On Nodal Point Sets for Flux Reconstruction

F. D. Witherden and P. E. Vincent

March 28, 2020

Nodal point sets, and associated collocation projections, play an important role in a range of high-order methods, including Flux Reconstruction (FR) schemes. Historically, efforts have focused on identifying nodal point sets that aim to minimise the L^∞ error of an associated interpolating polynomial. The present work combines a comprehensive review of known approximation theory results, with new results, and numerical experiments, to motivate that in fact point sets for FR should aim to minimise the L^2 error of an associated interpolating polynomial. New results include identification of a nodal point set that minimises the L^2 norm of an interpolating polynomial, and a proof of the equivalence between such an interpolating polynomial and an L^2 approximating polynomial with coefficients obtained using a Gauss–Legendre quadrature rule. Numerical experiments confirm that FR errors can be reduced by an order-of-magnitude by switching from popular point sets such as Chebyshev, Chebyshev–Lobatto and Legendre–Lobatto to Legendre point sets.

1 Introduction

Theoretical studies and numerical experiments suggest that unstructured high-order methods could provide a route to solving otherwise intractable fluid flow problems within the vicinity of complex geometries [1, 2]. In 2007 Huynh proposed the flux reconstruction (FR) approach [3], a unifying framework for high-order schemes that encompasses several existing methods while simultaneously admitting an efficient implementation. Using FR it is possible to

recover nodal discontinuous Galerkin (DG) methods of the type described by Hesthaven and Warburton [4], and the spectral difference schemes, original proposed by Kopriva and Kolias in 1996 [5] and later popularised by Sun et al. [6]. Unlike traditional DG methods, which are based on a weak formulation, and hence require integration, the FR approach is based on the differential form of the governing system. As a consequence, implementations of FR are typically *quadrature free*; a property which—if correctly exploited—can lead to reduced computational cost.

One problem with the FR approach is that for a non-linear flux function aliasing driven instabilities may develop. This arises as a consequence of the fact that a collocation projection is used to project the flux into the polynomial space of the solution. The severity of these aliasing driven instabilities depends upon the degree to which the flux is under resolved within each element. Aliasing is particularly problematic within the context of the compressible Euler and Navier–Stokes equations. Here the flux is not only non-linear with respect to the solution but also non-polynomial. Typically, these instabilities are mitigated by performing a least-squares or Galerkin projection of all relevant fluxes; something which has the potential to completely eliminate aliasing. However, in order to accomplish this it is necessary to accurately approximate integrals inside of each element—necessitating computationally-intensive numerical quadrature.

The high cost of least-squares projection has lead to efforts to reduce the amount of aliasing resulting from collocation projections. It has long been known that the polynomial resulting from a collocation projection depends heavily on the choice of nodal points. Perhaps the most famous example of this is Runge's phenomenon [7] which arises when one attempts to use equispaced points to construct a polynomial approximation of the function $f(x) = (1 + 25x^2)^{-1}$ with $x = [-1, 1]$. This has resulted in a substantial body of work which attempts to identify point sets which minimise the L^∞ error associated with a collocation projection of a function inside of a domain. Seminal papers include Chen and Babuška [8], Hesthaven [9], Caliari et al. [10], and Warburton [11]. More recent contributions include those of Chan and Warburton [12] and Hoang [13].

However, there is a growing body of evidence, both theoretical [14] and empirical [15, 16, 17, 18] which suggests that L^∞ performance of a nodal point set is not necessarily a good indicator of their performance as a point set for FR. The present work combines a comprehensive review of known approximation theory results with new analysis, and

numerical experiments, to demonstrate that, in fact, the L^2 norm provides a more reliable indicator of performance when evaluating nodal point sets for FR.

The remainder of this paper is structured as follows. Section 2 provides a brief overview of key theoretical results from Jameson et al. [14] on the non-linear stability of FR schemes and the role of solution points in aliasing. Section 3 provides background theory on approximating polynomials. Section 4 focuses on a class of approximating polynomials called interpolating polynomials, that have associated point sets. Specifically, Chebyshev and Legendre point sets are derived as those which aim to minimise the L^∞ and L^2 errors, respectively, of an associated interpolating polynomial. For completeness the Lebesgue constant is also introduced. Connections between Legendre point sets and least-squares approximation are then established. Section 5 presents a set of numerical experiments, results of which agree with the theoretical analysis, and demonstrate that overall solution accuracy can be improved by an order-of-magnitude by switching from L^∞ to L^2 optimised point sets. Finally, in section 6 conclusion are drawn.

2 Non-linear stability theory

Consider using FR to solve the following non-linear advection problem inside of a periodic domain

$$\frac{\partial u}{\partial t} + \frac{\partial f}{\partial x} = 0, \tag{1}$$

where $u(x,t)$ is the solution and $f(u)$ is a non-linear flux function. Let the domain be discretised into N elements with the numerical solution inside of each element being represented by a polynomial of degree k. If a linearly stable correction function is employed it has been shown by Jameson et al. [14] that energy stability is achieved if

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}||u^\delta||^2 = \Theta(t) + \sum_{i=0}^{N-1} \epsilon_i(t) < 0, \tag{2}$$

where u^δ denotes the discretised solution, $\Theta(t) \leq 0$ is a constant, and the $\epsilon_i(t)$ are per-element constants referred to as *error terms*. Clearly, if the inequality is to be satisfied then the sum of these terms must be less than or equal to zero.

The nature of the error term for the ith element can be understood by considering the error term for a reference element whose extent is between $r = [-1, 1]$. Inside such an element

Jameson et al. show that the error term is given by

$$\hat{\epsilon} = \sum_{i=0}^{k} e(\hat{u}^{\delta})\hat{\varepsilon}_i, \tag{3}$$

where $e(\hat{u}^{\delta})$ is an analytic function of the numerical solution inside of the reference element and

$$\hat{\varepsilon}_i = \int_{-1}^{1} \hat{f}^{\delta} \Upsilon_i \mathrm{d}r - \int_{-1}^{1} \hat{f} \Upsilon_i \mathrm{d}r, \tag{4}$$

where $\hat{f}$ corresponds to the transformed flux inside of the reference element, $\hat{f}^{\delta}$ corresponds to the projection of this transformed flux into the polynomial space of $\hat{u}^{\delta}$, and Υ_i corresponds to the ith Legendre polynomial. Neither the sign nor magnitude of $e(\hat{u}^{\delta})$ can be guaranteed. Therefore, in order to minimize $\hat{\epsilon}$ and thus ϵ_i, one should ensure the magnitude of all $\hat{\varepsilon}_i$ are as small as possible.

Hence, when employing FR for a non-linear problem our objective—in so far as stability is concerned—is to construct transformed numerical flux polynomials $\hat{f}^{\delta}$ which closely approximate the true transformed flux $\hat{f}$.

3 Approximating polynomials

A polynomial of degree m is said to approximate a function $f(x)$ with $x \in [a, b]$ if $\|f(x) - p(x)\|$ is small where $\|\cdot\|$ is a norm defined over $[a, b]$. In this section we shall review techniques for constructing polynomials which are optimal in the sense of the L^{∞} norm and the L^2 norm.

3.1 L^{∞} optimal

This polynomial is defined as having the 'minimum maximum error' over all such polynomials and hence is oft referred to as the *minimax* polynomial.

Theorem 3.1 (de la Vallée Poussin's theorem)**.** *Let $f(x)$ be a continuous function with $x \in [a, b]$ and take $q(x)$ to be an order m polynomial for which there exist $m + 2$ ordered points*

$$a \leq x'_0 < x'_1 < \cdots < x'_{m+1} \leq b,$$

at which the error $f(x) - q(x)$ oscillates signs. The error associated with the optimal order m polynomial $p^\star(x)$ is bounded from below as

$$\|f(x) - p^\star(x)\|_\infty \geq \min_{x_i'} |f(x_i') - q(x_i')|.$$

This theorem suggests that we may construct a minimax approximating polynomial by choosing $q(x)$ such that it attains its maximum deviation at $n + 2$ points, and does so with alternating signs. The existence and uniqueness of such a polynomial ensured through the following result.

Theorem 3.2 (Chebyshev equioscillation theorem)**.** *Let $f(x)$ be a continuous function with $x \in [a, b]$. An order m polynomial $q(x)$ is a minimax approximation to $f(x)$ if and only if it attains its maximum error at $m + 2$ points and does so with alternating signs. Furthermore, the minimax polynomial is unique.*

Proof. See [19]. □

Corollary 3.3. *The optimal order m approximating polynomial $p^\star(x)$ has the property that*

$$p^\star(x_i) = f(x_i), \quad 0 \leq i \leq m,$$

for some set of nodes $x_0, x_1, \ldots, x_m$.

Proof. Denote the points where $p^\star(x)$ attains its maximum deviation by $x_0' < x_1' < \cdots < x_{m+1}'$. We know that the sign of the deviation must change between points. Hence, between x_i' and x_{i+1}' there must exist a point x_i where the deviation is zero and thus $p^\star(x_i) = f(x_i)$. As there are $m + 1$ such points these serve to uniquely define $p^\star(x)$. □

Although it is not possible to obtain $p^\star(x)$ directly, using these results we are able to devise an iterative procedure. First proposed by Remez [20], the procedure starts with a guess polynomial $p_1(x)$ of order m and proceeds to generate a sequence of interpolating polynomials $p_k(x)$ that, in the limit $k \to \infty$, converge to the minimax polynomial $p^\star(x)$.

Whilst the Remez algorithm converges rapidly, the iterations themselves are relatively expensive. In addition to $m + 2$ evaluations of $f(x)$, we must also locate the $m + 2$ extrema of a degree m polynomial. This necessitates the use of a polynomial root-finding algorithm.

A consequence of this is that the use of minimax polynomials is typically restricted to applications where $f(x)$ is known ahead of time.

3.2 L^2 optimal

Take $p(x)$ to be a degree m polynomial approximation of a function $f(x)$ for $x \in [-1, 1]$. Expanding $p(x)$ we see that

$$p(x) = \sum_{i=0}^{m} \gamma_i \Upsilon_i(x), \tag{5}$$

where γ_i are a set of expansion coefficients and $\Upsilon_i(x)$ a set of polynomial basis functions which span the space. If $p(x)$ is the optimal polynomial in the sense of the Euclidean or L^2 norm then it must be the case that

$$\frac{\partial}{\partial \gamma_i} \|f(x) - p(x)\|_2^2 = \frac{\partial}{\partial \gamma_i} \int_{-1}^{1} \{f(x) - p(x)\}^2 \, \mathrm{d}x = 0. \tag{6}$$

Moving the partial derivative into the integral we see that

$$\begin{aligned} \frac{\partial}{\partial \gamma_i} \|f(x) - p(x)\|_2^2 &= \int_{-1}^{1} \frac{\partial}{\partial \gamma_i} p^2(x) - 2f(x) \frac{\partial}{\partial \gamma_i} p(x) \, \mathrm{d}x \\ &= \int_{-1}^{1} 2p(x)\Upsilon_i(x) - 2\Upsilon_i(x) f(x) \, \mathrm{d}x \\ &= 2 \sum_{j=0}^{m} \gamma_j \int_{-1}^{1} \Upsilon_i(x)\Upsilon_j(x) \, \mathrm{d}x - 2 \int_{-1}^{1} \Upsilon_i(x) f(x) \, \mathrm{d}x \\ &= 0, \quad \text{for } 0 \leq i \leq m, \end{aligned} \tag{7}$$

where in the final step we have substituted for $p(x)$. Dropping the constant factor of two and noting that the integrals all evaluate to constants we see that this is a linear equation with respect to the expansion coefficients. Hence, considering all of the expansion coefficients, we obtain a linear system of the form $G\gamma - b = 0$ where

$$g_{ij} = \int_{-1}^{1} \Upsilon_i(x)\Upsilon_j(x) \, \mathrm{d}x \quad \text{and} \quad b_i = \int_{-1}^{1} \Upsilon_i(x) f(x) \, \mathrm{d}x, \quad \text{for } 0 \leq i, j \leq m.$$

The matrix G is known as the Gram (or Gramian) matrix. So long as the basis functions are linearly independent it follows that no column can be expressed as a linear combination other columns. Thus $\det G \neq 0$ which guarantees the existence and uniqueness of a solution. The vector of expansion coefficients γ may therefore be obtained by solving this system. As the minimum of the L^2 norm is bounded from below by zero but not bounded from above the resulting polynomial must correspond to the global minimum inside of the domain.

Through a judicious choice of basis functions it is possible, however, to eliminate the need to solve a linear system. If the functions are chosen to satisfy the following orthogonality relationship

$$\int_{-1}^{1} \Upsilon_i(x)\Upsilon_j(x) = \delta_{ij},$$

then Equation (7) reduces down to

$$\gamma_i = \int_{-1}^{1} f(x)\Upsilon_i(x)\,\mathrm{d}x, \tag{8}$$

enabling us to directly evaluate the expansion coefficients. The set of basis polynomials satisfying this relationship are the orthonormal Legendre polynomials

$$\hat{P}_i(x) = \sqrt{\frac{2n+1}{2}} P_i(x),$$

where $P_i(x)$ is the ith order Legendre polynomial. For all but the simplest functions it is not usually possible to evaluate the integrals of Equation (8) analytically. Instead a numerical procedure—for example Gauss–Legendre quadrature—must be used. Although the cost of constructing the L^2 optimal approximating polynomial is not insignificant, it is typically an order of magnitude less than the cost of constructing the minimax polynomial.

Within the context of FR the L^2 optimal approximating polynomial is unique in that it results in all $\hat{\varepsilon}_i = 0$ for all i. To see this consider taking $\hat{f} = f$ and $\hat{f}^d = p^\star(r)$. Substituting into Equation (4) we find

$$\hat{\varepsilon}_i = \int_{-1}^{1} \sum_{j=0}^{m} \gamma_j \Upsilon_j \Upsilon_j \,\mathrm{d}r - \int_{-1}^{1} \hat{f}\Upsilon_i \,\mathrm{d}r = \int_{-1}^{1} f\Upsilon_i \,\mathrm{d}x - \int_{-1}^{1} \hat{f}\Upsilon_i \,\mathrm{d}r = 0,$$

where in the second step we have used the orthogonality relationship and in the third step have substituted for γ_i. An immediate consequence of this is that—so long as the integrals are evaluated with sufficient accuracy—all of the error terms will be zero.

4 Interpolating polynomials

A common class of approximating polynomials are *interpolating polynomials*. Rather than requiring the polynomial $p(x)$ to be optimal with respect to a norm we instead require that

$$p(x_i) = f(x_i), \quad \text{for } 0 \leq i \leq m, \tag{9}$$

where $\{x_i\}$ are a set of $m+1$ necessarily distinct points. The unique polynomial satisfying these constraints is known as the *Lagrange interpolating polynomial*

$$p(x) = \sum_{i=0}^{m} \ell_i(x) f(x_i), \tag{10}$$

where $\ell_i(x)$ is a *nodal basis function* with the property that $\ell_i(x_j) = \delta_{ij}$. Through direct substitution it may be readily verified that

$$\ell_i(x) = \prod_{\substack{0 \le j \le m \\ j \neq i}} \frac{x - x_j}{x_i - x_j}. \tag{11}$$

Another means of constructing an interpolating polynomial is by directly determining its coefficients. Consider expanding $p(x)$ in terms of its monomial coefficients as

$$\sum_{j=0}^{m} a_j x_i^j = y_i, \quad \text{for } 0 \le i \le m,$$

which can be seen to be a system of linear equations of dimension $m+1$ where the unknowns are the monomial coefficients a_j. This system can be expressed in matrix form as

$$\mathcal{V}\mathbf{a} = \mathbf{y},$$

with

$$\mathcal{V} = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^m \\ 1 & x_1 & x_1^2 & \dots & x_1^m \\ 1 & x_2 & x_2^2 & \dots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^m \end{bmatrix} \quad \text{and} \quad \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

The matrix $\mathcal{V}$ is known as the Vandermonde matrix. Moreover, it can be shown that

$$\det \mathcal{V} = \prod_{i=0}^{m-1} \prod_{j=i+1}^{m} (x_j - x_i). \tag{12}$$

The advantage of interpolation is that—given a suitable set of interpolation nodes—the corresponding polynomial can be formed with a fixed number of function evaluations. The efficacy of the resulting polynomial is closely connected to the choice of nodes.

Without loss of generality we shall, for the remainder of this section, restrict ourselves to the reference domain $\Omega = [-1, 1]$.

Theorem 4.1. *Let $f(x)$ be a function with an $m+1$ derivative and $p(x)$ an order m Lagrange type interpolating polynomial of $f(x)$. The remainder is*

$$f(x) - p(x) = \frac{f^{(m+1)}(\xi)}{(m+1)!} \prod_{0 \le i \le m} (x - x_i), \tag{13}$$

where ξ is an unknown constant in the interval defined by $x_0, x_1, \ldots, x_m$.

We remark here that $\xi = \xi(x)$. This, combined with the inherent difficulties associated with computing high-order derivatives, means that it is not practical to evaluate or bound the error term directly.

4.1 L^∞ optimisation

Heretofore much of the effort with regards to choosing good nodal points for Lagrange interpolation has been focused on minimising the L^∞ norm. Using the expression for the remainder we have

$$\begin{aligned} \|f(x) - p(x)\|_\infty &= \max_{x \in \Omega} \left| \frac{f^{(m+1)}(\xi)}{(m+1)!} \prod_{0 \le i \le m} (x - x_i) \right| \\ &\le \max_{\xi \in \Omega} \frac{|f^{(m+1)}(\xi)|}{(m+1)!} \cdot \max_{x \in \Omega} \left| \prod_{0 \le i \le m} (x - x_i) \right|. \end{aligned} \tag{14}$$

The product in question can be seen to be a polynomial of degree $m+1$. The polynomial is also monic. This result suggests that we can reduce the L^∞ error associated with our interpolating polynomial by choosing a set of nodes which minimise the maximal value of the product.

Theorem 4.2. *The polynomial $\tilde{T}_m(x) = 2^{1-m} T_m(x)$ where $T_m(x)$ is the mth Chebyshev polynomial of the first kind has the minimum maximum norm of all monic polynomials.*

Proof. By contradiction. Suppose that there exists a polynomial $p_m(x)$ such that $\|p_m(x)\|_\infty < \|\tilde{T}_m(x)\|_\infty$. From the properties of the Chebyshev polynomials we know that $T_m(x)$ has $m+1$ extrema, $-1 = x'_0 < x'_1 < \cdots < x'_m = 1$ and that these extrema alternate in sign. At these extrema it is necessarily the case that $|p_m(x'_i)| < |\tilde{T}_m(x'_i)|$. Without loss of generality we assume that x'_0 corresponds to a maxima and hence

$$p_m(x'_0) < \tilde{T}_m(x'_0), \quad p_m(x'_1) > \tilde{T}_m(x'_1), \quad p_m(x'_2) < \tilde{T}_m(x'_2), \quad \ldots.$$

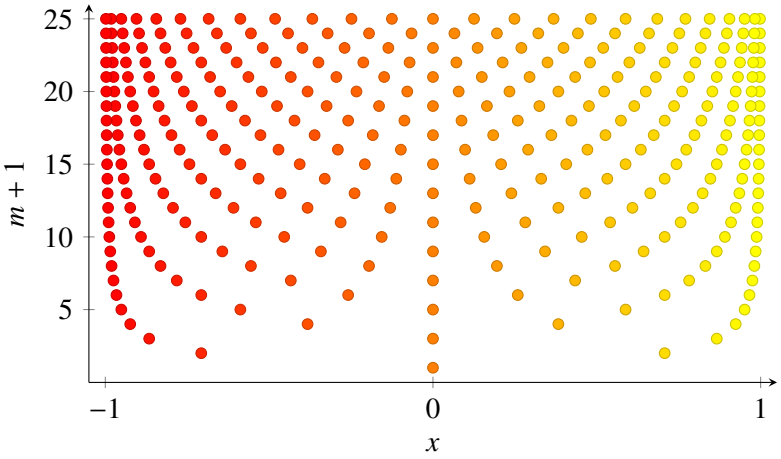


Figure 1. Roots of first 25 Chebyshev polynomials of the first kind $T_{m+1}(x)$.

Therefore, the polynomial $p_m(x) - \tilde{T}_m(x)$ must change sign m times; which implies that it has m roots. However, since $p_m(x)$ and $\tilde{T}_m(x)$ are monic their difference must be of degree $m - 1$ and hence can only have $m - 1$ roots. □

Using this result we see that

$$\min_{x_0,x_1,\ldots,x_m} \max_{x\in\Omega} \left| \prod_{0\le i\le m} (x - x_i) \right| = \tilde{T}_{m+1}(x),$$

with the nodes being given by the roots of $\tilde{T}_{m+1}(x)$ which are equivalent to the roots of $T_{m+1}(x)$. Hence

$$x_i = \cos \frac{2i+1}{2m+2}\pi, \quad \text{for } 0 \le i \le m. \tag{15}$$

Plots of the first 25 sets of nodes can be seen in Figure 1.

Looking at Equation (15) it is clear that the nodes are always located in the interior of the domain. However, in certain application it is desirable for there to be interpolation nodes at the endpoints. One means of satisfying this requirement is to take the nodes as the extrema of $T_m(x)$. Known as the Chebyshev–Lobatto nodes it may be readily verified that

$$x_i = \cos \frac{i}{m}\pi, \quad \text{for } 0 \le i \le m. \tag{16}$$

Another possibility is to simply stretch out the Chebyshev nodes such that they span the entire domain as

$$x_i = \cos\left[\frac{2i+1}{2m+2}\pi\right] \Big/ \cos\left[\frac{1}{2m+2}\pi\right], \quad \text{for } 0 \leq i \leq m. \tag{17}$$

This distribution is known as the *extended* Chebyshev nodes.

4.2 The Lebesgue constant

Another means of optimising the L^∞ performance of an interpolating polynomial is through the *Lebesgue constant*.

Theorem 4.3. *Denote the L^∞ optimal order m approximating polynomial of a function $f(x)$ inside of a domain Ω by $p^\star(x)$. The maximum error of another order m interpolating polynomial specified by $x_0, x_1, \ldots, x_m$ is bounded according to*

$$\|f(x) - p(x)\|_\infty \leq (1 + \Lambda_m)\|f(x) - p^\star(x)\|_\infty, \tag{18}$$

where

$$\Lambda_m = \max_{x\in\Omega} \Lambda_m(x) \quad \textit{and} \quad \Lambda_m(x) = \sum_{i=0}^{m} |\ell_i(x)|, \tag{19}$$

are the Lebesgue constant and the Lebesgue function, respectively.

Proof. Employing the triangle inequality we note that

$$\begin{aligned}\|f(x) - p(x)\|_\infty &= \|f(x) - p^\star(x) + p^\star(x) - p(x)\|_\infty \\ &\leq \|f(x) - p^\star(x)\|_\infty + \|p^\star(x) - p(x)\|_\infty.\end{aligned}$$

However, as $p^\star(x) - p(x)$ is a polynomial of maximal degree m we can rewrite it as a Lagrange interpolating polynomial specified by the same set of nodes as $p(x)$. Hence

$$\begin{aligned}\|f(x) - p(x)\|_\infty &\leq \|f(x) - p^\star(x)\|_\infty + \left\|\sum_{i=0}^{m} [p^\star(x_i) - f(x_i)]\ell_i(x)\right\|_\infty \\ &\leq \|f(x) - p^\star(x)\|_\infty + \|p^\star(x) - f(x)\|_\infty \cdot \left\|\sum_{i=0}^{m} \ell_i(x)\right\|_\infty \\ &\leq \|f(x) - p^\star(x)\|_\infty + \|p^\star(x) - f(x)\|_\infty \cdot \max_{x\in\Omega} \sum_{i=0}^{m} |\ell_i(x)| \\ &= (1 + \Lambda_m)\|f(x) - p^\star(x)\|_\infty,\end{aligned}$$

Table 1. Values of the Lebesgue constant for interpolating polynomials built using various node sets.

	Λ_m				
Nodes	2	5	10	20	100
Extended Chebyshev	1.25	1.69	2.07	2.48	3.47
Chebyshev–Lobatto	1.25	1.99	2.42	2.87	3.89
Chebyshev	1.67	2.10	2.49	2.90	3.90
Legendre–Lobatto	1.25	1.78	2.18	2.61	3.62
Legendre	1.73	3.32	5.19	7.89	19.39

as required. □

The utility of the Lebesgue constant comes from the fact that it enables us to place a bound on the performance of an order m interpolating polynomial $p(x)$ as a multiple of the performance of the L^∞ optimal polynomial $p^\star(x)$. This bound is independent of the function being approximated.

Although somewhat unwieldy the Lebesgue constant is easy to evaluate numerically. Values across a range of polynomial orders are tabulated in Table 1. Further, in the limit of $m \to \infty$ it is known that [21, 22]

$$\Lambda_m^{\text{Chebyshev}} \sim \log(m+1) \quad \text{and} \quad \Lambda_m^{\text{Legendre}} \sim \sqrt{m+1}$$

For both the Chebyshev and Legendre nodes the Lebesgue function obtains its maximum value at the endpoints of the domain. Indeed, away from the boundaries the Lebesgue function for the Legendre nodes behaves as $\log(m+1)$. Moreover, it can also be shown that the Lebesgue constant for the extended Chebyshev nodes is always smaller than that for the Chebyshev nodes [23]. Although the extended Chebyshev nodes result in a near-optimal Lebesgue constant small improvements are possible.

Variations on the Lebesgue coefficient have also been proposed. In [24] Erdős considers

$$\tilde{\Lambda}_m^2 = \int_{-1}^{1} \sum_{i=0}^{m} \ell_i^2(x)\,\mathrm{d}x, \tag{20}$$

and conjectures that the minimal coefficient is that obtained by the Legendre–Lobatto nodes. Subsequent numerical studies by Chen and Babuška [8] have found this not to be true.

Another variation considered by Chen and Babuška is

$$\hat{\Lambda}_m^2 = \max_{x\in\Omega} \hat{\Lambda}_m^2(x) \quad \text{and} \quad \hat{\Lambda}_m^2(x) = \sum_{i=0}^{m} \ell_i^2(x). \tag{21}$$

Both of these variations place an increased emphasis on the *mean* performance of the nodes and thus are potentially better suited for certain finite element applications.

4.3 L^2 optimisation

In Section 3.2 we established the importance of the L^2 norm for the non-linear stability of FR. Despite this importance, the impact of interpolation nodes on the L^2 norm of the residual is seldom considered. It is in fact possible to minimise the L^2 norm associated with a set of interpolation nodes. Starting from the definition of the L^2 norm of the residual we have

$$\begin{aligned}
\|f(x) - p(x)\|_2^2 &= \int_{-1}^{1} \{f(x) - p(x)\}^2 \,\mathrm{d}x \\
&= \int_{-1}^{1} \left[\frac{f^{(m+1)}(\xi)}{(m+1)!}\right]^2 \prod_{0\le i\le m} (x - x_i)^2 \,\mathrm{d}x \\
&= \left[\frac{f^{(m+1)}(\zeta)}{(m+1)!}\right]^2 \int_{-1}^{1} \prod_{0\le i\le m} (x - x_i)^2 \,\mathrm{d}x \\
&\le \max_{\zeta} \left[\frac{f^{(m+1)}(\zeta)}{(m+1)!}\right]^2 \int_{-1}^{1} \prod_{0\le i\le m} (x - x_i)^2 \,\mathrm{d}x \\
&= \max_{\zeta} \left[\frac{f^{(m+1)}(\zeta)}{(m+1)!}\right]^2 \left\|\Upsilon(x)\right\|_2^2,
\end{aligned}$$

where on the third line we have applied the weighted mean value theorem for integrals to move the first term outside of the integral. Here $-1 \le \zeta \le 1$ and $\Upsilon(x)$ is a monic polynomial of degree $m+1$. From this it is evident that we may bound the L^2 error by taking $\Upsilon(x)$ to be the monic polynomial with the minimal L^2 norm over the domain.

Theorem 4.4. *The polynomial $\tilde{P}_m(x) = P_m(x)/k_m$, where $P_m(x)$ is the order m Legendre polynomial and $k_m = (2m)!/2^m(m!)^2$, has the minimum L^2 norm of all monic polynomials defined in $[-1, 1]$.*

Proof. The minimal degree m polynomial

$$\Upsilon(x) = \prod_{i=0}^{m-1}(x - x_i),$$

has the property that

$$\frac{\partial}{\partial x_k}\|\Upsilon(x)\|_2^2 = 2\int_{-1}^{1}\Upsilon(x)\frac{\partial}{\partial x_k}\Upsilon(x)\,\mathrm{d}x = 0, \quad \text{for all } 0 \leq k \leq m-1.$$

Looking inside of the integral we see $\Upsilon(x)$ is of degree m whereas its derivative with respect to x_k is of degree $m-1$. We therefore require $\Upsilon(x)$ to be orthogonal to m distinct polynomials of degree m. These polynomials, however, form a basis over P_{m-1} and hence the condition is equivalent to requiring $\Upsilon(x)$ be orthogonal to *all* polynomials of degree $m-1$ or less. Such a property is the very definition of the Legendre polynomials. Therefore we conclude $\Upsilon(x) = \tilde{P}_m(x)$. □

Thus, to minimise the remainder of a Lagrange interpolating polynomial with respect to the L^2 norm it is suggested to choose the interpolation nodes to be the roots of $\tilde{P}_{m+1}(x)$. These nodes are equivalent to the roots of $P_{m+1}(x)$. They are commonly referred to as the Legendre or Gauss–Legendre nodes. Closed form expressions for the roots of $P_{m+1}(x)$ in terms of radicals are available up to $m = 8$. Beyond this the roots must be evaluated numerically or looked up in a table.

Until recently, the reference algorithm for calculating the nodes has been that of Golub and Welsch [25]. This reduces the task of finding the $m+1$ roots of $P_{m+1}(x)$ to that of finding the $m+1$ eigenvalues of a symmetric tridiagonal matrix. Hence, the computational cost of the Golub–Welsch algorithm is $O(m^2)$ flops. Whilst this is not an issue for moderate values of m it can pose a problem when m is large. In 2012 Bogaert [26, 27] proposed an iteration-free algorithm which enables all $m+1$ roots to be obtained at a cost of just $O(m)$ flops.

Plots of the first 25 sets of nodes can be seen in Figure 2. As with the Chebyshev nodes we observe a clustering of the nodes towards the endpoints but are never incident.

A popular variation on the Legendre nodes which include the endpoints are the Legendre–Lobatto nodes. Also known as the Gauss–Legendre–Lobatto nodes, they are given as the extrema of $P_m(x)$ inside the domain $[-1, 1]$. Whilst they do not minimise the L^2 error norm they do have the remarkable property of maximising the determinant of $\mathcal{V}$ [9].

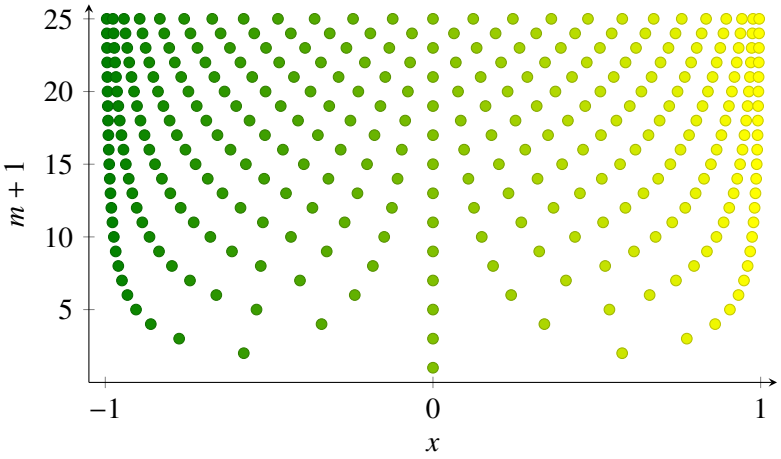


Figure 2. Roots of the first 25 Legendre polynomials $P_{m+1}(x)$.

Moreover, we further note that Lagrange interpolation using $m + 1$ Legendre nodes is equivalent to an L^2 optimal approximation polynomial where the coefficients are obtained using Legendre quadrature with $m + 1$ abscissa. The orthogonality relationship for the Legendre polynomials states that

$$\int_{-1}^{1} \hat{P}_i(x)\hat{P}_j(x)\,\mathrm{d}x = \delta_{ij},$$

with the integrand being a polynomial of degree $i + j$. If we require that $i, j \leq m$ then using a Legendre quadrature rule with $m + 1$ points we can evaluate this integral exactly as

$$\int_{-1}^{1} \hat{P}_i(x)\hat{P}_j(x)\,\mathrm{d}x = \sum_{k=0}^{m} \omega_k \hat{P}_i(x_k)\hat{P}_j(x_k) = \delta_{ij}.$$

where the ω_k are the quadrature weights and x_k the abscissa. This is the discrete form of the orthogonality relationship.

By Equation (8) we know that the expansion coefficients for the least squares optimal approximating polynomial are given by

$$\begin{aligned}\gamma_i &= \int_{-1}^{1} \hat{P}_i(x) f(x)\,\mathrm{d}x \\ &\approx \sum_{k=0}^{m} \omega_k \hat{P}_i(x_k) f(x_k) = \tilde{\gamma}_i,\end{aligned}$$

where in the second step we have approximated the integral using a Gauss–Legendre quadrature rule.

An alternative means of obtaining a degree m polynomial approximating to $f(x)$ is to use a Lagrange interpolating polynomial $p(x)$ where

$$p(x_k) = f(x_k), \quad \text{for } 0 \leq k \leq m.$$

The polynomial $p(x)$ can be expanded as

$$p(x) = \sum_{j=0}^{m} \bar{\gamma}_j \hat{P}_j(x),$$

with the expansion coefficients chosen so as to ensure that

$$\sum_{j=0}^{m} \bar{\gamma}_j \hat{P}_j(x_k) = f(x_k), \quad \text{for } 0 \leq k \leq m.$$

Multiplying both sides of this expression by $\omega_k \hat{P}_i(x_k)$ we obtain

$$\sum_{j=0}^{m} \omega_k \hat{P}_i(x_k) \hat{P}_i(x_k) \bar{\gamma}_j = \omega_k \hat{P}_i(x_k) f(x_k), \quad \text{for } 0 \leq k \leq m.$$

Summing these expressions we find

$$\sum_{k=0}^{m} \sum_{j=0}^{m} \omega_k \hat{P}_i(x_k) \hat{P}_j(x_k) \bar{\gamma}_j = \sum_{k=0}^{m} \omega_k \hat{P}_i(x_k) f(x_k),$$

which after applying the discrete orthogonality relationship gives

$$\bar{\gamma}_i = \sum_{k=0}^{m} \omega_k \hat{P}_i(x_k) f(x_k) = \tilde{\gamma}_i.$$

Thus the two are equivalent in the case where the least-squares expansion coefficients γ_i are evaluated using a Legendre quadrature rule with $m + 1$ abscissa.

5 Numerical experiments

Two dimensional isentropic vortex advection is a commonly used test case for assessing the accuracy of flow solvers for unsteady inviscid flows using the Euler equations [28]. In

this section we will use this problem to benchmark the performance of the nodal point sets outlined in the previous sections. Simulations will be undertaken on a domain $\mathbf{\Omega} = [-20, 20]^2$ using the three structured quadrilateral grids detailed in Table 2.

For reference, the two dimensional Euler equations are given according to

$$\frac{\partial u}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{f} = 0,$$

where $u = u(\mathbf{x}, t) = \left\{\rho, \rho v_x, \rho v_y, E\right\}$ with ρ corresponding to the mass density of the fluid, $\mathbf{v} = (v_x, v_y)^T$ the fluid velocity vector, and E the total energy per unit volume. For a perfect gas the pressure p and total energy can be related by the ideal gas law

$$E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho \|\mathbf{v}\|^2,$$

where γ is the ratio of specific heats. The flux is given by

$$\mathbf{f} = \begin{Bmatrix} \rho v_x & \rho v_y \\ \rho v_x^2 + p & \rho v_y v_x \\ \rho v_x v_y & \rho v_y^2 + p \\ v_x(E + p) & v_y(E + p) \end{Bmatrix},$$

which on account of the product terms and the definition of p is not only non-linear but also non-polynomial.

The initial conditions for isentropic vortex advection are given by

$$\rho(\mathbf{x}, t = 0) = \left\{1 - \frac{S^2 M^2 (\gamma - 1) \exp 2f}{8\pi^2}\right\}^{\frac{1}{\gamma - 1}}, \tag{22}$$

$$\mathbf{v}(\mathbf{x}, t = 0) = \frac{Sy \exp f}{2\pi R}\hat{\mathbf{x}} + \left\{1 - \frac{Sx \exp f}{2\pi R}\right\}\hat{\mathbf{y}}, \tag{23}$$

$$p(\mathbf{x}, t = 0) = \frac{\rho^\gamma}{\gamma M^2}, \tag{24}$$

where $f = (1 - x^2 - y^2)/2R^2$, $S = 13.5$ is the vortex strength, $M = 0.4$ is the free-stream Mach number, and $R = 1.5$ is the radius. These conditions correspond to a vortex advecting with a velocity of one in the y direction. The initial mass density can be seen in Figure 3. All meshes were configured with periodic boundary conditions along boundaries of constant y.

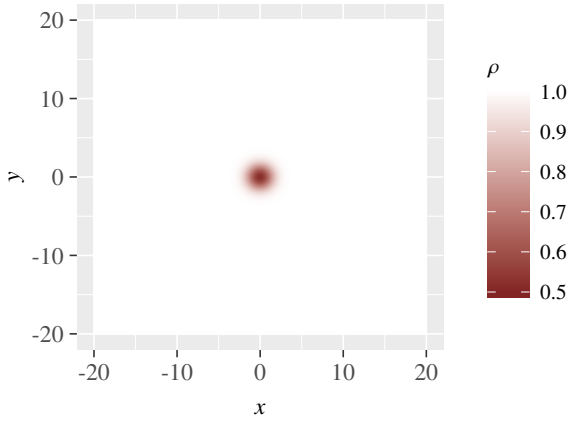


Figure 3. Initial density profile for the isentropic vortex experiment.

Along boundaries of constant x Riemann invariant free-stream conditions were imposed to ensure that

$$\begin{aligned}\rho(\mathbf{x} = y\hat{\mathbf{y}} \pm 20\hat{\mathbf{x}}, t) &\approx 1,\\ \mathbf{v}(\mathbf{x} = y\hat{\mathbf{y}} \pm 20\hat{\mathbf{x}}, t) &\approx \hat{\mathbf{y}},\\ p(\mathbf{x} = y\hat{\mathbf{y}} \pm 20\hat{\mathbf{x}}, t) &\approx \frac{1}{\gamma M^2},\end{aligned}$$

which are the limiting values of the initial conditions. Strictly speaking these conditions, on account of the periodicity, result in the modelling of an infinite array of coupled vortices. The impact of this is mitigated by the observation that the exponentially decaying vortex has a characteristic radius which is far smaller than the extent of the domain. Neglecting these effects the analytic solution of the system is a time t is simply a translation of the initial conditions. A single convective time is defined as the time required for the vortex to make one pass through domain; hence $t_c = 40$.

Using the analytical solution we can define an L^2 error as

$$\sigma(t)^2 = \iint_{\mathbf{\Omega}} \Big[\rho^{\delta}(\mathbf{x} + \Delta_y(t)\hat{\mathbf{y}}, t) - \rho(\mathbf{x}, 0)\Big]^2 \, \mathrm{d}^2\mathbf{x}, \tag{25}$$

where $\rho^{\delta}(\mathbf{x}, t)$ is the numerical mass density, $\rho(\mathbf{x}, 0)$ the analytic mass density at $t = 0$, and $\Delta_y(t)$ is an ordinate which accounts for the fact that the vortex is translating in a free stream velocity of unity in the y direction. Inside of the discretised domain this error can be

Table 2. Solver and mesh configurations for the isentropic vortex experiment. The total number of elements and solution points in the simulations are indicated by N_E and $\sum N_u$, respectively.

Order	Spacing	N_E	$\sum N_u$
$\wp = 2$	$h = 1/4$	80^2	240^2
$\wp = 3$	$h = 1/3$	60^2	240^2
$\wp = 4$	$h = 5/12$	48^2	240^2

computed by integrating difference between the analytical and numerical solutions inside of each element and then summing the residuals. If we restrict ourselves to times t which are integer multiples of the period we obtain

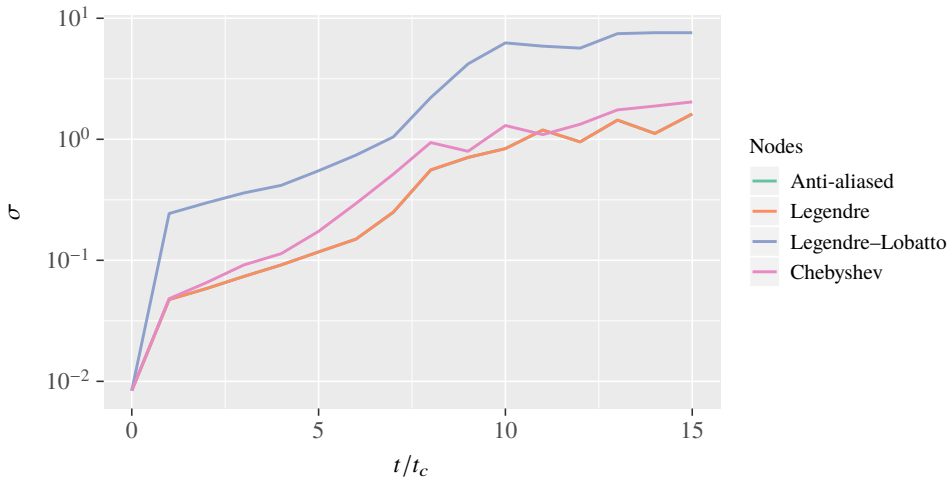
$$\sigma(t)^2 = \sum_i^{N_E} \iint_{\hat{\mathbf{\Omega}}} \left[\rho_i^\delta(\tilde{\mathbf{x}}, t) - \rho(\mathcal{M}_i(\tilde{\mathbf{x}}), 0)\right]^2 J_i(\tilde{\mathbf{x}}) \, \mathrm{d}^2\tilde{\mathbf{x}}, \tag{26}$$

where $\hat{\mathbf{\Omega}}$ is the reference quadrilateral, $\rho_i^\delta(\tilde{\mathbf{x}}, t)$ is the approximate mass density inside of the ith element, $\mathcal{M}_i(\tilde{\mathbf{x}})$ is the mapping from reference to physical space, and $J_i(\tilde{\mathbf{x}})$ the associated Jacobian. We can also define an L^∞ error as

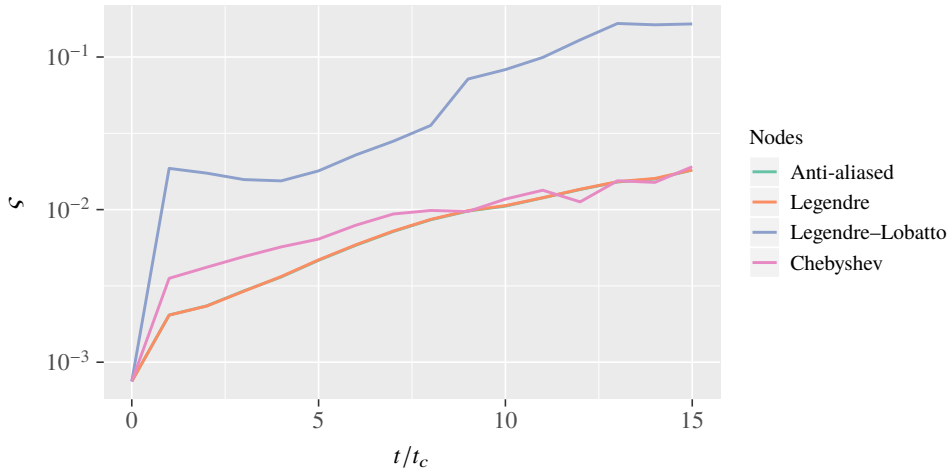
$$\varsigma(t) = \max_{1 \leq i \leq N_E} \max_{\tilde{\mathbf{x}} \in \hat{\mathbf{\Omega}}} \left|\rho_i^\delta(\tilde{\mathbf{x}}, t) - \rho(\mathcal{M}_i(\tilde{\mathbf{x}}), 0)\right|, \tag{27}$$

which can be approximated by computing the absolute error on a suitably fine grid inside of each element and then taking the pointwise maximum.

Simulations were run with five different sets of solution and flux points: Gauss–Legendre, Gauss–Legendre–Lobatto, Chebyshev, Chebyshev–Lobatto, and extended Chebyshev. In addition a set of reference simulations employing full anti-aliasing (via an L^2 projection using a 27th order quadrature rule) of all non-linear terms were also performed. Common interface fluxes were computed using a Rusanov Riemann solver. For all simulations the correction functions were chosen in order to recover a nodal DG scheme. To advance the solutions in time a classical fourth order Runge–Kutta method was used with $\Delta t = 0.005$. Initial conditions were laid down via an L^2 projection using a 27th order quadrature rule. This approach ensures that the initial condition is independent on the choice of solution points. All simulations were run until $t = 15t_c = 600$ with errors being computed every t_c.

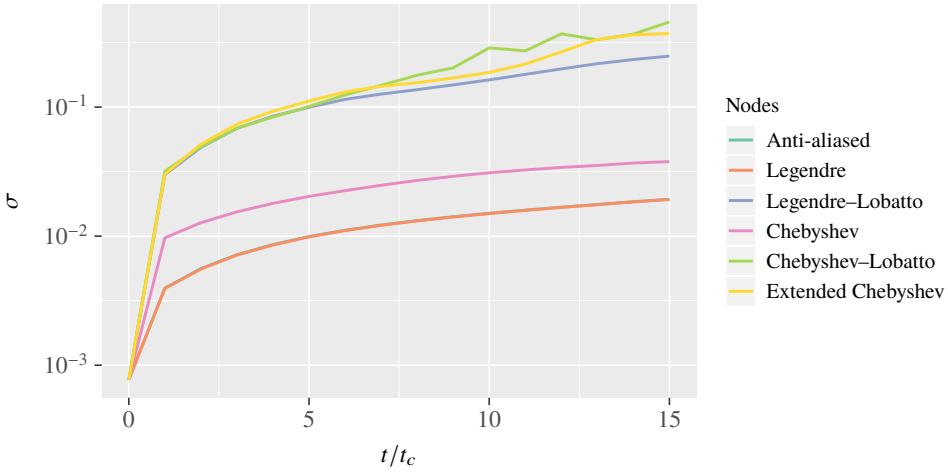


(a) L^2.

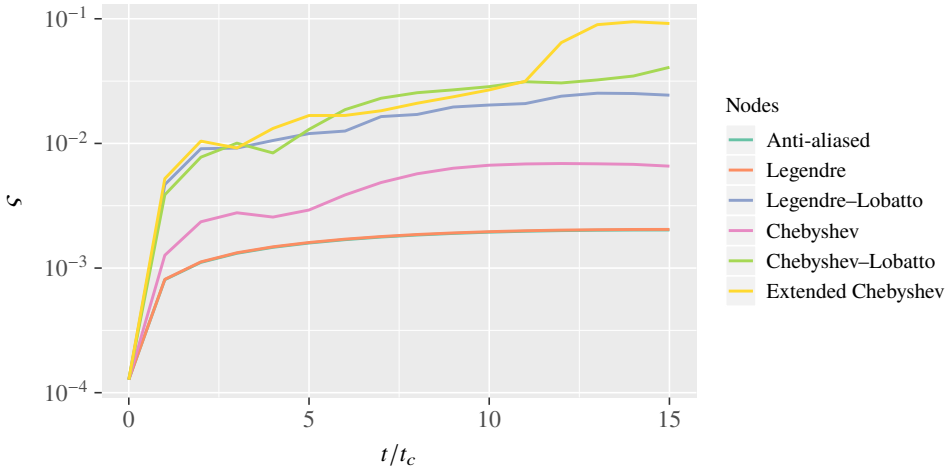


(b) L^∞.

Figure 4. L^2 and L^∞ errors in density for the $\wp = 2$ isentropic Euler vortex simulation. At $\wp = 2$ the Chebyshev–Lobatto and extended Chebyshev nodes are identical to the Gauss–Legendre–Lobatto nodes and are hence omitted. In both plots the Gauss–Legendre results are found to overlay those with anti-aliasing.

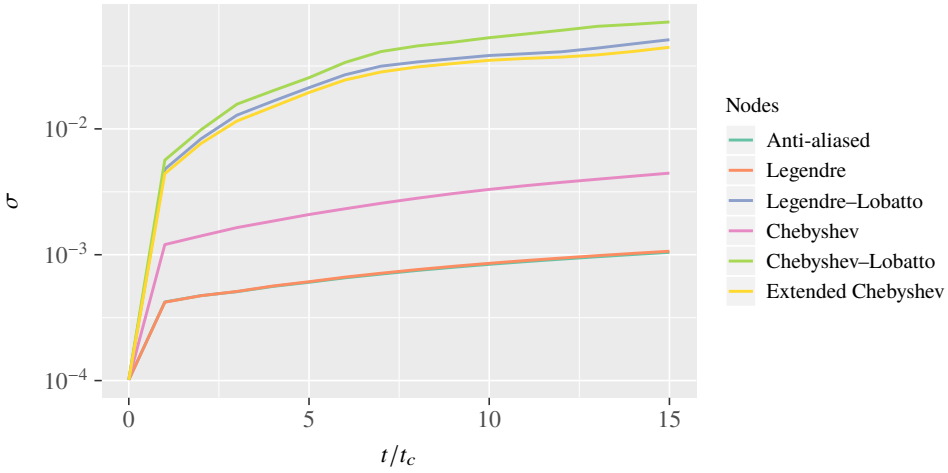


(a) L^2.

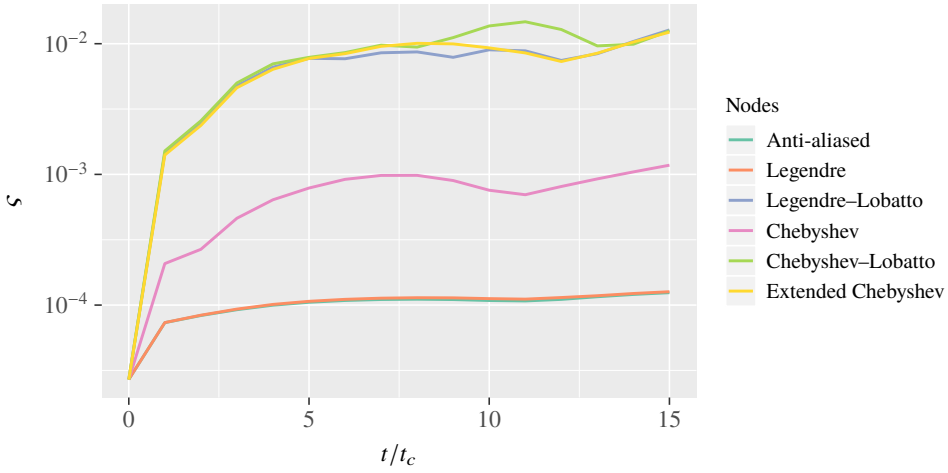


(b) L^∞.

Figure 5. L^2 and L^∞ errors in density for the $\wp = 3$ isentropic Euler vortex simulation. In both plots the Gauss–Legendre results are found to overlay those with anti-aliasing.



(a) L^2.



(b) L^∞.

Figure 6. L^2 and L^∞ errors in density for the $\wp = 4$ isentropic Euler vortex simulation. In both plots the Gauss–Legendre results are found to overlay those with anti-aliasing.

Density errors as a function of time for the three different orders of accuracy can be seen in Figures 4 to 6.

Looking at the figures we observe that for all three polynomial orders that the Legendre points result in errors which are almost identical to those of the fully anti-aliased simulations. This result is also interesting given that our choice of the Legendre nodes was motivated by considerations regarding the non-linear stability of FR as opposed to accuracy. Moreover, as a general trend we observe that point sets with nodes at ± 1 typically under-perform point sets whose nodes are strictly inside the domain. These results are contrary to what one would expect given the Lebesgue constants for these point sets.

6 Conclusion

In this paper we have presented a comprehensive review of known approximation theory results, along with new results, and numerical experiments, to motivate that in fact point sets for FR should aim to minimise the L^2 error of an associated interpolating polynomial. In particular we have shown that the Legendre nodes minimise the L^2 norm of an interpolating polynomial, and proved the equivalence between such an interpolating polynomial and an L^2 approximating polynomial whose coefficients have been obtained using a Gauss–Legendre quadrature rule. Numerical experiments confirm that FR errors can be reduced by an order-of-magnitude by switching from popular point sets such as Chebyshev, Chebyshev–Lobatto and Legendre–Lobatto to Legendre point sets.

Acknowledgements

The authors would like to thank the Engineering and Physical Sciences Research Council for their support via an Early Career Fellowship (EP/K027379/1 and EP/R030340/1).

References

[1] P E Vincent and A Jameson. Facilitating the adoption of unstructured high-order methods amongst a wider community of fluid dynamicists. *Mathematical Modelling of Natural Phenomena*, 6(03):97–140, 2011.

[2] Brian C Vermeire, Freddie D Witherden, and Peter E Vincent. On the utility of gpu accelerated high-order methods for unsteady flow simulations: A comparison with industry-standard tools. *Journal of Computational Physics*, 334:497–521, 2017.

[3] HT Huynh. A flux reconstruction approach to high-order schemes including discontinuous galerkin methods. *AIAA paper*, 4079:2007, 2007.

[4] Jan S Hesthaven and Tim Warburton. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*, volume 54. Springer, 2008.

[5] David A Kopriva and John H Kolias. A conservative staggered-grid chebyshev multidomain method for compressible flows. *Journal of computational physics*, 125(1):244–261, 1996.

[6] Yuzhi Sun, Zhi Jian Wang, and Yen Liu. High-order multidomain spectral difference method for the navier-stokes equations on unstructured hexahedral grids. *Communications in Computational Physics*, 2(2):310–333, 2007.

[7] Carl Runge. Über empirische funktionen und die interpolation zwischen äquidistanten ordinaten. *Zeitschrift für Mathematik und Physik*, 46(224-243):20, 1901.

[8] Qi Chen and Ivo Babuška. Approximate optimal points for polynomial interpolation of real functions in an interval and in a triangle. *Computer Methods in Applied Mechanics and Engineering*, 128(3-4):405–417, 1995.

[9] Jan S Hesthaven. From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex. *SIAM Journal on Numerical Analysis*, 35(2):655–676, 1998.

[10] Marco Caliari, Stefano De Marchi, and Marco Vianello. Bivariate polynomial interpolation on the square at new nodal sets. *Applied Mathematics and Computation*, 165(2):261–274, 2005.

[11] Tim Warburton. An explicit construction of interpolation nodes on the simplex. *Journal of engineering mathematics*, 56(3):247–262, 2006.

[12] Jesse Chan and T Warburton. A comparison of high order interpolation nodes for the pyramid. *SIAM Journal on Scientific Computing*, 37(5):A2151–A2170, 2015.

[13] N Hoang. On node distributions for interpolation and spectral methods. *Mathematics of Computation*, 85(298):667–692, 2016.

[14] Antony Jameson, Peter E Vincent, and Patrice Castonguay. On the non-linear stability of flux reconstruction schemes. *Journal of Scientific Computing*, 50(2):434–445, 2011.

[15] D Williams and A Jameson. Nodal points and the nonlinear stability of high-order methods for unsteady flow problems on tetrahedral meshes. In *43rd AIAA Fluid Dynamics Conference*, 2013.

[16] Freddie D Witherden and Peter E Vincent. An analysis of solution point coordinates for flux reconstruction schemes on triangular elements. *Journal of Scientific Computing*, 61(2):398–423, 2014.

[17] Freddie D Witherden, JS Park, and Peter E Vincent. An analysis of solution point coordinates for flux reconstruction schemes on tetrahedral elements. *Journal of Scientific Computing*, 69(2):905–920, 2016.

[18] Freddie David Witherden. *On the development and implementation of high-order flux reconstruction schemes for computational fluid dynamics*. PhD thesis, PhD thesis, Imperial College London, 2015.

[19] Günter Meinardus. *Approximation of functions: Theory and numerical methods*, volume 13. Springer Science & Business Media, 1967.

[20] Eugene Y Remez. Sur la détermination des polynômes d'approximation de degré donnée. *Comm. Soc. Math. Kharkov*, 10:41–63, 1934.

[21] Theodore Rivlin. *Chebyshev polynomials*. John Wiley & Sons, 1990.

[22] Gabor Szegö. *Orthogonal polynomials*. Amer. Math. Soc, Providence, RI, 1975.

[23] Simon J Smith. Lebesgue constants in polynomial interpolation. In *Annales Mathematicae et Informaticae*, volume 33, pages 1787–5021. Eszterházy Károly College, Institute of Mathematics and Computer Science, 2006.

[24] P Erdős. Problems and results on the theory of interpolation. ii. *Acta Mathematica Hungarica*, 12(1-2):235–244, 1961.

[25] Gene H Golub and John H Welsch. Calculation of gauss quadrature rules. *Mathematics of computation*, 23(106):221–230, 1969.

[26] Ignace Bogaert, Bart Michiels, and Jan Fostier. O(1) computation of legendre polynomials and gauss–legendre nodes and weights for parallel computing. *SIAM Journal on Scientific Computing*, 34(3):C83–C101, 2012.

[27] Ignace Bogaert. Iteration-free computation of gauss–legendre quadrature nodes and weights. *SIAM Journal on Scientific Computing*, 36(3):A1008–A1026, 2014.

[28] Z.J. Wang, Krzysztof Fidkowski, Rémi Abgrall, Francesco Bassi, Doru Caraeni, Andrew Cary, Herman Deconinck, Ralf Hartmann, Koen Hillewaert, H.T. Huynh, Norbert Kroll, Georg May, Per-Olof Persson, Bram van Leer, and Miguel Visbal. High-order CFD methods: current status and perspective. *International Journal for Numerical Methods in Fluids*, 72(8):811–845, 2013.