frac{1}{2}}(6) = \frac{1}{32m^4}(3\alpha^\dagger\phi^2\alpha-3\phi\alpha^\dagger\phi\alpha-3\alpha^\dagger\phi\alpha\phi+\phi^2\alpha^\dagger\alpha+\alpha^\dagger\alpha\phi^2+\phi\alpha^\dagger\alpha\phi)$$
$$+\frac{1}{256m^5}(-18\alpha^\dagger\alpha\alpha^\dagger\phi\alpha-18\alpha^\dagger\phi\alpha\alpha^\dagger\alpha+13\phi(\alpha^\dagger\alpha)^2+13(\alpha^\dagger\alpha)^2\phi+10\alpha^\dagger\alpha\phi\alpha^\dagger\alpha)$$
$$+\frac{49}{1024m^6}(\alpha^\dagger\alpha)^3$$

TABLE A8.4 $g_A^{-\frac{1}{2}}$ -- Non-relativistic Approximation

$$g_A^{-\frac{1}{2}}(0) = 1_A$$

$$g_A^{-\frac{1}{2}}(1) = 0$$

$$g_A^{-\frac{1}{2}}(2) = -\frac{1}{8m^2}\alpha^\dagger\alpha$$

$$g_A^{-\frac{1}{2}}(3) = 0$$

$$g_A^{-\frac{1}{2}}(4) = \frac{1}{16m^3}(-2\alpha^\dagger\phi\alpha+\phi\alpha^\dagger\alpha+\alpha^\dagger\alpha\phi) + \frac{11}{128m^4}(\alpha^\dagger\alpha)^2$$

$$g_A^{-\frac{1}{2}}(5) = 0$$

$$g_A^{-\frac{1}{2}}(6) = \frac{1}{32m^4}(-3\alpha^\dagger\phi^2\alpha+3\phi\alpha^\dagger\phi\alpha+3\alpha^\dagger\phi\alpha\phi-\phi^2\alpha^\dagger\alpha-\alpha^\dagger\alpha\phi^2-\phi\alpha^\dagger\alpha\phi)$$
$$\frac{1}{256m^5}(22\alpha^\dagger\alpha\alpha^\dagger\phi\alpha+22\alpha^\dagger\phi\alpha\alpha^\dagger\alpha-15\phi(\alpha^\dagger\alpha)^2-15(\alpha^\dagger\alpha)^2\phi-14\alpha^\dagger\alpha\phi\alpha^\dagger\alpha)$$
$$-\frac{69}{1024m^6}(\alpha^\dagger\alpha)^3$$

TABLE A8.5 Eriksen Hamiltonian (adapted from DeVries, (1970))

$$H_{Er}^{(0)} = m$$

$$H_{Er}^{(1)} = 0$$

$$H_{Er}^{(2)} = \phi + \frac{1}{2m}\alpha^{\dagger}\alpha$$

$$H_{Er}^{(3)} = 0$$

$$H_{Er}^{(4)} = -\frac{1}{8m^2}(\alpha^{\dagger}\alpha\phi - 2\alpha^{\dagger}\phi\alpha + \phi\alpha^{\dagger}\alpha) - \frac{1}{8m^3}(\alpha^{\dagger}\alpha)^2$$

$$H_{Er}^{(5)} = 0$$

$$H_{Er}^{(6)} = \frac{1}{16m^3}(\alpha^{\dagger}\alpha\phi^2 + \phi^2\alpha^{\dagger}\alpha - 2\alpha^{\dagger}\phi\alpha\phi - 2\phi\alpha^{\dagger}\phi\alpha + 2\alpha^{\dagger}\phi^2\alpha)$$

$$+\frac{1}{128m^4}(7(\alpha^{\dagger}\alpha)^2\phi + 7\phi(\alpha^{\dagger}\alpha)^2 - 12\alpha^{\dagger}\alpha\alpha^{\dagger}\phi\alpha - 12\alpha^{\dagger}\phi\alpha\alpha^{\dagger}\alpha + 10\alpha^{\dagger}\alpha\phi\alpha^{\dagger}\alpha)$$

$$+\frac{1}{16m^5}(\alpha^{\dagger}\alpha)^3$$

TABLE A8.6 Transformation Connecting H_{Pauli} and H_{Er}

(adapted from DeVries, (1970))

$$V_A^{(0)} = 1_A$$

$$V_A^{(1)} = 0$$

$$V_A^{(2)} = \frac{1}{8m^2}\alpha^{\dagger}\alpha$$

$$V_A^{(3)} = 0$$

$$V_A^{(4)} = \frac{1}{16m^3}(2\alpha^{\dagger}\phi\alpha - \alpha^{\dagger}\alpha\phi - \phi\alpha^{\dagger}\alpha) - \frac{9}{128m^4}(\alpha^{\dagger}\alpha)^2$$

APPENDIX 9

Additional Perturbation Series -- Non-orthonormal Basis

In this appendix, some alternative perturbation formulas, applicable in the case of a non-orthonormal basis, are derived and listed. In particular, the series for the metric g_A, and its powers, and two sets of alternative formulas for the operators $\tilde{H}_A^{(n)}$, are given here.

Formulas for the perturbation series for g_A in terms of f and S are obtained straightforwardly by expanding eq. (2.103a) to obtain

$$g_A = \sum_{n=0}^{\infty} g_A^{(n)},$$

where

$$g_A^{(n)} = S_{AA}^{(n)} + \sum_{j=1}^{n-1} [S_{AB}^{(n-j)} f^{(j)} + f^{(j)\dagger} (S_{BA}^{(n-j)} + \sum_{i=1}^{n-j} S_{BB}^{(n-j-i)} f^{(i)})] . \tag{A9.1}$$

Explicit expressions for several low order terms of (A9.1) are given in Table A9.1. It is seen that g_A now contains a nonzero first order term.

Similar explicit expressions could be obtained for the matrices g_A^{-1}, $g_A^{\frac{1}{2}}$, and $g_A^{-\frac{1}{2}}$. They rapidly become even more lengthy than those in Table A9.1, and lose their usefulness. However, eqs.(A7.4) - (A7.7) still hold, and can be used here to express the perturbation series for these powers of g_A in terms of the series for g_A itself. Such formulas, given in Tables A9.2 - A9.4, are seen to be very similar to the corresponding formulas in an orthonormal basis, given in Tables

A7.4 - A7.6. They are more lengthy generally, because of the presence of the first order term in g_A.

Finally, formulas such as those given in Tables A9.2 - A9.4 can again be used to obtain useful alternative formulas for the $\tilde{H}_A^{(n)}$ in terms of the $g_A^{(n)}$ and either $\hat{H}_A^{(n)}$ or $G_A^{(n)}$. Low order expressions of this type are given in Tables A9.5 and A9.6.

TABLE A9.1 $g_A^{(n)}$ -- Non-orthonormal Basis

$$g_A^{(0)} = 1_A$$

$$g_A^{(1)} = S_{AA}^{(1)}$$

$$g_A^{(2)} = f^{(1)} \dagger f^{(1)} + (S_{AB}^{(1)} f^{(1)} + f^{(1)} \dagger S_{BA}^{(1)}) + S_{AA}^{(2)}$$

$$g_A^{(3)} = f^{(1)} \dagger f^{(2)} + f^{(2)} \dagger f^{(1)} + (f^{(2)} \dagger S_{BA}^{(1)} + S_{AB}^{(1)} f^{(2)}) + f^{(1)} \dagger S_{BB}^{(1)} f^{(1)} + (S_{AB}^{(2)} f^{(1)} + f^{(1)} \dagger S_{BA}^{(2)}) + S_{AA}^{(3)}$$

$$g_A^{(4)} = f^{(1)} \dagger f^{(3)} + f^{(2)} \dagger f^{(2)} + f^{(3)} \dagger f^{(1)} + (f^{(3)} \dagger S_{BA}^{(1)} + S_{AB}^{(1)} f^{(3)}) + (f^{(2)} \dagger S_{BA}^{(2)} + S_{AB}^{(2)} f^{(2)}) + f^{(2)} \dagger S_{BB}^{(1)} f^{(1)} + f^{(1)} \dagger S_{BB}^{(1)} f^{(2)} + f^{(1)} \dagger S_{BB}^{(2)} f^{(1)} + (f^{(1)} \dagger S_{BA}^{(3)} + S_{AB}^{(3)} f^{(1)}) + S_{AA}^{(4)}$$

$$g_A^{(5)} = f^{(1)} \dagger f^{(4)} + f^{(2)} \dagger f^{(3)} + f^{(3)} \dagger f^{(2)} + f^{(4)} \dagger f^{(1)} + (f^{(4)} \dagger S_{BA}^{(1)} + S_{AB}^{(1)} f^{(4)}) + (f^{(3)} \dagger S_{BA}^{(2)} + S_{AB}^{(2)} f^{(3)}) + f^{(3)} \dagger S_{BB}^{(1)} f^{(1)} + f^{(1)} \dagger S_{BB}^{(3)} f^{(1)} + f^{(1)} \dagger S_{BB}^{(1)} f^{(3)} + (f^{(2)} \dagger S_{BA}^{(3)} + S_{AB}^{(3)} f^{(2)}) + f^{(2)} \dagger S_{BB}^{(2)} f^{(1)} + f^{(2)} \dagger S_{BB}^{(1)} f^{(2)} + f^{(1)} \dagger S_{BB}^{(2)} f^{(2)} + (f^{(1)} \dagger S_{BA}^{(4)} + S_{AB}^{(4)} f^{(1)}) + S_{AA}^{(5)}$$

TABLE A9.2 $g_A^{\frac{1}{2}(n)}$ -- Non-orthonormal Basis

$$g_A^{\frac{1}{2}(0)} = 1_A$$

$$g_A^{\frac{1}{2}(1)} = \tfrac{1}{2} g_A^{(1)}$$

$$g_A^{\frac{1}{2}(2)} = \tfrac{1}{2} g_A^{(2)} - \tfrac{1}{8} g_A^{(1)2}$$

$$g_A^{\frac{1}{2}(3)} = \tfrac{1}{2} g_A^{(3)} - \tfrac{1}{8}\{g_A^{(1)}, g_A^{(2)}\}_+ + \tfrac{1}{16} g_A^{(1)3}$$

$$g_A^{\frac{1}{2}(4)} = \tfrac{1}{2} g_A^{(4)} - \tfrac{1}{8}\{g_A^{(1)}, g_A^{(3)}\}_+ - \tfrac{1}{8} g_A^{(2)2} + \tfrac{1}{16}\{g_A^{(1)2}, g_A^{(2)}\}_+ + \tfrac{1}{16} g_A^{(1)} g_A^{(2)} g_A^{(1)} - \tfrac{5}{128} g_A^{(1)4}$$

$$g_A^{\frac{1}{2}(5)} = \tfrac{1}{2} g_A^{(5)} - \tfrac{1}{8}\{g_A^{(1)}, g_A^{(4)}\}_+ - \tfrac{1}{8}\{g_A^{(2)}, g_A^{(3)}\}_+ + \tfrac{1}{16}\{g_A^{(1)2}, g_A^{(3)}\}_+ + \tfrac{1}{16}\{g_A^{(1)}, g_A^{(2)2}\}_+ + \tfrac{1}{16} g_A^{(1)} g_A^{(3)} g_A^{(1)} + \tfrac{1}{16} g_A^{(2)} g_A^{(1)} g_A^{(2)} - \tfrac{5}{128}\{g_A^{(1)3}, g_A^{(2)}\}_+ - \tfrac{5}{128}(g_A^{(1)} g_A^{(2)} g_A^{(1)2} + g_A^{(1)2} g_A^{(2)} g_A^{(1)}) - \tfrac{1}{256} g_A^{(1)5}$$

TABLE A9.3 $g_A^{-\frac{1}{2}(n)}$ -- Non-orthonormal Basis

$$g_A^{-\frac{1}{2}(0)} = 1_A$$

$$g_A^{-\frac{1}{2}(1)} = -\tfrac{1}{2} g_A^{(1)}$$

$$g_A^{-\frac{1}{2}(2)} = -\tfrac{1}{2} g_A^{(2)} + \tfrac{3}{8} g_A^{(1)2}$$

$$g_A^{-\frac{1}{2}(3)} = -\tfrac{1}{2} g_A^{(3)} + \tfrac{3}{8}\{g_A^{(1)}, g_A^{(2)}\}_+ - \tfrac{5}{16} g_A^{(1)3}$$

$$g_A^{-\frac{1}{2}(4)} = -\tfrac{1}{2} g_A^{(4)} + \tfrac{3}{8}\{g_A^{(1)}, g_A^{(3)}\}_+ + \tfrac{3}{8} g_A^{(2)2} - \tfrac{5}{16}\{g_A^{(1)2}, g_A^{(2)}\}_+ - \tfrac{5}{16} g_A^{(1)} g_A^{(2)} g_A^{(1)} + \tfrac{35}{128} g_A^{(1)4}$$

$$g_A^{-\frac{1}{2}(5)} = -\tfrac{1}{2} g_A^{(5)} + \tfrac{3}{8}\{g_A^{(1)}, g_A^{(4)}\}_+ + \tfrac{3}{8}\{g_A^{(2)}, g_A^{(3)}\}_+ - \tfrac{5}{16}\{g_A^{(1)2}, g_A^{(3)}\}_+ - \tfrac{5}{16} g_A^{(1)} g_A^{(3)} g_A^{(1)} - \tfrac{5}{16}\{g_A^{(2)2}, g_A^{(1)}\}_+ - \tfrac{5}{16} g_A^{(2)} g_A^{(1)} g_A^{(2)} + \tfrac{35}{128}\{g_A^{(1)} g_A^{(2)} + g_A^{(2)} g_A^{(1)}, g_A^{(1)2}\}_+ - \tfrac{63}{256} g_A^{(1)5}$$

TABLE A9.4 $g_A^{-1(n)}$ -- Non-orthonormal Basis

$$g_A^{-1(0)} = 1_A$$

$$g_A^{-1(1)} = -g_A^{(1)}$$

$$g_A^{-1(2)} = -g_A^{(2)} + g_A^{(1)2}$$

$$g_A^{-1(3)} = -g_A^{(3)} + \{g_A^{(2)}, g_A^{(1)}\}_+ - g_A^{(1)3}$$

$$g_A^{-1(4)} = -g_A^{(4)} + \{g_A^{(3)}, g_A^{(1)}\}_+ + g_A^{(2)2} - \{g_A^{(1)2}, g_A^{(3)}\}_+ - g_A^{(1)} g_A^{(2)} g_A^{(1)} + g_A^{(1)4}$$

$$g_A^{-1(5)} = -g_A^{(5)} + \{g_A^{(4)}, g_A^{(1)}\}_+ + \{g_A^{(3)}, g_A^{(2)}\}_+ - \{g_A^{(1)2}, g_A^{(3)}\}_+ - g_A^{(1)} g_A^{(3)} g_A^{(1)}$$
$$- \{g_A^{(2)2}, g_A^{(1)}\}_+ - g_A^{(2)} g_A^{(1)} g_A^{(2)} + \{g_A^{(1)} g_A^{(2)} + g_A^{(2)} g_A^{(1)}, g_A^{(1)2}\}_+ - g_A^{(1)5}$$

TABLE A9.5 $\tilde{H}_A^{(n)}$ -- Non-orthonormal Basis

$$\tilde{H}_A^{(0)} = H_{AA}^{(0)}$$

$$\tilde{H}_A^{(1)} = \hat{H}_A^{(1)} + \tfrac{1}{2}[g_A^{(1)}, H_{AA}^{(0)}]_-$$

$$\tilde{H}_A^{(2)} = \hat{H}_A^{(2)} + \tfrac{1}{2}[g_A^{(2)}, H_{AA}^{(0)}]_- + \tfrac{1}{2}[g_A^{(1)}, \hat{H}_A^{(1)}]_-$$
$$- \tfrac{1}{8} g_A^{(1)2} H_{AA}^{(0)} - \tfrac{1}{4} g_A^{(1)} H_{AA}^{(0)} g_A^{(1)} + \tfrac{3}{8} H_{AA}^{(0)} g_A^{(1)2}$$

$$\tilde{H}_A^{(3)} = \hat{H}_A^{(3)} + \tfrac{1}{2}[g_A^{(3)}, H_{AA}^{(0)}]_- + \tfrac{1}{2}[g_A^{(2)}, \hat{H}_A^{(1)}]_- + \tfrac{1}{2}[g_A^{(1)}, \hat{H}_A^{(2)}]_- - \tfrac{1}{8} g_A^{(1)2} \hat{H}_A^{(1)}$$
$$- \tfrac{1}{4} g_A^{(1)} \hat{H}_A^{(1)} g_A^{(1)} + \tfrac{3}{8} \hat{H}_A^{(1)} g_A^{(1)2} - \tfrac{1}{8} \{g_A^{(1)}, g_A^{(2)}\}_+ H_{AA}^{(0)} + \tfrac{3}{8} H_{AA}^{(0)} \{g_A^{(1)}, g_A^{(2)}\}_+$$
$$- \tfrac{1}{4} g_A^{(1)} H_{AA}^{(0)} g_A^{(2)} - \tfrac{1}{4} g_A^{(2)} H_{AA}^{(0)} g_A^{(1)} + \tfrac{1}{16} g_A^{(1)3} H_{AA}^{(0)} - \tfrac{5}{16} H_{AA}^{(0)} g_A^{(1)3}$$
$$+ \tfrac{1}{16} g_A^{(1)2} H_{AA}^{(0)} g_A^{(1)} + \tfrac{3}{16} g_A^{(1)} H_{AA}^{(0)} g_A^{(1)2}$$

TABLE A9.6 $\tilde{H}_A^{(n)}$ -- Non-orthonormal Basis

$$\tilde{H}_A^{(0)} = H_{AA}^{(0)}$$

$$\tilde{H}_A^{(1)} = G_A^{(1)} - \tfrac{1}{2}\{g_A^{(1)}, G_A^{(0)}\}_+$$

$$\tilde{H}_A^{(2)} = G_A^{(2)} - \tfrac{1}{2}\{g_A^{(2)}, G_A^{(0)}\}_+ + \tfrac{1}{4} g_A^{(1)} G_A^{(0)} g_A^{(1)} - \tfrac{1}{2}\{g_A^{(1)}, G_A^{(1)}\}_+ + \tfrac{3}{8}\{g_A^{(1)2}, G_A^{(0)}\}_+$$

$$\begin{aligned}\tilde{H}_A^{(3)} = {} & G_A^{(3)} - \tfrac{1}{2}\{g_A^{(3)}, G_A^{(0)}\}_+ - \tfrac{1}{2}\{g_A^{(2)}, G_A^{(1)}\}_+ - \tfrac{1}{2}\{g_A^{(1)}, G_A^{(2)}\}_+ \\ & + \tfrac{3}{8}\{\{g_A^{(1)}, g_A^{(2)}\}_+, G_A^{(0)}\}_+ + \tfrac{3}{8}\{g_A^{(1)2}, G_A^{(1)}\}_+ - \tfrac{5}{16}\{g_A^{(1)3}, G_A^{(0)}\}_+ \\ & + \tfrac{1}{4} g_A^{(1)} G_A^{(1)} g_A^{(1)} + \tfrac{1}{4} g_A^{(2)} G_A^{(0)} g_A^{(1)} + \tfrac{1}{4} g_A^{(1)} G_A^{(0)} g_A^{(2)} \\ & - \tfrac{3}{16} g_A^{(1)2} G_A^{(0)} g_A^{(1)} - \tfrac{3}{16} g_A^{(1)} G_A^{(0)} g_A^{(1)2}\end{aligned}$$

APPENDIX 10

Self-Consistent Perturbation Theory When $F^{(0)}$ is not Block Diagonal

The requirement that the zero order part of the Fock matrix be at least block diagonal was imposed in section 7.4 for reasons of convenience rather than necessity. The basic changes in the formalism resulting from a relaxation of that requirement will be summarized here.

If $F^{(0)}$ has nonzero off-diagonal blocks, eq. (7.23) implies the existence of a zero order term in the series for f, given by the equation

$$D^{(0)}(f)=F_{BA}^{(0)}+F_{BB}^{(0)}f^{(0)}-f^{(0)}F_{AA}^{(0)}-f^{(0)}F_{AB}^{(0)}f^{(0)}=0. \quad \text{(A10.1)}$$

This equation has a non-zero solution $f^{(0)}$ in general, if $F_{BA}^{(0)} \neq 0$, because it is just the defining equation for the mapping $f^{(0)}$ corresponding to the non-block diagonal $F^{(0)}$.

In the coupled Hartree-Fock perturbation formalism, the n^{th} order equation (defining $f^{(n)}$) now becomes

$$D^{(n)}(f) = F_{BA}^{(n)}+\sum_{j=0}^{n}(F_{BB}^{(n-j)}f^{(j)}-f^{(j)}F_{AA}^{(n-j)})$$

$$+\sum_{i=0}^{n}\sum_{j=0}^{n-i} f^{(i)}F_{AB}^{(n-i-j)}f^{(j)}$$

$$=G_{BA}(f_1^{(n)})+G_{BB}(f_1^{(n)})f^{(0)}-f^{(0)}G_{AA}(f_1^{(n)})$$

$$-f^{(0)}G_{AB}(f_1^{(n)})f^{(0)}+$$

$$+F_{BB}^{(0)}f^{(n)}-f^{(n)}F_{AA}^{(0)}-f^{(0)}F_{AB}^{(0)}f^{(n)}-f^{(n)}F_{AB}^{(0)}f^{(0)}$$

$$+F_{BA}^{(n)}(\tilde{P}_A'^{(n)})+F_{BB}^{(n)}(\tilde{P}_A'^{(n)})f^{(0)}-f^{(0)}F_{AA}^{(n)}(\tilde{P}_A'^{(n)})$$

$$-f^{(0)}F_{AB}^{(n)}(\tilde{P}_A'^{(n)})f^{(0)}+\sum_{j=1}^{n-1}(F_{BB}^{(n-j)}f^{(j)}-f^{(j)}F_{AA}^{(n-j)})$$

$$-\sum_{i=1}^{n-1}\sum_{j=1}^{n-i}f^{(j)}F_{AB}^{(n-i-j)}f^{(j)}$$

$$= 0. \qquad \text{(A10.2)}$$

These equations can be written in the simplified form

$$D_{\tau s}^{(n)}(f) = \sum_{\sigma,r} B_{\tau s\sigma r}\, f_{\sigma r}^{(n)} - C_{\tau s}^{(n)} = 0, \qquad \text{(A10.3)}$$

$$(\tau=1, \ldots, n_B;\ \ s=1, \ldots, n_A),$$

but now,

$$B_{\tau s\sigma r} = A_{\tau sr\sigma} + \sum_{\rho} A_{\tau\rho r\sigma}f_{\rho s}^{(0)} - \sum_{t} A_{tsr\sigma}f_{\tau t}^{(0)} - \sum_{\rho,t} f_{\tau t}^{(0)}A_{t\rho r\sigma}f_{\rho s}^{(0)}$$

$$+(\hat{F}_B^{\dagger}(f^{(0)}))_{\tau\sigma}\delta_{sr} - (\hat{F}_A(f^{(0)}))_{rs}\delta_{\tau\sigma}, \qquad \text{(A10.4)}$$

and

$$C^{(n)} = F_{BA}^{(n)}(\tilde{P}_A'^{(n)})+F_{BB}^{(n)}(\tilde{P}_A'^{(n)})f^{(0)}-f^{(0)}F_{AA}^{(n)}(\tilde{P}_A'^{(n)})$$

$$-f^{(0)}F_{AB}^{(n)}(\tilde{P}_A'^{(n)})f^{(0)}+\sum_{j=1}^{n-1}(F_{BB}^{(n-j)}f^{(j)}-f^{(j)}F_{AA}^{(n-j)})$$

$$-\sum_{i=1}^{n-1}\sum_{j=1}^{n-i}f^{(i)}F_{AB}^{(n-i-j)}f^{(j)}. \qquad \text{(A10.5)}$$

The operators $\hat{F}_B$, and $\hat{F}_A$ are defined formally in eqs. (2.66a) and (2.65a), respectively.

The additional complexity of eqs. (A10.3) - (A10.5) over the corresponding equations given in section 7.4 for $f^{(0)}=0$, is easily seen. Nevertheless, there are situations in which it may be desirable to use this formalism. For example, if the calculation is to be carried out in a particular basis (for instance, localized orbitals of some sort), it is probable that the zero order Fock operator is not block diagonal. It may, however, be more efficient in such a case to carry out the calculation in a second basis in which $F^{(0)}$ is at least block diagonal, and then transform the results back to the desired basis. It must be remembered that the presence of a nonzero $f^{(0)}$ invalidates all the perturbation formulas derived in chapter 7, including those for P_A' and E.

APPENDIX 11

Minimization Algorithms

Details of two minimzation algorithms referred to in Chapter 8 are given here, with particular reference to direct energy minimization calculations for closed shell systems.

A11.1 Method of Conjugate Gradients

The conjugate gradients method is a descent optimization procedure. It can be regarded as a steepest descent algorithm with memory. As is true of any descent method, the value of the object function cannot diverge here if it is bounded from below. However, convergence is not guaranteed in general.

As applied to the closed shell case, when the energy is to be minimized with respect to the elements of the operator f, the algorithm is as follows:

1. Initialization--an initial estimate of the f-operator, leading to an initial estimate of the density matrix, R, is required. An initial estimate of the Fock matrix, F(R), is calculated from this initial density matrix.
2. The energy gradient is calculated,

$$\nabla_f E = 4\, F_{\tilde{e}_B^\dagger \tilde{e}_A} \quad , \quad \text{(all quantities real).}$$

3. Given $\nabla_f E$, and the search direction used in the previous iteration, $\underline{y}^{old}$, the current search direction is calculated as

$$\underset{\sim}{v} = -\nabla_f E + \beta \underset{\sim}{v}^{old},$$

where

$$\beta = \frac{\|\nabla_f E\|}{\|\nabla_f E\|^{old}} = \frac{\sum_{\sigma,r} |\nabla_f E_{\sigma r}|^2}{\sum_{\sigma,r} |\nabla_f E_{\sigma r}^{old}|^2} .$$

If this is the first iteration (or an iteration numbered a multiple of $n_A n_B$) take $\beta = 0$, that is

$$\underset{\sim}{v} = -\nabla_f E,$$

which is the steepest descent direction.

4. Minimize $E(f + \lambda v)$ as a function of the single parameter λ, representing a step length along the current search direction. This is usually done using a cubic interpolation procedure of Davidon (see Garton and Sutcliffe, 1974).

5. Update,

$$f \rightarrow f + \lambda_{min} v,$$

and re-evaluate R and F(R). If predetermined convergence criteria have not been satisfied, return to step 2. Otherwise, exit the procedure.

The linear search is the most costly step in the calculation. It is therefore important to use interpolation schemes which do not require a large number of energy evaluations, and which make maximal use of the information available. The cubic interpolation formula will give the exact minimum of a quadratic function, and is therefore quite suitable in direct energy minimization calculations, especially near the energy minimum. In the calculations

reported in section 8.2.c, a second interpolation procedure, based on the secant method for solving nonlinear equations, was used. Given values of $\partial E/\partial\lambda$ at two points along the search direction, an approximation to the minimizing step length is given by

$$\lambda_{min} = \frac{E'(\lambda_2)\lambda_1 - E'(\lambda_1)\lambda_2}{\lambda_2 - \lambda_1} \quad . \tag{A11.1}$$

While this interpolation formula does not make use of all the information available (it uses the energy derivatives, but not the energy itself), it does have the advantage of not requiring that the energy minimum be bracketed by λ_1 and λ_2. If $E(\lambda)$ is a quadratic function, λ_{min} given by (A11.1) is exact.

Since both the cubic interpolation formula and eq. (A11.1) locate the minimum along the search direction only approximately, it is necessary to ensure that $E(f + \lambda_{min}v)$ is indeed less than $E(f)$. If this is not so, then a second interpolation on one of the two subintervals of the original interval must be carried out.

Finally, it should be noted that components of the search direction v on surfaces where E is constant can only enter via the memory term. Therefore, if the calculation is converging (that is, if $\nabla_f E$ is decreasing), then $\beta < 1$, and these components are attenuated in succeeding iterations.

A11.2 The Newton-Raphson Method

The application of the Newton-Raphson method to the closed shell self-consistent field calculation involves a different

strategy for determining stationary values of the energy, namely, solving for the roots of the system of simultaneous nonlinear equations $F_{\tilde{e}_B^\dagger \tilde{e}_A} = 0$. This method is not a descent method, and does not necessarily converge to an energy minimum. The overall algorithm as applied to the closed shell case can be summarized as follows:

1. Initialization--same as for the conjugate gradient method.
2. The energy gradient is calculated,

$$\nabla_f E = 4\, F_{\tilde{e}_B^\dagger \tilde{e}_A}, \quad \text{(all quantities real)}.$$

3. The Jacobian matrix is calculated (the Hessian matrix of the energy),

$$J_{\sigma r,\tau s} = \frac{\partial^2 E}{\partial f_{\sigma r} \partial f_{\tau s}}.$$

4. The Newton-Raphson equations,

$$J\delta f = -\nabla_f E,$$

are solved for the elements of the correction δf to f.

5. The f-operator is updated, $f \rightarrow f + \delta f$, and new estimates of R and F(R) calculated. If the prescribed convergence criteria are satisfied at this point, the calculation is terminated. Otherwise, return to step 2.

The Newton-Raphson algorithm is conceptually simple to implement in the sense that there is no ambiguity present like

that associated with the linear search step in the conjugate gradient method. It is second order convergent; one Newton-Raphson iteration being roughly equivalent, in principle, to m conjugate gradient iterations, where m is the number of independent variables in the problem (Daniel, 1965). However, the large amount of computation required per iteration as m becomes large tends to offset the rapid rate of convergence, and it is generally considered inapplicable for application to self-consistent field calculations, as outlined above.

APPENDIX 12

Derivatives With Respect to Real and Imaginary Parts of f.

Most of the formulas derived in this chapter have been in terms of the elements of f and their complex conjugates. Under some circumstances, it is more useful to rewrite these formulas in terms of the real and imaginary parts of f, denoted here as f^R and f^I. If a real basis set is used, it is necessary to have derivatives of the energy only with respect to the real part of f. The formulas for obtaining these derivatives from the previously obtained ones are summarized here.

Writing

$$f_{\sigma r} = f^R_{\sigma r} + i f^I_{\sigma r} , \qquad f^*_{\sigma r} = f^R_{\sigma r} - i f^I_{\sigma r} ,$$

one has

$$\frac{\partial}{\partial f^R_{\sigma r}} = \frac{\partial}{\partial f_{\sigma r}} + \frac{\partial}{\partial f^*_{\sigma r}} , \qquad \frac{\partial}{\partial f^I_{\sigma r}} = i\left(\frac{\partial}{\partial f_{\sigma r}} - \frac{\partial}{\partial f^*_{\sigma r}}\right),$$

and

$$\frac{\partial^2}{\partial f^R_{\sigma r}\partial f^R_{\tau s}} = \frac{\partial^2}{\partial f_{\sigma r}\partial f_{\tau s}} + \frac{\partial^2}{\partial f_{\sigma r}\partial f^*_{\tau s}} + \frac{\partial^2}{\partial f^*_{\sigma r}\partial f_{\tau s}} + \frac{\partial^2}{\partial f^*_{\sigma r}\partial f^*_{\tau s}} ,$$

$$\frac{\partial^2}{\partial f^I_{\sigma r}\partial f^I_{\tau s}} = -\frac{\partial^2}{\partial f_{\sigma r}\partial f_{\tau s}} + \frac{\partial^2}{\partial f_{\sigma r}\partial f^*_{\tau s}} + \frac{\partial^2}{\partial f^*_{\sigma r}\partial f_{\tau s}} - \frac{\partial^2}{\partial f^*_{\sigma r}\partial f^*_{\tau s}} ,$$

and

$$\frac{\partial^2}{\partial f^R_{\sigma r}\partial f^I_{\tau s}} = \frac{\partial^2}{\partial f_{\sigma r}\partial f_{\tau s}} - \frac{\partial^2}{\partial f_{\sigma r}\partial f^*_{\tau s}} + \frac{\partial^2}{\partial f^*_{\sigma r}\partial f_{\tau s}} - \frac{\partial^2}{\partial f^*_{\sigma r}\partial f^*_{\tau s}} .$$

It is worth noting that if both E and f are real, then $\partial E/\partial f^I$ vanishes.

APPENDIX 13

Covariant and Contravariant Representations -- The General Case

An analysis of the metric properties of the non-orthonormal molecular orbitals defined im eq. (8.61) for a general multi-partitioning, can be carried out in a manner analogous to that of section 2.1.d. The major formulas only are summarized here.

We have

$$\tilde{e}^{(i)}_{JK} = (1 - R^{(i)}S)_{JK}$$

$$= 1_I\delta_{JK} - f^{(i)}_{JI}g^{(i)-1}_{I}\sum_{P=1}^{m+1} f^{(i)\dagger}_{PI}S_{PK}, \quad K\neq I, \qquad \text{(A13.1a)}$$

and

$$\tilde{e}^{(i)}_{JI} = R^{(i)}_{JI}$$

$$= f^{(i)}_{JI}g^{(i)-1}_{I}. \qquad \text{(A13.1b)}$$

Writing

$$\tilde{g}^{(i)} = \tilde{e}^{(i)\dagger}\tilde{e}^{(i)}, \qquad \text{(A13.2a)}$$

ome obtains

$$\tilde{g}^{(i)}_{II} = \sum_{L=1}^{m+1} g^{(i)-1}_{I} f^{(i)\dagger}_{LI} f^{(i)}_{LI} g^{(i)-1}_{I}$$

$$= (R^{(i)2})_{II},$$

$$\tilde{g}^{(i)}_{JI} = \Big[f^{(i)}_{JI} - \sum_{P=1}^{m+1} S_{JP}f^{(i)}_{PI}g^{(i)-1}_{I}\sum_{L=1}^{m+1} f^{(i)\dagger}_{LI}f^{(i)}_{LI}\Big]g^{(i)-1}_{I}$$

$$= (R^{(i)} - SR^{(i)2})_{JI} = \tilde{g}^{(i)\dagger}_{IJ}, \quad (J \neq I),$$

and

$$\tilde{g}^{(i)}_{JK} = \delta_{JK} - \sum_{P=1}^{m+1} [f^{(i)}_{JI} g^{(i)-1}_{I} f^{(i)\dagger}_{PI} S_{PK} - S_{JP} f^{(i)}_{PI} g^{(i)-1}_{I} f^{(i)\dagger}_{KI}]$$

$$+ \sum_{P,P',L=1}^{m+1} S_{JP} f^{(i)}_{PI} g^{(i)-1}_{I} f^{(i)\dagger}_{LI} f^{(i)}_{LI} g^{(i)-1}_{I} f^{(i)\dagger}_{P'I} S_{P'K}$$

$$= (1 - R^{(i)}S - SR^{(i)} - SR^{(i)2}S)_{JK}, \quad (J,K \neq I), \tag{A13.2b}$$

demonstrating the non-orthonormality of the $\tilde{e}^{(i)}$ with respect to the identity in general.

A set of vectors, $\underset{\sim}{e}^{(i)}$, dual to the $\tilde{e}^{(i)}$, are given by

$$\underset{\sim}{e}^{(i)}_{LM} = \delta_{LM} + \delta_{IM} \sum_{P=1}^{m+1} S_{LP} f^{(i)}_{PI}, \quad (L \neq I),$$

$$\underset{\sim}{e}^{(i)}_{IM} = -f^{(i)\dagger}_{MI}, \quad (M \neq I), \tag{A13.3}$$

and

$$\underset{\sim}{e}^{(i)}_{II} = \sum_{P=1}^{m+1} S_{IP} f^{(i)}_{PI}.$$

They are also non-orthonormal with respect to the identity,

$$\underset{\sim}{e}^{(i)\dagger} \underset{\sim}{e}^{(i)} = \underset{\sim}{g}^{(i)}, \tag{A13.4a}$$

where

$$\underset{\sim}{g}^{(i)}_{LM} = \delta_{LM} + f^{(i)}_{LI} f^{(i)\dagger}_{MI}, \quad (L,M \neq I),$$

$$\underset{\sim}{g}^{(i)}_{LI} = \sum_{P=1}^{m+1} S_{LP} f^{(i)}_{PI} - f^{(i)}_{LI} \sum_{P=1}^{m} S_{IP} f^{(i)}_{PI}$$

$$= g^{(i)\dagger}_{IL}, \quad (L \neq I), \tag{A13.4b}$$

and

$$\underset{\sim}{g}^{(i)}_{II} = \sum_{J,P,P'=1}^{m+1} f^{(i)\dagger}_{PI} S_{PJ} S_{JP'} f^{(i)}_{P'I}.$$

However, these two sets of contragredient vectors can be used to construct metric matrices, with respect to which they are orthonormal. In detail,

$$\tilde{e}^{(i)\dagger}\,\underset{\sim}{\Delta}^{(i)}\tilde{e}^{(i)} = 1, \tag{A13.5a}$$

where the blocks of $\underset{\sim}{\Delta}^{(i)} = \underset{\sim}{e}^{(i)}\underset{\sim}{e}^{(i)\dagger}$ are

$$\underset{\sim}{\Delta}^{(i)}_{LM} = \delta_{LM} + \sum_{P,P'=1}^{m+1} S_{LP} f^{(i)}_{PI} f^{(i)\dagger}_{P'I} S_{P'M}, \quad (L,M \neq I),$$

$$\underset{\sim}{\Delta}^{(i)}_{LI} = -f^{(i)}_{LI} + \sum_{P,P'=1}^{m+1} S_{LP} f^{(i)}_{PI} f^{(i)\dagger}_{P'I} S_{P'I} = \underset{\sim}{\Delta}^{(i)\dagger}_{IL}, \quad (L \neq I),$$

and (A13.5b)

$$\underset{\sim}{\Delta}^{(i)}_{II} = \sum_{\substack{J=1 \\ \neq I}}^{m+1} f^{(i)\dagger}_{JI} f^{(i)}_{JI} + \sum_{P,P'=1}^{m} S_{IP} f^{(i)}_{PI} f^{(i)\dagger}_{P'I} S_{P'I}.$$

Similarly,

$$\underset{\sim}{e}^{(i)\dagger}\,\tilde{\Delta}^{(i)}\underset{\sim}{e}^{(i)} = 1, \tag{A13.6a}$$

where the blocks of $\tilde{\Delta}^{(i)} = \tilde{e}^{(i)}\tilde{e}^{(i)\dagger}$ are,

$$\tilde{\Delta}^{(i)}_{LM} = \delta_{LM} - f^{(i)}_{LI} g^{(i)-1}_{I} \sum_{P=1}^{m+1} f^{(i)\dagger}_{PI} S_{PM} - \Big(\sum_{P=1}^{m+1} S_{LP} f^{(i)}_{PI}\Big) g^{(i)-1}_{I} f^{(i)\dagger}_{MI}$$

$$+ f^{(i)}_{LI} g^{(i)-1}_{I} \Big[1 + \sum_{\substack{J=1 \\ \neq I}}^{m+1} \sum_{P,P'=1} (f^{(i)\dagger}_{PI} S_{PJ})(S_{JP'} f^{(i)}_{P'I})\Big] g^{(i)-1}_{I} f^{(i)\dagger}_{MI}$$

$$= \big[1 - R^{(i)}S - SR^{(i)}\big]_{LM} + \sum_{J\neq I}^{m+1} (R^{(i)}S)_{LJ}(SR^{(i)})_{JM} + R^{(i)}_{LI} R^{(i)}_{IM},$$

$$(L, M \neq I),$$

(A13.6b)

$$\tilde{\Delta}_{LI}^{(i)} = f_{LI}^{(i)} g_I^{(i)-1} \sum_{\substack{JJ\neq I \\ P,P'=1}}^{m+1} (f_{PI}^{(i)\dagger} S_{PJ})(S_{JP'} f_{P'I}^{(i)}) g_I^{(i)-1}$$

$$- \sum_{P=1}^{m+1} S_{LP} f_{PI}^{(i)} g_I^{(i)-1} + f_{LI}^{(i)} g_I^{(i)-2}$$

$$= \tilde{\Delta}_{IL}^{(i)\dagger}, \quad (L \neq I),$$

and

$$\tilde{\Delta}_{II}^{(i)} = g_I^{(i)-1} \left[1 + \sum_{\substack{J\neq I \\ P,P'=1}}^{m+1} (f_{PI}^{(i)\dagger} S_{PJ})(S_{JP'} f_{P'I}^{(i)})\right] g_I^{(i)-1}.$$

It is seen that $\tilde{g}^{(i)} \neq \tilde{\Delta}^{(i)}$, and $\underset{\sim}{g}^{(i)} \neq \underset{\sim}{\Delta}^{(i)}$, so that the matrices $\tilde{e}^{(i)}$ and $\underset{\sim}{e}^{(i)}$ are not normal in this general case, where the fixed basis is non-orthonormal.

The above formulas simplify greatly if the fixed basis is orthonormal. One obtains,

$$\left.\begin{aligned} \tilde{e}_{JK}^{(i)} &= (1 - R^{(i)})_{JK}, \quad (K\neq I), \\ \tilde{e}_{JI}^{(i)} &= R_{JI}^{(i)}, \end{aligned}\right\} \begin{array}{l} (J,K=1, \ldots, m+1), \\ (i=1, \ldots, m). \end{array} \qquad (A13.7)$$

Then, one has,

$$\tilde{g}^{(i)} = \tilde{e}^{(i)\dagger}\tilde{e}^{(i)} = \tilde{e}^{(i)}\tilde{e}^{(i)\dagger} = \tilde{\Delta}^{(i)}, \qquad (A13.8a)$$

where

$$\tilde{g}_{LM}^{(i)} = (1 - R^{(i)})_{LM}, \quad (L,M \neq I),$$

$$\tilde{g}_{LI}^{(i)} = 0 = \tilde{g}_{IL}^{(i)\dagger}, \quad (L\neq I), \qquad (A13.8b)$$

and

$$\tilde{g}_{II}^{(i)} = R_{II}^{(i)}.$$

Similarly, for the dual vectors, one obtains,

$$\underset{\sim}{e}^{(i)}_{LL} = 1_L , \qquad (L = 1, \ldots, m+1),$$

$$\underset{\sim}{e}^{(i)}_{LI} = f^{(i)}_{LI} = -\underset{\sim}{e}^{(i)\dagger}_{IL}, \tag{A13.9}$$

and

$$\underset{\sim}{e}^{(i)}_{LM} = 0, \quad (M \neq L,\ L, M \neq I).$$

Then, one has,

$$\underset{\sim}{g}^{(i)} = \underset{\sim}{e}^{(i)\dagger}\underset{\sim}{e}^{(i)} = \underset{\sim}{e}^{(i)}\underset{\sim}{e}^{(i)\dagger} = \underset{\sim}{\Delta}^{(i)}, \tag{A13.10a}$$

where

$$g^{(i)}_{LM} = \delta_{LM} + f^{(i)}_{LI} f^{(i)\dagger}_{MI}, \quad (L, M \neq I),$$

$$g^{(i)}_{LI} = 0 = g^{(i)\dagger}_{IL}, \quad (L \neq I), \tag{A13.10b}$$

and

$$g^{(i)}_{II} = g^{(i)}_{I}.$$