TECHNICAL RESEARCH REPORT

Almost Symplectic Runge-Kutta Schemes for Hamiltonian Systems

by Xiaobo Tan

CDCSS TR 2002-7
(ISR TR 2002-51)

CENTER FOR DYNAMICS AND CONTROL OF SMART STRUCTURES

The Center for Dynamics and Control of Smart Structures (CDCSS) is a joint Harvard University, Boston University, University of Maryland center, supported by the Army Research Office under the ODDR&E MURI97 Program Grant No. DAAG55-97-1-0114 (through Harvard University). This document is a technical report in the CDCSS series originating at the University of Maryland.

Web site http://www.isr.umd.edu/CDCSS/cdcss.html

ALMOST SYMPLECTIC RUNGE-KUTTA SCHEMES FOR HAMILTONIAN SYSTEMS

Xiaobo Tan

Institute for Systems Research, University of Maryland

College Park, MD 20742, USA. email: xbtan@isr.umd.edu

Abstract.

Symplectic Runge-Kutta schemes for integration of general Hamiltonian systems are implicit. In practice the implicit equations are often approximately solved based on the Contraction Mapping Principle, in which case the resulting integration scheme is no longer symplectic. In this note we prove that, under suitable conditions, the integration scheme based on an n-step successive approximation is $\mathcal{O}(\delta^{n+2})$ away from a symplectic scheme with $\delta \in (0,1)$. Therefore, this scheme is "almost" symplectic when n is large.

AMS subject classification: 65L05, 65L06, 65P10.

Key words: Hamiltonian systems, symplectic Runge-Kutta methods, Contraction Mapping Principle, successive approximation.

1 Introduction.

Geometric integration methods, numerical methods that preserve geometric properties of the flow of a differential equation, outperform the off-the-shelf schemes (e.g., fourth order explicit Runge-Kutta method) in predicting the long-term qualitative behaviors of the original system [5]. An important class of geometric integrators are symplectic integration methods for Hamiltonian systems [9]. Consider a Hamiltonian system

$$\left\{ \begin{array}{l} \dot{p}(t) = -\frac{\partial H(p,q)}{\partial q} \\ \dot{q}(t) = \frac{\partial H(p,q)}{\partial p} \end{array} \right. , \tag{1.1}$$

with the Hamiltonian $H(p,q)$, where $(p,q) \in \mathbb{R}^d \times Q$ for some integer $d \geq 1$, and Q, the configuration space, is some d-dimensional manifold. For ease of discussion, in this note we assume $Q = \mathbb{R}^d$, but the results we present here apply to the case of a general Q directly. Let $z = (p, q)$, the system (1.1) can be rewritten as:

$$\dot{z}(t) = f(z(t)) \overset{\triangle}{=} J\nabla_z H(z(t)), \tag{1.2}$$

where

$$J = \begin{bmatrix} 0 & -I_d \\ I_d & 0 \end{bmatrix},$$

I_d denotes the d-dimensional identity matrix, and ∇_z stands for the gradient with respect to z.

When the Hamiltonian has a seperable structure, i.e., $H(q, p) = T(p) + V(q)$, explicit Runge-Kutta type algorithms exist which preserve the symplectic structure [4, 11, 3, 7]. However, this is not the case for general Hamiltonian systems.

An s-stage Runge-Kutta method to integrate (1.2) is as follows [6]:

$$\begin{cases} y_i = z_0 + \tau \sum_{j=1}^{s} a_{ij} f(y_j), \ i = 1, \cdots, s \\ z_1 = z_0 + \tau \sum_{i=1}^{s} b_i f(y_i) \end{cases}, \tag{1.3}$$

where τ is the time step, z_0 is the initial value at time t_0, z_1 is the numerical solution at time $t_0 + \tau$, a_{ij}, b_i are appropriate coefficients satisfying the order conditions of the Runge-Kutta method.

Let Ψ_τ be the one time-step flow associated with the algorithm (1.3), i.e., $z_1 = \Psi_\tau(z_0)$. From [8], the transformation Ψ_τ preserves the symplecticness of the original system (1.2) if

$$b_i a_{ij} + b_j a_{ji} - b_i b_j = 0, \ i, j = 1, \cdots, s. \tag{1.4}$$

Thus if (1.4) is satisfied, we have:

$$(\frac{\partial \Psi_\tau}{\partial z_0})^T J (\frac{\partial \Psi_\tau}{\partial z_0}) - J = 0, \tag{1.5}$$

where ^{T} denotes the transpose. The condition (1.4) forces the symplectic Runge-Kutta method (1.3) to be implicit. In the interest of computation efficiency, Aubry and Chartier investigated pseudo-symplectic Runge-Kutta methods, which are explicit and conserve the symplectic structure to a certain order [1]. We also note the closely related work in [2], where the error estimate for the Lie-Poisson structure is established for integration of Lie-Poisson systems using the mid-point rule.

In this note, we take a different approach from [1]. Successive approximation based upon the Contraction Mapping Principle is often used to obtain an approximate solution to y_i in (1.3). The resulting integration scheme based on the approximation is no longer symplectic. It's of interest to investigate, to what extent, the symplectic structure (1.5) is preserved by the approximation scheme. The rest of this note is devoted to answering this question, and it turns out that the scheme using an n-step approximation is $\mathcal{O}(\delta^{n+2})$ away from a symplectic one with $0 < \delta < 1$. Therefore, when n is large enough, the approximation scheme is "almost" symplectic.

2 A successive approximation method.

Denote

$$\mathbf{y} \triangleq \begin{pmatrix} y_1 \\ \vdots \\ y_s \end{pmatrix}, \quad \mathbf{F}(\mathbf{y}) \triangleq \begin{pmatrix} f(y_1) \\ \vdots \\ f(y_s) \end{pmatrix},$$

$\mathbf{b} = (b_1, \cdots, b_s)$, $A_0 = [a_{ij}]$, and $A = A_0 \otimes I_{2d}$, where "$\otimes$" denotes the Kronecker (tensor) product. We recall for two matrices $M = [m_{ij}]$ and $R = [r_{ij}]$, the Kronecker product

$$M \otimes R = \begin{bmatrix} m_{11}R & m_{12}R & \cdots \\ m_{21}R & m_{22}R & \cdots \\ \vdots & \vdots & \vdots \end{bmatrix}.$$

The algorithm (1.3) can now be written as

$$\left\{ \begin{array}{l} \mathbf{y} = \mathbf{G}(z_0, \mathbf{y}) \triangleq \mathbf{1} \otimes z_0 + \tau A \mathbf{F}(\mathbf{y}) \\ z_1 = z_0 + \tau \mathbf{b} \otimes I_{2d} \mathbf{F}(\mathbf{y}) \end{array} \right. , \tag{2.1}$$

where $\mathbf{1}$ is an s-dimensional column vector with 1 in every entry.

As noted in Section 1, when (1.4) is satisfied, the first equation in (2.1) is implicit for each fixed z_0. One algorithm often used to solve implicit equations, is the successive approximation scheme based on the Contraction Mapping Principle (see, e.g., [10]):

LEMMA 2.1 (CONTRACTION MAPPING PRINCIPLE). *Let S be a closed subset of a Banach space X and let φ be a mapping that maps S into S. If $\exists \rho \in (0, 1)$, such that*

$$\|\varphi(x) - \varphi(y)\| \leq \rho \|x - y\|, \forall x, y \in S,$$

then

1. *there exists a unique* $x^* \in S$ *satisfying* $x^* = \varphi(x^*)$;
2. x^* *can be obtained by the method of successive approximation*

$$x^{[n+1]} = \varphi(x^{[n]}),$$

   starting from an arbitrary $x^{[0]}$ *in* S; *and*

3. *the approximation error satisfies* $\|x^{[n]} - x^*\| \leq \rho^n \|x^{[0]} - x^*\|$.

In this note we will use $\|\cdot\|$ to denote the norm (or the induced norm) of a vector, matrix, or high order tensors, and the precise meaning should be clear from the context. The following proposition shows that when the step size τ is small enough, for each fixed z_0, the first equation in (2.1) has a unique solution $\mathbf{y}^*$:

Proposition 2.2. *Let* $\Omega \subset \mathbb{R}^{2d}$ *be a bounded open set. Let* f *be locally Lipschitz continuous. Then for any* $\delta \in (0,1)$, $\epsilon > 0$, *there exists* $\tau(\Omega, \epsilon, \delta) > 0$ *dependent on* Ω, ϵ *and* δ, *such that,* $\forall \tau \leq \tau(\Omega, \epsilon, \delta)$, $\forall z_0 \in \Omega$,

1. *there exists a unique solution* $\mathbf{y}^* = \mathbf{y}^*(z_0)$ *for the first equation in (2.1);*
2. $\mathbf{y}^*$ *can be approximated by successive approximation*

$$\left\{ \begin{array}{l} \mathbf{y}^{[n]} = \mathbf{G}(z_0, \mathbf{y}^{[n-1]}) \\ \mathbf{y}^{[0]} = \mathbf{1} \otimes z_0 \end{array} \right. ; \tag{2.2}$$

   and

3. $\|\mathbf{y}^{[n]} - \mathbf{y}^*\| \leq \delta^n \|\mathbf{y}^{[0]} - \mathbf{y}^*\|$.

Proof. Denote $N(\Omega, \epsilon)$ the $\epsilon-$neighbourhood of Ω, defined as

$$N(\Omega, \epsilon) \stackrel{\triangle}{=} \{z \in \mathbb{R}^{2d} : \min_{z_0 \in \bar{\Omega}} \|z - z_0\| \leq \epsilon\},$$

where $\bar{\Omega}$ denotes the closure of Ω. Denote $N^s(\Omega, \epsilon)$ the product of s copies of $N(\Omega, \epsilon)$, i.e.,

$$N^s(\Omega, \epsilon) = N(\Omega, \epsilon) \times \cdots \times N(\Omega, \epsilon).$$

Since $N(\Omega, \epsilon)$ is compact, f is bounded and Lipschitz continuous with some Lipschitz constant L_f on $N(\Omega, \epsilon)$. Thus there exists $\tau_1 > 0$, such that when $\tau \leq \tau_1$, for each fixed $z_0 \in \Omega$, $\mathbf{G}(z_0, \cdot)$ maps $N^s(\Omega, \epsilon)$ into itself.

For any $z_0 \in \Omega$, for $\mathbf{y}, \mathbf{y}' \in N^s(\Omega, \epsilon)$, by the definition of $\mathbf{G}$,

$$\begin{aligned} \|\mathbf{G}(z_0, \mathbf{y}) - \mathbf{G}(z_0, \mathbf{y}')\| &= \|\tau A \begin{pmatrix} f(y_1) - f(y_1') \\ \vdots \\ f(y_s) - f(y_s') \end{pmatrix}\| \\ &\leq \tau L_f \|A\| \, \|\mathbf{y} - \mathbf{y}'\|. \end{aligned}$$

For $\delta \in (0,1)$, let $\tau_2 = \frac{\delta}{L_f \|A\|}$. Now for $\tau \leq \tau(\Omega, \epsilon, \delta) \stackrel{\triangle}{=} \min\{\tau_1, \tau_2\}$, $\mathbf{G}(z_0, \cdot)$ is a contraction mapping for each fixed $z_0 \in \Omega$. All the claims then follow from Lemma 2.1. Note that $\tau(\Omega, \epsilon, \delta)$ depends on Ω, ϵ and δ. $\square$

Similarly we can prove:

PROPOSITION 2.3. *Let f be globally bounded and Lipschitz continuous. Then for any $\delta \in (0,1)$, there exists $\tau(\delta) > 0$ dependent on δ only, such that, $\forall \tau \leq \tau(\delta)$, for each fixed $z_0 \in \mathbb{R}^{2d}$, $\mathbf{G}(z_0, \cdot)$ is a contraction mapping and the claims in Proposition 2.2 hold.*

REMARK 2.1. *Traditionally implicit Runge-Kutta methods have been used mostly for stiff problems, where the Lipschitz constant for f is relatively large and the convergence of successive approximation based on the Contraction Mapping Principle is slow. However, in the new context of symplectic integration, we are dealing with implicit methods even for nonstiff problems. Hence the successive approximation plays an important role in solving the implicit equations.*

As we see from Proposition 2.2, when τ is sufficiently small, the solution $\mathbf{y}^*$ to the first equation in (2.1) is a function of z_0, and we can write it as $\mathbf{y}^*(z_0)$. If f is differentiable, we have from the Implicit Function Theorem that

$$\frac{\partial \mathbf{y}^*}{\partial z_0}(z_0) = [I_{2sd} - \tau A \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^*(z_0))]^{-1}(\mathbf{1} \otimes I_{2d}). \tag{2.3}$$

3 Main result.

An explicit but approximate algorithm to solve (2.1) is as follows: for some $n \geq 0$,

$$\left\{ \begin{array}{l} \mathbf{y}^{[k]} = \mathbf{G}(z_0, \mathbf{y}^{[k-1]}),\ k = 1, \cdots, n \\ \mathbf{y}^{[0]} = \mathbf{1} \otimes z_0 \\ z_1^{[n]} = z_0 + \tau \mathbf{b} \otimes I_{2d} \mathbf{F}(\mathbf{y}^{[n]}) \end{array} \right. . \tag{3.1}$$

REMARK 3.1. *The scheme (3.1) based on $n-$step successive approximation (to $\mathbf{y}^*$) is essentially an $s(n+1)$-stage explicit Runge-Kutta scheme with coefficients*

$\tilde{A}$ and $\tilde{\mathbf{b}}$, where

$$\tilde{A} = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \otimes A_0, \quad \tilde{\mathbf{b}} = (0, \cdots, 0, b_1, \cdots, b_s).$$

Note that in (2.1) and (3.1), $\mathbf{y}^*$, $\{\mathbf{y}^{[k]}\}_{k=0}^n$, z_1 and $z_1^{[n]}$ (and smooth functions of them) are all continuously differentiable functions of z_0 if f is differentiable and τ is sufficiently small. In the sequel when we write, e.g., $\frac{\partial \mathbf{y}^*}{\partial z_0}$ or $\frac{\partial}{\partial z_0}\mathbf{F}(\mathbf{y}^{[n]})$, *we think of $\mathbf{y}^*$ or $\mathbf{F}(\mathbf{y}^{[n]})$ as a function of z_0 although it is not explicitly written out.*

Denote by $\Psi_\tau^{[n]}$ the one time-step flow associated with the algorithm (3.1), i.e., $z_1^{[n]} = \Psi_\tau^{[n]}(z_0)$. We now want to study how far $\Psi_\tau^{[n]}$ is away from a symplectic transformation. The following lemma will be essential in the proof of our main result Theorem 3.3:

Lemma 3.1. *Let $\Omega \subset \mathbb{R}^{2d}$ be bounded, convex and open. For $\epsilon > 0$, let $N(\Omega, \epsilon)$ be the ϵ-neighbourhood of Ω, as defined in the proof of Proposition 2.2. Assume that f is twice continuously differentiable on $N(\Omega, \epsilon)$. Then for any $\delta \in (0, 1)$, there exists $\tau(\Omega, \epsilon, \delta) > 0$ dependent on Ω, ϵ and δ, such that when $\tau \leq \tau(\Omega, \epsilon, \delta)$, for each fixed $z_0 \in \Omega$, the first equations in (2.1) and (3.1) have (unique) solutions $\mathbf{y}^* \in N^s(\Omega, \epsilon)$ and $\mathbf{y}^{[n]} \in N^s(\Omega, \epsilon)$, respectively; and*

$$\left\|\frac{\partial \mathbf{y}^{[n]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0}\right\| \leq C(\Omega, \epsilon)\delta^{n+1}, \tag{3.2}$$

$$\left\|\frac{\partial}{\partial z_0}(\mathbf{F}(\mathbf{y}^{[n]}) - \mathbf{F}(\mathbf{y}^*))\right\| \leq C'(\Omega, \epsilon)\delta^{n+1}, \tag{3.3}$$

where $C(\Omega, \epsilon), C'(\Omega, \epsilon) > 0$ are constants dependent only on Ω and ϵ.

Proof. Since f is differentiable, it is Lipschitz continuous on the convex set $N(\Omega, \epsilon)$. By Proposition 2.2, there exists $\tau_1(\Omega, \epsilon, \delta) > 0$, such that when $\tau \leq \tau_1(\Omega, \epsilon, \delta)$, for any $z_0 \in \Omega$, $\mathbf{G}(z_0, \cdot)$ is a contraction mapping, $\mathbf{y}^*, \mathbf{y}^{[k]} \in N^s(\Omega, \epsilon)$, $\forall k \geq 0$, and (recall (2.3))

$$\left\|\frac{\partial \mathbf{y}^*}{\partial z_0}\right\| \leq C_1(\Omega, \epsilon), \tag{3.4}$$

$$\left\|\frac{\partial \mathbf{y}^{[0]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0}\right\| = \left\|\tau A \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^*)\frac{\partial \mathbf{y}^*}{\partial z_0}\right\| \leq \tau C_2(\Omega, \epsilon), \tag{3.5}$$

where $C_i(\Omega,\epsilon) > 0$, $i = 1, 2$, are constants dependent on Ω, ϵ.

From (2.1) and (3.1),

$$\mathbf{y}^{[n]} - \mathbf{y}^* = \tau A(\mathbf{F}(\mathbf{y}^{[n-1]}) - \mathbf{F}(\mathbf{y}^*)). \tag{3.6}$$

Taking derivative of both sides of (3.6) with respect to z_0 and re-arranging terms, we get

$$\begin{aligned} &\frac{\partial \mathbf{y}^{[n]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0} \\ = \;& \tau A[\frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^{[n-1]})(\frac{\partial \mathbf{y}^{[n-1]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0}) + (\frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^{[n-1]}) - \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^*))\frac{\partial \mathbf{y}^*}{\partial z_0}]. \end{aligned} \tag{3.7}$$

Denoting

$$\Theta^{[k]} \triangleq \tau A \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^{[k]}),\ \Gamma^{[k]} \triangleq \tau A(\frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^{[k]}) - \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^*))\frac{\partial \mathbf{y}^*}{\partial z_0},$$

we derive from (3.7)

$$\frac{\partial \mathbf{y}^{[n]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0} = (\prod_{k=0}^{n-1} \Theta^{[k]})(\frac{\partial \mathbf{y}^{[0]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0}) + \sum_{k=0}^{n-1}(\prod_{i=k+1}^{n-1} \Theta^{[i]})\Gamma^{[k]},$$

which implies

$$\|\frac{\partial \mathbf{y}^{[n]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0}\| \le (\prod_{k=0}^{n-1} \|\Theta^{[k]}\|)\|\frac{\partial \mathbf{y}^{[0]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0}\| + \sum_{k=0}^{n-1}(\prod_{i=k+1}^{n-1} \|\Theta^{[i]}\|)\|\Gamma^{[k]}\|. \tag{3.8}$$

The following two observations are in order:

1. $$\|\Theta^{[k]}\| \le \tau \|A\| \max_{\mathbf{y} \in N^s(\Omega,\epsilon)} \|\frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y})\| \le \tau C_3(\Omega,\epsilon), \tag{3.9}$$

   where $C_3(\Omega,\epsilon) > 0$ is a constant dependent only on Ω and ϵ.

2. When $\tau \le \tau_1(\Omega,\epsilon,\delta)$,

   $$\begin{aligned} \|\frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^{[k]}) - \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^*)\| &\le \max_{\mathbf{y} \in N^s(\Omega,\epsilon)} \|\frac{\partial^2 \mathbf{F}}{\partial \mathbf{y}^2}(\mathbf{y})\| \, \|\mathbf{y}^{[k]} - \mathbf{y}^*\| \\ &\le C_4(\Omega,\epsilon)\delta^k \|\mathbf{y}^{[0]} - \mathbf{y}^*\|, \end{aligned} \tag{3.10}$$

   where $C_4(\Omega,\epsilon) > 0$ is a constant dependent only on Ω and ϵ. Combining (3.4) and (3.10), and using

   $$\|\mathbf{y}^{[0]} - \mathbf{y}^*\| \le \tau \|A\| \max_{\mathbf{y} \in N^s(\Omega,\epsilon)} \|\mathbf{F}(\mathbf{y})\|, \tag{3.11}$$

we have

$$\|\Gamma^{[k]}\| \leq \tau^2 \delta^k C_5(\Omega, \epsilon), \tag{3.12}$$

for some constant $C_5(\Omega, \epsilon) > 0$.

Pluggin (3.5), (3.9) and (3.12) into (3.8), we obtain after some manipulations

$$\|\frac{\partial \mathbf{y}^{[n]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0}\| \leq \tau C_2(\Omega, \epsilon)(\tau C_3(\Omega, \epsilon))^n + \frac{\tau^2 \delta^{n-1} C_5(\Omega, \epsilon)}{1 - \frac{\tau C_3(\Omega, \epsilon)}{\delta}}.$$

We now let $\tau_2(\Omega, \epsilon, \delta) \triangleq \frac{\delta}{2C_3(\Omega, \epsilon)}$, and let

$$\tau(\Omega, \epsilon, \delta) = \min\{\tau_1(\Omega, \epsilon, \delta), \tau_2(\Omega, \epsilon, \delta)\}.$$

It's easy to verify that, $\forall \tau \leq \tau(\Omega, \epsilon, \delta)$,

$$\|\frac{\partial \mathbf{y}^{[n]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0}\| \leq C(\Omega, \epsilon)\delta^{n+1},$$

where $C(\Omega, \epsilon) \triangleq \frac{C_2(\Omega, \epsilon)}{2C_3(\Omega, \epsilon)} + \frac{C_5(\Omega, \epsilon)}{2C_3^2(\Omega, \epsilon)}$. This proves (3.2).

To show (3.3), we note that

$$\frac{\partial}{\partial z_0}(\mathbf{F}(\mathbf{y}^{[n]}) - \mathbf{F}(\mathbf{y}^*)) = \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^{[n]})(\frac{\partial \mathbf{y}^{[n]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0}) + (\frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^{[n]}) - \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^*))\frac{\partial \mathbf{y}^*}{\partial z_0},$$

and then use (3.2), (3.4), (3.10) and (3.11). $\square$

Similarly we can prove:

LEMMA 3.2. *Let f be globally bounded and twice continuously differentiable, with bounded first order and second order derivatives. Then for any $\delta \in (0, 1)$, there exists $\tau(\delta) > 0$ dependent on δ only, such that when $\tau \leq \tau(\delta)$, for each fixed $z_0 \in \mathbb{R}^{2d}$, the first equations in (2.1) and (3.1) have (unique) solutions $\mathbf{y}^*$ and $\mathbf{y}^{[n]}$, respectively; and*

$$\|\frac{\partial \mathbf{y}^{[n]}}{\partial z_0} - \frac{\partial \mathbf{y}^*}{\partial z_0}\| \leq C\delta^{n+1}, \tag{3.13}$$

$$\|\frac{\partial}{\partial z_0}(\mathbf{F}(\mathbf{y}^{[n]}) - \mathbf{F}(\mathbf{y}^*))\| \leq C'\delta^{n+1}, \tag{3.14}$$

for some constants $C, C' > 0$.

We are now ready to present the main result of this note:

THEOREM 3.3. *Let $\Omega \subset \mathbb{R}^{2d}$ be bounded, convex and open. For $\epsilon > 0$, let $N(\Omega, \epsilon)$ be the ϵ-neighbourhood of Ω. Assume that f is twice continuously* *differentiable on* $N(\Omega,\epsilon)$*. Consider the algorithm (3.1) and let* $\Psi_\tau^{[n]}$ *be the one time-step flow associated with (3.1). Let (1.4) be satisfied. Then for any* $\delta \in (0,1)$*, there exists* $\tau(\Omega,\epsilon,\delta) > 0$ *dependent on* Ω, ϵ *and* δ*, such that when* $\tau \leq \tau(\Omega,\epsilon,\delta)$*,*

$$\|(\frac{\partial \Psi_\tau^{[n]}(z_0)}{\partial z_0})^T J(\frac{\partial \Psi_\tau^{[n]}(z_0)}{\partial z_0}) - J\| \leq C(\Omega,\epsilon)\delta^{n+2}, \quad \forall z_0 \in \Omega, \tag{3.15}$$

where $C(\Omega,\epsilon)$ *is a constant dependent on* Ω *and* ϵ*.*

Proof. By Lemma 3.1, we can find $\tau(\Omega,\epsilon,\delta) > 0$, such that when $\tau \leq \tau(\Omega,\epsilon,\delta)$,

$$\|\frac{\partial}{\partial z_0}(\mathbf{F}(\mathbf{y}^{[n]}) - \mathbf{F}(\mathbf{y}^*))\| \leq C_1(\Omega,\epsilon)\delta^{n+1}, \tag{3.16}$$

for some constant $C_1(\Omega,\epsilon)$, where $\mathbf{y}^*$ and $\mathbf{y}^{[n]}$ are solutions to the first equations in (2.1) and (3.1), respectively. Let Ψ_τ be the one time-step flow associated with (2.1). From (2.1) and (3.1), we have

$$\Lambda^{[n]}(z_0) \stackrel{\triangle}{=} \Psi_\tau^{[n]}(z_0) - \Psi_\tau(z_0) = \tau\mathbf{b} \otimes I_{2d}(\mathbf{F}(\mathbf{y}^{[n]}) - \mathbf{F}(\mathbf{y}^*)),$$

which, by (3.16) and the definition of $\tau(\Omega,\epsilon,\delta)$ in the proof of Lemma 3.1, implies

$$\|\frac{\partial \Lambda^{[n]}(z_0)}{\partial z_0}\| \leq C_2(\Omega,\epsilon)\delta^{n+2}, \quad \forall z_0 \in \Omega, \tag{3.17}$$

where the constant $C_2(\Omega,\epsilon)$ depends only on Ω and ϵ. We now write

$$\begin{aligned} & \|(\frac{\partial \Psi_\tau^{[n]}(z_0)}{\partial z_0})^T J(\frac{\partial \Psi_\tau^{[n]}(z_0)}{\partial z_0}) - J\| \\ = \; & \|(\frac{\partial \Lambda^{[n]}(z_0)}{\partial z_0} + \frac{\partial \Psi_\tau(z_0)}{\partial z_0})^T J(\frac{\partial \Lambda^{[n]}(z_0)}{\partial z_0} + \frac{\partial \Psi_\tau(z_0)}{\partial z_0}) - J\| \\ \leq \; & \|(\frac{\partial \Lambda^{[n]}(z_0)}{\partial z_0})^T J(\frac{\partial \Lambda^{[n]}(z_0)}{\partial z_0})\| + \|(\frac{\partial \Lambda^{[n]}(z_0)}{\partial z_0})^T J(\frac{\partial \Psi_\tau(z_0)}{\partial z_0})\| \\ & + \|(\frac{\partial \Psi_\tau(z_0)}{\partial z_0})^T J(\frac{\partial \Lambda^{[n]}(z_0)}{\partial z_0})\| + \|(\frac{\partial \Psi_\tau(z_0)}{\partial z_0})^T J(\frac{\partial \Psi_\tau(z_0)}{\partial z_0}) - J\|, \end{aligned} \tag{3.18}$$

where the last term vanishes when (1.4) is satisfied.

Finally, we note that

$$\|\frac{\partial \Psi_\tau(z_0)}{\partial z_0}\| = \|I_{2d} + \tau\mathbf{b} \otimes I_{2d}\frac{\partial \mathbf{F}}{\partial \mathbf{y}}(\mathbf{y}^*)\frac{\partial \mathbf{y}^*}{\partial z_0}\| \leq C_3(\Omega,\epsilon) \tag{3.19}$$

for some constant $C_3(\Omega,\epsilon) > 0$, where (3.4) is used. Combining (3.17), (3.18), and (3.19) yields (3.15). $\square$

A global version of Theorem 3.3 can be proved analogously:

THEOREM 3.4. *Let f be bounded and twice continuously differentiable, with bounded first order and second order derivatives. Consider the algorithm (3.1) and let $\Psi_\tau^{[n]}$ be the one time-step flow associated with (3.1). Let (1.4) be satisfied. Then for any $\delta \in (0,1)$, there exists $\tau(\delta) > 0$ dependent only on δ, such that when $\tau \leq \tau(\delta)$,*

$$\|(\frac{\partial \Psi_\tau^{[n]}(z_0)}{\partial z_0})^T J (\frac{\partial \Psi_\tau^{[n]}(z_0)}{\partial z_0}) - J\| \leq C\delta^{n+2}, \ \ \forall z_0 \in \mathbb{R}^{2d}, \tag{3.20}$$

for some constant $C > 0$.

Acknowledgement.

The work reported in this note was originated during an independent study the author conducted under the supervision of Prof. P. S. Krishnaprasad at the University of Maryland, College Park.

REFERENCES

1. A. AUBRY AND P. CHARTIER, *Pseudo-symplectic Runge-Kutta methods*, BIT, 38 (1998), pp. 439–461.
2. M. A. AUSTIN, P. S. KRISHNAPRASAD, AND L. WANG, *Almost Poisson integration of rigid body systems*, Journal of Comput. Phys., 107 (1993), pp. 105–117.
3. J. CANDY AND W. ROZMUS, *A symplectic integration algorithm for seperable Hamiltonian functions*, Journal of Comput. Phys., 92 (1991), pp. 230–256.
4. E. FOREST AND R. D. RUTH, *Fourth-order symplectic integration*, Phys. D, 43 (1990), pp. 105–117.
5. E. HAIRER, C. LUBICH, AND G. WANNER, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer-Verlag, Berlin, New York, 2002.
6. E. HAIRER, S. P. NØRSETT, AND G. WANNER, *Solving Ordinary Differential Equations I: Nonstiff Problems*, Springer-Verlag, 1987.
7. R. I. MCLACHLAN AND P. ATELA, *The accuracy of symplectic integrators*, Nonlinearity, 5 (1992), pp. 541–562.
8. J. M. SANZ-SERNA, *Runge-Kutta schemes for Hamiltonian systems*, BIT, 28 (1988), pp. 877–883.
9. J. M. SANZ-SERNA AND M. P. CALVO, *Numerical Hamiltonian Problems*, Chapman & Hall, London, New York, 1994.
10. D. R. SMART, *Fixed Point Theorems*, Cambridge University Press, London, New York, 1974.
11. H. YOSHIDA, *Construction of higher order symplectic integrators*, Phys. Lett. A, 150 (1990), pp. 262–268.