

RICE UNIVERSITY

A Historical Development of the $(n+1)$ -point Secant Method

by

Joanna Maria Papakonstantinou

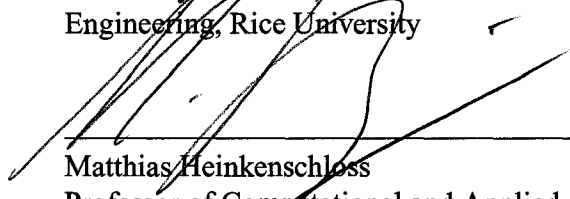
A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Masters of Arts

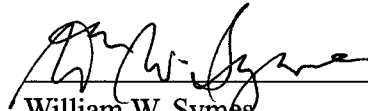
APPROVED, THESIS COMMITTEE:



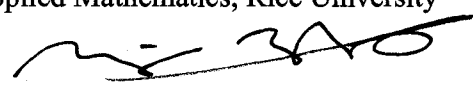
Richard Tapia, Chairman
University Professor, Maxfield-Oshman Chair in
Engineering, Rice University



Matthias Heinkenschloss
Professor of Computational and Applied
Mathematics, Rice University



William W. Symes
Noah Harding Professor of Computational and
Applied Mathematics, Rice University



Yin Zhang
Professor of Computational and Applied
Mathematics, Rice University

HOUSTON, TEXAS

DECEMBER, 2006

UMI Number: 1442093

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 1442093

Copyright 2007 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract
**A Historical Development of the $(n+1)$ -point Secant
Method**

by

Joanna Maria Papakonstantinou

Many finite-dimensional minimization problems and nonlinear equations can be solved using Secant Methods. In this thesis, we present a historical development of the $(n + 1)$ -point Secant Method tracing its evolution back before Newton's Method. Many believe the Secant Method arose out of the finite difference approximation of the derivative in Newton's Method. However, historical evidence reveals that the Secant Method predated Newton's Method by more than 3000 years, and it was most commonly referred to as the Rule of Double False Position.

The history of the Rule of Double False Position spans a period of several centuries and many civilizations. We describe the Rule of Double False Position and compare and contrast the Secant Method in 1-D with the Regula Falsi Method. We delineate the extension of the 1-D Secant Method to higher dimensions using two viewpoints, the linear interpolation idea and Discretized Newton Methods.

Acknowledgements

I would like to thank my advisor, Professor Richard Tapia, for the opportunity to work with him and benefit from his guidance. Professor Tapia always sets high standards for me. He encouraged me to work independently, welcomed my ideas and was patient and understanding during the process. I will always be indebted to him for all this.

I would also like to thank the other members of my committee, Professor Matthias Heinkenschloss, Professor William W. Symes, and Professor Yin Zhang for their continued support. I specially want to thank Mili for her friendship, support and for always being there for me and Layla for being such a caring person and good friend. I would also like to thank Fernando for his friendship. Finally, most of all, I would like to thank my parents, Anne and Taki, for caring so much, for believing in me, and for giving me the strength in difficult times.

Contents

Abstract	ii
Acknowledgements	iii
List of Figures	vii
List of Tables	ix
1 Introduction	1
2 Preliminaries	3
2.1 Two Classes of Problems	3
2.2 Iterative Methods	5
2.2.1 The Secant Equation	6
2.2.2 Convergence	7
3 Historical Development of the Secant Method in 1-D	9
3.1 Newton's Method	10

3.2	The Secant Method in 1-D	11
3.2.1	Geometric Approach	11
3.2.2	Newton's Waste Book	13
3.2.2.1	Whiteside's Annotation	13
3.2.2.2	Ypma's Similar Triangles	14
3.3	Regula Falsi Method	17
3.3.1	Modified Regula Falsi Method	20
3.4	The Rules of False Position	21
3.4.1	Egypt	22
3.4.1.1	Ahmes Papyrus	23
3.4.2	The Rule of Single False Position	25
3.4.3	Babylonia	28
3.4.4	The Rule of Double False Position	29
3.4.5	China	31
3.4.6	Arab Countries	33
3.4.7	Europe	34
3.4.8	India	35
3.4.9	Africa	36
3.4.10	Fibonacci	37
3.5	Evolution of the Name	38
3.6	Use as an Iterative Process	41

3.7	Convergence of the Secant Method in 1-D	42
3.8	Remarks	42
4	Secant Method in Higher Dimensions	44
4.1	General Position	45
4.2	1-D Discretized Newton Method	46
4.3	The Ortega and Rheinboldt General Framework	48
4.3.1	Wolfe's Interpolation Formulation of the $(n+1)$ -point Secant Method	51
4.4	n -D Discretized Newton Method	53
4.4.1	$(n+1)$ -Point Discretized Newton Method	55
4.4.1.1	Sequential $(n+1)$ -point Discretized Newton Method . .	56
4.4.2	Sequential 2-point Discretized Newton Method	56
4.5	Analysis	57
4.5.1	Properties of the $(n+1)$ -point Secant Method	58
4.5.2	Properties of the Sequential 2-point Method	59
4.5.3	Numerical Experiment	60
4.6	Conclusion	61
	Bibliography	63

List of Figures

3.1	The Secant Method	12
3.2	Whiteside's illustration of Newton's Secant Method.	13
3.3	Newton's Secant Method using similar triangles	15
3.4	The Regula Falsi Method	18
3.5	The Modified Regula Falsi Method	20
3.6	Egyptian hieroglyphic notation	22
3.7	Egyptian hieratic notation	22
3.8	Egyptian Rhind Mathematical Papyrus	23
3.9	Papyrus Problem 26 in hieroglyphic notation	24
3.10	Papyrus Problem 26 in hieratic notation	24
3.11	Hieroglyphic symbols for "aha"	25
3.12	Babylonian Clay Tablet	28
3.13	Cuneiform	29
3.14	Nine Chapters on the Mathematical Art	31
3.15	The Rule of Increase and Decrease	34

3.16 The Method of Scales	37
4.1 The Discretized Newton Method	47

List of Tables

3.1	Description of Problem 26	26
3.2	Evolution of the rule name	39
4.1	2-D examples	60

Chapter 1

Introduction

Numerous fundamental science and engineering problems may be modeled as the minimization of a real-valued function of several variables. To solve these minimization problems, we rely on various numerical techniques. A large part of the underlying theory behind the development of such numerical techniques originates from Newton's Method [27]. The current form of Newton's Method [12] is an iterative method to solve nonlinear equations, and we can formulate optimization problems in that framework. However, Newton's Method has several undesirable properties. To circumvent some of these shortcomings, Secant Methods, variations of Newton's Method, are often utilized today in large-scale nonlinear equations and optimization problems. In this thesis, we attempt to give a complete historical development of the $(n + 1)$ -point Secant Method.

This thesis is organized in the following manner. In Chapter 2, we present background information to motivate the interest in studying Secant Methods. We discuss the historical

development of the Secant Method in one dimension (1-D) in Chapter 3. In addition, we describe the Rules of Single and Double False Position and compare and contrast the Secant Method in 1-D with the Regula Falsi Method. In Chapter 4, we present the obvious extension of the 1-D Secant Method to higher dimensions using two viewpoints: the linear interpolation idea and the discretized Newton's method approach. The definition of general position is introduced, and formulations that define the basic $(n + 1)$ -point secant approximation are outlined. We discuss a plausible reason why the extension may fail numerically. Throughout the thesis, a historical timeline of the development of the methods and the roles of the contributors are included.

Chapter 2

Preliminaries

A goal of this thesis is to provide a chronological account of some of the methods (such as Newton's Method, the Secant Method, and the Regula Falsi Method) developed to solve nonlinear equations and unconstrained minimization problems commencing with the Rules of False Position and leading to $(n + 1)$ -point Secant Methods. While there exist numerous methods to solve nonlinear equations and unconstrained minimization problems, we restrict our discussion to those that we believe influenced the development of $(n + 1)$ -point Secant Methods.

2.1 Two Classes of Problems

I: Nonlinear Equations

The first problem of interest consists of finding $x^* \in \mathbb{R}^n$ that solves the system of nonlinear equations $F(x^*) = 0$ where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$. This solution point, x^* , is called a *zero*, or a

root, of the function F .

Recall, that the Jacobian of F at x , for the differentiable function $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, is given by

$$F'(x) = \left[\frac{\partial f_i(x)}{\partial x_j} \right] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}, \quad (2.1)$$

the matrix of partial derivatives of f_i with respect to the x_j s where $F(x) = (f_1(x), f_2(x), \dots, f_n(x))$.

II: Minimization

The second problem of interest is the unconstrained minimization problem denoted

$$\min_x f(x) \quad (2.2)$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is twice differentiable. We call x^* a *local minimizer* of f if there exists an open neighborhood D of x^* such that $f(x^*) \leq f(x) \quad \forall x \in D$.

Recall that the gradient of f at x is given by

$$\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right]^T, \quad (2.3)$$

the vector of first-order partial derivatives of f at x .

The Hessian of f at x is given by

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial^2 x_1} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & & & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 f(x)}{\partial^2 x_n} \end{bmatrix}, \quad (2.4)$$

the matrix of second-order partial derivatives of f at x with respect to x .

The solution to the unconstrained minimization problem (2.2), under mild conditions, satisfies the first-order necessary condition: $\nabla f(x) = 0$. Thus, we may solve the minimization problem (2.2) by viewing it as a nonlinear equation problem where the nonlinear equations are obtained by setting the gradient (2.3) of the real-valued function f equal to zero. Yet, this is not a sufficient condition to yield a minimum, as the gradient vector is also zero at a maximum or at a saddle point.

2.2 Iterative Methods

There are many iterative methods for solving the minimization problem (2.2). These iterative methods proceed from an initial guess x_0 , determine a direction in which to search for a better approximation to a solution, construct a sequence of points x_1, \dots, x_k , and terminate when x_k is sufficiently close to a solution x^* .

The general Quasi-Newton iteration for approximating a solution of $F(x^*) = 0$ for $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the iteration

$$x_{k+1} = x_k - \alpha_k A_k^{-1} F(x_k) \quad (2.5)$$

where x_k represents the k th approximation to the solution, the matrix A_k is viewed as an approximation to $F'(x_k)$, and α_k is called the *step-length parameter*. The choice of α_k is referred to as *step-length control*, and we often call the iteration (2.5) “damped” if $0 < \alpha_k \leq 1$. We refer to $p_k = -A_k^{-1} F(x_k)$ as the *search direction*. The success of these iterative methods depends on the choices of A_k and α_k .

In Newton's Method, A_k is chosen as $F'(x_k)$. In the case where $F(x) = \nabla f(x)$ for twice continuously differentiable function f , we have $F'(x) = \nabla^2 f(x)$, and the Newton iteration for approximating a solution of $\nabla f(x^*) = 0$ for $\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ becomes

$$x_{k+1} = x_k - \alpha_k \nabla^2 f(x_k)^{-1} \nabla f(x_k).$$

We can classify an iterative method based on the number of previous iterates that the method uses at each iteration. For example, if at the k th step, the iteration depends on precisely n of the previous iterates x_{k-n}, \dots, x_{k-1} , we call the method an *n-point method*. That is, an *n-point method* requires n previous iterates. Another way to classify an iterative method is based on whether or not it satisfies the Secant Equation; the topic of the next subsection.

2.2.1 The Secant Equation

We call the equation

$$A_{k+1}(x_{k+1} - x_j) = F(x_{k+1}) - F(x_j) \quad \text{for } j \leq k, \quad (2.6)$$

the *Level- j Secant Equation*. If $j = k$ in (2.6), we obtain the *Level-1 Secant Equation*:

$$A_{k+1}(x_{k+1} - x_k) = F(x_{k+1}) - F(x_k). \quad (2.7)$$

In the literature, this is known simply as the *Secant Equation* and more commonly written as

$$A_{k+1}s_k = y_k \quad (2.8)$$

where $s_k = x_{k+1} - x_k$ and $y_k = F(x_{k+1}) - F(x_k)$.

If $j = k - 1$ in (2.6), we obtain the *Level-2 Secant Equation*:

$$A_{k+1}(x_{k+1} - x_{k-1}) = F(x_{k+1}) - F(x_{k-1}). \quad (2.9)$$

We can continue in this fashion to obtain the *Level- n Secant Equation*:

$$A_{k+1}(x_{k+1} - x_{k-(n-1)}) = F(x_{k+1}) - F(x_{k-(n-1)}), \quad (2.10)$$

where $j = k - (n - 1)$ in the Level- j Secant Equation (2.6).

An iterative method of the form $x_{k+1} = x_k - A_k^{-1}F(x_k)$ is a Level- j Secant Method if it satisfies the Level- j Secant Equation (2.6). Often we do not specify the level because it is not known. As a result, we will abuse terminology and still call a method a Secant Method if it reduces to the Secant Method in 1-D. We should point out that Newton's Method is not a Secant Method.

2.2.2 Convergence

There are many criteria by which we can evaluate an iterative procedure, for example, the length of time taken to calculate a solution, or the amount of computer storage space used in the computation. The convergence rate of an algorithm is a key measure of performance. Broyden states [8], "the rate of convergence of a method is as important as the fact that it converges." We are aware that there could exist a situation when fast convergence occurs so late that it proves to be less beneficial than slower convergence.

In this section, we outline different convergence behaviors. In particular, we emphasize

the notion of superlinear convergence, as many of the methods we discuss in this thesis demonstrate this behavior. If a solution x^* exists, then we define *local convergence* by saying that there exists a neighborhood of x^* , such that for all initial vectors in the neighborhood, the iterates generated by an algorithm are well-defined and converge to x^* . This means that when our initial point x_0 is sufficiently close to x^* then $\lim_{k \rightarrow \infty} x_k = x^*$. *Global convergence* asserts convergence to a solution from any starting point.

Let $\{x_k\}$ be a sequence in \mathbb{R}^n that converges to x^* . We say the convergence is *Q-linear* if there is a constant $M \in (0, 1)$ such that

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \leq M$$

for all k sufficiently large. This indicates that eventually the error, the distance from the solution, decreases at each iteration by at least a constant factor M . We say the convergence is *Q-superlinear* if

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \leq r_k \tag{2.11}$$

holds for some sequence $\{r_k\}$ which converges to zero, i.e., $\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0$. We say the convergence has *Q-order*, or *Q-rate*, of at least p if

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^p} \leq M$$

for all k sufficiently large where M is a positive constant, not necessarily less than 1. In the case that $p = 2$, we use the term *quadratic*.

Chapter 3

Historical Development of the Secant

Method in 1-D

In 1669, Newton formally introduced the first formulation of his technique for solving a polynomial equation in *De analysi per aequationes numero terminorum infinitas* [41]. His original formulation was a purely algebraic procedure as calculus and the notion of the derivative had not yet been invented. Therefore, the original formulation differs in notation from the current iterative form stated explicitly in terms of the derivative. It is interesting that the correction term that Newton computed turned out to be the contemporary correction term which involves the derivative even though the derivative was not known at the time.

Secant Methods can be viewed as using a specific finite difference approximation to the derivative in Newton's Method instead of calculating it explicitly. In the first three sections of this chapter, we describe the methods which we reference throughout the chapter

- Newton's Method, the Secant Method, the Regula Falsi Method, and the Modified Regula Falsi Method. In §3.4, we begin our presentation of the historical development of the Secant Method in 1-D, originally known as the Rule of Double False Position. The evolution of the name Secant Method is outlined in §3.5.

3.1 Newton's Method

In \mathbb{R}^n , Newton's Method is a tool that allows us to approximate the solution of a square (number of equations equals number of variables) nonlinear system of equations by solving a sequence of square linear systems. Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and consider the square nonlinear system of equations

$$F(x) = 0. \tag{3.1}$$

The iteration

$$x_{k+1} = x_k - F'(x_k)^{-1} F(x_k) \tag{3.2}$$

is *Newton's Method*, and x_k represents the k th approximation to the solution of (3.1). Newton's Method is theoretically attractive, but it may be difficult to use in practice for various reasons including the need to calculate the derivative.

3.2 The Secant Method in 1-D

A popular way of obtaining the Secant Method in 1-D is to replace the derivative in the iteration for Newton's Method in 1-D

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad (3.3)$$

with the difference quotient

$$\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} \quad (3.4)$$

which can be viewed as an approximation to $f'(x_k)$. The resulting iteration

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k) \quad (3.5)$$

is the *Secant Method* in 1-D and can also be written as

$$x_{k+1} = \frac{x_{k-1}f(x_k) - x_k f(x_{k-1})}{f(x_k) - f(x_{k-1})}. \quad (3.6)$$

3.2.1 Geometric Approach

The Secant Method is a zero-finding algorithm that uses a succession of zeros of secant lines to better approximate a zero of a function. In Figure 3.1, $f(x)$ represents the nonlinear function whose zero, x^* , we are trying to find. Two initial estimates, x_0 and x_1 , are selected. For future comparisons to methods described in this chapter, it is important to point out that there are no restrictions on the choice of the two initial estimates. The secant line, $l(x)$, connecting the corresponding points $(x_0, f(x_0))$ and $(x_1, f(x_1))$ is used to approximate the nonlinear function $f(x)$. Using the slope of the line passing through $(x_0, f(x_0))$ and

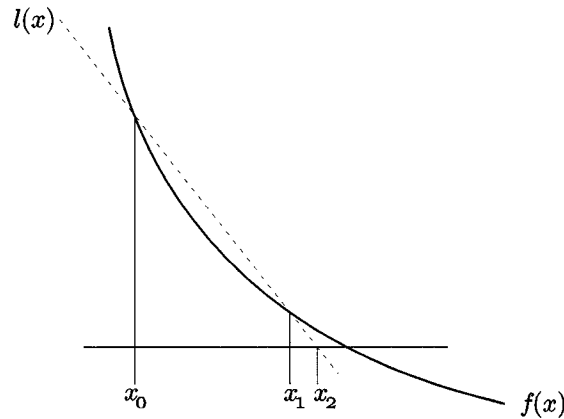


Figure 3.1: The Secant Method with the secant line, $l(x)$, passing through the points $(x_0, f(x_0))$ and $(x_1, f(x_1))$.

$(x_1, f(x_1))$:

$$\frac{f(x_0) - f(x_1)}{x_0 - x_1}, \quad (3.7)$$

we can write the secant line, $l(x)$, in the form:

$$l(x) = \frac{f(x_0) - f(x_1)}{x_0 - x_1}(x - x_1) + f(x_1). \quad (3.8)$$

To obtain an improved approximation to x^* , we solve $l(x) = 0$ for its x -intercept, $x = x_2$, to obtain

$$x_2 = x_1 - f(x_1) \frac{x_0 - x_1}{f(x_0) - f(x_1)}.$$

If we continue this process in the obvious fashion, we obtain the Secant Method in 1-D (3.5), and we can iterate until the approximate zero of the function is obtained to a desired precision.

3.2.2 Newton's Waste Book

Newton kept a notebook of early scientific and mathematical ideas. These unpublished papers, dated early 1665, remained in the possession of the family estate until 1872, when the fifth Earl of Portsmouth decided to donate many of the papers to Cambridge University where they still reside. Whiteside's collection of Newton's unpublished notes, entitled *Newton's Waste Book* [39], includes a zero-finding technique Newton used to solve for a zero of a nonlinear equation that can be identified as the iteration for the Secant Method in 1-D (3.5).

3.2.2.1 Whiteside's Annotation

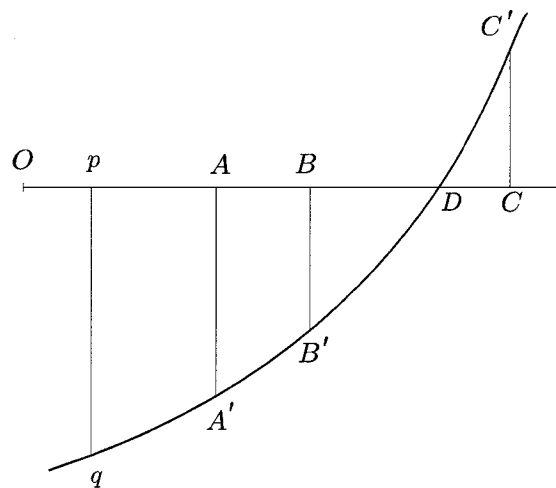


Figure 3.2: Whiteside's interpretation of Newton's Secant Method: A , B , and C are estimates of D . (Adapated from Whiteside [39].)

Figure 3.2 illustrates Whiteside's interpretation of Newton's argument. Given $OA = a$, $OB = b$, $OC = c$, which straddle $OD = X$, Whiteside [39] explained that if A and B are

on the same side of D , as Newton depicted in his original example, then

$$OD \approx OB + (OB - OA) \frac{BB'}{AA' - BB'},$$

which Whiteside represented as

$$X \approx b + (b - a) \frac{-f(b)}{-f(a) + f(b)}. \quad (3.9)$$

Similarly, if B and C lie on opposite sides of D , then

$$OD \approx OC - (OC - OB) \frac{CC'}{CC' + BB'}, \quad (3.10)$$

which Whiteside represented as

$$X \approx c - (c - b) \frac{f(c)}{f(c) - f(b)}. \quad (3.11)$$

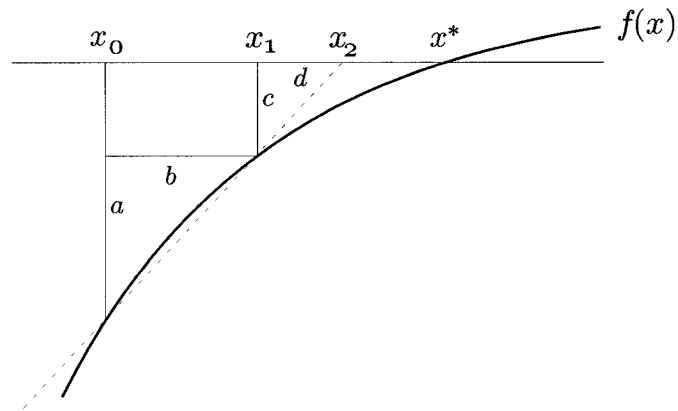
If either (3.9) or (3.11) are continued in an iterative fashion, although Whiteside did not iterate, then each can be represented by the iteration

$$x_{k+1} = x_k - (x_k - x_{k-1}) \frac{f(x_k)}{f(x_k) - f(x_{k-1})}, \quad (3.12)$$

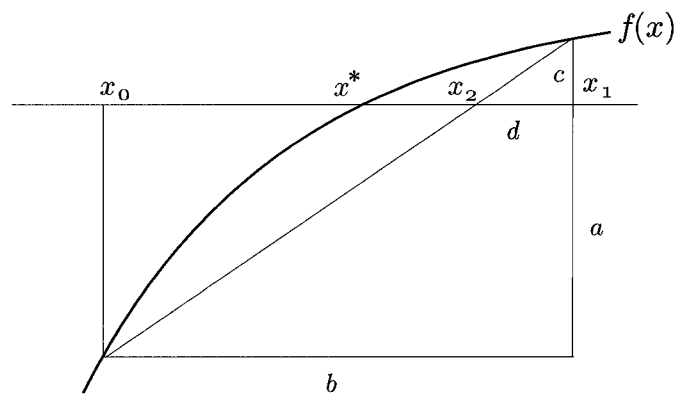
which is the Secant Method in 1-D (3.5).

3.2.2.2 Ypma's Similar Triangles

Ypma [41] constructed the similar triangles, illustrated in Figure 3.3, underlying Newton's geometric approach for approximating a zero, x^* , of a function $f(x)$ based on the phrasing in Newton's text and Whiteside's annotation of it [39]. Newton picked two arbitrary initial



(a) Both initial estimates, x_0 and x_1 , are on the same side of x^* , i.e., $f(x_0)f(x_1) > 0$.



(b) Each initial estimate, x_0 and x_1 , is on an opposite side of x^* , i.e., $f(x_0)f(x_1) < 0$.

Figure 3.3: Ypma's construction of the similar triangles underlying Newton's Secant Method. (Adapted from Ypma [41].)

estimates, x_0 and x_1 , of x^* . From the similarity of the labelled triangles in Figure 3.3,

Ypma used the relationship

$$\frac{a}{b} = \frac{c}{d} \quad (3.13)$$

to obtain

$$\begin{aligned} a &= \begin{cases} f(x_0) - f(x_1) & \text{if } f(x_0)f(x_1) > 0, \\ f(x_0) + f(x_1) & \text{if } f(x_0)f(x_1) < 0 \end{cases} \\ b &= x_1 - x_0 \\ c &= f(x_1) \\ d &= \begin{cases} x_2 - x_1 & \text{if } f(x_0)f(x_1) > 0, \\ x_1 - x_2 & \text{if } f(x_0)f(x_1) < 0. \end{cases} \end{aligned}$$

from which he claimed the proportion (3.13) can be written as

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} = \pm \frac{f(x_1)}{x_2 - x_1} \quad (3.14)$$

where the \pm accounts for the fact that there are no restrictions on the choice of the initial estimates. That is, if both initial estimates, x_0 and x_1 , are on the same side of x^* , then we have $d = x_2 - x_1$ which is of opposite sign than if each initial estimate, x_0 and x_1 , is on an opposite side of x^* . Using (3.14), we can obtain a better approximation, x_2 :

$$x_2 = x_1 - f(x_1) \frac{x_1 - x_0}{f(x_1) - f(x_0)}$$

to x^* . We can continue this process in an obvious fashion, as Ypma did, leading to the iteration

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k), \quad (3.15)$$

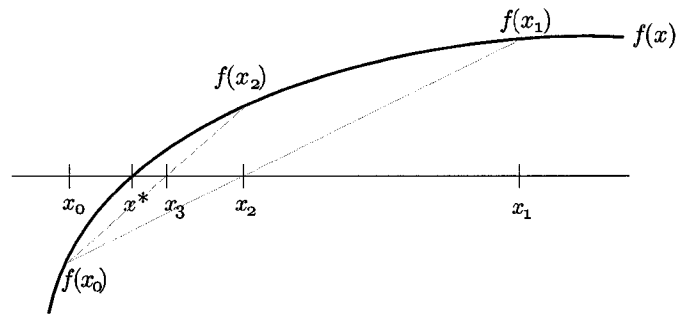
which is the Secant Method in 1-D (3.5).

Although Newton picked both initial estimates to be on the same side of a zero, both Whiteside and Ypma completed the description of Newton's Secant Method allowing for the initial estimates to be on either side of the zero. In addition, Ypma iterated the method. Newton's approach using similar triangles was not the origin of the Secant Method in 1-D. Instead, the Secant Method predated Newton's Method by over 3000 years and evolved from the Rule of Double False Position. We discuss the historical development of the Secant Method in §3.4.4 of this chapter.

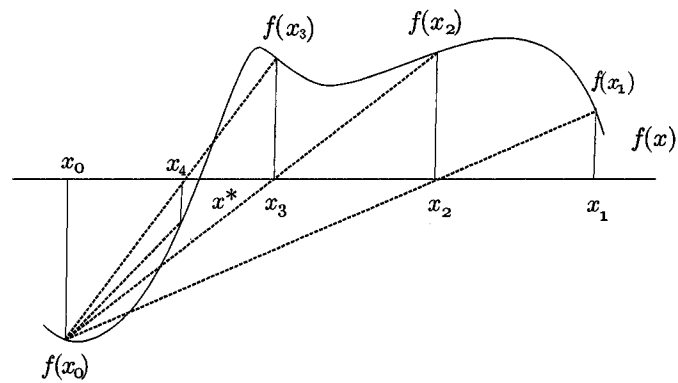
3.3 Regula Falsi Method

The Regula Falsi Method uses linear interpolation of $f(x)$ to approximate a zero, x^* , of $f(x)$. In Figure 3.4, $f(x)$ represents the nonlinear function whose zero we are trying to find. To begin the Regula Falsi Method, two initial estimates, x_0 and x_1 , are chosen such that $f(x_0)$ and $f(x_1)$ are of opposite signs ($f(x_0)f(x_1) < 0$). This restriction that the initial estimates bracket a zero is a key difference between the Regula Falsi Method and the Secant Method. By the Intermediate Value Theorem, there exists a zero in the interval $[x_0, x_1]$ if the function is continuous in that interval. Given x_0 and x_1 , the secant line of $f(x)$ connecting the points $(x_0, f(x_0))$ and $(x_1, f(x_1))$ is constructed. In point-slope form, the secant line can be defined as

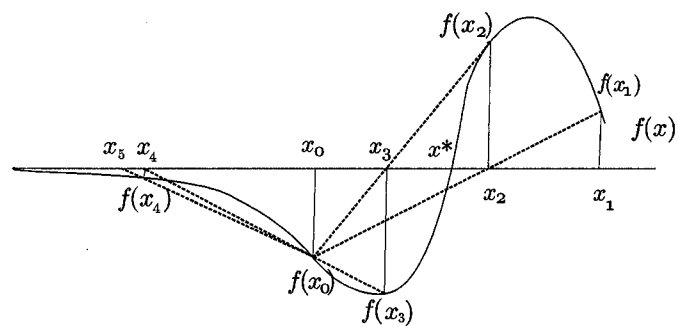
$$y - f(x_1) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_1). \quad (3.16)$$



(a) example where x^* remains bracketed at each step



(b) example where x^* does not remain bracketed at each step



(c) example where x^* does not remain bracketed at each step

and the method fails

Figure 3.4: The Regula Falsi Method: the two initial estimates, x_0 and x_1 , bracket x^* , i.e., $f(x_0)f(x_1) < 0$.

To obtain an improved approximation to x^* we solve for the x -intercept of (3.16), $x = x_2$, to obtain

$$x_2 = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)}. \quad (3.17)$$

The current approximation, x_2 , replaces the previous interval endpoint, either x_0 or x_1 , whose corresponding function value has the same sign as $f(x_2)$, the current best estimate of x^* , while the other interval endpoint is retained. A new secant line of $f(x)$ is constructed and the process is continued in an iterative fashion, always holding the same initial estimate, one endpoint of the original bracketing interval, fixed for all subsequent iterations while the other endpoint is always updated. The iteration

$$x_{k+1} = \frac{\bar{x} f(x_k) - x_k f(\bar{x})}{f(x_k) - f(\bar{x})} \quad (3.18)$$

is the *Regula Falsi Method*, where \bar{x} is an endpoint of the original bracketing interval that remains fixed.

In the examples of Figure 3.4, we begin with the two initial estimates x_0 and x_1 . Since $f(x_2)$ has the same sign as $f(x_1)$, x_2 replaces x_1 while x_0 is retained, and remains fixed throughout subsequent iterations, since $f(x_0)$ has the opposite sign as $f(x_2)$. In the example depicted in Figure 3.4(a), x^* remains bracketed at each step and the approach to x^* is in only one direction. However, in the example depicted in Figure 3.4(b), x^* does not remain bracketed at each step. Instead, x^* jumps in and out of the bracket from step to step. Moreover, in the example depicted in Figure 3.4(c), x^* does not remain bracketed at each step and, in this case, the Regula Falsi Method breaks down. In conclusion, when the Regula Falsi Method is used to approximate a zero, x^* , of a nonlinear function $f(x)$, x^*

does not necessarily remain bracketed at each step for all examples and, in some instances, the method fails.

3.3.1 Modified Regula Falsi Method

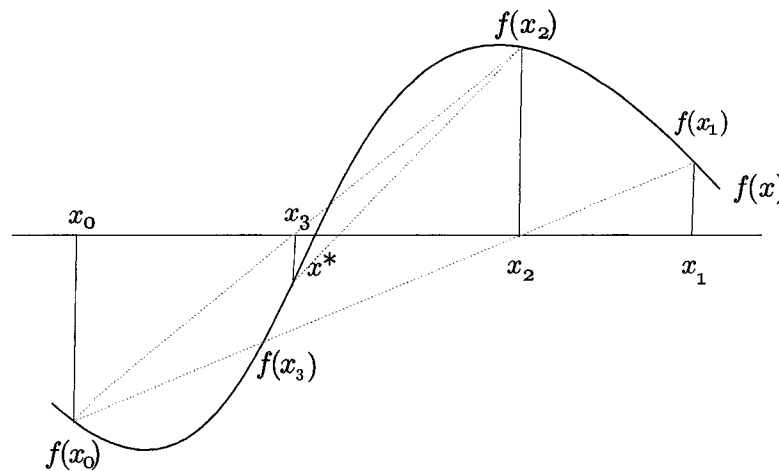


Figure 3.5: The Modified Regula Falsi Method: the two initial estimates, x_0 and x_1 , bracket x^* , and x^* remains bracketed at each step.

There is a modification of the Regula Falsi Method to ensure that x^* remains bracketed at each step. The Modified Regula Falsi Method (illustrated in Figure 3.5) begins with two initial estimates, x_0 and x_1 , chosen such that $f(x_0)$ and $f(x_1)$ are of opposite signs, just as in the Regula Falsi Method. The secant line of $f(x)$ connecting the points $(x_0, f(x_0))$ and $(x_1, f(x_1))$ is constructed and represented in point-slope form (3.16) to obtain an improved approximation, x_2 . Again, the x -value corresponding to the function value that has the same sign as the current best estimate of x^* , $f(x_2)$, is replaced while the x -value corresponding to the function value that has opposite sign as the current function value is retained. The

modification of the Regula Falsi Method occurs at this step. The process is continued in an iterative fashion with the modification that, at each step, the x -value corresponding to the function value that has opposite sign as the current function value is always retained, not just in the first step as in the Regula Falsi Method. Thus, the interval endpoints are changed to ensure that at each step, the new interval contains a zero of $f(x)$. It is worth noting that the Modified Regula Falsi Method and the Regula Falsi Method are identical if the function is concave, or convex, as illustrated in Figure 3.4(a).

There are other modifications of the Regula Falsi Method that circumvent the problem that one endpoint of the original bracketing interval remains fixed throughout the iterations and ensures that each new interval contains a zero. We refer the interested reader to Bronson [7] to learn more about other modifications. Next, we will present the historical development of the Secant Method in 1-D, more commonly referred to as the Rule of Double False Position.

3.4 The Rules of False Position

The history of the Rules of False Position spans a period of several centuries and many civilizations. The Rules of False Position were often presented within the context of a real-life situation. This approach to writing mathematics entirely in word is usually called *rhetorical*, in contrast to the symbolic style we use today. Problems that would be considered trivial to solve today posed a high degree of difficulty in ancient times as algebraic notation was not employed, nor was the notion of an equation known. We attempt to translate these

word problems into algebraic formulations although such usage was unknown at the time of the earliest examples. The earliest evidence of the Rules of False Position dates back to the 18th century B.C. in the Egyptian Rhind Papyrus and Babylonian clay tablets.

3.4.1 Egypt

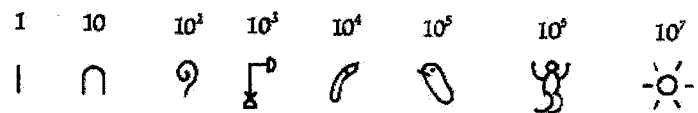


Figure 3.6: Egyptian hieroglyphic notation. (Joseph [22].)

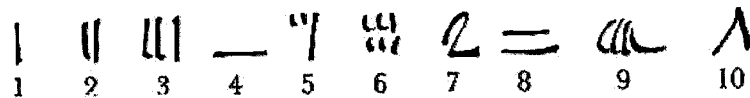


Figure 3.7: Egyptian hieratic notation. (Kline [24].)

The ancient Egyptians wrote with ink on papyrus. The Egyptian notational system consisted of three forms: hieroglyphic, hieratic, demotic [9]. The first two forms: hieroglyphic and hieratic, are from ancient Egypt. The hieroglyphic form was pictorial where each character represented an object. The Egyptians counted by powers of 10, and there was no zero [9]. The special symbols that represented each power of 10 from 1 to 10^7 are depicted in Figure 3.4.1. The hieratic form was symbolic. It replaced frequently used symbols with new symbols and, therefore, was more economical. The numbers 1 to 10 in hieratic form are depicted in Figure 3.4.1. As time passed and writing came into general use in Egypt,

even the hieratic form proved to be too cumbersome. This led to the invention of a type of shorthand, the demotic (popular) notation [9].

3.4.1.1 Ahmes Papyrus

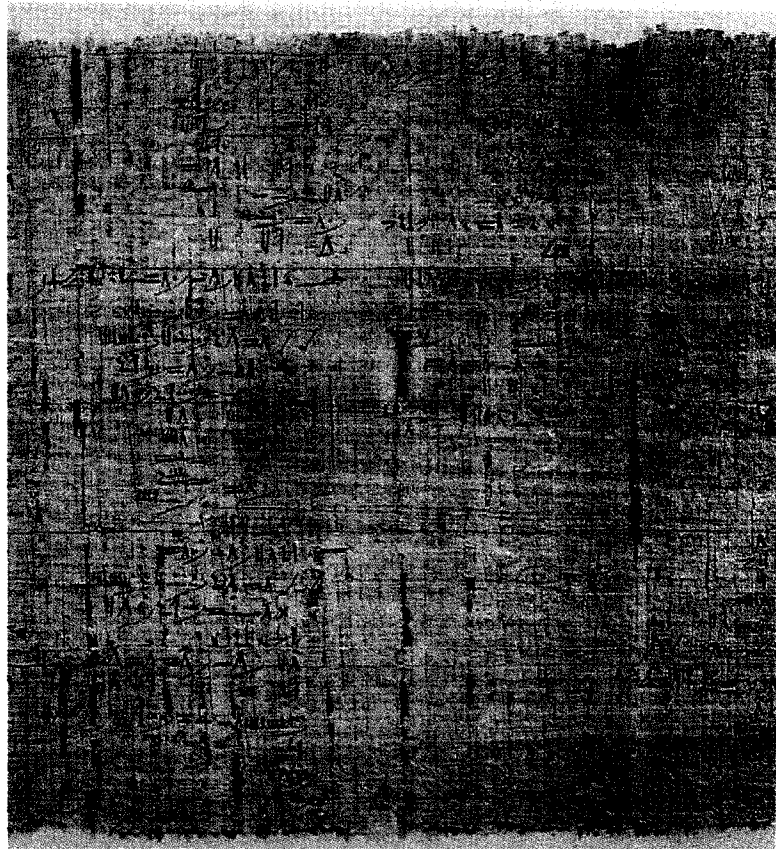


Figure 3.8: Egyptian *Rhind Mathematical Papyrus*: Problems 31-33 written in hieratic form. (Robins and Shute [33], copyright British Museum.)

The main mathematical texts from ancient Egypt date from the 18th century B.C. [11]. Among these, and most important, is the Ahmes Papyrus which was written by the scribe Ahmes in about 1659 B.C. (depicted in Figure 3.8). It is a copy of a document from

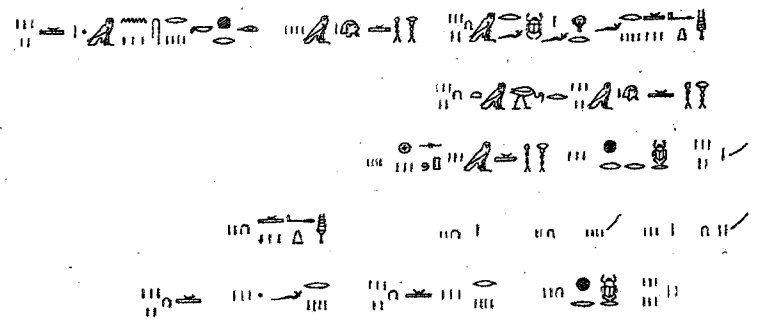


Figure 3.9: Egyptian *Rhind Mathematical Papyrus*: Problem 26 written in hieroglyphic notation. (Robins and Shute [33], copyright British Museum.)

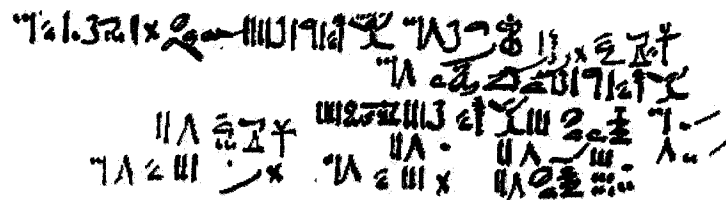


Figure 3.10: Egyptian *Rhind Mathematical Papyrus*: Problem 26 written in hieratic notation. (Robins and Shute [33], copyright British Museum.)

two centuries earlier and is named the *Rhind Mathematical Papyrus* after the archeologist, Henry A. Rhind, who brought it back to England in 1858, where it has resided in the British Museum since 1864 [11]. The *Rhind Mathematical Papyrus* was written using hieratic notation. This two-sided document contains tables to aid in computation on one side and a collection of 87 real-life word problems with solutions on the other. The examples cover a wide range of mathematical ideas that are needed for a scribe to fulfill his duties, and thus, the papyrus was used in the training of scribes. There is no evidence of the use of a general rule or procedure. Working with a general rule was unknown. Instead, each problem uses

specific numbers.

3.4.2 The Rule of Single False Position

Problems 24-34 of the *Rhind Mathematical Papyrus* are examples of problems in one unknown of the first degree. Problem 26 is written in hieroglyphic notation in Figure 3.9, and is written in hieratic notation in Figure 3.10. See Figure 3.9 for a transcription of Problem 26; for a description of the enumerated steps using current algebraic notation see Table 3.1. The technical word for unknown quantity, “aha,” was represented using either of the two hieroglyphic symbols depicted in boxes 1 and 6 of Figure 3.11.



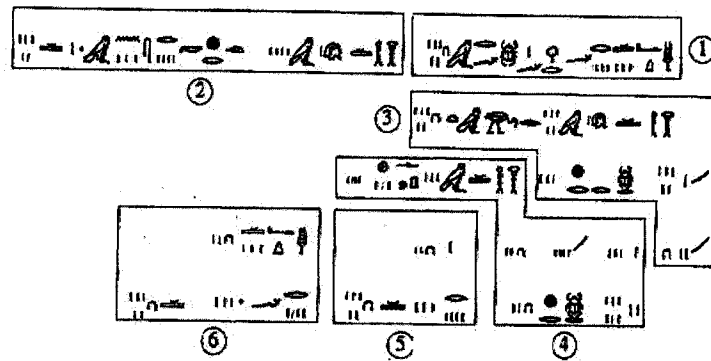
Figure 3.11: Hieroglyphic symbols for “aha” (technical word for unknown quantity). (Resnikoff and Wells [32].)

Problem 26 in the 18th-century B.C. Egyptian *Rhind Mathematical Papyrus* presents solutions to real-life arithmetic problems. Some of these problems can be represented using algebraic notation as finding a number x such that

$$a_1x + \dots + a_nx = c,$$

which we would simplify to

$$ax = c. \tag{3.19}$$



Problem 26 of the *Rhind Mathematical Papyrus* using hieroglyphic notation.
(Chabert [11].)

Transcription of hieroglyphics	Description using algebraic notation
1. A quantity, $\frac{1}{4}$ of it added to it, becomes 15	$x + \frac{1}{4}x = 15$
2. Operate on 4; make thou $\frac{1}{4}$ of them, namely 1; The total is 5.	Guess $x = 4$: $4 + 1 = 5$.
3. Operate on 5 for the finding of 15 $\begin{array}{r} \backslash 1 \quad 5 \\ \backslash 2 \quad 10 \end{array}$ There becomes 3.	Divide: $\frac{15}{5} = 3$
4. Multiply: 3 times 4. $\begin{array}{r} 1 \quad 3 \\ 2 \quad 6, \\ \backslash 4 \quad 12 \end{array}$ There becomes 12.	Multiply wrong answer ($x = 4$) by 3: $3 \times 4 = 12$.
5. $\begin{array}{r} 1 \quad 12, \\ \frac{1}{4} \quad 3 \\ \hline \text{Total } 15 \end{array}$	$12 + \frac{1}{4}(12) = 15$
6. The quantity is 12. $\frac{1}{4}$ of it is 3; the total is 15.	Thus, $x = 12$.

Table 3.1: Description of Problem 26 of the *Rhind Mathematical Papyrus*.

We can solve the equation:

$$x + \frac{1}{4}x = 15, \quad (3.20)$$

and deduce that $x = 12$. From a current mathematical point of view, this problem (3.20) is simple to solve if we sum the a_i s ($1 + \frac{1}{4}$) and solve for x using $x = \frac{c}{a}$ with $c = 15$, and $a = \frac{5}{4}$. However, the people of the time did not perform algebraic simplifications; for example, summing the a_i s. Moreover, since a in Equation (3.19) was not necessarily an integer in real-life problems, a method that avoided the possibility of dividing by a fraction was used instead of merely solving for x using $x = \frac{c}{a}$. Although the Egyptians could perform division, they preferred to avoid it as more difficult [11].

At the time, the first step of the method to solve for x in Equation (3.19) was to choose a (probably wrong, yet, not so arbitrary) estimate of x , to obtain a corresponding (probably wrong) value of c . If we suppose the estimate of the solution is $x = x_0$, then we get c_0 where

$$ax_0 = c_0 \neq c.$$

In our example (3.20), let the estimate of x be 4. If we carry out the calculations using $x_0 = 4$, we get $c_0 = 5$ (not 15). So if

$$ax_0 = c_0,$$

then, we can write the proportion

$$\frac{x}{x_0} = \frac{c}{c_0}.$$

Next, we multiply c by $\frac{x_0}{c_0}$ to get the solution

$$x = \frac{(c)(x_0)}{c_0}.$$

This method was called *Simple False Position* [11], *Process of Supposition*, or most commonly, the *Rule of Single False Position* [23] and avoids the need to explicitly determine a in the reduced equation (3.19).

3.4.3 Babylonia

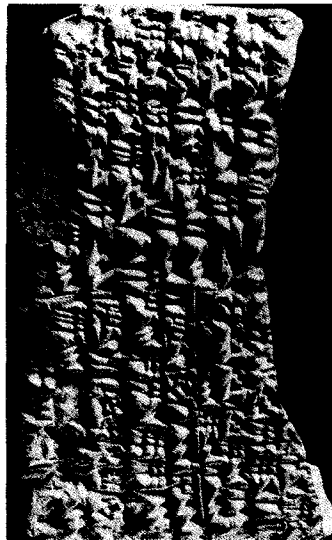


Figure 3.12: Babylonian Clay Tablet. (Resnikoff and Wells [32].)

The Babylonians existed at about the same time as the Egyptians, but there seems to be little evidence that either influenced the other's mathematics. The Babylonians wrote on clay tablets of various sizes (depicted in Figure 3.12) using cuneiform which utilized

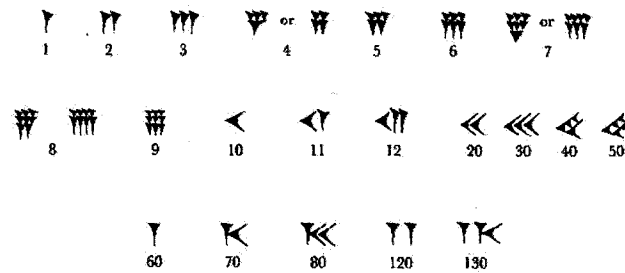


Figure 3.13: Cuneiform (Kline [24].)

wedge-shaped symbols formed using a wooden stylus with a triangular end. Their sexagesimal number system (represented in Figure 3.13) enabled the Babylonians to calculate with fractions as readily as with integers, impossible for the Egyptians. Of the estimated half a million inscribed clay tablets, 400 have been identified as having mathematical content [9]. Both the Egyptian *Rhind Mathematical Papyrus* and the Babylonian clay tablets contained problems that were solved using the Rule of Double False Position.

3.4.4 The Rule of Double False Position

Since the Babylonians and the Egyptians already had the Rule of Single False Position to solve for x of $ax = c$, they quite naturally tried to apply it to other real-life word problems which we would represent using algebraic notation as finding a number x such that

$$ax + b = c, \quad (3.21)$$

when $b \neq 0$. At the time, people did not know how to move terms from one side of an equation to the other [23]. Methods to solve the linear equation (3.21) began by choosing two

arbitrary (probably wrong) estimates of x (that is, no restrictions on the initial estimates) to get two different corresponding (probably wrong) values of c . If we suppose the first estimate of the solution is $x = x_0$, then we get c_0 which satisfies

$$ax_0 + b = c + c_0. \quad (3.22)$$

If we suppose the second estimate of the solution is $x = x_1$, we get c_1 which satisfies

$$ax_1 + b = c + c_1. \quad (3.23)$$

Subtract (3.23) from (3.22) and solve for a to obtain

$$a = \frac{c_0 - c_1}{x_0 - x_1}, \quad (3.24)$$

the ratio of the difference of the incorrect c -values to the difference of the initial estimates.

Add (3.23) and (3.22), use (3.24), and solve for $c - b$ to obtain

$$c - b = \frac{x_1 c_0 - x_0 c_1}{x_0 - x_1}. \quad (3.25)$$

Finally, from (3.21) we have

$$x = \frac{c - b}{a},$$

and following the substitution of a using (3.24), and the substitution of $c - b$ using (3.25),

we get the solution

$$x = \frac{x_1 c_0 - x_0 c_1}{c_0 - c_1}. \quad (3.26)$$

This method for solving for x was called the *Rule of Double False Position* [23]. If for the two arbitrary initial estimates, x_0 and x_1 , we write $c_0 = f(x_0)$ and $c_1 = f(x_1)$, where f is

given by (3.21) in the form $f(x) = ax + b - c$, then we obtain the solution

$$x = \frac{f(x_1)x_0 - f(x_0)x_1}{f(x_1) - f(x_0)} \quad (3.27)$$

which is the Secant Method in 1-D (3.6) for linear equations and not the Regula Falsi Method. The Rule of Double False Position and the Secant Method in 1-D are equivalent since it is arbitrary which estimate we call x_0 and which we call x_1 . Moreover, at the time, linear equations were being solved, thus, the solutions were attained in one step. Only in the case where $f(x_0)f(x_1) < 0$, i.e., the initial estimates bracket a solution, is the Rule of Double False Position the Regula Falsi Method, and only for linear equations.

3.4.5 China

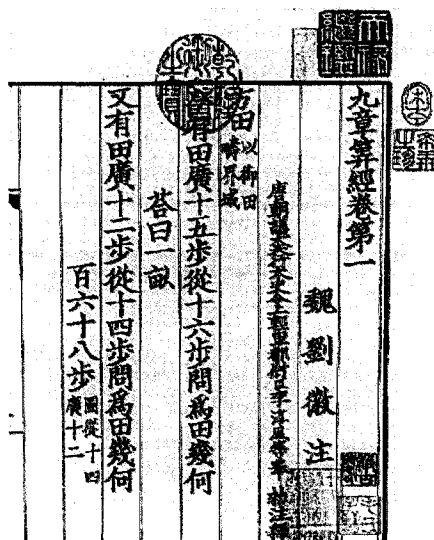


Figure 3.14: Beginning of Chapter 1 of the Nine Chapters on the Mathematical Art. (<http://www.math.sfu.ca/histmath/China/1stCenturyAD/NineChapIntro.html>.)

The earliest surviving Chinese mathematics text, *Jiǔ Zhāng Suàn Shù* (Computational Prescriptions in Nine Chapters) [11], a.k.a. *The Nine Chapters on the Mathematical Art* [23], dates back to the Hàn Dynasty around 200 B.C. and represents the collective efforts of many over several centuries. It contains a total of 246 problems in nine chapters with each chapter containing practical problems connected with everyday life, their solutions, and brief descriptions of the methods used to solve them. In 263 A.D., Liú Huī provided theoretical verifications of each of the problems from the fragments of the collection that were recovered [25]. *The Nine Chapters on the Mathematical Art*, as it survives today, is Liu's commentary. The purpose of *The Nine Chapters on the Mathematical Art* (the beginning of Chapter 1 depicted in Figure 3.14) was similar to that of the Egyptian *Rhind Mathematical Papyrus*. It served as a practical handbook with problems that the ruling officials of the state were likely to encounter [9].

The first millenium Chinese Rule of Double False Position was *yíng bù zú shu* and literally means 'too much and not enough' [11]. The initial estimate of the x -value, also referred to as the false position (or in Chinese, *jia she*, where *jia* = false and *she* = supposition), usually leads to a corresponding c -value that is too much (*yíng*) or not enough (*bù zú*) compared to the given c -value. In Chapter 7: "Excess and Deficit," twenty problems were solved using the Rule of Double False Position that we just described (later referred to in the West as the Regula Falsi Method but in fact was the Secant Method for linear equations). In fact, Problem 19 represented a quadratic equation that was solved using the Rule of Double False Position [25]. However, we should point out that the Chinese did

not use the rule in an iterative manner. Instead, they only performed one step, and thus, attained an approximation to the solution.

3.4.6 Arab Countries

Although Arab mathematicians used the Rule of Single False Position, they primarily used the Rule of Double False Position. In the 9th century, Abu Jafar Mohammad ibn-Mūsa al-Khwārizmī wrote two influential books which were translated into Latin in the 12th century and circulated throughout Europe [4]. His first book, *Hisab al-jabr w'al-muqābala*, 'Calculation by Restoration and Reduction,' on arithmetic was written approximately 820 A.D., and included a discussion of algebra of 1st- and 2nd-order equations. It was written only using words and with no algebraic symbolism. It was translated into Latin *Liber Algebrae et Almucabola*, by Robert of Chester in 1145, and then was shortened to al-jabr [9]. The word 'algebra' is the European corruption of the Arabic word al-jabr [9], [4].

The original Arabic version of Al-Khwārizmī's second book, *Algorithmi de numero indorum*, 'Calculation within Indian Numerals,' on solving equations did not survive. Only John of Seville's Latin translation from the beginning of the 12th century still exists [9]. The word 'algorithm' is a variant of 'algorism' which referred to a recipe for doing mathematics [4].

In the 9th century, Abū Kāmil wrote *Kitāb fil-jabr w'al muqābalah*, 'Book of Algebra.' It was a commentary on, and elaboration of, Al-Khwārizmī's work and was entirely devoted to *Hitāb al Khāta'ayn*, 'rule of the two errors,' which was the Rule of Double False Position

[11]. Abū Kāmil's book was much more extensive than Al-Khwārizmī's, containing a total of 69 problems compared to the 40 of his predecessor [9]. Abū Kāmil included many of the problems that Al-Khwārizmī had explained. Again, these Arabic books were entirely rhetorical, with all computations described in words [9].

3.4.7 Europe

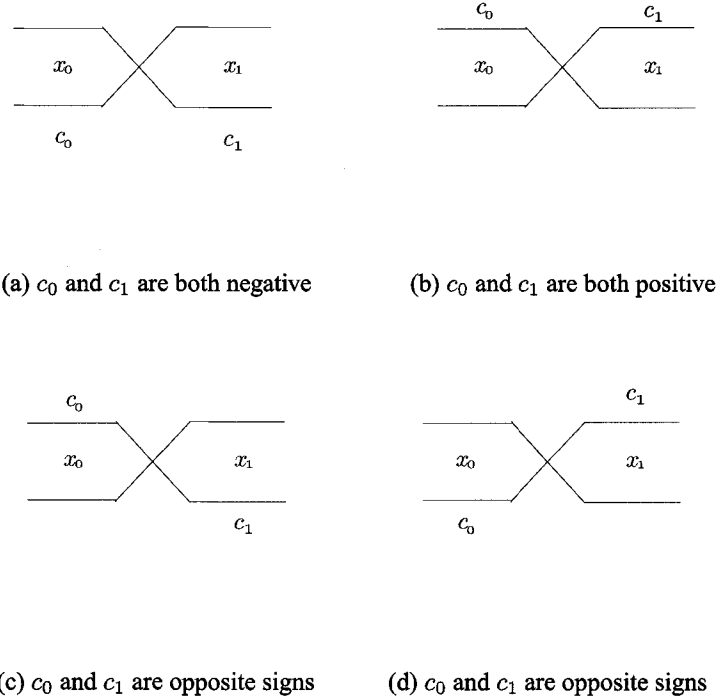


Figure 3.15: The Rule of Increase and Decrease: x_0 and x_1 are the initial estimates of x and c_0 and c_1 are the corresponding values of c . (Adapted from Lun [23].)

In 11th-century Europe, an anonymous Latin Book, *Liber Augmenti et Diminutionis*, 'Book of Increase and Decrease,' translated from Arabic presented the rule Hisab al

Khāta'ayn, (in Latin, 'Regula Augmenti et Diminutionis,' and in English, 'Rule of Increase and Decrease') to solve linear equations. This method was the rule found in "Excess and Deficit" of *Nine Chapters* [23], only presented differently. For example, given the two suppositions if $x = x_0$, then $ax_0 + b = c + c_0$, and if $x = x_1$, then $ax_1 + b = c + c_1$, then the two numbers x_0 and x_1 and the corresponding c_1 and c_2 were depicted as in Figure 3.15, using parallel lines with a cross, similar to the Method of Scales diagram illustrated in Figure 3.16.

The two initial estimates were placed in the middle section. If c_0 and c_1 were both of the same sign, then they were on the same level, either both below the middle section (illustrated in Figure 3.15(a)) or both above the middle section (illustrated in Figure 3.15(b)). If c_0 and c_1 were of different signs, then they were on different levels (one above and one below the middle section) as illustrated in Figures 3.15(c) and 3.15(d). Almost every problem in *Liber Augmenti et Diminutionis* is solved by the Rule of Double False Position.

3.4.8 India

Indian mathematicians as early as the 7th century, had worked with fractions and had been using names of colors to represent multiple unknowns in an expression [11]. Bhāskara was the leading Indian mathematician of the 12th century. In 1150, Bhāskara wrote his most reknown book, *Siddhānta Siromani*, 'Head Jewel of an Astronomical System.' The book was comprised of four parts, and its contents became known to western Europe through its Arabic translation in 1587 [9]. The first two parts, *Līlāvati*, 'The Beautiful,' named

after his daughter, dealt with arithmetic and many of the problems were posed in the form of questions addressed to her. The second two parts, *Bijaganita*, ‘Root Extraction,’ were entirely devoted to algebra. The first statement of the Rule of Single False Position, which Bhāskara called *ista karma*, ‘operating with a trial number,’ was stated explicitly in the third chapter of *Bijaganita*. Unlike the texts that previously contained mathematics, problems were posed simply for pleasure instead of for utilitarian function.

In contrast, the only evidence we have of the use of the Rule of Double False Position in India comes from a *Liber Augmenti et Diminutionis*. According to Chabert [11], Youschkevitch believes that the similarity of the expression ‘augmentation and diminution’ with the Chinese expression ‘too much and not enough’ implies that the introduction of the Rule of Double False Position into Arab literature came from the Chinese via a detour through India [11].

3.4.9 Africa

In the 13th century, Abu al-Abbas Ahmad ibn Muhammad, known as Ibn al-Bannā, in his treatise *Talkhis a'mal al-hisab*, ‘Summary of Operations of Computation,’ presented the *Method of Scales* as a technique for carrying out the Rule of Double False Position [11]. Since two false values (initial estimates) for x are considered, being either an excess or a deficit, the model of balance seems quite appropriate. As illustrated in Figure 3.16, the false positions (initial estimates) are placed on the two pans, the given answer in the middle (on the dome), and the corresponding false answers above if they are by excess, or below if

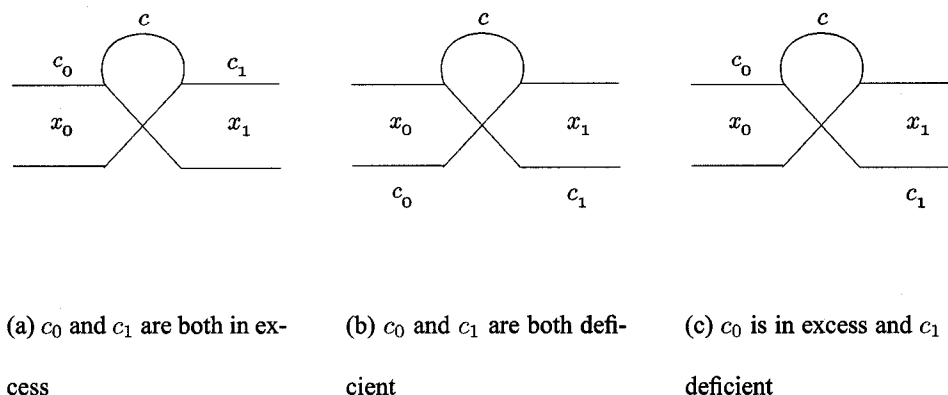


Figure 3.16: The Method of Scales: x_0 and x_1 are the initial estimates of x and c_0 and c_1 are the corresponding values of c . (Adapted from Chabert [11].)

they are by deficiency. In the Middle Ages, some authors adopted the diagrammatic form but without the dome [11] as illustrated in Figure 3.15. In that case, only the numbers needed for the calculation are recorded.

3.4.10 Fibonacci

Leonardo Pisano (Fibonacci) was the son of the representative of the Republic of Pisa at Bougie on the North African coast. Fibonacci joined his father there in 1192 to be initiated into the practice of business, and therefore, he received a solid mathematical foundation [11]. After his studies, he travelled extensively about the Mediterranean, visiting Egypt, Syria, Greece, Sicily, and Provence. On his return to Pisa, in 1202, he composed a book, *Liber Abaci*, 'Book of the Abacus,' later revised in 1228, containing fifteen chapters dealing with arithmetic and algebra including a mixture of Indian arithmetic methods and Arab al-

gebraic methods. Fibonacci wrote problems entirely in words and reproduced 29 problems, with little or no change, from the Arabic ‘Book of Algebra’ [9]. *Liber Abaci* circulated in manuscript form until it was printed in Italy in 1857, and was not translated into English until 2002 [9].

In Chapter 13 of *Liber Abaci*, Fibonacci described Elchataieym, the Arabic rule for solving a linear equation [23], [31]. Fibonacci concluded that the choice of the starting values in the Rule of Double False Position can be arbitrary; sometimes both starting values are smaller than the correct value, sometimes both are larger, or sometimes one may be larger while the other is smaller. This means that the starting values need not bracket the solution. Hence, what Fibonacci was really describing was the Secant Method.

3.5 Evolution of the Name

Throughout the centuries, the Secant Method has been referred to by different names. Even though the origin of the method dates back to the 18th century B.C., it was not thought of as a rule or a method at that time, and therefore, it was not given a specific name. It was not until 200 B.C., in China, that it was considered a rule and given a name. Since then, various civilizations had different names for this same rule (listed in Table 3.2). It has been most commonly referred to as the Rule of Double False Position since the 11th century A.D. In the 13th and 14th centuries, European mathematics was still almost all rhetorical, with occasional abbreviations from time to time [4]. In the 15th century, Al-Kashi, in §2,

Country	Century	Rule Name
Egypt	18th B.C.	-
Babylonia	18th B.C.	-
China	2nd B.C.	yíng bù zú (too much and not enough)
Arab	9th A.D.	hisab al-Khataayn (rule of two errors)
Europe	11th A.D.	elchataym (two errors)
India	12th A.D.	ista karma (operation with a trial number)
Africa	13th A.D.	method of scales
Europe	16th A.D.	rule of double false position/regula falsi
America	20th A.D.	secant method

Table 3.2: Evolution of the different names for the Secant Method

Chapter 5 of his *Key of Arithmetic* [1], used words to explain in detail the same rule as in the 11th-century *Book of Increase and Decrease*. He solved a problem which was an application of the Chinese rule of too much or not enough (Rule of Double False Position). In the late 15th century, some mathematicians started to use symbolic expressions in their work.

In the 16th century, Latin names and terms were introduced to describe existing methods. Peter Bienewitz used the term Regula Falsi, which translates to “rule of falseness,” to describe the Rule of Double False Position. He defined the Regula Falsi Method as an iterative method that “learns to produce truth from two lies” [26]. Bienewitz explained that the term ‘false’ is used because the solution is produced from two ‘false’ initial estimates and not because the method is wrong or false [2]. We believe this explains how the Rule of Double False Position acquired the name Regula Falsi Method. From this point on, the

Rule of Double False Position was referred to as both the Rule of Double False Position and Regula Falsi. However, we wish to reiterate that the Rule of Double False Position is the Secant Method for linear equations and not the Regula Falsi Method as we defined it in §3.3. Therefore, the term Regula Falsi Method came before the actual Regula Falsi Method as we defined it, and was used to describe the Rule of Double False Position.

In 1955, Booth [6] seems to be the first to describe the Regula Falsi Method as we defined it and call it the Regula Falsi Method. Hence, Booth used a pre-existing term to describe a new method. However, he stated that the method was also known as the Rule of Double False Position. This marked the start of the confusion in the use of the names Rule of Double False Position and Regula Falsi in the mathematical texts of the following decade.

Thomas Fincke [14] introduced the word “secant” in his 1583 treatise on geometry [4]. In 1958, Jeeves [21] seems to be the first to refer to the Rule of Double False Position as the Secant Method. However, subsequent mathematical texts did not perpetuate the use of this name. In 1960, Ostrowski [29] seems to be the first to distinguish between the Regula Falsi Method, as we defined it, and the Secant Method. He explained that the Secant Method uses the two last points instead of constantly using one of the initial points. However, he referred to the Secant Method as “iteration with successive adjacent points.” The confusion of names for describing the Regula Falsi Method and the Secant Method continued throughout the 1960s. For example, in 1964, Henrici [19] used the term Regula Falsi to describe the Secant Method. Two years later, in 1966, Isaacson and Keller [20]

used the term Method of False Position to describe the Secant Method and they used the term ‘Classical Regula Falsi Method’ to describe the Regula Falsi Method as we defined it. Later, in 1974, Dahlquist and Björck [13] distinguished between the Modified Regula Falsi Method (which they refer to as Regula Falsi) and the Secant Method. We should note that this 1974 English version is the translation and extension of the 1969 Swedish edition. We do not witness a consistent usage of names corresponding to our descriptions of the Regula Falsi Method and the Secant Method until the Gill, Murray and Wright’s 1981 book [17].

3.6 Use as an Iterative Process

Recall that the Rule of Double False Position (Secant Method) was originally defined for linear equations, with the first evidence of its use to solve an equation which was not linear but quadratic in China in 200 B.C. Therefore, it is surprising, based on the historical evidence provided, that in the 19th century, the Rule of Double False Position was still taught in schools to solve linear equations. Maas [26] claims it was not until the beginning of the 20th century that the literature recommended using the Rule of Double False Position to approximate solutions of nonlinear equations, including choosing the two initial estimates to bracket the solution and iterating the method. However, Cardano, in his 1545 *Artis Magnae*, described the Rule of Double False Position, calling it Regula Falsi, as an iterative process when he said multiple steps must be performed in order to improve the approximation [10]. He gave a complete explanation, with elaborate geometric proofs, of how to solve cubic equations [4].

3.7 Convergence of the Secant Method in 1-D

In 1954, Bachmann [3] claimed without proof that the Secant Method, which he referred to as the Regula Falsi Method, has golden mean convergence. The first proof we can locate of the golden mean convergence of the Secant Method in 1-D is by T. A. Jeeves [21] in 1958. He proved that at each iteration of the 1-D Secant Method, the increase in the number of significant digits is $\frac{1}{2}(1+\sqrt{5}) \approx 1.62$ (the golden mean) times the previous increase. Jeeves explained that this contradicted the claim made earlier that year by J. H. Wegstein [38] that the factor was 2, implying quadratic convergence. In 1966, Isaacson and Keller [20], and Ostrowski [30], independently presented a proof of how the golden mean convergence of the Secant Method in 1-D is obtained as a result of a Fibonacci-like recurrence.

3.8 Remarks

The Secant Method, the Regula Falsi Method, and the Modified Regula Falsi Method all use a secant line to approximate a function in a neighborhood of two initial estimates of a zero. The two initial estimates bracket the zero in the Regula Falsi Method (depicted in Figure 3.4) and the Modified Regula Falsi Method (depicted in Figure 3.5) but need not bracket a zero in the Secant Method (depicted in Figure 3.1). In the Regula Falsi Method, the current approximation replaces the previous interval endpoint whose corresponding function value has the same sign as the current best estimate of x^* , while the other interval endpoint is retained. However, in the Secant Method, the oldest point is always discarded

and replaced with the newer estimate, thus, at each iteration, the two most recent estimates are retained.

Hämmerlin and Hoffman [18], in 1991, stated that the Regula Falsi Method was one step of the Secant Method (which could mean they implicitly believed that the Regula Falsi Method, the Rule of Double False Position, was used for linear equations), and that the initial estimates did not need to bracket the solution. In addition, they stated that the Secant Method was a result of iterating the Regula Falsi Method (implying the Secant Method was used for nonlinear equations). We believe that the Regula Falsi Method, as we defined it, was introduced as a means to safeguard the Secant Method.

Many believe the Secant Method arose out of the finite difference approximation of the derivative in Newton's Method, as would be implied by our derivation, but historical evidence demonstrates the Secant Method predated Newton's Method by over 3000 years, and it was most commonly referred to in mathematical texts as the Rule of Double False Position and used to solve linear equations. Thus, we view the Secant Method as the first numerical linear algebra tool.

Chapter 4

Secant Method in Higher Dimensions

The interest to extend the 1-D Secant Method to higher dimensions was natural. Ostrowski [30] claims Gauss, in 1809 [16], [15] first extended the 1-D Secant Method to 2-D viewing it in the framework of linear interpolation of a function. Ostrowski [30], in 1966, refers to the Secant Method as the Regula Falsi Method, as Gauss did, and not the Rule of Double False Position (or “iteration with successive adjacent points,” as he did in 1960). Ortega and Rheinboldt [28] claim that the generalization of the 1-D Secant Method to higher dimensions is presented in Heinrich’s 1955 lectures. They credit Bittner [5] for the first analysis of the generalization in 1959. Wolfe [40], also in 1959, independently presented a generalization of the 1-D Secant Method to higher dimensions.

In 1970, Ortega and Rheinboldt [28] described a general framework underlying the construction of a basic secant approximation. In this chapter, we present this general framework and we demonstrate that the Wolfe and the Bittner formulations fit in this framework.

We describe $(n + 1)$ -point and 2-point Secant Methods, and discuss their properties to explain why there was motivation to find a Secant Method in higher dimensions that is numerically effective.

4.1 General Position

In this section, we present a definition and corresponding proposition that will aid in our discussion of the the n -D Secant Method. Our discussion follows §7.2 of Ortega and Rheinboldt [28], where the definition and proposition given below can be found. The i th component of the point $x_k \in \mathbb{R}^n$ is denoted

$$(x_k)_i \quad \text{for } i = 1, \dots, n, \quad (4.1)$$

and the n points used in the k th iteration are denoted

$$\{x_{k,j}\}_{j=1}^n = \{x_{k,1}, \dots, x_{k,n}\}.$$

Ortega and Rheinboldt utilize the term *general position* to describe a concept previously utilized, but not formally defined, by Bittner [5].

Definition 4.1.1. *Any $n + 1$ points x_0, \dots, x_n in \mathbb{R}^n are said to be in **general position** if the vectors $x_0 - x_j$, for $j = 1, \dots, n$, are linearly independent.*

The following proposition from Ortega and Rheinboldt [28] enlightens us with respect to points being in general position.

Proposition 4.1.2. *Let x_0, \dots, x_n be any $n+1$ points in \mathbb{R}^n . Then the following statements are equivalent:*

- (a) x_0, \dots, x_n are in general position.
- (b) For any j , $0 \leq j \leq n$, the vectors $x_j - x_m$, $m = 0, \dots, n$, $j \neq m$, are linearly independent.
- (c) The $(n+1) \times (n+1)$ matrix, (e, X^T) , where $e^T = (1, \dots, 1)$ and $X = (x_0, \dots, x_n)$ is nonsingular.
- (d) For any $y \in \mathbb{R}^n$, there exist scalars $\alpha_0, \dots, \alpha_n$ with $\sum_{j=0}^n \alpha_j = 1$ such that $y = \sum_{j=0}^n \alpha_j x_j$.

The geometric interpretation of general position given by Ortega and Rheinboldt [28] is that the points x_0, \dots, x_n are in general position if they do not lie in an affine subspace of dimension less than n . For example, for $n = 2$, the points x_0, x_1, x_2 are in general position if they are not colinear, that is, if they do not lie on a line in \mathbb{R}^2 .

4.2 1-D Discretized Newton Method

In the last chapter, we discussed methods that avoid explicitly calculating the derivative in the 1-D Newton iteration (3.3). By approximating the derivative using a finite difference quotient, we obtain the iteration

$$x_{k+1} = x_k - \left[\frac{f(x_k + h_k) - f(x_k)}{h_k} \right]^{-1} f(x_k), \quad k = 0, 1, \dots \quad (4.2)$$

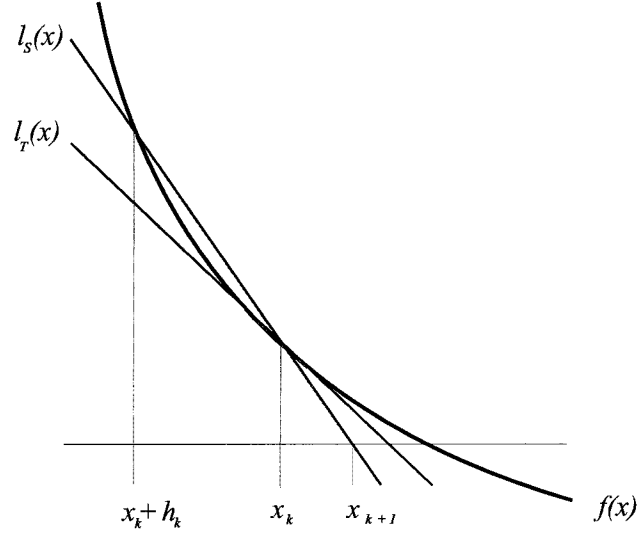


Figure 4.1: The Discretized Newton Method: two interpretations of $l(x)$.

which is the *1-D Discretized Newton Method* with discretization parameters h_k . The iteration (4.2) can also be written as

$$x_{k+1} = x_k - J(x_k, h_k)^{-1} f(x_k) \quad k = 0, 1, \dots, \quad (4.3)$$

where

$$J(x_k, h_k) = \frac{f(x_k + h_k) - f(x_k)}{h_k}.$$

The iterate x_{k+1} of iteration (4.2) is the solution of the linear equation

$$l(x) = \left[\frac{f(x_k + h_k) - f(x_k)}{h_k} \right] (x - x_k) + f(x_k) = 0. \quad (4.4)$$

Ortega and Rheinboldt [28] explain how we can view $l(x)$ (4.4) in two different ways (as depicted in Figure 4.1): either as an approximation of the tangent line, $l_T(x)$, or as a linear

interpolation, $l_S(x)$, of f passing through the points x_k and $x_k + h_k$.

If we let $h_k = x_{k-1} - x_k$ in the iteration (4.2) or (4.3), then we obtain the 1-D Secant Method that constructs $x_k \in \mathbb{R}$ as an approximation to a solution of $f(x) = 0$ for $f : \mathbb{R} \rightarrow \mathbb{R}$, using the iteration (3.5). Each new iteration of the 1-D Secant Method uses two recently computed iterates, x_{k-1} and x_k , thus we call it a 2-point method [36]. The 1-D Secant Method (i.e., $n = 1$) begins with two (i.e., $n + 1 = 2$) points in general position, and replaces $f(x)$ with the secant line, $l(x)$, connecting these two points. The subsequent iterate, x_{k+1} , is the solution of the linear equation

$$l(x) = \left[\frac{f(x_{k-1}) - f(x_k)}{x_{k-1} - x_k} \right] (x - x_k) + f(x_k) = 0, \quad (4.5)$$

that is, x_{k+1} is the point where the secant line intersects the x -axis. The general position of x_{k-1}, x_k and $f(x_{k-1}), f(x_k)$ reduces to $x_{k-1} \neq x_k$ and $f(x_{k-1}) \neq f(x_k)$, which ensures that there exists a secant line, it is unique, and it intersects the x -axis [28].

In the following section, we describe the extension of the 1-D Secant Method to higher dimensions using the second viewpoint (linear interpolation). Later, in §4.4 of this chapter, we show how to generalize the 1-D Discretized Newton Method (4.3) to n -D using the first viewpoint (discrete approximation).

4.3 The Ortega and Rheinboldt General Framework

Ortega and Rheinboldt [28] provide a general framework for obtaining a basic secant approximation. Their work embodies the theory presented, in 1959, by Bittner [5] and Wolfe

[40] independently. The Secant Method in n -D is used to approximate the solution of $F(x) = 0$ for $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \in \mathbb{R}^n$. Ortega and Rheinboldt begin with $n + 1$ points, x_0, \dots, x_n , and the corresponding $F(x_0), \dots, F(x_n)$, each of the two sets in general position. Recall that a function $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is called an *affine function* if it has the form

$$L(x) = Ax + b \quad \forall x \in \mathbb{R}^n, \quad (4.6)$$

where A is an $n \times n$ matrix and $b \in \mathbb{R}^n$. The function F is approximated by the affine function (4.6) that interpolates F at the $n + 1$ given points, i.e., $L(x_j) = F(x_j)$ for $j = 0, \dots, n$. The conditions $L(x_j) = F(x_j)$ for $j = 0, \dots, n$ can be represented in matrix form as

$$\begin{bmatrix} e, X^T \\ A^T \end{bmatrix} \begin{bmatrix} b^T \\ A^T \end{bmatrix} = \begin{bmatrix} F(x_0), \dots, F(x_n) \end{bmatrix}^T, \quad (4.7)$$

where $e^T = (1, 1, \dots, 1)$ and $X = (x_0, x_1, \dots, x_n)$, that is,

$$\begin{bmatrix} 1 & (x_0)_1 & (x_0)_2 & \dots & (x_0)_n \\ 1 & (x_1)_1 & (x_1)_2 & \dots & (x_1)_n \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & (x_n)_1 & (x_n)_2 & \dots & (x_n)_n \end{bmatrix} \begin{bmatrix} b_1 & \dots & b_n \\ A^T \end{bmatrix} = \begin{bmatrix} F(x_0)_1 & \dots & F(x_0)_n \\ \vdots & \dots & \vdots \\ F(x_n)_1 & \dots & F(x_n)_n \end{bmatrix}.$$

Clearly, the conditions

$$L(x_j) = Ax_j + b = F(x_j) \quad \text{for } j = 0, 1, \dots, n \quad (4.8)$$

imply

$$A(x_j - x_0) = F(x_j) - F(x_0) \quad \text{for } j = 1, \dots, n, \quad (4.9)$$

that is,

$$A \begin{pmatrix} x_1 - x_0 & x_2 - x_0 & \dots & x_n - x_0 \\ \vdots & \vdots & & \vdots \end{pmatrix}$$

$$= \begin{pmatrix} F(x_1) - F(x_0) & F(x_2) - F(x_0) & \dots & F(x_n) - F(x_0) \\ \vdots & \vdots & & \vdots \end{pmatrix}.$$

We point out that if the x_0, \dots, x_n are in general position, then we obtain a unique affine function (4.6) and the matrix A is unique. In addition, if the $F(x_0), \dots, F(x_n)$ are in general position, then A is invertible.

If we follow Ortega and Rheinboldt's presentation in a straightforward manner, we first solve the system (4.7) to obtain A and b that satisfy

$$Ax_j + b = F(x_j), \quad \text{for } j = 0, \dots, n. \quad (4.10)$$

Next, we solve for x in

$$Ax + b = 0,$$

that is, since A is invertible, we solve for the well-defined solution of the linear system

$$L(x) = 0. \quad (4.11)$$

Ortega and Rheinboldt call the point

$$x^s = -A^{-1}b \quad (4.12)$$

a *basic secant approximation* with respect to x_0, \dots, x_n [28]. It follows from (4.10) that at the k th iteration,

$$\begin{aligned} x_{k+1} &= -A_k^{-1}b_k \\ &= A_k^{-1}(-A_k x_k + F(x_k)) \\ &= x_k - A_k^{-1}F(x_k). \end{aligned}$$

Recall that we refer to an iterative method of the form $x_{k+1} = x_k - A_k^{-1}F(x_k)$, where A_k is viewed as an approximation to $F'(x_k)$, as a Secant Method if it satisfies a Level- j Secant Equation (2.6). From (4.9), we have

$$A_k(x_k - x_{k-1}) = F(x_k) - F(x_{k-1}).$$

Thus, the Level- j Secant Equation (2.6) for $j = 1, \dots, n$ is satisfied as long as the x_j s are in general position.

Ortega and Rheinboldt also present alternative formulations to compute a basic secant approximation that do not require computing the interpolating function (4.6) explicitly, and as a result, only require solving one linear system instead of two. We present an example of such a formulation in the next section and another example in §4.4.1.

4.3.1 Wolfe's Interpolation Formulation of the $(n + 1)$ -point Secant Method

In 1959, Wolfe [40] suggested and implemented an $(n + 1)$ -point method for simultaneous nonlinear equations that he viewed as a generalization of the 1-D Secant Method to n -D. Ortega and Rheinboldt present, on page 193 [28], Wolfe's formulation in the context of their general framework. Wolfe began his interpolation method with $n + 1$ points, x_0, \dots, x_n , with $x \in \mathbb{R}^n$, and the corresponding $F(x_0), \dots, F(x_n)$. He formed the matrix

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ F(x_0) & F(x_1) & \dots & F(x_n) \end{bmatrix} \quad (4.13)$$

and claimed that it is invertible which implies that $F(x_0), \dots, F(x_n)$ are in general position.

Wolfe then solved the $(n+1) \times (n+1)$ linear system

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ F(x_0) & F(x_1) & \cdots & F(x_n) \end{bmatrix} (z) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (4.14)$$

for the unique solution $z = (z_0, \dots, z_n)^T$ that clearly satisfies

$$\sum_{j=0}^n z_j = 1 \quad \text{and} \quad \sum_{j=0}^n z_j F(x_j) = 0. \quad (4.15)$$

Ortega and Rheinboldt [28] point out that from (4.10), we have

$$\begin{aligned} 0 &= \sum_{j=0}^n z_j F(x_j) \\ &= \sum_{j=0}^n z_j (Ax_j + b) \\ &= A \left(\sum_{j=0}^n z_j x_j \right) + b \end{aligned}$$

where $Ax + b = 0$ has a unique solution. Wolfe gave the point

$$x^s = \sum_{j=0}^n z_j x_j \quad (4.16)$$

as his basic secant approximation.

Wolfe [40] attained a basic secant approximation using the past $n+1$ iterates and the past $n+1$ function evaluations by solving the linear system of equations (4.14) followed by the calculation of the linear combination (4.16) of the vectors x_0, \dots, x_n . However, instead of discarding the oldest iterate, as in the Secant Method, Wolfe suggested [40] discarding the past iterate where $\|F(x_j)\|$ is a maximum since this iterate is most unlike a solution. This is a reasonable choice, since $\{\|F(x_k)\|\}$ is converging to zero and thus,

for subsequent iterations, one would expect monotone behavior in the quantity $\|F(x_k)\|$. Therefore, Wolfe's choice should eventually lead to the standard Secant Method.

Wolfe [40] solved a variety of problems in 2-D using FORTRAN II . With his programs, he witnessed, in his simple examples, that the error at a given step is proportional to the product of the errors at the two previous steps. Even though Wolfe did not provide convergence analysis, he believed his method to have the golden mean convergence that Jeeves [21] had demonstrated for the Secant Method in 1-D. In 1964, Tornheim [37] proved that the order of convergence of the $(n + 1)$ -point Secant Method is at least the value of the positive root of $r^{n+1} - r^n - 1 = 0$. Tornheim also demonstrated that the order of convergence of the $(n + 1)$ -point Secant Method using Wolfe's 2-D example complied with his convergence statement. Properties of the $(n + 1)$ -point Secant Method will be discussed in §4.5.1 of this chapter.

4.4 n -D Discretized Newton Method

We now describe how to generalize the 1-D Discretized Newton Method (4.2) to n -D following Ortega and Rheinboldt's second viewpoint of interpreting the linear equation $l(x)$ (4.5) as an approximation of the tangent line $l_T(x)$ depicted in Figure 4.1, Recall that we replaced the derivative in the 1-D Newton iteration

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

with a finite difference quotient to obtain the 1-D Discretized Newton iteration

$$x_{k+1} = x_k - \left[\frac{f(x_k + h_k) - f(x_k)}{h_k} \right]^{-1} f(x_k), \quad k = 0, 1, \dots, \quad (4.17)$$

which can also be written as

$$x_{k+1} = x_k - J(x_k, h_k)^{-1} f(x_k) \quad k = 0, 1, \dots,$$

where

$$J(x_k, h_k) = \frac{f(x_k + h_k) - f(x_k)}{h_k}.$$

If we set

$$H = (h_1, \dots, h_n) \quad \text{for given} \quad h_i \in \mathbb{R}^n, \quad (4.18)$$

we can similarly replace the derivative, $F'(x_k)$, in the n -D Newton iteration

$$x_{k+1} = x_k - F'(x_k)^{-1} F(x_k),$$

with $J(x_k, H_k)$, a matrix of difference quotients approximating it,

$$J(x_k, H_k) = (F(x_k + H_k e_1) - F(x_k), \dots, F(x_k + H_k e_n) - F(x_k)) H_k^{-1}, \quad (4.19)$$

where $e_i : 1 \leq i \leq n$ represents the set of natural basis vectors for \mathbb{R}^n , that is, e_i is the vector with a 1 in the i th coordinate and 0 elsewhere. The matrix

$$H_k = (x_{k,1} - x_k, \dots, x_{k,n} - x_k)$$

utilized in (4.19) is constructed from the **auxiliary points**, $\{x_{k,1}, \dots, x_{k,n}\}$. The basic secant approximation can be obtained from the n -D Discretized Newton iteration

$$x_{k+1} = x_k - J(x_k, H_k)^{-1} F(x_k), \quad k = 0, 1, \dots, \quad (4.20)$$

where the different choices of the auxiliary points determines the different iterative methods.

Next, we present examples of n -D Discretized Newton Methods: $(n + 1)$ -point and 2-point methods. We conclude the chapter with an analysis of these methods.

4.4.1 $(n + 1)$ -Point Discretized Newton Method

The $(n+1)$ -point Discretized Newton Method uses the iteration (4.20) where the n auxiliary points that form the H matrix are arbitrarily chosen from the set of previously iterates x_0, \dots, x_{k-1} . The $(n+1)$ -point Discretized Newton Method, as Ortega and Rheinboldt [28] present on page 194, begins with the x_0, \dots, x_n and the corresponding $F(x_0), \dots, F(x_n)$, each set in general position. If we set

$$H = (x_1 - x_0, \dots, x_n - x_0), \quad (4.21)$$

since $F(x_i) = F(x_0 + He_i)$, then it follows from (4.9) that

$$\begin{aligned} AH &= (F(x_0 + He_1) - F(x_0), \dots, F(x_0 + He_n) - F(x_0)) \\ &= (F(x_1) - F(x_0), \dots, F(x_n) - F(x_0)). \end{aligned}$$

Since H is nonsingular, $A = J(x_0, H)$. Moreover, the matrix of difference quotients, $J(x_0, H)$, is nonsingular. Using $b = F(x_0) - Ax_0$, we obtain

$$\begin{aligned} x^s &= -A^{-1}b \\ &= -A^{-1}(-Ax_0 + F(x_0)) \\ &= x_0 - A^{-1}F(x_0). \end{aligned}$$

Since $A = J(x_0, H)$, the point

$$x^s = x_0 - J(x_0, H)^{-1}F(x_0) \quad (4.22)$$

is a basic secant approximation. Hence, the basic secant approximation is obtained by solving the linear system of equations $AH(x) = F(x_0)$ followed by the calculation of the linear combination of the vectors x_0, \dots, x_n by means of (4.22). We wish to point out that both Wolfe's interpolation formulation and the discretized Newton formulation require the solution of only one linear system of equations followed by the linear combination of the vectors x_0, \dots, x_n .

4.4.1.1 Sequential $(n+1)$ -point Discretized Newton Method

The Sequential $(n+1)$ -point Discretized Newton Method uses the n -D Discretized Newton iteration (4.20) where the n auxiliary points that construct the H matrix are the n previously computed iterates

$$\{x_{k,j}\}_{j=1}^n = \{x_{k,k-1}, \dots, x_{k,k-n}\}.$$

4.4.2 Sequential 2-point Discretized Newton Method

The Sequential 2-point Discretized Newton Method uses the n -D Discretized Newton iteration (4.20) where the n auxiliary points that construct the H matrix depend on the 2 previously computed iterates. That is, we use

$$x_{k,j} = x_k + \sum_{i=1}^j \left[(x_{k-1})_i - (x_k)_i \right] e_i \quad \text{for } j = 1, \dots, n, \quad (4.23)$$

to obtain the auxiliary points

$$\begin{aligned}
 x_{k,1} &= x_k + \left[(x_{k-1})_1 - (x_k)_1 \right] e_1 \\
 x_{k,2} &= x_k + \sum_{i=1}^2 \left[(x_{k-1})_i - (x_k)_i \right] e_i \\
 &\vdots = \vdots \\
 x_{k,n} &= x_k + \sum_{i=1}^n \left[(x_{k-1})_i - (x_k)_i \right] e_i.
 \end{aligned}$$

4.5 Analysis

Recall that we refer to an iterative method of the form $x_{k+1} = x_k - A_k^{-1}F(x_k)$, where A_k is viewed as an approximation to $F'(x_k)$, as a Secant Method if it reduces to the Secant Method in 1-D, and thus satisfies the Level-1 Secant Equation. Both the Sequential $(n+1)$ -point Discretized Newton Method and the Sequential 2-point Discretized Newton Method reduce to the Secant Method for $n = 1$. In addition, these methods should satisfy a Level- j Secant Equation (2.6). Therefore, from this point on, we shall refer to them as the Sequential $(n+1)$ -point Secant Method and the Sequential 2-point Secant Method, respectively.

We believe that methods that satisfy a Level- j Secant Equation are good, however, under some circumstances, such methods may fail. In this section, we examine the Sequential $(n+1)$ -point Secant Method and the Sequential 2-point Secant Method. We state each method's order of convergence and whether or not it satisfies a Level- j Secant Equation (2.6).

4.5.1 Properties of the $(n + 1)$ -point Secant Method

In 1959, Bittner [5] was the first to provide a convergence statement of the $(n + 1)$ -point Secant Method for a continuously differentiable function $F(x)$ when there exists an x^* for which $F(x^*) = 0$ and $F'(x^*)$ is nonsingular. He let

$$K(\sigma) = \{H = (h_1, \dots, h_n) \in \mathbb{R}^{n \times n} \mid h_j \neq 0, \quad j = 1, \dots, n; \mid \det \left(\frac{h_1}{\|h_1\|}, \dots, \frac{h_n}{\|h_n\|} \right) \mid \geq \sigma\},$$

where H is nonsingular. Bittner assumed, further, that $\sigma > 0$ is chosen so that $K(\sigma)$ is not empty. Then, there is a constant $r_1 > 0$ such that, for any sequence $\{H_k\} \subset K(\sigma)$ with $\|H_k\| \leq r_1$, and $k = 0, 1, \dots$, the sequence (4.19) is well defined and $\lim_{k \rightarrow \infty} x_k = x^*$, provided that $\|x_0 - x^*\|$ is sufficiently small. Moreover, if $\lim_{k \rightarrow \infty} H_k = 0$, then the sequence $\{x_k\}$ converges to x^* Q -superlinearly (2.11). Further details of Bittner's convergence statement can be found in his article [5].

In 1964, Tornheim [37] proved that if, additionally, the function satisfies the Lipschitz condition

$$\|F'(x) - F'(x^*)\| \leq \gamma \|x - x^*\| \quad \text{for } \gamma \geq 0, \quad (4.24)$$

then the order of convergence of the Sequential $(n + 1)$ -point Secant Method, when the method does not fail, is at least the value of the largest root of $r^{n+1} - r^n - 1 = 0$. In 1970, Oregate and Rheinboldt [28] demonstrated that the order of convergence of the Sequential $(n + 1)$ -point Secant Method decreases rapidly with n . For example, if $n = 100$, then the largest root of $r^{n+1} - r^n - 1 = 0$ is $1.03 \dots$, hence as n increases, the value of the largest

root decreases, that is, the order of convergence decreases [28].

At the first step of the Sequential $(n + 1)$ -point Secant Method, we have x_0, \dots, x_n and must calculate $F(x_0), \dots, F(x_n)$. At subsequent steps, $F(x_{k-1}), \dots, F(x_{k-n})$ are already available, thus the Sequential $(n + 1)$ -point Secant Method requires the computation of only one new function evaluation per step - namely, $F(x_k)$. While this requirement of only one function evaluation per step is attractive, the disadvantage is that we must store $n + 1$ points and their corresponding function values at each step.

4.5.2 Properties of the Sequential 2-point Method

In 1961, Schmidt [34] derived the golden mean convergence of the Sequential 2-point Secant Method. Golden mean convergence is guaranteed, but at the expense of n new function evaluations per step since $F(x_{k-1})$ is available from the previous step.

In 1963, Schmidt [35] concluded that with the addition of the following three conditions:

1. $F(x)$ is sufficiently smooth, that is, $F'(x)$ satisfies the Lipschitz condition (4.24), and
2. $J(x, H)$ is a strongly consistent approximation to $F'(x)$, that is, $J(x, H)$ satisfies

$$\|F'(x) - J(x, H)\| \leq c\|H\| \quad \text{for } c > 0, \quad \text{and}$$

3. the rate of decrease of h_k , is sufficiently rapid,

the order of convergence of the Sequential 2-point Secant Method is at least golden mean.

Then, in 1970, Ortega and Rheinboldt [28] showed that the Sequential 2-point Secant

Wolfe example	Ortega and Rheinboldt example
$f_1(x, y) = x^2 + x - y^2 + 1$	$f_1(x, y) = x$
$f_2(x, y) = y(1 - 2x)$	$f_2(x, y) = x^2 - \frac{1}{2}y$
$(x_0, y_0) = (-.6, 1.1)$	$(x_0, y_0) = (0, -\alpha)$
$(x_1, y_1) = (-.3, 1.1)$	$(x_1, y_1) = (-\alpha, 2\alpha^2)$
$(x_2, y_2) = (-.6, 1.4)$	$(x_2, y_2) = (\alpha, 2\alpha^2)$
	for $0 < \alpha < 1$

Table 4.1: The Wolfe 2-D example with initial points, and the Ortega and Rheinboldt 2-D example with initial points.

Method satisfies the Level-1 Secant Equation (2.7).

4.5.3 Numerical Experiment

In this section, we present some observations that we made when we examined the Sequential $(n + 1)$ -point Secant Method, as described by Bittner and Ortega and Rheinboldt. For each formulation of the Sequential $(n + 1)$ -point Secant Method, we used high precision arithmetic in MATLAB to test the 2-D examples provided by Wolfe [40] and Ortega and Rheinboldt [28] that are presented in Table 4.1.

Our numerical experiments demonstrated that the Sequential $(n+1)$ -point Secant Method can fail due to the loss of general position of the x_j s. For example, using the Ortega and

Rheinboldt 2-D example, we begin with the three given initial points presented in Table 4.1, in general position, to obtain

$$H_0 = - \begin{pmatrix} 2\alpha & \alpha \\ 0 & \alpha + 2\alpha^2 \end{pmatrix}.$$

The first iterate, x_3 , of iteration (4.20) is

$$x_3 = \begin{pmatrix} \alpha \\ 2\alpha^2 \end{pmatrix} - \begin{pmatrix} 2\alpha & \alpha \\ 0 & \alpha + 2\alpha^2 \end{pmatrix} \begin{pmatrix} 2\alpha & \alpha \\ 0 & -\frac{\alpha}{2} \end{pmatrix}^{-1} \begin{pmatrix} \alpha \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2\alpha^2 \end{pmatrix}.$$

Note that x_3 is considerably closer to $x^* = 0$ than any of the initial points, however, x_1, x_2 , and x_3 are now colinear, thus H_1 is singular. This exemplifies how the Sequential $(n + 1)$ -point Secant Method can fail because H becomes singular. Our results also demonstrate that the Sequential $(n + 1)$ -point Secant Method, when it does not fail, satisfies the Level- j Secant Equation for $j \leq n$.

4.6 Conclusion

In summary, the Secant Method is numerically effective in 1-D but becomes less effective, and may fail in practice, as the dimension increases. As we significantly increase the dimension, the vectors $x_0 - x_k$ for $k = 1, \dots, n$ that are in general position become effectively numerically linearly dependent. As a result, the system we need to solve, $L(x) = 0$, becomes ill-conditioned (nearly singular) which makes it more difficult to solve. Thus, the algorithm becomes numerically unstable due to solving nearly singular systems. However, if the Sequential $(n + 1)$ -point Secant Method works, the Level- j Secant Equation is satisfied for $j \leq n$. In conclusion, historically, linear interpolation methods have not been

pursued, and there has been motivation to find a Secant Method in higher dimensions that is numerically effective.

Bibliography

- [1] AL-KASHI, *Miftah al-Hisab*, University of Damascus Press, 1977.
- [2] P. APIANUS, *Eyn Newe Unnd wolgegründte underweysun aller Kauffmanss Rechnung*, Ingolstadt, (1527).
- [3] K. BACHMANN, *Der Konvergenzgrad bei iterativer Lösung von Gleichungen durch inverse Interpolation*, Zeitschrift für angewandte Mathematik und Mechanik, 34 (1954), pp. 282–283.
- [4] W. P. BERLINGHOFF AND F. Q. GOUVÊA, *Math through the Ages - A Gentle History for Teachers and Others, Expanded Edition*, Oxten House and Mathematical Association of America, Washington, D.C. and Farmington, M.E., 2004.
- [5] L. BITTNER, *Eine Verallgemeinerung des Sekantenverfahrens (Regula Falsi) zur Näherungsweise Berechnung der Nullstellen eines nichtlinearen Gleichungssystems*, Wissenschaftliche Zeitschrift der Technischen Hochschule Dresden, 9 (1959), p. 325.
- [6] A. D. BOOTH, *Numerical Methods*, Academic Press, Inc., New York, N.Y., 1955.
- [7] G. J. BRONSON, *C++ for Engineers and Scientists*, PWS Publishing, 1999.
- [8] C. BROYDEN, *The Convergence of Single-Rank Quasi-Newton Methods*, Mathematics of Computation, 24 (1970), pp. 365–382.
- [9] D. M. BURTON, *The History of Mathematics: An Introduction, 6th Edition*, McGraw Hill, New York, N.Y., 2007.
- [10] M. CANTOR, *Vorlesungen über Geschichte der Mathematik*, Leipzig, 1900.
- [11] J.-L. CHABERT, *A History of Algorithms: From the Pebble to the Microchip*, Springer, 1998.
- [12] M. CONTRERAS AND R. TAPIA, *Sizing the BFGS and DFP updates*, Journal of Optimization Theory and Applications, 78 (1993), pp. 93–108.
- [13] G. DAHLQUIST AND Å. BJÖRCK, *Numerical methods*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1974. Translated from the 1969 Swedish by Ned Anderson, Prentice-Hall Series in Automatic Computation.

- [14] T. FINCKE, *Thomae Finkii Flenspurgensis Geometriae rotundi libri XIII*, 1583.
- [15] C. GAUSS, *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*, 120 (1809), pp. 136–138.
- [16] —, *Werke*, VII (1809), pp. 150–152.
- [17] P. GILL, W. MURRAY, AND W. WRIGHT, *Practical Optimization*, Academic Press, London, 1981.
- [18] G. HÄMMERLIN AND K.-H. HOFFMANN, *Numerical Mathematics*, Springer-Verlag, New York, 1991.
- [19] P. HENRICI, *Elements of Numerical Analysis*, John Wiley and Sons, New York, London, Sydney, 1964.
- [20] E. ISAACSON AND H. B. KELLER, *Analysis of Numerical Methods*, Dover, New York, N.Y., 1966.
- [21] T. JEEVES, *Secant Modification of Newton's Method*, Westinghouse Research Laboratories, (1958), pp. 9–10.
- [22] G. G. JOSEPH, *The Crest of the Peacock: Non-European Roots of Mathematics*, Princeton University Press, Princeton, N.J., 2000.
- [23] S. KANGSHEN, J. CROSSLEY, AND A. W.-C. LUN, *The Nine Chapters on the Mathematical Art: Companions and Commentary*, Oxford University Press, New York, 1999.
- [24] M. KLINE, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, New York, 1972.
- [25] L. MA, *The Rule of False: Early Applications and Conjectured Transmissions*, Technical Paper 1993-15, Chalmers University of Technology, University of Göteborg, 1993.
- [26] C. MAAS, *Was ist das Falsche an der Regula Falsi?*, Mitteilungen der Mathematischen Gesellschaft in Hamburg 11, 3 (1985), pp. 311–317.
- [27] J. J. MORÉ AND D. SORESENSEN, *Newton's Method*, Technical Paper ANL-82-8, Argonne National Laboratory, Argonne, Illinois, U.S.A., 1982.
- [28] J. M. ORTEGA AND W. C. RHEINBOLDT, *Studies in Numerical Analysis 2: Numerical Solutions of nonlinear Problems*, SIAM, Philadelphia, P.A., U.S.A., 1970.
- [29] A. OSTROWSKI, *Solution of Equations and Systems of Equations*, Academic Press, New York and London, 1960.

- [30] —, *Solution of Equations and Systems of Equations*, Academic Press, New York and London, second ed., 1966.
- [31] L. PISANO, *Liber Abaci*, (1857-1862).
- [32] H. RESNIKOFF AND R. WELLS JR., *Mathematics in Civilization*, Dover Publications, Inc., New York, 1984.
- [33] G. ROBINS AND C. SHUTE, *The Rhind Mathematical Papyrus: An ancient Egyptian text*, British Museum Publications, Ltd., London, 1987.
- [34] J. SCHMIDT, *Die Regula Falsi für Operatoren in Banachräumen*, Zeitschrift für angewandte Mathematik und Mechanik, 41 (1961), pp. 61–63.
- [35] —, *Eine Übertragung der Regula Falsi auf Gleichungen in Banach-Räumen*, Zeitschrift für angewandte Mathematik und Mechanik, (1963), pp. 1–8, 97–110.
- [36] G. STEWART, *Afternotes on Numerical Analysis*, SIAM, Philadelphia, 1996.
- [37] L. TORNHEIM, *Convergence of Multipoint Iterative Methods*, Journal of the Association for Computing Machinery, 11 (1964), pp. 210–220.
- [38] J. WEGSTEIN, *Accelerating Convergence of Iterative Processes*, Communications of the ACM, 1 (1958), pp. 9–13.
- [39] D. WHITESIDE, *The Mathematical Papers of Isaac Newton, Volumes I-VII*, Cambridge University Press, Cambridge, 1967-1976.
- [40] P. WOLFE, *The Secant Method for Simultaneous Nonlinear Equations*, Communications of the ACM, The RAND Corporation, Santa Monica, California., 1959.
- [41] T. J. YPMA, *Historical Development of the Newton-Raphson Method*, SIAM Review, 37 (1995), pp. 531–551.