Serdica J. Computing **8** (2014), No 1, 71–96

Serdica
Journal of Computing

Bulgarian Academy of Sciences
Institute of Mathematics and Informatics

ON OPTIMAL QUADRATIC LAGRANGE INTERPOLATION: EXTREMAL NODE SYSTEMS WITH MINIMAL LEBESGUE CONSTANT VIA SYMBOLIC COMPUTATION

Heinz-Joachim Rack, Robert Vajda

Abstract. We consider optimal Lagrange interpolation with polynomials of degree at most two on the unit interval $[-1, 1]$. In a largely unknown paper, Schurer (1974, Stud. Sci. Math. Hung. 9, 77-79) has analytically described the infinitely many zero-symmetric and zero-asymmetric extremal node systems $-1 \leq x_1 < x_2 < x_3 \leq 1$ which all lead to the minimal Lebesgue constant 1.25 that had already been determined by Bernstein (1931, Izv. Akad. Nauk SSSR 7, 1025-1050). As Schurer's proof is not given in full detail, we formally verify it by providing two new and sound proofs of his theorem with the aid of symbolic computation using quantifier elimination. Additionally, we provide an alternative, but equivalent, parameterized description of the extremal node systems for quadratic Lagrange interpolation which seems to be novel. It is our purpose to bring the computer-assisted solution of the first nontrivial case of optimal Lagrange interpolation to wider attention and to stimulate research of the higher-degree cases. This is why our style of writing is expository.

ACM Computing Classification System (1998): G.1.1, G.1.2.

Key words: extremal, interpolation nodes, Lagrange interpolation, Lebesgue constant, minimal, optimal, polynomial, quadratic, quantifier elimination, symbolic computation.

1. Introduction. Lagrange polynomial interpolation is a classical and feasible method to approximate continuous functions by algebraic polynomials of a given maximal degree. The goodness of this approximation method, as compared with the best possible approximation in Chebyshev's sense, is measured with the aid of the Lebesgue constants which can be viewed as operator norms or interpolating projection constants or condition numbers, see [34]. They depend solely on the chosen interpolation nodes and Lebesgue's lemma suggests that we choose the nodes in such a manner that the Lebesgue constants become minimal. Such optimal interpolation nodes are in a sense the opposite to the equidistant interpolation nodes which may lead to disastrous approximation results, see Runge's example in [1, p. 104], [17, p. 60].

The construction of an optimal Lagrange interpolation polynomial which minimizes the Lebesgue constant finds applications, for instance, in the method of Finite Elements when one has to describe the boundary of curvilinear domains in two or three dimensions and to construct the mapping for elements adjacent to the boundary: *... in this case the use of optimal points is essential because the use of a uniformly distributed set of points could introduce excessive error, which could affect the overall performance of the finite element method* [1, p. 105].

The search for an analytical determination of extremal node systems and corresponding minimal Lebesgue constants is still an intriguing topic in mathematics today, see e.g. [34]. In this paper we address the first nontrivial case of optimal Lagrange interpolation with quadratic polynomials, i.e., of maximal degree two, a case which had been considered earlier by [2], [6], [7], [13], [25], [29] and [30]. We collect known results on optimal quadratic interpolation on the unit interval $[-1, 1]$ and in particular draw attention to the largely unknown papers [25] and [29] which are referenced neither in dedicated books on interpolation theory nor in the survey paper [5], nor in the encyclopedia [34]. Tureckii seems to be the first to describe the infinitely many zero-symmetric node systems (and provides a proof in [30]). Schurer seems to be the first to describe all (zero-symmetric and zero-asymmetric) extremal node systems consisting of three interpolation nodes on $[-1, 1]$. However, his proof is only sketched, so we provide two new and sound proofs of his result (Section 3.4, Section 3.5). Our proofs are based on powerful symbolic computation using quantifier elimination. Additionally, we provide an alternative, but equivalent, parameterized description of all extremal node systems for quadratic interpolation (Theorem 3.6), which seems to be novel. Several examples of extremal node systems are given in the text. Furthermore we show, by example of a continuous function, that equality can be attained in Lebesgue's lemma (Proposition 3.2), and we verify a remark of

[25] stating that only on Bernstein's extremal node system the Lebesgue function equioscillates most (Proposition 3.10).

This paper is on the edge between computer algebra systems (Mathematica, QEPCAD) and theoretical approximation theory (in particular, polynomial interpolation) and leads to the computer-aided verification, by means of quantifier elimination, of the shape of the 2D set of optimal points for quadratic Lagrange interpolation. It thus provides another example of a successful link between symbolic and numerical methods. The established equivalence of Schurer's 2D set with our parameterized description of it (Theorem 3.6) will serve as a blueprint for future generalizations to the cubic and the quartic case assisted by more sophisticated symbolic tools such as Groebner bases and resultants, see [23] and [32]. We hope that this paper, along with our dedicated web repository

`www.math.u-szeged.hu/~vajda/Leb/`

will add to the dissemination of computer-aided optimal (quadratic) Lagrange interpolation and will facilitate its presentation and impartation.

2. Definitions and basic theoretical background. Let $C(\mathbf{I})$ denote the Banach space of continuous real functions f on the interval $\mathbf{I} = [-1, 1]$, equipped with the uniform norm (also called Chebyshev norm):

$$||f|| = \max_{x \in \mathbf{I}} |f(x)|. \tag{1}$$

Suppose we wish to approximate f by an algebraic polynomial of degree at most $n-1$, where $n \geq 3$. An old idea, going back to Waring and Euler but named after Lagrange, see [18], is to sample f at n distinct points of $\mathbf{I}$,

$$X_n : \, -1 \leq x_1 < x_2 < \ldots < x_{n-1} < x_n \leq 1, \tag{2}$$

and to construct an interpolating polynomial of degree at most $n-1$ as follows:

$$L_{n-1}(x) = L_{n-1}(f, X_n, x) = \sum_{j=1}^{n} f(x_j) \ell_{n-1,j}(X_n, x) \tag{3}$$

where

$$\ell_{n-1,j}(x) = \ell_{n-1,j}(X_n, x) = \prod_{i=1, i \neq j}^{n} \frac{x - x_i}{x_j - x_i}, \tag{4}$$

so that

$$\ell_{n-1,j}(X_n, x_i) = \delta_{j,i} \quad \text{(Kronecker delta)} \tag{5}$$

and hence

$$L_{n-1}(x_i) = f(x_i), \quad 1 \le i \le n. \tag{6}$$

Definition 2.1. *We call the* x_i*'s in* (2) *the* interpolation nodes *and the grid* X_n *the* node system.

Definition 2.2. *The unique polynomial* L_{n-1} *is called the* Lagrange interpolation polynomial *and the polynomials* $\ell_{n-1,j}$ (*of exact degree* $n-1$) *are called the* Lagrange fundamental polynomials.

If $\|f\| \le 1$ then (3) implies that $|L_{n-1}(x)|$ can be estimated from above by

$$\sum_{j=1}^{n} |\ell_{n-1,j}(x)| = \lambda_n(x) = \lambda_n(X_n, x). \tag{7}$$

Definition 2.3. *The non-negative function* λ_n*, which is independent of* f*, is called the* Lebesgue function (*named after Lebesgue, see* [33]).

Three properties of λ_n are summarized in the following statement, see [5], [16], [28, p. 95]:

Proposition 2.4.

(i) λ_n *is a piecewise polynomial satisfying* $\lambda_n(x) \ge 1$ *with equality only if* $x = x_i\,(1 \le i \le n)$.

(ii) λ_n *has precisely one local maximum, which we will denote by* $\mu_i = \mu_i(X_n)$*, in each open sub-interval* (x_i, x_{i+1}) *of* $X_n \quad (1 \le i \le n-1)$*. The extremum point in* (x_i, x_{i+1})*, at which the maximum* μ_i *is attained, we will denote by* $\xi_i = \xi_i(X_n)$ *so that* $\lambda_n(\xi_i) = \mu_i$ *holds.*

(iii) λ_n *is strictly decreasing and convex in* $(-\infty, x_1)$ *and strictly increasing and convex in* (x_n, ∞).

Definition 2.5. *The largest value of* λ_n *on* $\mathbf{I}$*, denoted by* $\Lambda_n = \Lambda_n(X_n)$*, is called the Lebesgue constant:*

$$\Lambda_n = \max_{x \in \mathbf{I}} \lambda_n(x). \tag{8}$$

We thus have either

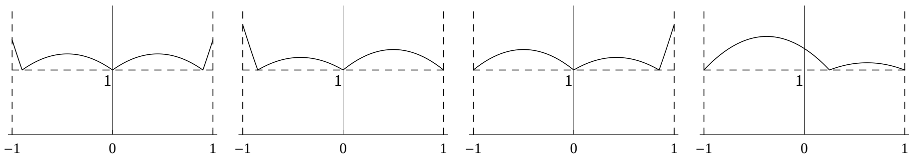

Fig. 1. Typical shapes of a Lebesgue function ($n = 3$)

$$\Lambda_n = \max(\lambda_n(-1), \mu_1, \ldots, \mu_{n-1}, \lambda_n(1)), \text{ if } -1 \neq x_1 \text{ and } 1 \neq x_n;$$

or

$$\Lambda_n = \max(\lambda_n(-1), \mu_1, \ldots, \mu_{n-1}), \text{ if } -1 \neq x_1 \text{ and } 1 = x_n;$$

or

$$\Lambda_n = \max(\mu_1, \ldots, \mu_{n-1}, \lambda_n(1)), \text{ if } -1 = x_1 \text{ and } 1 \neq x_n;$$

or

$$\Lambda_n = \max(\mu_1, \ldots, \mu_{n-1}), \text{ if } -1 = x_1 \text{ and } 1 = x_n.$$

The importance of Λ_n in interpolation theory stems from the following inequality which can be viewed as a version of Lebesgue's lemma, but can also be proved directly [24, Theorem 4.1]:

$$\|f - L_{n-1}\| \leq (1 + \Lambda_n)\|f - P^*_{n-1}\|, \tag{9}$$

where $f \in C(\mathbf{I})$, and P^*_{n-1} denotes the polynomial of best uniform approximation to f out of the linear space of all algebraic polynomials of degree at most $n-1$. Usually, P^*_{n-1} is much harder to determine than L_{n-1}, and of course there always holds $\|f - P^*_{n-1}\| \leq \|f - L_{n-1}\|$. The estimate (9), which is sharp for some f, tells us that a small Lebesgue constant implies that the approximation to f by the Lagrange interpolation polynomial is nearly as good as the best uniform approximation to f by means of P^*_{n-1}. Therefore, it is desirable to minimize Λ_n which can be achieved by a strategic placement of the interpolation nodes. However, Λ_n cannot be forced to be arbitrarily close to 1: as we shall see, the minimal value of Λ_n is 1.25, if $n = 3$, and in fact Λ_n grows at least logarithmically with n.

It is known [24, p. 100] that for each $n \geq 3$ a node system $X_n = X^*_n$ on $\mathbf{I}$ exists such that

(10) $\Lambda_n^* = \Lambda_n(X_n^*) \leq \Lambda_n(X_n)$ for all choices of node systems X_n according to (2).

Definition 2.6. *A node system X_n^* which satisfies* (10) *is called* extremal, *and the corresponding Lebesgue constant Λ_n^* is called* minimal.

It is furthermore known that for a given $n \geq 3$ an extremal node system is not unique, [16, Theorem 2], and in particular there exists an extremal node system which includes the endpoints of **I** as interpolation nodes, see [24, p. 100]. Obviously, all extremal node systems, for a given n, generate the same minimal Lebesgue constant.

Definition 2.7. *The construction of $L_{n-1}(f, X_n^*, x)$, the Lagrange interpolation polynomial on an extremal node system X_n^*, is called* optimal *Lagrange polynomial interpolation on* **I** *since it furnishes, for a given n, the minimal Lebesgue constant and hence the minimal interpolation error in the sense of* (9).

Definition 2.8. *Following standard usage* [5], [28] *we will call a node system, which includes the endpoints of* **I** *as interpolation nodes* (*that is, $x_1 = -1$ and $x_n = 1$*), *a* canonical *node system* (*CNS*).

In answering a conjecture which goes back to [2], it was proved by [9] and by [14] that the following deep result holds:

Proposition 2.9. *If a Lebesgue function corresponding to a CNS X_n satisfies the so-called equioscillation property*

$$\mu_1 = \mu_2 = \ldots = \mu_{n-2} = \mu_{n-1}, \tag{11}$$

then X_n is an extremal node system, i.e., $X_n = X_n^*$ with $\Lambda_n(X_n) = \Lambda_n^*$.

Thus, the fulfillment of (11) is a sufficient condition for a CNS to be extremal. Actually, it was additionally proved that a CNS which satisfies (11) is unique. This property answers part of a conjecture which goes back to [10]. However, extremal node systems are not given explicitly in [9], [14].

The search for the analytical (that is, not numerical) determination of X_n^* and Λ_n^* is among the most intriguing problems of interpolation theory. Here are some quotations on this subject:

- *In spite of this nice characterization, the optimal nodes as well as the optimal Lebesgue constants are not known explicitly... the problem of analytical* *description of the optimal matrix of nodes is considered by pure mathematicians as a great challenge.* [5]

- *The following questions are still open: 1. Is there a set of relatively simple functions f_n such that the roots of f_n are the optimal nodes for Lagrange interpolation?* [8, p. 21]
- *It is of interest to mention that no elegant general method, whether a formula or a special algorithm, has yet been discovered which serves to compute the nodes yielding $\mu_1 = \mu_2 = \ldots = \mu_{n-2} = \mu_{n-1}$.* [14]
- *The nature of the optimal set X^* remains a mystery* [15, p. xlvii]
- *In general, the location of optimal interpolation nodes is unknown and we have made some numerical computations which may indicate the direction in which such points should be sought.* [16]
- *It is an open problem to get the exact value of the optimal Lebesgue constants...* [17, p. 67]

In the present paper our main focus is on extremal (non-canonical) node systems consisting of $n = 3$ interpolation nodes on $\mathbf{I}$. The unique extremal CNS on $\mathbf{I}$ for $n = 3$ can be easily obtained with the aid of Proposition 2.9, see Section 3.1 below.

3. On optimal quadratic Lagrange polynomial interpolation.

Our goal is to describe all extremal node systems X_3^* on $\mathbf{I}$ which lead to the minimal Lebesgue constant $\Lambda_3^* = 1.25$ associated with optimal Lagrange polynomial interpolation by polynomials of a given maximal degree $n - 1 = 2$. In an attempt to contribute to the classical subject of Lagrange polynomial interpolation we

- collect known results on optimal quadratic interpolation (in particular, we refer to the largely unknown sources [29] and [25]),
- provide two complete proofs, by using symbolic computation, for Schurer's description [25] of the infinitely many (symmetric and asymmetric) extremal node systems consisting of three interpolation nodes on $\mathbf{I}$,
- add an alternative description of these node systems (Theorem 3.6), which is believed to be novel, and which was inspired by the proof given in [16, Theorem 2] for the non-uniqueness of extremal node systems,

- provide examples of $f \in C(\mathbf{I})$ showing that the estimate (9) can turn into an equation (Proposition 3.2), and
- verify an unproven remark in [25] on the maximal equioscillation of λ_3 on Bernstein's node system BNS (Proposition 3.10).

To this end, we consider first the unique extremal and canonical node system, then extremal zero-symmetric node systems (which encompass the canonical one), and finally extremal general node systems (which encompass the zero-symmetric ones).

We will skip the trivial linear case $n-1=1$ with the sole extremal (and canonical) node system $X_2^* : -1 = x_1^* < x_2^* = 1$ consisting of two interpolation nodes on $\mathbf{I}$ which yields the minimal Lebesgue constant $\Lambda_2^* = 1$. The quadratic case $n-1=2$ considered here is the first non-trivial one and it can be resolved completely. The investigation of the case $n-1=3$ (optimal cubic interpolation) we intend to expose in a separate paper [23], see also Remark 4.2 below.

3.1. The extremal and canonical node system. If X_3 is assumed to be canonical, we necessarily have $x_1=-1$ and $x_3=1$. That then $x_2 = \dfrac{x_1+x_3}{2} = 0$ must hold can be deduced from a remark in [14], see also [13, Theorem 2], stating that the interpolation nodes of an extremal CNS must be symmetric about the midpoint of $\mathbf{I}$. But we may simply take $x_2=0$ as a self-suggesting guess and look how the corresponding Lebesgue function behaves:

Example 3.1. Let $X_3 : x_1 = -1 < x_2 = 0 < x_3 = 1$. By their definition, the corresponding Lagrange fundamental polynomials read

$$\ell_{2,1}(X_3,x) = x(x-1)/2,\ \ell_{2,2}(X_3,x) = 1-x^2,\ \ell_{2,3}(X_3,x) = x(x+1)/2, \tag{12}$$

and the corresponding Lebesgue function is accordingly given by

$$\lambda_3(X_3,x) = -x^2-x+1 \ \text{ on } (-1,0) \ \text{ and } \lambda_3(X_3,x) = -x^2+x+1 \ \text{ on } (0,1). \tag{13}$$

Setting the first derivative of $\lambda_3(X_3,x)$ with respect to x equal to zero yields the extremum points $\xi_1=-0.5$ and $\xi_2=0.5$, and this eventually gives the equal local maxima

$$\lambda_3(\xi_1) = \mu_1 = 1.25 \ \text{ and } \lambda_3(\xi_2) = \mu_2 = 1.25. \tag{14}$$

According to Proposition 2.9, since the canonical grid X_3 with $x_2=0$ thus implies the equioscillation property of $\lambda_3(X_3,x)$, it is in fact the unique extremal

CNS, i.e., $X_3 = X_3^* : x_1^* = -1 < x_2^* = 0 < x_3^* = 1$, and, furthermore, $\Lambda_3^* = 1.25$ is the minimal Lebesgue constant.

The lowest estimate (9) hence reads $\|f - L_2\| \le \frac{9}{4}\|f - P_2^*\|$ for all $f \in C(\mathbf{I})$, if we interpolate on the grid X_3^*, or on any other extremal node system consisting of three interpolation nodes on $\mathbf{I}$. The sharpness of this estimate follows from the following statement:

Proposition 3.2. *Let $n = 3$. There exists some $f^* \in C(\mathbf{I})$ so that equality holds in (9) with $\Lambda_3 = \Lambda_3^* = 1.25$, that is*

$$\|f^* - L_2\| = 2.25\,\|f^* - P_2^*\|. \tag{15}$$

P r o o f. Consider the polygonal line $f^* \in C(\mathbf{I})$ defined by

$$f^*(x) = \begin{cases} 2x+1, & \text{if} \quad -1 \le x \le 0, \\ -4x+1, & \text{if} \quad 0 \le x \le 0.5, \\ 4x-3, & \text{if} \quad 0.5 \le x \le 1. \end{cases} \tag{16}$$

The quadratic polynomial L_2 with $L_2(x) = -x^2 + x + 1$ interpolates f^* on the CNS $X_3^* : x_1^* = -1 < x_2^* = 0 < x_3^* = 1$ since $f^*(-1) = L_2(-1) = -1$, $f^*(0) = L_2(0) = 1$, $f^*(1) = L_2(1) = 1$, i.e., L_2 is the (optimal) Lagrange interpolation polynomial of degree (at most) 2. The absolute value of the difference function $f^* - L_2$ reads

$$|f^*(x) - L_2(x)| = \begin{cases} |x^2 + x|, & \text{if} \quad -1 \le x \le 0, \\ |x^2 - 5x|, & \text{if} \quad 0 \le x \le 0.5, \\ |x^2 + 3x - 4|, & \text{if} \quad 0.5 \le x \le 1. \end{cases} \tag{17}$$

It follows from elementary calculus that the largest value in (17) is attained on $\mathbf{I}$ at $x = 0.5$:

$$\|f^* - L_2\| = \max_{x \in \mathbf{I}} |f^*(x) - L_2(x)| = |f^*(0.5) - L_2(0.5)| = |-1 - 1.25| = 2.25.$$

On the other hand, the best uniform approximation to f^* on $\mathbf{I}$ out of the linear space of (at most) quadratic polynomials is the zero polynomial P_2^* given by $P_2^*(x) = 0$ for $x \in \mathbf{I}$. This follows from Chebyshev's alternation theorem, see [24, Theorem 1.7], because the difference function $f^* - P_2^* = f^*$ has 4 alternation points on $\mathbf{I}$: $f^*(-1) = -1$, $f^*(0) = 1$, $f^*(0.5) = -1$, $f^*(1) = 1$. It follows from (16) that $|f^*(x)| \le 1$ for $x \in \mathbf{I}$ (and equality is attained), so that we have

$\|f^*\| = \|f^* - P_2^*\| = 1$. This eventually gives $\|f^* - L_2\| = 2.25 = 2.25\,\|f^* - P_2^*\|$, so that indeed (15) holds.

We point out that interpolation on any other extremal node system different from the above X_3^* would yield, after slight modifications, the same result (15). Take, for example, the extremal zero-asymmetric node system as given in Example 3.8 below, i.e., $X_3^* : x_1^* = -\frac{197}{207} < x_2^* = \frac{3}{207} < x_3^* = \frac{203}{207}$ with $\Lambda_3^* = \Lambda_3(X_3^*) = 1.25$ and consider now the polygonal line $f^* \in C(\mathbf{I})$ which connects the 5 pairs of points $(-1,0),(x_1^*,-1)$ and $(x_1^*,-1),(x_2^*,1)$ and $(x_2^*,1),\left(\frac{103}{207},-1\right)$ and $\left(\frac{103}{207},-1\right),(x_3^*,1)$ and $(x_3^*,1),(1,0)$. The (optimal) Lagrange polynomial of degree (at most) 2 which interpolates f^* on X_3^* is readily found to be L_2 with $L_2(x) = (-42849x^2 + 42642x + 39391)/40000$. We deduce, as before, that $\max\limits_{x\in\mathbf{I}} |f^*(x) - L_2(x)| = 2.25$ (this value is attained at $x = \frac{103}{207}$) and $\|f^*\| = 1$ and $P_2^*(x) = 0$ for $x \in \mathbf{I}$, so that $\|f^* - P_2^*\| = \|f^*\| = 1$, and hence (15) holds true. $\square$

The value $\Lambda_3^* = 1.25 = \frac{5}{4}$ was given first by [2] in a footnote on p. 1027, who considered a particular (non-canonical) extremal zero-symmetric node system, see Section 3.2 below. A proof of the equation $\Lambda_3^* = 1.25$ for the CNS X_3^* seems to have appeared first in the book by [30], as part of the solution to Problem 6.42.

3.2. Extremal zero-symmetric node systems. We consider next node systems X_3 having the property $x_1 = -x_3, x_2 = 0$ and $0 < x_3 \le 1$, where we may assume $0 < x_3 < 1$ in view of Section 3.1, but beyond that x_3 is left undetermined. The corresponding Lagrange fundamental polynomials $\ell_{2,j}$ then read, by their definition:

$$(18)\quad \ell_{2,1}(X_3,x) = \frac{x(x-x_3)}{2x_3^2},\quad \ell_{2,2}(X_3,x) = \frac{(x+x_3)(x_3-x)}{x_3^2},\quad \ell_{2,3}(X_3,x) = \frac{x(x+x_3)}{2x_3^2}.$$

It turns out that the infinitely many extremal node systems of this kind are given by

$$(19)\quad X_3^* : x_1^* = -x_3^* < x_2^* = 0 < x_3^* \text{ with } \frac{2\sqrt{2}}{3}(\approx 0.9428) \le x_3^* < 1 \quad \text{and } \Lambda_3(X_3^*) = 1.25.$$

A first proof of this statement seems to be the one given in [30, Problem 6.42]. Actually, Tureckii had stated (19) already in [29, p. 229], but without a proof. To the best of our knowledge, this rare source is not mentioned in the literature on optimal Lagrange interpolation, except for [30]. In [6, p. 65] the statement (19) is posed as advanced Problem 22, but without a solution (the condition $\frac{2\sqrt{2}}{3} < x_3^*$ given in the first edition of [6] was corrected in the second edition to $\frac{2\sqrt{2}}{3} \leq x_3^*$). See also [34] for a discussion of (19). An alternative proof of (19) we provide in Example 3.7 below. In his famous footnote [2] had obviously in mind the following particular extremal zero-symmetric node system $X_3^* : -\frac{2\sqrt{2}}{3} < 0 < \frac{2\sqrt{2}}{3}$, herein after referred to as BNS, which is a marginal case of X_3^* in (19), see also the remark in [17, p. 73]. Although he did not state BNS explicitly, Bernstein did provide both the four extremum points -1, $\xi_1 = -\frac{\sqrt{2}}{3}$, $\xi_2 = \frac{\sqrt{2}}{3}(\approx 0.4714)$, 1 of $\lambda_3(\text{BNS})$, and the minimal Lebesgue constant $\Lambda_3(\text{BNS}) = 1.25$. Since this historical node system is connected to Bernstein's conjecture on the equioscillation of the Lebesgue function, we provide the suppressed calculation of $\Lambda_3(\text{BNS})$:

Example 3.3. Let $X_3 : x_1 = -x_3 < x_2 = 0 < x_3$. Insert $x_3 = \frac{2\sqrt{2}}{3}$ into (18), yielding $\ell_{2,j}(\text{BNS}, x)$. It is then readily found, by inspecting the sign of $\ell_{2,j}$ on consecutive sub-intervals, that the Lebesgue function is given by $\lambda_3(\text{BNS}, x) = \lambda_3(x) =$

$$\begin{aligned}
&\ell_{2,1}(x) - \ell_{2,2}(x) + \ell_{2,3}(x) &&= \frac{9}{4}x^2 - 1, && \text{if } -1 \leq x \leq -x_3,\\
&\ell_{2,1}(x) + \ell_{2,2}(x) - \ell_{2,3}(x) &&= -\frac{9}{8}x^2 - \frac{3}{2\sqrt{2}}x + 1, && \text{if } -x_3 \leq x \leq 0,\\
&-\ell_{2,1}(x) + \ell_{2,2}(x) + \ell_{2,3}(x) &&= -\frac{9}{8}x^2 + \frac{3}{2\sqrt{2}}x + 1, && \text{if } 0 \leq x \leq x_3,\\
&\ell_{2,1}(x) - \ell_{2,2}(x) + \ell_{2,3}(x) &&= \frac{9}{4}x^2 - 1, && \text{if } x_3 \leq x \leq 1.
\end{aligned} \tag{20}$$

On the two boundary intervals, the largest value of λ_3 is $\lambda_3(-1) = \lambda_3(1) = 1.25$, according to Proposition 2.4 (iii). On the two interior intervals, the zeros of the first derivative of λ_3 are indeed the above indicated ξ_1 resp. ξ_2, as follows from (20), and insertion of these values immediately gives $\lambda_3(\xi_1) = \lambda_3(\xi_2) = 1.25$. It thus turns out that we have equioscillation on the four sub-intervals of $\mathbf{I}$ given in (20), see also Proposition 3.10 below.

We also take the opportunity to correct a misprint in the footnote of [2, p. 1027]: The node polynomial $(x-x_1)(x-x_2)(x-x_3)$, denoted there by $A_{n+1}(x)$, $n=2$, should have read

$$A_{n+1}(x) = x\left(x^2-\frac{8}{9}\right), \text{ not } A_{n+1}(x) = \left(x^2-\frac{8}{9}\right). \tag{21}$$

3.3. Extremal general node systems. We finally consider, for $n=3$, all extremal node systems on **I**, including in particular those which are not zero-symmetric. The infinitely many extremal general node systems can be described in terms of the position of the first node, $x_1 = x_1^*$, as follows:

Theorem 3.4. *Let $n=3$. The extremal general node systems on* **I** *are given by*

$$X_3^* : x_1^* < x_2^* = \frac{x_1^*+x_3^*}{2} < x_3^* \quad \textit{with} \quad \Lambda_3^* = \Lambda_3(X_3^*) = 1.25, \tag{22}$$

where

$$(i)\ -1 \le x_1^* \le -\frac{2\sqrt{2}}{3} (\approx -0.9428) \wedge (17-12\sqrt{2})x_1^* + 12\sqrt{2} - 16 \le x_3^* \le 1 \tag{23}$$

or

$$(ii)\ -\frac{2\sqrt{2}}{3} < x_1^* \le 33-24\sqrt{2} (\approx -0.9411) \wedge (17+12\sqrt{2})x_1^* + 12\sqrt{2} + 16 \le x_3^* \le 1. \tag{24}$$

This theorem seems to have appeared first in a paper by [25]. It has been reviewed, including the formulas (23) and (24), in the Mathematical Reviews (MR0374758) and in Zentralblatt MATH (Zbl 0306.41001), but [25] has been referenced neither in the dedicated textbooks on polynomial interpolation [3], [17], [19], [28] nor in the survey article on Lebesgue functions in polynomial interpolation [5]. Also, Tureckii's dedicated problem books on Lagrange interpolation [30], [31] do not cover (23) and (24), nor does the encyclopedia [34]. Schurer himself references to [25] in his valedictory lecture [26]. On the other hand, [25] does not cite Tureckii's result (19). The bound $\le x_3^*$ in (24) reads $< x_3^*$ in [25] (misprint).

To the best of our knowledge, [25] is the only source in literature where extremal zero-asymmetric node systems are mentioned (and provided). It is of course advantageous to have as many extremal node systems at one's disposal as possible: we can then flexibly choose the one which most appropriately fits a concrete approximation problem.

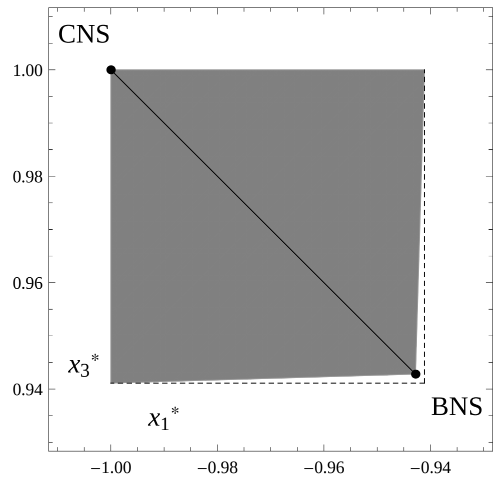

Fig. 2. Each point of the shaded quadrilateral region (which is not a square) gives rise to an extremal node system $x_1^* < x_2^* < x_3^*$ with $x_2^* = \dfrac{x_1^* + x_3^*}{2}$. The diagonal connecting CNS and BNS represents the zero-symmetric node systems (19)

Example 3.5. To give a zero-asymmetric example, we consider the upper right vertex point of the shaded region in Figure 2: Choose $x_1^* = 33 - 24\sqrt{2}(\approx -0.9411), x_3^* = 1$, and hence $x_2^* = 17 - 12\sqrt{2}(\approx 0.0294)$. Inserting these values into (4) with $n = 3$ yields

$$\begin{aligned}
\ell_{2,1}(x) &= \frac{(x-1)(x-(17-12\sqrt{2}))}{1088-768\sqrt{2}}, \\
\ell_{2,2}(x) &= \frac{(x-1)(x-(33-24\sqrt{2}))}{384\sqrt{2}-544}, \\
\ell_{2,3}(x) &= \frac{(x-(17-12\sqrt{2}))(x-(33-24\sqrt{2}))}{1088-768\sqrt{2}}.
\end{aligned} \tag{25}$$

Investigating the zeros of these functions on $\mathbf{I}$ leads to the following representation of the Lebesgue function: $\lambda_3(x) =$

$$\begin{array}{lcll} \ell_{2,1}(x) - \ell_{2,2}(x) + \ell_{2,3}(x), & \text{if} & x \in & [-1, x_1^*], \\ \ell_{2,1}(x) + \ell_{2,2}(x) - \ell_{2,3}(x), & \text{if} & x \in & [x_1^*, x_2^*], \\ -\ell_{2,1}(x) + \ell_{2,2}(x) + \ell_{2,3}(x), & \text{if} & x \in & [x_2^*, 1]. \end{array} \tag{26}$$

On the boundary interval $[-1, x_1^*]$ the largest value of λ_3 is $\lambda_3(-1)$, according to Proposition 2.4 (iii), and it is readily verified that we get $\lambda_3(-1) = 1.25$. On the remaining two intervals the zeros of the first derivative of λ_3 are $\xi_1 = 25 - 18\sqrt{2}(\approx -0.4558)$ and $\xi_2 = 9 - 6\sqrt{2}(\approx 0.5147)$, as is readily deduced from (26). And a straightforward insertion yields $\lambda_3(\xi_1) = \lambda_3(\xi_2) = 1.25$, so that we have equioscillation on the three sub-intervals in (26).

Schurer's proof [25] of Theorem 3.4 is not given in full detail but is sketched, using bridging phrases such as *it is easy to verify* or *we omit the calculational details* or *after elementary but somewhat tediuos calculations*, which make it hard to follow, for students and lecturers alike.

Our first humble contribution to optimal quadratic Lagrange interpolation is therefore to give a sound new proof of Schurer's result (23), (24). The first part of that proof follows the outline given by Schurer and leads us to the crucial two inequalities stated right before Theorem 1 in [25]. For the transformation of these inequalities to formulae (23), (24) we then use symbolic computation with quantifier elimination.

3.4. Alternative proof of Theorem 3.4 with the aid of symbolic computation (1).

Proof of Theorem 3.4. It suffices to consider node systems which do not contain an endpoint of $\mathbf{I}$, i.e., $X_3 : -1 < x_1 < x_2 < x_3 < 1$. For if both endpoints are included, then the only extremal node system is the extremal CNS $X_3^* : x_1^* = -1 < x_2^* = 0 < x_3^* = 1$. If only one endpoint is included, then the following argument can be traced down with obvious modifications (compare with Example 3.5). The Lebesgue function $\lambda_3(X_3, x) = \sum_{j=1}^{3} \prod_{i=1, i\neq j}^{3} \frac{|x - x_i|}{|x_j - x_i|}$ reads when $x \in (-1, x_1) = \mathbf{I1}$, in view of $\max(x-x_1, x-x_2, x-x_3, x_1-x_2, x_1-x_3, x_2-x_3) < 0$,

$$\lambda_3(X_3, x) = \frac{(x_2 - x)(x_3 - x)}{(x_2 - x_1)(x_3 - x_1)} + \frac{(x_1 - x)(x_3 - x)}{(x_2 - x_1)(x_3 - x_2)} + \frac{(x_1 - x)(x_2 - x)}{(x_3 - x_1)(x_3 - x_2)}. \tag{27}$$

For $x \in (x_1, x_2) = \mathbf{I2}$ we get, in view of

$$\max(x - x_2, x - x_3, x_1 - x_2, x_1 - x_3, x_2 - x_3) < 0 < x - x_1,$$

$$\lambda_3(X_3, x) = \frac{(x_2 - x)(x_3 - x)}{(x_2 - x_1)(x_3 - x_1)} + \frac{(x - x_1)(x_3 - x)}{(x_2 - x_1)(x_3 - x_2)} + \frac{(x - x_1)(x_2 - x)}{(x_3 - x_1)(x_3 - x_2)}. \tag{28}$$

For $x \in (x_2, x_3) = \mathbf{I3}$ we get, in view of

$$\max(x - x_3, x_1 - x_2, x_1 - x_3, x_2 - x_3) < 0 < \min(x - x_1, x - x_2),$$

$$\text{(29)}\quad \lambda_3(X_3, x) = \frac{(x - x_2)(x_3 - x)}{(x_2 - x_1)(x_3 - x_1)} + \frac{(x - x_1)(x_3 - x)}{(x_2 - x_1)(x_3 - x_2)} + \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}.$$

Finally, for $x \in (x_3, 1) = \mathbf{I4}$ we get, in view of

$$\max(x_1 - x_2, x_1 - x_3, x_2 - x_3) < 0 < \min(x - x_1, x - x_2, x - x_3),$$

$$\text{(30)}\quad \lambda_3(X_3, x) = \frac{(x - x_2)(x - x_3)}{(x_2 - x_1)(x_3 - x_1)} + \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_3 - x_2)} + \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}.$$

We consider next the maximum of $\lambda_3(X_3, x)$ on the sub-intervals **I1**, **I2**, **I3** and **I4** of **I**:
$\lambda(X_3, x)$ will be largest on **I1** at $x = -1$, according to Proposition 2.4 (iii). We thus obtain, inserting $x = -1$,

$$\text{(31)}\qquad \max_{x \in \mathbf{I1}} \lambda_3(X_3, x) = \frac{2 - x_2^2 + 2x_3 + x_2 x_3 + x_1(2 + x_2 + x_3)}{(x_1 - x_2)(x_2 - x_3)}.$$

The maximum of $\lambda_3(X_3, x)$ on **I2** will be attained at $x = \dfrac{x_1 + x_2}{2}$ since the first derivative of $\lambda_3(X_3, x)$, which reads $\dfrac{2(-2x + x_1 + x_2)}{(x_1 - x_3)(x_2 - x_3)}$, vanishes there. We thus obtain, inserting $x = \dfrac{x_1 + x_2}{2}$,

$$\text{(32)}\qquad \max_{x \in \mathbf{I2}} \lambda_3(X_3, x) = \frac{x_1^2 + x_2^2 - 2x_1 x_3 - 2x_2 x_3 + 2x_3^2}{2(x_1 - x_3)(x_2 - x_3)}.$$

The maximum of $\lambda_3(X_3, x)$ on **I3** will be attained at $x = \dfrac{x_2 + x_3}{2}$ since the first derivative of $\lambda_3(X_3, x)$, which reads $\dfrac{2(-2x + x_2 + x_3)}{(x_1 - x_2)(x_1 - x_3)}$, vanishes there. We thus obtain, inserting $x = \dfrac{x_2 + x_3}{2}$,

$$\text{(33)}\qquad \max_{x \in \mathbf{I3}} \lambda_3(X_3, x) = \frac{2x_1^2 + x_2^2 - 2x_1 x_2 - 2x_1 x_3 + x_3^2}{2(x_1 - x_2)(x_2 - x_3)}.$$

Finally, the maximum of $\lambda_3(X_3, x)$ on **I4** will be attained at $x = 1$ according to Proposition 2.4 (iii). We thus obtain, inserting $x = 1$,

$$\text{(34)}\qquad \max_{x \in \mathbf{I4}} \lambda_3(X_3, x) = \frac{2 - x_2^2 - 2x_3 + x_2 x_3 + x_1(-2 + x_2 + x_3)}{(x_1 - x_2)(x_2 - x_3)}.$$

We now map X_3 onto $\mathbf{I}$ by means of the linear function $f(x) = \dfrac{2}{x_3 - x_1}(x - x_3) + 1$, which satisfies $f(x_1) = -1$ and $f(x_3) = 1$. The mapped node system $X_3' : -1 = x_1' < f(x_2) = x_2' < 1 = x_3'$ leads to a Lebesgue constant which is not larger than $\Lambda_3(X_3)$. In fact, we obtain with $x' = \dfrac{2}{x_3 - x_1}(x - x_3) + 1$, following [24, p. 100],

$$
\begin{aligned}
\Lambda_3(X_3') &= \max_{x' \in \mathbf{I}} \sum_{j=1}^{3} \prod_{i=1, i \neq j}^{3} \frac{|x' - x_i'|}{|x_j' - x_i'|} \\
&= \max_{x_1 \leq x \leq x_3} \sum_{j=1}^{3} \prod_{i=1, i \neq j}^{3} \frac{|\frac{2}{x_3 - x_1}(x - x_3) + 1 - \frac{2}{x_3 - x_1}(x_i - x_3) - 1|}{|\frac{2}{x_3 - x_1}(x_j - x_3) + 1 - \frac{2}{x_3 - x_1}(x_i - x_3) - 1|} \\
&= \max_{x_1 \leq x \leq x_3} \sum_{j=1}^{3} \prod_{i=1, i \neq j}^{3} \frac{|x - x_i|}{|x_j - x_i|} \\
&\leq \max_{x \in \mathbf{I}} \sum_{j=1}^{3} \prod_{i=1, i \neq j}^{3} \frac{|x - x_i|}{|x_j - x_i|} = \Lambda_3(X_3).
\end{aligned}
\tag{35}
$$

If we assume that X_3 is an extremal node system with minimal Lebesgue constant $\Lambda_3^* = 1.25$, then we must have $\Lambda_3(X_3') = \Lambda_3(X_3) = 1.25$ since $\Lambda_3(X_3') < \Lambda_3(X_3)$ would contradict the extremality of X_3. This means that X_3' is actually an extremal CNS, so that we must necessarily have, by the uniqueness condition (see Section 3.1), $x_2' = f(x_2) = \dfrac{2}{x_3 - x_1}(x_2 - x_3) + 1 = 0$ which in turn implies $x_2 = \dfrac{x_1 + x_3}{2}$. Hence the assumption that X_3 is extremal allows us to eliminate the variable x_2 in the maxima defined above. We thus eventually get, inserting $x_2 = \dfrac{x_1 + x_3}{2}$:

$$
\max_{x \in \mathbf{I1}} \lambda_3(X_3, x) = \frac{8 + x_1^2 + 8x_3 + x_3^2 + x_1(8 + 6x_3)}{(x_1 - x_3)^2},
\tag{36}
$$

$$
\max_{x \in \mathbf{I2}} \lambda_3(X_3, x) = 1.25,
\tag{37}
$$

$$
\max_{x \in \mathbf{I3}} \lambda_3(X_3, x) = 1.25,
\tag{38}
$$

$$
\max_{x \in \mathbf{I4}} \lambda_3(X_3, x) = \frac{8 + x_1^2 - 8x_3 + x_3^2 + x_1(-8 + 6x_3)}{(x_1 - x_3)^2}.
\tag{39}
$$

To ensure that X_3 is indeed an extremal node system, we must impose on the values in (36) and (39) the condition that they, too, cannot exceed 1.25, that is

$$(40)\quad \frac{8+x_1^2+8x_3+x_3^2+x_1(8+6x_3)}{(x_1-x_3)^2} \le 1.25$$
$$\wedge\ \frac{8+x_1^2-8x_3+x_3^2+x_1(-8+6x_3)}{(x_1-x_3)^2} \le 1.25,$$

or

$$(41)\quad \frac{8+x_1^2+8x_3+x_3^2+x_1(8+6x_3)}{(x_1-x_3)^2} - 1 \le 0.25$$
$$\wedge\ \frac{8+x_1^2-8x_3+x_3^2+x_1(-8+6x_3)}{(x_1-x_3)^2} - 1 \le 0.25.$$

We readily observe that the last two inequalities coincide with the crucial two inequalities stated right before Theorem 1 in [25], that is with:

$$(42)\quad \frac{8(1+x_1)(1+x_3)}{(x_1-x_3)^2} \le 0.25 \text{ and simultaneously } \frac{8(1-x_1)(1-x_3)}{(x_1-x_3)^2} \le 0.25.$$

The final step is now to translate (42) into (23) and (24), a step whose calculation has not been revealed by Schurer. A convenient and powerful way to do so is by symbolic computation. We have employed the computer algebra system Mathematica, see [35], and its function **Reduce**[*expr, vars, dom*] which reduces, over the domain *dom*, the statement *expr* by solving equations or inequalities for *vars* and eliminating quantifiers, see e.g. [27]. On a standard PC, the required computation time is less than 1 second. The results have been cross-checked with the technical computing software QEPCAD, see [4]. Note that below the term beneath the radical sign is $(x_1-1)^2$ respectively $(x_1+1)^2$:

$$\mathbf{Reduce}[32(1+x_1)(1+x_3) \le (-x_1+x_3)^2 \wedge 32(1-x_1)(1-x_3) \le (-x_1+x_3)^2$$
$$\wedge -1 < x_1 < x_3 < 1, \{x_1, x_3\}, \mathrm{Reals}]$$

The output is a two-staged description of the solution set, which is obviously identical with Theorem 3.4 (where, however, the endpoints of $\mathbf{I}$ are also included):

(43)
$$\left(-1 < x_1 \le \frac{-2\sqrt{2}}{3} \wedge -16+17x_1+12\sqrt{2}\sqrt{1-2x_1+x_1^2} \le x_3 < 1\right)$$
$$\vee$$
$$\left(\frac{-2\sqrt{2}}{3} < x_1 < 3(11-8\sqrt{2}) \wedge 16+17x_1+12\sqrt{2}\sqrt{1+2x_1+x_1^2} \le x_3 < 1\right).\square$$

3.5. Alternative proof of Theorem 3.4 with the aid of symbolic computation (2). Our second contribution to optimal quadratic Lagrange interpolation is to give an alternative new proof of (22), (23), (24) that is based solely on symbolic computation, which underlines the power of the method. We have again employed the computer algebra system Mathematica, this time with its functions **CylindricalDecomposition** [*ineqs*, *x1, x2,*...], which finds a decomposition of the region represented by the inequalities *ineqs* into cylindrical parts *xi*, and **Resolve** [*expr*, *dom*], which attempts to resolve the expression *expr* into a form that eliminates, over the domain *dom*, the quantifiers ∀ and ∃. We so directly arrive at (22), (23), (24). On a standard PC, the required computation time is about 10 minutes. Again we have cross-checked the result with QEPCAD.

P r o o f o f T h e o r e m 3.4. We first introduce an auxiliary function to define the Lagrange fundamental polynomials:

$$\mathrm{LagrFund}[\mathrm{i\_}, \mathrm{x\_List}, \mathrm{var\_}] := \left(\prod_{i=1}^{j-1} \frac{var - x[[i]]}{x[[j]] - x[[i]]}\right)\left(\prod_{i=j+1}^{\mathrm{Length}[x]} \frac{var - x[[i]]}{x[[j]] - x[[i]]}\right).$$

By evaluating the expression

(44)

$$\textbf{CylindricalDecomposition}[\textbf{Resolve}[-1 \le x_1 < x_2 < x_3 \le 1 \wedge \forall x(-1 \le x \le 1) \Rightarrow \sum_{j=1}^{3} |\mathrm{LagrFund}(j, \{x_1, x_2, x_3\}, x)| \le \frac{5}{4}, \mathrm{Reals}], \{x_1, x_3, x_2\}],$$

we get a four-staged description of the solution set as

(45)

$$\left(x_1 = -1 \wedge 24\sqrt{2} - 33 \le x_3 \le 1 \wedge x_2 = \frac{x_3 - 1}{2}\right)$$
$$\vee$$
$$\left(-1 < x_1 \le \frac{-2\sqrt{2}}{3} \wedge -12\sqrt{2}x_1 + 17x_1 + 12\sqrt{2} - 16 \le x_3 \le 1 \wedge x_2 = \frac{x_1 + x_3}{2}\right)$$
$$\vee$$
$$\left(\frac{-2\sqrt{2}}{3} < x_1 < 33 - 24\sqrt{2} \wedge 12\sqrt{2}x_1 + 17x_1 + 12\sqrt{2} + 16 \le x_3 \le 1 \wedge x_2 = \frac{x_1 + x_3}{2}\right)$$
$$\vee$$
$$\left(x_1 = 33 - 24\sqrt{2} \wedge x_3 = 1 \wedge x_2 = \frac{1}{2}\left(34 - 24\sqrt{2}\right)\right).$$

This output is obviously identical with Theorem 3.4. □

3.6. New description of extremal general node systems with the aid of two parameters. Our third contribution to optimal quadratic Lagrange interpolation is the provision of an alternative, but equivalent, description of the extremal general node systems. In Schurer's description (Theorem 3.4), x_1^* is chosen from an interval with fixed endpoints, and x_3^* is chosen from an interval whose right endpoint is 1, whereas the left endpoint is a linear function of x_1^* (and we have $x_2^* = \dfrac{x_1^* + x_3^*}{2}$). In our description, two parameters (α and β) will be freely chosen from two disjoint intervals (which are zero-symmetric to each other) with fixed endpoints, and then x_1^* and x_3^* are computed as functions of α and β (we also have $x_2^* = \dfrac{x_1^* + x_3^*}{2}$). This description has been inspired by the proof of Theorem 2 in [16]. To avoid ambiguity, in our description we denote the extremal nodes by y_i^* rather than by x_i^*.

Theorem 3.6. *Set* $a = \dfrac{-3}{2\sqrt{2}}$ *and* $b = \dfrac{3}{2\sqrt{2}} (\approx 1.0607)$. *For any* $\alpha \in [a, -1]$ *and for any* $\beta \in [1, b]$, *the linear function* $S : [\alpha, \beta] \to \mathbf{I}$ *given by* $S(x) = \dfrac{1}{\beta - \alpha}(2x - \alpha - \beta)$ *determines an extremal node system* $Y_{\alpha,\beta}^* : -1 \leq y_1^* < y_2^* < y_3^* \leq 1$ *on* $\mathbf{I}$, *if one consecutively maps the members of the extremal CNS* X_3^* : $x_1^* = -1 < x_2^* = 0 < x_3^* = 1$, *i.e., if one inserts* $x = x_1^* = 1$, *and* $x = x_2^* = 0$ *and* $x = x_3^* = 1$, *which yields*

$$S(x_1^*) = y_1^* = \frac{1}{\beta - \alpha}(-2 - \alpha - \beta), \tag{46}$$

$$S(x_2^*) = y_2^* = \frac{y_1^* + y_3^*}{2} = \frac{-\alpha - \beta}{\beta - \alpha}, \tag{47}$$

$$S(x_3^*) = y_3^* = \frac{1}{\beta - \alpha}(2 - \alpha - \beta). \tag{48}$$

Proof. We have $S(\alpha) = -1$ and $S(\beta) = 1$, so that $S([\alpha, \beta]) = \mathbf{I}$, and the inverse function $S^{-1} : \mathbf{I} \to [\alpha, \beta]$ is given by $S^{-1}(y) = \dfrac{1}{2}((\beta - \alpha)y + \alpha + \beta)$ with $S^{-1}(-1) = \alpha$ and $S^{-1}(1) = \beta$. Since S is linear with positive slope $\dfrac{2}{\beta - \alpha}$, the ordering of $X_3^* : x_1^* = -1 < x_2^* = 0 < x_3^* = 1$ on $\mathbf{I}$ translates to $Y_{\alpha,\beta}^*$, so that indeed $y_1^* < y_2^* < y_3^*$ holds. The Lebesgue function corresponding to X_3^* reads, see (12) and (13),

(49) $\lambda_3(X_3^*, x) = |x(x-1)|/2 + |1-x^2| + |x(x+1)|/2$
with $\Lambda_3^* = \max_{x \in \mathbf{I}} \lambda_3(X_3^*, x) = 1.25$

We wish to determine the point $b > 1$ where $\lambda_3(X_3^*, x)$ intersects with the constant function $c(x) = \Lambda_3^* = 1.25$. According to Proposition 2.4, $\lambda_3(X_3^*, x)$ is monotone increasing on $(1, \infty)$, and in fact is represented there by $\lambda_3(X_3^*, x) = 2x^2 - 1$, see (49). The positive solution to the equation $2x^2 - 1 = 1.25$ is $x = b = \dfrac{3}{2\sqrt{2}} = \sqrt{\dfrac{9}{8}}$. In a similar fashion we determine the point $a < -1$, where $\lambda_3(X_3^*, x)$ intersects with the constant function $c(x) = \Lambda_3^* = 1.25$ and obtain $x = a = -\dfrac{3}{2\sqrt{2}} = -\sqrt{\dfrac{9}{8}}$. By construction, $\lambda_3(X_3^*, x) \leq 1.25$ on $[a, b]$, and hence $\lambda_3(X_3^*, x) \leq 1.25$ on any subinterval $[\alpha, \beta]$ of $[a, b]$, where $\alpha \in [a, -1]$ and $\beta \in [1, b]$. Equality $\lambda_3(X_3^*, x) = 1.25$ occurs, if $x \in \{a, -0.5, 0.5, b\}$.
We finally wish to verify that the node system $Y_{\alpha,\beta}^* : -1 \leq y_1^* < y_2^* < y_3^* \leq 1$ as given in (46)–(48) is an extremal node system on $\mathbf{I}$, for every choice of $\alpha \in [a, -1]$ and $\beta \in [1, b]$. To this end, we observe that, for $y \in \mathbf{I}$,

$$\lambda_3(Y_{\alpha,\beta}^*, y) = \sum_{j=1}^{3} \prod_{i=1, i\neq j}^{3} \frac{|y - y_i^*|}{|y_j^* - y_i^*|} = \sum_{j=1}^{3} \prod_{i=1, i\neq j}^{3} \frac{|S(x) - S(x_i^*)|}{|S(x_j^*) - S(x_i^*)|} =$$

$$\sum_{j=1}^{3} \prod_{i=1, i\neq j}^{3} \frac{|\frac{1}{\beta-\alpha}(2x - \alpha - \beta) - \frac{1}{\beta-\alpha}(2x_i^* - \alpha - \beta)|}{|\frac{1}{\beta-\alpha}(2x_j^* - \alpha - \beta) - \frac{1}{\beta-\alpha}(2x_i^* - \alpha - \beta)|} = \sum_{j=1}^{3} \prod_{i=1, i\neq j}^{3} \frac{|x - x_i^*|}{|x_j^* - x_i^*|} =$$

$$\sum_{j=1}^{3} \prod_{i=1, i\neq j}^{3} \frac{|S^{-1}(y) - S^{-1}(y_i^*)|}{|S^{-1}(y_j^*) - S^{-1}(y_i^*)|} = \lambda_3(X_3^*, x),$$

where $x \in [\alpha, \beta]$, and hence $\max_{y \in \mathbf{I}} \lambda_3(Y_{\alpha,\beta}^*, y) = \max_{x \in [\alpha,\beta]} \lambda_3(X_3^*, x) = 1.25 = \Lambda_3^*$. $\square$

Example 3.7. Set $\alpha = a = \dfrac{-3}{2\sqrt{2}}$. Then $y_1^* = \dfrac{-8 + 3\sqrt{2} - 4\beta}{3\sqrt{2} + 4\beta}$ and $y_3^* = \dfrac{8 + 3\sqrt{2} - 4\beta}{3\sqrt{2} + 4\beta}$, and hence
$y_1^* \in [-\dfrac{2\sqrt{2}}{3}, -(24\sqrt{2} - 33)]$ and $y_3^* \in [\dfrac{2\sqrt{2}}{3}, 1]$, if β varies in $[1, b]$.
Set $\alpha = -1$. Then $y_1^* = -1$ and $y_3^* = \dfrac{3 - \beta}{\beta + 1}$, and hence
$y_3^* \in [24\sqrt{2} - 33, 1]$, if β varies in $[1, b]$.

Set $\beta = 1$. Then $y_1^* = \dfrac{3+\alpha}{\alpha-1}$ and $y_3^* = 1$, and hence
$y_1^* \in [-1, -(24\sqrt{2}-33)]$, if α varies in $[a, -1]$.

Set $\beta = b = \dfrac{3}{2\sqrt{2}}$. Then $y_1^* = \dfrac{3+2\sqrt{2}(2+\alpha)}{2\alpha\sqrt{2}-3}$ and $y_3^* = \dfrac{3-2\sqrt{2}(2-\alpha)}{2\alpha\sqrt{2}-3}$, and hence
$y_1^* \in [-1, -\dfrac{2\sqrt{2}}{3}]$ and $y_3^* \in [24\sqrt{2}-33, \dfrac{2\sqrt{2}}{3}]$, if α varies in $[a, -1]$.

Set $\alpha = -\beta$. Then $y_1^* = -y_3^*$, $y_2^* = 0$, $y_3^* = \dfrac{1}{\beta}$, and hence $y_3^* \in [\dfrac{2\sqrt{2}}{3}, 1]$, if β varies in $[1, b]$. This provides an alternative proof for Tureckii's result (19) which describes the extremal zero-symmetric node systems.

Example 3.8. We now choose fixed values for both α and β. Set $\alpha = -1.05$ and $\beta = 1.02$, say. The resulting node system, which is asymmetric to zero and does not include an endpoint of $\mathbf{I}$, is
$Y^*_{-1.05,1.02} : y_1^* = -\dfrac{197}{207}(\approx -0.9517) < y_2^* = \dfrac{3}{207}(\approx 0.0145) < y_3^* = \dfrac{203}{207}(\approx 0.9807).$
We leave it to the reader to verify that

$$\begin{aligned}
\ell_{2,1} &= (42849y^2 - 42642y + 609)/80000,\\
\ell_{2,2} &= (-85698y^2 + 2484y + 79982)/80000,\\
\ell_{2,3} &= (42849y^2 + 40158y - 591)/80000,
\end{aligned}$$

and that $\max_{y\in\mathbf{I}} \lambda_3(Y^*_{-1.05,1.02}, y) = \max_{y\in\mathbf{I}} \sum_{j=1}^{3} |\ell_{2,j}(Y^*_{-1.05,1.02}, y)| = 1.25 = \Lambda_3^*$ is attained on $\mathbf{I}$ at $y = \xi_1 = -\dfrac{97}{207}(\approx -0.4686)$ and at $y = \xi_2 = \dfrac{103}{207}(\approx 0.4976)$, but not at $y = \pm 1$.

We are now going to verify that our description of extremal general node systems produces the same result as Schurer's:

Proposition 3.9. *Theorems* 3.4 *and* 3.6 *are equivalent.*

Proof. Since Theorem 3.4 describes all extremal general node systems, the set of extremal node systems y_1^*, y_2^*, y_3^* according to Theorem 3.6 must be a subset of the extremal node systems x_1^*, x_2^*, x_3^* according to Theorem 3.4. On the other hand, it suffices to show that all extremal node systems, as given in

Theorem 3.4, can be obtained as the image of the CNS under the linear map S. We will again use quantifier elimination. In view of $y_2^* = \dfrac{y_1^* + y_3^*}{2}$, it suffices to consider the image y_1^* of -1 and y_3^* of 1. By existential elimination of the variables α, β from the formula ϕ given in (50) below,

$$\begin{gathered} -\frac{3}{2\sqrt{2}} \le \alpha \le -1 \wedge 1 \le \beta \le \frac{3}{2\sqrt{2}} \\ \wedge \\ y_1^*(\beta - \alpha) = (-2 - \alpha - \beta) \wedge y_3^*(\beta - \alpha) = (2 - \alpha - \beta), \end{gathered} \tag{50}$$

i.e., after the Mathematica call **Resolve**[Exists[$\{\alpha, \beta\}, \phi$], $\{y_1^*, y_3^*\}$, Reals], we get exactly the same 2D region as given in (23)–(24):

$$\begin{gathered} \left(-1 \le y_1^* \le -\frac{2\sqrt{2}}{3} \wedge \frac{8 - 4y_1^* + 3\sqrt{2}y_1^*}{4 + 3\sqrt{2}} \le y_3^* \le 1\right) \\ \vee \\ \left(-\frac{2\sqrt{2}}{3} < y_1^* < \frac{-12 + 3\sqrt{2}}{4 + 3\sqrt{2}} \wedge \frac{8 + 4y_1^* + 3\sqrt{2}y_1^*}{-4 + 3\sqrt{2}} \le y_3^* \le 1\right) \\ \vee \\ \left(y_1^* = \frac{-12 + 3\sqrt{2}}{4 + 3\sqrt{2}} \wedge y_3^* = 1\right). \end{gathered} \tag{51}$$

□

Remark 2 in [25], given there without proof, states a maximal equioscillation property of λ_3 on the particular extremal node system BNS, see Section 3.2. We verify this statement here, for the reader's convenience.

Proposition 3.10. *The only extremal node system* $X_3 : -1 < x_1 < x_2 < x_3 < 1$ *having the property that the maximum of* $\lambda_3(X_3, x)$ *equals* $\Lambda_3^* = 1.25$ *on all four sub-intervals* **I1**, **I2**, **I3** *and* **I4** *of* **I**, *is the BNS* $X_3^* : -\dfrac{2\sqrt{2}}{3} < 0 < \dfrac{2\sqrt{2}}{3}$.

P r o o f. We know from Section 3.4 that $x_2 = \dfrac{x_1 + x_3}{2}$ and that on the two interior sub-intervals **I2** and **I3** the maximum of $\lambda_3(X_3, x)$ always coincides with $\Lambda_3^* = 1.25$ if $X_3 : -1 < x_1 < x_2 < x_3 < 1$ is extremal, see (37), (38). If we impose the additional condition that the values in (36) and (39) be equal then we get, after some elementary algebraic manipulation, that $x_1 = -x_3$. If we furthermore require that the values in (36) and (39) be equal and coincide with $\Lambda_3^* = 1.25$ then it follows from $\dfrac{8 + x_1^2 + 8x_3 + x_3^2 + x_1(8 + 6x_3)}{(x_1 - x_3)^2} = 1.25$ respectively from

$\dfrac{8 + x_1^2 - 8x_3 + x_3^2 + x_1(-8 + 6x_3)}{(x_1 - x_3)^2} = 1.25$, with $x_1 = -x_3$, that $x_3 = \dfrac{2\sqrt{2}}{3}$ must hold. Consequently, X_3 is the BNS. $\square$

4. Concluding remarks.

Remark 4.1. The minimal Lebesgue constant (for $n = 3$: $\Lambda_3^* = 1.25$) is also called minimal interpolating projection constant. It is surely an upper bound for the (general) minimal projection constant, but the minimal projection constant, for $n = 3$, is slightly smaller than 1.25, see [11], [12] for details. Incidentally, Theorem 18 in [12], stated there without proof, has been established already in [20].

Remark 4.2. If we consider optimal polynomial Lagrange interpolation with $n \geq 4$, then very little seems to be known about the analytic expressions of extremal node systems and minimal Lebesgue constants. At least for the cubic case $n = 4$ the analytic form of the unique extremal and canonical node system and of the minimal Lebesgue constant Λ_4^* has been determined, see [21], [22]. In a prospective paper [23] we intend, with the aid of symbolic computation, to shed some more light on the cubic case and we hope to achieve some progress for the next low-degree polynomial cases as well. Since quantifier elimination was completely sufficient for resolving the optimal quadratic interpolation problem, we did not consider other symbolic tools in this paper. However, in the course of the investigation of the cubic, quartic and quintic cases we will also use Groebner bases and resultants. The reader is invited to visit our online repository at `www.math.u-szeged.hu/~vajda/Leb/` where we maintain information on optimal low-degree Lagrange interpolation.

REFERENCES

[1] Babuška I., T. Strouboulis. The finite element method and its reliability. Oxford University Press, Oxford, UK, 2001.

[2] Bernstein S. Sur la limitation des valeurs d'un polynôme $P_n(x)$ de degré n sur tout un segment par ses valeurs en $(n+1)$ points du segment. *Izv. Akad. Nauk. SSSR*, **7** (1931), 1025–1050 (in French).

[3] Bojanov B. D. Interpolation. Narodna Prosveta, Sofia, 1984 (in Bulgarian).

[4] Brown C. W. An overview of QEPCAD B: A tool for real quantifier elimination and formula simplification. *J. JSSAC* , **10** (2003), 13–22.

[5] Brutman L. Lebesgue functions for polynomial interpolation – a survey. *Ann. Numer. Math*, **4** (1997), 111–127.

[6] Cheney E. W. Introduction to approximation theory. McGraw-Hill Book Co. (New York). Second Edition, AMS Chelsea Publishing (Providence, Rhode Island), 2002.

[7] Cheney E. W., P. D. Morris. The numerical determination of projection constants. In: Proceedings of the Conference Numerische Methoden und Approximationstheorie Band 2 (Eds L. Collatz, G. Meinardus), Oberwolfach, Birkhäuser, Basel, 1973, *ISNM*, **26** (1975), 29–40.

[8] Cheney E. W., W. A. Light. A course in approximation theory. AMS Publishing, Providence, Rhode Island, 2000.

[9] De Boor C., A. Pinkus. Proof of the conjectures of Bernstein and Erdős concerning the optimal nodes for polynomial interpolation. *J. Approx. Theory*, **24** (1978), 289–303.

[10] Erdős P. Problems and results in the theory of interpolation I. *Acta Math. Acad. Sci. Hung.*, **9** (1958), 381–388.

[11] Foucart S. On the best conditioned bases for quadratic polynomials. *J. Approx. Theory*, **130** (2004), 46–56.

[12] Foucart S. Allometry constants of finite-dimensional spaces: theory and computations. *Numer. Math.*, **112** (2009), 535–564.

[13] Kilgore T. A., E. W. Cheney. A theorem on interpolation in Haar subspaces. *Aequationes Math.*, **14** (1976), 391–400.

[14] Kilgore T. A. A characterization of the Lagrange interpolating projection with minimal Tchebycheff norm. *J. Approx. Theory.* **24** (1978), 273–288.

[15] Lorentz G. G., K. Jetter, S. D. Riemenschneider. Birkhoff interpolation. Cambridge University Press, Cambridge, UK, 1984.

[16] Luttmann F. W., Th. J. Rivlin. Some numerical experiments in the theory of polynomial interpolation. *IBM J. Res. Develop.*, **9** (1965), 187–191.

[17] Mastroianni G., G. V. Milovanović. Interpolation processes. Basic theory and applications, Springer, Berlin, 2008.

[18] Meijering E. A chronology of interpolation: From ancient astronomy to modern signal and image processing. In: Proceedings of the IEEE, vol. **90**, 2002, 319–342.

[19] Phillips G. M. Interpolation and approximation by polynomials. Springer, New York, 2003.

[20] Rack H.-J. Extreme Punkte in der Einheitskugel des Vektorraumes der trigonometrischen Polynome. *Elem. Math.*, **37** (1982), 164–165.

[21] Rack H.-J. An example of optimal nodes for interpolation. *Int. J. Math. Educ. Sci. Technol.*, **15** (1984), 355–357.

[22] Rack H.-J. An example of optimal nodes for interpolation revisited. In: Advances in applied mathematics and approximation theory, Contributions from AMAT 2012 (Eds G. A. Anastassiou, O. Duman), *Springer Proceedings in Mathematics and Statistics*, **41**, (2013), Springer, New York, 117–120.

[23] Rack H.-J., R. Vajda. Optimal cubic Lagrange interpolation: Extremal node systems with minimal Lebesgue constant via symbolic computation (submitted).

[24] Rivlin T. J. An introduction to the approximation of functions. Corrected republication of the 1969 original by Blaisdell Publishing Company, Dover Publications, New York, 1981.

[25] Schurer F. A remark on extremal sets in the theory of polynomial interpolation. *Stud. Sci. Math Hung.*, **9** (1974), 77–79.

[26] Schurer F. Omzien in tevredenheid. Afscheidscollege uitgesproken op 10 november 2000 aan de Technische Universiteit Eindhoven, 2000 (in Dutch). `http://repository.tue.nl/540800`

[27] Strzeboński A. Solving systems of strict polynomial inequalities. *J. Symb. Comput.*, **29** (2000), 471–480.

[28] Szabados J., P. Vértesi. Interpolation of functions. World Scientific, Singapore, 1990.

[29] Tureckii A. H. On certain extremal problems in the theory of interpolation. In: Proceedings of the Second All-Union Conference, 1962, Studies of contemporary problems in constructive function theory (Ed. I. I. Ibragimov), 220–232, Izdat. Akad. Nauk. Azerbaidzan SSR, Baku, 1965 (in Russian).

[30] Tureckii A. H. The theory of interpolation in problem form. Vol I, Izdat. Vyšejšaja Skola, Minsk, 1968 (in Russian).

[31] Tureckii A. H. The theory of interpolation in problem form. Vol II, Izdat. Vyšejšaja Skola, Minsk, 1977 (in Russian).

[32] Vajda R. Lebesgue constants and optimal node systems via symbolic computations (short paper). In: Proceedings of the Fifth International Symposium on Symbolic Computation (SCSS 2013) (Eds L. Kovacs, T. Kutsia), RISC-Linz Report Series No. 13-06, 2013, 125–125.

[33] Wikipedia (the free encyclopedia). Henri Lebesgue. `http://en.wikipedia.org/wiki/Henri_Lebesgue`

[34] Wikipedia (the free encyclopedia). Lebesgue constant(interpolation). `http://en.wikipedia.org/wiki/Lebesgue_constant_(interpolation)`

[35] Wolfram Research. Inc. Mathematica, Version 9.0, Champaign, IL., 2012.

Heinz-Joachim Rack
Steubenstrasse 26 a
D-58097 Hagen, Germany
email: `heinz-joachim.rack@drrack.com`

Robert Vajda
Bolyai Institute, University of Szeged
Aradi Vértanúk tére 1,
H-6720 Szeged, Hungary
email: `vajdar@math.u-szeged.hu`

Received July 30, 2014
Final Accepted November 6, 2014