

## ABSTRACT

HEWITT, CHRISTINA M. Real Roots of Polynomials with Real Coefficients. (Under the direction of Dr. Michael Singer).

Polynomial equations are used throughout mathematics. When solving polynomials many questions arise such as: Are there any real roots? If so, how many? Where are they located? Are these roots positive or negative? Depending on the problem being solved sometimes a rough estimate for the interval where a root is located is enough.

There are many methods that can be used to answer these questions. We will focus on Descartes' Rule of Signs, the Budan-Fourier theorem and Sturm's theorem. Descartes' Rule of Signs traditionally is used to determine the possible number of positive real roots of a polynomial. This method can be modified to also find the possible negative roots for a polynomial. The Budan-Fourier theorem takes advantage of the derivatives of a polynomial to determine the number of possible number of roots. While Sturm's theorem uses a blend of derivatives and the Euclidean Algorithm to determine the exact number of roots.

In some cases, an interval where a root of the polynomial exists is not enough. Two methods, Horner and Newton's methods, to numerically approximate roots up to a given precision are also discussed. We will also give a real world application that uses Sturm's theorem to solve a problem.

© Copyright 2009 by Christina M. Hewitt

All Rights Reserved

Real Roots of Polynomials with Real Coefficients

by  
Christina M. Hewitt

A thesis submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the Degree of  
Master of Science

Applied Mathematics

Raleigh, North Carolina

2009

APPROVED BY:

---

Dr. E. L. Stitzinger

---

Dr. A. Szanto

---

Dr. M. F. Singer  
Chair of Advisory Committee

## BIOGRAPHY

Christina Hewitt attended North Carolina State University as an undergraduate. She obtained two degrees in December of 2005, Bachelors of Science in Mathematics and Bachelors of Science in Math Education. After working in the private sector for a few years, Christina returned to North Carolina State University for graduate school to study Applied Mathematics.

## TABLE OF CONTENTS

<b>LIST OF TABLES</b> .....	<b>iv</b>
<b>LIST OF FIGURES</b> .....	<b>v</b>
<b>1 Introduction</b> .....	<b>1</b>
1.1 Descartes's Rule of Signs . . . . .	1
1.2 Budan-Fourier Theorem . . . . .	4
1.3 Euclidean Algorithm . . . . .	9
1.4 Sturm's Theorem . . . . .	16
<b>2 Numerical Approximations of Roots</b> .....	<b>24</b>
2.1 Horner's Method . . . . .	24
2.2 Newton's Method . . . . .	31
<b>3 Real World Applications</b> .....	<b>41</b>
3.1 Applications of Sturm sequences to bifurcation analysis of delay differential equation models. . . . .	41
<b>Bibliography</b> .....	<b>43</b>
<b>Appendices</b> .....	<b>44</b>
Appendix A . . . . .	45
Appendix B . . . . .	46

## LIST OF TABLES

Table 1.1 Budan-Fourier Sign Variations.....	5
Table 1.2 Budan-Fourier Sign Variations.....	6
Table 1.3 Budan-Fourier Sign Variations.....	8
Table 1.4 Sturm Function Sign Variations .....	18
Table 1.5 Sturm Function Sign Variations .....	22
Table 2.1 Horner's method on the interval $[2, 2.5]$ .....	30
Table 2.2 Horner's method on the interval $[2.5, 3]$ .....	31
Table 2.3 Derivative Sign Variations.....	38

## LIST OF FIGURES

Figure 2.1	$f(x) = x^3 - 3x^2 - 4x + 13$ .....	33
Figure 2.2	$f(x) = x^3 - 3x^2 - 4x + 13$ .....	34
Figure 2.3	$f(x) = x^3 - 3x^2 - 4x + 13$ .....	34
Figure 2.4	Newton Scenarios. ....	36

# Chapter 1

## Introduction

Polynomial equations are used throughout mathematics. When solving polynomials many questions arise such as: Are there any real roots? If so, how many? Where are they located? Are these roots positive or negative? Depending on the problem being solved sometimes a rough estimate for the interval where a root is located is enough.

There are many methods that can be used to answer these questions. We will focus on Descartes' Rule of Signs, the Budan-Fourier theorem and Sturm's theorem. Descartes' Rule of Signs traditionally is used to determine the possible number of positive real roots of a polynomial. This method can be modified to also find the possible negative roots for a polynomial. The Budan-Fourier theorem takes advantage of the derivatives of a polynomial to determine the number of possible number of roots. While Sturm's theorem uses a blend of derivatives and the Euclidean Algorithm to determine the exact number of roots.

In some cases, an interval where a root of the polynomial exists is not enough. Two methods, Horner and Newton's methods, to numerically approximate roots up to a given precision are also discussed. We will also give a real world application that uses Sturm's theorem to solve a problem.

### 1.1 Descartes's Rule of Signs

There are multiple ways to determine the number of roots of a polynomial. Descartes's Rule of Signs helps determine the number of positive or negative roots of a polynomial with real coefficients.



**Theorem 1.1.1.** (*Descartes's Rule of Signs*) Let  $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$  be a polynomial with real coefficients, where  $a_i \in \mathbb{R}$ ,  $i = 0, \dots, n$ , and  $a_n$  and  $a_0$  are nonzero. Let  $v$  be the number of changes of signs in the sequence  $(a_0, \dots, a_n)$  of its coefficients and let  $r$  be the number of its real positive roots, counted with their orders of multiplicity, then there exists some nonnegative integer  $m$  such that

$$r = v - 2m$$

(c.f. [Mig92]).

It is assumed that the polynomials are written in descending powers of  $x$ . *Proof of Descartes's Rule of Signs* can be found in the Appendix B. We will represent the coefficients of the polynomial  $f(x)$  in the sequence  $\{a_n, a_{n-1}, \dots, a_1, a_0\}$ . Some of the coefficients of the real polynomial may be zero. *We will only be considering the nonzero coefficients.* Let any coefficient  $a_i < 0$  be represented by  $-1$  and  $a_i > 0$  be represented by  $1$ . In order to gain a better understanding of Descartes's rule, we'll begin with an example.

**Example 1.1.1.** Consider the polynomial

$$f(x) = x^3 - 3x^2 - 4x + 13.$$

For this polynomial, the signs for the leading coefficients are represented by  $(1, -1, -1, 1)$  resulting in two sign changes, therefore  $v = 2$ . According to Descartes's Rule of Signs,  $f(x)$  has either

$$\begin{aligned} r &= 2 - 2(0) = 2 \text{ or} \\ r &= 2 - 2(1) = 0 \end{aligned}$$

positive roots. Hence, there is a maximum of 2 positive real roots and a minimum of 0 positive real root.

**Example 1.1.2.** Now, consider the polynomial

$$f(x) = x^5 - x^4 + 3x^3 + 9x^2 - x + 5.$$

The signs for the leading coefficients are represented by  $(1, -1, 1, 1, -1, 1)$ , resulting in four sign changes,  $v = 4$ . According to Descartes's Rule of Signs, the polynomial has either 4, 2, or no positive roots. So, there is a maximum of 4 positive roots.

**Example 1.1.3.** Consider the polynomial

$$f(x) = x^6 - x^4 + 2x^2 - 3x - 1.$$

For this polynomial, the signs for the leading coefficients are represented by  $(1, -1, 1, -1, -1)$  resulting in three sign changes, therefore  $v = 3$ . According to Descartes's Rule of Signs,  $f(x)$  has either

$$r = 3 - 2(0) = 3 \text{ or}$$

$$r = 3 - 2(1) = 1$$

positive roots. Hence, there is a maximum of 3 positive real roots and a minimum of 1 positive real root.

It was mentioned that Descartes's Rule of Signs can be used to determine the positive and the negative roots of a polynomials with real coefficients.

**Exercise 1.1.1.** How can the Descartes's Rule of Signs be used to determine the number of negative roots for a polynomial with real coefficients? Hint: Use  $f(-x)$ .

So far all of our examples have had at least two sign changes. Now, let us think about what happens when there are only one or no sign changes in the polynomial (i.e.  $v = 1$  or  $v = 0$ ). Recall, that we can only have a nonnegative number of roots,  $r \geq 0$ , and that  $r$  can only be reduced by even values.

**Exercise 1.1.2.** How many roots polynomials have with only one sign change? no sign change?

**Exercise 1.1.3.** Determine the number of possible positive and negative roots for the following polynomials using Descartes's Rule of Signs.

1.  $f(x) = x^4 - 2x^3 + 4x^2 - 3x + 1.$

2.  $q(x) = x^9 + 3x^8 - 5x^3 + 4x + 6.$

3.  $p(x) = 2x^3 + 5x^2 + x + 1.$

4.  $f(x) = 3x^4 + 10x^2 + 5x - 4.$

In the statement of Descartes's Rule of Signs, it is assumed that  $a_0 \neq 0$ . What happens if  $a_0 = 0$ ? Let us assume that  $a_0 = 0$ , then

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x$$

where  $a_i \in \mathbb{R}$ ,  $i = 1, \dots, n$ . Can we still use Descartes's Rule of Signs? Yes!

**Exercise 1.1.4.** Given that  $a_0 = 0$ , show how Descartes's Rule of Signs can still be used to determine the positive (and negative) real roots for the polynomial. Hint:  $x = 0$  is not considered a positive nor negative root. Divide.

## 1.2 Budan-Fourier Theorem

There are times when we want to know more than the number of real roots of a polynomial. We would prefer to know the exact values of the roots or an approximation. In order to determine real roots, we must know where they are located on the real line. A theorem used to determine an interval which contains a root is the Budan-Fourier theorem.

**Theorem 1.2.1.** (*Budan-Fourier Theorem*) Let

$$f(x) = 0$$

be a nonconstant polynomial of degree  $n$  with real coefficients. Designate by  $v(c)$  the number of changes of signs in the sequence

$$f(x), f'(x), f''(x), \dots, f^{(n)}(x)$$

when  $x = c$ , where  $c$  is any real number. Then the number of zeros of  $f$  in the interval  $(a, b]$ , counted with their orders of multiplicity is equal to

$$v(a) - v(b) - 2m, \quad \text{for some } m \in \mathbb{N}$$

(c.f. [Con43]).

The proof of the Budan-Fourier theorem can be found in the Appendix B. We will expand on previous examples to demonstrate the Budan-Fourier theorem.

**Example 1.2.1.** Consider the polynomial in Example 1.1.1

$$\begin{aligned} f(x) &= x^3 - 3x^2 - 4x + 13, \\ f'(x) &= 3x^2 - 6x - 4, \\ f''(x) &= 6x - 6, \\ f^{(3)}(x) &= 6. \end{aligned}$$

Hence, we get Table 1.1.

Table 1.1: Budan-Fourier Sign Variations

$x$	$f(x)$	$f'(x)$	$f''(x)$	$f^{(3)}(x)$	Variations
$-\infty$	—	+	—	+	3
−3	—	+	—	+	3
−2	+	+	—	+	2
−1	+	+	—	+	2
0	+	—	—	+	2
1	+	—	+	+	2
2	+	—	+	+	2
3	+	+	+	+	0
$\infty$	+	+	+	+	0

In Table 1.1, the rows  $x = -\infty$  and  $x = \infty$  can be interpreted as follows. As  $x$  approaches positive infinity, any nonzero polynomial is eventually always positive or always negative. We denote  $v(\infty)$  this sign. The variation  $v(-\infty)$  is defined in a similar manner.

By Descartes's Rule of Signs there are either 2 or no positive real roots and one negative real root. From Table 1.1 we get,

$$\begin{aligned} v(-3) - v(-2) - 2m &= 3 - 2 - 2m = 1, \\ v(-2) - v(-1) - 2m &= 2 - 2 - 2m = 0, \\ v(-1) - v(0) - 2m &= 2 - 2 - 2m = 0, \\ v(0) - v(1) - 2m &= 2 - 2 - 2m = 0, \\ v(1) - v(2) - 2m &= 2 - 2 - 2m = 0, \\ v(2) - v(3) - 2m &= 2 - 0 - 2m = 2 - 2m. \end{aligned}$$

Using the Budan-Fourier theorem, we can determine where these real roots are located. The Budan-Fourier theorem shows that there are no real roots on the interval  $(-\infty, -3]$  since  $v(a) - v(b) - 2m = 0$  for all  $a, b \in (-\infty, -3]$ . Since  $f^{(i)}(3) \geq 0$  for  $i = 0, \dots, 3$ , the polynomial has no real roots in the interval  $[3, \infty)$ . By the Budan-Fourier theorem, any of the possible 2 or no positive real roots are on the interval  $[2, 3]$ . There is also a negative real root on the interval  $[-3, -2]$ . Again note that the exact number and location of the real roots will take some additional investigation.

**Example 1.2.2.** Consider the polynomial in Example 1.1.2

$$\begin{aligned} f(x) &= x^5 - x^4 + 3x^3 + 9x^2 - x + 5, \\ f'(x) &= 5x^4 - 4x^3 + 9x^2 + 18x - 1, \\ f''(x) &= 20x^3 - 12x^2 + 18x + 18, \\ f^{(3)}(x) &= 60x^2 - 24x + 18, \\ f^{(4)}(x) &= 120x - 24, \\ f^{(5)}(x) &= 120. \end{aligned}$$

Hence, we get Table 1.2.

Table 1.2: Budan-Fourier Sign Variations

$x$	$f(x)$	$f'(x)$	$f''(x)$	$f^{(3)}(x)$	$f^{(4)}(x)$	$f^{(5)}(x)$	Variations
$-\infty$	—	+	—	+	—	+	5
$-3$	—	+	—	+	—	+	5
$-2$	—	+	—	+	—	+	5
$-1$	+	—	—	+	—	+	4
$0$	+	—	+	+	—	+	4
$1$	+	+	+	+	+	+	0
$\infty$	+	+	+	+	+	+	0

By Descartes's Rule of Signs there are either 4, 2, or no positive real roots and one negative real root. From Table 1.2 we get,

$$\begin{aligned}
 v(-3) - v(-2) - 2m &= 5 - 5 - 2m = 0, \\
 v(-2) - v(-1) - 2m &= 5 - 4 - 2m = 1, \\
 v(-1) - v(0) - 2m &= 4 - 4 - 2m = 0, \\
 v(0) - v(1) - 2m &= 4 - 0 - 2m = 4 - 2m.
 \end{aligned}$$

Since  $f^{(i)}(1) \geq 0$  for  $i = 0, 1, \dots, 5$ , the polynomial has no real roots in the interval  $[1, \infty)$ . By the Budan-Fourier theorem, any of the possible 4, 2, or no positive real roots are on the interval  $[0, 1]$ . Since  $v(a) - v(b) - 2m = 0$  for all  $a, b \in (-\infty, -3]$ , the Budan-Fourier theorem also shows that there are no real roots on the interval  $(-\infty, -3]$ . However, there is a negative real root on the interval  $[-2, -1]$ . Again note that the exact number and location of the real roots will take some additional investigation.

**Example 1.2.3.** Consider the polynomial in Example 1.1.3

$$\begin{aligned}
 f(x) &= x^6 - x^4 + 2x^2 - 3x - 1, \\
 f'(x) &= 6x^5 - 4x^3 + 4x - 3, \\
 f''(x) &= 30x^4 - 12x^2 + 4, \\
 f^{(3)}(x) &= 120x^3 - 24x, \\
 f^{(4)}(x) &= 360x^2 - 24, \\
 f^{(5)}(x) &= 720x, \\
 f^{(6)}(x) &= 720.
 \end{aligned}$$

Hence, we get Table 1.3.

Table 1.3: Budan-Fourier Sign Variations

$x$	$f(x)$	$f'(x)$	$f''(x)$	$f^{(3)}(x)$	$f^{(4)}(x)$	$f^{(5)}(x)$	$f^{(6)}(x)$	Variations
$-\infty$	+	-	+	-	+	-	+	6
-2	+	-	+	-	+	-	+	6
-1	+	-	+	-	+	-	+	6
0	-	-	+	+	-	+	+	3
1	-	+	+	+	+	+	+	1
2	+	+	+	+	+	+	+	0
$\infty$	+	+	+	+	+	+	+	0

By Descartes's Rule of Signs there are either 3 or 1 positive real roots and either 3 or 1 negative roots. From Table 1.3 we get,

$$v(-2) - v(-1) - 2m = 6 - 6 - 2m = 0 \quad (1.1)$$

$$v(-1) - v(0) - 2m = 6 - 3 - 2m = 3 - 2m \quad (1.2)$$

$$v(0) - v(1) - 2m = 3 - 1 - 2m = 2 - 2m, \quad (1.3)$$

$$v(1) - v(2) - 2m = 1 - 0 - 2m = 1. \quad (1.4)$$

Using the Budan-Fourier theorem, we can determine where these real roots are located. The Budan-Fourier theorem shows that there are no real roots on the intervals  $(-\infty, -1]$  since  $v(a) - v(b) - 2m = 0$  for all  $a, b \in (-\infty, -1]$ . Since,  $f^{(i)}(1) \geq 0$  for  $i = 0, 1, \dots, 6$  the polynomial also has no real roots on the interval  $[2, \infty)$ . The Budan-Fourier theorem also shows that

- equation (1.2) has either 3 or 1 real roots on the interval  $[-1, 0]$ ,
- equation (1.3) has either 2 or no real roots on the interval  $[0, 1]$ , and
- equation (1.4) has exactly one real root on the interval  $[1, 2]$ .

Note that in order to determine the exact number and location of the real roots, some additional investigations must be done.

**Exercise 1.2.1.** In Exercise 1.1.3 you determined the number of possible positive and negative roots for the following polynomials using Descartes's Rule of Signs. Now locate where these possible real roots are located on the number line.

1.  $f(x) = x^4 - 2x^3 + 4x^2 - 3x + 1.$

2.  $p(x) = 2x^3 + 5x^2 + x + 1.$

3.  $f(x) = 3x^4 + 10x^2 + 5x - 4.$

**Exercise 1.2.2.** Show that Descartes's Rule of Signs is a special case of Budan-Fourier's Theorem.

**Exercise 1.2.3.** The behavior of the roots of nonzero polynomials at  $x = \infty$  and  $x = -\infty$  was mentioned earlier. Show that  $v(\infty)$  is the sign of the coefficient of the highest degree term. Interpret  $v(-\infty)$  in a similar way.

## 1.3 Euclidean Algorithm

In Section 1.4 we will be discussing *Sturm's Algorithm* to determine the exact number of real roots of a polynomial on an interval. We will be assuming that this polynomial has no repeated roots on that interval. This method is based on a modification of the *Euclidean Algorithm*, which we shall explain in this section. In addition, to apply Sturm's Algorithm in practice, one will need to know if a polynomial has repeated roots and, if so, replace it with a polynomial which has the same roots but with no multiple roots. The Euclidean Algorithm is the key tool to do this as well.

There are some instances where integers and polynomials behave similarly. We will be discussing in this section the similarity of the Euclidean Algorithm for integers and polynomials. Before we dive too far into what the Euclidean Algorithm is and how it works, let us discuss the purpose of the algorithm. The Euclidean Algorithm is an efficient algorithm used to determine the greatest common divisor (GCD). Determining the greatest common divisor is used throughout elementary mathematics. Some simple examples include reducing fractions to simplest form and determining if two numbers are relatively prime.



**Definition 1.3.1.** (C.f. [Mig92]) Let  $a$  and  $b$  be two rational integers. Then a positive rational number  $d$  is called the *greatest common divisor* of  $a$  and  $b$  if

1.  $d$  divides  $a$  and  $b$ , and
2. any common divisor of  $a$  and  $b$  also divides  $d$ .

This integer  $d$  is denoted by  $\gcd(a, b)$ , or sometimes simply by  $(a, b)$ , when no confusion is possible. Moreover, there exists two rational integers  $u$  and  $v$  such that

$$d = ua + vb.$$

**Proposition 1.3.2.** (C.f. [Mig92]) Let  $a$  and  $b$  be two integers. Then  $a$  and  $b$  are *relatively prime* if, and only if, there are two integers  $u$  and  $v$  such that there is the relation

$$ua + vb = 1$$

In other words, the  $\gcd(a, b) = 1$ .

The Euclidean Algorithm uses a series of steps to determine the greatest common divisor of two integers  $a$  and  $b$ . Let us begin with a simple example.

**Example 1.3.1.** (C.f.[GG03], p.43) The Euclidean Algorithm for the two integers 126 and 35 works as follows:

$$\begin{aligned} 126 &= 3 \cdot 35 + 21, \\ 35 &= 1 \cdot 21 + 14, \\ 21 &= 1 \cdot 14 + 7, \\ 14 &= 2 \cdot 7, \end{aligned}$$

and 7 is the greatest common divisor of 126 and 35. So, if we wanted to reduce the fraction  $35/126$  to simplest form, it would be simplified to  $5/18$ .

ALGORITHM: Traditional Euclidean Algorithm (c.f. [GG03], p.45)

---

Input:  $f, g$  rational integers.

Output: A greatest common divisor of  $f$  and  $g$ .

```

1:  $r_0 \leftarrow f, r_1 \leftarrow g$ 
2:  $i \leftarrow 1$ 
3: while  $r_i \neq 0$  do
4:    $r_{i+1} \leftarrow r_{i-1} \bmod r_i$ 
5:    $i \leftarrow i + 1$ 
6: end while
7: return  $r_{i-1}$ 

```

---

**Exercise 1.3.1.** Every algorithm has an associated cost. The cost of an algorithm is defined to be the number of steps the algorithm requires multiplied by the number of additions and multiplications that occur at each step. If the Euclidean algorithm requires  $n$  steps for a pair of natural numbers  $a > b > 0$ , the smallest values of  $a$  and  $b$  for which this is true are the Fibonacci numbers  $F_{n+1}$  and  $F_n$ , respectively. Determine the cost of the Euclidean algorithm for  $n$  steps. Relate the cost of determining the greatest common divisor of two Fibonacci numbers to the cost of the Euclidean algorithm. What makes computing the GCD of two Fibonacci numbers significant? For a reference use Maurice Mignotte's *Mathematics for Computer Algebra*.

Definition 1.3.1 and Proposition 1.3.2 can be extended to include polynomials. The Euclidean algorithm can also be adapted to work for polynomials. For future reference we will be focusing on the polynomials with real coefficients,  $\mathbb{R}[x]$ .

**Example 1.3.2.** (C.f. [GG03]) Use the Euclidean Algorithm to determine the GCD for the two real coefficient polynomials  $f = 18x^3 - 42x^2 + 30x - 6$  and  $g = -12x^2 + 10x - 2$ . Using the traditional Euclidean algorithm we get

$$\begin{aligned} r_0 &= 18x^3 - 42x^2 + 30x - 6, \\ r_1 &= -12x^2 + 10x - 2, \\ r_2 &= r_1 \text{ rem } r_0 = \frac{9}{2}x - \frac{3}{2}, \\ r_3 &= r_2 \text{ rem } r_1 = 0. \end{aligned}$$

Since  $r_3 = 0$  the  $\gcd(f, g) = \frac{9}{2}x - \frac{3}{2}$ .

The extension the traditional Euclidean algorithm computes not only the GCD but also a representation of it as a linear combination of inputs. It generalizes the representation

$$7 = 21 - 1 \cdot 14 = 21 - (35 - 1 \cdot 21) = 2 \cdot (126 - 3 \cdot 35) - 35 = 2 \cdot 126 - 7 \cdot 35,$$

which is obtained by reading the lines of Example 1.3.1 from the bottom up. This method is called the **Extended Euclidean Algorithm**.

ALGORITHM: Traditional Extended Euclidean Algorithm ([GG03], p.46)

Input:  $f, g$  rational integers.

Output: A greatest common divisor of  $f$  and  $g$ .

```

1:  $r_0 \leftarrow f, r_1 \leftarrow g$ 
2:  $s_0 \leftarrow 1, s_1 \leftarrow 0$ 
3:  $t_0 \leftarrow 0, t_1 \leftarrow 1$ 
4:  $i \leftarrow 1$ 
5: while  $r_i \neq 0$  do
6:    $q_i \leftarrow r_{i-1} \text{ quo } r_i$ 
7:    $r_{i+1} \leftarrow r_{i-1} - q_i r_i$ 
8:    $s_{i+1} \leftarrow s_{i-1} - q_i s_i$ 
9:    $t_{i+1} \leftarrow t_{i-1} - q_i t_i$ 
10:   $i \leftarrow i + 1$ 
11: end while
```

12:  $\ell \leftarrow i - 1$

13: **return**  $\ell, r_i, s_i, t_i$  for  $0 \leq i \leq \ell + 1$ , and  $q_i$  for  $1 \leq i \leq \ell$

---

So, if the quotients are denoted by  $q_i(x)$  and the remainders by  $r_i(x)$  we have the following identities given by the Extended Euclidean Algorithm where  $r_0 = f(x)$  and  $r_1 = g(x)$ .

$$\begin{aligned}
 r_0(x) &= q_1(x)r_1(x) + r_2(x), \\
 r_1(x) &= q_2(x)r_2(x) + r_3(x), \\
 r_2(x) &= q_3(x)r_3(x) + r_4(x), \\
 &\vdots \\
 r_{n-2}(x) &= q_{n-1}(x)r_{n-1}(x) + r_n(x), \\
 r_{n-1}(x) &= q_n(x)r_n(x) + r_{n+1}(x).
 \end{aligned}$$

**Example 1.3.3.** (C.f. [GG03]) As in Example 1.3.1, let  $f = 126$  and  $g = 35$ . The following table illustrates the computation of the Extended Euclidean Algorithm for integers.

$i$	$q_i$	$r_i$	$s_i$	$t_i$
0		126	1	0
1	3	35	0	1
2	1	21	1	-3
3	1	14	-1	4
4	2	7	2	-7
5		0	-5	18

We can now read off row 4 that  $\gcd(126, 35) = 7 = 2 \cdot 126 + (-7) \cdot 35$ .

**Example 1.3.4.** (C.f. [GG03]) As in Example 1.3.2, let  $f = 18x^3 - 42x^2 + 30x - 6$  and  $g = -12x^2 + 10x - 2$ . Using the traditional Extended Euclidean algorithm we get

$i$	$q_i$	$r_i$	$s_i$	$t_i$
0		$18x^3 - 42x^2 + 30x - 6$	1	0
1	$-\frac{3}{2}x + \frac{9}{4}$	$-12x^2 + 10x - 2$	0	1
2	$-\frac{8}{3}x + \frac{4}{3}$	$\frac{9}{2}x - \frac{3}{2}$	1	$\frac{3}{2}x - \frac{9}{4}$
3		0	$\frac{8}{3}x - \frac{4}{3}$	$4x^2 - 8x + 4$

By reading off row 2 we can now find that a gcd of  $f$  and  $g$  is

$$\frac{9}{2}x - \frac{3}{2} = 1 \cdot (18x^3 - 42x^2 + 30x - 6) + \left(\frac{3}{2}x - \frac{9}{4}\right)(-12x^2 + 10x - 2).$$

**Exercise 1.3.2.** Determine the greatest common divisor for the following pairs of polynomials.

1.  $f(x) = x^5 - 4x^4 + 4x^3 + 2x^2 - 5x + 2$ ,  $g(x) = 20x^3 - 48x^2 + 24x + 4$ .
2.  $f(x) = 7x^5 - 30x^4 + 38x^3 - 4x^2 - 21x + 10$ ,  $g(x) = 7x^3 + 5x^2 - 28x - 20$ .
3.  $f(x) = 3x^4 - 4x^3 - 6x^2 + 12x - 5$ ,  $g(x) = 3x^2 + 12x + 12$ .

The next topic to be discussed is a method to replace a given polynomial with one having only simple roots. A simple root is defined to be a root with degree equal to 1. In general we have been focused on polynomials with real coefficients. Let  $F$  be a nonzero polynomial of this kind. We say that  $F$  is *squarefree* if  $F$  has only simple roots. In other words, the multiplicity of the roots of  $F$  is one. We do not need to know the roots of  $F$  to determine if it is squarefree. Let  $F'$  be the first derivative of  $F$  and let  $Q$  be the monic greatest common divisor of  $F$  and  $F'$ . For  $Q$  to be monic the leading coefficient is one. If  $\gcd(F, F') = 1$ ,

$$uF + vF' = 1,$$

then  $F$  and  $F'$  have no common roots, thus  $F$  must have only simple roots. Let's begin with a simple example of a polynomial  $F$  that is not squarefree.

**Example 1.3.5.** Define a polynomial of real coefficients that is not squarefree to be

$$F = (x - 2)^2(x + 1)(x - 1)^3.$$

Then by the chain rule

$$\begin{aligned} F' &= 2(x - 2)(x + 1)(x - 1)^3 + (x - 2)^2(x - 1)^3 + 3(x - 2)^2(x + 1)(x - 1)^2 \\ &= (x - 2)(x - 1)^2[2(x + 1)(x - 1) + (x - 2)(x - 1) + 3(x - 2)(x + 1)]. \end{aligned}$$

It follows that the greatest common divisor of  $F$  and  $F'$  is

$$Q = \gcd(F, F') = (x - 2)(x - 1)^2.$$

However, dividing  $F$  by  $Q$ ,

$$\tilde{F} = F/Q = (x-2)(x+1)(x-1) = x^3 - 2x^2 - x + 2,$$

we get a squarefree polynomial since for

$$\tilde{F}' = 3x^2 - 4x - 1$$

the greatest common divisor is

$$\tilde{Q} = \gcd(\tilde{F}, \tilde{F}') = 1.$$

We can generalize the idea exhibited in these last examples. Let

$$F = \lambda P_1^{e_1} \cdots P_l^{e_l},$$

where  $\lambda$  is a constant, the  $P_i$ 's are monic polynomials and the  $e_i \geq 1$ . We then have that

$$F' = \lambda P_1^{e_1-1} \cdots P_l^{e_l-1} R,$$

where  $R = \sum_{i=1}^k e_i P_1 \cdots P_{i-1} P_i' P_{i+1} \cdots P_k$ . Since  $P_i$  does not divide  $P_i'$ , these formulas imply that the

$$\gcd(F, F') = P_1^{e_1-1} \cdots P_l^{e_l-1}.$$

Therefore,  $G = F/\gcd(F, F') = P_1 \cdots P_k$ . From these considerations, one sees the following (c.f. [Mig92], p.242):

**Proposition 1.3.3.** Let  $F$  be a nonconstant polynomial with coefficients in a field containing the rational numbers. Then  $F$  is squarefree if, and only if,  $F$  is relatively prime to its derivative  $F'$ . Moreover, the polynomial  $G = F/\gcd(F, F')$  is always squarefree, divides  $F$  and has the same roots as  $F$ .

**Exercise 1.3.3.** Use Maple to determine if the following polynomials are squarefree. If not, determine the associated squarefree polynomial.

1.  $f(x) = x^3 - 3x^2 - 4x + 13$ .
2.  $f(x) = x^5 - 9x^4 + 30x^3 - 46x^2 + 33x - 9$ .
3.  $f(x) = x^6 - 3x^5 - 7x^4 + 19x^3 + 18x^2 - 28x - 24$ .
4.  $f(x) = x^5 - x^4 + 3x^3 + 9x^2 - x + 5$ .

## 1.4 Sturm's Theorem

Before stating Sturm's theorem, we will first explain how to obtain the functions used by the theorem.

Let  $f(x)$  be a polynomial with real coefficients and  $f_1(x)$  its first derivative. We will modify the Extended Euclidean Algorithm mentioned in Section 1.3 by exhibiting each remainder as the negative of a polynomial. In other words, the remainder resulting from the division of  $f(x)$  by  $f_1(x)$  is  $-f_2(x)$ ; the remainder resulting from the division of  $f_1(x)$  by  $f_2(x)$  is  $-f_3(x)$ ; and so on. Continue the procedure until a constant remainder  $-f_n$  is obtained. So, we have

$$\begin{aligned} f(x) &= q_1(x)f_1(x) - f_2(x), \\ f_1(x) &= q_2(x)f_2(x) - f_3(x), \\ f_2(x) &= q_3(x)f_3(x) - f_4(x), \\ &\vdots \\ f_{n-2}(x) &= q_{n-1}(x)f_{n-1}(x) - f_n, \quad \text{where } f_n = \text{constant}. \end{aligned}$$

Note that  $f_n \neq 0$  if the polynomial  $f(x)$  does not have a multiple roots. The sequence of polynomials

$$f(x), f_1(x), f_2(x), \dots, f_{n-1}(x), f_n, \tag{1.5}$$

is called the *Sturm sequence*. The  $f_i$  are referred to as the *Sturm functions*.

If  $x = c$  where  $c \in \mathbb{R}$ , then the number of variations of signs of sequence (1.5) is denoted by  $v(c)$ . Note that this  $v(c)$  is different than the one used for the Budan-Fourier theorem. Be careful not to confuse the two notations. Any terms of the sequence that become 0 are dropped before counting the variations in signs.

In practice the computation needed to obtain Sturm functions is costly, especially if  $f(x)$  is of high degree. However, once the functions are found, it is easy to locate the real roots of the equation  $f(x) = 0$  by the Sturm's theorem (c.f. [Con57]).

**Theorem 1.4.1.** (*Sturm's Theorem*) *Let  $f(x) = 0$  be an algebraic equation with real coefficients and without multiple roots. If  $a$  and  $b$  are real numbers,  $a < b$ , and neither a root of the given equation, then the number of real roots of  $f(x) = 0$  between  $a$  and  $b$  is equal to  $v(a) - v(b)$  (c.f. [Con57]).*

Notice that Sturm's theorem is limited to polynomials that do not have multiple roots. In other words the polynomials are squarefree (see Section 1.3). In the real world, it cannot be assumed that every polynomial we encounter will be squarefree. Even though it is very unlikely to have a polynomial that is not squarefree we must address the case. Given that a polynomial  $f$  with real coefficients has multiple roots, we must first perform the scheme in Section 1.3 to obtain a squarefree polynomial  $\tilde{f}$ . The resulting polynomial will have the same roots as the original polynomial but its roots are now simple.

The *proof of Sturm's Theorem* can be found in the Appendix B. So,  $v(a) - v(b)$  equals precisely the number of roots between  $a$  and  $b$ . In order to gain a better understanding of the theorem we will expand on previous examples.

**Example 1.4.1.** (C.f. [Con57], p.88) Consider the polynomial in Examples 1.1.1 and 1.2.1

$$f(x) = x^3 - 3x^2 - 4x + 13.$$

Use Sturm's theorem to isolate the roots of the equation  $f(x) = 0$ . We will be using **Maple** to calculate the remainders. The first derivative of  $f(x)$  is

$$f_1(x) = 3x^2 - 6x - 4.$$

Using **Maple**, we can determine  $f_2(x)$  which is the negative remainder of  $f(x)$  divided by  $f_1(x)$ .

**Maple Code:**

```
> r := rem(f, f_1, x, 'q');
```

$$r := -\frac{14}{3}x + \frac{35}{3}$$

```
> f_2 := -r;
```

$$f_2 := \frac{14}{3}x - \frac{35}{3}$$

```
> q;
```

$$q := \frac{1}{3}x - \frac{1}{3}$$

Given that  $q$  is the quotient of  $f$  divided by  $f_1$  and  $r$  is the remainder, we obtain

$$x^3 - 3x^2 - 4x + 13 = \left(\frac{1}{3}x - \frac{1}{3}\right)(3x^2 - 6x - 4) + \left(-\frac{14}{3}x + \frac{35}{3}\right).$$



Again using **Maple**, we can determine  $f_3(x)$  which is the negative remainder of  $f_1(x)$  divided by  $f_2(x)$ .

**Maple Code:**

```
> r:= rem(f_1,f_2,x,'q');
```

$$r := -\frac{1}{4}$$

```
> f_3 := -r;
```

$$f_3 := \frac{1}{4}$$

```
> q;
```

$$q := \frac{9}{14}x + \frac{9}{28}$$

Given that  $q$  is the quotient of  $f_1$  divided by  $f_2$  and  $r$  is the remainder, we obtain

$$3x^2 - 6x - 4 = \left(\frac{9}{14}x + \frac{9}{28}\right) \left(\frac{14}{3}x - \frac{35}{3}\right) - \frac{1}{4}.$$

Therefore, the list of Sturm functions is as follows

$$\begin{aligned} f &= x^3 - 3x^2 - 4x + 13, \\ f_1 &= 3x^2 - 6x - 4, \\ f_2 &= \frac{14}{3}x - \frac{35}{3}, \\ f_3 &= \frac{1}{4}. \end{aligned}$$

The signs of Sturm functions for selected  $x$  are in Table 1.4. We can see from Table 1.4

Table 1.4: Sturm Function Sign Variations

$x$	$f$	$f_1$	$f_2$	$f_3$	Variations
$-\infty$	-	+	-	+	3
-3	-	+	-	+	3
-2	+	+	-	+	2
0	+	-	-	+	2
2	+	-	-	+	2
3	+	+	+	+	0
$\infty$	+	+	+	+	0

that

$$\begin{aligned}v(-3) - v(-2) &= 3 - 2 = 1 \\v(2) - v(3) &= 2 - 0 = 2\end{aligned}$$

and the other combinations of  $v(a) - v(b)$  are equal to zero. Therefore, by Sturm's Theorem we know there is a single root between  $-3$  and  $-2$  and that there are exactly two roots between  $2$  and  $3$ .

**Example 1.4.2.** Consider the polynomial in Examples 1.1.2 and 1.2.2

$$f(x) = x^5 - x^4 + 3x^3 + 9x^2 - x + 5.$$

Use Sturm's theorem to isolate the roots of the equation  $f(x) = 0$ . We will be using **Maple** to calculate the remainders. The first derivative of  $f(x)$  is

$$f_1(x) = 5x^4 - 4x^3 + 9x^2 + 18x - 1.$$

Using **Maple** we will determine the remainder  $r$  and quotient  $q$  of  $f_i$  and  $f_{i+1}$  and assigning  $-r$  to  $f_{i+2}$ . We will denote  $f_0$  as  $f$  in all code. Dividing  $f(x)$  by  $f_1(x)$ , we get

**Maple Code:**

```
> r:= rem(f,f_1,x,'q');
```

$$r := \frac{26}{25}x^3 + \frac{144}{25}x^2 - \frac{2}{25}x + \frac{124}{25}$$

```
> f_2 := -r;
```

$$f_2 := -\frac{26}{25}x^3 - \frac{144}{25}x^2 + \frac{2}{25}x - \frac{124}{25}$$

```
> q;
```

$$q := \frac{1}{5}x - \frac{1}{25}$$

Given that  $q$  is the quotient of  $f$  divided by  $f_1$  and  $r$  is the remainder, we obtain

$$x^5 - x^4 + 3x^3 + 9x^2 - x + 5 = \left(\frac{1}{5}x - \frac{1}{25}\right) f_1(x) + \left(\frac{26}{25}x^3 + \frac{144}{25}x^2 - \frac{2}{25}x + \frac{124}{25}\right).$$

Now, divide  $f_1(x)$  by  $f_2(x)$  to get

**Maple Code:**

```
> r:= rem(f_1,f_2,x,'q');
```

$$r := \frac{25375}{169} + \frac{31250}{169}x^2 - \frac{1400}{169}x$$

```
> f_3 := -r;
```

$$f_3 := -\frac{25375}{169} - \frac{31250}{169}x^2 + \frac{1400}{169}x$$

```
> q;
```

$$q := -\frac{125}{26}x + \frac{5150}{169}$$

Given that  $q$  is the quotient of  $f_1$  divided by  $f_2$  and  $r$  is the remainder, we obtain

$$5x^4 - 4x^3 + 9x^2 + 18x - 1 = \left(-\frac{125}{26}x + \frac{5150}{169}\right)f_2(x) + \left(\frac{31250}{169}x^2 - \frac{1400}{169}x + \frac{25375}{169}\right).$$

Now, divide  $f_2(x)$  by  $f_3(x)$  to get

**Maple Code:**

```
> r:= rem(f_2,f_3,x,'q');
```

$$r := \frac{6487741}{9765625}x - \frac{478608}{1953125}$$

```
> f_4 := -r;
```

$$f_4 := -\frac{6487741}{9765625}x + \frac{478608}{1953125}$$

```
> q;
```

$$q := \frac{2197}{390625}x + \frac{7666516}{244140625}$$

Given that  $q$  is the quotient of  $f_2$  divided by  $f_3$  and  $r$  is the remainder, we obtain

$$-\frac{26}{25}x^3 - \frac{144}{25}x^2 + \frac{2}{25}x - \frac{124}{25} = \left(\frac{2197}{390625}x + \frac{7666516}{244140625}\right)f_3(x) + \left(\frac{6487741}{9765625}x - \frac{478608}{1953125}\right).$$

Lastly, divide  $f_3(x)$  by  $f_4(x)$  to get

**Maple Code:**

```
> r := rem(f_3, f_4, x, 'q');
```

$$r := -\frac{42900302734375}{249057889249}$$

```
> f_5 := -r;
```

$$f_5 := \frac{42900302734375}{249057889249}$$

```
> q;
```

$$q := \frac{305175781250}{1096428229}x + \frac{3796439453125000}{42090783283081}$$

Given that  $q$  is the quotient of  $f_3$  divided by  $f_4$  and  $r$  is the remainder, we obtain

$$-\frac{31250}{169}x^2 + \frac{1400}{169}x - \frac{25375}{169} = \left( -\frac{305175781250}{1096428229}x - \frac{3796439453125000}{42090783283081} \right) f_3(x) + \frac{42900302734375}{249057889249}.$$

Therefore, the list of Sturm functions is as follows

$$\begin{aligned} f &= x^5 - x^4 + 3x^3 + 9x^2 - x + 5, \\ f_1 &= 5x^4 - 4x^3 + 9x^2 + 18x - 1, \\ f_2 &= -\frac{26}{25}x^3 - \frac{144}{25}x^2 + \frac{2}{25}x - \frac{124}{25}, \\ f_3 &= -\frac{31250}{169}x^2 + \frac{1400}{169}x - \frac{25375}{169}, \\ f_4 &= -\frac{6487741}{9765625}x + \frac{478608}{1953125}, \\ f_5 &= \frac{42900302734375}{249057889249}. \end{aligned}$$

The signs of the Sturm functions for select significant values of  $x$  are in Table 1.5. We can see from Table 1.5 that

$$v(-2) - v(-1) = 3 - 2 = 1$$

and the other combinations of  $v(a) - v(b)$  are equal to zero. Therefore, by Sturm's Theorem we know there is a single root between  $-2$  and  $-1$  and no others.

**Maple** was used in all of the examples to compute Sturm functions. Depending on the degree of the polynomial we are given, this can be very tedious. There are calls in **Maple** that will generate the Sturm functions and another which will give the number of real roots of a polynomial on an interval.

Table 1.5: Sturm Function Sign Variations

$x$	$f$	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	Variations
$-\infty$	-	+	+	-	+	+	3
-3	-	+	-	-	+	+	3
-2	-	+	-	-	+	+	3
-1	+	-	-	-	+	+	2
0	+	-	-	-	+	+	2
1	+	+	-	-	-	+	2
2	+	+	-	-	-	+	2
3	+	+	-	-	-	+	2
$\infty$	+	+	-	-	-	+	2

- **sturm** - number of real roots of a polynomial in an interval
- **sturmseq** - Sturm sequence of a polynomial

#### Calling Sequence

```
> sturmseq(p,x);
> sturm(s,x,a,b);
```

#### Parameters

- $p$  - polynomial in  $x$  with rational or float coefficients
- $x$  - variable in polynomial  $p$
- $a, b$  - rationals or floats such that  $a \leq b$ ;  $a$  can be  $-\infty$  and  $b$  can be  $\infty$
- $s$  - Sturm sequence for polynomial  $p$

**Note:** The interval excludes the lower endpoint  $a$  and includes the upper endpoint  $b$  (unless it is  $\infty$ ). This is different from how the Sturm theorem is stated in Theorem 1.3.

**Exercise 1.4.1.** Use the call sequences in **Maple** mentioned above to generate Sturm functions for Examples 1.3.1, 1.3.2, and 1.3.3. How do the **Maple** generated Sturm functions differ from those in the examples? Does it change the intervals where the real roots can be found?

**Exercise 1.4.2.** Create a procedure in **Maple** that will output Sturm functions given a polynomial with real coefficients. Check your Sturm functions with those generated by the call sequence in **Maple**.

**Exercise 1.4.3.** In Exercise 1.2.1 you located the possible real roots on the number line using the Budan-Fourier theorem. Now, isolate the real roots of each of the following equations by means of Sturm's theorem:

1.  $f(x) = x^4 - 2x^3 + 4x^2 - 3x + 1.$

2.  $p(x) = 2x^3 + 5x^2 + x + 1.$

3.  $f(x) = 3x^4 + 10x^2 + 5x - 4.$

## Chapter 2

# Numerical Approximations of Roots

In most real world scenarios the real roots of polynomials are not usually integers nor easily determined. The quadratic formula is used to determine the roots of quadratic polynomials. There also exists formulas for cubic and fourth degree polynomials but are not commonly used. We have been able to estimate the location of the real roots of polynomials with real coefficients by using Descartes' Rule of Signs, the Budan-Fourier theorem and Sturm's theorem. After isolating the real roots there are multiple methods of numerically approximating the values of the roots up to a predetermined precision.

### 2.1 Horner's Method

Once the root has been isolated, the Horner's method can be used to compute the real root to any desired number of decimal places. The idea behind Horner's method is as follows. Assume that a root has been isolated using one of the previous methods such that there is only one root per interval. Suppose the root was determined to be in the interval  $[a, b]$  where  $a$  is an integer and  $b - a < 1$ . Then we know that the root will at least begin with the integer  $a$ . Next transform the original polynomial so as to diminish the root by  $a$  thus obtaining a transformed polynomial. What this means is that instead of the root being on the interval  $[a, b]$  is transformed to the interval  $[0, b - a]$ . Using this new transformed polynomial determine where the root is located on the interval  $[0, b - a]$ . Suppose that the

root is now on the interval  $[0.c, 0.d]$  where  $d - c < 1$ . Transform this polynomial so as to diminish the root by  $0.c$  so that the root is on the interval  $[0, 0.1]$ . We now have the first two values for the root approximation *a.c.* Continue in this process until the root has been approximated up to the desired precision. After some point there are some abbreviated steps that will be used.

We will illustrate Horner's method by computing the roots of the equation

$$f(x) \equiv x^3 - 3x^2 - 4x + 13 = 0 \quad (2.1)$$

which was first seen in Examples 1.1.1, 1.2.1, and 1.4.1. From previous examples, there are two roots between 2 and 3 and one root between  $-3$  and  $-2$ . We'll begin by approximating the roots between 2 and 3. It can be determined that one of the two roots is located in the interval  $[2, 2.5]$  and the other in  $[2.5, 3]$ . We will begin by approximating the root on the interval  $[2, 2.5]$ .

This equation will first be transformed so as to diminish the root by 2 (c.f. [Con57], p.34).

When a root is diminished the following scheme is performed. First divide equation (2.1) by  $x - 2$  to obtain

$$x^3 - 3x^2 - 4x + 13 = (x^2 - x - 6)(x - 2) + 1 \quad (2.2)$$

$$= [(x + 1)(x - 2) - 4] + 1 \quad (2.3)$$

where  $x^2 - x - 6$  is divided by  $x - 2$

$$= ([1(x - 2) + 3](x - 2) - 4) + 1 \quad (2.4)$$

where  $x + 1$  is divided by  $x - 2$

$$= 1 \cdot (x - 2)^3 + 3(x - 2)^2 - 4(x - 2) + 1. \quad (2.5)$$

The transformed equation is

$$x^3 + 3x^2 - 4x + 1 = 0, \quad (2.6)$$

where  $x = (\tilde{x} - 2)$ .



There is a short hand way of doing the same calculations using a table. Let 2 be the amount by which the root is being diminished. First place the coefficients of equation (2.1) underneath its respective term and place 2 underneath  $\tilde{x}$  as in the table below.

$\tilde{x}$	$x^3$	$x^2$	$x^1$	$x^0$
<span style="border: 1px solid black;">2</span>	1	-3	-4	+13

The first step is the divide the original equation (2.1) by the root  $\tilde{x} = 2$ . This is done by first bringing down the first coefficient as can be seen below.

$\tilde{x}$	$x^3$	$x^2$	$x^1$	$x^0$
<span style="border: 1px solid black;">2</span>	1	-3	-4	+13
	<hr/>			
	1			

Next, multiply number 1 by 2 and place it underneath -3. Add -3 and +2 together giving -1 as can be seen in the table below.

$\tilde{x}$	$x^3$	$x^2$	$x^1$	$x^0$
<span style="border: 1px solid black;">2</span>	1	-3	-4	+13
		+2		
	<hr/>			
	1	-1		

Now, multiply recently calculated -1 by 2 and place it beneath -4. Add -4 and -2 together to get -6 as seen below.

$\tilde{x}$	$x^3$	$x^2$	$x^1$	$x^0$
<span style="border: 1px solid black;">2</span>	1	-3	-4	+13
		+2	-2	
	<hr/>			
	1	-1	-6	

Finally, multiply the -6 by 2 and place it underneath +13. Add +13 and -12 together to get +1 as seen below.

$\tilde{x}$	$x^3$	$x^2$	$x^1$	$x^0$
<span style="border: 1px solid black;">2</span>	1	-3	-4	+13
		+2	-2	-12
	<hr/>			
	1	-1	-6	<span style="border: 1px solid black;">+1</span>

Note that the last number, +1, is the remainder of the division. Therefore, the

quotient of  $x^3 - 3x^2 - 4x + 13$  divided by  $x - 2$  is

$$1 \cdot x^2 + (-1) \cdot x + (-6) = x^2 - x - 6$$

with a remainder of  $+1$ . This is the same result as in (2.2).

Now let us compute (2.3) in a similar manner using the tables. We will be dividing the quotient from above by  $\tilde{x} = 2$ . The initial setup of the table is below.

$\tilde{x}$	$x^2$	$x^1$	$x^0$
$\boxed{2}$	1	-1	-6

Following the same process as before we get the table below.

$\tilde{x}$	$x^2$	$x^1$	$x^0$
$\boxed{2}$	1	-1	-6
		+2	+2
	1	+1	$\boxed{-4}$

The last number,  $-4$ , is the remainder of the division of  $x^2 - x - 6$  divided by  $x - 2$  and the quotient is

$$1 \cdot x + 1 = x + 1.$$

This is the same result as in (2.3). Last but not least, divide  $x + 1$  by  $x + 2$  as in (2.4) to get the final result. Following the same process as before we get the table below.

$\tilde{x}$	$x^1$	$x^0$
$\boxed{2}$	1	+1
		+2
	1	$\boxed{+3}$

If each table is created separately, this does not save us much work than the first process we used. However, if we use only one table then the process is shortened.

$\tilde{x}$	$x^3$	$x^2$	$x^1$	$x^0$
<span style="border: 1px solid black;">2</span>	1	-3	-4	+13
		+2	-2	-12
	1	-1	-6	<span style="border: 1px solid black;">+1</span>
		+2	+2	
	1	+1	<span style="border: 1px solid black;">-4</span>	
		+2		
	1	<span style="border: 1px solid black;">+3</span>		

The transformed equation can be obtained by selecting 1 as the coefficient of  $x^3$ , the remainder +3 will be the coefficient of  $x^2$ , -4 the coefficient of  $x$  and +1 will be the constant term. Thus obtaining the same transformed equation as before

$$x^3 + 3x^2 - 4x + 1 = 0,$$

where  $x = (\tilde{x} - 2)$ .

We must now compute the root of equation (2.6) which now lies between 0 and 1. It can be found by trial using Sturm's theorem that the root lies between 0.3 and 0.4. Hence, we transform (2.6) as to diminish the roots by 0.3, and thus obtain the equation

$$x^3 + 3.9x^2 - 1.93x + .097 = 0, \quad (2.7)$$

with a root between 0 and 0.1. Since the terms of higher degree than the first in (2.7) are relatively small when  $0 < x < 0.1$ , we get an approximate value for this root by solving

$$-1.93x + .097 = 0$$

for  $x$ . It thus appears that the root of (2.7) which we want to compute is approximately equal to 0.05. Actual substitution in (2.7) shows that the root is between 0.05 and 0.06. We can therefore diminish the roots of (2.7) by 0.05, and so obtain the equation

$$x^3 + 4.05x^2 - 1.5325x + .010375 = 0 \quad (2.8)$$

with a root between 0 and 0.01. Since the root of

$$-1.5325x + .010375 = 0$$

is approximately 0.006, we can infer that (2.8) has a root between 0.006 and 0.007. This can be verified using substitution. We can again diminish the roots of (2.8) by 0.006 to get

the equation

$$x^3 + 4.068x^2 - 1.483792x + .001326016 = 0 \quad (2.9)$$

with a root between 0 and 0.001. Again, the root can be approximated by determining the root of

$$-1.483792x + .001326016 = 0$$

which is about 0.0008. It is verified by substitution that the root is between 0.0008 and 0.0009. Thus the root of equation (2.1) in the interval  $[2, 2.5]$  is approximately 2.3568, correct to four decimal places. If more decimal places are required, this procedure can be continued until the desired precision is reached. In practice the computation outlined above should be tabulated in the following manner:

The boxed values are the numbers the roots are diminished by and they are also the values for the approximated root. Similarly, the root in the interval  $[2.5, 3]$  can be approximated up to four decimal places. Taking the values of the boxes in Table 2.2 below we get 2.6920. The last value is approximated from

$$x^3 + 5.076x^2 + 1.588592x - .000034122 = 0$$

by determining the root of

$$1.588592x - .000034122 = 0$$

which is approximately  $x \approx .00002$ . In order to find the negative roots of equation (2.1) by Horner's method, we compute the positive roots of  $f(-x) = 0$ , and then change the signs of the roots thus found (c.f. [Con57]) We can compare the roots approximated by Horner's method to those calculated in Maple.

**Maple Code:**

```
> f:= x^3 - 3*x^2 - 4*x + 13;
```

$$f := x^3 - 3x^2 - 4x + 13$$

```
> fsolve(f, x);
```

$$-2.048917340, 2.356895868, 2.692021472$$

The values of the positive roots approximated by Horner's method up to four decimal places agree with those determined by Maple.

Table 2.1: Horner's method on the interval  $[2, 2.5]$ .

$\tilde{x}$	$x^3$	$x^2$	$x^1$	$x^0$	Transformed Equation
<b>2</b>	1	-3	-4	+13	
		+2	-2	-12	
	1	-1	-6	+1	
		+2	+2		
	1	+1	-4		
		+2			
<b>0.3</b>	1	+3	-4	+1	yes
		+0.3	+0.99	-0.903	
	1	+3.3	-3.01	+0.097	
		+0.3	+1.08		
	1	+3.6	-1.93		
		+0.3			
<b>0.05</b>	1	+3.9	-1.93	+0.097	yes
		+0.05	+0.1975	-0.086625	
	1	+3.95	-1.7325	+0.010375	
		+0.5	+0.2		
	1	+4	-1.5325		
		+0.05			
<b>0.006</b>	1	+4.05	-1.5325	+0.010375	yes
		+0.006	+0.024336	-0.009048984	
	1	+4.056	-1.508164	+0.001326016	
		+0.006	+0.024372		
	1	+4.062	-1.483792		
		+0.006			

**Exercise 2.1.1.** Using Horner's method, find the root of the equation in Example 1.4.2 which lies in the interval  $[-2, -1]$ . Approximate the root up to four decimal places and compare it to the answer Maple calculates.

Table 2.2: Horner's method on the interval  $[2.5, 3]$ .

$\tilde{x}$	$x^3$	$x^2$	$x^1$	$x^0$	Transformed Equation
<b>2</b>	1	-3	-4	+13	
		+2	-2	-12	
	1	-1	-6	+1	
		+2	+2		
	1	+1	-4		
		+2			
<b>0.6</b>	1	+3	-4	+1	yes
		+0.6	+2.16	-1.104	
	1	+3.6	-1.84	-0.104	
		+0.6	+2.52		
	1	+4.2	+0.68		
		+0.6			
<b>0.09</b>	1	+4.8	+0.68	-0.104	yes
		+0.09	+0.4401	+0.100809	
	1	+4.89	+1.1201	-0.003191	
		+0.9	+0.4482		
	1	+4.98	+1.5683		
		+0.09			
<b>0.002</b>	1	+5.07	+1.5683	-0.003191	yes
		+0.002	+0.010144	+0.003156888	
	1	+5.072	+1.578444	-0.000034112	
		+0.002	+0.010148		
	1	+5.074	+1.588592		
		+0.002			

## 2.2 Newton's Method

Another method used to approximate roots of a real valued function is Newton's method. Similar to Horner's method, we must first isolate the root of an equation  $f(x) = 0$  to an interval  $[a, b]$  which is known to have one, and only one, root. Newton's method uses tangent lines to approximate the root between  $a$  and  $b$ . The idea is as follows:

1. Make an educated guess which is reasonably close to the actual root (or use Sturm sequences) to find an initial  $a$ .
2. Determine the tangent of the function at that point.

3. Compute the  $x$ -intercept of the tangent line from 2 and that is our next guess.
4. Go to Step 1.

From this idea we can derive a formula that will determine the  $x$ -intercepts of the tangent lines (c.f. [Con57]). Suppose that  $f(x)$  is differentiable on the interval  $[a, b]$ . Let the initial guess of the actual root  $a = x_0$  and let the  $x$ -intercept of the first tangent line be denoted by  $x_1$ . Then the slope of the first tangent line is

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

Therefore,

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

and consequently

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

We can generalize this formula to be

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \quad (2.10)$$

Using equation (2.10) we can compute the root on the interval  $[a, b]$  to any desired degree of accuracy.

We will illustrate Newton's method by computing the roots of the equation

$$f(x) \equiv x^3 - 3x^2 - 4x + 13 = 0. \quad (2.11)$$

The roots of equation (2.11) were also approximated by the Horner's method in a previous section. In Example 1.4.1 we used Sturm's theorem to isolate two positive roots on the interval  $[2, 3]$  and a negative root on the interval  $[-3, -2]$ . In order to use Newton's method we must go a step further and isolate the roots on the interval  $[2, 3]$  to intervals where there is only one root. Using the previously derived Sturm functions we were able to determine that there is one root on the interval  $[2, 2.5]$  and another root on the interval  $[2.5, 3]$ . The points of intersection can be seen in Figure 2.1.

Let us first approximate the root on the interval  $[2, 2.5]$ . We can let our initial guess of the real root on the interval be  $x_0 = 2$ . In order to use equation (2.10), we need to determine the derivative of equation (2.11) which is

$$f'(x) = 3x^2 - 6x - 4.$$

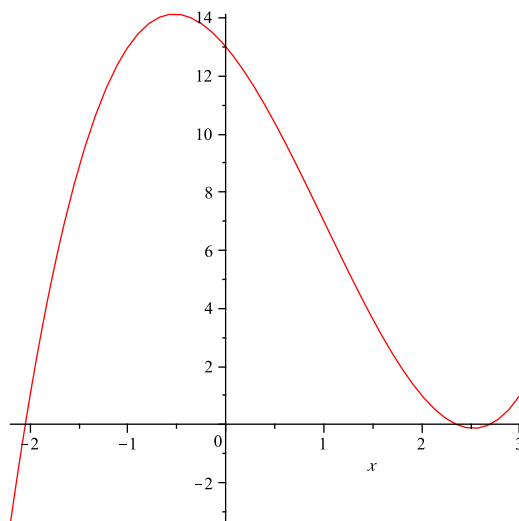


Figure 2.1:  $f(x) = x^3 - 3x^2 - 4x + 13$

Using equation (2.10), the first approximation of the root by Newton's method is

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 2 - \frac{f(2)}{f'(2)} = 2.25.$$

In Figure 2.2(a) the dashed blue line represents the initial guess,  $x_0$ , of the real root and the green line is the tangent of equation (2.11) at  $x_0$ . The brown dashed line in Figure 2.2(a) is the  $x$ -intercept,  $x_1$ , and the next approximation of the root. Figure 2.2(b) shows the next tangent line of equation (2.11) at  $x_1$ . The  $x$ -intercept of the yellow line is  $x_2$  which is also the next approximation of the real root.

Continuing with Newton's method we get the next three approximations for the real root on the interval  $[2, 2.5]$  to be

$$\begin{aligned} x_2 &= 2.25 - \frac{f(2.25)}{f'(2.25)} = 2.337837838, \\ x_3 &= 2.337837838 - \frac{f(2.337837838)}{f'(2.337837838)} = 2.355997619, \\ x_4 &= 2.355997619 - \frac{f(2.355997619)}{f'(2.355997619)} = 2.356893656. \end{aligned}$$



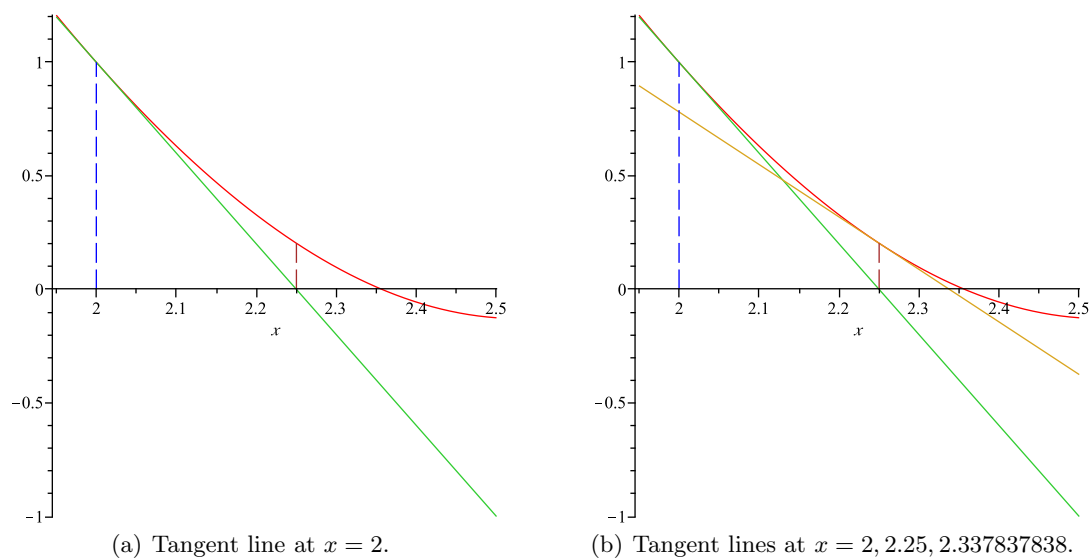


Figure 2.2:  $f(x) = x^3 - 3x^2 - 4x + 13$

Now, let us try to take the same approach for the root on the interval  $[2.5, 3]$ . The root can be seen on the interval in Figure 2.3.

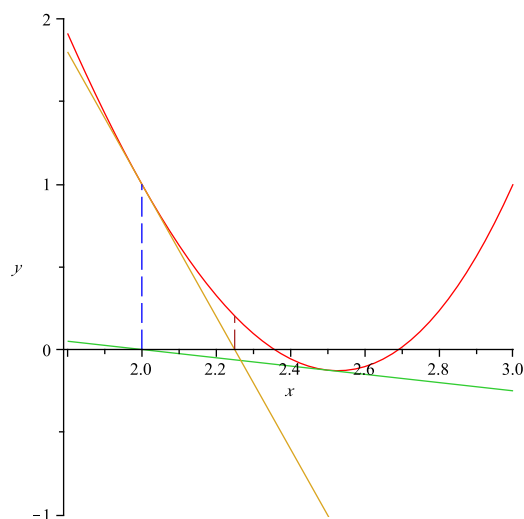


Figure 2.3:  $f(x) = x^3 - 3x^2 - 4x + 13$

Take  $x_0 = 2.5$  as the initial approximation to the root in the application of Newton's method. Then the root is approximated in the following manner:

$$\begin{aligned}x_1 &= 2.5 - \frac{f(2.5)}{f'(2.5)} = 2, \\x_2 &= 2 - \frac{f(2)}{f'(2)} = 2.25, \\x_3 &= 2.25 - \frac{f(2.25)}{f'(2.25)} = 2.337837838, \\x_4 &= 2.337837838 - \frac{f(2.337837838)}{f'(2.337837838)} = 2.355997619, \\x_5 &= 2.355997619 - \frac{f(2.355997619)}{f'(2.355997619)} = 2.356893656.\end{aligned}$$

It is clear we picked a poor initial approximation since we obtained the root in the interval  $[2, 2.5]$  and not the root we were approximating on the interval  $[2.5, 3]$ . This can also be observed by analyzing the first tangent lines in Figure 2.3.

Therefore, a certain amount of care must also be exercised in the employment of the Newton's method. The nature of the situations which must be taken into account is suggested by Figure 2.4.

First, let us recall from calculus the properties of derivatives.

1. Properties of first derivatives:

- At a point where  $f'(x) > 0$ , the slope is positive, so  $f(x)$  is increasing.
- At a point where  $f'(x) < 0$ , the slope is negative, so  $f(x)$  is decreasing.
- At a point where  $f'(x) = 0$ , the slope is zero, so  $f(x)$  has a critical point (maximum or minimum).

2. Properties of second derivatives

- At a point where  $f''(x) > 0$ ,  $f(x)$  is concave up.
- At a point where  $f''(x) < 0$ ,  $f(x)$  is concave down.
- At a point where  $f''(x) = 0$ ,  $f(x)$  has an inflection point (change of concavity).

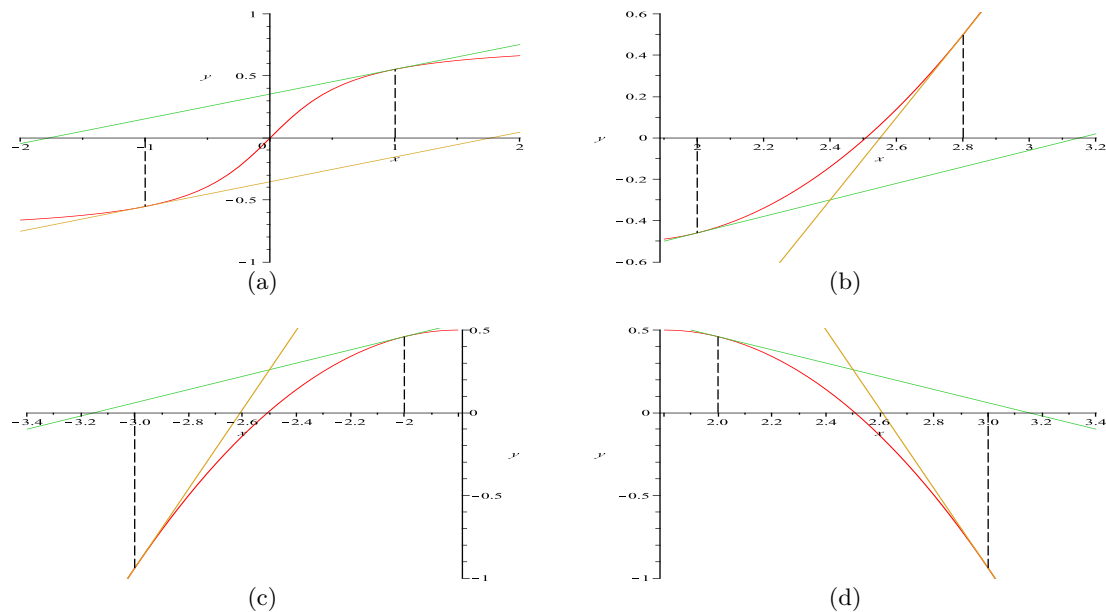


Figure 2.4: Newton Scenarios.

Analysis of figures in Figure 2.4.

- **Figure 2.4(a):** This figure is a plot of  $f(x) = \sqrt[3]{x} = x^{1/3}$ . Note that the first derivative is  $f'(x) = \frac{1}{3}x^{-2/3}$ . Let  $x_0$  be the initial approximation of the root and  $x_1$  be the approximation, then

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = x_0 - \frac{x_0^{1/3}}{1/3x_0^{-2/3}} = x_0 - 3x_0 = -2x_0.$$

As can be seen in the figure, Newton's method gives an over approximation and lands outside the set interval. Therefore, the tangent at the point  $x_0$  is much further than the original approximation. Applying Newton actually doubles the distance from the solution at each iteration.

- **Figure 2.4(b):** The black dotted lines represent the interval  $[a, b]$  on which the root can be found. By analyzing the figure, it can be seen that one tangent line overshoots the root approximation and lands outside of the interval while the other does not. Reviewing the properties of derivatives it can be determined that when the

original function  $f(x)$  and the second derivative  $f''(x)$  have the same sign the tangent line stays within the interval.

$x$	$f(x)$	$f'(x)$	$f''(x)$
$a$	—	+	+
$b$	+	+	+

- **Figure 2.4(c):** Similar to Figure 2.4(b), the tangent line at one end of the interval lands outside of the interval  $[a, b]$  while the other does not. Reviewing the properties of derivatives it can be determined that when the original function  $f(x)$  and the second derivative  $f''(x)$  have the same sign the tangent line stays within the interval.

$x$	$f(x)$	$f'(x)$	$f''(x)$
$a$	—	+	—
$b$	+	+	—

- **Figure 2.4(d):** Similar to Figure 2.4(b), the tangent line at one end of the interval lands outside of the interval  $[a, b]$  while the other does not. By the properties of derivatives it can be determined that when the original function  $f(x)$  and the second derivative  $f''(x)$  have the same sign the tangent line stays within the interval.

$x$	$f(x)$	$f'(x)$	$f''(x)$
$a$	+	—	—
$b$	—	—	—

An analysis of the various possibilities suggested by the figures leads to the following conclusion: *If neither  $f'(x)$  nor  $f''(x)$  vanishes between  $a$  and  $b$ , the root lies between  $\alpha$  and  $\beta - [f(\beta)/f'(\beta)]$ , where  $\beta$  is that one of the numbers  $a, b$  for which  $f(x)$  and  $f''(x)$  have the same sign, and  $\alpha$  is the other end point of the interval.*

There are other cases besides the cubic for which Newton does not work. One is when the derivative at  $x_0$  is zero. This can happen if the initial starting point chosen is a maximum or minimum of the original function. Another problem that can arise is if the starting point causes Newton's method to enter into an infinite cycle, preventing convergence. Sometimes the pitfalls of these examples can be avoided if a different starting point and interval is chosen, however, this might not always be the case.

A different type of problem that may come up involves the derivatives themselves. Problems arise when the polynomial is not continuously differentiable in the neighborhood of the root. In such cases it is possible that Newton's method may always diverge unless the root is guessed on the first attempt. A case was discussed earlier for Figure 2.4(a). In this case the derivative does not exist at the root. Another case is if there is not a second derivative, then convergence may fail. The last case we will concern ourselves with is if the derivative is zero at the root. In this case, the zero derivative implies there are multiple roots. This is handled by using the squarefree method and Sturm's theorem to eliminate any multiplicities.

**Exercise 2.2.1.** Consider the function the following function and initial starting points. Determine if Newton's method works for the case. Why or why not?

1. Let  $f(x) = x^3 - 2x + 2$  and  $x_0 = 0$ .
2. Let  $f(x) = 1 - x^2$  and  $x_0 = 0$ .

Let us quickly analyze the interval  $[2, 2.5]$  before continuing on with the interval  $[2.5, 3]$ . By reviewing Table 2.3,  $f(x)$  and  $f''(x)$  have like signs for  $x = 2$  and opposite signs for  $x = 2.5$ . Therefore, choosing  $x_0 = 2$  as the initial condition was the right decision.

Table 2.3: Derivative Sign Variations

$x$	$f$	$f'$	$f''$
-3	-	+	-
-2	+	+	-
2	+	-	+
2.5	-	-	+
3	+	+	+

However for the interval  $[2.5, 3]$ ,  $f(x)$  and  $f''(x)$  have like signs for  $x = 3$  and opposite signs for  $x = 2.5$ . Hence, we should have picked  $x_0 = 3$  instead of 2.5. Letting

$x_0 = 3$ , the root is approximated in the following manner:

$$\begin{aligned}x_1 &= 3 - \frac{f(3)}{f'(3)} = 2.8, \\x_2 &= 2.8 - \frac{f(2.8)}{f'(2.8)} = 2.714705882, \\x_3 &= 2.714705882 - \frac{f(2.714705882)}{f'(2.714705882)} = 2.693468980, \\x_4 &= 2.693468980 - \frac{f(2.693468980)}{f'(2.693468980)} = 2.692028111.\end{aligned}$$

Now, let us turn to the negative root on the interval  $[-3, -2]$ . We can see from Table 2.3  $f(x)$  and  $f''(x)$  have like signs for  $x = -3$  and opposite signs for  $x = -2$ . Therefore, we let  $x_0 = -3$  and approximate the root to be  $x_4 = -2.048917360$ . We can compare the roots approximated by Newton's method to those calculated in Maple.

**Maple Code:**

```
> f := x^3 - 3*x^2 - 4*x + 13;
```

$$f := x^3 - 3x^2 - 4x + 13$$

```
> fsolve(f, x);
```

$$-2.048917340, 2.356895868, 2.692021472$$

The pseudocode for the algorithm to determine  $x_i$  using Newton's method to approximate the roots is stated below.

ALGORITHM: Newton's Method

---

Input: The function  $f(x)$ , the initial guess  $x_0$ , and the number of iterations to be performed  $n$ .

Output: A vector with the approximated roots  $N$ .

- 1:  $g(x) := f'(x)$
- 2:  $N[1] := x_0$ ;
- 3: **for**  $i = 2, \dots, n$  **do**
- 4:    $x := N[i - 1]$ ;

```
5:    $N[i] := \text{evalf}\left(x - \frac{f(x)}{g(x)}\right)$   
6: end for  
7: return  $N$ ;
```

---

**Exercise 2.2.2.** Implement the pseudocode for Newton's method in Maple.

**Exercise 2.2.3.** In Exercise 1.4.3 the intervals on which the real roots were located were found using Sturm's theorem. Compute the root of each of the following equations using Newton's method on those intervals. Some extra calculation may be needed to isolate the roots to intervals with only one root.

1.  $f(x) = x^4 - 2x^3 + 4x^2 - 3x + 1.$
2.  $f(x) = 2x^3 + 5x^2 + x + 1.$
3.  $f(x) = 3x^4 + 10x^2 + 5x - 4.$

## Chapter 3

# Real World Applications

There are many applications of Sturm's theorem in the real world. We will discuss a few papers that use Sturm's theorem.

### 3.1 Applications of Sturm sequences to bifurcation analysis of delay differential equation models.

Prior to explaining how Sturm sequences are used in this case, let us go over some terminology so the explanation is comprehensible.

- Sturm sequences are the same as Sturm functions or chains.
- Bifurcation is the splitting into two different branches. For example, given a single line graph if it was bifurcated it would split into two different lines at a certain point.
- A delay differential equation is a special type of functional differential equation. Delay differential equations are similar to ordinary differential equations, but their evolution involves past values of the state variable. The solution of delay differential equations therefore requires knowledge of not only the current state, but also of the state a certain time previously.

Forde's paper formalizes a method used by several others in the analysis of biological models involving delay differential equations. In such a model, the characteristic equation about a steady state is transcendental. This paper shows that the analysis of the



bifurcation due to the introduction of the delay term can be reduced to finding whether a related polynomial equation has simple positive real roots. After this result as been established, we utilize Sturm sequences to determine whether a polynomial has positive real roots (c.f. [FN04]).

# Bibliography

- [Con43] N. B. Conkwright. An elementary proof of the budan-fourier theorem. *The American Mathematical Monthly*, 50(10):603–605, December 1943.
- [Con57] N. B. Conkwright. *Introduction to the Theory of Equations*. Ginn and Company, Boston, MA, 1957.
- [FN04] Jonathan Forde and Patrick Nelson. Applications of sturm sequences to bifurcation analysis of delay differential equation models. *Mathematical Analysis and Applications*, 300:273–284, 2004.
- [GG03] J. Gathen and J. Gerhard. *Modern Computer Algebra*. Cambridge University Press, Cambridge CB2 2RU, UK, 2 edition, 2003.
- [Lan93] Serge Lang. *Algebra*. Addison-Wesley, 3rd edition, 1993.
- [Mig92] M. Mignotte. *Mathematics for Computer Algebra*. Springer-Verlag, New York, NY, 1992.
- [Ste06] J. Stewart. *Calculus: Concepts and Contexts*. Thomson Brooks/Cole, Belmont, CA, 3rd edition, 2006.

# Appendices

## APPENDIX A

We will note a few facts from calculus that will be used in the proofs in Appendix B.

### Rolle's Theorem

**Theorem:** *Let  $f(x) = 0$  be an algebraic equation with real coefficients. Between two consecutive real roots  $a$  and  $b$  of this equation there is an odd number of roots of the equation  $f'(x) = 0$ . A root of multiplicity  $m$  is here counted as  $m$  roots (c.f. [Con57]).*

### Intermediate Value Theorem

**Theorem:** *Suppose that  $f$  is continuous on the closed interval  $[a, b]$  and let  $N$  be any number between  $f(a)$  and  $f(b)$ , where  $f(a) \neq f(b)$ . Then there exists a number  $c$  in  $(a, b)$  such that  $f(c) = N$  (c.f. [Ste06], p.124).*

One of general uses of the Intermediate Value Theorem is to find real solutions to equations. A simple example from calculus will help clarify any questions.

**Example:** (C.f. [Ste06], p. 125) Show there is a root of the equation

$$4x^3 - 6x^2 + 3x - 2 = 0$$

on the interval  $[1, 2]$ .

SOLUTION: Let  $f(x) = 4x^3 - 6x^2 + 3x - 2$ . We want to show there exists a number  $c$  between 1 and 2 that is a solution to  $f(x)$  such that  $f(c) = 0$ . Consider  $a = 1$ ,  $b = 2$  and  $N = 0$  in the Intermediate Value Theorem. Then

$$f(1) = 4 - 6 + 3 - 2 = -1 < 0$$

and

$$f(2) = 32 - 24 + 6 - 2 = 12 > 0.$$

Therefore,  $f(1) < 0 < f(2)$ , that is,  $N = 0$  is a number between  $f(1)$  and  $f(2)$ . Since  $f$  is continuous function the Intermediate Value Theorem says there exists a number  $c$  between 1 and 2 such that  $f(c) = 0$ . So, the equation  $4x^3 - 6x^2 + 3x - 2 = 0$  has at least root on the interval  $(1, 2)$ .

## APPENDIX B

### Proof of Descartes's Rule of Signs

Let the number of positive roots of the equation

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

be denoted by  $p$  and let  $v$  denote the number of sign variations of the coefficients. Since  $a_n$  is non zero, we can divide all the above coefficients by this number. This will not change the number of sign variations nor will it change the number of positive roots but it will allow us to assume that the leading coefficient of our polynomial is 1. We therefore can assume that  $f(x) = x^n + a_{n-1} x^{n-1} + \dots + a_0$ , where for some (possibly different)  $a_i$ . Let  $r_1, \dots, r_p$  be the positive roots (each listed as many times as its multiplicity). We then have

$$f(x) = g(x) \prod_{i=1}^p (x - r_i)$$

where  $g(x) = x^m + b_{m-1} x^{m-1} + \dots + b_0$  is a polynomial with no positive roots.

We will first show that  $p$  and  $v$  are both even or both odd. To see this first note that  $b_0$  must be positive. If  $b_0$  were negative, then  $g(0) = b_0$  would be negative while for  $x$  sufficiently large  $g(x)$  is positive. By the Intermediate Value Theorem,  $g(x)$  would then have a positive root, a contradiction. Therefore  $b_0$  is positive. We have that

$$a_0 = b_0 \left( (-1)^p \prod_{i=1}^p r_i \right)$$

so  $a_0$  is positive if  $p$  is even and  $a_0$  is negative if  $p$  is odd. Since the leading coefficient of  $f$  is 1, and so is positive, the number of sign changes must be even when  $a_0$  is positive and odd when  $a_0$  is negative. This allows us to conclude that  $p$  and  $v$  are both even or both odd.

We now will show that  $p \leq v$ . We will do this by induction on  $n$ , the degree of  $f$ . If  $n = 1$  we have that  $f(x) = x + a_0$  (recall we are assuming that the leading coefficient is 1). There is only one root of this polynomial,  $r_1 = -a_0$ . If  $a_0 > 0$ , then  $f$  has no sign changes,  $v = 0$ , and its only root is negative,  $p = 0$ , so  $v = p$ . If  $a_0 < 0$ , then  $f$  has one sign change,  $v = 1$ ,

and on positive root,  $p = 1$ , so  $v = p$  again.

Now assume that  $n > 1$ . Let  $q$  denote the number of positive roots of  $f'(x)$  and  $w$  denote the number of sign changes of the coefficients of  $f'$ . Since

$$f'(x) = nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \dots + a_1$$

the number of sign changes is at most  $v$ , so  $w \leq v$ . It is possible that  $a_1 = 0$  and so our induction hypothesis does not directly apply. If this is the case, we can divide  $f'(x)$  by some power  $x^k$  of  $x$  and yield a new polynomial  $h(x) = f'(x)/x^k$  whose constant term is nonzero. Note that the number of sign changes in  $h(x)$  is still  $w$  and the number of positive roots is still  $q$ . Applying the induction hypothesis we have  $q \leq w$ . Rolle's Theorem implies that  $p - 1 \leq q$ . Therefore

$$p - 1 \leq q \leq w \leq v.$$

We therefore have that  $p - 1 \leq v$ . We cannot have that  $p - 1 = v$  since  $p$  and  $q$  are both even or both odd. Therefore  $p \leq v$ , the desired conclusion (c.f. [Con57]).  $\square$

### Proof of Budan-Fourier's Theorem

Let the interval  $(a, b]$  be denoted by  $I$ . Let  $p$  equal the number of roots of  $f(x) = 0$  on the interval  $I$ .

It can be seen that  $p$  is equal to the number of positive non-zero roots of

$$f(x + a) = f(a) + f'(a)x + \frac{1}{2}f''(a)x^2 + \dots + \frac{1}{n!}f^{(n)}(a)x^n = 0 \quad (.1)$$

minus the number of positive non-zero roots of

$$f(x + b) = f(b) + f'(b)x + \frac{1}{2}f''(b)x^2 + \dots + \frac{1}{n!}f^{(n)}(b)x^n = 0. \quad (.2)$$

This follows from the fact that equations (.1) and (.2) are obtained by diminishing the roots of  $f(x) = 0$  by  $a$  and  $b$  respectively.

By Descartes' rule of signs, it follows that (.1) has  $v(a) - 2h_1$  positive roots and that (.2) has  $v(b) - 2h_2$  positive roots, for integers  $h_1 \geq 0$  and  $h_2 \geq 0$ . Therefore

$$p = (v(a) - 2h_1) - (v(b) - 2h_2) = v(a) - v(b) - 2m, \quad (.3)$$

where  $m = h_1 - h_2$ . We only need to prove that  $m \geq 0$ .

The proof that  $m \geq 0$  will be effected by mathematical induction. It can be verified that  $m \geq 0$  if  $f(x)$  has degree one. Hence it will be sufficient to show that this inequality holds for every equation of degree  $n$  if the corresponding inequality holds for every equation of degree  $n - 1$ . We shall therefore base the argument to follow upon the assumption that the specified inequality is valid for every equation of degree  $n - 1$ .

Now let  $q$  denote the number of roots of the equation  $f'(x) = 0$  in the interval  $I$ , and let  $v'(c)$  denote the number of variations of sign in the sequence

$$f'(x), f''(x), \dots, f^{(n)}(x),$$

when  $x = c$ , where  $c$  is any real number. By a similar argument to that which led to (.3), it can be shown that

$$q = v'(a) - v'(b) - 2m'$$

where  $m'$  is a positive integer or zero. It can also be determined that  $m' \geq 0$  by assumption.

It follows from Rolle's theorem that  $q \geq p - 1$ , say  $q = p - 1 + s$  where  $s \geq 0$  is an integer. We also note that

$$v(a) = v'(a) \quad \text{or} \quad v(a) = v'(a) + 1,$$

and

$$v(b) = v'(b) \quad \text{or} \quad v(b) = v'(b) + 1.$$

The possible combinations are

$$\begin{aligned} v(a) - v(b) &= v'(a) - v'(b) \\ v(a) - v(b) &= v'(a) - (v'(b) + 1) = v'(a) - v'(b) - 1 \\ v(a) - v(b) &= (v'(a) + 1) - v'(b) = v'(a) - v'(b) + 1 \\ v(a) - v(b) &= (v'(a) + 1) - (v'(b) + 1) = v'(a) - v'(b). \end{aligned}$$

Hence,

$$v(a) - v(b) \geq v'(a) - v'(b) - 1. \quad (.4)$$

Consider the first case in which  $v(a) - v(b) > v'(a) - v'(b) - 1$ . Then

$$v(a) - v(b) \geq v'(a) - v'(b) = q + 2m' = p - 1 + s + 2m'. \quad (.5)$$

Now if  $s + 2m' \neq 0$ , we have  $v(a) - v(b) \geq p$ . It follows from (.3) that  $m \geq 0$ , as was to be proved.

If  $s + 2m' = 0$ , we have  $v(a) - v(b) \geq p - 1$ . But it follows from (.3) that  $v(a) - v(b) - p$  is even, and hence  $v(a) - v(b) \neq p - 1$ . Consequently  $v(a) - v(b) > p - 1$ , and therefore  $v(a) - v(b) \geq p$ . Again it follows from (.3) that  $m \geq 0$ .

Now consider the case in which the two members of (.4) are equal. This is possible only if  $v(a) = v'(a)$ , and

$$v(b) = v'(b) + 1. \quad (.6)$$

It will be shown later that, if (.6) holds, then  $q \geq p$ . Assuming this fact for the moment, and letting  $q = p + t$ , where  $t \geq 0$  is an integer, we have

$$v(a) - v(b) = v'(a) - v'(b) - 1 = q + 2m' - 1 = p + t + 2m' - 1. \quad (.7)$$

The first and last members of (.7) are essentially the same in (.5). And by an argument similar to that developed in connection with (.5) it can be shown that  $m \geq 0$  in the case under consideration.

To complete the proof we need to show that  $q \geq p$  if (.6) holds. In this case  $f(b) \neq 0$ , and if  $f'(b) \neq 0$ , then  $f(b)$  and  $f'(b)$  must have different signs. Now let  $r$  be the greatest root of  $f(x) = 0$  in the interval  $I$ . Then  $f'(x)$  must vanish at least once in the interval  $r < x \leq b$ , or otherwise  $f(b)$  and  $f'(b)$  would be the same sign. Hence, the equation  $f'(x) = 0$  has in the interval  $I$  the  $p - 1$  roots vouched for by Rolle's theorem, and at least one additional root in the interval  $r < x \leq b$ . Therefore  $q \geq p$  (c.f. [Con43]).  $\square$



### Proof of Sturm's Theorem

We will follow the proof given in *Algebra, Third Edition* by S. Lang, Addison-Wesley 1993. First note that the Sturm Sequence  $\{f_0 = f, f_1 = f', \dots, f_n\}$  satisfies the following four properties that follow from the definitions of the  $f_i$ :

1. The last polynomial  $f_n$  is a non-zero constant.
2. There is no point  $x \in [a, b]$  such that  $f_j(x) = f_{j+1}(x) = 0$  for any value  $j = 0, \dots, n-1$ .
3. If  $x \in [a, b]$  and  $f_j(x) = 0$  for some  $j = 1, \dots, n-1$ , then  $f_{j-1}(x)$  and  $f_{j+1}(x)$  have opposite signs.
4. We have  $f_j(a) \neq 0$  and  $f_j(b) \neq 0$  for  $j = 0, \dots, n$ .

Item 1. follows from the fact that  $f$  has no multiple roots. To verify Item 2. note that

$$f_j = q_i f_{j+1} - f_{j+2}. \quad (.8)$$

Therefore, if  $f_j(x) = f_{j+1}(x) = 0$ , we would have  $f_{j+2}(x) = 0$ , and by induction,  $f_n(x) = 0$ , a contradiction. Item 3. follow from the fact that if  $f_j(x) = 0$ , then equation (.8) implies that  $f_{j-1}(x) = -f_{j+1}(x)$ . Item 4. is part of the definition of  $a$  and  $b$ .

Let  $r_1 < \dots < r_s$  be a list of the zeros of the Sturm functions on the interval  $[a, b]$ . Note that  $v(x)$  is constant on any interval  $(r_i, r_{i+1})$ . Therefore it is enough to show that if  $c < d$  are any numbers such that there is precisely one element  $r$  that is a root of some  $f_j$  with  $c < r < d$ , then  $v(c) - v(d) = 1$  if  $r$  is a root of  $f$  and  $v(c) - v(d) = 0$  if  $r$  is not a root of  $f$ . The conclusion of the theorem will then follow from the fact that

$$v(a) - v(b) = [v(x_0) - v(x_1)] + [v(x_1) - v(x_2)] + \dots + [v(x_{s-1}) - v(x_s)]$$

where the  $x_i$  are selected such that  $r_i < x_i < r_{i+1}$ .

Suppose that  $r$  is a root of some  $f_j$ . By Item 3. we have that  $f_{j-1}(r)$  and  $f_{j+1}(r)$  have opposite signs and by Item 2. these do not change when we replace  $r$  by  $c$  or  $d$ . Therefore, there will be one sign change in each of the triples

$$\{f_{j-1}(c), f_j(c), f_{j+1}(c)\} \quad \{f_{j-1}(d), f_j(d), f_{j+1}(d)\}.$$

If  $r$  is not a root of  $f$ , then we have that  $v(c) = v(d)$ . If  $r$  is a root of  $f$  then  $f(c)$  and  $f(d)$  have opposite signs. We also have that  $f'(r) \neq 0$ , so  $f'(c)$  and  $f'(d)$  have the same sign. If

$f(c) < 0$  then  $f$  must be increasing so  $f(d) > 0$ ,  $f'(c) > 0$  and  $f'(d) > 0$  and if  $f(c) > 0$ , then  $f$  must be decreasing so  $f(d) < 0$ ,  $f'(c) < 0$  and  $f'(d) < 0$ . Therefore there is one sing change in  $\{f(c), f'(c)\}$  and none in  $\{f(d), f'(d)\}$ . So we have  $v(c) = v(d) + 1$  (c.f. [Lan93])

□