Intervals Methods for Bounding the Range of Univariate Polynomials
— Confidential Acclaim Deliverable – *

Volker Stahl
Research Institute for Symbolic Computation
Johannes Kepler University
A-4040 Linz, Austria
vstahl@risc.uni-linz.ac.at

June 14, 1995

Contents

*The research was done in the framework of the ACCLAIM project sponsored by European Community Basic Research Action (ESPRIT 7195) and Austrian Science Foundation (P9374-PHY).

Abstract

The overestimation of the range of functions is a fundamental problem which arises in many interval methods for constraint solving. We will not treat the problem in its full generality, rather we restrict ourselves to polynomials. Apart from linear functions, for which range computation is trival, polynomials are the "simplest" subclass of elementary transcendental functions. We can therefore expect more efficient methods and more refined results for polynomials than for arbitrary functions. The practical importance of polynomials is out of question, therefore it seems worthwhile to study them separately. In this paper we consider univariate polynomials, the multivariate case is treated in another paper.

1 Introduction

In this paper we are working on the following problem: Given a univariate polynomial $f : \mathbb{R} \to \mathbb{R}$ and an interval X, find an interval Y such that $Y \supseteq \{f(x) \mid x \in X\}$. This problem arises in many interval methods for constraint solving, such as Krawczyk operator [Krawczyk, 1969], Hansen–Sengypta operator [Hansen and Sengupta, 1981], Tightening [Hong and Stahl, 1994b], etc.

Any range computation method is a compromise between efficiency and accuracy. The most efficient method is simply to return $[-\infty, \infty]$. The other extreme is to compute the range exactly by using infinite precision algebraic numbers. In this paper we review several methods, which seem to be good compromises.

In Section 2 we fix some notational conventions. In Section 3 we discuss the Horner form. Horner form is is very efficient and gives the exact range if the distance of the argument interval and zero is large enough. Further, we present some improvements and study bisection of the input interval. Section 4 and 5 contain quadratically convergent methods which are based on the mean value respectively Taylor theorem. The mean value form is cheaper and exploits sparsity of the input polynomial. Various improvements of both methods are presented. A completely different method based on Bernstein transformation is subject of Section 6. This method is rather costly but gives very tight inclusions and in many cases the exact range. A higher convergence order than quadratic is achieved by interpolation forms, which are discussed in Section 7. Finally, in Section 8 we compare the range computation algorithms by applying them in Newton's method and in a global optimization algorithm.

2 Notation

The set of real numbers is denoted by $\mathbb{R}$, the set of compact intervals over $\mathbb{R}$ is denoted by $\mathbb{IR}$. We use lower case letters to denote real numbers or functions between the reals and upper case letters to denote intervals or functions between intervals. The upper bound of an interval X is denoted by $\overline{X}$, the lower bound by $\underline{X}$. The width of an interval is defined as

$$\mathrm{w}(X) = \overline{X} - \underline{X}.$$

The distance between two intervals is defined as

$$\mathrm{q}(X) = \max(|\overline{X} - \overline{Y}|, |\underline{X} - \underline{Y}|).$$

The magnitude of an interval is defined as

$$\mathrm{mag}(X) = \max\{|x| \mid x \in X\},$$

the mignitude is defined as

$$\mathrm{mig}(X) = \min\{|x| \mid x \in X\}.$$

The interval arithmetic operations are defined elementwise, i.e.

$$X \circ Y = \{x \circ y \mid x \in X, y \in Y\}$$

for $\circ \in \{+, -, *, /\}$.

Throughout this paper let $f \in \mathbb{R}[x]$,

$$f(x) = a_n x^{d_n} + a_{n-1} x^{d_{n-1}} + \ldots + a_1 x^{d_1}, \quad a_i \neq 0, \quad d_{i+1} > d_i \tag{2.1}$$

be a polynomial and let

$$\begin{array}{rcl} \Delta_1 & = & d_1 \\ \Delta_i & = & d_i - d_{i-1}, \quad i = 2, \ldots, n. \end{array} \tag{2.2}$$

3 Horner Form

There are many ways to arrange a univariate arithmetic expression and it is well known, that expressions, which are equivalent as functions over the reals are usually different as functions over intervals. A standard arrangement of expressions, which contain only addition and multiplication, is the Horner form – it is numerically stable and allows efficient evaluation. Concerning the interval case, Horner form is optimal with respect to subdistributivity. Among all methods in this paper, Horner form is cheapest. On the other hand, there are many examples, where Horner form gives the exact range, where "more sophisticated" methods overestimate.

Section 3.1 contains a new criterion to decide whether Horner form returns the exact range or a proper overestimation. An algorithmic test of the criterion is straight forward. Sometimes it is sufficient to compute only the upper or the lower endpoint of the Horner form. A new algorithm for computing the endpoints separately is given in Section 3.2. Bisection of the input interval in connection with Horner form is studied in Section 3.3. We give some original Theorems to estimate the reduction of the overestimation error through bisection and devise an efficient algorithm for the special case when the input interval is bisected at zero.

If the Horner scheme of f is evaluated using interval arithmetic we obtain the Horner form of f.

Definition 3.1 (Horner Form) *The Horner form $H_f : \mathbb{IR} \to \mathbb{IR}$ of f is defined as*

$$H_f(X) = \Big(\big(\ldots (a_n X^{\Delta_n} + a_{n-1}) X^{\Delta_{n-1}} + \ldots + a_2 \big) X^{\Delta_2} + a_1 \Big) X^{\Delta_1}. \ \square \tag{3.1}$$

If iterated multiplication is used instead of the interval power function in the definition of the Horner form, we obtain the dense Horner form. Let a_i^* be the coefficient of x^i in f, i.e.

$$f(x) = \sum_{i=1}^{n} a_i x^{d_i} = \sum_{i=0}^{d_n} a_i^* x^i. \tag{3.2}$$

Definition 3.2 (Dense Horner Form) *The dense Horner form $H_f^* : \mathbb{IR} \to \mathbb{IR}$ of f is defined as*

$$H_f^*(X) = \Big(\big(\ldots (a_n^* X + a_{n-1}^*) X + \ldots + a_2^* \big) X + a_1^* \Big) X + a_0^*. \ \square \tag{3.3}$$

Theorem 3.3 *For all $X \in \mathbb{IR}$ it holds that*

$$H_f(X) \subseteq H_f^*(X). \tag{3.4}$$

If f is dense or $0 \notin \operatorname{int}(X)$ then

$$H_f(X) = H_f^*(X). \ \square \tag{3.5}$$

Proof. Let

$$\begin{array}{rclrcl} P_n(X) & = & a_n X^{\Delta_n}, & P_i(X) & = & (P_{i+1}(X) + a_i) X^{\Delta_i} \\ P_n^*(X) & = & a_n \underbrace{X \cdots X}_{\Delta_n \text{times}}, & P_i^*(X) & = & (P_{i+1}^*(X) + a_i) \underbrace{X \cdots X}_{\Delta_i \text{times}} \end{array} \tag{3.6}$$

for $i = 1, \ldots, n-1$. Note that $P_1(X) = H_f(X)$ and $P_1^*(X) = H_f^*(X)$. Inductively one shows for all i that $P_i(X) \subseteq P_i^*(X)$. If f is dense, then $\Delta_i \leq 1$, hence $P_i(X) = P_i^*(X)$. If $0 \notin \operatorname{int}(X)$ then X^{Δ_i} is equal to Δ_i times the product of X with itself, and $P_i(X) = P_i^*(X)$. $\square$

Theorem 3.4 (Convergence) *H_f and H_f^* converge linearly to f.* □

Proof. H_f and H_f^* are Lipschitz interval extensions of f Hence H_f and H_f^* converge linearly to f. □

The following theorem simplifies some considerations later on.

Theorem 3.5 *For all $c \in \mathbb{R}$, $X \in \mathbb{IR}$ the following holds.*

- *If $g(x) = f(x) + c$ then*
$$H_g(X) = H_f(X) + c.$$
- *If $g(x) = cf(x)$ then*
$$H_g(X) = cH_f(X).$$
- *If $g(x) = f(cx)$ then*
$$H_g(X) = H_f(cX). \ \square$$

Proof. Let $c \in \mathbb{R}$ and $X \in \mathbb{IR}$ arbitrary but fixed.

- Let $g(x) = f(x) + c$. Then
$$\begin{aligned} H_g(X) &= \Big((\ldots(a_n X^{\Delta_n} + a_{n-1})X^{\Delta_{n-1}} + \ldots + a_2)X^{\Delta_2} + a_1 \Big) X^{\Delta_1} + c \\ &= H_f(X). \end{aligned}$$
- Let $g(x) = c\, f(x)$. Then
$$\begin{aligned} H_g(X) &= \Big((\ldots(ca_n X^{\Delta_n} + ca_{n-1})X^{\Delta_{n-1}} + \ldots + ca_2)X^{\Delta_2} + ca_1 \Big) X^{\Delta_1} \\ &= \Big((\ldots c(a_n X^{\Delta_n} + a_{n-1})X^{\Delta_{n-1}} + \ldots + ca_2)X^{\Delta_2} + ca_1 \Big) X^{\Delta_1} \\ \ldots \\ &= c\Big((\ldots(a_n X^{\Delta_n} + a_{n-1})X^{\Delta_{n-1}} + \ldots + a_2)X^{\Delta_2} + a_1 \Big) X^{\Delta_1} \\ &= cH_f(X). \end{aligned}$$
- Let $g(x) = f(cx)$. Then
$$\begin{aligned} H_g(X) &= \Big((\ldots(a_n c^{d_n} X^{\Delta_n} + a_{n-1}c^{d_{n-1}})X^{\Delta_{n-1}} + \ldots + a_2 c^{d_2})X^{\Delta_2} + a_1 c^{d_1} \Big) X^{\Delta_1} \\ &= \Big((\ldots(a_n (cX)^{\Delta_n} + a_{n-1})c^{d_{n-1}}X^{\Delta_{n-1}} + \ldots + a_2 c^{d_2})X^{\Delta_2} + a_1 c^{d_1} \Big) X^{\Delta_1} \\ \ldots \\ &= \Big((\ldots(a_n (cX)^{\Delta_n} + a_{n-1})(cX)^{\Delta_{n-1}} + \ldots + a_2)(cX)^{\Delta_2} + a_1 \Big) (cX)^{\Delta_1} \\ &= H_f(cX). \ \square \end{aligned}$$

Remark. In general it is *not* true that $g(x) = f(x+c)$ implies $H_g(X) = H_f(X+c)$. For example, let $f(x) = x^2 - x$, $X = [0, 2]$ and $c = -1$. Then $g(x) = x^2 - 3x + 2$, $H_g(X) = [-4, 2]$ but $H_f(X + c) = [-2, 2]$. This observation can be exploited to reduce the overestimation error of the Horner form, see Section 5.

From Definition 3.1 we obtain the following algorithm for evaluating the Horner form.

Algorithm 3.1 (HF) [Horner Form]

In: $f(x) = a_n x^{d_n} + a_{n-1}x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{HF}(f, X) \in \mathbb{IF}$, $\mathrm{HF}(f, X) \supseteq H_f(X)$.

(1) [Initialize.]
$P_n \longleftarrow a_n$

(2) [Accumulate.]
for $i = n-1, \dots, 1$
$P_i \longleftarrow P_{i+1} X^{d_{i+1}-d_i} + a_i$

(3) [Last Power.]
$P_0 \longleftarrow P_1 X^{d_1}$

(4) [Return.]
return P_0

Remark. The specification of Algorithm 3.1 (HF) requires some explanation. It is clear how a polynomial $f(x) \in \mathbb{F}[x]$ is interpreted as a polynomial in $\mathbb{R}[x]$ if all coefficients are ordinary floating point numbers. As we allow $\perp$ and $\top$ as coefficients, the output must be specified more precisely as

$$\mathrm{HF}(f, X) \in \mathbb{IF}, \quad \mathrm{HF}(f, X) \supseteq \bigcup_{\tilde{f} \in \mathcal{F}} H_{\tilde{f}}(X),$$

where

$$\mathcal{F} = \left\{ \sum_{i=1}^{d} \tilde{a}_i x^{d_i} \;\middle|\; \begin{array}{cl} \tilde{a}_i = a_i & \text{if } a_i \in \mathbb{F}_o \\ \tilde{a}_i \geq \max \mathbb{F}_o & \text{if } a_i = \top \\ \tilde{a}_i \leq \min \mathbb{F}_o & \text{if } a_i = \perp \end{array} \right\}.$$

However, for the specification of the range computation algorithms in the following sections we write as in Algorithm 3.1 in order to simplify notation.

Theorem 3.6 (Complexity) *Algorithm 3.1* (HF) *costs*

$$\begin{array}{rl} n & \textit{interval power computations,} \\ n & \textit{interval multiplications and} \\ 2n-2 & \textit{number additions.} \;\square \end{array}$$

If $0 \notin \mathrm{int}(X)$ then every interval multiplication in Algorithm 3.1 (HF) costs only 2 number multiplications:

Theorem 3.7 (Complexity) *If* $0 \notin \mathrm{int}(X)$ *then Algorithm 3.1* (HF) *costs*

$$\begin{array}{rl} n & \textit{interval power computations,} \\ 2n & \textit{number multiplications and} \\ 2n-2 & \textit{number additions.} \;\square \end{array}$$

3.1 Non–Overestimation of Horner Form

Experimentally it was observed that there are many examples where Horner form does not overestimate, i.e. $H_f(X) = f(X)$.

- One such example is the Horner form of $f(x) = x^2 - x$ evaluated on $[1, 2]$. More generally, Horner form of $f(x)$ on X does not lead to overestimation for all intervals where $\underline{X} \geq 1$ or $\overline{X} \leq 0$.
- Another example is $f(x) = x^3 - 2x$ evaluated on $[-2, -1.5]$. Again, we can generalize: Horner evaluation of $f(x)$ on X does not lead to overestimation if $\underline{X} \geq \sqrt{2}$ or $\overline{X} \leq -\sqrt{2}$.

In this section we present a new criterion to decide whether Horner Form computes the exact range or a proper overestimation. If $\mathrm{mig}(X)$ is sufficiently large, i.e. if

$$\underline{X} \geq \overline{O}_f \text{ or } \overline{X} \leq \underline{O}_f$$

for some bounds

$$\overline{O}_f \geq 0 \geq \underline{O}_f$$

(Theorem 3.10), then Horner form does not lead to overestimation. The computation of the bounds $\overline{O}_f, \underline{O}_f$ is expensive. However, it can be tested very cheaply whether a given interval is outside the bounds, without actually knowing them (Algorithm 3.2).

It should be noted that this simple condition for non–overestimation is sufficient but *not* necessary. Counterexamples have been found, which could be generalized to whole classes of polynomials, where Horner form gives the range although X is not outside O_f (Section 3.1.4). Examples of two rather obvious classes are as follows:

- The Horner form of $f(x) = x^4 + x^2$ never overestimates.
- The Horner form of $x^3 + 2x$ gives the range for all argument intervals, which have midpoint zero.

A complete characterization of all cases where Horner form does not lead to overestimation is still unknown. The results of this section are original.

Before going into details we eliminate some trivial cases from the discussion.

Theorem 3.8 *For all $c \in \mathbb{R}$, $X \in \mathbb{IR}$, if $g(x) = f(x) + c$ or $g(x) = cf(x)$, $c \neq 0$, then*

$$H_f(X) = f(X) \text{ iff } H_g(X) = g(X). \ \square$$

Proof. Theorem 3.8 is an immediate consequence of Theorem 3.5. Let $c \in \mathbb{R}$ and $X \in \mathbb{IR}$ arbitrary but fixed.

- Let $g(x) = f(x) + c$. Then $H_f(X) = f(X)$ iff $H_f(X) + c = f(X) + c$ iff $H_g(X) = g(X)$.
- Let $g(x) = cf(x)$, $c \neq 0$. Then $H_f(X) = f(X)$ iff $cH_f(X) = cf(X)$ iff $H_g(X) = g(X)$. $\square$

Thus, we can make the following assumptions.

- Whether Horner form gives the range does not depend on an additive constant. Hence we can assume that f has no constant monomial, i.e. $d_1 > 0$.
- It is also independend of a constant factor whether Horner form overestimates. Therefore we may assume $a_n > 0$.
- Finally, if f is a monomial, then Horner form never overestimates. Thus, we assume $n > 1$.

Summarizing, throughout this section let

$$\begin{array}{rcl} f(x) & = & a_n x^{d_n} + a_{n-1} x^{d_{n-1}} + \ldots + a_1 x^{d_1} \\ a_i & \in & \mathbb{R}\backslash\{0\}, \quad i = 1, \ldots, n \\ a_n & > & 0 \\ d_1 & > & 0 \\ n & > & 1. \end{array} \tag{3.7}$$

3.1.1 A Sufficient Condition for Non–Overestimation

In this section we show that if the distance of X and 0 is sufficiently large, then Horner form does not overestimate. The key of the proof is that in this case certain "sub–polynomials" of f are monotone in X. Note however, that monotonicity of f in X is not sufficient for non–overestimation, for example $f(x) = x^2 - x$ is monotone in $X = [0.5, 1]$ but $H_f(X) = [-0.5, 0]$ whereas $f(X) = [-0.25, 0]$.

Let $p_i : \mathbb{R} \to \mathbb{R}$, $i = 0, \ldots, n$ be defined as

$$\begin{array}{rcrll} p_n(x) & = & & & a_n \\ p_{n-1}(x) & = & p_n(x) & x^{\Delta_n} & + \; a_{n-1} \\ p_{n-2}(x) & = & p_{n-1}(x) & x^{\Delta_{n-1}} & + \; a_{n-2} \\ & \vdots & & & \\ p_1(x) & = & p_2(x) & x^{\Delta_2} & + \; a_1 \\ p_0(x) & = & p_1(x) & x^{\Delta_1}. & \end{array} \tag{3.8}$$

Note that $p_0(x) = f(x)$.

Definition 3.9 (Overestimation Interval) *The overestimation interval $O_f \in \mathbb{IR}$ of f is defined as*

$$O_f = [\{x \in \mathbb{R} \mid p_i(x) = 0 \text{ for some } \; n \geq i \geq 0\}]. \; \square$$

Note that $p_0(x) = 0$ for $x = 0$, hence

$$\begin{array}{rcl} \overline{O}_f & \geq & 0 \\ \underline{O}_f & \leq & 0. \end{array}$$

We are now ready to formulate the main theorem of this section.

Theorem 3.10 (Non–Overestimation of Horner Form) *For all $X \in \mathbb{IR}$, if*

$$\operatorname{int}(X) \cap O_f = \emptyset$$

then

$$H_f(X) = f(X). \; \square$$

For the proof of Theorem 3.10 we need some monotonicity properties of the polynomials p_i, which are subject of Lemma 3.11 and 3.12.

In the sequel let

$$\begin{array}{rcl} s_n & = & 1 \\ s_{i-1} & = & (-1)^{\Delta_i} s_i, \quad i = n, \ldots, 1. \end{array}$$

For the inductive proofs below we define $a_0 = 0$.

Lemma 3.11 *For $i = 0, \ldots, n$ it holds that*

$$\left.\begin{array}{rcl} p_i'(x) & \geq & 0 \\ p_i(x) & \geq & 0 \end{array}\right\} \quad \text{for all } x \geq \overline{O}_f \tag{3.9}$$

$$\left.\begin{array}{rcl} s_i p_i'(x) & \leq & 0 \\ s_i p_i(x) & \geq & 0 \end{array}\right\} \quad \text{for all } x \leq \underline{O}_f. \; \square \tag{3.10}$$

Proof.

- We show (3.9) by proving inductively for $i = n, \ldots, 0$

$$\left.\begin{array}{rcl} p_i(u) & \geq & p_i(v) \\ p_i(u) & \geq & 0 \end{array}\right\} \text{ for all } \; u \geq v \geq \overline{O}_f. \tag{3.11}$$

  Obviously (3.11) holds for $i = n$. Assume (3.11) holds for some $i = \hat{i} > 0$.

$$\begin{array}{rcl} p_{\hat{i}-1}(u) & = & p_{\hat{i}}(u) u^{\Delta_{\hat{i}}} + a_{\hat{i}-1} \\ & \geq & p_{\hat{i}}(v) v^{\Delta_{\hat{i}}} + a_{\hat{i}-1} \\ & = & p_{\hat{i}-1}(v). \end{array}$$

  Assume $p_{\hat{i}-1}(u) < 0$. As $p_{\hat{i}-1}(x)$ is non–constant and increasing for $x \geq \overline{O}_f$, there exists $\hat{x} > u \geq \overline{O}_f$ such that $p_{\hat{i}-1}(\hat{x}) = 0$, which contradicts the definition of O_f.

- We show (3.10) by proving inductively for $i = n, \ldots, 0$

$$\left.\begin{array}{rcl} s_i p_i(u) & \geq & s_i p_i(v) \\ s_i p_i(u) & \geq & 0 \end{array}\right\} \text{ for all } u \leq v \leq \underline{O}_f. \tag{3.12}$$

Obviously (3.12) holds for $i = n$. Assume (3.11) holds for some $i = \hat{i} > 0$.

$$\begin{array}{rcl} s_{\hat{i}-1} p_{\hat{i}-1}(u) & = & s_{\hat{i}-1} p_{\hat{i}}(u) u^{\Delta_{\hat{i}}} + s_{\hat{i}-1} a_{\hat{i}-1} \\ & = & s_{\hat{i}} p_{\hat{i}}(u)(-u)^{\Delta_{\hat{i}}} + s_{\hat{i}-1} a_{\hat{i}-1} \\ & \geq & s_{\hat{i}} p_{\hat{i}}(v)(-v)^{\Delta_{\hat{i}}} + s_{\hat{i}-1} a_{\hat{i}-1} \\ & = & s_{\hat{i}-1} p_{\hat{i}}(v) v^{\Delta_{\hat{i}}} + s_{\hat{i}-1} a_{\hat{i}-1} \\ & = & s_{\hat{i}-1} p_{\hat{i}-1}(v) \end{array}$$

Assume $s_{\hat{i}-1} p_{\hat{i}-1}(u) < 0$. As $s_{\hat{i}-1} p_{\hat{i}-1}(x)$ is non–constant and decreasing for $x \leq \underline{O}_f$, there exists $\hat{x} < u \leq \underline{O}_f$ such that $s_{\hat{i}-1} p_{\hat{i}-1}(\hat{x}) = 0$, which contradicts the definition of O_f. □

We define multivariate polynomials $\tilde{p}_i : \mathbb{R}^{n-i} \to \mathbb{R}$, $i = 0, \ldots, n$ and $\tilde{f} : \mathbb{R}^n \to \mathbb{R}$ as

$$\begin{array}{rclllll} \tilde{p}_n() & = & & & & a_n & \\ \tilde{p}_{n-1}(x_n) & = & \tilde{p}_n() & x_n^{\Delta_n} & + & a_{n-1} & \\ \tilde{p}_{n-2}(x_{n-1}, x_n) & = & \tilde{p}_{n-1}(x_n) & x_{n-1}^{\Delta_{n-1}} & + & a_{n-2} & \\ & \vdots & & & & & \\ \tilde{p}_1(x_2, \ldots, x_n) & = & \tilde{p}_2(x_3, \ldots, x_n) & x_2^{\Delta_2} & + & a_1 & \\ \tilde{p}_0(x_1, \ldots, x_n) & = & \tilde{p}_1(x_2, \ldots, x_n) & x_1^{\Delta_1} & & & = \tilde{f}(x_1, \ldots, x_n) \end{array} \tag{3.13}$$

Note that

$$\{\tilde{f}(x_1, \ldots, x_n) \mid x_1 \in X, \ldots, x_n \in X\} = H_f(X)$$

Lemma 3.12

$$\left.\begin{array}{rcl} \tilde{f}(\boldsymbol{u}) & \geq & \tilde{f}(\boldsymbol{v}) \\ \tilde{f}(\boldsymbol{u}) & \geq & 0 \end{array}\right\} \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, u_1 \geq v_1 \geq \overline{O}_f, \ldots, u_n \geq v_n \geq \overline{O}_f \tag{3.14}$$

$$\left.\begin{array}{rcl} s_0\tilde{f}(\boldsymbol{u}) & \geq & s_0\tilde{f}(\boldsymbol{v}) \\ s_0\tilde{f}(\boldsymbol{u}) & \geq & 0 \end{array}\right\} \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, u_1 \leq v_1 \leq \underline{O}_f, \ldots, u_n \leq v_n \leq \underline{O}_f \ \square \tag{3.15}$$

Proof. For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, $i = 0, \ldots, n$ let

$$\begin{array}{rcl} \boldsymbol{u}_i & = & (u_{i+1}, u_{i+2}, \ldots, u_n) \\ \boldsymbol{v}_i & = & (v_{i+1}, v_{i+2}, \ldots, v_n). \end{array}$$

- We show (3.14) by proving inductively for $i = n, \ldots, 0$

$$\left.\begin{array}{rcl} \tilde{p}_i(\boldsymbol{u}_i) & \geq & \tilde{p}_i(\boldsymbol{v}_i) \\ \tilde{p}_i(\boldsymbol{u}_i) & \geq & 0 \end{array}\right\} \text{ for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, u_{i+1} \geq v_{i+1} \geq \overline{O}_f, \ldots, u_n \geq v_n \geq \overline{O}_f. \tag{3.16}$$

Obviously (3.16) holds for $i = n$. Assume (3.16) holds for some $i = \hat{i} > 0$.

$$\begin{array}{rcl} \tilde{p}_{\hat{i}-1}(\boldsymbol{u}_{\hat{i}-1}) & = & \tilde{p}_{\hat{i}}(\boldsymbol{u}_{\hat{i}}) u_{\hat{i}}^{\Delta_{\hat{i}}} + a_{\hat{i}-1} \\ & \geq & \tilde{p}_{\hat{i}}(\boldsymbol{v}_{\hat{i}}) v_{\hat{i}}^{\Delta_{\hat{i}}} + a_{\hat{i}-1} \\ & = & \tilde{p}_{\hat{i}-1}(\boldsymbol{v}_{\hat{i}-1}). \\ \tilde{p}_{\hat{i}-1}(\boldsymbol{u}_{\hat{i}-1}) & = & \tilde{p}_{\hat{i}}(\boldsymbol{u}_{\hat{i}}) u_{\hat{i}}^{\Delta_{\hat{i}}} + a_{\hat{i}-1} \\ & \geq & \tilde{p}_{\hat{i}}(\overline{O}_f, \ldots, \overline{O}_f)\overline{O}_f^{\Delta_{\hat{i}}} + a_{\hat{i}-1} \\ & = & p_{\hat{i}}(\overline{O}_f)\overline{O}_f^{\Delta_{\hat{i}}} + a_{\hat{i}-1} \\ & = & p_{\hat{i}-1}(\overline{O}_f) \\ & \geq & 0. \end{array}$$

- We show (3.15) by proving inductively for $i = n, \ldots, 0$

$$\left.\begin{array}{rcl} s_i\tilde{p}_i(\boldsymbol{u}_i) & \geq & s_i\tilde{p}_i(\boldsymbol{v}_i) \\ s_i\tilde{p}_i(\boldsymbol{u}_i) & \geq & 0 \end{array}\right\} \text{ for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, u_{i+1} \leq v_{i+1} \leq \overline{O}_f, \ldots, u_n \leq v_n \leq \underline{O}_f. \tag{3.17}$$

Obviously (3.17) holds for $i = n$. Assume (3.17) holds for some $i = \hat{\imath} > 0$.

$$\begin{array}{rcl} s_{\hat{\imath}-1}\tilde{p}_{\hat{\imath}-1}(\boldsymbol{u}_{\hat{\imath}-1}) & = & s_{\hat{\imath}-1}\tilde{p}_{\hat{\imath}}(\boldsymbol{u}_{\hat{\imath}})u_{\hat{\imath}}^{\Delta_{\hat{\imath}}} + s_{\hat{\imath}-1}a_{\hat{\imath}-1} \\ & = & s_{\hat{\imath}}\tilde{p}_{\hat{\imath}}(\boldsymbol{u}_{\hat{\imath}})(-u_{\hat{\imath}})^{\Delta_{\hat{\imath}}} + s_{\hat{\imath}-1}a_{\hat{\imath}-1} \\ & \geq & s_{\hat{\imath}}\tilde{p}_{\hat{\imath}}(\boldsymbol{v}_{\hat{\imath}})(-v_{\hat{\imath}})^{\Delta_{\hat{\imath}}} + s_{\hat{\imath}-1}a_{\hat{\imath}-1} \\ & = & s_{\hat{\imath}-1}\tilde{p}_{\hat{\imath}}(\boldsymbol{v}_{\hat{\imath}})v_{\hat{\imath}}^{\Delta_{\hat{\imath}}} + s_{\hat{\imath}-1}a_{\hat{\imath}-1} \\ & = & s_{\hat{\imath}-1}\tilde{p}_{\hat{\imath}-1}(\boldsymbol{v}_{\hat{\imath}-1}). \\ s_{\hat{\imath}-1}\tilde{p}_{\hat{\imath}-1}(\boldsymbol{u}_{\hat{\imath}-1}) & = & s_{\hat{\imath}-1}\tilde{p}_{\hat{\imath}}(\boldsymbol{u}_{\hat{\imath}})u_{\hat{\imath}}^{\Delta_{\hat{\imath}}} + s_{\hat{\imath}-1}a_{\hat{\imath}-1} \\ & = & s_{\hat{\imath}}\tilde{p}_{\hat{\imath}}(\boldsymbol{u}_{\hat{\imath}})(-u_{\hat{\imath}})^{\Delta_{\hat{\imath}}} + s_{\hat{\imath}-1}a_{\hat{\imath}-1} \\ & \geq & s_{\hat{\imath}}\tilde{p}_{\hat{\imath}}(\underline{O}_f, \ldots, \underline{O}_f)(-\underline{O}_f)^{\Delta_{\hat{\imath}}} + s_{\hat{\imath}-1}a_{\hat{\imath}-1} \\ & = & s_{\hat{\imath}}p_{\hat{\imath}}(\underline{O}_f)(-\underline{O}_f)^{\Delta_{\hat{\imath}}} + s_{\hat{\imath}-1}a_{\hat{\imath}-1} \\ & = & s_{\hat{\imath}-1}p_{\hat{\imath}-1}(\underline{O}_f) \\ & \geq & 0. \ \square \end{array}$$

Proof of Theorem 3.10. Let $X \in \mathbb{IR}$ such that $\operatorname{int}(X) \cap O_f = \emptyset$. Then either $\overline{X} \leq \underline{O}_f$ or $\underline{X} \geq \overline{O}_f$ and according to Lemma 3.11 f is monotone in X, i.e.

$$f(X) = [f(\underline{X}), f(\overline{X})].$$

By the definition of $\tilde{f}$ it holds that

$$\begin{array}{rcl} f(\underline{X}) & = & \tilde{f}(\underline{X}, \ldots, \underline{X}) \\ f(\overline{X}) & = & \tilde{f}(\overline{X}, \ldots, \overline{X}). \end{array}$$

Finally, it follows from Lemma 3.12 that

$$[\tilde{f}(\underline{X}, \ldots, \underline{X}), \tilde{f}(\overline{X}, \ldots, \overline{X})] = \{\tilde{f}(x_1, \ldots, x_n) \mid x_1 \in X, \ldots, x_n \in X\} = H_f(X)$$

Hence

$$f(X) = H_f(X). \square$$

3.1.2 Optimality of the Non–Overestimation Condition

In the previous section we proved the sufficient condition $\operatorname{int}(X) \cap O_f = \emptyset$ for $H_f(X) = f(X)$. In Section 3.1.4 we show that this condition is not necessary in general. For example let $f(x)$ such that $a_i > 0$ and Δ_i is even for all i. In this case $O_f = \{0\}$, but $H_f(X) = f(X)$ for all X in $\mathbb{IR}$, even if $\operatorname{int}(X) \cap O_f \neq \emptyset$.

However, O_f is minimal in the sense that there exists no interval Q_f such that $Q_f \not\supseteq O_f$ and $\operatorname{int}(X) \cap Q_f = \emptyset$ still implies $H_f(X) = f(X)$.

Throughout this section let $f(x)$ as in (3.7).

Theorem 3.13 (Optimality of the Overestimation Interval) *Assume $a_i < 0$ or Δ_i is odd for at least one $i \in \{1, \ldots, n\}$. Let $Q_f \in \mathbb{IR}$ such that $Q_f \not\supseteq O_f$. Then there exists $X \in \mathbb{IR}$ such that*

$$\operatorname{int}(X) \cap Q_f = \emptyset \text{ and } H_f(X) \supset f(X). \ \square \tag{3.18}$$

For the proof of Theorem 3.13 we distinguish the case $O_f = \{0\}$, which is treated in Lemma 3.14 and 3.15, and the case $O_f \neq \{0\}$, which is treated in Lemma 3.16, 3.17 and 3.18.

In the sequel let $Q_f \in \mathbb{IR}$ such that $Q_f \not\supseteq O_f$.

Lemma 3.14 *If $O_f = \{0\}$ then*

- $a_i > 0$ *for* $i = 1, \ldots, n$ *and*
- Δ_i *is even for* $i = 2, \ldots, n$. $\square$

Proof. Assume $O_f = \{0\}$ and let $n \geq i > 0$. From the definition of O_f it follows that $p_i(x) \neq 0$ for $x \neq 0$. Further, $p_i(0) = a_i \neq 0$, hence $p_i(x) \neq 0$ for all $x \in \mathbb{R}$. As the leading coefficient of p_i is $a_n > 0$, it holds that $p_i(x) > 0$ for sufficiently large x. Hence $p_i(x) > 0$ for all x, and $a_i = p_i(0) > 0$.

Assume Δ_i is odd for some $i > 1$ and let i be the largest index such that Δ_i is odd. Then the degree of

$$p_i(x) = \big(\ldots(a_n x^{\Delta_n} + a_{n-1})x^{\Delta_{n-1}} + \ldots + a_{i+1}\big)x^{\Delta_{i+1}} + a_i$$

is even and the degree of

$$p_{i-1}(x) = p_i(x)x^{\Delta_i} + a_{i-1}$$

is odd. Hence $p_{i-1}(x) < 0$ if x is small enough, which is a contradiction to $p_i(x) > 0$ for all x and $i > 0$. $\square$

Thus, in order to prove Theorem 3.13 for the case $O_f = \{0\}$ it remains to show the following lemma:

Lemma 3.15 *Assume $O_f = \{0\}$ and Δ_1 is odd. Then there exists $X \in \mathbb{IR}$ such that*

$$\mathrm{int}(X) \cap Q_f = \emptyset \quad \text{and} \quad H_f(X) \supset f(X). \ \square$$

Proof. Assume $O_f = \{0\}$ and Δ_1 is odd. From Lemma 3.14 it follows that $a_i > 0$ for $i = 1, \ldots, n$ and Δ_i is even for $i = 2, \ldots, n$. Hence the degree of f is odd and from Lemma 3.11 it follows that $f(x)$ is monotonically increasing. Thus

$$\underline{f(X)} = f(\underline{X})$$

for all $X \in \mathbb{IR}$. As $0 \notin Q_f$ there exists $X \in \mathbb{IR}$ such that

$$\begin{aligned} \mathrm{int}(X) \cap Q_f &= \emptyset \\ 0 &\in \mathrm{int}(X) \\ \overline{X} &> -\underline{X}. \end{aligned}$$

The monomials of $p_1(x)$ have even power and positive coefficient, hence

$$p_1(\overline{X}) > p_1(\underline{X}) > 0.$$

As $\underline{X}^{\Delta_1} < 0$, it holds that

$$\begin{aligned} f(\underline{X}) &= p_1(\underline{X})\underline{X}^{\Delta_1} \\ &> p_1(\overline{X})\underline{X}^{\Delta_1} \\ &= \tilde{f}(\underline{X}, \overline{X}, \overline{X}, \ldots, \overline{X}). \end{aligned}$$

Hence

$$\underline{H_f(X)} < \underline{f(X)},$$

which implies $H_f(X) \supset f(X)$. $\square$

In the following we consider the case $O_f \neq \{0\}$ in Theorem 3.13.

Lemma 3.16

- *If $x \geq 0$ and $p_i(x) \geq 0$ for all $i = 1, \ldots, n$ then $x \geq \overline{O}_f$.*

- *If $x \leq 0$ and $s_i p_i(x) \geq 0$ for all $i = 1, \ldots, n$ then $x \leq \underline{O}_f$.*

Proof.

- Assume $x \geq 0$ and $p_i(x) \geq 0$ for all $i = 1, \ldots, n$. Let $y > x$ arbitrary but fixed. We have to show that $p_i(y) \neq 0$ for $i = 0, \ldots, n$. Obviously $p_n(y) = a_n > 0$. By induction we show for $i = n-1, \ldots, 0$

$$p_i(y) > p_i(x). \tag{3.19}$$

As $p_{n-1}(y) = a_n y^{\Delta_n} + a_{n-1} > a_n x^{\Delta_n} + a_{n-1} = p_{n-1}(x)$, (3.19) holds for $i = n-1$. Assume (3.19) holds for some $i = \hat{i} > 0$.

$$\begin{aligned} p_{\hat{i}-1}(y) &= p_{\hat{i}}(y) y^{\Delta_{\hat{i}}} + a_{\hat{i}-1} \\ &> p_{\hat{i}}(x) x^{\Delta_{\hat{i}}} + a_{\hat{i}-1} \\ &= p_{\hat{i}-1}(x). \end{aligned}$$

Hence, $p_i(y) \neq 0$ for $i = 0, \ldots, n$.

- Assume $x \leq 0$ and $s_i p_i(x) \geq 0$ for all $i = 1, \ldots, n$. Let $y < x$ arbitrary but fixed. We have to show that $p_i(y) \neq 0$ for $i = 0, \ldots, n$. Obviously $p_n(y) = a_n > 0$. By induction we show for $i = n-1, \ldots, 0$

$$s_i p_i(y) > s_i p_i(x). \tag{3.20}$$

As $s_{n-1} p_{n-1}(y) = a_n (-y)^{\Delta_n} + s_{n-1} a_{n-1} > a_n (-x)^{\Delta_n} + s_{n-1} a_{n-1} = s_{n-1} p_{n-1}(x)$, (3.20) holds for $i = n-1$. Assume (3.20) holds for some $i = \hat{i} > 0$.

$$\begin{aligned} s_{\hat{i}-1} p_{\hat{i}-1}(y) &= s_{\hat{i}} p_{\hat{i}}(y) (-y)^{\Delta_{\hat{i}}} + s_{\hat{i}-1} a_{\hat{i}-1} \\ &> s_{\hat{i}} p_{\hat{i}}(x) (-x)^{\Delta_{\hat{i}}} + s_{\hat{i}-1} a_{\hat{i}-1} \\ &= s_{\hat{i}-1} p_{\hat{i}-1}(x). \end{aligned}$$

Hence, $p_i(y) \neq 0$ for $i = 0, \ldots, n$. □

In the sequel let

$$\tilde{f}_i(x_1, \ldots, x_n) = \frac{\partial}{\partial x_i} \tilde{f}(x_1, \ldots, x_n), \quad i = 1, \ldots, n$$

be the partial derivatives of $\tilde{f}$. Note that

$$\begin{aligned} \tilde{f}_i(x_1, \ldots, x_n) &= x_1^{\Delta_1} x_2^{\Delta_2} \ldots x_{i-1}^{\Delta_{i-1}} \frac{\partial}{\partial x_i} \tilde{p}_{i-1}(x_i, \ldots, x_n) \\ &= x_1^{\Delta_1} x_2^{\Delta_2} \ldots x_{i-1}^{\Delta_{i-1}} \Delta_i x_i^{\Delta_i - 1} \tilde{p}_i(x_{i+1}, \ldots, x_n), \end{aligned}$$

where $\tilde{p}_i$ is as in (3.13).

Lemma 3.17 *Let $X \in \mathbb{IR}, \underline{X} \neq \overline{X}$ and let $x \in X$. If there exist $i, j \in \{1, \ldots, n\}$ such that*

$$\tilde{f}_i(x, \ldots, x) \tilde{f}_j(x, \ldots, x) < 0$$

then $\tilde{f}(x, \ldots, x)$ is not an extremum of $\tilde{f}$ in X.

Proof. Assume $\tilde{f}(x, \ldots, x)$ is an extremum of $\tilde{f}$ in X, $\underline{X} \neq \overline{X}$ and $\tilde{f}_i(x, \ldots, x) > 0$, $\tilde{f}_j(x, \ldots, x) < 0$ for some $i, j \in \{1, \ldots, n\}$.

- As $\tilde{f}_i(x, \ldots, x) \neq 0$ it holds that $x \notin \mathrm{int}(X)$.
- Assume $x = \overline{X}$. Then $\tilde{f}(x, \ldots, x)$ cannot be a maximum of $\tilde{f}$ in X because $\tilde{f}_j(x, \ldots, x) < 0$. Further $\tilde{f}(x, \ldots, x)$ cannot be a minimum of $\tilde{f}$ in X because $\tilde{f}_i(x, \ldots, x) > 0$. Hence $x \neq \overline{X}$.

- Assume $x = \underline{X}$. Then $\tilde{f}(x,\dots,x)$ cannot be a maximum of $\tilde{f}$ in X because $\tilde{f}_i(x,\dots,x) > 0$. Further $\tilde{f}(x,\dots,x)$ cannot be a minimum of $\tilde{f}$ in X because $\tilde{f}_j(x,\dots,x) < 0$. Hence $x \neq \overline{X}$.

Lemma 3.18 *Assume $O_f \neq \{0\}$. Then there exists $X \in \mathbb{IR}$ such that*

$$\mathrm{int}(X) \cap Q_f = \emptyset$$

and for all $x \in X$ there exists $i \in \{1,\dots,n\}$ such that

$$\tilde{f}_i(x,\dots,x)\tilde{f}_n(x,\dots,x) < 0.$$

Proof. Assume $O_f \neq \{0\}$. Then $\underline{O}_f \neq \overline{O}_f$ and there exists an interval X such that

$$\begin{aligned} X &\subseteq \mathrm{int}(O_f), \\ 0 &\notin X \quad \text{and} \\ \mathrm{int}(X) \cap Q_f &= \emptyset. \end{aligned}$$

We distinguish the cases $\underline{X} > 0$ and $\overline{X} < 0$.

- Assume $\underline{X} > 0$. As $\overline{X} < \overline{O}_f$ and $\underline{X} > 0$ it follows from Lemma 3.16 that for all $x \in X$ there exists $i \in \{1,\dots,n\}$ such that $p_i(x) < 0$. Hence

$$\begin{aligned} \tilde{f}_i(x,\dots,x) &= x^{\Delta_1}x^{\Delta_2}\dots x^{\Delta_{i-1}}\Delta_i x^{\Delta_i - 1}\tilde{p}_i(x,\dots,x) \\ &= \Delta_i x^{d_i-1}p_i(x) \\ &< 0. \\ \tilde{f}_n(x,\dots,x) &= x^{\Delta_1}x^{\Delta_2}\dots x^{\Delta_{n-1}}\Delta_n x^{\Delta_n - 1}\tilde{p}_n() \\ &= \Delta_n x^{d_n-1}a_n \\ &> 0. \end{aligned}$$

- Assume $\overline{X} < 0$. As $\underline{X} > \underline{O}_f$ and $\overline{X} < 0$ it follows from Lemma 3.16 that for all $x \in X$ there exists $i \in \{1,\dots,n\}$ such that $s_i p_i(x) < 0$. Note that

$$\begin{aligned} s_0 &= (-1)^{d_n} \\ s_i &= (-1)^{d_n - d_i}, \quad i = 1,\dots,n. \end{aligned}$$

  Hence,

$$\begin{aligned} \tilde{f}_i(x,\dots,x) &= x^{\Delta_1}x^{\Delta_2}\dots x^{\Delta_{i-1}}\Delta_i x^{\Delta_i - 1}\tilde{p}_i(x,\dots,x) \\ &= \Delta_i x^{d_i-1}p_i(x) \\ &= \Delta_i(-1)^{d_n-d_i}x^{d_i-1}s_i p_i(x) \\ &= \Delta_i(-1)^{d_n-1}(-x)^{d_i-1}s_i p_i(x) \\ &\begin{cases} < 0 & \text{if } d_n \text{ is odd} \\ > 0 & \text{else.} \end{cases} \\ \tilde{f}_n(x,\dots,x) &= x^{\Delta_1}x^{\Delta_2}\dots x^{\Delta_{n-1}}\Delta_n x^{\Delta_n - 1}\tilde{p}_n() \\ &= \Delta_n x^{d_n-1}a_n \\ &\begin{cases} < 0 & \text{if } d_n \text{ is even} \\ > 0 & \text{else.} \end{cases} \end{aligned} \quad \square$$

The proof of Theorem 3.13 follows now easily from Lemmas 3.14 – 3.18.

Proof of Theorem 3.13. Let f and Q_f as in Theorem 3.13.

- Assume $O_f = \{0\}$. From Lemma 3.14 it follows that $a_i > 0$ for $i = 1,\dots,n$ and Δ_i is even for $i = 2,\dots,n$. Hence, Δ_1 must be odd and $H_f(X) \supset f(X)$ follows from Lemma 3.15.

- Assume $O_f \neq \{0\}$. According to Lemma 3.18 there exists $X \in \mathbb{IR}$ such that $\mathrm{int}(X) \cap Q_f = \emptyset$ and for all $x \in X$ there exists $i \in \{1, \ldots, n\}$ such that
$$\tilde{f}_i(x, \ldots, x)\tilde{f}_n(x, \ldots, x) < 0.$$
According to Lemma 3.17, for all $x \in X$, $\tilde{f}(x, \ldots, x)$ is not an extremum of $\tilde{f}$ in X. Thus there exist $u_1, \ldots, u_n, v_1, \ldots, v_n \in X$ such that
$$\begin{aligned} \tilde{f}(u_1, \ldots, u_n) &< \underline{f(X)} \\ \tilde{f}(v_1, \ldots, v_n) &> \overline{f(X)}, \end{aligned}$$
hence
$$H_f(X) \supset f(X). \ \square$$

3.1.3 Algorithmic Test of the Non–Overestimation Condition

If $\mathrm{int}(X) \cap O_f = \emptyset$, then Horner form gives the range of f on X without overestimation. As the computation of O_f is expensive, we are interested in methods to decide whether $\mathrm{int}(X)$ and O_f are disjoint, without actually computing O_f. The following Theorem shows how this can be done efficiently. Throughout this section let $f(x)$ as in (3.7).

Theorem 3.19 *Let $X \in \mathbb{IR}$, $\underline{X} \neq \overline{X}$. Then*
$$\mathrm{int}(X) \cap O_f = \emptyset$$
if and only if
$$\begin{aligned} \underline{X} &\geq 0 \\ p_i(\underline{X}) &\geq 0 \quad \text{for } i = 1, \ldots, n \end{aligned}$$
or
$$\begin{aligned} \overline{X} &\leq 0 \\ s_i p_i(\overline{X}) &\geq 0 \quad \text{for } i = 1, \ldots, n. \ \square \end{aligned}$$

Proof. Let $X \in \mathbb{IR}$, $\underline{X} \neq \overline{X}$.

$\Rightarrow$ Follows from Lemma 3.11. If $\mathrm{int}(X) \cap O_f = \emptyset$ then either $\underline{X} \geq \overline{O}_f$ or $\overline{X} \leq \underline{O}_f$.

- If $\underline{X} \geq \overline{O}_f$, then $\underline{X} \geq 0$ and $p_i(\underline{X}) \geq 0$ for all $i = 1, \ldots, n$.
- If $\overline{X} \leq \underline{O}_f$, then $\overline{X} \leq 0$ and $s_i p_i(\overline{X}) \geq 0$ for all $i = 1, \ldots, n$.

$\Leftarrow$ Follows from Lemma 3.16.

- If $\underline{X} \geq 0$ and $p_i(\underline{X}) \geq 0$ for $i = 1, \ldots, n$, then $\underline{X} \geq \overline{O}_f$, hence $\mathrm{int}(X) \cap O_f = \emptyset$.
- If $\overline{X} \leq 0$ and $s_i p_i(\overline{X}) \geq 0$ for $i = 1, \ldots, n$, then $\overline{X} \leq \underline{O}_f$, hence $\mathrm{int}(X) \cap O_f = \emptyset$. $\square$

In the sequel let
$$H_{p_i} : \mathbb{IR} \rightarrow \mathbb{IR}$$
be the Horner form of p_i and let O_{p_i} be the overestimation interval of p_i, $i = 1, \ldots, n$.

Theorem 3.20 *Let $X \in \mathbb{IR}$, $\underline{X} \neq \overline{X}$. Then*
$$\mathrm{int}(X) \cap O_f = \emptyset$$
if and only if
$$\begin{aligned} \underline{X} &\geq 0 \\ \underline{H_{p_i}(X)} &\geq 0 \quad \text{for } i = 1, \ldots, n \end{aligned}$$
or
$$\begin{aligned} \overline{X} &\leq 0 \\ \underline{s_i H_{p_i}(X)} &\geq 0 \quad \text{for } i = 1, \ldots, n. \ \square \end{aligned}$$

Proof. Let $X \in \mathbb{IR}$, $\underline{X} \neq \overline{X}$.

- $\Rightarrow$ Assume $\mathrm{int}(X) \cap O_f = \emptyset$. Then either $\underline{X} \geq \overline{O_f}$ or $\overline{X} \leq \underline{O_f}$. As $O_f \supseteq O_{p_i}$, it holds that $H_{p_i}(X) = p_i(X)$ for $i = 1, \ldots, n$.
  - If $\underline{X} \geq \overline{O_f}$ then $\underline{X} \geq 0$ and $\underline{p_i(X)} = p_i(\underline{X}) \geq 0$ by Lemma 3.11, hence $\underline{H_{p_i}(X)} \geq 0$.
  - If $\overline{X} \leq \underline{O_f}$ then $\overline{X} \leq 0$ and $\underline{s_i p_i(X)} = s_i p_i(\overline{X}) \geq 0$ by Lemma 3.11, hence $\underline{s_i H_{p_i}(X)} \geq 0$.
- $\Leftarrow$ Follows from Theorem 3.19. $\square$

Algorithm 3.2 is a modification of Algorithm 3.1 which, in addition to the computation of $H_f(X)$, decides whether $\mathrm{int}(X) \cap O_f = \emptyset$. The decision does not cost any additional arithmetic operations.

Algorithm 3.2 (HFOT) [Horner Form with Overestimation Test]

In: $f(x) = a_n x^{d_n} + a_{n-1} x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$
$X \in \mathbb{IF}$.

Out: $P_0 \in \mathbb{IF}, P_0 \supseteq H_f(X)$,
$t \in \{\texttt{true}, \texttt{false}\}$, if $\mathrm{int}(X) \cap O_f \neq \emptyset$ then $t = \texttt{false}$.

(1) [Initialize.]
$P_n \longleftarrow a_n$

(2) [Accumulate.]
for $i = n-1, \ldots, 1$
  $P_i \longleftarrow P_{i+1} X^{d_{i+1} - d_i} + a_i$

(3) [Last Power.]
$P_0 \longleftarrow P_1 X^{d_1}$

(4) [Trivial Cases for Disjointness Decision.]
if $0 \in \mathrm{int}(X)$ then $t \longleftarrow \texttt{false}$, goto 8
if $\underline{X} = \overline{X}$ then $t \longleftarrow \texttt{true}$, goto 8

(5) [If $d_1 = 0$ then ignore a_1.]
if $d_1 = 0$ then $l = 2$, else $l = 1$

(6) [Signs.]
$s_n = \mathrm{sign}(a_n)$
if $\underline{X} \geq 0$
  for $i = n-1, \ldots, l$ do $s_i \longleftarrow s_{i+1}$
if $\overline{X} \leq 0$
  for $i = n-1, \ldots, l$ do $s_i \longleftarrow (-1)^{d_{i+1} - d_i} s_{i+1}$

(7) [Decide Disjointness of $\mathrm{int}(X)$ and O_f.]
$t \longleftarrow \texttt{true}$
for $i = l, \ldots, n-1$
  if $\underline{s_i P_i} < 0$ then $t \longleftarrow \texttt{false}$

(8) [Return.]
return P_0, t

Remark. If exact arithmetic is used in Algorithm 3.2 then $t = \texttt{true}$ if and only if $\mathrm{int}(X) \cap O_f = \emptyset$.

If $Q_f \supseteq O_f$, then a sufficient condition for non–overestimation of the Horner form is $\mathrm{int}(X) \cap Q_f = \emptyset$. Thus, it is sometimes useful to know an overestimations of O_f. Overestimations of O_f can be obtained cheaply by the use of root bound theorems.

Theorem 3.21 *Let*

$$f(x) = \sum_{i=0}^{d_n} a_i^* x^i = \sum_{i=1}^{n} a_i x^{d_i}.$$

If $f(r) = 0$ for some $r \in \mathbb{R}$ then

$$\begin{aligned}
|r| &\leq \max\{|a_0^*/a_{d_n}^*|, 1 + |a_1^*/a_{d_n}^*|, \ldots, 1 + |a_{d_n-1}^*/a_{d_n}^*|\} \\
|r| &\leq \max\{1, \sum_{i=0}^{d_n-1} |a_i^*/a_{d_n}^*|\} \\
|r| &\leq 2\max\{|a_{d_n-1}^*/a_{d_n}^*|, |a_{d_n-2}^*/a_{d_n}^*|^{1/2}, |a_{d_n-3}^*/a_{d_n}^*|^{1/3}, \ldots, |a_0^*/a_{d_n}^*|^{1/d_n}\}. \ \square
\end{aligned}$$

Proof. See [Householder, 1970]. □

If $b \geq |r|$ for all real roots r of f, then b is called a root bound of f. If b is a common root bound of $p_1, \ldots, p_n$ then $[-b, b] \supseteq O_f$.

Corollary 3.22 (Bounds for Overestimation Interval) *If $b \in \mathbb{R}$ satisfies one of the conditions*

$$\begin{aligned}
b &\geq \max\{|a_1/a_n|, 1 + |a_2/a_n|, \ldots, 1 + |a_{n-1}/a_n|\} \\
b &\geq \max\{1, \sum_{i=1}^{n-1} |a_i/a_n|\} \\
b &\geq 2\max\{|a_{n-1}/a_n|^{1/(d_n-d_{n-1})}, |a_{n-2}/a_n|^{1/(d_n-d_{n-2})}, \ldots, |a_1/a_n|^{1/(d_n-d_1)}\}
\end{aligned}$$

then

$$[-b, b] \supseteq O_f. \ \square$$

Proof. If b satisfies one of the conditions of Corollary 3.22, then b is a common root bound of $p_1, \ldots, p_n$. □

3.1.4 Further Cases where Horner Form does not Overestimate

In Section 3.1.1 we proved the sufficient condition $\mathrm{int}(X) \cap O_f = \emptyset$ for $H_f(X) = f(X)$. In the following we study cases where $H_f(X) = f(X)$ but $\mathrm{int}(X) \cap O_f \neq \emptyset$. The results in this section are original.

Example. Consider the polynomial $f(x) = x^7 - 8x^4 + 64x$ and let $X = [0, 2]$. We obtain $O_f = [0, 2]$, hence $\mathrm{int}(X) \cap O_f \neq \emptyset$, i.e. the non–overestimation condition of Theorem 3.10 is not satisfied. Still,

$$H_f(X) = [0, 128] = f(X).$$

Figure 3.1 shows $\mathrm{q}(H_f(X), f(X))$ for $f(x) = x^7 - 8x^4 + 64x$ in dependence of $\mathrm{mid}(X)$, where $\mathrm{w}(X) \equiv 2$. According to Theorem 3.10, $H_f(X) = f(X)$ if $\mathrm{mid}(X) \geq 3$ or $\mathrm{mid}(X) \leq -1$. The observation that $H_f(X) = f(X)$ for $\mathrm{mid}(X) = 1$, i.e. $X = [0, 2]$ is generalized by the following Theorem.

Theorem 3.23 *If there exists $c \in \mathbb{R}$ and $\Delta \in \mathbb{N}$ such that*

$$a_i = c a_{i+1}, \ \ d_i = d_{i+1} - \Delta, \quad i = 1, \ldots, n-1,$$

$d_1 > 0$, n is odd, and

$$X = [0, \sqrt[\Delta]{|c|}] \text{ or } X = [-\sqrt[\Delta]{|c|}, 0],$$

then

$$H_f(X) = f(X). \ \square$$

Proof. Let $f(x)$ as in Theorem 3.23. Without loss of generality we may assume $a_n = 1$. Then

$$f(x) = \Big(\ldots((x^\Delta + c)x^\Delta + c^2)x^\Delta + \ldots + c^{n-1}\Big)x^\Delta x^{d_0}$$

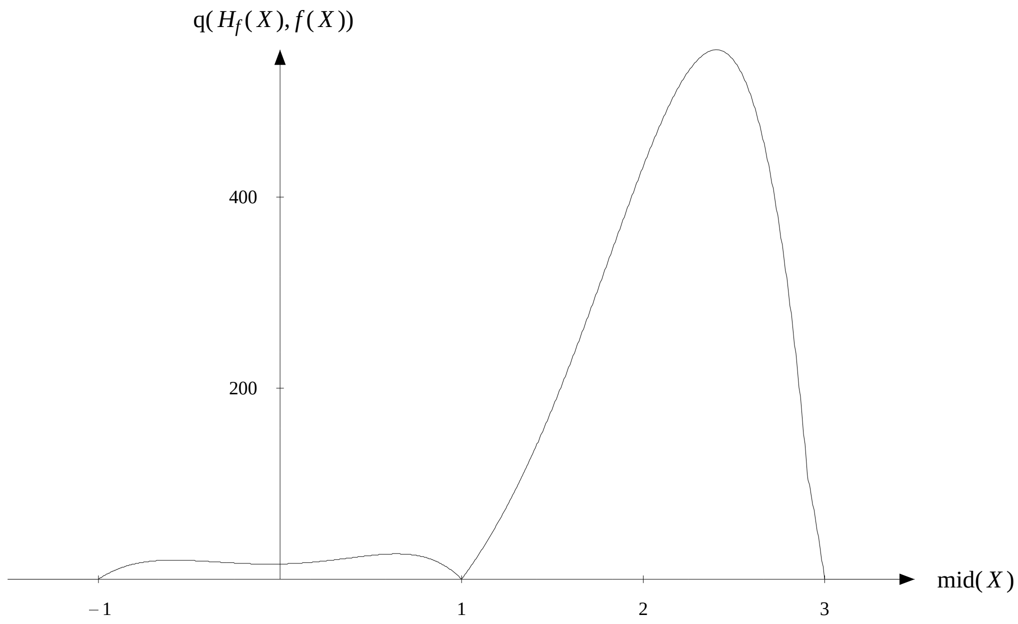

Figure 3.1: $\mathrm{q}(H_f(X), f(X))$ in dependence of $\mathrm{mid}(X)$, where $f(x) = x^7 - 8x^4 + 64x$ and $\mathrm{w}(X) \equiv 2$.

where $d_0 = d_1 - \Delta$. Let

$$\begin{aligned} g_1(x) &= x^{\Delta} \\ g_{i+1}(x) &= (g_i(x) + c^i)x^{\Delta}, \quad i = 1, \ldots, n-1. \end{aligned}$$

and let $H_{g_i}(X) : \mathbb{IR} \to \mathbb{IR}$ be the Horner form of $g_i(x)$, $i = 1, \ldots, n$. Then

$$\begin{aligned} f(x) &= g_n(x)x^{d_0} \\ H_f(X) &\subseteq H_{g_n}(X)X^{d_0}, \end{aligned}$$

and it suffices to show

$$H_{g_n}(X)X^{d_0} \subseteq f(X) \text{ for } \; X = [0, \sqrt[\Delta]{-c}] \; \text{ and } \; X = [0, -\sqrt[\Delta]{-c}].$$

We distinguish 8 cases:

- $c \geq 0$, $X = [0, \sqrt[\Delta]{c}]$, Δ even.
  $O_f = \{0\}$, hence $\mathrm{int}(X) \cap O_f = \emptyset$ and $H_f(X) = f(X)$ by Theorem 3.10.
- $c \geq 0$, $X = [0, \sqrt[\Delta]{c}]$, Δ odd.
  $\overline{O}_f = 0$, hence $\mathrm{int}(X) \cap O_f = \emptyset$ and $H_f(X) = f(X)$ by Theorem 3.10.
- $c \geq 0$, $X = [-\sqrt[\Delta]{c}, 0]$, Δ even.
  $O_f = \{0\}$, hence $\mathrm{int}(X) \cap O_f = \emptyset$ and $H_f(X) = f(X)$ by Theorem 3.10.
- $c \geq 0$, $X = [-\sqrt[\Delta]{c}, 0]$, Δ odd.
  By induction it follows for $i = 1, \ldots, n$

$$\begin{aligned} H_{g_i}(X) &= [-c^i, 0] \\ g_i(\underline{X}) &= \begin{cases} -c^i, & i \text{ odd} \\ 0, & i \text{ even} \end{cases} \\ g_i(\overline{X}) &= 0. \end{aligned}$$

Hence, as n is odd

$$\begin{aligned} H_{g_n}(X)X^{d_0} &= [-c^n, 0][\underline{X}^{d_0}, 0] \\ &= [g_n(\underline{X}), 0][\underline{X}^{d_0}, 0] \\ &= [g_n(\underline{X})\underline{X}^{d_0}, 0] \\ &= [f(\underline{X}), 0] \\ &\subseteq f(X). \end{aligned}$$

- $c < 0$, $X = [0, \sqrt[\Delta]{-c}]$, Δ even.
  By induction it follows for $i = 1, \ldots, n$

$$\begin{aligned} H_{g_i}(X) &= [0, -c^i] \\ g_i(\underline{X}) &= 0 \\ g_i(\overline{X}) &= \begin{cases} -c^i, & i \text{ odd} \\ 0, & i \text{ even.} \end{cases} \end{aligned}$$

  Hence, as n is odd

$$\begin{aligned} H_{g_n}(X)X^{d_0} &= [0, -c^n][0, \overline{X}^{d_0}] \\ &= [0, g_n(\overline{X})][0, \overline{X}^{d_0}] \\ &= [0, g_n(\overline{X})\overline{X}^{d_0}] \\ &= [0, f(\overline{X})] \\ &\subseteq f(X). \end{aligned}$$

- $c < 0$, $X = [0, \sqrt[\Delta]{-c}]$, Δ odd.
  This case is equivalent to the previous case.
- $c < 0$, $X = [-\sqrt[\Delta]{-c}, 0]$, Δ even.
  By induction it follows for $i = 1, \ldots, n$

$$\begin{aligned} H_{g_i}(X) &= [0, -c^i] \\ g_i(\underline{X}) &= \begin{cases} -c^i, & i \text{ odd} \\ 0, & i \text{ even} \end{cases} \\ g_i(\overline{X}) &= 0. \end{aligned}$$

  Hence, as n is odd

$$\begin{aligned} H_{g_n}(X)X^{d_0} &= [0, -c^n][\underline{X}^{d_0}, 0] \\ &= [0, g_n(\underline{X})][\underline{X}^{d_0}, 0] \\ &= [g_n(\underline{X})\underline{X}^{d_0}, 0] \\ &= [f(\underline{X}), 0] \\ &\subseteq f(X). \end{aligned}$$

- $c < 0$, $X = [-\sqrt[\Delta]{-c}, 0]$, Δ odd.
  $\underline{O}_f = 0$, hence $\mathrm{int}(X) \cap O_f = \emptyset$ and $H_f(X) = f(X)$ by Theorem 3.10. □

For the example $f(x) = x^7 - 8x^4 + 64x$ above, we obtain obtain $c = -8$, $\Delta = 3$, $d_0 = 1$ and by Theorem 3.23, $H_f(X) = f(X)$ for $X = [0, 2]$ and for $X = [-2, 0]$.

Example. Consider the polynomial $f(x) = x^4 + 2x^3 + 2x^2$ and let $X = [-0.5, 1.5]$. We obtain $O_f = [-2, 0]$, hence $\mathrm{int}(X) \cap O_f \neq \emptyset$, i.e. the non–overestimation condition of Theorem 3.10 is not satisfied. Still,

$$H_f(X) = [0, 16.3125] = f(X).$$

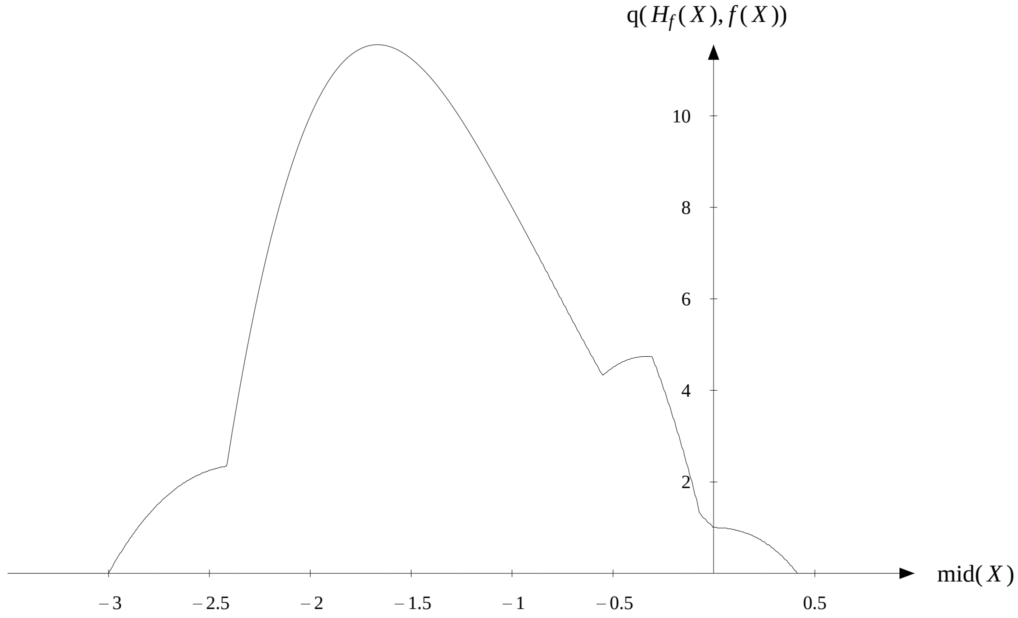

Figure 3.2: $\mathrm{q}(H_f(X), f(X))$ in dependence of $\mathrm{mid}(X)$, where $f(x) = x^4 + 2x^3 + 2x^2$ and $\mathrm{w}(X) \equiv 2$.

Figure 3.2 shows $\mathrm{q}(H_f(X), f(X))$ for $f(x) = x^4 + 2x^3 + 2x^2$ in dependence of $\mathrm{mid}(X)$, where $\mathrm{w}(X) \equiv 2$. According to Theorem 3.10, $H_f(X) = f(X)$ if $\mathrm{mid}(X) \geq 1$ or $\mathrm{mid}(X) \leq -3$. The observation that $H_f(X) = f(X)$ already for $\mathrm{mid}(X) \geq \sqrt{2} - 1 \approx 0.41$ is generalized by the following Theorem.

In the sequel let $p_i, \tilde{p}_i, i = 0, \ldots, n$ as in (3.8) respectively (3.13), and let $H_{p_1}(X)$ be the Horner form of $p_1(x)$.

Theorem 3.24 *If* $a_i > 0$ *for* $i = 1, \ldots, n$, $d_1 > 0$ *is even,* $\mathrm{mid}(X) \geq 0$ *and* $\underline{H_{p_1}(X)} \geq 0$, *then*

$$H_f(X) = f(X). \ \Box$$

For the proof of Theorem 3.24 we need the following two lemmas.

Lemma 3.25 *If* $a_i > 0$ *for* $i = 1, \ldots, n$, *then* $\overline{O}_f = 0$. $\Box$

Proof. Assume $a_i > 0$ for $i = 1, \ldots, n$. Then all coefficients of $p_i(x)$, are positive and $p_i(x)$ has no positive roots for $i = 0, \ldots, n$. Hence $\overline{O}_f = 0$. $\Box$

Lemma 3.26 *If* $a_i > 0$ *for* $i = 1, \ldots, n$ *and* $\mathrm{mid}(X) \geq 0$ *then*

$$\overline{H_{p_1}(X)} = p_1(\overline{X}). \ \Box$$

Proof. Assume $a_i > 0$ for $i = 1, \ldots, n$ and $\mathrm{mid}(X) \geq 0$. By induction we show for $i = n, \ldots, 1$

$$0 \leq |\tilde{p}_i(x_{i+1}, \ldots, x_n)| \leq \tilde{p}_i(\overline{X}, \ldots, \overline{X}) \text{ for all } \ x_{i+1}, \ldots, x_n \in X. \tag{3.21}$$

As $a_n > 0$, (3.21) holds for $i = n$. Assume (3.21) holds for some $i = \hat{i} > 1$ and let $x_{\hat{i}}, \ldots, x_n \in X$ arbitrary but fixed.

$$\begin{aligned} |\tilde{p}_{\hat{i}-1}(x_{\hat{i}}, \ldots, x_n)| &= |\tilde{p}_{\hat{i}}(x_{\hat{i}+1}, \ldots, x_n) x_{\hat{i}}^{\Delta_{\hat{i}}} + a_{\hat{i}}| \\ &\leq |\tilde{p}_{\hat{i}}(x_{\hat{i}+1}, \ldots, x_n)||x_{\hat{i}}^{\Delta_{\hat{i}}}| + a_{\hat{i}} \\ &\leq |\tilde{p}_{\hat{i}}(\overline{X}, \ldots, \overline{X})|\overline{X}^{\Delta_{\hat{i}}} + a_{\hat{i}} \\ &= \tilde{p}_{\hat{i}}(\overline{X}, \ldots, \overline{X})\overline{X}^{\Delta_{\hat{i}}} + a_{\hat{i}} \\ &= \tilde{p}_{\hat{i}-1}(\overline{X}, \ldots, \overline{X}). \end{aligned}$$

Hence

$$\begin{aligned}\overline{H_{p_1}(X)} &= \max_{x_i\in X} \tilde{p}_1(x_1,\ldots,x_n)\\ &\le \max_{x_i\in X} |\tilde{p}_1(x_1,\ldots,x_n)|\\ &\le \tilde{p}_1(\overline{X},\ldots,\overline{X})\\ &= p_1(\overline{X}).\end{aligned}$$

As $H_{p_1}(X) \supseteq p_1(X)$ it follows that $\overline{H_{p_1}(X)} = p_1(\overline{X})$. $\square$

Proof of Theorem 3.24. Assume $a_i > 0$ for $i = 1,\ldots,n$, $d_1 > 0$ is even, $\mathrm{mid}(X) \ge 0$ and $\underline{H_{p_1}(X)} \ge 0$. We distinguish the cases $0 \notin X$ and $0 \in X$.

- Assume $0 \notin X$. As $\mathrm{mid}(X) \ge 0$ it holds that $\underline{X} > 0$. From Lemma 3.25 it follows that $\overline{O}_f = 0$, hence $X \cap O_f = \emptyset$ and $H_f(X) = f(X)$ by Theorem 3.10.
- Assume $0 \in X$. As $\underline{H_{p_1}(X)} \ge 0$, $\overline{H_{p_1}(X)} = p_1(\overline{X})$ by Lemma 3.26 and $X^{d_1} = [0, \overline{X}^{d_1}]$, it holds that

$$\begin{aligned}H_f(X) &= H_{p_1}(X)X^{d_1}\\ &= [\underline{H_{p_1}(X)}, p_1(\overline{X})][0,\overline{X}^{d_1}]\\ &= [0, p_1(\overline{X})\overline{X}^{d_1}]\\ &= [f(0), f(\overline{X})]\\ &\subseteq f(X).\ \square\end{aligned}$$

Corollary 3.27

- *If $a_i < 0$ for $i = 1,\ldots,n$, $d_1 > 0$ is even, $\mathrm{mid}(X) \ge 0$ and $\overline{H_{p_1}(X)} \le 0$, then $H_f(X) = f(X)$.*
- *If $a_i > 0$ for all i where d_i is even, $a_i < 0$ for all i where d_i is odd, $d_1 > 0$ is even, $\mathrm{mid}(X) \le 0$ and $\underline{H_{p_1}(X)} \ge 0$, then $H_f(X) = f(X)$.*
- *If $a_i < 0$ for all i where d_i is even, $a_i > 0$ for all i where d_i is odd, $d_1 > 0$ is even, $\mathrm{mid}(X) \le 0$ and $\overline{H_{p_1}(X)} \le 0$, then $H_f(X) = f(X)$.* $\square$

Proof. Follows from Theorem 3.24 applied to $-f(x)$, $f(-x)$, respectively $-f(-x)$. $\square$ For the example $f(x) = x^4 + 2x^3 + 2x^2$, $X = [-0.5, 1.5]$ above, we obtain

$$H_{p_1}(X) = (X+2)X + 2 = [0.25, 7.25] \ge 0,$$

hence $H_f(X) = f(X)$ by Theorem 3.24.

Corollary 3.28 *If*

$$\mathrm{sign}(a_n) = \mathrm{sign}(a_{n-1}) = \ldots = \mathrm{sign}(a_1)$$

and d_i is even for $i = 1,\ldots,n$ then for all $X \in \mathbb{IR}$

$$H_f(X) = f(X).\ \square$$

Proof. Let $f(x)$ as in Corollary 3.28. Without loss of generality assume

$$\mathrm{sign}(a_n) = \mathrm{sign}(a_{n-1}) = \ldots = \mathrm{sign}(a_1) = 1,\ \ d_1 > 0.$$

Let $X \in \mathbb{IR}$ arbitrary but fixed. If $\mathrm{mid}(X) \ge 0$ we apply Theorem 3.24, if $\mathrm{mid}(X) \le 0$ we apply Corollary 3.27. In both cases we have to show that

$$\underline{H_{p_1}(X)} \ge 0.$$

As all coefficients of $\tilde{p}_1$ are positive and each x_i has even power in $\tilde{p}_1(x_1, \ldots, x_n)$, it holds that

$$\tilde{p}_1(x_1, \ldots, x_n) \geq 0 \text{ for all } \; x_i \in \mathbb{R},$$

hence $\underline{H_{p_1}(X)} \geq 0$.

Example. Consider the polynomial $f(x) = x^7 + 2x^5 + 4x^3 + x$ and let $X = [-1, 1]$. We obtain $O_f = \{0\}$, hence $\mathrm{int}(X) \cap O_f \neq \emptyset$, i.e. the non–overestimation condition of Theorem 3.10 is not satisfied. Still,

$$H_f(X) = [-8, 8] = f(X).$$

Figure 3.3 shows $\mathrm{q}(H_f(X), f(X))$ for $f(x) = x^7 + 2x^5 + 4x^3 + x$ in dependence of $\mathrm{mid}(X)$, where $\mathrm{w}(X) \equiv 2$. According to Theorem 3.10, $H_f(X) = f(X)$ if $\mathrm{mid}(X) \geq 1$ or $\mathrm{mid}(X) \leq -1$. The observation that $H_f(X) = f(X)$ for $\mathrm{mid}(X) = 0$, i.e. $X = [-1, 1]$ is generalized by the following Theorem.

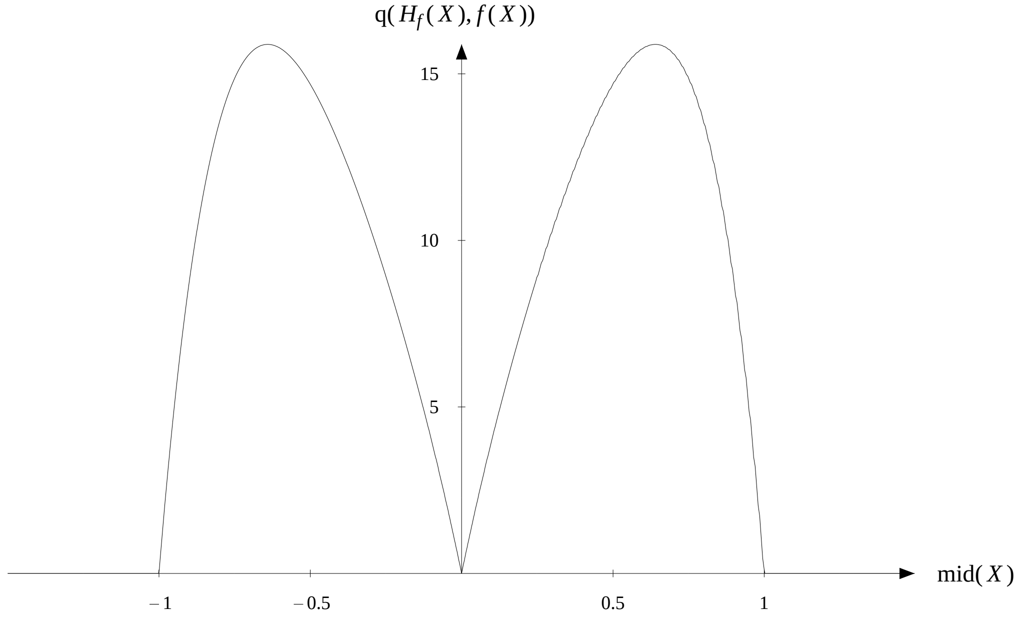

Figure 3.3: $\mathrm{q}(H_f(X), f(X))$ in dependence of $\mathrm{mid}(X)$, where $f(x) = x^7 + 2x^5 + 4x^3 + x$ and $\mathrm{w}(X) \equiv 2$.

Theorem 3.29 *If*

$$\mathrm{sign}(a_n) = \mathrm{sign}(a_{n-1}) = \ldots = \mathrm{sign}(a_1),$$

d_i is odd for $i = 1, \ldots, n$ and $\mathrm{mid}(X) = 0$, then

$$H_f(X) = f(X). \; \square$$

Proof. Let $f(x)$ as in Theorem 3.29 and $\mathrm{mid}(X) = 0$. Note that Δ_i is even for $i = 2, \ldots, n$ and $\Delta_1 = d_1$ is odd. Further, $f(\overline{X}) = -f(\underline{X})$. Without loss of generality assume

$$\mathrm{sign}(a_n) = \mathrm{sign}(a_{n-1}) = \ldots = \mathrm{sign}(a_1) = 1.$$

By induction we show for $i = n, \ldots, 0$ that

$$0 \leq |\tilde{p}_i(x_{i+1}, \ldots, x_n)| \leq \tilde{p}_i(\overline{X}, \ldots, \overline{X}) \text{ for all } \; x_{i+1}, \ldots, x_n \in X. \tag{3.22}$$

As $a_n > 0$, (3.22) holds for $i = n$. Assume (3.22) holds for some $i = \hat{i} > 0$.

$$\begin{aligned}
|\tilde{p}_{\hat{i}-1}(x_{\hat{i}}, \ldots, x_n)| &= |\tilde{p}_{\hat{i}}(x_{\hat{i}+1}, \ldots, x_n) x_{\hat{i}}^{\Delta_{\hat{i}}} + a_{\hat{i}}| \\
&\leq |\tilde{p}_{\hat{i}}(x_{\hat{i}+1}, \ldots, x_n)||x_{\hat{i}}^{\Delta_{\hat{i}}}| + a_{\hat{i}} \\
&\leq \tilde{p}_{\hat{i}}(\overline{X}, \ldots, \overline{X}) \overline{X}^{\Delta_{\hat{i}}} + a_{\hat{i}} \\
&= \tilde{p}_{\hat{i}-1}(\overline{X}, \ldots, \overline{X}).
\end{aligned}$$

Hence

$$\begin{aligned} rclH_f(X) &= \{\tilde{f}(x_1,\ldots,x_n) \mid x_1,\ldots,x_n \in X\} \\ &\subseteq [-f(\overline{X}), f(\overline{X})] \\ &= [f(\underline{X}), f(\overline{X})] \\ &\subseteq f(X). \ \square \end{aligned}$$

Remark. The theorems in this section do not give a complete characterization of all cases where $H_f(X) = f(X)$. For example, let

$$\begin{aligned} f(x) &= x^7 - 2x^4 + 4x, \\ X &= [-1,1] + (r^2-1)/r, \quad \text{where} \quad r \;=\; \sqrt[3]{\tfrac{1}{2}(1+\sqrt{5})}. \end{aligned}$$

Then $H_f(X) = f(X)$, but none of the criteria presented in this section applies. $\square$

3.2 Improvements if Input Interval does not Contain Zero

3.2.1 Efficiency Improvement

If $0 \notin \operatorname{int}(X)$ then some interval power computations of Algorithm 3.1 (HF) can be replaced by number power computations. More precisely, if $0 \in P_{i+1}$ and $\underline{X} \geq 0$, then

$$P_{i+1}X^{d_{i+1}-d_i} = P_{i+1}\overline{X^{d_{i+1}-d_i}} = P_{i+1} \text{ P}\hat{\text{O}}\text{W } (\overline{X}, d_{i+1}-d_i).$$

The case $\underline{X} \leq 0$ is treated similarly. Thus, we obtain the following algorithm.

Algorithm 3.3 (HFS) [Horner Form, Special Case for $0 \notin \operatorname{int}(X)$]

In: $f(x) = a_n x^{d_n} + a_{n-1}x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\text{HFS}(f,X) \in \mathbb{IF}$, $\text{HFS}(f,X) = \text{HF}(f,X)$.

(1) [Reduce to Case $\underline{X} \geq 0$.]
if $0 \in \operatorname{int}(X)$ then return $\text{HF}(f,X)$.
if $\overline{X} < 0$ then $f(x) \longleftarrow f(-x), X \longleftarrow -X$

(2) [Initialize.]
$P_n \longleftarrow a_n$

(3) [Accumulate.]
for $i = n-1,\ldots,1$
  if $0 \in P_{i+1}$ then
    $P_i \longleftarrow P_{i+1} \text{ P}\hat{\text{O}}\text{W } (\overline{X}, d_{i+1}-d_i) + a_i$
  else
    $P_i \longleftarrow P_{i+1}X^{d_{i+1}-d_i} + a_i$

(4) [Last Power.]
if $0 \in P_1$ then
  $P_0 \longleftarrow P_1 \text{ P}\hat{\text{O}}\text{W } (\overline{X}, d_1)$
else
  $P_0 \longleftarrow P_1 X^{d_1}$

(5) [Return.]
return P_0

3.2.2 Separate Computation of Upper and Lower Bound

There are many algorithms, in partiular algorithms for optimization, which require the computation of an upper (or lower) bound of $f(X)$. If u is an upper bound of $-f$ on X then $-u$ is a lower bound of f on X, hence it suffices to consider the computation of upper bounds. Obviously an upper bound of f on X can be obtained by evaluating $H_f(X)$ and selecting the upper bound of the result. In this section we present a more efficient method for computing $\overline{H_f(X)}$ if $0 \notin \mathrm{int}(X)$. The results in this section are original.

It suffices to consider the case $\underline{X} \geq 0$. If $\overline{X} \leq 0$, then we form the polynomial $g(x) = f(-x)$, compute $\overline{H_g(-X)}$, and according to Theorem 3.5, $\overline{H_g(-X)} = \overline{H_f(X)}$.

In the following let $\tilde{p}_i$ and $\tilde{f}$ as in (3.13).

Theorem 3.30 *Assume $\underline{X} \geq 0$ and define $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ recursively as*

$$\hat{x}_i = \begin{cases} \overline{X} & \text{if } \tilde{p}_i(\hat{x}_{i+1}, \ldots, \hat{x}_n) \geq 0 \\ \underline{X} & \text{else.} \end{cases}$$

Then

$$\tilde{f}(\hat{\boldsymbol{x}}) = \overline{H_f(X)}. \ \square$$

Proof. Let X and $\hat{\boldsymbol{x}}$ as in Theorem 3.30. By induction we show for $i = n, \ldots, 0$

$$\tilde{p}_i(\hat{x}_{i+1}, \ldots, \hat{x}_n) \geq \tilde{p}_i(x_{i+1}, \ldots, x_n) \text{ for all } \ x_{i+1}, \ldots, x_n \in X. \tag{3.23}$$

Obviously (3.23) holds for $i = n$. Assume (3.23) holds for some $i = \hat{\imath} > 0$ and let $x_{\hat{\imath}}, \ldots, x_n \in X$ arbitrary but fixed. Then

$$\begin{aligned} \tilde{p}_{\hat{\imath}-1}(x_{\hat{\imath}}, \ldots, x_n) &= \tilde{p}_{\hat{\imath}}(x_{\hat{\imath}+1}, \ldots, x_n) x_{\hat{\imath}}^{\Delta_{\hat{\imath}}} + a_{\hat{\imath}-1} \\ &\leq \tilde{p}_{\hat{\imath}}(\hat{x}_{\hat{\imath}+1}, \ldots, \hat{x}_n) x_{\hat{\imath}}^{\Delta_{\hat{\imath}}} + a_{\hat{\imath}-1} \\ &\leq \begin{cases} \tilde{p}_{\hat{\imath}}(\hat{x}_{\hat{\imath}+1}, \ldots, \hat{x}_n) \overline{X}^{\Delta_{\hat{\imath}}} + a_{\hat{\imath}-1} & \text{if } \tilde{p}_{\hat{\imath}}(\hat{x}_{\hat{\imath}+1}, \ldots, \hat{x}_n) \geq 0 \\ \tilde{p}_{\hat{\imath}}(\hat{x}_{\hat{\imath}+1}, \ldots, \hat{x}_n) \underline{X}^{\Delta_{\hat{\imath}}} + a_{\hat{\imath}-1} & \text{else} \end{cases} \\ &= \tilde{p}_{\hat{\imath}}(\hat{x}_{\hat{\imath}+1}, \ldots, \hat{x}_n) \hat{x}_{\hat{\imath}}^{\Delta_{\hat{\imath}}} + a_{\hat{\imath}-1} \\ &= \tilde{p}_{\hat{\imath}-1}(\hat{x}_{\hat{\imath}}, \ldots, \hat{x}_n). \ \square \end{aligned}$$

Algorithm and Complexity

Based on Theorem 3.30 we give algorithms for computing $\overline{H_f(X)}$ and $\underline{H_f(X)}$.

Algorithm 3.4 (HFUB) [Upper Bound of Horner Form]

In: $f(x) = a_n x^{d_n} + a_{n-1} x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{HFUB}(f, X) \in \mathbb{F}$, $\mathrm{HFUB}(f, X) \geq \overline{H_f(X)}$.

(1) [Initialize.]
$\tilde{p}_n \longleftarrow a_n$
clear invalid operation flag

(2) [Reduce to Case $\underline{X} \geq 0$]
if $0 \in \mathrm{int}(X)$ return $\overline{\mathrm{HF}(f, X)}$
if $\overline{X} < 0$ then $f(x) \longleftarrow f(-x)$, $X \longleftarrow -X$.

(3) [Accumulate.]
for $i = n-1, \ldots, 1$
if $\tilde{p}_{i+1} \geq 0$
then $\tilde{p}_i \longleftarrow \tilde{p}_{i+1} \mathbin{\hat{*}} \widehat{\mathrm{POW}}\,(\overline{X}, d_{i+1} - d_i) \mathbin{\hat{+}} a_i$,
else $\tilde{p}_i \longleftarrow \tilde{p}_{i+1} \mathbin{\hat{*}} \check{\mathrm{POW}}\,(\underline{X}, d_{i+1} - d_i) \mathbin{\hat{+}} a_i$

(4) [Last Power.]
if $\tilde{p}_1 \geq 0$
then $\tilde{p}_0 \longleftarrow \tilde{p}_1 \mathbin{\hat{*}} \widehat{\mathrm{POW}}\,(\overline{X}, d_1)$,
else $\tilde{p}_0 \longleftarrow \tilde{p}_1 \mathbin{\hat{*}} \check{\mathrm{POW}}\,(\underline{X}, d_1)$

(5) [Check Invalid Operation and Return.]
if invalid operation flag is raised, then $\tilde{p}_0 \longleftarrow \top$
return $\tilde{p}_0$

Algorithm 3.5 (HFLB) [Lower Bound of Horner Form]

In: $f(x) = a_n x^{d_n} + a_{n-1} x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{HFLB}(f, X) \in \mathbb{F}$, $\mathrm{HFLB}(f, X) \leq \underline{H_f(X)}$.

(1) [Reduce to Upper Bound.]
return $-\mathrm{HFUB}(-f, X)$

Theorem 3.31 (Complexity)

- *If* $0 \in \mathrm{int}(X)$ *then Algorithm 3.4 and 3.5 cost*

n *interval power computations,*
n *interval multiplications and*
$n-1$ *interval additions.*

- *If* $0 \notin \mathrm{int}(X)$ *then Algorithm 3.4 and 3.5 cost*

n *number power computations,*
n *number multiplications and*
$n-1$ *number additions.* □

3.3 Bisection of the Input Interval

The overestimation error of the Horner form can usually be reduced by dividing the input interval X at a point $c \in \mathrm{int}(X)$ into two subintervals $X_1 = [\underline{X}, c]$, $X_2 = [c, \overline{X}]$ and evaluating Horner form on both intervals separately. From the inclusion monotonicity of Horner form it follows that

$$H_f(X) \supseteq H_f(X_1) \cup H_f(X_2) \supseteq f(X).$$

It seems to be non–trivial to compute an optimal bisection point $\hat{c}$ for a given polynomial f and $X \in \mathbb{IR}$, for example in the sense that

$$\mathrm{w}\big(H_f([\underline{X}, \hat{c}]) \cup H_f([\hat{c}, \overline{X}])\big) \leq \mathrm{w}\big(H_f([\underline{X}, c]) \cup H_f([c, \overline{X}])\big) \text{ for all } c \in X. \tag{3.24}$$

Figure 3.4 shows $\mathrm{w}\big(H_f([\underline{X}, c]) \cup H_f([c, \overline{X}])\big)$ in dependence of c for $f(x) = x^5 + x^4 - 3x^3 - 5x^2 + 7x$ and $X = [-1, 2]$.

As the computation of an optimal or at least a "good" bisection point seems too expensive, we consider two special cases.

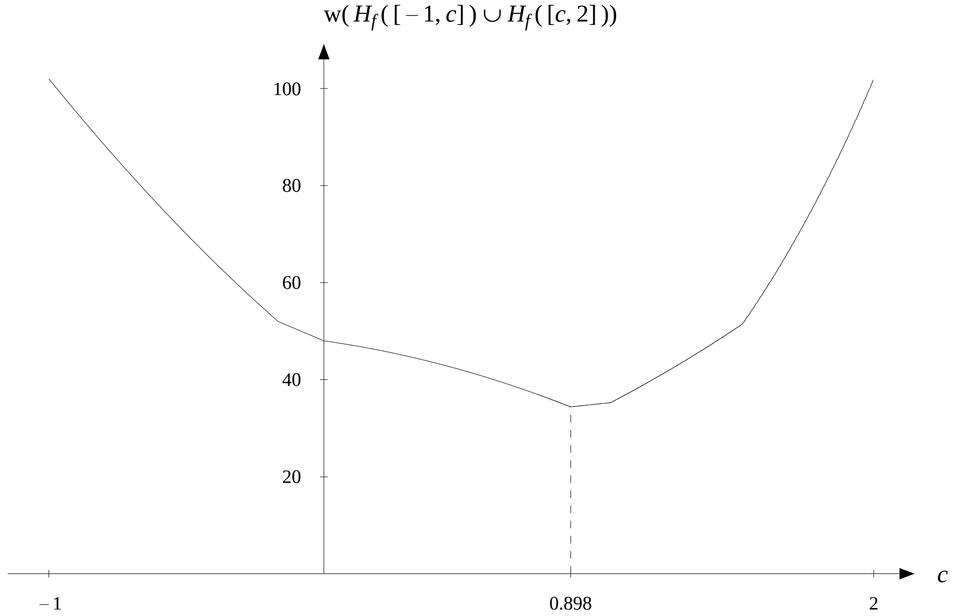

Figure 3.4: $\mathrm{w}(H_f([\underline{X}, c]) \cup H_f([c, \overline{X}]))$ in dependence of c for $f(x) = x^5 + x^4 - 3x^3 - 5x^2 + 7x$ and $X = [-1, 2]$. The optimal bisection point is $\hat{c} \approx 0.9$.

- *Bisection at zero.* In general, the computation of $H_f(X_1) \cup H_f(X_2)$ is twice as expensive as the computation of $H_f(X)$. However, the costs are comparable if $c = 0$ (Section 3.3.1). This means that if $0 \in \mathrm{int}(X)$, then the computation of $H_f([\underline{X}, 0]) \cup H_f([0, \overline{X}])$ is roughly as expensive as the computation of $H_f(X)$, but gives usually better inclusions of the range.
- *Bisection at the midpoint.* The midpoint is usually not an optimal bisection point, e.g. in the sense of (3.24). However, for the dense Horner form, some bounds for the reduction of the overestimation error through bisection at the midpoint can be given (Section 3.3.2).

As far as I am aware, the results in this section are original.

3.3.1 Bisection at Zero

In the sequel let $\check{H}_f : \mathbb{IR} \to \mathbb{IR}$ be defined as

$$\check{H}_f(X) = \begin{cases} H_f([\underline{X}, 0]) \cup H_f([0, \overline{X}]) & \text{if } 0 \in \mathrm{int}(X) \\ H_f(X) & \text{else.} \end{cases}$$

We give an algorithm for evaluating $\check{H}_f(X)$, which costs less than twice as many arithmetic floating point operations as Algorithm 3.1 (HF). The efficiency improvement is due to the fact, that an endpoint of $[\underline{X}, 0]$ and $[0, \overline{X}]$ is zero, and multiplications by zero need not be executed.

Algorithm 3.6 (HFBZ) [Horner Form with Bisection at Zero]

In: $f(x) = a_n x^{d_n} + a_{n-1} x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{HFBZ}(f, X) \in \mathbb{IF}$, $\mathrm{HFBZ}(f, X) \supseteq \check{H}_f(X)$.

(1) [Case $0 \notin \mathrm{int}(X)$.]
If $0 \notin \mathrm{int}(X)$ then return $\mathrm{HFS}(X)$

(2) [Bisect and Evaluate.]
$Y_1 \longleftarrow \mathrm{HFGEZ}(f(-x), [0, -\underline{X}])$
$Y_2 \longleftarrow \mathrm{HFGEZ}(f(x), [0, \overline{X}])$

(3) [Join.]
return $Y_1 \cup Y_2$

Algorithm 3.7 (HFGEZ) [Horner Form for $\underline{X} = 0$]

In: $f(x) = a_n x^{d_n} + a_{n-1} x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$, $\underline{X} = 0$.

Out: $\mathrm{HFGEZ}(f, X) \in \mathbb{IF}$, $\mathrm{HFGEZ}(f, X) \supseteq H_f(X)$.

(1) [Initialize.]
$P_n \longleftarrow a_n$
clear invalid operation flag

(2) [Accumulate.]
for $i = n-1, \ldots, 1$
$p \longleftarrow \hat{\mathrm{POW}}\ (\overline{X}, d_{i+1} - d_i)$
if $\underline{\underline{P_{i+1}}} < 0$ then $\underline{P_i} \longleftarrow \underline{\underline{P_{i+1}}} \check{*}\, p \check{+}\, a_i$, else $\underline{P_i} \longleftarrow a_i$
if $\overline{\overline{P_{i+1}}} > 0$ then $\overline{P_i} \longleftarrow \overline{\overline{P_{i+1}}} \hat{*}\, p \hat{+}\, a_i$, else $\overline{P_i} \longleftarrow a_i$

(3) [Last Power]
if $d_1 > 0$
$p \longleftarrow \hat{\mathrm{POW}}\ (\overline{X}, d_1)$
if $\underline{\underline{P_1}} < 0$ then $\underline{P_0} \longleftarrow \underline{\underline{P_1}} \check{*}\, p$, else $\underline{P_0} \longleftarrow 0$
if $\overline{\overline{P_1}} > 0$ then $\overline{P_0} \longleftarrow \overline{\overline{P_1}} \hat{*}\, p$, else $\overline{P_0} \longleftarrow 0$

(4) [Check Invalid Operation and Return.]
if invalid operation flag is raised, then $P_0 \longleftarrow [\bot, \top]$
return P_0

Theorem 3.32 (Complexity) *Algorithm 3.7* (HFGEZ) *costs*

n	*number power computations,*
$2n$	*number multiplications and*
$2n-2$	*number additions.* □

Theorem 3.33 (Complexity)

- *If* $0 \notin \mathrm{int}(X)$ *then Algorithm 3.6* (HFBZ) *costs*

n	*interval power computations,*
$2n$	*number multiplications and*
$2n-2$	*number additions.*

- *If* $0 \in \mathrm{int}(X)$ *then Algorithm 3.6* (HFBZ) *costs*

$2n$	*number power computations,*
$4n$	*number multiplications and*
$4n-4$	*number additions.* □

Theorem 3.33 does not justify the claim that the costs for computing $H_f(X)$ and $\check{H}_f(X)$ are comparable. The reason is, that the operations were counted for the worst case. For the average case, an experimental comparison seems to be more appropriate.

Table 3.1 shows the average cost and overestimation error of Algorithm 3.1 (HF) and Algorithm 3.6 (HFBZ) for 10^4 random polynomials f of degree 20 with different numbers of monomials and random intervals X. The polynomial coefficients are randomly chosen from $[-1, 1]$, $\underline{X}$ from $[-1, 0]$ and $\overline{X}$ from $[0, 1]$. All random numbers are uniformly distributed. The cost is the total number of arithmetic floating point instructions, including those which were executed during interval operations. The quantity $\mathrm{q}(H_f(X), f(X))/\mathrm{w}(X)$, respectively $\mathrm{q}(\check{H}_f(X), f(X))/\mathrm{w}(X)$ was chosen to measure the overestimation error.

Monomials	3	6	9	12	15	18	21
Flops for $H_f(X)$	25.59	40.86	52.65	63.19	73.79	85.87	102.14
Flops for $\check{H}_f(X)$	31.15	47.15	58.50	68.25	77.57	87.00	96.72
$\mathrm{q}(H_f(X)/\mathrm{w}(X)$	0.05	0.15	0.24	0.37	0.49	0.62	0.75
$\mathrm{q}(\check{H}_f(X)/\mathrm{w}(X)$	0.03	0.09	0.12	0.16	0.19	0.22	0.25

Table 3.1: Comparison of Horner form $H_f(X)$, computed by Algorithm HF, and Horner form with bisection at zero $\check{H}_f(X)$, computed by Algorithm HFBZ, for random polynomials of degree 20 and $0 \in X \subseteq [-1, 1]$.

Finally, one should mention that there are cases where $H_f(X) \neq f(X)$, $0 \in \mathrm{int}(X)$, but $\check{H}_f(X) = H_f(X)$, i.e. bisection at zero does not reduce overestimation. This can happen, even if 0 is the midpoint of X, as is shown in Figure 3.5. However, in the next section we show that this is not possible if f is dense.

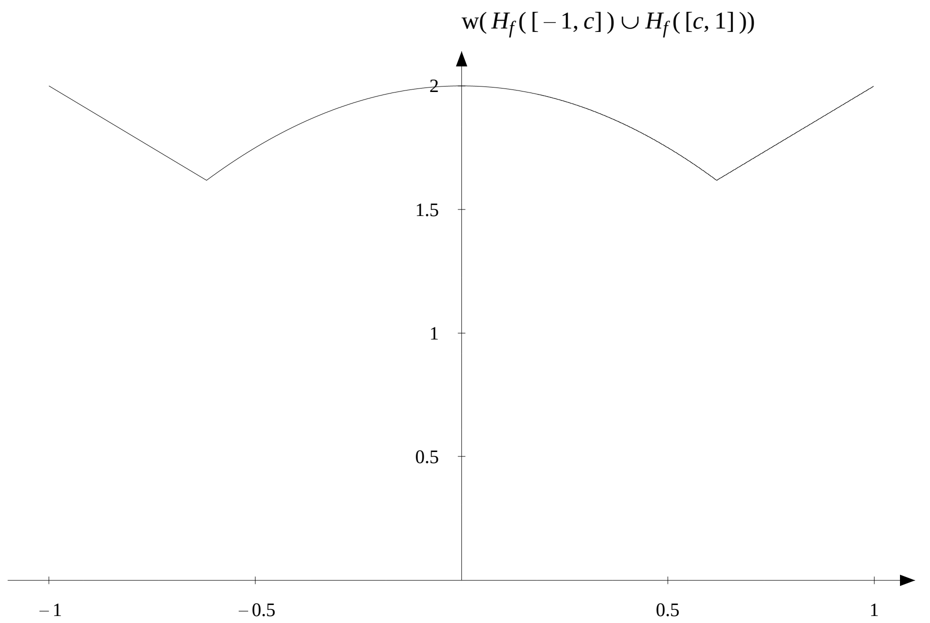

Figure 3.5: $\mathrm{w}(H_f([\underline{X}, c]) \cup H_f([c, \overline{X}]))$ in dependence of c for $f(x) = x^3 - x$ and $X = [-1, 1]$. Bisection at $c = 0$ does not reduce the overestimation error, hence in this case $\check{H}_f(X) = H_f(X)$.

3.3.2 Bisection at the Midpoint for Dense Horner Form

In this section we study the reduction of the overestimation error of the dense Horner form H_f^*, when the input interval X is bisected at its midpoint (Conjecture 3.39). In particular, if $\mathrm{mid}(X) = 0$, then the overestimation error is reduced at least by half (Theorem 3.37).

First, we show that if $H_f^*(X) \neq f(X)$, then bisection leads always to a reduction of the overestimation error (Corollary 3.36), i.e.

$$H_f^*([\underline{X},c]) \cup H_f^*([c,\overline{X}]) \subset H_f^*(X) \text{ for all } \ c \in \operatorname{int}(X).$$

Throughout this section let $X \in \mathbb{IR}$,

$$f(x) = \sum_{i=0}^{d_n} a_i^* x^i$$

and

$$\begin{array}{rcl} \tilde{f}(x_1,\ldots,x_{d_n}) & = & \displaystyle\sum_{i=0}^{d_n} a_i^* \prod_{j=1}^{i} x_j \ = \\ & & \begin{array}{lll} & a_{d_n}^* & x_1 x_2 \cdots x_{d_n-1} x_{d_n} \\ + & a_{d_n-1}^* & x_1 x_2 \cdots x_{d_n-1} \\ & \vdots & \quad \vdots \\ + & a_2^* & x_1 x_2 \\ + & a_1^* & x_1 \\ + & a_0^* & \end{array} \end{array}$$

Note that

$$H_f^*(X) = \{\tilde{f}(x_1,\ldots,x_{d_n}) \mid x_i \in X \text{ for all } \ i\}.$$

Let

$$\tilde{f}_i(x_1,\ldots,x_n) = \frac{\partial}{\partial x_i}\tilde{f}(x_1,\ldots,x_n), \quad i = 1,\ldots,n$$

be the partial derivatives of $\tilde{f}$. As $\tilde{f}$ is linear in each variable, $\tilde{f}_i$ does not depend on x_i.

The following lemma states that $\tilde{f}$ achieves its maximum and minimum in the $n-$cube $(X,\ldots,X)$ at a corner of $(X,\ldots,X)$.

Lemma 3.34 *There exist* $\hat{x}_1,\ldots,\hat{x}_{d_n},\check{x}_1,\ldots,\check{x}_{d_n} \in \{\underline{X},\overline{X}\}$ *such that*

$$\begin{array}{rcl} \max\limits_{x_1,\ldots,x_{d_n}\in X} \tilde{f}(x_1,\ldots,x_{d_n}) & = & \tilde{f}(\hat{x}_1,\ldots,\hat{x}_{d_n}) \\ \min\limits_{x_1,\ldots,x_{d_n}\in X} \tilde{f}(x_1,\ldots,x_{d_n}) & = & \tilde{f}(\check{x}_1,\ldots,\check{x}_{d_n}). \ \Box \end{array}$$

Proof. Follows immediately from the linearity of $\tilde{f}$ in each variable. $\Box$

If $H_f^*(X) \neq f(X)$ then bisection at any point in the interior of X reduces the overestimation error.

Theorem 3.35 *For all* $c \in \operatorname{int}(X)$ *it holds that*

- *If* $\overline{f(X)} < \overline{H_f^*(X)}$ *then*

$$\overline{H_f^*([\underline{X},c]) \cup H_f^*([c,\overline{X}])} < \overline{H_f^*(X)}. \tag{3.25}$$

- *If* $\underline{f(X)} > \underline{H_f^*(X)}$ *then*

$$\underline{H_f^*([\underline{X},c]) \cup H_f^*([c,\overline{X}])} > \underline{H_f^*(X)}. \ \Box \tag{3.26}$$

Proof. As (3.26) follows from (3.25) if f is replaced by $-f$, it suffices to prove (3.25). Assume

$$\overline{H_f^*([\underline{X},c]) \cup H_f^*([c,\overline{X}])} = \overline{H_f^*(X)}$$

and let

$$Y = \begin{cases} [\underline{X},c] & \text{if } \overline{H_f^*([\underline{X},c])} \geq \overline{H_f^*([c,\overline{X}])} \\ [c,\overline{X}] & \text{else.} \end{cases}$$

Then

$$\overline{H_f^*(Y)} = \overline{H_f^*(X)}.$$

Note that

$$\begin{aligned}\overline{H_f^*(Y)} &= \max_{x_1,\ldots,x_{d_n}\in Y} \tilde{f}(x_1,\ldots,x_{d_n})\\ \overline{H_f^*(X)} &= \max_{x_1,\ldots,x_{d_n}\in X} \tilde{f}(x_1,\ldots,x_{d_n}).\end{aligned}$$

According to Lemma 3.34 there exist $y_1,\ldots,y_{d_n} \in \{\underline{Y},\overline{Y}\}$ such that

$$\begin{aligned}\tilde{f}(y_1,\ldots,y_{d_n}) &= \max_{x_1,\ldots,x_{d_n}\in Y} \tilde{f}(x_1,\ldots,x_{d_n})\\ &= \max_{x_1,\ldots,x_{d_n}\in X} \tilde{f}(x_1,\ldots,x_{d_n}).\end{aligned}$$

From the linearity of $\tilde{f}$ it follows that if $y_i = c$ for some i, then

$$\tilde{f}(y_1,\ldots,y_{i-1},\underline{X},y_{i+1},\ldots,y_n) = \tilde{f}(y_1,\ldots,y_n) = \tilde{f}(y_1,\ldots,y_{i-1},\overline{X},y_{i+1},\ldots,y_n). \tag{3.27}$$

Repeated application of (3.27) yields

$$\tilde{f}(y_1,\ldots,y_{d_n}) = \begin{cases} \tilde{f}(\underline{X},\ldots,\underline{X}) & \text{if } Y=[\underline{X},c]\\ \tilde{f}(\overline{X},\ldots,\overline{X}) & \text{if } Y=[c,\overline{X}]\end{cases} \leq \overline{f(X)}$$

Hence,

$$\overline{f(X)} = \tilde{f}(y_1,\ldots,y_{d_n}) = H_f^*(X). \ \square$$

Corollary 3.36 *If $f(X) \subset H_f^*(X)$ then*

$$H_f^*([\underline{X},c]) \cup H_f^*([c,\overline{X}]) \subset H_f^*(X) \tag{3.28}$$

for all $c \in \mathrm{int}(X)$. $\square$

In the following we consider the special case $\mathrm{mid}(X) = 0$, i.e. the input interval is centered. Here, bisection at the midpoint reduces the overestimation error of the dense Horner form at least by half. This result is used later on for the reduction of the overestimation error of the Taylor form, see section 5.

Theorem 3.37 *Assume* $\mathrm{mid}(X) = 0$ *and let*

$$C = H_f^*([\underline{X},0]) \cup H_f^*([0,\overline{X}]).$$

Then

$$|\overline{f(X)} - \overline{C}| \leq 1/2|\overline{f(X)} - \overline{H_f^*(X)}| \tag{3.29}$$

$$|\underline{f(X)} - \underline{C}| \leq 1/2|\underline{f(X)} - \underline{H_f^*(X)}|. \ \square \tag{3.30}$$

Proof. Let X and C as in Theorem 3.37. As (3.30) follows from (3.29) if f is replaced by $-f$, it suffices to show (3.29). From

$$\overline{H_f^*(X)} \geq \overline{C} \geq \overline{f(X)}$$

it follows that (3.29) is equivalent to

$$\overline{H_f^*(X)} - \overline{C} \geq \overline{C} - \overline{f(X)}.$$

Further,

$$\overline{C} = \max\{\overline{H_f^*([0,\overline{X}])}, \overline{H_f^*([\underline{X},0])}\},$$

hence it suffices to show

$$\overline{H_f^*(X)} - \overline{H_f^*([0,\overline{X}]} \geq \overline{H_f^*([0,\overline{X}])} - \overline{f(X)} \tag{3.31}$$

$$\overline{H_f^*(X)} - \overline{H_f^*([\underline{X},0]} \geq \overline{H_f^*([\underline{X},0])} - \overline{f(X)}. \tag{3.32}$$

The proofs of (3.31) and (3.32) are analogous, but for the sake of completeness we give both of them explicitly.

- Proof of (3.31). According to Lemma 3.34 there exist $\hat{x}_1, \ldots, \hat{x}_{d_n} \in \{0, \overline{X}\}$ such that

$$\max_{x_1,\ldots,x_{d_n} \in [0,\overline{X}]} \tilde{f}(x_1, \ldots, x_{d_n}) = \tilde{f}(\hat{x}_1, \ldots, \hat{x}_{d_n}).$$

Let k be the largest index such that $\hat{x}_i = \overline{X}$ for all $i \leq k$. If no such k exists, let $k = 0$. Then

$$\begin{aligned}
\overline{H_f^*([0,\overline{X}])} &= \tilde{f}(\overline{X}, \ldots, \overline{X}, 0, \hat{x}_{k+2}, \ldots, \hat{x}_n) \\
&= a_k \overline{X}^k + a_{k-1}\overline{X}^{k-1} + \ldots + a_1 \overline{X} + a_0 \\
&\leq |a_k \overline{X}^k| + |a_{k-1}\overline{X}^{k-1}| + \ldots + |a_1 \overline{X}| + a_0.
\end{aligned}$$

As $\mathrm{mid}(X) = 0$, it holds that

$$\overline{H_f^*(X)} = |a_{d_n}\overline{X}^{d_n}| + |a_{d_n-1}\overline{X}^{d_n-1}| + \ldots + |a_1\overline{X}| + a_0.$$

Thus,

$$\begin{aligned}
\overline{H_f^*(X)} - \overline{H_f^*([0,\overline{X}])} &\geq |a_{d_n}|\overline{X}^{d_n} + |a_{d_n-1}|\overline{X}^{d_n-1} + \ldots + |a_{k+1}|\overline{X}^{k+1} \\
&\geq -(a_{d_n}\overline{X}^{d_n} + a_{d_n-1}\overline{X}^{d_n-1} + \ldots + a_{k+1}\overline{X}^{k+1}) \\
&= \overline{H_f^*([0,\overline{X}])} - f(\overline{X}) \\
&\geq \overline{H_f^*([0,\overline{X}])} - \overline{f(X)}.
\end{aligned}$$

- Proof of (3.32). According to Lemma 3.34 there exist $\hat{x}_1, \ldots, \hat{x}_{d_n} \in \{\underline{X}, 0\}$ such that

$$\max_{x_1,\ldots,x_{d_n} \in [\underline{X},0]} \tilde{f}(x_1, \ldots, x_{d_n}) = \tilde{f}(\hat{x}_1, \ldots, \hat{x}_{d_n}).$$

Let k be the largest index such that $\hat{x}_i = \underline{X}$ for all $i \leq k$. If no such k exists, let $k = 0$. Then

$$\begin{aligned}
\overline{H_f^*([\underline{X},0])} &= \tilde{f}(\underline{X}, \ldots, \underline{X}, 0, \hat{x}_{k+2}, \ldots, \hat{x}_n) \\
&= a_k \underline{X}^k + a_{k-1}\underline{X}^{k-1} + \ldots + a_1 \underline{X} + a_0 \\
&\leq |a_k \underline{X}^k| + |a_{k-1}\underline{X}^{k-1}| + \ldots + |a_1 \underline{X}| + a_0.
\end{aligned}$$

As $\mathrm{mid}(X) = 0$, it holds that

$$\overline{H_f^*(X)} = |a_{d_n}\underline{X}^{d_n}| + |a_{d_n-1}\underline{X}^{d_n-1}| + \ldots + |a_1\underline{X}| + a_0.$$

Thus,

$$\begin{aligned}
\overline{H_f^*(X)} - \overline{H_f^*([\underline{X},0])} &\geq |a_{d_n}\underline{X}^{d_n}| + |a_{d_n-1}\underline{X}^{d_n-1}| + \ldots + |a_{k+1}\underline{X}^{k+1}| \\
&\geq -(a_{d_n}\underline{X}^{d_n} + a_{d_n-1}\underline{X}^{d_n-1} + \ldots + a_{k+1}\underline{X}^{k+1}) \\
&= \overline{H_f^*([\underline{X},0])} - f(\underline{X}) \\
&\geq \overline{H_f^*([\underline{X},0])} - \overline{f(X)}. \ \square
\end{aligned}$$

Another special case, where bisection at the midpoint reduces the overestimation error of the dense Horner form at least by half, are parabolas.

Theorem 3.38 *Assume $d_n = 2$, let $c = \mathrm{mid}(X)$ and*

$$C = H_f^*([\underline{X}, c]) \cup H_f^*([c, \overline{X}]).$$

Then

$$|\overline{f(X)} - \overline{C}| \leq 1/2|\overline{f(X)} - \overline{H_f^*(X)}| \tag{3.33}$$

$$|\underline{f(X)} - \underline{C}| \leq 1/2|\underline{f(X)} - \underline{H_f^*(X)}|. \ \square \tag{3.34}$$

Proof. Let f, c and C as in Theorem 3.38. As (3.34) follows from (3.33) if f is replaced by $-f$, it suffices to show (3.33). From

$$\overline{H_f^*(X)} \geq \overline{C} \geq \overline{f(X)}$$

it follows that (3.33) is equivalent to

$$\overline{C} - \overline{f(X)} \leq \overline{H_f^*(X)} - \overline{C}.$$

Let

$$\mathcal{Y} = ([\underline{X}, c], [\underline{X}, c]) \cup ([c, \overline{X}], [c, \overline{X}]).$$

Note that $\overline{C} \geq \tilde{f}(y_1, y_2)$ for all $(y_1, y_2) \in \mathcal{Y}$, $\overline{f(X)} \geq \tilde{f}(x, x)$ for all $x \in X$, and $\overline{H_f^*(X)} \geq \tilde{f}(x_1, x_2)$ for all $x_1, x_2 \in X$. Hence, it suffices to show that for all $(y_1, y_2) \in \mathcal{Y}$ there exist $x, x_1, x_2 \in X$ such that

$$\tilde{f}(y_1, y_2) - \tilde{f}(x, x) \leq \tilde{f}(x_1, x_2) - \tilde{f}(y_1, y_2).$$

In fact, we show that for all $(y_1, y_2) \in \mathcal{Y}$ there exist $x, x_1, x_2 \in X$ such that

$$\tilde{f}(y_1, y_2) - \tilde{f}(x, x) = \tilde{f}(x_1, x_2) - \tilde{f}(y_1, y_2).$$

Let $(y_1, y_2) \in \mathcal{Y}$ arbitrary but fixed.

- If $y_1 \geq y_2 \geq c$ or $y_1 \leq y_2 \leq c$ then let $x = y_1$, $x_1 = y_1$ and $x_2 = 2y_2 - y_1$. Note that $x_2 \in X$. As $\tilde{f}$ is linear in its second argument, it holds that

$$\begin{aligned} \tilde{f}(y_1, y_2) - \tilde{f}(x, x) &= \tilde{f}(y_1, y_2) - \tilde{f}(y_1, y_1) \\ &= \tilde{f}(y_1, 2y_2 - y_1) - \tilde{f}(y_1, y_2) \\ &= \tilde{f}(x_1, x_2) - \tilde{f}(y_1, y_2). \end{aligned}$$

- If $y_2 \geq y_1 \geq c$ or $y_2 \leq y_1 \leq c$ then let $x = y_2$, $x_1 = 2y_1 - y_2$ and $x_2 = y_2$. Note that $x_1 \in X$. As $\tilde{f}$ is linear in its first argument, it holds that

$$\begin{aligned} \tilde{f}(y_1, y_2) - \tilde{f}(x, x) &= \tilde{f}(y_1, y_2) - \tilde{f}(y_2, y_2) \\ &= \tilde{f}(2y_1 - y_2, y_2) - \tilde{f}(y_1, y_2) \\ &= \tilde{f}(x_1, x_2) - \tilde{f}(y_1, y_2). \ \square \end{aligned}$$

Remark. In general it is not the case that bisection of the input interval at its midpoint reduces the overestimation error of the dense Horner form at least by half. For example, let $f(x) = -3x^3 - x^2 + 9x$, $X = [0, 1]$ and $c = \mathrm{mid}(X)$. Then

$$\begin{aligned} \overline{f(X)} - \overline{H_f^*(X)} &\approx 3.9 \\ \overline{f(X)} - \overline{H_f^*([\underline{X}, c]) \cup H_f^*([c, \overline{X}])} &\approx 2.6, \end{aligned}$$

i.e. the error of the upper bound was reduced by less than half.

According to experimental results, the reduction of the overestimation error of the dense Horner form, which is obtained through a bisection of the input interval at its midpoint, can be estimated as follows:

Conjecture 3.39 *Let $X \in \mathbb{IR}$ such that $\underline{X} \neq 0$, $\overline{X} \neq 0$, $\underline{X} \neq \overline{X}$, let $c = \mathrm{mid}(X)$ and let*

$$\begin{aligned} r &= \begin{cases} \overline{X}/\underline{X} & \text{if } c \geq 0 \\ \underline{X}/\overline{X} & \text{else} \end{cases} \\ t &= \frac{r+1}{2r}. \end{aligned}$$

Further

$$\begin{aligned} b_{\max} &= \begin{cases} 2/t^{d_n - 2} & \text{if } 0 \notin X \\ (r-1)/(rt^{d_n - 1} - 1) & \text{else} \end{cases} \\ b_{\min} &= \begin{cases} 1 + t & \text{if } d_n = 2 \\ 1 + 1/r^{d_n - 2} & \text{if } d_n > 2 \text{ and } 0 \notin X \\ 1 & \text{else} \end{cases} \end{aligned}$$

and

$$C = H_f^*([\underline{X}, c]) \cup H_f^*([c, \overline{X}]).$$

Then

$$\begin{array}{rcccl} b_{\min}(\overline{H_f^*(X)} - \overline{C}) & \leq & \overline{H_f^*(X)} - \overline{f(X)} & \leq & b_{\max}(\overline{H_f^*(X)} - \overline{C}) \\ b_{\min}(\underline{C} - \underline{H_f^*(X)}) & \leq & \underline{f(X)} - \underline{H_f^*(X)} & \leq & b_{\max}(\underline{C} - \underline{H_f^*(X)}). \quad \square \end{array}$$

3.4 Horner Form for Interval Polynomials

In the sequel let $F : \mathbb{R} \longrightarrow \mathbb{IR}$,

$$F(x) = A_n x^{d_n} + A_{n-1} x^{d_{n-1}} + \ldots + A_1 x^{d_1} \in \mathbb{IR}[x]$$

be an interval polynomial. F can be considered as a set of real polynomials, i.e.

$$F(x) = \left\{ a_n x^{d_n} + a_{n-1} x^{d_{n-1}} + \ldots + a_1 x^{d_1} \;\middle|\; a_n \in A_n, a_{n-1} \in A_{n-1}, \ldots, a_1 \in A_1 \right\}$$

Definition 3.40 *The Horner form $H_F : \mathbb{IR} \rightarrow \mathbb{IR}$ of F is defined as*

$$H_F(X) = \Big(\big(\ldots (A_n X^{\Delta_n} + A_{n-1}) X^{\Delta_{n-1}} + \ldots + A_2 \big) X^{\Delta_2} + A_1 \Big) X^{\Delta_1}. \ \square$$

The following theorem shows that splitting F into F_1, F_2 by splitting a coefficient and evaluating the Horner form of F_1 and F_2 separately does not lead to a better inclusion than evaluating the Horner form of F directly. More precisely, if $F = F_1 \cup F_2$ then

$$H_F(X) = H_{F_1}(X) \cup H_{F_2}(X). \tag{3.35}$$

Theorem 3.41

$$H_F(X) = \{ H_f(X) \mid f \in F \}. \ \square$$

Proof.

$$\begin{aligned} H_F(X) &= \Big(\big(\ldots (A_n X^{\Delta_n} + A_{n-1}) X^{\Delta_{n-1}} + \ldots + A_2 \big) X^{\Delta_2} + A_1 \Big) X^{\Delta_1} \\ &= \Big\{ \Big(\big(\ldots (a_n X^{\Delta_n} + a_{n-1}) X^{\Delta_{n-1}} + \ldots + a_2 \big) X^{\Delta_2} + a_1 \Big) X^{\Delta_1} \;\Big|\; a_i \in A_i, i = 1, \ldots, n \Big\} \\ &= \{ H_f(X) \mid f \in F \}. \ \square \end{aligned}$$

Algorithm and Complexity

Algorithm 3.8 (HFI) for evaluating $H_F(X)$ differs from Algorithm 3.1 (HF) only in that the the coefficients of the input polynomial are floating point intervals instead of floating point numbers.

Algorithm 3.8 (HFI) [Horner Form of Interval Polynomial]

In: $F(x) = A_n x^{d_n} + A_{n-1} x^{d_{n-1}} + \ldots + A_1 x^{d_1} \in \mathbb{IF}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{HFI}(F, X) \in \mathbb{IF}$, $\mathrm{HFI}(F, X) \supseteq H_F(X)$.

(1) [Initialize.]
$P_n \longleftarrow A_n$

(2) [Accumulate.]
for $i = n-1, \ldots, 1$
$P_i \longleftarrow P_{i+1} X^{d_{i+1}-d_i} + A_i$

(3) [Last Power.]
$P_0 \longleftarrow P_1 X^{d_1}$

(4) [Return.]
return P_0

Theorem 3.42 (Complexity) *Algorithm 3.8* (HFI) *costs*

$$\begin{array}{rl} n & \textit{interval power computations,} \\ n & \textit{interval multiplications and} \\ 2n-2 & \textit{number additions.} \ \square \end{array}$$

Theorem 3.43 (Complexity) *If* $0 \notin \mathrm{int}(X)$ *then Algorithm 3.8* (HFI) *costs*

$$\begin{array}{rl} n & \textit{interval power computations,} \\ 2n & \textit{number multiplications and} \\ 2n-2 & \textit{number additions.} \ \square \end{array}$$

4 Mean Value Form

The mean value form is a quadratially convergent interval extension and gives therefore tighter inclusions than the Horner form if the width of the argument interval is sufficiently small. The convergence of the mean value form is discussed e.g. in [Alefeld and Herzberger, 1974], [Skelboe, 1974], [Caprani and Madsen, 1980], and [Alefeld and Herzberger, 1983]. The inclusion monotonicity of the mean value form is proved in [Caprani and Madsen, 1980].

The mean value form can be derived easily from the mean value Theorem: For all $x, c \in X$ there exists $\xi \in X$ such that

$$f(x) = f(c) + f'(\xi)(x-c).$$

Hence

$$f(x) \in f(c) + f'(X)(x-c)$$

and

$$f(X) \subseteq f(c) + f'(X)(X-c).$$

Thus, every interval extension F' of f' and any choice of $c \in X$, give rise to an interval extension $F : \mathbb{IR} \to \mathbb{IR}$ of f:

$$F(X) = f(c) + F'(X)(X-c) \supseteq f(X). \tag{4.1}$$

The mean value form is a special case of (4.1), where F' is the Horner form of f' and $c = \mathrm{mid}(X)$. The choice $c = \mathrm{mid}(X)$ is justified later on by Corollary 4.21.

Definition 4.1 (Mean Value Form) *The mean value form* $M_f : \mathbb{IR} \to \mathbb{IR}$ *and the dense mean value form* $M_f^* : \mathbb{IR} \to \mathbb{IR}$ *are defined as*

$$\begin{array}{rcl} M_f(X) & = & f(\mathrm{mid}(X)) + H_{f'}(X)(X - \mathrm{mid}(X)) \\ M_f^*(X) & = & f(\mathrm{mid}(X)) + H_{f'}^*(X)(X - \mathrm{mid}(X)). \ \square \end{array}$$

Theorem 4.2 M_f *and* M_f^* *are interval extensions of* f. □

Proof. Follows from (4.1). □

The following theorem is taken from [Caprani and Madsen, 1980].

Theorem 4.3 (Inclusion Monotonicity) *M_f and M_f^* are inclusion monotone.* □

Proof. Let $X_\diamond \subseteq X$, $c_\diamond = \mathrm{mid}(X_\diamond)$, $c = \mathrm{mid}(X)$, $r_\diamond = \mathrm{rad}(X_\diamond)$, $r = \mathrm{rad}(X)$. We show inclusion montonicity of M_f, i.e. $M_f(X_\diamond) \subseteq M_f(X)$. The proof for M_f^* is analogous. As $X - c$ is a centered interval, it holds that

$$\begin{aligned} M_f(X) &= f(c) + \mathrm{mag}(H_{f'}(X))r[-1,1] \\ M_f(X_\diamond) &= f(c_\diamond) + \mathrm{mag}(H_{f'}(X_\diamond))r_\diamond[-1,1]. \end{aligned}$$

Hence, we have to show that

$$\begin{aligned} f(c_\diamond) + \mathrm{mag}(H_{f'}(X_\diamond))r_\diamond &\leq f(c) + \mathrm{mag}(H_{f'}(X))r \\ f(c_\diamond) - \mathrm{mag}(H_{f'}(X_\diamond))r_\diamond &\geq f(c) - \mathrm{mag}(H_{f'}(X))r. \end{aligned}$$

Note that

$$\begin{aligned} |c_\diamond - c| &= 1/2|(\overline{X_\diamond} + \underline{X_\diamond} - \overline{X} - \underline{X})| \\ &= 1/2\max\{\overline{X_\diamond} + \underline{X_\diamond} - \overline{X} - \underline{X}, \overline{X} + \underline{X} - \overline{X_\diamond} - \underline{X_\diamond}\} \\ &\leq 1/2(\overline{X} + \underline{X_\diamond} - \overline{X_\diamond} - \underline{X}) \\ &= r - r_\diamond. \end{aligned}$$

Further, $H_{f'}(X_\diamond) \subseteq H_{f'}(X)$. Let $\xi \in X$ such that

$$f(c_\diamond) = f(c) + f'(\xi)(c_\diamond - c).$$

Then

$$\begin{aligned} f(c_\diamond) + \mathrm{mag}(H_{f'}(X_\diamond))r_\diamond &= f(c) + f'(\xi)(c_\diamond - c) + \mathrm{mag}(H_{f'}(X_\diamond))r_\diamond \\ &\leq f(c) + |f'(\xi)||(c_\diamond - c)| + \mathrm{mag}(H_{f'}(X_\diamond))r_\diamond \\ &\leq f(c) + \mathrm{mag}(H_{f'}(X))(r - r_\diamond) + \mathrm{mag}(H_{f'}(X))r_\diamond \\ &= f(c) + \mathrm{mag}(H_{f'}(X))r \\ f(c_\diamond) - \mathrm{mag}(H_{f'}(X_\diamond))r_\diamond &= f(c) + f'(\xi)(c_\diamond - c) - \mathrm{mag}(H_{f'}(X_\diamond))r_\diamond \\ &\geq f(c) - |f'(\xi)||(c_\diamond - c)| - \mathrm{mag}(H_{f'}(X_\diamond))r_\diamond \\ &\geq f(c) - \mathrm{mag}(H_{f'}(X))(r - r_\diamond) - \mathrm{mag}(H_{f'}(X))r_\diamond \\ &= f(c) - \mathrm{mag}(H_{f'}(X))r. \ \square \end{aligned}$$

Theorem 4.4 (Convergence) *M_f and M_f^* are quadratically convergent.* □

Proof. Let $g(x,c) : \mathbb{R}^2 \to \mathbb{R}$ continuous, such that

$$f(x) = f(c) + g(x,c)(x - c) \text{ for all } \ x, c \in \mathbb{R}.$$

From the mean value Theorem it follows that

$$g(x, \mathrm{mid}(X)) \in f'(X) \subseteq H_{f'}^*(X) \text{ for all } \ x \in X.$$

$H_{f'}^*$ is Lipschitz. Hence M_f^* is a centered form and quadratically convergent by [Ratschek and Rokne, 1984]. As $M_f(X) \subseteq M_f^*(X)$ for all $X \in \mathbb{IR}$, the quadratic convergence of M_f follows. □

Algorithm and Complexity.

Algorithm 4.1 (MF) for evaluating the mean value form follows immediately from Definition 4.1. The algorithm requires evaluation of the Horner form of f'. As the coefficients of f' need not be floating point numbers we have to enclose f' by an interval polynomial.

Algorithm 4.1 (MF) [Mean Value Form]

In: $f(x) = a_n x^{d_n} + a_{n-1}x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{MF}(f, X) \in \mathbb{IF}$, $\mathrm{MF}(f, X) \supseteq M_f(X)$.

(1) [Midpoint.]
$c \longleftarrow \mathrm{MID}(X)$

(2) [Evaluate $f(c)$.]
$Y \longleftarrow \mathrm{HF}(f(x), [c])$

(3) [Inclusion of f'.]
if $d_1 \neq 0$
for $i = 1, \ldots, n$ do $A'_i \longleftarrow [d_i \overset{\vee}{*} a_i, d_i \overset{\wedge}{*} a_i]$
$F'(x) \longleftarrow A'_n x^{d_n - 1} + A'_{n-1} x^{d_{n-1}-1} + \ldots + A'_1 x^{d_1 - 1}$
else
for $i = 2, \ldots, n$ do $A'_i \longleftarrow [d_i \overset{\vee}{*} a_i, d_i \overset{\wedge}{*} a_i]$
$F'(x) \longleftarrow A'_n x^{d_n - 1} + A'_{n-1} x^{d_{n-1}-1} + \ldots + A'_2 x^{d_2 - 1}$

(4) [Evaluate $H_{f'}(X)$.]
$F' \longleftarrow \mathrm{HFI}(F'(x), X)$

(5) [Return.]
return $Y + F' * (X - c)$

Remark. No invalid operation can occur in Step 3, because $d_i \neq 0$.

Remark. As $F' * (X - c) = \mathrm{mag}(F') * (X - c)$, it would be sufficient to compute "only" $\mathrm{mag}(F')$ instead of F'. However, it seems that there is no algorithm, which computes $\mathrm{mag}(H_{F'}(X))$ faster than $H_{F'}(X)$ for arbitrary X. (An exception is the case $\mathrm{mid}(X) = 0$, see Section 5.)

Theorem 4.5 (Complexity) *Algorithm 4.1 (*MF*) costs*

$2n$	*interval power computations,*
$n + 1$	*interval multiplications,*
$4n + 1$	*number multiplications, and*
$4n + 2$	*number additions.* □

Proof.

- Step 1 costs 1 number multiplication and 2 number additions.
- Step 2 costs n interval power computations, $2n$ number multiplications and $2n - 2$ number additions (Theorem 3.7).
- Step 3 costs $2n$ number multiplications.
- Step 4 costs n interval power computations, n interval multiplications and $2n - 2$ number additions (Theorem 3.42).
- Step 5 costs 1 interval multiplication and 4 number additions. □

4.1 Slope Form

The slope form can be viewed as an improvement of the mean value form, see for example [Hansen, 1978b], [Alefeld, 1981], [Krawczyk and Neumaier, 1985] [Neumaier, 1990], [Hansen, 1992]. It has the same "shape" as the mean value form, but the overestimation of $f'(X)$ in the mean value form is replaced by an overestimation of the set of slopes between x and the midpoint of X, where x ranges over X.

It is well known that for every $c \in \mathbb{R}$, there exists a polynomial $g_c \in \mathbb{R}[x]$ such that

$$f(x) = f(c) + g_c(x)(x - c). \tag{4.2}$$

for all $x \in \mathbb{R}$. Thus, a whole class of interval extensions of f is given by

$$f(X) \subseteq f(c) + G(X)(X - c), \tag{4.3}$$

where G is an interval extension of g_c. The slope is a special case of (4.3), where G is the Horner form of g_c and c is the midpoint of X.

Definition 4.6 (Slope Form) *Let $c = \mathrm{mid}(X)$ and let g_c be the uniquely defined polynomial such that*

$$f(x) = f(c) + g_c(x)(x - c).$$

The slope form $M_f^{\leadsto} : \mathbb{IR} \to \mathbb{IR}$ is defined as

$$M_f^{\leadsto}(X) = f(c) + H_{g_c}(X)(X - c).$$

The dense slope form $M_f^{\overset{*}{\leadsto}} : \mathbb{IR} \to \mathbb{IR}$ is defined as

$$M_f^{\overset{*}{\leadsto}}(X) = f(c) + H_{g_c}^*(X)(X - c). \ \square$$

Obviously $M_f^{\leadsto}(X) \subseteq M_f^{\overset{*}{\leadsto}}(X)$ for all $X \in \mathbb{IR}$. In the sequel let c and g_c as in Definition 4.6.

For the sake of clarity we give an explicit formla for g_c and $H_{g_c}^*$. Let a_i^* be the coefficient of x^i in f and let

$$\begin{aligned} b_n(c) &= a_n^* \\ b_i(c) &= b_{i+1}(c)c + a_i^*, \quad i = n-1, \ldots, 1. \end{aligned}$$

Then

$$g_c(x) = \sum_{i=1}^{n} b_i(c) x^{i-1}.$$

We obtain

$$\begin{aligned} f(c) + g_c(x)(x - c) &= f(c) + \sum_{i=1}^{n} b_i(c) x^{i-1}(x - c) \\ &= f(c) + \sum_{i=1}^{n} b_i(c) x^i - \sum_{i=1}^{n} b_i(c) x^{i-1} c \\ &= f(c) + \sum_{i=1}^{n} b_i(c) x^i - \sum_{i=0}^{n-1} b_{i+1}(c) x^i c \\ &= f(c) + \sum_{i=1}^{n} b_i(c) x^i - \sum_{i=1}^{n-1} (b_i(c) - a_i) x^i - b_1(c) c \\ &= a_0^* + a_n^* x^n + \sum_{i=1}^{n-1} a_i x^i \\ &= f(x). \end{aligned}$$

Further,

$$H_{g_c}^*(X) = \Big(\big(b_0(\mathrm{mid}(X)) X + b_1(\mathrm{mid}(X)) \big) X + \ldots + b_{n-2}(\mathrm{mid}(X)) \Big) X + b_{n-1}(\mathrm{mid}(X)). \tag{4.4}$$

In contrast to the mean value form, neither the slope form nor the dense slope form are inclusion monotone.

Theorem 4.7 (Inclusion Monotonicity) *There exist polynomials f such that neither $M_f^{\leadsto}$ nor $M_f^{\overset{*}{\leadsto}}$ are inclusion monotone.* $\square$

Proof. Consider the example

$$f(x) = 2x^3 - x^2 - 2x, \quad X = [0, 0.9], \quad X_\diamond = [0, 0.5].$$

We obtain

$$\begin{array}{rcccl} M_f^{\leadsto}(f, X) & = & M_f^{\overset{*}{\leadsto}}(f, X) & = & [-1.881, 0.0405] \\ M_f^{\leadsto}(f, X_\diamond) & = & M_f^{\overset{*}{\leadsto}}(f, X_\diamond) & = & [-1.125, 0.0625], \end{array}$$

i.e.

$$\begin{array}{rcl} \overline{M_f^{\leadsto}(f, X)} & < & \overline{M_f^{\leadsto}(f, X_\diamond)} \\ \overline{M_f^{\overset{*}{\leadsto}}(f, X)} & < & \overline{M_f^{\overset{*}{\leadsto}}(f, X_\diamond)}, \end{array}$$

although $X \supset X_\diamond$. $\square$

Theorem 4.8 (Convergence) *$M_f^{\leadsto}$ and $M_f^{\overset{*}{\leadsto}}$ are quadratically convergent.* $\square$

Proof. From (4.4) it follows that $H^*_{g_c}$ is Lipschitz. Further, $g_c(x) \in H^*_{g_c}(X)$ for all $x \in X$. Hence $M_f^{\overset{*}{\leadsto}}$ is a centered form of f and quadratically convergent by [Ratschek and Rokne, 1984]. As $M_f^{\leadsto}(X) \subseteq M_f^{\overset{*}{\leadsto}}(X)$ for all $X \in \mathbb{IR}$, the quadratic convergenc of $M_f^{\overset{*}{\leadsto}}$ follows. $\square$

The similarity of the slope form and the mean value form motivates a comparison.

Lemma 4.9 *For all $X \in \mathbb{IR}$ it holds that*

$$g_c(X) \subseteq f'(X). \;\square \tag{4.5}$$

Proof. Assume $y \in g_c(X)$. Then $y = g_c(\hat{x})$ for some $\hat{x} \in X$.

- Assume $\hat{x} = c$. One easily checks that $g_c(c) = f'(c)$, hence $y \in f'(X)$.
- Assume $\hat{x} \neq c$. From the mean value Theorem it follows that there exists $\xi \in X$ such that

  $$f(\hat{x}) = f(c) + f'(\xi)(\hat{x} - c).$$

  According to (4.2),

  $$f(\hat{x}) = f(c) + y(\hat{x} - c).$$

  Hence $y = f'(\xi) \in f'(X)$. $\square$

Lemma 4.9 gives rise to the conjecture that the slope form yields always tighter inclusions than the mean value form. However, this is not true.

Theorem 4.10 *There exist polynomials f and intervals X such that*

$$M_f^{\leadsto}(X) \supset M_f(X). \;\square \tag{4.6}$$

Proof. Let

$$f(x) = x^3 - 2x, \quad X = [-0.5, 1] \tag{4.7}$$

We obtain

$$\begin{array}{rcl} M_f(X) & = & [-1.984375, 1.015625] \\ M_f^{\leadsto}(X) & = & [-2.40625, 1.4375]. \;\square \end{array}$$

The reason why (4.6) holds in the example above is that f' is sparse, whereas g_c is dense. Therefore, the superiority of the interval power function over interval multiplication reveals in the Horner evaluation of f', but not in the Horner evaluation of g_c. This gives rise to the following conjecture.

Conjecture 4.11 *For all $X \in \mathbb{IR}$ it holds that*

$$M_f^{\overset{*}{\sim}}(X) \subseteq M_f^*(X).$$

Justification. It suffices to show that

$$H_{g_c}^*(X) \subseteq H_{f'}^*(X). \tag{4.8}$$

Equation (4.8) can be written explicitly as

$$\begin{aligned} & \left((a_n^* X + a_n^* c + a_{n-1}^*)X + a_n^* c^2 + a_{n-1}^* c + a_{n-2}^*\right)X + \ldots + a_n^* c^{n-1} + a_{n-1}^* c^{n-2} + \ldots + a_2^* c + a_1^* \\ \subseteq\ & \left((n a_n^* X + (n-1)a_{n-1}^*)X + (n-2)a_{n-2}^*\right)X + \ldots + 2a_2^*)X + a_1^*, \end{aligned}$$

where a_i^* is the coefficient of x^i in f. Experimental results indicate that the following inclusions are valid for arbitrary $c \in X$. We give formulas for the case $n = 5$, the generalization is straigth forward.

$$\begin{aligned} & \Big(((a_5^* X + a_5^* c + a_4^*)X + a_5^* c^2 + a_4^* c + a_3^*)X + a_5^* c^3 + a_4^* c^2 + a_3^* c + a_2^*\Big)X \\ & + a_5^* c^4 + a_4^* c^3 + a_3^* c^2 + a_2^* c + a_1^* \\ \subseteq\ & \Big(((2a_5^* X + a_5^* c + 2a_4^*)X + a_5^* c^2 + a_4^* c + 2a_3^*)X + a_5^* c^3 + a_4^* c^2 + a_3^* c + 2a_2^*\Big)X + a_1^* \\ \subseteq\ & \Big(((3a_5^* X + a_5^* c + 3a_4^*)X + a_5^* c^2 + a_4^* c + 3a_3^*)X + 2a_2^*\Big)X + a_1^* \\ \subseteq\ & \Big(((4a_5^* X + a_5^* c + 4a_4^*)X + 3a_3^*)X + 2a_2^*\Big)X + a_1^* \\ \subseteq\ & \Big(((5a_5^* X + 4a_4^*)X + 3a_3^*)X + 2a_2^*\Big)X + a_1^* \end{aligned}$$

Similar inclusions have been proved by [Alefeld, 1981]. However, for the case above, it seems that no proof is known. □

A comparison between the slope form and the bicentered mean value form might also be of interest. The bicentered mean value form returns the range without overestimation if $0 \notin H_{f'}(X)$. This is not the case for the slope form, e.g. $f(x) = x^2 + 3x$, $X = [-1, 2]$. If $0 \in H_{f'}(X)$ sometimes the slope form is better than the bicentered mean value form, sometimes vice versa:

- $f(x) = -x^3 + x^2 - x$, $X = [-1, 1]$.

$$M_f^{\sim}(X) = [-3, 3] \subset [-4.968, 5.048] = \ddot{M}_f(X)$$

- $f(x) = -x^3 + x^2 + x$, $X = [-0.4, 1]$.

$$M_f^{\sim}(X) = [-1.254, 1.98] \supset [-0.551826, 1.35286] = \ddot{M}_f(X).$$

Algorithm and Complexity.

We give an algorithm for the dense slope form and not for the slope form because of the following reasons:

- Usually g_c is dense, even if f is sparse. This is in particluar the case if the coefficients of g_c are computed by rounded interval arithmetic. Hence, in general $M_f^{\sim}(X) = M_f^{\overset{*}{\sim}}(X)$.
- In a concrete implementation, an algorithm for $M_f^{\overset{*}{\sim}}$ is usually faster than a corresponding algorithm for $M_f^{\sim}$, because some integer operations and the power function calls are saved.

Algorithm 4.2 (DSF) evaluates $M_f^{\overset{*}{\sim}}(X)$. The coefficients of g_c and $f(c)$ are obtained simultaneously by the Horner scheme.

Algorithm 4.2 (DSF) [Dense Slope Form]

In: $f(x) = a_n x^{d_n} + a_{n-1}x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{DSF}(f, X) \in \mathbb{IF}$, $\mathrm{DSF}(f, X) \supseteq M_{\tilde{f}}^{*}(X)$.

(1) [Midpoint.]
$c \longleftarrow \mathrm{MID}(X)$

(2) [Coefficients of $g_c(x)$ and $f(c)$.]
$k \longleftarrow d_n$
$B_k \longleftarrow a_n$
for $i = n-1, \ldots, 1$
  for $j = d_{i+1}, \ldots, d_i + 1$
    $k \longleftarrow k-1$, $B_k \longleftarrow B_{k+1}c$
  $B_k \longleftarrow B_k + a_i$
for $j = d_1, \ldots, 1$
  $k \longleftarrow k-1$, $B_k \longleftarrow B_{k+1}c$

(3) [Evaluate $H^*_{g_c}(X)$.]
if $d_n < 1$ then $G \longleftarrow 0$, goto Step 4
$G \longleftarrow B_{d_n}$
for $i = d_n - 1, \ldots, 1$ do $G \longleftarrow GX + B_i$

(4) [Return.]
return $B_0 + G * (X - c)$

Theorem 4.12 (Complexity) *Algorithm 4.2 (*DSF*) costs*

d_n	*interval multiplications*
$2d_n + 1$	*number multiplication and*
$2d_n + 2n$	*number additions.* □

Proof.

- Step 1 costs 1 number multiplication and 2 number additions.
- Step 2 costs $2d_n$ number multiplications and $2n-2$ number additions.
- Step 3 costs $d_n - 1$ interval multiplications and $2d_n - 2$ number additions.
- Step 4 costs 1 interval multiplication and 2 number additions. □

4.2 Bicentered Mean Value Form

The bicentered mean value form is a modification of the mean value form, which was introduced by [Baumann, 1988]. The basic idea is to evaluate the mean value form twice with different centers $c^{\uparrow}$ and $c^{\downarrow}$ and intersect the results. The centers are optimal in the sense that $c^{\uparrow}$ minimizes the upper bound and $c^{\downarrow}$ maximizes the lower bound. The bicentered mean value form is therefore particularly suitable for the separate computation of an upper or lower bound of f in X.

In the sequel let

$$M_f(X, c) = f(c) + H_{f'}(X)(X - c).$$

Example. Let $f(x) = x^2 - x - 5$ and let $X = [0, 4]$. Then $H_{f'}(X) = [-1, 7]$ and we obtain

$$M_f(X, \mathrm{mid}(X)) = M_f(X) = [-17, 11].$$

Let $c^{\uparrow} = 3.5$ and $c^{\downarrow} = 0.5$. Then

$$\begin{aligned} M_f(X, c^{\uparrow}) &= [-20.75, 7.25] \\ M_f(X, c^{\downarrow}) &= [-8.75, 19.25] \\ M_f(X, c^{\uparrow}) \cap M_f(X, c^{\downarrow}) &= [-8.75, 7.25]. \end{aligned}$$

An illustration of this example is given in Figure 4.1. The dashed lines are the linear functions

$$f(c) + \overline{H_{f'}(X)}(x - c) \quad \text{and} \quad f(c) + \underline{H_{f'}(X)}(x - c) \tag{4.9}$$

where $c = \mathrm{mid}(X)$, $c = c^{\uparrow}$ and $c = c^{\downarrow}$ from left to right. From the picuture it is clear, why $c^{\uparrow}$ and $c^{\downarrow}$ can be called optimal centers. Further, we can see the following important property, which will be used later on:

$$\begin{aligned} f(c^{\uparrow}) + \underline{H_{f'}(X)}(\underline{X} - c^{\uparrow}) &= f(c^{\uparrow}) + \overline{H_{f'}(X)}(\overline{X} - c^{\uparrow}) \\ f(c^{\downarrow}) + \underline{H_{f'}(X)}(\overline{X} - c^{\downarrow}) &= f(c^{\uparrow}) + \overline{H_{f'}(X)}(\underline{X} - c^{\downarrow}). \end{aligned}$$

(Note that this is only true if $0 \in H_{f'}(X)$).

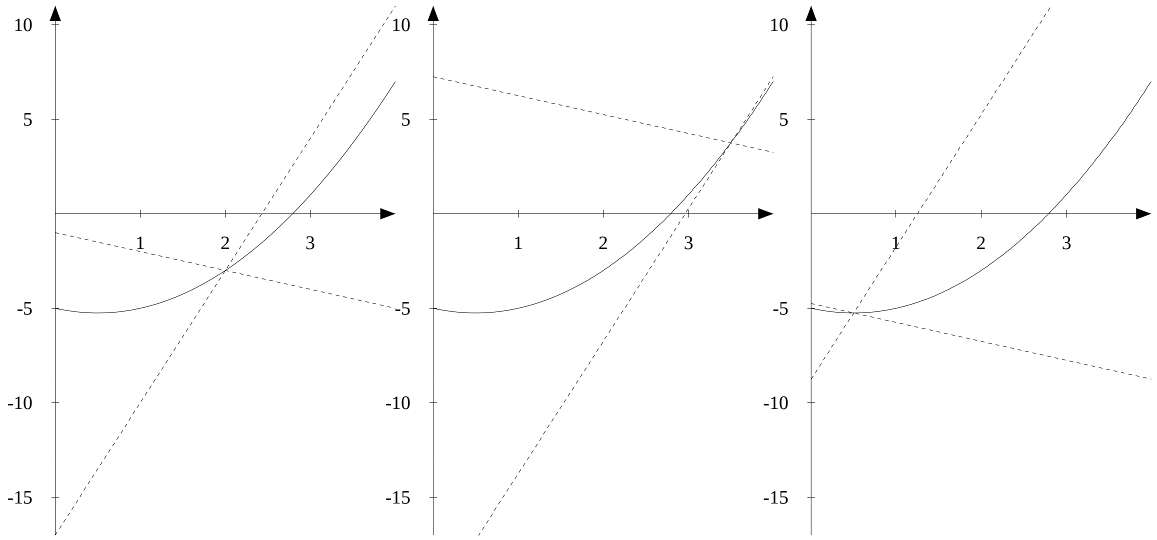

Figure 4.1: Geometric interpretation of the mean value form of $f(x) = x^2 - x - 5$ with different centers. From left to right: $M_f(X, \mathrm{mid}(X))$, $M_f(X, c^{\uparrow})$, $M_f(X, c^{\downarrow})$.

Definition 4.13 (Bicentered Mean Value Form) *The bicentered mean value form $\ddot{M}_f : \mathbb{IR} \to \mathbb{IR}$ is defined as*

$$\ddot{M}_f(X) = [\underline{M_f(X, c^{\downarrow})}, \overline{M_f(X, c^{\uparrow})}], \tag{4.10}$$

where the optimal centers $c^{\uparrow}$ and $c^{\downarrow}$ are

$$\begin{aligned} c^{\uparrow} &= \begin{cases} \overline{X} & \text{if } \underline{F'} \geq 0 \\ \underline{X} & \text{if } \overline{F'} \leq 0 \\ (\overline{F'}\,\overline{X} - \underline{F'}\,\underline{X})/\mathrm{w}(F') & \text{else} \end{cases} \\ c^{\downarrow} &= \begin{cases} \underline{X} & \text{if } \underline{F'} \geq 0 \\ \overline{X} & \text{if } \overline{F'} \leq 0 \\ (\overline{F'}\,\underline{X} - \underline{F'}\,\overline{X})/\mathrm{w}(F') & \text{else} \end{cases} \end{aligned}$$

and $F' = H_{f'}(X)$. $\square$

In the sequel let $c^\uparrow, c^\downarrow$ and F' as in Definition 4.13.

Theorem 4.14 *$\ddot{M}_f$ is an interval extension of f.* $\square$

Proof. As $c^\uparrow, c^\downarrow$ are convex linear combinations of $\overline{X}$ and $\underline{X}$, it holds that $c^\uparrow, c^\downarrow \in X$. From (4.10) it follows that

$$\ddot{M}_f(X) \supseteq M_f(X, c^\downarrow) \cap M_f(X, c^\uparrow).$$

According to (4.1), $f(X) \subseteq M_f(X, c)$ for all $c \in X$. Hence,

$$f(X) \subseteq M_f(X, c^\downarrow) \cap M_f(X, c^\uparrow) \subseteq \ddot{M}_f(X).$$

Finally, if $X \in \mathbb{R}$ then obviously $\ddot{M}_f(X) \in \mathbb{R}$.

Theorem 4.15 (Non–Overestimation) *If $\underline{F'} \geq 0$ or $\overline{F'} \leq 0$ then*

$$\ddot{M}_f(X) = f(X). \ \square$$

Proof.

- Assume $\underline{F'} \geq 0$.

$$\begin{aligned} \ddot{M}_f(X) &= [\underline{M_f(X, \underline{X})}, \overline{M_f(X, \overline{X})}] \\ &= [f(\underline{X}) + \underline{F' * (X - \underline{X})}, f(\overline{X}) + \overline{F' * (X - \overline{X})}] \\ &= [f(\underline{X}), f(\overline{X})]. \end{aligned}$$

- Assume $\overline{F'} \leq 0$.

$$\begin{aligned} \ddot{M}_f(X) &= [\underline{M_f(X, \overline{X})}, \overline{M_f(X, \underline{X})}] \\ &= [f(\underline{X}) + \underline{F' * (X - \overline{X})}, f(\underline{X}) + \overline{F' * (X - \underline{X})}] \\ &= [f(\underline{X}), f(\overline{X})]. \end{aligned}$$

In both cases f is monotone in X, hence $[f(\underline{X}), f(\overline{X})] = f(X)$. $\square$

Theorem 4.16 (Optimality of Centers) *For all $c \in X$ it holds that*

$$\begin{aligned} \overline{M_f(X, c^\uparrow)} &\leq \overline{M_f(X, c)} && (4.11) \\ \underline{M_f(X, c^\downarrow)} &\geq \underline{M_f(X, c)}. && (4.12) \end{aligned}$$

Proof. If $\underline{F'} \geq 0$ or $\overline{F'} \leq 0$ then (4.11) and (4.12) follow from Theorem 4.15. Hence assume $0 \in \mathrm{int}(F')$ and let $c \in X$ arbitrary but fixed. We show (4.11), the proof of (4.12) is similar.

$$\begin{aligned} \overline{M_f(X, c)} &= f(c) + \overline{F' * (X - c)} \\ &\geq f(c^\uparrow) + \underline{F' * (c - c^\uparrow)} + \overline{F' * (X - c)}. && (4.13) \end{aligned}$$

Next, we show that

$$\underline{F' * (c - c^\uparrow)} + \overline{F' * (X - c)} = \overline{F' * (X - c^\uparrow)}. \quad (4.14)$$

As $0 \in F'$ and $0 \in X - c$ we have

$$\overline{F' * (X - c)} = \max\{\underline{F'} * (\underline{X} - c), \overline{F'} * (\overline{X} - c)\} \quad (4.15)$$

and from the definition of $c^\uparrow$ it follows that

$$\overline{F' * (X - c^\uparrow)} = \underline{F'} * (\underline{X} - c^\uparrow) = \overline{F'} * (\overline{X} - c^\uparrow). \quad (4.16)$$

- Assume $c \geq c^{\uparrow}$. From (4.15) and (4.16) it follows that
$$\begin{aligned}\overline{F' * (X - c)} &= \underline{F'} * (\underline{X} - c)\\ \underline{F' * (c - c^{\uparrow})} &= \underline{F'} * (c - c^{\uparrow}).\end{aligned}$$
Hence,
$$\underline{F' * (c - c^{\uparrow})} + \overline{F' * (X - c)} = \underline{F'} * (\underline{X} - c^{\uparrow}) = \overline{F' * (X - c^{\uparrow})}.$$
- Assume $c < c^{\uparrow}$. From (4.15) and (4.16) it follows that
$$\begin{aligned}\overline{F' * (X - c)} &= \overline{F'} * (\overline{X} - c)\\ \underline{F' * (c - c^{\uparrow})} &= \overline{F'} * (c - c^{\uparrow}).\end{aligned}$$
Hence,
$$\underline{F' * (c - c^{\uparrow})} + \overline{F' * (X - c)} = \overline{F'} * (\overline{X} - c^{\uparrow}) = \overline{F' * (X - c^{\uparrow})}.$$

After having shown (4.14), we can continue with (4.13).
$$\begin{aligned}\overline{M_f(X, c)} &\geq f(c^{\uparrow}) + \overline{F' * (X - c^{\uparrow})}\\ &= \overline{M_f(X, c^{\uparrow})}. \ \Box\end{aligned}$$

Corollary 4.17
$$\ddot{M}_f(X) = \bigcap_{c \in X} M_f(X, c). \ \Box$$

Corollary 4.18 (Convergence) *$\ddot{M}_f$ is quadratically convergent.* $\Box$

Proof. According to Corollary 4.17, $\ddot{M}_f(X) \subseteq M_f(X)$ for all $X \in \mathbb{IR}$, and M_f is quadratically convergent by Theorem 4.4. $\Box$

Theorem 4.19 (Inclusion Monotonicity) *$\ddot{M}_f$ is inclusion monotone.* $\Box$

Proof. Let $X_\diamond \subseteq X$, let $c_\diamond^{\uparrow}, c_\diamond^{\downarrow}$ be the optimal centers of f in $X_\diamond$, and let $F'_\diamond = H_{f'}(X_\diamond)$. We have to show that
$$\ddot{M}_f(X_\diamond) \subseteq \ddot{M}_f(X), \tag{4.17}$$
or equivalently
$$\overline{M_f(X_\diamond, c_\diamond^{\uparrow})} \leq \overline{M_f(X, c^{\uparrow})} \tag{4.18}$$
$$\underline{M_f(X_\diamond, c_\diamond^{\downarrow})} \geq \underline{M_f(X, c^{\downarrow})} \tag{4.19}$$

We give a proof of (4.18), the proof of (4.19) is analogous. If $\underline{F'_\diamond} \geq 0$ or $\overline{F'_\diamond} \leq 0$ then $\ddot{M}_f(X_\diamond) = f(X)$ by Theorem 4.15, and (4.17) follows. Thus, in the sequel we assume $0 \in \mathrm{int}(F'_\diamond)$, and inclusion monotonicity of the Horner form implies $0 \in \mathrm{int}(F')$.

- Assume $c_\diamond^{\uparrow} \geq c^{\uparrow}$. Then
$$f(c_\diamond^{\uparrow}) \leq f(c^{\uparrow}) + \overline{F'} * (c_\diamond^{\uparrow} - c^{\uparrow}). \tag{4.20}$$
Hence,
$$\begin{aligned}\overline{M_f(X_\diamond, c_\diamond^{\uparrow})} &= f(c_\diamond^{\uparrow}) + \overline{F'_\diamond * (X_\diamond - c_\diamond^{\uparrow})}\\ &\overset{(4.16)}{=} f(c_\diamond^{\uparrow}) + \overline{F'_\diamond} * (\overline{X_\diamond} - c_\diamond^{\uparrow})\\ &\overset{(4.20)}{\leq} f(c^{\uparrow}) + \overline{F'} * (c_\diamond^{\uparrow} - c^{\uparrow}) + \overline{F'_\diamond} * (\overline{X_\diamond} - c_\diamond^{\uparrow})\\ &\leq f(c^{\uparrow}) + \overline{F'} * (c_\diamond^{\uparrow} - c^{\uparrow}) + \overline{F'} * (\overline{X} - c_\diamond^{\uparrow})\\ &= f(c^{\uparrow}) + \overline{F'} * (\overline{X} - c^{\uparrow})\\ &\overset{(4.16)}{=} f(c^{\uparrow}) + \overline{F' * (X - c^{\uparrow})}\\ &= \overline{M_f(X, c^{\uparrow})}. \ \Box\end{aligned}$$

- Assume $c_\diamond^\uparrow \le c^\uparrow$. Then

$$f(c_\diamond^\uparrow) \le f(c^\uparrow) + \underline{F'} * (c_\diamond^\uparrow - c^\uparrow). \tag{4.21}$$

  Hence,

$$\begin{aligned}
\overline{M_f(X_\diamond, c_\diamond^\uparrow)} &= f(c_\diamond^\uparrow) + \overline{F'_\diamond * (X_\diamond - c_\diamond^\uparrow)} \\
&\overset{(4.16)}{=} f(c_\diamond^\uparrow) + \underline{F'_\diamond} * (\underline{X_\diamond} - c_\diamond^\uparrow) \\
&\overset{(4.21)}{\le} f(c^\uparrow) + \underline{F'} * (c_\diamond^\uparrow - c^\uparrow) + \underline{F'_\diamond} * (\underline{X_\diamond} - c_\diamond^\uparrow) \\
&\le f(c^\uparrow) + \underline{F'} * (c_\diamond^\uparrow - c^\uparrow) + \underline{F'} * (\underline{X} - c_\diamond^\uparrow) \\
&= f(c^\uparrow) + \underline{F'} * (\underline{X} - c^\uparrow) \\
&\overset{(4.16)}{=} f(c^\uparrow) + \overline{F' * (X - c^\uparrow)} \\
&= \overline{M_f(X, c^\uparrow)}. \ \square
\end{aligned}$$

In Figure 4.1 one sees that the width of the mean value form is the same for all centers c between $c^\uparrow$ and $c^\downarrow$. In this case the width is determined by the steeper dashed line. If the center is chosen outside $[c^\downarrow, c^\uparrow]$ then the width of the mean value form is larger. The following Theorem is new.

Theorem 4.20 *Let $C = [c^\downarrow, c^\uparrow]$. For all $c \in X$ it holds that*

$$\mathrm{w}(M_f(X, c)) = \mathrm{mag}(F')\mathrm{w}(X) + \mathrm{w}(F')\mathrm{mig}(C - c). \ \square$$

Proof. Assume $0 \notin \mathrm{int}(F')$. Then $C = X$ and for all $c \in X$ it holds that $\mathrm{mig}(C - c) = 0$ and

$$\begin{aligned}
\mathrm{w}(M_f(X, c)) &= \mathrm{w}(F' * (X - c)) \\
&= \mathrm{w}(\mathrm{mag}(F')(X - c)) \\
&= \mathrm{mag}(F')\mathrm{w}(X).
\end{aligned}$$

Assume $0 \in \mathrm{int}(F')$. Further, assume $\overline{F'} \ge -\underline{F'}$, i.e. $\mathrm{mag}(F') = \overline{F'}$. The case $\overline{F'} < -\underline{F'}$ can be treated similarly. Note that

$$\begin{aligned}
\overline{F'}\,\overline{X} - \underline{F'}\,\underline{X} &\ge \overline{F'}\,\overline{X} - \underline{F'}\,\underline{X} - (\overline{F'} + \underline{F'})(\overline{X} - \underline{X}) \\
&= \overline{F'}\,\underline{X} - \underline{F'}\,\overline{X}.
\end{aligned}$$

Hence, $c^\downarrow \le c^\uparrow$. Further, for all $c \in X$

$$\overline{F' * (X - c)} = \max\{\underline{F'} * (\underline{X} - c), \overline{F'} * (\overline{X} - c)\} \tag{4.22}$$
$$\underline{F' * (X - c)} = \min\{\underline{F'} * (\overline{X} - c), \overline{F'} * (\underline{X} - c)\} \tag{4.23}$$

and

$$\underline{F'} * (\underline{X} - c^\uparrow) = \overline{F'} * (\overline{X} - c^\uparrow) \ \ge 0 \tag{4.24}$$
$$\underline{F'} * (\overline{X} - c^\downarrow) = \overline{F'} * (\underline{X} - c^\downarrow) \ \le 0. \tag{4.25}$$

- Assume $c \in C$. Note that $\mathrm{mig}(C - c) = 0$. As $c \le c^\uparrow$ and $c \ge c^\downarrow$, we obtain from (4.22) – (4.25)

$$\begin{aligned}
\overline{F' * (X - c)} &= \overline{F'} * (\overline{X} - c) \\
\underline{F' * (X - c)} &= \overline{F'} * (\underline{X} - c).
\end{aligned}$$

  Thus,

$$\begin{aligned}
\mathrm{w}(M_f(X, c)) &= \overline{F' * (X - c)} - \underline{F' * (X - c)} \\
&= \overline{F'} * (\overline{X} - \underline{X}) \\
&= \overline{F'}\mathrm{w}(X).
\end{aligned}$$

- Assume $c \notin C$. We give a proof only for the case $c > c^{\uparrow}$, the case $c < c^{\downarrow}$ is analogous. As $c > c^{\downarrow}$ and $c > c^{\uparrow}$ we obtain from (4.22) – (4.25)

$$
\begin{aligned}
\overline{F' * (X - c)} &= \underline{F'} * (\underline{X} - c) \\
\underline{F' * (X - c)} &= \overline{F'} * (\underline{X} - c).
\end{aligned}
$$

  Thus,

$$
\begin{aligned}
\mathrm{w}(M_f(X,c)) &= \overline{F' * (X - c)} - \underline{F' * (X - c)} \\
&= (\underline{F'} - \overline{F'})(\underline{X} - c) \\
&= (\underline{F'} - \overline{F'})(\underline{X} - c^{\uparrow}) + (\underline{F'} - \overline{F'})(c^{\uparrow} - c) \\
&= \underline{F'} * (\underline{X} - c^{\uparrow}) - \overline{F'} * (\underline{X} - c^{\uparrow}) + \mathrm{w}(F)\mathrm{mig}(C - c) \\
&= \overline{F'} * (\overline{X} - c^{\uparrow}) - \overline{F'} * (\underline{X} - c^{\uparrow}) + \mathrm{w}(F)\mathrm{mig}(C - c) \\
&= \overline{F'}\mathrm{w}(X) + \mathrm{w}(F)\mathrm{mig}(C - c). \ \square
\end{aligned}
$$

The midpoint of X is an optimal center for the mean value form in the following sense:

Corollary 4.21 (Optimality of the Midpoint for the Mean Value Form) *For all $c \in X$ it holds that*

$$\mathrm{w}(M_f(X)) \leq \mathrm{w}(M_f(X,c)). \ \square$$

Proof. According to Theorem 4.20, we have to show that $\mathrm{mid}(X) \in [c^{\downarrow}, c^{\uparrow}]$. In fact, we show that $\mathrm{mid}(X) = \mathrm{mid}([c^{\downarrow}, c^{\uparrow}])$. If $0 \notin \mathrm{int}(F')$ then $[c^{\downarrow}, c^{\uparrow}] = X$. If $0 \in \mathrm{int}(F')$ then

$$
\begin{aligned}
1/2(c^{\downarrow} + c^{\uparrow}) &= \frac{\overline{F'}\,\overline{X} - \underline{F'}\,\underline{X} + \overline{F'}\,\underline{X} - \underline{F'}\,\overline{X}}{2\mathrm{w}(F')} \\
&= \frac{(\overline{F'} - \underline{F'})(\overline{X} + \underline{X})}{2(\overline{F'} - \underline{F'})} \\
&= 1/2(\overline{X} + \underline{X}) \\
&= \mathrm{mid}(X). \ \square
\end{aligned}
$$

Algorithm and Complexity.

Algorithm 4.4 (BMF) for evaluating the bicentered mean value form follows immediately from Defintion 4.13. First, we give Algorithm 4.3 (OC) for computing the optimal centers. Note that an approximation of the optimal centers is sufficient for the correctness of Algorithm 4.4, provided that the approximations are elements of X.

Algorithm 4.3 (OC**)** [Optimal Centers]

In: $F' \in \mathbb{IF}$,
$X \in \mathbb{IF}$.

Out: $c^{\uparrow}, c^{\downarrow} \in X$, approximations of the optimal centers according to Definition 4.13.

(1) [Monotonicity Test.]
if $\underline{F'} \geq 0$ then return $\overline{X}, \underline{X}$
if $\overline{F'} \leq 0$ then return $\underline{X}, \overline{X}$

(2) [Approximate.]
clear invalid operation flag
$c^{\uparrow} \longleftarrow (\overline{F'} \stackrel{\approx}{*} \overline{X} \stackrel{\approx}{-} \underline{F'} \stackrel{\approx}{*} \underline{X}) \stackrel{\approx}{/} (\overline{F'} \stackrel{\approx}{-} \underline{F'})$
$c^{\downarrow} \longleftarrow (\overline{F'} \stackrel{\approx}{*} \underline{X} \stackrel{\approx}{-} \underline{F'} \stackrel{\approx}{*} \overline{X}) \stackrel{\approx}{/} (\overline{F'} \stackrel{\approx}{-} \underline{F'})$

(3) [Check Floating Point Errors.]
if invalid operation flag is raised then $c^{\uparrow} \longleftarrow \overline{X}$, $c^{\downarrow} \longleftarrow \underline{X}$
if $c^{\uparrow} < \underline{\underline{X}}$ then $c^{\uparrow} \longleftarrow \underline{\underline{X}}$
if $c^{\uparrow} > \overline{\overline{X}}$ then $c^{\uparrow} \longleftarrow \overline{\overline{X}}$
if $c^{\downarrow} < \underline{\underline{X}}$ then $c^{\downarrow} \longleftarrow \underline{\underline{X}}$
if $c^{\downarrow} > \overline{\overline{X}}$ then $c^{\downarrow} \longleftarrow \overline{\overline{X}}$

(4) [Return.]
return $c^{\uparrow}, c^{\downarrow}$

Remark. Note that in Step 2 $\overline{F} > 0$ and $\underline{F} < 0$, hence $\overline{F'} \overset{\approx}{-} \underline{F'} \neq 0$ and division by zero cannot occur.

Theorem 4.22 (Complexity) *Algorithm 4.3* (OC) *costs*

2	*number divisions,*
4	*number multiplications and*
3	*number additions.* □

Algorithm 4.4 (BMF) [Bicentered Mean Value Form]

In: $f(x) = a_n x^{d_n} + a_{n-1} x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{BMF}(f, X) \in \mathbb{IF}$, $\mathrm{BMF}(f, X) \supseteq \ddot{M}_f(X)$.

(1) [Inclusion of f'.]
$F'(x) \longleftarrow f'(x)$ (use interval arithmetic for coefficients)

(2) [Evaluate $f'(X)$.]
$F' \longleftarrow \mathrm{HFI}(F'(x), X)$

(3) [Optimal Centers.]
$c^{\uparrow}, c^{\downarrow} \longleftarrow \mathrm{OC}(F', X)$

(4) [Evaluate at Centers.]
$y^{\uparrow} \longleftarrow \mathrm{HFUB}(f(x), [c^{\uparrow}]) \mathbin{\hat{+}} \overline{F' * (X - c^{\uparrow})}$
$y^{\downarrow} \longleftarrow \mathrm{HFLB}(f(x), [c^{\downarrow}]) \mathbin{\check{+}} \underline{F' * (X - c^{\downarrow})}$
return $[y^{\downarrow}, y^{\uparrow}]$

Theorem 4.23 (Complexity) *Algorithm 4.4* (BMF) *costs*

n	*interval power computations,*
$n+2$	*interval multiplications,*
$2n$	*number power computations,*
2	*number divisions,*
$4n+4$	*number multiplications and*
$4n+5$	*number additions.* □

Proof.

- Step 1 costs $2n$ number multiplications.
- Step 2 costs n interval power computations, n interval multiplications and $2n-2$ number additions (Theorem 3.6).
- Step 3 costs 2 number divisions, 4 number multiplications and 3 number additions (Theorem 4.22).
- Step 4 costs $2n$ number power computations, $2n$ number multiplications and $2n-2$ number additions during the calls of HFLB and HFUB (Theorem 3.31). Further, it takes 2 interval multiplications and 6 number additions. □

4.3 Experimental Results

An experimental comparison of the mean value form M_f, computed by Algorithm 4.1 (MF), the dense linear factor form $M_f^{\overset{*}{\leadsto}}$, computed by Algorithm 4.2 (DSF) and the bicentered mean value form $\ddot{M}_f$, computed by Algorithm 4.4 (BMF) for dense and sparse polynomials with different degrees is given in Table 4.1, respectively Table 4.2. The coefficients of the polynomomials and the endpoints of the input intervals are uniformly distributed in $[-1, 1]$. For each degree d_n, the average cost and overestimation error of 10^4 random polynomials is reported. The cost is the total number of arithmetic floating point instructions, including those which were executed during interval operations. As the forms are quadratically convergent, the distance to the range divided by $\mathrm{w}(X)^2$ was chosen to measure accuracy.

Degree of f	2	6	10	14	18	22	26	30
Flops for $M_f(X)$	28.6	71.2	113.7	156.2	198.8	241.3	283.7	325.8
Flops for $M_f^{\overset{*}{\leadsto}}(X)$	20.3	54.2	88.5	122.7	157.1	191.5	225.8	259.7
Flops for $\ddot{M}_f(X)$	25.3	73.2	116.5	159.1	202.1	244.5	287.1	329.1
$\mathrm{q}(M_f(X), f(X))/\mathrm{w}(X)^2$	0.36	1.50	2.55	3.47	4.40	5.48	6.24	7.00
$\mathrm{q}(M_f^{\overset{*}{\leadsto}}(X), f(X))/\mathrm{w}(X)^2$	0.23	0.70	1.05	1.30	1.56	1.87	2.07	2.16
$\mathrm{q}(\ddot{M}_f(X), f(X))/\mathrm{w}(X)^2$	0.07	0.43	0.77	1.04	1.34	1.66	1.90	2.24

Table 4.1: Comparison of M_f, $M_f^{\overset{*}{\leadsto}}$ and $\ddot{M}_f$ for *dense* random polynomials f with different degrees and random intervals X. The coefficients of f and the endpoints of X are uniformly chosen in $[-1, 1]$.

Degree of f	2	6	10	14	18	22	26	30
Flops for $M_f(X)$	28.6	46.0	57.6	65.3	71.4	75.7	79.5	82.9
Flops for $M_f^{\overset{*}{\leadsto}}(X)$	20.3	47.3	74.6	102.4	130.3	158.2	186.3	213.9
Flops for $\ddot{M}_f(X)$	25.3	45.0	56.5	64.1	70.5	74.6	78.4	81.6
$\mathrm{q}(M_f(X), f(X))/\mathrm{w}(X)^2$	0.36	0.95	1.26	1.42	1.55	1.72	1.71	1.90
$\mathrm{q}(M_f^{\overset{*}{\leadsto}}(X), f(X))/\mathrm{w}(X)^2$	0.23	0.46	0.54	0.57	0.59	0.64	0.60	0.66
$\mathrm{q}(\ddot{M}_f(X), f(X))/\mathrm{w}(X)^2$	0.07	0.16	0.17	0.16	0.19	0.18	0.20	0.21

Table 4.2: Comparison of M_f, $M_f^{\overset{*}{\leadsto}}$ and $\ddot{M}_f$ for *sparse* random polynomials f with different degrees and random intervals X. Each polynomial has only 3 non–zero coefficients. The coefficients of f and the endpoints of X are uniformly chosen in $[-1, 1]$.

In all cases $\ddot{M}_f$ is more accurate thant $M_f^{\leadsto}$, which is significantly more accurate than M_f. The costs of $\ddot{M}_f$ and M_f are almost the same. If the polynomials are "sufficiently" sparse, then $\ddot{M}_f$ and M_f are more efficient than $M_f^{\overset{*}{\leadsto}}$. In the dense case $M_f^{\overset{*}{\leadsto}}$ is cheaper.

5 Taylor Form

The basic idea of the Taylor form is to translate f by $\mathrm{mid}(X)$ to the left and evaluate the Horner form of resulting polynomial on the centered interval $X - \mathrm{mid}(X)$. The Taylor form is a centered form and hence quadratically convergent.

The Taylor form for polynomials and rational functions was studied thoroughly in the literature, see e.g. [Alefeld and Herzberger, 1974], [Alefeld and Rokne, 1981], [Alefeld and Herzberger, 1983], [Hansen, 1969], [Moore, 1966], [Rall, 1983], [Ratschek, 1978], [Ratschek, 1980a], [Ratschek, 1980b], [Ratschek and Rokne, 1980b], [Ratschek and Rokne, 1980a], [Ratschek and Schröder, 1981], [Ratschek and Rokne, 1984]. Interval polynomials are considered in [Rokne, 1981] and complex polynomials in [Rokne and Wu, 1982], [Rokne and Wu, 1983].

According to Taylor's Theorem we can rewrite f for arbitrary $c \in \mathbb{R}$ as

$$f(x) = \sum_{i=0}^{d_n} \frac{f^{(i)}(c)}{i!}(x-c)^i.$$

If we translate f by c to the left, we obtain the polynomial

$$f^{(c)}(x) = \sum_{i=0}^{d_n} \frac{f^{(i)}(c)}{i!} x^i \tag{5.1}$$

and obviously

$$f(x) = f^{(c)}(x-c).$$

Thus, a whole class of interval extensions of f is given by

$$f(X) \subseteq F^{(c)}(X-c), \tag{5.2}$$

where $F^{(c)}$ is an interval extension of $f^{(c)}$. The Taylor form is a special case of (5.2), where $F^{(c)}$ is the Horner form of $f^{(c)}$ and c is the midpoint of X.

Definition 5.1 (Taylor Form) *The Taylor form $T_f : \mathbb{IR} \to \mathbb{IR}$ of f is defined as*

$$T_f(X) = H_{f(\mathrm{mid}(X))}(X - \mathrm{mid}(X)).$$

The dense Taylor form $T_f^* : \mathbb{IR} \to \mathbb{IR}$ of f is defined as

$$T_f^*(X) = H^*_{f(\mathrm{mid}(X))}(X - \mathrm{mid}(X)). \ \square$$

From Theorem 3.3 it follows that $T_f(X) \subseteq T_f^*(X)$ for all $X \in \mathbb{IR}$. The reason why it is still worthwhile to consider T_f^*, is that it can be evaluated more efficiently than T_f. This is surprising, because the evaluation of H_f is usually cheaper than the evaluation of H_f^*. Another reason is that T_f^* is inclusion montone, whereas T_f is not. Finally, $T_f(X) \subset T_f^*(X)$ only if a Taylor coefficient vanishes, which is "in general" not the case, especially if rounded interval arithmetic is used for their computation.

In the sequel let $c = \mathrm{mid}(X)$ and $f^{(c)}$ as in (5.1). Further, let

$$g^{(c)}(x) = \sum_{i=1}^{d_n} \frac{f^{(i)}(c)}{i!} x^{i-1}$$

and note that

$$T_f^*(X) = f(c) + H^*_{g^{(c)}}(X-c)(X-c). \tag{5.3}$$

A corresponding representation is not possible for T_f if $f'(c) = 0$.

The following Theorem shows, why T_f^* can be evaluated faster than T_f. Further, it will be used to prove inclusion monotonicity of T_f^*.

Theorem 5.2 *Let $b_0, \ldots, b_{d_n-1}$ such that*

$$g^{(c)}(x) = \sum_{i=0}^{d_n-1} b_i x^i,$$

let

$$g_{||}^{(c)}(x) = \sum_{i=0}^{d_n-1} |b_i| x^i,$$

and let $r = \mathrm{rad}(X)$. *Then*

$$H^*_{g^{(c)}}(X-c)\,(X-c) = g_{||}^{(c)}(r) r[-1,1]. \ \square$$

Proof. Let r and $b_0, \ldots, b_{d_n-1}$ as in Theorem 5.2. As $X - c$ is a centered interval, it holds that

$$H^*_{g^{(c)}}(X-c)\,(X-c) = \mathrm{mag}(H^*_{g^{(c)}}(X-c)) r[-1,1].$$

It remains to show that

$$\mathrm{mag}(H^*_{g^{(c)}}(X-c)) = g_{||}^{(c)}(r).$$

Let

$$\left.\begin{array}{rclrcl} P_{n-1} & = & b_{n-1}, & P_i & = & P_{i+1}(X-c) + b_i \\ p_{n-1} & = & |b_{n-1}|, & p_i & = & p_{i+1} r + |b_i| \end{array}\right\} i = n-2, \ldots, 0.$$

Note that $P_0 = H^*_{g^{(c)}}(X-c)$ and $p_0 = g_{||}^{(c)}(r)$. Obviously $\mathrm{mag}(P_{n-1}) = p_{n-1}$. If $\mathrm{mag}(P_i) = p_i$ for some $1 \leq i \leq n-1$ then

$$\mathrm{mag}(P_{i-1}) = \mathrm{mag}(P_i * (X-c) + b_{i-1}) = \mathrm{mag}(P_i) r + |b_{i-1}| = p_{i-1},$$

hence $\mathrm{mag}(P_0) = p_0$. $\square$

In [Caprani and Madsen, 1980] inclusion monotonicity of a centered form similar to the Taylor form (see [Ratschek, 1978], [Ratschek, 1980a]) is disproved by a counterexample. Apparently, inclusion monotonicity of the Taylor form was not studied before. In particular, Theorem 5.3 seems to be new.

Theorem 5.3 (Inclusion Monotonicity of Dense Taylor Form) T^*_f *is inclusion monotone.* $\square$

Proof. Let $X_\diamond \subseteq X$, $c_\diamond = \mathrm{mid}(X_\diamond)$, $r_\diamond = \mathrm{rad}(X_\diamond)$ and $r = \mathrm{rad}(X)$. Note that

$$\begin{array}{rcl} |c_\diamond - c| & = & 1/2 |(\overline{X_\diamond} + \underline{X_\diamond} - \overline{X} - \underline{X})| \\ & = & 1/2 \max\{\overline{X_\diamond} + \underline{X_\diamond} - \overline{X} - \underline{X}, \overline{X} + \underline{X} - \overline{X_\diamond} - \underline{X_\diamond}\} \\ & \leq & 1/2(\overline{X} + \underline{X_\diamond} - \overline{X_\diamond} - \underline{X}) \\ & = & r - r_\diamond. \end{array}$$

We have to show that $T^*_f(X_\diamond) \subseteq T^*_f(X)$, which can be rewritten according to Theorem 5.2 as

$$f(c_\diamond) + [-1,1] \sum_{i=1}^{d_n} \frac{|f^{(i)}(c_\diamond)|}{i!} r_\diamond^i \subseteq f(c) + [-1,1] \sum_{i=1}^{d_n} \frac{|f^{(i)}(c)|}{i!} r^i. \tag{5.4}$$

The corresponding inequalities for the endpoints are

$$f(c_\diamond) + \sum_{i=1}^{d_n} \frac{|f^{(i)}(c_\diamond)|}{i!} r_\diamond^i \quad \leq \quad f(c) + \sum_{i=1}^{d_n} \frac{|f^{(i)}(c)|}{i!} r^i \tag{5.5}$$

$$f(c_\diamond) - \sum_{i=1}^{d_n} \frac{|f^{(i)}(c_\diamond)|}{i!} r_\diamond^i \quad \geq \quad f(c) - \sum_{i=1}^{d_n} \frac{|f^{(i)}(c)|}{i!} r^i. \tag{5.6}$$

First, we express the derivatives of f at $c_\diamond$ in terms of derivatives of f at c.

$$f^{(i)}(c_\diamond) = \sum_{j=i}^{d_n} \frac{f^{(j)}(c)}{(j-i)!} (c_\diamond - c)^{j-i}.$$

Thus,

$$\begin{aligned}
\sum_{i=1}^{d_n} \frac{|f^{(i)}(c_\diamond)|}{i!} r_\diamond^i &= \sum_{i=1}^{d_n} \left| \sum_{j=i}^{d_n} \frac{f^{(j)}(c)}{i!(j-i)!}(c_\diamond - c)^{j-i} \right| r_\diamond^i \\
&\le \sum_{i=1}^{d_n} \sum_{j=i}^{d_n} \frac{|f^{(j)}(c)|}{i!(j-i)!} |(c_\diamond - c)|^{j-i} r_\diamond^i \\
&\le \sum_{i=1}^{d_n} \sum_{j=i}^{d_n} \frac{|f^{(j)}(c)|}{i!(j-i)!} (r - r_\diamond)^{j-i} r_\diamond^i \\
&= \sum_{j=1}^{d_n} \sum_{i=1}^{j} \frac{|f^{(j)}(c)|}{i!(j-i)!} (r - r_\diamond)^{j-i} r_\diamond^i \\
&= \sum_{j=1}^{d_n} \frac{|f^{(j)}(c)|}{j!} \sum_{i=1}^{j} \binom{j}{i} (r - r_\diamond)^{j-i} r_\diamond^i \\
&= \sum_{j=1}^{d_n} \frac{|f^{(j)}(c)|}{j!} \left(r^j - (r - r_\diamond)^j\right) \\
&= \sum_{j=1}^{d_n} \frac{|f^{(j)}(c)|}{j!} r^j - \sum_{j=1}^{d_n} \frac{|f^{(j)}(c)|}{j!} (r - r_\diamond)^j \\
&\le \sum_{j=1}^{d_n} \frac{|f^{(j)}(c)|}{j!} r^j - \sum_{j=1}^{d_n} \frac{|f^{(j)}(c)|}{j!} |c_\diamond - c|^j \\
&\le \sum_{j=1}^{d_n} \frac{|f^{(j)}(c)|}{j!} r^j - \left| \sum_{j=1}^{d_n} \frac{f^{(j)}(c)}{j!} (c_\diamond - c)^j \right| \\
&= \sum_{j=1}^{d_n} \frac{|f^{(j)}(c)|}{j!} r^j - |f(c_\diamond) - f(c)|.
\end{aligned}$$

Finally, (5.5) and (5.6) follow from

$$\begin{aligned}
f(c_\diamond) - |f(c_\diamond) - f(c)| &\le f(c). \\
f(c_\diamond) + |f(c_\diamond) - f(c)| &\ge f(c). \ \square
\end{aligned}$$

Theorem 5.4 *T_f is not inclusion monotone.* □

Proof. Let $f(x) = x^3 - 3 * x^2$.

- Let $X = [0, 2]$, $c = 1$. The Taylor coefficients are

$$\begin{aligned}
f(c) &= -2 \\
f'(c)/1! &= -3 \\
f''(c)/2! &= 0 \\
f'''(c)/3! &= 1,
\end{aligned}$$

  and

$$T_f(X) = -2 + ([-1, 1]^2 - 3)[-1, 1] = [-5, 1].$$

  Note that $f''(c)/2! = 0$, which caused that the interval square function was used instead of multiplication.

- Let $X_\diamond = [0, 1.8]$, $c_\diamond = 0.9$. The Taylor coefficients are

$$\begin{aligned} f(c_\diamond) &= -1.701 \\ f'(c_\diamond)/1! &= -2.97 \\ f''(c_\diamond)/2! &= -0.3 \\ f'''(c_\diamond)/3! &= 1, \end{aligned}$$

  and

$$T_f(X_\diamond) = -1.701 + \big(([-0.9, 0.9] - 0.3)[-0.9, 0.9] - 2.97\big)[-0.9, 0.9] = [-5.346, 1.944].$$

Thus $X_\diamond \subset X$ but $T_f(X_\diamond) \supset T_f(X)$. □

Theorem 5.5 (Convergence) *The Taylor form and the dense Taylor form are quadratically convergent.* □

Proof. According to (5.3) the dense Taylor form can be written as

$$T_f^*(X) = f(c) + H^*_{g(c)}(X - c)(X - c).$$

Let $G(X) = H^*_{g(c)}(X - c)$. $H^*_{g(c)}$ and hence G are Lipschitz. Further, $g^{(c)}(x - c) \in G(X)$ for all $x \in X$. Hence T_f^* is a centered form and quadratically convergent by [Ratschek and Rokne, 1984]. □

Algorithm and Complexity.

We give an algorithm for computing Taylor coefficients by the extended Horner scheme first.

Algorithm 5.1 (TC) [Taylor Coefficients]

In: $f(x) = a_n x^{d_n} + a_{n-1}x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$c \in \mathbb{F}$.

Out: $\mathrm{TC}(f, c) = (A_0, \ldots, A_{d_n})$, $A_i \in \mathbb{IF}$, $f^{(i)}(c)/i! \in A_i$, $i = 0, \ldots, d_n$.

(1) [Initialize.]
for $i = 0, \ldots, d_n$ do $A_i \longleftarrow$ coefficient of x^i in f
$S \longleftarrow cA_{d_n}$

(2) [Horner Scheme.]
for $i = 0, \ldots, d_n - 1$
  $A_{d_n-1} \longleftarrow A_{d_n-1} + S$
  for $j = d_n - 1, \ldots, i + 1$ do $A_{j-1} \longleftarrow A_{j-1} + cA_j$

(3) [Return.]
return $A_0, \ldots, A_{d_n}$

Theorem 5.6 (Comlexity) *Algorithm 5.1* (TC) *costs*

$$\begin{aligned} d_n^2 - d_n + 2 &\quad \textit{number multiplications and} \\ d_n^2 + d_n &\quad \textit{number additions.} \ \square \end{aligned}$$

Proof.

- Step 1 costs 2 number multiplication.
- The i-th iteration in step 2 costs $2(d_n - i - 1)$ number multiplications and $2(d_n - i)$ number additions. Hence step 2 costs $d_n(d_n - 1)$ number multiplications and $d_n(d_n + 1)$ number additions.

Algorithm 5.2 evaluates the dense Taylor form. For the evaluation of the translated polynomial we use Theorem 5.2 instead of Horner form. This allows to replace interval operations by floating point operations.

Algorithm 5.2 (DTF) [Dense Taylor Form]

In: $f(x) = a_n x^{d_n} + a_{n-1} x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{DTF}(f, X) \in \mathbb{IF}$, $\mathrm{DTF}(f, X) \supseteq T_f^*(X)$.

(1) [Special case $\overline{X} = \underline{X}$.]
if $\overline{X} = \underline{X}$ then return $H_f(X)$.

(2) [Midpoint and Radius of X.]
$c \longleftarrow \mathrm{MID}(X)$
$r \longleftarrow \max\{\overline{X} \overset{\wedge}{-} c, c \overset{\wedge}{-} \underline{X}\}$

(3) [Taylor coefficients.]
$(A_0, \ldots, A_{d_n}) \longleftarrow \mathrm{TC}(f, c)$

(4) [Magnitude of $H^*_{g^{(c)}}(X - c)(X - c)$.]
$m \longleftarrow \mathrm{mag}(A_{d_n}) \overset{\wedge}{*} r$
for $i = d_n - 1 \ldots 1$ do $m \longleftarrow (m \overset{\wedge}{+} \mathrm{mag}(A_i)) \overset{\wedge}{*} r$

(5) [Return.]
return $A_0 + [-m, m]$

Remark. The special case $\overline{X} = \underline{X}$ was excluded in Step 1, hence $r > 0$ in Step 4 and an invalid operation cannot occur.

Theorem 5.7 (Complexity) *If $\overline{X} \neq \underline{X}$ then Algorithm* DTF *costs*

$$\begin{array}{rl} d_n^2 + 3 & \textit{number multiplications and} \\ d_n^2 + 2d_n + 5 & \textit{number additions.} \ \square \end{array}$$

Proof.

- Step 2 costs 1 number multiplication and 4 number additions.
- Step 3 costs $d_n^2 - d_n + 2$ number multiplications and $d_n^2 + d_n$ number additions (Theorem 5.6).
- Step 4 costs d_n number multiplications and $d_n - 1$ number additions.
- Step 5 costs 2 number additions. □

5.1 Bisection at Zero

The Taylor form reduces the computation of the range of f on an arbitrary interval X to the evaluation of the Horner form of the Taylor polynomial $f^{(c)}$ on a centered interval $X - c$. According to Theorem 3.37, the overestimation error of the Horner form can be reduced at least by half if $X - c$ is bisected at zero and Horner form is evaluated on both halves. In Section 3.3 it was pointed out, that such a bisection is relatively inexpensive. The following modification of the Taylor form is therefore slightly more expensive than T_f^*, but gives significantly tighter inclusions. To my knowledge this modification is original.

In the sequel let

$$\breve{H}_f^*(X) = \begin{cases} H_f^*([\underline{X}, 0]) \cup H_f([0, \overline{X}]) & \text{if } 0 \in \mathrm{int}(X) \\ H_f^*(X) & \text{else.} \end{cases}$$

Definition 5.8 (Dense Taylor Form with Bisection) *The dense Taylor form with bisection* $\check{T}_f^* : \mathbb{IR} \to \mathbb{IR}$ *of* f *is defined as*

$$\check{T}_f^*(X) = \check{H}^*_{f(\mathrm{mid}(X))}(X - \mathrm{mid}(X)). \ \square$$

Theorem 5.9 *For all* $X \in \mathbb{IR}$ *it holds that*

$$\begin{aligned} |\overline{f(X)} - \overline{\check{T}_f^*(X)}| &\leq 1/2|\overline{f(X)} - \overline{T_f^*(X)}| \\ |\underline{f(X)} - \underline{\check{T}_f^*(X)}| &\leq 1/2|\underline{f(X)} - \underline{T_f^*(X)}|. \ \square \end{aligned}$$

Proof. Follows immediately from Theorem 3.37. $\square$

Algorithm 5.3 (DTFB) [Dense Taylor Form with Bisection]

In: $f(x) = a_n x^{d_n} + a_{n-1}x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{DTFB}(f, X) \in \mathbb{IF}$, $\mathrm{DTFB}(f, X) \supseteq \check{T}_f^*(X)$.

(1) [Midpoint of X.]
$c \longleftarrow \mathrm{MID}(X)$

(2) [Coefficients of $f^{(c)}(x)$]
$(A_0, \ldots, A_{d_n}) \longleftarrow \mathrm{TC}(f, c)$

(3) [Initialize Floating Point Operations.]
clear invalid operation flag

(4) [Right Half.]
$x \longleftarrow \overline{X} \overset{\wedge}{-} c$
$R \longleftarrow A_{d_n}$
for $i = d_n - 1, \ldots, 0$
  if $\overline{R} > 0$ then $\overline{R} \longleftarrow \overline{R} \overset{\wedge}{*} x \overset{\wedge}{+} \overline{A}_i$, else $\overline{R} \longleftarrow \overline{A}_i$
  if $\underline{R} < 0$ then $\underline{R} \longleftarrow \underline{R} \overset{\vee}{*} x \overset{\vee}{+} \underline{A}_i$, else $\underline{R} \longleftarrow \underline{A}_i$

(5) [Coefficients of $f^{(c)}(-x)$.]
for $i = 1, \ldots, d_n$ step 2 do $A_i \longleftarrow -A_i$

(6) [Left Half.]
$x \longleftarrow c \overset{\wedge}{-} \underline{X}$
$L \longleftarrow A_{d_n}$
for $i = d_n - 1, \ldots, 0$
  if $\overline{L} > 0$ then $\overline{L} \longleftarrow \overline{L} \overset{\wedge}{*} x \overset{\wedge}{+} \overline{A}_i$, else $\overline{L} \longleftarrow \overline{A}_i$
  if $\underline{L} < 0$ then $\underline{L} \longleftarrow \underline{L} \overset{\vee}{*} x \overset{\vee}{+} \underline{A}_i$, else $\underline{L} \longleftarrow \underline{A}_i$

(7) [Check Invalid Operation and Return.]
if invalid operation flag is raised, then return $[\perp, \top]$, else return $R \cup L$

Theorem 5.10 (Complexity) *Algorithm* 5.3 *(*DTFB*) costs*

$$\begin{array}{ll} d_n^2 + 3d_n + 3 & \textit{number multiplications and} \\ d_n^2 + 5d_n + 4 & \textit{number additions.} \ \square \end{array}$$

Proof.

- Step 1 costs 1 number multiplication and 2 number additions.
- Step 2 costs $d_n^2 - d_n + 2$ number multiplications and $d_n^2 + d_n$ number additions.
- Step 4 and 6 costs each $2d_n$ number multiplications and $2d_n + 1$ number additions. $\square$

Hence, Algorithm 5.3 (DTFB) costs $3d_n$ number multiplications and $3d_n - 1$ number additions more than Algorithm 5.2 (DTF).

5.2 Experimental Results

An experimental comparison between Algorithm 5.2 (DTF) and Algorithm 5.3 (DTFB) for dense polynomials with different degrees is given in Table 5.1. The coefficients of the polynomomials and the endpoints of the intervals are uniformly distributed in $[-1, 1]$. For each degree d_n, the average cost and overestimation error of 10^4 random polynomials is reported. The cost is the total number of arithmetic floating point instructions, including those which were executed during interval operations. As the forms are quadratically convergent, the distance to the range divided by $\mathrm{w}(X)^2$ was chosen to measure accuracy.

While the measured costs of Algorithm 5.2 (DTF) correspond exactly to Theorem 5.7, the measured costs of Algorithm 5.3 (DTFB) are significantly smaller than stated by Theorem 5.10. The reason is, that the worst case assumption of Theorem 5.10 does not hold in most cases.

Degree of f	2	6	10	14	18	22	26	30
Flops for $T_f^*(X)$	20.0	92.0	228.0	428.0	692.0	1020.0	1412.0	1868.0
Flops for $\check{T}_f^*(X)$	22.3	104.1	250.0	459.8	733.8	1071.7	1473.6	1939.1
$\mathrm{q}(T_f^*(X), f(X))/\mathrm{w}(X)^2$	0.23	0.70	1.06	1.31	1.58	1.90	2.10	2.22
$\mathrm{q}(\check{T}_f^*(X), f(X))/\mathrm{w}(X)^2$	0.11	0.26	0.36	0.43	0.52	0.60	0.66	0.69

Table 5.1: Comparison of T_f^* and $\check{T}_f^*$ for dense random polynomials f with different degrees and random intervals X. The coefficients of f and the endpoints of X are uniformly chosen in $[-1, 1]$.

6 Bernstein Form

It is well known that the $k+1$ Bernstein polynomials of order k form a basis of the vector space of polynomials of degree at most k. This means that every such polynomial f can be written as a linear combination of k-th order Bernstein polynomials. The coefficients of such a linear combination have an important property, namely the largest and the smallest coefficient bound the range of f in the interval $[0, 1]$. This observation can be generalized to bound the range of f on arbitrary intervals X.

Experimental results indicate that the Bernstein form gives very tight inclusions compared to the other forms described so far. There are many situations when Bernstein form yields the range without overestimation. An a posteriori test, whether such a situation is given, does not require any additional computation.

The idea to use Bernstein polynomials for range computation goes back to [Cargo and Shisha, 1966] and [Rivlin, 1970]. The case $X \neq [0, 1]$ was considered first by [Rokne, 1977]. Efficient algorithms for real and for interval polynomials are presented in [Rokne, 1979a] respectively [Rokne, 1982]. Complex polynomials are studied in [Rokne, 1979b] and [Grassmann and Rokne, 1979]. The multivariate case is treated in [Garloff, 1985]. An application to real root isolation can be found in [Lane and Riesenfeld, 1981]. The inclusion monotonicity of the Bernstein form (Theorem 6.17) was proved by [Hong and Stahl, 1994a]. A new criterion when the Bernstein form gives the range without overestimation is subject of Theorem 6.19.

Let us first review some basic properties of Bernstein polynomials.

Definition 6.1 (Bernstein Polynomial) *The j-th Bernstein polynomial of order k is defined as*

$$p_j^{(k)}(x) = \binom{k}{j} x^j (1-x)^{k-j}, \quad k \geq 0, \quad j = 0, \ldots, k. \;\square \tag{6.1}$$

The Bernstein polynomials $\{p_j^{(k)}(x) \mid j = 0, \ldots, k\}$ form a basis of the vector space of polynomials of degree $\leq k$. All we have to show is that the elements x^i, $i = 0, \ldots, k$ of the power basis are linear combinations of Bernstein polynomials.

Lemma 6.2 *For every $k \geq i \geq 0$ it holds that*

$$x^i = \sum_{j=i}^{k} \frac{\binom{j}{i}}{\binom{k}{i}} p_j^{(k)}(x). \ \Box$$

Proof.

$$\begin{aligned}
x^i &= x^i(x + (1-x))^{k-i} \\
&= \sum_{r=0}^{k-i} \binom{k-i}{r} x^{i+r}(1-x)^{k-i-r} \\
&= \sum_{j=i}^{k} \binom{k-i}{j-i} x^j (1-x)^{k-j} \\
&= \sum_{j=i}^{k} \frac{\binom{j}{i}}{\binom{k}{i}} \binom{k}{j} x^j (1-x)^{k-j} \\
&= \sum_{j=i}^{k} \frac{\binom{j}{i}}{\binom{k}{i}} p_j^{(k)}(x). \ \Box
\end{aligned}$$

In the sequel let $k \geq d_n$ and let a_i^* be the coefficient of x^i in f, $i = 0, \ldots, k$. According to Lemma 6.2 there exist coefficients $b_j^{(k)}$ such that

$$f(x) = \sum_{j=0}^{k} b_j^{(k)} p_j^{(k)}(x).$$

Using Lemma 6.2 we can derive the coefficients as follows:

$$\begin{aligned}
f(x) &= \sum_{i=0}^{d_n} a_i^* x^i \\
&= \sum_{i=0}^{d_n} a_i^* \sum_{j=i}^{k} \frac{\binom{j}{i}}{\binom{k}{i}} p_j^{(k)}(x) \\
&= \sum_{j=0}^{k} \sum_{i=0}^{j} a_i^* \frac{\binom{j}{i}}{\binom{k}{i}} p_j^{(k)}(x).
\end{aligned}$$

This motivates the following definition.

Definition 6.3 (Bernstein Coefficient) *The j-th Bernstein coefficient of f of order k is defined as*

$$b_j^{(k)} = \sum_{i=0}^{j} a_i^* \frac{\binom{j}{i}}{\binom{k}{i}}, \quad j = 0, \ldots, k. \ \Box \tag{6.2}$$

The following theorem shows the relationship between Bernstein coefficients and the range of polynomials over $[0, 1]$.

Theorem 6.4 (Range Overestimation)

$$f([0,1]) \subseteq [\{b_j^{(k)} \mid j = 0, \ldots, k\}]. \ \Box \tag{6.3}$$

Proof. Note that

$$p_j^{(k)}(x) \geq 0 \text{ for all } \; x \in [0,1], \quad j = 0, \ldots, k.$$

Further,

$$\begin{aligned} \sum_{j=0}^{k} p_j^{(k)}(x) &= \sum_{j=0}^{k} \binom{k}{j} x^j (1-x)^{k-j} \\ &= (x + (1-x))^k \\ &\equiv 1. \end{aligned}$$

Hence, for all $x \in X$

$$\begin{aligned} f(x) &= \sum_{j=0}^{k} b_j^{(k)} p_j^{(k)}(x) \\ &\leq \max_j b_j^{(k)} \sum_{j=0}^{k} p_j^{(k)}(x) \\ &= \max_j b_j^{(k)} \\ f(x) &= \sum_{j=0}^{k} b_j^{(k)} p_j^{(k)}(x) \\ &\geq \min_j b_j^{(k)} \sum_{j=0}^{k} p_j^{(k)}(x) \\ &= \min_j b_j^{(k)}. \; \Box \end{aligned}$$

The following theorem gives a necessary and sufficient condition when equality holds in (6.3).

Theorem 6.5 (Non–Overestimation)

$$\begin{aligned} \overline{f([0,1])} &= \max\{b_j^{(k)} \mid j = 0, \ldots, k\} && (6.4) \\ &\text{iff} \quad \max\{b_j^{(k)} \mid j = 0, \ldots, k\} \; \in \; \{b_0^{(k)}, b_k^{(k)}\} \\ \underline{f([0,1])} &= \min\{b_j^{(k)} \mid j = 0, \ldots, k\} && (6.5) \\ &\text{iff} \quad \min\{b_j^{(k)} \mid j = 0, \ldots, k\} \; \in \; \{b_0^{(k)}, b_k^{(k)}\}. \; \Box \end{aligned}$$

Proof. We give a proof of (6.4), the proof of (6.5) is analogous.

"$\Leftarrow$" Assume $\max\{b_j^{(k)} \mid j = 0, \ldots, k\} = \max\{b_0^{(k)}, b_k^{(k)}\}$. As

$$b_0^{(k)} = \sum_{i=0}^{0} a_i^* \frac{\binom{0}{i}}{\binom{k}{i}} = a_0^* = f(0)$$
$$b_k^{(k)} = \sum_{i=0}^{k} a_i^* \frac{\binom{k}{i}}{\binom{k}{i}} = \sum_{i=0}^{k} a_i^* = f(1),$$

it holds that $\max\{b_j^{(k)} \mid j = 0, \ldots, k\} \leq \overline{f([0,1])}$. Hence, by Theorem 6.4,

$$\max\{b_j^{(k)} \mid j = 0, \ldots, k\} = \overline{f([0,1])}.$$

"$\Rightarrow$" Assume $\overline{f([0,1])} = \max\{b_j^{(k)} \mid j = 0, \ldots, k\}$.

- If $b_0^{(k)} = b_1^{(k)} = \ldots = b_k^{(k)}$ then trivially $\max\{b_j^{(k)} \mid j = 0, \ldots, k\} = \max\{b_0^{(k)}, b_k^{(k)}\}$.
- Otherwise, if $0 < x < 1$ then $0 < p_j^{(k)}(x) < 1$ for all j and

$$\begin{aligned} f(x) &= \sum_{j=0}^{k} b_j^{(k)} p_j^{(k)}(x) \\ &< \max_j b_j^{(k)} \sum_{j=0}^{k} p_j^{(k)}(x) \\ &= \max_j b_j^{(k)}. \end{aligned}$$

Hence, f achieves its maximum in $[0,1]$ either for $x = 0$ or for $x = 1$ and

$$\max\{b_j^{(k)} \mid j = 0, \ldots, k\} = \overline{f([0,1])} = \max\{f(0), f(1)\} = \max\{b_0^{(k)}, b_k^{(k)}\}. \ \square$$

So far we restricted our considerations to the case $X = [0,1]$. In order to extend Theorem 6.4 and Theorem 6.5 to arbitrary intervals, we introduce Bernstein basis polynomials $p_j^{(k,X)}$ which depend on X. The relevant properties which were needed for the case $X = [0,1]$ are supposed to carry over to the generalized case. Hence, we expect

$$\begin{aligned} p_j^{(k,X)}(x) &\geq 0 \text{ for all } \ x \in X, \\ \sum_{j=0}^{k} p_j^{(k,X)}(x) &\equiv 1. \end{aligned}$$

In the sequel let $X \in \mathbb{IR}$, $\mathrm{w}(X) \neq 0$.

Definition 6.6 (Generalized Bernstein Polynomial) *The j-th Bernstein polynomial of order k over X is defined as*

$$p_j^{(k,X)}(x) = \binom{k}{j} \frac{(x - \underline{X})^j (\overline{X} - x)^{k-j}}{\mathrm{w}(X)^k}. \ \square$$

As in the case $X = [0,1]$, we show that $\{p_j^{(k,X)}(x) \mid j = 0, \ldots, k\}$ forms a basis of the vector space of polynomials of degree $\leq k$.

Lemma 6.7 *For every $k \geq i \geq 0$ it holds that*

$$\begin{aligned} (x - \underline{X})^i &= \mathrm{w}(X)^i \sum_{j=i}^{k} \frac{\binom{j}{i}}{\binom{k}{i}} p_j^{(k,X)}(x). \\ (x - \overline{X})^i &= -\mathrm{w}(X)^i \sum_{j=0}^{k-i} \frac{\binom{k-j}{i}}{\binom{k}{i}} p_j^{(k,X)}(x). \ \square \end{aligned}$$

Proof.

$$\begin{aligned} (x - \underline{X})^i &= \frac{1}{\mathrm{w}(X)^{k-i}} (x - \underline{X})^i (x - \underline{X} + \overline{X} - x)^{k-i} \\ &= \frac{1}{\mathrm{w}(X)^{k-i}} \sum_{r=0}^{k-i} \binom{k-i}{r} (x - \underline{X})^{i+r} (\overline{X} - x)^{k-i-r} \\ &= \frac{1}{\mathrm{w}(X)^{k-i}} \sum_{j=i}^{k} \binom{k-i}{j-i} (x - \underline{X})^j (\overline{X} - x)^{k-j} \\ &= \frac{1}{\mathrm{w}(X)^{k-i}} \sum_{j=i}^{k} \frac{\binom{j}{i}}{\binom{k}{i}} \binom{k}{j} (x - \underline{X})^j (\overline{X} - x)^{k-j} \end{aligned}$$

$$
\begin{aligned}
&= \mathrm{w}(X)^i \sum_{j=i}^{k} \frac{\binom{j}{i}}{\binom{k}{i}} p_j^{(k,X)}(x) \\
(x-\overline{X})^i &= \frac{(-1)^i}{\mathrm{w}(X)^{k-i}} (\overline{X}-x)^i (x-\underline{X}+\overline{X}-x)^{k-i} \\
&= \frac{(-1)^i}{\mathrm{w}(X)^{k-i}} \sum_{r=0}^{k-i} \binom{k-i}{r} (\overline{X}-x)^{i+r} (x-\underline{X})^{k-i-r} \\
&= \frac{(-1)^i}{\mathrm{w}(X)^{k-i}} \sum_{j=0}^{k-i} \binom{k-i}{k-j-i} (x-\underline{X})^j (\overline{X}-x)^{k-j} \\
&= \frac{(-1)^i}{\mathrm{w}(X)^{k-i}} \sum_{j=0}^{k-i} \frac{\binom{k-j}{i}}{\binom{k}{i}} \binom{k}{j} (x-\underline{X})^j (\overline{X}-x)^{k-j} \\
&= (-\mathrm{w}(X))^i \sum_{j=0}^{k-i} \frac{\binom{k-j}{i}}{\binom{k}{i}} p_j^{(k,X)}(x). \ \square
\end{aligned}
$$

Let us transform f from the power basis to the generalized Bernstein basis. For that purpose, it is useful to introduce Taylor coefficients first: In the sequel let

$$
\begin{aligned}
a_i^{(\underline{X})} &= f^{(i)}(\underline{X})/i! \\
a_i^{(\overline{X})} &= f^{(i)}(\overline{X})/i!
\end{aligned}
$$

be the i-th Taylor coefficient of f with developing point $\underline{X}$ respectively $\overline{X}$, i.e.

$$
\begin{aligned}
f(x) &= \sum_{i=0}^{d_n} a_i^{(\underline{X})} (x-\underline{X})^i \\
&= \sum_{i=0}^{d_n} a_i^{(\overline{X})} (x-\overline{X})^i .
\end{aligned}
$$

Thus,

$$
\begin{aligned}
f(x) &= \sum_{i=0}^{d_n} a_i^{(\underline{X})} (x-\underline{X})^i \\
&= \sum_{i=0}^{d_n} a_i^{(\underline{X})} \sum_{j=i}^{k} \frac{\binom{j}{i}}{\binom{k}{i}} \mathrm{w}(X)^i p_j^{(k,X)}(x) \\
&= \sum_{j=0}^{k} \sum_{i=0}^{j} \frac{\binom{j}{i}}{\binom{k}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^i p_j^{(k,X)}(x) \\
f(x) &= \sum_{i=0}^{d_n} a_i^{(\overline{X})} (x-\overline{X})^i \\
&= \sum_{i=0}^{d_n} a_i^{(\overline{X})} \sum_{j=0}^{k-i} \frac{\binom{k-j}{i}}{\binom{k}{i}} (-\mathrm{w}(X))^i p_j^{(k,X)}(x) \\
&= \sum_{j=0}^{k} \sum_{i=0}^{k-j} \frac{\binom{k-j}{i}}{\binom{k}{i}} a_i^{(\overline{X})} (-\mathrm{w}(X))^i p_j^{(k,X)}(x).
\end{aligned}
$$

This motivates the following definition.

Definition 6.8 (Generalized Bernstein Coefficient) *The j-th Bernstein coefficient of f of order k*

over X is defined as

$$
\begin{aligned}
b_j^{(k,X)} &= \sum_{i=0}^{j} \frac{\binom{j}{i}}{\binom{k}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^i &(6.6)\\
&= \sum_{i=0}^{k-j} \frac{\binom{k-j}{i}}{\binom{k}{i}} a_i^{(\overline{X})} (-\mathrm{w}(X))^i &(6.7)
\end{aligned}
$$

for $j = 0, \ldots, k$. □

Thus, Theorem 6.4 can be generalized as follows:

Theorem 6.9 (Range Overestimation)

$$f(X) \subseteq [\{b_j^{(k,X)} \mid j = 0, \ldots, k\}]. \; \square \tag{6.8}$$

Proof. Note that

$$p_j^{(k,X)}(x) \geq 0 \text{ for all } \; x \in X, \quad j = 0, \ldots, k.$$

Further,

$$
\begin{aligned}
\sum_{j=0}^{k} p_j^{(k,X)}(x) &= \frac{1}{\mathrm{w}(X)^k} \sum_{j=1}^{k} \binom{k}{j} (x - \underline{X})^j (\overline{X} - x)^{k-j}\\
&= \frac{1}{\mathrm{w}(X)^k} (x - \underline{X} + \overline{X} - x)^k\\
&= 1.
\end{aligned}
$$

Hence, for all $x \in X$

$$
\begin{aligned}
f(x) &= \sum_{j=0}^{k} b_j^{(k,X)} p_j^{(k,X)}(x)\\
&\leq \max_j b_j^{(k,X)} \sum_{j=0}^{k} p_j^{(k,X)}(x)\\
&= \max_j b_j^{(k,X)}\\
f(x) &= \sum_{j=0}^{k} b_j^{(k,X)} p_j^{(k,X)}(x)\\
&\geq \min_j b_j^{(k,X)} \sum_{j=0}^{k} p_j^{(k,X)}(x)\\
&= \min_j b_j^{(k,X)}. \; \square
\end{aligned}
$$

The generalization of Theorem 6.5 is as expected:

Theorem 6.10 (Non–Overestimation)

$$
\begin{aligned}
\overline{f(X)} &= \max\{b_j^{(k,X)} \mid j = 0, \ldots, k\}\\
\text{iff} \quad & \max\{b_j^{(k,X)} \mid j = 0, \ldots, k\} \;\in\; \{b_0^{(k,X)}, b_k^{(k,X)}\} &(6.9)\\
\underline{f(X)} &= \min\{b_j^{(k,X)} \mid j = 0, \ldots, k\}\\
\text{iff} \quad & \min\{b_j^{(k,X)} \mid j = 0, \ldots, k\} \;\in\; \{b_0^{(k,X)}, b_k^{(k,X)}\}. \; \square &(6.10)
\end{aligned}
$$

Proof. We give a proof of (6.9), the proof of (6.10) is analogous.

"$\Leftarrow$" Assume $\max\{b_j^{(k,X)} \mid j = 0, \ldots, k\} = \max\{b_0^{(k,X)}, b_k^{(k,X)}\}$. As

$$b_0^{(k,X)} = \sum_{i=0}^{0} \frac{\binom{0}{i}}{\binom{0}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^i = a_0^{(\underline{X})} = f(\underline{X}) \tag{6.11}$$

$$b_k^{(k,X)} = \sum_{i=0}^{k} \frac{\binom{k}{i}}{\binom{k}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^i = \sum_{i=0}^{k} a_i^{(\underline{X})} \mathrm{w}(X)^i = f(\overline{X}), \tag{6.12}$$

it holds that $\max\{b_j^{(k,X)} \mid j = 0, \ldots, k\} \le \overline{f(X)}$. Hence, by Theorem 6.9,

$$\max\{b_j^{(k,X)} \mid j = 0, \ldots, k\} = \overline{f(X)}.$$

"$\Rightarrow$" Assume $\overline{f(X)} = \max\{b_j^{(k,X)} \mid j = 0, \ldots, k\}$.

- If $b_0^{(k,X)} = b_1^{(k,X)} = \ldots = b_k^{(k,X)}$ then trivially $\max\{b_j^{(k,X)} \mid j = 0, \ldots, k\} = \max\{b_0^{(k,X)}, b_k^{(k,X)}\}$.
- Otherwise, if $\underline{X} < x < \overline{X}$ then $0 < p_j^{(k,X)}(x) < 1$ for all j and

$$\begin{aligned} f(x) &= \sum_{j=0}^{k} b_j^{(k,X)} p_j^{(k,X)}(x) \\ &< \max_j b_j^{(k,X)} \sum_{j=0}^{k} p_j^{(k,X)}(x) \\ &= \max_j b_j^{(k,X)}. \end{aligned}$$

  Hence, f achieves its maximum in X either for $x = \underline{X}$ or for $x = \overline{X}$ and

$$\max\{b_j^{(k,X)} \mid j = 0, \ldots, k\} = \overline{f(X)} = \max\{f(\underline{X}), f(\overline{X})\} = \max\{b_0^{(k,X)}, b_k^{(k,X)}\}. \ \square$$

The range of f is bounded by the smallest and the largest Bernstein coefficient. This defines an interval extension of f, which is called Bernstein form.

Definition 6.11 (Bernstein Form) *The k-th order Bernstein form $B_f^{(k)} : \mathbb{IR} \to \mathbb{IR}$ of f is defined as*

$$B_f^{(k)}(X) = [\{b_j^{(k,X)} \mid j = 0, \ldots, k\}]. \ \square$$

Theorem 6.12 (Interval Extension) *$B_f^{(k)}$ is an interval extension of f.* $\square$

Proof. From Theorem 6.9 it follows that $f(X) \subseteq B_f^{(k)}(X)$ for all $X \in \mathbb{IR}$. If $X \in \mathbb{R}$ then $b_j^{(k,X)} = f(\underline{X}) = f(\overline{X})$ for all j, hence $B_f^{(k)}(X) \in \mathbb{R}$. $\square$

For the parameter k it was required so far merely that it is greater or equal the degree of f. However, by incrementing k, usually better approximations of the range are obtained.

Theorem 6.13 (Overestimation Error Bounds)

$$\mathrm{q}(f(X), B_f^{(k)}(X)) \le \frac{1}{k} \sum_{i=2}^{d_n} (i-1)^2 |a_i^{(\underline{X})}| \mathrm{w}(X)^i. \tag{6.13}$$

$$\mathrm{q}(f(X), B_f^{(k)}(X)) \le \frac{1}{k} \sum_{i=2}^{d_n} (i-1)^2 |a_i^{(\overline{X})}| \mathrm{w}(X)^i. \ \square \tag{6.14}$$

The proof of Theorem 6.13 requires some preparation and will be given after Lemma 6.15. As the proofs of (6.13) and (6.14) are anlaogous, we give only the proof of (6.13).

Lemma 6.14 *Let* ${}^{i}c_j^{(k,X)} \in \mathbb{R}$, $i = 0, \ldots, k$, $j = 0, \ldots, k$ *such that*

$$\sum_{j=0}^{k} (j/k \ \mathrm{w}(X))^i p_j^{(k,X)} - (x - \underline{X})^i = \sum_{j=0}^{k} {}^{i}c_j^{(k,X)} p_j^{(k,X)}.$$

Then

$$0 \leq {}^{i}c_j^{(k,X)} \leq \frac{(i-1)^2}{k} \mathrm{w}(X)^i. \ \square$$

Proof. As $p_j^{(k,X)}$, $j = 0, \ldots, k$ span the space of polynomials of degree $\leq k$, the coefficients ${}^{i}c_j^{(k,X)}$ exist and are uniquely defined.

For $i = 0$ and $i = 1$ we obtain

$${}^{0}c_j^{(k,X)} = {}^{1}c_j^{(k,X)} = 0, \quad j = 0, \ldots, k.$$

For $i > 1$ it holds that

$$\begin{aligned}
&\sum_{j=0}^{k} (j/k \ \mathrm{w}(X))^i p_j^{(k,X)} - (x - \underline{X})^i \\
&= \sum_{j=0}^{k} (j/k \ \mathrm{w}(X))^i p_j^{(k,X)} - \sum_{j=i}^{k} \frac{\binom{j}{i}}{\binom{k}{i}} \mathrm{w}(X)^i p_j^{(k,X)} \\
&= \sum_{j=0}^{i-1} (j/k \ \mathrm{w}(X))^i p_j^{(k,X)} - \sum_{j=i}^{k} \left(j/k \ \mathrm{w}(X))^i - \frac{\binom{j}{i}}{\binom{k}{i}} \mathrm{w}(X)^i \right) p_j^{(k,X)} \\
&= \sum_{j=0}^{k} {}^{i}c_j^{(k,X)} p_j^{(k,X)}.
\end{aligned}$$

Comparing the coefficients of the Bernstein polynomials we get for $j < i$

$$0 \leq {}^{i}c_j^{(k,X)} = (j/k \ \mathrm{w}(X))^i \leq \frac{(i-1)^2}{k} \mathrm{w}(X)^i$$

and for $j \geq i$

$${}^{i}c_j^{(k,X)} = \left((j/k)^i - \frac{\binom{j}{i}}{\binom{k}{i}} \right) \mathrm{w}(X)^i.$$

It remains to show for $j \geq i > 1$

$$0 \leq (j/k)^i - \frac{\binom{j}{i}}{\binom{k}{i}} \leq \frac{(i-1)^2}{k}.$$

$$\begin{aligned}
(j/k)^i - \frac{\binom{j}{i}}{\binom{k}{i}} &= (j/k)^i - \frac{j!(k-i)!}{k!(j-i)!} \\
&= (j/k)^i - \frac{j(j-1)\cdots(j-i+1)}{k(k-1)\cdots(k-i+1)} \\
&= (j/k)^i \left(1 - \frac{\left(1 - \frac{1}{j}\right) \cdots \left(1 - \frac{i-1}{j}\right)}{\left(1 - \frac{1}{k}\right) \cdots \left(1 - \frac{i-1}{k}\right)} \right) \qquad (6.15) \\
&\leq (j/k)^i \left(1 - \left(1 - \frac{1}{j}\right) \cdots \left(1 - \frac{i-1}{j}\right) \right) \\
&\leq (j/k)^i \left(1 - \left(1 - \frac{i-1}{j}\right)^{i-1} \right).
\end{aligned}$$

Applying the mean value Theorem to $(1-x)^{i-1}$, we obtain

$$1-\left(1-\frac{i-1}{j}\right)^{i-1} \leq \frac{(i-1)^2}{j},$$

hence

$$(j/k)^i \frac{(i-1)^2}{j} = (j/k)^{i-1}\frac{(i-1)^2}{k} \leq \frac{(i-1)^2}{k}.$$

Finally,

$$0 \leq (j/k)^i - \frac{\binom{j}{i}}{\binom{k}{j}}$$

follows from comparison of the factors in the numerator and the denomiator of (6.15). □

The following lemma gives a bound for the difference between a Bernstein coefficient and $f(x)$ for certain $x \in X$.

Lemma 6.15 *Let*

$$c_j^{(k,X)} = f(j/k\ \mathrm{w}(X) + \underline{X}) - b_j^{(k,X)} \quad j = 0,\ldots,k.$$

Then

$$|c_j^{(k,X)}| \leq \frac{1}{k}\sum_{i=2}^{d_n}(i-1)^2 a_i^{(\underline{X})}\mathrm{w}(X)^i. \;\square$$

Proof. With the definition of $c_j^{(k,X)}$ as in Lemma 6.15 it holds that

$$\begin{aligned}
&\sum_{j=0}^{k} f(j/k\ \mathrm{w}(X)+\underline{X})p_j^{(k,X)}(x) - f(x)\\
&= \sum_{j=0}^{k}(c_j^{(k,X)} + b_j^{(k,X)})p_j^{(k,X)}(x) - \sum_{j=0}^{k} b_j^{(k,X)}p_j^{(k,X)}\\
&= \sum_{j=0}^{k} c_j^{(k,X)}p_j^{(k,X)}(x). \qquad (6.16)
\end{aligned}$$

Using the coefficient ${}^{i}c_j^{(k,X)}$ of Lemma 6.14 we obtain

$$\begin{aligned}
&\sum_{j=0}^{k} f(j/k\ \mathrm{w}(X)+\underline{X})p_j^{(k,X)}(x) - f(x)\\
&= \sum_{j=0}^{k}\sum_{i=0}^{d_n} a_i^*(j/k\ \mathrm{w}(X)+\underline{X})^i p_j^{(k,X)}(x) - \sum_{i=0}^{d_n} a_i^* x^i\\
&= \sum_{j=0}^{k}\sum_{i=0}^{d_n} a_i^{(\underline{X})}(j/k\ \mathrm{w}(X))^i p_j^{(k,X)}(x) - \sum_{i=0}^{d_n} a_i^{(\underline{X})}(x-\underline{X})^i\\
&= \sum_{i=0}^{d_n} a_i^{(\underline{X})}\sum_{j=0}^{k}(j/k\ \mathrm{w}(X))^i p_j^{(k,X)}(x) - (x-\underline{X})^i\\
&= \sum_{i=0}^{d_n} a_i^{(\underline{X})}\sum_{j=0}^{k} {}^{i}c_j^{(k,X)}p_j^{(k,X)}(x)\\
&= \sum_{j=0}^{k}\sum_{i=0}^{d_n} a_i^{(\underline{X})}\,{}^{i}c_j^{(k,X)}p_j^{(k,X)}(x). \qquad (6.17)
\end{aligned}$$

Comparing the coefficients of $p_j^{(k,X)}$ in (6.16) and (6.17) we obtain

$$c_j^{(k,X)} = \sum_{i=0}^{d_n} a_i^{(\underline{X})}\, {}^i c_j^{(k,X)}.$$

From Lemma 6.14 and ${}^0c_j^{(k,X)} = {}^1c_j^{(k,X)} = 0$ it follows that

$$\begin{aligned}
|c_j^{(k,X)}| &\leq \sum_{i=0}^{d_n} |a_i^{(\underline{X})}|\, {}^i c_j^{(k,X)} \\
&\leq \frac{1}{k}\sum_{i=2}^{d_n} |a_i^{(\underline{X})}|(i-1)^2 \mathrm{w}(X)^i \\
&= \frac{1}{k}\sum_{i=2}^{d_n} |a_i^{(\underline{X})}|(i-1)^2 \mathrm{w}(X)^i. \ \Box
\end{aligned}$$

Proof of Theorem 6.13. Let $b_{j_{\min}}^{(k,X)}$, $b_{j_{\max}}^{(k,X)}$ be the smallest respecively largest Bernstein coefficient of f. From Theorem 6.9 and Lemma 6.15 it follows that

$$\begin{aligned}
\overline{B_f^{(k)}(X)} - \overline{f(X)} &\leq |f(j_{\max}/k\ \mathrm{w}(X) + \underline{X}) - b_{j_{\max}}^{(k,X)}| \\
&\leq \frac{1}{k}\sum_{i=2}^{d_n} (i-1)^2 |a_i^{(\underline{X})}| \mathrm{w}(X)^i \\
\underline{f(X)} - \underline{B_f^{(k)}(X)} &\leq |f(j_{\min}/k\ \mathrm{w}(X) + \underline{X}) - b_{j_{\min}}^{(k,X)}| \\
&\leq \frac{1}{k}\sum_{i=2}^{d_n} (i-1)^2 |a_i^{(\underline{X})}| \mathrm{w}(X)^i. \ \Box
\end{aligned}$$

Corollary 6.16 (Convergence) $B_f^{(k)}$ *is quadratically convergent.* $\Box$

Proof. Let $A \in \mathbb{IR}$ arbitrary but fixed. Let

$$\hat{a}_i^{(A)} = \max\{|a_i^{(\underline{X})}| \mid X \in \mathbb{I}A\}.$$

The maximum exists because A is bounded and all derivatives of f are continuous. From Theorem 6.13 it follows that

$$\begin{aligned}
\mathrm{q}(f(X), B_f^{(k)}(X)) &\leq \left(\frac{1}{k}\sum_{i=2}^{d_n} (i-1)^2 \hat{a}_i^{(A)} \mathrm{w}(X)^{i-2}\right) \mathrm{w}(X)^2 \\
&\leq \left(\frac{1}{k}\sum_{i=2}^{d_n} (i-1)^2 \hat{a}_i^{(A)} \mathrm{w}(A)^{i-2}\right) \mathrm{w}(X)^2 \\
&= \ell_{B_f^{(k)},A} \mathrm{w}(X)^2,
\end{aligned}$$

where

$$\ell_{B_f^{(k)},A} = \frac{1}{k}\sum_{i=2}^{d_n} (i-1)^2 \hat{a}_i^{(A)} \mathrm{w}(A)^{i-2}. \ \Box$$

The following theorem is taken from [Hong and Stahl, 1994a].

Theorem 6.17 (Inclusion Monotonicity) $B_f^{(k)}$ *is inclusion monotone.* $\Box$

Before giving a formal proof of Theorem 6.17 we explain the key idea at an example. Consider the polynomial

$$f(x) = x^5 + 0.2x^4 + 0.5x^3 + 0.7x^2 - x,$$

and let $X, Y \in \mathbb{IR}$ such that $X \subseteq Y$. For simplicity, we first consider the special case where Y has one endpoint in common with X. Actually, this is already sufficient for the general case because the general case can be viewed as applying the special case twice on each endpoint. Precisely, consider the interval $[\underline{Y}, \overline{X}]$. This interval shares an endpoint with X and Y. Thus, if the special case is true, we have

$$B_f(X) \subseteq B_f([\underline{Y}, \overline{X}]) \subseteq B_f(Y).$$

In Figure 6.1 we show the Bernstein coefficients of order $k = 5$ of f over all X, where $\underline{X} = -1$ and $\overline{X}$ varies between -1 and 1. The curve for $b_j^{(5,X)}$ is labeled by j. By careful inspection of this picture one makes the following observations:

- The enveloping curves (bold lines) are monotone, i.e. $\max_j b_j^{(5,X)}$ increases and $\min_j b_j^{(5,X)}$ decreases as $\overline{X}$ grows. This precisely means inclusion monotonicity.
- Whenever the curve for some $b_j^{(5,X)}$ has a local extremum, it is intersected by the curve for $b_{j-1}^{(5,X)}$. Further, the curve for $b_0^{(5,X)}$ is constant. This observation is the key for the proof of inclusion monotonicity because, as one sees, it implies immediately monotonicity of the enveloping curves.

Next, we fix the right endpoint $\overline{X}$ to 1 and observe the Bernstein coefficients as $\underline{X}$ varies between -1 and 1 in Figure 6.2. Again, the enveloping curves are monotone, but this time, if $b_j^{(5,X)}$ has a local extremum, it is intersected by $b_{j+1}^{(5,X)}$ and $b_5^{(5,X)}$ is constant.

Lemma 6.18 makes the experiemental observations more precise. From now on we view $b_j^{(k,X)}$ as a function in the variables $\underline{X}$, $\overline{X}$.

Lemma 6.18

$$b_j^{(k,X)} - b_{j-1}^{(k,X)} = \frac{\mathrm{w}(X)}{j} \frac{\partial b_j^{(k,X)}}{\partial \overline{X}}, \quad j = 1, \ldots, k \tag{6.18}$$

$$b_{j+1}^{(k,X)} - b_j^{(k,X)} = \frac{\mathrm{w}(X)}{k-j} \frac{\partial b_j^{(k,X)}}{\partial \underline{X}}, \quad j = 0, \ldots, k-1. \ \Box \tag{6.19}$$

Proof.

- Proof of (6.18). Let $1 \leq j \leq k$ and note that

$$\begin{aligned} \binom{j}{i} - \binom{j-1}{i} &= \binom{j-1}{i-1} \\ i\binom{j}{i} &= j\binom{j-1}{i-1}. \end{aligned}$$

$$\begin{aligned} b_j^{(k,X)} - b_{j-1}^{(k,X)} &\overset{(6.6)}{=} \sum_{i=0}^{j} \frac{\binom{j}{i}}{\binom{k}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^i - \sum_{i=0}^{j-1} \frac{\binom{j-1}{i}}{\binom{k}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^i \\ &= \sum_{i=0}^{j-1} \frac{\binom{j}{i} - \binom{j-1}{i}}{\binom{k}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^i + \frac{1}{\binom{k}{j}} a_j^{(\underline{X})} \mathrm{w}(X)^j \\ &= \sum_{i=1}^{j-1} \frac{\binom{j-1}{i-1}}{\binom{k}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^i + \frac{1}{\binom{k}{j}} a_j^{(\underline{X})} \mathrm{w}(X)^j \\ &= \sum_{i=1}^{j} \frac{\binom{j-1}{i-1}}{\binom{k}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^i. \end{aligned} \tag{6.20}$$

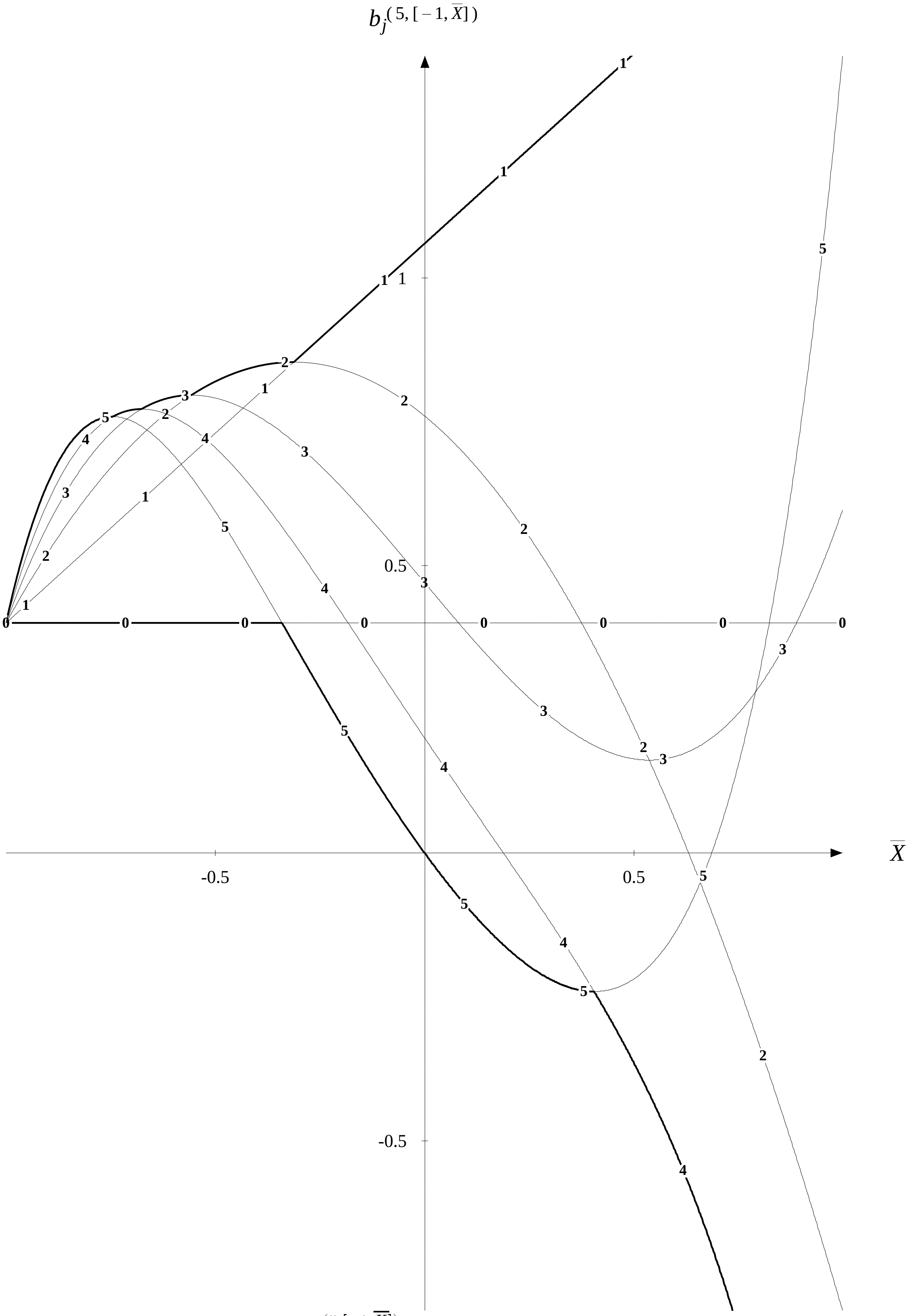

Figure 6.1: Bernstein coefficients $b_j^{(5,[-1,\overline{X}])}$ of $f(x) = x^5 + 0.2x^4 + 0.5x^3 + 0.7x^2 - x$ in dependence of $\overline{X}$.

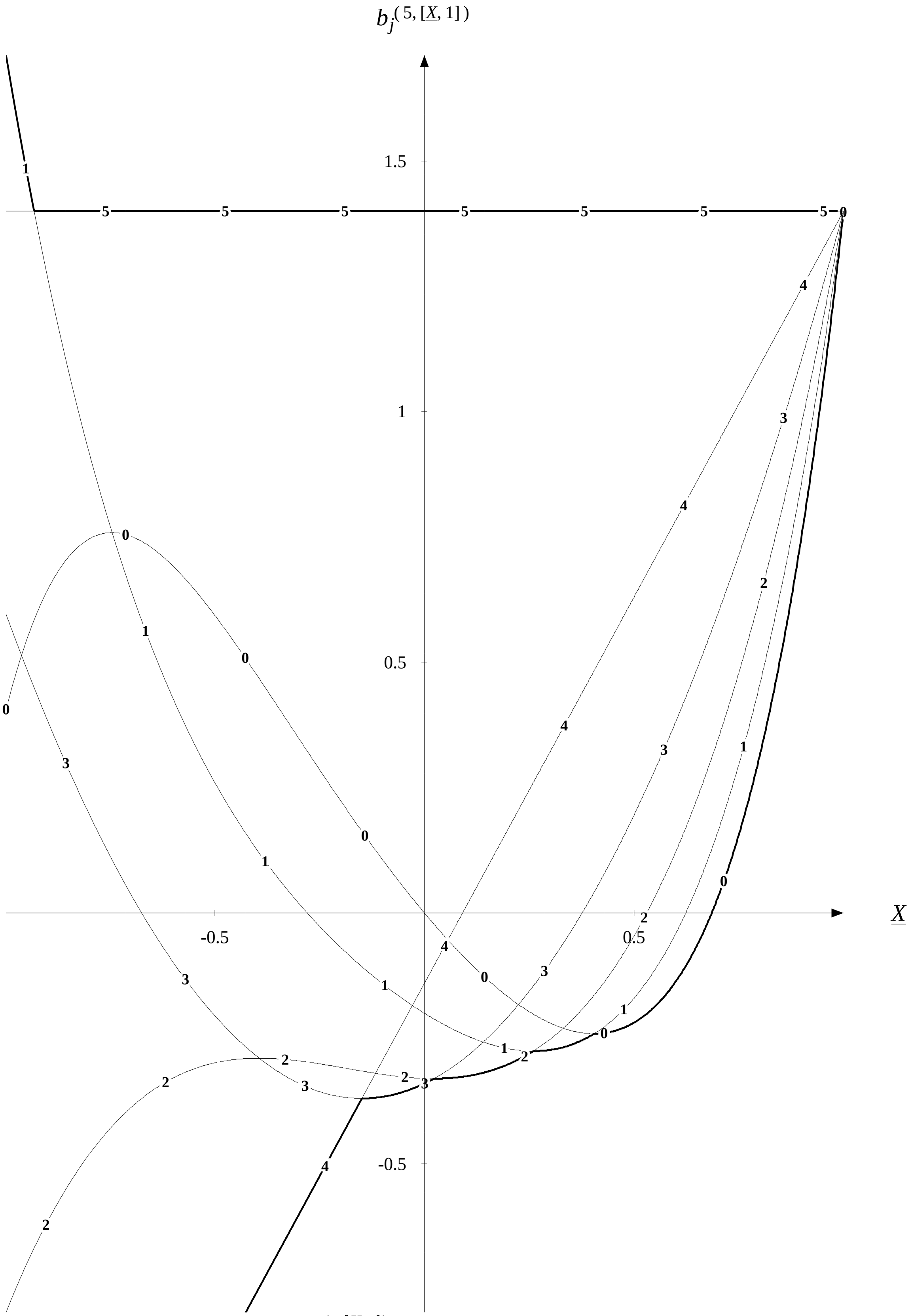

Figure 6.2: Bernstein coefficients $b_j^{(5,[\underline{X},1])}$, of $f(x) = x^5 + 0.2x^4 + 0.5x^3 + 0.7x^2 - x$ in dependence of $\underline{X}$.

$$
\begin{aligned}
\frac{\partial b_j^{(k,X)}}{\partial \overline{X}} &\overset{(6.6)}{=} \frac{\partial}{\partial \overline{X}} \sum_{i=0}^{j} \frac{\binom{j}{i}}{\binom{k}{i}} a_i^{(\underline{X})} (\overline{X} - \underline{X})^i \\
&= \sum_{i=1}^{j} \frac{\binom{j}{i}}{\binom{k}{i}} a_i^{(\underline{X})} i \mathrm{w}(X)^{i-1} \\
&= j \sum_{i=1}^{j} \frac{\binom{j-1}{i-1}}{\binom{k}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^{i-1}.
\end{aligned}
$$

- Proof of (6.19). Let $0 \le j \le k-1$ and note that

$$
\begin{aligned}
\binom{k-j-1}{i} - \binom{k-j}{i} &= -\binom{k-j-1}{i-1} \\
i\binom{k-j}{i} &= (k-j)\binom{k-j-i}{i-1}.
\end{aligned}
$$

$$
\begin{aligned}
b_{j+1}^{(k,X)} - b_j^{(k,X)} &\overset{(6.7)}{=} \sum_{i=0}^{k-j-1} \frac{\binom{k-j-1}{i}}{\binom{k}{i}} a_i^{(\overline{X})} (-\mathrm{w}(X))^i - \sum_{i=0}^{k-j} \frac{\binom{k-j}{i}}{\binom{k}{i}} a_i^{(\overline{X})} (-\mathrm{w}(X))^i \\
&= -\sum_{i=1}^{k-j-1} \frac{\binom{k-j-1}{i} - \binom{k-j}{i}}{\binom{k}{i}} a_i^{(\overline{X})} (-\mathrm{w}(X))^i + \frac{1}{\binom{k}{k-j}} a_j^{(\overline{X})} (-\mathrm{w}(X))^j \\
&= -\sum_{i=1}^{k-j} \frac{\binom{k-j-1}{i-1}}{\binom{k}{i}} a_i^{(\overline{X})} (-\mathrm{w}(X))^i. \qquad (6.21) \\
\frac{\partial b_j^{(k,X)}}{\partial \underline{X}} &\overset{(6.7)}{=} \frac{\partial}{\partial \underline{X}} \sum_{i=0}^{k-j} \frac{\binom{k-j}{i}}{\binom{k}{i}} a_i^{(\overline{X})} (\underline{X} - \overline{X})^i \\
&= \sum_{i=1}^{k-j} \frac{\binom{k-j}{i}}{\binom{k}{i}} a_i^{(\overline{X})} i (-\mathrm{w}(X))^{i-1} \\
&= (k-j) \sum_{i=1}^{k-j} \frac{\binom{k-j-1}{i-1}}{\binom{k}{i}} a_i^{(\overline{X})} (-\mathrm{w}(X))^{i-1}. \ \square
\end{aligned}
$$

Finally, we come to the proof of Theorem 6.17.

Proof of Theorem 6.17 According to the remarks above, we have to show the following:

$$
\begin{aligned}
&\min_j b_j^{(k,X)} && \text{is monotonically decreasing in } \overline{X} && (6.22) \\
&\max_j b_j^{(k,X)} && \text{is monotonically increasing in } \overline{X} && (6.23) \\
&\min_j b_j^{(k,X)} && \text{is monotonically increasing in } \underline{X} && (6.24) \\
&\max_j b_j^{(k,X)} && \text{is monotonically decreasing in } \underline{X}. && (6.25)
\end{aligned}
$$

Note that for all $X \in \mathbb{IR}$

$$
\begin{aligned}
b_0^{(k,X)} &= f(\underline{X}) \\
b_k^{(k,X)} &= f(\overline{X}).
\end{aligned}
$$

The proofs of (6.22) – (6.25) are similar, but for the sake of completeness we give them all in detail. Let $X \in \mathbb{IR}$ arbitrary but fixed.

- Proof of (6.22). Let $j_{\min}$ such that

$$
b_{j_{\min}}^{(k,X)} \le b_j^{(k,X)} \text{ for all } j. \qquad (6.26)
$$

We have to show that

$$\frac{\partial b_{j_{\min}}^{(k,X)}}{\partial \overline{X}} \leq 0.$$

As $b_0^{(k,X)}$ does not depend on $\overline{X}$, this is obviously true if $j_{\min} = 0$. Hence, assume $j_{\min} \neq 0$ and

$$\frac{\partial b_{j_{\min}}^{(k,X)}}{\partial \overline{X}} > 0.$$

From 6.18 we obtain

$$b_{j_{\min}}^{(k,X)} - b_{j_{\min}-1}^{(k,X)} = \frac{(\overline{X} - \underline{X})}{j_{\min}} \frac{\partial b_{j_{\min}}^{(k,X)}}{\partial \overline{X}} > 0,$$

which contradicts (6.26).

- Proof of (6.23). Let $j_{\max}$ such that

$$b_{j_{\max}}^{(k,X)} \geq b_j^{(k,X)} \text{ for all } \ j. \tag{6.27}$$

We have to show that

$$\frac{\partial b_{j_{\max}}^{(k,X)}}{\partial \overline{X}} \geq 0.$$

As $b_0^{(k,X)}$ does not depend on $\overline{X}$, this is obviously true if $j_{\max} = 0$. Hence, assume $j_{\max} \neq 0$ and

$$\frac{\partial b_{j_{\max}}^{(k,X)}}{\partial \overline{X}} < 0.$$

From 6.18 we obtain

$$b_{j_{\max}}^{(k,X)} - b_{j_{\max}-1}^{(k,X)} = \frac{(\overline{X} - \underline{X})}{j_{\max}} \frac{\partial b_{j_{\max}}^{(k,X)}}{\partial \overline{X}} < 0,$$

which contradicts (6.27).

- Proof of (6.24). Let $j_{\min}$ such that

$$b_{j_{\min}}^{(k,X)} \leq b_j^{(k,X)} \text{ for all } \ j. \tag{6.28}$$

We have to show that

$$\frac{\partial b_{j_{\min}}^{(k,X)}}{\partial \underline{X}} \geq 0.$$

As $b_k^{(k,X)}$ does not depend on $\underline{X}$, this is obviously true if $j_{\min} = k$. Hence, assume $j_{\min} \neq k$ and

$$\frac{\partial b_{j_{\min}}^{(k,X)}}{\partial \underline{X}} < 0.$$

From 6.19 we obtain

$$b_{j_{\min}+1}^{(k,X)} - b_{j_{\min}}^{(k,X)} = \frac{(\overline{X} - \underline{X})}{k - j_{\min}} \frac{\partial b_{j_{\min}}^{(k,X)}}{\partial \underline{X}} < 0,$$

which contradicts (6.28).

- Proof of (6.25). Let $j_{\max}$ such that

$$b_{j_{\max}}^{(k,X)} \geq b_j^{(k,X)} \text{ for all } \ j. \tag{6.29}$$

We have to show that

$$\frac{\partial b_{j_{\max}}^{(k,X)}}{\partial \underline{X}} \leq 0.$$

As $b_k^{(k,X)}$ does not depend on $\underline{X}$, this is obviously true if $j_{\max} = k$. Hence, assume $j_{\max} \neq k$ and

$$\frac{\partial b_{j_{\max}}^{(k,X)}}{\partial \underline{X}} > 0.$$

From 6.19 we obtain

$$b_{j_{\max}+1}^{(k,X)} - b_{j_{\max}}^{(k,X)} = \frac{(\overline{X} - \underline{X})}{k - j_{\max}} \frac{\partial b_{j_{\max}}^{(k,X)}}{\partial \underline{X}} > 0,$$

which contradicts (6.29). □

Figure 6.1 reveals another property of the Bernstein coefficients. Let $\xi \in X$. If $\overline{X} \approx -1$, then

$$b_0^{(k,X)} \leq b_j^{(k,X)} \leq b_k^{(k,X)} \text{ for all } j.$$

Similarly, Figure 6.2 shows that if $\underline{X} \approx 1$, then

$$b_0^{(k,X)} \leq b_j^{(k,X)} \leq b_k^{(k,X)} \text{ for all } j.$$

According to Theorem 6.10, this gives rise to the conjecture

$$B_f^{(k)}(X) = f(X) \text{ if } \mathrm{w}(X) \text{ is small enough.}$$

Yet, this conjecture holds only if we make the notion "small enough" more precise. This is done in the following Theorem, which is original.

Theorem 6.19

- *For all $\underline{X} \in \mathbb{R}$ there exists $\varepsilon > 0$ such that for all $\overline{X} \in [\underline{X}, \underline{X} + \varepsilon]$ it holds that*

$$B_f^{(k)}(X) = f(X).$$

- *For all $\overline{X} \in \mathbb{R}$ there exists $\varepsilon > 0$ such that for all $\underline{X} \in [\overline{X} - \varepsilon, \overline{X}]$ it holds that*

$$B_f^{(k)}(X) = f(X). \;\square$$

Proof.

- Let $\underline{X} \in \mathbb{R}$ arbitrary but fixed. According to Theorem 6.10 it suffices to show that there exists $\varepsilon > 0$ such that for all $\overline{X} \in [\underline{X}, \underline{X} + \varepsilon]$

$$b_j^{(k,X)} \geq b_{j-1}^{(k,X)} \text{ for all } j = 1, \ldots, k \quad \text{or} \tag{6.30}$$

$$b_j^{(k,X)} \leq b_{j-1}^{(k,X)} \text{ for all } j = 1, \ldots, k. \tag{6.31}$$

From (6.20) it follows that for all $\overline{X} \geq \underline{X}$

$$b_j^{(k,X)} - b_{j-1}^{(k,X)} = \sum_{i=1}^{j} \frac{\binom{j-1}{i-1}}{\binom{k}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^i.$$

Let $1 \leq l \leq k$ be the smallest index such that $a_l^{(\underline{X})} \neq 0$. If no such l exists, then let $l = k+1$. Thus,

$$\begin{aligned} b_j^{(k,X)} - b_{j-1}^{(k,X)} &= \sum_{i=l}^{j} \frac{\binom{j-1}{i-1}}{\binom{k}{i}} a_i^{(\underline{X})} \mathrm{w}(X)^i \\ &= \mathrm{w}(X)^l \sum_{i=0}^{j-l} \frac{\binom{j-1}{i+l-1}}{\binom{k}{i+l}} a_{i+l}^{(\underline{X})} \mathrm{w}(X)^i. \end{aligned}$$

Depending on the sign of $a_l^{(\underline{X})}$ we show that either (6.30) or (6.31) holds.

– Assume $a_l^{(\underline{X})} > 0$. Then for all $j = 1, \ldots, k$ there exists $\varepsilon_j > 0$ such that for all $\mathrm{w}(X) \leq \varepsilon_j$

$$\sum_{i=0}^{j-l} \frac{\binom{j-1}{i+l-1}}{\binom{k}{i+l}} a_{i+l}^{(\underline{X})} \mathrm{w}(X)^i \geq 0.$$

Let $\varepsilon = \min_j \varepsilon_j$. Then for all $\overline{X} \in [\underline{X}, \underline{X} + \varepsilon]$

$$b_j^{(k,X)} - b_{j-1}^{(k,X)} \geq 0.$$

– Assume $a_l^{(\underline{X})} < 0$. Then for all $j = 1, \ldots, k$ there exists $\varepsilon_j > 0$ such that for all $\mathrm{w}(X) \leq \varepsilon_j$

$$\sum_{i=0}^{j-l} \frac{\binom{j-1}{i+l-1}}{\binom{k}{i+l}} a_{i+l}^{(\underline{X})} \mathrm{w}(X)^i \leq 0.$$

Let $\varepsilon = \min_j \varepsilon_j$. Then for all $\overline{X} \in [\underline{X}, \underline{X} + \varepsilon]$

$$b_j^{(k,X)} - b_{j-1}^{(k,X)} \leq 0.$$

- Let $\overline{X} \in \mathbb{R}$ arbitrary but fixed. According to Theorem 6.10 it suffices to show that there exists $\varepsilon > 0$ such that for all $\underline{X} \in [\overline{X} - \varepsilon, \overline{X}]$

$$\begin{aligned} b_{j+1}^{(k,X)} &\geq b_j^{(k,X)} \text{ for all } j = 0, \ldots, k-1 \quad \text{or} && (6.32)\\ b_{j+1}^{(k,X)} &\leq b_j^{(k,X)} \text{ for all } j = 0, \ldots, k-1. && (6.33) \end{aligned}$$

From (6.21) it follows that for all $\underline{X} \leq \overline{X}$

$$b_{j+1}^{(k,X)} - b_j^{(k,X)} = -\sum_{i=1}^{k-j} \frac{\binom{k-j-1}{i-1}}{\binom{k}{i}} a_i^{(\overline{X})} (-\mathrm{w}(X))^i.$$

Let $0 \leq l \leq k-1$ be the smallest index such that $a_l^{(\overline{X})} \neq 0$. If no such l exists, then let $l = k+1$. Thus,

$$\begin{aligned} b_{j+1}^{(k,X)} - b_j^{(k,X)} &= -\sum_{i=l}^{k-j} \frac{\binom{k-j-1}{i-1}}{\binom{k}{i}} a_i^{(\overline{X})} (-\mathrm{w}(X))^i \\ &= -\mathrm{w}(X)^l \sum_{i=0}^{k-j-l} \frac{\binom{k-j-1}{i+l-1}}{\binom{k}{i+l}} (-1)^l a_{i+l}^{(\overline{X})} (-\mathrm{w}(X))^i. \end{aligned}$$

Depending on the sign of $(-1)^l a_l^{(\overline{X})}$ we show that either (6.32) or (6.33) holds.

– Assume $(-1)^l a_l^{(\overline{X})} > 0$. Then for all $j = 0, \ldots, k-1$ there exists $\varepsilon_j > 0$ such that for all $\mathrm{w}(X) \leq \varepsilon_j$

$$\sum_{i=0}^{k-j-l} \frac{\binom{k-j-1}{i+l-1}}{\binom{k}{i+l}} (-1)^l a_{i+l}^{(\overline{X})} (-\mathrm{w}(X))^i \geq 0.$$

Let $\varepsilon = \min_j \varepsilon_j$. Then for all $\underline{X} \in [\overline{X} - \varepsilon, \overline{X}$

$$b_{j+1}^{(k,X)} - b_j^{(k,X)} \leq 0.$$

– Assume $(-1)^l a_l^{(\overline{X})} < 0$. Then for all $j = 0, \ldots, k-1$ there exists $\varepsilon_j > 0$ such that for all $\mathrm{w}(X) \leq \varepsilon_j$

$$\sum_{i=0}^{k-j-l} \frac{\binom{k-j-1}{i+l-1}}{\binom{k}{i+l}} (-1)^l a_{i+l}^{(\overline{X})} (-\mathrm{w}(X))^i \leq 0.$$

Let $\varepsilon = \min_j \varepsilon_j$. Then for all $\underline{X} \in [\overline{X} - \varepsilon, \overline{X}]$

$$b_{j+1}^{(k,X)} - b_j^{(k,X)} \geq 0.$$

Remark. Note that it is *not* true that for all $c \in \mathbb{R}$ there exists $\varepsilon > 0$ such that for all $0 \le r \le \varepsilon$

$$B_f^{(k)}([c-r, c+r]) = f([c-r, c+r]).$$

As a counter example consider $f(x) = x^2$ and $c = 0$. The Bernstein coefficients in dependence of r are

$$\begin{aligned} b_0^{(2,[-r,r])} &= r^2 \\ b_1^{(2,[-r,r])} &= -r^2 \\ b_2^{(2,[-r,r])} &= r^2, \end{aligned}$$

Hence, $b_1^{(2,[-r,r])}$ is the smallest Bernstein coefficient, and according to Theorem 6.10,

$$B_f^{(2)}([-r,r]) \ne f([-r,r])$$

for all $r \in \mathbb{R}$.

Remark. If $|f'(\underline{X})|$ is sufficiently large relative to $|f^{(i)}(\underline{X})|$ for all i and $\mathrm{w}(X)$ is sufficiently small, then the sign of $b_j^{(k,X)} - b_{j-1}^{(k,X)}$ is the same for all j. Similarly, if $|f'(\overline{X})|$ is sufficiently large relative to $|f^{(i)}(\overline{X})|$ for all i and $\mathrm{w}(X)$ is sufficiently small, then the sign of $b_{j+1}^{(k,X)} - b_j^{(k,X)}$ is the same for all j. In both cases we obtain $B_f^{(k)}(X) = f(X)$. This qualitative result is illustrated by figure 6.3, where the overestimation error of the Bernstein form is computed for $f(x) = (x-1)(x-2)(x-3)(x-4)$, and various input intervals X, $\mathrm{w}(X) \equiv 0.5$ and $k = 4$, 8 and 16. One sees, that the Bernstein form is exact, if X is not close to an extremum of f. However, it is *not* true, that monotonicity of f in X is a sufficient condition for non–overestimation, although this is claimed sometimes in the literature. For example let $f(x) = -x^3 + 0.7x^2 - 0.2x + 0.9$, $k = 3$ and $X = [0,1]$. Note that f is monotone in X because $f'(x) < 0$ for all $x \in \mathbb{R}$. The Bernstein coefficients are

$$\begin{aligned} b_0^{(k,X)} &= 0.9, \\ b_1^{(k,X)} &= 0.8333, \\ b_2^{(k,X)} &= 1.0, \\ b_3^{(k,X)} &= 0.4. \end{aligned}$$

Thus, we obtain $B_f^{(k)}(X) = [0.4, 1.0]$, but $f(X) = [0.4, 0.9]$.

Algorithm and Complexity.

Algorithm 6.2 for evaluating the Bernstein form proceeds in 3 steps: Compute the Taylor coefficients $a_i^{(\underline{X})}$, compute the Bernstein coefficients $b_j^{(k,X)}$ according to (6.6), find the largest and the smallest Bernstein coefficient. Equivalently, one could compute the Taylor coefficients $a_i^{(\overline{X})}$ and the Bernstein coefficients according to (6.7). The Taylor coefficients are obtained by the extended Horner scheme. An efficient method for computing the Bernstein coefficients is given by the following lemma [Rokne, 1979a].

Lemma 6.20 *Let* ${}^{l}b_j^{(k,X)}$, $l = 0, \ldots, d_n$, $j = 0, \ldots, k-l$ *be defined recursively as*

$$\begin{aligned} {}^{l}b_0^{(k,X)} &= \frac{\mathrm{w}(X)^l}{\binom{k}{l}} a_l^{(\underline{X})} && \text{for } l = 0, \ldots, d_n && (6.34) \\ {}^{d_n}b_j^{(k,X)} &= {}^{d_n}b_0^{(k,X)} && \text{for } j = 1, \ldots, k - d_n && (6.35) \\ {}^{l}b_j^{(k,X)} &= {}^{l}b_{j-1}^{(k,X)} + {}^{l+1}b_{j-1}^{(k,X)} && \text{for } j = 1, \ldots, k, \quad l = 0, \ldots, \min\{d_n - 1, k - j\}. && (6.36) \end{aligned}$$

Then

$$b_j^{(k,X)} = {}^{0}b_j^{(k,X)} \text{ for } j = 0, \ldots, k. \ \square$$

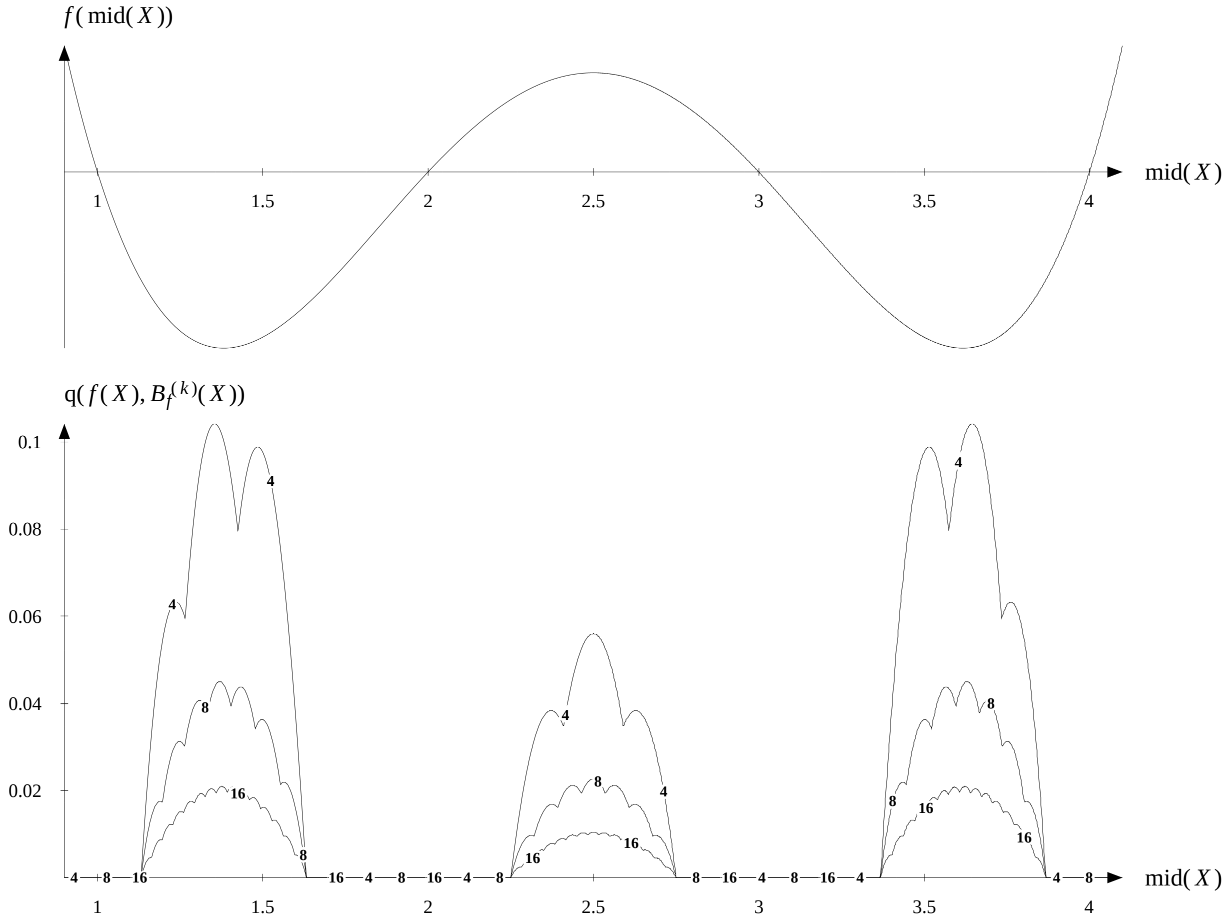

Figure 6.3: Overestimation error of Bernstein form for $f(x) = (x-1)(x-2)(x-3)(x-4)$, in dependence of $\mathrm{mid}(X)$, where $\mathrm{w}(X) \equiv 0.5$ and $k = 4, 8, 16$

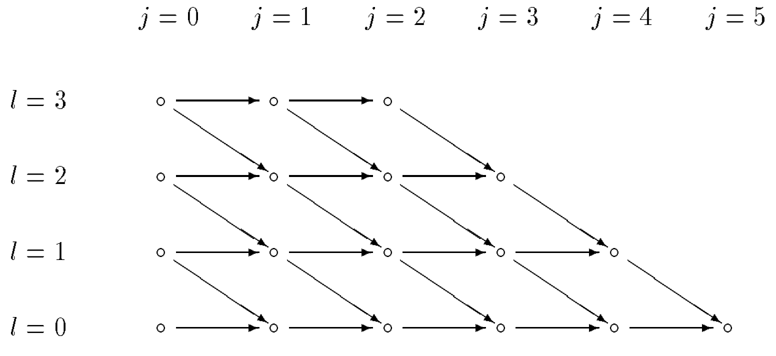

Figure 6.4: Computation of the Bernstein coefficients for $d_n = 3$ and $k = 5$.

Figure 6.4 illustrates the computation of the Bernstein coefficients according to Lemma 6.20. The left column is computed by (6.34), the upper row by (6.35) and the remaining points by (6.36). The Bernstein coefficients appear in the last row.

Proof of Lemma 6.20 Inductively we show for $l = 0, \ldots, d_n$, $j = 0, \ldots, k - l$ that

$$ {}^{l}b_j^{(k,X)} = \sum_{i=0}^{j} \frac{\binom{j}{i}}{\binom{k}{i+l}} a_{i+l}^{(\underline{X})} \mathrm{w}(X)^{i+l}. \tag{6.37} $$

Obviously (6.37) holds for $j = 0$, $l = 0, \ldots, d_n$ and for $l = d_n$, $j = 1, \ldots, k - d_n$. Further, for $j = 1, \ldots, k$, $l = 0, \ldots, \min\{d_n - 1, k - j\}$

$$
\begin{aligned}
{}^{l}b_j^{(k,X)} &= {}^{l}b_{j-1}^{(k,X)} + {}^{l+1}b_{j-1}^{(k,X)} \\
&= \sum_{i=0}^{j-1} \frac{\binom{j-1}{i}}{\binom{k}{i+l}} a_{i+l}^{(\underline{X})} \mathrm{w}(X)^{i+l} + \sum_{i=0}^{j-1} \frac{\binom{j-1}{i}}{\binom{k}{i+l+1}} a_{i+l+1}^{(\underline{X})} \mathrm{w}(X)^{i+l+1} \\
&= \sum_{i=0}^{j-1} \frac{\binom{j-1}{i}}{\binom{k}{i+l}} a_{i+l}^{(\underline{X})} \mathrm{w}(X)^{i+l} + \sum_{i=1}^{j} \frac{\binom{j-1}{i-1}}{\binom{k}{i+l}} a_{i+l}^{(\underline{X})} \mathrm{w}(X)^{i+l} \\
&= \sum_{i=1}^{j-1} \frac{\binom{j-1}{i} + \binom{j-1}{i-1}}{\binom{k}{i+l}} a_{i+l}^{(\underline{X})} \mathrm{w}(X)^{i+l} + \frac{1}{\binom{k}{l}} a_{l}^{(\underline{X})} \mathrm{w}(X)^{l} + \frac{1}{\binom{k}{j+l}} a_{j+l}^{(\underline{X})} \mathrm{w}(X)^{j+l} \\
&= \sum_{i=1}^{j-1} \frac{\binom{j}{i}}{\binom{k}{i+l}} a_{i+l}^{(\underline{X})} \mathrm{w}(X)^{i+l} + \frac{1}{\binom{k}{l}} a_{l}^{(\underline{X})} \mathrm{w}(X)^{l} + \frac{1}{\binom{k}{j+l}} a_{j+l}^{(\underline{X})} \mathrm{w}(X)^{j+l} \\
&= \sum_{i=0}^{j} \frac{\binom{j}{i}}{\binom{k}{i+l}} a_{i+l}^{(\underline{X})} \mathrm{w}(X)^{i+l}.
\end{aligned}
$$

In particular, if $l = 0$ we obtain from (6.37)

$$ {}^{0}b_j^{(k,X)} = \sum_{i=0}^{j} \frac{\binom{j}{i}}{\binom{k}{i}} a_{i}^{(\underline{X})} \mathrm{w}(X)^{i} = b_j^{(k,X)}. \ \square $$

Algorithm 6.1 (BCB) takes the Taylor coefficients as input and returns the smallest and the largest Bernstein coefficient.

Algorithm 6.1 (BCB**)** [Bernstein Coefficient Bounds]

In: $X \in \mathbb{IF}$,
$A_0, \ldots, A_{d_n} \in \mathbb{IF}$, $f^{(i)}(\underline{X})/i! \in A_i$, $i = 0, \ldots, d_n$,
$k \in \mathbb{N}$, $k \geq d_n$.

Out: $\mathrm{BCB}(X, A_0, \ldots, A_n, k) \in \mathbb{IF}$,
$\underline{\mathrm{BCB}(X, A_0, \ldots, A_n, k)} \leq b_j^{(k,X)} \leq \overline{\mathrm{BCB}(X, A_0, \ldots, A_n, k)}$, for all $j = 0, \ldots, k$.

(1) [Left Column.]
$w \longleftarrow \mathrm{WIDTH}(X)$
$W \longleftarrow w$
${}^{0}B \longleftarrow 1$
for $l = 1, \ldots, d_n$ do ${}^{l}B \longleftarrow {}^{l-1}B\, W/(k - l + 1)$, $W \longleftarrow W + w$
for $l = 0, \ldots, d_n$ do ${}^{l}B \longleftarrow {}^{l}B\, A_l$

(2) [Inner Points and Bounds for Bernstein Coefficients]
$F \longleftarrow {}^{0}B$
for $j = 1, \ldots, k$
  for $i = 0, \ldots, \min\{d_n - 1, k - j\}$ do ${}^{l}B \longleftarrow {}^{l}B + {}^{l+1}B$
  $\underline{F} \longleftarrow \min\{\underline{{}^{0}B}, \underline{F}\}$, $\overline{F} \longleftarrow \max\{\overline{{}^{0}B}, \overline{F}\}$

(3) [Return.]
return F

Theorem 6.21 (Complexity) *Algorithm 6.1 (BCB) costs*

$$\begin{array}{rl} 2d_n & \textit{number divisions,} \\ 4d_n+2 & \textit{number multiplications and} \\ d_n(2k-d_n+3)+1 & \textit{number additions.} \ \Box \end{array}$$

Proof.

- Step 1 costs $2d_n$ number divisions, $2d_n+1$ interval multiplications and $2d_n+1$ number additions. As $0 \notin \mathrm{int}({}^lB)$ for all l, each interval multiplication costs 2 number multiplications.
- Step 2 costs $d_n(2k-d_n+1)$ number additions.

An algorithm for computing the Bernstein form is now straight forward.

Algorithm 6.2 (BF) [Bernstein Form]

In: $f(x) = a_n x^{d_n} + a_{n-1}x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$,
$k \in \mathbb{N}$, $k \geq d_n$.

Out: $\mathrm{BF}(f, X, k) \in \mathbb{IF}, \mathrm{BF}(f, X, k) \supseteq B_f^{(k)}(X)$.

(1) [Taylor coefficients.]
$(A_0, \ldots, A_{d_n}) \longleftarrow \mathrm{TC}(f, \underline{X})$

(2) [Bounds for Bernstein coefficients.]
$F \longleftarrow \mathrm{BCB}(X, A_0, \ldots, A_{d_n}, k)$

(3) [Return.]
return F

Theorem 6.22 (Complexity) *Algorithm 6.2 (BF) costs*

$$\begin{array}{rl} 2d_n & \textit{number divisions,} \\ d_n^2+3d_n+4 & \textit{number multiplications and} \\ 2kd_n+4d_n+1 & \textit{number additions.} \ \Box \end{array}$$

Proof.

- Step 1 costs $d_n^2 - d_n + 2$ number multiplications $d_n^2 + d_n$ number additions (Theorem 5.6).
- Step 2 costs $2d_n$ number divisions, $4d_n+2$ number multiplications and $d_n(2k-d_n+3)+1$ number additions (Theorem 6.21). $\Box$

6.1 Bisection at Zero

If $0 \in X$ then we can avoid the expensive computation of Taylor coefficients in Algorithm 6.2 if we bisect X at 0 and compute the Bernstein coefficients separately for both halves of X. This reduces overall computing time if d_n is large and $k \approx d_n$. Further, due to the inclusion monotonicity of the Bernstein form, bisection leads usually to a reduction of the overestimation error.

Thus, in the sequel let

$$\breve{B}_f^{(k)}(X) = \begin{cases} B_f^{(k)}([\underline{X}, 0]) \cup B_f^{(k)}([0, \overline{X}]) & \text{if } 0 \in X \\ B_f^{(k)}(X) & \text{else.} \end{cases}$$

The following algorithm computes $\breve{B}_f^{(k)}(X)$.

Algorithm 6.3 (BFBZ) [Bernstein Form with Bisection at Zero]

In: $f(x) = a_n x^{d_n} + a_{n-1} x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$,
$k \in \mathbb{N}$, $k \geq d_n$.

Out: $\mathrm{BFBZ}(f, X, k) \in \mathbb{IF}$, $\mathrm{BFBZ}(f, X, k) \supseteq \breve{B}_f^{(k)}(X)$.

(1) [Test $0 \in X$.]
if $\underline{X} > 0$ or $\overline{X} < 0$ then return $\mathrm{BF}(f, X, k)$

(1) [Coefficients of $f(x)$.]
for $i = 0, \ldots, d_n$ do $A_i \longleftarrow$ coefficient of x^i in f

(2) [Bounds for Bernstein Coefficients in $[0, \overline{X}]$.]
$F^+ \longleftarrow \mathrm{BCB}([0, \overline{X}], A_0, \ldots, A_{d_n}, k)$

(3) [Coefficients of $f(-x)$.]
for $i = 1, \ldots, d_n$ step 2 do $A_i \longleftarrow -A_i$

(4) [Bounds for Bernstein Coefficients in $[0, -\underline{X}]$.]
$F^- \longleftarrow \mathrm{BCB}([0, -\underline{X}], A_0, \ldots, A_{d_n}, k)$

(5) [Return.]
return $F^+ \cup F^-$

Theorem 6.23 (Complexity)

- *If $0 \notin X$ then Algorithm 6.3 costs*

$$\begin{array}{rl} 2d_n & \textit{number divisions,} \\ d_n^2 + 3d_n + 4 & \textit{number multiplications and} \\ 2kd_n + 4d_n + 1 & \textit{number additions.} \end{array}$$

- *If $0 \in X$ then Algorithm 6.3 costs*

$$\begin{array}{rl} 4d_n & \textit{number divisions,} \\ 8d_n + 4 & \textit{number multiplications and} \\ 2d_n(2k - d_n + 3) + 2 & \textit{number additions.} \ \square \end{array}$$

Proof.

- If $0 \notin X$ then Algorithm 6.3 (BFBZ) and Algorithm 6.2 (BF) are identical.
- If $0 \in X$ then step 2 and step 4 cost each $2d_n$ number divisions, $4d_n + 2$ number multiplications and $d_n(2k - d_n + 3) + 1$ number additions. $\square$

Thus, if $0 \in X$ then Algorithm 6.3 (BFBZ) costs $d_n^2 - 5d_n$ number multiplications less, but $2d_n$ number divisions and $2d_n(k - d_n + 1) + 1$ number additions more than Algorithm BF. In particular, if $k = d_n$, the difference in number additions is only $2d_n + 1$. In this case BFBZ costs $d_n^2 - 9d_n - 1$ floating point operations less than BF, hence BFBZ is cheaper than BF if $d_n > 9$.

6.2 Experimental Results

An experimental comparison of Algorithm 6.2 (BF) and Algorithm 6.3 (BFBZ) for dense polynomials with different degrees is given in Table 5.1. The coefficients of the polynomomials are uniformly distributed in $[-1, 1]$ and the order of the Bernstein form is $k \equiv d_n$. Further, $\underline{X}$ and $\overline{X}$ are uniformly chosen in $[-1, 0]$ respectively $[0, 1]$, such that in all cases $0 \in X$. For each degree d_n the average cost and overestimation error of 10^4 random polynomials is reported. The cost is the total number of arithmetic floating point instructions, including those which were executed during interval operations. As the Bernstein form is quadratically convergent, the distance to the range divied by $\mathrm{w}(X)^2$ was chosen to measure accuracy.

According to Table 6.1, the error of $B_f^{(d_n)}$ is between one half and one third of the error of $\check{B}_f^{(d_n)}$. The measured costs of both algorithms correspond exactly to Theorem 6.22 respectively Theorem 6.23.

Degree d_n of f	2	6	10	14	18	22	26	30
Flops for $B_f^{(d_n)}(X)$	35.0	167.0	395.0	719.0	1139.0	1655.0	2267.0	2975.0
Flops for $\check{B}_f^{(d_n)}(X)$	50.0	186.0	386.0	650.0	978.0	1370.0	1826.0	2346.0
$\mathrm{q}(B_f^{(d_n)}(X), f(X))/\mathrm{w}(X)^2$	0.0523	0.0363	0.0252	0.0205	0.0174	0.0154	0.0137	0.0133
$\mathrm{q}(\check{B}_f^{(d_n)}(X), f(X))/\mathrm{w}(X)^2$	0.0194	0.0120	0.0092	0.0079	0.0069	0.0062	0.0057	0.0054

Table 6.1: Comparison of Bernstein form $B_f^{(d_n)}$ and Bernstein form with bisection at zero $\check{B}_f^{(d_n)}(X)$ for random polynomials f with different degrees and random intervals X, where $0 \in X$.

7 Interpolation Form

Interpolation forms were introduced by [Cornelius and Lohner, 1984]. They are the first interval extensions which have convergence order higher than 2. The original definition of interpolation forms comprises a whole class of interval extensions with different convergence orders. A particular instance, which is cubically convergent, and which allows efficient evaluation, is suggested in [Cornelius and Lohner, 1984]. In this section, we consider only this instance and call it "the" interpolation form. Section 7.1 and 7.2 contain some modifications, which seem to be new. Another cubially convergent interpolation form, the parabolic boundary value form, was introdued by [Neumaier, 1990], see Section 7.3.

The basic idea of interpolation forms is to approximate f by a low degree polynomial g such that

- the range of g can be computed without overestimation,
- the remainder $f - g$ can be enclosed with high convergence order.

Let us now derive the interpolation form. Using Taylor expansion of f at c we obtain

$$f(x) = f(c) + f'(c)(x-c) + 1/2 f''(\xi)(x-c)^2$$

for some ξ between c and x. For arbitrary $m \in \mathbb{R}$ we can continue with

$$f(x) = \underbrace{f(c) + f'(c)(x-c) + 1/2 m (x-c)^2}_{g_{c,m}(x)} + \underbrace{1/2(f''(\xi) - m)(x-c)^2}_{r_{c,m}(x)}.$$

Thus, a class of interval extensions is defined by

$$f(X) \subseteq g_{c,m}(X) + R_{c,m}(X), \tag{7.1}$$

where $R_{c,m}$ is an interval extension of $r_{c,m}$. Note that $g_{c,m}$ is a polynomial of degree two, hence $g_{c,m}(X)$ can be computed cheaply. The interpolation form is an instance of (7.1), where $c = \mathrm{mid}(X)$, $m = \mathrm{mid}(H_{f''}(X))$ and

$$R_{c,m}(X) = 1/2(H_{f''}(X) - m)(X-c)^2.$$

Definition 7.1 (Interpolation form) *The interpolation form $I_f : \mathbb{IR} \rightarrow \mathbb{IR}$ is defined as*

$$I_f(X) = g_{c,m}(X) + 1/2(H_{f''}(X) - m)(X-c)^2,$$

where

$$\begin{aligned} g_{c,m}(x) &= f(c) + f'(c)(x-c) + 1/2m(x-c)^2 \\ c &= \mathrm{mid}(X) \\ m &= \mathrm{mid}(H_{f''}(X)). \ \square \end{aligned}$$

In the sequel let c, m, and $g_{c,m}$ as in Definition 7.1 and let $r_{c,m} = f - g_{c,m}$.

Theorem 7.2 (Convergence) *I_f is cubically convergent.* $\square$

Proof. Let $A \in \mathbb{IR}$ arbitrary but fixed and let X range over subintervals of A. Note that for all $x \in X$ it holds that

$$\begin{aligned} |f(x) - g_{c,m}(x)| &= |r_{c,m}(x)| \\ &\leq \mathrm{mag}\left(1/2(H_{f''}(X) - m)(X-c)^2\right). \end{aligned}$$

Thus,

$$\mathrm{q}(f(X), g_{c,m}(X)) \leq \mathrm{mag}\left(1/2(H_{f''}(X) - m)(X-c)^2\right)$$

and

$$\begin{aligned} \mathrm{q}(f(X), I_f(X)) &= \mathrm{q}\big(f(X), g_{c,m}(X) + 1/2(H_{f''}(X) - m)(X-c)^2\big) \\ &\leq \mathrm{mag}\left((H_{f''}(X) - m)(X-c)^2\right). \end{aligned}$$

Hence, it sufficies to find a constant ℓ such that for all $X \in \mathbb{I}A$

$$\mathrm{mag}\left((H_{f''}(X) - m)(X-c)^2\right) \leq \ell \mathrm{w}(X)^3.$$

As $H_{f''}(X) - m$ is a centered interval we have

$$\begin{aligned} \mathrm{mag}\left((H_{f''}(X) - m)(X-c)^2\right) &= \mathrm{mag}([-1,1]\mathrm{rad}(H_{f''}(X))\mathrm{rad}(X)^2) \\ &= 1/2\mathrm{w}(H_{f''}(X))\mathrm{rad}(X)^2 \\ &\leq 1/2\ell_{H_{f''},A}\mathrm{w}(X)\mathrm{rad}(X)^2 \\ &= \frac{1}{8}\ell_{H_{f''},A}\mathrm{w}(X)^3, \end{aligned}$$

where $\ell_{H_{f''},A}$ is a Lipschitz constant of $H_{f''}$ in A. $\square$

The interpolation form requires exact range computation of parabolas. The following theorem ([Apostolatos and Kulisch, 1967]) shows, how this can be done efficiently.

Theorem 7.3 *Let $g(x) = a_2x^2 + a_1x + a_0$. Then*

$$g(X) = \begin{cases} a_1X + a_0 & \text{if } a_2 = 0 \\ \frac{1}{a_2}\left(\left(a_2X + \frac{a_1}{2}\right)^2 - \left(\frac{a_1}{2}\right)^2\right) + a_0 & \text{if } a_2 \neq 0. \ \square \end{cases} \tag{7.2}$$

Proof. The case $a_2 = 0$ is trivial, hence assume $a_2 \neq 0$. Note that

$$\frac{1}{a_2}\left(\left(a_2X + \frac{a_1}{2}\right)^2 - \left(\frac{a_1}{2}\right)^2\right) + a_0 = \left\{\frac{1}{a_2}\left(\left(a_2x + \frac{a_1}{2}\right)^2 - \left(\frac{a_1}{2}\right)^2\right) + a_0 \mid x \in X\right\},$$

because X occurs only once in the expression on the left hand side. Further

$$\begin{aligned} \frac{1}{a_2}\left(\left(a_2x + \frac{a_1}{2}\right)^2 - \left(\frac{a_1}{2}\right)^2\right) + a_0 &= \frac{1}{a_2}\left(a_2^2x^2 + a_2a_1x + \left(\frac{a_1}{2}\right)^2 - \left(\frac{a_1}{2}\right)^2\right) + a_0 \\ &= a_2x^2 + a_1x + a_0. \ \square \end{aligned}$$

An implementation of (7.2) using floating point numbers is numerically unstable if $a_2 \approx 0$. Therefore, we prefer to evaluate g at the endpoints of X, test whether g has a local extremum in X, and if so, evaluate g at the extremum. Algorithm 7.1 (PE) computes bounds for the range of parabolas $A_2x^2 + A_1x + A_0$, where A_2, A_1 and A_0 are intervals. However, it is assumed that A_2 and A_1 are tight, which will be the case when PE is called later on by other algorithms. Otherwise, the overestimation error of PE might be large. If $0 \in A_2$, then PE evaluates the Horner scheme. This does not lead to big overestimation errors if $\mathrm{w}(A_2)$ is small.

Algorithm 7.1 (PE) [Parabola Evaluation]

In: $A_2, A_1, A_0 \in \mathbb{IF}$,
$X \in \mathbb{IF}$.

Out: $\mathrm{PE}(A_2, A_1, A_0, X) \in \mathbb{IF}$, $\mathrm{PE}(A_2, A_1, A_0, X) \supseteq \{A_2x^2 + A_1x + A_0 \mid x \in X\}$.

(1) [Case $0 \in A_2$.]
if $0 \in A_2$ then return $(A_2X + A_1)X + A_0$

(2) [Evaluate at Endpoints.]
$Y \longleftarrow [(A_2\underline{X} + A_1)\underline{X}, A_2\overline{X} + A_1)\overline{X}]$

(3) [Extremum.]
$B_1 \longleftarrow 0.5A_1$
$E \longleftarrow (-B_1/A_2) \cap X$
if $E \neq \emptyset$ then $Y \longleftarrow [Y, B_1E]$

(4) [Return.]
return $Y + A_0$

Theorem 7.4 (Complexity) *Algorithm 7.1 costs*

1	*interval multiplication,*
2	*number divisions,*
10	*number multiplications and*
6	*number additions.* □

Proof. If $0 \in A_2$ then Algorithm 7.1 (PE) costs 2 interval multiplications and 2 interval additions. Assume $0 \notin A_2$.

- Step 2 costs 8 number multiplications and 4 number additions.
- Step 3 costs 1 interval multiplication, 2 number divisions and 2 number multiplications.
- Step 4 costs 2 number additions. □

Algorithm and Complexity

Algorithm 7.2 evaluates the interpolation form I_f. We assume that f is sparse, hence we do not use the extended Horner scheme for evaluating f' at c but compute the coefficients of f' explicitly. A modification of the algorithm for dense polynomials is given in Section 7.2. The coefficients of $g_{c,m}$ are

$$g_{c,m}(x) = 1/2mx^2 + (f'(c) - mc)x + f(c) - f'(c)c + 1/2mc^2.$$

Algorithm 7.2 (IF) [Interpolation Form]

In: $f(x) = a_nx^{d_n} + a_{n-1}x^{d_{n-1}} + \ldots + a_1x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{IF}(f, X) \in \mathbb{IF}$, $\mathrm{IF}(f, X) \supseteq I_f(X)$.

(1) [Midpoint.]
$c \longleftarrow \mathrm{MID}(X)$

(2) [Derivatives of f, use rounded interval arithmetic for the coefficients.]
$F'(x) \longleftarrow$ formal derivative of f
$F''(x) \longleftarrow$ formal derivative of $F'(x)$

(3) [$f(c), f'(c), f''(X)$.]
$F \longleftarrow \mathrm{HF}(f(x), [c])$
$F' \longleftarrow \mathrm{HFI}(F'(x), [c])$
$F'' \longleftarrow \mathrm{HFI}(F''(x), X)$

(4) [Coefficients of $g_{c,m}$.]
$M \longleftarrow [\mathrm{MID}(F'')]$
$A_2 \longleftarrow 0.5M$
$A_1 \longleftarrow F' - Mc$
$A_0 \longleftarrow (A_2c - F')c + F$

(5) [Evaluation of parabola.]
$Y \longleftarrow \mathrm{PE}(A_2, A_1, A_0, X)$

(6) [Add Remainder and Return.]
$r \longleftarrow \mathrm{mag}(X - c)$
return $Y + (F'' - M)[0, 0.5 \hat{*} r \hat{*} r]$

Theorem 7.5 (Complexity) *Algorithm 7.2 (IF) costs*

$3n$	*interval power computations,*
$n+1$	*interval multiplications,*
2	*number divisions,*
$8n+23$	*number multiplications and*
$6n+16$	*number additions.* □

Proof.

- Step 1 costs at 1 number multiplication and 2 number additions.
- Step 2 costs $4n$ number multiplications.
- Step 3 costs $3n$ interval power computations, n interval multiplications, $4n$ number multiplications and $6n-6$ number additions.
- Step 4 costs 9 number multiplication and 8 number additions.
- Step 5 costs 1 interval multiplication 2 number division, 10 number multiplications and 6 number additions.
- Step 6 costs 3 number multiplications and 6 number additions. □

7.1 Reduction of the Overestimation Error

In this section we present an improvement $\ddot{I}_f$ of the interpolation form I_f. The costs for evaluating $\ddot{I}_f$ and I_f are comparable (Theorem 7.9), but $\ddot{I}_f$ gives always at least as tight bounds as I_f (Theorem 7.8). The content of this section is original.

Definition 7.6 *The modified interpolation form* $\ddot{I}_f : \mathbb{IR} \to \mathbb{IR}$ *of* f *is defined as*

$$\ddot{I}_f(X) = [\underline{p_c^{\downarrow}(X)}, \overline{p_c^{\uparrow}(X)}],$$

where

$$\begin{aligned} p_c^{\uparrow}(x) &= f(c)+f'(c)(x-c)+1/2\overline{F''}*(x-c)^2 \\ p_c^{\downarrow}(x) &= f(c)+f'(c)(x-c)+1/2\underline{F''}*(x-c)^2, \end{aligned}$$

$F''=H_{f''}(X)$ *and* $c=\mathrm{mid}(X)$. $\square$

In the sequel let F'', c, $p_c^{\uparrow}$ and $p_c^{\downarrow}$ as in Definition 7.6.

Remark. If $I_f(X)$ is evaluated twice with $m=\underline{F''}$ and $m=\overline{F''}$ and the results are intersected, then $\ddot{I}_f(X)$ is obtained.

Theorem 7.7 *$\ddot{I}_f$ is an interval extension of f.* $\square$

Proof. Let $x\in X$ arbitrary but fixed and let $\xi\in X$ such that

$$f(x)=f(c)+f'(c)(x-c)+1/2f''(\xi)(x-c)^2.$$

Then

$$\begin{aligned} \overline{p_c^{\uparrow}(X)} &\geq f(c)+f'(c)(x-c)+1/2\overline{F''}*(x-c)^2 \\ &\geq f(c)+f'(c)(x-c)+1/2f''(\xi)(x-c)^2 \\ &= f(x) \\ \underline{p_c^{\downarrow}(X)} &\leq f(c)+f'(c)(x-c)+1/2\underline{F''}*(x-c)^2 \\ &\leq f(c)+f'(c)(x-c)+1/2f''(\xi)(x-c)^2 \\ &= f(x), \end{aligned}$$

i.e. $f(X)\subseteq\ddot{I}_f(X)$. Further, $\ddot{I}_f(X)\in\mathbb{R}$ if $X\in\mathbb{R}$. $\square$

Theorem 7.8 *For all $X\in\mathbb{IR}$ it holds that*

$$\ddot{I}_f(X)\subseteq I_f(X).\ \square$$

Proof. Note that

$$\overline{p_c^{\uparrow}(X)}\geq\underline{p_c^{\downarrow}(X)}.$$

Hence, we have to show that

$$\overline{I_f(X)} \geq \overline{p_c^{\uparrow}(X)} \tag{7.3}$$
$$\underline{I_f(X)} \leq \underline{p_c^{\downarrow}(X)}. \tag{7.4}$$

- We begin with (7.3). Let $g_{c,m}$ as in Definition 7.1 and let $a,b\in X$ such that

$$\begin{aligned} p_c^{\uparrow}(a) &= \max_{x\in X} p_c^{\uparrow}(x) \\ g_{c,m}(b) &= \max_{x\in X} g_{c,m}(x). \end{aligned}$$

  Then

$$\begin{aligned} \overline{p^{\uparrow}(X)} &= p^{\uparrow}(a) \\ &= f(c)+f'(a-c)+1/2\overline{F''}(a-c)^2 \\ &= f(c)+f'(a-c)+1/2\mathrm{mid}(F'')(a-c)^2+1/2\mathrm{rad}(F'')(a-c)^2 \\ &= g_{c,m}(a)+1/2\mathrm{rad}(F'')(a-c)^2 \\ &\leq g_{c,m}(b)+1/2\mathrm{rad}(F'')(\overline{X}-c)^2 \\ &= \overline{I_f(X)}. \end{aligned}$$

- Similarly, for (7.4), let $g_{c,m}$ as in Definition 7.1 and let $a, b \in X$, such that

$$\begin{aligned} p_c^{\downarrow}(a) &= \min_{x \in X} p_c^{\downarrow}(x) \\ g_{c,m}(b) &= \min_{x \in X} g_{c,m}(x). \end{aligned}$$

  Then

$$\begin{aligned} \underline{p^{\downarrow}(X)} &= p^{\downarrow}(a) \\ &= f(c) + f'(a-c) + 1/2\underline{F''}(a-c)^2 \\ &= f(c) + f'(a-c) + 1/2\mathrm{mid}(F'')(a-c)^2 - 1/2\mathrm{rad}(F'')(a-c)^2 \\ &= g_{c,m}(a) - 1/2\mathrm{rad}(F'')(a-c)^2 \\ &\geq g_{c,m}(b) - 1/2\mathrm{rad}(F'')(\overline{X}-c)^2 \\ &= \underline{I_f(X)}. \ \Box \end{aligned}$$

Algorithm and Complexity.

The following algorithm evaluates $\ddot{I}_f$. As in Algorithm 7.2 (IF) we assume that f is sparse. The coefficients of $p_c^{\uparrow}$ and $p_c^{\downarrow}$ are

$$\begin{aligned} p_c^{\uparrow}(x) &= 1/2\overline{F''}x^2 + (f'(c) - \overline{F''}c)x + f(c) - f'(c)c + 1/2\overline{F''}c^2 \\ p_c^{\downarrow}(x) &= 1/2\underline{F''}x^2 + (f'(c) - \underline{F''}c)x + f(c) - f'(c)c + 1/2\underline{F''}c^2. \end{aligned}$$

Algorithm 7.3 (MIF) [Modified Interpolation Form]

In: $f(x) = a_n x^{d_n} + a_{n-1}x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{MIF}(f, X) \in \mathbb{IF}$, $\mathrm{MIF}(f, X) \supseteq \ddot{I}_f(X)$.

(1) [Midpoint.]
$c \longleftarrow \mathrm{MID}(X)$

(2) [Derivatives of f, Use Rounded Interval Arithmetic for the Coefficients.]
$F'(x) \longleftarrow$ formal derivative of $f(x)$
$F''(x) \longleftarrow$ formal derivative of $F'(x)$

(3) [$f(c), f'(c), f''(X)$.]
$F \longleftarrow \mathrm{HF}(f(x), [c])$
$F' \longleftarrow \mathrm{HFI}(F'(x), [c])$
$F'' \longleftarrow \mathrm{HFI}(F''(x), X)$

(4) [Coefficients of $p_c^{\uparrow}$ and $p_c^{\downarrow}$.]
$A_2 \longleftarrow 0.5F''$
$A_2^{\uparrow} \longleftarrow \overline{A}_2$, $A_1^{\uparrow} \longleftarrow F' - [\overline{F''}] * c$, $A_0^{\uparrow} \longleftarrow (A_2^{\uparrow}c - F')c$
$A_2^{\downarrow} \longleftarrow \underline{A}_2$, $A_1^{\downarrow} \longleftarrow F' - [\underline{F''}] * c$, $A_0^{\downarrow} \longleftarrow (A_2^{\downarrow}c - F')c$

(5) [Evaluate parabolas.]
$Y \longleftarrow [\mathrm{PE}(A_2^{\uparrow}, A_1^{\uparrow}, A_0^{\uparrow}, X), \mathrm{PE}(A_2^{\downarrow}, A_1^{\downarrow}, A_0^{\downarrow}, X)] + F$

(6) [Return.]
return Y

Theorem 7.9 (Complexity) *Algorithm 7.3 (*MIF*) costs*

$3n$	*interval power computations,*
$n+2$	*interval multiplications,*
4	*number divisions,*
$8n+35$	*number multiplication and*
$6n+18$	*number additions.* $\Box$

Proof.

- Step 1 costs 1 number multiplication and 2 number additions.
- Step 2 costs $4n$ number multiplications.
- Step 3 costs $3n$ interval power computations, n interval multiplications, $4n$ number multiplications and $6n-6$ number additions.
- Step 4 costs 14 number multiplications and 8 number additions.
- Step 5 costs 2 interval multiplications, 4 number divisions, 20 number multiplications 14 number additions. □

7.2 Slopes instead of Derivatives

The comparison of the mean value form and the slope form in Section 4.1 turned out, that using slopes instead of derivatives improves efficiency and accuracy for dense polynomials. In the following, we apply this idea to the modified interpolation form. The content of this section is original.

Definition 7.10 (Slope Interpolation Form) *The slope interpolation form* $\ddot{I}_f^{(\mathrm{s})} : \mathbb{IR} \rightarrow \mathbb{IR}$ *of* f *is defined as*

$$\ddot{I}_f^{(\mathrm{s})}(X) = [\underline{p_c^{\downarrow}(X)}, \overline{p_c^{\uparrow}(X)}],$$

where $c = \mathrm{mid}(X)$,

$$\begin{aligned} p_c^{\uparrow}(x) &= f(c) + f'(c)(x-c) + \overline{H}(x-c)^2 \\ p_c^{\downarrow}(x) &= f(c) + f'(c)(x-c) + \underline{H}(x-c)^2 \\ H &= H_{h_c}^*(X) \end{aligned}$$

and h_c *is the uniquely defined polynomial such that*

$$f(x) = f(c) + f'(c)(x-c) + h_c(x)(x-c)^2. \ \square$$

In the sequel let $p_c^{\uparrow}$, $p_c^{\downarrow}$, h_c and H as in Definition 7.10. The proof that $\ddot{I}_f^{(\mathrm{s})}$ is an interval extension of f, is completely analogous to the proof of Theorem 7.7.

Theorem 7.11 $\ddot{I}_f^{(\mathrm{s})}$ *is cubically convergent.* □

Proof. Let $A \in \mathbb{IR}$ and let X range over subintervals of A. For all $x \in X$ it holds that

$$\begin{aligned} |f(x) - p_c^{\uparrow}(x)| &= |\overline{H} - h_c(x)|(x-c)^2 &\leq \mathrm{w}(H)\mathrm{w}(X)^2 \\ |f(x) - p_c^{\downarrow}(x)| &= |\underline{H} - h_c(x)|(x-c)^2 &\leq \mathrm{w}(H)\mathrm{w}(X)^2. \end{aligned}$$

$H_{h_{\mathrm{mid}(.)}}^*$ is Lipschitz, and there exists $\ell_{H,A} \in \mathbb{R}$ such that

$$\mathrm{w}(H_{h_c}^*(X) \leq \ell_{H,A}\mathrm{w}(X)$$

for all $X \in \mathbb{I}A$. Hence,

$$\begin{aligned} \mathrm{q}(f(X), \ddot{I}_f^{(\mathrm{s})}(X)) &\leq \mathrm{w}(H)\mathrm{w}(X)^2 \\ &\leq \ell_{H,A}\mathrm{w}(X)^3. \ \square \end{aligned}$$

Algorithm and Complexity.

The following algorithm evaluates $\ddot{I}_f^{(s)}$. For the computation of $f(c), f'(c)$ and the coefficients of h_c we use the extended Horner scheme. The coefficients of $p_c^{\uparrow}$ and $p_c^{\downarrow}$ are

$$\begin{aligned} p_c^{\uparrow}(x) &= \overline{H}x^2 + (f'(c) - 2\overline{H}c)x + f(c) - f'(c)c + \overline{H}c^2 \\ p_c^{\downarrow}(x) &= \underline{H}x^2 + (f'(c) - 2\underline{H}c)x + f(c) - f'(c)c + \underline{H}c^2. \end{aligned}$$

Algorithm 7.4 (SIF) [Slope Interpolation Form]

In: $f(x) = a_n x^{d_n} + a_{n-1}x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{SIF}(f, X) \in \mathbb{IR}$, $\mathrm{SIF}(f, X) \supseteq \ddot{I}_f^{(s)}(X)$.

(1) [Trival Cases.]
if $d_n < 2$ return $\mathrm{HF}(f, X)$

(2) [Midpoint.]
$c \longleftarrow \mathrm{MID}(X)$

(3) [Dense Coefficients]
for $i = 0, \ldots, d_n$ do $A_i \longleftarrow$ coefficient of x^i in f

(4) [Horner Scheme, $f(c)$, $f'(c)$, h_c.]
for $i = d_n, \ldots, 1$ do $A_{i-1} \longleftarrow A_{i-1} + cA_i$
for $i = d_n, \ldots, 2$ do $A_{i-1} \longleftarrow A_{i-1} + cA_i$

(5) [Compute $H^*_{h_c}(X)$.]
$H \longleftarrow A_{d_n}$
for $i = d_n - 1, \ldots, 2$ do $H \longleftarrow HX + A_i$

(6) [Coefficients of Parabolas.]
$H_c^{\uparrow} \longleftarrow [\overline{H}] * c$, $H_1^{\uparrow} \longleftarrow A_1 - H_c^{\uparrow}$, $P_1^{\uparrow} \longleftarrow H_1^{\uparrow} - H_c^{\uparrow}$, $P_0^{\uparrow} \longleftarrow -H_1^{\uparrow}c$
$H_c^{\downarrow} \longleftarrow [\underline{H}] * c$, $H_1^{\downarrow} \longleftarrow A_1 - H_c^{\downarrow}$, $P_1^{\downarrow} \longleftarrow H_1^{\downarrow} - H_c^{\downarrow}$, $P_0^{\downarrow} \longleftarrow -H_1^{\downarrow}c$

(7) [Evaluate Parabolas.]
$Y \longleftarrow [\mathrm{PE}(\overline{H}, P_1^{\uparrow}, P_0^{\uparrow}, X), \mathrm{PE}(\underline{H}, P_1^{\downarrow}, P_0^{\downarrow}, X)] + A_0$

(8) [Return.]
return Y

Theorem 7.12 (Complexity) *Algorithm 7.4 costs*

d_n	*interval multiplications,*
4	*number divisions,*
$4d_n + 27$	*number multiplication and*
$6d_n + 18$	*number additions.* □

Proof.

- Step 2 costs 1 number multiplication and 2 number additions.
- Step 4 costs $4d_n - 2$ number multiplications and $4d_n - 2$ number additions.
- Step 5 costs $d_n - 2$ interval multiplications and $2d_n - 4$ number additions.
- Step 6 costs 8 number multiplications and 8 number additions.
- Step 7 costs 4 number division, 2 interval multiplications, 20 number multiplications and 14 number additions. □

7.3 Parabolic Boundary Value Form

The parabolic boundary value form is a modification of the interpolation form, which was introduced by [Neumaier, 1990]. For the approximation of f, the values $f(\underline{X})$, $f(\overline{X})$ are used, whereas the interpolation forms in the previous sections used $f(c)$ and $f'(c)$.

Definition 7.13 (Parabolic Boundary Value Form) *The parabolic boundary value form* $\tilde{I}_f^{(s)} : \mathbb{IR} \to \mathbb{IR}$ *is defined as*

$$\tilde{I}_f^{(s)}(X) = [\underline{p_X^{\downarrow}(X)}, \overline{p_X^{\uparrow}(X)}],$$

where

$$\begin{aligned} p_X^{\uparrow}(x) &= a_X + b_X x + \underline{H}(x - \underline{X})(x - \overline{X}) \\ p_X^{\downarrow}(x) &= a_X + b_X x + \overline{H}(x - \underline{X})(x - \overline{X}) \\ H &= H^*_{h_X}(X) \end{aligned}$$

and a_X, b_X, h_X are uniquely defined by

$$f(x) = a_X + b_X x + h_X(x)(x - \underline{X})(x - \overline{X}). \;\square$$

In the sequel let a, b and h as in Definition 7.13

Theorem 7.14 $\tilde{I}_f^{(s)}$ *is a cubically convergent interval extension of f.* $\square$

Proof. See [Neumaier, 1990]. $\square$

Algorithm and Complexity.

For the computation of a_X, b_X and the coefficients of h_X we use the extended Horner scheme. More precisely, we compute first $f(\underline{X})$ and the coefficients of the polynomial g_X such that

$$f(x) = f(\underline{X}) + g_X(x)(x - \underline{X}).$$

Next, we compute $g_X(\overline{X})$ and the coefficients of h_X, where

$$g_X(x) = g_X(\overline{X}) + h_X(x)(x - \overline{X}).$$

Thus,

$$\begin{aligned} f(x) &= f(\underline{X}) + \big(g_X(\overline{X}) + h_X(x)(x - \overline{X})\big)(x - \underline{X}) \\ &= \underbrace{f(\underline{X}) - g_X(\overline{X})\underline{X}}_{a_X} + \underbrace{g_X(\overline{X})}_{b_X} x + h_X(x)(x - \underline{X})(x - \overline{X}). \end{aligned}$$

The coefficients of $p_X^{\uparrow}$ and $p_X^{\downarrow}$ are

$$\begin{aligned} p_X^{\uparrow}(x) &= f(\underline{X}) - g_X(\overline{X})\underline{X} + g_X(\overline{X})x + \underline{H}(x - \underline{X})(x - \overline{X}) \\ &= \underline{H}x^2 + \big(g_X(\overline{X}) - \underline{H}(\overline{X} + \underline{X})\big)x + f(\underline{X}) - g_X(\overline{X})\underline{X} + \underline{H}\,\underline{X}\overline{X} \\ p_X^{\downarrow}(x) &= f(\underline{X}) - g_X(\overline{X})\underline{X} + g_X(\overline{X})x + \overline{H}(x - \underline{X})(x - \overline{X}) \\ &= \overline{H}x^2 + \big(g_X(\overline{X}) - \overline{H}(\overline{X} + \underline{X})\big)x + f(\underline{X}) - g_X(\overline{X})\underline{X} + \overline{H}\,\underline{X}\overline{X}. \end{aligned}$$

Algorithm 7.5 (PBF) [Parabolic Boundary Value Form]

In: $f(x) = a_n x^{d_n} + a_{n-1}x^{d_{n-1}} + \ldots + a_1 x^{d_1} \in \mathbb{F}[x]$,
$X \in \mathbb{IF}$.

Out: $\mathrm{PBF}(f, X) \in \mathbb{IR}$, $\mathrm{PBF}(f, X) \supseteq \tilde{I}_f^{(s)}(X)$.

(1) [Trivial cases.]
if $d_n < 2$ return HF(f, X)

(2) [Dense coefficients]
for $i = 0, \ldots, d_n$ do $A_i \longleftarrow$ coefficient of x^i in f

(3) [Horner scheme, $f(\underline{X})$, $g_X(\overline{X})$, h_X.]
for $i = d_n, \ldots, 1$ do $A_{i-1} \longleftarrow A_{i-1} + \underline{X} A_i$
for $i = d_n, \ldots, 2$ do $A_{i-1} \longleftarrow A_{i-1} + \overline{X} A_i$

(4) [Compute $H^*_{h_X}(X)$.]
$H \longleftarrow A_{d_n}$
for $i = d_n - 1, \ldots, 2$ do $H \longleftarrow HX + A_i$

(5) [Coefficients of parabolas.]
$X^+ \longleftarrow [\underline{X}] + [\overline{X}]$
$X^* \longleftarrow [\underline{X}] * [\overline{X}]$
$P_1^\uparrow \longleftarrow A_1 - \underline{H}X^+$, $P_0^\uparrow \longleftarrow \underline{H}X^*$
$P_1^\downarrow \longleftarrow A_1 - \overline{H}X^+$, $P_0^\downarrow \longleftarrow \overline{H}X^*$

(6) [Evaluate parabolas.]
$Y \longleftarrow [\mathrm{PE}(\underline{H}, P_1^\uparrow, P_0^\uparrow, X), \mathrm{PE}(\overline{H}, P_1^\downarrow, P_0^\downarrow, X)] + A_0$

(7) [Return.]
return Y

Theorem 7.15 (Complexity) *Algorithm 7.5 (*PBF*) costs*

d_n	*interval multiplications,*
4	*number divisions,*
$4d_n + 28$	*number multiplications and*
$6d_n + 12$	*number additions.* □

Proof.

- Step 3 costs $4d_n - 2$ number multiplications and $4d_n - 2$ number additions.
- Step 4 costs $d_n - 2$ interval multiplications and $2d_n - 4$ number additions.
- Step 5 costs 10 number multiplications and 6 number additions.
- Step 6 costs 2 interval multiplications, 4 number divisions, 20 number multiplications and 12 number additions. □

7.4 Experimental Results

An experimental comparison of the interpolation form I_f, computed by Algorithm 7.2 (IF), the modified interpolation form $\ddot{I}_f$, computed by Algorithm 7.3 (MIF), the slope interpolation form $\ddot{I}_f^{(s)}$, computed by Algorithm 7.4 (SIF), and the parabolic boundary value form $\tilde{I}_f^{(s)}$, computed by Algorithm 7.5 (PBF), for dense and sparse polynomials with different degrees is given in Table 7.1, respectively Table 7.2. The coefficients of the polynomomials and the endpoints of the input intervals are uniformly distributed in $[-1, 1]$. For each degree d_n the average cost and overestimation error of 10^4 random polynomials is reported. The cost is the total number of arithmetic floating point instructions, including those which were executed during interval operations. As the forms are cubically convergent, the distance to the range divided by $\mathrm{w}(X)^3$ was chosen to measure accuracy.

- The accuracy of I_f and $\ddot{I}_f$ is almost the same, but $\ddot{I}_f$ is more expensive.

Degree of f (dense)	2	6	10	14	18	22	26	30
Flops for $I_f(X)$	68.2	137.2	204.2	273.2	338.4	398.2	467.4	542.0
Flops for $\ddot{I}_f(X)$	86.4	154.6	221.8	290.8	356.2	416.0	484.4	559.4
Flops for $\ddot{I}_f^{(\mathrm{s})}(X)$	68.4	118.8	168.8	222.2	271.6	315.2	367.6	425.2
Flops for $\tilde{I}_f^{(\mathrm{s})}(X)$	70.4	120.4	170.2	223.4	273.2	317.0	370.2	427.4
$\mathrm{q}(I_f(X), f(X))/\mathrm{w}(X)^3$	0.00	0.85	3.43	5.26	8.90	37.55	15.42	22.61
$\mathrm{q}(\ddot{I}_f(X), f(X))/\mathrm{w}(X)^3$	0.00	0.85	3.43	5.26	8.90	37.55	15.42	22.61
$\mathrm{q}(\ddot{I}_f^{(\mathrm{s})}(X), f(X))/\mathrm{w}(X)^3$	0.00	0.08	0.34	0.27	0.78	1.90	1.24	0.31
$\mathrm{q}(\tilde{I}_f^{(\mathrm{s})}(X), f(X))/\mathrm{w}(X)^3$	0.00	0.09	0.37	0.31	0.56	4.98	0.48	0.44

Table 7.1: Comparison of interpolation form I_f, modified interpolation form $\ddot{I}_f$, slope interpolation form $\ddot{I}_f^{(\mathrm{s})}$ and parabolic boundary value form $\tilde{I}_f^{(\mathrm{s})}$ for *dense* random polynomials f with different degrees and random intervals X. The coefficients of f and the endpoints of X are uniformly chosen in $[-1, 1]$.

Degree of f (sparse)	2	6	10	14	18	22	26	30
Flops for $I_f(X)$	68.6	101.4	120.9	133.5	143.4	150.5	156.4	162.1
Flops for $\ddot{I}_f(X)$	87.2	117.2	136.4	148.9	158.6	165.4	171.1	176.7
Flops for $\ddot{I}_f^{(\mathrm{s})}(X)$	69.2	119.8	171.0	222.8	274.7	326.5	378.6	430.2
Flops for $\tilde{I}_f^{(\mathrm{s})}(X)$	71.2	121.3	172.1	223.4	274.8	326.4	378.0	429.3
$\mathrm{q}(I_f(X), f(X))/\mathrm{w}(X)^3$	0.00	0.79	1.85	2.95	4.19	5.71	6.82	8.94
$\mathrm{q}(\ddot{I}_f(X), f(X))/\mathrm{w}(X)^3$	0.00	0.79	1.85	2.95	4.19	5.71	6.81	8.94
$\mathrm{q}(\ddot{I}_f^{(\mathrm{s})}(X), f(X))/\mathrm{w}(X)^3$	0.00	0.08	0.12	0.14	0.16	0.18	0.25	0.31
$\mathrm{q}(\tilde{I}_f^{(\mathrm{s})}(X), f(X))/\mathrm{w}(X)^3$	0.00	0.10	0.22	0.32	0.44	0.56	0.65	0.87

Table 7.2: Comparison of interpolation form I_f, modified interpolation form $\ddot{I}_f$, slope interpolation form $\ddot{I}_f^{(\mathrm{s})}$ and parabolic boundary value form $\tilde{I}_f^{(\mathrm{s})}$ for *sparse* random polynomials f with different degrees and random intervals X. Each polynomial has only 3 non–zero coefficients. The coefficients of f and the endpoints of X are uniformly chosen in $[-1, 1]$.

- In all cases $\ddot{I}_f^{(s)}$ and $\tilde{I}_f^{(s)}$ return significantly tighter inclusions than I_f and $\ddot{I}_f$. If f is dense, then $\ddot{I}_f^{(s)}$ is cheaper than I_f and $\ddot{I}_f$.
- It seems that $\ddot{I}_f^{(s)}$ gives slightly tighter inclusions than $\tilde{I}_f^{(s)}$. However, the following experiment shows that this depends on $\mathrm{w}(X)$.

Figure 7.1 shows a comparison of the slope interpolation form $\ddot{I}_f^{(s)}$ and the parabolic boundary value form $\tilde{I}_f^{(s)}$ for dense polynomials of degree 10. In the left graph, the input intervals X are chosen such that $\mathrm{mid}(X) \equiv 0.5$ and $\mathrm{w}(X)$ varies between 0 and 1, in the right graph $\mathrm{mid}(X) \equiv 0$ and $\mathrm{w}(X)$ varies between 0 and 2. The curves are averages over 1000 random polynomials with coefficients in $[-1, 1]$.

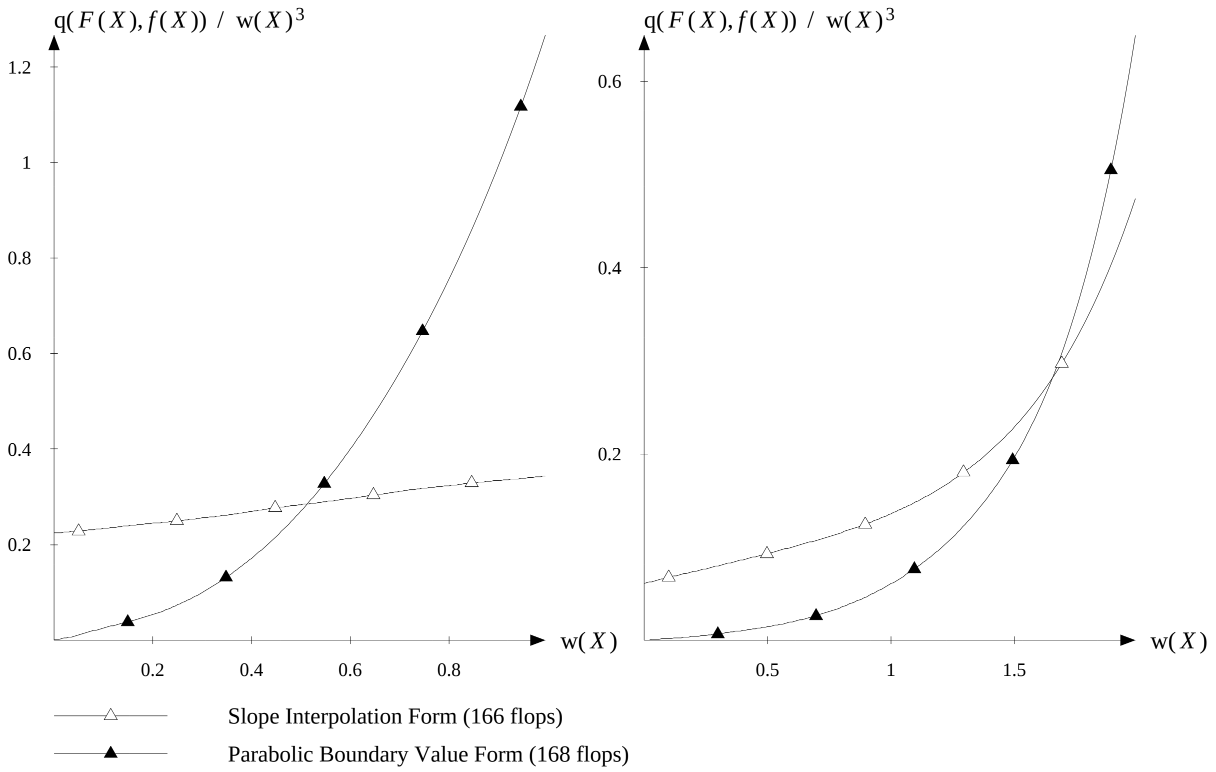

Figure 7.1: Comparison of $\ddot{I}_f^{(s)}$ and $\tilde{I}_f^{(s)}$ for dense random polynomials of degree 10. In the left graph the input intervals X are such that $\mathrm{mid}(X) \equiv 0.5$ and $\mathrm{w}(X)$ varies between 0 and 1, in the right graph $\mathrm{mid}(X) \equiv 0$ and $\mathrm{w}(X)$ varies between 0 and 2.

8 Experimental Comparison

After having defined several interval extensions of univariate polynomials in the previous sections, we now come to a comparison. The computational cost for the evaluation of the interval extensions can be compared easily because it depends basically only on the degree and on the sparsity of the input polynomial. A comparison of the overestimation errors is more involved, because the overestimation error depends on the actual value of the polynomial coefficients and on the input interval.

The main theoretical results of the previous sections can be summarized as follows.

- For input intervals with small width, interpolation forms are superior because of their cubic convergence (Theorem 7.2, 7.11, 7.14), the Horner form is worst because it is only linearly convergent

(Theorem 3.4), and the chances, that the bicentered mean value form and the Bernstein form are exact are high (Theorem 4.15, 6.19).

- If the Horner form of f' evaluated on X does not contain zero, then the bicentered mean value form is exact (Theorem 4.15). Note that every interval extension can be modified to be exact in this case, simply by executing the monotonicity test first. However, it depends on the application whether this test pays, hence we will not consider it further.
- The bicentered mean value form gives always at least as tight inclusions as the mean value form (Corollary 4.17).
- The linear factor form is at least as accurate as the dense mean value form (Conjecture 4.11).
- If $\mathrm{mig}(X)$ is large enough, then the Horner forms are exact (Theorem 3.10).
- If $\mathrm{mid}(X) = 0$, then the (dense) Horner form, the (dense) Taylor form and the (dense) linear factor form are equivalent.
- The overestimation error of the dense Taylor form with bisection is at most half as large as the overestimation error of the dense Taylor form (Theorem 5.9).
- If $0 \in \mathrm{int}(X)$ then a bisection at zero is inexpensive for the Horner forms (Algorithm 3.6) and the Bernstein form (Algorithm 6.3), and usually reduces the overestimation error.

Section 8.1 contains an experimental comparison of some interval extensions. The input interval X is fixed, and the average overestimation error of a class of random polynomials is computed. This kind of comparison shows at least statistically, which interval extension is best for a given input interval X.

Every interval extension is a compromise between computational cost and accuracy. For any algorithm, which evaluates interval extensions in a subroutine, we have to search for the optimal compromise. For example, in an iterative algorithm we would like to know, whether it is cheaper to choose an inaccurate and inexpensive interval extension (like the Horner form) and execute some more iterations, or to choose an accurate and expensive interval extension (like the Bernstein form) and save some iterations. In Section 8.2 and 8.3 we tackle this problem for two typical applications of interval arithmetic, namely finding the roots and the global minimum of univariate polynomials.

In the sequel, when we refer to the Bernstein form of a polynomial, we implicitly assume, that the order of the Bernstein form is equal to the degree of the polynomial.

8.1 Efficiency and Accuracy for Random Polynomials

In this section we compare the average overestimation error of some interval extensions for 3 classes of random polynomials and for 2 classes of input intervals.

- The classes of random polynomials f are:
  - Dense polynomials with degree 5 (Figure 8.1, 8.2).
  - Dense polynomials with degree 15 (Figure 8.3, 8.4).
  - Sparse polynomials with degree 10, where only 3 coefficients are non–zero (Figure 8.5, 8.6).

  The coefficients are uniformly chosen from $[-1, 1]$.
- The input intervals X are as follows:
  - $\mathrm{mid}(X) \equiv 0$, $0 \leq \mathrm{w}(X) \leq 1.0$ (Figure 8.1, 8.3, 8.5).
  - $\mathrm{mid}(X) \equiv 0.3$, $0 \leq \mathrm{w}(X) \leq 1.0$ (Figure 8.2, 8.4, 8.6).

The overestimation error of an interval extension F of f is defined as

$$\frac{\mathrm{q}(F(X), f(X))}{\mathrm{w}(X)}.$$

In each experiment we report the average overestimation error of 1000 random polynomials.

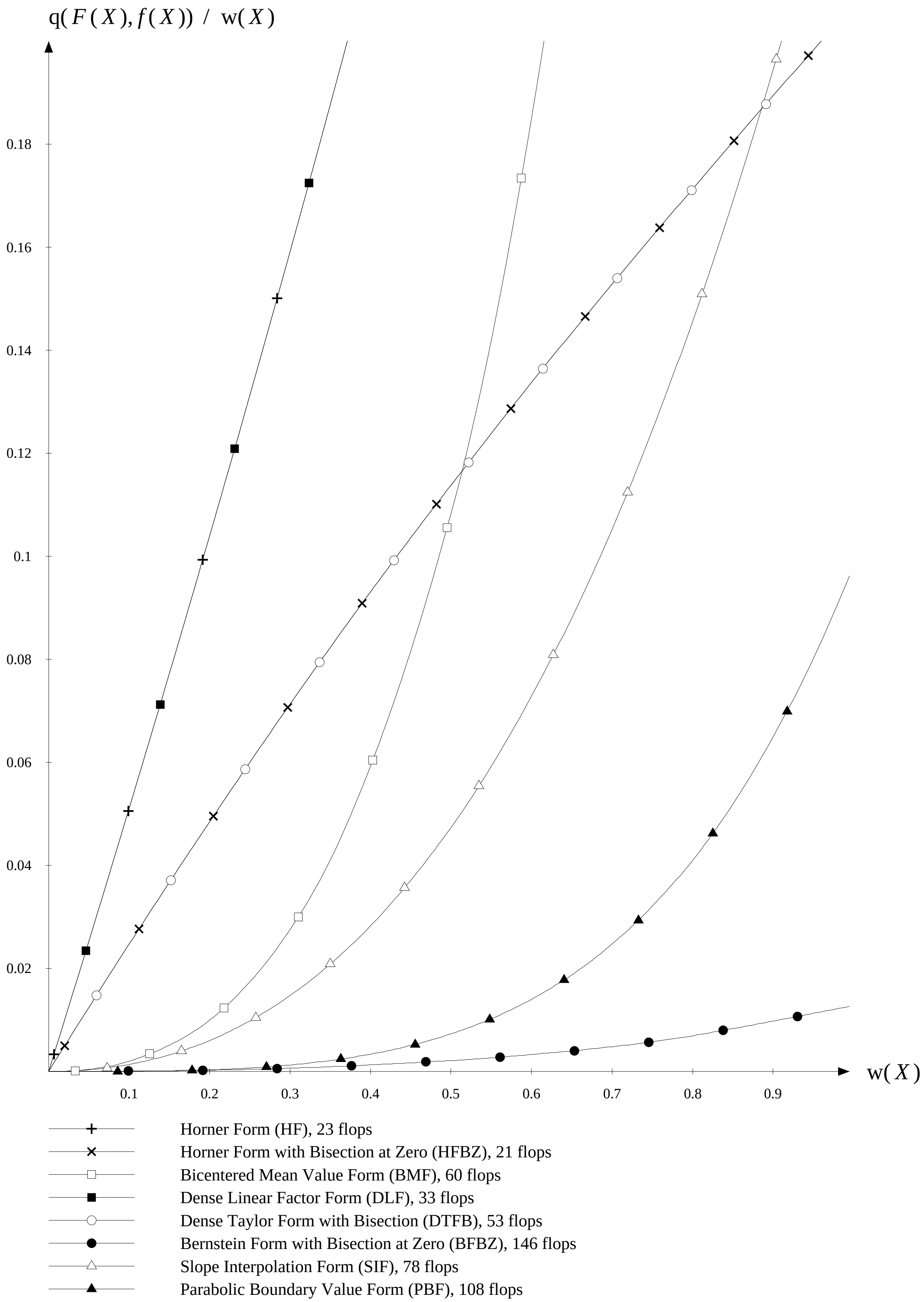

Figure 8.1: Average overestimation error for dense polynomials of degree 5, $\mathrm{mid}(X) \equiv 0$.

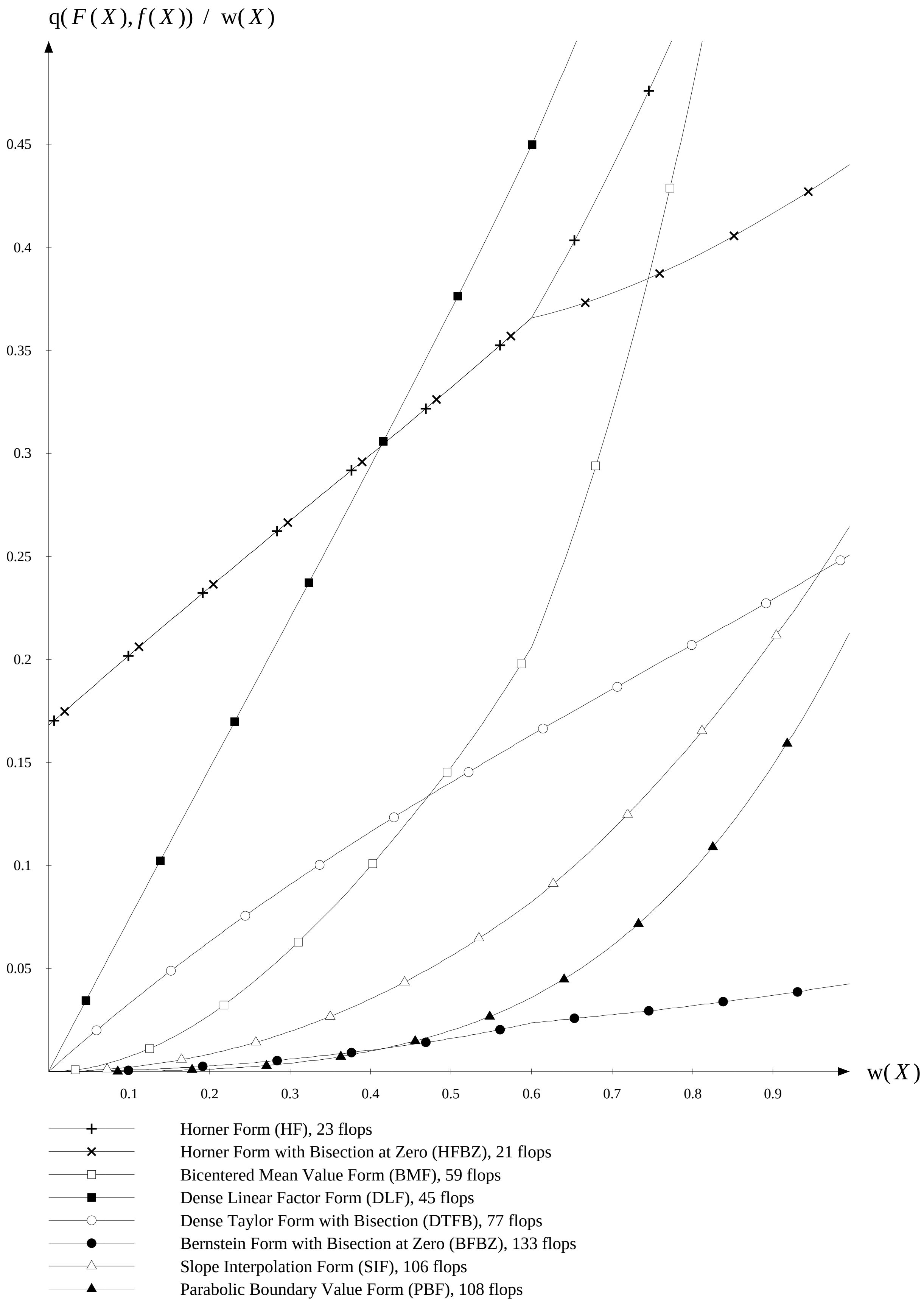

Figure 8.2: Average overestimation error for dense polynomials of degree 5, $\mathrm{mid}(X) \equiv 0.3$.

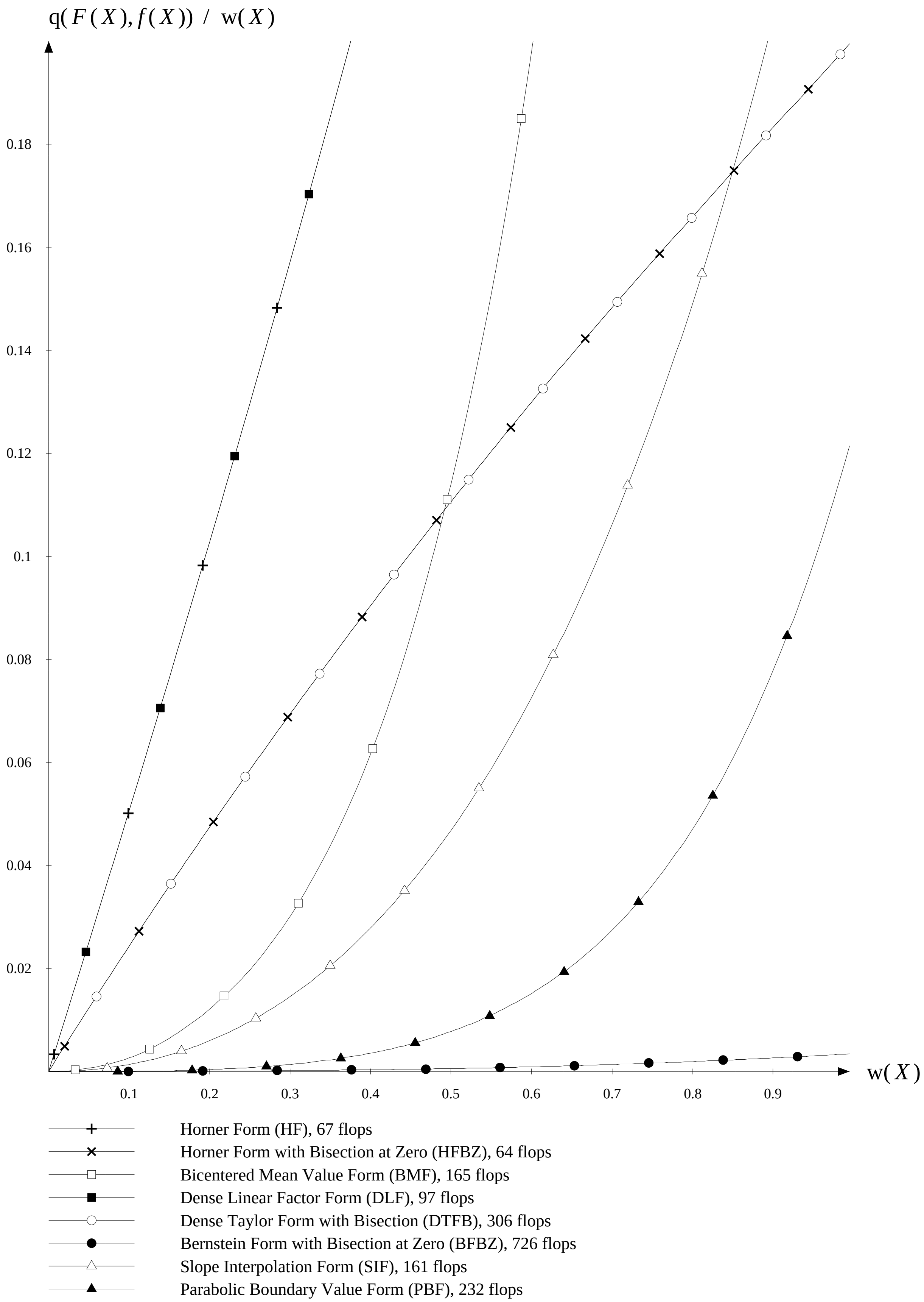

Figure 8.3: Average overestimation error for dense polynomials of degree 15, $\mathrm{mid}(X) \equiv 0.0$.

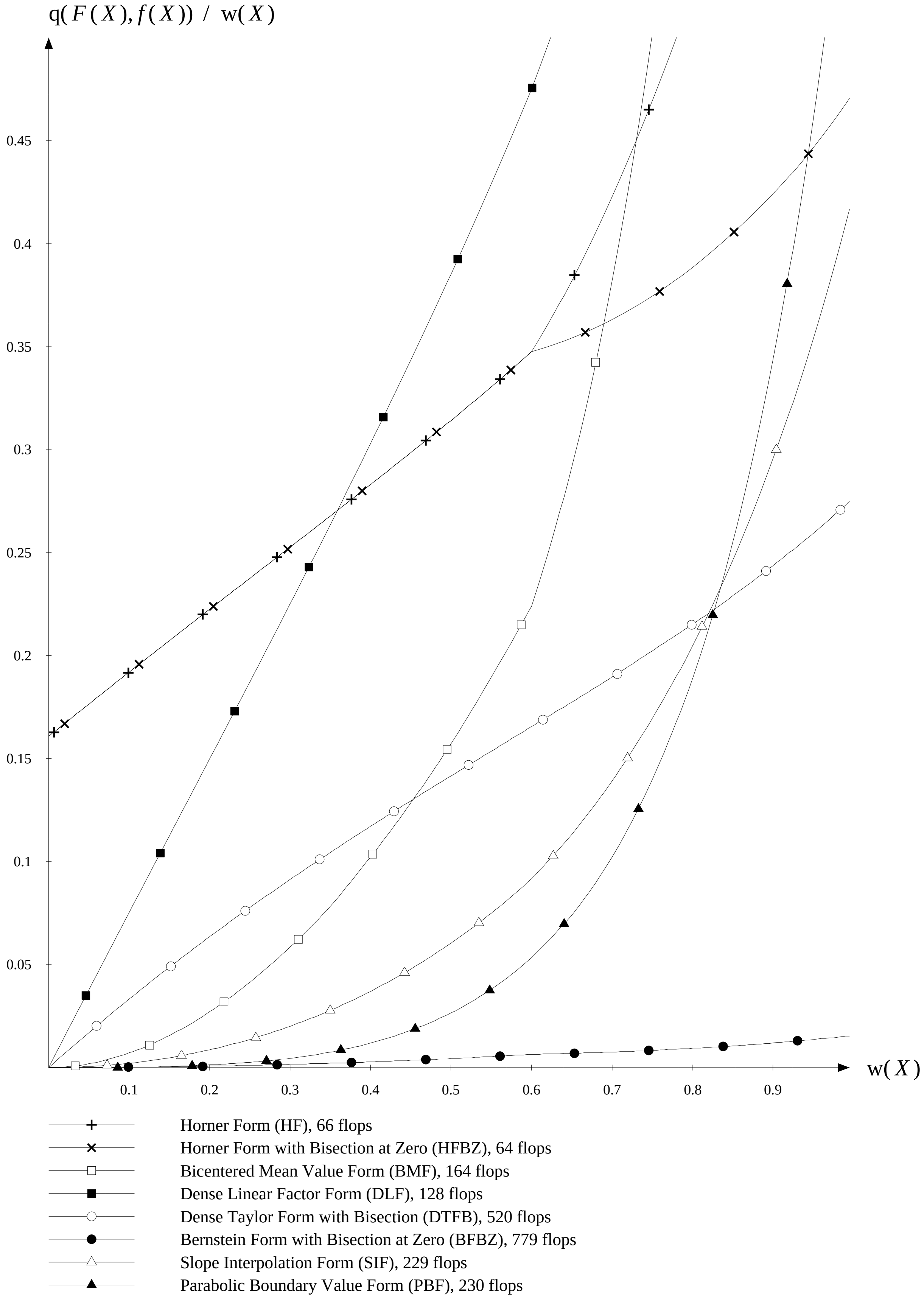

Figure 8.4: Average overestimation error for dense polynomials of degree 15, $\mathrm{mid}(X) \equiv 0.3$.

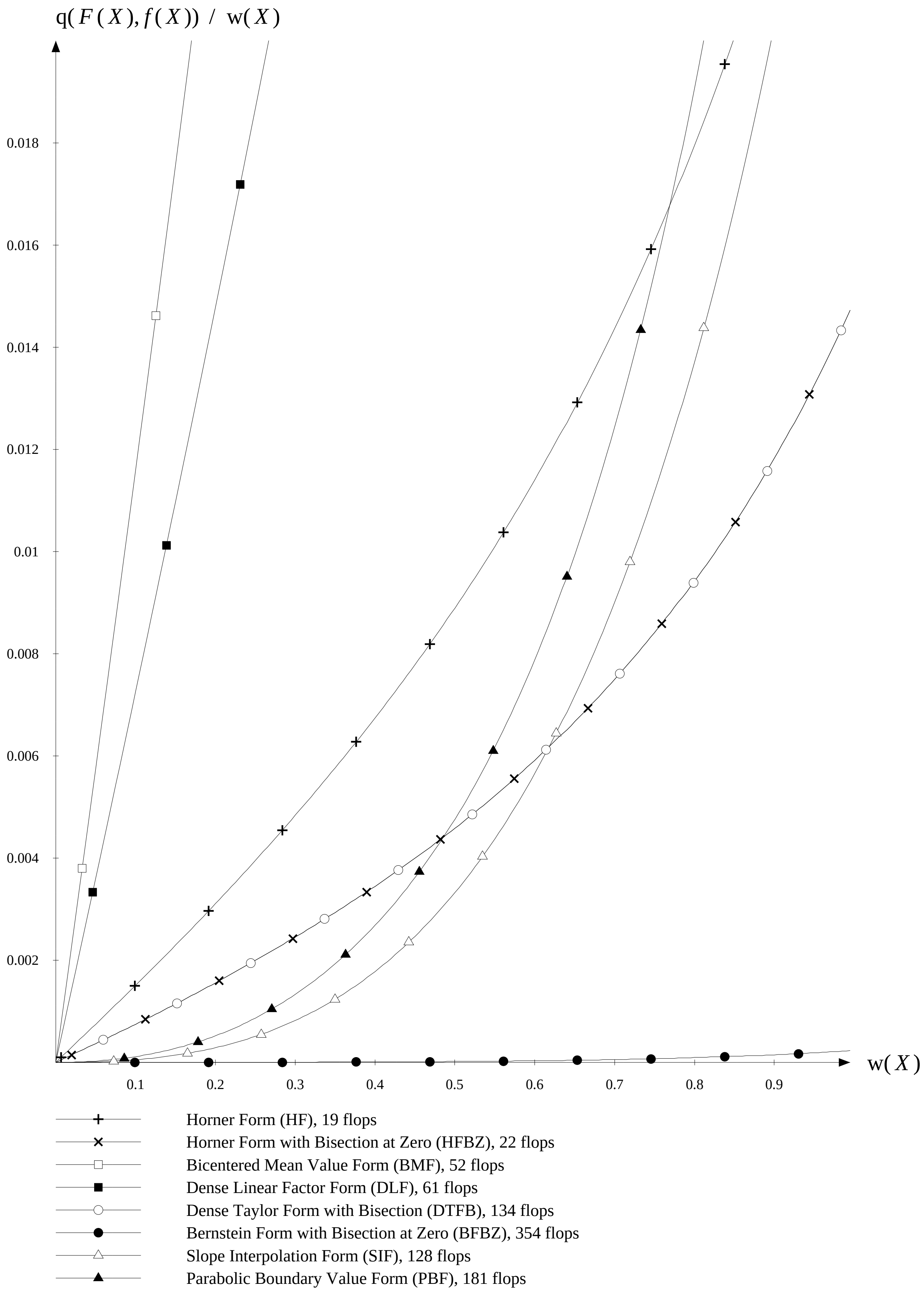

Figure 8.5: Average overestimation error for sparse polynomials of degree 10, mid$(X) \equiv 0.0$.

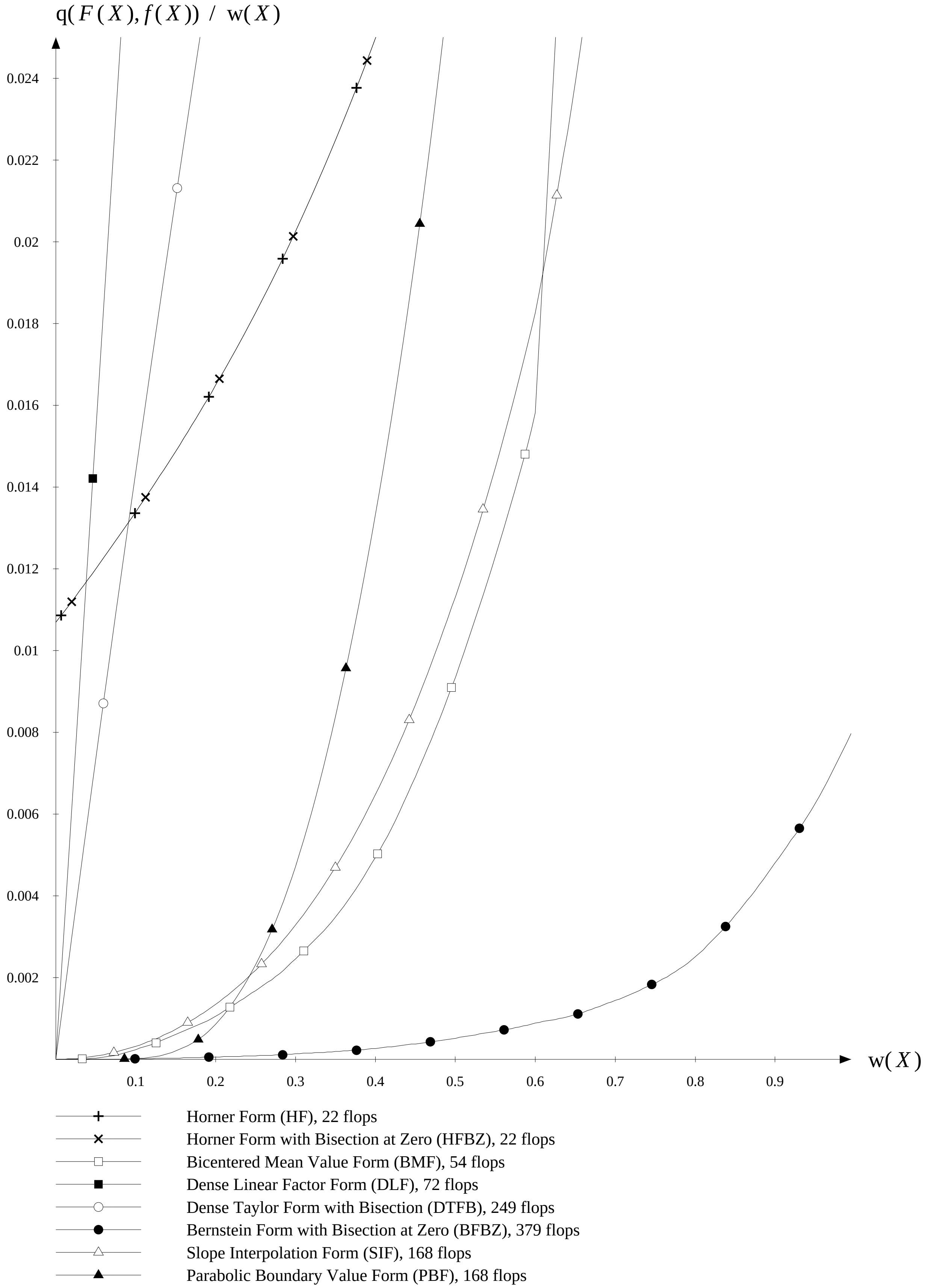

Figure 8.6: Average overestimation error for sparse polynomials of degree 10, $\mathrm{mid}(X) \equiv 0.3$.

8.2 Newton's Method

The (one–dimensional) interval Newton method is an iterative algorithm, which computes inclusions of all roots of a polynomial f in an interval A. It was studied thoroughly in the literature, see for example [Moore, 1966], [Krawczyk, 1969], [Alefeld, 1970], [Hansen, 1978a], [Hansen, 1978b], [Hansen, 1979], [Alefeld and Herzberger, 1983], [Hansen, 1992], [Hammer *et al.*, 1993] and many others. In each iteration, an interval extension of f' has to be evaluated. In this section we compare the efficiency of the interval Newton method in dependence of the chosen interval extension. The number of Newton iterations depends on the accuracy of the interval extension. Thus, a more accurate and more expensive interval extension reduces the number of Newton iterations, but increases the cost of each iteration.

In the experiments below, we compute the roots of random polynomials in $[-1, 1]$. We consider three classes of polynomials:

- Dense random polynomials of degree 10. The coefficients are uniformly chosen from $[-1, 1]$ (Table 8.1, left chart).
- Sparse random polynomials of degree 10. Only 3 coefficients are non–zero and chosen uniformly from $[-1, 1]$. Note that these polynomials have a multiple root at zero if the constant coefficient is zero (Table 8.1, middle chart).
- Dense random polynomials of degree 10 with at least 5 roots in $[0.5, 0.6]$. The polynomials are generated as follows: 5 points are randomly chosen from $[0.5, 0.6]$ and the unique polynomial of degree 5 is computed which vanishes on these points. The polynomial is multiplied by a random polynomial of degree 5, whose coefficients are uniformly chosen from $[-1, 1]$ (Table 8.1, right chart).

First, we give a skeleton algorithm for Newton's method. In the algorithm, F' denotes an interval extension of f'.

Algorithm 8.1 (NEWTON) [Newton Method for Univariate Polynomials]

In: $f(x) \in \mathbb{F}[x]$,
$A \in \mathbb{IF}$,
$\varepsilon \in \mathbb{F}$, $\varepsilon > 0$.

Out: $\mathcal{Z} \in \mathcal{P}(\mathbb{IF})$ such that

- for every $X \in \mathcal{Z}$ it holds that $\mathrm{w}(X) \leq \varepsilon$ or $\mathrm{w}(X)$ is so small that X cannot be bisected,
- if $f(x) = 0$ for some $x \in A$ then $x \in \bigcup \mathcal{Z}$.

(1) [Initialize.]
$\mathcal{S} \longleftarrow$ empty stack
$\mathcal{Z} \longleftarrow \{\}$
compute the coefficients of f'
push A on $\mathcal{S}$

(2) [Iterate.]
if $\mathcal{S}$ is empty, then return $\mathcal{Z}$
pop X from $\mathcal{S}$

(3) [Test if $\mathrm{w}(X)$ is too small.]
$c \longleftarrow \mathrm{MID}(X)$
if $\mathrm{WIDTH}(X) \leq \varepsilon$ or $c \notin \mathrm{int}(X)$ then $\mathcal{Z} \longleftarrow \mathcal{Z} \cup \{X\}$, goto Step 2

(4) [Compute interval extensions.]
$Y \longleftarrow \mathrm{HF}(f, [c])$
$D \longleftarrow F'(X)$

(5) [Newton step if $0 \notin D$.]
if $0 \notin D$

$X' \longleftarrow (c - Y/D) \cap X$
if $X' = \emptyset$ then goto Step 2
if $X' \neq X$ then push X' on $\mathcal{S}$ and goto Step 2
push $[\underline{X}, c]$ and $[c, \overline{X}]$ on $\mathcal{S}$ and goto Step 2

(6) [Split X if $0 \in D$.]
$X_1, X_2 \longleftarrow \text{GDIV}(-Y, D, X - c)$
if $X_1 \neq \emptyset$ then $X_1 \longleftarrow (X_1 + c) \cap X$
if $X_2 \neq \emptyset$ then $X_2 \longleftarrow (X_2 + c) \cap X$
if $X_1 \neq \emptyset$ and $X_2 \neq \emptyset$ and $X_1 \cap X_2 = \emptyset$
  if $X_1 \neq X$ and $X_2 \neq X$ then push X_1, X_2 on $\mathcal{S}$ and goto Step 2
  push $[\underline{X}, c]$ and $[c, \overline{X}]$ on $\mathcal{S}$ and goto Step 2
else
  $X' \longleftarrow X_1 \cup X_2$
  if $X' = \emptyset$ goto Step 2
  if $X' = X$ then push $[\underline{X}, c]$ and $[c, \overline{X}]$ on $\mathcal{S}$ and goto Step 2
  push X' on $\mathcal{S}$ and goto Step 2

Algorithm 8.1 (NEWTON) is modified slightly if the bicentered mean value form or the dense linear factor form is used.

- During the evaluation of the bicentered mean value form of f', the coefficients of f'' have to be computed. This is done only once in Step 1.
- As a side product of the evaluation of the the dense linear factor form of f on X, we obtain an overestimation of $f(c)$ and an overestimation of the set of slopes of f between x and c for all $x \in X$ and $c = \text{mid}(X)$. This set of slopes is used instead of $\overset{*}{M}_{\tilde{f}'}(X)$. Thus, in Step 4 the computation of $f(c)$ and $\overset{*}{M}_{\tilde{f}'}(X)$ is basically replaced by the evaluation of $\overset{*}{M}_{\tilde{f}}(X)$.

Finally, one should mention, that the cofficients of f' in Step 1 have to be computed by rounded interval arithmetic. Hence, for the computation of D in Step 4, we have to overestimate the range of an interval polynomial. The necessary modification of the algorithms is straight forward.

In Table 8.1 we report the number of floating point operations and iterations of Algorithm 8.1 (NEWTON), averaged over 1000 random input polynomials. The accuracy ε is 10^{-6} in all cases.

Dense

	Kflops	Iterations
$\check{H}_{f'}$	0.95	10.83
$\ddot{M}_{f'}$	1.61	12.08
$\overset{*}{M}_{\tilde{f}'}$	0.62	7.53
$\check{T}^*_{f'}$	2.67	10.68
$\check{B}^{(k)}_{f'}$	3.19	8.41
$\ddot{I}^{(s)}_{f'}$	1.91	9.56
$\tilde{I}^{(s)}_{f'}$	2.45	11.89

Sparse

	Kflops	Iterations
$\check{H}_{f'}$	1.32	30.21
$\ddot{M}_{f'}$	2.06	30.80
$\overset{*}{M}_{\tilde{f}'}$	1.52	25.45
$\check{T}^*_{f'}$	14.99	63.56
$\check{B}^{(k)}_{f'}$	9.14	28.71
$\ddot{I}^{(s)}_{f'}$	7.06	39.98
$\tilde{I}^{(s)}_{f'}$	10.11	58.09

Root Cluster

	Kflops	Iterations
$\check{H}_{f'}$	1224.54	13957.25
$\ddot{M}_{f'}$	99.01	686.76
$\overset{*}{M}_{\tilde{f}'}$	564.09	6596.54
$\check{T}^*_{f'}$	19.53	76.82
$\check{B}^{(k)}_{f'}$	20.68	54.57
$\ddot{I}^{(s)}_{f'}$	19.07	93.67
$\tilde{I}^{(s)}_{f'}$	18.91	91.84

Table 8.1: Newton's method for finding the roots of f in $[-1, 1]$ in dependence of the chosen interval extension for overestimating $f'(X)$. In each chart, the average of 1000 random polynomials of degree 10 is reported. In the left chart, the polynomials are dense, in the middle chart, the polynomials are sparse with only three non–zero coefficients, and in the right chart, the polynomials have a cluster of at least 5 roots in $[0.5, 0.6]$.

The results can be summarized and explained as follows.

- For dense random polynomials (left chart) the dense linear factor form ($M^{\overset{*}{\rightsquigarrow}}_{f'}$) is best. This is mainly due to the modification described above, i.e. instead of overestimating $f'(X)$ in Newton's method, merely an overestimation of the set of slopes of f between $\mathrm{mid}(X)$ and arbitrary points in X is used. This explains also the small number of iterations of the linear factor form.
- For sparse polynomials (middle chart), the Horner form with bisection at zero ($\check{H}_{f'}$) is comparable to $M^{\overset{*}{\rightsquigarrow}}_{f'}$. The reason is that $\check{H}_{f'}$ exploits the sparsity of f' for improving both efficiency and accuracy, whereas the accuracy of $M^{\overset{*}{\rightsquigarrow}}_{f'}$ is not affected, and the costs are reduced by a smaller amount compared to $\check{H}_{f'}$. In the dense case, the average costs per iteration are 88.9 flops for $\check{H}_{f'}$ and 82.3 flops for $M^{\overset{*}{\rightsquigarrow}}_{f'}$, whereas in the sparse case we obtain 48.6 flops for $\check{H}_{f'}$ and 59.7 flops for $M^{\overset{*}{\rightsquigarrow}}_{f'}$.
- The high accuracy of more expensive interval extensions pays only in the presence of root clusters (right chart). Here, most iterations are done when $\mathrm{w}(X)$ is very small. This explains the bad performance of the only linearly convergent Horner form. Further, it explains why the bicentered mean value form ($\ddot{M}_{f'}$) is now much better than $M^{\overset{*}{\rightsquigarrow}}_{f'}$: If $\mathrm{w}(X)$ is small, then the chances are high, that $0 \notin H_{f''}(X)$ and $\ddot{M}_{f'}$ is exact. The smallest number of iterations is achieved by the Bernstein form with bisection at zero ($\check{B}^{(k)}_{f'}$), but as it is more expensive than the interpolation forms and the dense Taylor form with bisection ($\check{T}^{*}_{f'}$), the methods perform equally well. It is interesting to note that $\check{T}^{*}_{f'}$, which is relatively inefficient for dense and sparse random polynomials, is one of the best in the presence of root clusters.

From our experimental observations we conclude that it is very important to choose an appropriate interval extension for overestimating $f'(X)$ in Newton's method. It seems that the following strategy can be recommended:

- If $\mathrm{w}(X)$ is large and f is dense, then use the dense linear factor form.
- If $\mathrm{w}(X)$ is large and f is sparse, then use the Horner form with bisection at zero.
- If $\mathrm{w}(X)$ is small and $\mathrm{mid}(X) \approx 0$ then use the Horner form with bisection at zero.
- If $\mathrm{w}(X)$ is small and $\mathrm{mid}(X) \not\approx 0$ then use the Bernstein form, the Taylor form with bisection or an interpolation form.

8.3 Global Optimization

In this section we compare interval extensions at the problem of finding the global minumum with given accuracy of a univariate polynomial f in an interval A. The problem is solved by an iterative algorithm, which evaluates an interval extension of f on a subinterval of A in each iteration. As in the case of Newton's method, the number of iterations and the cost of each iteration depend on the chosen interval extension.

Global optimization is one of the standard applications of interval arithmetic and was studied thoroughly in the literature, see for example [Skelboe, 1974], [Moore, 1976], [Ichida and Fujii, 1979], [Hansen, 1979], [Hansen, 1980], [Hansen and Sengupta, 1980], [Hansen, 1988], [Ratschek and Rokne, 1988], [Hansen, 1992], [Hammer *et al.*, 1993] and many others.

In order to faciliate the interpretation of the experimental results, we use a very simple optimization algorithm and omit all kinds of accelerating devices. In the experiments below, we compute the global minumum of random polynomials in $[-1, 1]$. We consider the same classes of random polynomials as in the previous chapter on Newtons method, but multiply each polynomial by $x^2 - 1$. Without this multiplication, the probability that f achieves its minmimum at the boundary of X is very high, and the experimental results might not be so meaningful.

The comparison is done at the following optimization algorithm, where F denotes an interval extension of f. The algorithm works by splitting the input interval A into two subintervals of same size in each iteration. The search for the minimum of f is done in these subintervals, but not uniformly which would be too expensive. Here, the branch and bound principle is involved. At each iteration, the search is continued in the subinterval X, where $\underline{F(X)}$ is smallest, since the chances are best for finding the minimum of f in this interval. The algorithm stops if $\mathrm{w}(F(X)) < \varepsilon$, where ε is the desired accuracy of the result, or if X cannot be bisected any more because of finite precision arithmetic. Thus, the central data structure of the algorithm is a list $\mathcal{L}$ of pairs of intervals $(X, F(X))$, which is sorted by $\underline{F(X)}$ in increasing order.

Algorithm 8.2 (MIN) [Global Minimum of Univariate Polynomials]

In: $f(x) \in \mathbb{F}[x]$,
$A \in \mathbb{IF}$,
$\varepsilon \in \mathbb{F}$, $\varepsilon > 0$.

Out: $Y \in \mathbb{IF}$ such that

- $\min_{x \in A} f(x) \in Y$
- $\mathrm{w}(Y) \leq \varepsilon$ if the precision of the floating point numbers is sufficiently high.

(1) [Initialize.]
$\mathcal{L} \longleftarrow \langle (A, F(A)) \rangle$

(2) [Iterate.]
pop (X, Y) from $\mathcal{L}$

(3) [Termination.]
if WIDTH$(Y) \leq \varepsilon$ then return Y
$c \longleftarrow$ MID(X)
if $c \notin \mathrm{int}(X)$ then return Y

(4) [Bisect.]
$X_1 \longleftarrow [\underline{X}, c]$
$X_2 \longleftarrow [c, \overline{X}]$

(5) [Insert.]
insert $(X_1, F(X_1))$ and $(X_2, F(X_2))$ in $\mathcal{L}$ such that the elements (X, Y) in $\mathcal{L}$ are sorted with increasing $\underline{Y}$.
goto Step 2

Algorithm 8.2 is modified slightly in dependence of the interval extension which is used to bound $f(X)$:

- The bicentered mean value form requires the computation of the coefficients of f'. Apparently, this should be done only once in Step 1. Further, instead of $\ddot{M}_f(X)$, the smaller interval $[\underline{M_f(X, c^{\downarrow})}, H_f(c^{\downarrow})]$ is used in Step 1 and Step 5. This saves the computation of $c^{\uparrow}$ and $H_f(c^{\uparrow})$, see Section 4.2.

- As a side product during the evaluation of $M_{\tilde{f}}^{*}(X)$, $\check{T}_f^{*}(X)$ and $\ddot{I}_f^{(\mathrm{s})}(X)$ we obtain an overestimation Y of $f(\mathrm{mid}(X))$. Thus, in Step 1 and Step 5 we use the smaller interval $[\underline{F(X)}, \overline{Y}]$ instead of $F(X)$, where F is $M_{\tilde{f}}^{*}$, $\check{T}_f^{*}$ and $\ddot{I}_f^{(\mathrm{s})}$ respectively.

- Similarly, during the evaluation of $B_f^{(k)}(X)$ and $\tilde{I}_f^{(\mathrm{s})}(X)$ we obtain overestimations L, R of $f(\underline{X})$ respectively $f(\overline{X})$. Thus, in Step 1 and Step 5 we use the smaller interval $[\underline{F(X)}, \min\{\overline{L}, \overline{R}\}]$ instead of $F(X)$, where F is $B_f^{(k)}$ and $\ddot{I}_f^{(\mathrm{s})}$ respectively.

In Table 8.2 we report the number of floating point operations and iterations of Algorithm 8.2 (MIN), averaged over 1000 random input polynomials. The accuracy ε is 10^{-3} in all cases. It does not seem to be reasonable to choose a higher accuracy unless quadratically convergent accelerators like Newton's method are used.

The results can be summarized and explained as follows.

Dense

	Kflops	Iterations
$\check{H}_f$	10.57	107.90
$\ddot{M}_f$	2.45	11.98
$M_f^{\overset{*}{\rightsquigarrow}}$	3.07	15.83
$\check{T}_f^*$	7.61	11.68
$B_f^{(k)}$	5.25	5.53
$\ddot{I}_f^{(\mathrm{s})}$	3.06	8.79
$\tilde{I}_f^{(\mathrm{s})}$	2.20	6.40

Sparse

	Kflops	Iterations
$\check{H}_f$	5.78	86.33
$\ddot{M}_f$	1.64	11.56
$M_f^{\overset{*}{\rightsquigarrow}}$	2.56	15.50
$\check{T}_f^*$	6.95	10.75
$B_f^{(k)}$	5.17	5.44
$\ddot{I}_f^{(\mathrm{s})}$	2.99	8.61
$\tilde{I}_f^{(\mathrm{s})}$	2.15	6.25

Root Cluster

	Kflops	Iterations
$\check{H}_f$	73.22	740.69
$\ddot{M}_f$	6.31	29.44
$M_f^{\overset{*}{\rightsquigarrow}}$	4.22	21.32
$\check{T}_f^*$	7.33	11.26
$B_f^{(k)}$	6.07	6.22
$\ddot{I}_f^{(\mathrm{s})}$	3.25	9.24
$\tilde{I}_f^{(\mathrm{s})}$	2.92	8.17

Table 8.2: Computation of the global minimum of f in $[-1, 1]$ in dependence of the chosen interval extension for overestimating $f(X)$. In each chart, the average of 1000 random polynomials of degree 12 is reported. In the left chart, the polynomials are dense, in the middle chart, the polynomials are sparse with at most 6 non–zero coefficients, and in the right chart, the polynomials have a cluster of at least 5 roots in $[0.5, 0.6]$. Every polynomial vanishes at -1 and 1.

- For dense random polynomials (left chart) the parabolic boundary value form $\tilde{I}_f^{(\mathrm{s})}$ is best, followed by the bicentered mean value form $\ddot{M}_f$. The good performance of $\ddot{M}_f$ is mainly due to the modification described above. The Horner form with bisection at zero $\check{H}_f$ is significantly worse than the other methods. Whereas all methods could be improved because an upper bound of f at the center or the boundary of X was available, this was not possible for the Horner form.
- For sparse random polynomials (middle chart), $\ddot{M}_f$ is better than $\tilde{I}_f^{(\mathrm{s})}$, because the former exploits sparsity to improve both efficiency and accuracy, whereas the latter does not. This explains also, why $\check{H}_f$ is much better compared to the dense case, whereas all other methods remain basically the same.
- In the presence of a root cluster, interpolation forms are best. A closer analysis shows that this is only true if the global minimum is near the root cluster. In this case many iterations are executed when the width of X is small, and the cubic convergence order of interpolation forms pays.

Note that the smallest number of iterations is achieved by the Bernstein form in all cases. However, as the evaluation of the Bernstein form is relatively expensive, its overall performance is not good. Further, the experiments show that $\check{H}_f$ is not a good choice for global optimization problems. Not only the number of floating point operations is large, but each iteration requires insertion of elements in a sorted list, which causes additional overhead.

Bibliography

[Alefeld and Herzberger, 1974]
G. Alefeld and J. Herzberger.
Einführung in die Intervallrechnung.
Bibliographisches Institut, Mannheim–Wien–Zürich, 1974.

[Alefeld and Herzberger, 1983]
G. Alefeld and J. Herzberger.
Introduction to Interval Computations.
Academic Press, N.Y., 1983.

[Alefeld and Rokne, 1981]
G. Alefeld and J. Rokne.
On the Evaluation of Rational Functions in Interval Arithmetic.

SIAM J. Numerical Analysis, 18:862–870, 1981
Keywords: Range Computation, Centered Forms.

[Alefeld, 1970]
G. Alefeld.
Eine Modifikation des Newtonverfahrens zur Bestimmung der reellen Nullstellen einer reellen Funktion.
ZAMM, 50:T32–T33, 1970
Keywords: Nonlinear Equation Systems.

[Alefeld, 1981]
G. Alefeld.
Bounding the Slope of Polynomial Operators and some Applications.
Computing, 26:227–237, 1981
Keywords: Nonlinear Equation Systems, Bounding Solutions.

[Apostolatos and Kulisch, 1967]
N. Apostolatos and U. Kulisch.
Grundlagen einer Maschinenintervallarithmetik.
Computing, 2:89–101, 1967
Keywords: Implementation.

[Baumann, 1988]
E. Baumann.
Optimal Centered Forms.
BIT, 28:80–87, 1988
Keywords: Range Computation, Centered Forms.

[Caprani and Madsen, 1980]
O. Caprani and K. Madsen.
Mean Value Forms in Interval Analysis.
Computing, 25:147–154, 1980
Keywords: Range Computation, Centered Forms.

[Cargo and Shisha, 1966]
G. T. Cargo and O. Shisha.
The Bernstein Form of a Polynomial.
Journal of Research of Nat. Bur. Standards, 70B:79–81, 1966
Keywords: Range Computation, Bernstein Form.

[Cornelius and Lohner, 1984]
H. Cornelius and R. Lohner.
Computing the Range of Values of Real Functions with Accuracy Higher than Second Order.
Computing, 33:331–347, 1984
Keywords: Range Computation, Interpolation Form.

[Garloff, 1985]
J. Garloff.
Convergent Bounds for the Range of Multivariate Polynomials.
In K. Nickel, editor, *Interval Mathematics*, pages 37–56. Springer, 1985
Keywords: Range Computation, Bernstein Form.

[Grassmann and Rokne, 1979]
E. Grassmann and J. Rokne.
The Range of Values of a Circular Complex Polynnomial over a Circular Complex Interval.
Computing, 23:139–169, 1979
Keywords: Range Computation, Bernstein Form.

[Hammer *et al.*, 1993]
R. Hammer, M. Hocks, U. Kulisch, and D. Ratz.
Numerical Toolbox for Verified Computing I.
Springer, 1993.

[Hansen and Sengupta, 1980]
E. R. Hansen and S. Sengupta.
Global Constrained Optimization using Interval Analysis.
In K. Nickel, editor, *Interval Mathematics*, pages 25–48. Springer, 1980
Keywords: Global Optimization.

[Hansen and Sengupta, 1981]
E. Hansen and S. Sengupta.
Bounding Solutions of Systems of Equations using Interval Analysis.
BIT, 21:203–211, 1981
Keywords: Nonlinear Equation Systems, Bounding Solutions.

[Hansen, 1969]
E. Hansen.
On the Centered Form.
In E. Hansen, editor, *Topics in Interval Analysis*, pages 102–106, Clarendon Press, Oxford, 1969
Keywords: Range Computation, Centered Forms.

[Hansen, 1978a]
E. Hansen.
A Globally Convergent Interval Method for Computing and Bounding Real Roots.
BIT, 18:415–424, 1978
Keywords: Nonlinear Equation Systems, Bounding Solutions.

[Hansen, 1978b]
E. Hansen.
Interval Forms of Newtons Method.
BIT, 20:153–163, 1978
Keywords: Range Computation, Nonlinear Equation Systems.

[Hansen, 1979]
E. Hansen.
Global Optimization using Interval Analysis: The One-Dimensional Case.
Optimization Theory and Applications, 29:331–344, 1979
Keywords: Global Optimization.

[Hansen, 1980]
E. Hansen.
Global Optimization using Interval Analysis: The Multidimensional Case.
Numerische Mathematik, 34:247–270, 1980
Keywords: Global Optimization.

[Hansen, 1988]
E. Hansen.
An Overview of Global Optimization using Interval Analysis.
In R. E. Moore, editor, *Reliability in Computing*, pages 289–307, Academic Press, London, 1988
Keywords: Range Computation.

[Hansen, 1992]
E. Hansen.
Global Optimization Using Interval Analysis.
Marcel Dekker, Inc., 1992.

[Hong and Stahl, 1994a]
H. Hong and V. Stahl.
Bernstein Form is Inclusion Monotone.
Technical Report 94-52, Research Institute for Symbolic Computation, Johannes Kepler University, Linz, Austria, 1994.

[Hong and Stahl, 1994b]
H. Hong and V. Stahl.
Safe Starting Regions by Fixed Points and Tightening.

Technical Report 94-29, Research Institute for Symbolic Computation, Johannes Kepler University, Linz, Austria, 1994.

[Householder, 1970]
A. S. Householder.
The Numerical Treatment of a Single Non-Linear Equation.
Mc Graw-Hill, 1970.

[Ichida and Fujii, 1979]
K. Ichida and Y. Fujii.
An Interval Arithmetic Method for Global Optimization.
Computing, 23:85–97, 1979
Keywords: Range Computation, Iterative Methods.

[Krawczyk and Neumaier, 1985]
R. Krawczyk and A. Neumaier.
Interval Slopes for Rational Functions and Associated Centered Forms.
SIAM J. Numerical Analysis, 22(3):604–615, 1985
Keywords: Range Computation, Centered Forms.

[Krawczyk, 1969]
R. Krawczyk.
Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken.
Computing, 4:187–201, 1969
Keywords: Nonlinear Equation Systems, Bounding Solutions.

[Lane and Riesenfeld, 1981]
J. M. Lane and R. F. Riesenfeld.
Bounds on a Polynomial.
BIT, 21:112–117, 1981
Keywords: Range Computation, Bernstein Form.

[Moore, 1966]
R. E. Moore.
Interval Anaysis.
Prentice-Hall, Englewood Cliffs, N.J., 1966.

[Moore, 1976]
R. E. Moore.
On Computing the Range of a Rational Function of n Variables over a Bounded Region.
Computing, 16:1–15, 1976
Keywords: Range Computation, Iterative Methods.

[Neumaier, 1990]
A. Neumaier.
Interval Methods for Systems of Equations.
Cambridge University Press, 1990.

[Rall, 1983]
L. B. Rall.
Mean Value and Taylor Forms in Interval Analysis.
SIAM, J. Math. An., 14:223–238, 1983
Keywords: Range Computation, Centered Forms, Nonlinear Equation Systems, Bounding Solutions.

[Ratschek and Rokne, 1980a]
H. Ratschek and J. Rokne.
About the Centered Form.
SIAM J. Numerical Analysis, 17(3):333–337, 1980
Keywords: Range Computation, Centered Forms.

[Ratschek and Rokne, 1980b]
H. Ratschek and J. Rokne.
Optimality of the Centered Form.

In K. Nickel, editor, *Interval Mathematics*, pages 499–508. Springer, 1980
Keywords: Range Computation, Centered Forms.

[Ratschek and Rokne, 1984]
H. Ratschek and J. Rokne.
Computer Methods for the Range of Functions.
Ellis Horwood Limited, 1984.

[Ratschek and Rokne, 1988]
H. Ratschek and J. Rokne.
New Computer Methods for Global Optimization.
Ellis Horwood Limited, 1988
Keywords: Global Optimization.

[Ratschek and Schröder, 1981]
H. Ratschek and G. Schröder.
Centered Forms for Functions in Several Variables.
Mathematical Analysis and Applications, 82:543–552, 1981
Keywords: Range Computation, Centered Forms.

[Ratschek, 1978]
H. Ratschek.
Zentrische Formen.
ZAMM, 58:434–436, 1978
Keywords: Range Computation, Centered Forms.

[Ratschek, 1980a]
H. Ratschek.
Centered Forms.
SIAM J. Numerical Analysis, 17(5):656–662, 1980
Keywords: Range Computation, Centered Forms.

[Ratschek, 1980b]
H. Ratschek.
Optimal Approximations in Interval Analysis.
In K. Nickel, editor, *Interval Mathematics*, pages 181–202. Springer, 1980
Keywords: Range Computation, Centered Forms.

[Rivlin, 1970]
T. J. Rivlin.
Bounds on a Polynomial.
Journal of Research of Nat. Bur. Standards, 74B:47–54, 1970
Keywords: Range Computation, Bernstein Form.

[Rokne and Wu, 1982]
J. G. Rokne and T. Wu.
The Circular Complex Centered Form.
Computing, 28:17–30, 1982
Keywords: Range Computation, Centered Forms.

[Rokne and Wu, 1983]
J. G. Rokne and T. Wu.
A Note on the Circular Complex Centered Form.
Computing, 30:201–211, 1983
Keywords: Range Computation, Centered Forms.

[Rokne, 1977]
J. G. Rokne.
Bounds for an Interval Polynomial.
Computing, 18:225–240, 1977
Keywords: Range Computation, Bernstein Form.

[Rokne, 1979a]
J. G. Rokne.
A note on the Bernstein Algorithm for Bounds for Interval Polynomials.
Computing, 21:159–170, 1979
Keywords: Range Computation, Bernstein Form.

[Rokne, 1979b]
J. G. Rokne.
The Range of Values of a Complex Polynomial over a Complex Interval.
Computing, 22:153–169, 1979
Keywords: Range Computation, Bernstein Form.

[Rokne, 1981]
J. G. Rokne.
The Centered Form for Interval Polynomials.
Computing, 27:339–348, 1981
Keywords: Range Computation, Centered Form.

[Rokne, 1982]
J. G. Rokne.
Optimal Computation of the Bernstein Algorithm for the Bound of an Interval Polynomial.
Computing, 28:239–246, 1982
Keywords: Range Computation, Bernstein Form.

[Skelboe, 1974]
S. Skelboe.
Computation of Rational Interval Functions.
BIT, 14:87–95, 1974
Keywords: Range Computation, Iterative Methods.