GENERALIZED STURM SEQUENCES

A Thesis

Presented to

The Faculty of the Department of Mathematics

California State University, Los Angeles

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Mathematics

By

Ibrahim Jaber Alotaibi

June 2015

The thesis of Ibrahim Jaber Alotaibi is approved.

Gary Brookfield, PhD., Committee Chair

Borislava Gutarts, PhD.

Kristin Webster, PhD.

Grant Fraser, PhD., Department Chair

California State University, Los Angeles

June 2015

ABSTRACT

GENERALIZED STURM SEQUENCES

By

Ibrahim Jaber Alotaibi

If you have a polynomial $f \in \mathbb{R}[x]$ and want to know about the number and values of its real roots, there is no better tool than a Sturm sequence. This is a sequence of polynomials $f_0, f_1, f_2, \ldots, f_n = f$ with the property that the values of these polynomials *at* real numbers $a < b$ determine the number of real roots of f that are *between* a and b. In this thesis we investigate a generalization of the definition of Sturm sequences. This more flexible definition allows us to make connections between our definition of Sturm sequences and interlacing of polynomials that have not been noticed before. New Sturm sequences can be obtained, for example, from families of orthogonal polynomials and characteristic polynomials of symmetric matrices.

ACKNOWLEDGMENTS

There are no proper words to convey my deep acknowledgement and appreciation for my thesis advisor, Dr. Gary Brookfield. One simply could not wish for a better supervisor. I have been extremely lucky to have a supervisor who cared so much about my work and progress, and who responded to my questions with nice smile. Without Dr. Brookfield this thesis would not have been completed or written. I will forever be thankful to him.

Also, many thanks to Dr. Borislava Gutarts and Dr. Kristin Webster for their considerations, time, and help. I am so glad that they are in my thesis committee. I will NEVER forget my uncle and my mother and my siblings for everything they have done for us. Special thanks for my love, my wife, and my kids for their support all the time. Also my love and appreciation to ALL of my friends in the math department. I will miss you guys.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

CHAPTER 1

Introduction

To find the number of real roots of polynomial with real coefficients over a given interval, there is no better tool than a Sturm sequence. This is a sequence of polynomials $f_0, f_1, f_2, \ldots, f_n = f$ with the property that the values of these polynomials *at* real numbers $a < b$ determine the number of real roots of f that are *between* a and b. This very important algebraic problem was solved in a surprisingly simple way in 1829 by the French mathematician Charles Sturm (1803-1855). The paper containing the famous Sturm's Theorem appeared in the eleventh volume of the Bulletin des sciences de Ferussac and bears the title, "Memoire sur la resolution des equations numeriques". Sturm's Theorem gives the number of real roots of a polynomial within an interval in terms of the number of changes of signs of the values of the Sturm's sequence at the bounds of the interval. Applying Sturm's Theorem to the interval of all the real numbers gives the total number of real roots of a polynomial [1]. In this thesis we show that there are many other sequences with the same desired properties. For instance, if $f \in \mathbb{R}[x]$ has degree $n \in \mathbb{N}$ and n real roots (counting multiplicities). Then the sequence of derivatives of f,

$$\big(\, f^{(n)}(x)\,,\; f^{(n-1)}(x)\,,\; \ldots\,,\; f''(x)\,,\; f'(x)\,,\; f(x)\,\big)$$

is a **generalized Sturm sequence**. Also, our Sturm sequences have new properties that are not in the standard Sturm sequence. This flexibility allows us to generalize the idea of Sturm theorem.

Definition 1.1. *Let* $S = (a_0, a_1, \ldots, a_n)$ *be a sequence of nonzero real numbers. Then the* ***number of variations*** *in* S *is the number of times the sign changes in the sequence as one reads from left to right. We write* $\mathcal{V}(S) = \mathcal{V}(a_0, a_1, \ldots, a_n)$ *for this number.*

Example 1.2. Here are some examples:

$$\mathcal{V}(1, 2, 3, 1) = 0 \qquad \mathcal{V}(1, -1, 5, -1) = 3 \qquad \mathcal{V}(1) = 0$$

$$\mathcal{V}(1, -\pi, -3, -4) = 1 \qquad \mathcal{V}(-1, 1, 1, -1) = 2 \qquad \mathcal{V}(5, -2, 5, -3) = 3.$$

Since only the signs of the terms of the sequence matter, we will say that two sequences are equivalent if the signs of their terms match up, and we will write

$$[\, a_0 \,,\, a_1 \,,\, \ldots \,,\, a_n \,]$$

for the equivalence class containing $(a_0, a_1, \ldots, a_n)$. Thus $[\, a \,] = [\, b \,]$ if and only if a and b have the same sign or are both zero, and

$$[\, a_0 \,,\, a_1 \,,\, \ldots \,,\, a_n \,] = [\, b_0 \,,\, b_1 \,,\, \ldots \,,\, b_n \,]$$

if $[\, a_i \,] = [\, b_i \,]$ for $i = 1, 2, \ldots, n$. For example, if $a, b, c \in \mathbb{R}$, then

$$\left[\, 1 + a^2 \,,\, b^3 \,,\, -e^c \,,\, 5d \,\right] = [\, 1 \,,\, b \,,\, -1 \,,\, d \,]\,,$$

and

$$\mathcal{V}\left[\, 1 + a^2 \,,\, b^3 \,,\, -e^c \,,\, 5d \,\right] = \mathcal{V}\,[\, 1 \,,\, b \,,\, -1 \,,\, d \,] = \begin{cases} 2 & \text{if } d > 0 \\ 1 & \text{if } d \le 0. \end{cases}$$

Notice that the number of variations in this sequence is independent of the value of b. This happens because the terms before and after b in the sequence have opposite signs. This property and several others are collected in the next lemma.

Lemma 1.3. *Let $S = (a_0, a_1, \ldots, a_n)$ be a sequence of nonzero real numbers.*

(1) *If $r \in \mathbb{R}$ is nonzero then $\mathcal{V}(S) = \mathcal{V}(ra_0, ra_1, \ldots, ra_n)$.*

(2) *If $0 \leq m \leq n$, then $\mathcal{V}(S) = \mathcal{V}(a_0, a_1, \ldots, a_m) + \mathcal{V}(a_m, a_{m+1}, \ldots, a_n)$.*

(3) *If $0 < m < n$, and $a_{m-1}a_{m+1} < 0$, then a_m can be changed arbitrarily without changing $\mathcal{V}(S)$. That is, the number of variations in S is independent of the sign of a_m.*

Our first example of a Sturm sequence is artificial in the sense that we construct it from the roots of the polynomial, rather than using a Sturm sequence to find the roots. Even so, it shows how the values of the polynomials in a Sturm sequence determine the location of the roots. Fix $x_1, x_2, x_3 \in \mathbb{R}$, and define

$$\begin{aligned} f_0(x) &= 1 \\ f_1(x) &= (x - x_1) \\ f_2(x) &= (x - x_1)(x - x_2) \\ f_3(x) &= (x - x_1)(x - x_2)(x - x_3) \end{aligned} \tag{1.1}$$

For $x \in \mathbb{R}$, the number of variations in the sequence

$$(\, f_0(x)\, ,\ f_1(x)\, ,\ f_2(x)\, ,\ f_3(x)\,)$$

is a function of x that we denote as $\mathcal{V}(x)$. Note $\mathcal{V}$ is not defined on $\{x_1, x_2, x_3\}$ since we have defined $\mathcal{V}$ only for sequences of nonzero real numbers.

The function $\mathcal{V}(x)$ has the remarkable property (see Theorem 2.7) that, so long as $a \notin \{x_1, x_2, x_3\}$, $\mathcal{V}(a)$ is the number of roots of f_3 that are greater than a. This is illustrated by the diagram below for which $x_2 < x_1 < x_3$.

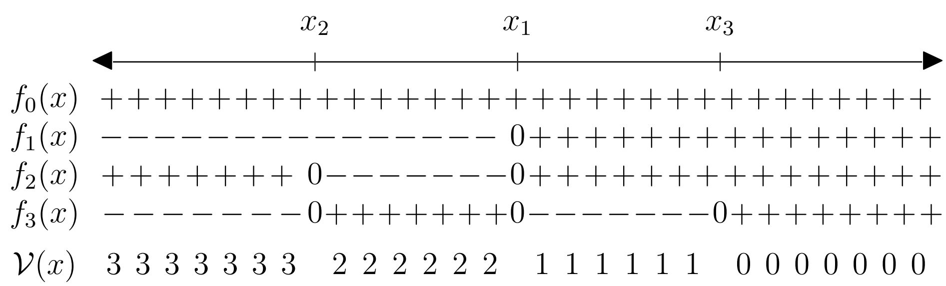

Figure 1.1: Interlacing that is not strict

For example, the diagram shows that $f_2(x) = (x - x_1)(x - x_2)$ is positive for all $x < x_2$, negative for $x_2 < x < x_1$, and then positive for all $x > x_1$. The value of $\mathcal{V}$ along the bottom of the diagram decreases from 3 to 0. This is, of course, in agreement with our claim that $\mathcal{V}(a)$ is the number of roots of f_3 that are greater than a.

The reader can easily check that this property of $\mathcal{V}(x)$ is independent of the order of the numbers x_1, x_2 and x_3 and even if some of these are equal.

Of course, if $a < b$ are real numbers that are not in $\{x_1, x_2, x_3\}$, then the number of roots of f between a and b is $\mathcal{V}(a) - \mathcal{V}(b)$.

This property of $\mathcal{V}(x)$ is the basis of the definition of a Generalized Sturm sequence in the following chapter.

CHAPTER 2

Generalized Sturm Sequences

Definition 2.1. *Let $f \in \mathbb{R}[x]$. If $a < b$ are real numbers or $\pm\infty$, then the number of real roots of f on the interval (a, b) (counted with multiplicities) is denoted $\mathcal{N}(f; a, b)$.*

Note that $\mathcal{N}(f; a, b) = \mathcal{N}(rf; a, b)$ for all $r \neq 0$. Also,

$$\mathcal{N}(f; a, b) = \mathcal{N}(f; a, \infty) - \mathcal{N}(f; b, \infty). \tag{2.1}$$

Definition 2.2. *A **Sturm sequence** is a sequence of polynomials in $\mathbb{R}[x]$,*

$$(\, f_0\,,\ f_1\,,\ \ldots\,,\ f_n\,)\,,$$

such that, if $a < b$ are real numbers neither of which is a root of any polynomial in the sequence, then

$$\mathcal{V}(a) - \mathcal{V}(b) = \mathcal{N}(f_n; a, b) - \mathcal{N}(f_0; a, b).$$

Here, as above, $\mathcal{V}(x)$ is defined by

$$\mathcal{V}(x) = \mathcal{V}\,(\, f_0(x)\,,\ f_1(x)\,,\ \ldots\,,\ f_n(x)\,)\,.$$

One might consider this the definition of a “relative” Sturm sequence since it involves the difference in the number of real roots for two polynomials f_0 and f_n. Usually, we are interested in Sturm sequences in which f_0 is a constant polynomial and so $\mathcal{V}(a) - \mathcal{V}(b) = \mathcal{N}(f_n; a, b)$.

Each polynomial in a Sturm sequence has only finitely many roots, so the set of these roots is bounded above, that is, there is $B \in \mathbb{R}$ such that all the roots of the polynomials are less than B. Then each polynomial has constant sign on the interval

(B,∞), and $\mathcal{V}$ is a constant function on that same interval. The value of this constant function is denoted $\mathcal{V}(\infty)$, since it could be defined by $\mathcal{V}(\infty) = \lim_{x\to\infty} \mathcal{V}(x)$. Then, $\mathcal{V}(\infty) = \mathcal{V}(B)$. Similarly, we define $\mathcal{V}(-\infty) = \lim_{x\to-\infty} \mathcal{V}(x)$. Then $\mathcal{V}(-\infty) - \mathcal{V}(\infty)$ is the number of real roots of f_n minus the number of roots of f_0 , counted with multiplicities, and, for all $a \in \mathbb{R}$, $\mathcal{V}(a) - \mathcal{V}(\infty)$ is the number of real roots of f_n minus the number of roots of f_0 that are greater than a and $\mathcal{V}(-\infty) - \mathcal{V}(a)$ is the number of real roots of f_n minus the number of roots of f_0 that are less than a.

Notice also that $\mathcal{V}(\infty)$ is simply the number of variations in the sequence of leading coefficients of the polynomials in the Sturm sequence. For example, if all polynomials have positive leading coefficients, then $\mathcal{V}(\infty) = 0$.

Perhaps an example of a Sturm sequence will help to make the definition clearer. Set

$$f_0(x) = 1$$

$$f_1(x) = x$$

$$f_2(x) = (x+2)(x-2)$$

$$f_3(x) = (x+3)x(x-3)$$

$$f_4(x) = (x+4)(x+1)(x-1)(x-4)$$

Figure 2.1 shows the signs of these functions on the number line and makes it easy to confirm that $(f_0, f_1, f_2, f_3, f_4)$ is a Sturm sequence.

Even though it is not part of the definition, in many of our applications we will have $\deg f_i = i$ for $i = 1, 2, \ldots, n$.

The bottom line of the diagram shows the number of variations in the sequence

x	$-5 \quad -4 \quad -3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5$
$f_0(x)$	$+++++++++++++++++++++++++++++++$
$f_1(x)$	$---------------\ 0+++++++++++++++$
$f_2(x)$	$++++++++++\ 0---------0++++++++++$
$f_3(x)$	$-------0+++++++0-------0++++++++$
$f_4(x)$	$+++++0-------0++++0-------0+++++$
$\mathcal{V}(x)$	4 4 4 4 4 3 3 3 3 3 3 3 2 2 2 2 1 1 1 1 1 1 1 0 0 0 0 0

Figure 2.1: An example of a Sturm sequence

$(f_0(x), f_1(x), f_2(x), f_3(x), f_4(x))$ as a function of x. For example,

$$\begin{aligned}\mathcal{V}(-2) &= \mathcal{V}(f_0(-2), f_1(-2), f_2(-2), f_3(-2), f_4(-2)) \\ &= \mathcal{V}(1, -2, 0, 10, -36) \\ &= 3\end{aligned}$$

Lemma 2.3. *Let $S = (f_0, f_1, \ldots, f_n)$ be a sequence of polynomials in $\mathbb{R}[x]$, then S is a Sturm sequence if and only if*

$$\mathcal{V}(a) - \mathcal{V}(\infty) = \mathcal{N}(f_n; a, \infty) - \mathcal{N}(f_0; a, \infty).$$

for all $a \in \mathbb{R}$ such that a is not a root of any of the polynomials in S. That means for all $a \in \mathbb{R}$, when a is not a root of any one of the polynomials in the sequence, $\mathcal{V}(a) - \mathcal{V}(\infty)$ is the number of real roots of f_n that are greater than a minus the number of real roots of f_0 that are greater than a.

Proof. This follows from (2.1). □

Lemma 2.4. *Let $S = (f_0, f_1, \ldots, f_n) \in \mathbb{R}[x]$ be a sequence of polynomials:*

(1) *If $g(x) \in \mathbb{R}[x]$ is nonzero polynomial, then S is a Sturm sequence if and only if $(gf_0, gf_1, \ldots, gf_n)$ is a Sturm sequence.*

(2) *If S is the concatenation of subsequences $S_1 = (f_0, f_1, \ldots, f_m)$ and $S_2 = (f_m, f_{m+1}, \ldots, f_n)$ with $0 < m < n$, then if any two of (S, S_1, S_2) are Sturm sequences, then so is the remaining sequence.*

Proof. This follows directly from Lemma 1.3 (1), (2), and the definition. □

Notice that segments of a Sturm sequence are not always Sturm sequences. That is, in the situation of the previous lemma, if S is a Sturm sequence then S_1 and S_2 may not be Sturm sequences.

Lemma 2.5. *(f, g) is a Sturm sequence if and only if $(g, -f)$ is a Sturm sequence.*

Proof. Since we have two potential Sturm sequences, we define $\mathcal{V}(x) = \mathcal{V}(f(x), g(x))$ and $\mathcal{V}'(x) = \mathcal{V}(g(x), -f(x))$. Note that $\mathcal{V}(x), \mathcal{V}'(x) \in 0, 1$ and

$$\mathcal{V}'(x) = \begin{cases} 0 & \text{if } \mathcal{V}(x) = 1 \\ 1 & \text{if } \mathcal{V}(x) = 0 \end{cases}.$$

Suppose (f, g) is a Sturm sequence. So, $\mathcal{V}(a) - \mathcal{V}(b) = \mathcal{N}(g; a, b) - \mathcal{N}(f; a, b)$ for all $a, b \in \mathbb{R}$ that are not roots of f or of g. We need to show that $(g, -f)$ is a Sturm sequence. That is, we need to show

$$\mathcal{V}'(a) - \mathcal{V}'(b) = \mathcal{N}[f; a, b] - \mathcal{N}[g; a, b]$$

Let us construct the table that shows all possible cases.

$\mathcal{V}(a)$	$\mathcal{V}(b)$	$\mathcal{V}'(a)$	$\mathcal{V}'(b)$	$\mathcal{V}(a) - \mathcal{V}(b)$	$\mathcal{V}'(a) - \mathcal{V}'(b)$
0	0	1	1	0	0
0	1	1	0	-1	1
1	0	0	1	1	-1
1	1	0	0	0	0

Figure 2.2: Table for Lemma 2.5

From the table we see that, in all cases,

$$\mathcal{V}'(a) - \mathcal{V}'(b) = -(\mathcal{V}(a) - \mathcal{V}(b))$$

Hence

$$\begin{aligned}\mathcal{V}'(a) - \mathcal{V}'(b) &= -(\mathcal{N}[g; a, b] - \mathcal{N}[f; a, b]) \\ &= \mathcal{N}[f; a, b] - \mathcal{N}[g; a, b]\end{aligned}$$

Therefore, $(\, g\,,\, -f\,)$ is a Sturm sequence.

Conversely, by symmetry, if $(\, g\,,\, -f\,)$ is a Sturm sequence then $(\, f\,,\, g\,)$ is a Sturm sequence. □

Now we can prove that our motivating example of a Sturm sequence in (1.1), actually is a Sturm sequence.

Lemma 2.6. *Let $f(x) \in \mathbb{R}[x]$ and $r \in \mathbb{R}$. Then $(\, f(x)\,,\, (x-r)f(x)\,)$ is a Sturm sequence.*

Proof. If $a \in \mathbb{R}$ is not a root of $(x-r)f(x)$, then

$$\mathcal{V}(a) = \mathcal{V}\,(\, f(a)\,,\, (a-r)f(a)\,) = \begin{cases} 0 & \text{if } a > r \\ 1 & \text{if } a < r \end{cases}.$$

Therefore, if $a < b \in \mathbb{R}$ are not roots of $(x-r)f(x)$, then $\mathcal{V}(a) - \mathcal{V}(b)$ will be 0 or 1, and will be 1, if and only if $a < r < b$, if and only if $(x-r)f(x)$ has exactly one more root on the interval (a, b) than $f(x)$ does. □

Theorem 2.7. *Let $x_1, x_2, \ldots, x_n \in \mathbb{R}$, and define*

$$\begin{aligned}
f_0(x) &= 1 \\
f_1(x) &= (x - x_1) \\
f_2(x) &= (x - x_1)(x - x_2) \\
&\vdots \\
f_n(x) &= \prod_{i=1}^{n}(x - x_i)
\end{aligned} \tag{2.2}$$

Then $(\, f_0\, ,\, f_1\, ,\, \ldots\, ,\, f_n\,)$ is a Sturm sequence.

Proof. This is an easy induction using Lemma 2.4(2) and Lemma 2.6. □

Theorem 2.8. *Suppose $f \in \mathbb{R}[x]$ has n real roots (counting multiplicities). Then there is a Sturm sequence $(\, f_0\, ,\, f_1\, ,\, \ldots\, ,\, f_n\,)$ with $f_0 = \pm 1$ and $f_n = f$.*

Proof. By assumption f can be written as

$$f(x) = g(x)(x - x_1)(x - x_2) \cdots (x - x_n)$$

with $x_1, x_2, \ldots, x_n \in \mathbb{R}$ and $g \in \mathbb{R}[x]$ having no real roots. Define

$$\begin{aligned}
f_0(x) &= g(x) \\
f_1(x) &= g(x)(x - x_1) \\
f_2(x) &= g(x)(x - x_1)(x - x_2) \\
&\vdots \\
f = f_n(x) &= g(x)(x - x_1)(x - x_2) \cdots (x - x_n)
\end{aligned} \tag{2.3}$$

By Lemma 2.4(1) and Theorem 2.7, $(\, f_0\, ,\, f_1\, ,\, \ldots\, ,\, f_n\,)$ is a Sturm sequence. Since g has a constant sign, f_0 can be changed to 1 or -1 without affecting $\mathcal{V}(x)$, so, with this change, $(\, f_0\, ,\, f_1\, ,\, \ldots\, ,\, f_n\,)$ is still a Sturm sequence. □

Note that Theorem 2.8 provides a minimal Sturm sequence in the following sense: any Sturm sequence that has f at one end and constant at the other end must have at least $n+1$ polynomials so that $\mathcal{V}(-\infty) - \mathcal{V}(\infty) = n$.

CHAPTER 3

Interlacing

Definition 3.1. *Suppose that all the real roots of* $f \in \mathbb{R}[x]$ *are* $x_1 \leq x_2 \leq \cdots \leq x_m$ *and all the real roots of* $g \in \mathbb{R}[x]$ *are* $y_1 \leq y_2 \leq \cdots \leq y_n$.

(1) *If* $m = n+1$ *and* $x_1 \leq y_1 \leq x_2 \leq y_2 \leq \cdots \leq x_{n-1} \leq y_{n-1} \leq x_n$*, we will write* $g \preceq f$*. If all the inequalities are strict, we write* $g \prec f$*. Note that we require that* g *to have exactly one less real root than* f*.*

(2) *If* $n = m$ *and* $x_1 \leq y_1 \leq x_2 \leq y_2 \leq \cdots \leq x_n \leq y_n$*, we will write* $g \trianglelefteq f$*. If all the inequalities are strict, we write* $g \triangleleft f$*. In this case, we require* f *to have exactly the same number of real roots as* g*.*

If $g \preceq f$ *or* $f \preceq g$ *or* $g \trianglelefteq f$ *or* $f \trianglelefteq g$*, we say that the roots of* f *and the roots of* g *are* ***interlaced****. If* $g \prec f$ *or* $f \prec g$ *or* $g \triangleleft f$ *or* $f \triangleleft g$*, we say that the roots of* f *and the roots of* g *are* ***strictly interlaced****.*

Strict interlacing is easy to characterize.

Lemma 3.2. *Let* $f, g \in \mathbb{R}[x]$ *then the roots of* f *and* g *are strictly interlaced if and only if*

(1) *all real roots of* f *and* g *are simple,*

(2) f *and* g *have no common roots,*

(3) *between any pair of adjacent roots of* f *there is a root of* g *and vice versa.*

Proof. This follows directly from the definition. □

In contrast to strict interlacing, if the roots of f and g are interlaced but not strictly, then f and g may have common roots and roots of multiplicity greater than 1.

For example, suppose that x_2 is a simple real root of f. Then $x_1 < x_2 < x_3$ and there are three possibilities for the corresponding root of g:

- If $y_1 < x_2 < y_2$, then x_2 is not a root of g.
- If $y_1 = x_2 < y_2$ or $y_1 < x_2 = y_2$, then x_2 is simple root of g.
- If $y_1 = x_2 = y_2$, then x_2 is double root of g.

Thus x_2 is a root of g with multiplicity 0, 1 or 2.

Similarly, suppose that $x_1 < x_2 = x_3 = x_4 < x_5$, that is, x_2 is a root of f with multiplicity 3. Then $y_2 = y_3 = x_2$, so x_2 is at least a double root of g. But it is also possible that $x_1 < y_1 = x_2$ and/or $x_4 = y_4 < x_5$, so x_2 could have multiplicity 3 or 4 as a root of g.

The general rule is easy to spot (but a little cumbersome to prove):

Lemma 3.3. *If $f, g \in \mathbb{R}[x]$ and the real roots of g and f are interlaced, then a real root of f with multiplicity m is a root of g with multiplicity $m-1$, m or $m+1$.*

Lemma 3.4. *If $g \preceq f$ ($g \trianglelefteq f$) and $h = \gcd(f, g)$, then $g/h \prec f/h$ ($g/h \triangleleft f/h$).*

Proof. Suppose $g \preceq f$ (or $g \trianglelefteq f$,) but $g \prec f$ (or $g \triangleleft f$) is not true. Then g and f have a common root a. By Lemma 3.3 , the multiplicity of a in f and g differ by at most one. Suppose the roots of f and g are indexed as in Definition 3.1 (1) or (2). Since a is a common root of f and g there is a segment of this sequence of inequalities in which the $\leq$ are replaced by $=$. We have two cases:

Case I: Suppose that the first root in the segment is a root of f and last root is a root of g. That is,

$$y_{k-1} < x_k = y_k = \cdots = x_{k+m} = y_{k+m} < x_{k+m+1}$$

for some $k, m \in \mathbb{N}$

Then at a, the multiplicity in f, g and h is m. So, f/h and g/h have no root at a. In addition, y_{k-1} is a root of g/h, x_{k+m+1} is a root of f/g and there are no other roots of g/h or f/h between y_{k-1} and x_{k+m+1}. So the g/h and f/h have the strict interlacing property on this interval.

Case II: Suppose that the first and last roots in the segment are roots of f. So,

$$y_{k-1} < x_k = y_k = \cdots = x_{k+m} < y_{k+m}.$$

Then at a, the multiplicity in f is m, and the multiplicity in g is $m-1$. So, the multiplicity in h is $m-1$. Thus, the multiplicity in f/h is one and in g/h is zero. Therefore, at a, f/h has simple root and g/h has no root. In addition, $y_{k-1} < x_k < y_{k+m}$ are roots of g/h,f/h,g/h, respectively so the strict interlacing property(see Lemma 3.2) holds on the interval.

Note that there are two other cases with f and g switched. These are proved similarly. Now we have shown that f/h and g/h have the strict interlacing property around any common root of f and g. This suffices to prove $g/h \prec f/h$ $(g/h \triangleleft f/h.)$ $\square$

Example 3.5. An easy example will clarify the proof of Lemma 3.4. Suppose f and g have three roots as shown in Figure 3.1, so that $g \trianglelefteq f$.

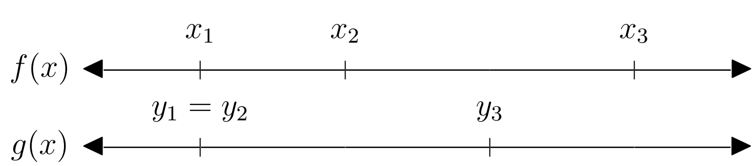

Figure 3.1: $g \trianglelefteq f$ interlacing

Note that x_1 is a simple root of f and a double root of g. Thus, $\gcd(f,g) = h = (x - x_1)$ Hence, the roots of f/h and g/h are as in Figure 3.2

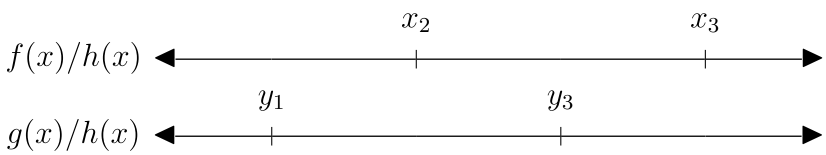

Figure 3.2: $g/h \lhd f/h$ interlacing

Now, x_1 is not a root of f/h and is a simple root of g/h. Thus, the roots of f/h and g/h are strictly interlaced, and $g/h \lhd f/h$.

Lemma 3.6. *If $g \prec f$ or $g \lhd f$, and f and g have leading coefficients of the same sign, then*

(1) $g(x_i)f'(x_i) > 0$ *at all roots x_i of f,*

(2) $f(y_i)g'(y_i) < 0$ *at all roots y_i of g.*

Proof. Suppose first that $g \prec f$ have roots as in Figure 3.3 and positive leading coefficients. Then, since all roots of g are simple, at each root, g changes sign. Specifically $g'(y_1) > 0$, $g'(y_2) < 0$ and $g'(y_3) > 0$. Because of the interlacing, the sign of f at roots of g is opposite to g'. Thus, $f(y_i)g'(y_i) < 0$ at all roots y_i of g. Similarly, $f'(x_i)g(x_i) > 0$ at all roots x_i of f. The proof of the claim for general f and g is similar but with cumbersome indexing. $\square$

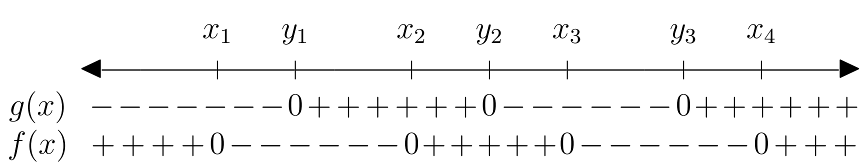

Figure 3.3: Strict interlacing.

Theorem 3.7. *(f, g) is a Sturm sequence if and only if $f, g \in \mathbb{R}[x]$ have leading coefficients of the same sign and $f \preceq g$ or $f \trianglelefteq g$, or $f, g \in \mathbb{R}[x]$ have leading coefficients of the opposite sign and $g \preceq f$ or $g \trianglelefteq f$.*

Proof. Suppose that (f, g) is a Sturm sequence and f and g have leading coefficients of the same sign. Because $\mathcal{V}(\infty) = 0$, for all $a \in \mathbb{R}[x]$, we have that $\mathcal{V}(a) = \mathcal{V}(g(a), f(a)) = \mathcal{N}[g; a, \infty] - \mathcal{N}[f; a, \infty]$.

Since $\mathcal{V}(g(a), f(a)) = 0$ or 1, the number of roots of g to the right of a is equal to the number of roots of f to the right of a or is one less. Either way, the roots of f and g are interlaced. So that $g \trianglelefteq f$ or $g \preceq f$.

When $\mathcal{V}(-\infty) = 0$, the number of roots of f is equal to the number of roots of g, and $f \trianglelefteq g$. When $\mathcal{V}(-\infty) = 1$, f has one more root than g, and the interlacing will be in the form $g \preceq f$.

Therefore, in all cases, the roots of f and g are interlaced. Hence if (f, g) is a Sturm sequence, then the roots are interlaced.

Conversely, suppose $f, g \in \mathbb{R}[x]$ have leading coefficients of the same sign and $f \preceq g$ or $f \trianglelefteq g$. Suppose that the real roots of f and of g are as indexed as described in Definition 3.1. Factoring these polynomials over $\mathbb{R}$ we get

$$f(x) = (x - x_1)(x - x_2) \cdots (x - x_m)\, F(x)$$

$$g(x) = (x - y_1)(x - y_2) \cdots (x - y_n)\, G(x)$$

where $F(x)$ and $G(x)$ are polynomials without real roots. Since $F(x)$ and $G(x)$ have no real roots, they have constant signs which coincide with the signs of the leading terms of f and g. Changing the signs of F and G does not change $\mathcal{V}(g(x), f(x))$, so it is harmless to assume that $F(x) = G(x) = 1$.

If a is not a root of f or of g, then $[g(a)] = [(-1)^k]$ where $k = \mathcal{N}[g; a, \infty]$ is the number of roots of g greater than a, and $[f(a)] = [(-1)^l]$ where $l = \mathcal{N}[f; a, \infty]$ is the number of roots of f greater than a.

Let $\mathcal{V}(x) = \mathcal{V}\,(\,g(x)\,,\,f(x)\,)$. Then $\mathcal{V}(\infty) = 0$ because the leading coefficients of f and g have the same sign. Because of the interlacing, for any $a \in \mathbb{R}$ that is not a root of f or of g, we have either $\mathcal{N}[g; a, \infty] - \mathcal{N}[f; a, \infty] = 0$ or $\mathcal{N}[g; a, \infty] - \mathcal{N}[f; a, \infty] = 1$. In the first case, $f(a)$ and $g(a)$ have the same sign and so $\mathcal{V}(a) = 0$. In the second case, $f(a)$ and $g(a)$ have opposite signs and $\mathcal{V}(a) = 1$. In either case we have $\mathcal{V}(a) = \mathcal{N}[g; a, \infty] - \mathcal{N}[f; a, \infty]$.

Thus, by Lemma 2.3, $(\,f\,,\,g\,)$ is a Sturm sequence. We have shown that the claim holds if f and g have leading coefficients of the same sign.

If f and g have leading coefficients of opposite sign, then the claim follows from the Lemma 2.5 with f and g switched. Specifically, if $f, g \in \mathbb{R}[x]$ have leading coefficients of the opposite sign and $g \preceq f$ or $g \trianglelefteq f$, then $(\,g\,,\,f\,)$ is a Sturm sequence. Therefore, if the roots of f and g are interlaced in one of the forms $f \preceq g$, $g \preceq f$, $f \trianglelefteq g$ or $g \trianglelefteq f$, then $(\,g\,,\,f\,)$ is a Sturm sequence.

To conclude, $(\,f\,,\,g\,)$ is a Sturm sequence if and only if the roots of f and g are interlaced in one of the forms $g \preceq f, f \preceq g$, $f \trianglelefteq g$ or $g \trianglelefteq f$ and $f, g \in \mathbb{R}[x]$. □

Theorem 3.8. *Let $S = (\,f_0\,,\,f_1\,,\,\ldots\,,\,f_n\,)$ be a sequence of polynomials in $\mathbb{R}[x]$, and all having leading coefficients of the same sign, and $f_i \preceq f_{i+1}$ for all i. Then S is a Sturm sequence.*

Proof. By Theorem 3.7, for each i, (f_i, f_i+1) is a Sturm sequence. By Lemma 2.4(2), we have that S is a Sturm Sequence. $\square$

Perhaps an example of a Sturm sequence will help to make Theorem 3.8 clearer. Set

$$f_0(x) = 1$$

$$f_1(x) = x$$

$$f_2(x) = (x+2)(x-2)$$

$$f_3(x) = (x+3)x(x-3)$$

$$f_4(x) = (x+4)(x+1)(x-1)(x-4)$$

We have chosen these polynomials so that $f_0 \prec f_1 \prec f_2 \prec f_3 \prec f_4$, as can be seen in Figure 3.4. The values of $\mathcal{V}(x)$ makes it clear that $(f_0, f_1, f_2, f_3, f_4)$ is a Sturm sequence.

	$-5 \quad -4 \quad -3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5$
$f_0(x)$	+++++++++++++++++++++++++++++++
$f_1(x)$	−−−−−−−−−−−−−−− 0+++++++++++++++
$f_2(x)$	++++++++++ 0−−−−−−−−−0++++++++++
$f_3(x)$	−−−−−−−0+++++++0−−−−−−−0++++++++
$f_4(x)$	+++++0−−−−−−−0++++0−−−−−−−0+++++
$\mathcal{V}(x)$	4 4 4 4 4 3 3 3 3 3 3 3 2 2 2 2 1 1 1 1 1 1 1 0 0 0 0 0

Figure 3.4: $f_{i-1} \prec f_i$

For more information about interlacing see [4].

CHAPTER 4

Rolle's Theorem

Theorem 4.1. *(Rolle's Theorem) [6, Section 4.2] Suppose that $a < b$ are real numbers, $f : \mathbb{R} \to \mathbb{R}$ is a continuous function on the closed interval $[a, b]$ that is differentiable on the open interval (a, b), and $f(a) = f(b) = 0$. Then there exists a real number c in the open interval (a, b) such that $f'(c) = 0$.*

Since polynomials in $\mathbb{R}[x]$ are continuous and differentiable functions on $\mathbb{R}$, this theorem says that between any two roots of a polynomial f there is a root of f'.

Rolle's Theorem doesn't say whether there is only one such root of f'. For example, the polynomial $f(x) = (x^2 - 1)(3x^2 + 1)$ has two real roots, namely ± 1, but $f'(x) = 4x(3x^2 - 1)$ has three real roots on the interval $[-1, 1]$ as seen in the graph:

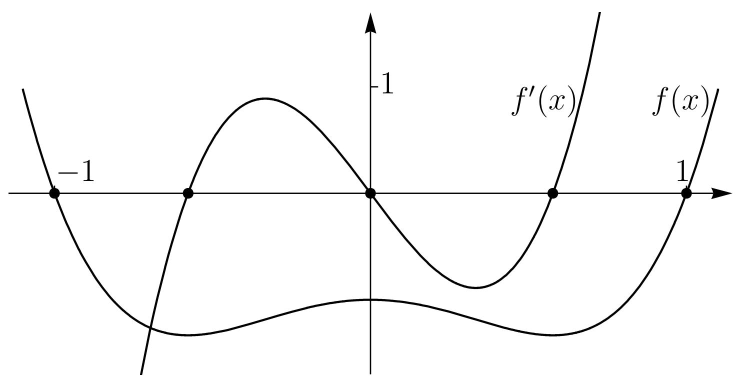

Figure 4.1: Example of Rolle's Theorem

Since we need Rolle's Theorem only for polynomials we provide a simpler proof in this special case.

Theorem 4.2. (*Rolle's Theorem*) *Let x_1 and x_2 be distinct real roots of a polynomial $f \in \mathbb{R}[x]$. Then f' has a real root strictly between x_1 and x_2.*

Proof. It suffices to show this for two adjacent roots of f such that $x_1 < x_2$. In this case, f has the form

$$f(x) = (x - x_1)^{m_1}(x - x_2)^{m_2} g(x)$$

where m_1 and m_2 are the multiplicities of the roots at x_1 and x_2, and $g(x) \in \mathbb{R}[x]$ has no roots in the interval $[x_1, x_2]$. The condition on g means that the sign of g is constant on $[x_1, x_2]$, so, in particular, $g(x_1)$ and $g(x_2)$ are nonzero and have the same sign.

Now we calculate the derivative of f using the product and chain rules:

$$\begin{aligned} f'(x) &= m_1(x - x_1)^{m_1 - 1}(x - x_2)^{m_2}\, g(x) + m_2(x - x_1)^{m_1}(x - x_2)^{m_2 - 1}\, g(x) \\ &\quad + (x - x_1)^{m_1}(x - x_2)^{m_2}\, g'(x) \qquad (4.1) \\ &= (x - x_1)^{m_1 - 1}(x - x_2)^{m_2 - 1}\, F(x) \end{aligned}$$

where $F(x) = m_1(x - x_2)\, g(x) + m_2(x - x_1)\, g(x) + (x - x_1)(x - x_2)\, g'(x) \in \mathbb{R}[x]$. Evaluating F at x_1 and x_2 we find

$$F(x_1) = m_1(x_1 - x_2)\, g(x_1) \text{ and } F(x_2) = m_2(x_2 - x_1)\, g(x_2).$$

Hence $F(x_1)$ and $F(x_2)$ are nonzero and have opposite signs. By the Intermediate Value Theorem, $F(c) = 0$ for some $c \in \mathbb{R}$ such that $x_1 < c < x_2$. By (4.1), c is also a root of f'. □

Lemma 4.3. *Let $f \in \mathbb{R}[x]$. If a is a root of f with multiplicity m, then a is a root of f' with multiplicity $m - 1$.*

Proof. Suppose f has root a of multiplicity m. That is,

$$f(x) = (x - a)^m g(x)$$

for some $g \in \mathbb{R}[x]$ such that $g(a) \neq 0$. Then,

$$\begin{aligned} f'(x) &= m(x-a)^{m-1}g(x) + (x-a)^m g'(x) \\ &= (x-a)^{m-1}(mg(x) + (x-a)g'(x)) \\ &= (x-a)^{m-1}G(x), \end{aligned}$$

where

$$G(x) = mg(x) + (x-a)g'(x)$$

Since $G(a) = mg(a) \neq 0$, a is a root of f' with multiplicity $m-1$. □

Two immediate consequences of Lemma 4.3 are worth mentioning.

Lemma 4.4. *Let $f \in \mathbb{R}[x]$*

(1) *f has a multiple root if and only if f and f' have a common root.*

(2) *$a \in \mathbb{R}$ is a simple root of f if and only if $f(a) = 0$ and $f'(a) \neq 0$.*

Lemma 4.5. *Suppose that $f \in \mathbb{R}[x]$ has n real roots (counting multiplicities). Then f' has at least $n-1$ real roots (counting multiplicities) between the smallest and greatest of the real roots of f. Specifically, if $x_1 \leq x_2 \leq \cdots \leq x_n$ are the real roots of f, then f' has real roots $x'_1, x'_2, \ldots, x'_{n-1}$ such that*

$$x_1 \leq x'_1 \leq x_2 \leq x'_2 \leq \cdots \leq x_{n-1} \leq x'_{n-1} \leq x_n. \tag{4.2}$$

Proof. Let $x_1 < x_2 < \cdots < x_k$ be the distinct real roots of f with multiplicities $m_1, m_2, \ldots, m_k$ respectively. Then $m_1 + m_2 + \cdots + m_k = n$. By Lemma 4.3, for $i = 1, 2, \ldots, k$, x_i is a root of f' with multiplicity $m_i - 1$. Moreover, by Theorem 4.2, there is a real root of f' between each pair of real roots of f, giving an additional $k-1$ roots of f'.

Thus f' has at least

$$(m_1 - 1) + (m_2 - 1) + \cdots + (m_k - 1) + (k - 1) = n - 1$$

real roots of these two types. $\square$

Lemma 4.6. *If $f \in \mathbb{R}[x]$ has degree $n \in \mathbb{N}$ and n real roots (counting multiplicities) then:*

(1) *f' has degree $n-1$ and $n-1$ real roots.*

(2) *$f' \preceq f$.*

(3) *$(f'(x), f(x))$ is a Sturm sequence.*

Proof. (1) Suppose f has a degree n and n real roots. By Lemma 4.5, f' has at least $n-1$ real roots (counting multiplicities). Since f' has degree $n-1$, f' has at most $n-1$ real roots.

(2) Since f' has exactly $n-1$ real roots, $x'_1 \leq x'_2 \leq \cdots \leq x'_{n-1}$, equation (4.2) of Lemma 4.5 implies that $f' \preceq f$.

(3) Since $f' \preceq f$, Lemma 3.7 implies that $(f'(x), f(x))$ is a Sturm sequence. $\square$

Theorem 4.7. *Suppose that $f \in \mathbb{R}[x]$ has degree $n \in \mathbb{N}$ and n real roots (counting multiplicities). Then the sequence of derivatives of f,*

$$\big(f^{(n)}(x) \, , \, f^{(n-1)}(x) \, , \, \ldots \, , \, f''(x) \, , \, f'(x) \, , \, f(x) \big)$$

is a Sturm sequence.

Proof. Suppose f has degree n and n real roots (counting multiplicities). Then f' has degree $(n-1)$ and $(n-1)$ real roots by Lemma 4.6. So, $(f'(x), f(x))$ is a Sturm sequence. Similarly, $(f''(x), f'(x))$ is a Sturm sequence. Repeating this process, $(f'''(x), f''(x))$,$(f''''(x), f'''(x))$,......, $(f^{(n)}(x), f^{(n-1)}(x))$ are also Sturm sequences. Hence, by Lemma 2.4,

$$\big(\, f^{(n)}(x)\,,\, f^{(n-1)}(x)\,,\,\ldots\,,\, f''(x)\,,\, f'(x)\,,\, f(x)\,\big)$$

is a Sturm sequence. □

Lemma 4.8. *If f has n real roots, then there exists a Sturm sequence $(f_0, f_1, \ldots, f_n = f)$ with f_0 having no real roots.*

Proof. Suppose $f(x)$ has n real roots. Then $f(x) = g(x)h(x)$ where g has the same roots as f and h has no real roots. Note that $\deg g = n$. By Theorem 4.7, we can choose g so that it's leading coefficients has the same sign as the leading coefficients as f. So

$$\big(\, g^{(n)}(x)\,,\, g^{(n-1)}(x)\,,\,\ldots\,,\, g''(x)\,,\, g'(x)\,,\, g(x)\,\big)$$

is a Sturm sequence. Therefore,

$$\big(\, g^{(n)}(x)\,,\, g^{(n-1)}(x)\,,\,\ldots\,,\, g''(x)\,,\, g'(x)\,,\, f(x)\,\big)$$

is a Sturm sequence. □

Example 4.9. Consider

$$f(x) = (x^2 - 1)^n.$$

Then, $f(x)$ has degree $2n$ and has two real roots each of multiplicity n. Therefore, the sequence of its derivatives is a Sturm sequence. Of particularly interest, is the n^{th} derivative of $f(x)$ which, except for multiplicative constant, is a Legendre polynomial. Specifically, for $n = 0, 1, 2, \ldots$,

$$P_n(x) = \frac{1}{n!\,2^n} \frac{d^n}{dx^n}(x^2 - 1)^n$$

is called the n^{th} **Legendre polynomial**. The first few of these are

$$\begin{aligned}
P_0(x) &= 1 \\
P_1(x) &= x \\
P_2(x) &= \frac{1}{2}(3x^2 - 1) \\
P_3(x) &= \frac{1}{2}(5x^3 - 3x) \\
P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3)
\end{aligned} \tag{4.3}$$

Their graphs on the interval $[-1, 1]$ suggest many interesting properties of these polynomials (see Figure 4.2). For example, $P_n(1) = 1$ and $P_n(-1) = (-1)^n$ for all $n \in \mathbb{N}$.

The property of interest here is that P_n has exactly n distinct real roots strictly between -1 and 1. Let us see why this holds.

The polynomial $(x^2 - 1)^n$ has degree $2n$ and two real roots 1 and -1, both with multiplicity n. By Lemma 4.3 and Theorem 4.2, $\dfrac{d}{dx}(x^2 - 1)^n$ has real roots 1 and -1 with multiplicity $n - 1$ as well as one real root strictly between -1 and 1. Comparing the degree of the polynomial with the number of real roots, we see that the root in the middle must be simple.

Similarly, $\dfrac{d^2}{dx^2}(x^2 - 1)^n$ has real roots 1 and -1 with multiplicity $n - 2$ and two distinct simple real roots strictly between -1 and 1.

Repeating this argument n times we find that $\dfrac{d^n}{dx^n}(x^2 - 1)^n$ no longer has roots -1 and 1, but has n distinct simple real roots strictly between -1 and 1.

Example 4.10. Let $f(x) = x^3 - 3x^2$. Note that f has three real roots, namely

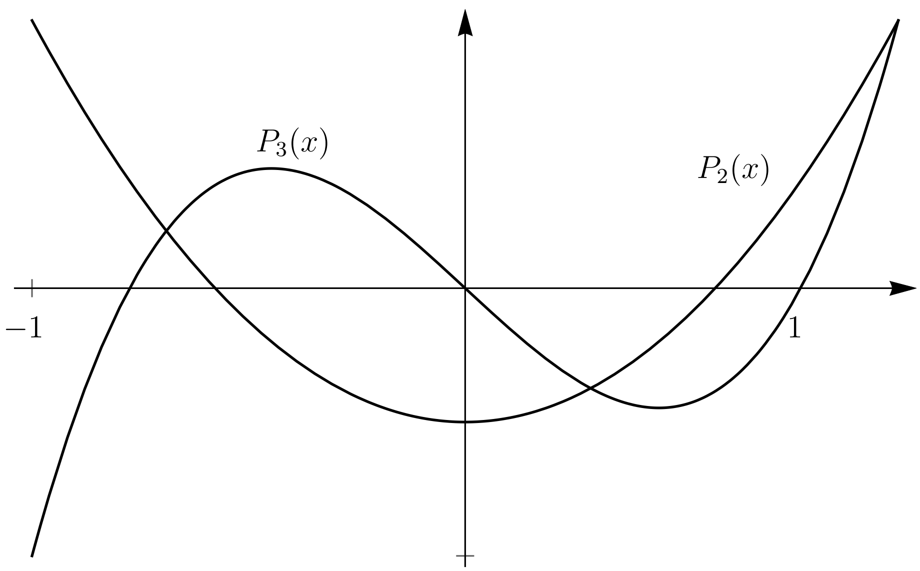

Figure 4.2: Legendre polynomials P_2 and P_3

$(0, 0, 3)$. Hence, we can construct a Sturm sequence from its derivatives.

$$f_0(x) = 6$$

$$f_1(x) = 6(x - 1)$$

$$f_2(x) = 3(x^2 - 2x)$$

$$f(x) = f_3(x) = x^3 - 3x^2$$

So, our Sturm sequence is $S(x) = (6, 6x - 6, 3x^2 - 6x, x^3 - 3x^2)$.

CHAPTER 5

The Wronskian

Definition 5.1. *Let f and g be polynomials in $R[x]$. The **Wronskian** of f and g is defined by $W(f,g) = f\,g' - g\,f'$.*

Lemma 5.2. *Let $f, g, h \in \mathbb{R}[x]$, $a \in \mathbb{R}$. Then,*

(1) $W(f,g) = -W(g,f)$

(2) $W(af,g) = aW(f,g)$

(3) $W(hf,hg) = h^2 W(f,g)$

(4) $W(f,g) = f^2 \dfrac{d}{dx}\left(\dfrac{g}{f}\right)$

Proof. All these properties follow directly from the definition of Wronskian. □

Note that we write $f \leq 0$ if and only if $f(x) \leq 0$ for all $x \in \mathbb{R}$. Similarly, we say $f < 0$ if and only if $f < 0$ for all $x \in \mathbb{R}$.

Lemma 5.3. *Let f be a polynomial in $\mathbb{R}[x]$. If f has same number of real roots as its degree, then $W(f,f') \leq 0$.*

Proof. For convenience, we will prove the claim assuming f has three roots. The general case is similar, but requires some cumbersome notation. Without loss of generality, suppose $f(x) = (x-x_1)(x-x_2)(x-x_3)$. Then,

$$f'(x) = (x-x_1)(x-x_2) + (x-x_2)(x-x_3) + (x-x_1)(x-x_3),$$

and so

$$\frac{f'(x)}{f(x)} = \frac{1}{(x-x_1)} + \frac{1}{(x-x_2)} + \frac{1}{(x-x_3)}.$$

Taking the derivative we get

$$\frac{d}{dx}\left(\frac{f'}{f}\right) = -\frac{1}{(x-x_1)^2} - \frac{1}{(x-x_2)^2} - \frac{1}{(x-x_3)^2},$$

and, with Lemma 5.2(4),

$$\begin{aligned} W(f,f') &= f^2 \frac{d}{dx}\left(\frac{f'}{f}\right) \\ &= -(x-x_2)^2(x-x_3)^2 - (x-x_1)^2(x-x_3)^2 - (x-x_1)^2(x-x_2)^2 \leq 0. \end{aligned}$$

Note that if all the roots of f are simple, then $W(f,f') < 0$. $\square$

For the proof of Lemma 5.4, it is convenient to introduce some new notation. If f has degree n and n real roots $x_1, x_2, \ldots, x_n$, we write $f_i(x) = \dfrac{f(x)}{x-x_i} \in \mathbb{R}[x]$ for $i = 1, 2, \ldots, n$. Then,

$$f_i(x_j) = \begin{cases} 0 & \text{if } i \neq j \\ f'(x_i) & \text{if } i = j \end{cases} \tag{5.1}$$

Lemma 5.4. *Let f and g be polynomials in $\mathbb{R}[x]$. If f and g have same number of roots as their degrees and $g \prec f$, then $W(f,g) < 0$ or $W(f,g) > 0$.*

Proof. Because of Lemma 5.2(1,2), it is harmless to assume that f and g are monic. By Lemma 3.2, all roots of f are simple and we can write

$$f(x) = (x-x_1)(x-x_2)\ldots(x-x_n)$$

for some $x_1, x_2, \ldots, x_n \in \mathbb{R}$. Because $\deg g < \deg f$, we can write the partial fraction expansion of $\dfrac{g}{f}$ as

$$\frac{g}{f} = \sum_i \frac{b_i}{x-x_i}$$

for some constants $b_1, b_2, \ldots, b_n \in \mathbb{R}$. Multiplying by f we get $g(x) = \sum_i b_i f_i(x)$. We show that $b_i > 0$ for all i. Plugging in x_j into $g(x) = \sum_i b_i f_i(x)$ and using (5.1), we get $g(x_j) = \sum_i b_i f_i(x_j) = b_j f'(x_j)$. Because of Lemma 3.6(1), b_i is positive, and so, with Lemma 5.2(4),

$$\frac{W(f,g)}{f^2} = \frac{d}{dx}\left(\frac{g}{f}\right) = \frac{d}{dx}\left(\sum_i \frac{b_i}{x - x_i}\right) = -\sum_i \frac{b_i}{(x - x_i)^2}.$$

Multiplying both sides by f^2 gives $W(f,g) = -\sum_i \dfrac{b_i f^2}{(x - a_i)^2} = -\sum_i b_i f_i^2(x) < 0$. $\square$

Example 5.5. From Example 4.9, the Legendre polynomials P_2 and P_3, given in (4.3), and graphed in Figure 4.2, have Wronskian

$$W(P_3, P_2) = -\frac{3}{4}(5x^4 - 2x^2 + 1).$$

The graph of $W(P_3, P_2)$ in Figure 5.1 shows that $W(P_3, P_2) < 0$ as claimed in Lemma 5.4.

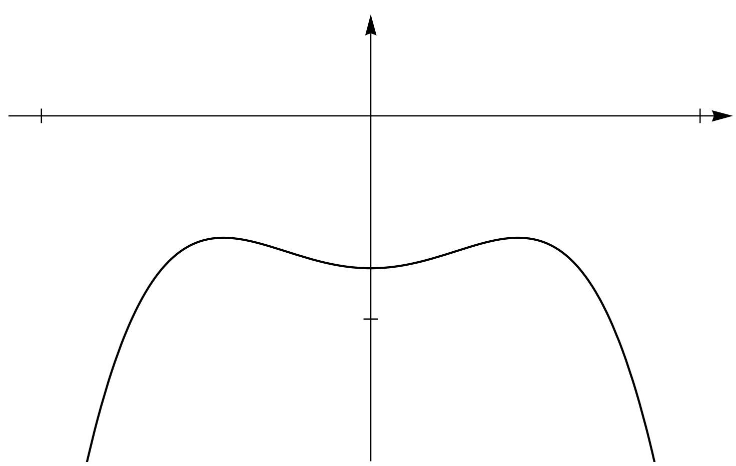

Figure 5.1: The Wronskian of $P_3(x)$ and $P_2(x)$

Lemma 5.6. *Let f and g be polynomials in $\mathbb{R}[x]$. If $W(f,g) < 0$ or $W(f,g) > 0$, then the roots of f and g are strictly interlaced.*

Proof. Without loss of generality, suppose $W(f,g) < 0$. Let x_1 and x_2 be adjacent roots of f. Then, directly from Definition 5.1,

$$W(f,g)(x_1) = -f'(x_1)g(x_1) < 0$$

$$W(f,g)(x_2) = -f'(x_2)g(x_2) < 0.$$

and so $f'(x_1) \neq 0$, $f'(x_2) \neq 0$ and x_1 and x_2 are simple roots. Since x_1 and x_2 are adjacent simple roots, $f'(x_1)f'(x_2) < 0$. Because $f'(x_1)g(x_1)f'(x_2)g(x_2) > 0$, this implies that $g(x_1)g(x_2) < 0$. Hence there is a root y_1 of g between x_1 and x_2. Similarly, all roots of g are simple and between each pair of adjacent roots of g there is a root of f. Finally, we notice that f and g cannot have any common roots since $W(f,g)$ would be zero at any common root. Hence by Lemma 3.2, the roots of f and g are strictly interlaced. $\square$

Example 5.7. Let $f(x) = x^3 - 9x$ and $g(x) = x^4 - 17x^2 + 16$. We want to know if the roots of f and g are interlaced or not. The Wronskian of f and g is

$$\begin{aligned} W(f,g) &= -x^6 + 10x^4 - 105x^2 - 144 \\ &= -x^2(x^2-5)^2 - 144 - 80x^2 \end{aligned}$$

Note that no matter which x you pick, the Wronskian is negative. Therefore, the roots of f and g are interlaced (see Figure 5.20.

$$\begin{array}{rl} & -5\ -4\ -3\ -2\ -1\ \ 0\ \ \ 1\ \ \ 2\ \ \ 3\ \ \ 4\ \ \ 5 \\ & \longleftarrow\!\!+\!\!-\!\!+\!\!-\!\!+\!\!-\!\!+\!\!-\!\!+\!\!-\!\!+\!\!-\!\!+\!\!-\!\!+\!\!-\!\!+\!\!-\!\!+\!\!-\!\!+\!\!-\!\!+\!\!\longrightarrow \\ f(x) & -\,-\,-\,-\,-\,0+\,+\,+\,+\,0\,-\,-\,-\,-\,0+\,+\,+\,+\,+\,+ \\ g(x) & +\,+\,+\,0\,-\,-\,-\,-\,0+\,+\,+\,0\,-\,-\,-\,-\,0+\,+\,+\,+ \end{array}$$

Figure 5.2: The Roots of $f(x) = x^3 - 9x$ and $g(x) = x^4 - 17x^2 + 16$

Theorem 5.8. *Let f and g be polynomials in $\mathbb{R}[x]$. If $\deg f = \deg g + 1$, g has as many real roots as its degree and $W(f, g) < 0$, then f has as many real roots as its degree and $g \prec f$, and g and f have leading coefficients of the same sign.*

Proof. Because of Lemmas 3.2 and 5.6, f and g have simple roots and no common roots. Let $\deg g = m$. Suppose f and g have leading terms $f_0 x^{m+1}$ and $g_0 x^m$, then the leading term of $W(f, g)$ is

$$(f_0 x^{m+1})(m g_0 x^{m-1}) - (g_0 x^m)(m+1)(f_0 x^m) = -f_0 g_0 x^{2m}$$

Because $W(f, g) < 0$, we must have $f_0\, g_0 > 0$. So f and g have leading coefficients of same sign. Without loss of generality, suppose f_0 and g_0 are positive. Let $y_1 < y_2 < \cdots < y_m$ be all the roots of g.

Claim : f has a root on (y_m, ∞).

Since $\lim_{x\to\infty} g(x) = \infty$, $g'(y_m) > 0$. Since $W(f, g)(y_m) = f(y_m) g'(y_m) < 0$, we have $f(y_m) < 0$. Because $\lim_{x\to\infty} f(x) = \infty$, f has a root on (y_m, ∞).

Claim: f has a root on $(-\infty, y_1)$.

(1) If m is odd, $\lim_{x\to-\infty} g(x) = -\infty$ and $g'(y_1) > 0$. Since

$$W(f, g)(y_1) = f(y_1) g'(y_1) < 0,$$

we have $f(y_1) < 0$. Because $\deg f$ is even, $\lim_{x\to-\infty} f(x) = \infty$. So f has a root on $(-\infty, y_1)$.

(2) If m is even, $\lim_{x\to-\infty} g(x) = \infty$ and $g'(y_1) < 0$. Since

$$W(f, g)(y_1) = f(y_1) g'(y_1) < 0,$$

we have that $f(y_1) > 0$. Because $\deg f$ is odd, $\lim_{x\to-\infty} f(x) = -\infty$. So f has a root on $(-\infty, y_1.)$

We now know that f has $m-1$ roots, one between each pair of adjacent root of g, and also a root on (y_m, ∞) and another on $(-\infty, y_1)$. Thus f has at least distinct $m+1$ roots. Since degree f is $m+1$, these are all roots of f. Moreover, the roots of g and f are interlaced so that $g \prec f$. □

Note that this theorem makes it easy to determine that the roots of f and g are interlaced without knowing what the roots are.

Theorem 5.9. *Let $S = (f_0, f_1, f_2, \ldots, f_n)$ be a sequence of polynomials in $\mathbb{R}[x]$ such that*

(1) *f_0 has as many real roots as its degree*

(2) $\deg f_{i+1} = \deg f_i + 1$ *for all i*

(3) $W(f_{i+1}, f_i) < 0$ *for all i*

Then S is a Sturm sequence. In particular, each f_i has as many real roots as its degree.

Proof. By induction from Theorem 5.8, $f_i \prec f_{i+1}$ for all i, each f_i has as many real root as its degree and all the polynomials have leading coefficients of the same sign. By Theorem 3.8, S is a Sturm sequence. □

As an aside, we prove that, if $W(f, g) = 0$, then f, g are linearly dependent. Suppose that $W(f, g) = 0$ with $f \neq 0$. We show g is a constant multiple of f. Since f is nonzero and has at most finitely many roots, there is an open interval I of the real line on which f is never zero. Then g/f is a real differentiable function on I.

The derivative of g/f on I is (Lemma 5.2(4))

$$\frac{d}{dx}\left(\frac{g}{f}\right) = \frac{f\,g' - g\,f'}{f^2} = \frac{W(f,g)}{f^2} = 0.$$

This implies that g/f is constant on I, or, equivalently, there is a constant c such that $cf(x) - g(x) = 0$ for all $x \in I$. This means that every number in I is a root of the polynomial $cf - g$. This is only possible if $cf - g$ is the zero polynomial. Hence $g = cf$ and $\{g, f\}$ is linearly dependent.

Example 5.10. Consider the polynomials

$$f_0(x) = 1$$

$$f_1(x) = x$$

$$f_2(x) = x^2 - 4$$

$$f_3(x) = x^3 - 9x$$

$$f_4(x) = x^4 - 17x^2 + 16$$

The Wronskians

$$W(f_4, f_3) = -x^2(x^2-5)^2 - 144 - 80x^2$$

$$W(f_3, f_2) = -(x^2 - \frac{3}{2})^2 - \frac{135}{4}$$

$$W(f_2, f_1) = -4 - x^2$$

$$W(f_1, f_0) = -1$$

are all negative. Therefore, $(f_0, f_1, f_2, f_3, f_4)$ is a Sturm sequence.

CHAPTER 6

Orthogonal Polynomials

Conveniently, many important sequences of polynomials are defined as Sturm sequences. Orthogonal polynomials have been studied for many years, since they appear in the solutions of many mathematical and physical problems. Examples of sequences of orthogonal polynomial are the Hermite polynomials, the Chebyshev polynomials, and the Legendre polynomials. A Russian mathematician called Pafnuty Lvovich Chebyshev developed the orthogonal polynomials concept in the late 19th century. Let us define the inner product of two polynomials f and g by

$$\langle f, g\rangle = \int_a^b f(x)g(x)h(x)dx$$

where $a, b \in \mathbb{R}$ or $\pm\infty$ and the weight function $h(x) \geq 0$ on the interval (a, b) or (when a and b are finite) on $[a, b]$.

For any such inner product we can construct a sequence $(P_0, P_1, P_2, \dots)$ of polynomials satisfying the conditions:

$\deg P_n = n$, $\langle P_n, P_m\rangle = 0$ for $n \neq m$. $(P_0, P_1, P_2, \dots)$ is called a sequence of orthogonal polynomials with respect to the weight function $h(x)$. It is harmless to assume all the polynomials in the sequence have leading coefficients of the same sign. In that circumstance, any three consecutive polynomials of a sequence of orthogonal polynomials are related by a recurrence formula

$$C_iP_{i-1}(x) + P_{i+1}(x) = (A_ix + B_i)P_i(x).$$

where $C_i, A_i, B_i > 0$, with $i \in \mathbb{Z}^+$ [7].

Example 6.1. Suppose $h(x) = 1$ and $a = -1$ and $b = 1$. Then you get Legendre polynomials:

$$P_0(x) = 1 \qquad P_1(x) = x \qquad P_2(x) = \frac{1}{2}(3x^2 - 1)$$
$$P_3(x) = \frac{1}{2}(5x^3 - 3x) \qquad P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$$

These polynomials satisfy the recurrence relation

$$(n+1)P_{n+1}(x) + nP_{n-1}(x) = (2n+1)xP_n(x).$$

This information can be found in [2].

Example 6.2. Suppose $h(x) = \dfrac{1}{\sqrt{1-x^2}}$ and $a = -1$ and $b = 1$. Then you get the Chebyshev polynomials:

$$T_0(x) = 1 \qquad T_1(x) = x \qquad T_2(x) = 2x^2 - 1 \qquad T_3(x) = 4x^3 - 3x \qquad T_4(x) = 8x^4 - 8x^2 + 1.$$

These polynomials satisfy the recurrence relation

$$T_{i+1}(x) = 2x\,T_i(x) - T_{i-1}(x) \tag{6.1}$$

for $i = 1, 2, 3, \ldots$. This information can be found in [2].

Lemma 6.3. *Let $f, g, h \in \mathbb{R}$ satisfy $af + bg = qh$ where $a, b \in \mathbb{R}$ are positive and $q \in \mathbb{R}[x]$ satisfies $q' \geqslant 0$. If $W(h, g) < 0$, then $W(f, h) < 0$.*

Proof. Using $af' + bg' = q'h + qh'$, we get

$$\begin{aligned} a\,W(f,h) &= afh' - af'h \\ &= (qh - bg)h' - (q'h + qh' - bq')h \\ &= b\,W(h,g) - q'h^2. \end{aligned}$$

since $a, b > 0$, $q' > 0$ and $W(h, q) < 0$, we get $W(f, g) < 0$. $\square$

Theorem 6.4. *Let $S = (f_0, f_1, f_2, \ldots, f_n)$ be a sequence of polynomials in $\mathbb{R}[x]$ such that*

(1) *f_0 has as many real roots as its degree*

(2) *$W(f_1, f_0) < 0$ and $\deg f_1 = \deg f_0 + 1$*

(3) *$a_i f_{i+1} + b_i f_{i-1} = q_i f_i$ for all i where $a_i, b_i > 0$ and q_i is a degree one polynomial with positive leading coefficient*

Then S is a Sturm sequence and for each i, f_i has as many real roots as its degree.

Proof. By induction from Lemma 6.3, $W(f_{i+1}, f_i) < 0$ for all i. Because of (3), $\deg f_2 = \deg q_1 f_1 = \deg f_1 + 1$. And by induction, $\deg f_{i+1} = \deg f_i + 1$ for all i. By Theorem 5.9, S is a Sturm sequence. $\square$

Lemma 6.5. *Any sequence of orthogonal polynomials is a Sturm Sequence.*

Proof. First, note that P_0 is constant. So, it has as many roots as its degree. Therefore, the first condition is satisfied. Now, P_1 is a degree one polynomial with positive leading coefficients. Thus, $W(P_1, 0) < 0$. By [3], the sequence of orthogonal polynomials satisfies the following recurrence relation

$$C_i P_{i-1}(x) + P_{i+1}(x) = (A_i x + B_i) P_i(x).$$

where $C_i, A_i, B_i > 0$, $i \in \mathbb{Z}^+$. Let $f_i = C_i P_i(x)$. Therefore, by Theorem 6.4, the sequence of orthogonal polynomials is a Sturm sequence. $\square$

Consequently, note that the sequence of orthogonal polynomials satisfies all Sturm sequence properties. For example, the n^{th} P_n has degree n and n real roots and the roots of P_n and P_{n-1} are interlaced.

CHAPTER 7

Standard Sturm Sequence

In this chapter, we discuss Sturm sequence that appeared in Sturm's original research we call it standard Sturm sequence. After we define this Sturm sequence, we will prove it is a Sturm sequence according to our "more" generalized definition. For more information about this standard Sturm sequence see [5].

Definition 7.1. *Let $f \in \mathbb{R}[x]$. The* ***standard Sturm sequence*** *for f is defined as follows:*

First set $f_0 = f$ and $f_1 = f'$. For $i = 1, 2, 3, \ldots$ define

$$f_{i+1}(x) = q_i(x) f_i(x) - f_{i-1}(x) \tag{7.1}$$

where q_i is the quotient when f_{i-1} is divided by f_i. Thus f_{i+1} is the negative of the remainder when f_{i-1} is divided by f_i. This sequence ends for some $k \in \mathbb{N}$ when $f_k = \gcd(f, f')$. Then $S = (f_k = gcd(f, f'), f_{k-1}, \ldots, f_1 = f', f_0 = f)$ is the standard Sturm sequence for f. Note the reversal of the order of the indexing from previous Sturm sequences.

Theorem 7.2. *Let $(f_0, f_1, f_2, \ldots, f_n)$ be a sequence of polynomials in $\mathbb{R}[x]$ with the following properties:*

(1) *f_0 has no real roots.*

(2) *If $f_i(r) = 0$ for some $r \in \mathbb{R}$ and $i = 1, 2, \ldots, n-1$, then $f_{i-1}(r) f_{i+1}(r) < 0$,*

(3) *$f_{n-1} \prec f_n$ and these polynomials have leading coefficients of the same sign.*

Then $(f_0, f_1, f_2, \ldots, f_n)$ is a Sturm sequence.

Proof. To prove the claim, it suffices to show that

(1) $\mathcal{V}(x)$ does not change at any root of f_i with $0 < i < n$

(2) $\mathcal{V}(x)$ decreases by one at any root of f_n

Case 1: If $f_i(r) = 0$ for some $r \in \mathbb{R}$, $0 < i < n$, then one of the intermediate polynomials has a root at r. Then, because $f_{i-1}(r)f_{i+1}(r) < 0$, f_{i-1} and f_{i+1} have opposite signs. Since f_{i-1} and f_{i+1} cannot have a zero in a sufficiently small neighborhood I containing r, they cannot change sign on I. Thus, the sign of $f_i(x)$ on I does not effect $\mathcal{V}(x)$. Hence, $\mathcal{V}(x)$ is constant on I. For example, consider the function diagrammed in Figure 2.1. we have that $f_2(-2) = 0$ and $f_1(x) < 0$ in the interval (-3,-1) and $f_3(x) > 0$ in the same interval.

Case 2: Suppose $f_n(r) = 0$ for some $r \in \mathbb{R}$. Since $f_{n-1} \prec f_n$, r is a simple root of f_n and by 3.6, $f_n'(r)f_{n-1}(r) > 0$. Let I be an interval containing r and no other roots of f_n or f_{n-1}. Since $f_n(r) = 0$, f_{n-1} has constant sign on I. If f_{n-1} is positive on I, then $f_n'(r) > 0$. So $f_n(r)$ is positive to the right of r and negative to the left. That is f_n and f_{n-1} have opposite signs to the left and same sign to the right of r. Thus $\mathcal{V}(x)$ decreases by one from the left of r to the right. □

Lemma 7.3. *Let $(f_0, f_1, f_2, \ldots, f_n)$ be a sequence of polynomials in $\mathbb{R}[x]$ such that*

(1) *f_0 is a nonzero constant polynomial,*

(2) *there are polynomials $q_1, q_2, \ldots, q_{n-1} \in \mathbb{R}[x]$ such that*

$$f_{i-1}(x) + f_{i+1}(x) = q_i(x)f_i(x) \tag{7.2}$$

for $i = 1, 2, \ldots, n-1$.

(3) *$f_{n-1} \prec f_n$ and have leading coefficients of same sign.*

Then $(f_0, f_1, f_2, \ldots, f_n)$ *is a Sturm sequence. In addition,* $\gcd(f_{i-1}, f_i) = 1$ *for all* $i = 1, 2, \ldots, n$.

Proof. We prove the last claim first. From Theorem 7.2 we get

$$\gcd(f_{i+1}, f_i) = \gcd(q_i f_i - f_{i-1}, f_i) = \gcd(f_{i-1}, f_i)$$

So, by induction, Using (7.2) repeatedly we get

$$\gcd(f_i, f_{i-1}) = \gcd(f_{i-1}, f_{i-2}) = \gcd(f_{i-1}, f_{i-2}) = \cdots = \gcd(f_1, f_0) = 1.$$

For the last of these equalities we have used the fact that f_0 is a nonzero constant polynomial.

Now suppose that $f_i(r) = 0$ for some $r \in \mathbb{R}$. Since $\gcd(f_{i-1}, f_i) = 1$, r cannot be a root of f_{i-1} (otherwise $x - r$ would be a common factor of f_i and f_{i-1}). Thus $f_{i-1}(r) \neq 0$. In this circumstance, (7.2) becomes $f_{i-1}(r) + f_{i+1}(r) = 0$ and so $f_{i+1}(r)$ is also nonzero and has opposite sign to $f_{i-1}(r)$. In particular, $f_{i-1}(r)f_{i+1}(r) < 0$. By Theorem 7.2, $(f_0, f_1, f_2, \ldots, f_n)$ is a Sturm sequence. $\square$

Theorem 7.4. *The standard Sturm sequence for* $f \in \mathbb{R}[x]$ *is a Sturm sequence.*

Proof. Suppose $h = gcd(f, f')$. Then by Theorem 7.2, h divides every polynomial in the standard Sturm sequence for f. Here we have $f = f_0 = hg_0$, $f' = f_1 = hg_1$, $f_2 = hg_2$, ... $f_k = hg_k$ for some polynomials $g_0, g_1, \ldots, g_k$. By Lemma 4.6, $f_1 \preceq f_0$. So by Lemma 3.4 $g_1 \prec g_0$. Moreover, we can cancel h from the recurrence (7.1) to get

$$g_{i+1}(x) = q_i(x)g_i(x) - g_{i-1}(x)$$

for all i. Finally, $f_k = \gcd(f, f') = h$. So $g_k = 1$ has no real roots. By Lemma 7.3,

$(g_k, g_{k-1}, \ldots, g_0)$ is a Sturm sequence. Then, by Lemma 2.4, $(f_k, f_{k-1}, \ldots, f_0)$= $(hg_k, hg_{k-1}, \ldots, hg_0)$ is a Sturm sequence. $\square$

Example 7.5. Let us see how Theorem 7.4 applies to a simple case, the quadratic polynomial $f(x) = x^2 + bx + c$.

To form the Sturm sequence we start with $f_0(x) = f(x)$ and $f_1(x) = f_0'(x) = 2x+b$. The remainder on dividing f_0 by f_1 is $-(b^2-4c)/4$, so we set $f_3(x) = b^2-4c = \Delta(f)$, a constant polynomial. If $b^2 - 4c \neq 0$ the standard Sturm sequence is

$$S(x) = [\, f_0\,,\ f_1\,,\ f_2\,] = \left[\, x^2 + bx + c\,,\ 2x + b\,,\ b^2 - 4c\,\right].$$

Evaluating this at $\pm\infty$ we get

$$S(-\infty) = \left[\, 1\,,\ -1\,,\ b^2 - 4c\,\right] \qquad S(\infty) = \left[\, 1\,,\ 1\,,\ b^2 - 4c\,\right].$$

If $b^2 - 4c > 0$, then $\mathcal{V}(-\infty) = 2$, $\mathcal{V}(\infty) = 0$, and so the number of real roots of f is $\mathcal{V}(-\infty) - \mathcal{V}(\infty) = 2$. If $b^2 - 4c < 0$, then $\mathcal{V}(-\infty) = 1$, $\mathcal{V}(\infty) = 1$, and so the number of real roots of f is $\mathcal{V}(-\infty) - \mathcal{V}(\infty) = 0$.

Evaluating the Sturm sequence at $x = 0$ gives

$$S(0) = \left[\, c\,,\ b\,,\ b^2 - 4c\,\right].$$

We suppose $c \neq 0$ so that $\mathcal{V}(0)$ is defined. If $b^2 - 4c < 0$, then c must be positive and so $\mathcal{V}(0) = 1$, independent of the sign of b. In this case the number of positive real roots of f is $\mathcal{V}(0) - \mathcal{V}(\infty) = 1 - 1 = 0$. This is no surprise since there are no real roots, positive or negative, if $b^2 - 4c < 0$.

If $b^2 - 4c > 0$, then, since $\mathcal{V}(\infty) = 0$, the number of positive real roots of f

is $\mathcal{V}(0)$:

$$\mathcal{V}(0) = \begin{cases} 1 & \text{if } c < 0 \\ 0 & \text{if } c > 0 \text{ and } b > 0 \\ 2 & \text{if } c > 0 \text{ and } b < 0. \end{cases}$$

Note that when $b = 0$, we can ignore the zero. Also if $c = 0$, that means we are calculating at a zero of the polynomial which is not allowed.

Example 7.6. Let us use Theorem 7.4 to determine the number of real, positive and negative roots of $g(x) = x^3 + p\,x + q$ with $p, q \in \mathbb{R}$. We also assume, for simplicity, that $q \neq 0$ so that 0 is not a root.

We set $g_0(x) = g(x)$ and $g_1(x) = g_0'(x) = 3x^2 + p$. The remainder on dividing g_0 by g_1 is $(2p\,x + 3q)/3$ so we set $g_2(x) = -2p\,x - 3q$. If $p = 0$, then we don't need to go further since g_2 is a constant polynomial. This special case is left to the reader, and we suppose here that $p \neq 0$.

The remainder on dividing g_1 by g_2 is $(4p^3 + 27q^2)/4p^2$ and so we choose $g_3(x) = -(4p^3 + 27q^2) = \Delta(g)$, a constant polynomial. Thus if $\Delta(g) \neq 0$, the standard Sturm sequence is

$$S(x) = [\,g_0(x)\,,\, g_1(x)\,,\, g_2(x)\,,\, g_3(x)\,] = \left[\,x^3 + p\,x + q\,,\, 3x^2 + p\,,\, -2p\,x - 3q\,,\, \Delta(g)\,\right].$$

Evaluating the Sturm sequence at $-\infty$ we get

$$S(-\infty) = [\,-1\,,\, 1\,,\, p\,,\, \Delta(g)\,]\,.$$

If $\Delta(g) > 0$ then $p < 0$, and consequently $S(-\infty) = [\,-1\,,\, 1\,,\, -1\,,\, 1\,]$ with 3 sign variations. If $\Delta(g) < 0$, then $S(-\infty) = [\,-1\,,\, 1\,,\, p\,,\, -1\,]$ with 2 sign variations, independent of the sign of p. Thus

$$\mathcal{V}(-\infty) = \begin{cases} 3 & \text{if } \Delta(g) > 0 \\ 2 & \text{if } \Delta(g) < 0. \end{cases}$$

Similarly, evaluating the Sturm sequence at ∞ we get

$$S(\infty) = [\,1\,,\,1\,,\,-p\,,\,\Delta(g)\,]$$

and

$$\mathcal{V}(\infty) = \begin{cases} 0 & \text{if } \Delta(g) > 0 \\ 1 & \text{if } \Delta(g) < 0. \end{cases}$$

Therefore the number of real roots of g is

$$\mathcal{V}(-\infty) - \mathcal{V}(\infty) = \begin{cases} 3 & \text{if } \Delta(g) > 0 \\ 1 & \text{if } \Delta(g) < 0. \end{cases} \qquad (7.3)$$

To find the number of positive real roots we evaluate the Sturm sequence at $x = 0$:

$$S(0) = [\,q\,,\,p\,,\,-q\,,\,\Delta(g)\,]\,.$$

Note that, independent of the value of p, the sequence $[\,q\,,\,p\,,\,-q\,]$ has exactly one sign change. So

$$\mathcal{V}(0) = \begin{cases} 2 & \text{if } q\Delta(g) > 0 \\ 1 & \text{if } q\Delta(g) < 0. \end{cases}$$

Thus the number of positive real roots of g is

$$\mathcal{V}(0) - \mathcal{V}(\infty) = \begin{cases} 2 & \text{if } q > 0 \text{ and } \Delta(g) > 0 \\ 0 & \text{if } q > 0 \text{ and } \Delta(g) < 0 \\ 1 & \text{if } q < 0. \end{cases} \qquad (7.4)$$

If $\Delta(g) > 0$, then, by (7.3), g has three real roots. This means that we can form another Sturm sequence from the derivatives of g, by Theorem 4.7,

$$S(x) = [\,g(x)\,,\,g'(x)\,,\,g''(x)\,,\,g'''(x)\,] = \left[\,x^3 + p\,x + q\,,\,3x^2 + p\,,\,6x\,,\,6\,\right].$$

Evaluating this at 0 and ∞, keeping in mind that $p < 0$ because $\Delta(g) > 0$, we get

$$S(0) = [\,q\,,\,-1\,,\,0\,,\,1\,] \text{ and } S(\infty) = [\,1\,,\,1\,,\,1\,,\,1\,],$$

implying that the number of positive roots of g is

$$\mathcal{V}(0) - \mathcal{V}(\infty) = \begin{cases} 2 & \text{if } q > 0 \\ 1 & \text{if } q < 0, \end{cases}$$

in agreement with (7.4).

Example 7.7. Here is an example of standard Sturm sequence of the polynomial $f = x^3 + 3x + 2$.

$$f_3 = x^3 + 3x + 2$$

$$f_2 = 3(1 + x^2)$$

$$f_1 = -2(1 + x)$$

$$f_0 = -6$$

So, standard Sturm sequence $S(x)$ is

$$S(x) = (-6, -2(1 + x), 3(1 + x^2), 2 + 3x + x^3)$$

Evaluating this sequence at -2, 0 and 10 gives

$$S(-2) = (-6, 2, 15, -12)$$

$$S(0) = (-6, -2, 3, 2)$$

$$S(10) = (-6, -22, 303, 1032)$$

$$\mathcal{V}(-2) = 2$$

$$\mathcal{V}(0) = 1$$

$$\mathcal{V}(10) = 1$$

Hence, there is one real root between-2 and 0 and no real roots between 0 and 10. In fact the graph of $f(x) = x^3 + 3x + 2$ shows that there is only one real roots of the polynomial at approximately -0.5

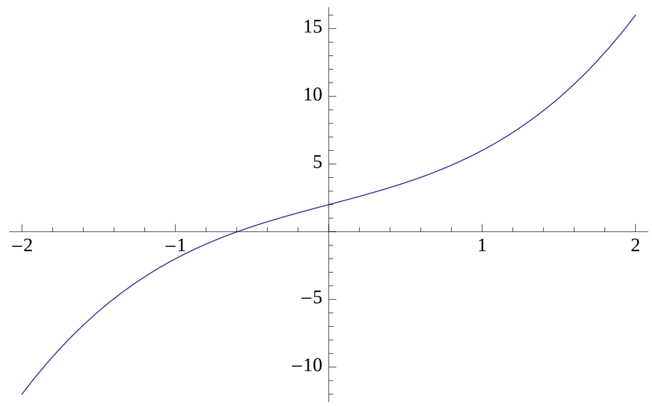

Figure 7.1: The graph of $f(x) = x^3 + 3x + 2$

REFERENCES

[1] H. Antin, Sturm's Problem of the Number of Roots. *in 100 great problems of elementary mathematics: Their history and solution*. New York: Dover Publications, 1965.

[2] W. Beyer, *CRC standard mathematical tables, 28th ed.* Boca Raton, Fla., CRC Press, (1987).

[3] T. Chihara,, A. W. Reid, *An introduction to orthogonal polynomials*, New York: Gordon and Breach, 1978.

[4] S. Fisk, Polynomials with all real roots. *Polynomials, roots, and interlacing, Version 2.* Cornell University, 2008.

[5] N. Jacobson, *Basic algebra,* 2^{nd}*ed.*, San Francisco: W.H. Freeman, 1989.

[6] J. Stewart, *Calculus, Early Transcendentals, 7ed.* Brooks/Cole, 2008.

[7] P. K. Suetin (originator), *Orthogonal polynomials*, Encyclopedia of Mathematics. `http://www.encyclopediaofmath.org`