ON MAXIMAL REGULARITY ESTIMATES
FOR DISCONTINUOUS GALERKIN TIME-DISCRETE METHODS

Georgios Akrivis† and Charalambos G. Makridakis‡

Abstract. We consider the discretization of differential equations satisfying the maximal parabolic $L^p$-regularity property in Banach spaces by discontinuous Galerkin methods. We use the maximal regularity framework to establish that the discontinuous Galerkin methods preserve the maximal $L^p$-regularity, satisfy corresponding a posteriori error estimates, and the estimators are of optimal asymptotic order of convergence. In our proofs, we use a suitable interpretation of the discontinuous Galerkin methods as modified Radau IIA methods.

Key words. A posteriori error estimates, discontinuous Galerkin methods, parabolic equations, maximal parabolic regularity, discrete maximal parabolic regularity, Radau IIA methods.

AMS subject classifications. 65M12, 65M15

1. Introduction. We consider the discretization of differential equations satisfying the maximal parabolic $L^p$-regularity property in Banach spaces by discontinuous Galerkin methods. In this paper we use the maximal regularity framework to establish that the discontinuous Galerkin methods (i) preserve the maximal parabolic $L^p$-regularity, (ii) satisfy corresponding a posteriori error estimates, and (iii) the estimators are of optimal asymptotic order of convergence.

1.1. Maximal parabolic regularity. We consider an initial value problem for a linear parabolic equation,

\[
\begin{aligned}
\frac{d}{dt} u(t) + Au(t) &= f(t), & 0 < t < T, \\
u(0) &= 0,
\end{aligned}
\]

in a Banach space $X$. Our structural assumption is that the operator $A$ is the generator of an analytic semigroup on $X$ having maximal $L^p$-regularity, i.e., the solution $u$ of (1.1) satisfies the stability estimate

\[
\|u\|_{L^p((0,T);X)} + \|Au\|_{L^p((0,T);X)} \leq c_{p,X} \|f\|_{L^p((0,T);X)} \quad \forall f \in L^p((0,T);X)
\]

for some, or, as it turns out, for all $p \in (1,\infty)$, with a constant $c_{p,X}$ independent of $T$, depending only on $p$ and $X$. In other words, $u'$ and $Au$ are well defined and have the same, i.e., maximal, regularity as their sum $u' + Au$, that is, the given forcing term $f$. 

\*Received by the editors December 1, 2020, accepted for publication (in revised form) June 18, 2021.
†Department of Computer Science and Engineering, University of Ioannina, 451 10 Ioannina, Greece, and Institute of Applied and Computational Mathematics, FORTH, 700 13 Heraklion, Crete, Greece.
Email address: akrivis@cse.uoi.gr
‡Modeling and Scientific Computing, Department of Mathematics & Applied Mathematics, University of Crete / Institute of Applied and Computational Mathematics, FORTH, 700 13 Heraklion, Crete, Greece, and MPS, University of Sussex, Brighton BN1 9QH, United Kingdom.
Email address: c.g.makridakis@iacm.forth.gr
It is known that every generator of a bounded analytic semigroup on a Hilbert space has maximal $L^p$-regularity and that a Banach space with an unconditional basis satisfying this property is a Hilbert space; see [8] and [10], respectively. We refer to [21] for a fundamental characterization of the maximal $L^p$-regularity property on $X = L^q(\Omega)$, with $1 < q < \infty$ and $\Omega$ a domain in $\mathbb{R}^d$, and more generally on unconditioned martingale differences (UMD) spaces, and to the lecture notes [14] for an excellent account of the theory. Coercive elliptic differential operators on $L^q(\Omega)$ with general boundary conditions possess the maximal $L^p$-regularity property; see [14] and references therein.

1.2. The numerical methods. We consider the discretization of the initial value problem (1.1) by discontinuous Galerkin methods.

Let $N \in \mathbb{N}, k = T/N$ be the constant time step, $t_n := nk, n = 0, \ldots, N$, be a uniform partition of the time interval $[0, T]$, and $J_n := (t_n, t_{n+1}]$. For $q \in \mathbb{N}$, with $0 < c_1 < \cdots < c_q = 1$ the Radau nodes in the interval $[0, 1]$, let $t_{ni} := t_n + c_ik, i = 1, \ldots, q$, be the intermediate nodes; we shall also use the notation $t_{n0} := t_n$.

For $s \in \mathbb{N}_0$, we denote by $\mathbb{P}(s)$ and $\mathbb{P}_{\mathcal{X}'}(s)$ the spaces of polynomials of degree at most $s$ with coefficients in the domain $\mathcal{D}(A)$ of the operator $A, \mathcal{D}(A) := \{v \in X : Av \in X\}$, and in the dual $X'$ of $X$, respectively, i.e., the elements $g$ of $\mathbb{P}(s)$ and of $\mathbb{P}_{\mathcal{X}'}(s)$, respectively, are of the form

$$g(t) = \sum_{j=0}^{s} t^j w_j, \quad w_j \in \mathcal{D}(A) \quad \text{and} \quad w_j \in X', \quad j = 0, \ldots, s.$$ 

With this notation, let $V_k^c(s)$ and $V_k^d(s)$ be the spaces of continuous and possibly discontinuous piecewise elements of $\mathbb{P}(s)$, respectively,

$$V_k^c(s) := \{v \in C([0, T]; \mathcal{D}(A)) : v|_{J_n} \in \mathbb{P}(s), \ n = 0, \ldots, N - 1\},$$

$$V_k^d(s) := \{v : [0, T] \to \mathcal{D}(A), \ v|_{J_n} \in \mathbb{P}(s), \ n = 0, \ldots, N - 1\}.$$ 

We denote by $\langle \cdot, \cdot \rangle$ the duality pairing between $X$ and $X'$.

For $q \in \mathbb{N}$, with starting value $U(0) = U_0 = 0$, we consider the discretization of the initial value problem (1.1) by the discontinuous Galerkin method $dG(q-1)$, i.e., we seek $U \in V_k^d(q-1)$ such that

$$\int_{J_n} (\langle U', v \rangle + \langle AU, v \rangle) \, dt + \langle U_n^+ - U_n, v_n^+ \rangle = \int_{J_n} \langle f, v \rangle \, dt \quad \forall v \in \mathbb{P}_{\mathcal{X}'}(q-1)$$

for $n = 0, \ldots, N - 1$. As usual, we use the notation $v_n := v(t_n), v_n^+ := \lim_{s \searrow 0} v(t_n + s)$.

Following [18], we define the reconstruction operator $\hat{I} : V_k^d(q-1) \to V_k^c(q), \hat{U} := \hat{I}U$, via

$$\hat{U}_n^+ = U_n,$$

$$\int_{J_n} \langle \hat{U}', v \rangle \, dt = \int_{J_n} \langle U', v \rangle \, dt + \langle U_n^+ - U_n, v_n^+ \rangle \quad \forall v \in \mathbb{P}_{\mathcal{X}'}(q-1).$$

This operator is in fact an extended interpolant at the Radau nodes, [18]: $\hat{U}$ is uniquely defined by (1.4) and satisfies

$$\hat{U}(t_{nj}) = U(t_{nj}), \quad j = 0, \ldots, q \quad (U(t_{n0}) = U_n).$$
Using this reconstruction, we can reformulate the discontinuous Galerkin method as
\[
\int_{J_n} \left( (\tilde{U}', v) + \langle AU, v \rangle \right) \, dt = \int_{J_n} \langle f, v \rangle \, dt \quad \forall v \in P_{X'}(q - 1).
\]
Denoting by \( P_{q-1} \) the piecewise \( L^2 \)-projection onto \( V_k^d(q - 1) \), relation (1.5) implies the pointwise equation
\[
\tilde{U}' + AU = P_{q-1} f.
\]
Relationship (1.6) will be important for the a posteriori error analysis in the sequel; it will allow us to apply the (continuous) maximal parabolic regularity property to the error equation.

1.3. Main results. First, we prove that the discontinuous Galerkin methods preserve the maximal parabolic regularity. Then, we combine the pointwise form (1.6) of the discontinuous Galerkin method and the maximal parabolic regularity of the differential equation to establish a posteriori error estimates. Finally, we use the discrete maximal parabolic regularity of the numerical method to show that the a posteriori error estimator is of optimal order.

Discrete maximal regularity. Utilizing the interpretation of discontinuous Galerkin methods as modified Radau IIA methods and the known maximal regularity property of Radau IIA methods, we prove that the discontinuous Galerkin methods preserve the maximal regularity property; more precisely:

**Theorem 1.1 (Discrete maximal regularity).** The discontinuous Galerkin approximations \( U_0, \ldots, U_N \in \mathcal{D}(A) \) are well defined by (1.3) and satisfy the maximal parabolic regularity stability estimates
\[
\| (\partial U_n)_{n=1}^N \|_{\ell^p(X)} + \| (AU_n)_{n=1}^N \|_{\ell^p(X)} \leq C_{p,X} \| f \|_{L^p((0,T);X)}
\]
and
\[
\sum_{i=1}^q \| (AU_{ni})_{n=0}^{N-1} \|_{\ell^p(X)} \leq C_{p,X} \| f \|_{L^p((0,T);X)}
\]
with \( U_{ni} := U(t_{ni}) \). Furthermore,
\[
\| \tilde{U}' \|_{L^p((0,T);X)} + \| A\tilde{U} \|_{L^p((0,T);X)} \leq C_{p,X} \| f \|_{L^p((0,T);X)},
\]
where \( C_{p,X} \) denotes a constant independent of \( N \) and \( T \).

Here, for a sequence \( (v_n)_{n \in \mathbb{N}} \subset X \), we used the notation
\[
\partial v_n := \frac{v_n - v_{n-1}}{k} \quad \text{and} \quad \| (v_n)_{n=1}^M \|_{\ell^p(X)} := \left( k \sum_{n=1}^M \| v_n \|_{X}^p \right)^{1/p}.
\]
Notice that \( \| (v_n)_{n=1}^M \|_{\ell^p(X)} \) is the \( L^p((0,t_M);X) \) norm of the piecewise constant function \( v \) taking the values \( v(t) = v_{n+1}, t_n < t < t_{n+1} \).
Our proof for the maximal parabolic regularity of discontinuous Galerkin methods hinges on the recent corresponding result for Radau IIA methods, see [12], and, as in [12], the result is valid for differential equations with the maximal parabolic regularity. This is possible through a precise reformulation of discontinuous Galerkin methods as modified Radau IIA methods, Lemma 2.2. This result will be instrumental in the a posteriori error analysis as well.

Logarithmically quasi-maximal parabolic regularity results for discontinuous Galerkin methods were recently established in [16, 17]. The approach taken in [16, 17] is fundamentally different from ours, and it does not rely on the maximal parabolic regularity of the underlined differential equation. On the one hand, this makes the results of [16, 17] more general as they are even valid for Banach spaces such as $L^1(\Omega)$ and $L^\infty(\Omega)$ as well as for nonconstant time steps; on the other hand, as opposed to the stability results proved herein, the bounds of [16, 17] contain a logarithmic factor depending on the time steps, a natural price to be paid for their generality.

**A posteriori error estimates.** We utilize the pointwise formulation of the discontinuous Galerkin method in combination with the maximal regularity of the differential equation to establish a posteriori error estimates. To this end, we denote by $R$ the residual of the reconstruction $\hat{U}$ of the discontinuous Galerkin approximation $U$,

$$R(t) := \hat{U}'(t) + A\hat{U}(t) - f(t), \quad t \in (t_n, t_{n+1}], \quad n = 0, \ldots, N - 1,$$

i.e., the amount by which $\hat{U}$ misses being exact solution of the differential equation in (1.1).

Then, the error $e := u - \hat{U}$ satisfies the error equation

$$e'(t) + Ae(t) = -R(t), \quad t \in (t_n, t_{n+1}], \quad n = 0, \ldots, N - 1.$$

Such pointwise error equations are instrumental in a posteriori error analyses using reconstruction operators; see, e.g., [2, 3, 4, 6, 19]. Now, the maximal $L^p$-regularity of the operator $A$ applied to the error equation (1.11) yields the desired a posteriori error estimate

$$\|e'\|_{L^p(0, t; X)} + \|Ae\|_{L^p(0, t; X)} \leq c_{p,X} \|R\|_{L^p(0, t; X)} \forall f \in L^p((0, \infty); X)$$

for all $0 < t \leq T$, for any $p \in (1, \infty)$, with a constant $c_{p,X}$ depending only on $p$ and $X$. Notice that since the residual $R$ is a computable quantity, depending only on the numerical solution $U$ and the given forcing term $f$, (1.12) is an a posteriori error estimate.

Our main task then in Section 3 is to show that the estimator $\|R\|_{L^p((0, t); X)}$ is of asymptotic optimal order of convergence as compared to a priori bounds for the error in the maximal regularity framework. Our main result is stated in Theorem 3.4. To do that, we have chosen to use a discrete consistency analysis similar to [1] for Radau IIA methods exploiting the precise connection provided by Lemma 2.2.

An outline of the paper is as follows. In Section 2, we prove Theorem 1.1. Section 3 is devoted to the a posteriori error analysis; furthermore, we briefly discuss the extension of our results to nonautonomous equations.

**2. Discrete maximal regularity.** In this section we prove that the discontinuous Galerkin methods preserve the maximal parabolic $L^p$-regularity property. Our proof utilizes the interpretation of discontinuous Galerkin methods as modified Radau IIA methods with the
modification concerning exclusively the forcing term \( f \); this in combination with the maximal parabolic regularity result for Radau IIA methods leads to the desired property for discontinuous Galerkin methods.

The discrete maximal regularity property is of independent interest; in addition, it will be our main tool to prove that the a posteriori estimator is of optimal order.

### 2.1. Radau IIA methods.

Our proof of the maximal parabolic regularity of discontinuous Galerkin methods is based on the recent corresponding result for Radau IIA methods; see \([12]\).

Here, we recall the Radau IIA methods.

The \( q \)-stage Radau IIA method is specified by the coefficients

\[
\begin{align*}
(2.1) \quad a_{ij} &= \int_0^{c_i} \ell_j(\tau) \, d\tau, \quad b_i = \int_0^1 \ell_i(\tau) \, d\tau \quad (= a_{qi}), \quad i, j = 1, \ldots, q;
\end{align*}
\]

here, \( \ell_1, \ldots, \ell_q \in \mathbb{P}_{q-1} \) are the Lagrange polynomials for the Radau nodes \( c_1, \ldots, c_q, \ell_i(c_j) = \delta_{ij} \). The coefficient matrix \( \mathcal{O} := (a_{ij})_{i,j=1,\ldots,q} \) is invertible and the stability function \( r \) of the method vanishes at infinity.

Relations (2.1) reflect the fact that the Radau IIA method is of collocation type, i.e., its stage order is \( q \). The first member of this family, for \( q = 1 \), is the implicit Euler method.

With starting value \( U_0 = 0 \), we consider the discretization of the initial value problem (1.1) by the \( q \)-stage Radau IIA method: we recursively define approximations \( U_n \in \mathcal{D}(\mathcal{A}) \) to the nodal values \( u(t_n) \), as well as internal approximations \( U_{ni} \in \mathcal{D}(\mathcal{A}) \) to the intermediate values \( u(t_{ni}) \), by

\[
\begin{align*}
(2.2) \quad U_{ni} = U_n - k \sum_{j=1}^q a_{ij} (\mathcal{A}U_{nj} - f(t_{nj})), \quad i = 1, \ldots, q, \\
U_{n+1} = U_n - k \sum_{i=1}^q b_i (\mathcal{A}U_{ni} - f(t_{ni})),
\end{align*}
\]

\( n = 0, \ldots, N - 1 \). Notice that \( U_{n+1} = U_{nq} \); this is a consequence of the fact that \( a_{qi} = b_i, i = 1, \ldots, q \).

The maximal parabolic regularity property for Radau IIA methods is:

**Lemma 2.1** ([12, Corollary 5.2, Theorem 5.1]; maximal regularity of Radau IIA methods).

The Radau IIA approximations \( U_0, \ldots, U_N \in \mathcal{D}(\mathcal{A}) \) are well defined by (2.2) and satisfy the maximal parabolic regularity stability estimates

\[
(2.3) \quad \|(\partial U_n)_{n=1}^N\|_{\ell^p(X)} + \| (\mathcal{A}U_n)_{n=1}^N\|_{\ell^p(X)} \leq C_{p,X} \sum_{i=1}^q \|(f(t_{ni}))_{n=0}^{N-1}\|_{\ell^p(X)}
\]

and

\[
(2.4) \quad \sum_{i=1}^q \|(\mathcal{A}U_{ni})_{n=0}^{N-1}\|_{\ell^p(X)} \leq C_{p,X} \sum_{i=1}^q \|(f(t_{ni}))_{n=0}^{N-1}\|_{\ell^p(X)}
\]

with a constant \( C_{p,X} \) independent of \( N \) and \( T \).
2.2. Discontinuous Galerkin methods as modified Radau IIA methods. Combining Lemma 2.1 with the interpretation of discontinuous Galerkin methods as modified Radau IIA methods, we shall prove Theorem 1.1, i.e., the discrete maximal regularity property of discontinuous Galerkin methods.

In addition to the Lagrange polynomials \( \ell_1, \ldots, \ell_q \in \mathbb{P}^{q-1} \) for the Radau nodes \( c_1, \ldots, c_q \), see (2.1), we shall use the Lagrange polynomials \( \hat{\ell}_0, \ldots, \hat{\ell}_q \in \mathbb{P}^q \) for the points \( c_0, c_1, \ldots, c_q \) with \( c_0 = 0 \). The corresponding Lagrange polynomials shifted to the interval \( \bar{J}_n \) are denoted by \( \ell_{n_i} \) and \( \hat{\ell}_{n_i} \), respectively.

It is known that the discontinuous Galerkin methods are related to Runge–Kutta methods; see \([15, 11]\) and \([20, \text{Chapter 12}]\). A connection through rational functions and nodal values at a uniform partition of each \( J_n \) was used in \([16, 17]\); see also \([9, \text{p. 1322}]\). The next lemma provides an explicit relationship of the values of the discontinuous Galerkin approximation \( U_n := U(t_n) \), \( U_{n_i} := U(t_{n_i}) \), \( i = 1, \ldots, q \), satisfying the modified Radau IIA method

\[
\begin{align*}
U_{n_i} &= U_n - k \sum_{j=1}^{q} a_{ij} (AU_{nj} - f_{nj}), \quad i = 1, \ldots, q, \\
U_{n+1} &= U_n - k \sum_{i=1}^{q} b_i (AU_{ni} - f_{ni}),
\end{align*}
\]

\( n = 0, \ldots, N - 1 \), with the modification consisting in the fact that the nodal values \( f(t_{n_i}) \) of the forcing term have been replaced by the averages

\[
f_{n_i} := \frac{1}{J_n} \int_{J_n} \ell_{n_i}(s) f(s) \, ds = \frac{1}{b_i k} \int_{J_n} \hat{\ell}_{n_i}(s) f(s) \, ds, \quad i = 1, \ldots, q.
\]

Again, \( U_{n+1} = U_{nq} \). Furthermore, (2.5) written in terms of \( \hat{U} \) is a perturbed collocation method in each interval \( J_n \) with starting value \( \hat{U}_{n0} = U_n \), namely

\[
\hat{U}'(t_{n_i}) + A\hat{U}_{ni} = f_{n_i}, \quad i = 1, \ldots, q.
\]

**Proof.** Using the fact that the Radau quadrature rule is exact for polynomials of degree up to \( 2q - 2 \), we see that (1.5) implies, for all \( v \in \mathbb{P}_{X'}^{(q-1)} \),

\[
k \sum_{j=1}^{q} b_j \langle \hat{U}'(t_{nj}) + AU_{nj}, v(t_{nj}) \rangle = \int_{J_n} \langle f, v \rangle \, ds.
\]

The choice \( v = \ell_{n_i}w \), with \( w \) any element of \( X' \), yields

\[
b_i k \hat{U}'(t_{n_i}) + b_i k A\hat{U}_{n_i} = \int_{J_n} \ell_{n_i}(s) f(s) \, ds,
\]
which is (2.7).

To prove (2.5), we observe that (2.7) implies that \( \hat{U}' \) has the representation

\[
\hat{U}'(t) = \sum_{j=1}^{q} \ell_{nj}(t)(f_{nj} - AU_{nj}), \quad t \in J_n,
\]

whence, integrating, we infer that

\[
U_{ni} - U_n = \int_{t_n}^{t_{ni}} \hat{U}'(t) \, dt = \sum_{j=1}^{q} \int_{t_n}^{t_{ni}} \ell_{nj}(t)(f_{nj} - AU_{nj}) \, dt = k \sum_{j=1}^{q} a_{ij}(f_{nj} - AU_{nj}),
\]

and the proof is complete.

\[\text{Remark 2.3 (Existence and uniqueness of approximations).} \]

The Radau IIA approximations \( U_{ni} \) are well defined by (2.2); see [12, (5.3)]. Consequently, the values \( U_{ni} = U(t_{ni}), i = 1, \ldots, q, \) of the discontinuous Galerkin approximation at the intermediate nodes are also well defined by (2.5). Since \( U \) is a polynomial of degree at most \( q - 1 \) in each subinterval \( J_n := (t_n, t_{n+1}] \), this argument yields existence and uniqueness of the discontinuous Galerkin approximation \( U \in V^{N}(q-1); \) see (1.3).

\[\text{2.3. Proof of Theorem 1.1.} \]

First, from Lemma 2.1 and the modified Radau IIA formulation (2.5) of the discontinuous Galerkin method, we obtain the preliminary stability estimates

\[
\|(\partial U_{ni})_{n=1}^{N}||_{\ell^{p}(X)} + \|(AU_{ni})_{n=1}^{N}||_{\ell^{p}(X)} \leq C_{p,X} \sum_{i=1}^{q} \|(f_{ni})_{n=0}^{N-1}||_{\ell^{p}(X)} \tag{2.10}
\]

and

\[
\sum_{i=1}^{q} \|(AU_{ni})_{n=0}^{N-1}||_{\ell^{p}(X)} \leq C_{p,X} \sum_{i=1}^{q} \|(f_{ni})_{n=0}^{N-1}||_{\ell^{p}(X)}, \tag{2.11}
\]

with a constant \( C_{p,X} \) independent of \( N \) and \( T \).

Therefore, to complete the proof of (1.7) and (1.8), it suffices to show that

\[
\sum_{i=1}^{q} \|(f_{ni})_{n=0}^{N-1}||_{\ell^{p}(X)} \leq \gamma \left\| f \right\|_{L^{p}((0,T);X)}, \tag{2.12}
\]

with a constant \( \gamma \) independent of \( N \) and the time step \( k \). Now, with \( p' \) the dual exponent to \( p, 1/p + 1/p' = 1 \), we have

\[
\left\| f_{ni} \right\|_{X}^{p} = \frac{1}{(b_{k})^{p}} \left\| \int_{J_n} \ell_{ni}(s)f(s) \, ds \right\|_{X}^{p} \leq \frac{1}{(b_{k})^{p}} \left( \int_{J_n} \left\| \ell_{ni}(s)f(s) \right\|_{X} \, ds \right)^{p} \leq \frac{1}{(b_{k})^{p}} \left( \int_{J_n} \left\| f(s) \right\|_{X}^{p'} \, ds \right)^{p/p'}.
\]

\[\text{\hfill \Box}\]
Notice that
\[
\int_{J_n} |\ell_{ni}(s)|^{p'} \, ds = k \int_0^1 |\ell_i(\tau)|^{p'} \, d\tau.
\]
In addition, \(k^{p'/p} = k^{-1}\). Therefore,
\[
(2.13) \quad k\|f_{ni}\|_X^p \leq \left( \frac{1}{b_i} \right) \int_0^1 |\ell_i(\tau)|^{p'} \, d\tau \int_{J_n} \|f(s)\|_X^p \, ds,
\]
and the proof of (2.12) is complete. To show \((1.9)\), we first notice that \((U_{n0} = U_n)\)
\[
\int_{J_n} \|A\hat{U}\|_X^p \, dt = \int_{J_n} \|A \sum_{i=0}^q \hat{\ell}_{ni}U_{ni}\|_X^p \, dt 
\leq \left( \sum_{i=0}^q k\|\hat{\ell}_{ni}\|_{L^\infty(J_n)} \|AU_{ni}\|_X \right)^p 
\leq \left( \sum_{i=0}^q \|\hat{\ell}_{ni}\|_{L^\infty(J_n)}^q \right)^{p/p'} \left( \sum_{i=0}^q k\|AU_{ni}\|_X^p \right).
\]
Using similar arguments for \(\|AU\|_{L^p((0,T);X)}\), we obtain, in view of \((1.7)\) and \((1.8)\),
\[
\|A\hat{U}\|_{L^p((0,T);X)} + \|AU\|_{L^p((0,T);X)} \leq C_{p,X} \|f\|_{L^p((0,T);X)}.
\]
To complete the proof of \((1.9)\), it suffices to note that
\[
\|A\hat{U}'\|_{L^p((0,T);X)} \leq \|AU\|_{L^p((0,T);X)} + \|P_{q-1}f\|_{L^p((0,T);X)} \leq C_{p,X} \|f\|_{L^p((0,T);X)}.
\]
In the last estimate we have also used the fact that \(\|P_{q-1}f\|_{L^p((0,T);X)} \leq C_{p,X} \|f\|_{L^p((0,T);X)}\). This is indeed true as the following simple argument shows: Let \(\ell_0, \ldots, \ell_{q-1} \in \mathbb{P}_{q-1}\) be the normalized Legendre polynomials in the interval \([0,1]\), i.e., \(\int_0^1 \ell_i(\tau)\ell_j(\tau) \, d\tau = \delta_{ij}\). Then, \(\hat{\ell}_{ni}(t) = \ell_{ni}(t_n + k\tau) := \frac{1}{\sqrt{k}}\ell_{ni}(\tau), i = 0, \ldots, q - 1\), are the corresponding polynomials shifted to the interval \(J_n\) and normalized. Now,
\[
P_{q-1}f = \sum_{i=1}^q \hat{\ell}_{ni} \int_{J_n} \hat{\ell}_{ni}(t) f(t) \, dt \text{ in } J_n
\]
and
\[
\left\| \int_{J_n} \hat{\ell}_{ni}(t) f(t) \, dt \right\|_{L^p(J_n;X)}^p = \int_{J_n} |\hat{\ell}_{ni}(s)|^p \, ds \left\| \int_{J_n} \hat{\ell}_{ni}(t) f(t) \, dt \right\|_X^p 
\leq \int_{J_n} |\hat{\ell}_{ni}(s)|^p \, ds \left( \int_{J_n} |\hat{\ell}_{ni}(t)|^{p'} \, dt \right)^{p/p'} \int_{J_n} \|f(t)\|_X^p \, dt 
\leq \int_0^1 |\ell_i(\tau)|^p \, d\tau \left( \int_0^1 |\ell_i(\tau)|^{p'} \, d\tau \right)^{p/p'} \int_{J_n} \|f(t)\|_X^p \, dt,
\]
which lead to the desired estimate.
3. A posteriori error estimates. In this section we prove that the a posteriori error estimator is of optimal order. The approach taken is a suitable modification of the analysis in [1].

3.1. A posteriori error estimates. Recall the piecewise $L^2$-projection $P_{q-1}$ onto $V^d_k(q-1)$ and the pointwise equation

$$\hat{U}' + AU = P_{q-1}f;$$

(3.1) see (1.6). This equation can be rewritten as

$$\hat{U}' + A\hat{U} = A\hat{U} - AU + P_{q-1}f = R + f,$$

(3.2) where $R$ is the residual of $\hat{U}$,

$$R(t) := \hat{U}'(t) + A\hat{U}(t) - f(t), \quad t \in (t_n, t_{n+1}], \quad n = 0, \ldots, N - 1.$$

Then, the error $e = u - \hat{U}$ satisfies the error equation

$$e'(t) + Ae(t) = -R(t), \quad t \in (t_n, t_{n+1}], \quad n = 0, \ldots, N - 1.$$

(3.4) Now, the maximal $L^p$-regularity of the operator $A$ applied to the error equation (3.4) yields the a posteriori error estimate

$$\|e'(t)\|_{L^p((0, T); X)} + \|Ae(t)\|_{L^p((0, T); X)} \leq c_{p,X}\|R\|_{L^p((0, T); X)},$$

(3.5) (cf. (1.12)), for all $0 < t \leq T$, for any $p \in (1, \infty)$, with a constant $c_{p,X}$ depending only on $p$ and $X$. Next, we shall prove that the a posteriori error estimator on the right-hand side of (3.5) is of optimal order.

3.1.1. Explicit representation of the residual. One of the fundamental properties of the reconstruction operator is the fact that $U$ and $\hat{U}$ coincide at the Radau nodes of each time interval. Therefore,

$$\tilde{U}(t) - U(t) = \tilde{U}(t) - I_{q-1}\tilde{U}(t).$$

Thus, (3.3) implies

$$R(t) = A[\tilde{U}(t) - I_{q-1}\tilde{U}(t)] - [f(t) - P_{q-1}f(t)], \quad t \in (t_n, t_{n+1}],$$

(3.6) $n = 0, \ldots, N - 1.$

Using the Kowalewski representation [7, Ex. 1, pp. 71–72] of the polynomial interpolation remainder, we have

$$\tilde{U}(t) - I_{q-1}\tilde{U}(t) = k^q\Phi_q(t - \frac{t_n}{k})\tilde{U}(q),$$

(3.7) for $t \in J_n$, with $\Phi_q(s) := \frac{1}{q!}\prod_{i=1}^d(s - c_i), s \in [0, 1]$, cf. [1, (3.2)], and thus

$$R(t) = k^q\Phi_q(t - \frac{t_n}{k})A\tilde{U}(q) - [f(t) - P_{q-1}f(t)],$$

(3.8) $t \in J_n, n = 0, \ldots, N - 1.$
3.2. Optimality of the estimator via a priori error analysis. Although discontinuous Galerkin is a finite element method, the presented analysis is a modification, at certain points only, of the proofs of [1]. The consistency estimates for the Radau IIA methods can be directly used. Then, the discrete maximal parabolic regularity property of the \( q \)-stage Radau IIA method leads to convergence rates for the discontinuous Galerkin method. These bounds will be instrumental to establish sharp asymptotic upper bounds for the estimator.

3.2.1. Error estimates. Let \( e_n := u(t_n) - \hat{U}_n = u(t_n) - U_n, n = 0, \ldots, N \), and \( e_{ni} := u(t_{ni}) - \hat{U}_{ni} = u(t_{ni}) - U_{ni}, n = 0, \ldots, N - 1 \). Notice that \( e_{n+1} = e_{nq}, n = 0, \ldots, N - 1 \).

The consistency errors \( E_{ni} \) and \( E_{n+1} \) of the \( q \)-stage Radau IIA method are determined by

\begin{equation}
\begin{cases}
  u(t_{ni}) = u(t_n) - k \sum_{j=1}^{q} a_{ij} (Au(t_{nj}) - f(t_{nj})) + E_{ni}, & i = 1, \ldots, q, \\
  u(t_{n+1}) = u(t_n) - k \sum_{i=1}^{q} b_i (Au(t_{ni}) - f(t_{ni})) + E_{n+1}.
\end{cases}
\end{equation}

Notice that \( E_{n+1} = E_{nq}, n = 0, \ldots, N - 1 \).

We recall an easy consistency estimate of the \( q \)-stage Radau IIA method from [1, Lemma 3.2] for \( q \geq 2 \). The case \( q = 1 \) is easy; see Remark 3.3.

Lemma 3.1 (Consistency estimate). If the solution \( u \) of (1.1) is sufficiently smooth, then the following consistency estimate holds, for \( q \geq 2 \),

\begin{equation}
\|E_{n+1}\|_X + k \sum_{i=1}^{q} \|E_{ni}\|_X \leqck^{q+2}, \quad n = 0, \ldots, N - 1.
\end{equation}

As in [13, §4.1], [1], we rewrite the first relation in (3.9) in a suitable for our purposes form. With \( \widetilde{E}_{ni}, i = 1, \ldots, q \), defined by

\begin{equation}
\sum_{j=1}^{q} a_{ij} \widetilde{E}_{nj} = E_{ni}, \quad i = 1, \ldots, q,
\end{equation}

it is easily seen that the first equation in (3.9) reads

\begin{equation}
u(t_{ni}) = u(t_n) - k \sum_{j=1}^{q} a_{ij} \left( Au(t_{nj}) - f(t_{nj}) - \widetilde{E}_{nj} \right), \quad i = 1, \ldots, q.
\end{equation}

Subtracting the first relation of (2.5) from (3.12), we obtain the error equations

\begin{equation}
e_{ni} = e_n - k \sum_{j=1}^{q} a_{ij} (Ae_{nj} - \widetilde{E}_{nj} - \rho_{nj}), \quad i = 1, \ldots, q,
\end{equation}

where

\begin{equation}
\rho_{ni} := f(t_{ni}) - f_{ni}, \quad i = 1, \ldots, q.
\end{equation}
Now, the discrete maximal parabolic regularity stability estimates of the \(q\)-stage Radau IIA method, Lemma 2.1, imply

\[
\|(\partial e_n)_{n=1}^N\|_{L^p(X)} + \|(Ae_n)_{n=1}^N\|_{L^p(X)} \leq C_{p,X} \sum_{i=1}^{q} \left( \|(\tilde{E}_{ni})_{n=0}^{N-1}\|_{L^p(X)} + \|\rho_{ni}\|_{n=0}^{N-1}\|_{L^p(X)} \right)
\]

and

\[
\sum_{i=1}^{q} \|(Ae_{ni})_{n=0}^{N-1}\|_{L^p(X)} \leq C_{p,X} \sum_{i=1}^{q} \left( \|(\tilde{E}_{ni})_{n=0}^{N-1}\|_{L^p(X)} + \|\rho_{ni}\|_{n=0}^{N-1}\|_{L^p(X)} \right),
\]

with a constant \(C_{p,X}\) (depending also on the specific method) independent of \(N\) and the time step \(k\).

The estimate

\[
\|\tilde{E}_{ni}\|_X \leq C k^q, \quad i = 1, \ldots, q,
\]

follows from (3.11) and the consistency estimate (3.10). Next, we prove that

\[
\|\rho_{ni}\|_X \leq C k^q, \quad i = 1, \ldots, q,
\]

and thus we conclude that the discontinuous Galerkin method satisfies the same a priori bounds as the Radau IIA method, i.e.,

\[
\|(\partial e_n)_{n=1}^N\|_{L^p(X)} + \|(Ae_n)_{n=1}^N\|_{L^p(X)} \leq \tilde{C}_{p,X,T} k^q
\]

and

\[
\|(Ae_{ni})_{n=0}^{N-1}\|_{L^p(X)} \leq \tilde{C}_{p,X,T} k^q, \quad i = 1, \ldots, q,
\]

respectively, with a constant \(\tilde{C}_{p,X,T}\) independent of \(N\) and the time step \(k\). It remains to prove the following lemma.

**Lemma 3.2.** Let \(\rho_{ni} = f(t_{ni}) - f_{ni}, i = 1, \ldots, q\), with \(f_{ni}\) given in (2.6). Then, (3.18) holds.

**Proof.** Let us define the functionals \(\rho_{ni}\),

\[
\rho_{ni}(g) = g(t_{ni}) - g_{ni} = g(t_{ni}) - \frac{1}{\int_{J_n} \ell_{ni}(s) \, ds} \int_{J_n} \ell_{ni}(s) g(s) \, ds, \quad i = 1, \ldots, q;
\]

cf. (2.6). Then, we observe that for any \(p \in \mathbb{P}(q-1)\) the exactness of the Radau quadrature rule implies

\[
\rho_{ni}(p) = p(t_{ni}) - p_{ni} = p(t_{ni}) - \frac{1}{b_{ij} k} \int_{J_n} \ell_{ni}(s) p(s) \, ds
\]

\[
= p(t_{ni}) - \frac{1}{b_{ij} k} k \sum_{j=1}^{q} b_{ij} \ell_{ni}(t_{nj}) p(t_{nj})
\]

\[
= p(t_{ni}) - \frac{1}{b_{ij} k} k b_{ij} \ell_{ni}(t_{ni}) p(t_{ni}) = 0, \quad i = 1, \ldots, q.
\]
Hence, for any \( p \in \mathcal{P}(q-1) \),
\[
\rho_{ni}(f) = \rho_{ni}(f - p) = (f - p)(t_{ni}) - \frac{1}{\int_{J_n} \ell_{ni}(s) \, ds} \int_{J_n} \ell_{ni}(s)(f - p)(s), \quad i = 1, \ldots, q.
\]
Therefore,
\[
|\rho_{ni}(f)| \leq 2 \| f - p \|_{L^\infty(J_n)}, \quad i = 1, \ldots, q,
\]
and the proof is complete.

**Remark 3.3 (The case \( q = 1 \)).** The case \( q = 1 \) is easy. The consistency error of the implicit Euler method is
\[
E_{n+1} := u(t_{n+1}) - u(t_n) + kAu(t_{n+1}) - kf(t_{n+1}),
\]
and the corresponding equation for the error \( e_m := u(t_m) - U_m \) for the discontinuous Galerkin method with piecewise constant elements reads
\[
e_{n+1} - e_n + kAe_{n+1} = E_{n+1} + k\rho_{n1},
\]
\( n = 0, \ldots, N - 1 \), with
\[
\rho_{n1} := f(t_{n+1}) - f_n = f(t_{n+1}) - \frac{1}{k} \int_{J_n} f(s) \, ds.
\]
Now, the discrete maximal parabolic regularity of the implicit Euler method, applied to the error equation (3.21), yields
\[
\|(\partial e_n)_{n=1}^{M}\|_{\ell^p(X)} + \|(Ae_n)_{n=1}^{M}\|_{\ell^p(X)} \leq C_{p,X}\|(\frac{1}{k}E_n + \rho_{n-1,1})_{n=1}^{M}\|_{\ell^p(X)},
\]
\( M = 1, \ldots, N \), with a constant \( C_{p,X} \) independent of \( M, T \) and the time step \( k \); see [5, Remark 5.2] and [12, Theorem 3.1]. Using in (3.22) the obvious estimates
\[
\|E_n\|_X \leq Ck^2, \quad \|\rho_{n1}\|_X \leq Ck,
\]
we arrive at the desired a priori error estimate
\[
\|(\partial e_n)_{n=1}^{N}\|_{\ell^p(X)} + \|(Ae_n)_{n=1}^{N}\|_{\ell^p(X)} \leq \tilde{C}_{p,X,T}k,
\]
which is (3.19) for \( q = 1 \).

**3.2.2. Optimality of the a posteriori error estimate (3.5).** We shall proceed as in [1].
We assume that the solution \( u \) is sufficiently smooth and we show that the estimator on the right-hand side of (3.5) is also of order \( q \).

The approximation properties of the \( L^2 \)-projection imply that the second term on the right-hand side of (3.8) is of optimal order \( O(k^q) \). It thus remains to show that \( \| A\hat{U}^{(q)} \|_X \) is bounded, uniformly in the time step \( k \). As in [1, Section 3.2.4], we see that \( e_n = u(t_n) - \hat{U}_n, e_0 = e_n \)
\[
\hat{U}(t) = -\sum_{i=0}^{q} \ell_{ni}(t)e_{ni} + \sum_{i=0}^{q} \ell_{ni}(t)u(t_{ni}), \quad t \in [t_n, t_{n+1}].
\]
Therefore,

\begin{equation}
\hat{U}^{(q)}(t) = -\sum_{i=0}^{q} \hat{\ell}^{(q)}_{ni} e_{ni} + \sum_{i=0}^{q} \hat{\ell}^{(q)}_{ni} u(t_{ni}), \quad t \in (t_n, t_{n+1}).
\end{equation}

We recall from [1, (3.23)] that

\begin{equation}
\|A \sum_{i=0}^{q} \hat{\ell}^{(q)}_{ni} u(t_{ni})\|_X \leq \hat{c}, \quad t \in (t_n, t_{n+1}).
\end{equation}

For the remaining part of the residual,

\begin{equation}
R_{dG}(t) := -k^q \Phi_q(t) \sum_{i=0}^{q} \hat{\ell}^{(q)}_{ni} A e_{ni},
\end{equation}

we obtain, as in [1, Section 3.2.4],

\[ \|R_{dG}\|_{L^p((0,T);X)} \leq c C_q \left( k \sum_{n=1}^{N} \|A e_{ni}\|_X^p \right)^{1/p} = c C_q \sum_{i=0}^{q} \|(A e_{ni})^N_{n=1}\|_{\ell^p(X)}. \]

Therefore, the a priori error estimates (3.19) and (3.20) imply

\begin{equation}
\|R_{dG}\|_{L^p((0,T);X)} \leq \tilde{C} k^q,
\end{equation}

with a constant \( \tilde{C} \) depending on \( T \), and the optimality of (3.5) follows. We have therefore proved the following.

**Theorem 3.4 (A posteriori error estimate).** Consider the discontinuous Galerkin approximations defined in (1.3). Let \( \hat{U} \in V_c^e(q) \) be the reconstruction of \( U \), i.e., the continuous piecewise polynomial function defined in (1.4), which coincides with \( U \) at the Radau nodes of each \( J_n \) and \( \hat{U}(t^+_n) = U(t_n) = \hat{U}(t_n) \). With \( u \) being the solution of (1.1), the following maximal regularity a posteriori error estimate holds

\begin{equation}
\|(u - \hat{U})^f\|_{L^p((0,T);X)} + \|A(u - \hat{U})\|_{L^p((0,T);X)} \leq c_{p,X} \|R\|_{L^p((0,T);X)},
\end{equation}

for \( 0 < t \leq T \), where the a posteriori estimator is given by

\begin{equation}
R(t) = \hat{U}'(t) + A \hat{U}(t) - f(t) = A(\hat{U}(t) - U(t)) - (I - P_{q-1}) f(t), \quad t \in (t_n, t_{n+1}].
\end{equation}

Furthermore, the estimator is of optimal asymptotic order of accuracy in the sense that, if \( u \) is sufficiently smooth, there exists a constant \( \tilde{C}_{p,X}(u) \) such that

\begin{equation}
\|R\|_{L^p((0,T);X)} \leq \tilde{C}_{p,X}(u) k^q.
\end{equation}

**Remark 3.5 (Variable time steps).** It is clear that the a posteriori error estimate (3.28), (3.29) holds as is for variable time steps as well. Furthermore, in the proof of (3.30) constant time steps are only assumed to guarantee the validity of Lemma 2.1, [12, Corollary 5.2, Theorem 5.1], i.e., the maximal regularity of Radau IIA methods. If one assumes that Lemma 2.1 holds for variable time steps as well, then (3.30) still holds true. Alternatively, by modifying some of the above arguments, using the a priori maximal regularity of discontinuous Galerkin methods established in [16, 17], it is possible to show that (3.30) holds also for variable time steps up to a logarithmic factor.
3.3. Extension to nonautonomous equations. Here, we briefly outline the extension of our results to nonautonomous parabolic equations,

\[
\begin{aligned}
    u'(t) + A(t)u(t) &= f(t), \quad 0 < t < T, \\
u(0) &= 0,
\end{aligned}
\]

in a Banach space \(X\). We assume that all operators \(A(t), t \in [0,T]\), share the same domain \(\mathcal{D}(A)\), \(A(t)\) is the generator of an analytic semigroup on \(X\) having maximal \(L^p\)-regularity, for every \(t \in [0,T]\), \(A(t)\) induce equivalent norms on \(\mathcal{D}(A)\), and \(A(t) : \mathcal{D}(A) \to X\) satisfies a suitable Lipschitz condition with respect to \(t\).

With starting value \(U(0) = U_0 = 0\), the discretization of the initial value problem (3.31) by the discontinuous Galerkin method \(dG(q-1)\) is to seek \(U \in V^h(q-1)\) such that

\[
\int_{J_n} \left( \langle U', v \rangle + \langle A(t)U, v \rangle \right) dt + \langle U_n^+ - U_n, v_n^+ \rangle = \int_{J_n} \langle f, v \rangle dt \quad \forall v \in P_X(q-1)
\]

for \(n = 0, \ldots, N-1\); cf. (1.3).

As before, one can use corresponding stability estimates for Radau IIA methods, see [1, (4.13), (4.14)], to treat the discontinuous Galerkin method. For this purpose, an analogue to Lemma 2.2 in the case of nonautonomous equations is needed. In fact, adopting the arguments of the proof of Lemma 2.2, we can prove the following: The discontinuous Galerkin approximations \(U_n = U(t_n), \ U_{nj} = U(t_{nj}), j = 1, \ldots, q,\) for the nonautonomous parabolic equation (3.31), given in (3.32), satisfy the modified Radau IIA method

\[
\begin{aligned}
    U_{ni} &= U_n - k \sum_{j=1}^{q} a_{ij}(A(t_{nj})U_{nj} - f_{nj} + \zeta_{nj}), \quad i = 1, \ldots, q, \\
    U_{n+1} &= U_n - k \sum_{i=1}^{q} b_i(A(t_n)U_{ni} - f_{ni} + \zeta_{ni}),
\end{aligned}
\]

\(n = 0, \ldots, N-1,\) with the averages \(f_{ni}\) given in (2.6) and

\[
\zeta_{ni} := \frac{1}{b_{i}k} \int_{J_n} \ell_{ni}(s)A(s)U(s) ds - A(t_n)U_{ni}, \quad i = 1, \ldots, q.
\]

Again, \(U_{n+1} = U_{nq}\). Furthermore, (3.33) written in terms of the reconstruction \(\hat{U}\) of \(U\) is a perturbed collocation method in each interval \(J_n\) with starting value \(\hat{U}_{n0} = U_n\), namely

\[
\hat{U}'(t_n) + A(t_n)\hat{U}_{ni} = f_{ni} - \zeta_{ni}, \quad i = 1, \ldots, q.
\]

Then, by adopting the preceding analysis in the nonautonomous case, and using arguments similar to the case of Radau IIA methods, see [1, §4.2], we can prove the following a priori and a posteriori bounds: Assume that the operator \(A(t)\) is the generator of an analytic semigroup on \(X\) having maximal \(L^p\)-regularity, for every \(t \in [0,T]\), and satisfies suitable Lipschitz conditions;
then, the discontinuous Galerkin approximations \( U_n = U(t_n), U_{ni} = U(t_{ni}), i = 1, \ldots, q, \) satisfy the maximal parabolic regularity stability estimates

\[
\begin{align*}
\|\partial U_n\|_{L^p(X)} + \|A(t_m)U_n\|_{L^p(X)} & \leq C_pX,T\|f\|_{L^p(0,t_m);X}, \\
\sum_{i=1}^{q} \|A(t_m)U_{ni}\|_{L^p(X)} & \leq C_pX,T\|f\|_{L^p(0,t_m);X},
\end{align*}
\]

(3.36) \( m = 1, \ldots, N. \) Furthermore, the optimal order a posteriori estimate

\[
\|e\|_{L^p((0,s);X)} + \|A(s)e\|_{L^p((0,s);X)} \leq c\|R\|_{L^p((0,s);X)}, \quad 0 < s \leq T,
\]

(3.37) with \( R \) the residual of the reconstruction \( \hat{U}, R(t) := \hat{U}'(t) + A(t)\hat{U}(t) - f(t), \) holds true.

We do not present the details. For similar ideas and results, we refer to \([17]\) and \([13, \S 3.6]\), \([1]\) for the discontinuous Galerkin method with piecewise constant elements and for Radau IIA methods, respectively.

REFERENCES

\[ \begin{align*}
[1] & \text{G. Akrivis and Ch. Makridakis, A posteriori error estimates for Radau IIA methods via maximal parabolic regularity, submitted for publication.} \\
\end{align*} \]


