Numerical Algorithms (2023) 93:157–177
https://doi.org/10.1007/s11075-022-01410-y

ORIGINAL PAPER

Error profile for discontinuous Galerkin time stepping of parabolic PDEs

William McLean[1] · **Kassem Mustapha**[1]

Received: 7 August 2022 / Accepted: 2 September 2022 / Published online: 7 October 2022

Abstract
We consider the time discretization of a linear parabolic problem by the discontinuous Galerkin (DG) method using piecewise polynomials of degree at most $r-1$ in t, for $r \geq 1$ and with maximum step size k. It is well known that the spatial L_2-norm of the DG error is of optimal order k^r globally in time, and is, for $r \geq 2$, superconvergent of order k^{2r-1} at the nodes. We show that on the nth subinterval (t_{n-1},t_n), the dominant term in the DG error is proportional to the local right Radau polynomial of degree r. This error profile implies that the DG error is of order k^{r+1} at the right-hand Gauss–Radau quadrature points in each interval. We show that the norm of the jump in the DG solution at the left end point t_{n-1} provides an accurate a posteriori estimate for the maximum error over the subinterval (t_{n-1},t_n). Furthermore, a simple post-processing step yields a *continuous* piecewise polynomial of degree r with the optimal global convergence rate of order k^{r+1}. We illustrate these results with some numerical experiments.

Keywords Superconvergence · Post-processing · Gauss–Radau quadrature · Legendre polynomials

Mathematics Subject Classification (2010) 65J08 · 65M15

1 Introduction

Consider an abstract, linear initial-value problem

$$u'(t) + Au(t) = f(t) \quad \text{for } 0 < t \leq T, \quad \text{with } u(0) = u_0. \tag{1}$$

✉ William McLean
w.mclean@unsw.edu.au

Kassem Mustapha
kassem.ahmad.mustapha@gmail.com

[1] School of Mathematics and Statistics, University of New South Wales, Kensington 2052, NSW, Australia

We assume a continuous solution $u : [0,T] \to \mathbb{L}$, with $u(t) \in \mathbb{H}$ if $t > 0$, for two Hilbert spaces $\mathbb{L}$ and $\mathbb{H}$ with a compact and dense imbedding $\mathbb{H} \subseteq \mathbb{L}$. By using the inner product $\langle \cdot, \cdot \rangle$ in $\mathbb{L}$ to identify this space with its dual $\mathbb{L}^*$, we obtain in the usual way an imbedding $\mathbb{L} \subseteq \mathbb{H}^*$. The linear operator $A : \mathbb{H} \to \mathbb{H}^*$ is assumed to be bounded and self-adjoint, as well as strictly positive-definite. For instance, if $A = -\nabla^2$ so that (1) is the classical heat equation on a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ where $d \geq 1$, and if we impose homogeneous Dirichlet boundary conditions, then in the usual way we can choose $\mathbb{L} = L_2(\Omega)$ and $\mathbb{H} = H_0^1(\Omega)$, in which case $\mathbb{H}^* = H^{-1}(\Omega)$.

For an integer $r \geq 1$, let U denote the discontinuous Galerkin (DG) time stepping solution to (1) using piecewise polynomials of degree at most $r-1$ with coefficients in $\mathbb{H}$. Thus, we consider only the time discretization with no additional error arising from a spatial discretization. Section 2 summarizes known results on the convergence properties of the DG solution U, and Section 3 introduces a local Legendre polynomial basis that is convenient for the practical implementation of DG time stepping as well as for our theoretical study. These sections serve as preparation for Section 4 where we show that

$$U(t) - u(t) = -a_{nr}(u)\big[p_{nr}(t) - p_{n,r-1}(t)\big] + O\big(k_n^{r+1}\big) \quad \text{for } t \in I_n. \tag{2}$$

Here, k_n denotes the length of the nth time interval $I_n = (t_{n-1}, t_n)$, the function p_{nr} denotes the Legendre polynomial of degree r, shifted to I_n, and $a_{nr}(u)$ denotes the coefficient of p_{nr} in the local Legendre expansion of u on I_n. Since $a_{nr}(u) = O(k_n^r)$, the result (2) shows that the dominant term in the DG error is proportional to the Gauss–Radau polynomial $p_{nr}(t) - p_{n,r-1}(t)$ for $t \in I_n$. However, the coefficient $a_{nr}(u)$ and the $O\big(k_n^{r+1}\big)$ term in (2) typically grow as $t \to 0$ at rates depending on the regularity of the solution u, which in turn depends on the regularity and compatibility of the data. A possible extension permitting a time-dependent operator $A(t)$ is discussed briefly in Remark 4.8.

In 1985, Eriksson, Johnson and Thomée [1] presented an error analysis for DG time stepping of (1), showing optimal $O(k^{r+1})$ convergence in $L_\infty\big((0,T);L_2(\Omega)\big)$ and $O(k^{2r-1})$ superconvergence for the nodal values $\lim_{t \to t_n^-} U(t)$, where $k = \max_{1 \leq n \leq N} k_n$. Subsequently, numerous authors [2–8] have refined these results, including a recent L_∞ stability result of Schmutz and Wihler [9] that we use in the proof of Theorem 4.4. Shortly before completing the present work we learned that the expansion (2) was proved by Adjerid et al. [10, 11] for a linear, scalar hyperbolic problem, and also for nonlinear systems of ODEs [12]; see Remark 4.7 for more details.

Section 5 discusses some practical consequences of (2), in particular the superconvergence of the DG solution at the right Radau points in each interval. This phenomenon was exploited by Springer and Vexler [13] in the piecewise-linear ($r = 2$) case to achieve higher-order accuracy for a parabolic optimal control problem. We will see in Lemma 5.1 how the norm of the jump in U at the break point t_{n-1} provides an accurate estimate of the maximum DG error over the interval I_n. Moreover, a simple, low-cost post-processing step yields a *continuous* piecewise polynomial U_* of degree at most r, called the *reconstruction* of U, that satisfies $U_*(t) - u(t) = O\big(k_n^{r+1}\big)$ for $t \in I_n$; see Corollary 5.3. Finally, Section 6 reports the results of some numerical

experiments for a scalar ODE and for heat equations in one and two spatial dimensions, confirming the convergence behavior from the theory based on (2).

Our motivation for the present study originated in a previous work [14] dealing with the implementation of DG time stepping for a subdiffusion equation $u'(t) + \partial_t^{1-\nu} Au(t) = f(t)$ with $0 < \nu < 1$, where $\partial_t^{1-\nu}$ denotes the Riemann–Liouville fractional time derivative of order $1 - \nu$. We observed in numerical experiments that (2) holds except with $O\big(k_n^{r+\nu}\big)$ in place of $O\big(k_n^{r+1}\big)$.

Treatment of the spatial discretization of (1) is beyond the scope of this paper, apart from its use in our numerical experiments. To make practical use of our result (2) it is necessary to ensure that the spatial error is dominated by the $O\big(k_n^{r+1}\big)$ term. Also, although we allow nonuniform time steps in our analysis, we will not consider questions such as local mesh refinement or adaptive step size control, which are generally required to resolve the solution accurately for t near 0.

2 Discontinuous Galerkin time stepping

As background and preparation for our results, we formulate in this section the DG time stepping procedure and summarize key convergence results from the literature. Our standard reference is the monograph of Thomée [15, Chapter 12].

Choosing time levels $0 = t_0 < t_1 < t_2 < \cdots < t_N = T$, we put

$$k = \max_{1 \le n \le N} k_n \quad \text{where} \quad k_n = t_n - t_{n-1}.$$

Let $\mathbb{P}_j(\mathbb{V})$ denote the space of polynomials of degree at most j with coefficients from a vector space $\mathbb{V}$. We fix an integer $r \ge 1$, put $\boldsymbol{t} = (t_n)_{n=0}^N$ and form the piecewise-polynomial space $\mathcal{X}_r = \mathcal{X}_r(\boldsymbol{t}, \mathbb{H})$ defined by

$$X \in \mathcal{X}_r \qquad \text{iff} \qquad X|_{I_n} \in \mathbb{P}_{r-1}(\mathbb{H}) \text{ for } 1 \le n \le N.$$

Denoting the one-sided limits of X at t_n by

$$X_+^n = \lim_{t \to t_n^+} X(t) \quad \text{and} \quad X_-^n = \lim_{t \to t_n^-} X(t),$$

we discretize (1) in time by seeking $U \in \mathcal{X}_r$ satisfying [15, p. 204]

$$\langle U_+^{n-1}, X_+^{n-1} \rangle + \int_{I_n} \langle U' + AU, X \rangle \, \mathrm{d}t = \langle U_-^{n-1}, X_+^{n-1} \rangle + \int_{I_n} \langle f, X \rangle \, \mathrm{d}t \tag{3}$$

for $X \in \mathcal{X}_r$ and $1 \le n \le N$, with $U_-^0 = u_0$. Section 3 describes how, given U_-^{n-1} and f, we can solve a linear system to obtain $U|_{I_n}$ and so advance the solution by one time step.

Remark 2.1 If the integral on the right-hand side of (3) is evaluated using the right-hand, r-point, Gauss–Radau quadrature rule on I_n, then the sequence of nodal values U_-^n coincides with the finite difference solution produced by the r-stage Radau IIA (fully) implicit Runge–Kutta method; see Vlasák and Roskovec [16, Section 3].

Let $\|\cdot\|$ denote the norm in $\mathbb{L}$ and let $u^{(\ell)}$ denote the ℓ th derivative of u with respect to t. It will be convenient to write

$$\|v\|_{I_n} = \sup_{t\in I_n} \|v(t)\|,$$

and to define the fractional powers of A in the usual way via its spectral decomposition [15, Chapter 3]. The DG time stepping scheme has the nodal error bound [15, Theorem 12.1]

$$\|U_-^n - u(t_n)\|^2 \le C \sum_{j=1}^{n} k_j^{2\ell} \int_{I_j} \|A^{1/2} u^{(\ell)}(t)\|^2 \, \mathrm{d}t \quad \text{for } 1 \le \ell \le r, \tag{4}$$

and the uniform bound [15, Theorem 12.2]

$$\|U - u\|_{I_n} \le \|U_-^n - u(t_n)\| + C\|U_-^{n-1} - u(t_{n-1})\| + Ck_n^{\ell} \|u^{(\ell)}\|_{I_n} \quad \text{for } 1 \le \ell \le r,$$

where in both cases $1 \le n \le N$. We therefore have optimal convergence

$$\|U(t) - u(t)\| = O(k^r) \quad \text{for } 0 \le t \le T, \tag{5}$$

provided $u^{(r)} \in L_\infty((0,T);\mathbb{L})$ and $A^{1/2}u^{(r)} \in L_2((0,T);\mathbb{L})$. In fact, U is superconvergent at the nodes [15, Theorem 12.3] when $r \ge 2$, with

$$\|U_-^n - u(t_n)\|^2 \le Ck^{2(\ell-1)} \sum_{j=1}^{n} k_j^{2\ell} \int_{I_j} \|A^{\ell-1/2} u^{(\ell)}(t)\|^2 \mathrm{d}t \quad \text{for } 1 \le \ell \le r.$$

Thus,

$$\|U_-^n - u(t_n)\| = O(k^{2r-1}), \tag{6}$$

provided $A^{r-1/2}u^{(r)} \in L_2((0,T);\mathbb{L})$.

Suppose for the remainder of this section that $f \equiv 0$, and consider error bounds involving the (known) initial data u_0 instead of the (unknown) solution u. By separating variables, one finds that [15, Lemma 3.2]

$$\|A^q u^{(\ell)}(t)\| \le Ct^{s-(q+\ell)} \|A^s u_0\| \quad \text{for } 0 \le s \le q + \ell \text{ and } 0 < t \le T, \tag{7}$$

assuming that u_0 belongs to the domain of A^s. It follows that, for sufficiently regular initial data, we have the basic error bound [1, Theorem 1],

$$\|U(t) - u(t)\| \le Ck^{\ell} \|A^{\ell} u_0\| \quad \text{for } 0 \le t \le T \text{ and } 0 \le \ell \le r. \tag{8}$$

For non-smooth initial data $u_0 \in L_2(\Omega)$, the full rate of convergence still holds but with a constant that blows up as t tends to zero [1, Theorem 3]: provided $k_n \le Ck_{n-1}$ for all $n \ge 2$,

$$\|U(t) - u(t)\| \le Ct^{-r}k^r \|u_0\| \quad \text{for } 0 < t \le T,$$

and hence, by interpolation,

$$\|U(t) - u(t)\| \le Ct^{s-r}k^r\|A^s u_0\| \quad \text{for } 0 < t \le T \text{ and } 0 \le s \le r. \tag{9}$$

At the nodes [1, Theorem 2],

$$\|U_-^n - u(t_n)\| \le Ck^s\|A^s u_0\| \quad \text{for } 1 \le n \le N \text{ and } 1 \le s \le 2r-1, \tag{10}$$

and [1, Theorem 3], provided $k_n \le Ck_{n-1}$ for all $n \ge 2$,

$$\|U_-^n - u(t_n)\| \le Ct_n^{-s}k^s\|u_0\| \quad \text{for } 1 \le n \le N \text{ and } 0 \le s \le 2r-1. \tag{11}$$

Taking $s = q$ in (10) and (11), we see by interpolation that

$$\|U_-^n - u(t_n)\| \le Ct_n^{s-q}k^q\|A^s u_0\| \quad \text{for } 1 \le n \le N \text{ and } 0 \le s \le q \le 2r-1. \tag{12}$$

3 Local Legendre polynomial basis

We now return to considering the general inhomogeneous problem and describe a practical formulation of the DG scheme using local Legendre polynomial expansions that will also play an essential role in our subsequent analysis.

Let P_j denote the Legendre polynomial of degree j with the usual normalization $P_j(1) = 1$, and recall that

$$\int_{-1}^{1} P_i(\tau)P_j(\tau)\,\mathrm{d}\tau = \frac{2\delta_{ij}}{2j+1}.$$

Using the affine map $\beta_n : [-1,1] \to [t_{n-1},t_n]$ given by

$$\beta_n(\tau) = \frac{1}{2}\big[(1-\tau)t_{n-1} + (1+\tau)t_n\big] \quad \text{for } -1 \le \tau \le 1, \tag{13}$$

we define local Legendre polynomials on the n th subinterval,

$$p_{nj}(t) = P_j(\tau) \quad \text{for } t = \beta_n(\tau) \text{ and } -1 \le \tau \le 1,$$

and note that

$$p_{nj}(t_n) = 1 \quad \text{and} \quad \int_{I_n} p_{ni}(t)\,p_{nj}(t)\,\mathrm{d}t = \frac{k_n\delta_{ij}}{2j+1}. \tag{14}$$

The local Fourier–Legendre expansion of a function v is then, for $t \in I_n$,

$$v(t) = \sum_{j=0}^{\infty} a_{nj}(v)p_{nj}(t) \quad \text{where} \quad a_{nj}(v) = \frac{2j+1}{k_n}\int_{I_n} v(t)p_{nj}(t)\,\mathrm{d}t.$$

In particular, for the DG solution U we put $U^{nj} = a_{nj}(U) \in \mathbb{H}$ so that

$$U(t) = \sum_{j=0}^{r-1} U^{nj} p_{nj}(t) \quad \text{for } t \in I_n.$$

Define [14, Lemma 5.1]

$$G_{ij} = P_j(-1)P_i(-1) + \int_{-1}^{1} P_j'(\tau)P_i(\tau)\,\mathrm{d}\tau = \begin{cases} (-1)^{i+j}, & \text{if } i \geq j, \\ 1, & \text{if } i < j, \end{cases}$$

and $H_{ij} = \int_{-1}^{1} P_j(\tau)P_i(\tau)d\tau = \delta_{ij}/(2j+1)$; e.g., if $r = 4$ then

$$\boldsymbol{G} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{H} = \begin{bmatrix} 1 & & & \\ & \frac{1}{3} & & \\ & & \frac{1}{5} & \\ & & & \frac{1}{7} \end{bmatrix}.$$

By choosing a test function of the form $X(t) = p_{ni}(t)\chi$, for $t \in I_n$ and $\chi \in \mathbb{H}$, we find that the DG equation (3) implies

$$\sum_{j=0}^{r-1} \left(G_{ij} + k_n H_{ij} A\right) U^{nj} = \check{U}^{n-1,i} + \int_{I_n} f(t) p_{ni}(t)\,\mathrm{d}t \tag{15}$$

for $0 \leq i \leq r-1$ and $1 \leq n \leq N$, where

$$\check{U}^{0i} = (-1)^i u_0 \quad \text{and} \quad \check{U}^{ni} = (-1)^i \sum_{j=0}^{r-1} U^{nj} \quad \text{for } n \geq 1.$$

Thus, given $U^{n-1,j}$ for $0 \leq j \leq r-1$, by solving the (block) $r \times r$ system (15) we obtain U^{nj} for $0 \leq j \leq r-1$, and hence $U(t)$ for $t \in I_n$. The existence and uniqueness of this solution follows from the stability of the scheme [15, p. 205]. Notice that

$$U_+^{n-1} = \sum_{j=0}^{r-1} (-1)^j U^{nj} \quad \text{and} \quad U_-^{n-1} = \begin{cases} u_0 & \text{if } n = 1, \\ \sum_{j=0}^{r-1} U^{n-1,j} & \text{if } 2 \leq n \leq N. \end{cases}$$

4 Behavior of the DG error

To prove our main results, we will make use of two projection operators. The first is just the orthogonal projector $\Pi_r : L^2((0,T);\mathbb{L}) \to \mathcal{X}_r$ defined by

$$\int_0^T \langle \Pi_r v, X \rangle\,\mathrm{d}t = \int_0^T \langle v, X \rangle\,\mathrm{d}t \quad \text{for all } X \in \mathcal{X}_r,$$

which has the explicit representation

$$(\Pi_r v)(t) = \sum_{j=0}^{r-1} a_{nj}(v) p_{nj}(t) \quad \text{for } t \in I_n \text{ and } 1 \le n \le N.$$

The second projector $\widetilde{\Pi}_r : C([0,T];\mathbb{L}) \to \mathcal{X}_r$ is defined by the conditions [15, Equation (12.9)]

$$(\widetilde{\Pi}_r v)_-^n = v(t_n) \qquad \text{and} \qquad \int_{I_n} \langle \widetilde{\Pi}_r v, X' \rangle \, \mathrm{d}t = \int_{I_n} \langle v, X' \rangle \, \mathrm{d}t \tag{16}$$

for all $X \in \mathcal{X}_r$ and for $1 \le n \le N$. The next lemma shows that $\widetilde{\Pi}_r u$ is in fact the DG solution of the trivial equation with $A = 0$; cf. Chrysafinos and Walkington [3, Section 2.2].

Lemma 4.1 If $u' : (0,T] \to \mathbb{L}$ is integrable, then

$$\langle (\widetilde{\Pi}_r u)_+^{n-1}, X_+^{n-1} \rangle + \int_{I_n} \langle (\widetilde{\Pi}_r u)', X \rangle \, \mathrm{d}t = \langle u(t_{n-1}), X_+^{n-1} \rangle + \int_{I_n} \langle u', X \rangle \, \mathrm{d}t.$$

Proof Integrating by parts and using the properties (16) of $\widetilde{\Pi}_r$, we have

$$\begin{aligned} \textstyle\int_{I_n} \langle (\widetilde{\Pi}_r u)', X \rangle \, \mathrm{d}t &= \langle (\widetilde{\Pi}_r u)_-^n, X_-^n \rangle - \langle (\widetilde{\Pi}_r u)_+^{n-1}, X_+^{n-1} \rangle - \textstyle\int_{I_n} \langle \widetilde{\Pi}_r u, X' \rangle \, \mathrm{d}t \\ &= \langle u(t_n), X_-^n \rangle - \langle (\widetilde{\Pi}_r u)_+^{n-1}, X_+^{n-1} \rangle - \textstyle\int_{I_n} \langle u, X' \rangle \, \mathrm{d}t, \end{aligned}$$

and a second integration by parts then yields the desired identity.

The Legendre expansion of $\widetilde{\Pi}_r v$ coincides with that of $\Pi_r v$, except for the coefficient of $p_{n,r-1}$. Below, we denote the closure of the nth time interval by $\bar{I}_n = [t_{n-1}, t_n]$.

Lemma 4.2 If $v : \bar{I}_n \to \mathbb{L}$ is continuous, then

$$(\widetilde{\Pi}_r v)(t) = \sum_{j=0}^{r-2} a_{nj}(v) p_{nj}(t) + \tilde{a}_{n,r-1}(v) p_{n,r-1}(t) \quad \text{for } t \in I_n,$$

where

$$\tilde{a}_{n,r-1}(v) = v(t_n) - (\Pi_{r-1} v)_-^n = v(t_n) - \sum_{j=0}^{r-2} a_{nj}(v).$$

Proof By choosing $X'|_{I_n} = p_{nj}$ in the second property of (16), we see that

$$a_{nj}(\widetilde{\Pi}_r v) = a_{nj}(v) \quad \text{for } 0 \le j \le r-2,$$

implying that $\widetilde{\Pi}_r v = \Pi_{r-1} v + \lambda p_{n,r-1}$ for some $\lambda \in \mathbb{H}$. Since $p_{nj}(t_n) = P_j(1) = 1$, the first property in (16) gives

$$v(t_n) = (\widetilde{\Pi}_r v)^n_- = (\Pi_{r-1} v)^n_- + \lambda \quad \text{with} \quad (\Pi_{r-1} v)^n_- = \sum_{j=0}^{r-2} a_{nj}(v),$$

showing that $\lambda = \tilde{a}_{n,r-1}(v)$.

By mapping to the reference element $(-1,1)$, applying the Peano kernel theorem, and then mapping back to I_n, we find [14, p. 137]

$$\|a_{nj}(v)\| \le C k_n^{j-1} \int_{I_n} \|v^{(j)}(t)\| \, \mathrm{d}t \le C k_n^j \|v^{(j)}\|_{I_n} \quad \text{for } j \ge 0, \tag{17}$$

and

$$\|v - \Pi_r v\|_{I_n} \le C k_n^{\ell-1} \int_{I_n} \|v^{(\ell)}(t)\| \, \mathrm{d}t \le C k_n^{\ell} \|v^{(\ell)}\|_{I_n} \quad \text{for } 1 \le \ell \le r. \tag{18}$$

Theorem 4.3 For $1 \le n \le N$, if $v : \bar{I}_n \to \mathbb{L}$ is C^{r+1} then

$$\left\| \widetilde{\Pi}_r v - v + a_{nr}(v)(p_{nr} - p_{n,r-1}) \right\|_{I_n} \le C k_n^{r+1} \|v^{(r+1)}\|_{I_n}.$$

Proof By Lemma 4.2, if $t \in I_n$ then

$$(\widetilde{\Pi}_r v)(t) = (\Pi_{r-1} v)(t) + \tilde{a}_{n,r-1}(v) p_{n,r-1}(t)$$

and

$$(\Pi_{r+1} v)(t) = (\Pi_{r-1} v)(t) + a_{n,r-1}(v) p_{n,r-1}(t) + a_{nr}(v) p_{nr}(t),$$

so

$$(\widetilde{\Pi}_r v)(t) - (\Pi_{r+1} v)(t) = [\tilde{a}_{n,r-1}(v) - a_{n,r-1}(v)] p_{n,r-1}(t) - a_{nr}(v) p_{nr}(t). \tag{19}$$

Taking the limit as $t \to t_n^-$, and recalling that $p_{nj}(t_n) = 1$, we see that

$$v(t_n) - (\Pi_{r+1} v)^n_- = \tilde{a}_{n,r-1}(v) - a_{n,r-1}(v) - a_{nr}(v). \tag{20}$$

Using (20) to eliminate $\tilde{a}_{n,r-1}(v)$ in (19), we find that

$$\widetilde{\Pi}_r v - v + a_{nr}(v)\big[p_{nr} - p_{n,r-1}\big] = (\Pi_{r+1} v - v) + \big[v(t_n) - (\Pi_{r+1} v)^n_-\big] p_{n,r-1}$$

on I_n, and the desired estimate follows at once from (18).

The following theorem and its corollary, together with the superconvergence result (6), show that

$$\|U - \widetilde{\Pi}_r u\|_{I_n} = O\big(k_n^{r+1}\big) \quad \text{for } r \ge 2, \tag{21}$$

provided u is sufficiently regular.

Theorem 4.4 For $1 \le n \le N$, if $Au : \bar{I}_n \to \mathbb{L}$ is C^r, then

$$\|U - \widetilde{\Pi}_r u\|_{I_n} \le C\|U_-^{n-1} - u(t_{n-1})\| + Ck_n^{r+1}\|Au^{(r)}\|_{I_n}.$$

Proof It follows from Lemma 4.1 that $\widetilde{\Pi}_r u$ satisfies

$$\langle(\widetilde{\Pi}_r u)_+^{n-1}, X_+^{n-1}\rangle + \int_{I_n}\langle(\widetilde{\Pi}_r u)' + A\widetilde{\Pi}_r u, X\rangle \,\mathrm{d}t = \langle u(t_{n-1}), X_+^{n-1}\rangle + \int_{I_n}\langle u' + A\widetilde{\Pi}_r u, X\rangle \,\mathrm{d}t,$$

whereas U satisfies

$$\langle U_+^{n-1}, X_+^{n-1}\rangle + \int_{I_n}\langle U' + AU, X\rangle \,\mathrm{d}t = \langle U_-^{n-1}, X_+^{n-1}\rangle + \int_{I_n}\langle u' + Au, X\rangle \,\mathrm{d}t,$$

for all $X \in \mathcal{X}_r$. Letting $\rho = A(u - \widetilde{\Pi}_r u)$ and noting $(\widetilde{\Pi}_r u)_-^{n-1} = u(t_{n-1})$, we see that the piecewise polynomial $\varepsilon = U - \widetilde{\Pi}_r u \in \mathcal{X}_r$ satisfies

$$\langle\varepsilon_+^{n-1}, X_+^{n-1}\rangle + \int_{I_n}\langle\varepsilon' + A\varepsilon, X\rangle \,\mathrm{d}t = \langle\varepsilon_-^{n-1}, X_+^{n-1}\rangle + \int_{I_n}\langle\rho, X\rangle \,\mathrm{d}t \tag{22}$$

for all $X \in \mathcal{X}_r$, with $\varepsilon_-^{n-1} = U_-^{n-1} - u(t_{n-1})$. A stability result of Schmutz and Wihler [9, Proposition 3.18] yields the estimate

$$\|\varepsilon\|_{I_n}^2 \le C\left(\|\varepsilon_-^{n-1}\|^2 + k_n\int_{I_n}\|\rho\|^2 \,\mathrm{d}t\right), \tag{23}$$

that is,

$$\|U - \widetilde{\Pi}_r u\|_{I_n}^2 \le C\left(\|U_-^{n-1} - u(t_{n-1})\|^2 + k_n\int_{I_n}\|\rho\|^2 \,\mathrm{d}t\right).$$

By putting $v = Au$ in (18) we find $k_n\int_{I_n}\|\rho\|^2 \,\mathrm{d}t \le k_n^2\|\rho\|_{I_n}^2 \le C(k_n^{r+1}\|Au^{(r)}\|_{I_n})^2$, and the desired estimate follows at once.

We are now able to establish the claim (2) from the Introduction.

Theorem 4.5 For $1 \le n \le N$, if $Au^{(r)}$ and $u^{(r+1)}$ are continuous on $\bar{I}_n$, then

$$\begin{aligned}\|U - u + a_{nr}(u)(p_{nr} - p_{n,r-1})\|_{I_n} \le\ & C\|U_-^{n-1} - u(t_{n-1})\| \\ & + Ck_n^{r+1}\left(\|Au^{(r)}\|_{I_n} + \|u^{(r+1)}\|_{I_n}\right).\end{aligned}$$

Proof Write

$$U - u + a_{nr}(u)(p_{nr} - p_{n,r-1}) = (U - \widetilde{\Pi}_r u) + \left(\widetilde{\Pi}_r u - u + a_{nr}(u)(p_{nr} - p_{n,r-1})\right),$$

and apply Theorem 4.3 and 4.4.

We therefore have the following estimate for the homogeneous problem expressed in terms of the initial data.

Corollary 4.6 Assume $k_n \le Ck_{n-1}$ for $2 \le n \le N$ so that (12) holds. If $f \equiv 0$, then for $0 \le s \le r+1$ and $2 \le n \le N$,

$$\|U - u + a_{nr}(u)(p_{nr} - p_{n,r-1})\|_{I_n} \le Ct_n^{s-(r+1)}k^{r+1}\|A^s u_0\|.$$

Proof Taking $q = r+1$ in (12) yields

$$\|U_-^{n-1} - u(t_{n-1})\| \le Ct_{n-1}^{s-(r+1)}k^{r+1}\|A^s u_0\|,$$

and using (7) we have $\|Au^{(r)}(t)\| = \|u^{(r+1)}(t)\| \le Ct^{s-(r+1)}\|A^s u_0\|$. The result follows for $n \ge 2$ after noting that $t_n = t_{n-1} + k_n \le t_{n-1} + Ck_{n-1} \le Ct_{n-1}$.

Remark 4.7 In their proof of (2) for the scalar linear problem

$$u' - au = 0 \quad \text{for } t > 0, \text{with } u(0) = u_0,$$

Adjerid et al. [10, Theorem 3] use an inductive argument to show an expansion of the form

$$U(t) - u(t) = \sum_{j=r}^{2r-2} Q_{nj}(t)k_n^j + O\big(k_n^{2r-1}\big) \quad \text{for } t \in I_n,$$

where $Q_{nj} \in \mathbb{P}_{j-1}$ and $Q_{nr}(t) = c_{np}[p_{nr}(t) - p_{n,r-1}(t)]$ for a constant c_{np}. They extend this result to a homogeneous linear system of ODEs $\boldsymbol{u}' - \boldsymbol{Au} = \boldsymbol{0}$, then a nonlinear scalar problem $u' - f(u) = 0$, and finally a nonlinear system $\boldsymbol{u}' - \boldsymbol{f}(\boldsymbol{u}) = \boldsymbol{0}$.

Remark 4.8 The proof of Theorem 4.4 is largely unaffected if the elliptic term is permitted to have time-dependent coefficients, resulting in a time-dependent operator $A(t)$. The main issue is to verify the stability property (23) for this more general setting. The only other complication is the estimation of $\rho(t)$. Consider, for example, $A(t)u(x,t) = -\nabla \cdot (a(x,t)\nabla u(x,t))$. Since $A(t)u(x,t)$ is of the form $\sum_{m=1}^{M} c_m(x,t)B_m u(x,t)$, where each B_m is a second-order linear differential operator involving only the spatial variables x, it follows that

$$\rho(t) = A(t)\Big(u(t) - \widetilde{\Pi}_r u(t)\Big) = \sum_{m=1}^{M} c_m(x,t)\Big(B_m u(t) - \widetilde{\Pi}_r B_m u(t)\Big),$$

and the final step of the proof becomes

$$k_n \int_{I_n} \|\rho\|^2 \mathrm{d}t \le Ck_n^{2(r+1)} \sum_{m=1}^{M} \|B_m u^{(r)}\|_{I_n}^2.$$

Of course, to exploit this generalization of Theorem 4.4, it would also be necessary to verify the superconvergent error bounds for U_-^n in this case.

5 Practical consequences

Throughout this section, we will assume that

$$\|U_-^{n-1} - u(t_{n-1})\| + \|U - u + a_{nr}(u)(p_{nr} - p_{n,r-1})\|_{I_n} \le C\phi(t_n, u)k_n^{r+1}, \qquad (24)$$

for $2 \le n \le N$, where the factor $\phi(t,u)$ will depend on the regularity of u, which in turn depends on the regularity and compatibility of the initial data u_0 and the source term f. Figure 1 plots the right-hand Gauss–Radau polynomials

$$p_{nr}(t) - p_{n,r-1}(t) = P_r(\tau) - P_{r-1}(\tau)$$

as functions of $\tau \in [-1,1]$ for $r \in \{1,2,3,4\}$. In general, there are $r+1$ points

$$-1 = \tau_0 < \tau_1 < \cdots < \tau_r = 1,$$

such that $\tau_1, \tau_2, \ldots, \tau_r$ are the r zeros of $P_r - P_{r-1}$, and hence are also the abscissas of the right-hand, r-point Gauss–Radau quadrature rule for the interval $[-1,1]$. Recalling our previous notation (13), let $t_{n\ell} = \beta_n(\tau_\ell)$ so that $t_{n-1} = t_{n0} < t_{n1} < \cdots < t_{nr} = t_n$ with

$$p_{nr}(t_{n\ell}) - p_{n,r-1}(t_{n\ell}) = 0 \quad \text{for } 1 \le \ell \le r.$$

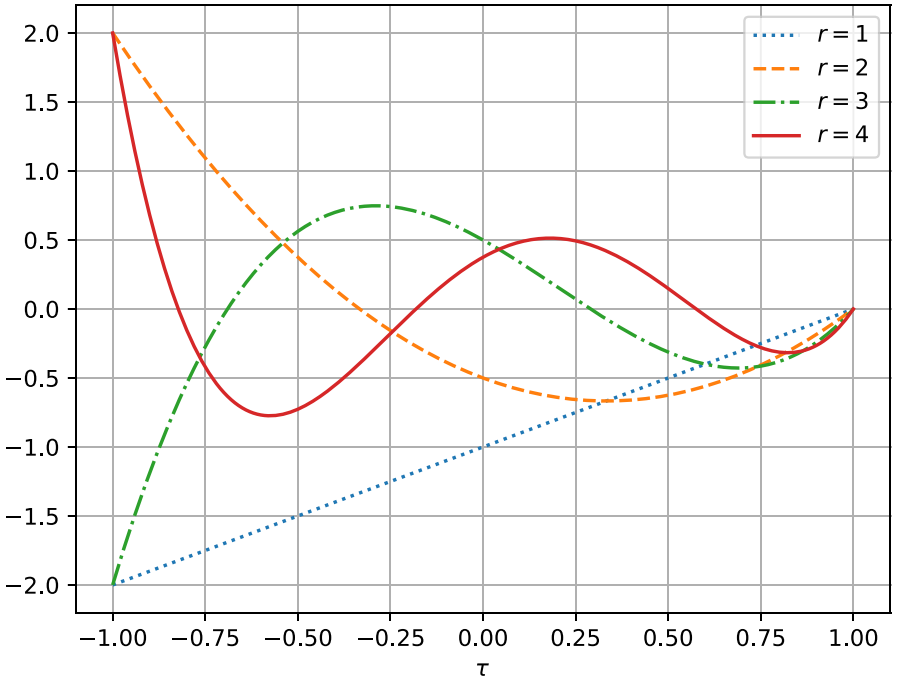

Fig. 1 The polynomials $P_r(\tau) - P_{r-1}(\tau)$

Thus, whereas $U(t) - u(t) = O(k_n^r)$ for general $t \in I_n$, the DG time stepping scheme is superconvergent at the r special points t_{n1}, t_{n2}, …, t_{nr} in the half-open interval $(t_{n-1},t_n]$. More precisely,

$$\|U(t_{n\ell}) - u(t_{n\ell})\| \le C\phi(t_n,u)k_n^{r+1} \quad \text{for } 1 \le \ell \le r.$$

Since $p_{nj}(t_{n-1}) = P_j(-1) = (-1)^j$, another consequence of (24) is that

$$\|U_+^{n-1} - u(t_{n-1}) + 2(-1)^r a_{nr}(u)\| \le C\phi(t_n,u)k_n^{r+1},$$

which, in combination with the estimate $\|U_-^{n-1} - u(t_{n-1})\| \le C\phi(t_n,u)k_n^{r+1}$, shows that the jump $[\![U]\!]^{n-1} = U_+^{n-1} - U_-^{n-1}$ in the DG solution at t_{n-1} satisfies

$$\left\| [\![U]\!]^{n-1} + 2(-1)^r a_{nr}(u) \right\| \le C\phi(t_n,u)k_n^{r+1}. \tag{25}$$

We are therefore able to show, in the following lemma, that $\|[\![U]\!]^{n-1}\|$ is a low-cost and accurate error indicator for the DG solution on I_n.

Lemma 5.1 For ϕ as in (24) and $2 \le n \le N$,

$$\left| \|U-u\|_{I_n} - \|[\![U]\!]^{n-1}\| \right| \le C\phi(t_n,u)k_n^{r+1}.$$

Thus,

$$\|U-u\|_{I_n} = 2\|a_{nr}(u)\| + O\big(k_n^{r+1}\big) = \left\|[\![U]\!]^{n-1}\right\| + O\big(k_n^{r+1}\big).$$

Proof First note that since

$$\max_{-1\le\tau\le1} |P_r(\tau) - P_{r-1}(\tau)| = |P_r(-1) - P_{r-1}(-1)| = 2,$$

we have

$$\|a_{nr}(u)(p_{nr} - p_{n,r-1})\|_{I_n} = |p_{nr}(t_{n-1}) - p_{n,r-1}(t_{n-1})|\,\|a_{nr}(u)\| = 2\|a_{nr}(u)\|. \tag{26}$$

Hence, for $t \in I_n$,

$$\begin{aligned}\|U(t) - u(t)\| \;&\le\; \|U(t) - u(t) + a_{nr}(u)\big[p_{nr}(t) - p_{n,r-1}(t)\big]\| \\ &\quad + \|a_{nr}(u)\big[p_{nr}(t) - p_{n,r-1}(t)\big]\| \le C\phi(t_n,u)k_n^{r+1} + 2\|a_{nr}(u)\|,\end{aligned}$$

and so $\|U-u\|_{I_n} \le 2\|a_{nr}(u)\| + C\phi(t_n,u)k_n^{r+1}$. Conversely,

$$\begin{aligned}2\|a_{nr}(u)\| \;&=\; \|a_{nr}(u)\big[p_{nr}(t_{n-1}) - p_{n,r-1}(t_{n-1})\big]\| \\ &\le\; \|U_+^{n-1} - u(t_{n-1}) + a_{nr}(u)\big[p_{nr}(t_{n-1}) - p_{n,r-1}(t_{n-1})\big]\| \\ &\quad + \|U_+^{n-1} - u(t_{n-1})\| \\ &\le\; C\phi(t_n,u)k_n^{r+1} + \|U-u\|_{I_n},\end{aligned}$$

and therefore

$$\Big| \|U-u\|_{I_n} - 2\|a_{nr}(u)\| \Big| \le C\phi(t_n,u)k_n^{r+1}.$$

Since, by (25),

$$\begin{aligned} |\|[\![U]\!]^{n-1}\| - 2\|a_{nr}(u)\|| &= |\|[\![U]\!]^{n-1}\| - \|2(-1)^{r+1}a_{nr}(u)\|| \\ &\le \Big\| [\![U]\!]^{n-1} + 2(-1)^r a_{nr}(u) \Big\| \le C\phi(t_n,u)k_n^{r+1}, \end{aligned}$$

the result follows.

A unique continuous function $U_* \in \mathcal{X}_{r+1}$ satisfies the $r+1$ interpolation conditions

$$U_*(t_{n\ell}) = \begin{cases} U_-^{n-1} & \text{if } \ell = 0, \\ U(t_{n\ell}) & \text{if } 1 \le \ell \le r-1, \\ U_-^{n} & \text{if } \ell = r, \end{cases}$$

for $1 \le n \le N$, and we see that

$$(U_* - u)(t_{n\ell}) = O\big(k_n^{r+1}\big) \quad \text{for } 0 \le \ell \le r. \tag{27}$$

Makridakis and Nochetto [4] introduced this interpolant in connection with *a posteriori* error analysis of diffusion problems, and called U_* the *reconstruction* of U. The next theorem provides a more explicit description of U_* that we then use to prove U_* achieves the optimal convergence rate of order k_n^{r+1} over the whole subinterval I_n.

Theorem 5.2 For $t \in I_n$ and $1 \le n \le N$, the reconstruction U_* of the DG solution U has the representation

$$U_*(t) = U(t) - \frac{(-1)^r}{2}[\![U]\!]^{n-1}(p_{nr} - p_{n,r-1})(t) = \sum_{j=0}^{r} U_*^{nj} p_{nj}(t),$$

where

$$U_*^{nj} = \begin{cases} U^{nj} & \text{if } 0 \le j \le r-2, \\ U^{n,r-1} + \frac{1}{2}(-1)^r[\![U]\!]^{n-1} & \text{if } j = r-1, \\ -\frac{1}{2}(-1)^r[\![U]\!]^{n-1} & \text{if } j = r. \end{cases}$$

Proof Since the polynomial $(U-U_*)|_{I_n} \in \mathbb{P}_r(\mathbb{H})$ vanishes at $t_{n\ell}$ for $1 \le \ell \le r$, there must be a constant γ such that $U(t) - U_*(t) = \gamma(p_{nr} - p_{n,r-1})(t)$ for $t \in I_n$. Taking the limit as $t \to t_{n-1}^+$, we have $U_+^{n-1} - U_-^{n-1} = \gamma[(-1)^r - (-1)^{r-1}] = 2(-1)^r\gamma$ and so $\gamma = (-1)^r[\![U]\!]^{n-1}/2$. It follows from (14) that

$$a_{nj}(U-U_*) = \frac{2j+1}{k_n}\int_{I_n}(U-U_*)(t)p_{nj}(t)\,\mathrm{d}t = \frac{(-1)^r}{2}[\![U]\!]^{n-1}(\delta_{jr} - \delta_{j,r-1}),$$

implying the formulae for $U_*^{nj} = a_{nj}(U_*)$.

Corollary 5.3 $\|U_* - u\|_{I_n} \le C\phi(t_n, u)k_n^{r+1}$ for $2 \le n \le N$.

Proof We see from the Theorem 5.2 and (26) that

$$\begin{aligned}\|U_* - u\|_{I_n} &= \|U - u - \tfrac{1}{2}(-1)^r[\![U]\!]^{n-1}(p_{nr} - p_{n,r-1})\|_{I_n} \\ &\le \|U - u + a_{nr}(u)(p_{nr} - p_{n,r-1})\|_{I_n} + \tfrac{1}{2}\|[\![U]\!]^{n-1} + 2(-1)^r a_{nr}(u)\|,\end{aligned}$$

so it suffices to apply (24) and (25).

Example 5.4 Let $f \equiv 0$ and let u_0 belong to the domain of A^s. By (9),

$$t^{r-s}\|U(t) - u(t)\| \le Ck^r\|A^s u_0\| \quad \text{if } 0 < t \le T \text{ and } 0 \le s \le r,$$

and by (12),

$$t_n^{2r-1-s}\|U_-^n - u(t_n)\| \le Ck^{2r-1}\|A^s u_0\| \quad \text{if } 1 \le n \le N \text{ and } 0 \le s \le 2r-1.$$

Furthermore, Corollary 4.6 shows that our assumption (24) is satisfied with

$$\phi(t, u) = t^{s-(r+1)}\|A^s u_0\|$$

so

$$t_n^{r+1-s}\|U_* - u\|_{I_n} \le Ck^{r+1}\|A^s u_0\| \quad \text{if } 2 \le n \le N \text{ and } 0 \le s \le r+1.$$

6 Numerical experiments

The computational experiments described in this section were performed in standard 64-bit floating point arithmetic using Julia v1.7.2 on a desktop computer having a Ryzen 7 3700X processor and 32GiB of RAM. The source code is available online [17]. In all cases, we use uniform time steps $k_n = k = T/N$.

6.1 A simple ODE

We begin with the ODE initial-value problem

$$u' + \lambda u = f(t) \quad \text{for } 0 \le t \le 2, \text{with } u(0) = 1,$$

where in place of a linear operator A we have just the scalar $\lambda = 1/2$, and where $f(t) = \cos(\pi t)$. For the piecewise-cubic case with $N = 5$ subintervals, Fig. 2 shows that $U - U_*$ provides an excellent approximation to the error $U - u$, and that the error profile is approximately proportional to $p_{nr} - p_{n,r-1}$ with $r = 4$; cf. (21) and Fig. 1. In particular, superconvergence at the Radau points is apparent. By sampling at 50 points in each subinterval, we estimated the maximum errors

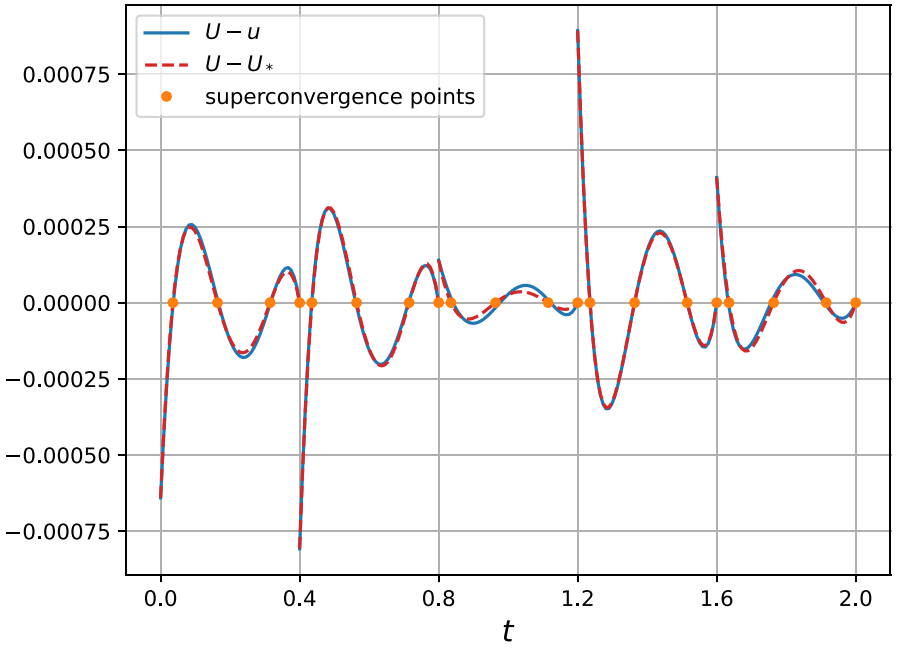

Fig. 2 The DG error $U-u$, the difference $U-U_*$ between the DG solution and its reconstruction, along with the superconvergence points t_{nj} $(1 \le j \le r)$, for the ODE of Section 6.1 using piecewise-cubics $(r=4)$

$$\max_{1\le n\le N} \sup_{t\in I_n} |U(t)-u(t)| \quad \text{and} \quad \max_{1\le n\le N} \sup_{t\in I_n} |U_*(t)-u(t)|,$$

and, as expected from (5) and Corollary 5.3, the values shown in Table 1 exhibit convergence rates $r=4$ and $r+1=5$, respectively. The table also shows a convergence rate $2r-1=7$ for the nodal error $\max_{1\le n\le N} |U_-^n - u(t_n)|$ up to the row where this error approaches the unit roundoff. By using Julia's BigFloat datatype, we were able

Table 1 Errors and convergence rates for the ODE of Section 6.1 using piecewise-cubics $(r=4)$

N	Error in U		Error in U_*		Error in U_-^n	
4	1.75e-03		6.15e-05		5.26e-09	
8	1.36e-04	3.684	2.26e-06	4.769	4.08e-11	7.010
16	8.85e-06	3.945	7.19e-08	4.973	3.27e-13	6.962
32	5.55e-07	3.996	2.26e-09	4.994	2.66e-15	6.941
64	3.48e-08	3.995	7.05e-11	5.000	7.77e-16	1.778
128	2.17e-09	3.999	2.20e-12	4.999	1.55e-15	-1.000
Theory		4		5		7

to observe $O(k^7)$ convergence of U^n_- up to $N = 128$, for which value the nodal error was `1.56e-19`.

6.2 A parabolic PDE in 1D

Now consider the 1D heat equation with constant thermal conductivity $\kappa > 0$,

$$u_t - \kappa u_{xx} = f(x,t) \quad \text{for } 0 < t \le T \text{ and } 0 \le x \le L, \tag{28}$$

subject to the boundary conditions $u(0,t) = 0 = u(L,t)$ for $0 \le t \le T$, and to the initial condition $u(x,0) = u_0(x)$ for $0 \le x \le L$. To obtain a reference solution, we introduce the Laplace transform $\hat{u}(x,z) = \int_0^\infty e^{-zt} u(x,t)\, \mathrm{d}t$, which satisfies the two-point boundary-value problem (with complex parameter z),

$$-\hat{u}_{xx} + \omega^2 \hat{u} = g(x,z) \quad \text{for } 0 \le x \le L, \quad \text{with } \hat{u}(0,z) = 0 = \hat{u}(L,z),$$

where $\omega = (z/\kappa)^{1/2}$ and $g(x,z) = \kappa^{-1}[u_0(x) + \hat{f}(x,z)]$. Consequently, the variation-of-constants formula yields the representation [14, Section 7.3]

$$\begin{aligned} \hat{u}(x,z) \;=\; & \tfrac{\sinh \omega(L-x)}{\omega \sinh \omega L} \textstyle\int_0^x g(\xi,z) \sinh \omega\xi \mathrm{d}\xi \\ & + \tfrac{\sinh \omega x}{\omega \sinh \omega L} \textstyle\int_x^L g(\xi,z) \sinh \omega(L-\xi) \mathrm{d}\xi, \end{aligned}$$

and we then invert the Laplace transform by numerical evaluation of the Bromwich integral [18],

$$u(x,t) = \frac{1}{2\pi i} \int_{\mathcal{C}} e^{zt} \hat{u}(x,z) dz,$$

for a hyperbolic contour $\mathcal{C}$ homotopic to the imaginary axis and passing to the right of all singularities of $\hat{u}(x,z)$.

To discretize in space, we introduce a finite difference grid

$$x_p = ph \quad \text{for } 0 \le p \le P, \quad \text{where } h = L/P,$$

and define $u_p(t) \approx u(x_p,t)$ via the method of lines, replacing u_{xx} with a second-order central difference approximation to arrive at the system of ODEs

$$u_p'(t) - \kappa \frac{u_{p+1}(t) - 2u_p(t) + u_{p-1}(t)}{h^2} = f_p(t) \quad \text{for } 1 \le p \le P-1, \tag{29}$$

where $f_p(t) = f(x_p,t)$ with the boundary conditions $u_0(t) = 0 = u_P(t)$ and the initial condition $u_p(0) = u_0(x_p)$. For our test problem, we choose

$$L = 2, \quad T = 2, \quad \kappa = (L/\pi)^2, \quad u_0(x) = x(L-x), \quad f(x,t) = (1+t)e^{-t}, \tag{30}$$

where the value of the thermal conductivity κ normalizes the time scale by making the smallest eigenvalue of $A = -\kappa(d/dx)^2$ equal 1. We will see below that $u_0 \in D(A^s)$ iff $s < 5/4$, so the regularity of the solution u is limited.

We apply DG to discretize $u_p(t)$ in time and denote the resulting fully discrete solution by $\boldsymbol{U}(t) = [U_p(t)] \approx \boldsymbol{u}(t) = [u_p(t)]$. Figure 3 plots the error in $\boldsymbol{U}$ and in its reconstruction $\boldsymbol{U}_*$ using piecewise-quadratics ($r = 3$) and $N = 8$ equal subintervals in time, with $P = 500$ for the spatial grid. The errors are measured in the discrete L_2-norm, that is,

$$\|\boldsymbol{U}(t) - \boldsymbol{u}(t)\|_h^2 = \sum_{p=0}^{P} |U_p(t) - u(x_p, t)|^2 h,$$

and we observe a clear deterioration in accuracy as t approaches zero.

To speed up the convergence as $h \to 0$, we compute also a second DG solution $U_p^{\text{fine}}(t)$ using a finer spatial grid with $P^{\text{fine}} = 2P$ subintervals, and then perform one step of Richardson extrapolation (on the coarser grid), defining

$$U_p^{\text{R}}(t) = U_{2p}^{\text{fine}}(t) + \frac{1}{3}\left[U_{2p}^{\text{fine}}(t) - U_p(t)\right] \quad \text{for } 0 \le p \le P.$$

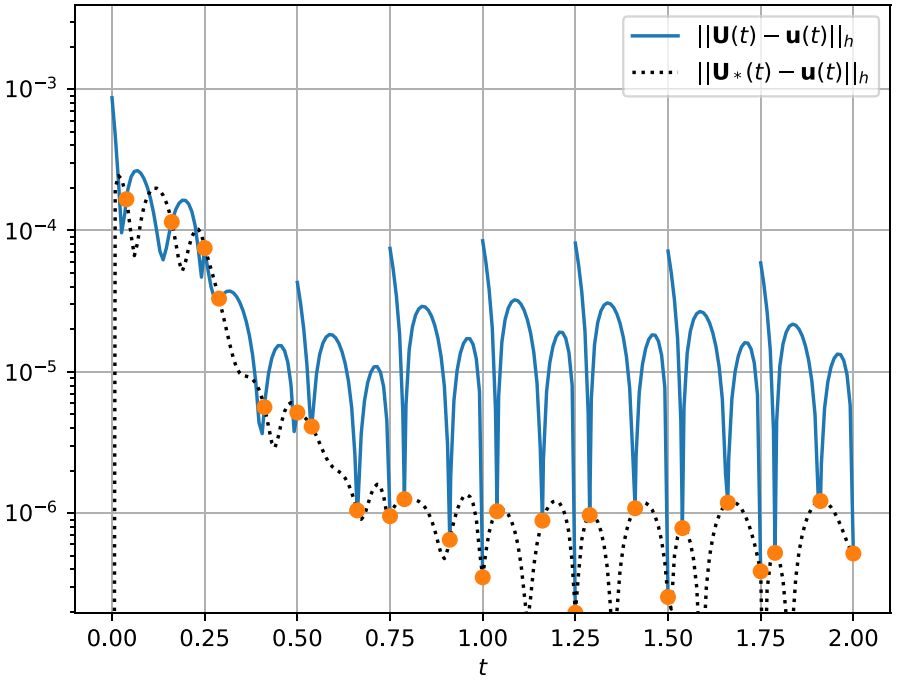

Fig. 3 Time dependence of the errors in the DG solution $U(t)$ and its reconstruction $U_*(t)$ for the 1D heat equation (28), using piecewise-quadratics ($r = 3$) over $N = 8$ time intervals

Table 2 Maximum errors over the time interval $[T/4,T]$ for the 1D heat equation of Section 6.2 using piecewise-quadratics ($r=3$)

N	P	Error in $\boldsymbol{U}^{\mathrm{R}}$		Error in $\boldsymbol{U}^{\mathrm{R}}_*$		Error in $(\boldsymbol{U}^{\mathrm{R}})^n_-$	
8	500	8.46e-05		7.51e-05		5.17e-06	
16	500	1.07e-05	2.978	5.37e-07	7.129	1.15e-07	5.490
32	500	1.35e-06	2.989	2.30e-08	4.544	3.99e-09	4.847
64	500	1.69e-07	2.995	1.21e-09	4.244	1.47e-10	4.762
128	500	2.12e-08	2.997	6.98e-11	4.121	5.80e-12	4.664
Theory			3		4		5

Table 2 shows errors in this spatially extrapolated DG solution over the time interval $[T/4,T]$, that is,

$$\max_{T/4\le t\le T}\|\boldsymbol{U}^{\mathrm{R}}(t)-\boldsymbol{u}(t)\|_h, \tag{31}$$

as well as the corresponding errors in the reconstruction $\boldsymbol{U}^{\mathrm{R}}_*(t)$ and the nodal values $(\boldsymbol{U}^{\mathrm{R}})^n_-$. Again, the observed convergence rates are as expected.

To investigate the time dependence of the error for t near zero, we consider the weighted error in the DG solution

Table 3 Weighted errors for the 1D heat equation of Section 6.2 using piecewise-quadratics ($r=3$) and the indicated exponent α in the weight function $w_\alpha(t)$. The top set of results is for the homogeneous equation ($f\equiv 0$). The bottom set is for the general case (both u_0 and f non-zero)

N	P	Error in $\boldsymbol{U}^{\mathrm{R}}$		Error in $\boldsymbol{U}^{\mathrm{R}}_*$		Error in $(\boldsymbol{U}^{\mathrm{R}})^n_-$	
		$\alpha=r-\frac{5}{4}$		$\alpha=r+1-\frac{5}{4}$		$\alpha=2r-1-\frac{5}{4}$	
8	500	8.52e-05		7.08e-06		1.77e-06	
16	500	1.15e-05	2.893	4.42e-07	4.001	5.53e-08	5.001
32	500	1.49e-06	2.946	2.76e-08	4.000	1.73e-09	5.000
64	500	1.90e-07	2.973	1.73e-09	4.000	5.40e-11	5.000
128	500	2.39e-08	2.987	1.08e-10	4.000	1.69e-12	5.000
Theory			3		4		5
N	P	Error in U^{R}		Error in $\boldsymbol{U}^{\mathrm{R}}_*$		Error in $(\boldsymbol{U}^{\mathrm{R}})^n_-$	
		$\alpha=r-\frac{5}{4}$		$\alpha=r+1-\frac{5}{4}$		$\alpha=2r-1-\frac{5}{4}$	
8	500	8.46e-05		1.66e-06		4.15e-07	
16	500	1.07e-05	2.978	1.03e-07	4.007	1.70e-08	4.606
32	500	1.35e-06	2.989	6.46e-09	3.999	7.89e-10	4.433
64	500	1.69e-07	2.995	4.04e-10	4.000	3.84e-11	4.362
128	500	2.12e-08	2.997	2.52e-11	4.000	1.93e-12	4.311
Theory			3		4		5

Table 4 Maximum errors over the time interval [*T*/4,*T*] for the spatially discrete, 2D heat equation of Section 6.3 using piecewise-quadratics ($r=3$)

N	P_x	P_y	Error in $\boldsymbol{U}_h$		Error in $(\boldsymbol{U}_h)_*$		Error in $(\boldsymbol{U}_h)^n_-$	
8	50	50	5.32e-04		4.70e-04		2.60e-05	
16	50	50	4.60e-05	3.533	1.48e-06	8.316	4.40e-07	5.888
32	50	50	5.15e-06	3.160	6.80e-08	4.440	1.43e-08	4.940
64	50	50	6.10e-07	3.078	4.16e-09	4.029	4.65e-10	4.944
128	50	50	7.42e-08	3.038	2.58e-10	4.010	1.49e-11	4.967
Theory				3		4		5

$$\max_{1\le n\le N}\sup_{t\in I_n} w_\alpha(t)\|\boldsymbol{U}^{\mathrm{R}}(t)-\boldsymbol{u}(t)\|_h \quad \text{where} \quad w_\alpha(t)=\min(t^\alpha,1),$$

and likewise incorporate the weight $w_\alpha(t)$ when measuring the reconstruction error and the nodal error. The top part of Table 3 shows results for the homogeneous problem, that is, with the same data as in (30) except $f(x,t)\equiv 0$. The mth Fourier sine coefficient of u_0 is proportional to m^{-3}, so $\|A^s u_0\|\le C\epsilon^{-1/2}$ for $s=\frac{5}{4}-\epsilon$ and $\epsilon>0$. Based on the estimates in Example 5.4, we choose the weight exponents $\alpha=r-\frac{5}{4}$ for the DG error, $r+1-\frac{5}{4}$ for the reconstruction error, and $2r-1-\frac{5}{4}$ for the nodal error, and observe excellent agreement in the top set of results in Table 3 with the expected convergence rates of order r, $r+1$ and $2r-1$, respectively.

Similar results are found if $u_0(x)\equiv 0$ with nonzero f. Curiously, in the bottom part of Table 3, choosing both u_0 and f as in (30) (so both nonzero) disturbs the observed convergence rates for $(\boldsymbol{U}^{\mathrm{R}})^n_-$, although not for $\boldsymbol{U}^{\mathrm{R}}$ or $\boldsymbol{U}^{\mathrm{R}}_*$.

6.3 A parabolic PDE in 2D

Now consider the 2D heat equation,

$$u_t-\kappa\nabla^2 u=f(x,y,t) \quad \text{for } 0<t\le T \text{ and } (x,y)\in\Omega=(0,L_x)\times(0,L_y), \tag{32}$$

subject to the boundary conditions $u(x,y,t)=0$ for $(x,y)\in\partial\Omega$, and to the initial condition $u(x,y,0)=u_0(x,y)$ for $(x,y)\in\Omega$. We introduce a spatial finite difference grid

$$(x_p,y_q)=(ph_x,qh_y) \quad \text{for } 0\le p\le P_x \text{ and } 0\le q\le P_y,$$

with $h_x=L_x/P_x$ and $h_y=L_y/P_y$. The semidiscrete finite difference solution $u_{pq}(t)\approx u(x_p,y_q,t)$ is then constructed using the standard 5-point approximation to the Laplacian, so that

$$u'_{pq}-\kappa\left(\frac{u_{p+1,q}-2u_{pq}+u_{p-1,q}}{h_x^2}+\frac{u_{p,q+1}-2u_{pq}+u_{p,q-1}}{h_y^2}\right)=f_{pq} \tag{33}$$

for $0 \le t \le T$ and $(x_p,y_q) \in \Omega$, where $f_{pq}(t) = f(x_p,y_q,t)$, together with the boundary condition $u_{pq}(t) = 0$ for $(x_p,y_q) \in \partial\Omega$, and the initial condition $u_{pq}(0) = u_0(x_p,y_q)$ for $(x_p,y_q) \in \Omega$. For $(x_p,y_q) \in \Omega$, we use column-major ordering to arrange the unknowns $u_{pq}(t)$, the source terms $f_{pq}(t)$ and initial data u_{0pq} into vectors $\boldsymbol{u}_h(t)$, $\boldsymbol{f}(t)$ and $\boldsymbol{u}_0 \in \mathbb{R}^M$ for $M = (P_x - 1)(P_y - 1)$. There is then a sparse matrix $\boldsymbol{A}$ such that the system of ODEs (33) leads to the initial-value problem

$$\boldsymbol{u}_h'(t) + \boldsymbol{A}\boldsymbol{u}_h = \boldsymbol{f}(t) \quad \text{for } 0 \le t \le T, \text{ with } \boldsymbol{u}_h(0) = \boldsymbol{u}_0. \tag{34}$$

For our test problem, we take $L_x = L_y = 2$ and $P_x = P_y = 50$ with

$$T = 2, \quad \kappa = 2/\pi^2, \quad u_0(x,y) = x(2-x)y(2-y), \quad f(x,y,t) = (1+t)e^{-t}, \tag{35}$$

where the choice of κ ensures that the smallest Dirichlet eigenvalue of $-\kappa\nabla^2$ on Ω equals 1. Table 4 compares the piecewise-quadratic ($r = 3$) DG solution $\boldsymbol{U}_h(t)$ of the semidiscrete problem (34) with $\boldsymbol{u}_h(t)$, evaluating the latter using numerical inversion of the Laplace transform as before except that now, instead of $\hat{u}(z)$, we work with the spatially discrete approximation $\hat{\boldsymbol{u}}_h(z)$ obtained by solving the (complex) linear system $(z\boldsymbol{I} + \boldsymbol{A})\hat{\boldsymbol{u}}_h(z) = u_0 + \hat{\boldsymbol{f}}(z)$. As with the 1D results in Table 2, we compute the maximum error over the time interval $[T/4,T]$, and observe the expected rates of convergence, keeping in mind that by treating $\boldsymbol{u}_h(t)$ as our reference solution we are ignoring the error from the spatial discretization.

Author contribution William McLean wrote an initial outline of the paper, which subsequently underwent multiple revisions arising from correspondence with Kassem Mustapha. William McLean carried out the numerical computations reported in the paper.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Data availability The paper does not make use of any data sets. The software used to generate the numerical results is available on github [17].

Declarations

Ethics approval and consent to participate Not applicable

Consent for publication Not applicable

Competing interests The authors declare no competing interests.

Human and animal ethics Not applicable

References

1. Eriksson, K., Johnson, C., Thomée, V.: Time discretization of parabolic problems by the discontinuous Galerkin method. ESAIM: M2AN **19**, 611–643 (1985). https://doi.org/10.1051/m2an/1985190406111
2. Schötzau, D., Schwab, C.: Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method. SIAM J. Numer. Anal. **38**, 837–875 (2000). https://doi.org/10.1137/S0036142999352394
3. Chrysafinos, K., Walkington, N.J.: Error estimates for the discontinuous Galerkin methods for parabolic equations. SIAM J. Numer. Anal. **44**, 349–366 (2006). https://doi.org/10.1137/030602289
4. Makridakis, C., Nochetto, R.H.: A posteriori error analysis for higher order dissipative methods for evolution problems. Numer. Math. **104**, 489–514 (2006). https://doi.org/10.1007/s00211-006-0013-6
5. Akrivis, G., Makridakis, C., Nochetto, R.H.: Galerkin and Runge–Kutta methods: unified formulation, a posterior error estimates and nodal superconvergence. Numer. Math. **118**, 429–456 (2011). https://doi.org/10.1007/s00211-011-0363-6
6. Richter, T., Springer, A., Vexler, B.: Efficient numerical realization of discontinuous Galerkin methods for temporal discretization of parabolic problems. Numer. Math. **124**, 151–182 (2013). https://doi.org/10.1007/s00211-012-0511-7
7. Saito, N.: Variational analysis of the discontinuous Galerkin time-stepping method for parabolic equations. IMA J. Numer. Anal. **41**, 1267–1292 (2021). https://doi.org/10.1093/imanum/draa017
8. Leykekhman, D., Vexler, B.: Discrete maximal parabolic regularity for Galerkin finite element methods. Numer. Math. **135**, 923–952 (2017). https://doi.org/10.1007/s00211-016-0821-2
9. Schmutz, L., Wihler, T.P.: The variable-order discontinuous Galerkin time stepping scheme for parabolic evolution problems is uniformly L^∞-stable. SIAM J. Numer. Anal. **37**, 293–319 (2019). https://doi.org/10.1137/17M1158835
10. Adjerid, S., Devin, K.D., Flahery, J.E., Krivodonova, L.: A posteriori error estimation for discontinuous Galerkin solutions of hyperbolic problems. Comput. Methods Appl. Mech. Engrg. **191**, 1097–1112 (2002). https://doi.org/10.1016/S0045-7825(01)00318-8
11. Adjerid, S., Baccouch, M.: Asymptotically exact a posteriori error estimates for a one-dimensional linear hyperbolic problem. Appl. Numer. Math. **60**, 903–914 (2010). https://doi.org/10.1016/j.apnum.2010.04.014
12. Baccouch, M.: The discontinuous Galerkin finite element method for ordinary differential equations. In: Petrova, R. (ed.) Perusal of the Finite Element Method. https://doi.org/10.5772/64967, pp 32–68. InTechOpen (2016)
13. Springer, A., Vexler, B.: Third order convergent time discretization for parabolic optimal control problems with control constraints. Comput. Optim. Appl. **57**, 205–240 (2014). https://doi.org/10.1007/s10589-013-9580-5
14. McLean, W.: Implementation of high-order, discontinuous Galerkin time stepping for fractional diffusion problems. ANZIAM J. **62**, 121–147 (2020). https://doi.org/10.1017/S1446181120000152
15. Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems. Springer, New York (2006)
16. Vlasák, M., Roskovec, F.: On Runge–Kutta, collocation and discontinuous Galerkin methods: mutual connections and resulting consequences to the analysis. In: Programs and Algorithms of Numerical Mathematics. https://eudml.org/doc/269918, pp 231–236. Institute of Mathematics AS CR, Prague (2015)
17. McLean, W.: DGErrorProfile. Github. https://github.com/billmclean/DGErrorProfile (2022)
18. Weideman, J.A.C., Trefethen, L.N.: Parabolic and hyperbolic contours for computing the Bromwich integral. Math. Comp. **76**, 1341–1356 (2007). https://doi.org/10.1090/S0025-5718-07-01945-X