STURM’S THEOREM

Given a univariate polynomial with simple roots \( p \) and the sequence of polynomials

\[
\begin{align*}
    p_0(x) &= p(x) \\
    p_1(x) &= p'(x) \\
    p_2(x) &= -\text{rem}(p_0, p_1) = p_1(x)q_0(x) - p_0(x) \\
    p_3(x) &= -\text{rem}(p_1, p_2) = p_2(x)q_1(x) - p_1(x) \\
    & \vdots \\
    p_m(x) &= -\text{rem}(p_{m-2}, p_{m-1})
\end{align*}
\]

denote the number of sign changes in the sequence \( p(\xi), p_1(\xi), p_2(\xi), \ldots, p_m(\xi) \) by \( \sigma(\xi) \).

Then for \( a < b \), both real and such that \( p(a), p(b) \neq 0 \), the number of real roots in \([a, b]\) is given by \( \sigma(a) - \sigma(b) \).

**Multiple root.** Consider a polynomial \( f \) with multiple roots. Then \((x - \alpha)^2\) divides \( f \), with \( \alpha \) being the root. Differentiating, we see that \((x - \alpha)\) divides \( f' \), hence \( f \) and \( f' \) have a common factor. From this it follows, that \( f \) and \( f' \) are relatively prime if and only if \( f \) has only simple roots.

Sturm’s theorem is still applicable in the multiple-root case, since the sequence above will yield this common factor and dividing \( f \) by it, results in a polynomial with the same, but only simple, root.

**Definition.** A Sturm sequence of a polynomial \( f \) in an interval \([a, b]\) is a sequence of polynomials \( f_0 = f, f_1, \ldots, f_m \) such that it holds

1. \( f_m \) has no zeros in \([a, b]\)
2. \( f_0(a), f_0(b) \neq 0 \)
3. for \( 0 < i < m - 1 \) and \( a < \gamma < b \), if \( f_i(\gamma) = 0 \) then \( f_{i-1} = -f_{i+1} \)
4. no two consecutive \( f_i \)’s vanish simultaneously at any point in the interval
5. within a sufficiently small neighbourhood of a root of \( f_0 \), \( f_1 \) has constant sign

The sequence \( p_i \) is a Sturm sequence. The algorithm given above to compute the sequence \( p_i \) is the Euclidean algorithm with a special way of defining the remainders. By assumption, \( f \) and \( f' \) are relatively prime, hence the final polynomial \( p_m \) is a constant non-zero polynomial and thus has no roots in \([a, b]\).

The second point is given by assumption and the third follows directly from the definition of the algorithm:

\[
p_i(x) = p_i(x)q_{i-1}(x) - p_{i-1}(x)
\]

If \( p_i = 0 \) then clearly \( p_{i+1}(x) = -p_{i-1}(x) \), for some \( x \) in the interval.

To show the forth point, suppose this was not true and \( p_i(x) = p_{i+1}(x) = 0 \) for some \( x \) in the interval. But then \( p_{i+2}(x) = \ldots = p_m(x) = 0 \) by the definition of the series. This contradicts the fact that \( p_m \) is a nonzero constant polynomial, and thus we have that no two consecutive \( p_i \)’s vanish simultaneously.

The last point is given by the continuity of polynomials and the fact that \( p \) has only simple roots. Then in a sufficiently small neighbourhood of a root, \( f \) is monotonously increasing or decreasing and thus \( p_i = p' \) has constant sign.

**Proof of main theorem.** Having established that our sequence \( p_i \) is a Sturm sequence, we can now proceed to prove the main theorem.

Evaluating the Sturm chain at some point \( x \), with \( x \) in the interval \([a, b]\), results in a sequence of values \( p_0(x), p_1(x), \ldots, p_m \). Let \( SC(x) \) denote the number of sign changes in the sequence at the point \( x \). That is, if we have \( ++++ \) or \( ---- \), \( SC(x) = 0 \) and for \( +++- \) for example \( SC(x) = 2 \).

The idea of the proof is to follow the changes in \( SC \) as \( x \) passes through the interval \([a, b]\). In particular, we will show that \( SC \) is a monotonically decreasing function and that each root of \( p \) and only a root of \( p \) makes \( SC \) drop by 1.
Clearly, $SC$ can change only if we pass through a root of one of the $p_i$, since only this will cause a change in sign in one of the values in the sequence. Here we have to consider two cases:

**Case 1:** $p_i(x) = 0, i > 0$: One of the intermediate polynomials passes through a zero. Then for $p_{i-1}, p_i, p_{i+1}$ we have by the definition of the Sturm sequence that $p_{i-1}$ and $p_{i+1}$ have opposite, but constant signs, since $p_{i-1}$ and $p_{i+1}$ cannot be zero in a sufficiently small neighborhood and thus cannot change sign. Hence, whatever the sign of $p_i$ is in this small neighborhood, it does not change the overall sign change count (To see this, note that $p_{i-1}$ and $p_{i+1}$ have opposite signs, hence if the sign sequence before is $+$ $-$ $-$, it is after $+$ $+$ $-$ and the number of sign changes remains the same. Similarly for the other cases.)

**Case 2:** $p_0(x) = 0$: By definition of the Sturm sequence, $p_1$ has constant sign in some small neighborhood, say $[α, β]$. Then there are two possibilities:
- $p_1 > 0$, an thus $p_0(α) < 0$ and $p_0(β) > 0$. The sign sequence before is $-$ $+$ and after $+$ $+$, hence $SC$ decreases by one.
- $p_1 < 0$, an thus $p_0(α) > 0$ and $p_0(β) < 0$. The sign sequence before is $+$ $-$ and after $-$ $-$, hence $SC$ decreases by one.

Thus, if (and only if) $x$ passes through a root of $p_0$, $SC$ looses one sign change. This implies that $SC$ is monotonically decreasing and that the number of sign-change-losses in the interval $[a, b]$ counts the number of real roots of the polynomial.

**References**