# Outline of Contents

This book is divided into eight chapters, a reference list, a software list, and an index. After an elementary introduction in Chapter 1, it splits into two parts: Part I, Chapter 2 to Chapter 5, on finite dimensional Newton methods for *algebraic equations*, and Part II, Chapter 6 to Chapter 8, on extensions to ordinary and partial *differential equations*. Exercises are added at the end of each chapter.

**Chapter 1.** This introductory chapter starts from the historical root, Newton's method for scalar equations (Section 1.1). The method can be derived either *algebraically*, which leads to *local* Newton methods only (presented in Chapter 2), or *geometrically*, which leads to *global* Newton methods via the concept of the Newton path (see Chapter 3).

The next Section 1.2 contains the *key to the basic understanding of this monograph.* First, four affine invariance classes are worked out, which represent the four basic strands of this treatise:

- *affine covariance*, which leads to *error* norm controlled algorithms,
- *affine contravariance*, which leads to *residual* norm controlled algorithms,
- *affine conjugacy*, which leads to *energy* norm controlled algorithms, and
- *affine similarity*, which may lead to *time* step controlled algorithms.

Second, the affine invariant local estimation of affine invariant Lipschitz constants is set as the central *paradigm* for the construction of adaptive Newton algorithms.

In Section 1.3, we give a roadmap of the large variety of Newton-type methods—essentially fixing terms to be used throughout the book such as ordinary and simplified Newton method, Newton-like methods, inexact Newton methods, quasi-Newton methods, Gauss-Newton methods, quasilinearization, or inexact Newton multilevel methods. In Section 1.4, we briefly collect details about iterative linear solvers to be used as inner iterations within finite dimensional inexact Newton algorithms; each affine invariance class is linked with a special class of inner iterations. In view of function space oriented inexact Newton algorithms, we also revisit linear multigrid methods. Throughout this section, we emphasize the role of adaptive error control.

**PART I.** The following Chapters 2 to 5 deal with *finite dimensional* Newton methods for algebraic equations.

**Chapter 2.** This chapter deals with *local* Newton methods for the numerical solution of systems of nonlinear equations with finite, possibly large dimension. The term 'local' refers to the situation that 'sufficiently good' initial guesses of the solution are assumed to be at hand. Special attention is paid to the issue of how to recognize, whether a given initial guess $x^0$ is 'sufficiently good'. Different affine invariant formulations give different answers to this question, in theoretical terms as well as by virtue of the algorithmic paradigm of Section 1.2.3. Problems of this structure are called 'mildly nonlinear'; their computational complexity can be bounded a-priori in units of the computational complexity of the corresponding linearized system.

As it turns out, different affine invariant Lipschitz conditions, which have been introduced in Section 1.2.2, lead to different characterizations of local convergence domains in terms of error oriented norms, residual norms, or energy norms, which, in turn, give rise to corresponding variants of Newton algorithms. We give three different, strictly affine invariant convergence analyses for the cases of affine covariant (error oriented) Newton methods (Section 2.1), affine contravariant (residual based) Newton methods (Section 2.2), and affine conjugate Newton methods for convex optimization (Section 2.3). Details are worked out for ordinary Newton algorithms, simplified Newton algorithms, and inexact Newton algorithms—synoptically for each of the three affine invariance classes. Moreover, affine covariance is naturally associated with Broyden's 'good' quasi-Newton method, whereas affine contravariance corresponds to Broyden's 'bad' quasi-Newton method.

Affine invariant *globalization*, which means global extension of the convergence domains of local Newton methods in the affine invariant frame, is possible along several lines:

- global Newton methods with damping strategy—see Chapter 3,
- parameter continuation methods—see Chapter 5,
- pseudo-transient continuation methods—see Section 6.4.

**Chapter 3.** This chapter deals with *global* Newton methods for systems of nonlinear equations with finite, possibly large dimension. The term 'global' refers to the situation that here, in contrast to the preceding chapter, 'sufficiently good' initial guesses of the solution are no longer assumed. Problems of this structure are called 'highly nonlinear'; their computational complexity depends on topological details of Newton paths associated with the nonlinear mapping and can typically not be bounded a-priori.

In Section 3.1 we survey globalization concepts such as

- steepest descent methods,
- trust region methods,
- the Levenberg-Marquardt method, and
- the Newton method with damping strategy.

In Section 3.1.4, a rather general geometric approach is taken: the idea is to derive a globalization concept without a pre-occupation to any iterative method, just starting from the requirement of affine covariance as a 'first principle'. Surprisingly, this general approach leads to a topological derivation of Newton's method with damping strategy via Newton paths.

In order to accept or reject a new iterate, *monotonicity tests* are applied. We study different such tests, according to different affine invariance requirements:

- the most popular *residual* monotonicity test, which is related to affine contravariance (Section 3.2),
- the error oriented so-called *natural* monotonicity test, which is related to affine covariance (Section 3.3), and
- the convex functional test as the natural requirement in convex optimization, which reflects affine conjugacy (Section 3.4).

For each of these three affine invariance classes, *adaptive trust region strategies* are designed in view of an efficient choice of damping factors in Newton's method. They are all based on the *paradigm* of Section 1.2.3. On a theoretical basis, details of algorithmic realization in combination with either *direct* or *iterative* linear solvers are worked out. As it turns out, an efficient determination of the steplength factor in global inexact Newton methods is intimately linked with the accuracy matching for affine invariant combinations of inner and outer iteration.

**Chapter 4.** This chapter deals with both *local* and *global Gauss-Newton* methods for *nonlinear least squares* problems in finite dimension—a method, which attacks the solution of the nonlinear least squares problem by solving a sequence of linear least squares problems. Affine invariance of both theory and algorithms will once again play a role, here restricted to *affine contravariance* and *affine covariance*. The theoretical treatment requires considerably more sophistication than in the simpler case of Newton methods for nonlinear equations.

In order to lay some basis, unconstrained and equality constrained *linear* least squares problems are first discussed in Section 4.1, introducing the useful calculus of generalized inverses. In Section 4.2, an affine contravariant convergence analysis of Gauss-Newton methods is given and worked out in the direction of *residual* based algorithms. Local convergence turns out to be only guaranteed for 'small residual' problems, which can be characterized in theoretical and algorithmic terms. Local and global convergence analysis as well as adaptive trust region strategies rely on some *projected residual* monotonicity test. Both *unconstrained* and *separable* nonlinear least squares problems are treated.

In the following Section 4.3, local convergence of *error* oriented Gauss-Newton methods is studied in affine covariant terms; again, Gauss-Newton methods are seen to exhibit guaranteed convergence only for a restricted problem class, named 'adequate' nonlinear least squares problems, since they are seen to be adequate in terms of the underlying statistical problem formulation. The globalization of these methods is done via the construction of two topological paths: the local and the global Gauss-Newton path. In the special case of nonlinear equations, the two paths coincide to one path, the Newton path. On this theoretical basis, adaptive trust region strategies (including rank strategies) combined with a natural extension of the *natural* monotonicity test are presented in detail for *unconstrained*, for *separable*, and—in contrast to the residual based approach—also for nonlinearly *constrained* nonlinear least squares problems. Finally, in Section 4.4, we study *underdetermined* nonlinear systems. In this case, a *geodetic Gauss-Newton path* exists generically and can be exploited to construct a quasi-Gauss-Newton algorithm and a corresponding adaptive trust region method.

**Chapter 5.** This chapter discusses the numerical solution of parameter dependent systems of nonlinear equations, which is the basis for parameter studies in systems analysis and systems design as well as for the globalization of local Newton methods. The key concept behind the approach is the (possible) existence of a *homotopy path* with respect to the selected parameter. In order to follow such a path, we here advocate *discrete continuation methods*, which consist of two essential parts:

- a *prediction* method, which, from given points on the homotopy path, produces some 'new' point assumed to be 'sufficiently close' to the homotopy path,
- an iterative *correction* method, which, from a given starting point close to, but not on the homotopy path, supplies some point on the homotopy path.

For the prediction step, *classical* or *tangent continuation* are the canonical choices. Needless to say that, for the iterative correction steps, we here concentrate on local Newton and (underdetermined) Gauss-Newton methods. Since the homotopy path is a mathematical object in the domain space of the nonlinear mapping, we only present the *affine covariant* approach.

In Section 5.1, we derive an adaptive *Newton continuation* algorithm with the ordinary Newton method as correction; this algorithm terminates locally in the presence of critical points including turning points. In order to follow the path beyond turning points, a *quasi-Gauss-Newton continuation* algorithm is worked out in Section 5.2, based on the preceding Section 4.4. This algorithm still terminates in the neighborhood of any higher order critical point. In order to overcome such points as well, we exemplify a scheme to construct *augmented systems*, whose solutions are just selected critical points of higher order—see Section 5.3. This scheme is an appropriate combination of Lyapunov-Schmidt reduction and topological universal unfolding. Details of numerical realization are only worked out for the computation of diagrams including simple bifurcation points.

**PART II.** The following Chapters 6 to 8 deal predominantly with *infinite dimensional*, i.e., function space oriented Newton methods. The selected topics are stiff initial value problems for ordinary differential equations (ODEs) and boundary value problems for ordinary and partial differential equations (PDEs).

**Chapter 6.** This chapter deals with *stiff* initial value problems for ODEs. The discretization of such problems is known to involve the solution of nonlinear systems per each discretization step—in one way or the other.

In Section 6.1, the contractivity theory for linear ODEs is revisited in terms of *affine similarity*. Based on an affine similar convergence theory for a simplified Newton method in *function space*, a *nonlinear contractivity* theory for stiff ODE problems is derived in Section 6.2, which is quite different from the theory given in usual textbooks on the topic. The key idea is to replace the Picard iteration in function space, known as a tool to show uniqueness in nonstiff initial value problems, by a simplified Newton iteration in function space to characterize stiff initial value problems. From this point of view, *linearly implicit* one-step methods appear as direct realizations of the simplified Newton iteration in function space. In Section 6.3, exactly the same theoretical characterization is shown to apply also to *implicit* one-step methods, which require the solution of a nonlinear system by some finite dimensional Newton-type method at each discretization step.

Finally, in a deliberately longer Section 6.4, we discuss *pseudo-transient continuation* algorithms, whereby steady state problems are solved via stiff integration. This type of algorithm is particularly useful, when the Jacobian matrix is singular due to hidden dynamical invariants (such as mass conservation). The (nearly) affine similar theoretical characterization permits the derivation of an *adaptive (pseudo-)time step strategy* and an accuracy matching strategy for a residual based inexact variant of the algorithm.

**Chapter 7.** In this chapter, we consider nonlinear two-point boundary value problems for ODEs. The presentation and notation is closely related to Chapter 8 in the textbook [71]. Algorithms for the solution of such problems can be grouped into two approaches: *initial value* methods such as multiple shooting and *global discretization* methods such as collocation. Historically, affine covariant Newton methods have first been applied to this problem class—with significant success.

In Section 7.1, the realization of Newton and discrete continuation methods within the standard multiple shooting approach is elaborated. Gauss-Newton methods for parameter identification in ODEs are discussed in Section 7.2, also based on multiple shooting. For periodic orbit computation, Section 7.3 presents Gauss-Newton methods, both in the shooting approach (Sections 7.3.1 and 7.3.2) and in a Fourier collocation approach, also called Urabe or harmonic balance method (Section 7.3.3).

In Section 7.4 we concentrate on *polynomial* collocation methods, which have reached a rather mature status including affine covariant Newton methods. In Section 7.4.1, the possible discrepancy between discrete and continuous solutions is studied including the possible occurrence of so-called 'ghost solutions' in the nonlinear case. On this basis, the realization of *quasilinearization* is seen to be preferable in combination with collocation. The following Section 7.4.2 is then devoted to the key issue that quasilinearization can be interpreted as an *inexact Newton method in function space*: the approximation errors in the infinite dimensional setting just replace the inner iteration errors arising in the finite dimensional setting. With this insight, an adaptive multilevel control of the collocation errors can be realized to yield an adaptive inexact Newton method in function space—which is the bridge to adaptive Newton multilevel methods for PDEs (compare Section 8.3).

**Chapter 8.** This chapter deals with Newton methods for boundary value problems in nonlinear PDEs. There are two principal approaches: (a) finite dimensional Newton methods applied to a given system of already discretized PDEs, also called *discrete Newton methods*, and (b) function space oriented Newton methods applied to the continuous PDEs, at best in the form of *inexact Newton multilevel methods*.

Before we discuss the two principal approaches in detail, we present an affine covariant analysis of *asymptotic mesh independence* that connects the finite dimensional and the infinite dimensional Newton methods, see Section 8.1. In Section 8.2, we assume the standard situation in industrial technology software, where the grid generation module is strictly separated from the solution module. Consequently, nonlinear PDEs arise there as discrete systems of nonlinear equations with fixed finite, but usually high dimension and large sparse ill-conditioned Jacobian matrix. This is the domain of applicability of finite dimensional inexact Newton methods. More advanced, but often less favored in the huge industrial software environments, are *function space* oriented inexact Newton methods, which additionally include the adaptive manipulation of discretization meshes within a multilevel or multigrid solution process. This situation is treated in Section 8.3 and compared there with *finite dimensional* inexact Newton techniques.