Iterative solution of large sparse systems of equations

Wolfgang Hackbusch.

Applied Mathematical Sciences

1. *John:* Partial Differential Equations, 4th ed.
2. *Sirovich:* Techniques of Asymptotic Analysis.
3. *Hale:* Theory of Functional Differential Equations, 2nd ed.
4. *Percus:* Combinatorial Methods.
5. *von Mises/Friedrichs:* Fluid Dynamics.
6. *Freiberger/Grenander:* A Short Course in Computational Probability and Statistics.
7. *Pipkin:* Lectures on Viscoelasticity Theory.
8. *Giacoglia:* Perturbation Methods in Non-linear Systems.
9. *Friedrichs:* Spectral Theory of Operators in Hilbert Space.
10. *Stroud:* Numerical Quadrature and Solution of Ordinary Differential Equations.
11. *Wolovich:* Linear Multivariable Systems.
12. *Berkovitz:* Optimal Control Theory.
13. *Bluman/Cole:* Similarity Methods for Differential Equations.
14. *Yoshizawa:* Stability Theory and the Existence of Periodic Solution and Almost Periodic Solutions.
15. *Braun:* Differential Equations and Their Applications, 3rd ed.
16. *Lefschetz:* Applications of Algebraic Topology.
17. *Collatz/Wetterling:* Optimization Problems.
18. *Grenander:* Pattern Synthesis: Lectures in Pattern Theory, Vol. I.
19. *Marsden/McCracken:* Hopf Bifurcation and Its Applications.
20. *Driver:* Ordinary and Delay Differential Equations.
21. *Courant/Friedrichs:* Supersonic Flow and Shock Waves.
22. *Rouche/Habets/Laloy:* Stability Theory by Liapunov's Direct Method.
23. *Lamperti:* Stochastic Processes: A Survey of the Mathematical Theory.
24. *Grenander:* Pattern Analysis: Lectures in Pattern Theory, Vol. II.
25. *Davies:* Integral Transforms and Their Applications, 2nd ed.
26. *Kushner/Clark:* Stochastic Approximation Methods for Constrained and Unconstrained Systems.
27. *de Boor:* A Practical Guide to Splines.
28. *Keilson:* Markov Chain Models—Rarity and Exponentiality.
29. *de Veubeke:* A Course in Elasticity.
30. *Shiatycki:* Geometric Quantization and Quantum Mechanics.
31. *Reid:* Sturmian Theory for Ordinary Differential Equations.
32. *Meis/Markowitz:* Numerical Solution of Partial Differential Equations.
33. *Grenander:* Regular Structures: Lectures in Pattern Theory, Vol. III.
34. *Kevorkian/Cole:* Perturbation Methods in Applied Mathematics.
35. *Carr:* Applications of Centre Manifold Theory.
36. *Bengtsson/Ghil/Källén:* Dynamic Meteorology: Data Assimilation Methods.
37. *Saperstone:* Semidynamical Systems in Infinite Dimensional Spaces.
38. *Lichtenberg/Lieberman:* Regular and Chaotic Dynamics, 2nd ed.
39. *Piccini/Stampacchia/Vidossich:* Ordinary Differential Equations in $\mathbf{R}^n$.
40. *Naylor/Sell:* Linear Operator Theory in Engineering and Science.
41. *Sparrow:* The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors.
42. *Guckenheimer/Holmes:* Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.
43. *Ockendon/Taylor:* Inviscid Fluid Flows.
44. *Pazy:* Semigroups of Linear Operators and Applications to Partial Differential Equations.
45. *Glashoff/Gustafson:* Linear Operations and Approximation: An Introduction to the Theoretical Analysis and Numerical Treatment of Semi-Infinite Programs.
46. *Wilcox:* Scattering Theory for Diffraction Gratings.
47. *Hale et al:* An Introduction to Infinite Dimensional Dynamical Systems—Geometric Theory.
48. *Murray:* Asymptotic Analysis.
49. *Ladyzhenskaya:* The Boundary—Value Problems of Mathematical Physics.
50. *Wilcox:* Sound Propagation in Stratified Fluids.
51. *Golubitsky/Schaeffer:* Bifurcation and Groups in Bifurcation Theory, Vol. I.

(continued following index)

Wolfgang Hackbusch

Iterative Solution of Large Sparse Systems of Equations

With 16 Illustrations

Springer-Verlag
New York Berlin Heidelberg London Paris
Tokyo Hong Kong Barcelona Budapest

Applied Mathematical Sciences

Volume 95

Applied Mathematical Sciences

(continued from page ii)

52. *Chipot:* Variational Inequalities and Flow in Porous Media.
53. *Majda:* Compressible Fluid Flow and System of Conservation Laws in Several Space Variables.
54. *Wasow:* Linear Turning Point Theory.
55. *Yosida:* Operational Calculus: A Theory of Hyperfunctions.
56. *Chang/Howes:* Nonlinear Singular Perturbation Phenomena: Theory and Applications.
57. *Reinhardt:* Analysis of Approximation Methods for Differential and Integral Equations.
58. *Dwoyer/Hussaini/Voigt (eds):* Theoretical Approaches to Turbulence.
59. *Sanders/Verhulst:* Averaging Methods in Nonlinear Dynamical Systems.
60. *Ghil/Childress:* Topics in Geophysical Dynamics: Atmospheric Dynamics, Dynamo Theory and Climate Dynamics.
61. *Sattinger/Weaver:* Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics.
62. *LaSalle:* The Stability and Control of Discrete Processes.
63. *Grasman:* Asymptotic Methods of Relaxation Oscillations and Applications.
64. *Hsu:* Cell-to-Cell Mapping: A Method of Global Analysis for Nonlinear Systems.
65. *Rand/Armbruster:* Perturbation Methods, Bifurcation Theory and Computer Algebra.
66. *Hlavácek/Haslinger/Necasl/Lovísek:* Solution of Variational Inequalities in Mechanics.
67. *Cercignani:* The Boltzmann Equation and Its Applications.
68. *Temam:* Infinite Dimensional Dynamical Systems in Mechanics and Physics.
69. *Golubitsky/Stewart/Schaeffer:* Singularities and Groups in Bifurcation Theory, Vol. II.
70. *Constantin/Foias/Nicolaenko/Temam:* Integral Manifolds and Inertial Manifolds for Dissipative Partial Differential Equations.
71. *Catlin:* Estimation, Control, and the Discrete Kalman Filter.
72. *Lochak/Meunier:* Multiphase Averaging for Classical Systems.
73. *Wiggins:* Global Bifurcations and Chaos.
74. *Mawhin/Willem:* Critical Point Theory and Hamiltonian Systems.
75. *Abraham/Marsden/Ratiu:* Manifolds, Tensor Analysis, and Applications, 2nd ed.
76. *Lagerstrom:* Matched Asymptotic Expansions: Ideas and Techniques.
77. *Aldous:* Probability Approximations via the Poisson Clumping Heuristic.
78. *Dacorogna:* Direct Methods in the Calculus of Variations.
79. *Hernández-Lerma:* Adaptive Markov Processes.
80. *Lawden:* Elliptic Functions and Applications.
81. *Bluman/Kumei:* Symmetries and Differential Equations.
82. *Kress:* Linear Integral Equations.
83. *Bebernes/Eberly:* Mathematical Problems from Combustion Theory.
84. *Joseph:* Fluid Dynamics of Viscoelastic Fluids.
85. *Yang:* Wave Packets and Their Bifurcations in Geophysical Fluid Dynamics.
86. *Dendrinos/Sonis:* Chaos and Socio-Spatial Dynamics.
87. *Weder:* Spectral and Scattering Theory for Wave Propagation in Perturbed Stratified Media.
88. *Bogaevski/Povzner:* Algebraic Methods in Nonlinear Perturbation Theory.
89. *O'Malley:* Singular Perturbation Methods for Ordinary Differential Equations.
90. *Meyer/Hall:* Introduction to Hamiltonian Dynamical Systems and the N-body Problem.
91. *Straughan:* The Energy Method, Stability, and Nonlinear Convection.
92. *Naber:* The Geometry of Minkowski Spacetime.
93. *Colton/Kress:* Inverse Acoustic and Electromagnetic Scattering Theory.
94. *Hoppensteadt:* Analysis and Simulation of Chaotic Systems.
95. *Hackbusch:* Iterative Solution of Large Sparse Systems of Equations.
96. *Marchioro/Pulvirenti:* Mathematical Theory of Incompressible Nonviscous Fluids.
97. *Lasota/Mackey:* Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics, 2nd ed.
98. *de Boor/Höllig/Riemenschneider:* Box Splines.
99. *Hale/Lunel:* Introduction to Functional Differential Equations.
100. *Sirovich:* Trends and Perspectives in Applied Mathematics.
101. *Nusse/Yorke:* Dynamics: Numerical Explorations.
102. *Chossat/Iooss:* The Couette-Taylor Problem.

Wolfgang Hackbusch
Institut für Informatik und Praktische Mathematik
Christian-Albrechts-Universität zu Kiel
Olshausenstr. 40
D 24098 Kiel
Germany

Library of Congress Cataloging-in-Publication Data
Hackbusch, W., 1948–
[Iterative Lösung grosser schwachbesetzter Gleichungssysteme, English]
Iterative solution of large sparse systems of equations / Wolfgang Hackbusch.
p. cm.—(Applied mathematical science; v. 95)
Includes bibliographical references and indexes.
ISBN 0-387-94064-2
1. Sparse matrices. 2. Iterative methods (Mathematics)
3. Differential equations, Partial—Numerical solutions. I. Title.
II. Series: Applied mathematical sciences (Springer-Verlag New York Inc.); v. 95.
QA1.A647 vol. 95
[QA188]
510 s—dc20
[512.9'434] 93-2098

Printed on acid-free paper.

Production managed by Henry Krell; manufacturing supervised by Jacqui Ashri.
Typeset by Asco Trade Typesetting Ltd., Hong Kong.
Printed and bound by R.R. Donnelley and Sons, Harrisonburg, VA.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-94064-2 Springer-Verlag New York Berlin Heidelberg
ISBN 3-540-94064-2 Springer-Verlag Berlin Heidelberg New York

Preface

C. F. Gauß in a letter from Dec. 26, 1823 to Gerling:

> Ich empfehle Ihnen diesen Modus zur Nachahmung. Schwerlich werden Sie je wieder direct eliminiren, wenigstens nicht, wenn Sie mehr als 2 Unbekannte haben. Das indirecte Verfahren lässt sich halb im Schlafe ausführen, oder man kann während desselben an andere Dinge denken.
>
> [C. F. Gauß: Werke vol. 9, Göttingen, p. 280, 1903]

What difference exists between solving large and small systems of equations? The standard methods well-known to any student of linear algebra are applicable to all systems, whether large or small. The necessary amount of work, however, increases dramatically with the size, so one has to search for algorithms that most efficiently and accurately solve systems of 1000, 10,000, or even one million equations. The choice of algorithms depends on the special properties the matrices in practice have. An important class of large systems arises from the discretisation of partial differential equations. In this case, the matrices are sparse (i.e., they contain mostly zeros) and well-suited to iterative algorithms. Because of the background in partial differential equations, this book is closely connected with the author's *Theory and Numerical Treatment of Elliptic Differential Equations*, whose English translation has also been published by Springer-Verlag.

This book grew out of a series of lectures given by the author at the Christian-Albrecht University of Kiel to students of mathematics. It tries to describe the recent state of iterative and related methods, without, however, delving into specialised areas. Even the volume's limitation to iterative techniques entails a selection: Various fast direct algorithms for special problems as well as the optimised versions of Gauß elimination, the Cholesky method, or band-width reduction are not taken into consideration.

Although special attention is devoted to the modern effective algorithms (conjugate gradients, multi-grid methods, parallel techniques), the theory of classical iterative methods should not be neglected. On the other hand, some effective algorithms are not or only marginally considered, if they are connected too closely with discretisation techniques that are not the subject of this book. A discussion of the iterative treatment of nonlinear problems or of eigenvalue problems is completely avoided. A chapter on saddle-point problems (special indefinite problems) arising in many interesting applications could not be realized because of the need to limit the size of the book.

This volume requires no basic mathematical knowledge other than courses on analysis and linear algebra. The principles of linear algebra are summarised in Chapter 2 of this book in order to provide as complete a presentation as possible and present the formulation and notation needed here.

With respect to a course of study, a selection of the given material is best suited to a full-semester course (four hours a week) in the second part of the study (between Vordiplom and Diplom). A partial selection can also be recommended for the second part of a course on numerical analysis.

The exercises cited, which may also be understood as remarks without proof, are an integral part of the presentation. Should this book be used as the basis for an academic course, they can be assigned as problems for students. However, the non-student reader should also try to test his comprehension by working on these exercises.

The discussion of the various methods is illustrated by very many numerical examples, mostly for the Poisson-model problem. To enable the interested reader to test the algorithms with other parameters, step sizes, etc., the PASCAL programs are explicitly given. A collection of the source codes required is available on disk (see [Prog] in the bibliography). The programs can also be used independently of the book for producing numerical examples for courses or seminars.

The present English version contains corrections of several misprints still in the original German edition. New publications have been added to the bibliography. Furthermore, we replaced references to German textbooks by English equivalents as much as possible.

The author would like to thank his colleagues, in particular Mr. J. Burmeister, for their help in proofreading. I am grateful for the stimulating conversations with W. Niethammer, G. Maeß, M. Dryja, G. Wittum, O. Widlund, and others. I also wish to express my gratitude to Teubner (publisher of the German version) and Springer-Verlag for their friendly cooperation on both the German and English editions of this book.

Kiel, April 1993 *W. Hackbusch*

Hints for Reading the Book:

§1: Prelude

§2: Thoughts to serve for reference. However, one should glance through §2.1.

§3: Read §§3.1–3 first. Rest ad libitum.

§4: Chapters 4.2–3 (classical iterations) are the basis of almost all other considerations. §4.4 deals with the corresponding convergence

analysis. §§4.1 and 4.7 refer to the Poisson-model problem and serve as illustration. First, §§4.5 and 4.8 may be passed over.

§5: Contains the SOR analysis and may be left out during the first reading.

§6: Independent chapter. The other parts refer only very seldom to §6.

§7: Necessary preparation to §9

§§8–11: Each chapter is independent.

Contents

Notations

Numbers of formulae: Equations in a subchapter x.y are numerated by (x.y.1), (x.y.2) etc. The equation **(3.2.1)** is quoted by **(1)** in the same Section 3.2, while we write **(2.1)** in the other sections of chapter 3.

Numeration of Theorems etc.: All theorems, definitions, and lemmata etc. are enumerated together. The reference to a theorem etc., is analogous to what is said above. **Lemma 3.2.7** is cited as «**Lemma 7**» in Section 3.2, while in the other sections of chapter 3 it is denoted by «**Lemma 2.7**». However, §1 indicates Chapter 1 and never the sections §3.1 or §3.2.1.

Special Symbols, Abbreviations, and Conventions

a, b	bounds for $\sigma(M)$; cf. (4.8.4c), Theorem 7.3.8
A, A_ℓ	matrix of the linear system; cf. (1.2.5), (10.1.8a)
$A^{\alpha\beta}, A^{ij}$	block of A; cf. (2.5.2b, c)
$A_{\alpha\beta}, a_{\alpha\beta}, A_{ij}, a_{ij}$	components of matrix A
b, b_ℓ	right-hand side of the linear system; cf. (1.2.5), (10.1.8a)
blockdiag$\{\dots\}$	block-diagonal matrix; cf. (2.5.3a, b)
blocktridiag$\{\dots\}$	block-tridiagonal matrix; cf. (2.5.4)
$\mathbb{C}$	complex numbers
cond	condition, cond_2: spectral condition; cf. (2.10.7)
$D, D', D_1, \dots$	(block-)diagonal matrix
det	determinant
diag$\{\dots\}$	diagonal matrix or diagonal part; cf. (2.1.5a–c)

e^m	error $x^m - x$ of the mth iterate
E	strictly lower triangular matrix; cf. (1.4.2)
F	strictly upper triangular matrix; cf. (1.4.2)
$G(A)$	graph of a matrix A; cf. (6.2.1)
h, h_ℓ	grid size; cf. (1.2.2)
i, j, k	indices of the ordered index set $I = \{1, \dots, n\}$
I	identity matrix
I	index set (not necessarily ordered)
I_α	subset of block indices; cf. (2.5.1)
$\mathbb{K}$	the field $\mathbb{R}$ or $\mathbb{C}$
$\mathbb{K}^I$	space of the vectors corresponding to the index set I
$\mathbb{K}^{I \times I}$	space of the matrices corresponding to the index set I
ℓ	level number in the discretisation hierarchy; cf. §10.1.2
$L, L', \hat{L}$	lower (block-)triangular matrix
log	natural logarithm
m	iteration number; cf. x^m
M, M^{xyz}	iteration matrix (of the iteration xyz); cf. §3.2.1
n, n_ℓ	dimension of the linear system; cf. §3.3, (10.1.8b)
N	number of the grid points per row or column; cf. (1.2.2)
N, N^{xyz}	matrix of the 2nd normal form (of iteration xyz); cf. (3.2.4)
$\mathbb{N}$	natural numbers $\{1, 2, 3, \dots\}$
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$
$O(\cdot)$	Landau symbol: $f(\alpha) = O(g(\alpha))$ if $\lvert f(\alpha)\rvert \leqslant C \lvert g(\alpha)\rvert$ for the underlying limit process $\alpha \to 0$ or $\alpha \to \infty$. The notation $f(h) = 1 - O(h^\kappa)$ is more special and means that $f(h) \leqslant 1 - Ch^\kappa$ with fixed $C > 0$ for $h \to 0$.
p	prolongation; §10.1.3, (11.2.1)
Q	often: unitary matrix
r	restriction; cf. §10.1.4, (11.2.8a)
$r(A)$	numerical radius of matrix A; cf. §2.9.5
$\mathbb{R}$	real numbers
range(Φ)	range (image space) of a mapping Φ
S_ℓ	iteration matrix of the smoother $\mathscr{S}_\ell$; cf. Lemma 10.2.1
$\mathscr{S}_\ell$	smoothing iteration; cf. §10.1.1, §10.2.1
span$\{\dots\}$	linear space spanned by $\{\dots\}$
T_ℓ, T_r	left- and right-sided transformation; cf. (8.1.4), (8.1.11)
tridiag$\{\dots\}$	tridiagonal matrix; cf. (2.1.6)
u_{ij}	components of the grid function $u = x$; cf. (1.2.6b)
$U, U', \hat{U}$	upper (block-)triangular matrix
W, W^{xyz}	matrix of the 3rd normal form (of iteration xyz); cf. (3.2.5)
x	vector; often solution of the equation $Ax = b$
x^*	solution of the equation $Ax = b$, if the symbol x is needed as a variable
x_ℓ, x_ℓ^*	vectors x, x^* at the level ℓ; cf. §10
x^0	starting value of the iteration

x^m	mth iterate
x^α, x^i	block of x corresponding to the index α or i; cf. (2.5.2a)
x_α, x_i	components of a vector x
$\mathbb{Z}$	integers $\{0, \pm 1, \pm 2, \dots\}$

Greek Letters

α, β, γ	indices of the index set; cf. §2.1
γ	in §10: number of the secondary multi-grid steps for the coarse-grid equation; cf. (10.4.2d$_2$)
γ, Γ	lower and upper eigenvalue bounds of $W^{-1}A$; cf. (4.8.4c)
δ_{ij}	Kronecker symbol: $\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ otherwise
ζ	often contraction number; cf. (10.3.21b), (10.5.3)
$\eta(\nu)$	zero sequence for smoothing property, cf. (10.6.4b)
$\eta_0(\nu)$	special function, defined in Lemma 10.6.1
ϑ, Θ	damping factor; cf. §§4.3.1–2
$\kappa(A)$	condition number (2.10.8)
λ, Λ	eigenvalue bounds of A; cf. Theorem 9.2.3, Theorem 4.4.8
$\lambda_{\max}(A)$	maximal eigenvalue of the matrix A, if $\sigma(A) \subset \mathbb{R}$
$\lambda_{\min}(A)$	minimal eigenvalue of the matrix A, if $\sigma(A) \subset \mathbb{R}$
ν, ν_1, ν_2	in §10: number of the smoothing steps; cf. §§10.2.1–2
$\rho(A)$	spectral radius of the matrix A; cf. Definition 2.4.13
$\rho_{m+k,m}$	convergence factors; cf. (3.2.18a, b)
$\sigma(A)$	spectrum of the matrix A; cf. (2.4.1)
ω	relaxation parameter; cf. §4.3.3.1
Ω_h	grid; cf. (1.2.3)

Symbols

$\mathbb{1}$	vector $(1, 1, \dots, 1)^T$
A^T, A^H, A^{-H}	cf. (2.1.1/2)
Δ	Laplace operator; cf. (1.2.1a)
$\langle \cdot, \cdot \rangle$	(Euclidean) scalar product; cf. (2.2.1a–c)
$\langle \cdot, \cdot \rangle_A$	energy scalar product; cf. (2.10.5b)
$\lVert \cdot \rVert, \vert\vert\vert \cdot \vert\vert\vert$	norm (of vectors or matrices)
$\lVert \cdot \rVert_A$	energy norm; cf. (2.10.5a)
$\lVert \cdot \rVert_2$	Euclidean norm; cf. (2.6.2). spectral norm; cf. (2.9.4a)
$\lVert \cdot \rVert_\infty$	maximum norm; cf. (2.6.2). row sum norm; cf. (2.6.8)
$\lVert \cdot \rVert_{Y \leftarrow X}$	cf. (2.6.11): norm of a mapping (matrix) from X into Y
$\lvert \cdot \rvert$	absolute value, in §6 also applied to matrices
$<, \leqslant, >, \geqslant$	in connection with matrices, the order relation from §2.10.2 is used; only in §6 (and parts of §8.5) is the order relation of (6.1.1a, b) meant

1
Introduction

1.1 Historical Remarks Concerning Iterative Methods

Iterative methods are almost 170 years old. The first iterative method for systems of linear equations is due to Carl Friedrich Gauß (the less correct spelling is: Gauss). His least squares method led him to a system of equations that was too large for the use of direct Gauß elimination. The iterative method described in "Supplementum theoriae combinationis observationum erroribus minime obnoxiae" (1819–1822) today would be called the blockwise Gauß-Seidel method. The value that Gauß attributed to his iterative method can be seen in the excerpt from his letter of 1823 preceding the foreword of this book.

A very similar method was described by Carl Gustav Jacobi in his 1845 paper "Über eine neue Auflösungsart der bei der Methode der kleinsten Quadrate vorkommenden linearen Gleichungen" ("On a new solution method of linear equations arising from the least square method"; *Astronom. Nachr.*). In 1874 Phillip Ludwig Seidel, a student of Jacobi, wrote about "a method, to solve the equations arising from the least squares method as well as general linear equations by successive approximation" ("Über ein Verfahren, die Gleichungen, auf welche die Methode der kleinsten Quadrate führt, sowie lineare Gleichungen überhaupt, durch successive Annäherung aufzulösen"; *Münch. Abh.*).

Since the time that electronic computers became available for solving systems of equations, the number of equations has increased many orders of magnitudes and the methods mentioned above have proved to be too slow. After 100 years of stagnation in this field, Southwell [1–3] experimented with variants of the Gauß-Seidel method («relaxation») and, in 1950, Young [1] succeeded in a breakthrough. His variant of the Gauß-Seidel method leads to an important acceleration of the convergence. This so-called SOR method will be described in §1.4 as an example of an iterative method. Since then, numerous other methods even more effective have been developed. Most of them are discussed in this book.

1.2 Model Problem (Poisson Equation)

During the time of Gauß, Jacobi and Seidel, the equations of the least square method have led to a larger number of equations. Today, in particular the boundary-value problems, i.e. the partial differential equations of elliptic type, lead to a number of equations of orders between 1000 and one million. In the following we will often refer to a model problem representing the simplest nontrivial example of a boundary-value problem. It is the *Poisson equation*

$$-\Delta u(x,y) = f(x,y) \quad \text{for } (x,y) \in \Omega, \qquad (1.2.1a)$$

$$u(x,y) = \varphi(x,y) \quad \text{on } \Gamma = \partial\Omega. \qquad (1.2.1b)$$

Here

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

abbreviates the *Laplace operator*. As domain Ω we choose the unit square

$$\Omega = (0,1) \times (0,1). \qquad (1.2.1c)$$

In (1a, b) the *source term* f and the *boundary value* φ are given, while the function u is unknown.

To discretise the differential equation (1a–c) the domain Ω is covered with a grid of the step size h (cf. Figure 1a). Each grid point (x,y) must have the representation $x = ih$, $y = jh$ $(0 \leq i,j \leq N)$, where

$$h = 1/N. \qquad (1.2.2)$$

By the term *grid* the set of *inner* grid points is meant:

$$\Omega_h := \{(x,y) = (ih,jh) \colon 1 \leq i,j \leq N-1\}. \qquad (1.2.3)$$

We abbreviate the desired values $u(x,y) = u(ih,jh)$ by u_{ij}. An approximation for the differential equation (1a) is given by the so-called *five-point formula*

$$h^{-2}[4u_{ij} - u_{i-1,j} - u_{i+1,j} - u_{i,j-1} - u_{i,j+1}] = f_{ij} \qquad (1.2.4a)$$

with $f_{ij} := f(ih,jh)$ for $1 \leq i,j \leq N-1$. The left side in (4a) coincides with $-\Delta u(ih,jh)$ up to a consistency error $O(h^2)$, when the solution u of (1a, b) is inserted (cf. Hackbusch [15]). For grid values on the boundary, i.e. for $i = 0$, $i = N$, $j = 0$ or $j = N$, u_{ij} is known from the boundary data (1b):

$$u_{ij} := \varphi(ih,jh) \quad \text{for } i = 0,\ i = N,\ j = 0 \text{ or } j = N. \qquad (1.2.4b)$$

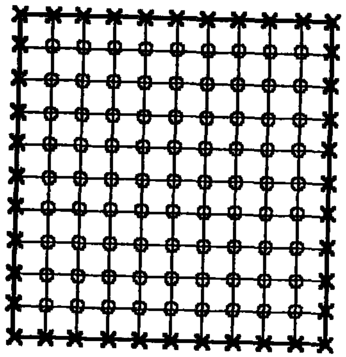

Fig. 1.2.1a Inner grid points (o) and boundary points (×)

The number of the unknown u_{ij} is $n := (N-1)^2$ and corresponds to the number of the inner grid points. In order to form the system of equations, one has to eliminate the boundary values (4b), which possibly may appear in (4a). For example for $N \geq 3$ the equation corresponding to the index $(i,j) = (1,1)$ reads

$$h^{-2}[4u_{11} - u_{12} - u_{21}] = g_{11} \quad \text{with } g_{11} := f(h,h) + h^{-2}[\varphi(0,h) + \varphi(h,0)].$$

In order to write the equations in the commonly used matrix formulation

$$Ax = b \qquad (1.2.5)$$

with an $n \times n$ matrix A and n-dimensional vectors x and b with $n = (N-1)^2$, one is forced to represent the twofold indexed unknowns u_{ij} by a singly indexed vector x. This implies that the (inner) grid points must be enumerated in some way. Figure 1b shows the lexicographic ordering. The exact definition of the matrix A and of the right-hand side b can be seen from the following

Definition of the matrix A and of the vectors b by lexicographical ordering for the Poisson-model problem: (1.2.6a)

```
     N1 := N − 1; n := N1 * N1; h2 := 1/(N * N);
k := 0; {1 ≤ k ≤ n is the index with respect to the lex. ordering}
A := 0; {all entries of A are initialised by zero}
for j := 1 to N1 do for i := 1 to N1 do
begin k := k + 1; a_kk := 4 * h2; b_k := f(ih,jh);
     if i > 1   then a_k−1,k   := −h2 else b_k := b_k + h2 * φ(0,jh);
     if i < N1 then a_k+1,k   := −h2 else b_k := b_k + h2 * φ(1,jh);
     if j > 1   then a_k,k−N1 := −h2 else b_k := b_k + h2 * φ(ih,0);
     if j < N1 then a_k,k+N1 := −h2 else b_k := b_k + h2 * φ(ih,1)
end;
```

Vice versa, the solution x of $Ax = b$ is to be interpreted as

$$x_k = u_{ij} = u(ih,jh) \quad \text{for } k = i + (j-1)*(N-1) \text{ with } 0 \leq i,j \leq N. \qquad (1.2.6b)$$

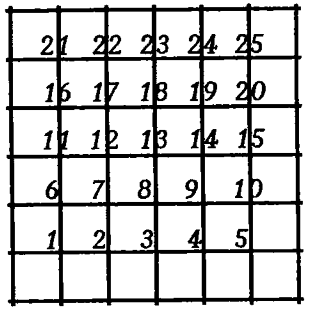

Fig. 1.2.1b Lexicographical ordering

When x is interpreted as the grid function, we shall use the notation u_{ij} or $u(x, y)$ with $x = ih$, $y = jh$.

Remark 1.2.1. The reformulation of the two-dimensionally ordered unknowns into a one-dimensionally ordered vector is rather unobvious. The reason should not be sought in the two-dimensional nature of the problem, but rather in the questionable concept of enumerating the components of a vector from 1 to n. We shall see that **the matrix A will never be needed in the presentation (2.6a).**

If, nevertheless, one wants to represent the matrix A in the usual form, A must be written as a *block matrix*. The vector x decomposes naturally into $N-1$ blocks

$$x^{(j)} := \begin{bmatrix} x_{k+1} \\ \vdots \\ x_{k+N-1} \end{bmatrix} = \begin{bmatrix} u_{1,j} \\ \vdots \\ u_{N-1,j} \end{bmatrix} \quad \begin{matrix} \text{with } k := (j-1)*(N-1) \\ \text{for } j = 1, \ldots, N-1, \end{matrix} \tag{1.2.7}$$

corresponding to the jth row in the grid Ω_h. Accordingly, A takes the form of a block-tridiagonal matrix built from $(N-1) \times (N-1)$ blocks T, which again are tridiagonal $(N-1) \times (N-1)$ matrices:

$$A = h^{-2} \begin{bmatrix} T & -I & & & \\ -I & T & -I & & \\ & -I & T & -I & \\ & & \ddots & \ddots & \ddots \\ & & -I & T & -I \\ & & & -I & T \end{bmatrix}, \quad T = \begin{bmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & -1 & 4 & -1 & \\ & & \ddots & \ddots & \ddots \\ & & -1 & 4 & -1 \\ & & & -1 & 4 \end{bmatrix}. \tag{1.2.8}$$

I is the $(N-1) \times (N-1)$ identity matrix. Unmarked matrix entries or blocks are zeros or zero blocks, respectively. The representation (8) proves

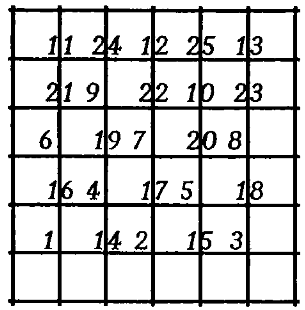

Fig. 1.2.1c Chequer-board ordering

Remark 1.2.2. For the lexicographical ordering of the unknowns, the matrix A has a block-tridiagonal structure.

Lexicographical ordering is by no means the only ordering one can think of. Another frequently used aproach is the chequer-board ordering (cf. Fig. 1c). In that case, the components u_{ij} with an even sum $i + j$ («black squares») are enumerated first and thereafter those with an odd sum $i + j$ («red squares») are numbered lexicographically. In the course of the next chapters further orderings will be mentioned. A broad collection of orderings of practical interest is given by Duff–Meurant [1].

Exercise 1.2.3. In the case of the chequer-board ordering, A decomposes into two blocks corresponding to the «red» and «black» indices. Prove that A has the block structure (9) with a rectangular submatrix B and identity matrices I_r, I_b, where the block size is given by the number of the red and black grid points:

$$A = \begin{bmatrix} D_r & B \\ B^T & D_b \end{bmatrix}, \qquad D_r = 4h^{-2} I_r, \qquad D_b = 4h^{-2} I_b. \tag{1.2.9}$$

1.3 Amount of Work for the Direct Solution of the System of Equations

Those methods are called *direct* that terminate after finitely many operations with an exact solution (up to rounding errors). The best known direct method is Gauß elimination. In the case of the model problem from §1.2, one may perform this method *without* pivoting (cf. §6.4.4). Concerning the valuation of the amount of computational work, we do not distinguish between additions, subtractions, multiplications, or divisions. Each is counted as one (arithmetical) operation. Traditionally, arithmetical operations for indices, data transfer, and similar activities are not counted.

Remark 1.3.1. In the general case the Gauß elimination for the solution of a system of n equations $Ax = b$ requires $2n^3/3 + O(n^2)$ operations. The storage amounts to $n^2 + n$.

Proof. During the ith elimination step, the ith row contains $n - i$ nonzero elements, whose multiples are to be subtracted from $n - i - 1$ matrix rows. Summation of these $2(n - i)^2 + O(n)$ operations over $1 \leqslant i \leqslant n$ yields the assertion. □

In the model case, n equals $(N - 1)^2 = h^{-2} + O(h^{-1})$ and implies

Remark 1.3.2. A naive application of the Gauß elimination to the model problem from §1.2 leads to $2N^6/3 + O(N^5) = 2h^{-6}/3 + O(h^{-5})$ operations and requires a storage of $N^4 + O(N^3) = h^{-4} + O(h^{-3})$.

Halving the grid size h yields the 64-fold amount of computational work. If we assume 1 CPU sec for the solution of grid size h, the same computation for the quartered grid size $h/4$ consumes more than 1 CPU hr!

However, the amount of work diminishes if the system matrix A is a band matrix.

Definition 1.3.3. A is a *band matrix* of *band width* w, if $a_{ij} = 0$ for all indices with $|i - j| > w$.

A band matrix has at maximum $2w$ side-diagonals besides the main diagonal. Concerning the analysis of the properties of band matrices, we refer to Berg [1].

Remark 1.3.4. The matrix A arising from the lexicographical ordering for the model problem according to (2.7) is a band matrix of band width $w = N - 1$.

The major part of the amount of work given in Remark 1 consists of unnecessary multiplications and additions by zeros. During the ith elimination step the ith row contains $w + 1$ nonzero elements. It suffices to eliminate the next w rows. This leads to $2w^2$ operations. In total one obtains

Remark 1.3.5. The amount of work for the Gauß elimination without pivoting for solving a system with an $n \times n$ matrix of band width w amounts to $2nw^2 + O(nw + w^3)$. The storage requirement reduces to $2n(w + 1)$, when only the $2w + 1$ diagonals of A and the right-hand side b are stored.

Remark 1.3.6. In the case of the model problem from §1.2, w equals $N - 1$. Therefore, the banded Gauß elimination requires $2N^4 + O(N^3) = 2h^{-4} + O(h^{-3})$ operations and a storage of $2N^3 + O(N^2)$.

In the latter version, $2w + 1$ diagonals of A are used, although the matrix A from (2.8) has only five diagonals: the main diagonal, two side-diagonals with distance 1, and two further ones with distance $N - 1$. Unfortunately, one cannot exploit this property for the Gauß elimination because of

Remark 1.3.7. The zeros in the second to $(N - 2)$th side-diagonals of the matrix A from (2.8) are completely filled during the elimination process with nonzeros (with the exception of the first block).

This occurrence is called *fill-in* and indicates a principal disadvantage of Gauß elimination when applied to *sparse* matrices. Here, we call an $n \times n$ matrix sparse, if the number of nonzero entries is by far smaller than n^2. Otherwise, the matrix is called a *full* matrix. Because of the equivalence of Gauß elimination to the *triangular or LU decomposition* (cf. Stoer [1, §4.1], Stoer-Bulirsch [1, §4.1], the same difficulties exist in the latter technique.

Remark 1.3.8. The decomposition $A = LU$ into a lower triangular matrix L and an upper triangular matrix U for the sparse matrix A from (2.8) yields factors L and U, which are full band matrices of width $w = N - 1$. The same holds for Cholesky decomposition.

There are special direct methods solving the system described in §1.2 with an amount of work between $O(n) = O(N^2)$ and $O(n \log n) = O(N^2 \log N)$. Examples are the Buneman algorithm and the method of total reduction, both described in Meis–Markowitz [1] (cf. also Buneman [1], Björstad [1], Duff–Erisman–Reid [1], Schröder–Trottenberg [1]).

1.4 Examples of Iterative Methods

For the iterative solution of a system of equations, one starts with an arbitrary *starting vector* x^0 and computes a sequence of *iterates* (*iterands*) x^m for $m = 1, 2, \ldots$:

$$x^0 \mapsto x^1 \mapsto x^2 \mapsto x^3 \mapsto \cdots \mapsto x^m \mapsto x^{m+1} \mapsto \cdots.$$

In the following x^{m+1} is dependent only on x^m, so that the mapping $x^m \mapsto x^{m+1}$ determines the iteration method. The choice of the starting value x^0 is not part of that method.

The already mentioned Gauß-Seidel iteration for solving the system (2.5): $Ax = b$ is as follows: