A Counterexample of a Conjecture Concerning the Root Location of a Real Univariate Polynomial and its Derivative

Lisa Lokteva
lisalokteva@gmail.com

under the direction of
Prof. Boris Shapiro
Department of Mathematics
Stockholm University

Research Academy for Young Scientists
July 10, 2013
Abstract

In this paper, we present a conjecture by Dr. Matthew Chasse related to and precising the Gauss-Lucas Theorem. It states that for every univariate polynomial with real coefficients and every line \( \text{Im}(z) = k > 0 \), there are no more roots of the derivative of the polynomial than roots of the polynomial itself above the line. We numerically test this conjecture using the programming language Mathematica and find counterexamples for polynomials of degree 6, 7 and 8, but not of degree 5. For degree 2, 3 and 4, the trueness of the conjecture can be readily deduced from the Gauss-Lucas theorem.
1 Introduction

It is well known that a polynomial of degree \( n \) has exactly \( n \) roots, counting multiplicity, as stated by the **Fundamental Theorem of Algebra** which was first proved by Gauss in 1799 [1]. For polynomials with degree \( \leq 4 \), formulae making use of the arithmetic operations, radicals and coefficients of the polynomial to obtain the roots exist, but it has been proven impossible to use such methods with polynomials of higher degree [2]. However, it is still sometimes possible to draw conclusions concerning the loci of the roots in the complex plane.

The well-known theorem of Gauss-Lucas claims that all roots of the first derivative of an arbitrary univariate polynomial lie inside the convex hull of the roots of the polynomial itself [3] (see Figure 1 and proof in Appendix A). This gives a geometrical relation between the roots of the polynomial and the roots of its first derivative. Recently, Dr. Matthew Chasse (private communication with Boris Shapiro) formulated the following Conjecture 1, related to the Gauss-Lucas theorem:

**Conjecture 1** (Chasse’s Conjecture). *For every line \( l : \text{Im}(z) = k > 0 \) and every polynomial \( P \in \mathbb{R}[z] \) there are no more zeros of \( P' \) than zeros of \( P \) above \( l \).*

This statement adds an extra constraint to the loci of the roots of \( P' \) giving further information specifying the Gauss-Lucas theorem. For polynomials of degree 2, 3 and 4,
Conjecture 1 can readily be deduced from Gauss-Lucas. This report aims to test the
conjecture for higher degrees and find counterexamples if there are any.

2 Method

The Chasse conjecture was tested by developing Mathematica code; see Appendix B for
full code. The main ideas used to develop the code were as follows:

The program starts by making a polynomial of the desired degree. All coefficients
are chosen at random as uniformly distributed integers between -100 and 100. The aim
is to test all lines $\text{Im}(z) = k > 0$ and count how many roots of the polynomial and its
derivative respectively lie above the line. However, instead of looking at lines we can just
look at the order of the imaginary parts of the roots of the polynomial and its derivative.

We make two counters, \texttt{counter1} for the number of regarded roots of the polynomial
and \texttt{counter2} for the regarded roots of its derivative, and initialise them to 0. We make
two lists sorted in decreasing order; one with the imaginary parts of the roots of the
polynomial and one with the imaginary parts of the roots of its derivative. The sorted
lists can be regarded as stacks with the greatest imaginary parts at the top.

As long as there are elements greater than 0 at the top of any one of the stacks and we
do not know whether the polynomial is a counterexample or not, we choose the greatest
of the top elements of the stacks and remove it, although in the code, we just let the
next root be regarded as the top element of the stack, and do not remove anything. In
case of equality between the top elements of the both stacks, the root of the polynomial
is prioritized. We update the counters based on whether it was a root of the polynomial
or the derivative that was removed. We then check if the number of regarded derivative
roots is greater than the number of regarded roots of the polynomial. In that case, the
loop is escaped and the polynomial is printed, since we know that that polynomial is a
counterexample. In the other case, the new top elements of the stacks are regarded and
the process is repeated. If we have gone through all positive imaginary parts without
counter2 being greater than counter1 any time, then we look at the next polynomial.

The polynomial and derivative roots are obtained using the Mathematica function NSolve. To make sure that the result is accurate, the working precision is set to 20 digits.

The conjecture was tested for degree 5, 6, 7 and 8, 200 000 polynomials of each degree.

3 Results

Figure 2: A counterexample for $n = 7$. The points represent the roots of the polynomial and the stars the roots of its derivative.

A number of counterexamples were found for degree 6, 7 and 8; Figure 3 shows one of them, of degree 7. It has the equation $p(x) = -28 - 30x - 97x^2 - 63x^3 - 77x^4 + 70x^5 + 43x^6 + 89x^7$. Appendix C contains lists of found polynomials of degree 6, 7 and 8 that are counterexamples. No counterexamples were found for degree 5. For details, see Table 1.

4 Discussion

From Table 1, it appears that the conjecture is false for degree 6, 7 and 8 while nothing can be established about degree 5. The limitation of the method is that it is numerical and can give unprecise solutions; however, setting the working precision to 20 digits should
Table 1: Number of tested polynomials and of found counterexamples.

<table>
<thead>
<tr>
<th>Degree</th>
<th>Number of Tested Polynomials</th>
<th>Number of Counterexamples</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>200 000</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>200 000</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>200 000</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>200 000</td>
<td>16</td>
</tr>
</tbody>
</table>

decrease the size of the error. For instance, if we look at the counterexample polynomial $p(x) = -56 + 52x - 97x^2 + 89x^3 - 84x^4 + 36x^5 - 18x^6$, whose numeric roots are listed in Appendix D, we see that the polynomial root with the second largest imaginary part is below the derivative root with the second largest imaginary part, their imaginary parts differing by more than 0.016. Since $0.016 \gg 10^{-19}$, the result should be accurate.

It is easy to deduce that Chasse’s conjecture is true for degree 2, 3 and 4 from the Gauss-Lucas Theorem. Our results show that it is false for degree 6, 7 and 8, and it is presumably false for higher degrees, since the number of counterexamples seems to grow with the degree of the polynomial. However, counterexamples have not yet been found for the fifth degree. The non-trivial case is when the polynomial has exactly one real root. Since there is a formula to solve fourth-degree polynomial equations we can always find the roots of the derivative of a fifth-degree polynomial, but those formulae are very complicated and it is almost impossible to draw conclusions from them by just looking at them, without use of a computer or any other tool. Thus, the problem remains open for the fifth degree, which could be a subject for further research.
Acknowledgements

I would like to give my biggest acknowledgements to Prof. Boris Shapiro for inventing this project and to Alexandre Loiko and my project partner Cesar Höjeberg for helping me with Mathematica. I would also like to thank Johan Henriksson and Mariam Andersson for proof-reading and criticizing this report. Finally, I would like to thank Mikael Ingemyr, Rays, Volvo and Stockholms matematikcentrum for making all this possible.
References


A Proof of Gauss-Lucas Theorem

We start off with some definitions and a lemma in order to prove the Gauss-Lucas theorem.

**Definition 1.** A set $M \subseteq \mathbb{R}^n$ is called convex if and only if the line segment between any two arbitrary points in $M$ is contained in $M$.

*Remark:* A convex set does not have to be bounded, e.g. the upper halfplane of $\mathbb{R}^2$ is convex.

**Definition 2.** A convex hull $K(S)$ to a finite set $S \subseteq \mathbb{R}^n$ is the intersection of all convex sets containing $S$.

Observe that $\mathbb{C} \cong \mathbb{R}^2$, since each complex number $z$ can be written in the form $x + iy$ and thus can be identified with the real vector $(x, y) \in \mathbb{R}^2$. Then we can talk about convexity of sets of complex numbers in exactly the same way as for $\mathbb{R}^2$. The same can be said about the convex hull.

**Lemma 1.** Let $S_n = \{z_1, z_2, z_3, ..., z_n\}$ be a finite set of points in $\mathbb{C}$. Then

$$H(S_n) = \left\{ \sum_{k=1}^{n} a_k z_k : a_k \in \mathbb{R}, a_k \geq 0 \ \forall k \text{ and } \sum_{k=1}^{n} a_k = 1 \right\}$$

is the convex hull of $S_n$.

*Proof.* For $n = 1$, the statement is obvious, since the convex hull of one point is the point itself. Instead, assume that $n \geq 2$.

Let us look at the set $H(S_n)$. Assume that we have two points: $A$ and $B$, belonging to $H(S_n)$. We can write them as linear combinations of $z_1, z_2, ..., z_n$ with some coefficients $a_1, ..., a_n$ and $b_1, ..., b_n$ respectively, all coefficients non-negative and summing up to 1.
The line segment between them is then the set of points

\[
P = \{tA + (1-t)B : 0 \leq t \leq 1\} = \\
= \left\{ t\left(\sum_{k=1}^{n} a_k z_k\right) + (1-t)\left(\sum_{k=1}^{n} b_k z_k\right) : 0 \leq t \leq 1 \right\} = \\
= \left\{ \sum_{k=1}^{n} z_k(ta_k + (1-t)b_k) : 0 \leq t \leq 1 \right\},
\]

which is another linear combination of \(z_1, \ldots, z_n\) with coefficients on the form \(ta_k + (1-t)b_k\).

Their sum is:

\[
\sum_{k=1}^{n} (ta_k + (1-t)b_k) = t\left(\sum_{k=1}^{n} a_k\right) + (1-t)\left(\sum_{k=1}^{n} b_k\right) = t \cdot 1 + (1-t) \cdot 1 = 1
\]

and \(ta_k + (1-t)b_k \geq 0\) for all \(t, a_k, b_k\). Thus, every point on \(P\) is a linear combination of \(z_1, \ldots, z_n\) with non-negative coefficients summing up to 1, which implies that \(P\) is a subset of \(H(S_n)\). Since this holds for arbitrary points \(A\) and \(B\), by Definition 1, the set \(H(S_n)\) is convex.

Now, note that for \(n = 2\), the definition of \(H(S_n)\) is exactly the definition of a line segment, and for two points, the convex hull is the line segment between them. This due to the fact that the line segment is a convex set of points which can not be smaller, since for the set to be convex and include \(S_n\) it has to include the line segment between the two points. Now assume that the statement holds for some \(n\) and \(H(S_n) = \left\{ \sum_{k=1}^{n} a_k z_k : a_k \in \mathbb{R}, a_k \geq 0 \ \forall k \text{ and } \sum_{k=1}^{n} a_k = 1 \right\}\) is the convex hull of the \(n\) points. We add an other point \(z_{n+1}\) and draw all line segments connecting \(z_{n+1}\) with points in the convex hull of \(S_n\). Note that all points lying on those line segments are essential for convex sets including \(S_{n+1}\), by Definition 1. The points \(Q\) on the line segments between a point \(A\) in the convex hull of \(S_n\) and \(z_{n+1}\) can be described as

\[
Q = tA + (1-t)z_{n+1} = t\left(\sum_{k=1}^{n} a_k z_k\right) + (1-t)z_{n+1} = \left(\sum_{k=1}^{n} a_k t z_k\right) + (1-t)z_{n+1}
\]
for a certain $t$. We see that $Q$ is a linear combination of $z_1, z_2, \ldots, z_{n+1}$ with coefficients that are non-negative and add up to 1 ($(\sum_{k=1}^{n} a_k t) + (1-t) = t(\sum_{k=1}^{n} a_k) + 1 - t = t \cdot 1 + 1 - t = 1$). Moreover, those coefficients ($ta_k$ for $z_k$ with $1 \leq k \leq n$, and $1 - t$ for $z_{n+1}$) can be any non-negative real numbers with sum 1, depending on how we choose $t$ and the $a_k$, or equivalently point $A$ and a place on the line segment between $A$ and $z_{n+1}$. That means that any point in the set $H(S_{n+1})$ can be represented as a point on a line segment between $z_{n+1}$ and a point $A$ in $H(S_n)$. Thus, the set of points on all those line segments is exactly the same set as $H(S_{n+1})$ and since we know both that $H(S_{n+1})$ is a convex set and that it only consists of points essential for including $S_{n+1}$ in it, $H(S_{n+1})$ is the convex hull of $S_{n+1}$. The result follows by induction. \hfill \Box

**Theorem 1** (Gauss-Lucas Theorem). *Let $P(z) \in \mathbb{C}[z]$ be a polynomial. The roots of $P'(z)$ lie inside the convex hull of the roots of $P(z)$.*

*Proof.* $P$ can be written as a product of prime factors, $P(z) = \alpha \prod_{k=1}^{n} (z - a_k)$, where $a_1, a_2, \ldots, a_n$ are (not necessarily distinct) complex roots of $P$, $\alpha$ is the leading coefficient of $P$ and $n$ is the degree of $P$. An easy calculation shows that

$$\frac{P'(z)}{P(z)} = \sum_{k=1}^{n} \frac{1}{z - a_k}.$$  

Let $z_0$ be a root of $P'(z)$ but not of $P(z)$. Then

$$\sum_{k=1}^{n} \frac{1}{z_0 - a_k} = \frac{P'(z_0)}{P(z_0)} = 0,$$

so

$$\sum_{k=1}^{n} \frac{1}{z_0 - a_k} = \sum_{k=1}^{n} \frac{z_0 - \overline{a_k}}{|z_0 - a_k|^2} = 0,$$

and hence

$$\sum_{k=1}^{n} \frac{z_0}{|z_0 - a_k|^2} = \sum_{k=1}^{n} \frac{\overline{a_k}}{|z_0 - a_k|^2}.$$
Taking the complex conjugate,

\[ \sum_{k=1}^{n} \frac{z_0}{|z_0 - a_k|^2} = \sum_{k=1}^{n} \frac{a_k}{|z_0 - a_k|^2}, \]

so

\[ z_0 \sum_{k=1}^{n} \frac{1}{|z_0 - a_k|^2} = \sum_{k=1}^{n} \frac{a_k}{|z_0 - a_k|^2}. \]

Denote

\[ x_k = \frac{1}{\sum_{j=1}^{n} \frac{1}{|z_0 - a_j|^2}}. \]

Then

\[ z_0 = \sum_{k=1}^{n} x_k a_k. \]

It is obvious that \( \sum_{k=1}^{n} x_k = 1 \) and that \( x_k \geq 0 \) for every \( k \), so, by Lemma 1, \( z_0 \) lies in the convex hull of the points \( a_1, ..., a_n \).

If \( z_0 \) is a root of both \( P'(z) \) and \( P(z) \), there is nothing to show, since it then is obvious that \( z_0 \) lies in the convex hull of the roots of \( P \).

\[ \square \]

B Mathematica code for testing the conjecture

This code is for testing polynomials of degree 8. However, it is easy to modify the degree by commenting away the first terms of \( \text{poly} \).

```mathematica
For[ii = 0, ii < 100000, ii++,
   isCounterexample = 0;
   a = RandomInteger[{-100, 100}]; b = RandomInteger[{-100, 100}];
   c = RandomInteger[{-100, 100}]; d = RandomInteger[{-100, 100}];
   e = RandomInteger[{-100, 100}]; f = RandomInteger[{-100, 100}];
   g = RandomInteger[{-100, 100}]; h = RandomInteger[{-100, 100}];
   \text{poly} = \ldots; \]
```

10
i = RandomInteger[{-100, 100}];
poly = a x^8 + b x^7 + c x^6 + d x^5 + e x^4 + f x^3 + g x^2 + h x + i;

(*Make lists of roots to the polynomial and its derivative*)
polyder = D[poly, x];
los1 = x /. NSolve[poly == 0, x, Complexes, WorkingPrecision -> 20];
los2 = x /. NSolve[polyder == 0, x, Complexes, WorkingPrecision -> 20];
(*Make lists of the imaginary parts and sort them in decreasing order.*)
listv = Im[los1];
listd = Im[los2];
listv = Sort[listv, Greater];
listd = Sort[listd, Greater];
counter1 = 0;
counter2 = 0;

(*Go through the lists and count roots. Print the polynomial if counterexample.*)
While[(listv[[counter1 + 1]] > 0 || listd[[counter2 + 1]] > 0) && bulle == 0,
    If[listv[[counter1 + 1]] >= listd[[counter2 + 1]], counter1++,
        counter2++ ];
    If[counter1 >= counter2, , isCounterexample = 1];
    If[isCounterexample == 1, Print[poly], ];

(*Print the number of tested polynomials with regular intervals*)
(*of 5000 polynomials.*)
If[Mod[ii, 5000] == 0, Print[ii]];]
Print[Done]
C Lists of found counterexamples

C.1 List of found counterexamples of degree 6

\[ p(x) = 74 + 4x + 92x^2 + 20x^3 + 86x^4 + 14x^5 + 37x^6 \]
\[ p(x) = 94 - 20x + 95x^2 - 43x^3 + 68x^4 - 18x^5 + 18x^6 \]
\[ p(x) = -56 + 52x - 97x^2 + 89x^3 - 84x^4 + 36x^5 - 18x^6 \]

C.2 List of found counterexamples of degree 7

\[ p(x) = -28 - 30x - 97x^2 - 63x^3 - 77x^4 + 70x^5 + 43x^6 + 89x^7 \]
\[ p(x) = -57 + 43x - 60x^2 - 46x^3 - 93x^4 + 50x^5 + 50x^6 + 86x^7 \]
\[ p(x) = 43 + 10x + 61x^2 - 37x^3 + 99x^4 - x^5 + 14x^6 + 54x^7 \]
\[ p(x) = -56 - 14x - 98x^2 - 40x^3 - 76x^4 - 91x^5 - 9x^6 - 55x^7 \]

C.3 List of found counterexamples of degree 8

\[ p(x) = -68 + 56x - 72x^2 - 76x^3 - 58x^4 + 56x^5 - 31x^6 - 88x^7 - 91x^8 \]
\[ p(x) = -71 + 72x - 14x^2 - 19x^3 - 99x^4 - 43x^6 - 23x^7 - 90x^8 \]
\[ p(x) = -68 + 73x - 20x^2 + 30x^3 - 70x^4 + 16x^5 - 73x^6 - 15x^7 - 82x^8 \]
\[ p(x) = 90 + 95x + 49x^2 - 36x^3 + 84x^4 + 43x^5 - 26x^6 + 36x^7 + 20x^8 \]
\[ p(x) = -18 - 53x^2 - 50x^3 - 51x^4 + 57x^5 + 24x^6 + 11x^7 - 87x^8 \]
\[ p(x) = -78 + 64x - 19x^2 + 32x^3 - 92x^4 + 28x^5 - 57x^6 + 64x^7 - 66x^8 \]
\[ p(x) = -35 + 8x - 76x^2 + 23x^3 - 86x^4 + 69x^5 + 21x^6 + 76x^7 + 62x^8 \]
\[ p(x) = 59 + 27x + 62x^2 + 32x^3 + x^4 + 75x^5 + 7x^6 + 32x^8 \]
\[ p(x) = -39 - 14x - 97x^2 - 74x^3 - 76x^4 + 76x^5 + 15x^6 + 23x^7 - 77x^8 \]
\[ p(x) = 25 + 35x + 45x^2 + 18x^3 + 87x^4 + 34x^5 + 72x^6 + 23x^7 + 68x^8 \]
\[ p(x) = -99 - 30x - 26x^2 - 10x^3 - 79x^4 + 45x^5 - 99x^6 + 51x^7 - 83x^8 \]
\[ p(x) = -82 - 26x - 49x^2 - 21x^3 - 78x^4 - 21x^5 - 57x^6 - 43x^7 - 72x^8 \]
\[ p(x) = 66 + 98x - 33x^2 + 88x^3 + 36x^4 + 92x^5 + 99x^7 + 55x^8 \]
$$p(x) = -74 - 56x - 71x^2 + 48x^3 - 93x^4 + 8x^5 - 90x^6 + 69x^7 - 83x^8$$

$$p(x) = -42 + 32x - 24x^2 + 99x^3 + 89x^4 + 68x^5 - 94x^6 - 96x^7 - 67x^8$$

$$p(x) = 78 + 66x - 12x^2 + 58x^3 + 32x^4 + 29x^5 + 52x^6 + 5x^7 + 62x^8$$

### D  Roots of a sixth degree polynomial

Roots of the polynomial $p(x) = -56 + 52x - 97x^2 + 89x^3 - 84x^4 + 36x^5 - 18x^6$ sorted in the decreasing order of the imaginary parts:

- 0.32234764254764610468 + 1.4736197677712608470i
- 0.24630822497500602758 + 0.90952324119780023564i
- 0.92396058242754498074 + 0.82834643878487342653i
- 0.92396058242754498074 - 0.82834643878487342653i
- 0.24630822497500602758 - 0.90952324119780023564i
- 0.32234764254764610468 - 1.4736197677712608470i

Roots of the derivative of the polynomial, sorted in the decreasing order of the imaginary parts:

- 0.24624138556228028560 + 1.07479498204642576627i
- 0.39113094204203928153 + 0.92599008315911425782i
- 0.39192201145802753242
- 0.39113094204203928153 - 0.92599008315911425782i
- 0.24624138556228028560 - 1.07479498204642576627i