
Contents

Preface	xi
Introduction	xv
Chapter 1. Computer Architectures	1
1.1. Different types of parallelism	1
1.1.1. Overlap, concurrency and parallelism	1
1.1.2. Temporal and spatial parallelism for arithmetic logic units	4
1.1.3. Parallelism and memory	6
1.2. Memory architecture	7
1.2.1. Interleaved multi-bank memory	7
1.2.2. Memory hierarchy	8
1.2.3. Distributed memory	13
1.3. Hybrid architecture	14
1.3.1. Graphics-type accelerators	14
1.3.2. Hybrid computers	16
Chapter 2. Parallelization and Programming Models	17
2.1. Parallelization	17
2.2. Performance criteria	19
2.2.1. Degree of parallelism	19
2.2.2. Load balancing	21
2.2.3. Granularity	21
2.2.4. Scalability	22

2.3. Data parallelism	25
2.3.1. Loop tasks	25
2.3.2. Dependencies	26
2.3.3. Examples of dependence	27
2.3.4. Reduction operations	30
2.3.5. Nested loops	31
2.3.6. OpenMP	34
2.4. Vectorization: a case study	37
2.4.1. Vector computers and vectorization	37
2.4.2. Dependence	38
2.4.3. Reduction operations	39
2.4.4. Pipeline operations	41
2.5. Message-passing	43
2.5.1. Message-passing programming	43
2.5.2. Parallel environment management	44
2.5.3. Point-to-point communications	45
2.5.4. Collective communications	46
2.6. Performance analysis	49
Chapter 3. Parallel Algorithm Concepts	53
3.1. Parallel algorithms for recurrences	54
3.1.1. The principles of reduction methods	54
3.1.2. Overhead and stability of reduction methods	55
3.1.3. Cyclic reduction	57
3.2. Data locality and distribution: product of matrices	58
3.2.1. Row and column algorithms	58
3.2.2. Block algorithms	60
3.2.3. Distributed algorithms	64
3.2.4. Implementation	66
Chapter 4. Basics of Numerical Matrix Analysis	71
4.1. Review of basic notions of linear algebra	71
4.1.1. Vector spaces, scalar products and orthogonal projection	71
4.1.2. Linear applications and matrices	74
4.2. Properties of matrices	79
4.2.1. Matrices, eigenvalues and eigenvectors	79
4.2.2. Norms of a matrix	80

4.2.3. Basis change	83
4.2.4. Conditioning of a matrix	85
Chapter 5. Sparse Matrices	93
5.1. Origins of sparse matrices	93
5.2. Parallel formation of sparse matrices: shared memory	98
5.3. Parallel formation by block of sparse matrices: distributed memory	99
5.3.1. Parallelization by sets of vertices	99
5.3.2. Parallelization by sets of elements	101
5.3.3. Comparison: sets of vertices and elements	101
Chapter 6. Solving Linear Systems	105
6.1. Direct methods	105
6.2. Iterative methods	106
Chapter 7. LU Methods for Solving Linear Systems	109
7.1. Principle of LU decomposition	109
7.2. Gauss factorization	113
7.3. Gauss–Jordan factorization	115
7.3.1. Row pivoting	118
7.4. Crout and Cholesky factorizations for symmetric matrices	121
Chapter 8. Parallelization of LU Methods for Dense Matrices	125
8.1. Block factorization	125
8.2. Implementation of block factorization in a message-passing environment	130
8.3. Parallelization of forward and backward substitutions	135
Chapter 9. LU Methods for Sparse Matrices	139
9.1. Structure of factorized matrices	139
9.2. Symbolic factorization and renumbering	142
9.3. Elimination trees	147
9.4. Elimination trees and dependencies	152
9.5. Nested dissections	153
9.6. Forward and backward substitutions	159

Chapter 10. Basics of Krylov Subspaces	161
10.1. Krylov subspaces	161
10.2. Construction of the Arnoldi basis	164
Chapter 11. Methods with Complete Orthogonalization for Symmetric Positive Definite Matrices	167
11.1. Construction of the Lanczos basis for symmetric matrices	167
11.2. The Lanczos method	168
11.3. The conjugate gradient method	173
11.4. Comparison with the gradient method	177
11.5. Principle of preconditioning for symmetric positive definite matrices	180
Chapter 12. Exact Orthogonalization Methods for Arbitrary Matrices	185
12.1. The GMRES method	185
12.2. The case of symmetric matrices: the MINRES method	193
12.3. The ORTHODIR method	196
12.4. Principle of preconditioning for non-symmetric matrices	198
Chapter 13. Biorthogonalization Methods for Non-symmetric Matrices	201
13.1. Lanczos biorthogonal basis for non-symmetric matrices	201
13.2. The non-symmetric Lanczos method	206
13.3. The biconjugate gradient method: BiCG	207
13.4. The quasi-minimal residual method: QMR	211
13.5. The BiCGSTAB	217
Chapter 14. Parallelization of Krylov Methods	225
14.1. Parallelization of dense matrix-vector product	225
14.2. Parallelization of sparse matrix-vector product based on node sets	227
14.3. Parallelization of sparse matrix-vector product based on element sets	229

14.3.1. Review of the principles of domain decomposition	229
14.3.2. Matrix-vector product	231
14.3.3. Interface exchanges	233
14.3.4. Asynchronous matrix-vector product with non-blocking communications	236
14.3.5. Comparison: parallelization based on node and element sets	236
14.4. Parallelization of the scalar product	238
14.4.1. By weight	239
14.4.2. By distributivity	239
14.4.3. By ownership	240
14.5. Summary of the parallelization of Krylov methods	241
Chapter 15. Parallel Preconditioning Methods	243
15.1. Diagonal	243
15.2. Incomplete factorization methods	245
15.2.1. Principle	245
15.2.2. Parallelization	248
15.3. Schur complement method	250
15.3.1. Optimal local preconditioning	250
15.3.2. Principle of the Schur complement method	251
15.3.3. Properties of the Schur complement method	254
15.4. Algebraic multigrid	257
15.4.1. Preconditioning using projection	257
15.4.2. Algebraic construction of a coarse grid	258
15.4.3. Algebraic multigrid methods	261
15.5. The Schwarz additive method of preconditioning	263
15.5.1. Principle of the overlap	263
15.5.2. Multiplicative versus additive Schwarz methods	265
15.5.3. Additive Schwarz preconditioning	268
15.5.4. Restricted additive Schwarz: parallel implementation	269
15.6. Preconditioners based on the physics	275
15.6.1. Gauss–Seidel method	275
15.6.2. Linelet method	276

Appendices	279
Appendix 1	281
Appendix 2	301
Appendix 3	323
Bibliography	339
Index	343