Preface

In 1970, my former academic teacher Roland Bulirsch gave an exercise to his students, which indicated the fascinating invariance of the *ordinary* Newton method under general affine transformation. To my surprise, however, nearly all *global* Newton algorithms used damping or continuation strategies based on residual norms, which evidently lacked affine invariance. Even worse, nearly all convergence theorems appeared to be phrased in not affine invariant terms, among them the classical Newton-Kantorovich and Newton-Mysovskikh theorem. In fact, in those days it was common understanding among numerical analysts that convergence theorems were only expected to give qualitative insight, but not too much of quantitative advice for application, apart from toy problems.

This situation left me deeply unsatisfied, from the point of view of both mathematical aesthetics and algorithm design. Indeed, since my first academic steps, my scientific guideline has been and still is that 'good' mathematical theory should have a palpable influence on the construction of algorithms, while 'good' algorithms should be as firmly as possible backed by a transparently underlying mathematical theory. Only on such a basis, algorithms will be efficient enough to cope with the enormous difficulties of real life problems.

In 1972, I started to work along this line by constructing global Newton algorithms with affine invariant damping strategies [59]. Early companions on this road were Hans-Georg Bock, Gerhard Heindl, and Tetsuro Yamamoto. Since then, the tree of affine invariance has grown lustily, spreading out in many branches of Newton-type methods. So the plan of a comprehensive treatise on the subject arose naturally. Florian Potra, Ekkehard Sachs, and Andreas Griewank gave highly valuable detailed advice. Around 1992, a manuscript on the subject with a comparable working title had already swollen to 300 pages and been distributed among quite a number of colleagues who used it in their lectures or as a basis for their research. Clearly, these colleagues put screws on me to 'finish' that manuscript.

However, shortly after, new relevant aspects came up. In 1993, my former coworker Andreas Hohmann introduced *affine contravariance* in his PhD thesis [119] as a further coherent concept, especially useful in the context of inexact Newton methods with GMRES as inner iterative solver. From then on, the former 'affine invariance' had to be renamed, more precisely, as *affine covariance*. Once the door had been opened, two more concepts arose: in 1996, myself and Martin Weiser formulated *affine conjugacy* for convex optimization [84]; a few years later, I found *affine similarity* to be important for steady state problems in dynamical systems. As a consequence, I decided to rewrite the whole manuscript from scratch, with these four affine invariance concepts representing the columns of a structural matrix, whose rows are the various Newton and Gauss-Newton methods. A presentation of details of the contents is postponed to the next section.

This book has two faces: the first one is that of a *textbook* addressing itself to graduate students of mathematics and computational sciences, the second one is that of a *research monograph* addressing itself to numerical analysists and computational scientists working on the subject.

As a *textbook*, selected chapters may be useful in classes on Numerical Analysis, Nonlinear Optimization, Numerical ODEs, or Numerical PDEs. The presentation is striving for structural simplicity, but not at the expense of precision. It contains a lot of theorems and proofs, from affine invariant versions of the classical Newton-Kantorovich and Newton-Mysovskikh theorem (with proofs simpler than the traditional ones) up to new convergence theorems that are the basis for advanced algorithms in large scale scientific computing. I confess that I did not work out all details of all proofs, if they were folklore or if their structure appeared repeatedly. More elaboration on this aspect would have unduly blown up the volume without adding enough value for the construction of algorithms. However, I definitely made sure that each section is self-contained to a reasonable extent. At the end of each chapter, exercises are included. Web addresses for related software are given.

As a *research monograph*, the presentation (a) quite often goes into the depth covering a large amount of otherwise unpublished material, (b) is open in many directions of possible future research, some of which are explicitly indicated in the text. Even though the experienced reader will have no difficulties in identifying further open topics, let me mention a few of them: There is no complete coverage of all possible combinations of local and global, exact and inexact Newton or Gauss-Newton methods in connection with continuation methods—let alone of all their affine invariant realizations; in other words, the above structural matrix is far from being full. Moreover, apart from convex optimization and constrained nonlinear least squares problems, general optimization and optimal control is left out. Also not included are recent results on interior point methods as well as inverse problems in L^2, even though affine invariance has just started to play a role in these fields.

Generally speaking, finite dimensional problems and techniques dominate the material presented here—however, with the declared intent that the finite dimensional presentation should filter out promising paths into the infinite dimensional part of the mathematical world. This intent is exemplified in several sections, such as

- Section 6.2 on ODE initial value problems, where stiff problems are analyzed via a simplified Newton iteration in function space—replacing the Picard iteration, which appears to be suitable only for nonstiff problems,
- Section 7.4.2 on ODE boundary value problems, where an adaptive multilevel collocation method is worked out on the basis of an inexact Newton method in function space,
- Section 8.1 on asymptotic mesh independence, where finite and infinite dimensional Newton sequences are synoptically compared, and
- Section 8.3 on elliptic PDE boundary value problems, where inexact Newton multilevel finite element methods are presented in detail.

The *algorithmic paradigm*, given in Section 1.2.3 and used all over the whole book, will certainly be useful in a much wider context, far beyond Newton methods.

Unfortunately, after having finished this book, I will probably lose all my scientific friends, since I missed to quote exactly that part of their work that should have been quoted by all means. I cannot but apologize in advance, hoping that some of them will maintain their friendship nevertheless. In fact, as the literature on Newton methods is virtually unlimited, I decided to not even attempt to screen or pretend to have screened all the relevant literature, but to restrict the references essentially to those books and papers that are either intimately tied to affine invariance or have otherwise been taken as direct input for the presentation herein. Even with this restriction the list is still quite long.

At this point it is my pleasure to thank all those coworkers at ZIB, who have particularly helped me with the preparation of this book. My first thanks go to Rainer Roitzsch, without whose high motivation and deep TEX knowledge this book could never have appeared. My immediate next thanks go to Erlinda Körnig and Sigrid Wacker for their always friendly cooperation over the long time that the manuscript has grown. Moreover, I am grateful to Ulrich Nowak, Andreas Hohmann, Martin Weiser, and Anton Schiela for their intensive computational assistance and invaluable help in improving the quality of the manuscript.

Nearly last, but certainly not least, I wish to thank Harry Yserentant, Christian Lubich, Matthias Heinkenschloss, and a number of anonymous reviewers for valuable comments on a former draft. My final thanks go to Martin Peters from Springer for his enduring support.

Berlin, February 2004

Peter Deuflhard

Table of Contents