Journal of Computational and Applied Mathematics 263 (2014) 45–58

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Strong-stability-preserving, Hermite–Birkhoff time-discretization based on k step methods and 8-stage explicit Runge–Kutta methods of order 5 and 4

Huong Nguyen-Thu [a], Truong Nguyen-Ba [b], Rémi Vaillancourt [b,*]

[a] Department of Mathematics, School of Education, Cantho University, Cantho, Vietnam

[b] Department of Mathematics and Statistics, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5

ARTICLE INFO

Article history:
Received 29 December 2012
Received in revised form 5 November 2013

MSC:
primary 65L06
secondary 65M20

Keywords:
Strong stability preserving
Hermite–Birkhoff method
SSP coefficient
Time discretization
Method of lines
Comparison with other SSP methods

ABSTRACT

Ruuth and Spiteri have shown, in 2002, that fifth-order strong-stability-preserving (SSP) explicit Runge–Kutta (RK) methods with nonnegative coefficients do not exist. One of the purposes of the present paper is to show that the Ruuth–Spiteri barrier can be broken by adding backsteps to RK methods. New optimal, 8-stage, explicit, SSP, Hermite–Birkhoff (HB) time discretizations of order p, $p = 5, 6, \ldots, 12$, with nonnegative coefficients are constructed by combining linear k-step methods of order $(p-4)$ with an 8-stage explicit RK method of order 5 (RK(8, 5)). These new SSP HB methods preserve the monotonicity property of the solution and prevent error growth; therefore, they are suitable for solving hyperbolic partial differential equations (PDEs) by the method of lines. Moreover, these new HB methods have larger effective SSP coefficients and larger maximum effective CFL numbers than Huang's hybrid methods and RK methods of the same order when applied to the inviscid Burgers equation. Generally, HB methods combined with RK(8, 5) have maximum stepsize 24% larger than HB combined with RK(8, 4).

1. Introduction

We are concerned with the numerical solution of initial value problems

$$\frac{dy}{dt} = f(t, y(t)), \qquad y(t_0) = y_0, \tag{1}$$

where the function $f : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$ is such that

$$\|y(t + \Delta t)\| \le \|y(t)\|, \tag{2}$$

for all $\Delta t \ge 0$, where $\|\cdot\|$ may be a norm, a semi-norm or, more generally, any convex functional. It is also assumed that f satisfies the discrete analog of (2),

$$\|y_n + \Delta t f(t_n, y_n)\| \le \|y_n\|, \qquad \Delta t \le \Delta t_{\mathrm{FE}}, \tag{3}$$

for the forward Euler (FE) method and y_n is a numerical approximation to $y(t_0 + n\Delta t)$. We are interested in higher-order accurate, multistep, Hermite–Birkhoff (HB) methods that preserve the monotonicity property

$$\|y_{n+1}\| \le \max_{0\le j\le k-1} \|y_{n-j}\|, \tag{4}$$

* Corresponding author. Tel.: +1 6135625800; fax: +1 6135625776.
E-mail addresses: nthuongctu@yahoo.com (H. Nguyen-Thu), trnguyen@uottawa.ca (T. Nguyen-Ba), remi@uottawa.ca (R. Vaillancourt).

0377-0427/$ – see front matter
http://dx.doi.org/10.1016/j.cam.2013.11.013

for $0 \le \Delta t \le \Delta t_{\max} = c\Delta t_{\mathrm{FE}}$ whenever condition (3) holds. Here k represents the number of previous steps used to compute the next solution value and c, called the strong stability preserving (SSP) coefficient, depends only on the numerical integration method but not on f. The monotonicity property (4) is suitable as it follows property (2) of the true solution and inhibits growth of errors.

SSP methods have been developed to satisfy the monotonicity property (4) for system (1) whenever condition (3) is fulfilled. Property (4) is guaranteed under the maximum time step $\Delta t_{\max} = c\Delta t_{\mathrm{FE}}$.

Substantial research efforts were dedicated to finding numerical methods with the highest c (see [1–3]). The main application of such monotonicity results are found in the numerical solution of hyperbolic PDEs, in particular, of conservation laws. For the one-dimensional equation

$$y_t + g(y)_x = 0, \qquad y(x, 0) = y_0(x), \tag{5}$$

the spatial derivative $g(y)_x$ can be approximated by a conservative finite difference or finite element at x_j, $j = 1, 2, \ldots, N$, (see [4–7]). This spatial semi-discretization leads to system (1) of ODEs.

It has been shown by Ruuth and Spiteri [8] that fifth-order SSP RK methods with nonnegative coefficients do not exist. One purpose of the present paper is to break the Ruuth–Spiteri barrier by adding backsteps to RK methods with nonnegative coefficients which exist.

To solve system (1), new optimal, 8-stage, explicit, strong-stability-preserving, Hermite–Birkhoff (SSP HB) time discretizations of order p, $p = 5, 6, \ldots, 12$, with nonnegative coefficients are constructed by combining linear k-step methods of order $(p-4)$ with an 8-stage explicit RK method of order 5 (RK(8, 5)).

All the HB(k, p) methods considered in the paper are based on the combination of a k-step method and an explicit 8-stage RK method. So the denomination "explicit 8-stage RK method" will be omitted.

These new methods are denoted HB(k, p), since HB interpolation polynomials enter in their construction as it is briefly sketched in Section 2 (see [9] for fuller developments).

The objective of such high-order methods is to maintain the monotonicity property (4) while achieving higher-order accuracy in time, perhaps with a modified time-step restriction, measured here with the SSP coefficient $c(\mathrm{HB}(k, p))$:

$$\Delta t \le c(\mathrm{HB}(k, p))\Delta t_{\mathrm{FE}}. \tag{6}$$

The SSP coefficient, historically called CFL coefficient, describes the ratio of the SSP time step to the strongly stable FE time step (see [1]). Since our arguments are based on convex decompositions of high-order methods in terms of FE, such high-order methods preserve SSP in any norm once FE is shown to be strongly stable.

It is found that the new 8-stage HB(k, p) are better than similar 7-, 9-, and 10-stage methods. Among our HB methods of order 11 and 12, the 8-stage HB(k, p) have an average SSP coefficient 24% higher than the average SSP coefficient of HB(k, p) which combine linear k-step methods and RK(8,4).

A brief review of the development of SSP methods will appear in Section 5 on the construction of HB(k, p).

The new HB(k, p) have larger effective SSP coefficients than known SSP hybrid methods HM(k, p) with the same k and p, especially when k is small. In particular, no counterparts of HB methods of order greater than 8 have been found in the literature among hybrid and general linear, multistep, multistage, methods.

The paper is organized as follows. Section 2 introduces notation and general formulae of 8-stage HB(k, p) which combine linear k-step methods and RK(8, 5). Order conditions are listed in Section 3. In Section 4, vector notation is used to describe the modified Butcher, modified Shu–Osher, and canonical Shu–Osher forms of our HB(k, p) as generalizations of similar forms of SSP RK(k, p). Section 5 presents the effective SSP coefficients of our HB(k, p) and compares them with other known SSP methods of the same order. The numerical verification of the order p of our methods is also shown in this section. In Section 6, numerical results are displayed for all our methods applied to Burgers' inviscid equation. Section 7 compares HB(k, p) obtained from combining k-step methods with RK(8, 5) and RK(8, 4) of order 5 and 4, respectively. Eight new SSP HB(s, p) with RK(8, 5) and nine new SSP HB(s, p) with RK(8, 4) are in Huong Nguyen-Thu's Ph.D. Thesis [10].

2. Notation and *k*-step SSP HB methods of order *p*

The following notation will be used.

Notation 1.
- HB(k, p) denotes 8-stage, k-step, SSP Hermite–Birkhoff methods of order p made of k-step linear methods and an 8-stage RK(8, 5) of order 5. In Section 7, when we compare HB(k, p) combined with RK(8, 5) and RK(8, 4), the extended notation $\mathrm{HB}_{\mathrm{RK5}}(k, p)$ and $\mathrm{HB}_{\mathrm{RK4}}(k, p)$ will be used.
- GL(k, p) denotes k-step, 4-stage, general linear methods of order p,
- HM(k, p) denotes Huang's k-step, SSP, hybrid methods of order p (see [3]),
- RK(s, p) denotes s-stage, SSP, explicit RK methods of order p,
- TSRK(s, p) denotes two-step, s-stage, SSP, explicit RK methods of order p.

All methods considered in this work are SSP and explicit, so the denomination "SSP" and "explicit" will generally be omitted. All our SSP HB methods have 8 stages; therefore, "8-stage" will not be repeated.

Notation 2. • The abscissa vector $\sigma = [c_1, c_2, c_3, \dots, c_8]^T$, $0 \le c_j \le 1$, defines the off-step points $t_n + c_j\Delta t$. For simplicity, set $c_1 = 0$ and c_1 raised to the zero power is 1, $c_1^0 = 1$, by convention.

- At each off-step point, let $F_j := f(t_n + c_j\Delta t, Y_j)$, $j = 2, 3, \dots, 8$, be the jth-stage derivative where Y_j is the jth-stage value and set $Y_1 = y_n$.

Definition 1. To perform an integration from t_n to t_{n+1}, an HB(k, p) method is defined by the following eight formulae:

Seven HB polynomials of degree $(2k + i - 3)$ are used as predictors to obtain the stage values Y_i,

$$Y_i = v_{\mathfrak{B},i} y_n + \sum_{j=1}^{k-1} A_{\mathfrak{B},ij} y_{n-j} + \Delta t \left[\sum_{j=1}^{i-1} a_{ij} F_j + \sum_{j=1}^{k-1} B_{\mathfrak{B},ij} f_{n-j} \right], \quad i = 2, 3, \dots, 8, \tag{7}$$

and an HB polynomial of degree $(2k + 6)$ is used as an integration formula to obtain y_{n+1} to order p,

$$y_{n+1} = v_{\mathfrak{B},9} y_n + \sum_{j=1}^{k-1} A_{\mathfrak{B},9,j} y_{n-j} + \Delta t \left[\sum_{j=1}^{8} b_j F_j + \sum_{j=1}^{k-1} B_{\mathfrak{B},9,j} f_{n-j} \right]. \tag{8}$$

In formulae (7) and (8), the constant coefficients $v_{\mathfrak{B},i}$, $A_{\mathfrak{B},ij}$, $B_{\mathfrak{B},ij}$, a_{ij} and b_j for $i = 2, 3, \dots, 9$ and $j = 1, 2, \dots, k - 1$ need to be constructed to approximate y_{n+1} to the solution $y(t_{n+1}) = y(t_n + \Delta t)$.

Here, the subscript $\mathfrak{B}$ refers to the Butcher form, while, later, the subscript SO will be used for the Shu–Osher form.

3. Order conditions for HB(k, p)

To construct the order conditions for HB(k, p), first we let:

$$\mathcal{B}_i(j) = \sum_{l=1}^{k-1} A_{\mathfrak{B},il} \frac{(-l)^j}{j!} + \sum_{l=1}^{k-1} B_{\mathfrak{B},il} \frac{(-l)^{j-1}}{(j-1)!}, \quad \begin{cases} i = 2, 3, \dots, 8, \\ j = 1, 2, \dots, p, \end{cases} \tag{9}$$

which come from the backsteps of the methods.

Expanding the numerical solution produced by formulae (7)–(8) to agree with a Taylor expansion of the true solution, we obtain the following multistep- and RK-type order conditions that must be satisfied by HB(k, p).

First, for consistency, we need to satisfy the multistep-type order conditions:

$$v_{\mathfrak{B},i} + \sum_{j=1}^{k-1} A_{\mathfrak{B},ij} = 1, \quad i = 2, 3, \dots, 9. \tag{10}$$

Second, for HB methods of order p, we impose the following $(p - 4)$ simplifying assumptions on the coefficient a_{ij}, backstep $\mathcal{B}_i(\cdot)$ and abscissa vector σ:

$$\sum_{j=1}^{i-1} a_{ij} c_j^m + m!\mathcal{B}_i(m+1) = \frac{1}{m+1} c_i^{m+1}, \quad \begin{cases} i = 2, 3, \dots, 8, \\ m = 0, 1, \dots, p-5. \end{cases} \tag{11}$$

These assumptions help reduce the large number of RK-type order conditions (see [9]) to the following 12 conditions (12)–(23), in the case $p > 5$,

$$\sum_{i=1}^{8} b_i c_i^m + m!\mathcal{B}(m+1) = \frac{1}{m+1}, \quad m = 0, 1, \dots, p-1, \tag{12}$$

$$\sum_{i=2}^{8} b_i \left[\sum_{j=1}^{i-1} a_{ij} \frac{c_j^{p-4}}{(p-4)!} + \mathcal{B}_i(p-3) \right] + \mathcal{B}(p-2) = \frac{1}{(p-2)!}, \tag{13}$$

$$\sum_{i=2}^{8} b_i \frac{c_i}{p-2} \left[\sum_{j=1}^{i-1} a_{ij} \frac{c_j^{p-4}}{(p-4)!} + \mathcal{B}_i(p-3) \right] + \mathcal{B}(p-1) = \frac{1}{(p-1)!}, \tag{14}$$

$$\sum_{i=2}^{8} b_i \left[\sum_{j=1}^{i-1} a_{ij} \frac{c_j^{p-3}}{(p-3)!} + \mathcal{B}_i(p-2) \right] + \mathcal{B}(p-1) = \frac{1}{(p-1)!}, \tag{15}$$

$$\sum_{i=2}^{8} b_i \left[\sum_{j=1}^{i-1} a_{ij} \left[\sum_{k=1}^{j-1} a_{jk} \frac{c_k^{p-4}}{(p-4)!} + \mathcal{B}_j(p-3) \right] + \mathcal{B}_i(p-2) \right] + \mathcal{B}(p-1) = \frac{1}{(p-1)!}, \tag{16}$$

$$\sum_{i=2}^{8} b_i \frac{c_i^2}{(p-2)(p-1)} \left[\sum_{j=1}^{i-1} a_{ij} \frac{c_j^{p-4}}{(p-4)!} + \mathcal{B}_i(p-3) \right] + \mathcal{B}(p) = \frac{1}{p!}, \tag{17}$$

$$\sum_{i=2}^{8} b_i \frac{c_i}{p-1} \left[\sum_{j=1}^{i-1} a_{ij} \frac{c_j^{p-3}}{(p-3)!} + \mathcal{B}_i(p-2) \right] + \mathcal{B}(p) = \frac{1}{p!}, \tag{18}$$

$$\sum_{i=2}^{8} b_i \frac{c_i}{p-1} \left[\sum_{j=1}^{i-1} a_{ij} \left[\sum_{k=1}^{j-1} a_{jk} \frac{c_k^{p-4}}{(p-4)!} + \mathcal{B}_j(p-3) \right] + \mathcal{B}_i(p-2) \right] + \mathcal{B}(p) = \frac{1}{p!}, \tag{19}$$

$$\sum_{i=2}^{8} b_i \left[\sum_{j=1}^{i-1} a_{ij} \frac{c_j^{p-2}}{(p-2)!} + \mathcal{B}_i(p-1) \right] + \mathcal{B}(p) = \frac{1}{p!}, \tag{20}$$

$$\sum_{i=2}^{8} b_i \left[\sum_{j=1}^{i-1} a_{ij} \frac{c_j}{p-2} \left[\sum_{k=1}^{j-1} a_{jk} \frac{c_k^{p-4}}{(p-4)!} + \mathcal{B}_j(p-3) \right] + \mathcal{B}_i(p-1) \right] + \mathcal{B}(p) = \frac{1}{p!}, \tag{21}$$

$$\sum_{i=2}^{8} b_i \left[\sum_{j=1}^{i-1} a_{ij} \left[\sum_{k=1}^{j-1} a_{jk} \frac{c_k^{p-3}}{(p-3)!} + \mathcal{B}_j(p-2) \right] + \mathcal{B}_i(p-1) \right] + \mathcal{B}(p) = \frac{1}{p!}, \tag{22}$$

$$\sum_{i=2}^{8} b_i \left\{ \sum_{j=1}^{i-1} a_{ij} \left[\sum_{k=1}^{j-1} a_{jk} \left[\sum_{\ell=1}^{k-1} a_{k\ell} \frac{c_\ell^{p-4}}{(p-4)!} + \mathcal{B}_k(p-3) \right] + \mathcal{B}_j(p-2) \right] + \mathcal{B}_i(p-1) \right\} + \mathcal{B}(p) = \frac{1}{p!}, \tag{23}$$

where the backstep parts, $\mathcal{B}(j)$, are defined by

$$\mathcal{B}(j) = \sum_{i=1}^{k-1} A_{\mathfrak{B},9,i} \frac{(-i)^j}{j!} + \sum_{i=1}^{k-1} B_{\mathfrak{B},9,i} \frac{(-i)^{j-1}}{(j-1)!}, \quad j = 1, \dots, p+1. \tag{24}$$

These order conditions are simply RK order conditions with backstep parts $\mathcal{B}_i(\cdot)$ and $\mathcal{B}(\cdot)$.

In the case $p = 5$, HB(k, p) has also to satisfy the additional condition:

$$\sum_{i=2}^{8} \frac{b_i}{6} \left[\sum_{j=1}^{i-1} a_{ij} c_j + \mathcal{B}_i(2) \right]^2 + \mathcal{B}(5) = \frac{1}{5!}, \tag{25}$$

besides order conditions (12)–(23). Without the backstep parts, these order conditions reduce to the order conditions of an 8-stage RK5.

4. Canonical Shu–Osher form of HB(k, p)

Gottlieb, Ketcheson and Shu presented canonical Shu–Osher forms in compact vector notation for RK methods (see [2, Sections 3.1–3.4] for details). Here, our construction proceeds in three steps in Sections 4.1–4.3.

4.1. Modified Shu–Osher form of HB(k, p)

The Butcher form (7)–(8) is transformed into the original Shu–Osher form [11] for HB(k, p) (see [12]) as follows:

$$\begin{aligned} Y_i &= \sum_{j=1}^{k-1} \left[A_{ij} y_{n-j} + \Delta t B_{ij} f_{n-j} \right] + \sum_{j=1}^{i-1} \left[\alpha_{ij} Y_j + \Delta t\, \beta_{ij} F_j \right], \quad i = 2, 3, \dots, 9, \\ y_{n+1} &= Y_9, \end{aligned} \tag{26}$$

where consistency requires that

$$\sum_{j=1}^{k-1} A_{ij} + \sum_{j=1}^{i-1} \alpha_{ij} = 1, \quad i = 2, 3, \dots, 9. \tag{27}$$

Definition 2. The *modified Shu–Osher form* of HB(k, p) is:

$$\begin{aligned} Y_i &= \left[v_i y_n + \Delta t\, w_i f_n\right] + \sum_{j=1}^{k-1}\left[A_{ij} y_{n-j} + \Delta t B_{ij} f_{n-j}\right] + \sum_{j=2}^{i-1}\left[\alpha_{ij} Y_j + \Delta t\, \beta_{ij} F_j\right], \quad i = 2, 3, \dots, 9, \\ y_{n+1} &= Y_9, \end{aligned} \tag{28}$$

which is obtained from formulae (26) by letting $v_i = \alpha_{i1}$ and $w_i = \beta_{i1}$, $i = 2, 3, \dots, 9$.

The consistency condition (27) becomes

$$v_i + \sum_{j=1}^{k-1} A_{ij} + \sum_{j=2}^{i-1} \alpha_{ij} = 1, \quad i = 2, 3, \dots, 9. \tag{29}$$

4.2. Vector notation

Vector and matrix notation will help represent HB(k, p) in canonical Shu–Osher form. Consider two 9-vectors

$$\boldsymbol{v} = [0, v_2, v_3, \dots, v_9]^T, \qquad \boldsymbol{w} = [0, w_2, w_3, \dots, w_9]^T,$$

two strictly lower triangular 9×9 matrices,

$$\boldsymbol{\alpha} = (\alpha_{ij}), \qquad \boldsymbol{\beta} = (\beta_{ij}),$$

and two $9 \times (k-1)$ rectangular matrices with zero first row,

$$\boldsymbol{A}_{\mathrm{SO}} = (A_{ij}), \qquad \boldsymbol{B}_{\mathrm{SO}} = (B_{ij}),$$

where the numbers $\alpha_{ij}, \beta_{ij}, A_{ij}, B_{ij}$ come from formulae (26). Moreover, the matrices $\boldsymbol{Y}, \boldsymbol{F} \in \mathbb{R}^{9\times N}$, $\boldsymbol{y}_{\mathrm{back}} \in \mathbb{R}^{(k-1)\times N}$ and $\boldsymbol{f}_{\mathrm{back}} \in \mathbb{R}^{(k-1)\times N}$ are:

$$\boldsymbol{Y} = [0, Y_2, \dots, Y_9]^T, \qquad \boldsymbol{F} = [0, F_2, \dots, F_9]^T,$$

$$\boldsymbol{y}_{\mathrm{back}} = [y_{n-1}, y_{n-2}, \dots, y_{n-(k-1)}]^T, \qquad \boldsymbol{f}_{\mathrm{back}} = [f_{n-1}, f_{n-2}, \dots, f_{n-(k-1)}]^T,$$

with the following N-vectors: Y_j, F_j for $j = 1, 2, \dots, 9$, y_j, f_j for $j = n-(k-1), \dots, n$, $Y_1 = y_n$, $F_1 = f_n$, $Y_9 = y_{n+1}$ and $F_9 = f_{n+1}$. Thus, the modified Shu–Osher form of HB(k, p) (28) in vector notation is:

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{v} y_n^T + \boldsymbol{\alpha} \boldsymbol{Y} + \boldsymbol{A}_{\mathrm{SO}} \boldsymbol{y}_{\mathrm{back}} + \Delta t \left(\boldsymbol{w} f_n^T + \boldsymbol{\beta} \boldsymbol{F} + \boldsymbol{B}_{\mathrm{SO}} \boldsymbol{f}_{\mathrm{back}}\right), \\ y_{n+1} &= Y_9. \end{aligned} \tag{30}$$

Here the consistency condition (29) becomes

$$\boldsymbol{v} + \boldsymbol{\alpha} \boldsymbol{e}_9 + \boldsymbol{A}_{\mathrm{SO}} \boldsymbol{e}_{\mathrm{back}} = \boldsymbol{e}_9, \tag{31}$$

where the 9- and $(k-1)$-vectors $\boldsymbol{e}_9$ and $\boldsymbol{e}_{\mathrm{back}}$ are

$$\boldsymbol{e}_9 = [0, 1, 1, \dots, 1]^T \in \mathbb{R}^9, \qquad \boldsymbol{e}_{\mathrm{back}} = [1, 1, \dots, 1]^T \in \mathbb{R}^{(k-1)}, \tag{32}$$

respectively.

Definition 3. The *modified Butcher form* of HB(k, p) in vector notation is:

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{v}_{\mathfrak{B}} y_n^T + \boldsymbol{A}_{\mathfrak{B}} \boldsymbol{y}_{\mathrm{back}} + \Delta t \left(\boldsymbol{w}_{\mathfrak{B}} f_n^T + \boldsymbol{\beta}_{\mathfrak{B}} \boldsymbol{F} + \boldsymbol{B}_{\mathfrak{B}} \boldsymbol{f}_{\mathrm{back}}\right), \\ y_{n+1} &= Y_9, \end{aligned} \tag{33}$$

where

$$\boldsymbol{v}_{\mathfrak{B}} = (\boldsymbol{I} - \boldsymbol{\alpha})^{-1} \boldsymbol{v}, \qquad \boldsymbol{w}_{\mathfrak{B}} = (\boldsymbol{I} - \boldsymbol{\alpha})^{-1} \boldsymbol{w}, \qquad \boldsymbol{A}_{\mathfrak{B}} = (\boldsymbol{I} - \boldsymbol{\alpha})^{-1} \boldsymbol{A}_{\mathrm{SO}}, \tag{34}$$

$$\boldsymbol{\beta}_{\mathfrak{B}} = (\boldsymbol{I} - \boldsymbol{\alpha})^{-1} \boldsymbol{\beta}, \qquad \boldsymbol{B}_{\mathfrak{B}} = (\boldsymbol{I} - \boldsymbol{\alpha})^{-1} \boldsymbol{B}_{\mathrm{SO}}, \tag{35}$$

and the consistency condition (31) reduces to

$$\boldsymbol{v}_{\mathfrak{B}} + \boldsymbol{A}_{\mathfrak{B}} \boldsymbol{e}_{\mathrm{back}} = \boldsymbol{e}_9. \tag{36}$$

The construction of relations (34)–(35) is in [13].

4.3. Canonical Shu–Osher form of HB(k, p)

To maximize $c(\mathrm{HB}(k, p))$, we consider a particular modified Shu–Osher form (30) where the elements of matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ satisfy the ratio

$$r = \frac{\alpha_{ij}}{\beta_{ij}}, \quad \forall i, j,\ i = 2, 3, 4, \dots, 9;\ j = 1, 2, 3, \dots, i-1,\ \text{such that } \beta_{ij} \neq 0,$$

or, in vector notation,

$$\boldsymbol{\alpha}_r = r\boldsymbol{\beta}_r. \tag{37}$$

Definition 4. The *canonical Shu–Osher form* of HB(k, p) is:

$$\boldsymbol{Y} = \left(\boldsymbol{v}_r y_n^T + \Delta t \boldsymbol{w}_r f_n^T\right) + \left(\boldsymbol{\alpha}_r \boldsymbol{Y} + \Delta t \boldsymbol{\beta}_r \boldsymbol{F}\right) + \left(\boldsymbol{A}_{\mathrm{SO},r} \boldsymbol{y}_{\mathrm{back}} + \Delta t \boldsymbol{B}_{\mathrm{SO},r} \boldsymbol{f}_{\mathrm{back}}\right), \tag{38}$$

where the coefficients are determined by the relations

$$\boldsymbol{v}_r = (\boldsymbol{I} - \boldsymbol{\alpha}_r)\, \boldsymbol{v}_{\mathfrak{B}} = \left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{v}_{\mathfrak{B}}, \tag{39}$$

$$\boldsymbol{w}_r = (\boldsymbol{I} - \boldsymbol{\alpha}_r)\, \boldsymbol{w}_{\mathfrak{B}} = \left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{w}_{\mathfrak{B}}, \tag{40}$$

$$\boldsymbol{\alpha}_r = r\boldsymbol{\beta}_{\mathfrak{B}} (\boldsymbol{I} - \boldsymbol{\alpha}_r) = r\left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{\beta}_{\mathfrak{B}}, \tag{41}$$

$$\boldsymbol{\beta}_r = \boldsymbol{\beta}_{\mathfrak{B}} (\boldsymbol{I} - \boldsymbol{\alpha}_r) = \left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{\beta}_{\mathfrak{B}}, \tag{42}$$

$$\boldsymbol{A}_{\mathrm{SO},r} = (\boldsymbol{I} - \boldsymbol{\alpha}_r)\, \boldsymbol{A}_{\mathfrak{B}} = \left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{A}_{\mathfrak{B}}, \tag{43}$$

$$\boldsymbol{B}_{\mathrm{SO},r} = (\boldsymbol{I} - \boldsymbol{\alpha}_r)\, \boldsymbol{B}_{\mathfrak{B}} = \left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{B}_{\mathfrak{B}}, \tag{44}$$

with the consistency condition

$$\boldsymbol{v}_r + \boldsymbol{\alpha}_r \boldsymbol{e}_9 + \boldsymbol{A}_{\mathrm{SO},r} \boldsymbol{e}_{\mathrm{back}} = \boldsymbol{e}_9. \tag{45}$$

See [13] for the relations (39)–(44). Like optimal SSP RK methods [14], the new sparse canonical Shu–Osher forms of HB(k, p) might allow for reduced-storage implementation.

The ratio $r = \frac{\alpha_{ij}}{\beta_{ij}}$ for i, j, $i = 3, 4, \ldots, 9$ and $j = 2, 3, \ldots, i-1$, becomes a *feasible SSP coefficient* of HB(k, p). Hence, this ratio r must satisfy two additional sets of conditions:

$$r \le \frac{v_i}{w_i}, \quad i = 2, 3, \ldots, 9,$$

and

$$r \le \frac{A_{ij}}{B_{ij}}, \quad \begin{cases} j = 1, 2, \ldots, k-1, \\ i = 2, 3, \ldots, 9. \end{cases}$$

Therefore, the following result is an extension of the corresponding result presented in [15,3].

Theorem 1. *If f satisfies the forward Euler condition (3), then the 8-stage* HB(k, p) (38) *satisfies the monotonicity property*

$$\|y_{n+1}\| \le \max_{0 \le j \le k-1} \|y_{n-j}\|$$

provided

$$\Delta t \le c(\boldsymbol{v}_r, \boldsymbol{w}_r, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_r, \boldsymbol{A}_{\mathrm{SO},r}, \boldsymbol{B}_{\mathrm{SO},r}) \Delta t_{\mathrm{FE}},$$

where $c(\boldsymbol{v}_r, \boldsymbol{w}_r, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_r, \boldsymbol{A}_{\mathrm{SO},r}, \boldsymbol{B}_{\mathrm{SO},r})$ is equal to

$$r = \left\{ \frac{\alpha_{ij}}{\beta_{ij}} \right\}, \quad \begin{cases} i = 3, 4, \ldots, 9, \\ j = 2, 3, \ldots, i-1, \end{cases} \tag{46}$$

and less than or equal to:

$$\min_{i=2,3,\ldots,9} \frac{v_i}{w_i}, \tag{47}$$

$$\min_{j=1,2,\ldots,k-1} \left\{ \frac{A_{ij}}{B_{ij}} \right\}, \quad i = 2, 3, \ldots, 9, \tag{48}$$

with the convention that $a/0 = +\infty$, under the assumption that all coefficients of (38) *are nonnegative.*

4.4. Optimizing c of HB(k, p)

To optimize c of HB(k, p) in canonical form (38), following Theorem 1, we maximize

$$c(\mathrm{HB}(k, p)) = \max_{\boldsymbol{v}_r, \boldsymbol{w}_r, \boldsymbol{\alpha}_r, \boldsymbol{\beta}_r, \boldsymbol{A}_{\mathrm{SO},r}, \boldsymbol{B}_{\mathrm{SO},r}} c(\boldsymbol{v}_r, \boldsymbol{w}_r, \boldsymbol{\alpha}_r, \boldsymbol{A}_{\mathrm{SO},r}, \boldsymbol{B}_{\mathrm{SO},r}).$$

In the optimization formulation with any feasible initial data, the ratio r becomes the variable r which satisfies the system of nonlinear equations in the variables $\alpha_{ij}, r, \beta_{ij}$,

$$\alpha_{ij} - r\beta_{ij} = 0, \quad i = 3, 4, \dots, 9,\ j = 2, 3, \dots, i-1,$$

together with conditions (47) and (48).

Hence, the problem of optimizing the canonical HB(k, p) can be formulated as

$$c(\mathrm{HB}(k,p)) = \max_{\boldsymbol{v}_{\mathfrak{B}}, \boldsymbol{w}_{\mathfrak{B}}, \boldsymbol{\beta}_{\mathfrak{B}}, \boldsymbol{A}_{\mathfrak{B}}, \boldsymbol{B}_{\mathfrak{B}}} r, \tag{49}$$

subject to the component-wise inequalities:

$$\left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{v}_{\mathfrak{B}} \geq 0, \tag{50}$$

$$\left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{w}_{\mathfrak{B}} \geq 0, \tag{51}$$

$$\boldsymbol{\beta}_{\mathfrak{B}} \left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \geq 0, \tag{52}$$

$$\left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{A}_{\mathfrak{B}} \geq 0, \tag{53}$$

$$\left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{B}_{\mathfrak{B}} \geq 0, \tag{54}$$

$$r\boldsymbol{\beta}_{\mathfrak{B}} \left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{e}_9 + \left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \boldsymbol{A}_{\mathfrak{B}} \boldsymbol{e}_{\text{back}} \leq \boldsymbol{e}_9, \tag{55}$$

$$\left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \left(-\boldsymbol{v}_{\mathfrak{B}} + r\boldsymbol{w}_{\mathfrak{B}}\right) \leq 0, \tag{56}$$

$$\left(\boldsymbol{I} + r\boldsymbol{\beta}_{\mathfrak{B}}\right)^{-1} \left(-\boldsymbol{A}_{\mathfrak{B}} + r\boldsymbol{B}_{\mathfrak{B}}\right) \leq 0, \tag{57}$$

together with the order conditions (12)–(23).

5. Construction of HB($\boldsymbol{k}$, $\boldsymbol{p}$)

In this section we obtain the SSP coefficient c and the *effective SSP coefficient* c_{eff} of HB(k, p) in canonical Shu–Osher form.

Since HB(k, p) contain many free parameters when k is sufficiently large, the MATLAB Optimization Toolbox has been used to search for the methods with largest $c(\mathrm{HB}(k,p))$ for different values of k. Several authors [16,17,2] have found optimal RK and hybrid methods by this technique. In this work, the "fmincon" function was used to tolerance 10^{-12} on the objective function $c(\mathrm{HB}(k,p))$ provided all constraints were satisfied to tolerance 10^{-14}.

It is not guaranteed that the obtained results are global because of the limitation of "fmincon" function. However, the obtained optimal HB(k, p) methods with nonnegative coefficients give really good SSP coefficients and show fairly good efficiency gain over well known methods. It is noted that all obtained c_{eff} are checked multiple times with different initial input solutions.

With time step limited by the monotonicity property (6), the computational efficiency of a method can be measured by its effective SSP coefficients, which provides a fair comparison between two methods of the same order.

Definition 5 (*See [18]*)**.** The *effective SSP coefficients* of an SSP method M is

$$c_{\text{eff}}(M) = \frac{c(M)}{\ell}, \tag{58}$$

where ℓ is the number of function evaluations used by M per time step and $c(M)$ is its SSP coefficient.

In this paper, ℓ is the number of stages of HB(k, p) or RK(s, p), and $\ell = 2$ for HM(k, p). By definition, $c_{\text{eff}}(\mathrm{FE}) = 1$.

Definition 6 (*See [16]*)**.** The *percentage efficiency gain* (PEG) of $c_{\text{eff}}(\mathrm{M2})$ of method 2 over $c_{\text{eff}}(\mathrm{M1})$ of method 1 is

$$\mathrm{PEG}(c_{\text{eff}}(\mathrm{M2}), c_{\text{eff}}(\mathrm{M1})) = \frac{c_{\text{eff}}(\mathrm{M2}) - c_{\text{eff}}(\mathrm{M1})}{c_{\text{eff}}(\mathrm{M1})}. \tag{59}$$

In the following subsections, we compare effective SSP coefficients of our HB methods with those of other methods such as RK(s, p), HM(k, p) and TSRK(s, p). Although our 2-step s-stage HB methods and TSRK(s, p) methods have the same step, the same stage and the same order, the two sets of conditions satisfied by the coefficients of 2-step s-stage HB methods and TSRK(s, p) methods are generally different. Moreover they have generally different formulae. Therefore, they will give generally different SSP coefficients.

5.1. Fifth-order methods

Ruuth and Spiteri [8] proved that there are no 5th-order SSP RK methods with nonnegative coefficients; they constructed methods of order five: RK(9, 5) and RK(10, 5) with negative coefficients in [18,19]. Ruuth and Hundsdorfer [20] pointed

Table 1
$c_{\text{eff}}(\text{HB}(k,p))$ as function of step number k and order p.

p	k						
	2	3	4	5	6	7	8
5	0.447						
6	0.241	0.328	0.341	0.345	0.347		
7	*0.040	0.248	0.284	0.285			
8		0.142	0.198	0.235	0.241	0.243	
9		*0.035	0.138	0.179	0.203	0.216	0.218
10			*0.043	0.121	0.156	0.182	0.186
11				*0.060	0.106	0.135	0.156
12					*0.025	0.100	0.116

* The superscript '∗' attached to the methods indicates these methods are not tested.

out that 5th-order linear multistep methods with nonnegative coefficients require at least $k = 7$ steps. Huang [3] found HM$(k, 5)$ with $k = 4, 5, 6, 7$. It is seen in Table 2 that HB of order five are better than RK methods, typically $\text{PEG}(c_{\text{eff}}(\text{HB}(2,5)), c_{\text{eff}}(\text{RK}(9,5))) = 49\%$ and $\text{PEG}(c_{\text{eff}}(\text{HB}(2,5)), c_{\text{eff}}(\text{RK}(10,5))) = 32\%$. TSRK$(s, 5)$ methods with nonnegative coefficients are found in [21] with the best $c_{\text{eff}}(\text{TSRK}(8,5)) = 0.447$. Our best method of order 5 is HB(2, 5) with $c(\text{HB}(2,5)) = 3.579$ and $c_{\text{eff}}(\text{HB}(2,5)) = 0.447$.

5.2. Sixth-order methods

Ketcheson [22] pointed out that LM$(k, 6)$ of order 6 with nonnegative coefficients requires at least $k = 10$ steps and $c_{\text{eff}}(\text{LM}(10,6)) = 0.052$. Huang [3] introduced HM$(k, 6)$ with $k = 5, 6, 7$ with largest $c_{\text{eff}}(\text{HM}(7,6)) = 0.220$.

Ketcheson, Gottlieb and Macdonald [21] found a 2-step 8-stage RK method of order 6 with $c_{\text{eff}}(\text{TSRK}(8,6)) = 0.242$. Our HB(2, 6) has similar $c_{\text{eff}}(\text{HB}(2,6)) = 0.241$. With $k = 3$, HB(3, 6) has larger $c_{\text{eff}}(\text{HB}(3,6)) = 0.328$ and, from data in Table 2, formula (59) gives $\text{PEG}(c_{\text{eff}}(\text{HB}(3,6)), c_{\text{eff}}(\text{TSRK}(8,6))) = 35.5\%$.

As seen in Table 1, our best HB$(k, 6)$ are with $k = 2, 3, \ldots, 6$; in fact, HB(6,6) has largest $c_{\text{eff}}(\text{HB}(6,6)) = 0.347$. Thus increasing k increases $c_{\text{eff}}(\text{HB}(k,6))$.

5.3. Higher-order methods

Table 1 lists $c_{\text{eff}}(\text{HB}(k,p))$ as function of step number k and order p. We see that c_{eff} increases as k increases and decreases as p increases. Only the fairly good unstarred methods are listed in Table 2.

Table 2 lists $c(\text{HB}(k,p))$, $c_{\text{eff}}(\text{HB}(k,p))$, and $c(\text{OM}(k,p))$, $c_{\text{eff}}(\text{OM}(k,p))$ for other methods on hand. Since $\text{PEG}(c_{\text{eff}}(\text{HB}(k,p)), c_{\text{eff}}(\text{OM}(k,p))) \geq 0$, it follows that our methods are better than the other methods on hand.

The superscript 'a' attached to eight HB(k, p) listed in Table 2 indicates methods with fairly good $c_{\text{eff}}(\text{HB}(k,p))$ and low step number k for reduced storage implementation. These methods can be found in [10, pp. 132–139] in their canonical Shu–Osher form with their $c(\text{HB}(k,p))$, $c_{\text{eff}}(\text{HB}(k,p))$ and abscissa vector σ.

Ketcheson, Gottlieb and Macdonald [21] found an 8-stage TSRK method of order 7 with $c_{\text{eff}}(\text{TSRK}(8,7)) = 0.071$. Table 1 shows that $c_{\text{eff}}(\text{HB}(2,7)) = 0.040$. However, increasing the step number to $k = 3, 4, 5$, we find HB$(k, 7)$ with much larger SSP coefficients than TSRK(8, 7) and formula (58) gives $\text{PEG}(c_{\text{eff}}(\text{HB}(3,7)), c_{\text{eff}}(\text{TSRK}(8,7))) = 249\%$. The c_{eff} of these methods are listed in Table 1 with numerically largest $c_{\text{eff}}(\text{HB}(5,7)) = 0.285$.

It is not mentioned in [21] that 8–10-stage TSRK methods of order 8 exist. However, we found HB(3, 8) with $c_{\text{eff}}(\text{HB}(3,8)) = 0.160$ and, by formula (59), $\text{PEG}(c_{\text{eff}}(\text{HB}(3,8)), c_{\text{eff}}(\text{TSRK}(12,8))) = 105\%$ even if our 8-stage HB(3, 8) has fewer stages than the 12-stage TSRK(12, 8).

We have not found, in the literature, general linear SSP methods of order 9–12 with nonnegative coefficients. However, Tables 1 and 2 show that HB(k, p) of order $p \geq 9$ with nonnegative coefficients exist.

5.4. Numerical verification of the order p of HB(k, p)

To show the relevance of the theoretical order of HB(k, p) when solving ODEs, we have applied these methods with various constant stepsizes, h, on the following initial value problem over $t_n \in [0, \pi + 8]$, with exact solution $y_i(t)$:

$$\begin{aligned}
&y_1' = -y_1, \qquad y_1(0) = 1, \qquad y_1(t) = e^{-t},\\
&y_2' = y_3, \qquad y_2(0) = 0, \qquad y_2(t) = \sin t,\\
&y_3' = -y_2, \qquad y_3(0) = 1, \qquad y_3(t) = \cos t,\\
&y_4' = 1, \qquad y_4(0) = 0, \qquad y_4(t) = t,\\
&y_5' = -y_1 + (y_2 + y_4 y_3), \qquad y_5(0) = 1, \qquad y_5(t) = e^{-t} + t\sin t.
\end{aligned} \tag{60}$$

Table 2
Row-wise PEG(c_{eff}(HB(k, p)), c_{eff}(OM(k, p))).

HB(k, p)	c(HB(k, p))	c_{eff}(HB(k, p))	OM(k, p)	c(OM(k, p))	c_{eff}(OM(k, p))	PEG
[a]HB(2, 5)	3.579	0.447	HM(4, 5)	0.371	0.185	141%
"	"	"	HM(5, 5)	0.525	0.262	70%
"	"	"	HM(6, 5)	0.657	0.328	36%
"	"	"	HM(7, 5)	0.746	0.373	20%
"	"	"	RK(9, 5)	2.696	0.300	49%
"	"	"	RK(10, 5)	3.395	0.339	32%
"	"	"	TSRK(8, 5)	3.576	0.447	0%
[a]HB(2, 6)	1.928	0.241	HM(5, 6)	0.209	0.104	131%
HB(3, 6)	2.621	0.328	"	"	"	213%
HB(4, 6)	2.732	0.341	"	"	"	227%
"	"	"	HM(6, 6)	0.362	0.181	89%
"	"	"	HM(7, 6)	0.440	0.220	55%
"	"	"	TSRK(8, 6)	1.936	0.242	41%
[a]HB(3, 7)	1.985	0.248	HM(7, 7)	0.234	0.117	112%
HB(4,7)	2.273	0.284	"	"	"	143%
"	"	"	TSRK(8, 7)	0.568	0.071	300%
[a]HB(3, 8)	1.277	0.160	HM(7, 7)	0.234	0.117	36%
HB(4, 8)	1.588	0.198	"	"	"	70%
HB(5, 8)	1.884	0.235	"	"	"	101%
HB(6, 8)	1.930	0.241	"	"	"	106%
HB(7, 8)	1.943	0.243	"	"	"	108%
"	"	"	TSRK(12, 8)	0.936	0.078	211%
[a]HB(4, 9)	1.107	0.138	HM(7, 7)	0.234	0.117	18%
HB(5, 9)	1.429	0.178	"	"	"	53%
HB(6, 9)	1.623	0.203	"	"	"	73%
HB(7, 9)	1.727	0.216	"	"	"	85%
HB(8, 9)	1.741	0.218	"	"	"	86%
[a]HB(5, 10)	0.971	0.121	HM(7, 7)	0.234	0.117	4%
HB(6, 10)	1.249	0.156	"	"	"	33%
HB(7, 10)	1.453	0.182	"	"	"	55%
HB(8, 10)	1.492	0.186	"	"	"	59%
[a]HB(7, 11)	1.078	0.135	HM(7, 7)	0.234	0.117	15%
HB(8, 11)	1.247	0.156	"	"	"	33%
[a]HB(7, 12)	0.801	0.100				
HB(8, 12)	0.930	0.116				

[a] The superscript 'a' attached to eight HB(k, p) indicates methods with fairly good c_{eff}(HB(k, p)) and low step number k for reduced storage implementation.

In Fig. 1, the global error of y_2 and y_5 at $t_n = \pi + 8$ is plotted for different stepsizes h in a log–log scale so that the curves appear as straight lines with slope p whenever the leading term of the global error at $t = t_n$ is of order p, that is

$$\max\{|y_{2,n} - y_2(t_n)|, |y_{5,n} - y_5(t_n)|\} = O(h^p).$$

For HB(k, p), we have straight lines of slope very close to p, thus confirming the orders of the methods.

6. Numerical results

Numerical results confirm the validity of the monotonicity preserving property of our new optimal schemes on two problems of inviscid Burgers' equation, with unit downstep and square-wave initial conditions, respectively. Similar to Huang [3], we only consider the validity of the monotonicity preserving property dependent on a fixed-order spatial discretization, the weighted essentially non-oscillatory finite difference scheme of order 5 (WENO5) of Jiang and Shu [23].

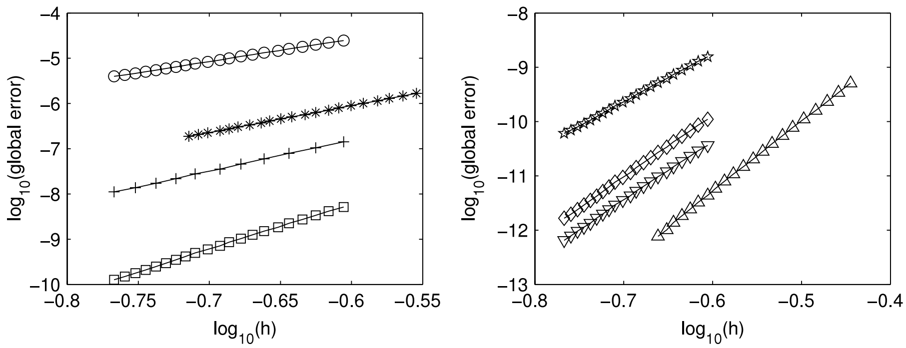
From now on, we shall use the total variation semi-norm $\|y_n\| = \text{TV}(y_n)$ where

$$\text{TV}(y_n) = \sum_j |y_{n,j+1} - y_{n,j}|, \tag{61}$$

and say that a method is *total variation diminishing* (TVD) if

$$\text{TV}(y_{n+1}) \leq \text{TV}(y_n). \tag{62}$$

The following two definitions will help compare different methods with different computational costs more easily and fairly (see more in [3]).



HB(2,5) $\circ$, HB(2,6) $*$, HB(3,7) $+$, HB(4,8) $\square$ (left figure),
HB(5,9) $\star$, HB(6,10) $\diamondsuit$, HB(7,11) $\bigtriangledown$, HB(7,12) $\triangle$ (right figure).

Fig. 1. $\log_{10}$(global error) versus $\log_{10} h$ at $t_n = \pi + 8$ for the listed HB(k, p) applied to problem (60) with constant stepsizes, h, over $t_n \in [0, \pi + 8]$, and $t_n = 4\pi$ for HB(2, 6).

Definition 7. The *largest effective CFL number* of method M denoted by $\mathrm{num}_{\mathrm{eff}}(M)$, for an error ϵ,

$$|\mathrm{TV}(u(x, t_{\mathrm{final}})) - \mathrm{TV}(u(x, t_0))| \le \epsilon, \tag{63}$$

is defined by

$$\mathrm{num}_{\mathrm{eff}}(M) = \max_{\Delta t} \left\{ \frac{\Delta t}{\Delta x} \frac{1}{\ell} \right\}, \tag{64}$$

where ℓ is the number of function evaluations of M per time step.

Then $\max \Delta t_{\mathrm{num}} = \ell \Delta x \mathrm{num}_{\mathrm{eff}}(M)$ is called the *maximum numerical stepsize*. We let $\max\{\Delta t_{\mathrm{theor}}\}$ be the *maximum theoretical time step* taken as

$$\max \Delta t_{\mathrm{theor}} = c(M) \Delta t_{\mathrm{FE}}, \tag{65}$$

and $R_{\mathrm{n/t}}$ be the ratio of the maximum numerical to the maximum theoretical stepsizes

$$R_{\mathrm{n/t}} = \frac{\max \Delta t_{\mathrm{num}}}{\max \Delta t_{\mathrm{theor}}}. \tag{66}$$

It is to be noted that $R_{\mathrm{n/t}}$ indicates how well a method behaves when applied to a given problem. According to Theorem 1, $\max \Delta t_{\mathrm{theor}}$ is the stepsize beyond which, the considered numerical method will theoretically breakdown while $\max \Delta t_{\mathrm{num}}$ is the maximum stepsize, at which the breakdown of the numerical method is observed. Hence, $R_{\mathrm{n/t}}$ can be regarded as a quality measurement of a numerical method for the considered problem. Given two methods with the same c_{eff} and same $\mathrm{num}_{\mathrm{eff}}$, it will be safer to use the method with larger $R_{\mathrm{n/t}}$.

Definition 8. The *percentage efficiency gain* of $\mathrm{num}_{\mathrm{eff}}(\mathrm{M2})$ of method 2 over $\mathrm{num}_{\mathrm{eff}}(\mathrm{M1})$ of method 1, is defined by

$$\mathrm{PEG}(\mathrm{num}_{\mathrm{eff}}(\mathrm{M2}), \mathrm{num}_{\mathrm{eff}}(\mathrm{M1})) = \frac{\mathrm{num}_{\mathrm{eff}}(\mathrm{M2}) - \mathrm{num}_{\mathrm{eff}}(\mathrm{M1})}{\mathrm{num}_{\mathrm{eff}}(\mathrm{M1})}. \tag{67}$$

Some starting values are computed by the optimal RK(5, 4) scheme [8] with small initial stepsize $h \approx 1.0\mathrm{e} - 04$.

6.1. Burgers' equation with a unit downstep initial condition

As in [3], we consider the following problem.

Problem 1. Burgers' equation with a unit downstep initial condition,

$$\frac{\partial}{\partial t} u(x, t) + \frac{\partial}{\partial x} \left[\frac{1}{2} u(x, t)^2 \right] = 0, \qquad u(x, 0) = \begin{cases} 1, & -1 \le x < 0, \\ 0, & 0 < x \le 1, \end{cases} \tag{68}$$

and boundary condition $u(-1, t) = 1$ for $t \ge 0$.

Table 3
Row-wise PEG($\mathrm{num_{eff}}(\mathrm{HB}(k,p))$, $\mathrm{num_{eff}}(\mathrm{OM}(k,p))$) and $R_{n/t}$ for Problem 1.

HB(k, p)	num_eff(HB(k, p))	$R_{n/t}$	OM(k, p)	num_eff(OM(k, p))	$R_{n/t}$	PEG
HB(2, 5)	0.336	2.311	HM(4, 5)	0.182	3.019	85%
"	"	"	HM(5, 5)	0.277	3.247	21%
"	"	"	HM(6, 5)	0.277	2.594	21%
"	"	"	HM(7, 5)	0.317	2.615	6%
"	"	"	RK(10, 5)	0.324	2.933	4%
HB(2, 6)	0.311	3.970	HM(5, 6)	0.174	5.123	79%
HB(3, 6)	0.316	2.968	"	"	"	82%
HB(4, 6)	0.306	2.757	"	"	"	76%
"	"	"	HM(6, 6)	0.169	2.873	81%
"	"	"	HM(7, 6)	0.189	2.643	62%
HB(3, 7)	0.309	3.831	HM(7, 7)	0.127	3.340	143%
HB(4, 7)	0.304	3.292	"	"	"	139%
HB(3, 8)	0.203	3.914	HM(7, 7)	0.127	3.340	60%
HB(4, 8)	0.234	3.628	"	"	"	84%
HB(5, 8)	0.249	3.253	"	"	"	96%
HB(6, 8)	0.289	3.686	"	"	"	128%
HB(7, 8)	0.319	4.041	"	"	"	151%
HB(4, 9)	0.148	3.290	HM(7, 7)	0.127	3.340	17%
HB(5, 9)	0.238	4.099	"	"	"	87%
HB(6, 9)	0.268	4.063	"	"	"	111%
HB(7, 9)	0.268	3.819	"	"	"	111%
HB(8, 9)	0.293	4.142	"	"	"	131%
HB(5, 10)	0.253	6.413	HM(7, 7)	0.127	3.340	99%
HB(6, 10)	0.163	3.213	"	"	"	28%
HB(7, 10)	0.288	4.879	"	"	"	127%
HB(8, 10)	0.288	4.751	"	"	"	127%
HB(7, 11)	0.165	3.766	HM(7, 7)	0.127	3.340	30%
HB(8, 11)	0.245	4.835	"	"	"	93%
HB(7, 12)	0.170	5.224	HM(7, 7)	0.127	3.340	34%
HB(8, 12)	0.180	4.764	"	"	"	42%

First, the spatial derivative of the flux $f(u) = u(x,t)^2/2$ is discretized by the weighted essentially non-oscillatory finite difference scheme of order 5 (WENO5) of Jiang and Shu [23] with stepsize $\Delta x = 1/150$ to obtain the semi-discrete system

$$\frac{d}{dt}u_j(t) = -\frac{1}{\Delta x}\left[f_{j+(1/2)} - f_{j-(1/2)}\right], \tag{69}$$

where $u_j(t) \approx u(x_j, t)$ with $x_j = j\Delta x$, $j = -150, -149, \ldots, 149, 150$, and $f_{j+(1/2)}$ is the numerical flux, which typically is a Lipschitz continuous function of several neighboring values $u_j(t)$ (see [23] for details). Now a time discretization can be applied to (69). For Problem 1, we consider the total variation norm of the numerical solution at $t_{\text{final}} = 1.8$ and take Δt sufficiently small, $\Delta t \le \max \Delta t_{\text{num}}$ such that (63) holds with error $\epsilon = 5.0\text{e} - 02$ for the methods listed in Table 3.

Numerical results show that FE satisfies the TVD property (62) under the time step restriction

$$\Delta t \le \Delta t_{\text{FE}} = 0.325\Delta x. \tag{70}$$

Table 3 presents $\mathrm{num_{eff}}(\mathrm{HB}(k,p))$ and $\mathrm{num_{eff}}(\mathrm{OM}(k,p))$ for other methods (OM) applied to Problem 1. The PEG($\mathrm{num_{eff}}(\mathrm{HB}(k,p))$, $\mathrm{num_{eff}}(\mathrm{OM}(k,p))$) is shown in column 7.

It is seen that:

(a) $\mathrm{num_{eff}}(\mathrm{HB}(k,p)) > \mathrm{num_{eff}}(\mathrm{OM}(k,p))$ for methods of the same order p and all k, so PEG($\mathrm{num_{eff}}(\mathrm{HB}(k,p))$, $\mathrm{num_{eff}}(\mathrm{OM}(k,p))$) > 0.
(b) quite remarkably, even though $c_{\text{eff}}(\mathrm{HB}(7,12)) = 0.100 < c_{\text{eff}}(\mathrm{HM}(7,7)) = 0.117$ and $c_{\text{eff}}(\mathrm{HB}(8,12)) = 0.116 < c_{\text{eff}}(\mathrm{HM}(7,7)) = 0.117$, in this test, HB(7, 12) and HB(8, 12) allow a larger time step since $\mathrm{num_{eff}}(\mathrm{HB}(7,12)) > \mathrm{num_{eff}}(\mathrm{HM}(7,7))$ and $\mathrm{num_{eff}}(\mathrm{HB}(8,12)) > \mathrm{num_{eff}}(\mathrm{HM}(7,7))$.

6.2. Burgers' equation with a square-wave initial condition

As a second comparison, we consider Burgers' equation with a square-wave initial value in Problem 2, which is one of Laney's five test problems [24, p. 311].

Table 4
Row-wise PEG($\text{num}_{\text{eff}}(\text{HB}(k,p))$, $\text{num}_{\text{eff}}(\text{OM}(k,p))$) and $R_{\text{n/t}}$ for Problem 2.

HB(k,p)	$\text{num}_{\text{eff}}(\text{HB}(k,p))$	$R_{\text{n/t}}$	OM(k,p)	$\text{num}_{\text{eff}}(\text{OM}(k,p))$	$R_{\text{n/t}}$	PEG
HB(2, 5)	0.366	4.463	HM(4, 5)	0.192	5.647	91%
"	"	"	HM(5, 5)	0.292	6.069	25%
"	"	"	HM(6, 5)	0.287	4.766	28%
"	"	"	HM(7, 5)	0.312	4.563	17%
"	"	"	RK(10, 5)	0.324	5.200	13%
HB(2, 6)	0.306	6.925	HM(5, 6)	0.179	9.345	71%
HB(3, 6)	0.331	5.512	"	"	"	85%
HB(4, 6)	0.296	4.729	"	"	"	65%
"	"	"	HM(6, 6)	0.174	5.244	70%
"	"	"	HM(7, 6)	0.194	4.811	53%
HB(3, 7)	0.334	7.343	HM(7, 7)	0.124	5.782	169%
HB(4, 7)	0.289	5.549	"	"	"	133%
HB(3, 8)	0.198	6.785	HM(7, 7)	0.124	5.782	60%
HB(4, 8)	0.229	6.295	"	"	"	85%
HB(5, 8)	0.234	5.421	"	"	"	89%
HB(6, 8)	0.329	7.439	"	"	"	165%
HB(7, 8)	0.309	6.940	"	"	"	149%
HB(4, 9)	0.148	5.853	HM(7, 7)	0.124	5.782	20%
HB(5, 9)	0.254	7.757	"	"	"	105%
HB(6, 9)	0.269	7.230	"	"	"	117%
HB(7, 9)	0.254	6.418	"	"	"	105%
HB(8, 9)	0.284	7.119	"	"	"	129%
HB(5, 10)	0.248	11.169	HM(7, 7)	0.124	5.782	100%
HB(6, 10)	0.149	5.207	"	"	"	20%
HB(7, 10)	0.279	8.381	"	"	"	125%
HB(8, 10)	0.289	8.453	"	"	"	133%
HB(7, 11)	0.153	6.191	HM(7, 7)	0.124	5.782	23%
HB(8, 11)	0.233	8.153	"	"	"	88%
HB(7, 12)	0.163	8.907	HM(7, 7)	0.124	5.782	32%
HB(8, 12)	0.178	8.377	"	"	"	44%

Problem 2. Burgers' equation with a square wave initial condition,

$$\frac{\partial}{\partial t}u(x,t)+\frac{\partial}{\partial x}\left[\frac{1}{2}\,u(x,t)^2\right]=0,\qquad u(x,0)=\begin{cases}1, & |x|\le \dfrac{1}{3},\\ 0, & \dfrac{1}{3}<|x|\le 1,\end{cases}\tag{71}$$

and boundary condition $u(-1,t)=u(1,t)$ for $t\ge 0$.

The spatial derivative of Problem 2 is discretized by WENO5 and the total variation of the numerical solution is computed as a function of the effective CFL number, $\text{num}_{\text{eff}}=\Delta t/(\ell\Delta x)$, at $t_{\text{final}}=0.6$. In this case, we take $\text{num}_{\text{eff}}=0.183$ in the time step restriction (70) instead of 0.325.

Table 4 lists num_{eff} in columns 2 and 5, the ratio $R_{\text{n/t}}$ for HB(k,p) in column 3 and for OM(k,p) (other methods) in column 6 applied to Problem 2. The PEG(num_{eff}) is in column 7.

It is seen that the results for Problem 2 listed in Table 4 confirm the observations (a) and (b) for Problem 1.

We observe that the ratio $R_{\text{n/t}}$ of HB(k,p) for Problems 1 and 2 are greater than 1, as with hybrid methods in [3]. The theoretical strong stability bounds of HB(k,p) are thus verified in the numerical comparison of the maximum time steps for these problems.

7. Comparing $\text{HB}_{\text{RK}q}(k,p)$ made of a k-step linear method and RKq of order $q=5$ and 4

This section shows that HB methods obtained by combining k-step methods with RK(8, 5) have larger c_{eff} and larger num_{eff} than HB methods obtained by combining k-step methods with RK(8, 4). (See [10, Section 4.3] for details about the construction of 8-stage HB methods obtained by combining k-step methods with RK(8, 4)).

For simplicity, in Table 5, we only compare $c_{\text{eff}}(\text{HB}_{\text{RK5}}(k,p))$ and $c_{\text{eff}}(\text{HB}_{\text{RK4}}(k,p))$. Since, in the last column, $\text{PEG}(\text{HB}_{\text{RK5}}(k,p),\text{HB}_{\text{RK4}}(k,p))\ge 0$, it is seen that $\text{HB}_{\text{RK5}}(k,p)$ is superior to $\text{HB}_{\text{RK4}}(k,p)$ for methods of the same order, especially for order 11 and order 12.

The eight entries with a superscript 'a', listed in Table 5, indicate $\text{HB}_{\text{RK4}}(k,p)$ with good $c_{\text{eff}}(\text{HB}(k,p))$ and low step number k for reduced storage implementation. These methods can be found in [10, pp. 124–131] in their canonical Shu–Osher form with their $c(\text{HB}(k,p))$, $c_{\text{eff}}(\text{HB}(k,p))$ and abscissa vector σ.

Table 5
$c_{\mathrm{eff}}(\mathrm{HB}_{\mathrm{RK}q}(k,p))$ for $p = 5, 6, \ldots, 12$, as functions of RKq, $q = 4, 5$, which are combined with a k-step method to obtain $\mathrm{HB}_{\mathrm{RK}q}(k,p)$ for the listed k, p taken row-wise. The last column is $\mathrm{PEG}(\mathrm{HB}_{\mathrm{RK5}}(k,p), \mathrm{HB}_{\mathrm{RK4}}(k,p))$.

p	k	c_{eff} of HB(k,p) and PEG		
		$\mathrm{HB}_{\mathrm{RK5}}(k,p)$	$\mathrm{HB}_{\mathrm{RK4}}(k,p)$	PEG
5	2	0.447	[a]0.447	0%
6	2	0.241	[a]0.240	0.4%
	3	0.328	0.328	0%
	4	0.341	0.336	1%
7	3	0.248	[a]0.229	8%
	4	0.284	0.280	1%
8	4	0.198	[a]0.190	4%
	5	0.235	0.231	2%
	6	0.241	0.233	3%
	7	0.243	0.233	4%
9	5	0.178	[a]0.153	16%
	6	0.203	0.191	8%
	7	0.216	0.206	5%
	8	0.218	0.210	4%
10	6	0.156	[a]0.141	11%
	7	0.182	0.170	7%
	8	0.186	0.176	6%
11	7	0.135	0.110	23%
	8	0.156	[a]0.127	23%
12	7	0.100	0.055	82%
	8	0.116	[a]0.091	27%

[a] The superscript 'a' attached to the eight entries indicates $\mathrm{HB}_{\mathrm{RK4}}(k,p)$ with good $c_{\mathrm{eff}}(\mathrm{HB}(k,p))$ and low step number k for reduced storage implementation.

Table 6
$\mathrm{num}_{\mathrm{eff}}(\mathrm{HB}_{\mathrm{RK}q}(k,p))$ for Burgers' equation as a function of RKq, $q = 4, 5$, for the listed k, p taken row-wise and $\mathrm{PEG}(\mathrm{HB}_{\mathrm{RK5}}(k,p), \mathrm{HB}_{\mathrm{RK4}}(k,p))$ for Problems 1 (left) and 2 (right).

p	k	$\mathrm{num}_{\mathrm{eff}}$ and PEG of HB(k,p) as a combination of a k-step linear method with RK5 and RK4					
		Problem 1			Problem 2		
		RK5	RK4	PEG	RK5	RK4	PEG
5	2	0.336	0.334	0.5%	0.336	0.335	0.5%
6	2	0.311	0.305	2%	0.306	0.265	15%
	3	0.316	0.315	0.3%	0.331	0.335	−1%
	4	0.306	0.300	2%	0.296	0.285	4%
7	3	0.309	0.295	5%	0.334	0.295	13%
	4	0.304	0.270	13%	0.289	0.265	9%
8	4	0.234	0.230	2%	0.229	0.219	5%
	5	0.249	0.280	−11%	0.234	0.279	−16%
	6	0.289	0.260	11%	0.329	0.254	30%
	7	0.319	0.270	18%	0.309	0.249	24%
9	5	0.238	0.240	−0.8%	0.254	0.229	11%
	6	0.268	0.260	3%	0.269	0.224	20%
	7	0.268	0.275	−3%	0.254	0.260	−2%
	8	0.293	0.270	9%	0.284	0.254	12%
10	7	0.288	0.215	34%	0.279	0.214	30%
	8	0.288	0.235	23%	0.289	0.232	25%
11	7	0.165	0.200	−18%	0.153	0.169	−9%
	8	0.245	0.195	26%	0.233	0.204	14%
12	7	0.170	0.105	62%	0.163	0.114	43%
	8	0.180	0.125	44%	0.178	0.110	62%

We also apply our $\mathrm{HB}_{\mathrm{RK}q}(k,p)$ on Problems 1 and 2 for Burgers' equation. Table 6 lists $\mathrm{num}_{\mathrm{eff}}$ of $\mathrm{HB}_{\mathrm{RK}q}(k,p)$ for $q = 4, 5$. The two columns PEG denote $\mathrm{PEG}(\mathrm{HB}_{\mathrm{RK5}}(k,p), \mathrm{HB}_{\mathrm{RK4}}(k,p))$ for Problems 1 and 2, respectively. It is seen that, generally, $\mathrm{HB}_{\mathrm{RK5}}(k,p)$ have higher $\mathrm{num}_{\mathrm{eff}}$ than $\mathrm{HB}_{\mathrm{RK4}}(k,p)$, especially with methods of order 12.

8. Conclusion

In this work, we construct new optimal, explicit, k-step, 8-stage, SSP Hermite–Birkhoff (HB) methods of order p, HB(k, p), for $p = 5, 6, \ldots, 12$, with nonnegative coefficients, in their Shu–Osher and canonical Shu–Osher forms, respectively, by combining linear k-step methods of order $(p - 4)$ with 8-stage explicit Runge–Kutta (RK) methods of order 5. We did not find in the literature any counterparts of HB(k, p) of order greater than eight among hybrid and general linear multistep and multistage methods. Moreover, compared to some other methods, such as hybrid methods and RK methods of the same order, HB methods generally have larger SSP coefficients and larger maximum effective CFL numbers when tested on Burgers' equation. Generally, HB methods of orders 11 and 12 combined with 8-stage, order 5 RK(8, 5) have maximum step size 24% larger than HB combined with RK(8, 4).

Acknowledgments

This work was supported in part by the Vietnam Ministry of Education and Training, and the Natural Sciences and Engineering Research Council of Canada.

References

[1] S. Gottlieb, D.I. Ketcheson, C.W. Shu, High order strong stability preserving time discretization, J. Sci. Comput. 38 (3) (2009) 251–289. http://dx.doi.org/10.1007/s10915-008-9239-z.
[2] S. Gottlieb, D.I. Ketcheson, C.W. Shu, Strong Stability Preserving Runge–Kutta and Multistep Time Discretizations, World Scientific, Singapore, 2011.
[3] C. Huang, Strong stability preserving hybrid methods, Appl. Numer. Math. 59 (2009) 891–904.
[4] A. Harten, High resolution schemes for hyperbolic conservation laws, J. Comput. Phys. 49 (1983) 357–393.
[5] S. Osher, S. Chakravarthy, High resolution schemes and the entropy condition, SIAM J. Numer. Anal. 21 (1984) 955–984.
[6] P.K. Sweby, High resolution schemes using flux limiters for hyperbolic conservation laws, SIAM J. Numer. Anal. 21 (1984) 995–1011.
[7] B. Cockburn, C.W. Shu, TVB Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws II: general framework, Math. Comp. 52 (1989) 411–435.
[8] S.J. Ruuth, R.J. Spiteri, Two barriers on strong-stability-preserving time discretization methods, J. Sci. Comput. 17 (2002) 211–220.
[9] T. Nguyen-Ba, V. Bozic, E. Kengne, R. Vaillancourt, One-step 7-stage Hermite–Birkhoff–Taylor ODE solver of order 11, J. Comput. Appl. Math. 234 (2010) 192–208. http://dx.doi.org/10.1016/j.cam.2009.12.015.
[10] Huong Nguyen-Thu, aka Thu Huong Nguyen, Strong stability preserving Hermite–Birkhoff time-discretization methods, Ph.D. Thesis, University of Ottawa, Ottawa, Ontario, Canada, 2012, p. 146. http://www.ruor.uottawa.ca/en/bitstream/handle/10393/23491/Nguyen_Thu_Huong_2012_thesis.pdf?sequence=3.
[11] C.W. Shu, S. Osher, Efficient implementation of essentially non-oscillatory shock-capturing schemes, J. Comput. Phys. 77 (1988) 439–471.
[12] T. Nguyen-Ba, H. Nguyen-Thu, T. Giordano, R. Vaillancourt, Strong-stability-preserving 7-stage Hermite–Birkhoff time-discretization methods, J. Sci. Comput. 50 (1) (2011) 63–90. http://dx.doi.org/10.1007/s10915-011-9473-7.
[13] T. Nguyen-Ba, H. Nguyen-Thu, T. Giordano, R. Vaillancourt, Strong-stability-preserving, 6-stage, Hermite–Birkhoff time-discretization methods of order 4 to 12, Automat. Comput. Appl. Math. 21 (1) (2012) 5–28.
[14] D.I. Ketcheson, Highly efficient strong stability preserving Runge–Kutta methods with low-storage implementations, SIAM J. Sci. Comput. 30 (4) (2008) 2113–2136.
[15] S. Gottlieb, C.W. Shu, E. Tadmor, Strong stability-preserving high-order time discretization methods, SIAM Rev. 43 (2001) 89–112.
[16] R.J. Spiteri, S.J. Ruuth, A new class of optimal high-order strong-stability-preserving time-stepping schemes, SIAM J. Numer. Anal. 40 (2007) 469–491.
[17] R.J. Spiteri, S.J. Ruuth, Nonlinear evolution using optimal fourth-order strong-stability-preserving Runge–Kutta methods, Math. Comput. Simul. 62 (2003) 125–135.
[18] S.J. Ruuth, R.J. Spiteri, High-order strong-stability-preserving Runge–Kutta methods with down-biased spatial discretizations, SIAM J. Numer. Anal. 42 (2004) 974–996.
[19] S.J. Ruuth, Global optimization of explicit strong-stability-preserving Runge–Kutta methods, Math. Comp. 75 (2006) 183–207.
[20] S.J. Ruuth, W. Hundsdorfer, High-order linear multistep methods with general monotonicity and boundedness properties, J. Comput. Phys. 209 (2005) 226–248.
[21] D.I. Ketcheson, S. Gottlieb, C.B. Macdonald, Strong stability two-step Runge–Kutta methods, in: Report Number 11/7, Oxford Centre for Collaborative Applied Mathematics, Mathematical Institute, University Oxford, 2011.
[22] D.I. Ketcheson, Computation of optimal monotonicity preserving general linear methods, Math. Comp. 78 (2009) 1497–1513.
[23] G. Jiang, C.W. Shu, Efficient implementation of weighted ENO schemes, J. Comput. Phys. 126 (1996) 202–228.
[24] C. Laney, Computational Gasdynamics, Cambridge University Press, Cambridge, UK, 1998.