Numerical Path Following

Eugene L. Allgower[1] and Kurt Georg[1]
Department of Mathematics
Colorado State University
Ft. Collins, CO 80523, USA
E-mail: Georg@Math.ColoState.Edu

August 1994

[1]Partially supported by the National Science Foundation via grant # DMS-9104058

Contents

Chapter I

Introduction

Because of their versatility and robustness, numerical continuation or path following methods have been finding ever wider use in scientific applications. Our aim here is to give an introduction to the main ideas of numerical path following and to present some of the recent advances in this subject regarding new adaptations, applications, and analysis of efficiency and complexity. Introductions into aspects of the subject of numerical path following may be found in the books by Garcia and Zangwill [1981], Gould and Tolle [1983], Keller [1987], Rheinboldt [1986], Seydel [1988], and Todd [1976a]. The philosophy and notation of the present paper will be that of our book Allgower and Georg [1990], which also contains an extensive bibliography up to 1990. An updating of the literature is given in Allgower and Georg [1993].

The viewpoint which will be adopted here is that numerical continuation methods are techniques for numerically approximating a solution curve c which is implicitly defined by an underdetermined system of equations. In the literature of numerical analysis, the terms *numerical continuation* and *path following* are used interchangeably.

There are various objectives for which the numerical approximation of c can be used, and depending upon the particular objective, the approximating technique ought to be adapted accordingly. In fact, continuation is a unifying concept under which various numerical methods that may seem to have very little in common can be interpreted and related. For example, simplicial fixed point methods for solving problems in mathematical economics, the generation of bifurcation diagrams of nonlinear eigenvalue problems involving partial differential equations, and the recently developed interior point methods for solving linear programming problems seem to be quite unrelated. Nevertheless, there is some benefit in considering

them as specific examples of path following. We personally are struck by the remarkable fact that a technique which was initially developed for overcoming difficulties involved with nonlinear problems now turns out to be extremely useful even for treating various problems which are essentially linear: e.g., linear eigenvalue problems, linear programming, and linear complementarity problems.

For readers who wish to skip directly to topics which are of particular interest to them, we give a brief outline of the contents of the paper.

Chapter 1 presents the basic ideas of predictor corrector path following methods and some convergence results are given. For theoretical and basic considerations, it is convenient to consider a solution curve parametrized according to (standard) arclength, since this parametrization is globally valid. However, specific implementations practically always deal with approximations of the arclength, the precise arclength is never calculated (except for using initial value solvers on the equation (1.8) which is not advisable). A typical example is the concept of pseudo arclength which is discussed in Section 2.

In Chapter III predictor steps and steplength strategies are studied. Several alternatives for higher order predictors are explored. Devices for calculating special points on a curve are also discussed, since they can be viewed in the context of steplength adaptation.

In Chapter IV the technical aspects of implementing predictor corrector methods are addressed, e.g., the numerical linear algebra involved in performing the predictor and corrector steps, approximating the derivative terms which arise in the course of numerical path following, update methods, special considerations for large scale problems. In particular, it is pointed out that the use of the pseudo inverse $H'(u)^+$ and use of the bordering algorithm, i.e., the use of $\tilde{H}'(u)^{-1}$ for an augmented sytem $\tilde{H}$, are closely related (also in terms of computational expense).

Chapter V deals with various applications of path following methods. We begin with a brief discussion of homotopy methods for fixed point problems and global Newton methods. Here the relationship can be seen of path following methods to celebrated results of Poincaré [1881–1886], Klein [1882–1883] and Bernstein [1910]. In fact, the path following methods involving homotopy deformations are essentially constructive implementations of degree arguments. The problem of finding multiple solutions is examined in detail. Of particular interest are recent homotopy methods for finding all solutions of polynomial systems of equations. We survey some path following aspects of nonlinear eigenvalue problems, and address the question of handling bifurcations. Also three rather new developments in path following are discussed: (1) The solution of linear eigenvalue problems via special homotopy approaches, (2) the handling of parametric program-

ming problems by following certain branches of critical points via active set strategies, (3) the path following aspects involved in the interior point methods for solving linear and quadratic programming problems.

Chapter VI presents an introduction to the principles of piecewise linear methods. These methods view path following in a different light: Instead of approximately following a solution curve of a smooth system $H(u) = 0$, they exactly follow the solution curve (i.e., a polygonal path) of a piecewise affine system $\tilde{H}(u) = 0$ which may or may not be related to a smooth problem $H(u) = 0$. Some instances where these methods are useful are discussed, e.g., linear complementarity problems or homotopy methods where predictor corrector methods are not implementable, because of lack of smoothness. We also briefly address the related topic of approximating implicitly defined surfaces. Several error estimate results are given.

The issue of the computational complexity of path following is considered in Chapter VII. This issue is related to the Newton-Kantorovich theory and is currently of considerable interest in the context of interior point methods.

We conclude by listing some available software related to path following and indicate how the reader might access these codes. We make no attempt to compare or evaluate the various codes. Our opinion is that path following codes ought to be adapted to the special purposes for which they are to be used. Although there are some general purpose codes, probably none will slay every dragon and by making appropriate adaptatations for the special purpose at hand, improved efficiency can usually be achieved.

Many colleagues helped us with suggestions and references. In particular, we are grateful to Hubert Schwetlick for valuable contributions.

Chapter II

Basics

1 Predictor Corrector Path Following

In the context of numerical continuation methods, one considers curves which are implicitly defined by an underdetermined system of equations

$$H(u) = 0 \tag{1.1}$$

where $H : \mathbb{R}^{N+1} \to \mathbb{R}^N$ is a smooth map.

When we say that a map is smooth we shall mean that it has as many continuous derivatives as the context of the subsequent discussion requires. For convenience, the reader may assume smoothness means C^∞.

More generally, if $H : \mathbb{R}^{N+K} \to \mathbb{R}^N$, we would expect the solution set of $H(u) = 0$ to describe a K-dimensional manifold. However, to ensure this, we need a non-degeneracy condition: We call u a *regular point* of H if the Jacobian $H'(u)$ has maximal rank. We call y a *regular value* of H if u is a regular point of H whenever $H(u) = y$. In particular, if $y \notin \operatorname{range} H$, then y is a regular value. If a point or value is not regular, then it is called *singular*.

With these conditions, we may now state the classical implicit function theorem in a global setting, see, e.g., Milnor [1969]:

Theorem 1.1 (Implicit Function Theorem) *Let $H : \mathbb{R}^{N+K} \to \mathbb{R}^N$ be a smooth map such that $0 \in \operatorname{range}(H)$. Then*

$$M = \{x \in \mathbb{R}^{N+K} : H(x) = 0,\ x \text{ is a regular point of } H\}$$

is a smooth K-dimensional manifold.

Unless we explicitly say otherwise, we now assume hereafter that $H : \mathbb{R}^{N+1} \to \mathbb{R}^N$ is smooth. In particular, if $u_0 \in \mathbb{R}^{N+1}$ is a regular point of H such that $H(u_0) = 0$, it follows from the Implicit Function Theorem that the solution set $H^{-1}(0)$ can be locally parametrized about u_0 with respect to some parameter, say s. We thus obtain the *solution curve* $c(s)$ of the equation $H(u) = 0$.

For convenience of the theoretical discussions, we usually consider parametrization with respect to arclength, though in practical implementations it is often more convenient to use related parametrizations which are easier to incorporate, see in particular KELLER's [1977] pseudo arclength parametrization. In terms of the intrinsic properties of the curve there is no difference which underlying parametrization is used.

Hence, by a re-parametrization (according to arclength), we obtain a smooth solution curve $c : J \to \mathbb{R}^{N+1}$ for some open interval J containing zero such that for all $s \in J$:

$$c(0) = u_0 \tag{1.2}$$

$$H'\big(c(s)\big)\dot{c}(s) = 0, \tag{1.3}$$

$$||\dot{c}(s)|| = 1, \tag{1.4}$$

$$\det \begin{pmatrix} H'\big(c(s)\big) \\ \dot{c}(s)^* \end{pmatrix} > 0. \tag{1.5}$$

The above conditions uniquely determine the tangent $\dot{c}(s)$ with a specific orientation. Here and in the following, B^* denotes the Hermitian transpose of B, $||u||$ the Euclidean norm of u, H' the total derivative (e.g., the Jacobian) of H, and $\dot{c}$ the derivative of c with respect to arclength. Condition (1.4) normalizes the parametrization to arclength. This is only for theoretical convenience, and it is not an intrinsic restriction. Condition (1.5) chooses one of the two possible orientations.

The preceding discussion motivates the following definition: Let A be an $(N, N+1)$-matrix with maximal rank. For the purpose of our exposition, the unique vector $t(A) \in \mathbb{R}^{N+1}$ satisfying the conditions

$$At = 0, \quad ||t|| = 1, \quad \det \begin{pmatrix} A \\ t^* \end{pmatrix} > 0, \tag{1.6}$$

will be called the *tangent vector induced by* A.

The vector $\dot{c}(s)$ will hereafter be referred to as the *oriented unit tangent vector to the solution curve* c. At a point $u \in \mathbb{R}^{N+1}$ the matrix

$$\begin{pmatrix} H'(u) \\ t\big(H'(u)\big)^* \end{pmatrix} \tag{1.7}$$

is called the *augmented Jacobian* of H at u.

Making use of this definition, the above solution curve $c(s)$ is characterized as the solution of the initial value problem

$$\dot{u} = t\big(H'(u)\big), \quad u(0) = u_0 \tag{1.8}$$

which in this context is sometimes attributed to DAVIDENKO [1953], see also BRANIN [1972]. Note that the domain

$$\{u \in \mathbb{R}^{N+1} : u \text{ is a regular point of } H\}$$

is open, and thus the right-hand side of the differential equation is defined on an open domain. The initial value problem (1.8) is not used in efficient path following algorithms, but it serves as a useful device for analyzing the path. The following two statements illustrate this:

THEOREM 1.2 *Let (a, b) be the maximal interval of existence for the solution $c(s)$ of the initial value problem (1.8). If a is finite, then $c(s)$ converges to a singular zero point of H as $s \to a$, $s > a$. An analogous statement holds if b is finite.*

PROOF Since $c(s)$ satisfies the defining initial value problem (1.8), we have

$$c(s_1) - c(s_2) = \int_{s_2}^{s_1} t\Big(H'\big(c(\xi)\big)\Big)\, d\xi \quad \text{for} \quad s_1, s_2 \in (a, b).$$

Because the integrand has unit norm, it follows that

$$||c(s_1) - c(s_2)|| \le |s_1 - s_2| \quad \text{for} \quad s_1, s_2 \in (a, b).$$

If $\{s_n\}_{n=1}^{\infty} \subset (a, b)$ is a sequence such that $s_n \to a$ as $n \to \infty$, then the above inequality shows that the sequence $\{c(s_n)\}_{n=1}^{\infty}$ is Cauchy. Hence it converges to a point $\tilde{u}$. By continuity it follows that $H(\tilde{u}) = 0$. We will show by contradiction that $c(a)$ is a singular point of H. Suppose that $\tilde{u}$ is a regular point of H. Then using the initial point $u(0) = \tilde{u}$ in the defining initial value problem (1.8), we obtain a local solution $\tilde{c}(s)$. This solution is unique in some neighborhood U of $\tilde{u}$ (*local uniqueness of solutions*). If $\xi > 0$ is small enough, then $c(a + \xi) \in U$, and hence $c(a + \xi) = \tilde{c}(\xi)$ for $\xi > 0$ holds by the local uniqueness of solutions. Thus it follows that c can be extended beyond a by setting $c(a + \xi) := \tilde{c}(\xi)$ for $\xi \le 0$, contradicting the maximality of the interval (a, b). □

THEOREM 1.3 *Let zero be a regular value of H. Then the solution curve c is defined on the real line and satisfies one of the following two conditions:*

(a) The curve c is diffeomorphic to a circle. More precisely, there is a period $T > 0$ such that $c(s_1) = c(s_2)$ if and only if $s_1 - s_2$ is an integer multiple of T.

(b) The curve c is diffeomorphic to the real line. More precisely, c is injective, and $c(s)$ has no accumulation point for $s \to \pm\infty$.

PROOF Since zero is a regular value, no zero point of H is singular, and by the preceding theorem, c is defined on all of $\mathbb{R}$. Furthermore, since the defining differential equation $\dot{u} = t\big(H'(u)\big)$ is autonomous, the following translation invariance holds: for all $s_0 \in \mathbb{R}$, the curve $s \mapsto c(s_0 + s)$ is also a solution of $\dot{u} = t\big(H'(u)\big)$. Let us now consider the two possibilities:

1. The curve c is not injective. We define $T := \min\{s > 0 : c(s) = c(0)\}$. By the local uniqueness of the solutions of initial value problems and by the above mentioned translation invariance, the assertion (a) follows.

2. The curve c is injective. We show assertion (b) by contradiction. Let us assume without loss of generality that $\tilde{u}$ is an accumulation point of $c(s)$ as $s \to \infty$. By continuity, $H(\tilde{u}) = 0$. Since $\tilde{u}$ is a regular point of H, we can use the initial point $u(0) = \tilde{u}$ in (1.8) to obtain a solution $\tilde{c}$. By local uniqueness, the two curves c and $\tilde{c}$ must coincide up to a translation, and hence there exists an $s_1 > 0$ such that $c(s_1) = \tilde{u}$. Since $\tilde{u}$ is also an accumulation point of $c(s_1 + s)$ as $s \to \infty$, and since the curve $s \mapsto c(s_1 + s)$ is also a solution curve, the above argument can be repeated to obtain an $s_2 > 0$ such that $c(s_1 + s_2) = \tilde{u}$. This contradicts the injectivity of c.

□

Since the solution curve c is characterized by the initial value problem (1.8), it is evident that the numerical methods for solving initial value problems could immediately be used to numerically trace c. However, in general this is not an efficient approach, since it ignores the strong contractive properties which the curve c has relative to corrector steps in view of the fact that it satisfies the equation $H(u) = 0$. In fact, a typical path following method consists of a succession of two different steps:

Predictor step. An approximate step along the curve, usually in the general direction of the tangent of the curve. The initial value problem (1.8) provides motivation for generating predictor steps in the spirit of the techniques of numerical solution of initial value problems.

Corrector steps. One or more iterative steps for solving $H(u) = 0$ (typically of Newton or gradient type) which bring the predicted point back to the curve.

It is usual to call such procedures *predictor corrector* path following methods. However, let us emphasize that this name should not be confused with the predictor corrector multistep methods for initial value problems, since the latter employ correctors which do not converge back to the solution curve.

The following pseudocode shows the basic steps of a generic predictor corrector method, see also Figure 1.1.

ALGORITHM 1.1 **(Generic Predictor Corrector Method)**

INPUT $u_0 \in \mathbb{R}^{N+1}$ such that $H(u_0) \approx 0$ (initial point)
$h > 0$ initial steplength
OUTPUT u_i , $i = 0, 1, 2, \ldots$, approximating the solution curve

FOR $i = 1, 2, \ldots$
% *predictor step*
predict v_i such that $H(v_i) \approx 0$ and $||v_i - u_{i-1}|| \approx h$ ...
and $v_i - u_{i-1}$ points in the direction of traversing
% *corrector step*
let $u_i \in \mathbb{R}^{N+1}$ approximately solve ...
$\min_w \{ ||v_i - w|| : H(w) = 0\}$
% *steplength adaptation*
choose a new steplength $h > 0$
END FOR

The predictor corrector type of algorithms for curve following seem to date to HASELGROVE [1961]. In contrast to the present predictor corrector methods, the classical embedding methods assume that the solution path is parametrized with respect to an explicit parameter which is identified with the last variable in H. Hence, we consider the equation (1.1) in the form

$$H(x, \lambda) = 0. \tag{1.9}$$

If we assume that the partial derivative $H_x(x, \lambda)$ is nonsingular, then the solution curve can be parametrized in the form $(x(\lambda), \lambda)$. This assumption has the drawback that *folds* are excluded, i.e., points such that $H(x, \lambda) = 0$ and $\det H_x(x, \lambda) = 0$. Such points are sometimes called *turning points* in the literature. However, the assumption has the advantage that the corrector steps can be more easily handled, in particular if H_x is sparse. In some applications it is known a priori that no folds are present, and then

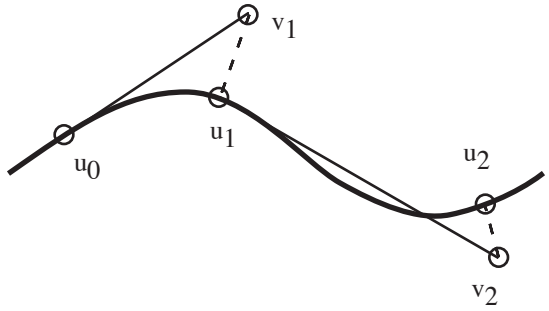

Figure 1.1. A Generic Predictor Corrector Method

the embedding method is applicable. But even in this case, the predictor corrector methods of the type in Algorithm 1.1 are likely to be more efficient since it may be necessary to use very small steps in λ if det H_x is near zero. For purposes of illustration we present a generic embedding method which is an analogue of the above predictor corrector method:

ALGORITHM 1.2 **(Generic Embedding Method)**

INPUT $(x_0, \lambda_0) \in \mathbb{R}^{N+1}$ such that $H(x_0, \lambda_0) \approx 0$ (initial point)
$h > 0$, an initial steplength
OUTPUT (x_i, λ_i), $i = 0, 1, 2, \ldots$,
approximating the solution curve

FOR $i = 1, 2, \ldots$
$\lambda_i \leftarrow \lambda_{i-1} + h$
find $x_i \in \mathbb{R}^N$ such that $H(x_i, \lambda_i) \approx 0$
choose a new steplength $h > 0$
END FOR

In the preceding algorithm the predictor step is hidden; the predictor point would correspond to the starting point of an iterative method for solving $H(x, \lambda_i) = 0$. The most commonly used starting point is the previous point x_{i-1}.

2 Pseudo Arclength Algorithms

It is common to blend aspects of the above two algorithms 1.1,1.2. A simple example is to use a predictor tangent to the curve $(x(\lambda), \lambda)$ in the embedding algorithm 1.2. A more sophisticated example is the use of a so-called

pseudo arclength parametrization in connection with a bordering algorithm (see Section 13) in the corrector phase of the predictor corrector method, as introduced by KELLER [1977, 1983, 1987]. To avoid dealing with the arclength parameter, one can also adopt a strategy of parameter switching, see, e.g., RHEINBOLDT [1980, 1981]. In this approach, one changes the embedding parameter when nearing a fold point with respect to the current parameter during the traversing of the curve.

Let us give a brief account of these ideas. We assume that 0 is a regular value of H. If $\bar{u}$ is a current solution point, i.e., $H(\bar{u}) = 0$, a (local) *pseudo arclength parameter* s is introduced via an additional equation

$$F(u, s) = 0 \tag{2.1}$$

where $F : \mathbb{R}^{N+1} \times \mathbb{R} \to \mathbb{R}$ is sufficiently smooth and such that $F(\bar{u}, 0) = 0$. We denote by F_u the (row of) partial derivatives of F with respect to u and assume that the matrix

$$\begin{pmatrix} H'(\bar{u}) \\ F_u(\bar{u}, 0) \end{pmatrix}$$

is nonsingular, i.e., $F_u(\bar{u}, 0)\, t\big(H'(\bar{u})\big) \neq 0$, or better $F_u(\bar{u}, 0)\, t\big(H'(\bar{u})\big) > 0$ if we want to follow the solution curve in the direction of positive orientation. The Implicit Function Theorem guarantees that the solution curve can be parametrized near $\bar{u}$ according to s. Differentiating the system

$$\begin{pmatrix} H\big(u\big) \\ F\big(u, s\big) \end{pmatrix} = 0$$

with respect to the pseudo arclength parameter s yields the initial value problem

$$\dot{u} = \begin{pmatrix} H'(u) \\ -F_u(u, s)/F_s(u, s) \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad u(0) = \bar{u}. \tag{2.2}$$

We compare this characterization of the solution curve with (1.8) which can also be written in the form

$$\dot{u} = t\big(H'(u)\big) = \begin{pmatrix} H'(u) \\ t\big(H'(u)\big)^* \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad u(0) = \bar{u}. \tag{2.3}$$

By applying a standard sensitivity analysis to the initial value problems (2.2), (2.3) we obtain the following

THEOREM 2.1 *Let $c(s)$ be a solution to the initial value problem (2.3) and $d(s)$ the solution of (2.2). Hence, the two curves differ only in the choice*

of the parameter, $c(s)$ having exact arclength parameter and $d(s)$ a pseudo arclength parameter. Let

$$\left\| \lim_{s\to 0+} -F_u\big(d(s),s\big)^* / F_s\big(d(s),s\big) - t\big(H'(\bar u)\big) \right\| \le \varepsilon.$$

Then

$$\begin{aligned} ||\dot c(s) - \dot d(s)|| &\le \mathcal{O}(\varepsilon) + \mathcal{O}(s), \\ ||c(s) - d(s)|| &\le \mathcal{O}(\varepsilon s) + \mathcal{O}(s^2). \end{aligned}$$

The above theorem justifies the name "pseudo arclength" for the parameter introduced in (2.1) for the case that ε is small or even zero. In most applications, a convenient choice of the pseudo arclength parameter is easier to handle than the exact arclength parameter and leads to an efficient algorithm. The exact arclength parametrization is mainly useful for theoretical considerations, since it is global and simple. In fact, very few implementations of numerical continuation methods really maintain an arclength parametrization, the typical implementation uses an approximation in one way or another.

The two different parametrizations also suggest different corrector iterations. The pseudo arclength parametrization suggests a Newton iteration of the form

$$u_{n+1} = u_n - \begin{pmatrix} H'(u_n) \\ F_u(u_n,s) \end{pmatrix}^{-1} \begin{pmatrix} H(u_n) \\ F(u_n,s) \end{pmatrix}. \tag{2.4}$$

The exact arclength parametrization suggests the Gauss-Newton iteration

$$u_{n+1} = u_n - H'(u_n)^+ H(u_n), \tag{2.5}$$

see (3.3), (5.2) and Sections 4–5. However, implementations of numerical continuation methods are free to choose any corrector iteration, including the above two conceptually simplest ones. Furthermore, we will see in Sections 12,13 that the numerical linear algebra and the computational expense involved in a step of one of the two methods (2.4) or (2.5) are very similar.

Let us list a few pseudo arclength choices which are common. A very popular choice is linked to the Euler predictor (5.1) of Section 5.

$$F(u,s) = t\big(H'(\bar u)\big)^*(u-\bar u) - s. \tag{2.6}$$

It follows that

$$-F_u\big(u,s\big)^* / F_s\big(u,s\big) = t\big(H'(\bar u)\big)$$

and hence $\varepsilon = 0$ in the above Theorem 2.1.

Another popular choice is a secant predictor, where the tangent $t\big(H'(\bar{u})\big)$ in (2.6) which may be computationally expensive to obtain is replaced by the secant $||\bar{u}-p||^{-1}\big(\bar{u}-p\big)$ which essentially comes for free if we use a previous point p on the solution curve. It is easy to see that $\varepsilon = \mathcal{O}(||\bar{u}-p||)$ in this case.

Another choice of a pseudo arclength parameter is linked to the idea of finding a point on the solution curve at Euclidean distance s from $\bar{u}$. This leads to

$$F(u,s) = ||u-\bar{u}|| - s.$$

It follows that

$$-F_u\big(d(s),s\big)^*/F_s\big(d(s),s\big) = (d(s)-\bar{u})/||d(s)-\bar{u}|| = t\big(H'(\bar{u})\big) + \mathcal{O}(s)$$

and hence again $\varepsilon = 0$. This parametrization has been successfully used by Menzel and Schwetlick [1985].

Finally, the simplest choice is a parametrization according to some chosen co-ordinate (usually the eigenvalue parameter, if possible):

$$F(u,s) = e_i^*(u-\bar{u}) - s$$

where e_i is a unit basis vector, $i \in \{1,\ldots,N+1\}$. In this case, the iteration (2.4) can be reduced by setting $e_i^* u = e_0^*\bar{u} + s$ and iterating

$$E_i u_{n+1} = E_i u_n - (H'(u_n)E_i)^{-1}H(u_n)$$

where E_i is obtained from the (N+1,N+1) identity matrix by deleting column i. This choice of parametrization (at least for the corrector phase) in connection with a suitable strategy for switching co-ordinates is very popular in structural engineering, see, e.g., Rheinboldt [1980, 1981].

Let us conclude this section by giving a pseudo code.

Algorithm 2.1 **(Generic Pseudo Arclength Method)**

INPUT $u_0 \in \mathbb{R}^{N+1}$ such that $H(u_0) \approx 0$ (initial point)
$h > 0$ initial steplength
OUTPUT u_i , $i = 0,1,2,\ldots$, approximating the solution curve
FOR $i = 1,2,\ldots$
% *choose a local pseudo arclength parameter*
find F such that $F_u(u_{i-1},0)\,t\big(H'(u_{i-1})\big) > 0$
% *predictor step*
predict v_i such that $H(v_i) \approx 0$ and $F(v_i,h) \approx 0$
% *corrector step*
let $u_i \in \mathbb{R}^{N+1}$ approximately solve ...

$H(u) = 0$ and $F(u,h) = 0$ starting an iteration at v_i
% *steplength adaptation*
choose a new steplength $h > 0$
END FOR

3 The Moore-Penrose Inverse

The corrector step in Algorithm 1.1 tries to approximately solve an underdetermined nonlinear system of equations from a given starting point (predictor point). An analogue of Newton's method for underdetermined nonlinear systems can be formulated by using the Moore-Penrose inverse. This is often called the *Gauss-Newton method*. It has been discussed in several monographs, see, e.g., Ortega and Rheinboldt [1970] or Ben-Israel and Greville [1974]. More precisely, a Gauss-Newton step approximately solves the minimization problem in Algorithm 1.1. Since this concept will occur on several occasions in the following discussions, we begin with a brief review of the role of the Moore-Penrose inverse in this regard.

Let us consider the minimization problem in Algorithm 1.1. If u_i is a solution, then it satisfies the Lagrangian equations

$$H(u_i) = 0, \quad u_i - v_i = H'(u_i)^*\lambda \tag{3.1}$$

for some vector of multipliers $\lambda \in \mathbb{R}^N$. The second condition is equivalent to $u_i - v_i \in \text{range}\big(H'(u_i)^*\big) = \big\{t\big(H'(u_i)\big)\big\}^\perp$. Thus a necessary condition for u_i to solve the minimization problem is that it satisfies the equations

$$H(u_i) = 0, \quad t\big(H'(u_i)\big)^*(u_i - v_i) = 0. \tag{3.2}$$

In Newton's method, this nonlinear system is solved approximately via a linearization about v_i. To illustrate this, let us consider the Taylor expansion about v_i:

$$\begin{aligned} H(u_i) &= H(v_i) + H'(v_i)(u_i - v_i) + \mathcal{O}\big(||u_i - v_i||^2\big), \\ t\big(H'(u_i)\big)^*(u_i - v_i) &= t\big(H'(v_i)\big)^*(u_i - v_i) + \mathcal{O}\big(||u_i - v_i||^2\big). \end{aligned}$$

Hence, the linearization of (3.2) consists of neglecting the higher order terms $\mathcal{O}(||u_i - v_i||^2)$. As is usual in Newton's method, we obtain an approximation $\mathcal{N}(v_i)$ to the solution u_i, which has a truncation error of second order. Hence, the Newton point $\mathcal{N}(v_i)$ satisfies the following equations:

$$\begin{aligned} &H(v_i) + H'(v_i)\big(\mathcal{N}(v_i) - v_i\big) = 0, \\ &t\big(H'(v_i)\big)^*\big(\mathcal{N}(v_i) - v_i\big) = 0. \end{aligned}$$

As can be seen from the discussion below, see (3.4), the solution of the latter equations can be written in terms of the Moore-Penrose inverse $H'(v_i)^+$ of $H'(v_i)$:

$$\mathcal{N}(v_i) := v_i - H'(v_i)^+ H(v_i). \tag{3.3}$$

The map $\mathcal{N}$ defined on the regular points of H will for our purposes be called the *Gauss-Newton map*. Note that the Gauss-Newton method is analogous to the classical Newton's method, with the only formal difference being that the Moore-Penrose inverse $H'(v)^+$ replaces the standard inverse.

For completeness, let us review some features of the Moore-Penrose inverse which pertain to our special context. Let A be an $(N, N+1)$ matrix with maximal rank, and let $t(A)$ be its tangent vector, see (1.6). Then it is easy to see that the following statements are equivalent for all $b \in \mathbb{R}^N$ and $x \in \mathbb{R}^{N+1}$:

$$Ax = b \quad \text{and} \quad t(A)^* x = 0, \tag{3.4}$$

$$x = A^+ b \text{ where } A^+ := A^*(AA^*)^{-1}, \tag{3.5}$$

$$x \text{ solves the problem } \min_w \{\, ||w|| : Aw = b \,\}. \tag{3.6}$$

The last problem is one of the standard linear least squares problems. The matrix A^+ is the *Moore-Penrose inverse* of A, see, e.g., GOLUB and VAN LOAN [1989] for a general discussion on how to implement the action of A^+. Note that we are using the Moore-Penrose inverse only for the special case of an $(N, N+1)$ matrix with maximal rank.

Some further properties of the Moore-Penrose inverse are:

$$\begin{aligned} A^+A &= \mathrm{Id} - t(A)t(A)^* && (3.7)\\ &\quad \text{(orthogonal projection onto range } A^*), \\ AA^+ &= \mathrm{Id}, && (3.8)\\ A^+ &= \big(\mathrm{Id} - t(A)t(A)^*\big)B && (3.9)\\ &\quad \text{holds for any right inverse } B \text{ of } A. \end{aligned}$$

4 The Gauss-Newton Method

In (3.3) the Gauss-Newton map was introduced as a corrector step. Full theoretical discussions of the Gauss-Newton method can be found in several monographs, see, e.g., ORTEGA and RHEINBOLDT [1970] or BEN-ISRAEL and GREVILLE [1974], see also DEUFLHARD and HEINDL [1979]. Let us discuss some important features of the Gauss-Newton map $v \mapsto \mathcal{N}(v)$.

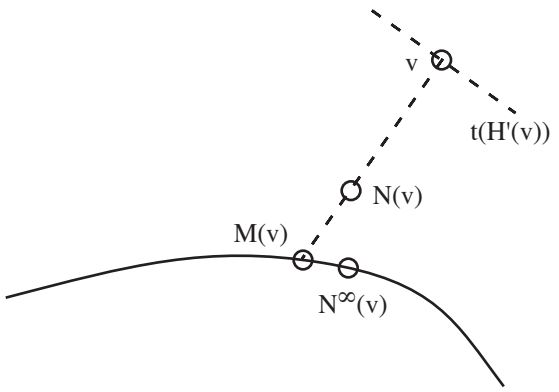

Figure 4.1. Diagram for the Proof of Theorem 4.1

Since our context deals with the curve-following problem, we confine our discussion to the case $H : \mathbb{R}^{N+1} \to \mathbb{R}^N$. Generalizations to the case $H : \mathbb{R}^{N+K} \to \mathbb{R}^N$ for $K > 1$ are straightforward.

THEOREM 4.1 *Let $H : \mathbb{R}^{N+1} \to \mathbb{R}^N$ be a smooth map having zero as a regular value. Then there exists an open neighborhood $U \supset H^{-1}(0)$ such that the following assertions hold:*

1. *Points in U are regular.*
2. *U is stable under the Gauss-Newton map, i.e., $\mathcal{N}(U) \subset U$.*
3. *For $v \in U$ the sequence $\{\mathcal{N}^i(v)\} \subset U$ converges. Let us call the limit $\mathcal{N}^\infty(v)$. Clearly, $\mathcal{N}^\infty(v) \in H^{-1}(0)$.*
4. *The following quadratic convergence*

$$||\mathcal{N}(v) - \mathcal{N}^\infty(v)|| \leq \mathcal{O}(||v - \mathcal{N}^\infty(v)||^2)$$

holds uniformly for v in compact subsets of U.

PROOF We will only give a brief sketch of the main points. See ALLGOWER and GEORG [1990, (3.4.1)] for further details.

For $v \in \mathbb{R}^{N+1}$ we define $\mathcal{M}(v)$ as the solution of the minimization problem

$$\min_w \{||w - v|| : H(w) = 0\},$$

see Figure 4.1. It is clear that we can find an open neighborhood $U_1 \supset H^{-1}(0)$ consisting of regular points such that the map $\mathcal{M} : U_1 \to H^{-1}(0)$ is uniquely defined. Recall, see (3.2) and following, that a Gauss-Newton step was motivated by this minimization. It is hence clear from classical results on Newton's method that

$$||\mathcal{N}(v) - \mathcal{M}(v)|| \leq \mathcal{O}(||v - \mathcal{M}(v)||^2)$$

holds uniformly for v in compact subsets of U_1.

Let us introduce the following notation: $B_\rho(w) := \{u : ||u - w|| < \rho\}$. Let us choose for each $w \in H^{-1}(0)$ an $\varepsilon(w) > 0$ such that $B_{\varepsilon(w)}(w) \subset U_1$. From the above estimates it is not hard, but somewhat technical to show that for each $w \in H^{-1}(0)$ there is a $\delta(w) > 0$, $\delta(w) < \varepsilon(w)$, such that $\{\mathcal{N}^i(v)\} \subset B_{\varepsilon(w)}(w)$ and $\mathcal{N}^\infty(v) \subset B_{\varepsilon(w)}(w)$ for $v \in B_{\delta(w)}(w)$.

We now define

$$U_2 := \bigcup_{w \in H^{-1}(0)} B_{\delta(w)}(w)$$

and

$$U := \{u : \{\mathcal{N}^i(u)\}_{i=0,1,2,\ldots} \subset U_2\}.$$

Let us show that U is open. Choose $v \in U$. We set $w := \mathcal{N}^\infty(v)$. It is possible to find a ball $B_\rho(w)$ such that $\mathcal{N}^i(B_\rho(w)) \subset B_{\delta(w)}(w)$ for $i = 1, 2, \ldots$. Let $k > 0$ be such that $\mathcal{N}^i(v) \in B_\rho(w)$ for $i \geq k$. Then

$$\{u : u, \mathcal{N}(u), \ldots, \mathcal{N}^k(u) \in U_2\,,\, \mathcal{N}^k(u) \in B_\rho(w)\}$$

is an open neighborhood of v contained in U.

The remaining assertions of the theorem are easy to obtain. □

For a sensitivity analysis and for certain applications it is worthwhile to note that the map $v \mapsto \mathcal{N}^\infty(v)$ is smooth. This fact is surprisingly difficult to show:

THEOREM 4.2 *An additional conclusion to Theorem 4.1 is that the map $\mathcal{N}^\infty : U \to H^{-1}(0)$ is smooth.*

PROOF We only sketch the main steps of the proof which is due to BEYN [1992]. In view of the fact that $\mathcal{N} : U \to U$ is smooth, we only need to find for each $u_0 \in H^{-1}(0)$ an open neighborhood U_0 of u_0 such that the restriction $\mathcal{N}^\infty : U_0 \to H^{-1}(0)$ is smooth. Furthermore, without loss of generality, we assume $u_0 = 0$. We abbreviate $t := t(H'(0))$.

We next observe that near $u_0 = 0$, the solution curve can be parametrized by $c(h)$ which satisfies the system

$$H(c(h)) = 0, \qquad t^*(ht - c(h)) = 0.$$

For $0 < \rho < 1$ we introduce a Banach space of sequences which converge to zero with at least the geometrical rate ρ:

$$\mathcal{B}_\rho := \{\{v_n\} \subset \mathbb{R}^{N+1} : \underbrace{\sup ||\rho^{-n} v_n||}_{\text{norm!}} < \infty\}.$$

Let us denote $W := (\ker H'(0))^\perp$. For a sufficiently small open neighborhood $\mathcal{V}$ of $\mathcal{B}_\rho \times W \times \mathbb{R}$ the map

$$\Gamma : \mathcal{V} \to \mathcal{B}_\rho \times W$$

is defined by

$$\begin{pmatrix} \{v_n\} \\ w \\ h \end{pmatrix} \mapsto \begin{pmatrix} \{\mathcal{N}(v_n + c(h)) - (v_{n+1} + c(h))\} \\ (\mathrm{Id} - tt^*)v_0 - w \end{pmatrix}.$$

Note that $\Gamma(0) = 0$.

Next, it is possible to show that the derivative

$$\partial_1 \Gamma : \mathcal{B}_\rho \to \mathcal{B}_\rho \times W$$

is an isomorphism. Hence, by the implicit function theorem, the solution of the equation

$$\Gamma \begin{pmatrix} \{v_n\} \\ w \\ h \end{pmatrix} = 0$$

can be (locally) parametrized by w and h, see Figure 4.2. From the definition of Γ it is clear that $\{v_n(w,h) + c(h)\}$ is obtained via Gauss-Newton steps, such that the starting point $v_0(w,h) + c(h)$ satisfies the condition $(\mathrm{Id} - tt^*)v_0(w,h) = w$.

Finally, we consider the map

$$\Phi : W \times \mathbb{R} \to \mathbb{R}^{N+1}$$

defined by

$$(w,h) \mapsto c(h) + v_0(w,h).$$

Clearly $\Phi(0) = 0$, and it can be seen that $\Phi'(0)(w,h) = w + ht$ which is nothing else but the natural isomorphism between $W \times \mathbb{R}$ and $\mathbb{R}^{N+1}$. Hence, by the inverse function theorem, Φ can be locally inverted.

Let z with $||z||$ sufficiently small be a starting point for a Gauss-Newton iteration. Applying Φ^{-1} we obtain $(w,h) := \Phi^{-1} z$. Hence, we can identify z with the starting point of a Gauss-Newton sequence $\{v_n(w,h) + c(h)\}$

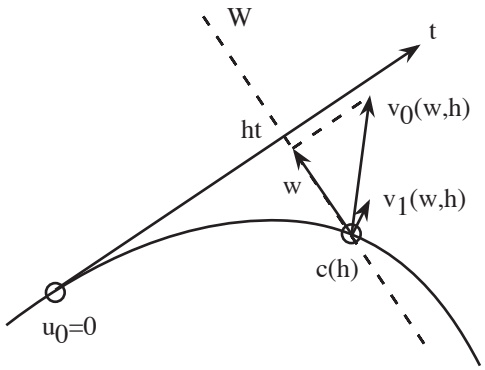

Figure 4.2. Diagram for the Proof of Theorem 4.2

converging to $c(h)$. The smoothness of $\mathcal{N}^\infty$ now becomes evident from the following composition of smooth maps:

$$\mathcal{N}^\infty : z \mapsto (w, h) \mapsto c(h).$$

□

5 Euler-Newton Continuation

A straightforward way of approximating a solution curve is to alternatingly perform a predictor step along the tangent of the curve

$$v = u + ht\big(H'(u)\big), \tag{5.1}$$

and a Gauss-Newton corrector step:

$$u := v - H'(v)^+ H(v). \tag{5.2}$$

The predictor step along the tangent can also be viewed as an Euler step for the differential equation (1.8). Here $h > 0$ represents a current stepsize.

The following algorithm sketches this procedure where approximate Euler and Newton steps are used.

ALGORITHM 5.1 **(Euler-Newton Method)**

INPUT $u_0 \in \mathbb{R}^{N+1}$ such that $H(u_0) \approx 0$ (initial point)
$h > 0$ initial steplength
OUTPUT u_i , $i = 0, 1, 2, \ldots$, approximating the solution curve
FOR $i = 1, 2, \ldots$
approximate $A_{i-1} \approx H'(u_{i-1})$
$v_i \leftarrow u_{i-1} + ht(A_{i-1})$ % *predictor step*
$u_i \leftarrow v_i - A_{i-1}^{+} H(v_i)$ % *corrector step*
choose a new steplength $h > 0$
END FOR

Let us first state a convergence result, see ALLGOWER and GEORG [1990, (5.2.1)], which ensures that the above algorithm safely follows the solution curve under reasonable assumptions.

THEOREM 5.1 *Let $H : \mathbb{R}^{N+1} \to \mathbb{R}^N$ be a smooth map having zero as a regular value and let $H(u_0) = 0$. Denote by $c_h(s)$ the polygonal path, starting at u_0, going through all points u_i generated by Algorithm 5.1 with fixed steplength $h > 0$. Denote by $c(s)$ the corresponding curve in $H^{-1}(0)$ given by the initial value problem (1.8). For definiteness, we assume that $c_h(0) = c(0) = u_0$, and that both curves are parametrized with respect to arclength. If the estimate $||A_i - H'(u_i)|| = \mathcal{O}(h)$ holds uniformly for the approximation in the loop of the algorithm, then, for a given maximal arclength s_0, the following quadratic bounds hold uniformly for $0 \le s \le s_0$:*

$$||H\big(c_h(s)\big)|| \le \mathcal{O}(h^2), \quad ||c_h(s) - c(s)|| \le \mathcal{O}(h^2).$$

PROOF Let us sketch the main arguments of the proof, see also ALLGOWER and GEORG [1990, Theorem 5.2.1]. Let U be a compact neighborhood of $c([0, s_0])$ which consists only of regular points of H. We define the following constants for U:

$$\alpha := \max\{||H'(v)|| : v \in U\}; \tag{5.3}$$
$$\beta := \max\{||H'(v)^{+}|| : v \in U\}; \tag{5.4}$$
$$\gamma := \max\{||H''(v)|| : v \in U\}. \tag{5.5}$$

From the estimates below it is evident that the Algorithm 5.1 generates only predictor and corrector points in U for sufficiently small steplength h and so long as the maximal arclength s_0 is not exceeded. To make the proof more brief, we use the Landau symbol $\mathcal{O}$ when we actually describe

global constants in terms of α, β, γ. The readers may convince themselves that in fact all asymptotic estimates are uniform with respect to U.

Let us first show

$$H(u_i) = \mathcal{O}(h^2). \tag{5.6}$$

We proceed by induction. Let us assume that the estimate (5.6) is true for a current corrector point u_i. For the next predictor step

$$v_{i+1} = u_i + ht(A_i)$$

we obtain

$$H(v_{i+1}) = H(u_i) + hH'(u_i)t(A_i) + \mathcal{O}(h^2),$$

and since $A_i t(A_i) = 0$, this implies

$$H(v_{i+1}) = H(u_i) + h\underbrace{(H'(u_i) - A_i)}_{\mathcal{O}(h)} t(A_i) + \mathcal{O}(h^2) = \mathcal{O}(h^2).$$

For the next corrector step

$$u_{i+1} = v_{i+1} - A_i^+ H(v_{i+1})$$

we obtain

$$H(u_{i+1}) = H(v_{i+1}) - H'(v_{i+1})A_i^+ H(v_{i+1}) + \underbrace{\mathcal{O}(||A_i^+ H(v_{i+1})||^2)}_{\mathcal{O}(h^4)}. \tag{5.7}$$

Next we observe that

$$A_i - H'(v_{i+1}) = [A_i - H'(u_i)] + [H'(u_i) - H'(v_{i+1})]$$

and hence

$$A_i = H'(v_{i+1}) + \mathcal{O}(h).$$

Since the map $A \mapsto A^+$ is smooth, it follows that

$$A_i^+ = H'(v_{i+1})^+ + \mathcal{O}(h).$$

Substituting this into (5.7) yields

$$\begin{aligned} H(u_{i+1}) &= \underbrace{H(v_{i+1}) - H'(v_{i+1})H'(v_{i+1})^+ H(v_{i+1})}_{=0} + \mathcal{O}(h^3) + \mathcal{O}(h^4) \\ &= \mathcal{O}(h^3). \end{aligned}$$

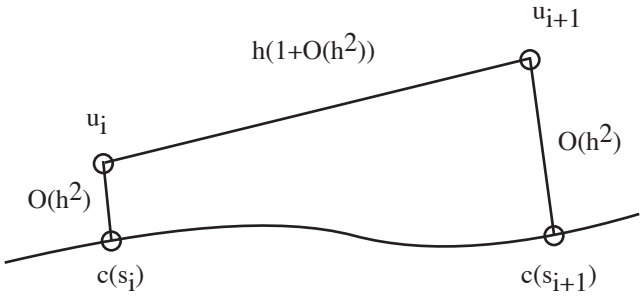

Figure 5.1. Diagram for the Estimates

The additional order (as compared to (5.6)) serves to keep the above mentioned constants uniform, and assertion (5.6) is shown.

Let us next show that

$$||u_i - u_{i+1}|| = h(1 + \mathcal{O}(h^2)). \tag{5.8}$$

In fact, $u_i - v_{i+1} = ht(A_i)$ and $v_{i+1} - u_{i+1} = -A_i^+ H(v_{i+1})$ are orthogonal to each other, and hence, combining the two differences yields

$$||u_i - u_{i+1}||^2 = h^2 + \mathcal{O}(h^4) = h^2(1 + \mathcal{O}(h^2))$$

and thus (5.8) is obtained by taking the square root.

For $\tau \in [0, 1]$, an easy consequence of Taylor's formula is

$$H(\tau u_i + (1 - \tau)u_{i+1}) = \tau H(u_i) + (1 - \tau)H(u_{i+1}) + \mathcal{O}(||u_{i+1} - u_i||^2).$$

Together with (5.6) and (5.8), this immediately implies the first assertion of the theorem.

For sufficiently small step sizes h, it is clear that for each u_i there is a unique s_i such that $||u_i - c(s_i)|| = \min_s ||u_i - c(s)||$. By the choice of the s_i, it follows that the orthogonality

$$u_i - c(s_i) \perp \dot{c}(s_i) \tag{5.9}$$

holds, see Figure 5.1. Furthermore, it is not difficult to see that

$$c(s_i) - u_i \ \rightarrow \ 0 \quad \text{as} \quad h \rightarrow 0. \tag{5.10}$$

Let us next show that in fact

$$||c(s_i) - u_i|| = \mathcal{O}(h^2). \tag{5.11}$$

From Taylor's formula it follows that

$$H(u_i) = H(c(s_i)) + H'(c(s_i))(u_i - c(s_i)) + \mathcal{O}(||u_i - c(s_i)||^2).$$

Now $H(c(s_i)) = 0$, and if we multiply through with $H'(c(s_i))^+$, then (5.9) and (3.7) imply

$$H'(c(s_i))^+ H(u_i) = (u_i - c(s_i)) + \mathcal{O}(||u_i - c(s_i)||^2).$$

Now assertion (5.11) follows from (5.6) and (5.10).

Let us define $\Delta s_i := s_{i+1} - s_i$. Differentiating $||\dot{c}(s)||^2 \equiv 1$ implies $\dot{c}(s) \perp \ddot{c}(s)$, and Taylor's formula yields

$$\begin{aligned} ||c(s_{i+1}) - c(s_i)||^2 &= \left\| \int_0^1 \dot{c}(s_i + \xi \Delta s_i) \Delta s_i \, d\xi \right\|^2 \\ &= ||\dot{c}(s_i)\Delta s_i + \frac{1}{2}\ddot{c}(s_i)(\Delta s_i)^2 + \mathcal{O}\big((\Delta s_i)^3\big)||^2 \\ &= (\Delta s_i)^2 + \mathcal{O}\big((\Delta s_i)^4\big) \end{aligned}$$

and consequently

$$||c(s_{i+1}) - c(s_i)|| = \Delta s_i \Big(1 + \mathcal{O}\big((\Delta s_i)^2\big)\Big). \tag{5.12}$$

From (5.8), (5.11) and (5.12) it is straightforward to obtain

$$\Delta s_i [1 + \mathcal{O}\big((\Delta s_i)^2\big)] = h[1 + \mathcal{O}(h)]$$

and consequently

$$\begin{aligned} h &= \Delta s_i (1 + \mathcal{O}\big((\Delta s_i)\big), \\ \Delta s_i &= h(1 + \mathcal{O}(h)). \end{aligned}$$

Hence the terms $\mathcal{O}(h)$ and $\mathcal{O}(\Delta s_i)$ can be used interchangeably.

From the orthogonality relation $\big(u_i - c(s_i)\big) \perp \dot{c}(s_i)$ and Taylor's formula we obtain

$$\begin{aligned} \big(u_i - c(s_i)\big)^* \big(c(s_{i+1}) - c(s_i)\big) &= \big(u_i - c(s_i)\big)^* \Big(\dot{c}(s_i)\Delta s_i + \mathcal{O}\big((\Delta s_i)^2\big)\Big) \\ &= \mathcal{O}\big(h^2\big)\mathcal{O}\big((\Delta s_i)^2\big) \\ &= \mathcal{O}(h^4). \end{aligned}$$

Therefore

$$\begin{aligned} ||u_{i+1} - u_i||^2 &= ||\big(u_{i+1} - c(s_{i+1})\big) + \big(c(s_{i+1}) - c(s_i)\big) + \big(c(s_i) - u_i\big)||^2 \\ &= ||c(s_{i+1}) - c(s_i)||^2 + \mathcal{O}(h^4). \end{aligned}$$

Taking square roots and using (5.12) yields

$$\begin{aligned} ||u_{i+1} - u_i|| &= ||c(s_{i+1}) - c(s_i)|| + \frac{1}{2}\frac{\mathcal{O}(h^4)}{||c(s_{i+1}) - c(s_i)||} \\ &= ||c(s_{i+1}) - c(s_i)|| + \mathcal{O}(h^3). \end{aligned} \tag{5.13}$$

If we sum up over the terms $||u_{i+1} - u_i||$ (the nodes of c_h), $||c(s_{i+1} - c(s_i)||$ and the Δs_i, respectively, then the summation drops the order of h by one degree, and hence the corresponding summations (performed to a certain arclength) differ by $\mathcal{O}(h^2)$:

$$\sum ||u_{i+1} - u_i|| = \sum ||c(s_{i+1}) - c(s_i)|| + \mathcal{O}(h^2)$$

by (5.13) and

$$\sum ||c(s_{i+1}) - c(s_i)|| = \sum \Delta s_i + \mathcal{O}(h^2)$$

by (5.12). This implies

$$||c_h(s_i) - u_i|| = \mathcal{O}(h^2). \tag{5.14}$$

Now we are in a position to show the second assertion of the theorem: Let $s_i \le s \le s_{i+1}$ and $s = \tau s_i + (1-\tau)s_{i+1}$ for a suitable $\tau \in [0,1]$. Taylor's formula and the estimates (5.14) and (5.11) imply:

$$\begin{aligned} & ||c(s) - c_h(s)|| \\ = \;& ||c(\tau s_i + (1-\tau)s_{i+1}) - c_h(\tau s_i + (1-\tau)s_{i+1})|| \\ \le \;& ||[\tau c(s_i) + (1-\tau)c(s_{i+1})] - c_h(\tau s_i + (1-\tau)s_{i+1})|| + \mathcal{O}(h^2) \\ \le \;& ||[\tau c(s_i) + (1-\tau)c(s_{i+1})] - [\tau u_i + (1-\tau)u_{i+1}]|| + \mathcal{O}(h^2) \\ \le \;& \tau||c(s_i) - u_i|| + (1-\tau)||c(s_{i+1}) - u_{i+1}|| + \mathcal{O}(h^2) \\ \le \;& \mathcal{O}(h^2). \end{aligned}$$

□

Similar methods and results hold for pseudo arclength continuation or other forms of continuation methods.

Some major points which remain to be clarified are:

— How do we formulate efficient steplength strategies? See Chapter III for details.

— How do we efficiently handle the numerical linear algebra involved in the calculation of the quantities $t(A)$ and $A^+H(v)$ which are essentially involved in the predictor and corrctor steps respectively? See Chapter IV for details.

Chapter III

Predictors

The convergence considerations of Theorem 5.1 were carried out under the assumption that the steplength of the Algorithm 5.1 was uniformly constant throughout. This assumption is also typical for complexity studies, see Chapter VII. Such an approach is of course not efficient for any practical implementation. An efficient algorithm needs to incorporate an automatic strategy for controlling the steplength. In this respect the predictor corrector methods are similar to the methods for numerically integrating initial value problems in ordinary differential equations. To some extent of course, the steplength strategy depends upon the accuracy with which it is desired to numerically trace a solution curve. Path following methods usually split into two categories:

— Either it is wished to approximate the solution curve with some given accuracy, e.g., for plotting purposes, or if one wants to calculate a contour integral along the curve,

— or the objective is just to safely follow the curve as fast as possible, until a certain point is reached, e.g., a zero point or critical point with respect to some additional functional defined on the curve.

We briefly sketch some ideas which are used to adjust the steplength.

6 Steplength Control Via Error Models.

One method, due to Den Heijer and Rheinboldt [1981], is based upon an error model for the corrector iteration. For Gauss-Newton corrector steps, such error models can be obtained by analyzing the Newton-Kantorovich

theory. The steplength is controlled by the number of steps which are taken in the corrector iteration until a given stopping criterion is fulfilled.

We sketch a somewhat modified and simplified version of this steplength strategy. Let us assume that u is a point on the solution curve, and consider, for simplicity, an Euler predictor $v_0(h) = u + ht(H'(u))$. Let $v_0(h)$, $v_1(h)$, ..., $v_k(h)$ be an iterative corrector process for approximating a nearby point on the curve. Suppose a certain stopping criterion is met after k iterations. The exact nature of the criterion is not important in this context. We assume theoretical convergence to $v_\infty(h)$.

It is assumed that there exists a constant $\gamma > 0$ (which is independent of h) such that the *modified error*

$$\varepsilon_i(h) := \gamma||v_\infty(h) - v_i(h)||$$

satisfies inequalities of the following type

$$\varepsilon_{i+1}(h) \le \psi\big(\varepsilon_i(h)\big),$$

where $\psi : \mathbb{R} \to \mathbb{R}$ is a known monotone function such that $\psi(0) = 0$. For example, if Newton's method is employed, Den Heijer and Rheinboldt suggest two models:

$$\psi(\varepsilon) = \frac{\varepsilon^2}{3 - 2\varepsilon}, \quad 0 \le \varepsilon \le 1, \tag{6.1}$$

$$\psi(\varepsilon) = \frac{\varepsilon + \sqrt{10 - \varepsilon^2}}{5 - \varepsilon^2}\varepsilon^2, \quad 0 \le \varepsilon \le 1. \tag{6.2}$$

We may evaluate a posteriori the quotient

$$\omega(h) := \frac{||v_k(h) - v_{k-1}(h)||}{||v_k(h) - v_0(h)||} \approx \frac{||v_\infty(h) - v_{k-1}(h)||}{||v_\infty(h) - v_0(h)||} = \frac{\varepsilon_{k-1}(h)}{\varepsilon_0(h)}.$$

Using the estimate $\varepsilon_{k-1}(h) \le \psi^{k-1}\big(\varepsilon_0(h)\big)$, we obtain

$$\omega(h) \le \frac{\psi^{k-1}\big(\varepsilon_0(h)\big)}{\varepsilon_0(h)}.$$

This motivates taking the solution ε of the equation

$$\omega(h) = \frac{\psi^{k-1}(\varepsilon)}{\varepsilon}$$

as an estimate for $\varepsilon_0(h)$.

We now try to choose the steplength $\tilde{h}$ so that the corrector process satisfies the stopping criterion after a chosen number (say $\tilde{k}$) of iterations.

Such a steplength leads to the modified error $\varepsilon_0(\tilde{h})$. Hence, we want the modified error $\varepsilon_{\tilde{k}}(\tilde{h})$ after $\tilde{k}$ iterations to be so small that the stopping criterion is satisfied. Using the inequality $\varepsilon_{\tilde{k}}(\tilde{h}) \leq \psi^{\tilde{k}}\big(\varepsilon_0(\tilde{h})\big)$, we accept the solution ε of the equation

$$\psi^{\tilde{k}}(\varepsilon) = \psi^k\big(\varepsilon_0(h)\big)$$

as an estimate for $\varepsilon_0(\tilde{h})$. Now we use the asymptotic expansion

$$||v_\infty(h) - v_0(h)|| = Ch^2 + \mathcal{O}(h^3)$$

to obtain the approximation

$$\left(\frac{h}{\tilde{h}}\right)^2 \approx \frac{\varepsilon_0(h)}{\varepsilon_0(\tilde{h})},$$

which can be used to determine $\tilde{h}$. This steplength $\tilde{h}$ will now be used in the next predictor step. It is usually safeguarded by some additional considerations such as limiting the steplength to some interval $h_{\min} \leq \tilde{h} \leq h_{\max}$, or limiting the factor $.5 \leq h/\tilde{h} \leq 2$, etc.

7 Steplength Control Via Asymptotic Expansion.

Another method, based upon asymptotic estimates, is due to Georg [1983]. The basic idea in this approach is to observe the performance of the corrector procedure in dependence on a given steplength h. This dependence is typically expressed by some asymptotic expansions in h. The aim is then to adapt the steplength in such a way as to strive toward a desired performance.

Let us discuss this approach by means of a simple, but typical example. Suppose that a point u on the solution curve has been approximated. Suppose further that a steplength $h > 0$ and a predictor point

$$v_0(h) = u + ht(H'(u)) \tag{7.1}$$

are given. Then a Gauss-Newton type iterative corrector process is performed:

$$v_{i+1}(h) = v_i(h) - H'(v_0(h))^+ H(v_i(h)) \tag{7.2}$$

The sequence $\{v_i(h)\}$ converges to the next point $v_\infty(h)$ on the curve.

The steplength strategy is motivated by the following question: Given the observed performance of the corrector process for the steplength h,

which steplength $\tilde{h}$ would have been "best" for obtaining $v_\infty(\tilde{h})$ from u? This "ideal" steplength $\tilde{h}$ is determined via asymptotic estimates, and it is then taken as the steplength for the next predictor step. This strategy depends primarily upon two factors: the particular predictor corrector method being utilized, and the criteria used in deciding what performance is considered "best".

Let us illustrate this technique with the case of the *contraction rate* of the corrector process (7.2):

$$\kappa(u,h) := \frac{||H'\big(v_0(h)\big)^+ H\big(v_1(h)\big)||}{||H'\big(v_0(h)\big)^+ H\big(v_0(h)\big)||}.$$

Since Newton's method is locally quadratically convergent, it is plain that $\kappa(u,h)$ will decrease (and hence Newton's method will become faster) as h decreases. The following theorem characterizes the asymptotic behavior of $\kappa(u,h)$ with respect to h, see also (6.1.2) in ALLGOWER and GEORG [1990].

THEOREM 7.1 *Suppose that*

$$H''(u)\big[t\big(H'(u)\big), t\big(H'(u)\big)\big] \neq 0.$$

Then

$$\kappa(u,h) = \kappa_2(u)h^2 + \mathcal{O}(h^3)$$

for some constant $\kappa_2(u) \geq 0$ which is independent of h and depends smoothly on u.

Let us first point out that the above assumption implies that the solution curve has non-zero curvature at u. In fact, if $c(s)$ denotes the solution curve (parametrized with respect to arclength), then, by differentiating $H(c(s)) = 0$ twice with respect to s, we obtain

$$H'(c(s))\ddot{c}(s) + H''(c(s))[\dot{c}(s), \dot{c}(s)] = 0.$$

Note that $\ddot{c}(s) \perp \dot{c}(s)$, and hence from (3.4) we obtain

$$\ddot{c}(s) = -H'(c(s))^+ H''(c(s))[\dot{c}(s), \dot{c}(s)].$$

PROOF For convenience we hereafter abbreviate $H''(u)[b]^2 := H''(u)[b,b]$. Expanding $H(u + ht(H'(u))$ about u we obtain

$$\begin{aligned} H\big(v_0(h)\big) &= C_1(u)h^2 + \mathcal{O}(h^3), \\ \text{where } C_1(u) &:= \frac{1}{2}H''(u)\Big[t\big(H'(u)\big)\Big]^2, \end{aligned}$$

since $H(u) = 0$ and $H'(u)t\big(H'(u)\big) = 0$. Because the maps $u \mapsto H'(u)$, $u \mapsto H'(u)^+$ and $u \mapsto H''(u)$ are smooth, we have

$$\begin{aligned} H'\big(v_0(h)\big) &= H'(u) + \mathcal{O}(h) \\ H'\big(v_0(h)\big)^+ &= H'(u)^+ + \mathcal{O}(h) \\ H''\big(v_0(h)\big) &= H''(u) + \mathcal{O}(h) \end{aligned}$$

and hence

$$\begin{aligned} H'\big(v_0(h)\big)^+ H\big(v_0(h)\big) &= C_2(u)h^2 + \mathcal{O}(h^3), \\ \text{where } C_2(u) &:= H'(u)^+ C_1(u). \end{aligned}$$

Now we expand $H\Big(v_0(h) - H'(v_0(h))^+ H(v_0(h))\Big)$ about $v_0(h)$:

$$\begin{aligned} H\big(v_1(h)\big) &= \underbrace{H\big(v_0(h)\big) - H'(v_0(h))H'(v_0(h))^+ H\big(v_0(h)\big)}_{=0} \\ &\quad + \frac{1}{2} H''\big(v_0(h)\big)\Big[H'(v_0(h))^+ H\big(v_0(h)\big)\Big]^2 + \mathcal{O}(h^5) \\ &= C_3(u)h^4 + \mathcal{O}(h^5), \\ \text{where } C_3(u) &:= \frac{1}{2} H''(u)\big[C_2(u)\big]^2. \end{aligned}$$

Furthermore we have

$$\begin{aligned} H'\big(v_0(h)\big)^+ H\big(v_1(h)\big) &= C_4(u)h^4 + \mathcal{O}(h^5), \\ \text{where } C_4(u) &:= H'(u)^+ C_3(u). \end{aligned}$$

Finally we obtain

$$\begin{aligned} \kappa(u,h) &= \kappa_2(u)h^2 + \mathcal{O}(h^3), \\ \text{where } \kappa_2(u) &:= \frac{||C_4(u)||}{||C_2(u)||}. \end{aligned} \tag{7.3}$$

Note that the assumption implies that $C_1(u) \neq 0$ and hence $C_2(u) \neq 0$. The smoothness of $\kappa_2(u)$ follows from the smoothness of the vectors $C_2(u)$ and $C_4(u)$. □

In view of asymptotic relation (7.3), the steplength modification $h \to \tilde{h}$ is now easy to explain. Assume that an Euler predictor (7.1) and a Gauss-Newton iteration (7.2) has been performed with steplength h. Then $H'\big(v_0(h)\big)^+ H\big(v_0(h)\big)$ and $H'\big(v_0(h)\big)^+ H\big(v_1(h)\big)$ will have been calculated

and thus $\kappa(u,h)$ can be obtained without any significant additional cost. Now an a posteriori estimate

$$\kappa_2(u) = \frac{\kappa(u,h)}{h^2} + \mathcal{O}(h)$$

is available.

In order to have a robust and efficient method we want to continually adapt the steplength h so that a nominal prescribed contraction rate $\tilde{\kappa}$ is maintained. The choice of $\tilde{\kappa}$ will generally depend upon the nature of the problem at hand, and on the desired security with which we want to traverse the curve. That is, the smaller $\tilde{\kappa}$ is chosen, the greater will be the security with which the method will follow the curve. When using the term *securely* or *safely* following the curve we mean that a safeguard prevents the method from jumping to a different part of the curve (at a significantly different arclength value) or to a different connected component of $H^{-1}(0)$. Depending on the structure of the solution manifold $H^{-1}(0)$, this may be an important issue.

Once $\tilde{\kappa}$ has been chosen, we will consider a steplength $\tilde{h}$ to be appropriate if $\kappa(u,\tilde{h}) \approx \tilde{\kappa}$. By using the above equation and neglecting higher order terms we obtain the formula

$$\tilde{h} = h\,\sqrt{\frac{\tilde{\kappa}}{\kappa(u,h)}}$$

as the steplength for the next predictor step.

In a similar way, other quantities which are important for the performance of the path following method can be taken into account, e.g., the angle of two successive predictor directions, the size of the first Gauss-Newton step (which gives an approximation of the distance of the predictor point to the curve), or the function value $H(v_0(h))$. All these quantities admit asymptotic expansions in h (with varying order). For example, Algorithm (6.1.10) and Program 1 in Allgower and Georg [1990] incorporate such features in the steplength strategy. Let us summarize the above discussion in a basic Euler-Newton algorithm where the steplength is only monitored via a nominal contraction rate $\tilde{\kappa}$.

Algorithm 7.1 **(Basic Steplength Euler-Newton Method)**

INPUT $u_0 \in \mathbb{R}^{N+1}$ such that $H(u_0) \approx 0$ (initial point)
$1 > \tilde{\kappa} > 0$ nominal contraction rate
$h > 0$ initial steplength
OUTPUT u_i , $i = 0, 1, 2, \ldots$, approximating the solution curve

FOR $i = 1, 2, \ldots$
  approximate $A_{i-1} \approx H'(u_{i-1})$
  % *predictor step*
  $v_{i,0} \leftarrow u_{i-1} + ht(A_{i-1})$
  approximate $B_i \approx H'(v_{i,0})$
  % *corrector iteration*
  FOR $k = 1, 2, \ldots$ until convergence
    $v_{i,k} \leftarrow v_{i,k-1} - B_i^+ H(v_{i,k-1})$
  END FOR
  $u_i \leftarrow v_{i,k}$
  $\kappa \leftarrow \dfrac{||B_i^+ H(v_{i,1})||}{||B_i^+ H(v_{i,0})||}$
  $h \leftarrow h\ \sqrt{\tilde{\kappa}/\kappa}$
END FOR

KEARFOTT [1989, 1990] proposes interval arithmetic techniques to determine a first order predictor which stresses secure path following.

8 Special Points on the Curve

One of the main purposes of numerical continuation methods concerns the accurate determination of certain points on the solution curve $c(s)$ which are of special interest. The following are some examples.

1. In the applications dealing with homotopy methods, the equation $H(x, \lambda) = 0$ for $x \in \mathbb{R}^N$ and $\lambda \in \mathbb{R}$ generally has a known starting point (x_0, λ_0). The homotopy path $c(s)$ passes through this point, and we seek a point $(\bar{x}, \bar{\lambda})$ on $c(s)$ such that $H(\bar{x}, \bar{\lambda}) = 0$ for a certain value $\bar{\lambda}$ of the homotopy parameter λ. Examples of applications of homotopy methods are given in Chapter V.

2. Fold points in $H^{-1}(0)$ may be of interest when the equation represents a branch of solutions for a nonlinear eigenvalue problem involving the eigenvalue parameter λ, see also Section 24. Such points are characterized by the fact that λ has a local extremum on $H^{-1}(0)$. In physics and engineering applications, a fold point can signify a change in the stability of the solutions. A vast literature exists for calculating such points, the following papers are a sample: BOLSTAD and KELLER [1986], CHAN [1984b], FINK and RHEINBOLDT [1986, 1987], MELHEM and RHEINBOLDT [1982], PÖNISCH and SCHWETLICK [1981, 1982], SCHWETLICK [1984b, 1984a], USHIDA and CHUA [1984].

3. Simple bifurcation points will be discussed in Section 24. There we show how to detect the presence of such points along the curve c. It is of interest to accurately approximate a bifurcation point. They are of great interest since they represent points at which the stability of the solutions changes.

To unify our discussion, let $f : \operatorname{range} c \to \mathbb{R}$ be a smooth functional. There are two general types of special points on the curve c which we shall consider:

Zero Points. In this case we seek points $c(s)$ such that $f\big(c(s)\big) = 0$. The homotopy method is such a case if we set $f(x, \lambda) := \lambda - \bar{\lambda}$. Simple bifurcation points are another such case if we set, e.g.,

$$f\big(c(s)\big) := \det \begin{pmatrix} H'\big(c(s)\big) \\ \dot{c}(s)^* \end{pmatrix}.$$

Extremal Points. In this case we seek extremal points (usually maxima or minima) of $f\big(c(s)\big)$. Fold points are such a case if we set $f(x, \lambda) := \lambda$. Certain regularization methods may also be formulated as determining a fold point on an implicitly defined curve.

We now treat these two general cases in greater detail.

8.1 Calculating Zero Points

Let $H : \mathbb{R}^{N+1} \to \mathbb{R}^N$ be a smooth map, let $c(s)$ be a smooth solution curve parametrized with respect to arclength (for the sake of convenience), and let $f : \operatorname{range} c \to \mathbb{R}$ be a smooth functional. Suppose that some point $c(s_n)$ has been found which is an approximate zero point of f. For example, it would be reasonable to take $c(s_n)$ as an approximate zero point if a predictor corrector method produced two successive points $c(s_{n-1})$ and $c(s_n)$ such that $f\big(c(s_{n-1})\big) f\big(c(s_n)\big) < 0$. Then it is reasonable to replace the usual steplength adaptation used to traverse the curve c by a *Newton steplength adaptation* which is motivated by the following one-dimensional Newton method for solving the equation $f\big(c(s)\big) = 0$:

$$s_{n+1} = s_n - \frac{f\big(c(s_n)\big)}{f'\big(c(s_n)\big)\dot{c}(s_n)}. \tag{8.1}$$

This suggests that we can take the new steplength

$$h := -\frac{f\big(c(s_n)\big)}{f'\big(c(s_n)\big)\dot{c}(s_n)} \tag{8.2}$$

at $u := c(s_n)$ in order to obtain a predictor point $v = u + ht\big(H'(u)\big)$, which should lead to a better approximation of a zero point of f on c.

The following algorithm illustrates for a simple Euler-Newton method how a standard steplength adaptation can be switched to the above Newton steplength adaptation in order to approximate a zero point of f on c while traversing c.

ALGORITHM 8.1 **(Newton Steplength Adaptation)**

INPUT $u \in \mathbb{R}^{N+1}$ such that $H(u) \approx 0$ (initial point)
$h > h_{\min} > 0$ (initial, minimal steplength)
OUTPUT u such that $H(u) = 0$ and $f(u) = 0$

$\nu \leftarrow$ FALSE % *switch*
REPEAT $v \leftarrow u + ht\big(H'(u)\big)$ % *predictor step*
REPEAT $v \leftarrow v - H'(v)^+ H(v)$ % *corrector loop*
UNTIL convergence
IF $f(u)f(v) \le 0$; $\nu \leftarrow$ TRUE; END IF
IF $\nu =$ TRUE
$$h \leftarrow -\frac{f(v)}{f'(v)\, t\big(H'(v)\big)} \quad \% \textit{ Newton steplength}$$
ELSE choose a new $h > 0$ % *standard steplength*
END IF
$u \leftarrow v$ % *new point along the curve*
UNTIL $|h| < h_{\min}$

A sufficient condition for a sequence of points u produced by Algorithm 8.1 to converge to a solution $\bar{u}$ of

$$H(u) = 0, \quad f(u) = 0,$$

is that the predictor point is sufficiently near $\bar{u}$ and that

$$\det \begin{pmatrix} H'(\bar{u}) \\ f'(\bar{u}) \end{pmatrix} \neq 0\,. \tag{8.3}$$

Under these assumptions quadratic convergence can be shown.

Algorithm 8.1 requires the quantity

$$\frac{d}{ds} f\big(c(s)\big) = f'\big(c(s)\big)\dot{c}(s),$$

and this may be inconvenient to obtain. As an example, we will see in Section 24 that bifurcation points $c(\bar{s})$ are points where

$$f\big(c(s)\big) = \det \begin{pmatrix} H'\big(c(s)\big) \\ \dot{c}(s)^* \end{pmatrix} = 0 \tag{8.4}$$

holds. In this case, furnishing $\frac{d}{ds} f\big(c(s)\big)$ would be undesirable, since it would require, at least formally, the calculation of H''. Thus it is reasonable to formulate the secant analogue of (8.1) which leads to the following Newton steplength adaptation in Algorithm 8.1:

$$h := -\frac{f(v)}{f(v) - f(u)}\ h\,.$$

This reduces the above mentioned quadratic convergence to superlinear convergence.

For the case of calculating a simple bifurcation point, care should be taken since the augmented Jacobian

$$\begin{pmatrix} H'(u) \\ t\big(H'(u)\big)^* \end{pmatrix}$$

is ill-conditioned near the bifurcation point, and hence the evaluation of (8.4) encounters instabilities. But the above mentioned superlinear convergence of the secant method generally overcomes this difficulty since the instability generally only manifests itself at a predictor point which can already be accepted as an adequate approximation of the bifurcation point.

Obviously, if one zero point $\bar{u}$ of the functional f on the curve c has been approximated, the predictor corrector method can be restarted in order to seek additional zero points.

8.2 Calculating Extremal Points

The aim now is to give some specific details for calculating an extremal point on the solution curve c for a smooth functional $f : \operatorname{range} c \to \mathbb{R}$. Clearly, a necessary condition which must hold at a local extremum $c(\bar{s})$ of f is that the equation

$$f'\big(c(s)\big)\dot{c}(s) = 0 \tag{8.5}$$

holds. Following the same motivation as in the previous section, we can formulate the analogous switchover to a Newton steplength adaptation:

$$h := -\frac{f'(v)\dot{v}}{f'(v)\ddot{v} + f''(v)[\dot{v}, \dot{v}]}\ , \tag{8.6}$$

where $v = c(s)$, $\dot{v} = \dot{c}(s) = t\big(H'(v)\big)$, $\ddot{v} = \ddot{c}(s)$, and $v = c(s)$ is the point currently approximated on c in Algorithm 8.1.

The quantities v and $\dot{v}$ are readily obtained. Let us discuss how to calculate an approximation of $\ddot{v}$. By differentiating the equation $H\big(c(s)\big) = 0$ we obtain $H'(v)\dot{v} = 0$ and

$$H''(v)[\dot{v}, \dot{v}] + H'(v)\ddot{v} = 0. \tag{8.7}$$

Differentiating $||\dot{c}(s)||^2 = 1$ yields the orthogonality

$$\dot{v}^* \perp \ddot{v},$$

and we obtain from (8.7) and (3.4):

$$\ddot{v} = -H'(v)^+ H''(v)[\dot{v}, \dot{v}]. \tag{8.8}$$

To approximate $H''(v)[\dot{v}, \dot{v}]$ we can use the centered difference formula

$$\frac{H(v + \varepsilon\dot{v}) - 2H(v) + H(v - \varepsilon\dot{v})}{\varepsilon^2} = H''(v)[\dot{v}, \dot{v}] + \mathcal{O}(\varepsilon^2), \tag{8.9}$$

see Section 16.2 for more details. Now (8.8), (8.9) provide an approximation of $\ddot{v}$ in the Newton steplength adaptation (8.6). If necessary, higher order formulae and, in particular, an extrapolation method may be used to obtain higher precision approximations of $\ddot{v}$ since the truncation error in (8.9) is expandable in powers of ε^2.

The following example illustrates this approach for the case of calculating a fold point with respect to the last co-ordinate, i.e., $f(x, \lambda) = \lambda$ in the case of a nonlinear eigenvalue problem $H(x, \lambda) = 0$. Then $f(v) = v(N + 1)$ (i.e., the last co-ordinate), and (8.6) takes the particular form

$$h := -\frac{\dot{v}[N + 1]}{\ddot{v}[N + 1]} \, .$$

A second possible method for calculating a local extremal point of $f\big(c(s)\big)$ is to use a secant steplength adaptation applied to equation (8.5). Analogously to the discussion in the previous section we obtain the corresponding Newton steplength adaptation

$$h := -\frac{f'(v)\dot{v}}{f'(v)\dot{v} - f'(u)\dot{u}} \; h \, ,$$

where v is the latest point and u the previous point on the solution curve. The advantage of using this in Algorithm 9.2.3 is that the need to calculate $\ddot{v}$ is avoided. Under assumptions analogous to (8.3) superlinear convergence of the sequence u generated by the algorithm to a local extremum of $f\big(c(s)\big)$ can be shown.

An alternative way for choosing the stepsize based on Hermite interpolation of $\dot{\lambda}$ at the latest two points is described in PÖNISCH and SCHWETLICK [1982].

8.3 Calculating Fold and Bifurcation Points

As we already mentioned above, calculating a fold point may be viewed as calculating a special point on the curve which is characterized as an extremal point with respect to a particularly simple functional (e.g., the last co-ordinate or bifurcation parameter). A more efficient approach is to extend the equation $H(u) = 0$ in order to directly characterize such points. A rapidly convergent Newton-like iteration can then be applied to the extended system of equations. A starting value is obtained from the curve following process. We refer to MOORE and SPENCE [1980] as a pioneer of this method, see also PÖNISCH and SCHWETLICK [1981].

The case of bifurcation points is more complex and will be described in more detail in Section 24. In particular, the above mentioned technique of extending the equations $H(u) = 0$ in order to characterize and approximate such points is discussed in Section 24.2.

9 Higher Order Predictors

The steplength strategies we have discussed up to now have been based upon the Euler predictor, which is only of local order two. This is very often satisfactory since it is usually used in conjunction with rapidly converging correctors such as Newton type correctors. However, for large systems, often less rapidly convergent iterative methods such as Krylov subspace methods are used. Hence, one may expect to obtain improved efficiency by using higher order predictors.

Another instance where it is desirable to use higher order predictors occurs when the solution curve needs to be approximated very well at all points, e.g., for plotting purposes or numerical integration, see Section 30 for the latter. Since in such cases a polynomial approximation or something similar needs to be performed anyway, one may as well expend this numerical effort at the outset by performing higher order predictors.

Variable order predictors and corresponding steplength strategies may be generated in a way reminiscent of the ones used in multistep methods for solving initial value problems, see, e.g., SHAMPINE and GORDON [1975]. Such a method was suggested by GEORG [1982, 1983]. LUNDBERG and POORE [1991] have made an implementation using variable order Adams-Bashforth predictors. Their numerical results show that there is often a definite benefit to be derived by using higher order predictors.

Let us sketch a general philosophy for monitoring the order and steplength of higher order predictors. Let u_n be a current point on the solution curve c which can be locally parametrized via the parameter s (not neces-

sarily arclength). For simplicity, let us assume that $c(0) = u_n$. Consider a polynomial predictor of the form

$$c(h) \approx p_k(h) = u + \sum_{i=1}^{k} c_i h^i, \tag{9.1}$$

$$c_i \approx \frac{c^{(i)}(0)}{i!}, \tag{9.2}$$

which represents an approximation of the Taylor formula. We see essentially two different ways for obtaining the coefficients c_i: (1) by divided differences or polynomial interpolation making use of previously calculated points on the curve, (2) by successive numerical differentiation at u. The former is less expensive to calculate, but the latter is likely to be more accurate and stable. We elaborate on this in the next sections.

We sketch one possible way of determining the next steplength and the next order in the predictor. Let $\varepsilon > 0$ be a given tolerance. The term $||c_k||h^k$ can be viewed as a rough estimate for the truncation error of the predictor $p_{k-1}(h)$. Hence, by solving $||c_k||h^k = \varepsilon$ for h, we estimate

$$h_k := \left(\frac{\varepsilon}{||c_k||}\right)^{\frac{1}{k}} \tag{9.3}$$

as the steplength for the predictor p_{k-1} in order to remain within the given tolerance. Due to instabilities of various kinds, we anticipate that eventually

$$h_2 < h_3 \ldots < h_q \geq h_{q+1} \tag{9.4}$$

will hold for some q. Hence, the predictor p_{q-1} with steplength h_q is our next choice.

This idea can be implemented and modified in various ways, and needs some stabilizing safeguards, such as setting a maximum increase in steplength and in the order and a maximal steplength and order. The strategy to be developed depends on the objective of the application at hand.

10 Interpolation Predictors

Let us consider an example of a polynomial interpolation predictor. Assume that the points u_0, u_1, …, u_n along the solution curve c have already been generated. In certain versions of the continuation method, also the corresponding tangents $t_0 := t\big(H'(u_0)\big)$, …, $t_n := t\big(H'(u_n)\big)$ are available. The idea is to use an interpolating polynomial p_q of degree q (with coefficients in $\mathbb{R}^{N+1}$) satisfying $p_q(0) = u_n$ as a predicting polynomial. We

need to express the interpolating polynomial in terms of a suitable parameter ξ. Naturally, the arclength parameter s which we always consider for theoretical discussions would be ideal to use. This is done by LUNDBERG and POORE [1991]. However, for purposes of exposition, we shall avoid the additional complexity of obtaining precise numerical approximations of the arclength s_i such that $c(s_i) = u_i$. We therefore propose to use a local parametrization ξ induced by the current approximate tangent $t \approx t\big(H'(u_n)\big)$, which does not need to be very accurate. We assume however, that the normalization $||t|| = 1$ holds. This local parametrization $c(\xi)$ is defined as the locally unique solution of the system

$$\begin{aligned} &H(u) = 0, \\ &t^*\big(u_n + \xi t - u\big) = 0, \end{aligned} \tag{10.1}$$

for ξ in some open interval containing zero. It follows immediately that

$$c(\xi_i) = u_i \quad \text{where} \quad \xi_i = t^*\big(u_i - u_n\big).$$

Differentiating (10.1) with respect to ξ yields

$$\frac{dc(\xi)}{d\xi} = \frac{\dot{c}(s)}{t^*\dot{c}(s)}.$$

If the tangents t_i at the points u_i are available for use, we may form a Hermite interpolating polynomial p_q. Otherwise, a standard interpolating polynomial using Newton's formula is generated.

Hence, the interpolation polynomial is of the form (Newton)

$$\begin{aligned} p_q(h) = {} & c[\xi_n] + c[\xi_n, \xi_{n-1}](h - \xi_n) + c[\xi_n, \xi_{n-1}, \xi_{n-2}](h - \xi_n)(h - \xi_{n-1}) \\ & + \ldots + c[\xi_n, \ldots, \xi_{n-q}](h - \xi_n) \cdots (h - \xi_{n-q+1}), \end{aligned}$$

or (Hermite)

$$\begin{aligned} p_{2i+2}(h) = {} & c[\xi_n] + c[\xi_n, \xi_n](h - \xi_n) + c[\xi_n, \xi_n, \xi_{n-1}](h - \xi_n)^2 + \ldots \\ & + c[\xi_n, \xi_n, \ldots, \xi_{n-i}, \xi_{n-i}, \xi_{n-i-1}](h - \xi_n)^2 \cdots (h - \xi_{n-i})^2, \\ p_{2i+1}(h) = {} & c[\xi_n] + c[\xi_n, \xi_n](h - \xi_n) + c[\xi_n, \xi_n, \xi_{n-1}](h - \xi_n)^2 + \ldots \\ & + c[\xi_n, \xi_n, \ldots, \xi_{n-i}, \xi_{n-i}](h - \xi_n)^2 \cdots (h - \xi_{n-i-1})^2 (h - \xi_{n-i}), \end{aligned}$$

where the coefficients are obtained via the formulae

$$\begin{aligned} \xi_i &:= t^*(u_i - u_n) \\ c[\xi_i] &:= u_i \\ c[\xi_i, \xi_i] &= \frac{dc(\xi_i)}{d\xi} := \frac{t_i}{t^* t_i} \\ c[\xi_i, \ldots, \xi_j] &:= \frac{c[\xi_i, \ldots, \xi_{j+1}] - c[\xi_{i-1}, \ldots, \xi_j]}{\xi_i - \xi_j} \quad \text{for} \quad i > j,\ \xi_i \neq \xi_j. \end{aligned}$$

11 Taylor Polynomial Predictors

Inexpensive higher order predictors are generally based on polynomial interpolation. It is known that such predictors are not very stable, in particular for larger stepsizes. In view of the stability of Newton's method as a corrector, it may be advantageous to also use more stable predictors. MACKENS [1989] has proposed such predictors which are based on Taylor's formula and which are obtained by successive numerical differentiation, see also SCHWETLICK and CLEVE [1987] as a predecessor. However, the gain in stability has to be paid for by additional evaluations of the map H and additional applications of the Moore-Penrose inverse of the Jacobian H' (where it may be assumed that H' has already been decomposed into some factorization).

Let us give a brief account of Mackens' approach. We assume again that $u_n = c(0)$ is a current point on the solution curve, and consider a local parametrization $c(\xi)$ with respect to the current tangent $t := t(H'(u_n))$, i.e., $c(\xi)$ solves

$$H(u) = 0, \qquad t^*(u - (u_n + \xi t)) = 0.$$

Note that by differentiating with respect to ξ we obtain immediately

$$c'(0) = t \quad \text{and} \quad c^{(k)}(0) \perp t \quad \text{for} \quad k > 1. \tag{11.1}$$

Consider the Taylor expansion

$$c(\xi) = \sum_{i=0}^{k} \frac{1}{i!} c^{(i)}(0)\xi^i + \mathcal{O}(\xi^{k+1}).$$

In principle, the derivatives of c can be obtained by differentiating the equation $H(c(\xi)) = 0$ repeatedly and evaluating at zero:

$$\begin{array}{rcccl} H'(u_n)c'(0) & = & 0 & =: & R_1 \\ H'(u_n)c''(0) & = & -H''(u_n)[c'(0), c'(0)] & =: & R_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ H'(u_n)c^{(k)}(0) & = & \cdots & =: & R_k \end{array} \tag{11.2}$$

Note that the orthogonality (11.1) and (3.4) implies that

$$c^{(k)}(0) = H'(u_n)^+ R_k \tag{11.3}$$

for $k > 1$.

The crucial observation is to note that the values R_k can also be obtained via

$$\left(\frac{d}{d\xi}\right)^{(k)} H\left(\sum_{i=0}^{k-1}\frac{1}{i!}c^{(i)}(0)\xi^i\right)\Bigg|_{\xi=0} = -R_k \tag{11.4}$$

for $k > 1$.

PROOF Let $p(\xi) := \sum_{i=0}^{k-1}\frac{1}{i!}c^{(i)}(0)\xi^i$. On the one hand we have

$$\left(\frac{d}{d\xi}\right)^{(k)} H(c(\xi)) = H'(c(\xi))c^{(k)}(\xi) - R_k(\xi) = 0,$$

and on the other hand

$$\left(\frac{d}{d\xi}\right)^{(k)} H(p(\xi)) = H'(p(\xi))p^{(k)}(\xi) + S_k(\xi).$$

The expressions $-R_k(\xi)$ and $S_k(\xi)$ contain higher derivatives of H at c and p, respectively, and derivatives of c and p up to order $k-1$. Except for replacing the c's by the p's, the expressions $-R_k(\xi)$ and $S_k(\xi)$ are the same. Since $c^{(i)}(0) = p^{(i)}(0)$ for $i < k$, and since $p^{(k)} = 0$, we obtain $-R_k(0) = S_k(0)$. □

By combining (11.2–11.4), it becomes evident that the derivatives of c can be obtained recursively. We summarize the discussion with the following

ALGORITHM 11.1 **(Taylor Coefficients)**

INPUT $u_n \in \mathbb{R}^{N+1}$ such that $H(u_n) = 0$
OUTPUT $c_k = c^{(k)}(0)$ for $k = 0, 1, 2, \ldots$

$c_0 \leftarrow u_n$
$c_1 \leftarrow t(H'(u_n))$
FOR $k = 2, 3, \ldots$
$\quad R_k \leftarrow -\left(\frac{d}{d\xi}\right)^{(k)} H\left(\sum_{i=0}^{k-1}\frac{1}{i!}c_i\xi^i\right)\Big|_{\xi=0}$
$\quad c_k \leftarrow H'(u_n)^+ R_k$
END FOR

The k-th derivative employed in the above algorithm for obtaining the coefficient R_k can be numerically approximated by difference formulae (preferably by central difference formulae). The formulae make use of a certain meshsize, say δ. The resulting truncation errors are then propagated to the higher derivatives in the above algorithm. MACKENS [1989] gives a careful discussion of this truncation error propagation.

Chapter IV

Special Topics

12 Implementation of the Moore-Penrose Inverse

The main ingredients of the predictor corrector method from the linear algebra standpoint consist of obtaining an approximation of the tangent $t(A)$ and either an implementation of an action of the Moore-Penrose inverse A^+ for a given approximation $A \approx H'(u_i)$ of the Jacobian, or an action of

$$\begin{pmatrix} A \\ t^* \end{pmatrix}^{-1} \tag{12.1}$$

where t^* is an appropriate additional row. In this section we concentrate on the Moore-Penrose inverse, and in the next section on the bordering algorithm for an efficient implementation of the action of (12.1).

A straightforward and simple (but not the most efficient) way to handle the numerical linear algebra would be to use a QR factorization:

$$A^* = Q \begin{pmatrix} R \\ 0^* \end{pmatrix}, \tag{12.2}$$

where Q is an $(N{+}1,N{+}1)$ orthogonal matrix, and R is a nonsingular (N, N) upper triangular matrix. We assume that A is an $(N, N{+}1)$ matrix with maximal rank. If q denotes the last column of Q, then $t(A) = \sigma q$, where the orientation defined in (1.5) leads to the choice

$$\sigma = \operatorname{sign}\Big(\det Q \, \det R \Big). \tag{12.3}$$

Hence σ is easy to determine. The Moore-Penrose inverse of A can be obtained from the same decomposition in the following way:

$$A^+ = A^*(AA^*)^{-1} = Q \begin{pmatrix} (R^*)^{-1} \\ 0^* \end{pmatrix}. \tag{12.4}$$

Of course, one does not generally calculate A^+, but instead the action of A^+ involved in the above equation, i.e., a backsolving and a matrix multiplication. The matrix Q will probably be available in factored form.

Similar ideas apply if an LU decomposition is employed:

$$PA^* = L \begin{pmatrix} U \\ 0^* \end{pmatrix}, \tag{12.5}$$

where L is a lower triangular (N+1,N+1) matrix, U is an (N, N) upper triangular matrix, and P is a permutation matrix corresponding to partial pivoting of the columns of A, which is in general necessary to improve the numerical stability.

Let us first consider the calculation of the tangent vector $t(A)$. From (12.5) it follows that

$$A = (U^*, 0)L^*P. \tag{12.6}$$

Hence, if we set

$$y := P^*(L^*)^{-1}e_{N+1}\,,$$

then it is readily seen from (12.6) that $Ay = 0$. Of course $y \neq 0$, and can be calculated by one backsolving and a permutation of its co-ordinates. Hence $t(A) = \sigma y/||y||$, where the sign σ is determined by evaluating the sign of the determinant of

$$\begin{aligned} (A^*, y) &= \left(P^*L \begin{pmatrix} U \\ 0^* \end{pmatrix}, P^*(L^*)^{-1}e_{N+1} \right) \\ &= P^*L \left(\begin{pmatrix} U \\ 0^* \end{pmatrix}, L^{-1}(L^*)^{-1}e_{N+1} \right). \end{aligned}$$

The last entry $||(L^*)^{-1}e_{N+1}||^2$ of $L^{-1}(L^*)^{-1}e_{N+1}$ must be positive, and hence

$$\sigma = \text{sign } \det(A^*, y) = \text{sign } \det(P)\det(L)\det(U).$$

The right hand side is easily determined.

Let us now turn to the problem of determining the Moore-Penrose inverse. From (12.6) it follows that

$$B := P^*(L^*)^{-1} \begin{pmatrix} (U^*)^{-1} \\ 0^* \end{pmatrix}$$

is a right inverse of A, and hence $A^+ = \big(\mathrm{Id} - t(A)t(A)^*\big)B$ by (3.9). Finally, let us note that one action of A^+ amounts to essentially one forward solving with U^*, one backsolving with L^*, and one scalar product for the orthogonal projection with $\big(\mathrm{Id} - t(A)t(A)^*\big)$.

The above methods are useful for small dense matrices A. However, in many applications of path following methods, the corresponding matrix A is large and sparse, and then the above procedure is inefficient. Among such applications are the approximation of branches of nonlinear eigenvalue problems or the central path methods of linear and nonlinear programming. Let us discuss some ideas which are useful in dealing with such situations.

In many applications of path following, one encounters matrices A with the following structure:

$$A = \big(A_0\,, b\big)\,, \tag{12.7}$$

where equations of the form $A_0 x = y$ permit a fast linear solver. We will refer to one solving of such a sytem as one *action* of A_0^{-1}. In the path following it is convenient to consider an augmented matrix of the form:

$$\tilde{A} := \begin{pmatrix} A_0 & b \\ c^* & d \end{pmatrix} \tag{12.8}$$

where the additional row $\begin{pmatrix} c^* & d \end{pmatrix}$ is typically generated via the last predictor direction. See also the discussion of the bordering algorithm in Section 13.

A standard block elimination

$$\begin{pmatrix} A_0 & b \\ c^* & d \end{pmatrix} \to \begin{pmatrix} A_0 & b \\ 0 & d - c^* A_0^{-1} b \end{pmatrix}$$

may be employed to generate the Schur complement

$$s := d - c^* A_0^{-1} b$$

of A_0 in the augmented matrix $\tilde{A}$. Clearly,

$$\det \tilde{A} = s \, \det A_0. \tag{12.9}$$

It can be checked that if $\tilde{A}$ is nonsingular, then

$$\tilde{A}^{-1} = \begin{pmatrix} A_0^{-1} + A_0^{-1} b s^{-1} c^* A_0^{-1} & -A_0^{-1} b s^{-1} \\ -s^{-1} c^* A_0^{-1} & s^{-1} \end{pmatrix}. \tag{12.10}$$

As an easy consequence, the tangent $t(A)$ is obtained via

$$t(A) = \sigma y / ||y||, \tag{12.11}$$

where y denotes the last column of $\tilde{A}^{-1}$. The sign $\sigma \in \{\pm 1\}$ can either be obtained from an angle test with the previous predictor direction or from (12.9). In fact,

$$\begin{aligned}
\det \begin{pmatrix} (A_0, b) \\ y* \end{pmatrix} &= \det \begin{pmatrix} A_0 & b \\ -(A_0^{-1} b s^{-1})^* & s^{-1} \end{pmatrix} \\
&= s^{-1} \det \begin{pmatrix} A_0 & b \\ -b^*(A_0^*)^{-1} & 1 \end{pmatrix} \\
&= s^{-1} \det \begin{pmatrix} A_0 & b \\ 0 & 1 + b^*(A_0^*)^{-1} A_0^{-1} b \end{pmatrix} \\
&= s^{-1} \det(A_0)(1 + ||A_0^{-1} b||^2)
\end{aligned}$$

shows that

$$\sigma = \operatorname{sign}\Big(s \det A_0\Big). \tag{12.12}$$

To obtain the Moore-Penrose inverse, let us first note that the matrix $(\tilde{A}^{-1})_N$ consisting of the first N columns of $\tilde{A}^{-1}$ is a right inverse of A. Therefore (3.9) implies

$$A^+ = \big(I - t(A) t(A)^*\big)\big(\tilde{A}^{-1}\big)_N. \tag{12.13}$$

Note that after an initial computational expense of one action of A_0^{-1}, i.e., $A_0^{-1} b$, and one scalar product (to calculate s), the cost of calculating $t(A)$ amounts to essentially one additional scalar product, the cost of one action of $\tilde{A}^{-1}$ amounts to essentially one additional action of A_0^{-1} and one additional scalar product, and the cost of one action of A^+ amounts to essentially one additional call of the action of A_0^{-1} and two additional scalar products (assuming that $t(A)$ has been calculated).

Among the fast solvers which are of importance here are direct solvers for sparse linear systems, or preconditioned iterative solvers such as conjugate-gradient or other Krylov subspace methods, see, e.g., Freund, Golub, and Nachtigal [1992].

The above Schur complement construction is also valid if b, c and d are matrices (of appropriate size). This is of interest in parametric optimization, see Lundberg and Poore [1993].

If A_0 is symmetric, then an Arnoldi method may be used as a fast linear solver. As a byproduct, the generated Hessenberg matrix may be used to approximate the largest and smallest eigenvalues of A_0. In typical applications, this will often be sufficient to determine the sign of $\det A_0$, see Huitfeldt and Ruhe [1990].

13 The Bordering Algorithm

The popular bordering algorithm of KELLER [1977, 1983], see also CHAN [1984a], CHAN and SAAD [1985], MENZEL and SCHWETLICK [1978, 1985], is related to the ideas of the previous section. These approaches are akin to KELLER's [1977] pseudo arclength method, see Section 2, in which the equation $H(u) = 0$ is extended by an additional condition $F(u, s) = 0$ which models a pseudo arclength parametrization. This viewpoint is often convenient, in particular for structured problems such as those arising from discretizations of nonlinear eigenvalue problems.

Let us give a brief account of the bordering algorithm. We describe the implicit equations in the form $H(u) = H(x, \lambda) = 0$ where $H : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^N$ is sufficiently smooth and $\lambda \in \mathbb{R}$ is an exceptional parameter such as a nonlinear eigenvalue parameter in discretizations of boundary value problems. We assume that 0 is a regular value of H. The derivative $A_0 := H_x$ of H with respect to the space variables $x \in \mathbb{R}^N$ usually leads to some sparse matrix, and solving equations with this sparse matrix can typically be accomplished by some special solver such as an iterative method or a multigrid method, see, e.g., CHAN and KELLER [1982], HACKBUSCH [1994], FREUND, GOLUB, and NACHTIGAL [1992]. Hence, it is desirable to exploit this special solver in a predictor corrector method.

Introducing the pseudo arclength parameter s via $F(u, s) = 0$, see (2.1), we consider the Newton iteration (2.4) where we need to solve a linear system of equations involving the matrix

$$\tilde{A} := \begin{pmatrix} A_0 & b \\ c^* & d \end{pmatrix} = \begin{pmatrix} H'(u) \\ F_u(u, s) \end{pmatrix} = \begin{pmatrix} H_x(\bar{x}, \bar{\lambda}) & H_\lambda(\bar{x}, \bar{\lambda}) \\ F_x(\bar{x}, \bar{\lambda}, 0) & F_\lambda(\bar{x}, \bar{\lambda}, 0) \end{pmatrix}.$$

We can now follow a discussion very similar to the one starting at (12.8). We refer to the computation of an action of $\tilde{A}^{-1}$ according to (12.10) as the *bordering algorithm*. In particular, we have seen that after an initial computational cost of one action of A_0^{-1} and one scalar product, one action of $\tilde{A}^{-1}$ can be obtained essentially for the cost of one action of A_0^{-1} and one scalar product. It is also pointed out there that the tangent $t(A_0, b)$ can be obtained for essentially one additional scalar product. Hence, solving for a bordered system of equations, i.e., a system involving the matrix $\tilde{A}$, is only slightly cheaper than performing one action of the pseudo inverse $H'(u)^+ = (A_0, b)^+$.

A critical case is the situation where A_0 is nearly singular. One would then expect that the bordering algorithm would fail, since the inversion of A_0 is badly conditioned. However, KELLER [1982, 1987] points out that under certain circumstances the bordering algorithm can still be used. More

precisely, let us assume that $\tilde{A}$ is nonsingular (i.e., $\mathrm{cond}(\tilde{A}) = \mathcal{O}(1)$), and

$$PA_0Q = \begin{pmatrix} L & 0 \\ l^* & 1 \end{pmatrix} \begin{pmatrix} U & u \\ 0 & \varepsilon \end{pmatrix}$$

is an LU factorization of A_0 where column and row pivots (indicated by the permutation matrices Q and P) are used in such a way that ε catches the order of singularity of $\tilde{A}$, i.e, $\mathrm{cond}(A_0) = \mathcal{O}(\varepsilon^{-1})$. If this factorization is used to calculate an action of $\tilde{A}^{-1}$ according to (12.10), then even for small ε this bordering algorithm is acceptable, though possibly some cancellation of digits must be expected. Numerical experience confirms this analysis.

One last remark involves the weighting of the space variables x against the eigenvalue parameter λ. If the nonlinear system $H(x, \lambda) = 0$ is obtained via a discretization of some boundary value problem over a domain in $\mathbb{R}^n$ with stepsize h, then the standard Euclidean norm $||(x, \lambda)|| = \sqrt{||x||^2 + \lambda^2}$ and the corresponding scalar product would lead to a very bad scaling problem for λ, e.g., when calculating the tangent or approximating the arclength. A better choice would be to try to approximate the L^2 norm of the underlying function space, e.g.,

$$||(x, \lambda)|| := \sqrt{h^n||x||^2 + \lambda^2},$$

see KELLER [1987, p. 86].

14 Update Methods

In this section we will briefly describe how to incorporate an analogue of Broyden's update method into a predictor corrector algorithm. A more extensive discussion may be found in ALLGOWER and GEORG [1990, Chap. 7].

Let us first recall the Broyden update method for solving a zero-point problem $F(x) = 0$ where $F : \mathbb{R}^N \to \mathbb{R}^N$ is smooth. For a general reference on update methods we suggest the book of DENNIS and SCHNABEL [1983]. Suppose that $F(\bar{x}) = 0$ with $F'(\bar{x})$ having maximal rank N. It is well-known that Newton's method

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n), \qquad n = 0, 1, \ldots$$

is locally quadratically convergent. Even when an adequate starting point x_0 has been chosen, there remains the drawback that after every iteration the Jacobian matrix $F'(x_n)$ needs to be calculated and a new matrix decomposition has to be obtained in order to solve the linear system

$$F'(x_n)s_n = -F(x_n)$$

for $s_n := x_{n+1} - x_n$. On the other hand, if an approximate Jacobian is held fixed, say for example $A := F'(x_0)$, a familiar Newton-Chord method is obtained:

$$x_{n+1} = x_n - A^{-1}F(x_n).$$

This method offers the advantage that A stays fixed. Thus, once the matrix decomposition for A has been obtained, further iterations may be cheaply carried out. The drawback of the Newton-Chord method is that the local convergence is only linear.

The method of BROYDEN [1965] involves the use of previously calculated data to iteratively improve the quality of the approximation $A \approx F'(x_n)$ via successive rank-one updates. A more general class of update methods, usually called Quasi-Newton methods, have since been developed which take into account possible special structure of the Jacobian F' such as positive definiteness, symmetry or sparseness. It is possible to prove local superlinear convergence of a large class of these methods under the standard hypotheses for Newton's method.

For general purpose updates, i.e., when no special structure is present, the so-called "good formula" of Broyden appears to be rank-one update of choice. For this reason we will confine our discussion of update methods for curve-following to an analogue of this formula. Similar but more complicated extensions of the discussion below can be given for the case that special structure is present, see also BOURJI and WALKER [1990], WALKER [1990].

Let us motivate our discussion by reviewing the Broyden update formula for solving $F(x) = 0$ via a Newton-type method. From Taylor's formula, we have

$$F'(x_n)(x_{n+1} - x_n) = F(x_{n+1}) - F(x_n) + \mathcal{O}\big(||x_{n+1} - x_n||^2\big). \tag{14.1}$$

By neglecting the higher order term in (14.1), and setting

$$s_n := x_{n+1} - x_n\,, \qquad y_n := F(x_{n+1}) - F(x_n),$$

we obtain the *secant equation*

$$As_n = y_n\,, \tag{14.2}$$

which should be satisfied (at least to first order) by an approximate Jacobian $A \approx F'(x_n)$.

When Newton-type steps

$$x_{n+1} = x_n - A_n^{-1}F(x_n)$$

are performed using some approximate Jacobian $A_n \approx F'(x_n)$, it is natural to require that the next approximate Jacobian A_{n+1} should satisfy the secant equation

$$A_{n+1}s_n = y_n\,, \tag{14.3}$$

since the data s_n and y_n are already available.

Clearly, equation (14.3) does not uniquely determine A_{n+1}, since it involves N equations in N^2 unknowns. An additional natural consideration is that if A_n was a good approximation to $F'(x_n)$, then this quality ought to be incorporated in formulating subsequent approximations. This leads to the idea of obtaining A_{n+1} from A_n by the *least change principle*, i.e., among all matrices A satisfying the secant equation, we choose the one with the smallest distance from A_n. Thus we are led to the following definition: We define the updated approximate Jacobian A_{n+1} as the solution of the problem

$$\min_A \{||A - A_n||_F : As_n = y_n\} \tag{14.4}$$

where the norm $||\cdot||_F$ is the *Frobenius norm*:

$$||A||_F = \Big(\sum_{i,j=1}^{N} \big(A(i,j)\big)^2\Big)^{\frac{1}{2}}.$$

A straightforward calculation shows that the solution to (14.4) is given explicitly by

$$A_{n+1} = A_n + \frac{y_n - A_n s_n}{||s_n||^2}\, s_n^*\,, \tag{14.5}$$

which is generally referred to as *Broyden's good update formula*.

Let us now describe the analogue of Broyden's method for updating the Jacobian along the solution curve $c(s)$. Suppose we are given two approximate zero-points $u_n, u_{n+1} \in \mathbb{R}^{N+1}$ of H and the corresponding values $H(u_n), H(u_{n+1})$. Motivated by the preceding discussion we again set

$$\begin{aligned} s_n &:= u_{n+1} - u_n\,, \\ y_n &:= H(u_{n+1}) - H(u_n), \end{aligned}$$

and we consider the analogous secant equation (14.2) where $A \approx H'(u_n)$ is an approximate Jacobian. The corresponding *Broyden update on the points* u_n, u_{n+1} is again given by (14.5). The following algorithm illustrates how this update formula may be incorporated into an Euler-Newton method. Let us stress that updates are performed for both predictor and corrector steps. Note that, for convenience, the corrector points are now the even numbered u_n, and the predictor points the odd numbered u_n.

ALGORITHM 14.1 **(Euler-Newton Method Using Updates)**

INPUT $u_0 \in \mathbb{R}^{N+1}$ such that $H(u_0) \approx 0$ (initial point)
$h > 0$ initial steplength
OUTPUT u_i, $i = 0, 2, 4, \ldots$, approximating the solution curve

approximate $A_0 \approx H'(u_0)$
FOR $i = 1, 3, 5, \ldots$
$u_i \leftarrow u_{i-1} + ht(A_{i-1})$ % *predictor step*
update A_i from A_{i-1} on u_i, u_{i-1} % *predictor update*
$u_{i+1} \leftarrow u_i - A_i^+ H(u_i)$ % *corrector step*
update A_{i+1} from A_i on u_{i+1}, u_i % *corrector update*
choose a new steplength $h > 0$
END FOR

Let us make a few remarks concerning the above algorithm. It does not necessarily generate a reliable approximation of the Jacobians $H'(u_i)$, i.e., a relation such as

$$||A_i - H'(u_i)|| = \mathcal{O}(h),$$

(which we assumed in Theorem 5.1) does not hold in general. The reason behind this is that we are in general not assured that the update data spans the whole space $\mathbb{R}^{N+1}$ sufficiently well. To put this more precisely, let S denote the $(N+1, k)$ matrix (for some fixed $k \geq 2N$), whose columns have unit norm and indicate the last k directions in which the update formula was used in Algorithm 14.1. Then the condition

$$\text{cond}(SS^*) < C$$

for some $C > 0$ independent of h and of whichever last k directions S are taken, is sufficient to ensure the convergence of Algorithm 14.1 for h sufficiently small, in the sense of Theorem 5.1. See GEORG [1982, Chap. 4] for a sketch of a proof. However, it is unrealistic to expect that this condition will in general be satisfied. Let us briefly describe how this difficulty may be circumvented. For details we refer to ALLGOWER and GEORG [1990, Chap. 7].

Setting

$$\begin{aligned} t_n &:= t(A_n), \\ w_n &:= \frac{H(u_{n+1}) - H(u_n) - A_n(u_{n+1} - u_n)}{||u_{n+1} - u_n||}, \\ v_n &:= \frac{u_{n+1} - u_n}{||u_{n+1} - u_n||}, \end{aligned}$$

we see that the update formula (14.5) may be written as

$$A_{n+1} := A_n + w_n v_n^* \,.$$

Two types of updates arise, namely predictor updates and corrector updates. For the predictor update we have $u_{n+1} = u_n + ht_n$ and consequently

$$\begin{aligned} w_n &= \frac{H(u_{n+1}) - H(u_n)}{h} \,, \\ v_n &= t_n \,. \end{aligned}$$

It can be verified that

$$\begin{aligned} t_n &\perp A_n^+ w_n \,, \\ t_{n+1} &= \frac{t_n - A_n^+ w_n}{||t_n - A_n^+ w_n||} \,, \\ A_{n+1}^+ &= (\mathrm{Id} - t_{n+1} t_{n+1}^*) A_n^+ \,. \end{aligned}$$

From this it is clear that updates based *only* upon predictor steps cannot in general maintain good approximations of the Jacobian H'.

For the corrector update we have $u_{n+1} = u_n - A_n^+ H(u_n)$ and consequently

$$\begin{aligned} w_n &= \frac{H(u_{n+1})}{||A_n^+ H(u_n)||} \,, \\ v_n &= -\frac{A_n^+ H(u_n)}{||A_n^+ H(u_n)||} \,. \end{aligned} \tag{14.6}$$

Setting

$$D_n := 1 + v_n^* A_n^+ w_n \,,$$

it can be verified in this case that

$$\begin{aligned} t_{n+1} &= \operatorname{sign} D_n t_n \,, \\ A_{n+1}^+ &= \left(\mathrm{Id} - \frac{A_n^+ w_n v_n^*}{D_n} \right) A_n^+ \,. \end{aligned}$$

From this it is again clear that updates based *only* upon corrector steps cannot maintain a good approximation of the Jacobian H'.

A consequence of (14.6) is that

$$||A_n^+ w_n|| = \frac{||A_n^+ H(u_{n+1})||}{||A_n^+ H(u_n)||} \,. \tag{14.7}$$

The vector $A_n^+ w_n$ arises naturally in the corrector update formula, and its norm gives a reasonable measure for the contraction rate of the corrector step. If this rate is large, then this may be attributed to one of two factors: either the predictor point was too far from the solution curve, or the Jacobian H' is poorly approximated by A_n. Often the first of these two possibilities is easy to check, e.g., by monitoring one of the quantities $||H(u_n)||$, $||A_n^+ H(u_n)||$. Hence $||A_n^+ w_n||$ affords us an empirical measure of the quality of the approximation of the Jacobian.

Let us return to the basic issue of Algorithm 14.1, namely to assure that a good approximation of the Jacobian H' is maintained. To do this, we may be guided by a device proposed by POWELL [1970a, 1970b] in which he suggests monitoring the directions of the differences contributing to the updates. In our context, this amounts to monitoring whether the condition number $\mathrm{cond}(SS^*)$ discussed above is sufficiently small. Instead of following Powell's approach of introducing additional directions, we combine a stepsize adaptation with several tests, the most important being the measuring of the above contraction rate (14.7). If the criteria of the tests are satisfied, the algorithm performs a successive predictor corrector step with increased stepsize. If however, the criteria are not satisfied, the predictor corrector step is repeated with reduced stepsize. In both cases, the predictor and corrector updates which were already performed are *not* rejected. This enables the method to generally update an approximation of the Jacobian in directions which are most needed. The following detailed version of Algorithm 14.1 incorporates these considerations. It is given as a possible implementation of the above discussion.

ALGORITHM 14.2 **(Euler-Newton Method Monitoring Updates)**

INPUT $u_0 \in \mathbb{R}^{N+1}$ such that $H(u_0) \approx 0$ (initial point)
$h > 0$ initial steplength
$\delta_0 > 0$ minimal residual
$\kappa \in (0,1)$ maximal contraction rate
OUTPUT u_i , $i = 0, 2, 4, \ldots$, approximating the solution curve

approximate $A \approx H'(u_0)$
$i \leftarrow 1$
WHILE a stopping criterion is not met
% *predictor step*
$u_i \leftarrow u_{i-1} + ht(A)$
$A \leftarrow A + h^{-1}\big(H(u_i) - H(u_{i-1})\big)t(A)^*$
% *perturbation*
IF $||A^+ H(u_i)|| \leq \delta_0$
choose p such that $||A^+\big(H(u_i) - p\big)|| = \delta_0$

ELSE $p \leftarrow 0$
END IF
% *corrector step*
$u_{i+1} \leftarrow u_i - A^+(H(u_i) - p)$
$A \leftarrow A + ||u_{i+1} - u_i||^{-2}\big(H(u_{i+1}) - p\big)(u_{i+1} - u_i)^*$
% *contraction test*
IF $||A^+\big(H(u_i) - p\big)||^{-1}\,||A^+\big(H(u_{i+1}) - p\big)|| > \kappa$
$h \leftarrow h/2$; BREAK LOOP
END IF
$h \leftarrow 2h$
$i \leftarrow i + 2$
END WHILE

Let us emphasize that usually an implementation of the above algorithm updates some decomposition of the matrix A_n at each step. For details on such update methods we refer to GOLUB and VAN LOAN [1989, Sec. 12.6] or ALLGOWER and GEORG [1990, Chap. 16]. The perturbation term p in the above algorithm serves several purposes. The main purpose is to prevent numerical instability in the corrector update formula due to cancellation of digits. As a general rule for choosing δ_0, one may require that $v - w$ should carry at least half as many significant digits as the computer arithmetic carries. Another purpose of the perturbation p is to safeguard the algorithm against intrinsic instabilities such as those arising from singular points on the curve, e.g., bifurcation points. In fact, the above algorithm will usually bifurcate off at simple bifurcation points.

It is sometimes also useful to incorporate more tests in the step "contraction test" of the above algorithm, e.g., so that a maximal distance to the curve or a maximal angle between successive secants is not exceeded.

Several versions of the above algorithm have been implemented and tested on a variety of non-sparse problems, see, e.g., GEORG [1981a, 1981b, 1982], and they have turned out to be very efficient. It should be noted that an iteration of the corrector step has not been incorporated. This needs only to be done if it is wished to follow the curve closely. However, one additional corrector step could easily be performed at low cost since $A^+\big(H(w) - p\big)$ is calculated in the "contraction test".

Let us again emphasize that it is important that although a predictor or corrector point may not be accepted, because it may not satisfy the test criteria, the update information which it contributes to approximating the Jacobian is nevertheless utilized. Finally, let us point out that the essential feature of the above algorithm is that it updates *along* the curve as well as in the corrector steps. This is to be distinguished from Euler-Newton methods where the Jacobian is precisely calculated at the predictor point,

and updates are only used to accelerate the Newton-type corrector iteration. In this case, proofs of superlinear convergence have been given in Bourji and Walker [1990] and Walker [1990].

15 Linear Solvers Exploiting Symmetry

This section uses some technical concepts from group theory, and can be skipped by the reader without loss of continuity.

Many problems in science and mathematics may be transformed via a linear or nonlinear change of coordinates into related problems. A tranformation which carries a problem or equation into itself is called a *symmetry* of the equation. Usually the symmetry is derived from underlying geometric symmetries of the domain or body on which the problem is considered. The incorporation of symmetry considerations in the analysis and solution of such problems may result in enormous gains in insight and efficiency.

The numerical treatment of problems such as partial differential equations and integral equations generally involves discretizations. In this setting a general principle may be formulated: if a domain enjoys a symmetry group, and an operator equation which expresses a coordinate-free physical law is to be solved on that domain, and if the discretization is chosen to incorporate or respect the symmetries, then the resulting numerical problem (which is usually a matrix equation) is amenable to a block decomposition which greatly reduces the computational cost of determining the solution.

Symmetry groups have long played a significant role in physics, engineering and other areas of science. As Bossavit [1986] points out, any randomly selected issue of the Journal of Mathematical Physics will generally contain four or five papers with entries "group", "symmetry" or "representation theory" in their list of key words. Also in nonlinear analysis, there is considerable literature involving symmetry aspects, more recently in bifurcation phenomena, see, e.g., Golubitsky and Schaeffer [1985], Golubitsky, Stewart, and Schaeffer [1988], Healey [1988a, 1988b, 1989], Vanderbauwhede [1982]. A similar search in any numerical analysis journal will typically turn up no such papers at all. It is surprising to see how little systematic use of symmetry groups has been made in the numerical treatment of operator equations exhibiting geometric symmetries.

A start in this direction was made in Fässler [1976] and the book Fässler and Stiefel [1979]. In fact, there is a very early forerunner by Stiefel [1952]. Thereafter several papers began to appear, see, e.g., Guy and Mangeot [1981], Ballisti, Hafner, and Leuchtman [1982] and Bossavit [1986, 1993]. Recently symmetry reduction methods for linear systems have been investigated, see Allgower, Böhmer, Georg, and

MIRANDA [1992a], ALLGOWER and FÄSSLER [1993], ALLGOWER, GEORG, and MIRANDA [1993a], ALLGOWER, GEORG, and WALKER [1993b], GEORG and MIRANDA [1992, 1993].

The usefulness of exploiting the actions of symmetry groups in continuation methods is illustrated with the example of bifurcation in Section 24.3, see also Section 25. In this section we want to show how the exploitation of symmetry can significantly reduce the computational cost of solving linear systems of equations (which is one of the basic cost factors in continuation methods).

To briefly illustrate the approach, let us consider an operator equation

$$\mathcal{L}f = g \tag{15.1}$$

where f and g are elements of a Banach space $\mathcal{F}$ of functions defined on a domain $\mathcal{D}$ in $\mathbb{R}^3$. For example, if the operator equation describes a partial differential equation, then $\mathcal{D}$ is a three-dimensional body, and if the operator equation describes a boundary integral formulation, then $\mathcal{D}$ is the surface of a three-dimensional body.

Often, $\mathcal{D}$ exhibits geometric symmetries (e.g., $\mathcal{D}$ may be a cube) which may be described by the action of a certain finite group Γ of orthogonal transformations. For many classical equations, these symmetries are respected by the operator equations. This situation is usually described by saying that the operator $\mathcal{L}$ is *equivariant* with respect to the group action (i.e., commutes with the group actions).

More precisely, if $\gamma \in \Gamma$ is an orthogonal transformation which is an isometry of the body $\mathcal{D}$, then γ acts on the Banach space $\mathcal{F}$ by the formula

$$(\gamma f)(x) = f\big(\gamma^{-1}x\big)$$

for $f \in \mathcal{F}$ and $x \in \mathcal{D}$. The equivariance of the operator $\mathcal{L}$ is expressed by requiring that

$$\mathcal{L}\gamma = \gamma\mathcal{L}$$

for all isometries $\gamma \in \Gamma$.

We are now in a position to discuss the numerical solution of the operator equation $\mathcal{L}f = g$ where $\mathcal{L}$ is equivariant with respect to some group Γ of isometries, and in particular, to exploit the symmetry in order to reduce the computational effort as much as possible. It should be stressed that no symmetry assumptions are made concerning either g or f in the equation $\mathcal{L}f = g$.

If the linear operator equation $\mathcal{L}f = g$ is suitably discretized (respecting the action of a symmetry group Γ), then a matrix equation $Lu = v$ results,

where L is an equivariant (N, N)-matrix with respect to Γ which now acts as a group of permutations on $\{1, \ldots, N\}$, i.e.,

$$L(\gamma i, \gamma j) = L(i, j) \text{ for } \gamma \in \Gamma. \tag{15.2}$$

For example, let us consider a collocation method (e.g., with constant elements) for the operator equation $\mathcal{L}f = g$ based on a choice of basis functions $\{\phi_j\}_{j\in J} \subset \mathcal{F}$ and collocation points $\{s_j\}_{j\in J}$ indexed by the same set J, such that interpolation is possible, i.e., the matrix $\phi_j(s_k)$ has nonzero determinant. Higher order collocation or Galerkin methods can be treated in a very similar way. Following standard techniques we can write the collocation equation as

$$Lu = v,$$

where the entries of the square matrix L are given by

$$L(k, j) := (\mathcal{L}\phi_j)(s_k)\,,$$

the column v is given by $v(k) := g(s_k)$, and the unknown column u represents the coefficients of the approximate solution, i.e., $f \approx \sum_{j\in J} u(j)\phi_j$. The standard assumption in collocation methods is that L is a nonsingular matrix.

In order to exploit the symmetry, we need to make the assumption that for all $\gamma \in \Gamma$ we have

$$\{\gamma\phi_j\}_{j\in J} \text{ is a permutation of } \{\phi_j\}_{j\in J}\,,$$

and

$$\{\gamma s_j\}_{j\in J} \text{ is a permutation of } \{s_j\}_{j\in J}\,,$$

which signifies that the chosen configuration for the collocation method respects the symmetry structure. This will usually be easy to satisfy in practical cases. For simplicity, let us assume that the induced permutations of the index set J are the same for the two actions. These conditions lead immediately to the equivariance of the system matrix L as described in (15.2).

Let us look at the simplest case, i.e., Γ is a cyclic group of order N and $\gamma : i \mapsto i+1$ is the generating element of Γ. Here the indices i of the matrix entries are viewed modulo N. It can be easily seen that the matrix L has the circulant form (illustrated for $N = 5$)

$$L = \begin{pmatrix} a_1 & a_5 & a_4 & a_3 & a_2 \\ a_2 & a_1 & a_5 & a_4 & a_3 \\ a_3 & a_2 & a_1 & a_5 & a_4 \\ a_4 & a_3 & a_2 & a_1 & a_5 \\ a_5 & a_4 & a_3 & a_2 & a_1 \end{pmatrix}.$$

Circulant matrices have been studied and used extensively, see, e.g., DAVIS [1979] and VAN LOAN [1992]. The symmetry reduction method for solving the equation $Lu = v$ is well known in this case and consists of an application of the finite Fourier Transform (illustrated for $N = 5$)

$$F = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \omega^4 \\ 1 & \omega^2 & \omega^4 & \omega^6 & \omega^8 \\ 1 & \omega^3 & \omega^6 & \omega^9 & \omega^{12} \\ 1 & \omega^4 & \omega^8 & \omega^{12} & \omega^{16} \end{pmatrix}, \tag{15.3}$$

where ω is a primitive N-th root of unity, and the exponents could (of course) be modified modulo N. It is seen that F is unitary. Hence we consider the transformed equation

$$\hat{L}\hat{u} = \hat{v} \text{ where } \hat{L} := FLF^*,\ \hat{u} := Fu,\ \hat{v} := Fv.$$

If we denote the first column of L by a, then it is seen that $\hat{L} = \sqrt{N}\,\mathrm{diag}(\hat{a})$, i.e., the finite Fourier transform reduces the problem of solving $Lu = v$ to the (trivial) solving of N scalar equations $\sqrt{N}\,\hat{a}(i)\hat{u}(i) = \hat{v}(i)$ and three finite Fourier transforms, namely $a \mapsto \hat{a}$, $v \mapsto \hat{v}$ and $\hat{u} \mapsto u$. Systems involving block circulant matrices are handled analogously, e.g., the column a now becomes a column of block matrices.

Interestingly enough, the above finite Fourier transform method has not been generalized to arbitrary equivariant matrices until very recently, i.e., TAUSCH [1993] has interpreted the symmetry reduction approach in GEORG and MIRANDA [1993] in this way. Let us give a brief sketch of this generalization for the case of block structures.

Hence, we assume that the matrix L splits into (k,k)-blocks $L(i,j)$, $i,j = 1,\ldots,N$. The equivariance (15.2) is now written for these blocks. Let us further assume (for simplicity) that the actions of the permutations $\gamma \in \Gamma$ are fixed point free, i.e.,

$$\gamma j = j \tag{15.4}$$

for some $\gamma \in \Gamma$ and some $j \in \{1,\ldots,N\}$ implies $\gamma = 1$, the identity element of Γ. If we take the above described block structure as large as possible (possibly rearranging the ordering), it is clear that the order of Γ equals N, hence, for the purposes of this exposition, we index L via Γ. The equivariance now leads to

$$L(\alpha,\beta) = L(\beta^{-1}\alpha, 1) =: a(\beta^{-1}\alpha) \text{ where } \alpha,\beta \in \Gamma.$$

The matrix L is now determined by its first block column a.

To describe the decomposition of L via equivariance, we recall the concept of the irreducible representations of a group. Let $U(i)$ denote the group of unitary (i,i)-matrices. An irreducible representation (of dimension i) $\rho : \Gamma \to U(i)$ is a group homomorphism without a proper invariant subspace, i.e., there is no proper linear subspace $V \subset \mathbb{C}^i$, $\dim V \notin \{0, i\}$, such that $\rho(\gamma)V = V$ for all $\gamma \in \Gamma$. Two such representations are equivalent if they result from each other via a unitary similarity transformation. It is well known, see, e.g., SERRE [1977], that any finite group admits a complete list $\mathcal{R}(\Gamma)$ of irreducible representations which are mutually non-equivalent. Furthermore,

$$\sum_{\rho \in \mathcal{R}(\Gamma)} \dim \rho^2 = \operatorname{order}(\Gamma) = N. \tag{15.5}$$

Let us now construct the analogue F of the finite Fourier transform. A typical block row of F has the elements

$$\sqrt{\frac{\dim \rho}{N}} \Big(\rho(\gamma)(i,j)\, I_k \Big)_{\gamma \in \Gamma}$$

where i, j are some indices in $\{1, \ldots, \dim \rho\}$ and I_k is the (k,k) identity matrix. By (15.5), F is a (kn, kn)-matrix, and the well-known orthogonality relations in group representation theory, see, e.g., SERRE [1977], imply that F is unitary.

Now we can repeat for equivariant matrices the program described above for circulant matrices:

$$\hat{L}\hat{u} = \hat{v} \text{ where } \hat{L} := FLF^*,\ \hat{u} := Fu,\ \hat{v} := Fv.$$

However, $\hat{L}$ is not a simple block diagonal matrix generated by $\hat{a}$. To describe its structure, let us index the block entries of $\hat{L}$ by $\rho(i,j)$ where $\rho \in \mathcal{R}(\Gamma)$ and $i, j \in \{1, \ldots, \dim \rho\}$. This is possible by (15.5). Then

$$\hat{L}\Big(\rho(i,j), \rho'(i',j')\Big) = \delta_{\rho\rho'}\, \delta_{ii'}\, \hat{a}\Big(\rho(j',j)\Big).$$

Hence, a typical block diagonal entry of $\hat{L}$ referring to ρ consists of $\dim \rho$ identical blocks each of size $k \dim \rho$.

By applying this "generalized" finite Fourier transform, we thus obtain a subdivision of the linear problem $Lu = v$ into much smaller subproblems.

As an illustration, let us consider the most simple example of an equivariant matrix with respect to a non-abelian group: let

$$D_3 = \{1, R, R^2, F, FR, FR^2\}$$

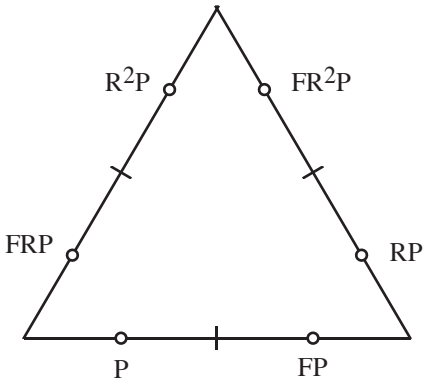

Figure 15.1. The Group D_3

be the symmetry group of an equilateral triangle. It is generated, e.g., by a $2\pi/3$ rotation R and a reflection F across the y-axis, see Figure 15.1. It has 2 irreducible representations of dimension one and 1 irreducible representation of dimension two.

The multiplication table of the group (the elements ordered as above) has the following form:

	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6
γ_1	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6
γ_2	γ_2	γ_3	γ_1	γ_6	γ_4	γ_5
γ_3	γ_3	γ_1	γ_2	γ_5	γ_6	γ_4
γ_4	γ_4	γ_5	γ_6	γ_1	γ_2	γ_3
γ_5	γ_5	γ_6	γ_4	γ_3	γ_1	γ_2
γ_6	γ_6	γ_4	γ_5	γ_2	γ_3	γ_1

A complete list of irreducible representations is captured in the following generalized Fourier transform:

$$F = \begin{pmatrix} \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{3}}{6} & -\frac{\sqrt{3}}{6} & -\frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{6} \\ 0 & -1/2 & 1/2 & 0 & 1/2 & -1/2 \\ 0 & 1/2 & -1/2 & 0 & 1/2 & -1/2 \\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{3}}{6} & -\frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{3} & -\frac{\sqrt{3}}{6} & -\frac{\sqrt{3}}{6} \end{pmatrix}.$$

Then the equivariant matrix

$$L = \begin{pmatrix} 1 & 3 & 4 & 6 & 7 & 9 \\ 4 & 1 & 3 & 7 & 9 & 6 \\ 3 & 4 & 1 & 9 & 6 & 7 \\ 6 & 7 & 9 & 1 & 3 & 4 \\ 7 & 9 & 6 & 4 & 1 & 3 \\ 9 & 6 & 7 & 3 & 4 & 1 \end{pmatrix}$$

is transformed into the matrix

$$\hat{L} = \begin{pmatrix} 30 & 0 & 0 & 0 & 0 & 0 \\ 0 & -14 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1/2 & -\frac{\sqrt{3}}{2} & 0 & 0 \\ 0 & 0 & -\frac{3\sqrt{3}}{2} & -9/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1/2 & -\frac{\sqrt{3}}{2} \\ 0 & 0 & 0 & 0 & -\frac{3\sqrt{3}}{2} & -9/2 \end{pmatrix}.$$

Let us give an idea of the computational savings the method provides for the simple case that all linear problems are solved via a direct solver. Not exploiting the symmetry results in CN^3k^3 flops. On the other hand, using the symmetry reduction method leads to an overhead of 3 applications of the generalized finite Fourier transform, i.e., $3N^2k^2$ flops, and additional flops for solving the various subproblems, i.e.,

$$Ck^3 \sum_{\rho \in \mathcal{R}(\Gamma)} (\dim \rho)^3.$$

Neglecting the overhead gives a reduction factor of

$$\frac{1}{N^3} \sum_{\rho \in \mathcal{R}(\Gamma)} (\dim \rho)^3.$$

For example, if the full symmetry group of the three-dimensional cube is considered, then the reduction factor equals approximately 0.00116.

Further advantages are that the reduced problems remain sparse if the original problem is sparse. Furthermore, the diagonal blocks have at most the condition number of the large problem, and often the condition numbers are significantly reduced.

Another advantage of the symmetry reduction method is that it greatly facilitates the study of bifurcation phenomena, since the bifurcation structures of the reduced problems are much simpler, owing again to the fact that the multiple eigenvalues of the large problem have become distributed over the various subproblems. See Section 24.3 for more details.

16 Derivatives

An integral part of continuation methods involves the calculation of the Jacobian of a vector valued function, and occasionally also higher derivatives. Let us sketch in this section some ideas for implementing such derivatives.

Of course, the update methods described in Section 14 can be viewed as a method to numerically approximate the Jacobian H'. Since this is explicitly incorporated into a path following algorithm, we have treated these ideas separately in a special section.

A very special case of numerical differentiation (of the implicitly defined solution curve) is discussed in Section 11. Let us just note in passing that automatic differentiation (which will be discussed next) does not handle implicit function definitions.

An obvious way to obtain derivatives is to resort to a symbolic software such as MAPLE or MATHEMATICA. However, especially for higher dimensions and more complicated equations $H(u) = 0$, these methods become cumbersome, and the exact analytic expressions for the derivatives may become unmanagable.

16.1 Automatic Differentiation

Recently, a new topic in computer science has evolved dealing with automatic differentiation, see, e.g., the surveys of Griewank [1989] and Iri and Kubota [1987] or the proceedings edited by Corliss and Griewank and Corliss [1991]. This is a technique which is reasonable to consider integrating into a continuation code. In fact, this has been done for HOMPACK in connection with systems of equations arising in the analysis of integrated circuits, see Melville, Trajkovic, Fang, and Watson [1993b] and Melville, Moinian, Feldmann, and Watson [1993a]. The authors report that simple timing studies show that for each Jacobian matrix evaluated at a point of the solution curve, considerably more time is spent in HOMPACK performing linear algebra computations on the matrix than is spent in constructing the derivatives.

Let us sketch the basic idea of automatic differentiation. Assume that a function $f(x_1, \ldots, x_n)$ can be expressed by means of a code of the following form:

Algorithm 16.1 **(Function)**

INPUT $x_1, \ldots, x_n$
OUTPUT $y = f(x_1, \ldots, x_n)$
FOR $i = n+1, \ldots, m$

$x_i \leftarrow f_i\left(\{x_j\}_{j\in I_i}\right)$
END FOR
$y \leftarrow x_m$

The variables $x_{n+1}, \ldots, x_{m-1}$ store intermediate results and can be discarded when they are no longer needed. The functions f_i depend explicitly on the quantities $\{x_j\}_{j\in I_i}$, which have already been computed, hence $j < i$ if $j \in I_i$. Moreover, the f_i are assumed to represent basic computational steps in the sense that f_i is one of the basic operations $\{+, -, *, /\}$ possibly with one constant operand as, e.g., $x_1 + 3$, or else an elementary function such as $\{\sin, \cos, \exp, \log, \ldots\}$. Then the partial derivatives $\partial_j f_i = \frac{\partial f_i}{\partial x_j}$ are available.

For example, consider $y = f(x_1, x_2) = (x_1 + 2x_2)\sin(x_1 x_2)$ which gives

$$\begin{aligned} x_3 &\leftarrow 2x_2 \\ x_4 &\leftarrow x_1 + x_3 \\ x_5 &\leftarrow x_1 x_2 \\ x_6 &\leftarrow \sin x_5 \\ x_7 &\leftarrow x_4 x_6 \end{aligned}$$

A straightforward way of obtaining the gradient of f would be to differentiate all the ingredients of Algorithm 16.1.

ALGORITHM 16.2 **(Function and Gradient, Forward Mode)**

INPUT $x_1, \ldots, x_n$
OUTPUT $y = f(x_1, \ldots, x_n)$, $g = \nabla f(x_1, \ldots, x_n)$
FOR $i = 1, \ldots, n$
$\nabla x_i \leftarrow e_i$ % *initialization with unit base vectors*
END FOR
FOR $i = n+1, \ldots, m$
$x_i \leftarrow f_i\left(\{x_j\}_{j\in I_i}\right)$
$\nabla x_i \leftarrow \sum_{j\in I_i} \partial_j f_i\left(\{x_j\}_{j\in I_i}\right) \nabla x_j$
END FOR
$y \leftarrow x_m$, $g \leftarrow \nabla x_m$

This is the so-called forward mode of automatic differentiation. It has the advantage of being easy to implement, but it increases the computational effort of evaluating the gradient compared to evaluating the original function (Algorithm 16.1) by a factor $\mathcal{O}(n)$.

For our example we would obtain the following sequence:

$$
\begin{aligned}
\nabla x_3 &\leftarrow 2\nabla x_2 = (0,2)^* \\
\nabla x_4 &\leftarrow \nabla x_1 + \nabla x_3 = (1,0)^* + (0,2)^* = (1,2)^* \\
\nabla x_5 &\leftarrow x_2\nabla x_1 + x_1\nabla x_2 = (x_2, x_2)^* \\
\nabla x_6 &\leftarrow \cos x_5 \nabla x_5 = \cos x_5 (x_2, x_2)^* \\
\nabla x_7 &\leftarrow x_6\nabla x_4 + x_4\nabla x_6 = (x_6 + x_2 x_4 \cos x_5, 2x_6 + x_1 x_4 \cos x_5)^*
\end{aligned}
$$

Let us now outline the reverse mode of automatic differentiation. For this purpose, we view the output y with varying dependencies, namely

$$y = h_i(x_1, \ldots, x_i) \quad \text{for} \quad i = m, m-1, \ldots, n\,,$$

where h_i denotes the result of eliminating the variables in Algorithm 16.1 in reverse order. In our example:

$$
\begin{aligned}
h_7 &= x_7 \\
h_6 &= x_6 x_4 \\
h_5 &= x_4 \sin x_5 \\
h_4 &= x_4 \sin(x_1 x_2) \\
h_3 &= (x_1 + x_3) \sin(x_1 x_2) \\
h_2 &= (x_1 + 2x_2) \sin(x_1 x_2).
\end{aligned}
$$

Note that

$$h_{i-1}(x_1, \ldots, x_{i-1}) = h_i(x_1, \ldots, x_{i-1}, f_i\left(\{x_j\}_{j \in I_i}\right),$$

and therefore the chain rule gives

$$\partial_j h_{i-1} = \partial_j h_i + (\partial_i h_i)(\partial_j f_i)$$

for $j = 1, \ldots, i-1$. Of course, the second term vanishes for $j \notin I_i$. This leads to the following algorithm:

ALGORITHM 16.3 **(Function and Gradient, Reverse Mode)**

INPUT $x_1, \ldots, x_n$
OUTPUT $y = f(x_1, \ldots, x_n)$, $g = \nabla f(x_1, \ldots, x_n)$
FOR $i = n+1, \ldots, m$
    $x_i \leftarrow f_i\left(\{x_j\}_{j \in I_i}\right)$
END FOR

$y \leftarrow x_m$ % *first output*
FOR $j = 1, \ldots, m$
$\partial_j h_m = \delta_{mj}$ % *initialization*
END FOR
FOR $i = m, m-1, \ldots, n+1$
FOR $j = 1, \ldots, i-1$

$$\partial_j h_{i-1} \leftarrow \begin{cases} \partial_j h_i + \partial_j f_i\left(\{x_j\}_{j \in I_i}\right) \partial_i h_i & j \in I_i \\ \partial_j h_i & \text{otherwise} \end{cases}$$

END FOR
END FOR
$g \leftarrow (\partial_1 h_n, \ldots, \partial_n h_n)^*$ % *second output*

The starting quantities $\{\partial_j h_m\}$ can be overwritten by the $\{\partial_j h_{i-1}\}$ such that only the elements with $j \in I_i$ have to be updated in the i-loop. So the computational cost of the reverse mode is reduced by the factor n compared to the forward mode. Indeed, it can be shown that the computational effort of evaluating the gradient compared to evaluating the original function is multiplied only by a number $C \leq 5$ (which is independent of n). On the other hand, the values $\{x_i\}$ from the forward sweep and the information about the arguments of f_i have to be saved for the reverse sweep which, in general, causes high additional bookkeeping costs. In fact, the decision of what mode should be used depends on the class of functions, especially on the relation between n and m, and on the specific implementation.

In our example, the calculation of

$$(\partial_1 h_i, \ldots, \partial_i h_i)$$

for $i = 7, 6, \ldots, 2$ in reverse mode takes the form:

i	$\partial_1 h_i$	$\partial_2 h_i$	$\partial_3 h_i$	$\partial_4 h_i$	$\partial_5 h_i$	$\partial_6 h_i$	$\partial_7 h_i$
7	0	0	0	0	0	0	1
6	0	0	0	x_6	0	x_4	
5	0	0	0	x_6	$x_4 \cos x_5$		
4	$x_2 x_4 \cos x_5$	$x_1 x_4 \cos x_5$	0	x_6			
3	$x_6 + x_2 x_4 \cos x_5$	$x_1 x_4 \cos x_5$	x_6				
2	$x_6 + x_2 x_4 \cos x_5$	$2x_6 + x_1 x_4 \cos x_5$					

Of course, the end result is the same as for the forward mode.

Forward evaluations of higher order derivatives can be performed similarly to Algorithm 16.1. The reverse mode is not so simple. However, often in applications only an action of a higher derivative on vectors $v_1, v_2, \ldots, v_p$

is needed, and these can be obtained via successive applications of the gradient, i.e.,

$$\begin{aligned}\nabla^{1+p} f(x) v_1 v_2 \dots v_p &= \nabla\big(\nabla^p f(x) v_1 v_2 \dots v_p\big) \\ &= v_p^* \Big(\nabla\big(\nabla^p f(x) v_1 v_2 \dots v_{p-1}\big)\Big).\end{aligned}$$

The above ideas could be implemented by a user whenever implementing a function subroutine. However, though the implementation is straightforward, typical users are likely to find this onerous. Therefore, in recent years, extensive efforts have been made to let a pre-compiler handle the extensions, i.e., a user only needs to furnish a function in some standard way (say, in a C^{++} or FORTRAN function subroutine), and a precompiler automatically extends the subroutine to include either a forward or a reverse mode calculation of the derivatives. For details and an extensive bibliography on the subject including available software such as, e.g., ADOL-C or ADIFOR, we refer to Corliss and Griewank [1991].

16.2 Finite Difference Approximations

A straightforward finite difference approximation of the Jacobian H' would be quite expensive ($N + 1$ evaluations of H even for the coarse forward difference formula). However, in some cases, it is only necessary to calculate an action $H'(u)p$ of H' on a vector p. To evaluate this action, we may use a central difference approach

$$H'(u)p = \frac{H(u + \varepsilon p) - H(u - \varepsilon p)}{2\varepsilon} + \mathcal{O}(\varepsilon^2). \tag{16.1}$$

This formula involves only two evaluations of H. Decreasing ε reduces the truncation error, but unfortunately increases the cancellation error.

For $||H(u)|| = \mathcal{O}(1)$ a balance is reached at $\varepsilon \approx \sqrt[3]{\delta}$ where δ indicates the unit round-off of the computer. This is seen from the following argument. The cancellation error can be estimated by $\mathcal{O}(\delta/\varepsilon)$. Setting the cancellation error equal to the truncation error leads to $\mathcal{O}(\delta/\varepsilon) = \mathcal{O}(\varepsilon^2)$ from which $\varepsilon = \mathcal{O}(\sqrt[3]{\delta})$ is readily obtained.

The resulting precision is approximately $\mathcal{O}(\delta^{2/3})$ which means that about 2/3 of the significant digits are valid in the approximation of the directional derivative (16.1). If this precision is still unsatisfactory, a higher order formula can be employed which is easily obtained from (16.1) via Richardson extrapolation, since the truncation error in (16.1) has an expansion in powers of ε^2. If, for example, a formula of order $2k$ is employed,

then $2k$ evaluations of H are necessary, a balanced stepsize for the finite difference formula would lead to $\varepsilon \approx \delta^{\frac{1}{2k+1}}$, and the resulting precision would mean that about $\frac{2k}{2k+1}$ of the significant digits are valid.

A similar analysis can be done for the second derivative using the central difference approximations:

$$\begin{aligned} H''(u)[p,p] &= \frac{H(u+\varepsilon p)+H(u-\varepsilon p)-2H(u)}{\varepsilon^2}+\mathcal{O}(\varepsilon^2), \qquad (16.2)\\ H''(u)[p,q] &= \frac{1}{4}\big(H''(u)[p+q,p+q]-H''[p-q,p-q]\big)\\ &= \frac{1}{4}\frac{H(u+\varepsilon p+\varepsilon q)+H(u-\varepsilon p-\varepsilon q)}{\varepsilon^2}\\ &\quad -\frac{1}{4}\frac{H(u+\varepsilon p-\varepsilon q)+H(u-\varepsilon p+\varepsilon q)}{\varepsilon^2}+\mathcal{O}(\varepsilon^2). \end{aligned}$$

The truncation errors are again expandable in terms of powers of ε^2, and the corresponding statements about choosing the adequate stepsize ε and about Richardson's extrapolation are the same as those above for the first derivative. There is only one change: the cancellation error can now be estimated by $\mathcal{O}(\delta/\varepsilon^2)$, from which $\varepsilon = \mathcal{O}(\sqrt[4]{\delta})$ is obtained, see, e.g., MACKENS [1989, p. 248].

17 Large Scale Problems

One of the primary applications of continuation methods involves the numerical solution of nonlinear eigenvalue problems. Such problems are likely to have arisen from a discretization of an operator equation in a Banach space context, which also involves an additional bifurcation parameter. For some specific examples see Section 25. As a result of the discretization and the wish to maintain a reasonably low truncation error, the corresponding finite dimensional problem $H(u) = 0$ where $H : \mathbb{R}^{N+1} \to \mathbb{R}^N$, may require that N be quite large. This then leads to the task of solving large scale continuation problems.

An area in which a considerable amount of experience concerning large scale continuation methods exists is structural mechanics, see, e.g., RHEINBOLDT [1981, 1986] and the further references cited therein. There has also been work done on combining continuation methods with multigrid methods for solving large scale continuation problems arising from discretization of elliptic problems via finite differences, see, e.g., CHAN and KELLER [1982], BANK and CHAN [1986], MITTELMANN and ROOSE [1990], and further literature cited therein. Another area where large scale continuation problems have been treated concerns finite element discretizations of elliptic

problems, which are then combined with a conjugate gradient solver in the continuation algorithm, see GLOWINSKI, KELLER, and REINHART [1985].

It is clear that an endless variety of combinations can be made of continuation algorithms and sparse solvers. In view of this, following equation (12.7) in Section 12 we have discussed how any sparse solver process can be incorporated into the general scheme of continuation methods which we have been describing. Here we will indicate more specifically how to incorporate a conjugate gradient method or a Krylov subspace method. The following discussion is still at an early stage of development.

Let us recall the nonlinear conjugate gradient method of POLAK and RIBIÈRE [1969]. This choice of nonlinear minimizer is based upon reports that in numerical practice it has generally yielded the best results, see, e.g., POWELL [1977] or BERTSEKAS [1984]. To outline the method, let us assume that the problem to be solved is

$$\min_u \{\varphi(u) : u \in \mathbb{R}^{N+1}\} \tag{17.1}$$

where $\varphi : \mathbb{R}^{N+1} \to \mathbb{R}$ is a smooth nonlinear functional, usually having an isolated local minimal point $\bar{u}$ which we desire to approximate.

The following is an outline of the conjugate gradient method due to POLAK and RIBIÈRE [1969].

ALGORITHM 17.1 **(Nonlinear CG)**

INPUT: $u_0 \in \mathbb{R}^{N+1}$ % *initial point*
OUTPUT: $u_0, u_1, \ldots$ % *converging to* $\bar{u}$ *(hopefully)*
$g_0 \leftarrow \nabla\varphi(u_0)$; $d_0 \leftarrow g_0$; % *initial gradients*
FOR $n = 0, 1, \ldots$
$\rho_n \leftarrow \arg\min_{\rho>0}\varphi(u_n - \rho d_n)$; % *line search*
$u_{n+1} \leftarrow u_n - \rho_n d_n$
$g_{n+1} \leftarrow \nabla\varphi(u_{n+1})$
$\gamma_n \leftarrow \dfrac{[(g_{n+1} - g_n)^* g_{n+1}]}{||g_n||^2}$
$d_{n+1} \leftarrow g_n + \gamma_n d_n$ % *new conjugate gradient*
END FOR

Since our aim here is merely to make an application of a conjugate gradient method, we will not give a detailed account concerning conjugate gradient methods for nonlinear problems. However, we shall recall some of their properties. For more details we suggest, e.g., the book of FLETCHER [1987].

The main theoretical justification of the (nonlinear) conjugate gradient algorithm lies in its properties when $\varphi(u)$ is a uniformly convex quadratic functional. In this special case the algorithm becomes theoretically equivalent to the familiar (linear) conjugate gradient method due to HESTENES and STIEFEL [1952]. Furthermore, in this case the steplength ρ_n in Algorithm 17.1 is readily calculated to be

$$\rho_n = \frac{\varphi'(u_n)d_n}{d_n^* \nabla\varphi'(u_n)d_n} \tag{17.2}$$

where $\nabla\varphi'$ denotes the Hessian of φ.

The following two statements summarize the main convergence results on conjugate gradient methods, see GOLUB and VAN LOAN [1989, Sec. 10.2.8] or STOER [1983, Sec. 1].

Let φ be a uniformly convex quadratic form. If the Hessian $\nabla\varphi'$ has exactly k distinct eigenvalues, then the conjugate gradient method stops at the solution $\bar{u}$ after k steps.

Let φ be a uniformly convex quadratic form. If κ denotes the condition of the Hessian $\nabla\varphi'$, then

$$||u_n - \bar{u}|| \leq \frac{||u_0 - \bar{u}||}{T_n\left(\frac{\kappa+1}{\kappa-1}\right)} \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n ||u_0 - \bar{u}||, \tag{17.3}$$

where $T_n(x) = \cos(n \arccos(x))$ is the n-th Chebyshev polynomial.

The conclusion to be drawn from the above results is that initially, the convergence of the conjugate gradient method may be slow because of (17.3), but by the k-th step a very substantial improvement in the approximation of the solution has been obtained. This appears to hold even in the general case where the functional φ is no longer quadratic. To be more specific, let us assume that $\bar{u}$ is a local minimal solution point of the problem (17.1) at which the Hessian $\nabla\varphi'(\bar{u})$ is positive definite. There are several results concerning the convergence of the conjugate gradient method which essentially state that local superlinear convergence towards $\bar{u}$ holds, see, e.g., COHEN [1972] or MCCORMICK and RITTER [1974]. However it appears that as of this date, the convergence results are somewhat unsatisfactory. One of the difficulties is that there are various possibilities for obtaining the factors γ_n in Algorithm 17.1. Another difficulty is that in practice, we do not want to perform a very precise one-dimensional minimization in the line search of Algorithm 17.1 in order to obtain an acceptable ρ_n since this is costly. Most of the convergence rate proofs require cyclic reloading, i.e., setting $\gamma_n = 0$ after every $N + 1$ steps. The general idea of such proofs involves the approximation of $\varphi(u)$ via Taylor's formula

by

$$\varphi(u) \approx \varphi(\bar{u}) + \varphi'(\bar{u})(u-\bar{u}) + (u-\bar{u})^* \nabla\varphi'(\bar{u})(u-\bar{u}),$$

and then to use the convergence results for the quadratic case. Actually, even in the quadratic case, because of the presence of rounding errors, we cannot expect that stopping will occur after k steps. Instead, we should regard the conjugate gradient method even in this case as an iterative method which makes a substantial improvement after k steps.

The ideal convergence (one step!) would occur when the condition number $\kappa = 1$, i.e., when all the eigenvalues of $\nabla\varphi'$ are equal. Intuitively, the next best situation would occur when the eigenvalues have as few cluster points as possible. This observation may be used to motivate the idea of preconditioning for the conjugate gradient method. For more details (in case of a quadratic functional) the reader may refer to GOLUB and VAN LOAN [1989, Sec. 10.3].

Let us now turn to the case which actually concerns us, namely minimizing the functional

$$\varphi(u) := \frac{1}{2}||H(u)||^2. \tag{17.4}$$

The minimal points of (17.4) obviously form the solution curve $c(s)$ which we are considering throughout the paper. We have

$$\begin{aligned} \nabla\varphi(u) &= H'(u)^* H(u); \\ \nabla\varphi'(u) &= H'(u)^* H'(u) + \mathcal{O}\big(||H(u)||\big). \end{aligned}$$

The gradient $\nabla\varphi(u) = H'(u)^* H(u)$ is orthogonal to the tangent vector $t\big(H'(u)\big)$. This motivates the idea for implementing the conjugate gradient method (Algorithm 17.1) as a corrector into a continuation method. Analogously to the case when the minimization problem has isolated solutions at which the Hessian is positive definite, we may also expect local superlinear convergence of the conjugate gradient method for the functional (17.4).

The solution will be a point $\bar{u} \in H^{-1}(0)$ which is essentially nearest to a predictor point v which is taken as the starting point for the (nonlinear) conjugate gradient method. Numerical experience suggests that local superlinear convergence holds, as in the case of functionals φ with an isolated minimal point. However, to our knowledge even in the latter case a general proof of this has not been given. We propose the above conjugate gradient method as a reasonable corrector procedure nevertheless, provided once again, that an effective preconditioning is incorporated. In the present context, we propose a preconditioning of the dependent variables, which we now describe in more detail.

Let us consider the following transformation of φ:

$$\tilde{\varphi}(u) = \frac{1}{2}||L^{-1}H(u)||^2$$

where L is an as yet unspecified nonsingular (N, N)-matrix. Then we have

$$\begin{aligned} \nabla\tilde{\varphi}(u) &= H'(u)^*(LL^*)^{-1}H(u); \\ \nabla\tilde{\varphi}'(u) &= H'(u)^*(LL^*)^{-1}H'(u) + \mathcal{O}\big(||H(u)||\big). \end{aligned} \tag{17.5}$$

If we assume that our continuation method furnishes predictor points which are already near $H^{-1}(0)$, we may neglect the $\mathcal{O}\big(||H(u)||\big)$ term in (17.5).

Furthermore, if $H(u) \approx 0$, then an ideal choice would be to take L such that

$$LL^* = H'(u)H'(u)^*$$

is the Cholesky decomposition. We then have

$$\begin{aligned} \nabla\tilde{\varphi}(u) &= H'(u)^*(LL^*)^{-1}H(u) \\ &= H'(u)^*(H'(u)H'(u)^*)^{-1}H(u) \\ &= H'(u)^+H(u). \end{aligned}$$

Hence in this case, the gradient $\nabla\tilde{\varphi}(u) = H'(u)^+H(u)$ coincides with the usual Newton direction which we have discussed as the standard corrector.

Of course, if we really want to use the Cholesky decomposition, then we would relinquish whatever advantage sparseness may have offered. Thus, we want to determine L also with a small computational expense and in such a way that linear equations $LL^*x = y$ are cheaply solved for x. This leads naturally to a recommendation of using an incomplete Cholesky factorization, see, e.g., GOLUB and VAN LOAN [1989, Sec. 10.3.2].

The paper of GLOWINSKI, KELLER, and REINHART [1985] treats nonlinear elliptic eigenvalue problems via a combining of finite element discretization, formulating a pseudo-arclength continuation method via an augmented system of equations and applying the Polak-Ribière nonlinear conjugate gradient method to a least squares formulation of the augmented systems of equations. In this situation the minimization problem generally has isolated solutions.

In the paper by ALLGOWER, CHIEN, and GEORG [1989] a version of a continuation algorithm incorporating a nonlinear conjugate gradient method as a corrector is given. Applications are made to a nonlinear elliptic boundary value problem on a variety of domains having symmetries.

Several of the recently developed large scale continuation methods deal with discretizations of nonlinear elliptic eigenvalue problems. In these cases,

usually the $N \times N$-submatrix of $H'(u)$ obtained by deleting the column corresponding to the eigenvalue parameter is positive definite or even diagonally dominant. Let us stress here that Algorithm 17.2 does not make use of any such property and is meant to apply to more general situations. This greater versatility may be obtained at the cost of a greater computational effort.

We now outline an algorithm which incorporates a nonlinear conjugate gradient method as corrector.

ALGORITHM 17.2 **(CG Predictor Corrector Method)**

INPUT $u_0 \in \mathbb{R}^{N+1}$ % *such that* $H(u_0) \approx 0$ *(initial point)*
$t \in \mathbb{R}^{N+1}$ % *initial approximation to* $t\big(H'(u_0)\big)$
$h > 0$ % *initial steplength*
OUTPUT u_i , $i = 0, 1, 2, \ldots$ % *approximating the solution curve*
FOR $i = 1, 2, \ldots$
$v \leftarrow u_{i-1} + ht$ % *predictor step*
calculate $LL^* \approx H'(v)H'(v)^*$ % *preconditioner*
$g_v \leftarrow H'(v)^*(LL^*)^{-1}H(v), \quad d \leftarrow g_v$ % *gradients*
REPEAT % *corrector loop*
$\bar{\rho} \leftarrow \arg\min_{\rho \geq 0} ||L^{-1}H(v - \rho d)||^2$
$w \leftarrow v - \bar{\rho} d$ % *corrector step, nonlinear CG*
$g_w \leftarrow H'(w)^*(LL^*)^{-1}H(w)$ % *new gradient*
$\gamma \leftarrow \dfrac{(g_w - g_v)^* g_w}{||g_v||^2}$ % *Polak-Ribière*
$d \leftarrow g_w + \gamma d$ % *new conjugate gradient*
$v \leftarrow w, \quad g_v \leftarrow g_w$
UNTIL convergence
adapt steplength $h > 0$
$t \leftarrow \dfrac{u_{i-1} - w}{||u_{i-1} - w||}$ % *approximation to* $t\big(H'(w)\big)$
$u_i \leftarrow w$ % *new point approximately on* $H^{-1}(0)$
END FOR

Let us make some remarks concerning details and modifications of the above algorithm. First of all, if the evaluation of $H'(w)$ is very costly, one may prefer to hold it fixed in the corrector loop. Furthermore, let us mention several possibilities for solving the line search problem

$$\min_{\rho \geq 0} ||L^{-1}H(v - \rho d)||^2. \tag{17.6}$$

Recalling (17.2), let us approximate the functional $\tilde{\varphi}(v - \rho d)$, which is to be minimized, by its truncated Taylor expansion:

$$\tilde{\varphi}(v) - \rho\tilde{\varphi}'(v)d + \frac{1}{2}\rho^2 d^* \nabla\tilde{\varphi}(v)'d.$$

This is minimized exactly when

$$\bar{\rho} = \frac{\tilde{\varphi}'(v)d}{d^*\nabla\tilde{\varphi}(v)'d} \tag{17.7}$$

provided $d^*\nabla\tilde{\varphi}(v)'d > 0$. For the case at hand, this is a reasonable assumption, because of (17.9) and the fact that the conjugate gradient d is chosen essentially orthogonal to $\ker H'(v)$. Furthermore, for

$$\tilde{\varphi}(v - \rho d) = \frac{1}{2}||L^{-1}H(v - \rho d)||^2$$

we have

$$\tilde{\varphi}'(v)d = H(v)^*(LL^*)^{-1}H'(v)d = g_v^* d \tag{17.8}$$

and

$$\begin{aligned} d^*\nabla\tilde{\varphi}(v)'d &= d^*H'(v)^*(LL^*)^{-1}H'(v)d + \mathcal{O}(\,||H(v)||\;||d||\,) \\ &\approx ||L^{-1}H'(v)d||^2. \end{aligned} \tag{17.9}$$

Since the evaluation of $H'(v)d$ may be costly for large scale problems, an inexpensive approximation of $H'(v)d$ may be made by using the central difference formula

$$H'(v)d = \frac{H(v + \varepsilon\frac{d}{||d||}) - H(v - \varepsilon\frac{d}{||d||})}{2\varepsilon}||d|| + \mathcal{O}(\varepsilon^2) \tag{17.10}$$

for an appropriate discretization step $\varepsilon > 0$, see Section 16.2. Usually, the predictor corrector steps of a continuation method are performed in such a way that all generated points are close to the solution curve in $H^{-1}(0)$. Hence, the quadratic approximation considered in (17.7) will give good results in the situation which we are presently considering. Thus we recommend substituting (17.8)–(17.10) into (17.7).

A second possibility for solving (17.6) is to merely use a standard line search algorithm which does not require the evaluation of $\nabla\tilde{\varphi}$ such as a quadratic fit or golden section algorithm, see, e.g., FLETCHER [1987]. The disadvantage of this approach is that it may require many evaluations of $L^{-1}H$.

Recently, some classes of generalized conjugate direction methods or Krylov subspace methods have been developed to solve systems of linear

equations $Mx = b$ where the matrix M is not necessarily assumed to be positive definite or even symmetric, see, e.g., the survey FREUND, GOLUB, and NACHTIGAL [1992]. The generalized minimal residual algorithm of SAAD and SCHULTZ [1986] seems to be of particular interest in our context, since it only uses multiplications by M. If we take

$$M = \begin{pmatrix} H'(u) \\ t^* \end{pmatrix},$$

where t is some suitable approximation of $t\big(H'(u)\big)$, e.g., given by a secant, then it is easy to program a multiplication Mx. In fact, the multiplication $H'(u)x$ may be approximated by a forward or central difference formula for the directional derivative as in (17.10), see also Section 16.2, so that one multiplication by M essentially involves one scalar product and one or two evaluations of the map H. Note that for this choice of M the solution s of the equation

$$Ms = \begin{pmatrix} H(u) \\ 0 \end{pmatrix}$$

is the Gauss-Newton increment $s = H'(u)^+ H(u)$ which is used in our path following algorithms.

An early survey of the incorporation of fast solvers into continuation codes has been given by CHAN [1984c]. Some implementations and applications are discussed in ALLGOWER, CHIEN, GEORG, and WANG [1991d]. DESA, IRANI, RIBBENS, WATSON, and WALKER [1992] have incorporated Craig's variant of the conjugate gradient method and the SYMMLQ algorithm for symmetric indefinite problems into a homotopy program. However, one may prefer an incorporation of a transpose-free method, see FREUND, GOLUB, and NACHTIGAL [1992, Sec. 3.4], since we are not aware of any cheap finite difference approximation of the transpose action $H'(u)^* x$.

Chapter V

Applications

In this chapter we present a selection of examples where path following methods have been used. Many more specific examples exist in the literature. Our discussion of applications concentrates to a large extent on cases in which the predictor corrector methods apply. Applications in which the dimension is relatively low but smoothness does not hold can be handled by the piecewise linear methods discussed in Chapter VI.

18 Sard's Theorem

One frequent application of path following methods involves homotopy methods. These methods are often resorted to in zero point or fixed point problems when no suitable starting point is available for an iterative method such as Newton's method, or when contraction methods do not apply. In these situations, one constructs a homotopy map which deforms from a map which is trivial (or at least its solution points are known) to the map of interest. By the implicit function theorem, a curve emanates from each of the trivial starting points. These curves are then numerically traced in the expectation that a solution point of the map of interest will be reached. In many applications of the numerical homotopy methods, it is possible to avoid degeneracies in the solution curve by introducing suitable parameters (perturbations). The theoretical basis of this approach lies in Sard's theorem for maps with additional (perturbation) parameters, see, e.g., Abraham and Robbin [1967] or Hirsch [1976]. We consider the following general form:

Theorem 18.1 (Sard) *Let A, B, C be smooth manifolds of finite dimensions with* $\dim A \geq \dim C$, *and let $F : A \times B \to C$ be a smooth map.*

Assume that $c \in C$ is a regular value of F, i.e., for $F(a,b) = c$ we have that the total derivative $F'(a,b) : T_aA \times T_bB \to T_cC$ has maximal rank. Here T_aA denotes the tangent space of A at a, etc. Then for almost all $b \in B$ (in the sense of some Lebesgue measure on B) the restricted map $F(\cdot, b) : A \to C$ has c as a regular value.

The global version of the Implicit Function Theorem, see Theorem 1.1, now implies that for such b and c the set $\{a \in A : F(a,b) = c\}$ is a smooth manifold of dimension $\dim A - \dim C$. In particular, for the case $\dim A = \dim C + 1$, this perturbation technique is a basic tool in homotopy continuation methods to guarantee the existence of a smooth solution curve (without singularities) which is then traced numerically. The parameter set B is typically considered to consist of perturbation parameters in this case, and a choice of b as in the above theorem is considered to be a *generic* choice.

19 Fixed Point Problems

To illustrate the use of Sard's theorem, let us consider a homotopy arising from a fixed point problem. Let $f : \mathbb{R}^N \to \mathbb{R}^N$ be a smooth map which is bounded. According to the theorem of Brouwer [1912], the map f has at least one fixed point. To simplify the discussion, let us make the assumption that the map $x \mapsto x - f(x)$ has zero as a regular value. This implies that the fixed points of f are isolated, and that Newton's method converges locally. However, the global convergence of Newton's method is by no means guaranteed.

We therefore consider the homotopy

$$H(x, \lambda, p) = x - p - \lambda(f(x) - p). \tag{19.1}$$

For the *trivial level* $\lambda = 0$, we obtain the *trivial map* $H(x, 0, p) = x - p$ which has the unique zero point p, our *starting point*. On the *target level* $\lambda = 1$, we obtain the *target map* $H(x, 1, p) = x - f(x)$ whose zero points are our points of interest, i.e., the fixed points of f.

Let us illustrate via this example how Sard's theorem is typically employed: The Jacobian of H is given by

$$H'(x, \lambda, p) = (\mathrm{Id} - \lambda f'(x) \ , \ p - f(x) \ , \ (\lambda - 1)\mathrm{Id}).$$

Due to our assumptions, the first N columns of the Jacobian are linearly independent for $H(x, \lambda, p) = 0$ and $\lambda = 1$, and clearly the last N columns are linearly independent for $\lambda \neq 1$. Consequently, by Sard's theorem we can

conclude that for almost all $p \in \mathbb{R}^N$ zero is a regular value of the restricted map $H(\cdot, \cdot, p)$.

For such a generic choice of p, the solution manifold $H(\cdot, \cdot, p)^{-1}(0)$ consists of smooth curves which are either diffeomorphic to the circle or to the real line, see Theorem 1.3. Consider the solution curve $c(s) = (x(s), \lambda(s))$ (parametrized for convenience with respect to arclength) such that $c(0) = (p, 0)$. It is easy to see that the initial tangent vector in the direction of increasing λ has the form

$$\dot{c}(0) = \left(1 + ||f(p) - p||^2\right)^{-\frac{1}{2}} \begin{pmatrix} f(p) - p \\ 1 \end{pmatrix},$$

and hence the curve is not tangent to the plane $\lambda = 0$.

Since the solution point $(p, 0)$ is unique for $\lambda = 0$, it follows that c cannot be closed and hence is diffeomorphic to the real line. Furthermore, the boundedness of f implies that $x(s)$ is bounded for $0 \leq \lambda(s) \leq 1$.

Let us assume (for contradiction) that $\lambda(s) < 1$ for all $s > 0$. By the above boundedness, the theorem of Bolzano-Weierstrass implies that there is a sequence $s_i \to \infty$ and a $\bar{u} \in \mathbb{R}^{N+1}$ such that $c(s_i) \to \bar{u}$. It follows that $H(\bar{u}, p) = 0$, and by the Implicit Function Theorem the connected component $\mathcal{C} \subset \{u \in \mathbb{R}^{N+1} : H(u, p) = 0\}$ containing c also has to contain $\bar{u}$. It follows that for each i there is a parameter value $\bar{s}_i > s_i$ such that $c(\bar{s}_i) = \bar{u}$. This implies that the curve c is closed, a contradiction.

It follows that the curve c reaches the level $\lambda = 1$ after a finite arclength s_0, i.e., $c(s_0) = (x_0, 1)$, and hence x_0 is a fixed point of f which can be approximated by tracing the curve c, see Figure 19.1.

Let us note that

$$(\mathrm{Id} - f'(x_0))\dot{x}(s_0) = \dot{\lambda}(s_0)(f(x_0) - p),$$

and our above assumption on f implies that $(\mathrm{Id} - f'(x_0))$ cannot have a nontrivial kernel, and hence $\dot{\lambda}(s_0) \neq 0$, i.e., the curve c is tranversal to the level $\lambda = 1$ at any solution.

The above discussion is in the spirit of Chow, Mallet-Paret, and Yorke [1978]. An earlier approach based on the nonretraction principle of Hirsch [1963] was given by Kellogg, Li, and Yorke [1976]. Since the appearence of the constructive proofs of the Brouwer fixed point theorem many other constructive existence proofs have been described. Further references may be found in Allgower and Georg [1990, Sec. 11.1].

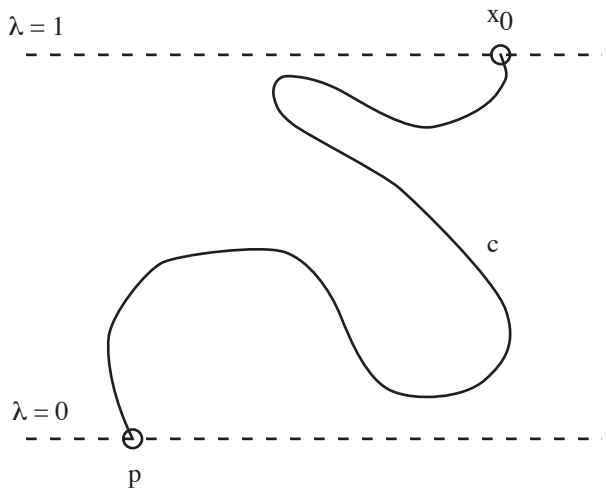

Figure 19.1. Homotopy Fixed Point Method

20 Global Newton Methods

Newton's method is a popular method for numerically calculating a zero point of a smooth map $G : \mathbb{R}^N \to \mathbb{R}^N$. As is well known, this method may diverge if the starting point p is not sufficiently near to a zero point $\bar{x}$ of G. Often one would like to determine whether a certain open bounded region $\Omega \subset \mathbb{R}^N$ contains a zero point $\bar{x}$ of G and furthermore, for which starting values p this solution $\bar{x}$ can be obtained by Newton's method. The so-called global Newton methods offer a possibility of answering such questions.

One may interpret Newton's method as the numerical integration of the differential equation

$$\dot{x} = -G'(x)^{-1}G(x) \tag{20.1}$$

using Euler's method with unit step size. The idea of using the above flow to find zero points of G was exploited by Branin [1972]. Smale [1976] gave conditions on $\partial\Omega$ under which the flow leads to a zero point of G in Ω. Such numerical methods have been referred to as *global Newton methods*. However, we note that this method becomes numerically unstable near singularities, i.e., at points x where the condition number of $G'(x)$ is large. Keller [1978] observed that the above flow can also be obtained in a *numerically stable* way from a homotopy equation which he consequently named the *global homotopy* method. Independently, Garcia and Gould

[1978, 1980] discussed this flow.

We briefly sketch Keller's approach. The global homotopy method involves tracing the curve defined by the equation $G(x) - (1-\lambda)G(p) = 0$ starting from $(x,\lambda) = (p,0) \in \partial\Omega \times \{0\}$ inward into $\Omega \times \mathbb{R}$. If the level $\Omega \times \{1\}$ is encountered, then a zero point of G has been found. Hence, the *global homotopy* $H : \mathbb{R}^N \times \mathbb{R} \times \partial\Omega \longrightarrow \mathbb{R}^N$ is defined by

$$H(x,\lambda,p) := G(x) - (1-\lambda)G(p). \tag{20.2}$$

The following theorem is essentially due to SMALE [1976].

THEOREM 20.1 *Let the following conditions be satisfied for the smooth map* $G : \mathbb{R}^N \to \mathbb{R}^N$*:*

1. $\Omega \subset \mathbb{R}^N$ *is open and bounded and* $\partial\Omega$ *is a connected smooth submanifold of* $\mathbb{R}^N$*,*
2. *zero is a regular value of* G*,*
3. $G(q) \neq 0$ *for* $q \in \partial\Omega$*,*
4. *the Jacobian* $G'(q)$ *is nonsingular for* $q \in \partial\Omega$*,*
5. *the Newton direction* $-G'(q)^{-1}G(q)$ *is not tangent to* $\partial\Omega$ *at* q *for* $q \in \partial\Omega$*.*

Then for almost all initial points $p \in \partial\Omega$*, there is a smooth curve* $c(s) = (x(s),\lambda(s))$ *emanating from* $c(0) := (p,0)$*. Here we assume for simplicity that* s *is arclength. The derivative* $\dot{x}(0)$ *is not tangent to* $\partial\Omega$*, and hence we orient* c *in such a way that* $\dot{x}(0)$ *points into* Ω*. Furthermore,* c *hits the target level* $\Omega \times \{1\}$ *at an odd number of points* $x(s_1),\ldots,x(s_k)$ *with* $s_1 < \ldots < s_k$*.*

This possibility of obtaining more than one solution was first observed by BRANIN and HOO [1972].

PROOF We give a modified version of KELLER's [1978] proof. Since p varies over the $(N-1)$-dimensional surface $\partial\Omega$, it is somewhat difficult to apply Sard's theorem. This task was achieved by PERCELL [1980] who showed that for almost all $p \in \partial\Omega$ the global homotopy H has 0 as a regular value. Let p be such a generic choice.

We consider the solution curve $c(s) = (x(s),\lambda(s))$ in $H(\cdot,\cdot,p)^{-1}(0)$ such that $c(0) = (p,0)$ and such that s represents arclength, see Figure 20.1.

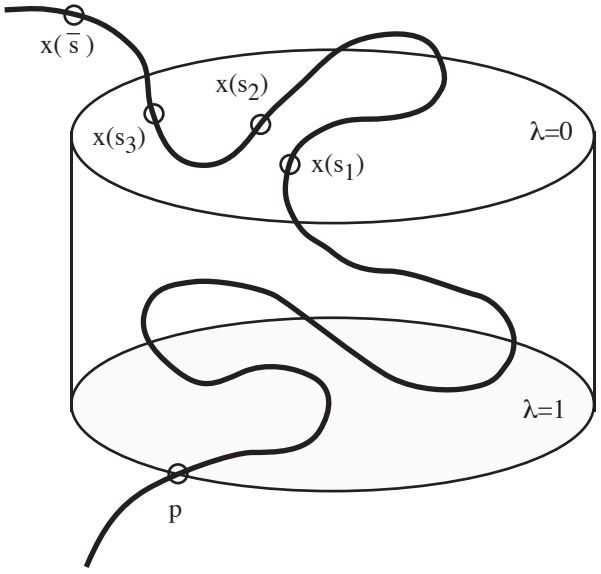

Figure 20.1. Illustration of Global Homotopy

Since $\partial\Omega$ is connected, we can infer from assumption 5 without loss of generality that the Newton direction

$$G'(q)^{-1}G(q) \quad \text{always points out of} \quad \Omega \quad \text{for} \quad q \in \partial\Omega. \tag{20.3}$$

The other case, i.e., that the Newton direction $G'(q)^{-1}G(q)$ always points into Ω for $q \in \partial\Omega$, is treated in a similar way by switching some signs.

We differentiate the homotopy equation

$$G\big(x(s)\big) - (1 - \lambda(s))G(p) = 0 \tag{20.4}$$

and obtain

$$G'\big(x(s)\big)\dot{x}(s) + \dot{\lambda}(s)G(p) = 0, \tag{20.5}$$

and by substituting $(1 - \lambda(s))^{-1}G\big(x(s)\big)$ for $G(p)$ we obtain

$$\lambda(s) \neq 1 \Rightarrow \; G'\big(x(s)\big)\,\dot{x}(s) + \frac{\dot{\lambda}(s)}{1 - \lambda(s)}\,G\big(x(s)\big) = 0. \tag{20.6}$$

If $\dot{\lambda}(s) = 0$, then $||\dot{x}(s)|| = 1$ since s is arclength, and the above equation would imply that $G'(x(s))$ is singular. Hence, from assumption 4 it follows

that $\dot\lambda(s) \neq 0$ for $x(s) \in \partial\Omega$. Furthermore, assumption 3 implies that $\lambda(s) \neq 1$ for $x(s) \in \partial\Omega$. And finally, from (20.3) and the above equation, it now follows that $\dot x(s)$ points into Ω or out of Ω for $x(s) \in \partial\Omega$, according as the factor $\frac{\dot\lambda(s)}{1-\lambda(s)}$ is positive or negative.

In particular, initially, we choose the orientation of c in such a way that

$$\dot\lambda(0) > 0, \tag{20.7}$$

and hence the above considerations imply that $\dot x(0)$ points into Ω.

Since $G(p) \neq 0$ and Ω is bounded, we see that the set

$$\{\lambda : G(x) = (1-\lambda)G(p),\ x \in \overline{\Omega}\}$$

is bounded. Hence the curve c must exit from $\Omega \times \mathbb{R}$ at some parameter value $\bar s > 0$, i.e., $x(\bar s) \in \partial\Omega$. Since $x(\bar s)$ points out of Ω, it follows that

$$\frac{\dot\lambda(\bar s)}{1-\lambda(\bar s)} < 0. \tag{20.8}$$

Now consider the augmented Jacobian

$$A(s) := \begin{pmatrix} G'\big(x(s)\big) & G(p) \\ \dot x(s)^* & \dot\lambda(s) \end{pmatrix}$$

of the homotopy (20.2). We obtain

$$A(s) \begin{pmatrix} \mathrm{Id} & \dot x(s) \\ 0^* & \dot\lambda(s) \end{pmatrix} = \begin{pmatrix} G'\big(x(s)\big) & 0 \\ \dot x(s)^* & 1 \end{pmatrix}$$

and consequently

$$\det A(s)\, \dot\lambda(s) = \det G'\big(x(s)\big). \tag{20.9}$$

By assumption 4, and since $\partial\Omega$ is connected, the function $\det G'(x)$ does not change sign on $\partial\Omega$. On the other hand, the function $\det A(s)$ does not change sign along the path c. Consequently, $\mathrm{sign}\dot\lambda(0) = \mathrm{sign}\dot\lambda(\bar s)$, and hence (20.7) implies $\dot\lambda(\bar s) > 0$. From (20.8) we obtain

$$\lambda(\bar s) > 1, \tag{20.10}$$

and hence there exists at least one s_1 with $0 < s_1 < \bar s$ such that $\lambda(s_1) = 1$.

Let $\lambda(s_i) = 1$ for $0 < s_i < \bar s$. If $\dot\lambda(s_i) = 0$, then $G(x(s_i)) = 0$ by (20.4), and (20.5) would imply $G'\big(x(s)\big)\dot x(s) = 0$ with $||\dot x(s)|| = 1$, contradictory to assumption 2. Hence, $\dot\lambda(s_i) \neq 0$.

Since $\lambda(0) = 0$ and $\lambda(\bar s) > 1$, it now follows that there are an odd number of values $0 < s_1 < \ldots < s_k < \bar s$ such that $\lambda(s_i) = 1$ for $i = 1, \ldots, k$. □

If $\tilde{x}$ is a zero point of G such that $G'(\tilde{x})$ is nonsingular, then the assumptions of Theorem 20.1 hold for a sufficiently small ball Ω around $\tilde{x}$. Thus, in a certain sense the global homotopy extends the well-known Newton-Kantorovich type theorems concerning the local convergence of Newton's method, see, e.g., ORTEGA and RHEINBOLDT [1970].

20.1 Some Applications to Chemical Engineering and Circuit Analysis

The principal application considered in the papers by BRANIN [1972] and BRANIN and HOO [1972] was in the area of electrical circuit design and analysis. Since then homotopy continuation methods have frequently been employed as a tool in the computer aided design of electronic circuits and computer chips. The paper of CHAO and SAEKS [1977] surveys continuation methods in circuit analysis and cites many papers on the topic up to 1977. Recently, MELVILLE, TRAJKOVIC, FANG, and WATSON [1993b] have applied a homotopy continuation algorithm to compute multiple solutions of systems of equations arising in the simulation of integrated circuits, see also MCQUAIN, RIBBENS, WATSON, and MELVILLE [1993] and MELVILLE, MOINIAN, FELDMANN, and WATSON [1993a]. The paper MELVILLE, TRAJKOVIC, FANG, and WATSON [1993b] also contains an updated survey of the literature in which continuation methods are used in circuit analysis. Watson, the principal author of HOMPACK, see WATSON, BILLUPS, and MORGAN [1987], reports that homotopy methods are on the verge of replacing damped Newton methods in the circuit simulation programs of at least one of the major industrial producers of integrated circuits. In addition, homotopy methods have become a standard tool in the numerical simulation of robotics in several major industrial firms. The latter application deals primarily with polynomial systems and so we will discuss this application further in Section 22.

The fundamental nonlinear term arising in the network problems, is of the form $\alpha(e^{v/\beta}-1)$ where α and β are constants and v denotes voltage. In CHUA and USHIDA [1976] an example is given of a system of four nonlinear equations modelling a four-transistor multi-state circuit, which after some simplifications assumes the following general form:

$$A\begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} + B\begin{pmatrix} e^{v_1/\beta} \\ \vdots \\ e^{v_N/\beta} \end{pmatrix} + \begin{pmatrix} c_1 \\ \vdots \\ c_N \end{pmatrix} = 0$$

where N is the number of nodes of the circuit, A, B are (N, N)-matrices, the c_i are constants and the unknown v_i are voltages.

In general, A and B are large sparse nonsymmetric matrices. The example of this form which was numerically treated in CHUA and USHIDA [1976] (for $N = 4$) essentially by means of the global Newton method was found to have nine solutions.

In addition to the application of the predictor corrector homotopy methods in the area of circuit analysis, there is also a body of work in which piecewise linear methods have been applied. Such works can be traced back at least to the paper of KATZENELSON [1965]. We will discuss this topic further in Chapter VI. In any event, the application of continuation methods in the area of electronic circuit analysis seems currently to be a busy scene of activity.

Another active area in which path following methods are being currently used concerns the numerical simulation of chemical processes. The paper of SEIDER, BRENGEL, and WIDAGDO [1991] reviews the use of nonlinear analysis in chemical process design and includes a comprehensive bibliography concerning the use of continuation methods in the field of chemical engineering. The paper of BYRNE and BAIRD [1985] appears to be one of the earliest in which a continuation method was applied to solve difficult distillation problems where the determination of a suitable starting point for a Newton type method presents considerable trouble. The paper by WAYBURN and SEADER [1987] gives a general discussion outlining how and when homotopy continuation might be applied to chemical engineering problems. WAYBURN [1982] reviews the literature on the application of continuation methods to separation problems up to 1988. The paper of KOVACH and SEIDER [1987] presents an example of a distillation system in which five steady state solutions were found via a homotopy continuation where previously only three solutions had been found by Newton type methods. The paper of CHAVEZ, SEADER, and WAYBURN [1986] demonstrated via homotopy continuation the presence of multiple solutions in interlinked separation systems.

It should perhaps be noted that in the application of homotopy methods there is in general no distinction made as to the stability of the solutions which are found. Thus a user may still need to do some analysis to sort out the unstable solutions among the solutions which are found by the homotopy method. In the next section we will see that successive solutions which lie on a homotopy curve are necessarily of opposite topological index, see Theorem 21.3. Hence in the case of critical points, a homotopy method cannot yield successive minima or successive maxima. This suggests that in some cases successive solutions found on a continuation curve may not both be stable.

21 Multiple Solutions

In the previous section it was observed that the global homotopy method might in fact yield more than one zero point of the map G in a bounded region Ω. This raises the question whether one might be able to compute more zero points of G in Ω in addition to those which lie on the global homotopy path. To be more precise, let us suppose that $\Omega \subset \mathbb{R}^N$ is an open bounded region, and that $G : \mathbb{R}^N \to \mathbb{R}^N$ is a smooth map having a zero point $z_0 \in \Omega$. The task is now to find additional zero points of G in Ω, provided they exist. One method which has often been used for handling this problem is deflation, see, e.g., BROWN and GEARHART [1971]. In this method a *deflated map* $G_1 : \mathbb{R}^N \setminus \{z_0\} \to \mathbb{R}^N$ is defined by

$$G_1(x) = G(x)/||x - z_0||. \tag{21.1}$$

One then applies an iterative method to try to find a zero point of G_1. Numerical experience with deflation has shown that it is often a matter of seeming chance whether one obtains an additional solution and if one is obtained, it is very often not the one which is nearest to z_0.

By utilizing homotopy-type methods we can give some conditions which will guarantee the existence of an additional solution and yield insights into the behavior of deflation. This additional solution will lie on a homotopy path. We illustrate this approach with a discussion of the *d-homotopy*. Let us consider the homotopy map $H_d : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}$ defined by

$$H_d(x, \lambda) := G(x) - \lambda d$$

where $d \in \mathbb{R}^N$ is some fixed vector with $d \neq 0$. Since we assume that a zero point z_0 is already given, we have $H_d(z_0, 0) = 0$. Let us further assume zero is a regular value of G. Then it follows from Sard's theorem that zero is also a regular value of H_d for almost all $d \in \mathbb{R}^N$. In order to assure that the solution curve c in $H_d^{-1}(0)$ which contains $(z_0, 0)$ again reaches the level $\lambda = 0$, we need to impose a boundary condition. The following theorem uses a boundary condition which is motivated by a simple degree consideration.

THEOREM 21.1 *Let the following hypotheses hold:*

1. $G : \mathbb{R}^N \to \mathbb{R}^N$ *is a smooth map with zero as a regular value;*
2. $d \in \mathbb{R}^N \setminus \{0\}$ *is a point such that the homotopy* H_d *also has zero as a regular value;*
3. $\Omega \subset \mathbb{R}^N$ *is a bounded open set which contains a (known) initial zero point* z_0 *of* G*;*

4. *the* boundary condition $H_d(x,\lambda) = G(x) - \lambda d \neq 0$ *holds for all* $x \in \partial\Omega$, $\lambda \in \mathbb{R}$.

Then the curve c *in* $H_d^{-1}(0)$ *which contains* $(z_0, 0)$ *intersects the level* $\Omega\times\{0\}$ *an even number of times at points* $(z_i, 0)$, $i = 0, \ldots, n$, *at which* $G(z_i) = 0$.

PROOF The boundary condition (item 4) implies that the curve c lies strictly inside the cylinder $\partial\Omega \times \mathbb{R}$. A point (x, λ) on the curve c satisfies $G(x) = \lambda d$ and $x \in \Omega$, and hence $|\lambda| = ||G(x)||/||d||$ remains bounded. Recalling that c is homeomorphic either to the line $\mathbb{R}$ or to the circle S^1, the boundedness of c implies that $c \simeq S^1$. Since 0 is a regular value of G, it is easily seen that c intersects the level $\Omega \times \{0\}$ transversely, and the assertion follows immediately. □

The boundary condition (item 4) can be relaxed:

THEOREM 21.2 *The conclusion of the above theorem remains true if the boundary condition (item 4) is replaced by either of*

4-1. $H_d(x,\lambda) = G(x) - \lambda d \neq 0$ *for all* $x \in \partial\Omega$, $\lambda \geq 0$;

4-2. $H_d(x,\lambda) = G(x) - \lambda d \neq 0$ *for all* $x \in \partial\Omega$, $\lambda \leq 0$.

PROOF We consider only the case 4-1. If (x, λ) on c is a solution with $\lambda \geq 0$, the same argument as in the above proof shows that $\lambda = ||G(x)||/||d||$ remains bounded. Hence, starting at a solution point and traversing the curve initially in the positive λ-direction gives the desired assertion. □

To continue our discussion, we now introduce the following definition: Let z_0 be a zero point of the smooth map $G : \mathbb{R}^N \to \mathbb{R}^N$ such that $\det G'(z_0) \neq 0$. Then the *index* of z_0 is defined to be the sign of $\det G'(z_0)$.

THEOREM 21.3 *Under the hypotheses of Theorem 21.1 or 21.2, any two zero points of* G *which are consecutively obtained by traversing the curve* c *have opposite index.*

PROOF Let $(x(s)\,,\,\lambda(s)) = c(s)$ be parametrized according to arclength. Differentiating

$$H_d\big(x(s),\lambda(s)\big) = G\big(x(s)\big) - \lambda(s)d = 0$$

we obtain

$$G'\big(x(s)\big)\dot{x}(s) - \dot{\lambda}(s)d = 0$$

and hence

$$\begin{pmatrix} G'\big(x(s)\big) & -d \\ \dot{x}(s)^* & \dot{\lambda}(s) \end{pmatrix} \begin{pmatrix} \mathrm{Id} & \dot{x}(s) \\ 0^* & \dot{\lambda}(s) \end{pmatrix} = \begin{pmatrix} G'\big(x(s)\big) & 0 \\ \dot{x}(s)^* & 1 \end{pmatrix}.$$

We note that the left matrix in the latter equation is the augmented Jacobian of H_d. Its determinant is of constant sign $\varepsilon \in \{-1, 1\}$ on c. Hence by the product rule of determinants we obtain that

$$\operatorname{sign}\dot{\lambda}(s) \cdot \operatorname{sign}\det G'\big(x(s)\big) = \varepsilon.$$

Let $\big(x(s_1), 0\big)$ and $\big(x(s_2), 0\big)$ be two consecutive solution points on c. It is clear that they are traversed in opposite λ-directions, i.e., $\dot{\lambda}(s_1)\dot{\lambda}(s_2) < 0$ and the assertion follows. □

The above theorems enable us to make conclusions concerning the relationship between deflation and homotopy methods, see ALLGOWER and GEORG [1983b] for a more detailed discussion. In the context of the above discussion of deflation, see Equation (21.1), let us consider the global homotopy $H : \big(\mathbb{R}^N \setminus \{z_0\}\big) \times \mathbb{R} \to \mathbb{R}^N$ defined by

$$H(x, \lambda) := G_1(x) - \lambda G_1(x_0). \tag{21.2}$$

In view of our discussions following Equation (20.1), performing Newton's method on G_1 starting at x_0 amounts to a particular integration of the global homotopy (21.2) and also starting at $(x_0, 1)$. Thus we see that in general successive deflation will at best produce the zeros of G which lie on $H^{-1}(0)$. However, because the Newton steps represent an Euler method for integrating Equation (20.1) with possibly large steps, some iterate may get far enough away from $H^{-1}(0)$ that the Newton method might diverge or possibly accidently converge to a zero point not on $H^{-1}(0)$. Numerical experience, see BROWN and GEARHART [1971], with deflation confirms the above analysis in the sense that, in general, zero points which are successively obtained via deflation have opposite index, and zero points even in close proximity are not successively obtained by deflation if they have the same index.

In a recent paper LAGARIAS, MELVILLE, and GEOGHEGAN [1993] make the observation based on the proof of Theorem 21.3 that $\deg G \in \{0, 1, -1\}$ is a necessary condition for the solution curve of a homotopy H to pass through all of the zero points of G, provided that zero is a regular value of G. Conversely, as a corollary of the Whitney Lemma (see, e.g., the paper of JEZIERSKI [1993]), if $\deg G \in \{0, 1, -1\}$, and if $N \neq 2$, then there exists a homotopy H of the same smoothness class as G for which the solution

curve passes through all of the zero points of G. Except for the trivial case $N = 1$, there does not seem to be any explicit construction of the desired homotopy available. The authors discuss some applications to the DC operating point problem for nonlinear circuits (see also Section 20.1). For certain types of circuits a coercivity condition can be verified from which the condition $\deg G \in \{0, 1, -1\}$ can in turn be concluded.

21.1 A Nonlinear Boundary Value Problem

To give an illustration of how Theorem 21.1 can be applied, we consider a system of equations arising from a discretization of a nonlinear elliptic boundary value problem

$$\begin{aligned} \mathcal{L}u(\xi) &= \mu f(u(\xi)), & \xi \in \mathcal{D}\,; \\ u(\xi) &= 0, & \xi \in \partial\mathcal{D}\,. \end{aligned}$$

Here $\mathcal{D} \subset \mathbb{R}^m$ is a bounded domain, $\mathcal{L}$ is a linear elliptic differential operator, and f is a smooth nonlinear function which is bounded from below and satisfies

$$\lim_{u\to\infty} \frac{f(u)}{u} = \infty.$$

Problems of this general form are discussed in the survey paper of Amann [1976], the problem from which our particular example derives has been discussed by Ambrosetti and Hess [1980].

Discretizations of the above problem generally take the form

$$G(x) := Ax - \mu F(x) = 0 \tag{21.3}$$

where A is a positive definite (N, N)-matrix such that A^{-1} has only positive entries, and $F : \mathbb{R}^N \to \mathbb{R}^N$ is a smooth map whose co-ordinates are bounded from below by a constant $-C < 0$ and satisfy

$$\lim_{x[i]\to\infty} \frac{F(x)[i]}{x[i]} = \infty$$

for each co-ordinate $x[i]$ of the vector x.

From the contraction principle it follows that for small $\mu > 0$ the fixed point iteration

$$x_{k+1} = \mu A^{-1} F(x_k), \quad x_0 = 0$$

converges to a zero point z_0 of G. We choose a directional vector $d \in \mathbb{R}^N$ whose co-ordinates are all positive, and a set

$$\Omega := \{x \in \mathbb{R}^N : ||x||_\infty < \beta\}$$

where $\beta > 0$ is chosen so large that the following conditions are satisfied:

$$\begin{aligned} \beta &> \mu C||A^{-1}||_\infty\,; \\ F(x)[i] &> \frac{\beta}{\mu a} \quad \text{for } x[i] = \beta \text{ and all } i\,. \end{aligned}$$

In the latter inequality $a > 0$ denotes the smallest entry of A^{-1}. Both inequalities can be satisfied, because of the assumptions on F.

Let us show that for the above choices the assumptions of Theorem 21.2 and in particular the boundary condition 4-1 are satisfied. If $x \in \partial\Omega$, then either $x[i] = -\beta$ or $x[i] = \beta$ for some co-ordinate i. In the first case we estimate

$$\begin{aligned} A^{-1}\big(G(x) - \lambda d\big)[i] &= -\beta - \mu A^{-1}F(x)[i] - \lambda A^{-1}d[i] \\ &\leq -\beta + \mu C||A^{-1}||_\infty < 0. \end{aligned}$$

In the second case we obtain

$$\begin{aligned} A^{-1}\big(G(x) - \lambda d\big)[i] &= \beta - \mu A^{-1}F(x)[i] - \lambda A^{-1}d[i] \\ &\leq \beta - \mu a F(x)[i] < 0. \end{aligned}$$

Hence, in both cases we have $G(x) - \lambda d \neq 0$ for $\lambda \geq 0$.

This application of the d-homotopy makes it possible to reach an additional solution z_1 via the homotopy equation

$$Ax - \mu F(x) - \lambda d = 0$$

for a fixed μ and varying λ. We emphasize that z_0, z_1 do not necessarily lie on the same solution branch of the equation

$$Ax - \mu F(x) = 0$$

for varying μ. Hence, the d-homotopy can permit moving between disjoint solution branches of the nonlinear eigenvalue problem (21.3).

21.2 Periodic Solutions of a Duffing Equation

Recently a constructive proof of the Poincaré-Birkhoff theorem has been given by Li and Lin [1994]. Their proof is in the spirit of the methods discussed in this chapter. The statement of the theorem is as follows: Let A denote an annular region in $\mathbb{R}^2 - \{0\}$, whose inner boundary Γ_1 and outer boundary Γ_2 are two disjoint simple closed curves. Let D_i denote the open region bounded by Γ_i , $i = 1, 2$. Then $A = \bar{D}_2 - D_1$ and $0 \in D_1 \subset D_2$.

THEOREM 21.4 (Poincaré-Birkhoff Theorem) *Let $T : A \to T(A) \subset \mathbb{R}^2 - \{0\}$ be an area preserving homeomorphism. Suppose:*

(i) $T : (r, \theta) \mapsto (r^*, \theta^*)$ has the polar co-ordinate representation $r^* = f(r, \theta), \theta^* = \theta + g(r, \theta)$ such that $g(r, \theta) > 0$ on Γ_1 and $g(r, \theta) < 0$ on Γ_2, where f and g are $\mathcal{C}^2$ and 2π periodic in θ.

(ii) There exists a continuous area preserving map $T_1 : \bar{D}_2 \to \mathbb{R}^2$ such that $T_1|_A = T$ and $0 \in T_1(D_1)$.

Then T has at least two fixed points in A.

The Poincaré-Birkhoff theorem has been applied on many ocassions in the study of dynamical systems and periodic solutions of Duffing's equation. Let us briefly show how the Poincaré-Birkhoff Theorem together with the homotopy method is applied to calculate periodic solutions of a Duffing equation. We consider an equation of the form

$$x'' + F(x, t) = 0 \tag{21.4}$$

where $F : \mathbb{R}^2 \to \mathbb{R}$ is a continuous function which is twice continuously differentiable and 2π periodic in the second variable. Equation (21.4) has the equivalent form

$$x' = y, \; y' = -F(x, t). \tag{21.5}$$

Let $(x(t, x_0, y_0), y(t, x_0, y_0))$ denote the unique solution of (21.5) for the initial value $(x(0), y(0)) = (x_0, y_0)$. Then the Poincaré map $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T(x_0, y_0) = (x(2\pi, x_0, y_0), y(2\pi, x_0, y_0))$$

is an area preserving homeomorphism (see JACOBOWITZ [1976]). More precisely, T satisfies the assumptions of the Theorem 21.4.

Upon setting $x = r\cos\theta$, $y = r\sin\theta$, (21.5) assumes the form

$$\begin{aligned} r' &= r\sin\theta\cos\theta - \sin\theta \; F(r\cos\theta, t), \\ \theta' &= -\sin^2\theta - \frac{1}{r}\cos\theta \; F(r\cos\theta, t). \end{aligned} \tag{21.6}$$

Denote by $(r(t, r_0, \theta_0), \theta(t, r_0, \theta_0))$ the unique solution of (21.6) with the initial value $(r(0), \theta(0)) = (r_0, \theta_0)$. Using the notation of Theorem 21.4, we have

$$\begin{aligned} r(2\pi, r_0, \theta_0) &= f(r_0, \theta_0), \\ \theta(2\pi, r_0, \theta_0) &= g(r_0, \theta_0) + \theta_0 \pmod{2\pi}. \end{aligned}$$

To compute the periodic solutions of (21.6), we may fix a value $\theta_0 \in [0, 2\pi)$ and find a point r_0 near the set $\{r : g(r, \theta_0) = 0\}$ by, e.g., a Newton-type method. By Theorem 18.1 (Sard), the probability is unity that (r_0, θ_0)

is such a point that zero is a regular value of the global homotopy

$$H(r,\theta,\lambda) := \begin{pmatrix} f(r,\theta) - r \\ g(r,\theta) \end{pmatrix} - (1-\lambda) \begin{pmatrix} f(r_0,\theta_0) - r_0 \\ 0 \end{pmatrix} = 0. \quad (21.7)$$

When $\lambda = 0$, the solution of $H(r,\theta,\lambda) = 0$ is $(r_0,\theta_0,0)$ and when $\lambda \to 1$, any limit point (r^*,θ^*) of the solutions of (21.7) satisfies $f(r^*,\theta^*) = r^*$, $g(r^*,\theta^*) = \theta^* \pmod{2\pi}$.

Li and Lin [1994] show (in the general context of Theorem 21.4) that the solution curve of (21.7), when started from the initial point $(r_0,\theta_0,0)$ in the two possible directions leads to two different solutions at $\lambda = 1$.

For actual numerical implementations the function values $f(r,\theta), g(r,\theta)$ can be calculated via an initial value solving routine, such as a Runge-Kutta method.

22 Polynomial Systems

In the preceding section we considered the task of computing multiple zero points of general smooth maps. In the case of complex polynomial systems it is possible to compute (at least in principle) all of the zero points by means of homotopy methods. This subject has received considerable attention in recent years. The book of Morgan [1987] deals extensively with this topic, in connection with the path following approach. It also contains a number of interesting applications to robotics and other fields.

We consider a system of complex polynomials $P : \mathbb{C}^N \to \mathbb{C}^N$. The task is to find *all* solutions of the equation $P(z) = 0$. Let us review some of the standard terminology in this context. If a term of the k-th component P_k of P has the form

$$a z_1^{r_1} z_2^{r_2} \cdots z_N^{r_N} ,$$

then its degree is $r_1 + r_2 + \ldots + r_N$. The degree d_k of P_k is the maximum of the degrees of its terms. The *homogeneous part* $\hat{P}$ of P is obtained by deleting in each component P_k all terms having degree less than d_k. The *homogenization* $\tilde{P}$ of P is obtained by multiplying each term of each component P_k with an appropriate power z_0^r such that its degree is d_k. Note that the homogenization $\tilde{P} : \mathbb{C}^{N+1} \to \mathbb{C}^N$ involves one more variable z_0. If

$$(w_0, \ldots, w_N) \neq 0$$

is a zero point of $\tilde{P}$, then the entire ray

$$[w_0 : \cdots : w_N] := \{(\xi w_0, \ldots, \xi w_N) \mid \xi \in \mathbb{C}\}$$

consists of zero points of $\tilde{P}$. Usually, $[w_0 : \cdots : w_N]$ is regarded as a point in the complex projective space $\mathbf{CP}^N$. There are two cases to consider:

1. The solution $[w_0 : \cdots : w_N]$ intersects the hyperplane $z_0 = 0$ transversely, i.e., $w_0 \neq 0$, and hence, without loss of generality, we may take $w_0 = 1$. This corresponds to a zero point $(w_1, \ldots, w_N)$ of P. Conversely, each zero point $(w_1, \ldots, w_N)$ of P corresponds to a solution $[1 : w_1 : \cdots : w_N]$ of $\tilde{P}$.

2. The solution $[w_0 : \cdots : w_N]$ lies in the hyperplane $z_0 = 0$, i.e., $w_0 = 0$. This corresponds to a *nontrivial* solution $[w_1 : \cdots : w_N]$ of the homogeneous part $\hat{P}$, and such solutions are called *zero points of P at infinity*.

As in the case of one variable, it is possible to define the multiplicity of a solution. However, this is a more complicated matter than it is in one dimension and requires some deeper ideas of algebra and analysis. We will give a brief sketch. If $[w_0 : \cdots : w_N]$ is an isolated solution of the homogenization $\tilde{P}(z_0, \ldots, z_N) = 0$ with respect to the topology of $\mathbf{CP}^N$, we can define a multiplicity of $[w_0 : \cdots : w_N]$ in two different ways. However, it is not a trivial exercise to show that these definitions are equivalent.

(1) Consider a co-ordinate w_k which is different from zero. Without loss of generality we can assume $w_k = 1$. If we fix the variable $z_k = 1$ in the the homogenization $\tilde{P}(z_0, \ldots, z_N) = 0$, then we have N complex equations in N complex variables or $2n$ real equations in $2n$ real variables with the complex solution $z_j = w_j$, $j = 0, \ldots, N$, $j \neq k$. The multiplicity is now defined by the local topological degree of this solution, see, e.g., MILNOR [1968]. It can be shown that this definition is independent of the special choice of the non-vanishing co-ordinate w_k.

(2) As above, consider a co-ordinate $w_k = 1$. Again, we fix the variable $z_k = 1$ in the homogenization $\tilde{P}(z_0, \ldots, z_N) = 0$, and after a translation obtain new equations

$$F(z_0, \ldots, \widehat{z_k}, \ldots, z_N) := \tilde{P}(z_0 + w_0, \ldots, 1, \ldots, z_N + w_N) = 0$$

in the variables $z_0, \ldots, \widehat{z_k}, \ldots, z_N$ where $\widehat{\ }$ denotes omission of the term beneath it. These new equations have a zero point at the origin. Now, the multiplicity of the solution is defined as the dimension of the quotient

$$\frac{\mathbb{C}\big[[z_0, \ldots, \widehat{z_k}, \ldots, z_N]\big]}{(F_1, \ldots, \widehat{F_k}, \ldots, F_N)}$$

where $\mathbb{C}\big[[z_0,\dots,\widehat{z_k},\dots,z_N]\big]$ is the usual power series ring and the symbol $(F_1,\dots,\widehat{F_k},\dots,F_N)$ denotes the ideal generated by the corresponding polynomials $F_1,\dots,\widehat{F_k},\dots,F_N$, see, e.g., FULTON [1984] or VAN DER WAERDEN [1953]. It can be shown that also this definition is independent of the special choice of the non-vanishing co-ordinate w_k.

The higher dimensional analogue of the fundamental theorem of algebra is Bezout's theorem, which states that the number of zero points of P (counting their multiplicities and zeros at infinity) equals the product $d = d_1 d_2 \cdots d_N$, provided all solutions are isolated. This number d is often called the *Bezout number* of the system.

As an illustration, let us examine the polynomial system

$$\begin{aligned} z_1^3 - z_1 &= 0, \\ z_1^2 z_2 + 1 &= 0, \end{aligned}$$

which has the homogeneous part

$$\begin{aligned} z_1^3 &= 0, \\ z_1^2 z_2 &= 0, \end{aligned}$$

and the homogenization

$$\begin{aligned} z_1^3 - z_0^2 z_1 &= 0, \\ z_1^2 z_2 + z_0^3 &= 0. \end{aligned}$$

It has the three isolated solutions $[1:\pm 1:-1]$ and $[0:0:1]$. According to Bezout's theorem, this system has nine roots. It is routine to see that $\det P'(z) \neq 0$ at the zero points $(\pm 1,-1)$. Hence they are simple, and by Bezout's theorem the zero point at infinity has multiplicity seven. Let us show that this is true by using the second definition of multiplicity. Setting $z_2 = 1$ in the homogenization, we obtain

$$\begin{aligned} z_1^3 - z_0^2 z_1 &= 0, \\ z_1^2 + z_0^3 &= 0. \end{aligned}$$

Since $z_1^3 - z_0^2 z_1 = z_1(z_1+z_0)(z_1-z_0)$, a factorization theorem yields

$$\begin{aligned} \dim \frac{\mathbb{C}\big[[z_0,z_1]\big]}{(z_1^3 - z_0^2 z_1\,,\, z_1^2 + z_0^3)} &= \dim \frac{\mathbb{C}\big[[z_0,z_1]\big]}{(z_1\,,\, z_1^2 + z_0^3)} + \dim \frac{\mathbb{C}\big[[z_0,z_1]\big]}{(z_1+z_0\,,\, z_1^2 + z_0^3)} \\ &\qquad + \dim \frac{\mathbb{C}\big[[z_0,z_1]\big]}{(z_1-z_0\,,\, z_1^2 + z_0^3)} \\ &= 3 + 2 + 2. \end{aligned}$$

GARCIA and ZANGWILL [1979] and CHOW, MALLET-PARET, and YORKE [1979] introduced homotopy methods in $\mathbf{C}^N \times \mathbb{R}$ for finding all solutions of the equation $P = 0$. A difficulty they had to overcome was the handling of solution paths which connect to zero points at infinity. WRIGHT [1985] realized that their approaches could be simplified by going into the complex projective space $\mathbf{CP}^N$. We use his approach to illustrate the homotopy idea for polynomial systems.

Define a homotopy $H = (H_1, \ldots, H_N)$ by involving the homogenization $\tilde{P}$ of P via

$$H_k(z_0, \ldots, z_N, \lambda) = (1-\lambda)\big(a_k z_k^{d_k} - b_k z_0^{d_k}\big) + \lambda \tilde{P}_k(z_0, \ldots, z_N).$$

Wright shows by Sard-type arguments that for almost all coefficients a_k, b_k in $\mathbb{C}$ the restricted homotopies $H^{(j)}$ which are obtained from H by fixing $z_j = 1$ for $j = 0, \ldots, N$ have zero as a regular value for $\lambda < 1$. Given such a generic choice of coefficients, he concludes that for $\lambda < 1$, the homogeneous system of polynomials H has exactly d simple zero point curves $c_i(\lambda) \in \mathbf{CP}^N$, $i = 1, \ldots, d$, in complex projective N-space. On the trivial level $\lambda = 0$, the d solutions are obvious, and it is possible to trace the d curves emanating from these solutions into the direction of increasing λ. The solution curves are monotone in λ, and hence all have to reach the target level $\lambda = 1$ on the compact manifold $\mathbf{CP}^N$. Thus, in this approach solutions at infinity are treated no differently than finite solutions. The solution curves are traced in the projective space $\mathbf{CP}^N$, and from the numerical point of view we have the slight drawback that occasionally a chart in $\mathbf{CP}^N$ has to be switched.

Recently, attention has been given to the task of trying to formulate homotopies which eliminate the sometimes wasteful effort involved in tracing paths which go to solutions of $P(z_1, \ldots, z_N) = 0$ at infinity. Work in this direction has been done in MORGAN [1986], SAUER, and YORKE [1987] LI, SAUER, and YORKE [1989] and LI and WANG [1993a, 1993b]. MORGAN and SOMMESE [1987] describe the easily implemented "projective transformation" which allows the user to avoid the drawback of changing co-ordinate charts on $\mathbf{CP}^N$. MORGAN and SOMMESE [1989] show how to exploit relations among the system coefficients, via "coefficient parameter continuation". Such relations occur commonly in engineering problems, as described in WAMPLER and MORGAN [1991], WAMPLER, MORGAN, and SOMMESE [1990, 1992]. The papers MORGAN, SOMMESE, and WAMPLER [1991, 1992, 1993] combine a homotopy method with contour integrals to calculate singular solutions to polynomial and nonlinear analytic systems. MORGAN, SOMMESE, and WATSON [1989] documented that `HOMPACK`, see WATSON, BILLUPS, and MORGAN [1987], in the case of polynomial systems

has some stability issues that CONSOL8, see MORGAN [1987], does not have. The path following approach to systems of polynomial equations is particularly suited for parallel processing, see ALLISON, HARIMOTO, and WATSON [1989].

23 Sparse Polynomial Systems

In many applications, the Bezout number of a system of polynomial equations is prohibitively high, so that the path following codes mentioned in the previous section run into serious problems because of the many paths that need to be followed. Recently, a new approach has been initiated to solve sparse polynomial systems via path following, see HUBER and STURMFELS [1993]. This approach looks very promising and is based on Bernstein's theorem, see Theorem 23.1 below. The method is too technical even to sketch, but let us try to at least give a few hints. Related

Let us denote by $x^a := x_1^{a_1} \cdots x_N^{a_N}$ a monomial in N variables having integer exponents $a = (a_1, \ldots, a_N) \in \mathbb{Z}^N$. A sparse system $f = (f_1, \ldots, f_N)$ is a collection of Laurent polynomials

$$f_i(x) = \sum_{a \in A_i} c_{i,a} x^a, \qquad i = 1, \ldots, N$$

where the A_i are (usually small) finite subsets of $\mathbb{Z}^N$ and are called the support of f_i.

We denote by $Q_i := \text{conv} A_i$ the convex hull of A_i in $\mathbb{R}^N$, and by "vol" the usual Euclidean volume in $\mathbb{R}^N$. It can be seen that

$$R(\lambda_1, \ldots, \lambda_N) := \text{vol}(\lambda_1 Q_1 + \ldots + \lambda_N Q_N)$$

is a polynomial in $\{\lambda_1, \ldots, \lambda_N\}$, and the coefficient of $\lambda_1 \cdots \lambda_N$ is called the *mixed volume* $\mathcal{M}(Q_1, \ldots, Q_N)$. If $\mathbb{C}^*$ denotes the non-zero complex numbers, then the theorem of BERNSTEIN [1975] can be stated in the following way, see also CANNY and ROJAS [1991].

THEOREM 23.1 *For almost all coefficients $c_{i,a} \in \mathbb{C}^*$, the number of zeros of $f(x) = 0$ in $(\mathbb{C}^*)^N$ equals the mixed volume $\mathcal{M}(Q_1, \ldots, Q_N)$.*

If each polynomial f_i has only a few terms, which is typical for applications, then the mixed volume is a much smaller number than the Bezout number. For example, if $N = 2$ and

$$\begin{aligned} f_1(x) &= a_0 + a_1 x_1 + a_2 x_1^n x_2^n \ , \\ f_2(x) &= b_0 + b_1 x_1 + b_2 x_1^n x_2^n \ , \end{aligned}$$

then the mixed volume is $2n$ and the Bezout number is $(2n)^2$.

The qualifier *almost all* in the theorem is very technical, it cannot be explained by a simple perturbation as in Sard's theorem. However, it is easily overcome numerically.

The point of departure for constructing a homotopy is a *lifting* function $\omega = (\omega_1, \ldots, \omega_N)$ such that $\omega_i : A_i \to \mathbb{R}$. This defines a homotopy $\hat{f}(x,t)$ via

$$\hat{f}_i(x,t) := \sum_{q \in A_i} c_{i,q} x^q t^{w_i(q)}, \qquad i = 1, \ldots, N.$$

The lifting has to be chosen in a *sufficiently generic* way so that a certain finite subdivision $\{S_\omega = (A_{1,j}, \ldots, A_{N,j})\}_{j \in J}$ of the set $A := (A_1, \ldots, A_N)$ which is generated via ω has adequate properties (called a *fine mixed subdivision*).

In principle, the path following technique consists in following a solution branch of $f(x,t) = 0$ from $t = 0$ to $t = 1$. However, the branches are singular at $t = 0$, and transformations depending on the member $j \in J$ of the subdivision are used. The main point of the approach is that the number of branches to be followed equals the mixed volume. Furthermore, it is possible to exploit additional special structure in the polynomial system which arises from some of the exponent sets A_i being equal.

Other recent methods related to this approach have been discussed in Verlinden and Haegemans [1994], Verschelde and Cools [1992, 1994a, 1994b], Verschelde and Gatermann [1994], Verschelde and Haegemans [1993, 1994], Verschelde, Verlinden and Cools [1994].

24 Nonlinear Eigenvalue Problems, Bifurcation

Path following methods are frequently applied in numerical studies of bifurcation problems. For purposes of curve tracing we have so far assumed that zero is a regular value of the smooth mapping $H : \mathbb{R}^{N+1} \to \mathbb{R}^N$. However, in the context of nonlinear eigenvalue problems, bifurcation points play an important role, and they are singular points on $H^{-1}(0)$. Hence, if path following algorithms are applied, some special adaptations are required. In typical applications, bifurcation points are defined in a Banach space context, see for example the book of Chow and Hale [1982]. In the case that H represents a mapping arising from a discretization of an operator of the form $\mathcal{H} : E \times \mathbb{R} \to F$ where E and F represent appropriate Banach spaces, it is usually of interest to approximate bifurcation points of the operator equation $\mathcal{H} = 0$. Often one can make the discretization H in such

a way that the resulting discretized equation $H = 0$ also has a corresponding bifurcation point. Under reasonable non-degeneracy assumptions it is possible to obtain error estimates for the bifurcation point of the original problem $\mathcal{H} = 0$. Such studies are made in the papers BREZZI, RAPPAZ, and RAVIART [1980a, 1980b, 1981], CROUZEIX and RAPPAZ [1990], FINK and RHEINBOLDT [1983, 1984, 1985], LIU and RHEINBOLDT [1991].

Since we are primarily concerned with bifurcation in the numerical curve following context, we confine our discussion to the case of the finite dimensional (discretized) equation $H = 0$. However, we note that the theoretical discussion below will essentially extend to the Banach space context if we assume that H is a Fredholm operator of index one. We will discuss how certain types of bifurcation points along a solution curve c can be detected, and having detected a bifurcation point, how one can numerically switch from c onto a bifurcating branch.

Some of the fundamental results on the numerical solution of bifurcation problems are due to KELLER [1970], see also KEENER and KELLER [1974] and KELLER [1977]. The recent literature on the numerical treatment of bifurcation is very extensive. For an introduction into the field we suggest the lecture notes of KELLER [1987], see also the two articles DOEDEL, KELLER, and KERNÉVEZ [1991a, 1991b] which discuss the use of the software package `AUTO`. For surveys and bibliography we suggest the recent book SEYDEL [1988] and the recent proceedings MITTELMANN and ROOSE [1990], ROOSE, DE DIER, and SPENCE [1990], SEYDEL, SCHNEIDER, KÜPPER, and TROGER [1991].

Most authors study bifurcation problems in the context of a nonlinear eigenvalue problem

$$H(x, \lambda) = 0$$

where λ is the eigenvalue parameter which usually has some physical significance. Conventionally, the solution branches are parametrized according to λ. We have taken the viewpoint that the solution branches c_i are parametrized with respect to arclength. There is only one essential difference, namely that the former approach considers folds with respect to λ also as singularities, i.e., points such that $H(x, \lambda) = 0$ and $H'(x, \lambda)$ has full rank, but $H_x(x, \lambda)$ is rank deficient.

Such folds frequently are of intrinsic interest, and there are special algorithms for detecting and calculating them. We refer the interested reader to, e.g., BOLSTAD and KELLER [1986], CHAN [1984b], FINK and RHEINBOLDT [1986, 1987], MELHEM and RHEINBOLDT [1982], PÖNISCH and SCHWETLICK [1981], SCHWETLICK [1984a, 1984b], USHIDA and CHUA [1984]. See also the remarks on this topic in Section 8 and in Section 25.

A standard approach to the determination of bifurcation or other sin-

gular points is to directly characterize such points by adjoining additional equations to $H = 0$ and handling the resulting new set of equations by some special iterative method. In this context, continuation methods often are used to obtain starting points for these direct methods, see, e.g., GRIEWANK [1985], MOORE and SPENCE [1980], YANG and KELLER [1986], see also Section 25. Hybrid and recursive projection methods for handling unstable branches have been developed by SHROFF and KELLER [1991, 1993]. See also JARAUSCH and MACKENS [1984, 1987].

In view of the extensive literature we can only touch upon the problem here, and we will confine our discussion primarily to the task of detecting a simple bifurcation point along a solution curve c and effecting a branch switching numerically. We will see that the detection of simple bifurcation points requires only minor modifications of predictor corrector algorithms.

Suppose that $c : J \to \mathbb{R}^{N+1}$ is a smooth curve, defined on an open interval J containing zero, and parametrized (for reasons of simplicity) with respect to arclength such that $H\big(c(s)\big) = 0$ for $s \in J$. The point $c(0)$ is called a *bifurcation point* of the equation $H = 0$ if there exists an $\varepsilon > 0$ such that every neighborhood of $c(0)$ contains zero-points z of H which are not on $c(-\varepsilon, \varepsilon)$.

An immediate consequence of this definition is that a bifurcation point of $H = 0$ must be a singular point of H. Hence the Jacobian $H'\big(c(0)\big)$ must have a kernel of dimension at least two. We consider the simplest case:

A point $\bar{u} \in \mathbb{R}^{N+1}$ is called a *simple bifurcation point* of the equation $H = 0$ if the following conditions hold:

$$H(\bar{u}) = 0, \quad \dim\ker H'(\bar{u}) = 2, \quad \text{and} \quad \bar{e}^* H''(\bar{u})\Big|\big(\ker H'(\bar{u})\big)^2 \qquad (24.1)$$

has one positive and one negative eigenvalue, where $\bar{e}$ spans $\ker H'(\bar{u})^*$. For convenience we normalize $||\bar{e}|| = 1$.

Using the well-known Liapunov-Schmidt reduction, the following theorem can be shown, which is essentially a restatement of a famous result of CRANDALL and RABINOWITZ [1971].

THEOREM 24.1 *Let $\bar{u} \in \mathbb{R}^{N+1}$ be a simple bifurcation point of the equation $H = 0$. Then there exist two smooth curves $c_1(s), c_2(s) \in \mathbb{R}^{N+1}$, parametrized with respect to arclength s, defined for $s \in (-\varepsilon, \varepsilon)$ and ε sufficiently small, such that the following holds:*

1. *$H\big(c_i(s)\big) = 0$, $i \in \{1, 2\}$, $s \in (-\varepsilon, \varepsilon)$,*
2. *$c_i(0) = \bar{u}$, $i \in \{1, 2\}$,*
3. *$\dot{c}_1(0)$, $\dot{c}_2(0)$ are linearly independent,*

4. $H^{-1}(0)$ *coincides locally with* $\operatorname{range}(c_1) \cup \operatorname{range}(c_2)$*, more precisely:* $\bar{u}$ *is not in the closure of* $H^{-1}(0) \setminus \big(\operatorname{range}(c_1) \cup \operatorname{range}(c_2)\big)$.

PROOF Let us introduce the decompositions

$$\mathbb{R}^{N+1} = E_1 \oplus E_2 \quad \text{and} \quad \mathbb{R}^N = F_1 \oplus F_2,$$

where

$$E_1 := \ker\, H'(\bar{u}), \quad E_2 := E_1^{\perp}, \quad F_2 := \operatorname{range} H'(\bar{u}), \quad F_1 := F_2^{\perp}.$$

Consequently,

$$\dim E_1 = 2,\ \dim E_2 = N-1,\ \dim F_1 = 1,\ \dim F_2 = N-1.$$

By introducing bases in $\mathbb{R}^{N+1}$ and $\mathbb{R}^N$ which respect these direct sums, we may view H and its derivatives in the following block form:

$$H(u) = H(u_1, u_2) = \begin{pmatrix} H_1(u_1, u_2) \\ H_2(u_1, u_2) \end{pmatrix}$$

where $u_i \in E_i$ and $H_i : E_i \to F_i$, $i = 1, 2$. From the above choice of decompositions, we have

$$H'(u) = \begin{pmatrix} \partial_1 H_1(u_1, u_2) & \partial_2 H_1(u_1, u_2) \\ \partial_1 H_2(u_1, u_2) & \partial_2 H_2(u_1, u_2) \end{pmatrix}$$

and in particular,

$$H'(\bar{u}) = \begin{pmatrix} 0 & 0 \\ 0 & \partial_2 H_2(\bar{u}_1, \bar{u}_2) \end{pmatrix}. \tag{24.2}$$

Here ∂_1, ∂_2 denote the partial derivative operators with respect to the parameters of E_1 and E_2, respectively.

Note that $\partial_2 H_2(\bar{u})$ is a nonsingular (N-1,N-1)-matrix. Since the equation $H_2(u_1, u_2) = 0$ has the solution point $(\bar{u}_1, \bar{u}_2)$, by the implicit function theorem, there exist neighborhoods U_1 of $\bar{u}_1$ in E_1 and U_2 of $\bar{u}_2$ in E_2 and a smooth map $\varphi : U_1 \to U_2$ such that

$$H_2(u_1, u_2) = 0 \quad \text{if and only if} \quad u_2 = \varphi(u_1) \tag{24.3}$$

holds for all $u_1 \in U_1$, $u_2 \in U_2$. Thus, we have a local parametrization of the equation $H_2(u_1, u_2) = 0$ in terms of the variable u_1 in the 2-dimensional space E_1. Consequently, for all $u_1 \in U_1$, $u_2 \in U_2$, the equation $H(u) = 0$ is equivalent to $u_2 = \varphi(u_1)$ and $H_1\big(u_1, u_2\big) = 0$, or

$$b(u_1) := H_1\big(u_1, \varphi(u_1)\big) = 0. \tag{24.4}$$

This is called the *bifurcation equation* for $H(u) = 0$ at the singular point $\bar{u}$.

By differentiating the equation $H_2\big(u_1, \varphi(u_1)\big) = 0$ arising from (24.3) we have by the chain rule

$$\partial_1 H_2(\bar{u}) + \partial_2 H_2(\bar{u})\varphi'(\bar{u}_1) = 0.$$

Since $\partial_1 H_2(\bar{u}) = 0$ and $\partial_2 H_2(\bar{u})$ is nonsingular, it follows that

$$\varphi'(\bar{u}_1) = 0.$$

Differentiating $b(u_1) = H_1\big(u_1, \varphi(u_1)\big)$ twice we obtain for $u = \big(u_1, \varphi(u_1)\big)$,

$$\begin{aligned} b'(u_1) &= \partial_1 H_1(u) + \partial_2 H_1(u)\varphi'(u_1), \\ b''(u_1) &= \partial_1^2 H_1(u) + 2\partial_1\partial_2 H_1(u)\varphi'(u_1) \\ &\quad + \partial_2^2 H_1(u)\big[\varphi'(u_1), \varphi'(u_1)\big] + \partial_2 H_1(u)\varphi''(u_1). \end{aligned}$$

Setting $u_1 := \bar{u}_1$ and taking into account that $\varphi'(\bar{u}_1) = 0$, $\partial_1 H_1(\bar{u}) = 0$, $\partial_2 H_1(\bar{u}) = 0$, we obtain

$$b(\bar{u}_1) = 0, \qquad b'(\bar{u}_1) = 0, \qquad b''(\bar{u}_1) = \partial_1^2 H_1(\bar{u}).$$

The simplest (generic) case is that the $(2,2)$ Hessian matrix $b''(\bar{u}_1)$ is nonsingular, i.e., both eigenvalues are different from zero. We use the following 2-dimensional version of a celebrated theorem of Morse, see, e.g., the book HIRSCH [1976, p.145], in order to characterize the local structure of the solution set $b^{-1}(0)$.

Let $\bar{u}_1 \in \mathbb{R}^2$, and let $b : \mathbb{R}^2 \to \mathbb{R}$ be a smooth function such that $b(\bar{u}_1) = 0$, $b'(\bar{u}_1) = 0$ and the Hessian $b''(\bar{u}_1)$ has nonzero eigenvalues λ_1, λ_2. Then there are open neighborhoods U of $0 \in \mathbb{R}^2$ and V of $\bar{u}_1 \in \mathbb{R}^2$ and a diffeomorphism $\psi : U \to V$ such that $\psi(0) = \bar{u}_1$ and $b\big(\psi(\xi_1, \xi_2)\big) = \lambda_1\xi_1^2 + \lambda_2\xi_2^2$ where $(\xi_1, \xi_2) \in U$.

If both eigenvalues have the same sign, then $\bar{u}_1$ is an isolated zero point of b and consequently $\bar{u}$ is an isolated zero point of H. Such points are of no interest to us, since they cannot be obtained by traversing a solution curve of the equation $H = 0$.

If the eigenvalues are of opposite sign, then the local structure of $b^{-1}(0)$ near $\bar{u}_1$ and consequently the local structure of $H^{-1}(0)$ near $\bar{u}$ are described by two curves, intersecting transversely at $\bar{u}_1$ and $\bar{u}$, respectively. Note that definition (24.1) generates exactly this case since the component map H_1 corresponds to $\bar{e}^* H$. □

By differentiating the equation $\bar{e}^* H\big(c_i(s)\big) = 0$ twice and evaluating the result at $s = 0$, we obtain the following equations:

$$\ker H'(\bar{u}) = \operatorname{span}\{\dot{c}_1(0), \dot{c}_2(0)\}, \tag{24.5}$$

$$\bar{e}^* H''(\bar{u})\big[\dot{c}_i(0), \dot{c}_i(0)\big] = 0 \text{ for } i \in \{1, 2\}. \tag{24.6}$$

The following theorem reflects the well-known fact, see KRASNOSEL'SKIĬ [1964] or RABINOWITZ [1971], that simple bifurcation points cause a switch of orientation along the solution branches. This furnishes a numerically implementable criterion for detecting a simple bifurcation point when traversing one of the curves c_i.

THEOREM 24.2 *Let $\bar{u} \in \mathbb{R}^{N+1}$ be a simple bifurcation point of the equation $H = 0$. Then the determinant of the augmented Jacobian*

$$\det \begin{pmatrix} H'\big(c_i(s)\big) \\ \dot{c}_i(s)^* \end{pmatrix}$$

changes sign at $s = 0$ for $i \in \{1, 2\}$.

PROOF We treat the case $i = 1$. It is more convenient for the proof to use the permuted matrix:

$$A(s) := \begin{pmatrix} \dot{c}_1(s)^* \\ H'\big(c_1(s)\big) \end{pmatrix}.$$

Consider an orthogonal $(N{+}1,N{+}1)$-matrix $V = (v_1, \ldots, v_{N+1})$ where $v_1 := \dot{c}_1(0)$, $\operatorname{span}\{v_1, v_2\} = \ker H'(\bar{u})$, and an orthogonal (N, N)-matrix $W = (w_1, \ldots, w_N)$ where $w_1 := e$ spans $\ker H'(\bar{u})^*$ as in definition (8.1.11). Since

$$\begin{aligned} \dot{c}_1(s)^* v_j &= \dot{c}_1(0)^* v_j + \mathcal{O}(s), \\ w_k^* H'\big(c_1(s)\big) v_j &= w_k^* H'(\bar{u}) v_j + w_k^* H''(\bar{u})\big[\dot{c}_1(0), v_j\big] s + \mathcal{O}(s^2), \end{aligned}$$

we obtain:

$$\begin{pmatrix} 1 & 0^* \\ 0 & W^* \end{pmatrix} A(s) V = \begin{pmatrix} 1 + \mathcal{O}(s) & \mathcal{O}(s) & \mathcal{O}(s) \\ \mathcal{O}(s^2) & \rho s + \mathcal{O}(s^2) & \mathcal{O}(s) \\ \mathcal{O}(s) & \mathcal{O}(s) & B + \mathcal{O}(s) \end{pmatrix}. \tag{24.7}$$

The $(N{-}1,N{-}1)$ block matrix B in (24.7) is nonsingular, see (24.2) and the remarks thereafter. The scalar ρ in (24.7) is given as the off-diagonal entry of the following symmetric $(2, 2)$-matrix

$$\begin{pmatrix} \bar{e}^* H''(\bar{u})[v_1, v_1] & \bar{e}^* H''(\bar{u})[v_1, v_2] \\ \bar{e}^* H''(\bar{u})[v_2, v_1] & \bar{e}^* H''(\bar{u})[v_2, v_2] \end{pmatrix}.$$

Since this matrix is nonsingular, see (24.1), and since the diagonal entry $\bar{e}^* H''(\bar{u})[v_1, v_1]$ vanishes, see (24.6), it follows that $\rho \neq 0$. Now by performing Gaussian elimination upon the first two columns of (24.7), we obtain a reduced form

$$\begin{pmatrix} 1 + \mathcal{O}(s) & \mathcal{O}(1) & \mathcal{O}(s) \\ 0 & \rho s + \mathcal{O}(s^2) & \mathcal{O}(s) \\ 0 & 0 & B + \mathcal{O}(s) \end{pmatrix}$$

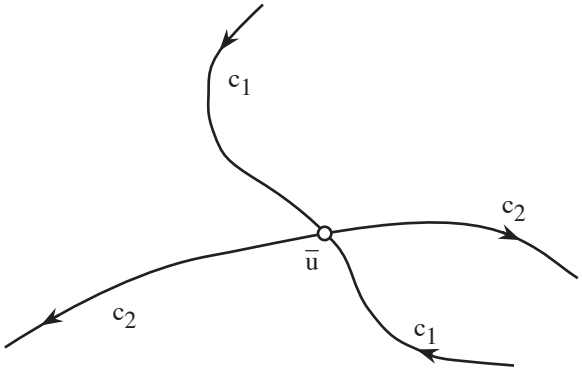

Figure 24.1. Change of Orientation

which clearly has a determinant of the form

$$\rho \det(B)s + \mathcal{O}(s^2).$$

It follows that the determinant of $A(s)$ changes sign at $s = 0$. □

The above theorem implies that when traversing a solution curve c, a simple bifurcation point is detected by a change in orientation, see Figure 24.1. Depending upon the method used to perform the decomposition of the Jacobian during path following, the above orientation can often be calculated at very small additional cost. A predictor corrector algorithm generally has no difficulty in *jumping over*, i.e., proceeding beyond the bifurcation point $\bar{u}$. That is, KELLER [1977] has shown that for sufficiently small steplength h, the predictor point will fall into the *cone of attraction* of the Newton corrector. See JEPSON and DECKER [1986] for further studies. See also ALLGOWER and GEORG [1990, Lemma (8.1.16)].

Conversely, suppose that a smooth c in $H^{-1}(0)$ is traversed and that $c(0)$ is an isolated singular point of H such that the above determinant changes sign at $s = 0$. Then, using a standard argument in degree theory, see KRASNOSEL'SKIĬ [1964] or RABINOWITZ [1971], it can be shown that $c(0)$ is a bifurcation point of $H = 0$. However, $c(0)$ is not necessarily a simple bifurcation point.

When using a predictor corrector method with Newton-type correctors along the solution curve, then generally the Jacobian H' my be replaced by an approximation, e.g., a finite difference approximation, or an approximation generated by an update method, see Section 14. Thus a chord method

may be implemented. We emphasize, however, that it is necessary to obtain a good approximation of the Jacobian at least once at the predictor point since otherwise the local convergence of the Newton corrector iterations cannot be guaranteed when jumping over a simple bifurcation point. This is explained in ALLGOWER and GEORG [1990, p. 84].

The determinant in Theorem 24.2 is only the simplest example of a so-called *test function*. Such test functions are real functions defined on a neighborhood of the curve c and are monitored during path following to reveal certain types of singular points by a change of sign. In the case of Hopf bifurcation, the determinant is not an adequate test function. Recently, several authors have proposed and studied classes of test functions for various types of singular points, see, e.g., CHU, GOVAERTS, and SPENCE [1994], DAI and RHEINBOLDT [1990], GARRATT, MOORE, and SPENCE [1991], GRIEWANK and REDDIEN [1984], SEYDEL [1991b], WERNER [1992]. A different approach for the prediction of singular points along the path c has been given by HUITFELDT and RUHE [1990].

24.1 Switching Branches Via Perturbation.

In the previous section we have seen that it is possible to detect and jump over simple bifurcation points while numerically tracing a solution curve c via a predictor corrector method. The more difficult task is to numerically branch off onto the second solution curve at the detected bifurcation point $\bar{u}$. The simplest device for branching off numerically rests upon Sard's theorem (18.1). If a small perturbation vector $d \in \mathbb{R}^N$ is chosen at random, then the probability that d is a regular value of H is unity. Of course, in this case $H^{-1}(d)$ has no bifurcation point. Since $d \in \mathbb{R}^N$ is chosen so that $||d||$ is small, the solution sets $H^{-1}(0)$ and $H^{-1}(d)$ are close together. On $H^{-1}(d)$, no change of orientation can occur. Therefore, corresponding solution curves in $H^{-1}(d)$ must branch off near the bifurcation point $\bar{u}$, see Figure 24.2. It is easy to implement this idea, see, e.g., ALLGOWER and CHIEN [1986], ALLGOWER, CHIEN, GEORG, and WANG [1991d], CHIEN [1989], GEORG [1981b], GLOWINSKI, KELLER, and REINHART [1985].

Recently, an interesting variation on this idea has been proposed by HUITFELDT [1991]. He introduces an additional parameter on the perturbation and an additional constraint equation to obtain the *branch connecting equation*

$$\mathcal{B}(u,\tau) := \begin{pmatrix} H(u) + \tau d \\ ||u - \hat{u}||^2 + \tau^2 - \varepsilon^2 \end{pmatrix} = 0, \tag{24.8}$$

where $\hat{u}$ is an approximation to the bifurcation point $\bar{u}$. Such approximations are easily obtained via path following together with test function

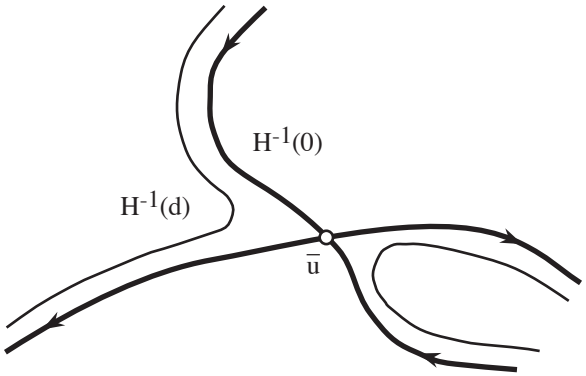

Figure 24.2. Perturbation of a Simple Bifurcation Point

monitoring as described above. Note the relationship between the above homotopy and the d-homotopy discussed in Section 21 in connection with finding multiple solutions.

It is not difficult to see that for almost all d and $\varepsilon > 0$, zero is a regular value of $\mathcal{B}$, provided that $\bar{u}$ is an isolated singular point of H in $H^{-1}(0)$. Let us assume that such a generic choice of d and ε has been made.

Then the solution manifold $\mathcal{B}^{-1}(0)$ splits into one or more simple closed curves of the form $(b(s), \tau(s))$. For $\tau(s) = 0$ we obtain $H(b(s)) = 0$. Hence the curves connect points in the intersection of $H^{-1}(0)$ with the sphere $||u - \hat{u}||^2 = \varepsilon^2$. Starting points for a path following of $(b(s), \tau(s))$ are available from the tracing of the current solution curve c of $H = 0$. Let $b_i = b(s_i)$, $i = 0, 1, \ldots$, be successively obtained points such that $\tau(s_i) = 0$. It remains to be demonstrated that b_i and b_{i+1} are on different solution branches of the equation $H = 0$.

Since this seems to have been omitted in the paper of HUITFELDT [1991], we sketch a proof. It is easily seen that the determinant of the matrix

$$\begin{pmatrix} H'(b(s)) & d \\ (b(s) - \hat{u})^* & \tau(s) \\ \dot{b}(s)^* & \dot{\tau}(s) \end{pmatrix}$$

never changes sign since it never becomes singular. By multiplying the above matrix on the right with

$$\begin{pmatrix} \mathrm{Id} & \dot{b}(s) \\ 0^* & \dot{\tau}(s) \end{pmatrix}$$

we obtain

$$\begin{pmatrix} H'(b(s)) & 0 \\ (b(s)-\hat{u})^* & 0 \\ \dot{b}(s)^* & 1 \end{pmatrix}.$$

Since $\dot{\tau}(s_i)$ changes sign for successive i, we obtain that the determinant of

$$\begin{pmatrix} H'(b_i) \\ (b_i-\hat{u})^* \end{pmatrix}$$

changes sign for successive i. Under reasonable assumptions this implies that $t(H'(b_{i+1}))$ points out of the sphere $||u-\hat{u}||^2 = \varepsilon^2$ if $t(H'(b_i))$ points into it. For a simple bifurcation point (or more generally for a bifurcation point which is detected by a change of determinant in the sense of Theorem 24.2), this means that b_i and b_{i+1} cannot lie on the same solution branch.

Huitfeldt reports very successful numerical tests on some interesting problems of applied mathematics: the Taylor problem, and the von Karman plate equations. In his experiments he succeeded in obtaining all of the bifurcating branches at several multiple bifurcation points, i.e., the 1-manifold $\mathcal{B}^{-1}(0)$ was connected in all cases he considered. However, it does not seem that this should always be the case. Advantages of this approach are that no a priori information concerning the multiplicity of the bifurcation is needed, and that it enjoys better numerical stability properties than ordinary perturbation. It should, however, be emphasized that any existing symmetries leading to higher multiplicities ought to be taken into account initially, i.e., by using group actions in the formulation of the problem, see GOLUBITSKY, STEWART, and SCHAEFFER [1988] and the discussion in Section 24.3.

24.2 Branching Off Via Minimally Extended Systems

Although the branching off via perturbation techniques works effectively, this approach has shortcommings. In general, it cannot be decided in advance which of the two possible directions along the bifurcating branch will be taken. Furthermore, if the perturbation vector d is not chosen correctly (and it is not always clear how this is to be done), one may still have some difficulty in tracing the resulting path (because of instabilities). Furthermore, the solution set $H^{-1}(0)$ can be approximated near the bifurcation point $\bar{u}$ only after an additional bifurcating branch has been approximated.

To obtain an approximation of $H^{-1}(0)$ near a simple bifurcation point $\bar{u}$, the alternative is a direct approach. After a coarse (initial) approximation of the bifurcation point has been obtained via a determinant (or some other) test, the following methods could be employed:

1. The nonlinear system $H(u) = 0$ is extended in order to directly characterize bifurcation points. A rapidly convergent Newton-like iteration can then be applied to the extended system of equations. We refer to SEYDEL [1979] and WEBER [1981] as early papers on this method.

2. A numerical model for the so-called bifurcation equation is constructed in order to approximate all tangents of the bifurcating branches in $\bar{u}$. Equation 24.6 describes such an equation for the case of a simple bifurcation point. The approaches, e.g., in KELLER [1977, 1987] and RHEINBOLDT [1978, 1986] deal with this idea.

We will discuss in some detail the first approach, in particular for the case that only a minimal number of auxiliary equations are added. The first minimally extended system for bifurcation points appears to go back to GRIEWANK and REDDIEN [1984] where auxiliary quantities are defined by nonsingular $(n+2)$-dimensional linear systems. It seems, however, that this system (which occurs there as a special case of a more general approach) has not fully been recognized as a tool for computing simple bifurcation points.

A different system has been proposed by PÖNISCH [1985] where the auxiliary quantities are obtained from $(n+1)$-dimensional linear systems. However, there the restricting assumption

$$\operatorname{rank} \partial_x H(\underbrace{\bar{x}, \bar{\lambda}}_{\bar{u}}) = N - 1 \tag{24.9}$$

is posed. Later JANOVSKÝ [1989] gave a simplification of Pönisch's system which reduces the cost of the linear algebra involved. Since in both systems the freedom in the choice of certain parameters which specify the method was unnecessarily restricted and, in particular, excluded bifurcation points with $\operatorname{rank} \partial_x F(\bar{x}, \bar{\lambda}) = N-2$, ALLGOWER and SCHWETLICK [1993] proposed a generalization of the systems considered by Pönisch and Janovský which does not require assumption (24.9). This allows one to compute arbitrary simple bifurcation points and to choose the parameters in an optimal way.

Though there seems to be no difference in the computational cost between the method of Griewank and Reddien and the one of Allgower and Schwetlick, we follow the latter approach, because it seems to be somewhat better suited to be integrated into the path following situation.

Let us assume in the following that $\bar{u}$ is a simple bifurcation point of the equation $H(u) = 0$. In particular, see (24.1), there exist vectors $\bar{v}_1, \bar{v}_2$ spanning $\ker H'(\bar{u})$, and a vector $\bar{\psi}$ spanning $\ker H'(\bar{u})^*$.

Given a fixed $d \in \mathbb{R}^N$ and sufficiently smooth $f_i : \mathbb{R}^{N+1} \to \mathbb{R}$, $i = 1, 2$, we call the system

$$G(u, \mu) := \begin{pmatrix} H(u) + \mu d \\ f_1(u) \\ f_2(u) \end{pmatrix} = 0 \tag{24.10}$$

a *minimally extended system* for $\bar{u}$ of $H(u) = 0$ if $u = \bar{u}$ and $\mu = 0$ is a regular zero point of this system. The idea is that rapidly convergent Newton-like iterations can be employed on (24.10) to approximate $\bar{u}$.

Given fixed vectors $w, r \in \mathbb{R}^{N+1}$, it is not difficult to see that the matrices

$$C_1(u) := \begin{pmatrix} H'(u) + dr^* \\ w^* \end{pmatrix}, \quad C_2(u) := \begin{pmatrix} H'(u) + dw^* \\ r^* \end{pmatrix} \tag{24.11}$$

are nonsingular at $u = \bar{u}$ if the conditions

$$d^* \bar{\psi} \neq 0, \ \det \left[\begin{pmatrix} r^* \\ w^* \end{pmatrix} (\bar{v}_1 \, , \bar{v}_2) \right] \neq 0 \tag{24.12}$$

hold.

Let us assume that these (generic) conditions hold for d, w, r. Then the auxiliary variables $v_i(u) \in \mathbb{R}^{N+1}$ are implicitly defined and smooth on some neighborhood of $\bar{u}$ as solutions of the linear systems

$$C_i(u) v_i(u) = e_{N+1} \, . \tag{24.13}$$

It is not difficult to see that the quantities d and

$$f_1(u) := r^* v_1(u), \quad f_2(u) := w^* v_2(u) \tag{24.14}$$

minimally extend $H(u) = 0$ in the sense of (24.10). Furthermore, the $v_i(u)$ are linearly independent at $u = \bar{u}$ and therefore on some neighborhood of $\bar{u}$.

Let us consider Newton's method for solving (24.10). If we denote by (u, μ) the current approximation of the solution, and by $(u + \Delta u, \mu + \Delta \mu)$ its Newton correction, then we obtain the following equations for Δu and $\Delta \mu$:

$$H'(u) \Delta u + d \Delta \mu = -H(u) - \mu d, \tag{24.15}$$

$$f_i'(u) \Delta u = -f_i(u). \tag{24.16}$$

If we denote by $v_0(u)$ the solution of the equation

$$C_2(u) v_0(u) = \begin{pmatrix} -H(u) - d\mu \\ 0 \end{pmatrix} \tag{24.17}$$

(which is defined on some neighborhood of $\bar{u}$), then we see that $(\Delta u, \Delta\mu) = \big(v_0(u), w^* v_0(u)\big)$ satisfies (24.15). Furthermore, the $v_i(u)$ are solutions of the homogeneous part of (24.15), and therefore a solution of (24.15–24.16) has the form

$$\begin{pmatrix} \Delta u \\ \Delta\mu \end{pmatrix} = \begin{pmatrix} v_0(u) + \xi_1 v_1(u) + \xi_2 v_2(u) \\ w^* v_0(u) + \xi_1 f_1(u) + \xi_2 f_2(u) \end{pmatrix}. \tag{24.18}$$

It remains to find the values ξ_1, ξ_2. They are determined by (24.16).

To do this, it is helpful to see that (24.13–24.14) implies

$$\begin{aligned} f_1'(u)\Delta u = - \quad & \underbrace{r^* C_1(u)^{-1}}_{=: (\psi_1(u)^*, g_1(u))} \quad C_1'(u)\Delta u \quad \underbrace{C_1(u)^{-1} e_{N+1}}_{= v_1(u)} \\ f_2'(u)\Delta u = - \quad & \underbrace{w^* C_2(u)^{-1}}_{=: (\psi_2(u)^*, g_2(u))} \quad C_2'(u)\Delta u \quad \underbrace{C_2(u)^{-1} e_{N+1}}_{= v_2(u)} \end{aligned} \tag{24.19}$$

and consequently by (24.11):

$$f_i'(u)\Delta u = -\psi_i^* H''(u)[v_i(u), \Delta u]. \tag{24.20}$$

By using the definitions in (24.13–24.14) and (24.19), it can be shown $f_i(u) = g_i(u)$, $i = 1, 2$, and

$$\begin{aligned} \psi_2(u)^* H'(u) &= \frac{f_1(u)(f_2(u) - 1)}{1 - f_1(u)} w^* - f_2(u) r^*, \\ \psi_2(u)^* d &= \frac{1 - f_1(u) f_2(u)}{1 - f_1(u)}. \end{aligned} \tag{24.21}$$

Finally, we note that

$$C_1(u) - C_2(u) = \begin{pmatrix} d \\ -1 \end{pmatrix} (r^* - w^*)$$

has rank one, and hence by the Sherman-Morrison formula

$$C_1(u)^{-1} = C_2(u)^{-1} - \frac{\overbrace{C_2(u)^{-1} \begin{pmatrix} d \\ -1 \end{pmatrix}}^{=: \hat{v}(u)} (r^* - w^*) C_2(u)^{-1}}{\underbrace{1 + (r^* - w^*) C_2(u)^{-1} \begin{pmatrix} d \\ -1 \end{pmatrix}}_{= -w^* \hat{v}(u)}}. \tag{24.22}$$

It can be seen from the definition of the $v_i(u)$ in (24.13) and $\hat{v}(u)$ in (24.22) that $\hat{v}(\bar{u}) = v_1(\bar{u}) - v_2(\bar{u})$ holds, and consequently $w^*\hat{v}(\bar{u}) = 1$. Hence the update formula (24.22) for $C_1(u)^{-1}$ is numerically stable on some neighborhood of $\bar{u}$.

Multiplying (24.22) from the right with e_{N+1} gives

$$v_1(u) = v_2(u) + \frac{1 - f_2(u)}{w^*\hat{v}(u)}\hat{v}(u).$$

Multiplying (24.22) from the left with r^* and taking the first n components gives

$$\psi_1(u)^* = \frac{\psi_2(u)^*}{w^*\hat{v}(u)}.$$

Hence the calculation of $v_1(u)$ and $\psi_1(u)$ can be cheaply obtained from $v_2(u)$ and $\psi_2(u)$, respectively.

Since it is known that the solution of (24.10) is $u = \bar{u}$ and $\mu = 0$, it makes sense to reset $\mu = 0$ after each Newton step. This clearly can only improve the rate of convergence. The above ideas are summarized in the following algorithm for computing a simple bifurcation point with Newton's method, see also ALLGOWER and SCHWETLICK [1993].

ALGORITHM 24.1 **(Simple Bifurcation Point)**

INPUT initial point $u \in \mathbb{R}^{N+1}$
parameters $d \in \mathbb{R}^N$, $w, r \in \mathbb{R}^{N+1}$

REPEAT until convergence

$J \leftarrow H'(u),\ C \leftarrow \begin{pmatrix} J + dw^* \\ r^* \end{pmatrix}$

solve $C \begin{pmatrix} v_2 & \hat{v} & v_0 \end{pmatrix} = \begin{pmatrix} 0 & d & -H(u) \\ 1 & -1 & 0 \end{pmatrix}$ for v_2, $\hat{v}$, v_0

$v_1 \leftarrow v_2 + \dfrac{1 - w^*v_2}{w^*\hat{v}}\,\hat{v}$

solve $C^* \begin{pmatrix} \psi_2 \\ g_2 \end{pmatrix} = w$ for ψ_2

$\eta_i \leftarrow \psi_2^* H''(u)[v_i, v_0]$, $i = 1, 2$

$M_{ij} \leftarrow \psi_2^* H''(u)[v_i, v_j]$, $M \leftarrow (M_{ij})$, $i, j = 1, 2$

solve $M \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} (w^*\hat{v})(r^*v_2) - \eta_1 \\ w^*v_2 - \eta_2 \end{pmatrix}$

$u \leftarrow u + v_0 + \xi_1 v_1 + \xi_2 v_2$

END REPEAT

Note that the second derivatives involved in the calculation of the η and M values can be obtained via finite difference formulas, i.e., requiring only a few evaluations of H, see (16.2).

The original approach of GRIEWANK and REDDIEN [1984] can be implemented in a similar way. The computational cost is essentially the same as in Algorithm 24.1. However, when a bifurcation point is approached by path following along one of the paths using predictor corrector methods, then usually linear systems with bordered matrices, e.g.,

$$A := \begin{pmatrix} J \\ r^* \end{pmatrix} \quad \text{with} \quad J \approx H'(u) \text{ and } r \approx t\big(H'(u)\big) \tag{24.23}$$

have to be solved. So switching to Algorithm 24.1 requires only modifying the upper block J of C by the rank one term dw^*, but not to change the dimension of the linear systems. Therefore, the approach of ALLGOWER and SCHWETLICK [1993] is somewhat more suitable in the context of path following than the method of GRIEWANK and REDDIEN [1984].

Switching Branches

Let us now turn to the problem of switching branches after a bifurcation point has been located with the above method. Let us recall the definition (24.1) of a simple bifurcation point, and note that $\bar{e}$ is now called $\bar{\psi}$. Furthermore, (24.21) implies that $\psi_2(\bar{u})$ is a multiple of $\bar{\psi}$. Hence, the above algorithm furnishes a good approximation of $\bar{\psi}$.

Let us also recall that the tangents t_i of the two bifurcating branches at $\bar{u}$ were characterized in (24.6). This leads to

$$\bar{\psi}^* H''(\bar{u})\big[t_i, t_i\big] = 0.$$

Since the matrix M in Algorithm 24.1 approximates $\bar{\psi}^* H''(\bar{u})$ on $\ker H'(\bar{u})$ (which is spanned by $\bar{v}_1, \bar{v}_2$), we have

$$\tilde{t}_i := \beta_1 \bar{v}_1 + \beta_2 \bar{v}_2, \; t_i \approx \frac{\tilde{t}_i}{||\tilde{t}_i||}$$

for a non-trivial solution of the quadratic equations $\beta^* M \beta = 0$. Hence, the above algorithm furnishes without further computational expense the starting direction for following the bifurcating branch.

Choosing the Parameters

Because of (24.12), the parameters d and $\{r, w\}$ should be chosen in such a way that $\ker\big(H'(\bar{u})^*\big) \approx \operatorname{span}\{d\}$ and $\ker H'(\bar{u}) \approx \operatorname{span}\{r, w\}$ with r, w approximately orthogonal, respectively. When starting Algorithm 24.1 from

a predictor corrector method, we have the following situation: One of the two bifurcating curves is being traced, say c_1. An approximation u of $\bar{u}$ is known, e.g., by interpolating between the last two points on the curve c_1 for which a change of sign in the sense of Theorem 24.2 was detected. More generally, some test function might have been used for this detection, see the remarks before Section 24.1. Also, the choices $r = t\big(H'(u)\big)$ and $J = H'(u)$ are already available. Then, setting A as in (24.23), the missing parameters d and w can be obtained by $A^*d \approx 0$, $Aw \approx 0$. Since u is only an approximation of $\bar{u}$, these conditions can be implemented in a numerically stable way, see ALLGOWER and GEORG [1990, (8.3.6)], by approximating $\min_d\{||A^*d|| : ||d|| = 1\}$ and $\min_w\{||Aw|| : ||w|| = 1\}$ either via inverse iteration or via a Lanczos method for approximating the smallest eigenvalues and corresponding eigenvectors of AA^* and A^*A, see GOLUB and VAN LOAN [1989]. Both problems can be solved simultaneously at the same cost.

24.3 Multiple Bifurcation and Symmetry

Multiple bifurcations often arise from symmetries with respect to certain group actions, i.e., H satisfies an *equivariance* condition

$$\mathcal{H}(\gamma x, \lambda) = \gamma \mathcal{H}(x, \lambda) \tag{24.24}$$

for γ in a group Γ. See the books of GOLUBITSKY and SCHAEFFER [1985], GOLUBITSKY, STEWART, and SCHAEFFER [1988], and VANDERBAUWHEDE [1982], and the recent monograph of IZE [1993]. These symmetries can also be exploited numerically, see, e.g., ALLGOWER, BÖHMER, and MEI [1991c, 1991a], ALLGOWER, BÖHMER, GEORG, and MIRANDA [1992a], CLIFFE and WINTERS [1986], DELLNITZ and WERNER [1989], GEORG and MIRANDA [1992, 1993], JEPSON, SPENCE, and CLIFFE [1991], HEALEY [1988b, 1988a, 1989], HEALEY and TREACY [1991], HONG [1993], see also the proceedings ALLGOWER, BÖHMER, and GOLUBITSKY [1992b]. As the above list suggests, there is currently very much interest in this topic.

Let us sketch briefly the main ideas involved in equivariant branching. We consider a nonlinear eigenvalue problem

$$\mathcal{H}(u, \lambda) = 0 \tag{24.25}$$

where $\mathcal{H} : E \times \mathbb{R} \to F$ is sufficiently smooth. Here E, F denote Banach spaces of functions which are defined on a domain Ω. We further assume that Γ is a finite group of transformations leaving Ω invariant such that $\mathcal{H}$ has the equivariance property (24.24) for all $\lambda \in \mathbb{R}$ and all $\gamma \in \Gamma$. Here the

group Γ acts on both spaces E and F according to $\gamma u(z) := u(\gamma^{-1}z)$ for $z \in \Omega$.

Under a suitable discretization of (24.25) which respects these symmetry properties a corresponding finite-dimensional nonlinear eigenvalue problem is obtained:

$$H(x, \lambda) = 0 \tag{24.26}$$

where $H : \mathbb{R}^{n+1} \to \mathbb{R}^n$ has an equivariance property

$$H(gx, \lambda) = gH(x, \lambda) \tag{24.27}$$

for all $g \epsilon G$, where now G is a group of permutations.

An immediate conclusion which can be drawn from the equivariance property is that if $\mathcal{H}(u, \lambda) = 0$, then $\mathcal{H}(\gamma u, \lambda) = 0$. That is, if we define the *orbit* of u under Γ by $\text{or}(u) := \{\gamma u : \gamma \in \Gamma\}$, then $\mathcal{H}$ vanishes on the entire orbit. On the other hand, we only gain a new solution in $\text{or}(u)$ if $\gamma u \neq u$ for some $\gamma \in \Gamma$. The elements of Γ which leave u fixed form the *isotropy subgroup*

$$\Sigma_u := \{\gamma \in \Gamma : \gamma u = u\}.$$

The elements of E which are fixed under a subgroup Σ of Γ are the *fixed-point subspace*

$$E^\Sigma := \{u \in E : \gamma u = u \text{ for } \gamma \in \Sigma\}.$$

A subspace $X_1 \subset E$ is called Σ-*invariant* for a subgroup Σ of Γ if $\sigma u \in X_1$ for every $\sigma \in \Sigma$ and $u \in X_1$. An example is the above subspace E^Σ.

Now $\mathcal{H}$ maps $E^\Sigma \times \mathbb{R}$ into F^Σ, since

$$\gamma \mathcal{H}(u, \lambda) = \mathcal{H}(\gamma u, \lambda) = \mathcal{H}(u, \lambda)$$

for $\gamma \in \Sigma, u \in E^\Sigma$, and $\lambda \in \mathbb{R}$. Hence it is often convenient to replace the original problem (24.25) by corresponding reduced problems

$$\mathcal{H}^\Sigma(u^\Sigma, \lambda) = 0 \tag{24.28}$$

where $\mathcal{H}^\Sigma : E^\Sigma \times \mathbb{R} \to F^\Sigma$ denotes the restriction of $\mathcal{H}$ on $E^\Sigma \times \mathbb{R}$. A projection $P^\Sigma : E \to E^\Sigma$ is given by

$$P^\Sigma u := \frac{1}{|\Sigma|} \sum_{\sigma \in \Sigma} \sigma u$$

where $|\Sigma|$ denotes the order of the subgroup Σ.

Now let (u_0, λ_0) be a bifurcation point of (24.25) and let $\mathcal{H}_0$, $D_u\mathcal{H}_0$, $D_\lambda\mathcal{H}_0$, etc. represent the evaluations of $\mathcal{H}$, $D_u\mathcal{H}$, $D_\lambda\mathcal{H}$ at (u_0, λ_0), respectively. We assume further that E, F are Hilbert spaces and $D_u\mathcal{H}_0$ is a Fredholm operator which hence admits the following orthogonal decompositions:

$$E = N(D_u\mathcal{H}_0) \oplus R((D_u\mathcal{H}_0)^*), \tag{24.29}$$

$$F = N((D_u\mathcal{H}_0)^*) \oplus R(D_u\mathcal{H}_0), \tag{24.30}$$

where * denotes forming the adjoint operator, and N, R denote the null space and range, respectively.

From the equivariance of $\mathcal{H}$ and the chain rule we see that $N(D_u\mathcal{H}_0)$ and $R((D_u\mathcal{H}_0)^*)$ are Σ_{u_0}-invariant spaces and hence also

$$E^\Sigma = N(D_u\mathcal{H}_0^\Sigma) \oplus R((D_u\mathcal{H}_0^\Sigma)^*)$$

for any subgroup Σ of Σ_{u_0}.

If $D_\lambda\mathcal{H}_0 \in R(D_u\mathcal{H}_0)$, then there is a $w_0 \in E$ such that $D_u\mathcal{H}_0 w_0 + D_\lambda\mathcal{H}_0 = 0$. If $v_0 \in R((D_u\mathcal{H}_0)^*)$ is the second component of w_0 with respect to (24.29), then

$$D_u\mathcal{H}_0 v_0 + D_\lambda\mathcal{H}_0 = 0$$

with $v_0 \in R((D_u\mathcal{H}_0)^*)$ being unique.

On the other hand, by the Γ-equivariance of $\mathcal{H}$, we have $D_u\mathcal{H}_0\sigma v_0 + D_\lambda\mathcal{H}_0 = 0$ for every $\sigma \in \Sigma_{u_0}$. Since $R((D_u\mathcal{H}_0^\Sigma)^*)$ is Σ_{u_0}-invariant, the uniqueness of v_0 yields $\sigma v_0 = v_0$ for all $\sigma \in \Sigma_{u_0}$, i.e., $v_0 \in E^{\Sigma_{u_0}}$.

If $D_\lambda\mathcal{H}_0 \notin R(D_u\mathcal{H}_0)$, then the null space of $D\mathcal{H}(u_0, \lambda_0)$ is given by

$$N\Big(D\mathcal{H}(u_0, \lambda_0)\Big) = N(D_u\mathcal{H}_0) \times \{0\},$$

and hence (u_0, λ_0) is only a fold point.

A subgroup $\Sigma \subset \Sigma_{u_0}$ is called a *bifurcation subgroup* of Σ_{u_0} with respect to a subspace $Y \subset E \times \mathbb{R}$ if:

(i) Y is Σ_{u_0}-invariant;

(ii) there is a pair $(y, c) \in Y$ such that $\Sigma_y = \Sigma$;

(iii) $\dim\,[Y \cap (E^\Sigma \times \mathbb{R})] = 2$.

The following results are shown in Allgower, Böhmer and Mei [1991c] and may be shown in a fashion similar to the proofs in Crandall and Rabinowitz or in Section 8.3 of Allgower and Georg [1990].

THEOREM 24.3 *Let Σ be a bifurcation subgroup of Σ_{u_0} with respect to $N(D_u\mathcal{H}_0)$ and let $D_\lambda\mathcal{H}_0 \in R(D_u\mathcal{H}_0)$. Then there are $(\phi,0),(v_0,1) \in N(D\mathcal{H}_0)$ such that*

$$N(D\mathcal{H}_0) \cap (E^\Sigma \times \mathbb{R}) = \operatorname{span}\{(\phi,0),(v_0,1)\} \tag{24.31}$$

and

(a) $N(D_u\mathcal{H}_0^\Sigma) = \operatorname{span}\{\phi\}$,

(b) $E^\Sigma = \operatorname{span}\{\phi\} \oplus R((D_u\mathcal{H}_0^\Sigma)^*)$.

For ϕ, v_0 as in (24.31) and $\tilde{\phi} \neq 0$ in $N((D_u\mathcal{H}_0^\Sigma)^*)$, define:

$$\begin{aligned} a &:= (\tilde{\phi}, D_{uu}\mathcal{H}_0^\Sigma\phi\phi), \\ b &:= (\tilde{\phi}, DD_u\mathcal{H}_0^\Sigma(v_0,1)\phi), \\ c &:= (\tilde{\phi}, D^2\mathcal{H}_0^\Sigma(v_0,1)^2). \end{aligned}$$

THEOREM 24.4 *Let (u_0,λ_0) be a bifurcation point of $\mathcal{H}$ and let $D_\lambda\mathcal{H}_0 \in R(D_u\mathcal{H}_0)$. Let Σ be a bifurcation subgroup of Σ_{u_0} with respect to $N(D_u\mathcal{H}_0)$ for the Γ-equivariant operator $\mathcal{H}$ such that (24.29) holds and $b^2 - ac \geq 0$. Then (u_0,λ_0) is a simple bifurcation point of the reduced problem (24.28).*

Theorem 24.4 is a version of the equivariant branching lemma of CICOGNA [1981] and VANDERBAUWHEDE [1982], see also GOLUBITSKY, STEWART, and SCHAEFFER [1988] or DELLNITZ and WERNER [1989]. The result shows that for every bifurcation subgroup $\Sigma \subset$ of the isotropy subgroup Σ_{u_0} the point (u_0,λ_0) is a simple bifurcation point in the invariant subspace E^Σ. Hence the equivariance property (24.24) often splits multiple bifurcations into simple bifurcations, which can be more easily detected by path following, see the preceding sections.

The above discussion can also be applied to finite-dimensional maps $H : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$. As mentioned above, the corresponding discretization (24.26) then usually has an equivariance (24.27) provided the discretization is chosen suitably, i.e., conforming to the symmetry structure. Now the numerical linear algebra calculations in the corrector steps of the path following can be decomposed into linear subproblems as outlined in Section 15. These subproblems can possibly be handled in parallel. Such an implementation has been given by GATERMANN and HOHMANN [1991].

It has been observed in the case of a number of semi-linear elliptic problems (see, e.g., ALLGOWER,BÖHMER and MEI [1991b, 1991c]) that the subproblems often correspond to the discrete analogues arising from the symmetries corresponding to the bifurcation subgroups of Γ. In this way the multiple bifurcations are split up so that the subproblems have simple bifurcations which are again easily detected via, e.g., the determinant criterion (24.2).

25 Applications to Fluid Dynamics

Researchers in computational fluid dynamics, computational physics and combustion have in recent years made extensive use of numerical continuation methods. In these applications one generally deals with systems of nonlinear partial differential equations. Typically, the systems are of the Navier-Stokes type. A research group at Harwell Laboratory involving Cliffe, Winters et al has given numerous reports of their experiences, see, e.g., WINTERS [1987, 1991], WINTERS and CLIFFE [1985], WINTERS, CLIFFE, and JACKSON [1987] WINTERS, MYERSCOUGH, and MAINI [1990]. We will sketch here a sampling of such examples.

After the partial differential equations have been discretized, for example by finite elements, one obtains a system of equations of the form:

$$H(x, \lambda, \alpha) = 0, \tag{25.1}$$

where $H : \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}^p \to \mathbb{R}^N$ is a smooth map. In general, x is a state variable, N is the number of degrees of freedom in the discretization, λ is a bifurcation parameter, and α is a vector of control parameters of dimension p. The *bifurcation parameter* λ is singled out, because one wants to study the change in behavior of the solution as this parameter is varied. Thus this parameter is regarded as free to be varied, e.g., as in an experiment, whereas the x and α parameters are more geometric in nature, or at least more or less fixed at the outset. The bifurcation parameter may for example, be the Reynolds number or the Rayleigh number in problems arising from fluid mechanics.

In many of the applications the points of greatest physical interest on the solution curves are thoses at which a change of stability occurs. Typically, these points are fold points and bifurcation points. Very frequently the physical problem also exhibits geometrical symmetries arising from cell structure or geometrical symmetry of vessels, pipes, etc. In such cases of geometric symmetry, if the differential operator also enjoys equivariance properties, see (24.24), with respect to the elements of the symmetry group, multiple bifurcation may arise. Situations of this sort may be theoretically described via the theory of equivariant branching which stems from VANDERBAUWHEDE [1982] and CICOGNA [1981], see also GOLUBITSKY, STEWART, and SCHAEFFER [1988] for an extensive discussion of the equivariant branching lemma. See also the discussion in Section 24.3.

The numerical treatment of the discretized problem (25.1) involves in addition to the handling of relatively large degrees of freedom the need to detect the points where changes of stability, i.e., fold or bifurcation points, occur. At the end of Section 24 we touched upon some functionals for detecting and calculating special points on the solution curve.

The paper WINTERS, CLIFFE, and JACKSON [1987] summarizes a number of applications in fluid dynamics. In their numerical studies the authors report experience using an Euler-Newton method involving LU factorizations of $H_x = \frac{\partial H}{\partial x}$ and a simple steplength control of halving or doubling governed by the behavior of the corrector. Since the bifurcation parameter λ has the greatest physical meaning in these applications, singularity and stability is regarded with respect to it. As an alternative to switching to Newton steplength control as discussed in Algorithm 8.1, singular points are calculated directly by introducing augmented maps in the spirit of MOORE and SPENCE [1980], GRIEWANK and REDDIEN [1984], RABIER and REDDIEN [1986]. Section 24.2 contains details concerning such a method.

Let us briefly illustrate the main points in this application. A point $(x_0, \lambda_0, \alpha_0)$ is a simple singular point of the equation (25.1) if the null space of $H_x(x_0, \lambda_0, \alpha_0)$ is of dimension 1, i.e., is spanned by a vector $\varphi_0 \neq 0$. The range is then given by

$$\text{range}H_x(x_0, \lambda_0, \alpha_0) := \{y \in \mathbb{R}^N : \psi_0^* y = 0\}$$

for some $\psi_0 \neq 0$.

A fold point of H is a simple singular point which satisfies

$$\psi_0^* H_\lambda(x_0, \lambda_0, \alpha_0) \neq 0.$$

If in addition

$$\psi_0^* H_{xx}(x_0, \lambda_0, \alpha_0)[\varphi_0, \varphi_0] \neq 0,$$

then it is called a *quadratic fold point*. The augmented system of equations used for calculating a fold point is that given by MOORE and SPENCE [1980]:

$$F(x, \varphi, \lambda, \alpha) := \begin{pmatrix} H(x, \lambda, \alpha) \\ H_x(x, \lambda, \alpha)\varphi \\ l^*\varphi - 1 \end{pmatrix} = 0 \tag{25.2}$$

where $l \in \mathbb{R}^N$ is not in the null space of $H_x(x_0, \lambda_0, \alpha_0)$. In general, from the curve following technique, good starting values for a Newton-type method for solving (25.2) are already available.

The equation (25.2) has a regular zero point at $(x_0, \varphi_0, \lambda_0, \alpha_0)$ if and only if it is a quadratic fold point of H in the above sense. A path of fold points can now be numerically traced by applying a continuation method with respect to one of the α parameters, see, e.g., RHEINBOLDT [1986, 1992a].

The above described approach was applied by WINTERS and CLIFFE [1985] to the equations modelling thermal ignition in a finite cylinder with aspect ratio $\gamma = \frac{R}{L}$ where R and L denote the radius and length of the

cylinder. Assume the cylinder has the origin as center and the z-axis as its rotational axis. Let us sketch their discussion as an example.

The non-dimensional form of the thermal equation takes the form

$$-\Delta u = \delta \exp\left(\frac{u}{1+\varepsilon u}\right)$$

where Δ is the Laplacian operator in cylindrical co-ordinates, δ is the exothermicity, and ε is the inverse activation energy. The conditions on the boundary of the cylinder are $u = 0$. In addition, the following symmetry conditions are imposed: $\frac{\partial u}{\partial z} = 0$ on the plane $z = 0$, and $\frac{\partial u}{\partial r} = 0$ on the axis $r = 0$. The replacement of z by γz leads to the equation

$$-\frac{1}{r}\frac{\partial}{\partial r} r \frac{\partial u}{\partial r} - \frac{1}{\gamma^2}\frac{\partial^2 u}{\partial z^2} = \delta \exp\left(\frac{u}{1+\varepsilon u}\right).$$

For a fixed value of γ, a starting point $\delta = 0$ and $u = 0$ is available on the solution curve for the discretized problem.

The peak temperature $\hat{u}$ of the discretized solution when plotted against δ has a characteristic S-shaped curve whose two fold points correspond to the points of ignition and extinction of combustion. The Euler-Newton continuation method was used to calculate the $(\hat{u}, \delta)$-solution curve for fixed values of ε, γ. The Moore-Spence system was used to calculate the fold points. Then in turn the Euler-Newton method was used to calculate curves of fold points.

An analogous approach may be used to handle many bifurcation problems. Let us illustrate this in the case of symmetry breaking bifurcation points. As a simple example we take the symmetry group $\mathbb{Z}_2 = \{1, S\}$ where S represents, e.g., a reflection. We assume that equivariance holds:

$$H(Sx, \lambda, \alpha) = SH(x, \lambda, \alpha).$$

The reflection mapping induces a natural decomposition

$$\mathbb{R}^N = X_s \oplus X_a$$

where $X_s := \{x \in \mathbb{R}^N : Sx = x\}$, $X_a := \{x \in \mathbb{R}^N : Sx = -x\}$ represent symmetric and anti-symmetric vectors of $\mathbb{R}^N$.

Now X_s is invariant under H and H_x, and so the analysis of the symmetric solutions of (25.1) can be carried out in this restricted context. However, the stability can be affected by the unstable eigenvalues corresponding to an anti-symmetric eigenvector. At a symmetric solution, the eigenvectors are either symmetric or anti-symmetric. A simple singular point having an

anti-symmetric eigenvector is called a *symmetry-breaking bifurcation point* provided

$$\psi_0^*(H_{x\lambda}(x_0,\lambda_0,\alpha_0)\varphi_0 + H_{xx}(x_0,\lambda_0,\alpha_0)[\varphi_0, v_\lambda] \neq 0$$

where v_λ satisfies $H_x(x_0,\lambda_0,\alpha_0)v_\lambda + H_\lambda(x_0,\lambda_0,\alpha_0) = 0$. At a symmetry-breaking bifurcation point a curve of anti-symmetric solutions bifurcates off from a curve of symmetric solutions.

The extended system of equations for calculating a symmetry-breaking bifurcation point proposed by WERNER and SPENCE [1984] takes the form

$$G(x,\varphi,\lambda,\alpha) := \begin{pmatrix} H(x,\lambda,\alpha) \\ H_x(x,\lambda,\alpha)\varphi \\ l^*\varphi - 1 \end{pmatrix} = 0 \tag{25.3}$$

where

$$G : X_s \times X_a \times \mathbb{R} \times \mathbb{R}^p \to X_s \times X_a \times \mathbb{R}$$

and $l \in \mathbb{R}^N$ is not in the null space of $H_x(x_0,\lambda_0,\alpha_0)$. The main difference between the maps in (25.3) and (25.2) is that now in (25.3), $x \in X_s$ and $\phi \in X_a$.

In WINTERS, CLIFFE, and JACKSON [1987] the successful application of the above described techniques to problems modelling Bénard convection, flow in curved tubes, Taylor flow between concentric rotating cylinders, periodic flow in semi-conductor crystal melt, and flow past a cylinder are reported. In general, the extended system (25.3) is used to accurately locate a symmetry-breaking bifurcation point (in a few Newton iterations), and then the continuation method is again employed to track a bifurcating solution curve.

Other recent numerical experience involving the application of path following techniques to the Navier-Stokes equations and nonlinear biharmonic equations have been given by GLOWINSKI, KELLER, and REINHART [1985] and HUITFELDT and RUHE [1990], see also Sections 17, 12.

26 Critical Points

Among the applications of numerical continuation methods is the calculation of critical points of a smooth mapping $f : \mathbb{R}^N \to \mathbb{R}$. In general, one chooses a smooth mapping $g : \mathbb{R}^N \to \mathbb{R}$ with known regular *critical points* $a \in \mathbb{R}^N$, i.e., $\nabla g(a) = 0$ for the gradient and the Hessian $\nabla g'(a)$ has full rank. One then formulates a smooth homotopy map $H : \mathbb{R}^{N+1} \to \mathbb{R}^N$ such that

$$H(0,x) = \nabla g(x) \quad \text{and} \quad H(1,x) = \nabla f(x).$$

Typically, one uses the convex homotopy

$$H(\lambda, x) := (1-\lambda)\nabla g(x) + \lambda \nabla f(x). \tag{26.1}$$

The numerical aspect then consists of tracing a smooth curve

$$c(s) = \big(\lambda(s), x(s)\big) \in H^{-1}(0)$$

with starting point $c(0) = (0, a)$ for some given critical point a of g, and starting tangent $\dot{c}(0) = \big(\dot{\lambda}(0), \dot{x}(0)\big)$ with $\dot{\lambda}(0) > 0$. The aim of course is to trace the curve c until the homotopy level $\lambda = 1$ is reached, at which a critical point of f is obtained. If all critical points of f are regular, then by Sard's Theorem 18.1 it is generally possible to make a choice of g such that zero is a regular value of H. The following result of Allgower and Georg [1980, Sec. 2.17] indicates that the continuation method has an appealing property which can permit targeting critical points having a specific Morse index.

Theorem 26.1 *Let $f, g : \mathbb{R}^N \to \mathbb{R}$ be smooth functions and let H be the convex homotopy of (26.1) which has zero as a regular value. Let $c(s) = \big(\lambda(s), x(s)\big) \in H^{-1}(0)$ be the smooth solution curve such that $c(0) = (0, a)$ where a is a regular critical point of g. Suppose that $\lambda(s)$ is increasing for $s \in [0, \bar{s}]$, $\lambda(\bar{s}) = 1$, and that the critical point $b := x(\bar{s})$ of ∇f is regular. Then the critical points a, b of g and f, respectively, have the same Morse index, i.e., the Hessians $\nabla g'(a)$ and $\nabla f'(b)$ have the same number of negative eigenvalues.*

Proof By differentiating $H(c(s)) = 0$ we have that the augmented Jacobian, see (1.7), satisfies

$$\begin{pmatrix} \dot{\lambda}(s) & \dot{x}(s)^* \\ H_\lambda\big(c(s)\big) & H_x\big(c(s)\big) \end{pmatrix} \begin{pmatrix} \dot{\lambda}(s) \\ \dot{x}(s) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} .$$

This implies that

$$\begin{pmatrix} \dot{\lambda}(s) & \dot{x}(s)^* \\ H_\lambda\big(c(s)\big) & H_x\big(c(s)\big) \end{pmatrix} \begin{pmatrix} \dot{\lambda}(s) & 0^* \\ \dot{x}(s) & \mathrm{Id} \end{pmatrix} = \begin{pmatrix} 1 & \dot{x}(s)^* \\ 0 & H_x\big(c(s)\big) \end{pmatrix} .$$

Since the determinant of the augmented Jacobian never changes sign, it follows from the above equation that $\det H_x\big(c(s)\big)$ changes sign exactly when $\dot{\lambda}(s)$ changes sign. The latter does not occur, and $H_x\big(c(s)\big)$ can have at most one eigenvalue equal to zero, since all points $c(s)$ are regular points of H. Using a result of perturbation theory, see, e.g., Dunford and Schwartz

[1963, p. 922], namely that an isolated eigenvalue of the symmetric matrix $H_x\big(c(s)\big)$ depends smoothly on s, we conclude that no eigenvalue of $H_x\big(c(s)\big)$ can change sign for increasing s. Hence $H_x(0,a) = \nabla g'(a)$ and $H_x(1,b) = \nabla f'(b)$ have the same number of negative eigenvalues. □

From the above proof it is clear that the convex homotopy (26.1) can be replaced by any other homotopy linking f and g.

More general studies of the behavior of critical points in parametric optimization are to be found in the books of JONGEN, JONKER, and TWILT [1983, 1986]. Also semi-infinite problems can in principle be regarded as parametric optimization problems, see, e.g., JONGEN and ZWIER [1985].

27 Complex Bifurcation

It has been observed by ALLGOWER [1984] and ALLGOWER and Georg [1983a] that folds in the λ co-ordinate of solution curves of $H(x,\lambda) = 0$ lead to bifurcation points in a setting of complex extension. This observation can be used to connect separated real components of $H^{-1}(0)$, and hence may serve as a tool to find additional solutions of the equation $H = 0$. HENDERSON [1985] and HENDERSON and KELLER [1990] study complex bifurcation in a general Banach space setting.

For simplicity, let us discuss some results for the finite-dimensional setting, see also ALLGOWER and GEORG [1990, Sec. 11.8].

Consider a smooth nonlinear problem of the form

$$H(\lambda, z) = 0, \quad H : \mathbb{R} \times \mathbb{C}^N \to \mathbb{C}^N,$$

where H is analytic in the complex variable $z = x + iy \in \mathbb{C}^N$. We denote the real and imaginary parts of H by H^r, H^i, respectively. Hence,

$$H^r = \frac{1}{2}(H + \bar{H}), \quad H^i = \frac{1}{2i}(H - \bar{H}).$$

THEOREM 27.1 *Let $c(s) = (\lambda(s), z(s)) = (\lambda(s), x(s) + iy(s))$ (s = arc-length) be a smooth solution curve of the equation $H = 0$, and let $c(0)$ be a* fold*, i.e., $\dot{\lambda}(s)$ changes signs at $s = 0$ such that $c(s)$ is a regular point of H for $s \neq 0$ and sufficiently small. Then $c(0)$ is a bifurcation point of the equation $H = 0$.*

PROOF We use the Cauchy-Riemann equations

$$H^r_y = -H^i_x \quad \text{and} \quad H^i_y = H^r_x\,.$$

In the following matrix representations we consider the block structure implied by the λ-axis, the real part of z, and the imaginary part of z, respectively.

The augmented Jacobian satisfies the equations

$$\begin{pmatrix} \dot\lambda & \dot x^* & \dot y^* \\ H_\lambda & H_x & H_y \end{pmatrix} \begin{pmatrix} \dot\lambda & 0^* & 0^* \\ \dot x & \mathrm{Id} & 0 \\ \dot y & 0 & \mathrm{Id} \end{pmatrix} = \begin{pmatrix} 1 & \dot x^* & \dot y^* \\ 0 & H_x^r & H_x^i \\ 0 & -H_x^i & H_x^r \end{pmatrix}$$

and

$$\dot\lambda \det \begin{pmatrix} \dot\lambda & \dot x & \dot y \\ H_\lambda & H_x & H_y \end{pmatrix} = \det \begin{pmatrix} H_x^r & H_x^i \\ -H_x^i & H_x^r \end{pmatrix} = \det H_{(x,y)} \ .$$

A standard argument shows that the last determinant is always nonnegative. Hence, the last equation implies that $\dot\lambda\big(c(s)\big)$ changes sign at $s = 0$, and consequently the determinant of

$$\begin{pmatrix} \dot\lambda & \dot x^* & \dot y^* \\ H_\lambda & H_x & H_y \end{pmatrix}$$

evaluated at $c(s)$ does the same.

Now, using a standard argument in degree theory, see KRASNOSEL'SKIĬ [1964] or RABINOWITZ [1971], it can be shown that $c(0)$ is a bifurcation point of $H = 0$. However, $c(0)$ is not necessarily a simple bifurcation point. □

Let us now also assume that H is real for real arguments, i.e.,

$$\overline{H(\lambda, z)} = H(\lambda, \bar z),$$

and denote the restriction of H to real arguments by $\tilde H$.

THEOREM 27.2 *Let the curve $s \longmapsto c(s) = \big(\lambda(s), x(s), 0\big)$ be a "real" solution curve of $H^{-1}(0)$ such that the point $\big(\lambda(s), x(s)\big)$ is a regular point of the real homotopy $\tilde H$ for all s* (including $s = 0$). *Suppose that $\big(\lambda(0), x(0)\big)$ is a* simple fold *of the equation $H = 0$, i.e., $\dot\lambda(0) = 0$ and $\ddot\lambda(0) \neq 0$. Then $c(0)$ is a simple bifurcation point of the equation $H = 0$.*

PROOF Since H is real, it is easy to see that

$$H_x^i(\lambda, x, 0) = 0\,, \quad H_{xx}^i(\lambda, x, 0) = 0\,, \quad H_\lambda^i(\lambda, x, 0) = 0 \tag{27.1}$$

holds for $x \in \mathbb{R}^N$ and $\lambda \in \mathbb{R}$. Hence, using the Cauchy-Riemann equations, the augmented Jacobian satisfies the equation

$$\begin{pmatrix} \dot\lambda & \dot x^* & \dot y^* \\ H_\lambda & H_x & H_y \end{pmatrix} = \begin{pmatrix} 0 & \dot x^* & 0 \\ H_\lambda^r & H_x^r & 0 \\ 0 & 0 & H_x^r \end{pmatrix} \tag{27.2}$$

at $c(0)$.

Since the rank of

$$\begin{pmatrix} 0 & \dot{x}^* \\ \tilde{H}_\lambda & \tilde{H}_x \end{pmatrix}$$

at $\big(\lambda(0), x(0)\big)$ is $N+1$, we conclude that the matrix

$$H_x^r\big(c(0)\big) = \tilde{H}_x\big(\lambda(0), x(0)\big)$$

has rank at least $N-1$. On the other hand, differentiating $H\big(c(s)\big) = 0$ at $s=0$ implies that $H_x^r\big(c(0)\big)\dot{x}(0) = 0$, and hence $\mathrm{rank} H_x^r\big(c(0)\big) = N-1$. Therefore the matrix in (27.2) has rank $2N$, and the Jacobian $H'\big(c(0)\big)$ has a two-dimensional kernel spanned by the vectors

$$\begin{pmatrix} 0 \\ \dot{x}(0) \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \dot{x}(0) \end{pmatrix} .$$

It remains to show that the non-degeneracy conditions for the second derivatives hold, see (24.1). Let e span the kernel of $H_x^r\big(c(0)\big)^*$. Then $\binom{0}{e}$ spans the kernel of $H'\big(c(0)\big)^*$. Furthermore,

$$e^* H_\lambda^r\big(c(0)\big) \neq 0 \tag{27.3}$$

since otherwise the above kernel would not have dimension one.

We now investigate whether the bilinear form

$$(\xi,\eta) \quad \longmapsto \quad (0,e^*)H''\big(c(0)\big)\left[\xi\begin{pmatrix} 0 \\ \dot{x}(0) \\ 0 \end{pmatrix}, \eta\begin{pmatrix} 0 \\ 0 \\ \dot{x}(0) \end{pmatrix}\right]$$

has one positive and one negative eigenvalue. Using (27.1) and the Cauchy-Riemann equations, a straightforward calculation shows that the above bilinear form reduces to

$$(\xi,\eta) \quad \longmapsto \quad 2\xi\eta\; e^* H_{xy}^i\big(c(0)\big)[\dot{x}(0), \dot{x}(0)] \,. \tag{27.4}$$

It is clear that the simple bilinear form (27.4) has one positive and one negative eigenvalue if and only if

$$e^* H_{xy}^i\big(c(0)\big)[\dot{x}(0), \dot{x}(0)] \neq 0 \,. \tag{27.5}$$

To show this, let us differentiate the equation $e^*\tilde{H}\big(\lambda(s), x(s)\big) = 0$ twice. Using the facts $e^*\tilde{H}_x\big(\lambda(0), x(0)\big) = 0$, $\tilde{H}_x\big(\lambda(0), x(0)\big)\dot{x}(0) = 0$ and $\dot{\lambda}(0) = 0$, we obtain

$$e^*\tilde{H}_\lambda\big(\lambda(0), x(0)\big)\ddot{\lambda}(0) + e^*\tilde{H}_{xx}\big(\lambda(0), x(0)\big)[\dot{x}(0), \dot{x}(0)] = 0 \,.$$

Since $\ddot{\lambda}(0) \neq 0$, we can conclude from (27.3) that $e^* \tilde{H}_{xx}\big(\lambda(0), x(0)\big) \neq 0$. Now (27.5) follows from the Cauchy-Riemann equations. □

Let us finally show that at bifurcation points at which one of the two solution branches is real, the corresponding curvatures are of opposite sign, and hence at such points a choice of following branches is available so that the λ co-ordinate is increasing.

THEOREM 27.3 *Under the assumptions of the previous thorem, let us now denote the "real" solution curve in $H^{-1}(0)$ by $c_1(s) =: \big(\lambda_1(s), x_1(s), 0\big)$ and the bifurcating solution curve by $c_2(s) =: \big(\lambda_2(s), x_2(s), y_2(s)\big)$. The curves are defined for s near 0, and $\bar{u} := c_1(0) = c_2(0)$ is the bifurcation point. Then $\ddot{\lambda}_1(0) = -\ddot{\lambda}_2(0)$.*

PROOF Let us denote by $c(s) =: \big(\lambda(s), x(s), y(s)\big)$ either of the two solution curves c_1 or c_2. Differentiating $H\big(c(s)\big) = 0$ twice with respect to s and taking $\dot{\lambda}(0) = 0$ into account yields

$$\begin{aligned} &H_\lambda(\bar{u})\ddot{\lambda}(0) + H_{xx}(\bar{u})[\dot{x}(0), \dot{x}(0)] + 2H_{xy}(\bar{u})[\dot{x}(0), \dot{y}(0)] \; + \\ &H_{yy}(\bar{u})[\dot{y}(0), \dot{y}(0)] + H_x(\bar{u})\ddot{x}(0) + H_y(\bar{u})\ddot{y}(0) = 0\,. \end{aligned} \tag{27.6}$$

Let e span the kernel of $H^r_x\big(c(0)\big)^*$ as in the previous proof. Multiplying (27.6) from the left with $(e^*, 0)$ and taking the Cauchy-Riemann equations and (27.1) into account, we obtain

$$e^* H^r_\lambda(\bar{u})\ddot{\lambda}(0) + e^* H^r_{xx}(\bar{u})[\dot{x}(0), \dot{x}(0)] + e^* H^r_{yy}(\bar{u})[\dot{y}(0), \dot{y}(0)] = 0\,. \tag{27.7}$$

Since

$$\dot{c}_1(0) = \big(0, \dot{x}_1(0), 0\big) \tag{27.8}$$

holds, it can be seen from (27.4) and (24.6) that

$$\dot{c}_2(0) = \big(0, 0, \pm\dot{x}_1(0)\big)\,. \tag{27.9}$$

Substituting (27.8), (27.9) into (27.7) we obtain

$$\begin{aligned} e^* H^r_\lambda(\bar{u})\ddot{\lambda}_1(0) + e^* H^r_{xx}(\bar{u})[\dot{x}_1(0), \dot{x}_1(0)] = 0\,, \\ e^* H^r_\lambda(\bar{u})\ddot{\lambda}_2(0) + e^* H^r_{yy}(\bar{u})[\dot{x}_1(0), \dot{x}_1(0)] = 0\,, \end{aligned}$$

respectively. Since $e^* H^r_\lambda(\bar{u}) \neq 0$, see (27.3), and since by the Cauchy-Riemann equations $H^r_{xx}(\bar{u}) = -H^r_{yy}(\bar{u})$ holds, the assertion follows. □

The last two theorems have been generalized by LI and WANG [1993b, Prop. 2.1] for complex folds.

28 Linear Eigenvalue Problems

In recent years many of the classical problems of numerical linear algebra have been re-examined in the context of homotopies and path following. One of the earliest contributors has been CHU [1984a, 1984b, 1986, 1988, 1990, 1991] In these papers iterative processes and matrix factorizations have been studied in the context of flows satisfying various differential equations. A typical example is the Toda flow which has been studied as a continuous analogue of the QR algorithm. A survey of these ideas has been given by WATKINS [1984]. According to Watkins, although it seems that the Toda flow and related flows yield insight into the workings of algorithms, they do not necessarily directly offer algorithms which are competitive with standard library algorithms that have been developed and refined over numerous years.

Surprisingly, LI and LI [1993], LI and RHEE [1989], LI, ZENG, and CONG [1992], LI, ZHANG, and SUN [1991] have been able to construct special implementations of homotopy methods which are now at least competitive with the library routines of EISPACK and IMSL for linear eigenvalue problems.

The versatility of homotopy methods permits their application also to generalized eigenvalue problems, see CHU, LI, and SAUER [1988] and non-symmetric matrices, see LI and ZENG [1992], LI, ZENG, and CONG [1992]. In this case complex eigenvalues are likely to arise, and it is necessary to invoke the idea of complex bifurcation, see Section 27.

As an example, let us briefly discuss the homotopy approach given by LI, ZHANG, and SUN [1991]. Consider a real symmetric tridiagonal matrix A. We assume that A is irreducible, since otherwise one off-diagonal element $A[i+1,i] = A[i,i+1]$ would vanish and the matrix A would split into two blocks which can be treated independently. We consider a homotopy $H : \mathbb{R}^N \times \mathbb{R} \times [0,1] \to \mathbb{R}^N \times \mathbb{R}$ defined by

$$H(x,\lambda,s) = \begin{pmatrix} \lambda x - [(1-s)D + sA]x \\ x^*x - 1 \end{pmatrix}.$$

Here D is a real symmetric reducible tridiagonal matrix which is generated from A by setting some of the off-diagonal entries of A to zero. The simplest example for D would be to set all off-diagonal entries to zero. However, it is advantageous to only reduce D to tridiagonal block structure with relatively small blocks, e.g., of size < 50. This technique is referred to as *divide and conquer*.

Since $A(s) := (1-s)D + sA$ is irreducible for all $s > 0$, the solution set of $H = 0$ consists of $2n$ disjoint smooth curves c (*eigenpaths*) which can be parametrized with respect to s. Note that s is not the arclength,

but the homotopy parameter. Hence $c(s) = (\pm x(s), \lambda(s))$ for $0 \leq s \leq 1$. The curves obviously occur in pairs, and of course only one of each pair needs to be traced. At the level $s = 0$, initial values on the curves can be obtained by approximating all eigenvectors and eigenvalues of all small blocks in D. If D is diagonal, this is trivial, and otherwise a QR routine has to be employed.

Let us sketch a typical step of the predictor corrector method. We note first that it follows from differentiation of $H(c(s)) = 0$ with respect to s that

$$\dot{\lambda}(s) = x(s)^*(A - D)x(s).$$

Assume that $\big(x(s), \lambda(s)\big)$ is (approximately) known. After having decided on a stepsize h (we are not going to discuss this feature), a predicted eigenvalue $\tilde{\lambda}(s + h)$ is obtained from the above differential equation by a two-step ODE method. Now a predicted eigenvector $\tilde{x}(s + h)$ is obtained by one step of the inverse power method with shift, i.e., solve $\big(A(s + h) - \tilde{\lambda}(s + h)\mathrm{Id}\big)y = x(s)$ for y and set $\tilde{x}(s + h) = y/||y||$. Then a Rayleigh quotient iteration is performed as a corrector to approximate $\big(x(s + h), \lambda(s + h)\big)$.

There are some stability problems for the case that different eigenvalues become close. Sturm sequences are computed to stabilize the procedure.

Let us finally note that the above homotopy method has an order preserving property, i.e., different λ-paths can never cross. Hence the j-th eigenvalue of A can be calculated without calculating any other eigenvalues. This is very often an advantageous feature for applications. On the other hand, the homotopy method lends itself conveniently to parallelization, since each solution path can be traced independently of the others and hence also simultaneously.

29 Mathematical Programming

29.1 Generalized Equations

The paper of Reinoza [1985] deals with homotopy methods for generalized equations. The concept of generalized equations has been introduced to describe inclusion relations involving multivalued functions, in particular normal cone operators. The type of *generalized equations* which are considered are of the form:

$$0 \in f(x) + N_C(x), \tag{29.1}$$

where $f : \Omega \to \mathbb{R}^N$ is a C^2 function on an open set $\Omega \subset \mathbb{R}^N$, C is a convex polyhedral set in $\mathbb{R}^N$ and $N_C(x)$ is the *normal cone operator*

$$N_C(x) := \begin{cases} \{v \in \mathbb{R}^N : \forall_{c \in C} \ \ (v, c - x) \leq 0\}, \\ \phi \text{ otherwise.} \end{cases}$$

Many problems arising in nonlinear programming can be formulated as generalized equations. To illustrate this, note that if (29.1) holds, then the sum is nonempty since it contains 0, and so $x \in C$. Also then $-f(x)$ must belong to $N_C(x)$ and so for each $c \in C$,

$$(-f(x), c - x) \leq 0.$$

Thus we see that (29.1) is equivalent to the *variational inequality*

$$x \in C, \text{ and } \forall_{c \in C} \ \ (f(x), c - x) \ \geq 0.$$

It is sometimes useful to formulate problems in mathematical programming in terms of generalized equations to facilitate their analysis. The survey paper by ROBINSON [1983] gives an extensive review of the theory and applications of generalized equations.

A generalized homotopy equation is a relation of the form

$$0 \in \Phi(x, \lambda) := H(x, \lambda) + N_C(x) \tag{29.2}$$

where $H : \Omega \to \mathbb{R}^N$ is a C^2 function on an open set $\Omega \subset \mathbb{R}^{N+1}$, C is a convex polyhedral set in $\mathbb{R}^N$ and $N_C(x)$ is the normal cone operator.

It turns out that under appropriate modifications and extensions of the ideas in Section 1 an analogue of Theorem 1.3 can be obtained. Let us give a brief synopsis of Reinoza's results. Let $K \subset \mathbb{R}^{N+1}$ be a cylinder defined by $K := (C \cap P) \times [-l, l]$ where $P \subset \mathbb{R}^N$ is a compact convex set with smooth boundary ∂P and $l \in \mathbb{R}$ is a positive number. A curve $\Gamma \subset K$ is called a *path* if it starts and ends on the boundary of K. The curve Γ is called a *closed loop* if it does not intersect the boundary of K. Under a somewhat technical definition for $0 \in \mathbb{R}^N$ to be a regular value of Φ, the following theorem is shown:

THEOREM 29.1 *Let* 0 *be a regular value of* Φ. *Then* $\Phi^{-1}(0) \cap K$ *is a finite union of disjoint paths and closed loops in* K.

In this more general context an analogue of Theorem 20.1 is also shown. An analogue of Newton's method for generalized equations can be formulated by considering the analogue of the linearization of (29.1):

$$0 \in f(x_0) + f'(x_0)(x - x_0) + N_C(x). \tag{29.3}$$

For such a Newton's method JOSEPHY [1979] gave a version of the Newton-Kantorovich theory, see also ROBINSON [1983]. REINOZA [1985] shows for the generalized global homotopy

$$0 \in \Phi(x, \lambda) := f(x) + \lambda f(x_0) + N_C(x) \tag{29.4}$$

that a sequence of points generated by the generalized homotopy method coincides with the sequence of points generated by Josephy's method. SELLAMI and ROBINSON [1993] recently presented an implementation of a continuation method for normal maps.

The *differentiable* techniques for generalized equations discussed above represent a bridge linking the predictor corrector methods to the piecewise linear methods which we will discuss in Chapter VI. It will be seen that piecewise linear methods also apply to certain multi-valued maps.

29.2 Parametric Programming Problems

Parametric programming problems and sensitivity analysis also can be studied in the context of continuation methods. Consider the problem

$$\min\{f(x, \alpha) : c_i(x, \alpha) = 0,\ i \in E,\ c_i(x, \alpha) \leq 0,\ i \in I\}, \tag{29.5}$$

where $f, c_i : \mathbb{R}^{N+1} \to \mathbb{R}$ are smooth functions. Here

$$E = \{1, \ldots, q\} \quad \text{and} \quad I = \{q+1, \ldots, q+p\}$$

denote the index sets for the equality and inequality constraints, respectively. The local sensitivity of such systems has been analysed, e.g., in FIACCO [1983, 1984] and ROBINSON [1987]. Many authors have used bifurcation and singularity theory to investigate the local behavior and persistence of minima at the singular points of the above system, see, e.g., BANK, GUDDAT, KLATTE, KUMMER, and TAMMER [1983], GFRERER, GUDDAT, and WACKER [1983], GFRERER, GUDDAT, WACKER, and ZULEHNER [1985], GUDDAT, GUERRA VASQUEZ, and JONGEN [1990], GUDDAT, JONGEN, KUMMER, and NOŽIČKA [1987], JONGEN, JONKER, and TWILT [1983, 1986], JONGEN and WEBER [1990], KOJIMA and HIRABAYASHI [1984], POORE and TIAHRT [1987, 1990]. RAKOWSKA, HAFTKA, and WATSON [1991] discuss algorithms for tracking paths of optimal solutions. LUNDBERG and POORE [1993] report on a numerical implementation of a path following method for the above problem. Our discussion is motivated by their exposition.

The Fritz John first order necessary conditions for (29.5) imply the existence of $(\lambda, \nu) \in \mathbb{R}^{p+q} \times \mathbb{R}$ such that

$$\mathcal{L}_x(x, \lambda, \nu, \alpha) = 0, \tag{29.6}$$

$$c_i(x,\alpha) = 0,\ i \in E, \tag{29.7}$$

$$\lambda_i c_i(x,\alpha) = 0,\ i \in I, \tag{29.8}$$

$$\nu \geq 0,\ c_i(x,\alpha) \leq 0,\ \lambda_i \geq 0,\ i \in I, \tag{29.9}$$

where $\mathcal{L}(x,\lambda,\nu,\alpha) = \nu f(x,\alpha) + \sum \lambda_i c_i(x,\alpha)$ is the Lagrangian.

Now an active set strategy is implemented by using the following homotopy equation for a path following algorithm:

$$H\left(x,\{\lambda_i\}_{i\in\mathcal{A}},\nu,\alpha\right) = \begin{pmatrix} \mathcal{L}_x\left(x,\{\lambda_i\}_{i\in\mathcal{A}},\nu,\alpha\right) \\ c_i(x,\alpha),\ i \in \mathcal{A} \\ \nu^2 + \sum \lambda_i^2 - 1 \end{pmatrix} = 0, \tag{29.10}$$

where $\mathcal{A}$ is the set of active constraints. Hence $\mathcal{A}$ includes all of the indices E and some of the indices I. During the path following procedure, this active set is adapted in such a way that the inequalities (29.9) are respected.

There are various technical difficulties (such as handling singularities or efficiently adapting the active set) which have to be overcome in a successful implementation.

Continuous deformations of semi-infinite optimization problems have been studied by Hettich and Jongen [1994].

29.3 Interior Point Methods

Khachiyan [1979] started a new class of polynomial time algorithms for solving the linear programming problem. Karmarkar [1984] subsequently gave a much noted polynomial time algorithm based upon projective rescaling. Gill, Murray, Saunders, Tomlin, and Wright [1986] noted that Karmarkar's algorithm is equivalent to a projected Newton barrier method which in turn is closely related to a recent class of polynomial time methods involving a continuation method, namely the tracing of the "central path". This last technique can be extended to quadratic programming problems, and both linear and nonlinear complementarity problems. Typically, algorithms of this nature are now referred to as *interior point* methods.

The presentation of a continuous trajectory (central path) of the iterative Karmarkar method was extensively studied by Bayer and Lagarias [1989], see also Sonnevend [1985]. Megiddo [1988] related this path to the classical barrier path of nonlinear optimization, Fiacco and McCormick [1968]. Several authors have proposed algorithms that generally follow the central path to a solution, see, e.g., Gonzaga [1988], Kojima, Mizuno, and Yoshise [1988, 1989], Monteiro and Adler [1989], Nazareth [1986, 1991, 1994b, 1994a], Renegar [1988a], Vaidya [1990].

To make the algorithms more efficient, variable steplength and/or higher order predictor algorithms have been proposed in Adler, Resende, Veiga,

and KARMARKAR [1989], MIZUNO, TODD, and YE [1992], SONNEVEND, STOER, and ZHAO [1989, 1991]. The algorithm of MIZUNO, TODD, and YE [1992] has subsequently been shown by YE, GÜLER, TAPIA, and ZHANG [1991] to have both polynomial time complexity and quadratic convergence. KOJIMA, MEGIDDO, and MIZUNO [1991a] believe that there still remain differences between the theoretical primal-dual algorithms which enjoy global and/or polynomial-time convergence and the efficient implementations of primal-dual algorithms, e.g., MARSTEN, SUBRAMANIAN, SALTZMAN, LUSTIG, and SHANNO [1990] and MCSHANE, MONMA, and SHANNO [1989].

ADLER, RESENDE, VEIGA, and KARMARKAR [1989] report extensive computational experiments for an interior point implementation with solution times being in most cases less than those required by a state-of-the-art simplex method MINOS, see MURTAGH and SAUNDERS [1987]. KARMARKAR and RAMAKRISHNAN [1991] report computational experience on large scale problems which are representative of large classes of applications of current interest. Their interior point implementation incorporates a preconditioned conjugate gradient method as a corrector step and is consistently faster than MINOS by orders of magnitude. Further computational experience comparing an interior point method `OB1` and a simplex method CPLEX is reported in technical reports BIXBY, GREGORY, LUSTIG, MARSTEN, and SHANNO [1991], CARPENTER and SHANNO [1991] and LUSTIG, MARSTEN, and SHANNO [1991]. POLAK, HIGGINS, and MAYNE [1992] have given an algorithm for solving semi-infinite minimax problems which bears a resemblance to the interior penalty function methods. They report numerical results which show that the algorithm is extremely robust and its performance is at least comparable to that of current first-order minimax algorithms.

There is currently immense activity in studying and developing implementations of interior point algorithms. It is to be expected that our brief account will be outdated in a few years. For further details and literature, we refer to the recent surveys of GONZAGA [1992], KOJIMA, MEGIDDO, NOMA, and YOSHISE [1991c], TODD [1989], WRIGHT [1992], and the proceedings edited by ROOS and VIAL [1991]. As an example, we outline the central path approach for a primal-dual linear programming problem, following the introductory parts of MONTEIRO and ADLER [1989] and MIZUNO, TODD, and YE [1992].

Consider the following linear programming problem and its corresponding dual form:

PROBLEM 29.1

$$\min_x \{c^* x : Ax = b,\ x \geq 0\}, \tag{29.11}$$

$$\max_y \{b^* y : A^* y + z = c,\ z \geq 0\}, \tag{29.12}$$

We make the following standard

ASSUMPTION 29.1 *The rank of A equals the number of its rows, and the interior feasible set of the primal-dual problem*

$$\mathcal{F}^o := \{(x, z) : x, z > 0,\ Ax = b,\ A^* y + z = c \text{ for some } y\}$$

is not empty.

It is well-known that the linear programming problem has a unique solution under the above assumptions. The logarithmic barrier function method associated with Problem 29.1 is

$$\min_x \{c^* x - \mu \sum_j \ln x_j : Ax = b,\ x > 0\}, \tag{29.13}$$

where $\mu > 0$ is the barrier penalty parameter. Under Assumption 29.1, the logarithmic barrier function is strictly convex and has a unique minimal point $x(\mu)$ for all $\mu > 0$. Moreover, $x(\mu)$ tends to the unique solution of Problem 29.1 as μ tends to zero.

The Karush-Kuhn-Tucker optimality condition which characterizes the solution $x(\mu)$ can be expressed in the following way: $(x(\mu), z)$ must belong to the set

$$\mathcal{C} := \{(x, z) \in \mathcal{F}^o : \text{diag}(x) z = \mu e\}, \tag{29.14}$$

where e denotes the column of ones. In fact, $\mathcal{C}$ is parametrized by μ and is commonly called the *central path* of the problem. It turns out that μ is related to the so-called *duality gap*: $c^* x - b^* y = x^* z$ via

$$\mu = \frac{x^* z}{N} \tag{29.15}$$

for $(x, z) \in \mathcal{C}$, where N is the number of columns of A.

By the above remarks, it is clear that the objective now is to follow the central path $\mathcal{C}$ as μ tends to zero. In fact, most interior point methods can be viewed, one way or another, as a special path following method along these lines. The methods differ in the choice of predictor step, corrector procedure (usually one or several Newton type iterations) and predictor steplength control. Many papers discussing such methods or introducing new methods also contain a sophisticated complexity analysis, see, e.g., Chapter VII.

The above interior point algorithms typically require a phase I in which a feasible starting point is generated. A somewhat different approach is taken

by FREUND [1991] who introduces a shifted barrier function approach so that the need for phase I is obviated.

Finally, the above technique is quite general and can be extended to quadratic programming problems and linear and nonlinear complementarity problems, see, e.g., KOJIMA, MEGIDDO, and YE [1992]. The literature on interior methods is rapidly increasing, and the subject has become one of the major topics of mathematical programming. In our opinion, it is only a question of time until the venerable simplex methods will be superceded by interior point implementations.

30 Numerical Integration over Curves

Although the predictor corrector curve tracing method has found a wide variety of significant applications, the calculation of line integals over implicitly defined curves seems to have been overlooked. In view of the interesting possiblities of applications involving Stoke's theorem, complex integration and boundary integrals, this omission seems somewhat surprising. The recent dissertation of SIYYAM [1994] investigates the efficient blending of predictor corrector methods and numerical quadrature methods. In this section we outline some of the ideas for this task and discuss some of the results of the Siyyam dissertation.

One of the interesting cases for the line integral is the integration over a closed curve. In order to handle this case for an implicitly defined curve, it is necessary to develop a reliable numerical method for determining when the curve has been completely traversed. This can be achieved by means of an adaptation of the methods for computing special points on an implicitly defined curve which were discussed in Section 8.

Another aspect which has to be addressed is how to deal with the lack of an explicit parametrization of the curve in the numerical quadrature for the line integral. We will discuss two basic approaches for handling this aspect. One involves essentially constructing a local high order approximation via, e.g., the ideas given by MACKENS [1989], see also Section 11. The other approach involves adapting a modified trapezoidal rule to the numerical integration over curves. Such a modified trapezoidal rule was developed by GEORG [1991] to approximate surface integrals for implicitly defined surfaces, see also Section 41, but it can obviously be adapted to the easier task of handling curve integrals.

Let us first deal briefly with the task of giving a numerical method for determining when an implicitly defined closed curve has been completely traversed. We suppose that $H : \mathbb{R}^{N+1} \to \mathbb{R}^N$ is a smooth map and $\mathcal{C} \subset H^{-1}(0)$ is a smooth closed curve with a given starting point $u \in \mathcal{C}$. We

suppose further that t is the tangent to $\mathcal{C}$ at the starting point u in the direction in which the traversing of $\mathcal{C}$ begins. The points z on the hyperplane which is normal to $\mathcal{C}$ at u satisfy the equation

$$f(z) := t^*(z - u) = 0. \tag{30.1}$$

Thus the functional f defined in equation (30.1) changes sign each time the curve $\mathcal{C}$ passes through the hyperplane $f(z) = 0$ during the traversing process. Each time f changes sign for two successive points u_n, u_{n+1}, the curve tracing algorithm is switched over to a Newton steplength control, see (8.2), using the functional $f(z)$ in the context of the discussion in Section 8.1. In general, the Newton steplength control algorithm will converge within very few iterations to a point w which satisfies both $H(w) = 0$ and $f(w) = 0$. When in addition, $||w - u||$ is within a chosen tolerance, then u_{n+1} is replaced by the starting point u and the traversing is stopped. Otherwise the curve tracing is resumed at u_{n+1} in the same direction as before.

30.1 The Modified Trapezoidal Rule for Curves

We assume that F is a continuous vector field in $\mathbb{R}^{N+1}$ which is defined on a neighborhood of $\mathcal{C}$. The line integral which we want to numerically approximate is:

$$\int_{\mathcal{C}} F \cdot d\mathcal{C}, \tag{30.2}$$

where a particular orientation of $\mathcal{C}$ will hereafter be assumed. If u_i and u_{i+1} are two consecutive points on $\mathcal{C}$ which have been generated by a numerical traversing, and $\mathcal{C}_i$ is the section of $\mathcal{C}$ between u_i and u_{i+1}, the *modified trapezoidal rule* for curves is given by:

$$\int_{\mathcal{C}_i} F \cdot d\mathcal{C} \approx (1/2)[F(u_i) + F(u_{i+1})] \cdot [u_{i+1} - u_i]. \tag{30.3}$$

For a sequence of points $\{u_i\}_{i=1}^n$ obtained via a traversing of $\mathcal{C}$, the composite version of the modified trapezoidal rule takes the form:

$$\int_{\mathcal{C}} F \cdot d\mathcal{C} \approx (1/2) \sum_{i=1}^{n-1} [F(u_i) + F(u_{i+1})] \cdot [u_{i+1} - u_i], \tag{30.4}$$

where u_1 and u_n are the end points of $\mathcal{C}$, and both are identified with the starting point if $\mathcal{C}$ is a closed curve.

Let us point out the distinction between the modified trapezoidal rule and the standard trapezoidal rule for integrating over a curve $\mathcal{C}$ which has

an explicit parametrization, say $\mathcal{C} = \{c(s) : 0 \leq s \leq 1\}$. Then the line integral (30.2) can be written as an ordinary integral, i.e.,

$$\int_{\mathcal{C}} F \cdot d\mathcal{C} = \int_0^1 F(c(s)) \cdot c'(s)\, ds. \tag{30.5}$$

For the integral on the right hand side of (30.5) the standard composite trapezoidal rule can be applied over any partitioning of the interval $[0, 1]$. For this case, if $c'(s)$ is smooth on $[0, 1]$ and if the interval is uniformly partitioned with a mesh length, say $1/n$, it is well known that the truncation error has an expansion in even powers of $1/n$, and hence extrapolation methods (in the sense of Romberg integration) can be applied.

The advantage of the modified trapezoidal rule (30.3) is that it obviates the need for derivatives, which are generally only available at an additional computational expense. However, it is now no longer clear that the use of extrapolation methods are justified for the modified trapezoidal rule.

Let us begin by formulating what needs to be shown. We suppose again that u and w are two consecutive points on $\mathcal{C}$ which have been obtained in the numerical curve tracing process. More precisely, we suppose that $v = u + ht$ is the predictor point for which the corrector process yields the limit point w. Similarly, the corrector process applied to the points $v_j := u + t(jh/n)$ for any positive integer n and $j = 0, \ldots, n$ yields corresponding limit points w_j, $j = 1, \ldots, n$ on $\mathcal{C}$ to which the composite version of the modified trapezoidal rule (30.4) can then be applied for integrating over the section $\mathcal{C}_i$ of the curve $\mathcal{C}$ which lies between the points u and w. Now the truncation formula which needs to be shown is essentially

$$\int_{\mathcal{C}_i} F \cdot d\mathcal{C} - \frac{1}{2} \sum_{j=1}^{n-1} [F(w_j) + F(w_{j+1})] \cdot [w_{j+1} - w_j] \sim \sum_{k=1}^{\infty} c_k \left(\frac{1}{n}\right)^{2k} \tag{30.6}$$

for some constants c_k which are independent of n. In fact, it turns out that the c_k are polynomials in h. This was proven by Siyyam [1994], whose proof is a modification for the one-dimensional case of a result of Verlinden and Cools [1993].

Siyyam has made a Matlab implementation of a Romberg like method for the modified trapezoidal rule and performed successful tests on a number of examples. As one would expect, it is not crucial to obtain an intial approximation with many points on the implicitly defined curve over which the integration is to be carried out, since successive refinements are made anyway in performing the Romberg procedure.

30.2 Taylor Approximation Methods

Other approaches to the task of numerically integrating over an implicitly defined curve generally involve obtaining a high order local approximation to the curve and then performing a highly accurate integration method such as Gaussian quadrature. We mention two such approximation methods. One is to consider the Taylor expansion of $c(h)$ about $h = 0$, where $c(h)$ is defined as the point on the curve such that $c(h) - (u + ht)$ is orthogonal to t. In Section 11 we describe how the Taylor coefficients in

$$c(h) \sim \sum_{i=0}^{k} \frac{1}{i!} c^{(i)}(0) h^i \tag{30.7}$$

can be successively approximated using finite difference formulae for higher derivatives. Finally, the Taylor approximation (30.7) is inserted into the line integral (30.5) and this is in turn approximated via Gaussian quadrature.

This method has been implemented and studied by Siyyam. He also implemented a steplength and order control in the spirit of Section 9. The general conclusions reached in his study were as follows:

1. Both of the above mentioned implementations succeeded in yielding highly accurate approximations to line integrals over smooth implicitly defined curves.

2. Romberg extrapolation in connection with the modified trapezoidal rule generally required the least amount of computational effort to obtain a highly accurate numerical approximation of the line integral.

3. The method employing the variable order predictor and variable steplength control in connection with the Taylor polynomial approximation generally performed well for a moderately accurate numerical approximation of the line integral.

A second possible approach is to use interpolation predictors, see Section 10, instead of Taylor polynomials. We do not have any numerical experience with this approach.

Chapter VI

Piecewise Linear Methods

Up to now we have assumed that the map $H : \mathbb{R}^{N+1} \to \mathbf{R}^N$ was smooth. Next we will discuss piecewise linear methods which can again be viewed as curve tracing methods, but which can be applied to non-smooth situations. The piecewise linear methods trace a polygonal path which is obtained by successively stepping through certain "transversal" cells of a piecewise linear manifold. The first and most prominent example of a piecewise linear algorithm was designed by LEMKE and HOWSON [1964] and LEMKE [1965] to calculate a solution of the linear complementarity problem, see Section 38. This algorithm played a crucial role in the development of subsequent piecewise linear algorithms. SCARF [1967] gave a numerically implementable proof of the Brouwer fixed point theorem, based upon Lemke's algorithm. EAVES [1972] observed that a related class of algorithms can be obtained by considering piecewise linear approximations of homotopy maps. Thus the piecewise linear continuation methods began to emerge as a parallel to the classical embedding or predictor corrector methods.

The piecewise linear methods require no smoothness of the underlying equations and hence have, at least in theory, a more general range of applicability than classical embedding methods. In fact, they can be used to calculate fixed points of set-valued maps. They are more combinatorial in nature and are closely related to the topological degree, see PEITGEN and SIEGBERG [1981]. piecewise linear continuation methods are usually considered to be less efficient than the predictor corrector methods when the latter are applicable, especially in higher dimensions. The reasons for

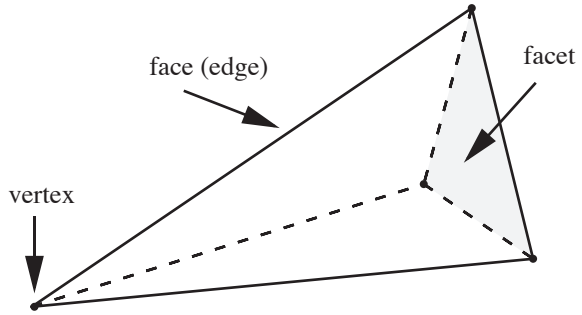

Figure 31.1. A 3-simplex, a special 3-cell

this lie in the fact that steplength adaptation and exploitation of special structure are more difficult to implement for piecewise linear methods.

EAVES [1976] has given a very elegant geometric approach to general piecewise-linear methods, see also EAVES and SCARF [1976]. We adopt this point of view and cast the notion of piecewise linear algorithms into the general setting of subdivided manifolds which we will call *piecewise linear manifolds*. Our exposition follows the introduction of GEORG [1990] to some extent.

31 Basic Facts

Let $\mathbf{E}$ denote some ambient finite dimensional Euclidean space which contains all points arising in the sequel. A *half-space* η and the corresponding *hyperplane* $\partial\eta$ are defined by $\eta = \{y \in \mathbf{E} : x^*y \leq \alpha\}$ and $\partial\eta = \{y \in \mathbf{E} : x^*y = \alpha\}$, respectively, for some $x \in \mathbf{E}$ with $x \neq 0$ and some $\alpha \in \mathbb{R}$. A finite intersection of half-spaces is called a *cell*. If σ is a cell and ξ a half-space such that $\sigma \subset \xi$ and $\tau := \sigma \cap \partial\xi \neq \emptyset$, then the cell τ is called a *face* of σ. For reasons of notation we consider σ also to be a face of itself, and all other faces are *proper* faces of σ. The *dimension* of a cell is the dimension of its affine hull. In particular, the dimension of a singleton is 0 and the dimension of the empty set is -1. If the singleton $\{v\}$ is a face of σ, then v is called a *vertex* of σ. If τ is a face of σ such that $\dim\tau = \dim\sigma - 1$, then τ is called a *facet* of σ, see Figure 31.1. The *interior* of a cell σ consists of all points of σ which do not belong to a proper face of σ.

A *piecewise linear manifold* of dimension N is a system $\mathcal{M} \neq \emptyset$ of cells of dimension N such that the following conditions hold, see Figure 31.2:

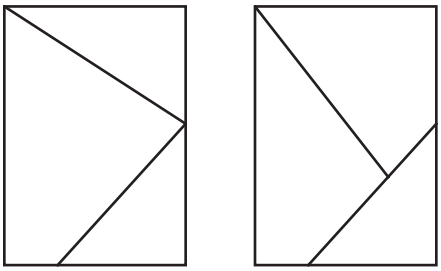

Figure 31.2. Condition 1 is not satisfied

1. If $\sigma_1, \sigma_2 \in \mathcal{M}$, then $\sigma_1 \cap \sigma_2$ is a common face of σ_1 and σ_2.
2. A cell τ of dimension $N-1$ can be a facet of at most two cells in $\mathcal{M}$.
3. The family $\mathcal{M}$ is *locally finite*, i.e., any relatively compact subset of

$$|\mathcal{M}| := \bigcup_{\sigma \in \mathcal{M}} \sigma \tag{31.1}$$

meets only finitely many cells $\sigma \in \mathcal{M}$.

The simplest example of a piecewise linear manifold is $\mathbb{R}^N$ subdivided into unit cubes with integer vertices.

We introduce the *boundary* $\partial\mathcal{M}$ of $\mathcal{M}$ as the system of facets which are common to exactly one cell of $\mathcal{M}$. Generally, we cannot expect $\partial\mathcal{M}$ to again be a piecewise linear manifold. However, this is true for the case that $|\mathcal{M}|$ is convex. Two cells which have a common facet τ are called *adjacent*. We say that one cell is *pivoted* into the other cell across the facet τ. We will see that piecewise linear algorithms perform pivoting steps.

Typical for piecewise linear path following is that only one current cell is stored in the computer, along with some additional data, and the pivoting step is performed by calling a subroutine which makes use of the data to determine an adjacent cell which then becomes the new current cell.

A cell of particular interest is a *simplex* $\sigma = [v_1, v_2, \ldots, v_{N+1}]$ of dimension N which is defined as the convex hull of $N+1$ affinely independent points $v_1, v_2, \ldots, v_{N+1} \in \mathbf{E}$. These points are the vertices of σ. If a piecewise linear manifold $\mathcal{M}$ of dimension N consists only of simplices, then $\mathcal{M}$ is called a *pseudo manifold* of dimension N. Such manifolds are of special importance, see, e.g., GOULD and TOLLE [1983], TODD [1976a]. If a pseudo

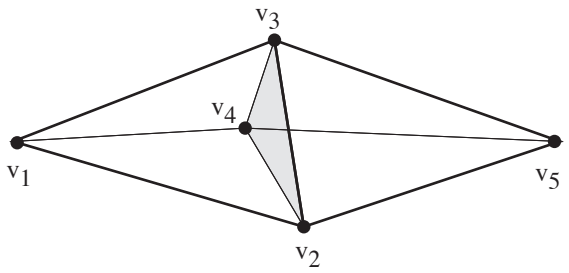

Figure 31.3: The simplex $[v_1\,,\,v_2\,,\,v_3\,,\,v_4]$ is pivoted into the simplex $[v_5\,,\,v_2\,,\,v_3\,,\,v_4]$ across the facet $[v_2\,,\,v_3\,,\,v_4]$, and the vertex v_1 is pivoted into the vertex v_5

manifold $\mathcal{T}$ subdivides a set $|\mathcal{T}|$, then we also say that $\mathcal{T}$ *triangulates* $|\mathcal{T}|$. We will use the notions pseudo manifold and triangulation somewhat synonymously. Some triangulations of $\mathbb{R}^N$ of practical importance were already considered by COXETER [1934] and FREUDENTHAL [1942], see also TODD [1976a]. EAVES [1984] gave an overview of standard triangulations.

If σ is a simplex in a pseudo manifold $\mathcal{T}$ and τ a facet of σ which is not in the boundary of $\mathcal{T}$, then there is exactly one simplex $\tilde{\sigma}$ in $\mathcal{T}$ which is different from σ but has the same facet τ, and there is exactly one vertex v of σ which is not a vertex of $\tilde{\sigma}$. We call v the vertex of σ opposite τ. There is also exactly one vertex $\tilde{v}$ of $\tilde{\sigma}$ opposite τ. We say that σ is pivoted across τ into $\tilde{\sigma}$, and that the vertex v of σ is pivoted into $\tilde{v}$. See Figure 31.3.

A simple triangulation can be generated by the following pivoting rule (pivoting by reflection), see ALLGOWER and GEORG [1979] or COXETER [1963]: if

$$\sigma = [v_1, v_2, \ldots, v_i \ldots, v_{N+1}]$$

is a simplex in $\mathbf{R}^N$, and τ is the facet opposite a vertex v_i, then σ is pivoted across τ into $\tilde{\sigma} = [v_1, v_2, \ldots, \tilde{v}_i \ldots, v_{N+1}]$ by setting

$$\tilde{v}_i = \begin{cases} v_{i+1} + v_{i-1} - v_i & \text{for} \quad 1 < i < N+1, \\ v_2 + v_{N+1} - v_1 & \text{for} \quad i = 1, \\ v_N + v_1 - v_{N+1} & \text{for} \quad i = N+1. \end{cases} \tag{31.2}$$

In fact, a minimal (non-empty) system of N-simplices in $\mathbb{R}^N$ which is closed under the above pivoting rule is a triangulation of $\mathbb{R}^N$.

32 Special Triangulations

In this section we present the details for pivoting in two important examples of triangulations:

1. The Coxeter-Freudenthal-Kuhn (CFK) Triangulation K_1, see Coxeter [1934], Freudenthal [1942], Kuhn [1968],
2. The Union Jack triangulation J_1 of Todd [1976a].

We will describe these triangulations and formulate their pivoting rules via pseudo codes in the general context of the space $\mathbb{R}^q$. For an extensive discussion of the triangulations which are used in the context of piecewise linear homotopy methods, we refer the reader to the monograph of Eaves [1984]. For the purpose of our discussion, we will ignore the mesh size and present the triangulations via integer nodes $v \in \mathbb{Z}^q$.

A simplex $\sigma \in K_1$ (the CFK Triangulation) of $\mathbb{R}^q$ can be expressed as the convex hull of affinely independent vertices v_i, $i = 1, \ldots, q+1$, which satisfy a recursion formula

$$v_{i+1} = v_i + e_{\pi(i)}, \ i = 1, \ldots, q \tag{32.1}$$

for a permutation $\pi : \{1, \ldots, q\} \to \{1, \ldots, q\}$. The simplex σ can be compactly stored via the pair of integer tuples: $(v_1; \pi)$. We denote the q-simplex σ with the vertices v_i, $i = 1, \ldots, q+1$ by $\sigma = [v_1, \ldots, v_{q+1}]$. It is easy to see that the set of all q-simplices σ of this form is a pseudo manifold which triangulates $\mathbb{R}^q$, and that the simplices are all geometrically isometric. See Figure 32.1.

The pivoting steps in K_1 can be equivalently described in terms of either reflections, or in terms of the permutation representation (32.1).

The pseudo codes below are formulated under the assumption that at each step a decision has been made for determining which vertex is to be pivoted next, e.g., the linear programming steps of Section 34 may furnish such a decision. Our first pseudo code is based on pivoting by reflection. We denote the cyclic permutation $(1, 2, \ldots, q+1)$ by ρ.

Algorithm 32.1 **(Pivoting in K_1 by Reflection)**

INPUT $[v_1, v_2, \ldots, v_{q+1}] \subset \mathbb{R}^q$ % *starting simplex*

REPEAT

ENTER $i \in \{1, 2, \ldots, q+1\}$ % *index of vertex to be pivoted next*

$v_i \leftarrow v_{\rho^{-1}(i)} - v_i + v_{\rho(i)}$ % *reflection rule*

UNTIL pivoting is stopped

The pivoting rules in K_1 can also be equivalently performed by interchange permutations.

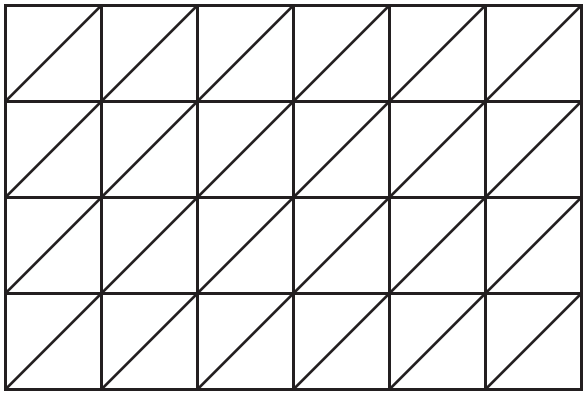

Figure 32.1. The CFK Triangulation in $\mathbb{R}^2$

ALGORITHM 32.2 **(Pivoting in K_1 by Interchange Permutations)**

INPUT $[v_1, v_2, \ldots, v_{q+1}] \subset \mathbb{R}^q$ % *starting simplex*

$u_j \leftarrow \begin{cases} v_{j+1} - v_j & \text{for } j = 1, \ldots, q \\ v_1 - v_{q+1} & \text{for } j = q+1 \end{cases}$ % *standard axes*

FOR $j = 1, \ldots, q+1$ DO $\pi(j) \leftarrow j$ % *initial permutation*

REPEAT

ENTER $i \in \{1, 2, \ldots, q+1\}$ % *index of vertex to be pivoted next*

$v_i \leftarrow v_{\rho^{-1}(i)} + u_{\pi(i)}$ % *pivoting rule*

interchange $\pi\big(\rho^{-1}(i)\big)$ and $\pi(i)$

UNTIL pivoting is stopped

The above pseudo codes illustrate that pivoting in K_1 can be performed in integer arithmetic, and in the case of a piecewise linear homotopy algorithm, at any given stage, only one simplex has to be stored. On the other hand, in the case of a piecewise linear algorithm for approximating a surface or manifold of dimension greater than 1, the generated simplices can be cheaply stored in terms of integer vectors (see, e.g., Section 40.2).

In our next example, we give similar descriptions of the triangulation J_1. One of the advantageous features of J_1 over K_1 is that it carries less directional bias and offers more symmetry. This is useful in the piecewise linear approximation of surfaces which have symmetry. The nodes of J_1 are again given by the points $v \in \mathbb{Z}^q$, i.e., having integer co-ordinates. A q-simplex $\sigma \subset \mathbb{R}^q$ belongs to the triangulation J_1 if the following rules are obeyed:

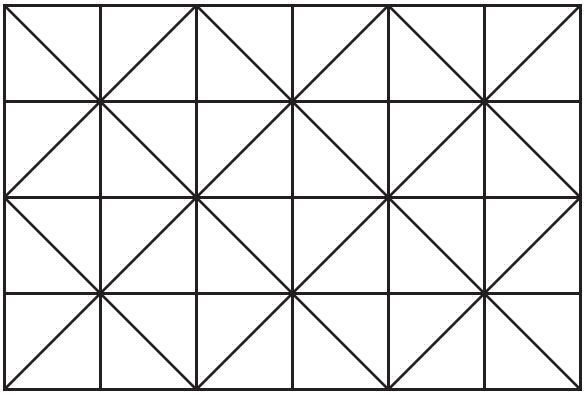

Figure 32.2. The Union Jack Triangulation in $\mathbb{R}^2$

1. the vertices of σ are nodes of J_1 in the above sense;

2. the vertices of σ can be ordered in such a way that they are given by the following recursion formula

$$v_{j+1} = v_j + s(j)\, e_{\pi(j)}, \quad j = 1, \ldots, q\,, \tag{32.2}$$

where $e_1, \ldots, e_q$ is the standard unit basis of $\mathbb{R}^q$, $\pi : \{1, 2, \ldots, q\} \to \{1, 2, \ldots, q\}$ is a permutation and $s : \{1, 2, \ldots, q\} \to \{+1, -1\}$ is a sign function;

3. the *central vertex* v_1 has odd integer co-ordinates.

From Figure 32.2 it is apparent why TODD [1976a] names this triangulation the *Union Jack Triangulation*.

From the description of the pivoting rules below it is evident that these conditions define a pseudo manifold $\mathcal{T}$ which triangulates $\mathbb{R}^q$. The formal proof is however somewhat technical, and we refer the reader to TODD [1976a]. Analogously to Algorithms 32.1 and 32.2, let us now describe the pivoting rules in J_1:

ALGORITHM 32.3 **(Pivoting in J_1 by Reflections)**

INPUT $[v_1, v_2, \ldots, v_{q+1}] \subset \mathbb{R}^q$ % *starting simplex*

REPEAT

ENTER $i \in \{1, 2, \ldots, q+1\}$ % *index of vertex to be pivoted next*

$$v_i \leftarrow \begin{cases} 2v_2 - v_1 & \text{for } i = 1 \\ 2v_q - v_{q+1} & \text{for } i = q+1 \\ v_{i-1} - v_i + v_{i+1} & \text{else} \end{cases}$$

UNTIL pivoting is stopped

Similarly to the discussion for pivoting in Freudenthal's triangulation, the pivoting rules for J_1 can also be obtained by interchange permutations:

ALGORITHM 32.4 **(Pivoting in J_1 by Interchange Permutations)**

INPUT $[v_1, v_2, \ldots, v_{q+1}] \subset \mathbb{R}^q$ % *starting simplex*
FOR $j = 1$ to q DO
  $u_j \leftarrow v_{j+1} - v_j$ % *standard axes*
  $\pi(j) \leftarrow j$ % *initial permutation*
  $s(j) \leftarrow 1$ % *initial sign function*
END FOR
REPEAT
  ENTER $i \in \{1, 2, \ldots, q+1\}$ % *index of vertex to be pivoted next*
  IF $i = 1$ % *consider different cases*
    $v_1 \leftarrow v_2 + s(1)\, u_{\pi(1)}$
    $s(1) \leftarrow -s(1)$
  ELSE IF $i = q+1$
    $v_{q+1} \leftarrow v_q - s(q)\, u_{\pi(q)}$
    $s(q) \leftarrow -s(q)$
  ELSE $v_i \leftarrow v_{i-1} + s(i)\, u_{\pi(i)}$
    interchange $s(i-1)$ and $s(i)$
    interchange $\pi(i-1)$ and $\pi(i)$
  END IF
UNTIL pivoting is stopped

It is possible to compare different triangulations via various ways of measuring their efficiency. Such results can be found in TODD [1976a], SAIGAL [1977, 1979], VAN DER LAAN and TALMAN [1980], ALEXANDER and SLUD [1983], ALEXANDER [1987], EAVES and YORKE [1984], EAVES [1984]. One of these measures is the *thickness of a triangulation* (see also (40.6)).

33 Piecewise Linear Algorithms

Let $\mathcal{M}$ be a piecewise linear manifold of dimension $N+1$. We call $H : |\mathcal{M}| \to \mathbb{R}^N$ a *piecewise linear map* if the restriction $H_\sigma : \sigma \to \mathbb{R}^N$ of H to σ is an affine map for all $\sigma \in \mathcal{M}$. In this case, H_σ can be uniquely extended

to an affine map on the affine space spanned by σ. The Jacobian H'_σ has the property $H'_\sigma(x-y) = H_\sigma(x) - H_\sigma(y)$ for x, y in this affine space. Note that under an appropriate choice of basis H'_σ corresponds to an $(N, N+1)$-matrix which has a one-dimensional kernel in case of non-degeneracy, i.e., if its rank is maximal.

If $\mathcal{M}$ is a pseudo manifold triangulating a set $X = |\mathcal{M}|$, and if $\tilde{H} : X \to \mathbb{R}^k$ is a map, then the *piecewise linear approximation* of $\tilde{H}$ (with respect to $\mathcal{M}$) is defined as the unique piecewise linear map $H : X \to \mathbb{R}^k$ which coincides with $\tilde{H}$ on all vertices of $\mathcal{M}$, i.e., $\tilde{H}$ is affinely interpolated on the simplices of $\mathcal{M}$.

A piecewise linear algorithm is a method for following a polygonal path in $H^{-1}(0)$. To avoid degeneracies, we introduce a concept of regularity, see EAVES [1976]. A point $x \in |\mathcal{M}|$ is called a *regular point* of H if x is not contained in any face of dimension $< N$, and if H'_τ has maximal rank N for all facets τ. A value $y \in \mathbb{R}^N$ is a *regular value* of H if all points in $H^{-1}(y)$ are regular. By definition, y is vacuously a regular value if it is not contained in the range of H. If a point or value is not regular it is called *singular*.

The following analogue of Sard's theorem 18.1 holds for piecewise linear maps, see, e.g., EAVES [1976] or PEITGEN and SIEGBERG [1981] for details. This enables us to confine ourselves to regular values. We note that degeneracies could be handled via the concept of lexicographical ordering, see DANTZIG [1963], TODD [1976a].

THEOREM 33.1 (Perturbation Theorem) *Let $H : \mathcal{M} \to \mathbb{R}^N$ be a piecewise linear map where $\mathcal{M}$ is a piecewise linear manifold of dimension $N+1$. Then for any relatively compact subset $C \subset |\mathcal{M}|$ there are at most finitely many $\varepsilon > 0$ such that $C \cap H^{-1}(\vec{\varepsilon})$ contains a singular point of H. Consequently, $\vec{\varepsilon}$ is a regular value of H for almost all $\varepsilon > 0$. Here we use the notation*

$$\vec{\varepsilon} := \begin{pmatrix} \varepsilon \\ \varepsilon^2 \\ \vdots \\ \varepsilon^N \end{pmatrix}.$$

PROOF The proof will be given by contradiction. Let us assume that there is a strictly decreasing sequence $\{\varepsilon_i\}_{i\in\mathcal{N}}$ of positive numbers, converging to zero, for which a bounded sequence $\{x_i\}_{i\in\mathcal{N}} \subset |\mathcal{M}|$ of singular points can be found such that the equations

$$H(x_i) = \vec{\varepsilon}_i \tag{33.1}$$

for $i \in \mathcal{N}$ are satisfied. For any subset $I \subset \mathcal{N}$ of cardinality $N+1$ we see that the $\{\vec{\varepsilon}_i\}_{i\in I}$ are affinely independent, and by (33.1) and the piecewise linearity of H the $\{x_i\}_{i\in I}$ cannot all be contained in the same lower dimensional face $\tau \in \mathcal{M}^k$ for $k < N$. Since this holds for all index sets I, we use this argument repeatedly, and the local finiteness of $\mathcal{M}$ permits us to find a strictly increasing function $\nu : \mathcal{N} \to \mathcal{N}$ (to generate a subsequence), and to find a face $\sigma \in \mathcal{M}^{N+1} \cup \mathcal{M}^N$ such that the subsequence $\{x_{\nu(i)}\}_{i\in\mathcal{N}}$ is contained in the interior of σ. But now we can again use the above argument: for an index set $I \subset \nu(\mathcal{N})$ of cardinality $N+1$ the $\{\vec{\varepsilon}_i\}_{i\in I}$ are affinely independent, and we conclude that H'_σ has maximal rank N. However, this means that all points $\{x_{\nu(i)}\}_{i\in\mathcal{N}}$ are regular, a contradiction to the choice of $\{x_i\}_{i\in\mathcal{N}}$. The last assertion of the perturbation theorem follows since $|\mathcal{M}|$ can be written as a countable union of relatively compact subsets. □

Let 0 be a regular value of H. This implies that $H^{-1}(0)$ consists of polygonal paths whose vertices are always in the interior of some facet. If σ is a cell, then $\sigma \cap H^{-1}(0)$ is a segment (two end points), a ray (one end point), a line (no end point) or empty. The latter two cases are not of interest for piecewise linear path following. A step of the method consists of following the ray or segment from one cell into a uniquely determined adjacent cell. The method is typically started at a point of the boundary or on a ray (coming from infinity), and it is typically terminated at a point of the boundary or in a ray (going to infinity), see Figure 33.1. The numerical linear algebra required to perform one step of the method is typical for linear programming and usually involves $\mathcal{O}(N^2)$ operations for dense matrices.

On the other hand, even if 0 is not a regular value of H, the above theorem helps us to do something similar. Namely, $\sigma \cap H^{-1}(\vec{\varepsilon})$ is a segment (two end points) for all sufficiently small $\varepsilon > 0$, a ray (one end point) for all sufficiently small $\varepsilon > 0$, a line (no end point) for all sufficiently small $\varepsilon > 0$ or empty for all sufficiently small $\varepsilon > 0$. This leads us to the following definition: we call a facet τ *completely labeled* with respect to H, if $\tau \cap H^{-1}(\vec{\varepsilon}) \neq \emptyset$ for all sufficiently small $\varepsilon > 0$. We call a cell σ *transverse* with respect to H, if $\sigma \cap H^{-1}(\vec{\varepsilon}) \neq \emptyset$ for all sufficiently small $\varepsilon > 0$. Instead of following the paths $H^{-1}(0)$ for a regular value 0, we now follow more specifically the *regularized paths*

$$\bigcup\{H^{-1}(0) \cap \sigma : \sigma \text{ transverse}\}.$$

Of course, this set coincides with $H^{-1}(0)$ for the case that 0 is a regular value of H.

For $\varepsilon > 0$ sufficiently small and $\vec{\varepsilon}$ a regular value of H, a node of the polygonal paths $H^{-1}(\vec{\varepsilon})$ corresponds to a completely labeled facet (which

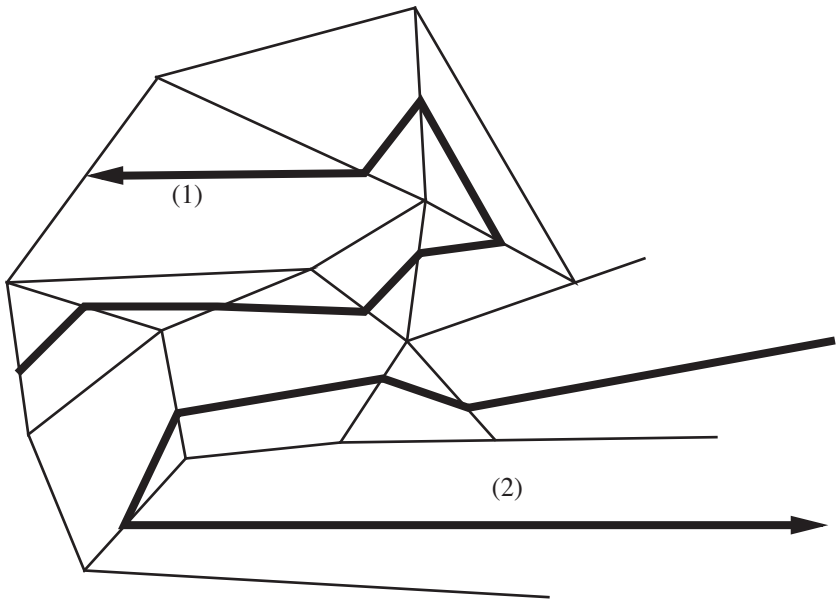

Figure 33.1:
(1) Boundary Start with Boundary Termination
(2) Ray Start with Ray Termination

is intersected), and hence the piecewise linear algorithm traces such completely labeled facets belonging to the same cell. The method is usually started either on the boundary, i.e., in a completely labeled facet $\tau \in \partial\mathcal{M}$, or on a ray, i.e., in a transverse cell $\sigma \in \mathcal{M}$ which has only one completely labeled facet. We are thus led to the following two generic versions of a piecewise linear algorithm.

ALGORITHM 33.1 **(PL Algorithm with Boundary Start)**

INPUT $\tau_1 \in \partial\mathcal{M}$ completely labeled % *starting facet*
FIND the unique $\sigma_1 \in \mathcal{M}$ such that $\tau_1 \subset \sigma_1$
FOR $i = 1, 2, 3, \ldots$
  IF τ_i is the only completely labeled facet of σ_i
    STOP % *ray termination*
  ELSE % *piecewise linear step*
    FIND the other completely labeled facet τ_{i+1} of σ_i
  END IF
  IF $\tau_{i+1} \in \partial\mathcal{M}$
    STOP % *boundary termination*
  ELSE % *pivoting step*
    pivot σ_i across τ_{i+1} into σ_{i+1}
  END IF
END FOR

ALGORITHM 33.2 **(PL Algorithm with Ray Start)**

INPUT $\sigma_1 \in \mathcal{M}$ transverse % *starting cell*
  having exactly one completely labeled facet
FIND the completely labeled facet τ_2 of σ_1 % *initial PL step*
FOR $i = 2, 3, \ldots$
  IF $\tau_i \in \partial\mathcal{M}$
    STOP % *boundary termination*
  ELSE % *pivoting step*
    pivot σ_{i-1} across τ_i into σ_i
  END IF
  IF τ_i is the only completely labeled facet of σ_i
    STOP % *ray termination*
  ELSE % *piecewise linear step*
    FIND the other completely labeled facet τ_{i+1} of σ_i
  END IF
END FOR

34 Numerical Considerations

From a numerical point of view, two steps of a piecewise linear algorithm have to be efficiently implemented. Usually, a current cell σ and a completely labeled facet τ of σ is stored via some characteristic data.

A pivoting step consists of finding the adjacent cell $\tilde{\sigma}$ with the same facet τ. The implementation of this step is dependent on the special piecewise linear manifold under consideration. But typically this step is performed by only a few operations. The pivoting rule (31.2) is a simple example, see also Section 32.

A piecewise linear step consists of finding a second completely labeled facet $\tilde{\tau}$ of σ (if it exists, otherwise we have ray termination). This is usually computationally more expensive than the pivoting rule and typically involves some numerical linear algebra.

Let us consider an example. We assume that a cell of dimension $N+1$ is given by

$$\sigma := \{x \in \mathbb{R}^{N+1} : Lx \geq c\},$$

where $L : \mathbb{R}^{N+1} \to \mathbb{R}^m$ is a linear map and $c \in \mathbb{R}^m$ is a given value. Furthermore, let us assume that

$$\tau_i := \{x \in \mathbb{R}^{N+1} : Lx \geq c,\ e_i^* Lx = e_i^* c\},$$

for $i = 1, 2, \ldots, m$ is a numbering of all the facets of σ.

On the cell σ, the piecewise linear map $H : \mathcal{M} \to \mathbb{R}^N$ reduces to an affine map, and hence there is a linear map $A : \mathbb{R}^{N+1} \to \mathbb{R}^N$ and a vector $b \in \mathbb{R}^N$ such that the segment of the path in σ can be written as

$$\sigma \cap H^{-1}(0) = \{x \in \mathbb{R}^{N+1} : Ax = b,\ Lx \geq c\}. \tag{34.1}$$

Let τ_i be completely labeled. This implies that the rank of A is N. If we exclude degeneracies, then $\tau_i \cap H^{-1}(0) = \{x_0\}$ is a singleton, and there is a unique vector t in the one-dimensional kernel $A^{-1}(0)$ such that $e_i^* Lt = -1$. Since x_0 is in the interior of τ_i (by excluding degeneracies), we have $e_j^* Lx_0 > e_j^* c$ for $j = 1, \ldots, m$, $j \neq i$, and hence $x_0 - \lambda t$ is in the interior of σ for small $\lambda > 0$.

If (34.1) is a ray, then $e_j^* L(x_0 - \lambda t) > e_j^* c$ for all $\lambda > 0$. Otherwise we have $e_j^* Lt > 0$ for at least one index j, and since we are excluding degeneracies, the minimization

$$k := \arg\min \left\{ \frac{e_j^*(Lx_0 - c)}{e_j^* Lt} : j = 1, \ldots, m,\ e_j^* Lt > 0 \right\} \tag{34.2}$$

yields the unique completely labeled facet τ_k of σ with $k \neq i$. For the minimum

$$\lambda_0 := \frac{e_k^*(Lx_0 - c)}{e_k^* Lt} > 0$$

we obtain: $\sigma \cap H^{-1}(0) = \{x_0 - \lambda t : 0 \leq \lambda \leq \lambda_0\}$.

Minimizations such as (34.2) are typical for linear programming, and the numerical linear algebra can be efficiently handled by standard routines. Successive linear programming steps can often make use of previous matrix factorizations via update methods, see, e.g., GILL, GOLUB, MURRAY, and SAUNDERS [1974]. In the case of a pseudo manifold $\mathcal{M}$ where the cell σ is a simplex, it is convenient to handle the numerical linear algebra with respect to the barycentric co-ordinates based on the vertices of σ. Then the equations become particularly simple, see, e.g., ALLGOWER and GEORG [1990, Sec. 12.2–12.4] or TODD [1976a] for details.

We describe the simplest case (without taking degeneracies into account) which is sufficient for most implementations. Let $\sigma = [v_1, \ldots, v_{N+2}]$ in $\mathbb{R}^{N+1}$ be a simplex of dimension $N+1$, and let us denote by τ_i the facet obtained via deleting vertex v_i in σ. We further introduce the labeling matrix

$$L := \begin{pmatrix} 1 & \ldots & 1 & 1 \\ H(v_1) & \ldots & H(v_{N+1}) & H(v_{N+2}) \end{pmatrix}$$

and the submatrix L_i obtained via deleting column i in L. Then τ_i is completely labeled with respect to H if and only if L_i^{-1} exists and is *lexicographically positive*, i.e., the first non-zero entry in each row is positive. For numerical purposes, it can usually be assumed that the situation is non-degenerate, i.e., that the first column of L_i^{-1} consists of positive entries. In this case, we have a solution $L\alpha = e_1$ such that $\alpha(j) > 0$ for $j \neq i$ and $\alpha(i) = 0$. Let γ be a solution of $L\gamma = 0$ such that $\gamma(i) < 0$. The idea now is to find a largest possible $t > 0$ such that $\alpha - t\gamma$ has non-negative entries. This leads to the minimization

$$k := \arg\min_j \left\{ \frac{\alpha(j)}{\gamma(j)} : \gamma(j) > 0 \right\}.$$

If this minimization yields a unique k, then τ_k is the other completely labeled facet of σ.

We now give some examples of how the piecewise linear path following methods are used.

35 Piecewise Linear Homotopy Algorithms.

Let us see how the above ideas can be used to approximate zero points of a map $G : \mathbb{R}^N \to \mathbb{R}^N$ by applying piecewise linear methods to an appropriate homotopy map. In order to also allow for applications to optimization problems or other nonlinear programming problems, we consider the case where G is not necessarily continuous, e.g., G might be a selection of a multi-valued map. For the case that $\bar{x}$ is a point of discontinuity of G, we have to generalize the notion of a zero point in an appropriate way, i.e., $0 \in G_\Sigma(\bar{x})$ in the terminology introduced below.

EAVES [1972] presented the first piecewise linear homotopy method. A restart method based on somewhat similar ideas was developed by MERRILL [1972]. A number of authors have studied the efficiency and complexity of piecewise linear homotopy algorithms, see, e.g., ALEXANDER [1987], EAVES and YORKE [1984], SAIGAL [1977, 1984], SAIGAL and TODD [1978], SAUPE [1982], TODD [1982, 1986].

As an example of a piecewise linear homotopy algorithm, let us sketch the algorithm of EAVES and SAIGAL [1972]. We consider a triangulation $\mathcal{T}$ of $\mathbb{R}^N \times (0, 1]$ into $(N+1)$-simplices σ such that every simplex is contained in some slab $\mathbb{R}^N \times [2^{-k}, 2^{-k-1}]$ for $k = 0, 1, \ldots$. Let us call the maximum of the last co-ordinates of all vertices of σ the *level* of σ. We call $\mathcal{T}$ a *refining* triangulation if for $\sigma \in \mathcal{T}$, the diameter of σ tends to zero as the level of σ tends to zero. Of course, the main point here is to obtain a triangulation which is easily implemented. The first such triangulation was proposed by EAVES [1972]. TODD [1976a] gave a triangulation with refining factor 1/2, see Figure 35.1. Subsequently, many triangulations with arbitrary refining factors were developed, see EAVES [1984]. To insure success (i.e., convergence) of the algorithms, it is necessary to assume a boundary condition. For this we will follow a presentation of GEORG [1982] which uses a quite general boundary condition extending somewhat that of MERRILL [1972].

Let us first introduce some notation. For $x \in \mathbb{R}^N$ we denote by $\mathcal{U}(x)$ the system of neighborhoods of x. By $\overline{\text{co}}(X)$ we denote the closed convex hull of a set $X \subset \mathbb{R}^N$. By $\mathbb{R}^N_\Sigma$ we denote the system of compact convex non-empty subsets of $\mathbb{R}^N$. We call the map $G : \mathbb{R}^N \to \mathbb{R}^N$ *asymptotically linear* if the following three conditions hold:

1. G is *locally bounded*, i.e., each point $x \in \mathbb{R}^N$ has a neighborhood $U \in \mathcal{U}(x)$ such that $G(U)$ is a bounded set.

2. G is *differentiable at* ∞, i.e., there exists a linear map $G'_\infty : \mathbb{R}^N \to \mathbb{R}^N$ such that $||x||^{-1}||G(x) - G'_\infty x|| \to 0$ for $||x|| \to \infty$.

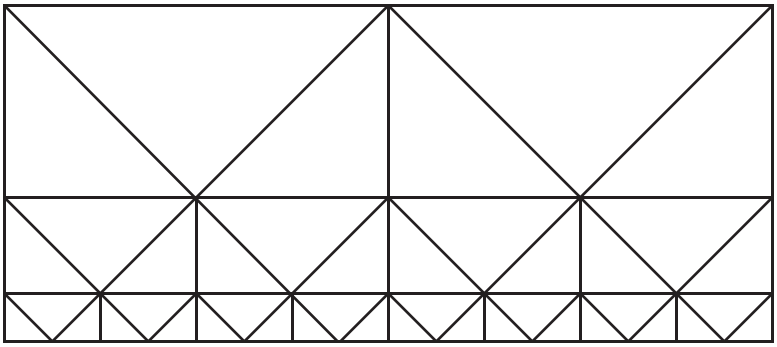

Figure 35.1. Refining Triangulation

3. G'_∞ is nonsingular.

If a map $G : \mathbb{R}^N \to \mathbb{R}^N$ is locally bounded, then we can define its *set-valued hull* $G_\Sigma : \mathbb{R}^N \to \mathbb{R}^N_\Sigma$ by setting

$$G_\Sigma(x) := \bigcap_{U \in \mathcal{U}(x)} \overline{\text{co}}\Big(G(U)\Big)\,.$$

It is not difficult to see that G_Σ is upper semi-continuous, and that G is continuous at x if and only if $G_\Sigma(x)$ is a singleton. By using a degree argument, see, e.g., Górniewicz [1976], on the set-valued homotopy

$$H_\Sigma(x, \lambda) := (1 - \lambda)G'_\infty x + \lambda G_\Sigma(x)\,,$$

it can be seen that G_Σ has at least one zero point, i.e., a point $\bar{x}$ such that $0 \in G_\Sigma(\bar{x})$. It is known, see, e.g., Peitgen and Prüfer [1979], Peitgen and Siegberg [1981], that degree arguments in nonlinear analysis are essentially constructive. Our aim here is to approximate this solution numerically.

We now construct a piecewise linear homotopy for an asymptotically linear map $G : \mathbb{R}^N \to \mathbb{R}^N$. First we define $\tilde{H} : \mathbb{R}^N \times [0, \infty) \to \mathbb{R}^N$ by setting

$$\tilde{H}(x, \lambda) := \begin{cases} G'_\infty(x - x_1) & \text{for } \lambda = 1, \\ G(x) & \text{for } \lambda < 1. \end{cases}$$

Here x_1 is a chosen starting point of the method. Then we consider a refining triangulation $\mathcal{T}$ of $\mathbb{R}^N \times (0, 1]$ as above, and we use the piecewise linear approximation H of $\tilde{H}$ (with respect to $\mathcal{T}$) to trace the polygonal path in $H^{-1}(0)$ which contains the starting point $(x_1, 1)$.

The boundary $\partial\mathcal{T}$ is a pseudo manifold which triangulates the sheet $\mathbb{R}^N \times \{1\}$. If we assume that the starting point $u_1 := (x_1, 1)$ is in the interior of a facet $\tau_1 \in \partial\mathcal{T}$, then it is immediately clear that τ_1 is the only completely labeled facet of $\partial\mathcal{T}$. Hence, the piecewise linear algorithm started in τ_1 can not terminate in the boundary, and since all cells of $\mathcal{T}$ are compact, it cannot terminate in a ray. Hence, it has no termination. Thus the piecewise linear algorithm generates a sequence $\tau_1, \tau_2, \ldots$ of completely labeled facets of $\mathcal{T}$. Let us also consider the polygonal path generated by the piecewise linear algorithm. This path is characterized by the nodes $(x_1, \lambda_1), (x_1, \lambda_2), \ldots$ such that (x_i, λ_i) is the unique zero point of the piecewise linear homotopy H in τ_i for $i = 1, 2, \ldots$. The resulting algorithm, i.e., applying Algorithm 33.1 to the above homotopy H, is due to EAVES [1972] and EAVES and SAIGAL [1972].

We call $\bar{x} \in \mathbb{R}^N$ an *accumulation point* of the algorithm if

$$\liminf_{i\to\infty} ||x_i - \bar{x}|| = 0.$$

The following convergence theorem holds:

THEOREM 35.1 *The set A of accumulation points of the Eaves-Saigal algorithm is compact, connected and non-empty. Each point $\bar{x} \in A$ is a zero point of G_Σ, i.e., we have $0 \in G_\Sigma(\bar{x})$*

PROOF From the construction of the piecewise linear map H it follows that

$$\lim_{||x||\to\infty} ||x||^{-1} ||H(x, \lambda) - G'_\infty x|| = 0$$

uniformly for $\lambda \in (0, 1]$. Since G'_∞ is non-singular, $H(x_i, \lambda_i) = 0$ implies that the sequence x_i is bounded. Hence the set A is non-empty and compact.

Let us assume that A can be written as a disjoint union of two non-empty compact sets A_1 and A_2. Then $\operatorname{dist}(A_1, A_2) > 0$, and

$$\liminf_{i\to\infty} \operatorname{dist}(x_i, A_j) = 0 \text{ for } j = 1, 2.$$

On the other hand, $\lim_{i\to\infty} \operatorname{dist}(x_i, A) = 0$, and the refining property of the triangulation $\mathcal{T}$ implies that $\lim_{i\to\infty} ||x_i - x_{i+1}|| = 0$. This leads to a contradiction, and hence A is connected.

Since a piecewise linear manifold is locally bounded and the τ_i stay in a bounded set, it follows that the level of the τ_i tends to 0 for $i \to \infty$. Hence, for i sufficiently large, the definition of H and the fact that the facets τ_i are completely labeled implies that $0 \in \overline{\text{co}} G\big(\pi_N(\tau_i)\big)$. Here $\pi_N : \mathbb{R}^N \times (0, 1] \to \mathbb{R}^N$ denotes the canonical projection.

Since a point $\bar{x} \in A$ is an accumulation point of the sequence x_i, and since $\lim_{i\to\infty} \operatorname{diam}(\tau_i) = 0$, we have that for each neighborhood $U \in \mathcal{U}(\bar{x})$ there is an i (arbitrarily large) such that $\pi_N(\tau_i) \subset U$ and hence

$$0 \in \overline{\text{co}} G\big(\pi_N(\tau_i)\big) \subset \overline{\text{co}} G(U).$$

Intersecting over all $U \in \mathcal{U}(\bar{x})$ gives $0 \in G_\Sigma(\bar{x})$. □

As a consequence, if the set-valued hull G_Σ has only isolated zero points, then the sequence x_i generated by the Eaves-Saigal algorithm converges to a zero point of G_Σ.

As a simple example, we consider the situation of the celebrated Brouwer fixed point theorem, see BROUWER [1912]. Let $F : C \to C$ be a continuous map on a convex, compact, non-empty subset $C \subset \mathbb{R}^N$ with non-empty interior. We define an asymptotically linear map $G : \mathbb{R}^N \to \mathbb{R}^N$ by setting

$$G(x) := \begin{cases} x - F(x) & \text{for } x \in C, \\ x - x_1 & \text{for } x \notin C. \end{cases}$$

Here, a point x_1 in the interior of C is used as a starting point. The above piecewise linear algorithm generates a point $\bar{x} \in \mathbb{R}^N$ such that $0 \in G_\Sigma(\bar{x})$. If $\bar{x} \notin C$, then $G_\Sigma(\bar{x}) = \{\bar{x} - x_1\}$, but $\bar{x} \neq x_1$ implies that this case is impossible. If $\bar{x}$ is an interior point of C, then $G_\Sigma(\bar{x}) = \{\bar{x} - F(\bar{x})\}$, and hence $\bar{x}$ is a fixed point of F. If $\bar{x}$ is in the boundary ∂C, then $G_\Sigma(\bar{x})$ is the convex hull of $\bar{x} - x_1$ and $\bar{x} - F(\bar{x})$, and hence $\bar{x} = (1-\lambda)x_1 + \lambda F(\bar{x})$ for some $0 \leq \lambda \leq 1$. But $\lambda < 1$ would imply that $\bar{x}$ is an interior point of C, and hence we have $\lambda = 1$, and again $\bar{x}$ is a fixed point of F. Hence, the above piecewise linear homotopy algorithm generates a fixed point of F in either case. Many similar asymptotically linear maps can be constructed which correspond to important nonlinear problems, see, e.g., ALLGOWER and GEORG [1990, Chap. 13].

36 Mixing PL and Newton Steps

As we have seen above, the refining triangulation used in the Eaves-Saigal algorithm is very convenient for discussing the question of convergence. If no stopping is allowed, it generates a sequence of nodes (x_n, λ_n) for $n = 0, 1, 2, \ldots$. We have seen that x_n converges to a zero point $\bar{x}$ of G_Σ under reasonable and very weak assumptions. Without additional assumptions on G however, nothing can be said about the rate of convergence of the x_n. BROOKS [1980] has shown that infinite retrogression (yoyoing) can occur. To ensure linear convergence, assumptions in the spirit of the

Newton-Kantorovitch theorems, see ORTEGA and RHEINBOLDT [1970], are necessary. Such convergence discussions have been given by SAIGAL [1977]. It can be seen that there is a close interrelationship between piecewise linear steps and Newton's method. Several other papers discuss techniques of mixing piecewise linear and Newton steps in order to accelerate a piecewise linear homotopy algorithm, see, e.g., SAIGAL and TODD [1978] and TODD [1978b, 1980]. In the context of piecewise linear continuation methods, i.e., when a whole curve $c(s)$ is to be approximated by a polygonal path. SAUPE [1982] has considered a mixing of piecewise linear and predictor corrector steps.

An elementary way of mixing piecewise linear and Newton steps was given by GEORG [1982], see also ALLGOWER and GEORG [1990, Sec. 13.5]. It is based on the simple observation that a modified Newton's method expressed in barycentric co-ordinates leads to a system of linear equations which is closely related to the linear equations obtained in a piecewise linear step described in Section 34.

More precisely, the following can be seen. Let $G : \mathbb{R}^N \to \mathbb{R}^N$ be a map and $\tau = [z_1, z_2, \ldots, z_{N+1}] \subset \mathbb{R}^N$ an N-simplex. Let $B := G'_\tau$ denote the Jacobian of the piecewise linear approximation G_τ of G with respect to τ, i.e., B is the finite difference approximation of G' using the values of G on $z_1, z_2, \ldots, z_{N+1}$. We assume that B is nonsingular and define a modified Newton step $\mathcal{N} : \mathbb{R}^N \to \mathbb{R}^N$ by

$$\mathcal{N}(x) := x - B^{-1}G(x)\,. \tag{36.1}$$

Then

$$G_\tau \mathcal{N}(z_i) = 0 \quad \text{for} \quad i = 1, 2, \ldots, N+1\,.$$

Furthermore, for any $z_{N+2} \in \mathbb{R}^N$, let

$$L(\tau, z_{N+2}) := \begin{pmatrix} 1 & \ldots & 1 & 1 \\ G(z_1) & \ldots & G(z_{N+1}) & G(z_{N+2}) \end{pmatrix}$$

and consider the barycentric co-ordinates β such that

$$L\beta = e_1, \qquad \beta(N+2) = -1.$$

Then

$$\mathcal{N}(z_{N+2}) = \sum_{j=1}^{N+2} \beta(j) z_j\,.$$

The above equations show that the piecewise linear steps of the Eaves-Saigal algorithm, see the end of Section 34, can be combined with Newton-type steps (36.1), since they are based on the same labeling matrix. In

particular, no initialization is necessary when switching between the two methods. Thus it is possible to design an algorithm which enjoys both the global convergence features mentioned in Theorem 35.1 and the fast local convergence of Newton-like methods (if the map G is sufficiently smooth and regular at the solution point). See ALLGOWER and GEORG [1990, Algorithm 13.5.2] for more details.

37 Index and Orientation

Nearly all piecewise linear manifolds $\mathcal{M}$ which are of importance for practical implementations, are orientable. If $\mathcal{M}$ is orientable and of dimension $N+1$, and if $H : \mathcal{M} \to \mathbb{R}^N$ is a piecewise linear map, then it is possible to introduce an index for the piecewise linear solution manifold $H^{-1}(0)$ which has important invariance properties and occasionally yields some useful information, see EAVES [1976], EAVES and SCARF [1976], LEMKE and GROTZINGER [1976], SHAPLEY [1974], TODD [1976c]. It should be noted that this index is closely related, see, e.g., PEITGEN [1982], to the topological index which is a standard tool in topology and nonlinear analysis. Occasionally, index arguments are used to guarantee a certain qualitative behavior of the solution path. There are many ways to introduce the index. Our discussion is similar to that in ALLGOWER and GEORG [1990, Sec. 14.2].

We begin with some basic definitions. Let $\mathbf{F}$ be a linear space of dimension k. An *orientation* of $\mathbf{F}$ is a function $\mathrm{or} : \mathbf{F}^k \to \{-1, 0, 1\}$ such that the following conditions hold:

1. $\mathrm{or}(b_1, \ldots, b_k) \neq 0$ if and only if $b_1, \ldots, b_k$ are linearly independent.
2. $\mathrm{or}(b_1, \ldots, b_k) = \mathrm{or}(c_1, \ldots, c_k) \neq 0$ if and only if the transformation matrix between $b_1, \ldots, b_k$ and $c_1, \ldots, c_k$ has positive determinant.

It is clear from the basic facts of linear algebra that any finite dimensional linear space permits exactly two orientations.

Let σ be a cell of dimension k and $\operatorname{aff}\sigma$ its affine hull. We introduce the k-dimensional linear space $\operatorname{tng}\sigma := \{x - y : x, y \in \operatorname{aff}\sigma\}$ as the *tangent space* of σ. The cell σ is oriented by orienting this tangent space. Such an orientation or_σ of σ *induces an orientation* $\mathrm{or}_{\tau,\sigma}$ on a facet τ of σ by the following convention:

$$\mathrm{or}_{\tau,\sigma}(b_1, \ldots, b_{k-1}) := \mathrm{or}_\sigma(b_1, \ldots, b_k)$$

whenever b_k points from τ into the interior of the cell σ. It is routine to check that the above definition of $\mathrm{or}_{\tau,\sigma}$ indeed satisfies the definition of an orientation.

If $\mathcal{M}$ is a piecewise linear manifold of dimension $N+1$, then an *orientation of* $\mathcal{M}$ is a choice of orientations $\{\mathrm{or}_\sigma\}_{\sigma\in\mathcal{M}}$ such that

$$\mathrm{or}_{\tau,\sigma_1} = -\mathrm{or}_{\tau,\sigma_2} \tag{37.1}$$

for each τ which is a facet of two different cells $\sigma_1, \sigma_2 \in \mathcal{M}$. By making use of the standard orientation

$$\mathrm{or}\big(b_1,\dots,b_N\big) := \mathrm{sign}\det\big(b_1,\dots,b_N\big)$$

of $\mathbb{R}^N$, it is clear that any piecewise linear manifold of dimension N which subdivides a subset of $\mathbb{R}^N$ is oriented in a natural way.

If $H : \mathcal{M} \to \mathbb{R}^N$ is a piecewise linear map on a piecewise linear manifold of dimension $N+1$ such that zero is a regular value of H, then it is clear that the system

$$\ker H := \{\sigma \cap H^{-1}(0)\}_{\sigma\in\mathcal{M}}$$

is a 1-dimensional piecewise linear manifold which subdivides the solution set $H^{-1}(0)$. For the case that $\mathcal{M}$ is oriented, the orientation of $\mathcal{M}$ and the natural orientation of $\mathbb{R}^N$ induce an orientation of $\ker H$. Namely, for $\xi \in \ker H$, $v \in \mathrm{tng}(\xi)$ and $\sigma \in \mathcal{M}$ such that $\xi \subset \sigma$, the definition

$$\mathrm{or}_\xi(v) := \mathrm{or}_\sigma(b_1,\dots,b_N,v)\ \mathrm{sign}\det\big(H'_\sigma b_1\,,\ \dots\ ,\ H'_\sigma b_N\big) \tag{37.2}$$

is independent of the special choice of $b_1,\dots,b_N \in \mathrm{tng}(\sigma)$, provided the $b_1,\dots,b_N$ are linearly independent. Clearly, an orientation of the 1-dimensional manifold $\ker H$ is just a rule which indicates a direction for traversing each connected component of $\ker H$. Keeping this in mind, we now briefly indicate why the above definition indeed yields an orientation for $\ker H$.

Let τ be a facet of $\mathcal{M}$ which meets $H^{-1}(0)$ and does not belong to the boundary $\partial\mathcal{M}$, let $\sigma_1, \sigma_2 \in \mathcal{M}$ be the two cells containing τ, and let $\xi_j := H^{-1}(0) \cap \sigma_j \in \ker H$ for $j = 1, 2$. If $b_1,\dots,b_N$ is a basis of $\mathrm{tng}(\tau)$, and if $a_j \in \mathrm{tng}(\xi_j)$ points from τ into σ_j, then from condition (37.1) it follows that

$$\mathrm{or}_{\sigma_1}(b_1,\dots,b_N,a_1) = -\mathrm{or}_{\sigma_2}(b_1,\dots,b_N,a_2)\,,$$

and hence (37.2) implies that

$$\mathrm{or}_{\xi_1}(a_1) = -\mathrm{or}_{\xi_2}(a_2)\,,$$

which is exactly the right condition in the sense of (37.1) to ensure that the manifold $\ker H$ is oriented.

Let $H : \mathcal{M} \to \mathbb{R}^N$ be a piecewise linear map on an oriented piecewise linear manifold $\mathcal{M}$ of dimension $N+1$. Given a facet τ of a cell $\sigma \in \mathcal{M}$, we can define the *index* of H at τ with respect to σ by setting

$$\mathrm{index}_{\tau,\sigma}(H) := \mathrm{or}_{\tau,\sigma}(b_1,\dots,b_N)\,\mathrm{sign}\det(H'_\sigma b_1,\dots,H'_\sigma b_N)$$

if τ is completely labeled with respect to H, and $\text{index}_{\tau,\sigma}(H) := 0$ otherwise. It is clear that this definition is independent of the special choice of the basis $b_1, \ldots, b_N$ of $\text{tng}(\tau)$. Furthermore, if zero is a regular value of H, τ is completely labeled and $\xi := H^{-1}(0) \cap \sigma \in \ker H$, then (37.2) implies that $\text{index}_{\tau,\sigma}(H) = 1$ if and only if a positively oriented vector in $\text{tng}(\xi)$ points from τ into σ, i.e., the path runs from τ into σ. By using the perturbation technique (if necessary), see Theorem 33.1 and the definitions thereafter, we obtain

$$\text{index}_{\tau_1,\sigma}(H) = -\text{index}_{\tau_2,\sigma}(H)$$

for the case that τ_1 and τ_2 are two different completely labeled facets of a cell σ (see the piecewise linear step in Algorithms 33.1,33.2). Similarly we obtain

$$\text{index}_{\tau,\sigma_1}(H) = -\text{index}_{\tau,\sigma_2}(H)$$

for the case that τ is a completely labeled facet of two different cells σ_1 and σ_2 (see the pivoting step in Algorithms 33.1). A case of special importance is a facet τ in the boundary $\partial\mathcal{M}$, since then we do not have to specify the cell σ which contains τ because σ is unique. If the piecewise linear Algorithm 33.1 starts on the boundary in a completely labeled facet τ_1 and stops again in the boundary in a completely labeled facet τ_k , then the above formulae imply that

$$\text{index}_{\tau_1}(H) = -\text{index}_{\tau_k}(H)$$

holds. Hence, for a compact piecewise linear manifold (i.e., if $|\mathcal{M}|$ is compact) where only a boundary start or termination is possible, we obtain the following celebrated *index formula*

$$\sum_{\tau \in \partial\mathcal{M}} \text{index}_{\tau}(H) = 0\,.$$

38 Lemke's Algorithm

The first and most prominent example of a piecewise linear algorithm was designed by Lemke [1965] and Lemke and Howson [1964] to calculate a solution of the linear complementarity problem. Subsequently, several authors have studied complementarity problems from the standpoint of piecewise linear homotopy methods, see, e.g., Kojima [1974, 1979], Kojima, Nishino, and Sekine [1976], Saigal [1971, 1976], Todd [1976b]. Complementarity problems can also be considered from an interior point algorithm viewpoint, see Section 29.3, hence by following a smooth path, see, e.g., Kojima, Mizuno, and Noma [1990b], Kojima, Mizuno, and Yoshise

[1991d], KOJIMA, MEGIDDO, and NOMA [1991b], KOJIMA, MEGIDDO, and MIZUNO [1990a], MIZUNO [1992].

We present the Lemke algorithm as an example of a piecewise linear algorithm since it played a crucial role in the development of subsequent piecewise linear algorithms. Let us consider the following *linear complementarity problem*: Given an affine map $g : \mathbb{R}^N \to \mathbb{R}^N$, find an $x \in \mathbb{R}^N$ such that

$$x \in \mathbb{R}^N_+; \quad g(x) \in \mathbb{R}^N_+; \quad x^* g(x) = 0. \tag{38.1}$$

Here $\mathbb{R}_+$ denotes the set of non-negative real numbers, and in the sequel we also denote the set of positive real numbers by $\mathbb{R}_{++}$. If $g(0) \in \mathbb{R}^N_+$, then $x = 0$ is a trivial solution to the problem. Hence this trivial case is always excluded and the additional assumption

$$g(0) \notin \mathbb{R}^N_+$$

is made.

Linear complementarity problems arise in quadratic programming, bimatrix games, variational inequalities and economic equilibria problems. Hence numerical methods for their solution have been of considerable interest, see, e.g., COTTLE [1974], COTTLE and DANTZIG [1968], COTTLE, GOLUB, and SACHER [1978], LEMKE [1980]. See also the recent book COTTLE, PANG, and STONE [1992] for further references.

For $x \in \mathbb{R}^N$ we introduce the positive part $x_+ \in \mathbb{R}^N_+$ by setting $e_i^* x_+ := \max\{e_i^* x, 0\}$, $i = 1, \ldots, N$ and the negative part $x_- \in \mathbb{R}^N_+$ by $x_- := (-x)_+$. The following formulae are then obvious: $x = x_+ - x_-$, $(x_+)^*(x_-) = 0$.

It is not difficult to show the following: Define $f : \mathbb{R}^N \to \mathbb{R}^N$ by $f(z) := g(z_+) - z_-$. If x is a solution of the linear complementarity problem, then $z := x - g(x)$ is a zero point of f. Conversely, if z is a zero point of f, then $x := z_+$ solves the linear complementarity problem.

The advantage which f provides is that it is obviously a piecewise linear map if we subdivide $\mathbb{R}^N$ into orthants. This is the basis for our description of Lemke's algorithm. For a fixed $d \in \mathbb{R}^N_{++}$ we define the homotopy $H : \mathbb{R}^N \times [0, \infty) \to \mathbb{R}^N$ by

$$H(x, \lambda) := f(x) + \lambda d\,. \tag{38.2}$$

For a given subset $I \subset \{1, 2, \ldots, N\}$ an orthant can be written in the form

$$\sigma_I := \{\, (x, \lambda) : \lambda \geq 0,\ e_i^* x \geq 0 \text{ for } i \in I,\ e_i^* x \leq 0 \text{ for } i \in I' \,\}\,, \tag{38.3}$$

where I' denotes the complement of I. The collection of all such orthants forms a piecewise linear manifold $\mathcal{M}$ (of dimension $N+1$) which subdivides $\mathbb{R}^N \times [0, \infty)$. Furthermore it is clear that $H : \mathcal{M} \to \mathbb{R}^N$ is a piecewise linear

map since $x \mapsto x_+$ switches its linearity character only at the co-ordinate hyperplanes.

Let us assume for simplicity (as usual) that zero is a regular value of H. We note however, that the case of a singular value is treated in the same way by using the perturbation techniques. Lemke's algorithm is started on a ray: if $\lambda > 0$ is sufficiently large, then

$$\big(-g(0)-\lambda d\big)_+ = 0 \quad \text{and} \quad \big(-g(0)-\lambda d\big)_- = g(0)+\lambda d \in \mathbb{R}^N_{++}\,,$$

and consequently

$$H\big(-g(0)-\lambda d\,,\,\lambda\big) = 0\,.$$

Hence, the ray defined by

$$\lambda \in [\lambda_0, \infty) \longmapsto -g(0)-\lambda d \in \sigma_\emptyset \tag{38.4}$$

$$\text{for} \qquad \lambda_0 := \max_{i=1,\dots,N} \frac{-g(0)[i]}{d[i]} \tag{38.5}$$

is used (for decreasing λ-values) to start the path following. Since the piecewise linear manifold $\mathcal{M}$ consists of the orthants of $\mathbb{R}^N \times [0,\infty)$, it is finite, and there are only two possibilities:

1. The algorithm terminates on the boundary $|\partial\mathcal{M}| = \mathbb{R}^N \times \{0\}$ at a point $(z,0)$. Then z is a zero point of f, and hence z_+ solves the linear complementarity problem.

2. The algorithm terminates on a secondary ray. Then it can be shown, see COTTLE [1974], that the linear complementarity problem has no solution, at least if the Jacobian g' belongs to a certain class of matrices.

Let us illustrate the use of index and orientation by showing that the algorithm generates a solution in the sense that it terminates on the boundary under the assumption that all principle minors of the Jacobian g' are positive. Note that the Jacobian g' is a constant matrix since g is affine.

For $\sigma_I \in \mathcal{M}$, see (38.3), we immediately calculate the Jacobian

$$H'_{\sigma_I} = (f'_{\sigma_I}\,,\,d)\,,$$

where

$$f'_{\sigma_I} e_i = \begin{cases} g'e_i & \text{for } i \in I, \\ e_i & \text{for } i \in I'. \end{cases} \tag{38.6}$$

If $\xi \in \ker H$ is a solution path in σ_I, then formula (37.2) yields

$$\mathrm{or}_\xi(v) = \mathrm{sign}\det f'_{\sigma_I}\ \mathrm{or}_{\sigma_I}(e_1,\dots,e_N,v)\,,$$

and since $\mathrm{or}_{\sigma_I}(e_1, \ldots, e_N, v) = \mathrm{sign}\big(v^* e_{N+1}\big)$ by the standard orientation in $\mathbb{R}^{N+1}$, we have that $\det f'_{\sigma_I}$ is positive or negative if and only if the λ-direction is increasing or decreasing, respectively, while ξ is traversed according to its orientation. It is immediately seen from (38.6) that $\det f'_{\sigma_I}$ is obtained as a *principle minor of* g', i.e., by deleting all columns and rows of g' with index $i \in I'$ and taking the determinant of the resulting matrix (where the determinant of the "empty matrix" is assumed to be 1). Since we start in the negative orthant $\sigma_\emptyset$ where the principle minor is 1, we see that the algorithm traverses the primary ray against its orientation, because the λ-values are initially decreased. Hence, the algorithm continues to traverse $\ker H$ against its orientation. For the important case that all principle minors of g' are positive, the algorithm must continue to decrease the λ-values and thus stops at the boundary $|\partial\mathcal{M}| = \mathbb{R}^N \times \{0\}$. Hence, in this case the algorithm finds a solution. Furthermore, it is clear that this solution is unique, since $\ker H$ can contain no other ray than the primary ray.

39 Variable Dimension Algorithms

In recent years, a new class of piecewise linear algorithms has attracted considerable attention. They are called *variable dimension algorithms* since they all start from a single point, a zero dimensional simplex, and successively generate simplices of varying dimension, until a completely labeled simplex is found. Numerical results of KOJIMA and YAMAMOTO [1984] indicate that these algorithms improve the computational efficiency of piecewise linear homotopy methods. The first variable dimension algorithm is due to KUHN [1969]. However, this algorithm had the disadvantage that it could only be started from a vertex of a large triangulated standard simplex S, and therefore piecewise linear homotopy algorithms were preferred. By increasing the sophistication of Kuhn's algorithm considerably, VAN DER LAAN and TALMAN [1979] developed an algorithm which could start from any point inside S. It soon became clear, see TODD [1978a], that this algorithm could be interpreted as a homotopy algorithm. Numerous other variable dimension algorithms were developed. Some of the latest are due to DAI, SEKITANI, and YAMAMOTO [1992], DAI and YAMAMOTO [1989], KAMIYA and TALMAN [1990], TALMAN and YAMAMOTO [1989]. Two unifying approaches have been given, one due to KOJIMA and YAMAMOTO [1982], the other due to FREUND [1984a, 1984b]. A variable dimension algorithm which is easy to comprehend and may serve the reader as a gateway is the octrahedral algorithm of WRIGHT [1981], see also Section 39.2.

We present here a modified version of the approach described by KOJIMA

and YAMAMOTO [1982]. The modification consists of introducing a cone construction for dealing with the homotopy parameter. In a special case, this construction was also used by Kojima and Yamamoto, see their lemma 5.13.

Before we can give a description of these algorithms, we introduce the notion of a primal-dual pair of piecewise linear manifolds due to Kojima and Yamamoto. In fact, we only need a special case. Let $\mathcal{P}$ and $\mathcal{D}$ be two piecewise linear manifolds of dimension N. Let us denote the set of k-dimensional faces of cells in $\mathcal{P}$ by $\mathcal{P}^k$. In particular, $\mathcal{P}^0$ is the set of vertices of $\mathcal{P}$, and $\mathcal{P}^N = \mathcal{P}$.

We call $(\mathcal{P}, \mathcal{D})$ a *primal-dual pair* if there is a bijective map

$$\tau \in \mathcal{P}^k \longmapsto \tau^d \in \mathcal{D}^{N-k}, \quad k = 0, 1, \ldots, N\,,$$

such that

$$\tau_1 \subset \tau_2 \iff \tau_2^d \subset \tau_1^d \text{ holds for all } \tau_1, \tau_2 \in \bigcup_{k=0}^{N} \mathcal{P}^k.$$

We will deal with a homotopy parameter via the following cone construction. Throughout the rest of this chapter, ω denotes a point which is affinely independent from all cells under consideration. The introduction of ω is only formal and may be obtained by, e.g., increasing the dimension of the ambient finite dimensional Euclidean space $\mathbf{E}$ introduced at the beginning of Section 31. If σ is a cell, then

$$\sigma^\omega := \Big\{(1-\lambda)\omega + \lambda x : x \in \sigma,\ \lambda \geq 0\Big\}$$

denotes the cone with vertex ω generated by σ. Clearly, σ^ω is again a cell and $\dim \sigma^\omega = \dim \sigma + 1$. If $H : \sigma \to \mathbb{R}^k$ is an affine map, then the affine extension $H^\omega : \sigma^\omega \to \mathbb{R}^k$ is defined by

$$H^\omega\big((1-\lambda)\omega + \lambda x\big) := \lambda H(x)$$

for $x \in \sigma$ and $\lambda \geq 0$, i.e., H is extended in such a way that $H^\omega(\omega) = 0$.

If $\mathcal{M}$ is a piecewise linear manifold of dimension N, then

$$\mathcal{M}^\omega := \Big\{\sigma^\omega : \sigma \in \mathcal{M}\Big\}$$

is a piecewise linear manifold of dimension $N+1$, and a piecewise linear map $H : \mathcal{M} \to \mathbb{R}^k$ is extended to a piecewise linear map $H^\omega : \mathcal{M}^\omega \to \mathbb{R}^k$.

We will be interested below in rays traversing a cone σ^ω, and we therefore collect some formulae. A ray in σ^ω is given as

$$\left\{(1-\varepsilon)z_1 + \varepsilon z_2 : \varepsilon \geq 0\right\} \subset \sigma^\omega\,,$$
$$\text{where} \qquad z_j = (1-\lambda_j)\omega + \lambda_j x_j\,, \quad j = 1,2$$

for some suitable $\lambda_1, \lambda_2 \geq 0$ and $x_1, x_2 \in \sigma$. A simple calculation using the affine independence of ω from σ yields

$$(1-\varepsilon)z_1 + \varepsilon z_2 = (1-\lambda_\varepsilon)\omega + \lambda_\varepsilon x_\varepsilon\,,$$
$$\text{where} \qquad \lambda_\varepsilon = (1-\varepsilon)\lambda_1 + \varepsilon\lambda_2$$
$$\text{and} \qquad x_\varepsilon = \frac{(1-\varepsilon)\lambda_1 x_1 + \varepsilon\lambda_2 x_2}{\lambda_\varepsilon}\,.$$

Since $\lambda_\varepsilon \geq 0$ for all $\varepsilon \geq 0$, it follows that $\lambda_2 \geq \lambda_1$. This leaves two cases to consider:

$$\lambda_2 > \lambda_1 \geq 0 \quad \Rightarrow \quad \lim_{\varepsilon\to\infty} x_\varepsilon = \frac{\lambda_2 x_2 - \lambda_1 x_1}{\lambda_2 - \lambda_1} \in \sigma\,, \tag{39.1}$$
$$\lambda_2 = \lambda_1 > 0 \quad \Rightarrow \quad x_1 \neq x_2\,,$$
$$x_\varepsilon = (1-\varepsilon)x_1 + \varepsilon x_2 \in \sigma \quad \text{for} \quad \varepsilon \geq 0\,.$$

The second case is only possible if the cell σ is unbounded.

Let $\mathcal{T}$ and $\mathcal{M}$ be piecewise linear manifolds of dimension N. We call $\mathcal{T}$ a *refinement* of $\mathcal{M}$ if for all $\sigma \in \mathcal{M}$ the restricted piecewise linear manifold $\mathcal{T}_\sigma := \{\xi : \xi \in \mathcal{T},\ \xi \subset \sigma\}$ subdivides σ.

We are now in a position to introduce primal-dual manifolds. Let $(\mathcal{P}, \mathcal{D})$ be a primal-dual pair of N-dimensional piecewise linear manifolds, and let $\mathcal{T}$ be a refinement of $\mathcal{P}$. Then

$$\mathcal{T} \otimes \mathcal{D} := \{\xi \times \tau^d : \xi \in \mathcal{T}^k,\ \tau \in \mathcal{M}^k,\ \xi \subset \tau,\ 0 \leq k \leq N\},$$

is an N-dimensional piecewise linear manifold with empty boundary. A proof of this and related results was given by Kojima and Yamamoto. We call $\mathcal{T} \otimes \mathcal{D}$ the *primal-dual manifold* generated by $\mathcal{T}$ and $\mathcal{D}$. In particular, $\mathcal{T} = \mathcal{P}$ is one of the standard choices. An essential part of the proof consists of discussing the possible pivoting steps:

Let $\xi \times \tau^d \in \mathcal{T} \otimes \mathcal{D}$ with $k = \dim \xi = \dim \tau$ as above, and let κ be a facet of $\xi \times \tau^d$. We now describe the pivoting of $\xi \times \tau^d$ across the facet κ, i.e., we have to find a cell $\eta \in \mathcal{T} \otimes \mathcal{D}$ such that $\eta \neq \xi \times \tau^d$ and $\kappa \subset \eta$. There are three possible cases:

Increasing the dimension. Let $\kappa = \xi \times \sigma^d$ such that $\sigma \in \mathcal{M}^{k+1}$ contains τ. Then there is exactly one $\rho \in \mathcal{T}^{k+1}$ such that $\xi \subset \rho$ and $\rho \subset \sigma$. This is a consequence of the fact that $\mathcal{T}$ refines $\mathcal{P}$ and is not difficult to prove. Then $\eta := \rho \times \sigma$ is the desired second cell. In this case the dimension k of the primal cell ξ is increased when performing the pivoting step.

Decreasing the dimension. Let $\kappa = \delta \times \tau^d$ such that $\delta \in \mathcal{T}^{k-1}$ is a facet of ξ. If $\delta \subset \partial\tau$, then there exists exactly one facet $\nu \in \mathcal{M}^{k-1}$ of τ such that $\delta \subset \nu$, and $\eta := \delta \times \nu^d$ is the desired second cell. In this case the dimension k of the primal cell ξ is decreased when performing the pivoting step.

Keeping the dimension. Let $\kappa = \delta \times \tau^d$ such that $\delta \in \mathcal{T}^{k-1}$ is a facet of ξ. If $\delta \not\subset \partial\tau$, then there exists exactly one cell $\xi' \in \mathcal{T}^k$ such that $\xi' \neq \xi$, $\xi' \subset \tau$ and $\delta \subset \xi'$. This is again a consequence of the fact that $\mathcal{T}$ refines $\mathcal{P}$ and is not difficult to prove. Now $\eta := \xi' \times \tau$ is the desired second cell. In this case the dimension k of the primal cell ξ is left invariant when performing the pivoting step.

The main point for practical purposes is that the above three different kinds of pivoting steps must be easy to implement. This is of course mainly a question of choosing a simple primal-dual pair $(\mathcal{P}, \mathcal{D})$ and either $\mathcal{T} = \mathcal{P}$ or some standard refinement $\mathcal{T}$ of $\mathcal{P}$ which can be handled conveniently.

We now slightly modify the construction of primal-dual manifolds to include cones for the refinement $\mathcal{T}$ of the primal manifold:

$$\mathcal{T}^\omega \otimes \mathcal{D} := \{\xi^\omega \times \tau^d : \xi \in \mathcal{T}^k,\ \tau \in \mathcal{M}^k,\ \xi \subset \tau,\ 0 \leq k \leq N\}.$$

If $\dim \xi = k > 0$, then the facets of ξ^ω are simply the ρ^ω where $\rho \in \mathcal{T}^{k-1}$ is a facet of ξ, and it is readily seen that the above pivoting steps apply correspondingly. The only exception is the case $\dim \xi = k = 0$. In this case it follows that $\xi = \tau$, and ξ is a vertex of the piecewise linear manifold $\mathcal{D}$, but ξ^ω is a ray which has one vertex, namely $\{\omega\}$. Hence, we now have a boundary

$$\partial(\mathcal{T}^\omega \otimes \mathcal{D}) = \left\{ \{\omega\} \times \{v\}^d : \{v\} \in \mathcal{P}^0 \right\}.$$

Clearly, such a boundary facet $\{\omega\} \times \{v\}^d$ belongs to the $(N+1)$-dimensional cell $\{v\}^\omega \times \{v\}^d \in \mathcal{T}^\omega \otimes \mathcal{D}$. We will later see that such boundary facets are used for starting a piecewise linear algorithm. This corresponds to starting a homotopy method on the trivial level $\lambda = 0$ at the point v. We will now apply the above concept of primal-dual manifolds in order to describe some piecewise linear algorithms.

In particular, we want to illustrate that the concept of primal-dual manifolds allows for a unifying description of many variable dimension algorithms. The homotopy methods where the homotopy parameter induces a cone construction serve as one example, see Section 39.2. An important feature of primal-dual manifolds is that a complementarity property of the variables (x, y) may be incorporated into the construction of the primal-dual manifolds so that this property needs not to be assumed by extra conditions or constructions. This is a very convenient trick for dealing with complementarity problems or related questions, and will be illustrated here for the case of the linear complementarity problem in the next section. Many more applications have been considered, see the literature cited in ALLGOWER and GEORG [1990].

39.1 Lemke's Algorithm Revisited

We consider again the linear complementarity problem (38.1) and choose a primal-dual pair $(\mathcal{P}, \mathcal{D})$ by defining for $I \subset \{1, 2, \ldots, N\}$ and $I' := \{1, 2, \ldots, N\} \setminus I$ the primal and dual faces

$$\begin{aligned} \alpha_I &:= \{x \in \mathbb{R}^N : e_i^* x \geq 0 \text{ for } i \in I,\ e_i^* x = 0 \text{ for } i \in I'\}, \\ \alpha_I^d &:= \alpha_{I'}. \end{aligned}$$

The primal and dual manifolds consist of just one cell: $\mathcal{P} = \mathcal{D} = \{\mathbb{R}_+^n\}$. We now define a piecewise linear map $H : \mathcal{P} \otimes \mathcal{D} \times [0, \infty) \longrightarrow \mathbb{R}^N$ by $H(x, y, \lambda) := y - g(x) - \lambda d$ where $d \in \mathbb{R}_{++}^N$ is fixed. Note that the variables x and y are placed into complementarity with each other by the construction of $\mathcal{P} \otimes \mathcal{D}$, and hence a more complex definition of H as in (38.2) is not necessary. For sufficiently large $\lambda > 0$ the solutions of $H(x, y, \lambda) = 0$ are given by the primary ray $(x, y, \lambda) = (0\,,\, g(0) + \lambda d\,,\, \lambda)$. Here the piecewise linear algorithm following $H^{-1}(0)$ is started in the negative λ-direction. If the level $\lambda = 0$ is reached, a solution $H(x, y, 0) = 0$ solves the linear complementarity problem since the complementarity $x \in \mathbb{R}_+^N$, $y = g(x) \in \mathbb{R}_+^N$, $x^* y = 0$ holds by the construction of $\mathcal{P} \otimes \mathcal{D}$.

39.2 The Octrahedral Algorithm

As a typical representative of the class of variable dimension algorithms for approximating a zero point of a map we present the *octrahedral algorithm* of WRIGHT [1981], since numerical experiments indicate that it performs favorably, see KOJIMA and YAMAMOTO [1984], and since it can be described in a reasonably simple way. Let us point out that similar discussions also

hold for many other algorithms where the refinement $\mathcal{T}$ of the primal manifold $\mathcal{P}$ is a pseudo manifold which triangulates $\mathbb{R}^N$, and where the dual manifold $\mathcal{D}$ subdivides a compact subset of $\mathbb{R}^N$, see ALLGOWER and GEORG [1990], FREUND [1984a, 1984b], KOJIMA and YAMAMOTO [1982, 1984], VAN DER LAAN and TALMAN [1979].

We denote by $\Sigma := \{+1, 0, -1\}^N \setminus \{0\}$ the set of all nonzero sign vectors. For two vectors $s, p \in \Sigma$ we introduce the relation

$$s \prec p \quad :\Longleftrightarrow \quad \bigvee_{i=1,\ldots,N} \Big(e_i^* s \neq 0 \Rightarrow e_i^* s = e_i^* p \Big),$$

i.e., s and p coincide except that s may have additional zeros. Then we define a primal-dual pair $(\mathcal{P}, \mathcal{D})$ of N-dimensional manifolds by introducing the following duality:

$$\alpha_0 := \{0\}, \quad \alpha_0^d := \{\, y \in \mathbb{R}^N : ||y||_1 \leq 1 \,\},$$

and for $s \in \Sigma$ we consider

$$\begin{aligned} \alpha_s &:= \left\{ \sum_{\substack{p \in \Sigma \\ s \prec p}} \lambda_p \, p \;:\; \lambda_p \geq 0 \right\}, \\ \alpha_s^d &:= \{\, y \in \mathbb{R}^N : ||y||_1 \leq 1,\; s^* y = 1 \,\}, \end{aligned}$$

i.e., α_s is a cone spanned by the unit basis vectors $\{s(i)\, e_i\}_{s(i) \neq 0}$, and $y \in \alpha_s^d$ if and only if

$$\begin{aligned} y(i) = 0 \quad &\Longleftrightarrow \quad s(i) = 0, \\ s(i) y(i) > 0 \quad &\Longleftrightarrow \quad s(i) \neq 0, \\ \sum_i |y(i)| &= 1. \end{aligned}$$

Hence, the primal manifold $\mathcal{P}$ subdivides $\mathbb{R}^N$ into 2^N cones spanned by the unit basis vectors $\pm e_i$ for $i = 1, 2, \ldots, N$, and the dual manifold $\mathcal{D}$ just consists of the unit ball with respect to the $||.||_1$-norm. We easily check that

$$y \in \alpha_s^d, \; s \prec p \quad \Rightarrow \quad y^* p \geq 0$$

and hence

$$(x, y) \in \mathcal{P} \otimes \mathcal{D} \quad \Rightarrow \quad x^* y \geq 0\,. \tag{39.2}$$

We now consider a pseudo manifold $\mathcal{T}$ which is a refinement of $\mathcal{P}$, for example it is easy to see that the Union Jack triangulation of $\mathbb{R}^N$, see TODD [1976a], has this property, see Section 32.

The aim now is to find an approximate zero point of an asymptotically linear map $G : \mathbb{R}^N \to \mathbb{R}^N$. To do this, let $G_{\mathcal{T}}$ be the piecewise linear approximation of G with respect to the pseudo manifold $\mathcal{T}$ triangulating $\mathbb{R}^N$, see the beginning of Section 33. If $\sigma = [v_1, \ldots, v_{N+1}] \in \mathcal{T}$, and if a point $u \in \sigma$ is expanded into its barycentric co-ordinates $u = \sum_{i=1}^{N+1} c_i v_i$, then $G_{\mathcal{T}}(u) = \sum_{i=1}^{N+1} c_i G(v_i)$. It is clear that $G_{\mathcal{T}}$ is also asymptotically linear and $G'_{\mathcal{T}}(\infty) = G'(\infty)$. A homotopy $\tilde{H} : \mathcal{T} \otimes \mathcal{D} \times [0, \infty) \to \mathbb{R}^N$ is introduced by setting

$$\tilde{H}(x, y, \lambda) := G'(\infty)y + \lambda G_{\mathcal{T}}(x)\,. \tag{39.3}$$

Here, for simplicity, $y = 0$ plays the role of a starting point. Unfortunately, $\tilde{H}$ is not piecewise linear. Hence, we use the cone construction to identify $\tilde{H}$ with a piecewise linear map $H : \mathcal{P}^\omega \otimes \mathcal{D} \longrightarrow \mathbb{R}^N$ by collecting the variables in a different way:

$$H(z, y) := G'(\infty)y + G^\omega_{\mathcal{T}}(z)\,. \tag{39.4}$$

For $z = \omega$, which corresponds to $\lambda = 0$, there is exactly one solution of $H(z, y) = 0$, namely $(z, y) = (\omega, 0)$. Hence $H^{-1}(0)$ intersects the boundary $\partial(\mathcal{P}^\omega \otimes \mathcal{D})$ at just a single point. This is the starting point for our piecewise linear algorithm which traces $H^{-1}(0)$.

Let us first show that there is a constant $C > 0$ such that $\tilde{H}(x, y, \lambda) = 0$ implies $||x|| < C$. Indeed, otherwise we could find a sequence

$$\big\{(x_k, y_k, \lambda_k)\big\}_{k=1,2,\ldots} \subset H^{-1}(0)$$

such that $\lim_{k\to\infty} ||x_k|| = \infty$. It follows from $\tilde{H}(x_k, y_k, \lambda_k) = 0$ and (39.3) that

$$\lambda_k^{-1} y_k + G'(\infty)^{-1} G_{\mathcal{T}}(x_k) = 0\,. \tag{39.5}$$

If we multiply this equation from the left with x_k^* and divide by $||x_k||^2$, the asymptotic linearity of $G_{\mathcal{T}}$ yields

$$\lim_{k\to\infty} ||x_k||^{-2} x_k^* G'(\infty)^{-1} G_{\mathcal{T}}(x_k) = 1\,,$$

and the boundedness $||y_k|| \leq 1$ implies that

$$x_k^* G'(\infty)^{-1} G_{\mathcal{T}}(x_k) > 0 \tag{39.6}$$

for all sufficiently large k. Now by (39.5), (39.6) we have that $x_k^* y_k > 0$ for all sufficiently large k, which is a contradiction to (39.2).

Hence the above boundedness implies that the algorithm can only traverse finitely many cells, and since the solution on the boundary $\partial(\mathcal{T}^\omega \otimes \mathcal{D})$ is unique, it can only terminate in a ray

$$\big\{\big((1-\varepsilon)z_1 + \varepsilon z_2\,,\, (1-\varepsilon)y_1 + \varepsilon y_2\big) : \varepsilon \geq 0\big\} \subset \tau^\omega \times \alpha_I^d \in \mathcal{T}^\omega \otimes \mathcal{D}\,,$$

where $\tau \in \mathcal{T}^k$ such that $\tau \subset \alpha_I$ and k is the number of elements of I. We refer to (39.1) and the notation and remarks preceeding it. It follows from

$$H\big((1-\varepsilon)z_1 + \varepsilon z_2\,,\,(1-\varepsilon)y_1 + \varepsilon y_2\big) = 0 \tag{39.7}$$

and (39.3–39.4) that

$$(1-\varepsilon)y_1 + \varepsilon y_2 + \lambda_\varepsilon G'(\infty)^{-1} G_{\mathcal{T}}(x_\varepsilon) = 0 \quad \text{for} \quad \varepsilon \geq 0\,.$$

Since the k-cell τ is bounded, we only have to consider the case $\lambda_2 > \lambda_1 \geq 0$, see (39.1). Dividing equation (39.7) by $\varepsilon > 0$ and letting $\varepsilon \to \infty$ yields

$$G_{\mathcal{T}}(x) = 0\,, \quad \text{where} \quad x := \frac{\lambda_2 x_2 - \lambda_1 x_1}{\lambda_2 - \lambda_1} \in \tau$$

is the desired approximate zero point of G.

40 Approximating Manifolds

The ideas of predictor corrector and piecewise linear curve tracing can be extended to the approximation of implicitly defined manifolds $\tilde{H}^{-1}(0)$ where $\tilde{H} : \mathbb{R}^{N+K} \to \mathbb{R}^N$.

For simplicity, we assume in this section that zero is a regular value of the smooth map $\tilde{H} : \mathbb{R}^{N+K} \to \mathbb{R}^N$. Hence $\tilde{\mathcal{M}} := \tilde{H}^{-1}(0)$ is a smooth K-dimensional manifold. Before we discuss the methods for obtaining piecewise linear approximations of $\tilde{\mathcal{M}}$, let us briefly indicate that the Gauss-Newton method can be used as a corrector in the sense of Section 4 also in this more general setting.

If B is an $N \times (N+K)$-matrix with maximal rank, then in analogy to Section 3, the Moore-Penrose inverse B^+ of B is given by, e.g., $B^+ = B^*(BB^*)^{-1}$. The product BB^+ is the identity on $\mathbb{R}^N$, and $\mathrm{Id} - B^+B$ is the orthogonal projection onto $\ker(B)$.

In analogy to Theorem 4.1, there exists an open neighborhood U of $\tilde{\mathcal{M}}$ such that the Gauss-Newton method

$$v_{i+1} = v_i - \tilde{H}'(v_i)^+ \tilde{H}(v_i),\ i = 0, 1, \ldots \tag{40.1}$$

converges quadratically to a point $v_\infty \in \tilde{\mathcal{M}}$ whenever the starting point v_0 is in U. Since the evaluation and decomposition of the Jacobian matrix $\tilde{H}'(v_i)$ may be costly, one often modifies (40.1) to the so-called *chord method*

$$v_{i+1} = v_i - B^+ \tilde{H}(v_i),\ i = 0, 1, \ldots \tag{40.2}$$

where B is some fixed approximation of $\tilde{H}'(v_0)$. It is well known that the above mentioned quadratic convergence reduces to linear convergence in the latter case.

Orthogonal decompositions are particularly useful in this context. If Q is an orthogonal $(N+K, N+K)$-matrix such that $BQ = (L, 0)$ for some lower triangular (N, N)-matrix L, and if we split the orthogonal matrix $Q = (Q_N, Q_K)$ into the first N and the last K columns, then it is straightforward to see that $B^+ = Q_N L^{-1}$, and the columns of Q_K provide an orthonormal basis for $\ker(B)$.

Unfortunately, for all known decomposition methods, this basis matrix Q_K does not depend continuously on the choice of the matrix B, and this is a fact which complicates matters in constructing the moving frame algorithm, see Section 40.1. The remedy is to introduce a reference $(N+K, K)$-matrix T_K whose columns form an orthonormal system (i.e., $T_K^* T_K = \mathrm{Id}$) and to use the singular value decomposition

$$V_1^* T_K^* Q_K V_2 = \Sigma \tag{40.3}$$

see, e.g., GOLUB and VAN LOAN [1989]. RHEINBOLDT [1987] shows that the map

$$B \mapsto W_K := Q_K V_1 V_2^* \tag{40.4}$$

is smooth if B varies over the open set of $(N+K, K)$-matrices which have full rank and a kernel such that $T_K^* Q_K$ is non-singular. We simplify our discussion by slightly abusing the notation of Rheinboldt and calling the new matrix W_K the *moving frame* of the kernel of B with respect to the *reference matrix* T_K.

There are two basic types of algorithms for approximating manifolds with $K > 1$: one is the moving frame algorithm of RHEINBOLDT [1987, 1988b], see also BRODZIK and RHEINBOLDT [1994], which is a higher dimensional analogue of the predictor corrector method, the other is a piecewise linear algorithm which has been developed in ALLGOWER and GNUTZMANN [1987], ALLGOWER and SCHMIDT [1985], GNUTZMANN [1989], WIDMANN [1990a, 1990b], see also Chapter 15 of ALLGOWER and GEORG [1990].

40.1 The Moving Frame Algorithm

The moving frame algorithm involves predictors arising from a local triangulation of the tangent space at a current point. The corrector consists of using the Gauss-Newton method for projecting the generated mesh back to the manifold. This method is well-suited for smooth manifolds in which the dimension N is large, such as in multiple parameter nonlinear eigenvalue problems, see, e.g., RHEINBOLDT [1988b, 1992a]. It has been applied to the

calculation of fold curves and to differential-algebraic equations, see DAI and RHEINBOLDT [1990], RHEINBOLDT [1986, 1991, 1992b].

To motivate the idea of the moving frame algorithm, we first give a very heuristic description. At some starting point $p \in \tilde{\mathcal{M}}$, we triangulate the tangent space of $\tilde{\mathcal{M}}$ at p using some standard triangulation $\mathcal{T}$. The tangent space is of course isometric to $\mathbb{R}^K$. We now imagine that the manifold is "rolled" over the tangent space, thereby "imprinting" a triangulation on $\tilde{\mathcal{M}}$. This is used to provide an approximation of some part of $\tilde{\mathcal{M}}$ by a pseudo manifold. Actually, the "imprinting" is carried out via a chord method (40.2) which must eventually be restarted from time to time with a new tangent space. The moving frame idea keeps the fitting together of the imprinted triangulation in the "rolling" process consistent when restarts are made. The following pseudo algorithm sketches the essential ideas of Rheinboldt's method. Given a triangulation $\mathcal{T}$ of $\mathbb{R}^K$, the algorithm constructs an "imprint" $\varphi : \mathcal{X} \to \mathcal{M}$ where $\mathcal{X} \subset \mathcal{T}^0$ is a subset of "marked" nodes of $\mathcal{T}$, which is successively enlarged.

ALGORITHM 40.1 **(Moving Frame Algorithm)**

INPUT $s \in \mathcal{T}^0$ % *initial vertex*
$\varphi(s) \in \tilde{\mathcal{M}}$ % *starting point on $\tilde{\mathcal{M}}$, imprint of s*
T_K % *reference matrix*
$h > 0$ % *steplength,*
% *should be much bigger than the meshsize of the triangulation*

$\mathcal{X} \leftarrow \{s\}$ % *initial set of marked nodes*
REPEAT
% *begin building a new frame:*
GET $x \in \mathcal{X}$ such that dist $\big(x, \mathcal{T}^0 \setminus \mathcal{X}\big) < h$
% *new Jacobian which will generally be decomposed at this point:*
$B \leftarrow \tilde{H}'(\varphi(x))$
% *moving frame of* $\ker(B)$ *with respect to T_K:*
calculate W_K according to (40.4)
WHILE dist $\big(x, \mathcal{T}^0 \setminus \mathcal{X}\big) < h$ DO
% *new marked node:*
GET $y \in \mathcal{T}^0 \setminus \mathcal{X}$ such that $||x - y|| < h$
$v \leftarrow W_K(y - x) + \varphi(x)$ % *predictor for imprint of y*
REPEAT $v \leftarrow v - B^+\tilde{H}(v)$ % *chord corrector method*
UNTIL convergence
$\varphi(y) \leftarrow v$ % *imprint of y*
$\mathcal{X} \leftarrow \mathcal{X} \cup \{y\}$ % *set of marked nodes is augmented*
END WHILE
UNTIL a stopping criterion is satisfied.

By examining the construction of the moving frame at the end of the previous section it becomes evident that we have to make the following technical restriction for nodes $x \in \mathcal{X}$ where we begin a new frame: let $\tilde{\mathcal{M}}_0$ be the set of points where the reference matrix T_K induces a local co-ordinate system on $\tilde{\mathcal{M}}$, i.e.,

$$\tilde{\mathcal{M}}_0 := \left\{ z \in \tilde{\mathcal{M}} : \det \begin{pmatrix} \tilde{H}'(z) \\ T_K^* \end{pmatrix} \neq 0 \right\}.$$

Then a point x is only permitted if its imprint $\varphi(x)$ is in the connected component of $\tilde{\mathcal{M}}_0$ which contains the starting point $\varphi(s)$. It is possible to relax this restriction, but this is usually done at the cost of having an overlapping approximation of $\tilde{\mathcal{M}}$ by a pseudo manifold.

In typical applications of the above method, the dimension N will be significantly larger than K, and hence the computational cost of the singular value decomposition (40.3) is comparatively small.

The above algorithm can be regarded as a higher dimensional analogue of the predictor corrector continuation methods. The predictor step is more complicated than for $K = 1$, since a triangulation of $\mathbb{R}^K$ is mapped onto the tangent space of $\tilde{\mathcal{M}}$ at x via the moving frame device. For the case $K = 1$, the moving frame idea coincides with the concept of orientation as described, e.g., in (1.5). For $K > 1$ however, the moving frame device induces more structure than just orientation. The corrector process is quite analogous to the case $K = 1$. Some topics which remain to be investigated further are:

Globalization: If $\tilde{\mathcal{M}}$ is a compact manifold without boundary, it would be desirable to adapt the construction of the marked nodes $\mathcal{X}$ and the imprint $\varphi(\mathcal{X})$ in such a way that $\varphi(\mathcal{X})$ can be regarded as a compact pseudo manifold without boundary (by declaring appropriate K-simplices). HOHMANN [1991] has made a start on this for $K = 2$.

Steplength Adaptation: As in the case $K = 1$, it is possible to vary the steplength h in the above algorithm according to the performance of the Gauss-Newton corrector (and possibly other factors.)

Handling Singular Points: It would be desirable to incorporate techniques for detecting, classifying and handling singularities on the manifold (e.g., bifurcation points). This is a much more complex problem than for the case $K = 1$.

40.2 Approximating Manifolds via PL Methods

The difficulty of obtaining a global approximation of an implicitly defined compact manifold without boundary does not exist for piecewise linear algorithms. That is, it is possible to obtain an approximating piecewise linear manifold having no holes or overlappings. However, these algorithms become extremely costly for large N. The piecewise linear algorithms have been applied to the visualization of body surfaces, see ALLGOWER and GNUTZMANN [1991], and to the approximation of surface and body integrals, see ALLGOWER, GEORG, and WIDMANN [1991e, 1993]. They can also be used as automatic mesh generators for boundary element methods, see GEORG [1991]. For software for surface and volume approximation via piecewise linear methods, see the software "pla_s_k" in Chapter VIII.

We begin with a description of the underlying ideas. Let us suppose that the space $\mathbb{R}^{N+K}$ is triangulated by a triangulation $\mathcal{T}$. In our earlier piecewise linear algorithms there was not much reason to store any simplices. In the present situation, however, we will need for certain reasons to store simplices. An important advantage of the usual standard triangulations is that any simplex can be very compactly stored and cheaply recovered by means of an $(N + K)$-tuple of integers corresponding to its barycenter. Let us illustrate this for the example of the Coxeter-Freudenthal-Kuhn triangulation, see Section 32 and the notation therein.

Let $\sigma = [v_1, \ldots, v_{N+K+1}]$ be an $(N+K)$-simplex whose vertices have integer co-ordinates such that $v_{i+1} = v_i + e_{\pi(i)}$, $i = 1, \ldots, N + K$, where π is some permutation of the numbers $\{1, \ldots, N+K\}$. Adding up all vertices defines the vector $m = \sum_{i=1}^{N+K+1} v_i$, which can be uniquely decomposed in the following way: $m = (N + K + 1)z + \lambda$ where z has integer co-ordinates, and the co-ordinates of λ are a permutation of $\{1, \ldots, N+K\}$. It is easy to see that the vertex v_1 and the permutation π can be easily recovered from m via $v_1 = z$ and $\lambda\big(\pi(1)\big) > \lambda\big(\pi(2)\big) > \ldots > \lambda\big(\pi(N+K)\big)$. Note also that $(N + K + 1)^{-1}m$ is the barycenter of σ. Hence, it is possible to compactly store the information characterizing the simplex σ in the integer vector m.

It is also possible to perform the pivoting steps directly on the integer vector m and thereby to save some arithmetic operations. The following rules are immediately recovered by translating the pivoting rules of the Coxeter-Freudenthal-Kuhn triangulation for m:

1. Pivoting the leading vertex v_1 of σ generates a simplex $\tilde{\sigma}$ whose integer vector $\tilde{m}$ is obtained by adding 1 to all components of m and an additional 1 to the component $m_{\pi(1)}$, which otherwise would have a remainder 0 modulo $N + K + 1$.

2. Conversely, if the last vertex v_{N+K+1} of σ is pivoted, a simplex $\tilde{\sigma}$ is

generated whose integer vector $\tilde{m}$ is obtained by subtracting 1 from all components of m and an additional 1 from the component $m_{\pi(N+K)}$, which otherwise would have a remainder 0 modulo $N+K+1$.

3. Pivoting one of the other vertices v_q, $1 < q < N+K+1$, of σ generates a simplex $\tilde{\sigma}$ whose integer vector $\tilde{m}$ is obtained by adding 1 to the component $m_{\pi(q)}$ and subtracting 1 from the component $m_{\pi(q-1)}$.

The Union Jack triangulation of TODD [1976a], see also COXETER [1963], similarly offers compact storing and pivoting in integer arithmetic.

As in Section 33, let H denote the piecewise linear approximation of $\tilde{H}$ with respect to $\mathcal{T}$. The definitions of regular points and regular values extend analogously to this context. We again obtain a perturbation theorem, i.e., the proof of Theorem 33.1 involving ϵ-perturbations, generalizes verbatim if 1 is replaced by K.

If zero is a regular value of H, the zero set $H^{-1}(0)$ carries the structure of a K-dimensional piecewise linear manifold. We formulate this last remark more precisely:

THEOREM 40.1 *Let zero be a regular value of H. If $\sigma \in \mathcal{T}$ has a non-empty intersection with $H^{-1}(0)$, then $\mathcal{M}_\sigma := \sigma \cap H^{-1}(0)$ is a K-dimensional polytope, and the family*

$$\mathcal{M} := \{\mathcal{M}_\sigma : \sigma \in \mathcal{T},\ \sigma \cap H^{-1}(0) \neq \emptyset\}$$

is a K-dimensional piecewise linear manifold.

The following algorithm describes the fundamental steps for obtaining the piecewise linear manifold $\mathcal{M}$ approximating $\tilde{\mathcal{M}}$. We again make the assumptions that $\tilde{H} : \mathbb{R}^{N+K} \to \mathbb{R}^N$ is a smooth map, $\mathcal{T}$ is a triangulation of $\mathbb{R}^{N+K}$, and zero is a regular value of both $\tilde{H}$ and its piecewise linear approximation H. Analogously to the definitions preceding Algorithm 33.1, we call a simplex $\sigma \in \mathcal{T}$ *transverse* if it contains an N-face which is completely labeled with respect to H. In the algorithm below, (see the update step), the dynamically varying set $V(\sigma)$ keeps track of all vertices of the transverse simplex σ which remain to be checked in order to find all possible new transverse simplices by pivoting.

ALGORITHM 40.2 **(PL Approximation of a Manifold)**

INPUT $\sigma \in \mathcal{T}$ transverse % *starting simplex*
$D \subset \mathbb{R}^{N+K}$ compact % *bounds the search*

$\Sigma := \{\sigma\}$ % *list of transverse simplices*

$V(\sigma) :=$ set of vertices of σ
WHILE $V(\sigma) \neq \emptyset$ for some $\sigma \in \Sigma$ DO
    GET $\sigma \in \Sigma$ such that $V(\sigma) \neq \emptyset$
    GET $v \in V(\sigma)$
    pivot $\sigma \to \sigma'$ via $v \to v'$
    IF σ' is not transverse OR $\sigma' \cap D = \emptyset$
        delete v from $V(\sigma)$ % *update step*
    ELSE IF $\sigma' \in \Sigma$ % σ' *is not new*
        delete v from $V(\sigma)$ and v' from $V(\sigma')$ % *update step*
    ELSE % σ' *is added to the list* Σ *in this case*
        $\Sigma := \Sigma \cup \{\sigma'\}$
        $V(\sigma') :=$ set of vertices of σ'
        delete v from $V(\sigma)$ and v' from $V(\sigma')$ % *update step*
    END IF
END WHILE

For purposes of exposition we have formulated the above algorithm in a very general way. One may regard the algorithm as a draft for the "outer loop" of the method. A number of items remain to be discussed. We will show below how a starting simplex can be obtained in the neighborhood of a point $x \in \tilde{\mathcal{M}}$. The list Σ can be used to generate a K-dimensional connected piecewise linear manifold

$$\mathcal{M} := \{\mathcal{M}_\sigma\}_{\sigma \in \Sigma} ,$$

see Theorem 40.1. This piecewise linear manifold approximates $\tilde{\mathcal{M}}$ quadratically in the mesh size of $\mathcal{T}$, as will be seen from the error estimates which will be given in the next section. If $\tilde{\mathcal{M}}$ is compact, the generated piecewise linear manifold will be compact without boundary, provided the mesh of the triangulation is sufficiently small. It is not really necessary to perform the pivot $\sigma \to \sigma'$ if σ' is not transverse, since it will already be known from the current data whether the facet $\sigma \cap \sigma'$ is transverse. In the above comparing process called "σ' is not new", it is crucial that compact exact storing is possible for standard triangulations. The list searching can be performed via efficient binary tree searching. Implementations using such ideas has been given by Gnutzmann [1989] and Widmannn [1990a, 1990b].

The piecewise linear manifold $\mathcal{M}$ furnishes an initial coarse piecewise linear approximation of $\tilde{\mathcal{M}}$. Several improvements are possible. The first is that a Gauss-Newton type method (see (40.1) or (40.2)) can be used to project the nodes of $\mathcal{M}$ onto $\tilde{\mathcal{M}}$. Thus a new piecewise linear manifold $\mathcal{M}_1$ is generated which inherits the adjacency structure of the nodes from $\mathcal{M}$ and has nodes on $\tilde{\mathcal{M}}$.

In many applications (e.g., boundary element methods) it is desirable to uniformize the mesh $\mathcal{M}_1$. A very simple and successful means of doing this is "mesh smoothing". One such possible method consists of replacing each node of the mesh by the average of the nodes with which it shares an edge and by using the resulting point as a starting value for a Gauss-Newton type process to iterate back to $\tilde{\mathcal{M}}$. The edges or nodal adjacencies are maintained as before. Three or four sweeps of this smoothing process over all of the nodes of $\mathcal{M}_1$ generally yields a very uniform piecewise linear approximation of $\tilde{\mathcal{M}}$. Mesh smoothing has been implemented in the programs developed by Gnutzmann and by Widmann.

Another step which is useful for applications such as boundary elements is to locally subdivide the cells of the piecewise linear manifolds $\mathcal{M}$ or $\mathcal{M}_1$ into simplices in such a way that the resulting manifold can be given the structure of a pseudo manifold $\mathcal{M}_2$. This is a technical problem which for $K = 2$ is easy to implement, and this has been done in the above mentioned programs.

Once an approximating pseudo manifold $\mathcal{M}_2$ has been generated, it is easy to refine it by, e.g., the well-known construction of halving all edges of each simplex $\tau \in \mathcal{M}_2$, triangulating it into 2^K subsimplices and projecting the new nodes back onto $\tilde{\mathcal{M}}$.

We have assumed that zero is a regular value of H. In fact, as in the Perturbation Theorem 33.1 and following remarks, $\vec{\varepsilon}$-perturbations and the corresponding general definition "completely labeled" automatically resolves singularities even if zero is not a regular value of H. The situation is similar to the case $K = 1$ which has been discussed by Peitgen [1982], Peitgen and Schmitt [1983], see also Gnutzmann [1989] where the general case is treated.

Let us next address the question of obtaining a transverse starting simplex. If we assume that a point x on $\tilde{\mathcal{M}}$ is given, then it can be shown that any $(N + K)$-simplex with barycenter x and sufficiently small diameter is transverse, see Theorem 40.5.

There is a better way, namely to construct a completely labeled N-face and a transverse simplex containing it. Let us assume that x is a point in $\tilde{\mathcal{M}}$. The normal space of $\tilde{\mathcal{M}}$ at x is given by the orthogonal complement $\ker \tilde{H}'(x)^\perp$ of $\ker \tilde{H}'(x)$. From the Inverse Function Theorem, it is clear that the restriction of $\tilde{H}$ to the affine subspace $x + \ker \tilde{H}'(x)^\perp$ has x as a regular isolated zero point. Hence, if $\tau \subset x + \ker \tilde{H}'(x)^\perp$ is an N-simplex with barycenter x and sufficiently small diameter, then it is completely labeled. Error estimates implying this will be given in the next section. Hence, we only have to construct an affine map Φ sending an N-face of some standard triangulation $\tilde{\mathcal{T}}$ of $\mathbb{R}^{N+K}$ onto τ, and then using the triangulation $\mathcal{T} = \Phi(\tilde{\mathcal{T}})$, see Allgower and Georg [1990, p. 242] for more details.

Algorithm 40.2 merely generates a list Σ of transverse simplices. For particular purposes such as boundary element methods, computer graphics, etc., a user will wish to have more information concerning the structure of the piecewise linear manifold $\mathcal{M}$, e.g., all nodes of the piecewise linear manifold $\mathcal{M}$ together with their adjacency structure. Hence, to meet such requirements, it is necessary to "customize" the above algorithm by, e.g., incorporating inner loops which serve to yield such information. This is illustrated in ALLGOWER and GEORG [1990, (15.4.7–8)].

40.3 Approximation Estimates

We conclude this section with some error estimates concerning the quality of the above piecewise linear approximations. Although some of the piecewise linear algorithms are useful under much weaker assumptions on the map $\tilde{H}$, in order to obtain error estimates, it is necessary to make some smoothness assumptions regarding the first and second derivatives of $\tilde{H}$. The results in this section are analogous to results given in GNUTZMANN [1989] and ALLGOWER and GEORG [1989, 1990]. For reasons of simplicity, in this section we make the following assumptions:

$$\begin{aligned} &\tilde{H} : \mathbb{R}^{N+K} \to \mathbb{R}^N \text{ is a smooth map with zero a regular value,} \\ &||\tilde{H}'(x)^+|| \leq \kappa \text{ for all } x \in \tilde{\mathcal{M}} := \tilde{H}^{-1}(0), \\ &||\tilde{H}''(x)|| \leq \alpha \text{ for all } x \in \mathbb{R}^{N+K}. \end{aligned} \tag{40.5}$$

Actually, the bounds in (40.5) need only to hold in a convex region containing all of the points considered in the following discussion. We remark also that it would be sufficient to assume that the Jacobian $\tilde{H}'(x)$ is Lipschitz continuous with constant α. The above assumptions only serve to make our proofs less technical, however the results are essentially the same.

Let $\mathcal{T}$ be a triangulation of $\mathbb{R}^{N+K}$ having mesh size $\delta > 0$. As in the preceding section we let H denote the piecewise linear approximation of $\tilde{H}$ with respect to $\mathcal{T}$. Our first result concerns the accuracy with which H approximates $\tilde{H}$.

THEOREM 40.2 *Under the assumptions (40.5), there holds:*

$$||\tilde{H}(x) - H(x)|| \leq \frac{1}{2}\alpha\delta^2 \quad \textit{for} \quad x \in \mathbb{R}^{N+K}.$$

PROOF Let $\sigma = [v_1, v_2, \ldots, v_{N+K+1}] \in \mathcal{T}$ be an $(N+K)$-simplex such that

$$x = \sum_{i=1}^{N+K+1} \gamma_i v_i \in \sigma.$$

From Taylor's formula we have

$$\tilde{H}(v_i) = \tilde{H}(x) + \tilde{H}'(x)(v_i - x) + \frac{1}{2}A_i[v_i - x, v_i - x]$$

for $i = 1, 2, \ldots, N + K + 1$, where we use the mean values

$$A_i := \int_0^1 \tilde{H}''\big(x + t(v_i - x)\big)2(1 - t)dt$$

of $\tilde{H}''$. Multiplying these equations with the corresponding barycentric co-ordinates γ_i, summing and taking norms yields

$$||\tilde{H}(x) - \sum_{i=1}^{N+K+1} \gamma_i \tilde{H}(v_i)|| \leq \frac{1}{2}\alpha\delta^2$$

as a consequence of (40.5). The result now follows since

$$H(x) = \sum_{i=1}^{N+K+1} \gamma_i \tilde{H}(v_i).$$

□

In the next estimate the thickness of a simplex has a meaningful role. One possible measure of thickness is the following definition:

Let σ be a simplex with diameter δ and barycenter x. Let ρ be the radius of the largest ball having center x and being contained in σ. Then the *measure of thickness* of σ is defined by

$$\theta(\sigma) := \frac{\rho}{\delta}. \tag{40.6}$$

The *measure of thickness of a triangulation* $\mathcal{T}$ is defined by

$$\theta(\mathcal{T}) := \inf\{\theta(\sigma) : \sigma \in \mathcal{T}\}.$$

For standard triangulations, such measures are well-known and > 0, see, e.g., KOJIMA [1978] or SAIGAL [1979]. For example, the Coxeter-Freudenthal-Kuhn triangulation of $\mathbb{R}^q$ has thickness $\theta = 1/\big((q+1)\sqrt{2}\big)$.

The next estimate gives a measure of how well H'_σ approximates the derivative $\tilde{H}'$.

THEOREM 40.3 *Let $\sigma \subset \mathbb{R}^{N+K}$ be an $(N+K)$-simplex having diameter δ and thickness θ. If $x \in \sigma$, then $||\tilde{H}'(x) - H'_\sigma(x)|| \leq \delta\alpha/\theta$.*

PROOF Let $\sigma = [v_1, v_2, \ldots, v_{N+K+1}]$. From Taylor's formula we have

$$\begin{aligned}\tilde{H}'(x)(v_i - v_j) &= \tilde{H}'(x)(v_i - x) - \tilde{H}'(x)(v_j - x) \\ &= \tilde{H}(v_i) - \tilde{H}(v_j) \\ &\quad -\frac{1}{2}A_i[v_i - x, v_i - x] + \frac{1}{2}A_j[v_j - x, v_j - x]\end{aligned}$$

for $i, j = 1, 2, \ldots, N + K + 1$, where the mean values A_i of $\tilde{H}''$ are defined as in the previous proof. From the definition of the piecewise linear approximation we immediately obtain

$$H'_\sigma(x)(v_i - v_j) = \tilde{H}(v_i) - \tilde{H}(v_j).$$

Subtracting corresponding sides of the above equations and taking norms and using Theorem 40.2 yields

$$||\big(\tilde{H}'(x) - H'_\sigma(x)\big)(v_i - v_j)|| \leq \alpha\delta^2.$$

By making convex combinations with this last estimate, we obtain

$$||\big(\tilde{H}'(x) - H'_\sigma(x)\big)(u - v)|| \leq \alpha\delta^2$$

for all $u, v \in \sigma$. From the definition of thickness (40.6) it follows that the set $\{u - v : u, v \in \sigma\}$ contains the ball with radius $\theta\delta$ and center zero. Thus the above estimate extends to the corresponding matrix norms

$$\theta\delta||\big(\tilde{H}'(x) - H'_\sigma(x)\big)|| \leq \alpha\delta^2,$$

and the assertion follows. □

The next theorem is a useful characterization of transverse simplices. We employ the perturbation notation of Theorem 33.1.

THEOREM 40.4 *A simplex $\sigma \in \mathcal{T}$ is transverse if and only if it contains solutions v_ε of $H(v) = \vec{\varepsilon}$ for sufficiently small $\varepsilon > 0$.*

PROOF The proof is obtained by modifying the arguments in the Perturbation Theorem 33.1 and the piecewise linear step (see, e.g., Algorithm 33.1). If σ does not contain the asserted solutions v_ε for sufficiently small $\varepsilon > 0$, then by the definition of a completely labeled facet and a transverse cell, it cannot be transverse. On the other hand, if σ contains solutions v_ε for sufficiently small $\varepsilon > 0$, then by an obvious generalization of Theorem 33.1, the solution set consists of regular points of H for sufficiently small $\varepsilon > 0$. Hence, if ε varies, no faces of σ of dimension $< N$ can be intersected, and hence always the same N-faces of σ have to be intersected by this solution set. Clearly, those are the completely labeled N-faces of σ. □

The following theorem guarantees that all regular zero points of $\tilde{H}$ can be approximated by transverse simplices. In particular, such estimates as these may be used for obtaining the starting simplices for the piecewise linear algorithms discussed in this chapter.

THEOREM 40.5 *Let $\sigma \subset \mathbb{R}^{N+K}$ be an $(N+K)$-simplex having vertices $v_i, i = 1 \cdots N+K+1$, diameter δ, thickness θ and barycenter x such that $\tilde{H}(x) = 0$. If*

$$\frac{\kappa\alpha\delta}{\theta} < \frac{1}{2},$$

then σ is transverse.

PROOF In view of the previous theorem, it suffices to show that the affine approximation H_σ has a solution point $x_\varepsilon \in \sigma$ such that

$$H_\sigma(x_\varepsilon) = \vec{\varepsilon} \tag{40.7}$$

for sufficiently small $\varepsilon > 0$. Since H_σ is affine, any point given by a generalized Newton step

$$x_\varepsilon := x - B\big(H_\sigma(x) - \vec{\varepsilon}\big)$$

satisfies the equation (40.7), provided that B is a right inverse of H'_σ. If we show that the essential part of the Newton term satisfies the estimate

$$||BH_\sigma(x)|| < \theta\delta \tag{40.8}$$

for a particular B, then we conclude from the definition of thickness of a simplex that $x_\varepsilon \in \sigma$ for sufficiently small $\varepsilon > 0$, and the assertion follows. From Theorem 40.3 we have

$$||\tilde{H}'(x) - H'_\sigma(x)|| \leq \frac{\delta\alpha}{\theta}$$

and hence by the bounds (40.5) and the hypothesis,

$$||\tilde{H}'(x)^+\big(\tilde{H}'(x) - H'_\sigma(x)\big)|| \leq \frac{\kappa\delta\alpha}{\theta} < \frac{1}{2}.$$

We can now define B via the Neumann series

$$B := \sum_{i=0}^{\infty} \Big(\tilde{H}'(x)^+\big(\tilde{H}'(x) - H'_\sigma(x)\big)\Big)^i \tilde{H}'(x)^+.$$

Multiplying the identity

$$H'_\sigma(x) = \tilde{H}'(x)\Big(\mathrm{Id} - \tilde{H}'(x)^+\big(\tilde{H}'(x) - H'_\sigma(x)\big)\Big)$$

from the right by B verifies that B is indeed a right inverse of H'_σ. From the Neumann series we can also see that the estimate

$$||B|| \leq \frac{\kappa}{1 - \frac{\kappa\alpha\delta}{\theta}} < 2\kappa$$

holds. On the other hand, Theorem 40.2 implies

$$||H_\sigma(x)|| = ||H_\sigma(x) - \tilde{H}(x)|| \leq \frac{1}{2}\alpha\delta^2.$$

Combining the last two estimates yields the estimate (40.8) and hence the assertion follows. □

The next theorem shows that the piecewise linear manifold $\mathcal{M} = H^{-1}(0)$ approximates the given manifold $\tilde{\mathcal{M}} = \tilde{H}^{-1}(0)$ quadratically in the mesh-size.

THEOREM 40.6 *Let $x \in \mathbb{R}^{N+K}$ be such that $\operatorname{dist}(x, \tilde{\mathcal{M}}) < (\kappa\alpha)^{-1}$. Let $w \in \tilde{\mathcal{M}}$ be a nearest point to x, i.e., $||x - w|| = \operatorname{dist}(x, \tilde{\mathcal{M}})$. If $H(x) = 0$, then $||x - w|| \leq \kappa\alpha\delta^2$.*

PROOF Since w satisfies the optimization problem

$$\min_w \{\, ||x - w|| : \tilde{H}(w) = 0 \,\},$$

the Lagrange equations yield

$$x - w \in \operatorname{range}\big(\tilde{H}'(w)^*\big) \quad \text{or equivalently,} \quad (x - w) \perp \ker\big(\tilde{H}'(w)\big).$$

From Taylor's formula we have

$$\tilde{H}(x) - \tilde{H}(w) = \tilde{H}'(w)(x - w) + \frac{1}{2}A[x - w, x - w],$$

where

$$A = \int_0^1 \tilde{H}''\big(w + t(x - w)\big)2(1 - t)\, dt$$

again denotes a mean value of $\tilde{H}''$. Since $(x-w) \perp \ker\big(\tilde{H}'(w)\big)$, and since the Moore-Penrose inverse performs the inversion orthogonally to $\ker\big(\tilde{H}'(w)\big)$, we have

$$\tilde{H}'(w)^+\tilde{H}(x) = x - w + \frac{1}{2}\tilde{H}'(w)^+A[x - w, x - w].$$

From Theorem 40.2 we have

$$||\tilde{H}(x)|| = ||\tilde{H}(x) - H(x)|| \leq \frac{1}{2}\alpha\delta^2.$$

From these last two statements and the assumptions (40.5) we obtain

$$\begin{aligned} ||x - w|| &\leq \frac{1}{2}\kappa\alpha\delta^2 + \frac{1}{2}\kappa\alpha||x - w||^2 \\ &\leq \frac{1}{2}\kappa\alpha\delta^2 + \frac{1}{2}||x - w||, \end{aligned}$$

and the assertion follows. □

Up to now our approximation estimates have been of a local nature. In order to obtain global approximation results we need to apply more sophisticated tools and technical arguments. One such tool is the Brouwer degree, which for $K = 1$ may be used in a manner similar to that of Rabinowitz [1971] to obtain the existence of global continua. Peitgen and Prüfer [1979] and also Peitgen [1982] have given extensive discussions of the constructive role the piecewise linear methods play in connection with such arguments.

For our purpose the continuous Newton method seems to be a suitable tool. We consider the autonomous differential equation

$$\dot{x} = -\tilde{H}'(x)^+\tilde{H}(x). \tag{40.9}$$

If an initial point x_0 for (40.9) is sufficiently close to $\tilde{\mathcal{M}} = \tilde{H}^{-1}(0)$, then the flow initiating at x_0 has an exponentially asymptotic limit $x_\infty \in \tilde{\mathcal{M}}$, and the map $x_0 \mapsto x_\infty$ is smooth, see, e.g., Tanabe [1979]. Analogously, if zero is a regular value of H and the meshsize of $\mathcal{T}$ is sufficiently small, then we may consider the flow defined by

$$\dot{x} = -H'(x)^+H(x). \tag{40.10}$$

Note that the right hand of (40.10) is piecewise affine, but not continuous, and that a solution path consists of a polygonal path having nodes on lower dimensional faces $\tau \in \mathcal{T}^{N+K-1}$.

To see this, we concentrate on one simplex σ where the map is affine and consider the auxiliary flow

$$\dot{x} = -A^+(Ax - b)$$

for some $(N, N+K)$-matrix A with maximal rank. Integrating this equation leads to

$$x = e^{-A^+At}(x_0 - A^+b) + A^+b$$

where x_0 is an initial point on one facet of σ. Of course, the solution is only valid until the simplex σ is exited. We see that the above solution describes a straight line, though the parametrization is not linear.

It is possible by use of some technical arguments to show that the piecewise linear case (40.10) has results analogous to (40.9), i.e., if an initial point x_0 for (40.10) is sufficiently close to $\mathcal{M} = H^{-1}(0)$, then the flow initiating at x_0 has an exponentially asymptotic limit $x_\infty \in \tilde{\mathcal{M}}$, and the map $x_0 \mapsto x_\infty$ is absolutely continuous. The detailed arguments are omitted. We shall only sketch how this technique may be used to obtain the following two theorems.

THEOREM 40.7 *If $x_0 \in \tilde{\mathcal{M}}$ and the meshsize divided by the measure of thickness δ/θ of $\mathcal{T}$ is sufficiently small, then there exists a transverse $\sigma \in \mathcal{T}$ such that* $\operatorname{dist}(x_0, \sigma) \leq \kappa\alpha\delta^2$.

PROOF Consider the initial value problem (40.10) with initial value x_0 and asymptotic limit $x_\infty \in \mathcal{M}$. A full initial Newton step is given by

$$x_1 = x_0 - H'(x_0)^+ H(x_0).$$

From Theorem 40.2 we obtain the estimate $||H(x_0)|| \leq \frac{1}{2}\alpha\delta^2$. From Theorem 40.3 and (40.5) we obtain $||H'(x_0)^+|| \approx ||\tilde{H}'(x_0)^+|| \leq \kappa$. Thus a rough bound for the stepsize of the full initial Newton step is given by $\frac{1}{2}\kappa\alpha\delta^2$. Hence to obtain the assertion we estimate $||x_0 - x_\infty||$ by twice this steplength. □

The algorithms in this section generate connected components of the piecewise linear manifold $\mathcal{M}$. The following theorem assures that such a connected component approximates the entire manifold $\tilde{\mathcal{M}}$ if it is compact and connected.

THEOREM 40.8 *Let zero also be a regular value of H. Let $C \subset \tilde{\mathcal{M}}$ be a compact connected subset (which could be all of $\tilde{\mathcal{M}}$). Then for any triangulation $\mathcal{T}$ for which the meshsize divided by the measure of thickness δ/θ is sufficiently small, there is a connected compact piecewise linear submanifold $C_\mathcal{T} \subset \mathcal{M}$ such that for every $x_0 \in C$ there is an $x_\infty \in C_\mathcal{T}$ for which $||x_0 - x_\infty|| < \kappa\alpha\delta^2$ holds.*

PROOF Consider the Newton map $x_0 \in C \mapsto x_\infty \in \mathcal{M}$ introduced above. Since this map is continuous, and since the continuous image of a compact and connected set is compact and connected, the piecewise linear submanifold

$$C_\mathcal{T} := \{\mathcal{M}_\sigma : \sigma \in \mathcal{T} \text{ and } x_\infty \in \sigma \text{ for some } x_0 \in C\}$$

is compact and connected. Now the assertion follows from estimates in Theorem 40.7. □

It is now clear from the preceding discussion that if $\tilde{\mathcal{M}}$ is compact and connected, then a connected component of $\mathcal{M}$ approximates $\tilde{\mathcal{M}}$ globally and quadratically for sufficiently small meshsize, provided the measure of thickness of $\mathcal{T}$ stays bounded away from zero.

It is also possible to formulate measures of efficiency for piecewise linear approximations of K-manifolds. Analogously to corresponding results for $K = 1$ as cited in the section on piecewise linear homotopy methods, ALEXANDER [1987] has studied the average intersection density for several triangulations in the context of piecewise linear approximations of K-manifolds.

If zero is a regular value of $\tilde{H}$ and H, then the smooth manifold $\tilde{\mathcal{M}}$ and the approximating manifold $\mathcal{M}$ inherit a natural orientation which in the former case is a basic concept of differential geometry and in the latter case is analogous to the orientation described in Section 37. It can be shown that these orientations are consistent with each other for sufficiently fine mesh size, see GNUTZMANN [1989].

41 Numerical Integration over Surfaces

Recently, GEORG [1991] introduced a new approach to the numerical quadrature of surface integrals for compact surfaces $\mathcal{B}$ which are only implicitly defined, for example, by $\mathcal{B} = \{x \in \mathbb{R}^3 : H(x) = 0\}$ where $H : \mathbb{R}^3 \to \mathbb{R}$. It is assumed that the surface $\mathcal{B}$ is globally approximated via a pseudo manifold $\mathcal{T}$ (piecewise linear approximation) such as was described in the preceding section. Such piecewise linear approximations are typically used in the panel method for solving boundary integral equations related to three-dimensional partial differential equations, see, e.g., BALLMAN, EPPLER, and HACKBUSCH [1988] or HACKBUSCH [1989]. Since (in the context of the last section) the approximation is given automatically, there is typically no explicit parametrization of $\mathcal{B}$ available.

Hence, Georg assumed that a parametrization of the surface $\mathcal{B}$ is only indirectly given via a piecewise smooth isomorphism

$$m : \mathcal{T} \to \mathcal{B},$$

e.g., via an iterative method such as Newton's method.

For example, the boundary element package of ATKINSON [1993] asks the user to write a subroutine for defining a parametrization m which seems quite different from the above assumptions. However, it can be shown that both concepts are equivalent. Atkinson handles the difficulty that the parametrization is not given explicitly by using cubic interpolation. The purpose of Georg's approach was to avoid the handling of the partial

derivatives of m at all, even via finite differences or interpolation. This leads to a quadrature method which is potentially as accurate as wanted or needed.

Let σ denote a triangle of $\mathcal{T}$. Hence $m(\sigma)$ denotes a smooth piece of the surface $\mathcal{B}$ and $m : \sigma \to m(\sigma)$ denotes a smooth parametrization of this piece. For the purposes of our discussion we can assume without loss of generality that σ is the standard triangle $\sigma = \{(s,t) : 0 \le s, t, s+t \le 1\}$. We consider the task of numerically approximating

$$\int_{m(\sigma)} f(x)\,\mu(dx). \tag{41.1}$$

Here μ indicates the usual measure on $\mathcal{B}$ (i.e., the so-called surface element) and f is a given integrand.

The standard approach to the numerical quadrature of (41.1) is to consider the equivalent integral in planar coordinates

$$\int_{m(\sigma)} f(s)\,\mu(dx) = \int_{\sigma} f(m(s,t))\,\|m_s \times m_t\|\,ds\,dt. \tag{41.2}$$

Here m_s, m_t denote the partial derivatives of m with respect to the parameters s, t. Many quadrature rules are known for the right-hand side of (41.2), see, e.g., Davis and Rabinowitz [1984] or Stroud [1971]. A simple approach to numerically approximate (41.2) is to subdivide σ into small triangles σ_i (this corresponds to a subdivision of the surface $m(\sigma)$ into the pieces $m(\sigma_i)$) and then to approximate the integral via one of the composite rules:

$$\begin{aligned}
\int_{m(\sigma)} f(x)\,\mu(dx) &= \sum_i \int_{\sigma_i} f(m(s,t))\,\gamma(s,t)\,ds\,dt \\
&\approx \frac{1}{3}\sum_i \sum_{j=0}^{2} f(m(v_{i,j}))\,\gamma(v_{i,j})\,\mathcal{A}(\mathcal{V}_i), \qquad (41.3)\\
\int_{m(\sigma)} f(x)\,\mu(dx) &\approx \sum_i f(m(b_i))\,\gamma(b_i)\,\mathcal{A}(\mathcal{V}_i). \qquad (41.4)
\end{aligned}$$

Here, the vertices and the barycenter of σ_i are denoted by $\mathcal{V}_i := \{v_{i,j} : j = 0,1,2\}$ and b_i, respectively, and $\gamma := \|m_s \times m_t\|$. $\mathcal{A}(\mathcal{V}_i)$ denotes the area of a triangle with vertices $\mathcal{V}_i$. Methods (41.3) and (41.4) are the respective extensions of the composite trapezoidal or midpoint rule for integrating over a triangle.

Under adequate smoothness assumptions on the integrand f, the local error, i.e., the error of each summand, is $\mathcal{O}(h^4)$ where $h = \max_i \operatorname{diam} \sigma_i$. Under the summation of the composite rule, these local errors lead to a global error $\mathcal{O}(h^2)$. It is also known (in the case of an equidistant subdivision) that the global error can be expanded in terms of h^2, see Lyness [1978]. This is important for applying extrapolation methods to increase the accuracy (e.g., Romberg's scheme).

An alternative approach to approximate (41.1), which does not use the partial derivatives of m, was proposed by Georg [1991]:

$$\int_{m(\sigma)} f(x)\,\mu(dx) \approx \frac{1}{3}\sum_i \sum_{j=0}^{2} f(m(v_{i,j}))\,\mathcal{A}(m(\mathcal{V}_i)), \tag{41.5}$$

$$\int_{m(\sigma)} f(x)\,\mu(dx) \approx \sum_i f(m(b_i))\,\mathcal{A}(m(\mathcal{V}_i)). \tag{41.6}$$

We refer to the quadrature formulae (41.5), (41.6) as the *modified trapezoidal rule* and the *modified mid-point rule* for surface integrals, respectively.

The analogues of these quadrature rules for line integrals were discussed in Section 30. As in the case of the line integrals, the modified quadrature versions offer the significant advantage of avoidance of the need to evaluate derivatives.

Until very recently, it was not known whether the modified rules would also permit an asymptotic expansion of the global error in terms of h^2 (in the case of an equidistant subdivision). Georg [1991] conjectured this expansion, Georg and Tausch [1993] established that at least the leading error term is $\mathcal{O}(h^2)$. Very recently, Verlinden and Cools [1993], and Lyness [1993] have independently proven the full validity of the asymptotic expansion.

This justifies the incorporation of extrapolation steps in an adaptive numerical integration method for surface integrals which was suggested by Georg and Widmann [1993]. The authors assumed an integrand with a weak singularity in an unknown position and developed a strategy which mixes cautious extrapolation steps (when indicators signal smoothnes) with an adaptive strategy (when indicators signal the vicinity of a singularity).

We present the third example taken from their paper. The numerical integration method was programmed in C and was run on a PC with the 80386/387 processors. We integrate over the surface of a ring cyclide, i.e., a torus with varying radius. The surface $\mathcal{B}$ is defined by the equation

$$(x_1^2 + x_2^2 + x_3^2 + R^2 - b^2 - k^2)^2 - 4(Rx_1 + kb)^2 - 4(R^2 - b^2)x_2^2 = 0.$$

Here, R defines the "big" radius, and the small radius r satisfies $k - b \le r \le$

$k+b$. We choose the parameters $R = 1, k = 0.3$ and $b = 0.15$. The approximating pseudo manifold $\mathcal{T}$ consists of 4486 triangles and was generated with our software pla_s_k written by WIDMANN [1990a], see Chapter VIII, using a mesh size of $\delta = 0.15$.

We use the integrand

$$f(x) = \frac{\nu(x) \cdot (x - e)}{||x - e||^3}$$

with the weak singularity $e := (1 + k + b)e_1$, where $\nu(x)$ denotes the outer normal of length 1 at $x \in \mathcal{B}$. The resulting integral is typical for boundary element methods. In fact, this particular integral has an interpretation as a solid angle and hence is known explicitly, i.e.,

$$\int_{\mathcal{B}} f(x)\, \mu(dx) = 2\pi.$$

The table below lists the performance of the numerical integration with increasing precision. "Tol" is a tolerance parameter used to monitor the adaptive and extrapolation steps. "Rel. Error" gives the relative error of the numerical integral, "Time" gives the run time of the integration in seconds.

Tol	Rel. Error	Time
1e-2	2.7e-03	117.10
1e-3	8.5e-05	117.70
1e-4	2.4e-05	120.89
1e-5	6.6e-06	129.30
1e-6	4.2e-06	151.82
1e-7	2.5e-06	208.83
1e-8	4.3e-08	357.46
1e-9	3.2e-08	722.33
1e-10	1.4e-08	1363.09
1e-11	1.4e-08	3801.83
1e-12	3.0e-09	10872.55

Chapter VII

Complexity

42 Smale's Approach

In modern complexity investigations of continuation-type methods the so-called α-theory of SMALE [1986] is a convenient tool. This theory is closely related to the classical Kantorovich estimates for Newton iterations, see, e.g., ORTEGA and RHEINBOLDT[1970] and DEUFLHARD and HEINDL [1979]. In contrast to the Kantorovich estimates, Smale's estimates are based on information at only one point, involving however all derivatives. The maps under consideration have to be analytic.

On the other hand, an analytic map is characterized by all its derivatives at one point. In fact, RHEINBOLDT [1988a] showed that Smale's estimates can be derived from the Kantorovich estimates. However, for complexity considerations, it is more convenient to have all the relevant information situated at only one point. Let us briefly present Smale's estimates and show how they are used for complexity discussions. Our presentation is based on the introductory parts of the papers of SHUB and SMALE [1991] and RENEGAR and SHUB [1992].

Let E, F be complex Banach spaces and $f : E \to F$ an analytic map. It would be possible to assume that f is given only on some open domain, but for reasons of simplicity of exposition we assume f to be defined on all of E. Then for each point $x \in E$ such that $Df(x) : E \to F$ is an isomorphism the following quantities are defined:

$$\beta(f,x) = ||Df(x)^{-1}f(x)|| \tag{42.1}$$

$$\gamma(f,x) = \sup_{k>1} \frac{1}{k!}||Df(x)^{-1}D^k f(x)||^{\frac{1}{k-1}} \tag{42.2}$$

$$\alpha(f,x) = \beta(f,x)\gamma(f,x) \tag{42.3}$$

$$\mathcal{N}_f(x) = x - Df(x)^{-1}f(x). \tag{42.4}$$

Note that $\mathcal{N}_f(x)$ is the Newton iterate of x. It it also convenient to introduce the notation

$$\mathcal{N}_f^\infty(x) = \lim_{i\to\infty} \mathcal{N}_f^i(x) \tag{42.5}$$

provided Newton's method (started at x) is convergent.

A related one-dimensional "control" Newton method is occasionally generated from the following family of functions

$$h_{\beta,\gamma}(t) = \beta - t + \frac{\gamma t^2}{1-\gamma t}. \tag{42.6}$$

For $0 < \alpha < 3 - 2\sqrt{2} \approx .1716$, the function $h_{\beta,\gamma}$ has two real positive roots, the smaller one being

$$\frac{\tau(\alpha)}{\gamma} = \frac{(\alpha+1) - \sqrt{(\alpha+1)^2 - 8\alpha}}{4\gamma}. \tag{42.7}$$

Moreover, $h''_{\beta,\gamma} > 0$ on the interval $\left(0, \frac{1}{\gamma}\right)$. Thus, Newton's method starting at zero generates a strictly increasing sequence $t_i(\beta,\gamma) = \mathcal{N}^i_{h_{\beta,\gamma}}(0)$ converging to this root.

Occasionally, a slightly smaller upper bound for α is used, namely $\alpha_0 = \frac{1}{4}(13 - 3\sqrt{17}) \approx .1577$.

The following is a modification of Smale's α-theorem.

THEOREM 42.1 *Let $x_0 \in E$, $\alpha = \alpha(f,x_0)$, $\gamma = \gamma(f,x_0)$. If $\alpha \le \alpha_0 \approx .1577$, then the iterates $x_{i+1} = \mathcal{N}_f(x_i)$ are defined and converge to a zero point $x_\infty = \mathcal{N}_f^\infty(x_0) \in E$ with the rate*

$$||x_{i+1} - x_i|| \le \left(\frac{1}{2}\right)^{2^i - 1} ||x_1 - x_0||.$$

Moreover, the following estimates hold:

$$||x_\infty - x_0|| \le \frac{\tau(\alpha)}{\gamma}, \quad ||x_\infty - x_1|| \le \frac{\tau(\alpha) - \alpha}{\gamma}.$$

An easy consequence is

COROLLARY 42.1 *$||x_\infty - x_i|| \le \varepsilon$ for $i \ge 1 + \log\left|\log \frac{\tau(\alpha)}{\varepsilon\gamma}\right|$.*

Furthermore, by using the control Newton iterates $t_i = t_i(\beta, \gamma)$, a stricter estimate can be obtained under the same hypotheses:

THEOREM 42.2 $||x_i - x_{i-1}|| \leq t_i - t_{i-1}$.

Another property which is important for complexity discussions is the fact that α is upper semi-continuous, more precisely:

THEOREM 42.3 *Let* $\psi(u) := 2u^2 - 4u + 1$ *and* $u := \gamma(f, x_0)||x_0 - x||$. *Then*

$$\alpha(f, x) \leq \frac{\alpha(f, x_0)(1-u) + u}{\psi(u)^2}.$$

From the previous theorem it is possible to obtain a uniform estimate for Newton steps:

THEOREM 42.4 *There are universal constants* $\bar{\alpha} \approx .0802$ *and* $\bar{u} \approx .0221$ *with the following property: Let* $\bar{\gamma} > 0$ *and* $x, \zeta \in E$. *If* $\beta(f, \zeta) \leq \frac{\bar{\alpha}}{\bar{\gamma}}$ *and* $||x - \zeta|| \leq \frac{\bar{u}}{\bar{\gamma}}$, *then* $||\mathcal{N}_f(x) - \mathcal{N}_f^\infty(\zeta)|| \leq \frac{\bar{u}}{\bar{\gamma}}$.

This theorem is used to investigate the complexity of path following in the following way: Let $H : [0,1] \times E \to F$ be a continuous (homotopy) map which is analytic in the second argument. We further assume that a continuous solution path $\zeta : [0,1] \to E$ exists, i.e., $H(t, \zeta(t)) = 0$ for $t \in [0,1]$, such that the derivative $H_\zeta(t, \zeta(t))$ is an isomorphism. The following crude path-following method can be designed: choose a subdivision $0 = t_0 < t_1 < \ldots < t_k = 1$ and define

$$x_i := \mathcal{N}_{H(t_i, .)}(x_{i-1}) \text{ for } i = 1, \ldots, k. \tag{42.8}$$

It is clear that this method follows the solution curve if $||x_0 - \zeta(0)||$ and $|t_i - t_{i-1}|$ are small enough. Of course, the crucial number for complexity considerations is the number k of Newton steps involved in the above embedding method. If it is wished to obtain some points of the solution curve with high accuracy, then the complexity described in Theorem 42.1 has to be added.

The preceding analysis immediately furnishes a tool to determine the estimates necessary for a successful tracing of the solution curve:

THEOREM 42.5 *Let* $||x_0 - \zeta(0)|| \leq \frac{\bar{u}}{\bar{\gamma}}$, *and let the mesh* t_i *be so fine that* $\beta(H(t_i, .), \zeta(t_{i-1})) \leq \frac{\bar{\alpha}}{\bar{\gamma}}$ *and* $\gamma(H(t_i, .), \zeta(t_{i-1})) \leq \bar{\gamma}$. *Then the embedding method (42.8) follows the solution path* ζ. *In fact,* $||x_i - \zeta(t_i)|| \leq \frac{\bar{u}}{\bar{\gamma}}$.

43 Notes and Remarks

To summarize, we have outlined a program for approaching complexity investigations when Newton steps are the primary tool of path following methods. As can be seen from the last theorem, the success of the approach depends heavily on the availability of estimates $\beta(H(t,\,.),\zeta(s)) \le C_1|t-s|$ and $\gamma(H(t,\,.),\zeta(s)) \le C_2|t-s|$ with explicit constants C_1 and C_2.

This program was carried out by SHUB and SMALE [1991] for the case of a homotopy method for calculating all solutions of a system of polynomial equations (Bezout's theorem). A previous effort along similar lines was given by RENEGAR [1987].

Recently, this approach has also been used by RENEGAR and SHUB [1992] for a unified complexity analysis of various interior methods designed for solving linear and convex quadratic programming problems. They obtain and rederive various "polynomial time" estimates. The linear programming barrier method was first analysed by GONZAGA [1988]. The quadratic programming barrier method was analysed by GOLDFARB and LIU [1991]. A primal-dual linear programming algorithm was investigated by KOJIMA, MIZUNO, and YOSHISE [1988] and MONTEIRO and ADLER [1989]. The algorithm has roots in MEGIDDO [1988]. Primal-dual linear complementarity and quadratic programming algorithms were discussed by KOJIMA, MIZUNO, and YOSHISE [1989] and MONTEIRO and ADLER [1989]. All of the above algorithms follow the *central trajectory* studied by BAYER and LAGARIAS [1989] and MEGIDDO and SHUB [1989]. For the case of the linear complementarity problem, MIZUNO, YOSHISE, and KIKUCHI [1989] present several implementations and report computational experience which confirms the polynomial complexity.

The above discussion involved path following methods of Newton type. RENEGAR [1985, 1988b] has made complexity investigations for piecewise linear path following methods.

Chapter VIII

Available Software

We conclude the paper by listing some available software related to path following and indicate how the reader might access these codes. No attempt to compare or evaluate the various codes is offered. In any case, our opinion is that path following codes always need to be considerably adapted to the special purposes for which they are designed. The path following literature offers various tools for accomplishing such tasks. Although there are some general purpose codes, probably none will slay every dragon.

Rheinboldt, Roose, and Seydel [1990] present a list of features and options that appear to be necessary or desirable for continuation codes. This should be viewed as a guideline for people who want to create a new code.

Several of the codes can be accessed via *netlib*: The best way to obtain them is to ftp into `netlib@research.att.com`, login as netlib, password = your e-mail address. It is also possible to e-mail to netlib by writing *send index*. Information on how to proceed will then be e-mailed back to you.

ABCON

This is a predictor corrector continuation algorithm using variable order Adams-Bashforth predictors written by Lundberg and Poore [1991], see also Section 9. It can be obtained via anonymous ftp from

netlib@research.att.com
(ftp 192.20.225.2, ..., bin, get contin/abcon.f.Z).

ALCON

This package has been written by DEUFLHARD, FIEDLER, and KUNKEL [1987]. It is a continuation method for algebraic equations $f(x, \tau) = 0$, based on QR factorization as a solver for the equations arising in the Gauss-Newton iteration of the corrector step. Turning points and simple bifurcations can be computed on demand. It can be found in the electronic library of the Konrad Zuse Zentrum für Informationstechnik in Berlin. The reader may telnet or ftp to sc.ZIB-Berlin.de (130.73.108.11) and login under the user identification elib, no password is required. The sources can be found in the directory /pub/ELIB/codelib either in unpacked form or as a tar.Z file.

AUTO

This is a software package written by E. Doedel. It is mainly intended to investigate bifurcation phenomena. There is a charge of $175 for the software, a manual by DOEDEL and KERNÉVEZ [1986] is also available, contact:

S. K. Shull
Applied Mathematics, 217-50
California Institute of Technology
Pasadena, CA 91125
phone: (818) 356-4560

BIFPACK

This package has been written by SEYDEL [1991a]. It is meant primarily for bifurcation analysis of ODE's. This is not a public domain software. However, as a research tool, it is freely distributed for *noncommercial* use, except for a $20 contribution for handling. Indicate whether you prefer `BIFPACK` on 5.25" or on 3.5" diskette (1.4 MB, DOS double-density). Contact:

Prof. Rüdiger Seydel
Abt. Mathematik VI, Universität Ulm
Postfach 4066
W - 7900 Ulm, Fed. Rep. of Germany
e-mail: `seydel@rz.uni-ulm.dbp.de`

CANDYS/QA

This is a software system for the qualitative analysis of nonlinear dynamical systems by Feudel and Jansen [1992] of the Arbeitsgruppe 'Nichtlineare Dynamik', Max-Planck-Gesellschaft, Universität Potsdam, Germany.

CONKUB

This is an interactive program for continuation and bifurcation of large systems of nonlinear equations written by Mejia [1986], see also Mejia[1990]. It is currently available from him via e-mail: `ray@helix.nih.gov`.

DERPAR

This package was written by Kubíček [1976] and Holodniok and Kubíček [1984]. This is a Fortran subprogram for the evaluation of the dependence of the solution of a nonlinear system on a parameter. The modified method of Davidenko, which applies the implicit function theorem, is used in combination with Newton's method and Adam's integration formulas. The program can be accessed via netlib, see number 502 in the directory *toms*.

DSTOOL

This is a computer assisted exploration package for dynamical systems by Back, Guckenheimer, Myers, Wiclin, and Worfolk [1992] of the Center for Applied Mathematics, Cornell University, Ithaca.

DYNAMICS

This is a software for the numerical exploration of chaotic systems developed by Nusse and Yorke [1992] of the University of Maryland.

HOMPACK

This is a suite of FORTRAN 77 subroutines for solving nonlinear systems of equations by homotopy methods, written by L. T. Watson, see Watson, Billups, and Morgan [1987]. There are subroutines for fixed point, zero finding, and general homotopy curve tracking problems, utilizing both dense and sparse Jacobian matrices, and implementing three different algorithms:

ODE-based, normal flow, and augmented Jacobian. The program can be accessed via netlib under the directory *hompack*. See also number 652 in the directory *toms*.

INSITE

This is a set of practical numerical algorithms for chaotic systems written by PARKER and CHUA [1989] of INSITE Software, Berkeley.

KAOS

This is a computational environment for exploring dynamical systems developed by GUCKENHEIMER and KIM [1991] of the Center for Applied Mathematics, Cornell University, Ithaca.

LOCBIF

A. Khibnik and collaborators in Moscow have developed several codes for path following and bifurcation analysis. `CYCLE` is a one-parameter continuation program for limit cycles. `LINLBF` has been designed for multi-parameter bifircation analysis of equilibrium points, limit cycles, fixed points of maps, respectively. `LOCBIF`, developed by KHIBNIK, KUZNETSOV, LEVITIN, and NIKOLAEV [1993], is a package involving contination techniques and interactive software for bifurcation analysis of ODEs and iterated maps, built originally on top of `LINLBF`. People interested in trying this software should contact A. Khibnik via e-mail: `na.khibnik@na-net.ornl.gov` or `khibnik@impb.serpukhov.su`.

MacMath

This is a dynamical systems software package for the Macintosh, developed by HUBBARD and WEST [1992] of Cornell University, Ithaca.

OB1

This interior point method has been written by I. J. Lustig, R. E. Marsten, and D. F. Shanno. The version of `OB1` that implements a primal-dual algorithm for linear programming is available in source code form to academics from Roy Marsten at Georgia Tech. This is the December 1989 version, also

known as the WRIP (Workshop on Research in Programming) version. The current version of `OB1` is commercial. It implements a primal-dual predictor corrector algorithm for linear programming and is available from XMP Software at prices ranging from $15,000 to $100,000.

XMP Software
Suite 279, Bldg 802
930 Tahoe Blvd
Incline Village, NV 89451
phone: (702) 831- 4XMP
e-mail: `tlowe@mcimail.com`

PATH

This software package for dynamical systems was originally coded in FORTRAN 77 by Kaas-Petersen [1989], and is currently modified to include a graphical interface. According to the workers at the Technical University of Denmark, it seems to be able to handle much larger systems of ODE's than `AUTO`. For more details and availability, readers may contact Michael Rose via e-mail: `lamfmr@lamf.dth.dk`.

Phase Plane — XPP

This is a dynamical systems tool developed by Ermentrout [1990] of the University of Pittsburgh.

PITCON

This is a Fortran subprogram for continuation and limit points, written by Rheinboldt and Burkardt [1983a, 1983b]. It is used for computing solutions of a nonlinear system of equations containing a parameter. The location of target points where a given variable has a specified value can be located. Limit points are also identified. It uses a local parameterization based on curvature estimates to control the choice of parameter value. The program can be accessed via netlib under the directory *contin*. See also number 596 in the directory *toms*.

PLALGO

This is a software for piecewise linear homotopy methods developed by TODD [1981]. It can be obtained from him via e-mail:

`miketodd@orie.cornell.edu`.

No support is available, and he says that on-line documentation is weak, although he can send a hard copy.

pla_s_k

This is a C program, written by WIDMANN [1990a], for triangulating surfaces in $\mathbb{R}^3$ which are implicitly defined, see Section 40. It incorporates mesh smoothing and some other features. It is particularly suited for mesh generation (e.g., for boundary element methods) and for visualization purposes. The program can be obtained via e-mail (`Georg@Math.ColoState.Edu`).

PLTMG

This package has been written by R. E. Bank, see also the paper of BANK and CHAN[1986]. It solves elliptic partial differential equations in general regions of the plane. It features adaptive local mesh refinement, multigrid iteration, and a pseudo arclength continuation option for parameter dependencies. The package includes an initial mesh generator and several graphics packages. Full documentation can be obtained in the `PLTMG` User's Guide by R. E. Bank, available from SIAM publications

(e-mail: `SIAMPUBS@wharton.upenn.edu`).

The program can be accessed via netlib under the directory *pltmg*.

SYMCON

This is a special path following program featuring symbolic exploitation of symmetry developed by GATERMANN and HOHMANN [1991] of the Konrad-Zuse-Zentrum für Informationstechnik, Berlin, Germany. SYMCON is available by anonymous ftp from elib.zib-berlin.de (130.73.108.11) in the subdirectory pub/symcon.

TraX

This is a program for the simulation and analysis of dynamical systems developed by Levitin and Khibnik of the Institute of Mathematical Problems in Biology, Russian Academy of Sciences, Pushchino, see KHIBNIK [1990]. It is available from Exeter Software, Setauket, NY.

Last and Least

The book ALLGOWER and GEORG [1990] contains several Fortran codes for path following which are to be regarded primarily as illustrations. The intention was to encourage the readers to experiment and be led to make improvements and adaptations suited to their particular applications. We emphasize that these programs should not be regarded as programs of library quality. They can be obtained via e-mail (`Georg@Math.ColoState.Edu`).

Bibliography

Abraham, R. and J. Robbin, (1967), *Transversal Mappings and Flows*, W. A. Benjamin, New York.

Adler, I., G. C. Resende, G. Veiga, and N. Karmarkar, (1989), An implementation of Karmarkar's algorithm for linear programming, *Math. Programming*, 44:297–335.

Alexander, J. C. and E. V. Slud, (1983), Global convergence rates of piecewise-linear continuation methods: a probabilistic approach, in: Eaves, B. C., F. J. Gould, H.-O. Peitgen, and M. J. Todd, editors, *Homotopy Methods and Global Convergence*, pages 15–30, New York, Plenum Press.

Alexander, J. C., (1987), Average intersection and pivoting densities, *SIAM J. Numer. Anal.*, 24:129–146.

Allgower, E. L. and C.-S. Chien, (1986), Continuation and local perturbation for multiple bifurcations, *SIAM J. Sci. Statist. Comput.*, 7:1265–1281.

Allgower, E. L. and A. F. Fässler, (1993), Blockstructure and equivariance of matrices, submitted.

Allgower, E. L. and K. Georg, (1979), Generation of triangulations by reflections, *Utilitas Math.*, 16:123–129.

Allgower, E. L. and K. Georg, (1980), Homotopy methods for approximating several solutions to nonlinear systems of equations, in: Forster, W., editor, *Numerical Solution of Highly Nonlinear Problems*, pages 253–270, Amsterdam, North-Holland.

Allgower, E. L. and K. Georg, (1983a), Predictor-corrector and simplicial methods for approximating fixed points and zero points of nonlinear mappings, in: Bachem, A., M. Grötschel, and B. Korte,

editors, *Mathematical Programming: The State of the Art*, pages 15–56, Berlin, Heidelberg, New York, Springer Verlag.

Allgower, E. L. and K. Georg, (1983b), Relationships between deflation and global methods in the problem of approximating additional zeros of a system of nonlinear equations, in: Eaves, B. C., F. J. Gould, H.-O. Peitgen, and M. J. Todd, editors, *Homotopy Methods and Global Convergence*, pages 31–42, New York, Plenum Press.

Allgower, E. L. and K. Georg, (1989), Estimates for piecewise linear approximations of implicitly defined manifolds, *Appl. Math. Lett.*, 1(5):1–7.

Allgower, E. L. and K. Georg, (1990), *Numerical Continuation Methods: An Introduction*, volume 13 of *Series in Computational Mathematics*, Springer Verlag, Berlin, Heidelberg, New York, pp 388.

Allgower, E. L. and K. Georg, (1993), Continuation and path following, *Acta Numerica*, 2:1–64.

Allgower, E. L. and S. Gnutzmann, (1987), An algorithm for piecewise linear approximation of implicitly defined two-dimensional surfaces, *SIAM J. Numer. Anal.*, 24:452–469.

Allgower, E. L. and S. Gnutzmann, (1991), Simplicial pivoting for mesh generation of implicitly defined surfaces, *Computer Aided Geometric Design*, 8:305–325.

Allgower, E. L. and P. H. Schmidt, (1985), An algorithm for piecewise-linear approximation of an implicitly defined manifold, *SIAM J. Numer. Anal.*, 22:322–346.

Allgower, E. L. and H. Schwetlick, (1993), On minimally extended systems for simple bifurcation points, Preprint.

Allgower, E. L., C.-S. Chien, and K. Georg, (1989), Large sparse continuation problems, *J. Comput. Appl. Math.*, 26:3–21.

Allgower, E. L., K. Böhmer, and Z. Mei, (1991a), A complete bifurcation scenario for the 2-D nonlinear Laplacian with Neumann boundary conditions, in: Seydel, R., F. W. Schneider, T. Küpper, and H. Troger, editors, *Bifurcation and Chaos: Analysis, Algorithms, Applications*, volume 97 of *ISNM*, pages 1–18, Basel, Switzerland, Birkhäuser.

Allgower, E. L., K. Böhmer, and Z. Mei, (1991b), On a problem decomposition for semi-linear nearly symmetric elliptic problems, in: Hackbusch, W., editor, *Parallel Algorithms for PDE's*, volume 31 of *Notes on Numerical Fluid Mechanics*, pages 1–17, Braunschweig, Fed. Rep. Germany, Vieweg Verlag.

Allgower, E. L., K. Böhmer, and Z. Mei, (1991c), On new bifurcation results for semi-linear elliptic equations with symmetries, in: Whiteman, J., editor, *The Mathematics of Finite Elements and Applications VII*, pages 487–494, Brunel, England, MAFELAP 1990, Academic Press.

Allgower, E. L., C.-S. Chien, K. Georg, and C.-F. Wang, (1991d), Conjugate gradient methods for continuation problems, *J. Comput. Appl. Math.*, 38:1–16.

Allgower, E. L., K. Georg, and R. Widmann, (1991e), Volume integrals for boundary element methods, *J. Comput. Appl. Math.*, 38:17–29.

Allgower, E. L., K. Böhmer, K. Georg, and R. Miranda, (1992a), Exploiting symmetry in boundary element methods, *SIAM J. Numer. Anal.*, 29:534–552.

Allgower, E. L., K. Böhmer, and M. Golubitsky, editors, (1992b), *Bifurcation and Symmetry*, volume 104 of *ISNM*, Basel, Switzerland, Birkhäuser Verlag.

Allgower, E. L., K. Georg, and R. Miranda, (1993a), Exploiting permutation symmetry with fixed points in linear equations, in: Allgower, E. L., K. Georg, and R. Miranda, editors, *Exploiting Symmetry in Applied and Numerical Analysis*, volume 29 of *Lectures in Applied Mathematics*, pages 23–36, Providence, RI, American Mathematical Society.

Allgower, E. L., K. Georg, and J. Walker, (1993b), Exploiting symmetry in 3D boundary element methods, in: Agarwal, R. P., editor, *Contributions in Numerical Mathematics*, volume 2 of *World Scientific Series in Applicable Analysis*, pages 15–25, World Scientific Publ. Comp., Singapore.

Allgower, E. L., (1984), Bifurcations arising in the calculation of critical points via homotopy methods, in: Küpper, T., H. D. Mittelmann, and H. Weber, editors, *Numerical Methods for Bifurcation Problems*, volume 70 of *ISNM*, pages 15–28, Basel, Birkhäuser Verlag.

Allison, D. C. S., S. Harimoto, and L. T. Watson, (1989), The granularity of parallel homotopy algorithms for polynomial systems of equations, *Intern. J. Comput. Math.*, 29:21–37].

Amann, H., (1976), Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces, *SIAM Rev.*, 18:620–709.

Ambrosetti, A. and P. Hess, (1980), Positive solutions of asymptotically linear elliptic eigenvalue problems, *J. Math. Anal. Appl.*, 73:411–422.

Atkinson, K. E., (1993), *User's Guide to a Boundary Element Package for Solving Integral Equations on Piecewise Smooth Surfaces*, Report 44, Univ. of Iowa.

Back, A., J. Guckenheimer, M. Myers, F. Wiclin, and P. Worfolk, (1992), Dstool; computer assisted exploration of dynamical systems, *Notices of the ACM*, 39:303–309.

Ballisti, R., C. Hafner, and P. Leuchtman, (1982), Application of the representation theory of finite groups to field computation problems with symmetrical boundaries, *IEEE Trans. On Magnetics*, 18:584–587.

Ballman, J., R. Eppler, and W. Hackbusch, (1988), Panel method in fluid mechanics with emphasis on aerodynamics, in: Hackbusch, W., editor, *Notes on Numerical Fluid Dynamics*, volume 21, Braunschweig, Vieweg.

Bank, R. E. and T. F. Chan, (1986), PLTMGC: A multi-grid continuation program for parameterized nonlinear elliptic systems, *SIAM J. Sci. Statist. Comput.*, 7:540–559.

Bank, B., J. Guddat, D. Klatte, B. Kummer, and K. Tammer, (1983), *Non-Linear Parametric Optimization*, Birkhäuser, Basel.

Bayer, D. and J. C. Lagarias, (1989), The nonlinear geometry of linear programming, I: Affine and projective scaling trajectories, II: Legendre transform coordinates and central trajectories, *Trans. Amer. Math. Soc.*, 314:499–581.

Ben-Israel, A. and T. N. E. Greville, (1974), *Generalized Inverses: Theory and Applications*, John Wiley and Sons, New York.

Bernstein, S., (1910), Sur la généralisation du problème de Dirichlet., *Math. Ann.*, 69:82–136.

Bernstein, D. N., (1975), The number of roots of a system of equations, *Functional Analysis and Appl.*, 9:1–4.

Bertsekas, D. P., (1984), *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, London.

Beyn, W.-J., (1992), On smoothness and invariance properties of the Gauss-Newton method, Preprint, Univ. of Bielefeld, Germany.

Bixby, R. E., J. W. Gregory, I. J. Lustig, R. E. Marsten, and D. F. Shanno, (1991), Very large-scale linear programming: A case study in combining interior point and simplex methods, Technical Report RRR 34-91, Rutgers Univ., New Brunswick, NJ.

Bolstad, J. H. and H. B. Keller, (1986), A multigrid continuation method for elliptic problems with folds, *SIAM J. Sci. Statist. Comput.*, 7:1081–1104.

Bossavit, A., (1986), Symmetry, groups, and boundary value problems. A progressive introduction to noncommutative harmonic analysis of partial differential equations in domains with geometrical symmetry, *Comput. Methods Appl. Mech. and Engrg.*, 56:167–215.

Bossavit, A., (1993), Boundary value problems with symmetry and their approximation by finite elements, *SIAM J. Appl. Math.*, 53:1352–1380.

Bourji, S. K. and H. F. Walker, (1990), Least-change secant updates of nonsquare matrices, *SIAM J. Numer. Math.*, 27:1263–1294.

Branin, F. H. and S. K. Hoo, (1972), A method for finding multiple extrema of a function of n variables, in: Lootsma, F. A., editor, *Numerical Methods for Non-linear Optimization*, pages 231–237, New York, London, Academic Press.

Branin, F. H., (1972), Widely convergent method for finding multiple solutions of simultaneous nonlinear equations, *IBM J. Res. Develop.*, 16:504–522.

Brezzi, F., J. Rappaz, and P. A. Raviart, (1980a), Finite dimensional approximation of nonlinear problems. Part 1: Branches of nonsingular solutions, *Numer. Math.*, 36:1–25.

Brezzi, F., J. Rappaz, and P. A. Raviart, (1980b), Finite dimensional approximation of nonlinear problems. Part 2: Limit points, *Numer. Math.*, 37:1–28.

Brezzi, F., J. Rappaz, and P. A. Raviart, (1981), Finite dimensional approximation of nonlinear problems. Part 3: Simple bifurcation points, *Numer. Math.*, 38:1–30.

Brodzik, M. L. and W. C. Rheinboldt, (1994), On the computation of simplicial approximations of implicitly defined two-dimensional manifolds, preprint University of Pittsburgh.

Brooks, P. S., (1980), Infinite retrogression in the Eaves-Saigal algorithm, *Math. Programming*, 19:313–327.

Brouwer, L. E. J., (1912), Über Abbildung von Mannigfaltigkeiten, *Math. Ann.*, 71:97–115.

Brown, K. M. and W. B. Gearhart, (1971), Deflation techniques for the calculation of further solutions of a nonlinear system, *Numer. Math.*, 16:334–342.

Broyden, C. G., (1965), A class of methods for solving nonlinear simultaneous equations, *Math. Comp.*, 19:577–593.

Byrne, G. D. and L. A. Baird, (1985), Distillation calculations using a locally parametrized continuation method, *Comput. chem. Engng.*, 9:593–599.

Canny, J. and J. M. Rojas, (1991), An optimal condition for determining the exact number of roots of a polynomial system, in: *Proceedings of the ISSAC 91, Bonn, Germany*, pages 96–102.

Carpenter, T. J. and D. F. Shanno, (1991), An interior point method for quadratic programs based on conjugate projected gradients, Technical Report RRR 55-91, Rutgers Univ., New Brunswick, NJ.

Chan, T. F. and H. B. Keller, (1982), Arc-length continuation and multi-grid techniques for nonlinear eigenvalue problems, *SIAM J. Sci. Statist. Comput.*, 3:173–194.

Chan, T. F. and Y. Saad, (1985), Iterative methods for solving bordered systems with applications to continuation methods, *SIAM J. Sci. Statist. Comput.*, 6:438–451.

Chan, T. F., (1984a), Deflation techniques and block-elimination algorithms for solving bordered singular systems, *SIAM J. Sci. Stat. Comp.*, 5:121–134.

CHAN, T. F., (1984b), Newton-like pseudo-arclength methods for computing simple turning points, *SIAM J. Sci. Statist. Comput.*, 5:135–148.

CHAN, T. F., (1984c), Techniques for large sparse systems arising from continuation methods, in: KÜPPER, T., H. D. MITTELMANN, and H. WEBER, editors, *Numerical Methods for Bifurcation Problems*, volume ISNM 70, pages 116–128, Basel, Birkhäuser Verlag.

CHAO, K. S. and R. SAEKS, (1977), Continuation methods in circuit analysis, *Proc. IEEE*, 65:1187–1194.

CHAVEZ, R., J. D. SEADER, and T. L. WAYBURN, (1986), Multiple steady state solutions for interlinked separation systems, *Ind. Eng. Chem. Fundam.*, 25:566–576.

CHIEN, C.-S., (1989), Secondary bifurcations in the buckling problem, *J. Comput. Appl. Math.*, 25:277–287.

CHOW, S. N. and J. K. HALE, (1982), *Methods of Bifurcation Theory*, Springer Verlag, New York.

CHOW, S. N., J. MALLET-PARET, and J. A. YORKE, (1978), Finding zeros of maps: Homotopy methods that are constructive with probability one, *Math. Comp.*, 32:887–899.

CHOW, S. N., J. MALLET-PARET, and J. A. YORKE, (1979), A homotopy method for locating all zeros of a system of polynomials, in: PEITGEN, H.-O. and H.-O. WALTHER, editors, *Functional Differential Equations and Approximation of Fixed Points*, volume 730 of *Lecture Notes in Math.*, pages 77–88, Berlin, Heidelberg, New York, Springer Verlag.

CHU, M. T., T.-Y. LI, and T. SAUER, (1988), Homotopy methods for general λ-matrix problems, *SIAM J. Matrix Anal. Appl.*, 9:528–536.

CHU, K.-W. E., W. GOVAERTS, and A. SPENCE, (1994), Matrices with rank deficiency two in eigenvalue problems and dynamical systems, *SIAM J. Num.Anal.*, 31:524–539.

CHU, M. T., (1984a), The generalized Toda flow, the QR algorithm and the center manifold theory, *SIAM J. Alg. Disc. Meth.*, 5:187–201.

CHU, M. T., (1984b), A simple application of the homotopy method to symmetric eigenvalue problems, *Linear Algebra Appl.*, 59:85–90.

Chu, M. T., (1986), A continuous approximation to the generalized Schur's decomposition, *Linear Algebra Appl.*, 78:119–132.

Chu, M. T., (1988), On the continuous realization of iterative processes, *SIAM Rev.*, 30:375–387.

Chu, M. T., (1990), Solving additive inverse eigenvalue problems for symmetric matrices by the homotopy method, *IMA J. Numer. Anal.*, 9:331–342.

Chu, M. T., (1991), A continuous Jacobi-like approach to the simultaneous reduction of real matrices, *Linear Algebra Appl.*, 147:75–96.

Chua, L. O. and A. Ushida, (1976), A switching-parameter algorithm for finding multiple solutions of nonlinear resistive circuits, *Circuit Theory and Appl.*, 4:215–239.

Cicogna, G., (1981), Symmetry breakdown from bifurcation, *Lett. Nuovo Cimento*, 31:600–602.

Cliffe, K. A. and K. H. Winters, (1986), The use of symmetry in bifurcation calculations and its application to the Bénard problem, *J. Comput. Physics*, 67:310–326.

Cohen, A. I., (1972), Rate of convergence of several conjugate gradient algorithms, *SIAM J. Numer. Anal.*, 9:248–259.

Cottle, R. W. and G. B. Dantzig, (1968), Complementary pivot theory of mathematical programming, *Linear Algebra Appl.*, 1:103–125.

Cottle, R. W., G. H. Golub, and R. S. Sacher, (1978), On the solution of large structured linear complementarity problems: The block partitioned case, *Appl. Math. Optim.*, 4:347–363.

Cottle, R. W., J.-S. Pang, and R. E. Stone, (1992), *The Linear Complementarity Problem*, Academic Press.

Cottle, R. W., (1974), Solution rays for a class of complementarity problems, *Math. Programming Study*, 1:58–70.

Coxeter, H. S. M., (1934), Discrete groups generated by reflections, *Ann. of Math.*, 6:13–29.

Coxeter, H. S. M., (1963), *Regular Polytopes*, MacMillan, New York, Third edition.

Crandall, M. G. and P. H. Rabinowitz, (1971), Bifurcation from simple eigenvalues, *J. Funct. Anal.*, 8:321–340.

Crouzeix, M. and J. Rappaz, (1990), *On Numerical Approximation in Bifurcation Theory*, RMA, Masson, Paris.

Dai, R.-X. and W. C. Rheinboldt, (1990), On the computation of manifolds of fold points for parameter-dependent problems, *SIAM J. Numer. Anal.*, 27:437–446.

Dai, Y. and Y. Yamamoto, (1989), The path following algorithm for stationary point problems on polyhedral cones, *J. Op. Res. Soc. of Japan*, 32:286–309.

Dai, Y., K. Sekitani, and Y. Yamamoto, (1992), A variable dimension algorithm with the Dantzig-Wolfe decomposition for structured stationary point problems, *ZOR — Methods and Models of Operations Research*, 36:23–53.

Dantzig, G. B., (1963), *Linear Programming and Extensions*, Princeton Univ. Press, Princeton, NJ.

Davidenko, D., (1953), On a new method of numerical solution of systems of nonlinear equations, *Dokl. Akad. Nauk USSR*, 88:601–602, in Russian.

Davis, P. J. and P. Rabinowitz, (1984), *Methods of Numerical Integration*, Academic Press, Orlando, Second edition.

Davis, P. J., (1979), *Circulant Matrices*, John Wiley and Sons, New York.

Dellnitz, M. and B. Werner, (1989), Computational methods for bifurcation problems with symmetries — with special attention to steady state and Hopf bifurcation points, *J. Comput. Appl. Math.*, 26:97–123.

Den Heijer, C. and W. C. Rheinboldt, (1981), On steplength algorithms for a class of continuation methods, *SIAM J. Numer. Anal.*, 18:925–948.

Dennis, J. E. and R. B. Schnabel, (1983), *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice Hall, Englewood Cliffs, NJ.

deSa, C., K. M. Irani, C. G. Ribbens, L. T. Watson, and H. F. Walker, (1992), Preconditioned iterative methods for homotopy curve tracking, *SIAM J. Sci. Stat. Comput.*, 13:30–46.

DEUFLHARD, P. and G. HEINDL, (1979), Affine invariant convergence theorems for Newton's method and extensions to related methods, *SIAM Numer. Anal.*, 16:1–10.

DEUFLHARD, P., B. FIEDLER, and P. KUNKEL, (1987), Efficient numerical pathfollowing beyond critical points, *SIAM J. Numer. Anal.*, 24:912–927.

DOEDEL, E. and J. P. KERNÉVEZ, (1986), *AUTO: Software for Continuation and Bifurcation Problems in Ordinary Differential Equations*, California Institute of Technology.

DOEDEL, E., H. B. KELLER, and J. P. KERNÉVEZ, (1991a), Numerical analysis and control of bifurcation problems. Part I: Bifurcation in finite dimensions, *Int. J. of Bifurcation and Chaos*, 1.

DOEDEL, E., H. B. KELLER, and J. P. KERNÉVEZ, (1991b), Numerical analysis and control of bifurcation problems. Part II: Bifurcation in infinite dimensions, *Int. J. of Bifurcation and Chaos*, 1.

DUNFORD, N. and J. T. SCHWARTZ, (1963), *Linear Operators. Part II: Spectral Theory*, Interscience Publ., New York.

EAVES, B. C. and R. SAIGAL, (1972), Homotopies for computation of fixed points on unbounded regions, *Math. Programming*, 3:225–237.

EAVES, B. C. and H. SCARF, (1976), The solution of systems of piecewise linear equations, *Math. Oper. Res.*, 1:1–27.

EAVES, B. C. and J. A. YORKE, (1984), Equivalence of surface density and average directional density, *Math. Oper. Res.*, 9:363–375.

EAVES, B. C., (1972), Homotopies for the computation of fixed points, *Math. Programming*, 3:1–22.

EAVES, B. C., (1976), A short course in solving equations with PL homotopies, in: COTTLE, R. W. and C. E. LEMKE, editors, *Nonlinear Programming*, volume 9 of *SIAM-AMS Proc.*, pages 73–143, Providence, RI, AMS.

EAVES, B. C., (1984), *A Course in Triangulations for Solving Equations with Deformations*, volume 234 of *Lecture Notes in Economics and Mathematical Systems*, Springer Verlag, Berlin, Heidelberg, New York.

ERMENTROUT, B., (1990), *PhasePlane: The Dynamical Systems Tool*, Brooks/Cole Publishing.

FÄSSLER, A., (1976), *Application of Group Theory to the Method of Finite Elements*, PhD thesis, ETH Zürich, Switzerland.

FEUDEL, U. and W. JANSEN, (1992), CANDYS/QA — A software system for the qualitative analysis of nonlinear dynamical systems, *Bifurcation & Chaos*, to appear.

FIACCO, A. V. and G. MCCORMICK, (1968), *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley and Sons, New York, NY.

FIACCO, A. V., (1983), *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York.

FIACCO, A. V., editor, (1984), *Sensitivity, Stability and Parametric Analysis*, volume 21 of *Mathematical Programming Study*, Amsterdam, North-Holland.

FINK, J. P. and W. C. RHEINBOLDT, (1983), On the discretization error of parametrized nonlinear equations, *SIAM J. Numer. Anal.*, 20:732–746.

FINK, J. P. and W. C. RHEINBOLDT, (1984), Solution manifolds and submanifolds of parametrized equations and their discretization errors, *Numer. Math.*, 45:323–343.

FINK, J. P. and W. C. RHEINBOLDT, (1985), Local error estimates for parametrized nonlinear equations, *SIAM J. Numer. Anal.*, 22:729–735.

FINK, J. P. and W. C. RHEINBOLDT, (1986), Folds on the solution manifold of a parametrized equation, *SIAM J. Numer. Anal.*, 23:693–706.

FINK, J. P. and W. C. RHEINBOLDT, (1987), A geometric framework for the numerical study of singular points, *SIAM J. Numer. Anal.*, 24:618–633.

FLETCHER, R., (1987), *Practical Methods of Optimization*, J. Wiley, New York, second edition.

FREUDENTHAL, H., (1942), Simplizialzerlegungen von beschränkter Flachheit, *Ann. of Math.*, 43:580–582.

FREUND, R. W., G. H. GOLUB, and N. M. NACHTIGAL, (1992), Iterative solution of linear systems, *Acta Numerica*, 1:57–100.

FREUND, R. M., (1984a), Variable dimension complexes. Part I: Basic theory, *Math. Oper. Res.*, 9:479–497.

FREUND, R. M., (1984b), Variable dimension complexes. Part II: A unified approach to some combinatorial lemmas in topology, *Math. Oper. Res.*, 9:498–509.

FREUND, R. M., (1991), A potential-function reduction algorithm for solving a linear program directly from an infeasible "warm start", in: ROOS, C. and J.-P. VIAL, editors, *Interior Point Methods for Linear Programming: Theory and Practice*, volume 52 of *Math. Programming, Ser. B*, pages 441–446, Amsterdam, Mathematical Programming Society, North-Holland.

FULTON, W., (1984), *Intersection Theory*, Springer Verlag.

GARCIA, C. B. and F. J. GOULD, (1978), A theorem on homotopy paths, *Math. Oper. Res.*, 3:282–289.

GARCIA, C. B. and F. J. GOULD, (1980), Relations between several path following algorithms and local and global Newton methods, *SIAM Rev.*, 22:263–274.

GARCIA, C. B. and W. I. ZANGWILL, (1979), Finding all solutions to polynomial systems and other systems of equations, *Math. Programming*, 16:159–176.

GARCIA, C. B. and W. I. ZANGWILL, (1981), *Pathways to Solutions, Fixed Points, and Equilibria*, Prentice-Hall, Englewood Cliffs, NJ.

GARRATT, T. J., G. MOORE, and A. SPENCE, (1991), Two methods for the numerical detection of Hopf bifurcations, in: SEYDEL, R., F. W. SCHNEIDER, T. KÜPPER, and H. TROGER, editors, *Bifurcation and Chaos: Analysis, Algorithms, Applications*, volume 97 of *ISNM*, pages 129–134, Basel, Switzerland, Birkhäuser.

GATERMANN, K. and A. HOHMANN, (1991), Symbolic exploitation of symmetry in numerical pathfollowing, *Impact of Computing in Science and Engineering*, 3:330–365.

GEORG, K. and R. MIRANDA, (1992), Exploiting symmetry in solving linear equations, in: ALLGOWER, E. L., K. BÖHMER, and M. GOLUBITSKY, editors, *Bifurcation and Symmetry*, volume 104 of *ISNM*, pages 157–168, Basel, Switzerland, Birkhäuser Verlag.

GEORG, K. and R. MIRANDA, (1993), Symmetry aspects in numerical linear algebra with applications to boundary element methods, in: ALLGOWER, E. L., K. GEORG, and R. MIRANDA, editors, *Exploiting Symmetry in Applied and Numerical Analysis*, volume 29 of *Lectures in Applied*

Mathematics, pages 213–228, Providence, RI, American Mathematical Society.

GEORG, K. and J. TAUSCH, (1993), Some error estimates for the numerical approximation of surface integrals, *Math. Comp.*, 62:755–763.

GEORG, K. and R. WIDMANN, (1993), Adaptive quadratures over surfaces, Preprint, Colorado State University.

GEORG, K., (1981a), A numerically stable update for simplicial algorithms, in: ALLGOWER, E. L., K. GLASHOFF, and H.-O. PEITGEN, editors, *Numerical Solution of Nonlinear Equations*, volume 878 of *Lecture Notes in Math.*, pages 117–127, Berlin, Heidelberg, New York, Springer Verlag.

GEORG, K., (1981b), On tracing an implicitly defined curve by quasi-Newton steps and calculating bifurcation by local perturbation, *SIAM J. Sci. Stat. Comput.*, 2:35–50.

GEORG, K., (1982), Zur numerischen Realisierung von Kontinuitätsmethoden mit Prädiktor-Korrektor- oder simplizialen Verfahren, Habilitationsschrift, Univ. Bonn, Germany.

GEORG, K., (1983), A note on stepsize control for numerical curve following, in: EAVES, B. C., F. J. GOULD, H.-O. PEITGEN, and M. J. TODD, editors, *Homotopy Methods and Global Convergence*, pages 145–154, New York, Plenum Press.

GEORG, K., (1990), An introduction to PL algorithms, in: ALLGOWER, E. L. and K. GEORG, editors, *Computational Solution of Nonlinear Systems of Equations*, volume 26 of *Lectures in Applied Mathematics*, pages 207–236, Providence, RI, American Mathematical Society.

GEORG, K., (1991), Approximation of integrals for boundary element methods, *SIAM J. Sci. Stat. Comput.*, 12:443–453.

GFRERER, H., J. GUDDAT, and H.-J. WACKER, (1983), A globally convergent algorithm based on imbedding and parametric optimization, *Computing*, 30:225–252.

GFRERER, H., J. GUDDAT, H.-J. WACKER, and W. ZULEHNER, (1985), Path-following for Kuhn-Tucker curves by an active set strategy, in: BAGHI, A. and H. T. JONGEN, editors, *System and Optimization*, volume 66 of *Lect. Notes Contr. Inf. Sc.*, pages 111–132, Springer-Verlag.

GILL, P. E., G. H. GOLUB, W. MURRAY, and M. A. SAUNDERS, (1974), Methods for modifying matrix factorizations, *Math. Comp.*, 28:505–535.

Gill, P. E., W. Murray, M. A. Saunders, J. A. Tomlin, and M. H. Wright, (1986), On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method, *Math. Programming*, 36:183–209.

Glowinski, R., H. B. Keller, and L. Reinhart, (1985), Continuation-conjugate gradient methods for the least squares solution of nonlinear boundary value problems, *SIAM J. Sci. Statist. Comput.*, 6:793–832.

Gnutzmann, S., (1989), *Stückweise lineare Approximation implizit definierter Mannigfaltigkeiten*, PhD thesis, University of Hamburg, Fed. Rep. Germany.

Goldfarb, D. and S. Liu, (1991), An algorithm for solving linear programming problems in $O(n^3L)$ operations, *Math. Programming*, 49:325–340.

Golub, G. H. and C. F. van Loan, (1989), *Matrix Computations*, J. Hopkins University Press, Baltimore, London, second edition.

Golubitsky, M. and D. G. Schaeffer, (1985), *Singularities and Groups in Bifurcation Theory*, volume 1, Springer-Verlag, Berlin, Heidelberg, New York.

Golubitsky, M., I. Stewart, and D. G. Schaeffer, (1988), *Singularities and Groups in Bifurcation Theory*, volume 2, Springer-Verlag, Berlin, Heidelberg, New York.

Gonzaga, C. C., (1988), An algorithm for solving linear programming problems in $O(n^3L)$ operations, in: Megiddo, N., editor, *Progress in Mathematical Programming, Interior Point and Related Methods*, pages 1–28, New York, Springer.

Gonzaga, C. C., (1992), Path-following methods for linear programming, *SIAM Rev.*, 34:167–224.

Górniewicz, L., (1976), Homological methods in fixed point theory of multivalued maps, *Diss. Math.*, 129.

Gould, F. J. and J. W. Tolle, (1983), *Complementary Pivoting on a Pseudomanifold Structure with Applications on the Decision Sciences*, volume 2 of *Sigma Series in Applied Mathematics*, Heldermann Verlag, Berlin.

GRIEWANK, A. and G. CORLISS, editors, (1991), *Automatic Differentiation of Algorithms: Theory, Implementation, and Application*, SIAM, Philadelphia, 1991.

GRIEWANK, A. and G. W. REDDIEN, (1984), Characterization and computation of generalized turning points, *SIAM J. Numer. Anal.*, 21:176–185.

GRIEWANK, A., (1985), On solving nonlinear equations with simple singularities or nearly singular solutions, *SIAM Rev.*, 27:537–563.

GRIEWANK, A., (1989), On automatic differentiation, in: IRI, M. and K. TANABE, editors, *Mathematical Programming: Recent Developments and Applications*, pages 83–108, Kluwer.

GUCKENHEIMER, J. and S. KIM, (1991), Computational environments for exploring dynamical systems, *Bifurcation & Chaos*, 1:269–276.

GUDDAT, J., H. T. JONGEN, B. KUMMER, and F. NOŽIČKA, editors, (1987), *Parametric Optimization and Related Topics*, Berlin, Akademie Verlag.

GUDDAT, J., F. GUERRA VASQUEZ, and H. T. JONGEN, (1990), *Parametric Optimization: Singularities, Path Following, and Jumps*, John Wiley and Sons, Chichester, England.

GUY, J. and B. MANGEOT, (1981), Use of group theory in various integral equations, *SIAM J. Appl. Math.*, 40:390–399.

HACKBUSCH, W., (1989), *Integralgleichungen. Theorie und Numerik*, B. G. Teubner, Stuttgart, Germany.

HACKBUSCH, W., (1994), *Iterative Solution of Large Sparse Systems of Equations*, volume 95 of *Applied Mathematical Sciences*, Springer.

HASELGROVE, C. B., (1961), The solution of nonlinear equations and of differential equations with two-point boundary conditions, *Comput. J.*, 4:255–259.

HEALEY, T. J. and J. A. TREACY, (1991), Exact block diagonalization of large eigenvalue problems for structures with symmetry, *Internat. J. Numer. Methods Engrg.*, 31:265–285.

HEALEY, T. J., (1988a), Global bifurcation and continuation in the presence of symmetry with an application to solid mechanics, *SIAM J. Math. Anal.*, 19:824–840.

HEALEY, T. J., (1988b), A group theoretic approach to computational bifurcation problems with symmetry, *Comput. Meth. Appl. Mech. Eng.*, 67:257–295.

HEALEY, T. J., (1989), Symmetry and equivariance in nonlinear elastostatics. Part I, *Arch. Rational Mech. Anal.*, 105:205–228.

HENDERSON, M. E. and H. B. KELLER, (1990), Complex bifurcation from real paths, *SIAM J. Appl. Math.*, 50:460–482.

HENDERSON, M. E., (1985), *Complex Bifurcation*, PhD thesis, CALTECH, Pasadena.

HESTENES, M. R. and E. STIEFEL, (1952), Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Standards*, 49:409–436.

HETTICH, R. and H. T. JONGEN, (1994), On continuous deformations of semi-infinite optimization problems, Preprint 94-02, Mathematics, University of Trier, Trier, Germany.

HIRSCH, M. W., (1963), A proof of the nonretractibility of a cell onto its boundary, *Proc. Amer. Math. Soc.*, 14:364–365.

HIRSCH, M. W., (1976), *Differential Topology*, Springer Verlag, Berlin, Heidelberg, New York.

HOHMANN, A., (1991), An adaptive continuation method for implicitly defined surfaces, Preprint, Konrad-Zuse-Zentrum Berlin, SC-91-20.

HOLODNIOK, M. and M. KUBÍČEK, (1984), DERPER — An algorithm for the continuation of periodic solutions in ordinary differential equations, *J. Comp. Phys.*, 55:254–267.

HONG, B., (1993), Computational methods for bifurcation problems with symmertries on the manifold, *SIAM J. Num. Anal.*, 30:1134–1154.

HUBBARD, J. and B. WEST, (1992), *A Dynamical Systems Software Package for the Macintosh*, Springer Verlag.

HUBER, B. and B. STURMFELS, (1993), A polyhedral method for solving sparse polynomial systems, Preprint, Cornell Univ.

HUITFELDT, J. and A. RUHE, (1990), A new algorithm for numerical path following applied to an example from hydrodynamical flow, *SIAM J. Sci. Stat. Comput.*, 11:1181–1192.

HUITFELDT, J., (1991), Nonlinear eigenvalue problems — prediction of bifurcation points and branch switching, Technical Report 17, Univ. of Göteborg, Sweden.

IRI, M. and K. KUBOTA, (1987), Methods of fast automatic differentiation and applications, Research Memorandum RMI 87 – 02, Department of Mathematical Engineering and Information Physics, Faculty of Engineering, University of Tokyo.

IZE, J., (1993), Topological bifurcation, in: MATZEU, M. and A. VIGNOLI, editors, *Nonlinear Analysis*, Boston, MA, Birkhäuser.

JACOBOWITZ, H., (1976), Periodic solutions of $x'' + f(x,t) = 0$ via the poincaré-birkhoff theorem, *J. Diff. Eqs.*, 20:37–52.

JANOVSKÝ, V., (1989), A note on computing simple bifurcation points, *Computing*, 43:27–36.

JARAUSCH, H. and W. MACKENS, (1984), Numerical treatment of bifurcation branches by adaptive condensation, in: KÜPPER, T., H. D. MITTELMANN, and H. WEBER, editors, *Numerical Methods for Bifucation Problems*, volume 70 of *ISNM*, Basel, Birkhäuser.

JARAUSCH, H. and W. MACKENS, (1987), Solving large nonlinear systems of equations by an adaptive condensation process, *Numer. Math.*, 50:633–653.

JEPSON, A. D. and D. W. DECKER, (1986), Convergence cones near bifurcation, *SIAM J. Numer. Anal.*, 23:959–975.

JEPSON, A. D., A. SPENCE, and K. A. CLIFFE, (1991), The numerical solution of nonlinear equations having several parameters. Part III: Equations with Z_2-symmetry, *SIAM J. Numer. Anal.*, 28:809–832.

JEZIERSKI, J., (1993), One codimensional Wecken type theorems, to appear in Forum Mathematicum.

JONGEN, H. T. and G.-W. WEBER, (1990), On parametric nonlinear programming, *Ann. Oper. Res.*, 27:253–284.

JONGEN, H. T. and G. ZWIER, (1985), On the local structure of the feasible set in semi-infinite optimization, volume 72 of *ISNM*, pages 185–202, Basel, Birkhäuser Verlag.

JONGEN, H. T., P. JONKER, and F. TWILT, (1983), *Nonlinear Optimization in R^N. I. Morse Theory, Chebyshev Approximation*, Peter Lang Verlag, New York.

JONGEN, H. T., P. JONKER, and F. TWILT, (1986), *Nonlinear Optimization in R^N. II. Transversality, Flows, Parametric Aspects*, Peter Lang Verlag, New York.

JOSEPHY, N. H., (1979), Newton's method for generalized equations, Technical Report 1965, University of Wisconsin, Madison, WI, available from National Technical Information Service under accession number AD A077 096.

KAAS-PETERSEN,, (1989), *PATH — User's Guide*, University of Leeds, England.

KAMIYA, K. and A. J. J. TALMAN, (1990), Variable dimension simplicial algorithm for balanced games, Technical Report 9025, Tilburg Univ., The Netherlands.

KARMARKAR, N. K. and K. G. RAMAKRISHNAN, (1991), Computational results of an interior point algorithm for large scale linear programming, in: ROOS, C. and J.-P. VIAL, editors, *Interior Point Methods for Linear Programming: Theory and Practice*, volume 52 of *Math. Programming, Ser. B*, pages 555–586, Amsterdam, Mathematical Programming Society, North-Holland.

KARMARKAR, N. K., (1984), A new polynomial-time algorithm for linear programming, *Combinatorica*, 4:373–395.

KATZENELSON, J., (1965), An algorithm for solving nonlinear resistive networks, *Bell System Tech. J.*, 44:1605–1620.

KEARFOTT, R. B., (1989), An interval step control for continuation methods, *Math. Comp.*, to appear.

KEARFOTT, R. B., (1990), Interval arithmetic techniques in the computational solution of nonlinear systems of equations: Introduction, examples and comparisons, in: ALLGOWER, E. L. and K. GEORG, editors, *Computational Solution of Nonlinear Systems of Equations*, volume 26 of *Lectures in Applied Mathematics*, pages 337–357, Providence, RI, American Mathematical Society.

KEENER, J. P. and H. B. KELLER, (1974), Perturbed bifurcation theory, *Arch. Rational Mech. Anal.*, 50:159–175.

KELLER, H. B., (1970), Nonlinear bifurcation, *J. Diff. Equ.*, 7:417–434.

KELLER, H. B., (1977), Numerical solution of bifurcation and nonlinear eigenvalue problems, in: RABINOWITZ, P. H., editor, *Applications of Bifurcation Theory*, pages 359–384, New York, London, Academic Press.

KELLER, H. B., (1978), Global homotopies and Newton methods, in: BOOR, C. DE and G. H. GOLUB, editors, *Recent Advances in Numerical Analysis*, pages 73–94, New York, Academic Press.

KELLER, H. B., (1982), Practical procedures in path following near limit points, in: GLOWINSKI, R. and J.-L. LIONS, editors, *Computing Methods in Applied Sciences and Engineering*, Amsterdam, New York, North-Holland.

KELLER, H. B., (1983), The bordering algorithm and path following near singular points of higher nullity, *SIAM J. Sci. Statist. Comput*, 4:573–582.

KELLER, H. B., (1987), *Lectures on Numerical Methods in Bifurcation Problems*, Springer Verlag, Berlin, Heidelberg, New York.

KELLOGG, R. B., T.-Y. LI, and J. A. YORKE, (1976), A constructive proof of the Brouwer fixed point theorem and computational results, *SIAM J. Numer. Anal.*, 13:473–483.

KHACHIYAN, L. G., (1979), A polynomial algorithm in linear programming, *Soviet Math. Dokl.*, 20:191–194.

KHIBNIK, A. I., Y. A. KUZNETSOV, V. V. LEVITIN, and E. V. NIKOLAEV, (1993), Continuation techniques and interactive software for bifurcation analysis of odes and iterated maps, *Physica D*, pages 360–371.

KHIBNIK, A., (1990), *Using TraX, a Tutorial to Accompany TraX, a Program for the Simulation and Analysis of Dynamical Systems*, Exeter Software, Setauket, NY.

KLEIN, F., (1882–1883), Neue Beiträge zur Riemannschen Funktionentheorie, *Math. Ann.*, 21.

KOJIMA, M. and R. HIRABAYASHI, (1984), Sensitivity, stability and parametric analysis, volume 21 of *Mathematical Programming Study*, pages 150–198, Amsterdam, North-Holland.

KOJIMA, M. and Y. YAMAMOTO, (1982), Variable dimension algorithms: Basic theory, interpretation, and extensions of some existing methods, *Math. Programming*, 24:177–215.

KOJIMA, M. and Y. YAMAMOTO, (1984), A unified approach to the implementation of several restart fixed point algorithms and a new variable dimension algorithm, *Math. Programming*, 28:288–328.

KOJIMA, M., H. NISHINO, and T. SEKINE, (1976), An extension of Lemke's method to the piecewise linear complementarity problem, *SIAM J. Appl. Math.*, 31:600–613.

KOJIMA, M., S. MIZUNO, and A. YOSHISE, (1988), A primal-dual interior point algorithm for linear programming, in: MEGIDDO, N., editor, *Progress in Mathematical Programming, Interior Point and Related Methods*, pages 29–47, New York, Springer.

KOJIMA, M., S. MIZUNO, and A. YOSHISE, (1989), A polynomial-time algorithm for a class of linear complementarity problems, *Math. Programming*, 44:1–26.

KOJIMA, M., N. MEGIDDO, and S. MIZUNO, (1990a), A general framework of continuation methods for complementarity problems, Technical report, IBM Almaden Research Center, San Jose, CA.

KOJIMA, M., S. MIZUNO, and T. NOMA, (1990b), Limiting behavior of trajectories generated by a continuation method for monotone complementarity problems, *Math. Oper. Res.*, 15:662–675.

KOJIMA, M., N. MEGIDDO, and S. MIZUNO, (1991a), Theoretical convergence of large-step primal-dual interior point algorithms for linear programming, Preprint, Tokyo Inst. of Techn.

KOJIMA, M., N. MEGIDDO, and T. NOMA, (1991b), Homotopy continuation methods for nonlinear complementarity problems, *Math. Oper. Res.*, 16:754–774.

KOJIMA, M., N. MEGIDDO, T. NOMA, and A. YOSHISE, (1991c), *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, volume 538 of *Lecture Notes in Computer Science*, Springer-Verlag.

KOJIMA, M., S. MIZUNO, and A. YOSHISE, (1991d), An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems, *Math. Programming*, 50:331–342.

KOJIMA, M., N. MEGIDDO, and Y. YE, (1992), An interior point potential reduction algorithm for the linear complementarity problem, *Math. Programming*, 54:267–279.

KOJIMA, M., (1974), Computational methods for solving the nonlinear complementarity problem, *Keio Engrg. Rep.*, 27:1–41.

KOJIMA, M., (1978), Studies on piecewise-linear approximations of piecewise-C^1 mappings in fixed points and complementarity theory, *Math. Oper. Res.*, 3:17–36.

KOJIMA, M., (1979), A complementarity pivoting approach to parametric programming, *Math. Oper. Res.*, 4:464–477.

KOVACH, J. W. and W. D. SEIDER, (1987), Heterogeneous azeotropic distillation — Homotopy-continuation methods, *Comput. chem. Engng.*, 11:593–605.

KRASNOSEL'SKIĬ, M. A., (1964), *Topological Methods in the Theory of Nonlinear Integral Equations*, Pergamon Press, New York.

KUBÍČEK, M., (1976), Algorithm 502. Dependence of solutions of nonlinear systems on a parameter, *ACM Trans. Math. Software*, 2:98–107.

KUHN, H. W., (1968), Simplicial approximation of fixed points, *Proc. Nat. Acad. Sci. U.S.A.*, 61:1238–1242.

KUHN, H. W., (1969), Approximate search for fixed points, in: ZADEK, L. A., L. W. NEUSTAT, and A. V. BALAKRISHNAN, editors, *Computing Methods in Optimization Problems 2*, pages 199–211, New York, Academic Press.

LAGARIAS, J. C., R. C. MELVILLE, and R. GEOGHEGAN, (1993), λ-threading homotopies and DC operating points of nonlinear circuits, Preprint.

LEMKE, C. E. and S. J. GROTZINGER, (1976), On generalizing Shapley's index theory to labelled pseudo manifolds, *Math. Programming*, 10:245–262.

LEMKE, C. E. and J. T. HOWSON, (1964), Equilibrium points of bimatrix games, *SIAM J. Appl. Math.*, 12:413–423.

LEMKE, C. E., (1965), Bimatrix equilibrium points and mathematical programming, *Management Sci.*, 11:681–689.

LEMKE, C. E., (1980), A survey of complementarity theory, in: COTTLE, R. W., F. GIANESSI, and J. L. LIONS, editors, *Variational Inequalities and Complentarity Problems*, London, John Wiley and Sons.

Li, K. and T.-Y. Li, (1993), An algorithm for symmetric tridiagonal eigenproblems — Divide and conquer with homotopy continuation, *SIAM J. Sci. Stat. Comput.*, 14:735–751.

Li, Y. and Z. Lin, (1994), Constructive proof of the Poincaré-Birkhoff theorem, Preprint, Jilin University, China.

Li, T.-Y. and N. H. Rhee, (1989), Homotopy algorithm for symmetric eigenvalue problems, *Numer. Math.*, 55:265–280.

Li, T.-Y. and X. Wang, (1993a), Nonlinear homotopies for solving deficient polynomial systems with parameters, *SIAM J. Numer. Anal.*, 29:1104–1118.

Li, T.-Y. and X. Wang, (1993b), Solving real polynomial systems with real homotopies, *Math. Comp.*, 60:669–680.

Li, T.-Y. and Z. Zeng, (1992), Homotopy-determinant algorithm for solving nonsymmetric eigenvalue problems, *Math. Comp.*, to appear.

Li, T.-Y., T. Sauer, and J. A. Yorke, (1987), Numerical solution of a class of deficient polynomial systems, *SIAM J. Numer. Anal.*, 24:435–451.

Li, T.-Y., T. Sauer, and J. A. Yorke, (1989), The cheater's homotopy: An efficient procedure for solving systems of polynomial equations, *SIAM J. Numer. Anal.*, 26:1241–1251.

Li, T.-Y., H. Zhang, and X.-H. Sun, (1991), Parallel homotopy algorithm for the symmetric tridiagonal eigenvalue problem, *SIAM J. Sci. Stat. Comput.*, 12:469–487.

Li, T.-Y., Z. Zeng, and L. Cong, (1992), Solving eigenvalue problems of real nonsymmetric matrices with real homotopies, *SIAM J. Numer. Anal.*, 29:229–248.

Liu, J. L. and W. C. Rheinboldt, (1991), A posteriori error estimates for parametrized nonlinear equations, in: Wriggers, P. and W. Wagner, editors, *Nonlinear Computational Mechanics*, pages 31–46, Heidelberg, Germany, Springer Verlag.

Lundberg, B. N. and A. B. Poore, (1991), Variable order Adams-Bashforth predictors with an error-stepsize control for continuation methods, *SIAM J. Sci. Stat. Comput.*, 12:695–723.

LUNDBERG, B. N. and A. B. POORE, (1993), Numerical continuation and singularity detection methods for parametric nonlinear programming, *SIAM J. Optim.*, 3:134–154.

LUSTIG, I. J., R. E. MARSTEN, and D. F. SHANNO, (1991), The interaction of algorithms and architectures for interior point methods, Technical Report RRR 36-91, Rutgers Univ., New Brunswick, NJ.

LYNESS, J. N., (1978), Quadrature over a simplex: Part 2. A representation for the error functional, *SIAM J. Numer. Anal.*, 15:870–887.

LYNESS, J. N., (1993), Quadrature over curved surfaces by extrapolation, *Math. Comp.*, to appear.

MACKENS, W., (1989), Numerical differentiation of implicitly defined space curves, *Computing*, 41:237–260.

MARSTEN, R., R. SUBRAMANIAN, M. SALTZMAN, I. J. LUSTIG, and D. F. SHANNO, (1990), Interior point methods for linear programming: Just call Newton, Lagrange and Fiacco and Mccormick!, *Interfaces*, 20:105–116.

MCCORMICK, G. P. and K. RITTER, (1974), Alternate proofs of the convergence properties of the conjugate-gradient method, *J. Optim. Theory Appl.*, 13:497–518.

MCQUAIN, W. D., C. J. RIBBENS, L. T. WATSON, and R. C. MELVILLE, (1993), Preconditioned iterative methods for sparse linear algebra problems arising in circuit simulation, *Comput. Math. Appl.*, to appear.

MCSHANE, K. A., C. L. MONMA, and D. F. SHANNO, (1989), An implementation of a primal-dual interior point method for linear programming, *ORSA J. Comput.*, 1:70–83.

MEGIDDO, N. and M. SHUB, (1989), Boundary behavior of interior point algorithms in linear programming, *Math. Oper. Res.*, 14:97–146.

MEGIDDO, N., (1988), Pathways to the optimal set in linear programming, in: MEGIDDO, N., editor, *Progress in Mathematical Programming, Interior Point and Related Methods*, pages 131–158, New York, Springer.

MEJIA, R., (1986), CONKUB: A conversational path-follower for systems of nonlinear equations, *J. Comp. Phys.*, 63:67–84.

MEJIA, R., (1990), Interactive program for continuation of solutions of large systems of nonlinear equations, in: ALLGOWER, E. L. and K. GEORG, editors, *Computational Solution of Nonlinear Systems of Equations*, volume 26 of *Lectures in Applied Mathematics*, pages 429–449, Providence, RI, American Mathematical Society.

MELHEM, R. G. and W. C. RHEINBOLDT, (1982), A comparison of methods for determining turning points of nonlinear equations, *Computing*, 29:201–226.

MELVILLE, R., S. MOINIAN, P. FELDMANN, and L. T. WATSON, (1993a), Sframe: An efficient system for detailed DC simulation of bipolar analog integrated circuits using continuation methods, *Analog Integrated Circuits Signal Processing*, 3:163–180.

MELVILLE, R. C., L. TRAJKOVIC, S.-C. FANG, and L. T. WATSON, (1993b), Globally convergent homotopy methods for the DC operating point problem, *IEEE Trans. Comput. Aided Design*, to appear.

MENZEL, R. and H. SCHWETLICK, (1978), Zur Lösung parameterabhängiger nichtlinearer Gleichungen mit singulären Jacobi-Matrizen, *Numer. Math.*, 30:65–79.

MENZEL, R. and H. SCHWETLICK, (1985), Parametrization via secant length and application to path following, *Numer. Math.*, 47:401–412.

MERRILL, O., (1972), *Applications and Extensions of an Algorithm that Computes Fixed Points of a Certain Upper Semi-Continuous Point to Set Mapping*, PhD thesis, Univ. of Michigan, Ann Arbor, MI.

MILNOR, J. W., (1968), *Singular Points of Complex Hypersurfaces*, Princeton University Press.

MILNOR, J. W., (1969), *Topology from the Differentiable Viewpoint*, Univ. Press of Virginia, Charlottesville, VA.

MITTELMANN, H. D. and D. ROOSE, editors, (1990), *Continuation Techniques and Bifurcation Problems*, volume 92 of *ISNM*, Birkhäuser.

MIZUNO, S., A. YOSHISE, and T. KIKUCHI, (1989), Practical polynomial time algorithms for linear complementarity problems, *J. Oper. Res. Soc. Japan*, 32:75–92.

MIZUNO, S., M. J. TODD, and Y. YE, (1992), On adaptive-step primal-dual interior-point algorithms for linear programming, *Math. Oper. Res.*, to appear.

MIZUNO, S., (1992), A new polynomial time method for a linear complementarity problem, *Math. Programming*, 56:31–43.

MONTEIRO, R. C. and I. ADLER, (1989), Interior path following primal-dual algorithms, I: Linear programming, II: Convex quadratic programming, *Math. Programming*, 44:27–66.

MOORE, G. and A. SPENCE, (1980), The calculation of turning points of nonlinear equations, *SIAM J. Numer. Anal.*, 17:567–576.

MORGAN, A. P. and A. J. SOMMESE, (1987), A homotopy for solving general polynomial systems that respects m-homogeneous structures, *Appl. Math. Comput.*, 24:101–113.

MORGAN, A. P. and A. J. SOMMESE, (1989), Coefficient parameter polynomial continuation, *Appl. Math. Comput*, 29:123–160.

MORGAN, A. P., A. J. SOMMESE, and L. T. WATSON, (1989), Finding all isolated solutions to polynomial systems using HOMPACK, *ACM Trans. Math. Software*, 15:93–122.

MORGAN, A. P., A. J. SOMMESE, and C. W. WAMPLER, (1991), Computing singular solutions to nonlinear analytic systems, *Numer. Math.*, 58:669–684.

MORGAN, A. P., A. J. SOMMESE, and C. W. WAMPLER, (1992), Computing singular solutions to polynomial systems, *Adv. in Appl. Math.*, 13:305–327.

MORGAN, A. P., A. J. SOMMESE, and C. W. WAMPLER, (1993), A power series method for computing singular solutions to nonlinear analytic systems, *Numer. Math.*, 63:391–409.

MORGAN, A. P., (1986), A transformation to avoid solutions at infinity for polynomial systems, *Appl. Math. Comput.*, 18:77–86.

MORGAN, A. P., (1987), *Solving Polynomial Systems Using Continuation for Engineering and Scientific Problems*, Prentice-Hall, Englewood Cliffs, NJ.

MURTAGH, B. A. and M. A. SAUNDERS, (1987), *MINOS 5.1 Users Guide*, Stanford Univ., CA.

NAZARETH, J. L., (1986), Homotopy techniques in linear programming, *Algorithmica*, 1:529–535.

NAZARETH, J. L., (1991), The homotopy principle and algorithms for linear programming, *SIAM J. Optimization*, 1:316–332.

NAZARETH, J. L., (1994a), A concise introduction to intrior point methods for linear programming, Preprint.

NAZARETH, J. L., (1994b), The implementation of linear programming algorithms based on homotopies, *Algorithmica*, to appear.

NUSSE, H. E. and J. A. YORKE, (1992), Numerical explorations of chaotic systems using DYNAMICS, to appear, Springer.

ORTEGA, J. M. and W. C. RHEINBOLDT, (1970), *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, London.

PARKER, T. and L. CHUA, (1989), *Practical Numerical Algorithms for Chaotic Systems*, Springer Verlag.

PEITGEN, H.-O. and M. PRÜFER, (1979), The leray-schauder continuation method is a constructive element in the numerical study of nonlinear eigenvalue and bifurcation problems, in: PEITGEN, H.-O. and H.-O. WALTHER, editors, *Functional Differential Equations and Approximation of Fixed Points*, volume 730 of *Lecture Notes in Math.*, pages 326–409, Berlin, Heidelberg, New York, Springer Verlag.

PEITGEN, H.-O. and K. SCHMITT, (1983), Global topological perturbations in the study of nonlinear eigenvalue problems, *Math. Methods Appl. Sci.*, 5:376–388.

PEITGEN, H.-O. and H. W. SIEGBERG, (1981), An $\overline{\varepsilon}$-perturbation of Brouwer's definition of degree, in: FADELL, E. and G. FOURNIER, editors, *Fixed point theory*, volume 886 of *Lecture Notes in Math.*, pages 331–366, Berlin, Heidelberg, New York, Springer Verlag.

PEITGEN, H.-O., (1982), Topologische Perturbationen beim globalen numerischen Studium nichtlinearer Eigenwert- und Verzweigungsprobleme, *Jahresber. Deutsch. Math.-Verein.*, 84:107–162.

PERCELL, P., (1980), Note on a global homotopy, *Numer. Funct. Anal. Optim.*, 2:99–106.

POINCARÉ, H., (1881–1886), *Sur les Courbes Definé par une Équation Differentielle. I–IV*, Oeuvres I. Gauthier-Villars, Paris.

Polak, E. and G. Ribière, (1969), Note sur la convergence de méthodes de directions conjugées, *Rev. Française Informat. Recherche Opérationelle*, 3:35–43.

Polak, E., J. E. Higgins, and D. Q. Mayne, (1992), A barrier function method for minimax problems, *Math. Programming*, 54:155–176.

Pönisch, G. and H. Schwetlick, (1981), Computing turning points of curves implicitly defined by nonlinear equations depending on a parameter, *Computing*, 26:107–121.

Pönisch, G. and H. Schwetlick, (1982), Ein überlinear konvergentes Verfahren zur Bestimmung von Rückkehrpunkten implizit definierter Raumkurven, *Numer. Math.*, 38:455–466.

Pönisch, G., (1985), Computing simple bifurcation points using a minimally extended system of nonlinear equations, *Computing*, 35:277–294.

Poore, A. B. and C. A. Tiahrt, (1987), Bifurcation problems in nonlinear parametric programming, *Math. Programming*, 39:189–205.

Poore, A. B. and C. A. Tiahrt, (1990), A bifurcation analysis of the nonlinear parametric programming problem, *Math. Programming*, 47:117–141.

Powell, M. J. D., (1970a), A Fortran subroutine for solving nonlinear algebraic equations, in: Rabinowitz, P., editor, *Numerical Methods for Nonlinear Algebraic Equations*, pages 115–161, Gordon and Breach, New York.

Powell, M. J. D., (1970b), A hybrid method for nonlinear equations, in: Rabinowitz, P., editor, *Numerical Methods for Nonlinear Algebraic Equations*, pages 87–114, Gordon and Breach, New York.

Powell, M. J. D., (1977), Restart procedures for the conjugate gradient method, *Math. Programming*, 12:241–254.

Rabier, P. J. and G. W. Reddien, (1986), Characterization and computation of singular points with maximum rank deficiency, *SIAM J. Numer. Anal.*, 23:1040–1051.

Rabinowitz, P. H., (1971), Some global results for nonlinear eigenvalue problems, *J. Funct. Anal.*, 7:487–513.

Rakowska, J., R. T. Haftka, and L. T. Watson, (1991), An active set algorithm for tracing parametrized optima, *Structural Optimization*, 3:29–44.

Reinoza, A., (1985), Solving generalized equations via homotopies, *Math. Programming*, 31:307–320.

Renegar, J. and M. Shub, (1992), Unified complexity analysis for Newton LP methods, *Math. Programming*, 53:1–16.

Renegar, J., (1985), On the complexity of a piecewise linear algorithm for approximating roots of complex polynomials, *Math. Programming*, 32:301–318.

Renegar, J., (1987), On the efficiency of Newton's method in approximating all zeros of systems of complex polynomials, *Math. Oper. Res.*, 12:121–148.

Renegar, J., (1988a), A polynomial-time algorithm, based on Newton's method, for linear programming, *Math. Programming*, 40:59–93.

Renegar, J., (1988b), Rudiments of an average case complexity theory for piecewise-linear path following algorithms, *Math. Programming*, 40:113–163.

Rheinboldt, W. C. and J. V. Burkardt, (1983a), Algorithm 596: A program for a locally-parametrized continuation process, *ACM Trans. Math. Software*, 9:236–241.

Rheinboldt, W. C. and J. V. Burkardt, (1983b), A locally-parametrized continuation process, *ACM Trans. Math. Software*, 9:215–235.

Rheinboldt, W. C., D. Roose, and R. Seydel, (1990), Aspects of continuation software, in: Roose, D., B. de Dier, and A. Spence, editors, *Continuation and Bifurcations: Numerical Techniques and Applications*, volume 313 of *NATO ASI Series C*, pages 261–268, Dordrecht, Kluwer.

Rheinboldt, W. C., (1978), Numerical methods for a class of finite dimensional bifurcation problems, *SIAM J. Numer. Anal.*, 15:1–11.

Rheinboldt, W. C., (1980), Solution fields of nonlinear equations and continuation methods, *SIAM J. Numer. Anal.*, 17:221–237.

Rheinboldt, W. C., (1981), Numerical analysis of continuation methods for nonlinear structural problems, *Comput. and Structures*, 13:103–113.

Rheinboldt, W. C., (1986), *Numerical Analysis of Parametrized Nonlinear Equations*, J. Wiley, New York.

RHEINBOLDT, W. C., (1987), On a moving-frame algorithm and the triangulation of equilibrium manifolds, in: KÜPPER, T., R. SEYDEL, and H. TROGER, editors, *Bifurcation: Analysis, Algorithms, Applications*, volume 79 of *ISNM*, pages 256–267, Basel, Birkhäuser Verlag.

RHEINBOLDT, W. C., (1988a), On a theorem of S. Smale about Newton's method for analytic mappings, *Applied Math. Letters*, 1:69–72.

RHEINBOLDT, W. C., (1988b), On the computation of multi-dimensional solution manifolds of parametrized equations, *Numer. Math.*, 53:165–182.

RHEINBOLDT, W. C., (1991), On the sensitivity of solutions of parametrized equations, Technical Report ICMA-91-158, Univ. of Pittsburgh.

RHEINBOLDT, W. C., (1992a), On the sensitivity of parametrized equations, *SIAM J. Numer. Anal.*, to appear.

RHEINBOLDT, W. C., (1992b), On the theory and numerics of differential-algebraic equations, in: LIGHT, W., editor, *Advances in Numerical Analysis*, pages 237–175, Oxford, UK, Oxford University Press.

ROBINSON, S. M., (1983), Generalized equations, in: BACHEM, A., M. GRÖTSCHEL, and B. KORTE, editors, *Mathematical Programming — The State of the Art*, pages 346–367, Springer, Berlin.

ROBINSON, S. M., (1987), Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity, *Math. Programming Study*, 30:45–66.

ROOS, C. and J.-P. VIAL, editors, (1991), *Interior Point Methods for Linear Programming: Theory and Practice*, volume 52 of *Math. Programming, Ser. B*, Amsterdam, Mathematical Programming Society, North-Holland.

ROOSE, D., B. DE DIER, and A. SPENCE, editors, (1990), *Continuation and Bifurcations: Numerical Techniques and Applications*, volume 313 of *NATO ASI Series C*, Dordrecht, Kluwer.

SAAD, Y. and M. SCHULTZ, (1986), GMRES: A generalized minimal residual method for solving nonsymmetric linear systems, *SIAM J. Sci. Stat. Comp.*, 7:856–869.

SAIGAL, R. and M. J. TODD, (1978), Efficient acceleration techniques for fixed point algorithms, *SIAM J. Numer. Anal.*, 15:997–1007.

SAIGAL, R., (1971), Lemke's algorithm and a special linear complementarity problem, *Oper. Res.*, 8:201–208.

SAIGAL, R., (1976), Extension of the generalized complementarity problem, *Math. Oper. Res.*, 1:260–266.

SAIGAL, R., (1977), On the convergence rate of algorithms for solving equations that are based on methods of complementary pivoting, *Math. Oper. Res.*, 2:108–124.

SAIGAL, R., (1979), On piecewise linear approximations to smooth mappings, *Math. Oper. Res.*, 4:153–161.

SAIGAL, R., (1984), Computational complexity of a piecewise linear homotopy algorithm, *Math. Programming*, 28:164–173.

SAUPE, D., (1982), On accelerating PL continuation algorithms by predictor-corrector methods, *Math. Programming*, 23:87–110.

SCARF, H. E., (1967), The approximation of fixed points of a continuous mapping, *SIAM J. Appl. Math.*, 15:1328–1343.

SCHWETLICK, H. and J. CLEVE, (1987), Higher order predictors and adaptive stepsize control in path following algorithms, *SIAM J. Numer. Anal.*, 24:1382–1393.

SCHWETLICK, H., (1984a), Algorithms for finite-dimensional turning point problems from viewpoint to relationships with constrained optimization methods, in: KÜPPER, T., H. MITTELMANN, and H. WEBER, editors, *Numerical methods for bifurcation problems*, pages 459–479, Birkhäuser, Basel.

SCHWETLICK, H., (1984b), Effective methods for computing turning points of curves implicitly defined by nonlinear equations, in: WAKULICZ, A., editor, *Computational Mathematics*, volume 13 of *Banach Center Publications*, pages 623–645, PWN (Polish Scientific Publ.), Warsaw.

SEIDER, W. D., D. D. BRENGEL, and S. WIDAGDO, (1991), Nonlinear analysis in process design, *AIChE Journal*, 37:1–38.

SELLAMI, H. and S. M. ROBINSON, (1993), A continuation method for normal maps, Technical Report 93-8, Department of Industrial Engineering, University of Wisconsin-Madison, Madison, WI.

SERRE, J.-P., (1977), *Linear Representations of Finite Groups*, volume 42 of *Graduate Texts in Mathematics*, Springer Verlag, Berlin, Heidelberg, New York.

SEYDEL, R., F. W. SCHNEIDER, T. KÜPPER, and H. TROGER, editors, (1991), *Bifurcation and Chaos: Analysis, Algorithms, Applications*, volume 97 of *ISNM*, Basel, Switzerland, Birkhäuser.

SEYDEL, R., (1979), Numerical computation of branch points of nonlinear equations, *Numer. Math.*, 33:339–352.

SEYDEL, R., (1988), *From Equilibrium to Chaos. Practical Bifurcation and Stability Analysis*, Elsevier, New York.

SEYDEL, R., (1991a), *BIFPACK: A Program Package for Continuation, Bifurcation and Stability Analysis, Version 2.3+*, University of Ulm, Germany.

SEYDEL, R., (1991b), On detecting stationary bifurcations, *Internat. J. Bifurcation and Chaos*, 1:335–337.

SHAMPINE, L. F. and M. K. GORDON, (1975), *Computer Solutions of Ordinary Differential Equations. The Initial Value Problem*, W. H. Freeman and Co., San Francisco.

SHAPLEY, L. S., (1974), A note on the Lemke-Howson algorithm, in: BALINSKI, M. L., editor, *Pivoting and Extensions: In Honor of A. W. Tucker*, volume 1 of *Math. Programming Study*, pages 175–189, New York, North-Holland.

SHROFF, G. M. and H. B. KELLER, (1991), Stabilization of unstable procedures: A hybrid algorithm for continuation, *SIAM J. Numer. Anal.*, to appear.

SHROFF, G. M. and H. B. KELLER, (1993), Stabilization of unstable procedures: The recursive projection method, *SIAM J. Numer. Anal.*, 30:1099–1120.

SHUB, M. and S. SMALE, (1991), Complexity of Bezout's theorem, I, geometric aspects, IBM Research Report.

SIYYAM, H. I., (1994), *Numerical Integration over Implicitly Defined Curves*, PhD thesis, Colorado State University, Fort Collins, CO.

SMALE, S., (1976), A convergent process of price adjustement and global Newton methods, *J. Math. Econom*, 3:1–14].

SMALE, S., (1986), Newton's method estimates from data at one point, in: *The Merging of Disciplines in Pure, Applied and Computational Mathematics*, pages 185–196, New York, Springer.

Sonnevend, G., J. Stoer, and G. Zhao, (1989), On the complexity of following the central path of linear programs by linear extrapolation, *Methods of Operations Research*, 63:19–31.

Sonnevend, G., J. Stoer, and G. Zhao, (1991), On the complexity of following the central path of linear programs by linear extrapolation II, in: Roos, C. and J.-P. Vial, editors, *Interior Point Methods for Linear Programming: Theory and Practice*, volume 52 of *Math. Programming, Ser. B*, pages 527–553, Amsterdam, Mathematical Programming Society, North-Holland.

Sonnevend, G., (1985), An analytical center for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming, volume 84 of *Lecture Notes in Control and Information Sciences*, pages 866–876, New York, Springer.

Stiefel, E. and A. Fässler, (1979), *Gruppentheorethische Methoden und ihre Anwendung*, Teubner, Stuttgart, Fed. Rep. Germany.

Stiefel, E., (1952), Two applications of group characters to the solution of boundary-value problems, *J. Res. Nat. Bureau of Standards*, 6:424–427.

Stoer, J., (1983), Solution of large linear systems of equations by conjugate gradient type methods, in: Bachem, A., M. Grötschel, and B. Korte, editors, *Mathematical Programming: The State of the Art*, pages 540–565, Berlin, Heidelberg, New York, Springer Verlag.

Stroud, A. H., (1971), *Approximate Calculation of Multiple Integrals*, Prentice Hall, Englewood Cliffs, New Jersey.

Talman, A. J. J. and Y. Yamamoto, (1989), A simplicial algorithm for stationary point problems on polytopes, *Math. of Operations Research*, 14:383–399.

Tanabe, K., (1979), Continuous Newton-Raphson method for solving an underdetermined system of nonlinear equations, *Nonlinear Anal.*, 3:495–503.

Tausch, J., (1993), A generalization of the discrete Fourier transformation, in: Allgower, E. L., K. Georg, and R. Miranda, editors, *Exploiting Symmetry in Applied and Numerical Analysis*, volume 29 of *Lectures in Applied Mathematics*, pages 405–412, Providence, RI, American Mathematical Society.

TODD, M. J., (1976a), *The Computation of Fixed Points and Applications*, volume 124 of *Lecture Notes in Economics and Mathematical Systems*, Springer Verlag, Berlin, Heidelberg, New York.

TODD, M. J., (1976b), Extensions of Lemke's algorithm for the linear complementarity problem, *J. Optim. Theory Appl.*, 20:397–416.

TODD, M. J., (1976c), Orientation in complementary pivot algorithms, *Math. Oper. Res.*, 1:54–66.

TODD, M. J., (1978a), Fixed-point algorithms that allow restarting without extra dimension, Technical Report, Cornell University, Ithaca, NY.

TODD, M. J., (1978b), On the Jacobian of a function at a zero computed by a fixed point algorithm, *Math. Oper. Res.*, 3:126–132.

TODD, M. J., (1980), A quadratically convergent fixed point algorithm for economic equilibria and linearly constrained optimization, *Math. Programming*, 18:111–126.

TODD, M. J., (1981), *PLALGO: A FORTRAN implementation of a piecewise-linear homotopy algorithm for solving systems of nonlinear equations*, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY.

TODD, M. J., (1982), On the computational complexity of piecewise-linear homotopy algorithms, *Math. Programming*, 24:216–224.

TODD, M. J., (1986), Polynomial expected behavior of a pivoting algorithm for linear complementarity and linear programming problems, *Math. Programming*, 35:173–192.

TODD, M. J., (1989), Recent developments and new directions in linear programming, in: IRI, N. and K. TANABE, editors, *Mathematical Programming, Recent Developments and Applications*, pages 109–157, London, Kluwer Academic Publishers.

USHIDA, A. and L. O. CHUA, (1984), Tracing solution curves of nonlinear equations with sharp turning points, *Internat. J. Circuit Theory Appl.*, 12:1–21.

VAIDYA, P. M., (1990), An algorithm for linear programming which requires $O(((m+n)n^2+(m+n)^{1.5}n)L)$ arithmetic operations, *Math. Programming*, 47:175–202.

Laan, G. van der and A. J. J. Talman, (1979), A restart algorithm for computing fixed points without an extra dimension, *Math. Programming*, 17:74–84.

Laan, G. van der and A. J. J. Talman, (1980), An improvement of fixed point algorithms by using a good triangulation, *Math. Programming*, 18:274–285.

Waerden, B. L. van der, (1953), *Modern Algebra*, Ungar, New York, Volumes I and II.

Van Loan, C., (1992), *Computational Frameworks for the Fast Fourier Transform*, volume 10 of *Frontiers in Applied Mathematics*, SIAM.

Vanderbauwhede, A., (1982), *Local Bifurcation Theory and Symmetry*, Pitman, London.

Verlinden, P. and R. Cools, (1993), Proof of a conjectured asymptotic expansion for the approximation of surface integrals, *Math. Comp.*, to appear.

Verlinden, P. and A. Haegemans, (1994), Solving very deficient systems of polynomial equations by implicit elimination, Report TW 207, Katholieke Universiteit Leuven, Department of Computing Science, Leuven, Belgium.

Verschelde, J. and R. Cools, (1992), Nonlinear reduction for solving deficient polynomial systems by continuation methods, *Numer. Math.*, 63:263–282.

Verschelde, J. and R. Cools, (1994a), Symbolic homotopy construction, *Applicable Algebra in Engineering, Communication and Computing*, 4:169–183.

Verschelde, J. and R. Cools, (1994b), Symmetric homotopy construction, *J. of Comput. and Appl. Math.*, to appear.

Verschelde, J. and K. Gatermann, (1994), Symmetric newton polytopes for solving sparse polynomial systems, Preprint, Konrad-Zuse-Zentrum Berlin, SC 94-3.

Verschelde, J. and A. Haegemans, (1993), The GBQ-algorithm for constructing start systems of homotopies for polynomial systems, *SIAM J. Numer. Anal.*, 30:583–594.

VERSCHELDE, J. and A. HAEGEMANS, (1994), Homotopies for solving polynomial systems within a bounded domain, in: *Theoretical Computer Science A (Algorithms, Automata, Complexity and Games)*, to appear.

VERSCHELDE, J., P. VERLINDEN, and R. COOLS, (1994), Homotopies exploiting Newton polytopes for solving sparse polynomial systems, *SIAM J. Numer. Anal.*, 31:915–930.

WALKER, H. F., (1990), Newton-like methods for underdetermined systems, in: ALLGOWER, E. L. and K. GEORG, editors, *Computational Solution of Nonlinear Systems of Equations*, volume 26 of *Lectures in Applied Mathematics*, pages 679–700, Providence, RI, American Mathematical Society.

WAMPLER, C. W. and A. P. MORGAN, (1991), Solving the $6R$ inverse position problem using a generic-case solution methodology, *Mech. Mach. Theory*, 26:91–106.

WAMPLER, C. W., A. P. MORGAN, and A. J. SOMMESE, (1990), Numerical continuation methods for solving polynomial systems arising in kinematics, *ASME J. on Design*, 112:59–68.

WAMPLER, C. W., A. P. MORGAN, and A. J. SOMMESE, (1992), Complete solution of the nine-point path synthesis problem for four-bar linkages, *ASME J. on Mechanical Design*, 114:153–159.

WATKINS, D. S., (1984), Isospectral flows, *SIAM Rev.*, 26:379–391.

WATSON, L. T., S. C. BILLUPS, and A. P. MORGAN, (1987), HOMPACK: A suite of codes for globally convergent homotopy algorithms, *ACM Trans. Math. Software*, 13:281–310.

WAYBURN, T. L. and J. D. SEADER, (1987), Homotopy continuation methods for computer-aided process design, *Comput. chem. Engng.*, 11:7–25.

WAYBURN, T. L., (1982), A review of continuation methods and their application to separations problems, *AIChE Comput. and Sys, Tech.*, 11:8–22.

WEBER, H., (1981), On the numerical approximation of secondary bifurcation problems, in: ALLGOWER, E. L., K. GLASHOFF, and H.-O. PEITGEN, editors, *Numerical Solution of Nonlinear Equations*, volume 878 of *Lecture Notes in Math.*, pages 407–425, Springer Verlag, New York.

WERNER, B. and A. SPENCE, (1984), The computation of symmetry-breaking bifurcation points, *SIAM J. Numer. Anal.*, 21:388–399.

WERNER, B., (1992), Test functions for bifurcation points and Hopf points in problems with symmetries, in: ALLGOWER, E. L., K. BÖHMER, and M. GOLUBITSKY, editors, *Bifurcation and Symmetry*, volume 104 of *ISNM*, pages 317–327, Basel, Switzerland, Birkhäuser Verlag.

WIDMANN, R., (1990a), An efficient algorithm for the triangulation of surfaces in $\mathbf{R}^3$, Preprint, Colorado State University.

WIDMANN, R., (1990b), Efficient triangulation of 3-dimensional domains, Preprint, Colorado State University.

WINTERS, K. H. and C. K. A., (1985), The prediction of critical points for thermal explosions in a finite volume, *Combustion and Flame*, 62:13–20.

WINTERS, K. H., K. A. CLIFFE, and C. P. JACKSON, (1987), The prediction of instabilities using bifurcation theory, in: LEWIS, R. W., E. P. HINTON, P. BETTESS, and B. A. SCHREFLER, editors, *Numerical Methods for Transient and Coupled Problems*, pages 179–198, New York, John Wiley.

WINTERS, K. H., M. R. MYERSCOUGH, and P. K. MAINI, (1990), Tracking bifurcating solutions of a model biological pattern generator, *Impact of Computing in Sci. & Eng.*, 2:355–371.

WINTERS, K. H., (1987), A bifurcation study of laminar flow in a curved rectangular cross-section, *J. Fluid Mech.*, 180:343–369.

WINTERS, K. H., (1991), Bifurcation and stability: A computational approach, *Computer Physics Communications*, 65:299–309.

WRIGHT, A. H., (1981), The octahedral algorithm, a new simplicial fixed point algorithm, *Math. Programming*, 21:47–69.

WRIGHT, A. H., (1985), Finding all solutions to a system of polynomial equations, *Math. Comp*, 44:125–133.

WRIGHT, M. H., (1992), Interior methods for constrained optimization, *Acta Numerica*, 1:341–407.

YANG, Z.-H. and H. B. KELLER, (1986), A direct method for computing higher order folds, *SIAM J. Sci. Statist. Comput.*, 7:351–361.

Ye, Y., O. Güler, R. A. Tapia, and Y. Zhang, (1991), A quadratically convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming, *Math. Programming*, 59:151–162.

Index