Journal of Computational Mathematics, Vol.18, No.6, 2000, 645–656.

ORDER PROPERTIES AND CONSTRUCTION OF SYMPLECTIC RUNGE-KUTTA METHODS*1)

Shou-fu Li
(*Institute for Computational and Applied Mathematics, Xiangtan University, Xiangtan 411105, China*)

Abstract

The main results of this paper are as follows: (1) Suppose an s stage Runge-Kutta method is consistent, irreducible, non-confluent and symplectic. Then this method is of order at least $2p+l(1 \le p \le s-1)$ provided that the simplifying conditions $C(p)$ (or $D(p)$ with non-zero weights) and $B(2p+l)$ hold, where $l = 0, 1, 2$. (2) Suppose an s stage Runge-Kutta method is consistent, irreducible and non-confluent, and satisfies the simplifying conditions $C(p)$ and $D(p)$ with $0 < p \le s$. Then this method is symplectic if and only if either $p = s$ or the nonlinear stability matrix M of the method has an $(s-p)\times(s-p)$ chief submatrix $\hat{M} = 0$. (3) Using the results (1) and (2) as bases, we present a general approach for the construction of symplectic Runge-Kutta methods, and a software has been designed, by means of which, the coefficients of s stage symplectic Runge-Kutta methods satisfying $C(p), D(p)$ and $B(2p+l)$ can be easily computed, where $1 \le p \le s, 0 \le l \le 2, s \le 2p+l \le 2s$.

Key words: Numerical analysis, Symplectic Runge-Kutta methods, Simplifying conditions, Order results.

1. Introduction

For a given s stage Runge-Kutta method

$$\begin{array}{c|c} \mu & \text{A} \\ \hline & \gamma^T \end{array} \tag{1.1}$$

with $A = [a_{ij}], \mu = [\mu_1, \mu_2, \cdots, \mu_s]^T$ and $\gamma = [\gamma_1, \gamma_2, \cdots, \gamma_s]^T \ne 0$, we introduce the following simplifying conditions as in Butcher [1]

$$\begin{cases} B(p): \ i\gamma^T\mu^{i-1} = 1, & i = 1, 2, \cdots, p, \\ C(p): \ iA\mu^{i-1} = \mu^i, & i = 1, 2, \cdots, p, \\ D(p): \ iA^T\text{diag}(\gamma)\mu^{i-1} = \gamma - \text{diag}(\gamma)\mu^i, & i = 1, 2, \cdots, p, \end{cases}$$

and make the notational convension

$$\begin{cases} M = [m_{ij}] := \text{diag}(\gamma)A + A^T\text{diag}(\gamma) - \gamma\gamma^T, \\ U_{lm} := [\rho_l(\mu), \rho_{l+1}(\mu), \cdots, \rho_m(\mu)], \\ V_{lm} := [\rho_l'(\mu), \rho_{l+1}'(\mu), \cdots, \rho_m'(\mu)], \\ B_{lm} := [b_l, b_{l+1}, \cdots, b_m], \ \ C_{lm} := [c_l, c_{l+1}, \cdots, c_m], \\ D_{lm} := [d_l, d_{l+1}, \cdots, d_m], \end{cases}$$

* Received February 15, 1998.
1)Project supported by the National Natural Science Foundation of China.

where $l \le m, \rho_i(x), i = 1,2,3,\cdots$, are arbitrarily given i-th polynomials with the property that $\rho_i(0) = 0$,

$\rho_i(\nu) := [\rho_i(\nu_1), \rho_i(\nu_2), \cdots, \rho_i(\nu_N)]^T$,
$\rho_i'(\nu) := [\rho_i'(\nu_1), \rho_i'(\nu_2), \cdots, \rho_i'(\nu_N)]^T$, for $\nu = [\nu_1, \nu_2, \cdots, \nu_N]^T \in \mathbf{R}^N$,
$b_i := \gamma^T \rho_i'(\mu) - \rho_i(1)$,
$c_i := A\rho_i'(\mu) - \rho_i(\mu)$,
$d_i := A^T \mathrm{diag}(\gamma)\rho_i'(\mu) - \gamma\rho_i(1) + \mathrm{diag}(\gamma)\rho_i(\mu)$, for $i = 1,2,3,\cdots$.

Note that $B(p), C(p)$ and $D(p)$ are equivalent to $B_{1,p} = 0, C_{1,p} = 0$ and $D_{1,p} = 0$ respectively. We shall always denote $B_{1,s}, C_{1,s}, D_{1,s}$ and $V_{1,s}$ by B, C, D and V respectively, and frequently refer the following two theorems in the sequel.

Theorem 1.1. *(cf. Butcher [1]) $B(p), C(\eta)$ and $D(\xi)$ with* $\min\{\eta+\xi+1, 2\eta+2\} \ge$ *p implies that the method has order at least p.*

Theorem 1.2. *(cf. Sanz-Serna [2] and Lasagni [3]) An irreducible Runge-Kutta method is symplectic if and only if $M = 0$.*

2. Order Properties

Lemma 2.1. *Suppose $B(q)$ holds. Then we have*

$$d_i^T \rho_j'(\mu) - c_j^T \mathrm{diag}(\gamma)\rho_i'(\mu) = 0 \qquad for \quad i+j \le q. \tag{2.1}$$

Lemma 2.2. *Suppose that*

$$d_{i_k}^T \rho_{j_k}'(\mu) - c_{j_k}^T \mathrm{diag}(\gamma)\rho_{i_k}'(\mu) = 0, \;\; i_k + j_k = k \quad for \quad k = 2,3,\cdots,q. \tag{2.2}$$

Then $B(1)$ implies $B(q)$.

Corollary 2.3. *The following implications hold.*

(1) $B(q)$ *and* $C(p) \Longrightarrow d_i^T \rho_j'(\mu) = 0$ *for* $j \le p$, $i + j \le q$,
(2) $B(q)$ *and* $D(p) \Longrightarrow c_i^T \mathrm{diag}(\gamma)\rho_j'(\mu) = 0$ *for* $j \le p$, $i + j \le q$,
(3) $B(p+q)$ *and* $C(p) \Longrightarrow D_{1,q}^T V_{1,p} = 0$,
(4) $B(p+q)$ *and* $D(p) \Longrightarrow C_{1,q}^T \mathrm{diag}(\gamma) V_{1,p} = 0$,
(5) $B(1)$, $C_{1,p}^T \mathrm{diag}(\gamma) V_{1,q} = 0$ *and* $D_{1,q}^T V_{1,p} = 0 \Longrightarrow B(p+q)$,
(6) $B(1)$, $C(p)$ *and* $D(q) \Longrightarrow B(p+q)$.

Proof. Corollary 2.3 follows directly from Lemmas 2.1 and 2.2. Lemmas 2.1 and 2.2 can be easily verified by using the following identity and simple induction.

$$d_i^T \rho_j'(\mu) - c_j^T \mathrm{diag}(\gamma)\rho_i'(\mu) = b_{i+j}^{(i,j)} - \rho_i(1) b_j, \quad i,j = 1,2,3,\cdots, \tag{2.3}$$

where

$$b_{i+j}^{(i,j)} = \gamma^T \frac{d}{dx}(\rho_i(x)\rho_j(x))_{x=\mu} - \rho_i(1)\rho_j(1).$$

Theorem 2.4. *Suppose the method (1.1) is irreducible and symplectic. Then we have the following implications for $1 \le p \le s$.*

(1) $B(1)$ *and* $C(p) \Longrightarrow B(2p)$;
(2) $B(1)$ *and* $D(p) \Longrightarrow B(2p)$;
(3) $B(1)$ *and* $C(p)$ *with distinct abscissae* $\Longrightarrow B(2p)$ *and* $D(p)$;
(4) $B(1)$ *and* $D(p)$ *with distinct abscissae and nonzero weights* $\Longrightarrow B(2p)$ *and* $C(p)$;
(5) $C(1)$ *and* $D(p)$ *with distinct abscissae* $\Longrightarrow B(2p)$;
(6) $D(1)$ *and* $C(p)$ *with distinct abscissae* $\Longrightarrow B(2p)$ *and* $D(p)$;

(7) *$C(1)$ and $D(p)$ with distinct abscissae and nonzero weights* $\Longrightarrow B(2p)$ *and* $C(p)$.

Proof. Since

$$\begin{aligned} M\rho_i'(\mu) &= A^T \mathrm{diag}(\gamma)\rho_i'(\mu) - \gamma\rho_i(1) + \mathrm{diag}(\gamma)\rho_i(\mu) \\ &+ \mathrm{diag}(\gamma)[A\rho_i'(\mu) - \rho_i(\mu)] - \gamma[\gamma^T\rho_i'(\mu) - \rho_i(1)] \\ &= d_i + \mathrm{diag}(\gamma)c_i - \gamma b_i, \end{aligned}$$

we have

$$V^T M V = (D + \mathrm{diag}(\gamma)C - \gamma B)^T V, \tag{2.4}$$

and therefore, in view of Theorem 1.2, irreducibility and symplecticness implies

$$Q := (d_i + \mathrm{diag}(\gamma)c_i - \gamma b_i)^T \rho_j'(\mu) = 0, \quad i,j = 1,2,\cdots,s. \tag{2.5}$$

On the other hand, from (2.3) we find that

$$\begin{aligned} Q &= d_i^T\rho_j'(\mu) + d_j^T\rho_i'(\mu) - b_{i+j}^{(i,j)} + \rho_j(1)b_i - b_i(b_j + \rho_j(1)) \\ &= d_i^T\rho_j'(\mu) + d_j^T\rho_i'(\mu) - b_{i+j}^{(i,j)} - b_ib_j, \end{aligned} \tag{2.6}$$

and that

$$\begin{aligned} Q &= c_j^T\mathrm{diag}(\gamma)\rho_i'(\mu) + b_{i+j}^{(i,j)} - \rho_i(1)b_j + c_i^T\mathrm{diag}(\gamma)\rho_j'(\mu) - b_i(b_j + \rho_j(1)) \\ &= c_i^T\mathrm{diag}(\gamma)\rho_j'(\mu) + c_j^T\mathrm{diag}(\gamma)\rho_i'(\mu) + b_{i+j}^{(i,j)} - \rho_i(1)b_j - \rho_j(1)b_i - b_ib_j. \end{aligned} \tag{2.7}$$

Thus, by simple induction, (1) follows from (2.5) and (2.7), (2) follows from (2.5) and (2.6). (3)–(7) are direct consequences of (1), (2) and (2.5). Note that the results (3) and (4) of Theorem 2.4 have also been obtained in paper [4].

From Theorem 2.4 we see that for any consistent irreducible symplectic Runge-Kutta method with distinct abscissae and non-zero weights, the maximum natural number p such that $C(p)$ holds is the same as that such that $D(p)$ holds, and the maximum natural number q such that $B(q)$ holds is not less than $2p$. So without loss of generality we can expect a symplectic Runge-Kutta method to satisfy $C(p), D(p)$ and $B(q)$ with $q \geq 2p$.

Theorem 2.5. *Suppose an s stage Runge-Kutta method of the form (1.1) is consistent, irreducible, non-confluent and symplectic, and satisfies $C(p)$ (or $D(p)$ with non-zero weights) and $B(2p+2)$, where $1 \leq p \leq s-1$. Then this method is of order at least $2p+2$.*

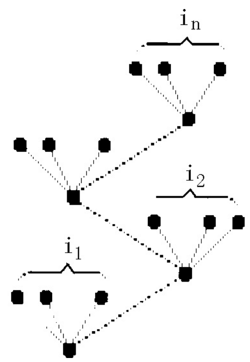
Proof. We first note that in view of Theorem 2.4, the assumptions of Theorem 2.5 implies that the simplifying conditions $C(p), D(p)$ and $B(2p+2)$ are all satisfied. Let T denote the set of all rooted trees with no more than $2p+2$ nodes (cf. [5,6]). With each tree $t \in T$ we associate two numbers $r(t)$ and $\delta(t)$, and an $r(t)$-th polynomial $\varphi(t)$ in the coefficients a_{ij}, γ_j and μ_j. Here $r(t)$ denotes the number of nodes in t, $\delta(t)$ denotes the product of $r(u)$ over u where for each node of t, u is the subtree formed from that node and all nodes that can be reached from it by following upward growing branches. To form $\varphi(t)$, attach labels $i, j, k, \cdots$ to each of the nodes of t, with i the label attached to the root, form the product of γ_i and of a_{jk} for each upward growing branch from j to k, and then sum over each label from 1 to s. Note that in writing formula for $\varphi(t)$ we can always use the abreviation $\sum_k a_{jk} = \mu_j$ since $C(1)$ holds. Following Butcher [5], we thus need to prove that $\varphi(t) = 1/\delta(t)$ for all $t \in T$. Suppose a tree $t \in T$ has a node, other than the root, from which exactly k nodes branch and each of these k

nodes is a terminal, where $k < p$. Then, as $C(p)$ holds, we have

$$\begin{cases} \varphi(t) = \sum_{l=1}^{s} \varphi_l(t) \sum_{m=1}^{s} a_{lm}\mu_m^k = \frac{1}{k+1} \sum_{l=1}^{s} \varphi_l(t)\mu_l^{k+1} = \frac{1}{k+1}\varphi(\bar{t}), \\ r(t) = r(\bar{t}), \qquad \delta(t) = (k+1)\delta(\bar{t}), \end{cases}$$

where $\varphi_l(t)$ are some $(2p+1-k)$-th polynomials, rooted tree $\bar{t}$ is the same as t except that the k nodes referred to above are all moved one step closer to the root. This means that $\varphi(t) = 1/\delta(t)$ is equivalent to $\varphi(\bar{t}) = 1/\delta(\bar{t})$. Thus all trees with the property that characterised t can be removed from consideration. With all the trees removed in aforementioned way, there remain only trees of the form

$t_{i_1,i_2,\cdots,i_n} =$



where $i_1, i_2, \cdots, i_n$ are nonnegative integers, $i_n \geq p$ whenever $n > 1$. It is easy to varify that

$$\begin{cases} r(t_{i_1,i_2,\cdots,i_n}) = \sum_{l=1}^{n}(i_l+1) \leq 2p+2, \\ \delta(t_{i_1,i_2,\cdots,i_n}) = \prod_{j=1}^{n}\sum_{l=j}^{n}(i_l+1), \\ \varphi(t_{i_1,i_2,\cdots,i_n}) = \gamma^T \prod_{l=1}^{n-1}(\mathrm{diag}(\mu^{i_l})A)\mu^{i_n}. \end{cases} \tag{2.8}$$

We thus only need to prove that

$$\varphi(t_{i_1,i_2,\cdots,i_n}) = \frac{1}{\delta(t_{i_1,i_2,\cdots,i_n})} \quad \text{for } \sum_{l=1}^{n}(i_l+1) \leq 2p+2, \tag{2.9}$$

where $i_n \geq p$ whenever $n > 1$. The truth of (2.9) for $n = 1$ is a direct consequence of $B(2p+2)$, and it is also easy to show the truth of (2.9) for $n = 2$, i. e.

$$\gamma^T \mathrm{diag}(\mu^{i_1})A\mu^{i_2} = \tfrac{1}{(i_2+1)(i_1+i_2+2)} \quad \text{for } i_1+i_2 \leq 2p, i_2 \geq p. \tag{2.10}$$

In fact, since $i_1 + i_2 \leq 2p$ and $i_2 \geq p$, we have $i_1 \leq p$. (2.10) with $i_1 < p$ follows directly from $D(p)$ and $B(2p+2)$. Furthermore, since the method is irreducible and symplectic, (2.5) and (2.7) hold. Choose $\rho_i(\mu) = \mu^i$, $i = 1, 2, \cdots, s$. Then (2.5), (2.7) and $B(2p+2)$ yield

$$(\mu^{i-1})^T \mathrm{diag}(\gamma)A\mu^{j-1} + (\mu^{j-1})^T \mathrm{diag}(\gamma)A\mu^{i-1} = \frac{1}{ij} \;\text{ for }\; i, j = 1, 2, \cdots, p+1.$$

Letting $i = j = p+1$, we find that (2.10) with $i_1 = i_2 = p$ also holds. We now only need to prove that

$$\varphi(t_{i_1,\cdots,i_{m+1}}) = \frac{1}{\delta(t_{i_1,\cdots,i_{m+1}})} \quad \text{for} \sum_{l=1}^{m+1}(i_l+1) \le 2p+2,$$

under the inductive assumption that (2.9) with $n = m$ holds, where $2 \le m < 2p+2, i_{m+1} \ge p$. Since

$$i_{m+1} \ge p, \quad i_1 + i_{m+1} + 2 < \sum_{l=1}^{m+1}(i_l+1) \le 2p+2,$$

we have $i_1 < p$, and therefore in view of $D(p)$ and the inductive assumption

$$\begin{aligned}\varphi(t_{i_1,\cdots,i_{m+1}}) &= (A^T\mathrm{diag}(\gamma)\mu^{i_1})^T \prod_{l=2}^{m}(\mathrm{diag}(\mu^{i_l})A)\mu^{i_{m+1}} \\ &= \tfrac{1}{i_1+1}(\gamma - \mathrm{diag}(\gamma)\mu^{i_1+1})^T \prod_{l=2}^{m}(\mathrm{diag}(\mu^{i_l})A)\mu^{i_{m+1}} \\ &= \tfrac{1}{i_1+1}[\gamma^T \prod_{l=2}^{m}(\mathrm{diag}(\mu^{i_l})A)\mu^{i_{m+1}} - \gamma^T\mathrm{diag}(\mu^{i_1+i_2+1})A\prod_{l=3}^{m}(\mathrm{diag}(\mu^{i_l})A)\mu^{i_{m+1}}] \\ &= \frac{1}{i_1+1}\{[\prod_{j=2}^{m+1}\sum_{l=j}^{m+1}(i_l+1)]^{-1} - [\sum_{q=1}^{m+1}(i_q+1)\prod_{j=3}^{m+1}\sum_{l=j}^{m+1}(i_l+1)]^{-1}\} \\ &= \frac{1}{\delta(t_{i_1,\cdots,i_{m+1}})}.\end{aligned}$$

This completes the proof of Theorem 2.5.

Using Theorems 2.4, 2.5 and 1.1, we get the following Corollaries:

Corollary 2.6. *Suppose an s stage Runge-Kutta method of the form (1.1) is consistent, irreducible, non-confluent and symplectic. Then this method is of order at least $2p(1 \le p \le s)$ provided that $C(p)$ holds, of order at least $2p+i(1 \le p \le s-1)$ provided that both $C(p)$ and*

$$b_{2p+j} = 0, \qquad j = 1, 2, \cdots, i \tag{2.11}$$

hold, where $i = 1, 2$.

Corollary 2.7. *Suppose an s stage Runge-Kutta method of the form (1.1) is consistent, irreducible, non-confluent and symplectic, and the weights $\gamma_i \ne 0$, $i = 1, 2, \cdots, s$. Then this method is of order at least $2p(1 \le p \le s)$ provided that $D(p)$ holds, of order at least $2p+i(1 \le p \le s-1)$ provided that both $D(p)$ and (2.11) hold, where $i = 1, 2$.*

Note that The first part of the results of Corollary 2.6 (resp. 2.7) has also been obtained by Aiguo Xiao and Shoufu Li [4].

Theorem 2.8. *An irreducible Runge-Kutta method satisfying $B(1), C(p)$ and $D(p)$, with $0 < p \le s$ and $\mu_1, \mu_2, \cdots, \mu_s$ distinct, is symplectic if and only if either p=s or*

$$V_{p+1,s}^T M V_{p+1,s} = 0. \tag{2.12}$$

Proof. Since $B(1), C(p)$ and $D(p)$ implies B(2p) (cf. Corollary 2.3), it follows from (2.4) and the symmetricness of the matrix M that

$$V^TMV = \begin{cases} \begin{bmatrix} 0 & 0 \\ 0 & V_{p+1,s}^T M V_{p+1,s} \end{bmatrix} & \text{for } 0 < p < s, \\ 0 & \text{for } p = s, \end{cases}$$

where V is non-singular because of $\mu_1, \mu_2, \cdots, \mu_s$ distinct. Hence, $M = 0$ is equivalent to either $p = s$ or $V_{p+1,s}^T M V_{p+1,s} = 0$, and therefore in view of Theorem 1.2 the conclusion follows.

Theorem 2.9. *An irreducible Runge-Kutta method satisfying $B(1), C(p)$ and $D(p)$, with $0 < p < s$ and $\mu_1, \mu_2, \cdots, \mu_s$ distinct, is symplectic if and only if the matrix M has an $(s-p) \times (s-p)$ chief submatrix $\hat{M} = 0$.*

Proof. In view of Theorems 1.2 and 2.8, we only need to prove that $\hat{M} = 0$ implies that (2.12) holds. Suppose the matrix $\hat{M}$ is formed by removing the $i_1, i_2, \cdots, i_p$-th rows and $i_1, i_2, \cdots, i_p$-th colums from the matrix M. Thus we can reorder the elements of the set $\{1, 2, \cdots, s\}$ such that $\{1, 2, \cdots, s\} = \{i_1, i_2, \cdots, i_p; i_{p+1}, \cdots, i_s\}$. Since $\hat{M} = 0$, we have

$$m_{i_l i_r} = 0 \qquad \text{for } l, r = p+1, p+2, \cdots, s. \tag{2.13}$$

Letting

$$\rho_q'(x) = \prod_{l=1}^{q-1} (x - \mu_{i_l}), \qquad q = p+1, p+2, \cdots, s, \tag{2.14}$$

we get

$$\rho_q'(\mu_{i_l}) = 0 \qquad \text{for } l = 1, 2, \cdots, q-1. \tag{2.15}$$

It follows from (2.13) and (2.15) that the (j,k)-element of the matrix

$$V_{p+1,s}^T M V_{p+1,s}$$

is equal to

$$\begin{aligned} &\rho_{p+j}'(\mu)^T M \rho_{p+k}'(\mu) = \sum_{l,r=1}^{s} m_{i_l i_r} \rho_{p+j}'(\mu_{i_l}) \rho_{p+k}'(\mu_{i_r}) \\ &= \Big(\sum_{l=1}^{p+j-1} \sum_{r=1}^{s} + \sum_{l=p+j}^{s} \sum_{r=1}^{p+k-1} + \sum_{l=p+j}^{s} \sum_{r=p+k}^{s} \Big) m_{i_l i_r} \rho_{p+j}'(\mu_{i_l}) \rho_{p+k}'(\mu_{i_r}) \\ &= 0, \qquad\qquad j, k = 1, 2, \cdots, s-p. \end{aligned}$$

This means that (2.12) holds and completes the proof of Theorem 2.9.

3. Construction of Symplectic Methods

Using the results obtained in the present paper, we can easily make a thorough discussion and establish general approach for the construction of symplectic Runge-Kutta methods. In fact, Theorem 2.9, Corollaries 2.3, 2.6 and 2.7 lead directly to

Proposition 3.1. *Suppose the method (1.1) with $\gamma_1, \gamma_2, \cdots, \gamma_s \neq 0$ and $\mu_1, \mu_2, \cdots, \mu_s$ distinct is irreducible and satisfies $C(p), B(2p+l)$ and*

$$D_{1,p}^T \hat{E}_p = 0, \qquad \hat{E}_p^T A \hat{E}_p = \hat{A} \mathrm{diag}(\gamma_{p+1}, \cdots, \gamma_s), \tag{3.1}$$

where $0 < p < s$, $l = 0, 1, 2$,

$$\hat{E}_p = [e_{p+1}, \cdots, e_s], \quad \hat{A} = \begin{bmatrix} \alpha_{p+1,p+1} & \cdots & \alpha_{p+1,s} \\ \cdots & \cdots & \cdots \\ \alpha_{s,p+1} & \cdots & \alpha_{s,s} \end{bmatrix},$$

where α_{ij} are any given real numbers satisfying $\alpha_{ij} + \alpha_{ji} = 1$, $i, j = p+1, \cdots, s$, e_j denotes the s-dimensional vector with all components equal to zero except the j-th *component equal to 1. Then this method is symplectic, satisfies $D(p)$, and has consistency order $2p+l$. Furthermore, for the case of $\gamma_1, \gamma_2, \cdots, \gamma_s \geq 0$, the method is also algebraically stable.*

Let

$$p(x) = \prod_{k=1}^{s} (x - \mu_k).$$

Then it is readily shown that for any given integer $q \in [0, s]$, the simplifying condition $B(2s-q)$ holds if and only if the equations

$$\begin{cases} i \sum_{j=1}^{s} \gamma_j \mu_j^{i-1} = 1, i = 1, 2, \cdots, s, & (3.2a) \\ \int_0^1 x^{k-1} p(x) dx = 0, k = 1, 2, \cdots, s-q & (3.2b) \end{cases}$$

are satisfied (cf. [7]). For $q = 0$ equation (3.2b) leads to

$$p(x) = \left(\det \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{s} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{s+1} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{s} & \frac{1}{s+1} & \cdots & \frac{1}{2s-1} \end{bmatrix} \right)^{-1} \det \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{s+1} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{s+2} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{s} & \frac{1}{s+1} & \cdots & \frac{1}{2s} \\ 1 & x & \cdots & x^s \end{bmatrix}, \tag{3.3}$$

and for $0 < q < s$, to

$$p(x) = \hat{p}(x) \prod_{m=1}^{q} (x - \mu_{i_m}), \tag{3.4}$$

where

$$\begin{cases} \hat{p}(x) = c^{-1} \det \begin{bmatrix} g_1 & g_2 & \cdots & g_{s-q+1} \\ g_2 & g_3 & \cdots & g_{s-q+2} \\ \cdots & \cdots & \cdots & \cdots \\ g_{s-q} & g_{s-q+1} & \cdots & g_{2(s-q)} \\ 1 & x & \cdots & x^{s-q} \end{bmatrix}, \\ g_l = \int_0^1 x^{l-1} \prod_{m=1}^{q} (x - \mu_{i_m}) dx, \quad l = 1, 2, 3, \cdots, \\ c = \det \begin{bmatrix} g_1 & g_2 & \cdots & g_{s-q} \\ g_2 & g_3 & \cdots & g_{s-q+1} \\ \cdots & \cdots & \cdots & \cdots \\ g_{s-q} & g_{s-q+1} & \cdots & g_{2(s-q)-1} \end{bmatrix}. \end{cases} \tag{3.5}$$

For the case of $q > 0$, the real numbers $\mu_{i_1}, \mu_{i_2}, \cdots, \mu_{i_q}$, which are q distinct roots of $p(x)$, should be appropriately choosen in advance such that the polynomial $p(x)$ defined by (3.4) and (3.5) is of degree s and has s distinct real roots (cf. [7]). Now we are in a position to present the following general approach for the construction of s stage symplectic Runge-Kutta methods of the form (1.1) satisfying $C(p), D(p)$ and $B(2p+l)$, where the natural numbers p and l satisfies $1 \le p \le s, 0 \le l \le 2, s \le 2p+l \le 2s$.

Step 1. Determine abscissas. For $2p+l=s$, $\mu_1, \mu_2, \cdots, \mu_s$ may be any given distinct real numbers, for $2p+l=2s$, choose $\mu_1, \mu_2, \cdots, \mu_s$ as s roots of the polynomial $p(x)$ defined by (3.3), and for $s < 2p+l < 2s$, we first choose appropriately $\mu_{i_1}, \mu_{i_2}, \cdots, \mu_{i_q}$ with $q = 2s-2p-l$, and then compute all the roots $\mu_1, \mu_2, \cdots, \mu_s$ of the polynomial $p(x)$ defined by (3.4) and (3.5).

Step 2. Determine weights. Compute γ from (3.2a).

Step 3. Determine Runge-Kutta matrix. The matrix A can be determined by condition $C(s)$ for $p=s$, and by $C(p)$ and (3.1) for $p<s$. For more detail, $C(p)$ means that

$$A[\mu^0, 2\mu^1, \cdots, p\mu^{p-1}] = [\mu^1, \mu^2, \cdots, \mu^p],$$

equation (3.1) is equivalent to

$$BA\hat{E}_p = \eta,$$

where

$$\begin{cases} B = \begin{bmatrix} \gamma^T \\ \gamma^T \text{diag}(\mu) \\ \vdots \\ \gamma^T \text{diag}(\mu^{p-1}) \\ e_{p+1}^T \\ \vdots \\ e_s^T \end{bmatrix}, \quad \eta = \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}, \\ \eta_1 = \begin{bmatrix} \frac{1-\mu_1}{1} & \cdots & \frac{1-\mu_s}{1} \\ \cdots & \cdots & \cdots \\ \frac{1-\mu_1^p}{p} & \cdots & \frac{1-\mu_s^p}{p} \end{bmatrix} \text{diag}(\gamma)\hat{E}_p, \quad \eta_2 = \hat{A}\text{diag}(\gamma_{p+1}, \cdots, \gamma_s), \end{cases} \tag{3.6}$$

and therefore we have

$$AV = Q, \tag{3.7}$$

where

$$V = [\mu^0, 2\mu^1, \cdots, p\mu^{p-1}, \hat{E}_p], \quad Q = [\mu^1, \mu^2, \cdots, \mu^p, B^{-1}\eta]. \tag{3.8}$$

Since V is evidently nonsincular, the matrix A can be uniquely determined by (3.7),(3.8) and (3.6).

In view of Proposition 3.1, for any given natural numbers s, p, l satisfying $1 \le p \le s, 0 \le l \le 2, s \le 2p+l \le 2s$, the relations (3.3)—(3.8) and (3.2a) thus determine a class of s stage simplectic Runge-Kutta methods of order at least $2p+l$ with $\frac{1}{2}(s-p)(s-p+3)-l$ free parameters μ_{i_m}, $m = 1, 2, \cdots, 2(s-p)-l, and\ \alpha_{ij}$, $i = p+1, \cdots, s-1$, $j = i+1, \cdots, s$.

It should be pointed out that Geng Sun [8] has presented techniques based on the W-transformation of Hairer and Wanner, by means of which he has also constructed two classes of high order symplectic Runge-Kutta methods.

Example 3.1. It is easily seen from Corollary 2.6 and Theorem 1.1 that the one stage symplectic Runge-Kutta method

$$\begin{array}{c|c} \mu & \frac{1}{2} \\ \hline & 1 \end{array}$$

is of order 2 for $\mu = \frac{1}{2}$, of order 1 for μ taking any other values.

Example 3.2. It is easily seen that the 2-stage Runge-Kutta method

$$\begin{array}{c|cc} \mu_1 & a_{11} & a_{12} \\ \mu_2 & a_{21} & a_{22} \\ \hline & \gamma_1 & \gamma_2 \end{array} \quad , \tag{3.9}$$

with $\gamma_1, \gamma_2 \neq 0$ is consistent, irreducible, non-confluent and symplectic if and only if the following conditions are satisfied:

$$\begin{cases} \gamma_1 + \gamma_2 = 1, \ \ \gamma_1, \gamma_2 \neq 0, \ \ \mu_1 \neq \mu_2, \\ \gamma_1 = 2a_{11}, \ \ \gamma_2 = 2a_{22}, \ \ \gamma_1 a_{12} + \gamma_2 a_{21} = \gamma_1 \gamma_2. \end{cases} \tag{3.10}$$

The simplifying condition $C(1)$:

$$\mu_1 = a_{11} + a_{12}, \qquad \mu_2 = a_{21} + a_{22} \tag{3.11}$$

together with (3.10) is evidently equivalent to

$$\begin{cases} a_{11} = \frac{1-2\mu_2}{4(\mu_1-\mu_2)}, \quad a_{12} = \mu_1 - \frac{1-2\mu_2}{4(\mu_1-\mu_2)}, \\ a_{21} = \frac{(1-2\mu_2)(1-4\mu_1+2\mu_2+4\mu_1^2-4\mu_1\mu_2)}{4(1-2\mu_1)(\mu_1-\mu_2)}, \quad a_{22} = \frac{2\mu_1-1}{4(\mu_1-\mu_2)}, \\ \gamma_1 = 2a_{11}, \quad \gamma_2 = 2a_{22}, \quad \mu_1, \mu_2 \neq \frac{1}{2}, \ \ \mu_1 \neq \mu_2. \end{cases} \tag{3.12}$$

The relations (3.9)-(3.12) can be regarded as a class of 2 stage symplectic Runge-Kutta methods with μ_1, μ_2 being two free parameters, and in view of Corollary 2.6 methods of this class are all of order at least 2. For example, letting $\mu_1 = 1/4, \mu_2 = 3/4$, we get the 2nd order diagonally implicit symplectic Runge-Kutta method

$$\begin{array}{c|cc} \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \ .$$

For the methods of aforementioned class to satisfy simplifying condition $B(3)$, it is necessary and sufficient that

$$\int_0^1 (x-\mu_1)(x-\mu_2)dx = 0,$$

or equivalently

$$\mu_2 = \frac{2-3\mu_1}{3(1-2\mu_1)}, \quad \mu_1 \neq \frac{1}{2}. \tag{3.13}$$

In view of Corollary 2.6 the relations (3.9)-(3.12)-(3.13) with free parameter μ_1 determine a class of symplectic Runge-Kutta methods of order at least 3. For example, by letting $\mu_1 = 1$, (3.9)-(3.12)-(3.13) yields the 3rd order symplectic Runge-Kutta method

$$\begin{array}{c|cc} 1 & \frac{1}{8} & \frac{7}{8} \\ \frac{1}{3} & \frac{-1}{24} & \frac{3}{8} \\ \hline & \frac{1}{4} & \frac{3}{4} \end{array} .$$

Especially, letting $\mu_1 = 1/2 + \sqrt{3}/6$, a root of the shift Legendre polynomial $6x^2 - 6x + 1$, we get the well known 2 stage Gauss-Legendre Runge-Kutta method, which is symplectic and of order 4.

Example 3.3. To construct 3 stage symplectic Runge-Kutta methods of order at least 4, we would use the general approach based on Proposition 3.1 with $s = 3$ and $p = 2$. For $l = 2$, we have $2p + l = 2s$, and (3.3) leads to

$$p(x) = \prod_{k=1}^{s} (x - \mu_k) = \frac{1}{20}(20x^3 - 30x^2 + 12x - 1)$$

and

$$\mu_1 = \frac{1}{2} - \frac{\sqrt{15}}{10}, \ \ \mu_2 = \frac{1}{2}, \quad \mu_3 = \frac{1}{2} + \frac{\sqrt{15}}{10}, \tag{3.14}$$

for $l = 1$, we have $2p + l = 2s - 1$, (3.4)-(3.5) with $q = 1$ leads to

$$p(x) = (x - \mu_1)(x^2 - \frac{2(15\mu_1^2 - 16\mu_1 + 3)}{5(6\mu_1^2 - 6\mu_1 + 1)}x + \frac{10\mu_1^2 - 12\mu_1 + 3}{10(6\mu_1^2 - 6\mu_1 + 1)})$$

and

$$\mu_{2,3} = \frac{15\mu_1^2 - 16\mu_1 + 3}{5(6\mu_1^2 - 6\mu_1 + 1)} \pm \frac{\sqrt{300\mu_1^4 - 600\mu_1^3 + 384\mu_1^2 - 84\mu_1 + 6}}{10(6\mu_1^2 - 6\mu_1 + 1)}, \tag{3.15}$$

and for $l = 0, 2p + l = 2s - 2$, (3.4)-(3.5) with $q = 2$ leads to

$$p(x) = (x - \mu_1)(x - \mu_2)(x - \frac{6\mu_1\mu_2 - 4(\mu_1 + \mu_2) + 3}{2(6\mu_1\mu_2 - 3(\mu_1 + \mu_2) + 2)})$$

and

$$\mu_3 = \frac{6\mu_1\mu_2 - 4(\mu_1 + \mu_2) + 3}{2(6\mu_1\mu_2 - 3(\mu_1 + \mu_2) + 2)}. \tag{3.16}$$

From (3.2a) we get

$$\begin{cases} \gamma_1 + \gamma_2 + \gamma_3 = 1, \\ \mu_1\gamma_1 + \mu_2\gamma_2 + \mu_3\gamma_3 = \frac{1}{2}, \\ \mu_1^2\gamma_1 + \mu_2^2\gamma_2 + \mu_3^2\gamma_3 = \frac{1}{3}, \end{cases}$$

and therefore

$$\begin{cases} \gamma_1 = \frac{1}{\Delta}[\mu_2\mu_3(\mu_3-\mu_2) - \frac{1}{2}(\mu_3^2-\mu_2^2) + \frac{1}{3}(\mu_3-\mu_2)], \\ \gamma_2 = \frac{1}{\Delta}[\mu_3\mu_1(\mu_1-\mu_3) - \frac{1}{2}(\mu_1^2-\mu_3^2) + \frac{1}{3}(\mu_1-\mu_3)], \\ \gamma_3 = \frac{1}{\Delta}[\mu_1\mu_2(\mu_2-\mu_1) - \frac{1}{2}(\mu_2^2-\mu_1^2) + \frac{1}{3}(\mu_2-\mu_1)], \\ \Delta = (\mu_3-\mu_2)(\mu_3-\mu_1)(\mu_2-\mu_1). \end{cases} \tag{3.17}$$

From (3.6),(3.7) and (3.8) we get

$$\begin{cases} V^{-1} = \frac{1}{\mu_2-\mu_1}\begin{bmatrix} \mu_2 & -\mu_1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ \mu_3-\mu_2 & \mu_1-\mu_3 & \mu_2-\mu_1 \end{bmatrix}, \\ B^{-1}\eta = \begin{bmatrix} \delta_{13} \\ \delta_{23} \\ \frac{\gamma_3}{2} \end{bmatrix}, \\ Q = \begin{bmatrix} \mu_1 & \mu_1^2 & \delta_{13} \\ \mu_2 & \mu_2^2 & \delta_{23} \\ \mu_3 & \mu_3^2 & \frac{\gamma_3}{2} \end{bmatrix}, \end{cases}$$

where

$$\begin{cases} \delta_{13} = \frac{\gamma_3}{2\gamma_1(\mu_2-\mu_1)}[(1-\mu_3)(2\mu_2-\mu_3-1) + \gamma_3(\mu_3-\mu_2)], \\ \delta_{23} = \frac{\gamma_3}{2\gamma_2(\mu_2-\mu_1)}[(1-\mu_3)(-2\mu_1+\mu_3+1) + \gamma_3(\mu_1-\mu_3)], \end{cases} \tag{3.18}$$

and therefore

$$A = QV^{-1} = \frac{1}{\mu_2-\mu_1}\begin{bmatrix} \mu_1 & \mu_1^2 & \delta_{13} \\ \mu_2 & \mu_2^2 & \delta_{23} \\ \mu_3 & \mu_3^2 & \frac{\gamma_3}{2} \end{bmatrix}\begin{bmatrix} \mu_2 & -\mu_1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ \mu_3-\mu_2 & \mu_1-\mu_3 & \mu_2-\mu_1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{2\mu_1\mu_2-\mu_1^2+2(\mu_3-\mu_2)\delta_{13}}{2(\mu_2-\mu_1)} & \frac{-\mu_1^2+2(\mu_1-\mu_3)\delta_{13}}{2(\mu_2-\mu_1)} & \delta_{13} \\ \frac{\mu_2^2+2(\mu_3-\mu_2)\delta_{23}}{2(\mu_2-\mu_1)} & \frac{-2\mu_1\mu_2+\mu_2^2+2(\mu_1-\mu_3)\delta_{23}}{2(\mu_2-\mu_1)} & \delta_{23} \\ \frac{2\mu_2\mu_3-\mu_3^2+\gamma_3(\mu_3-\mu_2)}{2(\mu_2-\mu_1)} & \frac{-2\mu_1\mu_3+\mu_3^2+\gamma_3(\mu_1-\mu_3)}{2(\mu_2-\mu_1)} & \frac{\gamma_3}{2} \end{bmatrix}. \tag{3.19}$$

Thus, for $l = 0$, (3.16)-(3.17)-(3.18)-(3.19) determines a class of 3 stage 4-th order symplectic Runge-Kutta methods with two free parameters μ_1 and μ_2 satisfying $6\mu_1\mu_2 - 3(\mu_1+\mu_2)+2 \neq 0$, as an example, letting $\mu_1 = 0, \mu_2 = 1$, we get the 3 stage Runge-Kutta method

$$\begin{array}{c|ccc} 0 & \frac{1}{12} & \frac{1}{12} & -\frac{1}{6} \\ 1 & \frac{1}{12} & \frac{1}{12} & \frac{5}{6} \\ \frac{1}{2} & \frac{5}{24} & -\frac{1}{24} & \frac{1}{3} \\ \hline & \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{array}$$

which is symplectic and of order 4. For $l = 1$, (3.15)-(3.17)-(3.18)-(3.19) determines a class of 3 stage 5-th order symplectic Runge-Kutta methods with one free parameter μ_1 satisfying $6\mu_1^2 - 6\mu_1 + 1 \neq 0$, as an example, letting $\mu_1 = 0$, we get the Radau1B Runge-Kutta method

$$\begin{array}{c|ccc} 0 & \frac{1}{18} & \frac{-1-\sqrt{6}}{36} & \frac{-1+\sqrt{6}}{36} \\ \frac{6-\sqrt{6}}{10} & \frac{52+3\sqrt{6}}{450} & \frac{16+\sqrt{6}}{72} & \frac{472-217\sqrt{6}}{1800} \\ \frac{6+\sqrt{6}}{10} & \frac{52-3\sqrt{6}}{450} & \frac{472+217\sqrt{6}}{1800} & \frac{16-\sqrt{6}}{72} \\ \hline & \frac{1}{9} & \frac{16+\sqrt{6}}{36} & \frac{16-\sqrt{6}}{36} \end{array}$$

which is well known to be symplectic and of order 5 (cf. [8]). At last, we note that for the case of $l = 2$, (3.14)-(3.17)-(3.18)-(3.19) leads to the 3 stage Gauss-Legendre Runge-Kutta method, which is symplectic and of order 6.

Recently, based on the formulas (3.2a),(3.3)–(3.8), we have designed a software, by means of which, the coefficients of s stage $(2p + l)$-th order symplectic Runge-Kutta methods (with $\frac{1}{2}(s-p)(s-p+3)-l$ free parameters) satisfying $C(p), D(p)$ and $B(2p+l)$ can be easily computed, where $1 \leq p \leq s, 0 \leq l \leq 2, s \leq 2p + l \leq 2s, s = 1, 2, 3, \cdots$.

References

[1] J.C. Butcher, On Runge-Kutta Processes of High Order, *J. Austral. Math. Soc.,* **4** (1964), Part 2, 179–194.

[2] J.M. Sanz-Serna, Runge-Kutta Schemes for Hamiltonian Systems, *BIT,* **28** (1988), 877–883.

[3] F. Lasagni, Canonical Runge-Kutta Methods, *Z. Angow. Math. Phys.,* **39** (1988), 952–953.

[4] Aiguo Xiao & Shoufu Li, Characterization and Construction of Symplectic Runge-Kutta Methods, *Numer. Math., J. Chin. Univ.,* **17**:3 (1995), 213–222.

[5] J.C. Butcher, The Numerical Analysis of Ordinary Differential Equations, John Wiley & Sons Ltd, 1987.

[6] E. Hairer, S.p. Nørsett & G. wanner, Solving Ordinary Differential Equations I, 2nd Edition, Springer-Verlag Berlin Heidelberg, 1993.

[7] Shoufu Li, Theory of Computational Methods for Stiff Differential Equations, Hunan Sci. Tech. Pub., 1997.

[8] Geng Sun, Construction of High Order Symplectic Runge-Kutta Methods, *J. Comput. Math.,* **3** (1993), 250–260.