

# UC Davis

## UC Davis Previously Published Works

### Title

Computing the roots of sparse high-degree polynomials that arise from the study of random simplicial complexes

### Permalink

<https://escholarship.org/uc/item/1708692f>

### Journal

Numerical Algorithms, 83(4)

### ISSN

1017-1398

### Authors

Farouki, Rida T  
Strom, Jeffrey A

### Publication Date

2020-04-01

### DOI

10.1007/s11075-019-00745-3

Peer reviewed

# Computing the roots of sparse high-degree polynomials that arise from the study of random simplicial complexes

Rida T. Farouki

Department of Mechanical and Aerospace Engineering,  
University of California, Davis, CA 95616, USA

Jeffrey A. Strom

Department of Mathematics,  
Western Michigan University, Kalamazoo, MI 49008, USA

## Abstract

The problem of computing the roots of a particular sequence of sparse polynomials  $p_n(t)$  is considered. Each instance  $p_n(t)$  incorporates only the  $n+1$  monomial terms  $t, t^2, t^4, \dots, t^{2^n}$  associated with the binomial coefficients of order  $n$  and alternating signs. It is shown that  $p_n(t)$  has (in addition to the obvious roots  $t = 0$  and  $1$ ) precisely  $n - 1$  simple roots on the interval  $(0, 1)$  with no roots greater than  $1$ , and a recursion relating  $p_n(t)$  and  $p_{n+1}(t)$  is used to show that they possess interlaced roots. Closed-form expressions for the Bernstein coefficients of  $p_n(t)$  on  $[0, 1]$  are derived, and employed to compute the roots in double-precision arithmetic. Despite the severe variation of the graph of  $p_n(t)$ , and tight clustering of roots near  $t = 1$ , it is shown that for  $n \leq 10$  the roots on  $(0, 1)$  are remarkably well-conditioned, and can be computed to high accuracy using both the power and Bernstein forms.

Keywords: random simplicial complex, Euler characteristic, polynomial roots, Descartes' Law of Signs, interlaced roots, Bernstein basis, condition number.

e-mail addresses: farouki@ucdavis.edu, jeff.strom@wmich.edu

# 1 Introduction

The accurate and efficient computation of the real roots of polynomials is a fundamental problem in numerical analysis, of great importance in numerous applications. Although this problem has been intensively investigated since the advent of digital computers, a “universal” root-finding algorithm based on floating-point arithmetic, that infallibly computes to a specified precision the roots of polynomials with arbitrary coefficients, remains an elusive ideal. A primary factor in this state of affairs is the high variability observed in the sensitivity of polynomial roots to perturbations in the polynomial coefficients, and the sometimes counter-intuitive trends of this sensitivity.

For polynomials with exactly-specified integer coefficients, the vagaries of floating-point arithmetic can be circumvented by invoking arbitrary-precision rational arithmetic to determine isolating intervals for each root, smaller than any prescribed width. However, this approach can become very demanding in computing time and memory requirements, and for many practical problems the initial polynomial coefficients will not be exactly known.

Floating-point arithmetic, based on rounding or truncation of significant digits to ensure a fixed memory size for numbers, has become widely adopted in scientific computing since the 1960s. In the late 1950s, it was implemented in software rather than in the cpu, and the two-part paper [13] by the British numerical analyst J. H. Wilkinson describes one of its earliest applications in computing polynomial roots. Subsequently [15], he characterized his choice of an “easy” test polynomial (with the 20 equally-spaced real roots  $1, 2, \dots, 20$ ) as *the most traumatic experience in my career as a numerical analyst*, due to the extreme difficulty in computing any accurate digits for most of the roots. The polynomial with the 20 roots  $2^{-1}, 2^{-2}, \dots, 2^{-20}$  in geometric progression is also considered in [13], and despite the close proximity of successive roots, it was found that they could be computed to a high accuracy.

The choice of basis plays an important role in determining the sensitivity of the roots of a polynomial to uncertainties in its coefficients, and it is known that the usual monomial (power) form  $a_0 + a_1t + \dots + a_mt^m$  is not, in general, a favorable choice. For computing the real roots on a finite interval  $t \in [a, b]$  the Bernstein form is preferred, and in fact it is “optimally stable” among all bases that are non-negative over that interval [4].

In the present paper, we consider the roots of a sequence of sparse high-degree polynomials that arise from the study of random simplicial complexes (see Section 2), wherein each instance  $p_n(t)$  includes only the  $n + 1$  monomial

terms  $t^k$  for  $k = 1, 2, \dots, 2^n$  with binomial coefficients of alternating sign. Because of the probabilistic interpretation of the independent variable  $t$ , the main focus is on real roots  $t \in [0, 1]$  although we also briefly discuss positive roots  $> 1$ , negative roots, and complex conjugate root pairs.

The remainder of this paper is organized as follows. Section 2 introduces the topological problem that motivates the study of the positive roots of the family of sparse high-degree polynomials considered herein. In Section 3, we elucidate some of the distinctive properties of the roots of these polynomials. Section 4 then considers numerical computation of the roots on  $(0, 1)$  using double-precision floating-point arithmetic for  $n \leq 10$ , based upon the power and Bernstein representations, and Section 5 discusses the accuracy of the computed roots in relation to the root condition numbers. Finally, Section 6 summarizes the principal results of the present study.

## 2 Euler characteristic of random simplicial complexes

We begin by explaining how the sequence of polynomials investigated herein arises quite naturally in the context of a fundamental question concerning the homotopy type of randomly constructed simplicial complexes.

### 2.1 Random simplicial complexes

Erdős and Rényi introduced random graphs [1] in 1959. In the years since their introduction, random graphs have proven to be an extremely powerful tool in graph theory, and their study has opened up a huge collection of new research directions. The most useful and deeply-studied probability model starts with  $n$  vertices and a fixed probability  $t$ . Then, for each pair of vertices  $u$  and  $v$ , we flip a “ $t$ -coin” (that comes up heads with probability  $t$ ) to decide whether or not they should be connected by an edge.

Even though, to a topologist, a graph is just a 1-dimensional simplicial complex, the study of random simplicial complexes has attracted substantial attention only in the past 15 years or so. The generalization from the random models of Erdős and Rényi to probability models for random simplicial complexes is straightforward. Recall that an  $n$ -dimensional simplex  $\Delta^n$  is the convex hull of a set of  $n + 1$  points  $\{\mathbf{x}_1, \dots, \mathbf{x}_{n+1}\}$  in general position in  $\mathbb{R}^m$  (with  $m \geq n$ ). The points  $\mathbf{x}_i$  are called the *vertices* of the simplex. Any

subset with  $k+1$  of these vertices defines a subsimplex of dimension  $k$ . Thus, there are  $\binom{n+1}{k+1}$   $k$ -dimensional simplices in  $\Delta^n$ , and a total of  $2^{n+1}$  simplices, including the empty simplex. Combinatorially, a simplicial complex  $K$  of  $\Delta^n$  is simply the union of some (perhaps all) of the simplices of  $\Delta^n$ , with the proviso that if a simplex is in  $K$ , then so are all its subsimplices.

Several probability models for constructing random simplicial complexes have been studied, which are all specializations of the following extremely general procedure. List all the non-empty simplices of  $\Delta^n$  with an ordering  $\sigma_1, \dots, \sigma_{2^{n+1}-1}$  such that their dimensions are weakly increasing, and for each simplex  $\sigma$ , assign a probability  $p_\sigma \in [0, 1]$ . Then, beginning with  $K = \emptyset$  and  $\sigma = \sigma_1$ , check if all the subsimplices of  $\sigma$  are already in  $K$ . If they are, then  $\sigma$  is “available for inclusion into  $K$ ” and we flip a coin that comes up heads with probability  $p_\sigma$  to decide whether or not to include it. Once that decision is made, we consider the next simplex in our list; the construction ends once all of the simplices of  $\Delta^n$  have been considered.

We focus henceforth on the specialization of this general model in which we choose a fixed value  $t \in [0, 1]$  and set  $p_\sigma = t$  for all simplices  $\sigma$  of  $\Delta^n$ .

## 2.2 Euler characteristic

A topologist is naturally interested in the homotopy-theoretical properties of random simplicial complexes. The most fundamental question is whether or not these spaces are contractible (i.e, homotopically trivial). One way to distinguish the homotopy types of spaces is through the *Euler characteristic*, defined as

$$\chi(K) := \sum_{k=0}^{\dim(K)} (-1)^k \# \{k\text{-simplices in } K\}.$$

If  $K$  and  $L$  are homotopy equivalent, then  $\chi(K) = \chi(L)$ , and since  $\chi(K) = 1$  if  $K$  is a single point, we see that  $K$  cannot be contractible when  $\chi(K) \neq 1$ . Let the random variable  $\chi(n)$  be the Euler characteristic of the complex  $K$  constructed by the above procedure, with expected value  $E(\chi(n))$ .

**Proposition 1.**  $E(\chi(n)) = -\frac{1}{t} \sum_{j=1}^{n+1} (-1)^j \binom{n+1}{j} t^{2^j}.$

This follows directly from the discussion above, together with the linearity of expected value. Straightforward manipulation of the equation  $E(\chi(n)) = 1$  yields the following result.

**Corollary 1.** *The “expected random subcomplex” of  $\Delta^{n-1}$  is not contractible if  $t$  is not a root of the polynomial*

$$p_n(t) := \sum_{k=0}^n (-1)^k \binom{n}{k} t^{2^k}. \quad (1)$$

**Motivation for Studying the Roots of  $p_n(t)$ .** Corollary 1 indicates that the random process generally produces homotopically non-trivial complexes — with possible exceptions when the probability  $t$  is one of the finite number of real roots of the polynomial  $p_n(t)$  in the interval  $[0, 1]$ . Furthermore, our study of these polynomials suggests that they always change sign at the roots, which means that the algebraic topology of the random complexes may have a phase transition or threshold of some kind at these particular  $t$  values. The study of thresholds separating different kinds of behavior is a major strand in the well-established theory of random graphs.

### 3 Anatomy of the polynomials

We are interested in the roots of the sparse polynomials  $p_n(t)$  of degree  $2^n$  defined for  $n \geq 1$  by (1). The first few instances of these polynomials are

$$\begin{aligned} p_1(t) &= t - t^2, \\ p_2(t) &= t - 2t^2 + t^4, \\ p_3(t) &= t - 3t^2 + 3t^4 - t^8, \\ p_4(t) &= t - 4t^2 + 6t^4 - 4t^8 + t^{16}, \\ p_5(t) &= t - 5t^2 + 10t^4 - 10t^8 + 5t^{16} - t^{32}, \\ p_6(t) &= t - 6t^2 + 15t^4 - 20t^8 + 15t^{16} - 6t^{32} + t^{64}, \end{aligned}$$

and one can easily verify that they satisfy the simple recursion relation

$$p_{n+1}(t) = p_n(t) - p_n(t^2). \quad (2)$$

For all  $n \geq 1$ ,  $t = 0$  is obviously a root of  $p_n(t)$ , and  $t = 1$  is also a root, since

$$p_n(1) = \sum_{k=0}^n (-1)^k \binom{n}{k} = 0.$$

These are simple roots of  $p_n(t)$ , since  $p'_n(0) = 1$  and  $p'_n(1) = (-1)^n$ . Thus, the factor  $(1 - t)t$  can be extracted from  $p_n(t)$  to obtain a polynomial  $\tilde{p}_n(t)$  of the modestly lower degree  $2^n - 2$ , although  $\tilde{p}_n(t)$  is no longer sparse. From an abstract algebraic point of view, we have verified using **Maple** that  $\tilde{p}_n(t)$  is irreducible over the rational numbers for  $2 \leq n \leq 10$ .

Clearly  $t = 0$  and  $t = 1$  are the only roots of  $p_1(t)$ , while  $p_2(t)$  has the additional real roots  $t = \frac{1}{2}(-1 \pm \sqrt{5})$ , one negative and one positive. Using **Maple**, we have verified that the Galois group of the splitting field  $E$  of  $\tilde{p}_3(t)$  is isomorphic to the full symmetric group  $\text{Sym}(6)$ , and since this group is not solvable, the roots of  $p_3(t)$  (apart from 0 and 1) cannot be expressed in terms of radicals. It seems likely that this is also true for all  $n \geq 3$ , although the degree of  $p_4(t)$  is already too high for **Maple** to cope with.

Since the polynomial (1) has  $n$  sign changes in its coefficients, Descartes' Law of Signs [12] indicates that the number of its positive real roots is less than  $n$  by an even amount. We show below that  $p_n(t)$  has exactly  $n$  positive roots ( $n - 1$  roots on  $(0, 1)$  together with the root  $t = 1$ ) and identify an “interlacing” property of the roots for successive  $n$  values, i.e., the roots of  $p_{n+1}(t)$  on  $(0, 1)$  lie within intervals delineated by the roots of  $p_n(t)$ .

**Lemma 1.** *For  $n \geq 1$  the polynomials (1) have at most  $n + 1$  roots on  $[0, 1]$  and their derivatives have at most  $n$  roots on  $(0, 1)$ .*

*Proof.* Since  $p_n(t)$  exhibits  $n$  sign changes in its coefficients, Descartes' Law of Signs indicates that it has at most  $n$  positive roots, including  $t = 1$ . Since  $t = 0$  is also a root of  $p_n(t)$ , it has at most  $n + 1$  roots on  $[0, 1]$ . The derivative of  $p_n(t)$ , namely

$$p'_n(t) = \sum_{k=0}^n (-1)^k \binom{n}{k} 2^k t^{2^k - 1},$$

also exhibits  $n$  coefficient sign changes, so  $p'_n(t)$  has at most  $n$  roots on  $(0, 1)$ . Note that 0 and 1 are not roots, since  $p'_n(0) = 1$  and  $p'_n(1) = (-1)^n$ . ■

**Proposition 2.** *Consider a polynomial  $f(t)$  satisfying*

- *$f(t)$  has only simple roots on  $[0, 1]$  including 0 and 1;*
- *$r < s^2$  for consecutive root pairs  $r, s$  of  $f(t)$  on  $[0, 1]$ ;*
- *$f'(t)$  has exactly one root between  $r$  and  $s$ .*

Then the polynomial  $g(t) = f(t) - f(t^2)$  inherits the following properties

- $g(t)$  has one simple root  $w$  between the roots  $r, s$  of  $f(t)$ ;
- the root  $w$  of  $g(t)$  actually lies between  $\sqrt{r}$  and  $s$ ;
- any two consecutive roots  $v, w$  of  $g(t)$  satisfy  $v < w^2$ .

*Proof.* Let  $q, r, s$  be consecutive roots of  $f(t)$  on  $t \in [0, 1]$ . Then  $r < \sqrt{r} < s$  and  $q < r^2 \leq t^2$  for  $r \leq t \leq s$ . Now the values of  $f(t)$  on  $[r, s]$  are replicated by  $f(t^2)$  on  $[\sqrt{r}, \sqrt{s}]$ . In particular,  $f(t)$  and  $f(t^2)$  must have opposite signs on  $[r, \sqrt{r}]$  and  $[s, \sqrt{s}]$  so  $g(t)$  cannot vanish on these intervals (see Figure 1). For simplicity, we assume that  $f(t) < 0$  for  $r < t < s$  (an analogous argument holds when  $f(t) > 0$ ). Then  $f(t) < 0 = f(t^2)$  at  $t = \sqrt{r}$ , and  $0 = f(t) > f(t^2)$  at  $t = s$ . Thus, since  $g(t) < 0$  at  $t = \sqrt{r}$  and  $g(t) > 0$  at  $t = s$ , the interval  $(\sqrt{r}, s)$  must contain a root  $w$  of  $g(t)$ . Moreover, since

$$g'(t) = f'(t) - 2t f'(t^2),$$

and  $f'(t) > 0$  and  $f'(t^2) < 0$  at  $t = w$  when  $f(t) < 0$  for  $r < t < s$ , we have  $g'(w) > 0$ , so  $w$  is a simple root of  $g(t)$ .

Finally, consecutive roots  $v, w$  of  $g(t)$  lying between successive root pairs  $q, r$  and  $r, s$  of  $f(t)$  must satisfy

$$q < \sqrt{q} < v < r < \sqrt{r} < w < s,$$

and consequently it is clear that  $v < w^2$ . ■

**Theorem 1.** *For  $n \geq 1$ , the root structure of the polynomials (1) on  $[0, 1]$  may be characterized as follows:*

- $p_n(t)$  has exactly  $n+1$  simple roots on  $[0, 1]$  including  $t = 0$  and  $t = 1$ ;
- the roots exhibit a “super-quadratic distribution” — i.e.,  $r < s^2$  for any pair of consecutive roots  $r, s$  of  $p_n(t)$ .

*Proof.* We argue by induction. First, the stated properties are easily verified by direct calculation for small  $n$  values. Now suppose that they hold for some larger  $n$ . Under this supposition, Lemma 1 indicates that  $p_n(t)$  has exactly  $n$  extrema on  $(0, 1)$  that isolate its  $n+1$  roots on  $[0, 1]$ . Consequently,  $p_n(t)$  satisfies the conditions stipulated for the polynomial  $f(t)$  in Proposition 2,



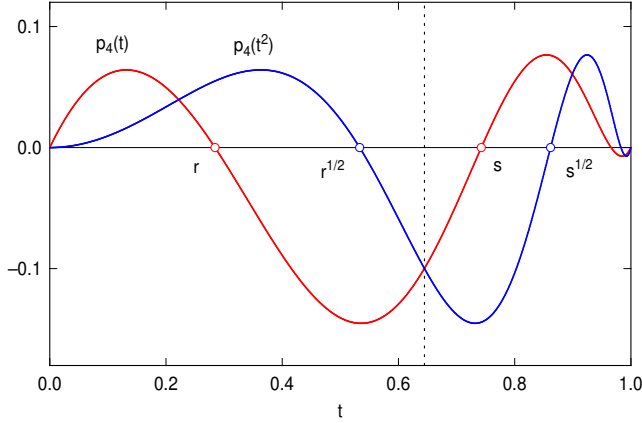


Figure 1: The graphs (red) of  $p_4(t)$  and (blue) of  $p_4(t^2)$ , showing consecutive roots  $r, s$  and  $\sqrt{r}, \sqrt{s}$ . Observe that  $p_4(t), p_4(t^2)$  have opposite signs on the intervals  $[r, \sqrt{r}]$  and  $[s, \sqrt{s}]$  but the same sign on  $[\sqrt{r}, s]$ , which contains a root of  $p_5(t)$ , indicated by the dashed line, where the graphs of  $p_4(t), p_4(t^2)$  cross. Note also that  $p_4(t), p_4(t^2)$  have derivatives of opposite sign on  $[\sqrt{r}, s]$ .

and  $p_{n+1}(t)$  — as defined by the recursion relation (2) — has the properties of the polynomial  $g(t)$  in Proposition 2. This implies that  $p_{n+1}(t)$  has  $n + 2$  simple roots on  $[0, 1]$  (including  $t = 0$  and  $t = 1$ ) with a super-quadratic distribution, i.e.,  $v < w^2$  for any pair of consecutive roots  $v, w$  of  $p_{n+1}(t)$ . This completes the induction proof. ■

As a direct consequence of Theorem 1, we make the following observations.

**Remark 1.** By Descartes' Law of Signs, the polynomials  $p_n(t)$  have no real roots for  $t > 1$ .

**Corollary 2.** *The  $n$  roots  $t_{n+1,1}, \dots, t_{n+1,n}$  of  $p_{n+1}(t)$  on  $(0, 1)$  interlace the  $n - 1$  roots  $t_{n,1}, \dots, t_{n,n-1}$  of  $p_n(t)$  on  $(0, 1)$  — i.e., with  $t_{n,0} = 0$  and  $t_{n,n} = 1$  they satisfy*

$$t_{n,k-1} < t_{n+1,k} < t_{n,k} \quad \text{for } k = 1, \dots, n. \quad (3)$$

*The ordering (3) is just a weaker form of the relation  $\sqrt{t_{n,k-1}} < t_{n+1,k} < t_{n,k}$  for  $k = 1, \dots, n$  established in Theorem 1.*

The remarkable behavior of the graph of  $p_n(t)$ , illustrated in Figure 2 for  $t \in [0, 1]$  in the case  $n = 10$ , makes it an interesting test case for polynomial

root solvers. Using a linear scale in  $t$ , only the first 6 of its 9 roots on  $(0, 1)$  are clearly discernible. As  $n$  increases, the graph of  $p_n(t)$  on  $[0, 1]$  oscillates with increasing frequency but decreasing amplitude between negative and positive values as  $t \rightarrow 1$ , and the roots become tightly clustered. For  $t > 1$ ,  $p_n(t)$  rapidly diverges to  $+\infty$  or  $-\infty$  according to whether  $n$  is even or odd.

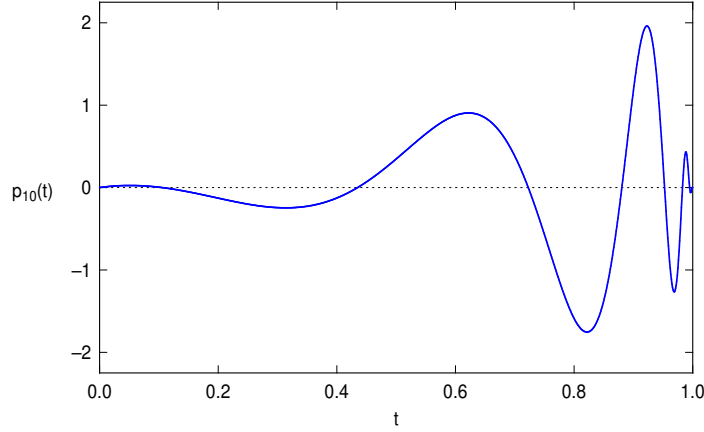


Figure 2: Graph of the polynomial  $p_{10}(t)$  over the interval  $t \in (0, 1)$ .

## 4 Computing the roots on $(0, 1)$

At the cost of sacrificing certainty, numerical computations permit greater scope and efficiency than symbolic computations in computing the real roots of (1) on  $t \in (0, 1)$ . In particular, the Bernstein form

$$f(t) = \sum_{i=0}^m c_i \binom{m}{i} (1-t)^{m-i} t^i$$

of a degree- $m$  polynomial  $f(t)$  on  $t \in [0, 1]$  offers significant computational advantages, including:

- interpolation of end-point values:  $f(0) = c_0$  and  $f(1) = c_m$ ;
- convex hull property:  $\min_i c_i \leq f(t) \leq \max_i c_i$  for  $t \in [0, 1]$ ;
- variation-diminishing property:  $\#$  of roots on  $(0, 1) = \#$  sign variations of  $(c_0, \dots, c_m) - 2k$  for some integer  $k \geq 0$ ;

- subdivision algorithm: for any point  $\tau \in (0, 1)$  the Bernstein coefficients on the subintervals<sup>1</sup>  $[0, \tau]$  and  $[\tau, 1]$  can be computed through recursive convex combinations of the coefficients  $c_0, \dots, c_m$ ;
- numerical stability:  $f(t)$  on  $t \in [0, 1]$  is always less sensitive to uniform perturbations in its Bernstein coefficients than in its power coefficients.

A scheme to isolate the real roots of a polynomial on  $t \in (0, 1)$  may be based on the variation–diminishing property and subdivision algorithm. See [3] for a comprehensive review of the properties of the Bernstein form — further details on the numerical stability properties may be found in [4, 5].

#### 4.1 Bernstein form of $p_n(t)$

The change of variables  $t \in (0, \infty) \rightarrow u \in (0, 1)$  specified by

$$t = \frac{u}{1-u} \quad (4)$$

defines a bijective relation between  $t$  and  $u$ , since  $dt/du = (1-u)^{-2} > 0$  for  $u \in (0, 1)$ . Note also that  $t \in (0, 1)$  and  $t \in (1, \infty)$  correspond to  $u \in (0, \frac{1}{2})$  and  $u \in (\frac{1}{2}, 1)$ . Substituting (4) into (1) yields

$$p_n(u) = \frac{1}{(1-u)^{2^n}} \sum_{k=0}^n (-1)^k \binom{n}{k} (1-u)^{2^n-2^k} u^{2^k},$$

In computing the positive roots of (1), we omit the factor  $1/(1-u)^{2^n}$  since it is non-zero and finite on  $(0, 1)$ . The remaining term  $f_n(u) = (1-u)^{2^n} p_n(u)$  can be expressed in the Bernstein basis of degree  $2^n$  on  $u \in [0, 1]$  as

$$f_n(u) = \sum_{i=0}^{2^n} c_i \binom{2^n}{i} (1-u)^{2^n-i} u^i, \quad (5)$$

where  $c_i = 0$  if  $i \neq 2^k$  for  $k = 0, \dots, n$  and

$$c_i = (-1)^k \frac{\binom{n}{k}}{\binom{2^n}{2^k}} \quad \text{if } i = 2^k \text{ for } k = 0, \dots, n. \quad (6)$$

---

<sup>1</sup>It is understood here that these sub-intervals are both mapped to  $[0, 1]$  through the transformations  $t \rightarrow t/\tau$  and  $t \rightarrow (t-\tau)/(1-\tau)$ , respectively.

The form (5) retains the sparsity of (1), with only  $n+1$  non-zero coefficients.

Although the Bernstein coefficients of the restriction of (5) to  $u \in (0, \frac{1}{2})$  can be computed by the de Casteljau subdivision algorithm [2], it is preferable to derive them from first principles. Setting

$$p_n(t) = \sum_{l=0}^{2^n} c_l \binom{2^n}{l} (1-t)^{2^n-l} t^l, \quad (7)$$

the coefficients  $c_0, \dots, c_{2^n}$  of (1) on  $t \in [0, 1]$  may be determined as follows. For each  $k$ , the  $k^{\text{th}}$  term in (7) is multiplied by  $1 = [(1-t) + t]^{2^n-2^k}$  to give

$$p_n(t) = \sum_{k=0}^n (-1)^k \binom{n}{k} t^{2^k} [(1-t) + t]^{2^n-2^k},$$

and through binomial expansions of the factors  $[(1-t) + t]^{2^n-2^k}$  we obtain

$$\begin{aligned} p_n(t) &= \sum_{k=0}^n (-1)^k \binom{n}{k} t^{2^k} \sum_{j=0}^{2^n-2^k} \binom{2^n-2^k}{j} (1-t)^{2^n-2^k-j} t^j \\ &= \sum_{k=0}^n \sum_{j=0}^{2^n-2^k} (-1)^k \binom{n}{k} \binom{2^n-2^k}{j} (1-t)^{2^n-2^k-j} t^{2^k+j}. \end{aligned}$$

By setting  $l = 2^k + j$ , re-arranging the order of summation, and performing some manipulations on the binomial coefficients, the Bernstein form of  $p_n(t)$  on  $t \in [0, 1]$  is determined to have coefficients defined by  $c_0 = 0$  and

$$c_l = \frac{1}{\binom{2^n}{l}} \sum_{k=0}^{k_{\max}} (-1)^k \binom{n}{k} \binom{2^n-2^k}{l-2^k}, \quad l = 1, \dots, 2^n, \quad (8)$$

$k_{\max} = \lfloor \log_2 l \rfloor$  being the largest integer  $k$  with  $2^k \leq l$ . Note that  $c_{2^n} = 0$ , since  $t = 1$  is a root of  $p_n(t)$ . Because of the restricted range of summation, and the appearance of the summation index as an exponent<sup>2</sup> in the binomial coefficients, there is no obvious further reduction of the sum (8).

---

<sup>2</sup>The authoritative compilation of summation formulas [10] contains no examples that include the summation index as an exponent within the binomial coefficients.

The binomial coefficients appearing in (8) can be very large, because of the exponential dependence on  $n$ . Goetgheluck [8, 9] describes an algorithm to represent any binomial coefficient through its factorization

$$\binom{m}{l} = \prod_{i=1}^r p_i^{e_i}, \quad (9)$$

where  $\{p_1, p_2, p_3, \dots\} = \{2, 3, 5, \dots\}$  is the ordered set of prime numbers, and  $e_1, e_2, e_3, \dots$  are their corresponding exponents. Since primes larger than  $m$  cannot appear in (9), the index  $r$  is such that  $p_r \leq m$  and  $p_{r+1} > m$ . The form (9) allows exact representation of high-order binomial coefficients, and also their products and ratios, since multiplication and division are performed by adding or subtracting their prime exponents. Having determined the sum in (8) exactly as an integer, the coefficient  $c_l$  is obtained by a single floating-point division. Nevertheless, using extended-precision integer arithmetic (the `long long` data type in the C programming language), this approach is only feasible for  $n \leq 6$ , since larger values incur overflow in evaluating the sums in (8). Thus, for  $n \leq 6$ , the prime factorization (9) is used in evaluating the binomial coefficients in (8), and for  $n > 6$  they are evaluated using double-precision floating-point arithmetic.

The Bernstein representation (7) is “dense” in the sense that, apart from  $c_0 = c_{2^n} = 0$  since  $p_n(0) = p_n(1) = 0$ , the coefficients  $c_l$  for  $1 \leq l \leq 2^n - 1$  are, in general, non-zero. For  $n \leq 3$ , the non-zero coefficients of  $p_n(t)$  are

$$\begin{aligned} n = 1 : \quad c_1 &= \frac{1}{2}, & n = 2 : \quad c_1, c_2, c_3 &= \frac{1}{4}, \frac{1}{6}, -\frac{1}{4}, \\ n = 3 : \quad c_1, c_2, c_3, c_4, c_5, c_6, c_7 &= \frac{1}{8}, \frac{2}{14}, \frac{3}{56}, -\frac{1}{10}, -\frac{13}{56}, -\frac{3}{14}, \frac{1}{8}. \end{aligned}$$

For  $n > 3$  it is impractical to list the coefficients, and they are best displayed graphically. The Bernstein form (7) of  $p_n(t)$  has *control points* specified by

$$(l/2^n, c_l) \quad \text{for } l = 0, \dots, 2^n$$

and the *control polygon* is the piecewise-linear graph obtained by connecting them in order. The control polygon mimics the graph of  $p_n(t)$ , as illustrated in Figures 3 for  $n = 5$  and 6, and it can be refined by degree elevation or subdivision [2] so as to converge monotonically to  $p_n(t)$ .

The roots of  $p_n(t)$  on  $(0, 1)$  can be isolated through a hierarchical binary subdivision, such that each subinterval exhibits exactly 0 or 1 sign changes in

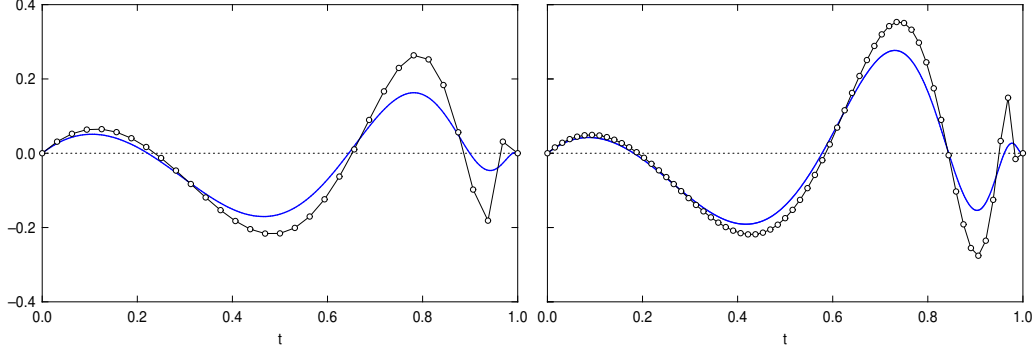


Figure 3: Graphs of the polynomials  $p_5(t)$  and  $p_6(t)$  over the interval  $[0, 1]$ , on the left and right, together with their Bernstein-form control polygons.

its associated Bernstein coefficients. Moreover, a simple check for guaranteed convergence of Newton–Raphson iterations [11] to the unique real root on each isolating interval can be expressed in terms of these Bernstein coefficients and their first and second forward differences.

Using initial estimates obtained from the graphs of  $p_n(t)$  for  $n \leq 6$ , the roots on  $(0, 1)$  were computed by Newton–Raphson iterations using both the power form (1) and Bernstein form (7). To evaluate  $p_n(t)$  and its derivative, Horner’s method and the de Casteljau algorithm were used in the former and latter cases. To ensure the greatest accuracy, the Bernstein coefficients (8) were computed using the binomial coefficient factorizations (9).

$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
0.6180339888	0.3940092283	0.2843286564	0.2217578136	0.1816228563
	0.8860953981	0.7425684029	0.6446586494	0.5776201937
		0.9659707987	0.8991950734	0.8439600927
			0.9895661392	0.9608761970
				0.9967274894

Table 1: Computed roots of  $p_n(t)$  on  $(0, 1)$  for  $2 \leq n \leq 6$ .

The roots on  $t \in (0, 1)$  for  $n \leq 6$  computed using the power and Bernstein forms were found to differ by  $< 10^{-15}$ . They are enumerated in Table 1 to 10 decimal places. For  $n > 6$ , double-precision floating-point arithmetic was used to compute the binomial coefficients in (8), to avoid integer overflow. However, the roots computed using the power and Bernstein forms still differ

by  $< 10^{-15}$  in most cases, and  $10^{-13}$  for the roots clustered near  $t = 1$  when  $n = 10$ . These roots are listed in Table 2 to 10 decimal places.

$n = 7$	$n = 8$	$n = 9$	$n = 10$
0.1537506418	0.1332813706	0.1176164554	0.1052439966
0.5285323068	0.4905966981	0.4601069756	0.4348788819
0.8026230006	0.7705178543	0.7445053722	0.7227374666
0.9335597010	0.9120061995	0.8947784829	0.8805076463
0.9846749330	0.9717430537	0.9611371063	0.9525258901
0.9989563611	0.9939230694	0.9878924924	0.9827779812
	0.9996632097	0.9975629411	0.9947648577
		0.9998903961	0.9990133089
			0.9999641141

Table 2: Computed roots of  $p_n(t)$  on  $(0, 1)$  for  $7 \leq n \leq 10$ .

## 4.2 Distribution of the positive real roots

The computed roots of  $p_n(t)$  on  $(0, 1)$  for  $2 \leq n \leq 10$ , as listed in Tables 1 and 2, are shown in Figure 4 with linear scales for  $t$  and  $n$ , and in Figure 5 with logarithmic scales. For  $n > 10$ , the computations are very cumbersome, and the clustering of roots near  $t = 1$  makes them difficult to resolve. Figure 5 suggests a power-law dependence of the roots on  $n$ , but this is not exact. A least-squares fit to the smallest root of  $p_n(t)$  for  $2 \leq n \leq 10$  gives

$$t \approx a n^b, \quad a \approx 1.146625, \quad b \approx -1.036283. \quad (10)$$

As seen in Table 3, there is a noticeable discrepancy at the smaller  $n$  values, although the accuracy improves significantly if the fit is restricted to  $n \geq 4$ .

Since the tightly-clustered roots near  $t = 1$  are not easily distinguishable for the larger  $n$  values in Figures 4 and 5, an alternative plot is presented in Figure 6, showing their distances from  $t = 1$  on a logarithmic scale in  $t$ .

Figure 7 illustrates the super-quadratic distribution of the roots of  $p_n(t)$  identified in Theorem 1 for the case  $n = 10$ . The interlacing property of the roots of  $p_n(t)$  and  $p_{n+1}(t)$  is evident in Tables 1 and 2, and through a close inspection of Figure 4.

$n$	actual root	least-squares fit
2	0.618033988750	0.559073785600
3	0.304009228317	0.367272762062
4	0.284328656449	0.272594325571
5	0.221757813626	0.216316972336
6	0.181622856260	0.179075595125
7	0.153750641776	0.152637263660
8	0.133281370572	0.132912091834
9	0.117616455363	0.117640265263
10	0.105243996624	0.105472266277

Table 3: Comparison of smallest roots of the polynomials  $p_n(t)$  for  $2 \leq n \leq 10$  with the values obtained from a power law least-squares fit of the form (10).

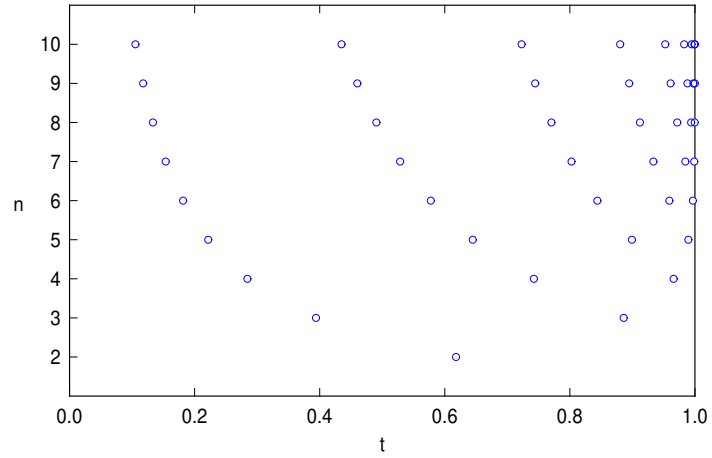


Figure 4: Distribution of real roots of the polynomials  $p_n(t)$  over the interval  $t \in (0, 1)$  for the cases  $n \leq 2 \leq 10$  (the roots at  $t = 0$  and  $t = 1$  are omitted).



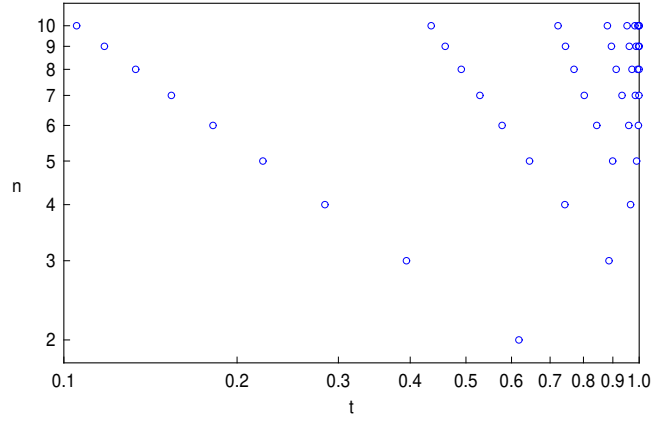


Figure 5: Distribution of real roots of the polynomials  $p_n(t)$  over the interval  $t \in (0, 1)$  for the cases  $2 \leq n \leq 10$ , employing logarithmic scales for  $t$  and  $n$ .

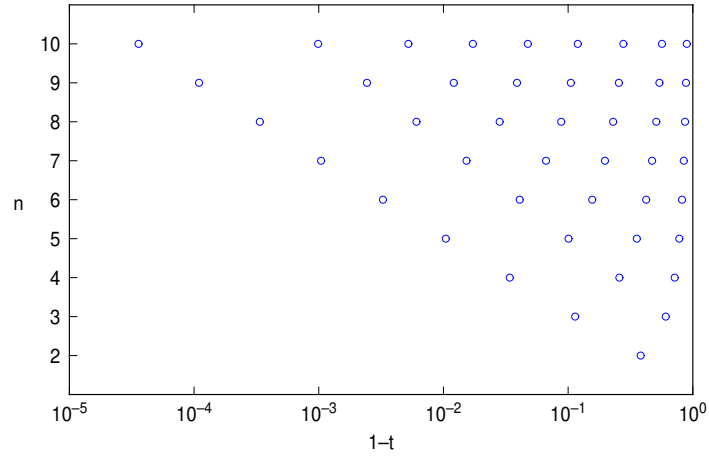


Figure 6: Distance of the real roots of the polynomials  $p_n(t)$  on  $(0, 1)$  from  $t = 1$  for the cases  $n \leq 2 \leq 10$ , plotted using a logarithmic scale for  $1 - t$ .

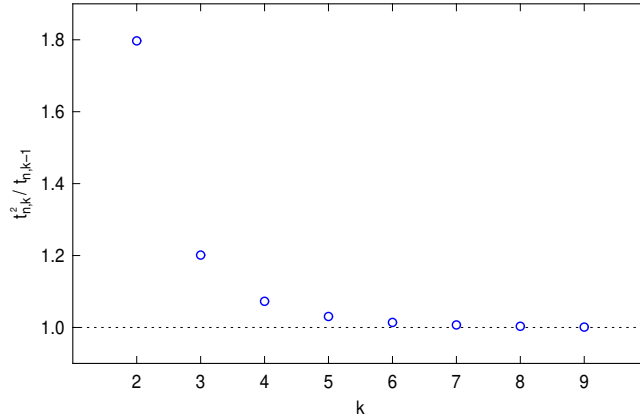


Figure 7: The “super-quadratic” distribution property  $t_{n,k}^2 > t_{n,k-1}$  for  $k = 2, \dots, n-1$  of the roots  $t_{n,1}, \dots, t_{n,n-1}$  of the polynomial  $p_n(t)$  for  $n = 10$ .

### 4.3 The complete set of roots

Although we are mainly interested in the positive real roots of  $p_n(t)$  on  $(0, 1)$ , we can easily identify some basic facts about its entire set of roots. First, as a consequence of Theorem 1 and Descartes’ Law of Signs, we observe that  $p_n(t)$  has no real roots  $t > 1$ . Concerning the negative roots we note that, whereas  $p_n(t)$  has  $n$  coefficient sign variations,  $p_n(-t)$  has only  $n-1$ . Thus, Descartes’ Law of Signs implies that  $p_n(t)$  has at least one negative real root for even  $n$ . Moreover,  $p_n(-t) = p_n(t) - 2t$ , and we have  $p_n(-1) = -2$  for any  $n$ , since  $p_n(1) = 0$ . Hence, noting that  $p_n(t) \rightarrow +\infty$  as  $t \rightarrow -\infty$  for even  $n$ , there must be a real root on  $(-\infty, -1)$  if  $n$  is even. The negative real roots of  $p_n(t)$  can be identified as the negatives of the values  $t > 0$  where the graphs of  $p_n(t)$  and  $2t$  intersect. For  $t \in (0, 1)$  this first occurs when  $n = 10$ , as can be seen by noting in Figure 2 that the graph of  $2t$  crosses  $p_{10}(t)$  twice.

Finally, we briefly comment on the complete set of  $2^n$  (real and complex) roots of  $p_n(t)$  which (as computed in **Maple**) are illustrated in Figure 8 for the cases  $n = 4, 6, 8, 10$ . It is seen that (most of) the complex roots cluster near a circle centered at the origin, whose radius approaches 1 as  $n$  increases.

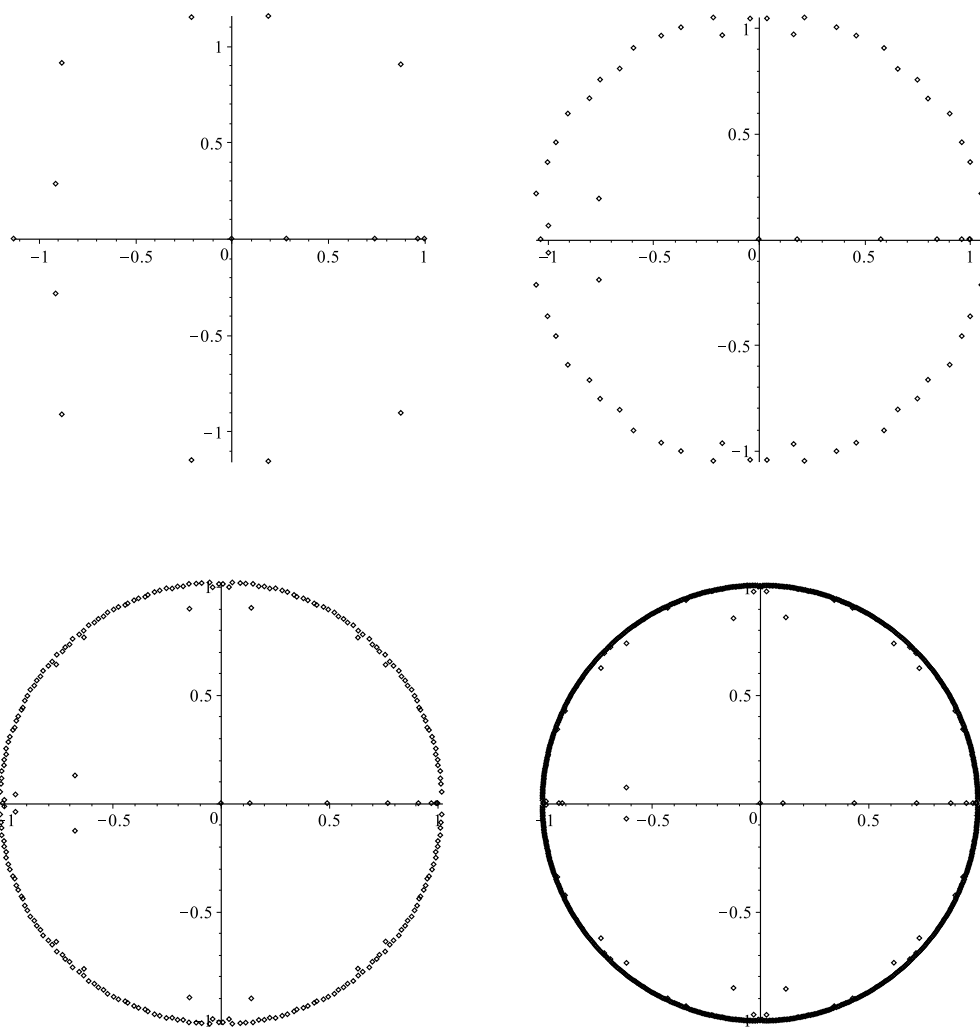


Figure 8: The entire set of (real and complex) roots of  $p_n(t)$  for  $n = 4, 6, 8, 10$ .

## 5 Numerical stability of roots

It is known that, for a polynomial specified in the power and Bernstein forms

$$f(t) = \sum_{i=0}^m a_i t^i = \sum_{i=0}^m c_i \binom{m}{i} (1-t)^{m-i} t^i,$$

on  $t \in [0, 1]$ , the values (and roots) of  $f(t)$  are systematically less sensitive to uncertainties of a uniform maximum relative magnitude  $\epsilon$  in the Bernstein coefficients than in the power coefficients [5]. In fact, it has been shown [4] that among all non-negative polynomial bases on  $[0, 1]$  the Bernstein basis is “optimally stable” with respect to such coefficient perturbations.

For the power form (1), the  $n + 1$  non-zero coefficients have magnitudes  $\geq 1$  that (except for the lowest- and highest-order terms) are increasing with  $n$ . On the other hand, the Bernstein form (7) has  $2^n - 1$  non-zero coefficients, of magnitude  $\leq 1$  that are decreasing with  $n$ . For both forms, the non-zero coefficients have alternating signs. A somewhat surprising result is that both representations admit computation of the roots of  $p_n(t)$  with similar accuracy — it is only for the tightly-clustered roots near  $t = 1$  in the case  $n = 10$  that the power form begins to exhibit slower and more erratic convergence of the Newton–Raphson iterations than the Bernstein form (see Table 4).

iteration	Bernstein	power
1	0.999964010117222	0.999964010117263
2	0.999964113848415	0.999964113848471
3	0.999964114129893	0.999964114129906
4	0.999964114129895	0.999964114129931
5	0.999964114129895	0.999964114129896
6	0.999964114129895	0.999964114129898
7	0.999964114129895	0.999964114129856
8	0.999964114129895	0.999964114129880

Table 4: Newton–Raphson iterations for the largest real root of  $p_{10}(t)$  from the initial value  $t = 0.999966$ , using the Bernstein and power representations. Note that the computed values differ only in the 14<sup>th</sup> and 15<sup>th</sup> decimal places.

This phenomenon may be understood as follows. For the power form (1),

the  $n + 1$  non-zero coefficients are known exactly<sup>3</sup> *a priori* as integers. For the Bernstein form (7), on the other hand, the  $2^n - 1$  non-zero coefficients must be computed from the expression (8). This can be done almost exactly for  $n \leq 6$  using “long” (64-bit) integers, but floating-point arithmetic must be employed for  $n > 6$ , incurring “initial” coefficient errors.

However, initial coefficient errors are not the only factors influencing the accuracy of the computed roots. In *backward error analysis* [14, 16] the result of each floating-point arithmetic operation is interpreted as an exact outcome for perturbed operands. By “backward” propagation of these perturbations through the algorithm, from the final result to the input values, the outcome of the computation can be regarded as exact for certain perturbed inputs. Knowing these input perturbations and the *condition number* for the problem (i.e., the sensitivity of the output values to changes in the input values), the accuracy of the final result can (in principle) be assessed.

The condition number of a simple real root  $t = \tau$  of a polynomial  $p_m(t)$  with coefficients  $a_0, \dots, a_m$  in the basis  $\{\phi_0(t), \dots, \phi_m(t)\}$  is conventionally defined [6, 7] as

$$C(\tau) := \frac{\sum_{i=0}^m |a_i \phi_i(\tau)|}{|p'_m(\tau)|}.$$

In the limit  $\epsilon \rightarrow 0$ , the perturbation  $\delta\tau$  in the root  $\tau$  incurred by errors in the coefficients of uniform maximum relative magnitude  $\epsilon$  satisfies

$$|\delta\tau| \leq C(\tau) \epsilon.$$

Table 5 lists the condition numbers  $C_B$  and  $C_P$  in the Bernstein and power bases for the 9 roots of the instance  $n = 10$  of the polynomial (1). It is seen that the inequality  $C_B < C_P$  holds for all the roots, and the growth of  $C_P$  as the roots approach 1 is noteworthy (whereas  $C_B$  actually decreases). This explains the slower convergence to the largest root seen in Table 4.

The condition number values listed in Table 5 are actually very modest. This fact, coupled with the remarkable degree of agreement between the roots computed using the power and Bernstein forms, imparts confidence in the roots quoted herein. In proceeding beyond  $n = 10$ , it is likely that Bernstein form will furnish better results for the roots tightly clustered near  $t = 1$ , although at a significantly higher computational cost.

---

<sup>3</sup>Note, however, that not all of these coefficients admit a finite binary representation, and small errors may be incurred in converting them to floating-point numbers.

root	$C_B$	$C_P$
0.105243996624	0.0077	0.2475
0.434878881858	0.0124	0.9521
0.722737466605	0.0112	1.4939
0.880507646268	0.0082	1.8829
0.952525890137	0.0055	2.4949
0.982777981150	0.0036	4.2066
0.994764857698	0.0024	11.4735
0.999013308936	0.0015	69.9995
0.999964114130	0.0001	1085.2778

Table 5: Root condition numbers for the Bernstein and power forms of  $p_{10}(t)$ .

## 6 Closure

In this study, several interesting theoretical properties of the polynomials  $p_n(t)$  defined by (1) have been elucidated — including (i) the existence of a simple recursion relating  $p_{n+1}(t)$  to  $p_n(t)$ ; (ii) the fact that these polynomials have, for each  $n > 1$ , precisely  $n - 1$  roots on  $t \in (0, 1)$  and none for  $t > 1$ ; (iii) the super-quadratic distribution of these roots, i.e., each root is smaller than the *square* of the next-largest root; and (iv) the interlacing of the roots of  $p_n(t)$  and  $p_{n+1}(t)$ , wherein the roots of the latter on  $(0, 1)$  lie within the intervals delineated by the roots of former.

The unusual nature of the polynomials  $p_n(t)$  poses formidable challenges for root computation algorithms on account of their high degrees and strong scale variations of their graphs. Nevertheless, it was observed that for  $n \leq 10$  their roots are remarkably well-conditioned, and admit accurate computation in ordinary double-precision arithmetic using both the power and Bernstein forms. Although the Bernstein form is more stable than the power form, it lacks the sparsity of the power form, and its advantages are only apparent with the tightly-clustered roots of  $p_{10}(t)$  near  $t = 1$ . Due to the exponential growth in the degree, root computations for  $p_n(t)$  with  $n > 10$  have not been attempted: these tax the capabilities of both symbolic and numeric methods, and special procedures may be needed to reliably address them.

## References

- [1] P. Erdős and A. Rényi (1959), On random graphs. I, *Publ. Math. Debrecen* **6**, 290–297.
- [2] G. E. Farin, *Curves and Surfaces for Computer-Aided Geometric Design: a practical guide*, 3rd ed., Academic Press, Boston, 1993.
- [3] R. T. Farouki, The Bernstein polynomial basis: a centennial retrospective, *Computer Aided Geometric Design* **29**, 379–419, 2012.
- [4] R. T. Farouki and T. N. T. Goodman, On the optimal stability of the Bernstein basis, *Mathematics of Computation*, **65**, 1553–1566, 1996.
- [5] R. T. Farouki and V. T. Rajan, On the numerical condition of polynomials in Bernstein form, *Computer Aided Geometric Design* **4**, 191–216, 1987.
- [6] W. Gautschi (1973), On the condition of algebraic equations, *Numerische Mathematik* **21**, 405–424.
- [7] W. Gautschi (1984), Questions of numerical condition related to polynomials, *Studies in Numerical Analysis* (G. H. Golub, ed.), MAA Studies in Mathematics **24**, 140–177.
- [8] P. Goetgheluck, Computing binomial coefficients, *American Mathematical Monthly*, **94**, 360–365, 1987.
- [9] P. Goetgheluck (1988), On prime divisors of binomial coefficients, *Mathematics of Computation* **51**, 325–329.
- [10] H. W. Gould (1972), *Combinatorial Identities: A Standardized Set of Tables Listing 500 Binomial Coefficient Summations*, Gould Publications, Morgantown, W. Va.
- [11] P. Henrici, *Elements of Numerical Analysis*, Wiley, New York, 1964.
- [12] J. V. Uspensky, *Theory of Equations*, McGraw–Hill, New York, 1948.
- [13] J. H. Wilkinson (1959), The evaluation of the zeros of ill-conditioned polynomials, Parts I & II, *Numerische Mathematik* **1**, 150–166 & 167–180.

- [14] J. H. Wilkinson (1963), *Rounding Errors in Algebraic Processes*, Prentice–Hall, Englewood Cliffs, NJ.
- [15] J. H. Wilkinson (1984), The perfidious polynomial, *Studies in Numerical Analysis* (G. H. Golub, ed.), MAA Studies in Mathematics, Vol. 24, pp. 1–28.
- [16] J. H. Wilkinson (1988), *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford.