

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

THE KWIK  
ALGORITHM FOR  
COULOMB  
INTERACTIONS AND  
ITS APPLICATIONS

by

Jeremy P. Dombroski  
B.Sc.(Hons)

A thesis presented in partial fulfilment  
of the requirements for the degree  
of Doctor of Philosophy

Department of Chemistry  
Massey University  
1997

## ABSTRACT

The KWIK algorithm is introduced, generalised and applied to the problem of determining the Coulomb energy of  $N$  localised charge distributions. Coulomb interactions are typical of  $N$ -body problems which require the exhaustive pairing of all distributions, which leads to prohibitive computational cost scaling characteristics for large  $N$ .

The KWIK algorithm for Coulomb interactions begins by optimally separating the Coulomb operator into rapidly decaying real- and Fourier- space partitions yielding a hybrid technique not dissimilar in concept to other approximations methods. KWIK's superiority lies in that its efficiency increases with distribution size, so that large distributions become computationally advantageous for increasing accuracy.

Model calculations on a distribution consisting of one million particles using KWIK afforded energies, to high accuracy, within minutes compared with days for quadratic methods. The extension of such a feat to even larger distributions is now limited by machine hardware configurations.

Particular emphasis is placed on the application of the algorithm to Molecular Quantum Mechanics where it is illustrated that the algorithm may be applied to linearise single-point self consistent field calculations. In particular, KWIK can be used to form the Exchange matrix in linear computational cost. This has previously only been achieved by crude approximation techniques and cannot be achieved using Coulomb multipole based methods.

## ACKNOWLEDGMENTS

There are many people to whom I owe a great deal of thanks.

Firstly, I offer my sincere thanks to my supervisor Dr Peter M. W. Gill. Peter is a scholar and a gentleman, an outstanding scientist and a friend. He has guided me well throughout the last three years, maintaining an open door (or email) offering clear and well thought out assistance for all problems, and a clever knack for simplifying the complex. His enthusiasm for science, particularly quantum chemistry, has been a great inspiration. Thanks also to the second in command, Dr Ken Jolley, who took over the reins after Peter departed for Cambridge.

There have been others who have aided in increasing my understanding of research, quantum chemistry, mathematics and computing which has had a direct benefit on this thesis. Dr Benny Johnson of Q-Chem Inc., Dr Steve Taylor of the Mathematics Department at the University of Auckland and Chris White, presently at University of California at Berkeley, have been particularly helpful in this regard.

I thank my family for their continued encouragement and support which tends to be required more through difficult times, rather than during the more productive and successful ones. I imagine this thesis will end up on the coffee table as an answer to the question "... and what *exactly* is Jeremy doing at Massey?" continually posed by extended family and visitors.

To the members of Peter's research group who have had to put up with me hacking my way through system administration and bringing workstations to a standstill with my large scale calculations, I thank you for your patience. Of the group members, Dr Terry Adams is one who has had the largest positive effect on broadening my knowledge and improving my outlook. Thanks Terry for your friendship. I would also like to thank Mrs 'Bot' Gill for many fine lunches and offer my apologies to Malcolm and Stephanie who weren't allowed seconds at dinner time because the remainder was set aside to become my lunch the next day. Thanks also goes to Gwen Adams for providing me with a far superior poster than I could have possibly even attempted to construct for my trip to Paris and Cambridge.

There are also many other friends who have helped maintain a level of sanity (and often reality) outside of Massey whom I have chosen not to name for fear forgetting anyone, or placing them in the incorrect order. I must however, thank Gaile Peddie for putting up with endless sleepless nights, grumpy moods, hogging the phone-line with the modem and general ungentlemanly behaviour, especially during the write-up phase. Past flatmates have also had to suffer through similar agony.

I have been most fortunate in my bids for funding grants and scholarships which allowed me to attend four conferences, three of these international. I offer my humble thanks to the New Zealand Vice Chancellors Committee for a William Georgetti Scholarship and to Massey University for a Doctoral Scholarship. For assistance towards travel to conferences and for computer software, I wish to thank: - Q-Chem Inc., The New Zealand Institute of Chemistry, DFT-95 Conference organisers, Massey University Research Fund, Department of Chemistry, Massey University Graduate Research Fund and the Science Faculty Dean's Fund for generous contributions.

# TABLE OF CONTENTS

ABSTRACT .....	ii
ACKNOWLEDGMENTS .....	iii
TABLE OF CONTENTS .....	v
LIST OF TABLES, SCHEMES, FIGURES AND THEOREMS .....	ix
LIST OF ABBREVIATIONS .....	xiv
LIST OF PUBLICATIONS .....	xvii
NOTE ON CALCULATIONS .....	xvii
CHAPTER SUMMARIES .....	xviii
<i>Chapter One</i> .....	xviii
<i>Chapter Two</i> .....	xviii
<i>Chapter Three</i> .....	xviii
<i>Chapter Four</i> .....	xviii
<i>Chapter Five</i> .....	xix
<i>Chapter Six</i> .....	xix
<i>Chapter Seven</i> .....	xix
<i>Chapter Eight</i> .....	xix
<i>Chapter Nine</i> .....	xix
INTRODUCTION .....	1
<b>CHAPTER ONE .....</b>	<b>3</b>
THEORETICAL MOTIVATION .....	3
1.0. <i>Introduction</i> .....	3
1.1. <i>Classical Beginnings</i> .....	3
1.2. <i>Schrödinger's Equation</i> .....	4
1.3. <i>Solving Schrödinger's Equation</i> .....	5
1.4. <i>The Variation Principle and Variation Method</i> .....	7
1.5. <i>The Hartree-Fock Approximation</i> .....	8
1.5.1. <i>Restricted Closed-Shell Hartree-Fock (RHF)</i> .....	8
1.5.2. <i>Unrestricted Open-Shell Hartree-Fock (UHF)</i> .....	10
1.6. <i>Matrix Elements</i> .....	10
1.6.1. <i>Unrestricted Hartree-Fock Matrix Elements</i> .....	13
1.7. <i>The Correlation Problem</i> .....	14
1.8. <i>Conventional Methods</i> .....	15
1.8.1. <i>Configuration Interaction</i> .....	16
1.8.2. <i>Quadratic Configuration Interaction</i> .....	16
1.8.3. <i>Coupled Cluster</i> .....	17
1.8.4. <i>Perturbation Theory</i> .....	18
1.8.5. <i>Other approaches</i> .....	20
1.8.6. <i>Additional Comments</i> .....	20
1.9. <i>Density Functional Theory</i> .....	22
1.10. <i>Commentary</i> .....	25
<b>CHAPTER TWO .....</b>	<b>27</b>
TECHNOLOGICAL MOTIVATION .....	27
2.0. <i>Introduction</i> .....	27
2.1. <i>High-Performance Computing</i> .....	27
2.2. <i>Historical Prospective</i> .....	28
2.2.1. <i>Computing Facility</i> .....	28
2.2.2. <i>Programming Languages</i> .....	28
2.3. <i>Algorithm Development</i> .....	30
2.4. <i>Modern-day Computer Architectures</i> .....	31
2.4.1. <i>Vector Machines</i> .....	32
2.4.2. <i>CISC and RISC</i> .....	33
2.4.3. <i>Parallel Computers</i> .....	34
2.5. <i>Summary</i> .....	35
<b>CHAPTER THREE .....</b>	<b>36</b>

HISTORY OF THE QUANTUM CHEMICAL COULOMB PROBLEM .....	36
3.0. <i>Introduction</i> .....	36
3.1. <i>The Quantum Chemical Coulomb Problem</i> .....	37
3.1.0. <i>Introduction</i> .....	37
3.1.1. <i>Gaussian Basis Functions</i> .....	37
3.1.2. <i>Electron Repulsion Integrals Over GTO's</i> .....	39
3.1.3. <i>Linear Methods</i> .....	40
3.2. <i>The Coulomb Problem</i> .....	43
3.2.1. <i>The N Body Problem</i> .....	43
3.2.2. <i>The Ewald Summation Technique</i> .....	44
3.2.3. <i>Particle-Particle-Particle-Mesh</i> .....	45
3.2.4. <i>The Fast Multipole Method</i> .....	46
3.2.5. <i>The Fast Wavelet Transform</i> .....	47
3.3. <i>Summary</i> .....	47
<b>CHAPTER FOUR .....</b>	<b>49</b>
PARTITIONING THE COULOMB OPERATOR .....	49
4.0. <i>Introduction</i> .....	49
4.1. <i>The KWIK Algorithm</i> .....	50
4.2. <i>Adapting KWIK for Coulomb Interactions</i> .....	51
4.3. <i>Separating the operator</i> .....	52
4.4. <i>The Ultimate Separator</i> .....	55
4.5. <i>The Error Function Separator</i> .....	59
4.6. <i>Summary</i> .....	61
<b>CHAPTER FIVE .....</b>	<b>63</b>
SHORT-RANGE INTERACTIONS .....	63
5.0. <i>Introduction</i> .....	63
5.1. <i>Computational Methods</i> .....	64
5.1.0. <i>Introduction</i> .....	64
5.1.1. <i>Computing the Interaction</i> .....	65
5.1.2. <i>Verlet Neighbour Lists</i> .....	65
5.1.3. <i>Linked-Cell Method</i> .....	66
5.2. <i>Improvements to the Linked-Cell Method</i> .....	68
5.2.0. <i>Introduction</i> .....	68
5.2.1. <i>Ordering an Ordinate</i> .....	69
5.2.2. <i>Changing the Cell Geometry</i> .....	73
5.3. <i>Approximating Spheres</i> .....	73
5.3.0. <i>Introduction</i> .....	73
5.3.1. <i>New Basic Cell Structure</i> .....	73
5.3.2. <i>Cell Size Reduction</i> .....	75
5.4. <i>Maintaining Vectorisability</i> .....	81
5.5. <i>Conclusions</i> .....	84
<b>CHAPTER SIX .....</b>	<b>85</b>
LONG RANGE KWIK .....	85
6.0. <i>Introduction</i> .....	85
6.1. <i>Fourier Transform of the Long-Range Partition</i> .....	85
6.2. <i>Numerical Quadrature</i> .....	87
6.2.0. <i>Introduction</i> .....	87
6.2.1. <i>Simple Quadrature</i> .....	88
6.2.2. <i>Gauss Type Quadrature Rules</i> .....	90
6.2.3. <i>Basis Spanning</i> .....	91
6.3. <i>Integration of FT Long-Range KWIK</i> .....	91
6.3.0. <i>Introduction</i> .....	91
6.3.1. <i>The Separator Transform</i> .....	91
6.3.2. <i>Simple Quadrature</i> .....	92
6.3.3. <i>Gauss Type Rules</i> .....	94
6.3.4. <i>COP's and ROP's</i> .....	97
6.3.5. <i>The Fourier Intensity I(k) and the Random Walk Advantage</i> .....	98
6.3.6. <i>Discussion</i> .....	100
6.4. <i>Fourier Series</i> .....	101

6.4.0. Introduction .....	101
6.4.1. Fourier Coefficients .....	102
6.4.2. Recasting KWIK using a Fourier Series .....	104
6.4.3. Evaluating Fourier Coefficients in KWIK .....	105
6.4.3.1. Fourier Coefficients in One Dimension .....	105
6.4.3.2. Fourier Coefficients in Three Dimensions .....	109
6.4.4. Gibbs Phenomenon .....	111
6.5. Summary .....	113
<b>CHAPTER SEVEN .....</b>	<b>115</b>
APPLICATION OF KWIK TO DISTRIBUTIONS OF PARTICLES .....	115
7.0. Introduction .....	115
7.1. Particles in One Dimension .....	115
7.1.1. Derivation .....	116
7.1.2. Short-Range .....	117
7.1.2.0. Introduction .....	117
7.1.2.1. Interpolation .....	118
7.1.2.2. Cut-offs .....	120
7.1.2.3. Memory Requirements .....	121
7.1.2.4. Distribution Dependence .....	122
7.1.3. Long-Range .....	122
7.1.3.0. Introduction .....	122
7.1.3.1. MOP's vs FLOP's .....	123
7.1.3.2. Memory Requirements .....	123
7.1.3.3. Cut-offs .....	124
7.1.4. Charged Systems .....	125
7.1.4.0. Introduction .....	125
7.1.4.1. Illustrative Timings .....	125
7.1.4.2. The Random Walk and $I(k)$ .....	127
7.1.5. Neutral Systems .....	134
7.1.5.0. Introduction .....	134
7.1.5.1. A Random Walk Advantage? .....	135
7.1.5.2. Illustrative Timings .....	137
7.2. Particles in Three dimensions .....	141
7.2.0. Introduction .....	141
7.2.1. Derivation .....	141
7.2.2. Short-Range .....	142
7.2.2.1. Interpolation .....	142
7.2.2.2. Cut-offs .....	142
7.2.2.3. Memory Storage .....	143
7.2.3. Long-Range .....	143
7.2.3.0. Introduction .....	143
7.2.3.1. Memory Requirements .....	143
7.2.3.2. 2-D Fourier Coefficients .....	144
7.2.4. Charged Systems .....	146
7.2.4.0. Introduction .....	146
7.2.4.1. Illustrative Timings .....	146
7.3. Summary and Discussion .....	148
<b>CHAPTER EIGHT .....</b>	<b>150</b>
KWIK APPLIED TO QUANTUM CHEMISTRY .....	150
8.0. Introduction .....	150
8.1. KWIK for $O(N)$ Fock Matrix Formation .....	150
8.1.0. Introduction .....	150
8.1.1. KWIK in Dirac's Notation .....	151
8.1.2. Core Hamiltonian Matrix .....	152
8.1.3. Coulomb and Exchange Matrices .....	153
8.2. Evaluating Short-Range KWIK .....	155
8.2.0. Introduction .....	155
8.2.1. Integrals Over Gaussian S Type Functions .....	155
8.2.2. Screening Insignificant Integrals in $O(N)$ Work .....	156
8.3. Evaluating Long-Range KWIK .....	157
8.3.1. CASE and CAP's .....	157
8.3.2. Approximating the Coulomb Operator .....	159



8.4. <i>Self Consistent Field Calculations</i> .....	160
8.5. <i>Quantum Chemistry Using Slater Functions</i> .....	160
8.5.1. STO's.....	160
8.5.2. Slater Function Density.....	162
8.5.3. Stewart-Slater Atoms.....	163
8.6. <i>KWIK in Correlated methods</i> .....	164
<b>CHAPTER NINE</b> .....	<b>166</b>
CRITICAL ANALYSIS.....	166
9.0. <i>Introduction</i> .....	166
9.1. <i>Scaling</i> .....	166
9.2. <i>Significant Findings</i> .....	168
9.3. <i>KWIK vs The Rest</i> .....	169
9.3.1. KWIK vs Ewald.....	169
9.3.2. KWIK vs FMM.....	170
9.3.3. KWIK vs CFMM.....	171
9.4. <i>Further Work Directions</i> .....	172
9.5. <i>Summary</i> .....	175
<b>APPENDIX A</b> .....	<b>177</b>
DATA FOR THE SHORT-RANGE INVESTIGATIONS.....	177
<b>APPENDIX B</b> .....	<b>179</b>
DERIVATION OF NUMERICAL APPROXIMATION FOR DETERMINING THREE DIMENSIONAL FOURIER COEFFICIENTS.....	179
<b>APPENDIX C</b> .....	<b>186</b>
QUADRATURE RULES AND POLYNOMIALS FOR THE EXPONENTIAL INTEGRAL WEIGHT FUNCTION.....	186
<b>APPENDIX D</b> .....	<b>190</b>
DETAILS OF KWIK TIMINGS PRESENTED IN CHAPTER SEVEN.....	190
<b>REFERENCES</b> .....	<b>194</b>

## LIST OF TABLES, SCHEMES, FIGURES AND THEOREMS

<b>Table 1.1.</b>	Cost scaling behaviour of Conventional Correlated methods with the size (N) of the molecular system. ....	22
<b>Table 4.1.</b>	Value of the functional $J[g(\mathbf{r})]$ (4.21.3) for various separators with their respective optimal decay parameters, $\omega$ . The difference in the value of the functional between that of the optimal separator and error function is minimal. ....	59
<b>Table 5.1.</b>	Efficiencies of the linked cell method in all three dimensions. Note in 2-D and 3-D the majority of the computational effort is put into calculating insignificant interactions. ....	69
<b>Table 5.2.</b>	Efficiencies of the linked cell method with $x$ -direction ordinate ordering for all three dimensions. Note in 3-D the majority of the computational effort is still put into calculating insignificant interactions and that 2-D efficiency is still unacceptable. ....	70
<b>Table 5.3.</b>	Efficiencies of alternative cell geometries in two dimensions. ....	75
<b>Table 5.4.</b>	The efficiency of cell size reduction. * denotes the total number of cells required relative to the $k=1$ case. ....	76
<b>Table 5.5.</b>	Effectiveness of ordinate ordering is reduced on increased dimension and cell size reduction. * ordinate ordering has been applied, † some outlying cells become completely insignificant and are neglected as the simulation cell becomes more representative of a sphere/circle. ....	76
<b>Table 5.6.</b>	Theoretical efficiencies and numbers of significant near neighbour cells for subdivided cubes. ....	80
<b>Table 7.1.</b>	CPU times (seconds) for evaluating 400,000 interactions in the interval (0,4) involving the erfc function. Key: Function - denotes evaluating the erfc function itself; 1-D - the one-dimensional interaction potential; 3-D - the three-dimensional interaction including a square root; Standard - is the time for evaluating the same potential using the standard mathematical library. ....	120
<b>Table 7.2.</b>	CPU times (seconds) for a short-range KWIK calculation ( $\omega=0.008$ ) on distributions of 100,000 unit point charges. Ordered implies a binary search has been used and division indicates the level of cell reduction. ....	122
<b>Table 7.3.</b>	CPU times (seconds) for determining the Coulomb energy using KWIK for collinear, uniformly distributed, positive unit point charges for various predetermined accuracies. ....	126
<b>Table 7.4.</b>	CPU times (seconds) for determining the Coulomb energy using KWIK for collinear, uniformly distributed, positive unit point charges for various predetermined accuracies. (As reported in reference [148]). ....	127
<b>Table 7.5.</b>	CPU times (seconds) for determining the Coulomb energy for various system sizes (N) using the classical double sum quadratic method. ....	128
<b>Table 7.6.</b>	CPU times (seconds) for determining the Coulomb energy using KWIK for collinear, uniformly distributed, net-neutral systems of unit point charges for various predetermined accuracies. Decay parameter and degree of Fourier expansion determined as for charged systems. ....	138
<b>Table 7.7.</b>	Total KWIK timings for various predetermined accuracies for net-neutral, uniformly distributed particles using a fixed decay parameter of $\omega = 0.01$ . ....	139
<b>Table 7.8.</b>	Total KWIK timings for various predetermined accuracies for net-neutral, uniformly distributed particles using a fixed decay parameter of $\omega = 0.005$ . ....	139
<b>Table 7.9.</b>	Short-range KWIK energies for net-neutral and charged uniformly distributed unit point charges. A zero decay parameter corresponds to the total coulomb energy. The decay parameter of $\omega = 0.01$ corresponds to a real space cut-off of $\approx 302$ for $\epsilon = 10^{-6}$ . ....	140

<b>Table 7.10.</b>	Short-range KWIK energies for net-neutral, uniformly distributed unit point charges. The decay parameters of $\omega = 0.01$ and $0.005$ have real space cut-offs of $\approx 302$ and $603$ respectively for $\epsilon = 10^{-6}$ .....	140
<b>Table 7.11.</b>	Two dimensional KWIK Fourier coefficients for $\omega = 1$ . $k_x$ and $k_y$ correspond to the x and y components of the k-space sum.....	144
<b>Table 7.12.</b>	Two dimensional KWIK Fourier coefficients for $\omega = 0.1$ . $k_x$ and $k_y$ correspond to the x and y components of the k-space sum.....	145
<b>Table 7.13.</b>	Two dimensional KWIK Fourier coefficients for $\omega = 0$ . $k_x$ and $k_y$ correspond to the x and y components of the k-space sum.....	145
<b>Table 7.14.</b>	KWIK timings for charged systems of three dimensional, uniformly distributed unit point charges as previously reported [148].....	146
<b>Table 7.15.</b>	KWIK timings for charged systems of three dimensional, uniformly distributed unit point charges with incorporation of interpolation and a crude implementation of long-range cut-offs. ....	147
<b>Table A.1.</b>	Theoretical efficiencies and number of near neighbour cells for the use of Hexagons as the parent cell in 2-D short-range calculations. ....	177
<b>Table A.2.</b>	Theoretical efficiencies and number of near neighbour cells for the use of Squares as the parent cell in 2-D short-range calculations. ....	178
<b>Table A.3.</b>	Theoretical efficiencies and number of near neighbour cells for the use of Triangles as the parent cell in 2-D short-range calculations. ....	178
<b>Table B.1.</b>	Table of approximation coefficients for equation (B.22) as obtained using <i>Mathematica</i> [168]. ....	185
<b>Table C.1.</b>	Roots and weights for the Gauss type quadrature scheme over the weight function $E_1(k^2/4)$ . ....	188
<b>Table D.1.</b>	Details for KWIK calculations conducted on collinear charged uniformly distributed systems as presented in Table 7.3. Key: NBOX - the number of cells in the short-range calculation; M - the degree of the Fourier expansion; $\omega$ - the decay parameter; $R_c$ - the real space cut-off radius; RLONG - the relative contribution of the long range energy to the total energy; E - the total Coulomb energy; N - the size (and range) of the distribution. ....	190
<b>Table D.2.</b>	Details for KWIK calculations conducted on collinear, net-neutral uniformly distributed systems as presented in Table 7.6. Key: NBOX - the number of cells in the short-range calculation; M - the degree of the Fourier expansion; $\omega$ - the decay parameter; $R_c$ - the real space cut-off radius; RLONG - the relative contribution of the long range energy to the total energy; E - the total Coulomb energy; N - the size (and range) of the distribution. ....	191
<b>Table D.3.</b>	Details for KWIK calculations conducted on collinear, net-neutral uniformly distributed systems for fixed decay parameter $\omega = 0.01$ , as presented in Table 7.7. Key: $R_c$ - the real space cut-off radius; M - the degree of the Fourier expansion; RTIME - the relative contribution of the long-range time to the total time; RLONG - the relative contribution of the long range energy to the total energy; N - the size (and range) of the distribution. ....	192
<b>Table D.4.</b>	Details for KWIK calculations conducted on collinear, net-neutral uniformly distributed systems for fixed decay parameter $\omega = 0.005$ , as presented in Table 7.7. Key: $R_c$ - the real space cut-off radius; M - the degree of the Fourier expansion; RTIME - the relative contribution of the long-range time to the total time; RLONG - the relative contribution of the long range energy to the total energy; N - the size (and range) of the distribution. ....	193
<b>Scheme 5.1.</b>	Sorting particles into Verlet neighbour lists requires work that scales with the square of the number of particles which can be seen by the nested loop over particles above. Note also that the memory requirements for storing lists of neighbours for each particle is significant. ....	66

<b>Scheme 5.2.</b>	Sorting particles/molecules into cells in linear work using integers. ICOUNT(I) indicates the number of particles in cell I, and IPOINT(I) points to the first particle in cell I in a cell-ordered coordinate array XORD. The first three elements of X are particle coordinates, and the final entry is the charge. ....	67
<b>Scheme 5.3.</b>	The sorting of an ordinate to improve efficiency is achieved similarly to that without the sort and simply requires an extra loop over all Cells and a call to the sorting routine XSORT. This routine sorts the x ordinate into increasing value maintaining the integrity of the other coordinates and associated charge.....	71
<b>Scheme 5.4.</b>	Algorithm structure for short-range interactions with an ordered ordinate.....	72
<b>Scheme 7.1.</b>	Partial FORTRAN code illustrating the use of interpolating the erfc function in a 3-D simulation. The array T contains the interpolating constants $f_k$ which are a function of the polynomial coefficients. ....	120
<b>Scheme 7.2.</b>	Illustration of memory requirements for the one dimensional sorting process. ....	121
<b>Scheme 7.3.</b>	Illustrating the use of recursion to save FLOP's and unrolling the k loop to save on MOP's. Unrolling beyond 5 gave little improvement in speed-up on the IBM. ....	124
<b>Scheme 7.4.</b>	Illustration of memory requirements for the three dimensional sorting process. ....	143
<b>Scheme 7.5.</b>	Array storage for long-range KWIK.....	144
<b>Figure 1.1.</b>	Molecular energy level diagram for H <sub>2</sub> . ....	6
<b>Figure 1.2.</b>	Pople diagram illustrating the conventional approach to improved quality of calculation. The quality of a calculation can be improved either by an increase in the size of the basis set or by considering a greater number of configurations.....	21
<b>Figure 3.1.</b>	Radial dependence of a Gaussian vs Slater basis function. Note that the Slater has a cusp at the origin, whereas the Gaussian is flat and that the long-range behaviour of a Gaussian is flatter than a Slater.....	38
<b>Figure 4.1.</b>	Graphical representation of the short and long range characteristics of the Coulomb, short- and long- range operators. ....	53
<b>Figure 5.1.</b>	The particle denoted P need only interact with particles within the cut-off radius R <sub>C</sub> . Using the linked-cell method, particle P interacts with all the particles within the region depicted by the bold square, which is significantly more than necessary. ....	68
<b>Figure 5.2.</b>	As for Figure 5.1, the particle denoted P need only interact with particles within the cut-off radius R <sub>C</sub> . Using the ordinate ordered linked-cell method, particle P interacts only with those particles within the unshaded region. This is somewhat less than the previous requirement.....	72
<b>Figure 5.3.</b>	Dimensions of the squares, triangles and hexagons used in our investigations of alternative two dimensional cell geometries. One arrow length is equal to the cut-off distance R <sub>C</sub> .....	74
<b>Figure 5.4.</b>	Efficiency vs relative total number of cells contrasting the use of squares, triangles and hexagons as the basic geometric structure with which to divide a 2-D simulation system. ....	78
<b>Figure 5.5.</b>	A closer view of efficiency vs relative total number of cells contrasting the use of squares, triangles and hexagons over a practical range of relative total number of cells. ....	79
<b>Figure 5.6.</b>	Efficiency vs relative total number of cells for the case of using cubes and the division of cells method in 3-D. ....	81
<b>Figure 5.7.</b>	Illustration of the organisation of particles into boxes in the mega-array. The square represents the simulation cell divided into squares. Particles are organised into the mega-array (the tall column) by row with each square representing a list of particles. ....	82
<b>Figure 5.8.</b>	Scheme illustrating the importance of ordering the particles into their respective boxes. ....	83

<b>Figure 6.1.</b>	Fourier transform of the long-range operator obtained from using the error function separator. The logarithmic singularity and rapid long-range decay typify the behaviour of all the trialed separator functions.....	92
<b>Figure 6.2.</b>	Plots of the 'even' series representation of the erfc long-range operator. The true operator is the smooth line, whereas the series approximations are the oscillating lines. Note the changing frequency and amplitude of the 'tail' with increasing $M$ . .....	96
<b>Figure 6.3.</b>	Plots of the 'odd' series representation of the erfc long-range operator. The true operator is the smooth line, whereas the series approximations are the oscillating lines. Note the changing amplitude and frequency of the 'tail' with increasing $M$ . .....	97
<b>Figure 6.4.</b>	Log of the absolute value of the 1-D Fourier coefficients vs $k$ for $\omega=10$ . Note that $A_x(k,10)$ decreases by approximately 5 orders of magnitude between $k=0$ and $k=60$ . It is not until $k=10,000$ that a second such reduction is achieved. ....	108
<b>Figure 6.5.</b>	Log of the absolute value of the 1-D Fourier coefficients vs $k$ for $\omega=100$ . ....	109
<b>Figure 6.6.</b>	Illustrating the Gibbs phenomenon for the unit step function. Note that near the discontinuities the Fourier series approximation overshoots the true function significantly more than in the centre of the steps. ....	112
<b>Figure 6.7.</b>	1-D Fourier series representation of the erfc long-range operator superimposed upon a plot of the true function.....	113
<b>Figure 7.1.</b>	CPU time vs $N$ (the number of particles) for determining the Coulomb energy using the KWIK algorithm for increasing accuracy. Note the circled data point indicating a higher efficiency than expected obtained due to statistical fluctuations. Actual timings are tabulated in Table 7.3. ....	126
<b>Figure 7.2.</b>	$I(k)$ vs $k$ for discrete values of $k$ for 1,000,000 collinear <u>uniformly</u> distributed, positive, unit point charges. ....	128
<b>Figure 7.3.</b>	$I(k)$ vs $k$ for discrete values of $k$ for 100,000 collinear <u>uniformly</u> distributed, positive, unit point charges. ....	129
<b>Figure 7.4.</b>	Histogram of the uniformly distributed particles used in Figure 7.3. ....	130
<b>Figure 7.5.</b>	$I(k)$ vs $k$ for discrete values of $k$ for 100,000 collinear <u>exponentially</u> distributed, positive, unit point charges. ....	131
<b>Figure 7.6.</b>	Histogram of the exponentially distributed particles as used for Figure 7.5.....	132
<b>Figure 7.7.</b>	$I(k)$ vs $k$ for discrete values of $k$ for 100,000 collinear <u>normally</u> distributed, positive, unit point charges. ....	133
<b>Figure 7.8.</b>	Histogram of the normally distributed particles used in Figure 7.7. ....	134
<b>Figure 7.9.</b>	$I(k)$ vs $k$ for a <u>uniform</u> , net neutral distribution of 1,000,000 unit point charges. ....	135
<b>Figure 7.10.</b>	$I(k)$ vs $k$ for a <u>uniform</u> , net-neutral distribution of 100,000 unit point charges.....	136
<b>Figure 7.11.</b>	$I(k)$ vs $k$ for 100,000 <u>exponentially</u> distributed, net-neutral system of unit point charges. ....	137
<b>Figure 7.12.</b>	CPU time (seconds) vs $N$ (number of particles) for determining the total Coulomb energy using KWIK for a net-neutral, uniformly distributed system of unit point charges. The circled data point indicates a higher accuracy than expected was determined due to statistical fluctuations. Times are tabulated in Table 7.6.....	138
<b>Figure 7.13.</b>	CPU Time (seconds) vs $N$ (the number of particles). The filled symbols correspond to the computationally cheaper 1-D systems, while the open symbols the slightly more expensive 3-D systems. Like symbols represent similar accuracies. ....	148
<b>Figure 9.1.</b>	A plot of linear and $x\log(x)$ functions to illustrate that the scaling behaviour of the two are remarkably similar .....	167

**Theorem 6.1.** If a periodic function  $f(x)$  with period  $2l$  is piece-wise continuous in the interval  $-l \leq x \leq l$  and has a left-hand derivative and right-hand derivative at each point of that interval, then the Fourier series representation of  $f(x)$  is convergent. Its sum is  $f(x)$ , except at a point  $x_0$  at which  $f(x)$  is discontinuous and the sum of the series is the average of the left- and right-hand limits of  $f(x)$  at  $x_0$ . ..... 102

## LIST OF ABBREVIATIONS

1-D	One Dimensional Space
2-D	Two Dimensional Space
3-D	Three Dimensional Space
AO	Atomic Orbital
BLAS	Basis Linear Algebra Subroutines
C++	(object oriented programming language)
CAP	Coulomb Attenuated Potential
CASE	Coulomb Attenuated Schrödinger Equation
CC	Coupled Cluster
CCD	Coupled Cluster with Doubles excitations
CCSD(T)	Coupled Cluster with Singles and Doubles excitations, incorporating Triples perturbatively
CCSDT	Coupled Cluster with Singles, Doubles and Triples excitations
CFMM	Continuous Fast Multipole Method
CI	Configuration Interaction
CISC	Complex Instruction Set Computer
CISD	Configuration Interaction with Singles and Doubles excitations
COP	Coulomb Orthonormal Polynomial
CNDO	Complete Neglect of Differential Overlap
CPU	Central Processing Unit
DE	Differential Equation
DFT	Density Functional Theory
DIIS	Direct Inversion in the Iterative Subspace
ERI	Electron (two) Repulsion Integral
FCI	Full Configuration Interaction
FFT	Fast Fourier Transform
FLOP	Floating Point Operation

FMA	Floating-point Multiply-Add
FMM	Fast Multipole Method
FORTRAN	Formula Translation (programming language)
FT	Fourier Transform
G2	Gaussian-2 theory
GTO	Gaussian type Orbital
GvFMM	Gaussian very Fast Multipole Method
HF	Hartree-Fock
HOMO	Highest Occupied Molecular Orbital
INDO	Intermediate Neglect of Differential Overlap
I/O	Input/Output
KWIK	Name given to the algorithm originally devised by Gill for combinatorial problems which are highly dependent. See Chapter 6 equation (6.9).
LAPACK	Linear Algebra Package
LCAO	Linear Combination of Atomic Orbitals
LHS	Left Hand Side
LUMO	Lowest Unoccupied Molecular Orbital
MA	Multiply-Add
MASPAR	Massively Parallel
MINDO	Modified Intermediate Neglect of Differential Overlap
MO	Molecular Orbital
MOP	Memory Operation
MP	Møller-Plesset
MP <sub>n</sub>	n <sup>th</sup> -order Møller-Plesset
MQM	Molecular Quantum Mechanics
NDDO	Neglect of Diatomic Differential Overlap
PC	Personal Computer
PNDO	Partial Neglect of Differential Overlap
PPPM	Particle-Particle-Particle-Mesh



QCI	Quadratic Configuration Interaction
QCID	Quadratic Configuration Interaction with Doubles excitations
QCISD	Quadratic Configuration Interaction with Singles and Doubles excitations
QCISD(T)	Quadratic Configuration Interaction with Singles and Doubles excitations, incorporating Triples perturbatively
RHF	Restricted Hartree-Fock
RHS	Right Hand Side
RISC	Reduced Instruction Set Computer
ROP	Repulsive Orthonormal Polynomial
SCF	Self-Consistent Field
STO	Slater Type Orbital
UHF	Unrestricted Hartree-Fock
WS	Well-Separated

## LIST OF PUBLICATIONS

The following papers have been published as a direct consequence of the work undertaken for this thesis:

1. “KWIK: Coulomb Energies in  $O(N)$  Work.” Jeremy P. Dombroski, Stephen W. Taylor and Peter M. W. Gill. *Journal of Physical Chemistry*, (1996) **100**, 6272.
2. “Chemistry without Coulomb Tails.” Ross D. Adamson, Jeremy P. Dombroski and Peter M. W. Gill. *Chemical Physics Letters*, (1996) **254**, 329.
3. “The Optimal Partition of the Coulomb Operator.” Aaron M. Lee, Stephen W. Taylor, Jeremy P. Dombroski and Peter M. W. Gill. *Physical Review A*, (1997), in press.

## NOTE ON CALCULATIONS

All calculations and efficiency developments reported herein, were obtained on an IBM RS/6000 model 355 with 64 MB RAM. All programming was written in FORTRAN 77 and compiled with the IBM XL FORTRAN for AIX compiler, versions 3.2.5 and 4.1.0 with O2 optimisation. Code was linked against the BLAS, LAPACK and Q-CHEM libraries when required.