CONSTRUCTION OF CANONICAL DIFFERENCE SCHEMES FOR HAMILTONIAN FORMALISM VIA GENERATING FUNCTIONS * 1)

Feng Kang Wu Hua-mo Qin Meng-shao Wang Dao-liu
(*Computing Center, Academia Sinica, Beijing, China*)

Abstract

This paper discusses the relationship between canonical maps and generating functions and gives the general Hamilton-Jacobi theory for time-independent Hamiltonian systems. Based on this theory, the general method–*the generating function method*–of the construction of difference schemes for Hamiltonian systems is considered. The transition of such difference schemes from one time-step to the next is canonical. So they are called the canonical difference schemes. The well known Euler centered scheme is a canonical difference scheme. Its higher order canonical generalisations and other families of canonical difference schemes are given. The construction method proposed in the paper is also applicable to time-dependent Hamiltonian systems.

Contents

§0. Introduction

As is well known, Hamiltonian systems have many intrinsic properties: the preservation of phase areas of even dimension and the phase volume, the conservation laws of energy and momenta and other symmetries. The *canonicity* of the phase flow for time-independent Hamiltonian systems is the most important property. It

* Received January 17, 1988.
1) The Project Supported by National Natural Science Foundation of China.

ensures the preservation of phase areas and the phase volume. Thus we can hope that preserving the canonicity of transition of difference schemes from one time step to the next is also important in numerical solution of Hamiltonian systems. The first author in [1] has proposed a notion–*canonical difference schemes*. Just as its name implies, the transition of such difference schemes is canonical. In this paper, we give a general method –*generating function method*–for the construction of canonical difference schemes via generating functions. We first establish the relationship between canonical maps and generating functions and then give the general Hamilton–Jacobi theory for time–independent Hamiltonian systems. Given a matrix $\alpha \in \mathrm{CSp}(\tilde{J}_{4n}, J_{4n})$, then canonical maps and generating functions can be determined each other under so called transversality conditions. Moreover, to the phase flow of the system with Hamiltonian H there corresponds a time-dependent generating function which satisfies the Hamilton-Jacobi equation related to the given α and H. If the Hamiltonian function is analytic, then the generating function can be expressed as a power series in t, and the series can be determined recursively(Theorem 20). So truncating or approximating it in some way, we can get certain canonical map which approximates the phase flow of the Hamiltonian system. Fixing t as the time step, we obtain difference schemes. In general, such difference scheme is implicit.

In Sec. 1, we review some notions and facts about symplectic geometry which can be found in the standard texts, e.g., [2], [3], [4]. Sec. 2 concerns linear fractional transformations. This theory is important for next sections. In Sec. 3, we discuss the relationship between linear canonical maps and generating functions. It shows the outline of our idea. Sec. 4 is the continuation and deepening of section 3. It gives the relationship between nonlinear canonical maps and generating functions. In Sec. 5, we give the general Hamilton Jacobi theory. With the aid of the theory, generating functions can be represented as power series in t. It makes the preparation for constructing canonical difference schemes. In Sec. 6, it shows the general method for the construction of canonical difference schemes. Many canonical difference schemes, such as Euler centered scheme, 4–th order centered scheme, staggered explicit scheme and others are presented.

We shall limit ourselves to the local case throughout the paper. Moreover, in this paper we use the older terminologies such as canonicity, canonical maps, etc., in stead of the modern ones such as symplecticity, symplectic maps, etc. So the *canonical* difference schemes can also be called synonymously or even more preferably *symplectic* difference schemes.

§1. Preliminary Facts about Symplectic Geometry

We now review some notions and facts of symplectic geometry [2],[3],[4].

Let $\mathbf{R}^{2n}$ be a $2n$-dim real linear space. The elements of $\mathbf{R}^{2n}$ are $2n$-dim column vectors $z = (z_1, \cdots, z_n, z_{n+1}, \cdots, z_{2n})^T = (p_1, \cdots, p_n, q_1, \cdots, q_n)^T$. The superscript T represents the matrix transpose.

A symplectic form ω_K is a bilinear form defined by the anti-symmetric matrix $K \in \mathbf{GL}(2n)$ as

$$\omega_K(z_1, z_2) = z_1^T K z_2, \qquad \text{for all } z_1, z_2 \in \mathbf{R}^{2n}. \tag{1}$$

The symplectic form ω_J,

$$\omega_J(z_1, z_2) = z_1^T J z_2, \qquad J = \begin{bmatrix} 0_n & I_n \\ -I_n & 0_n \end{bmatrix}, \qquad \text{for all } z_1, z_2 \in \mathbf{R}^{2n}, \tag{2}$$

where I_n resp. 0_n is the $n \times n$ unit resp. zero matrix, is called the standard symplectic form of $\mathbf{R}^{2n}$, briefly denoted by ω.

Every $2n \times n$ matrix

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \in \mathbf{M}(2n, n) \quad \text{of rank } n, \quad A_1, A_2 \in \mathbf{M}(n),$$

defines an n-dim subspace $\{A\}$, spanned by its n column vectors. Evidently, $\{A\} = \{B\}$ iff $\exists P \in \mathbf{GL}(n)$ such that

$$AP = B, \qquad i.e., \qquad \begin{bmatrix} A_1 P \\ A_2 P \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}.$$

An n-dim subspace $\{A\}$ is called K-Lagrangian if

$$\omega_K(z_1, z_2) = z_1^T K z_2 = 0, \qquad \text{for all } z_1, z_2 \in \{A\}.$$

Evidently, $\{A\}$ is K-Lagrangian iff

$$A^T K A = 0.$$

Define

$$\mathbf{Sp}(K_1, K_2; 2n) = \{M \in \mathbf{GL}(2n) \mid M^T K_2 M = K_1\};$$
$$\mathbf{CSp}(K_1, K_2; 2n) = \{M \in \mathbf{GL}(2n) \mid M^T K_2 M = \mu K_1, \text{ for some } \mu = \mu(M) \neq 0\};$$
$$\mathbf{Sp}(K; 2n) = \mathbf{Sp}(K, K; 2n) = \{M \in \mathbf{GL}(2n) \mid M^T K M = K\};$$
$$\mathbf{CSp}(K; 2n) = \mathbf{CSp}(K, K; 2n) = \{M \in \mathbf{GL}(2n) \mid M^T K M = \mu K, \text{ for some } \mu = \mu(M) \neq 0\}.$$

$\mathbf{Sp}(K; 2n)$ and $\mathbf{CSp}(K; 2n)$ are groups and called the K-symplectic group and the conformal K-symplectic group respectively. $\mathbf{Sp}(2n) \equiv \mathbf{Sp}(J; 2n)$ and $\mathbf{CSp}(2n) \equiv \mathbf{CSp}(J; 2n)$. They are usually called the symplectic group and the conformal symplectic group respectively.

Proposition 1.

$$M \in \mathbf{Sp}(K_1, K_2; 2n) \text{ iff } M^{-1} \in \mathbf{Sp}(K_2, K_1; 2n).$$

$$M \in \mathbf{CSp}(K_1, K_2; 2n) \text{ iff } M^{-1} \in \mathbf{CSp}(K_2, K_1; 2n).$$

Proposition 2. Let $M_0 \in \mathbf{Sp}(K_1, K_2; 2n)$. Then

$$\mathbf{Sp}(K_1, K_2; 2n) = \mathbf{Sp}(K_2; 2n) \cdot M_0 = M_0 \cdot \mathbf{Sp}(K_1; 2n), \tag{3}$$

$$\mathbf{CSp}(K_1, K_2; 2n) = \mathbf{CSp}(K_2; 2n) \cdot M_0 = M_0 \cdot \mathbf{CSp}(K_1; 2n), \tag{4}$$

where

$$\mathbf{Sp}(K_2; 2n) \cdot M_0 = \{M \cdot M_0 | M \in \mathbf{Sp}(K_2; 2n)\}$$

and others are similar.

§2. Linear Fractional Transformations

Definition 3. *Let* $\alpha = \begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix} \in \mathbf{GL}(2m)$. *Define a linear fractional transformation*

$$\begin{aligned} \sigma_\alpha : \ & \mathbf{M}(m) \longrightarrow \mathbf{M}(m), \\ & M \longrightarrow N = \sigma_\alpha(M) = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1} \end{aligned} \tag{5}$$

under the transversality condition

$$|C_\alpha M + D_\alpha| \neq 0. \tag{6}$$

Proposition 4. Let $\alpha \in \mathbf{GL}(2m)$. Denote $\alpha^{-1} = \begin{bmatrix} A^\alpha & B^\alpha \\ C^\alpha & D^\alpha \end{bmatrix}$. Then

$$|C_\alpha M + D_\alpha| \neq 0 \quad \text{iff} \quad |MC^\alpha - A^\alpha| \neq 0, \tag{7}$$

$$|A_\alpha M + B_\alpha| \neq 0 \quad \text{iff} \quad |B^\alpha - MD^\alpha| \neq 0. \tag{8}$$

Proof. From equations

$$\begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix} \begin{bmatrix} A^\alpha & B^\alpha \\ C^\alpha & D^\alpha \end{bmatrix} = \begin{bmatrix} A^\alpha & B^\alpha \\ C^\alpha & D^\alpha \end{bmatrix} \begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix} = I_{2m} \tag{9}$$

i.e.,

$$\begin{aligned} A_\alpha A^\alpha + B_\alpha C^\alpha &= A^\alpha A_\alpha + B^\alpha C_\alpha = I_m, \\ C_\alpha B^\alpha + D_\alpha D^\alpha &= C^\alpha B_\alpha + D^\alpha D_\alpha = I_m, \\ A_\alpha B^\alpha + B_\alpha D^\alpha &= A^\alpha B_\alpha + B^\alpha D_\alpha = 0, \\ C_\alpha A^\alpha + D_\alpha C^\alpha &= C^\alpha A_\alpha + D^\alpha C_\alpha = 0, \end{aligned} \tag{10}$$

we obtain the identities

$$\begin{bmatrix} I & -M \\ C_\alpha & D_\alpha \end{bmatrix} \begin{bmatrix} A^\alpha & B^\alpha \\ C^\alpha & D^\alpha \end{bmatrix} = \begin{bmatrix} A^\alpha - MC^\alpha & B^\alpha - MD^\alpha \\ 0 & I \end{bmatrix},$$
$$\begin{bmatrix} I & -M \\ A_\alpha & B_\alpha \end{bmatrix} \begin{bmatrix} A^\alpha & B^\alpha \\ C^\alpha & D^\alpha \end{bmatrix} = \begin{bmatrix} A^\alpha - MC^\alpha & B^\alpha - MD^\alpha \\ I & 0 \end{bmatrix}.$$

In addition,

$$\begin{bmatrix} I & -M \\ C_\alpha & D_\alpha \end{bmatrix} = \begin{bmatrix} I & 0 \\ C_\alpha & I \end{bmatrix} \begin{bmatrix} I & -M \\ 0 & C_\alpha M + D_\alpha \end{bmatrix},$$
$$\begin{bmatrix} I & -M \\ A_\alpha & B_\alpha \end{bmatrix} = \begin{bmatrix} I & 0 \\ A_\alpha & I \end{bmatrix} \begin{bmatrix} I & -M \\ 0 & A_\alpha M + B_\alpha \end{bmatrix}.$$

Taking their determinant, we get

$$\begin{aligned} |C_\alpha M + D_\alpha|\ |\alpha|^{-1} &= |A^\alpha - MC^\alpha|, \\ |A_\alpha M + B_\alpha|\ |\alpha|^{-1} &= (-1)^m |B^\alpha - MD^\alpha|. \end{aligned}$$

Q.E.D.

Proposition 5. The linear fractional transformation σ_α in (5) can be represented as

$$\sigma_\alpha(M) = (MC^\alpha - A^\alpha)^{-1}(B^\alpha - MD^\alpha). \tag{11}$$

Proof. By (7), (11) is well defined. Therefore we need only to verify the identity

$$(MC^\alpha - A^\alpha)^{-1}(B^\alpha - MD^\alpha) = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1},$$

i.e.,

$$(B^\alpha - MD^\alpha)(C_\alpha M + D_\alpha) = (MC^\alpha - A^\alpha)(A_\alpha M + B_\alpha). \tag{12}$$

Expanding it and using the conditions (10), we know that (12) holds. Q.E.D.

Proposition 6.

$$(C^\alpha N + D^\alpha)(C_\alpha M + D_\alpha) = I, \tag{13}$$

hence

$$|C^\alpha N + D^\alpha| \neq 0 \quad \text{iff} \quad |C_\alpha M + D_\alpha| \neq 0, \tag{14}$$

where $N = \sigma_\alpha(M)$. So under the transversality condition (6) σ_α has an inverse transformation $\sigma_\alpha^{-1} = \sigma_{\alpha^{-1}}$.

$$M = \sigma_{\alpha^{-1}}(N) = (A^\alpha N + B^\alpha)(C^\alpha N + D^\alpha)^{-1} \tag{15}$$

$$= (NC_\alpha - A_\alpha)^{-1}(B_\alpha - ND_\alpha). \tag{16}$$

Proof.

$$\begin{aligned}
&(C^\alpha N + D^\alpha)(C_\alpha M + D_\alpha)\\
&= (C^\alpha(A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1} + D^\alpha)(C_\alpha M + D_\alpha)\\
&= (C^\alpha A_\alpha + D^\alpha C_\alpha)M + C^\alpha B_\alpha + D^\alpha D_\alpha\\
&= I \qquad (by (10)).
\end{aligned}$$

This is (13). In Prop.4, substituting α^{-1} for α and N for M, we get

$$|C^\alpha N + D^\alpha| \neq 0 \quad \text{iff} \quad |NC_\alpha - A_\alpha| \neq 0.$$

Hence the equation

$$N = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1}$$

is solvable for M and the solution is

$$M = (NC_\alpha - A_\alpha)^{-1}(B_\alpha - ND_\alpha).$$

This is (16). (15) can be got from (11). Q.E.D.

Putting together the above stated, the following four transversality conditions are equivalent mutually:

$$|C_\alpha M + D_\alpha| \neq 0 \tag{17.1}$$

$$|MC^\alpha - A^\alpha| \neq 0 \tag{17.2}$$

$$|C^\alpha N + D^\alpha| \neq 0 \tag{17.3}$$

$$|NC_\alpha - A_\alpha| \neq 0, \tag{17.4}$$

where

$$N = \sigma_\alpha(M) = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1}$$

$$M = \sigma_{\alpha^{-1}}(N) = (A^\alpha N + B^\alpha)(C^\alpha N + D^\alpha)^{-1}.$$

§3. Linear Symplectic and Linear Gradient Maps and Generating Functions

In this section we study linear fractional transformation relating the symplectic group $\mathrm{Sp}(2n)$ and the space $\mathrm{Sm}(2n)$ of symmetric matrices of order $2n$.

Consider the two different symplectic forms on $\mathbf{R}^{4n}$, the natural symplectic form $J_{4n} = \begin{bmatrix} 0 & I_{2n} \\ -I_{2n} & 0 \end{bmatrix}$ and the natural product symplectic form $\tilde{J}_{4n} = \begin{bmatrix} J_{2n} & 0 \\ 0 & -J_{2n} \end{bmatrix}$. Denote $\mathbf{R}^{4n} = (\mathbf{R}^{4n}, J_{4n})$, $\tilde{\mathbf{R}}^{4n} = (\mathbf{R}^{4n}, \tilde{J}_{4n})$.

A $2n$-dim subspace $\{X\} = \begin{Bmatrix} X_1 \\ X_2 \end{Bmatrix}$ of $\mathbf{R}^{4n}$, $X_1, X_2 \in \mathbf{M}(2n)$, is J_{4n}-Lagrangian if

$$X^T J_{4n} X = 0,$$

i.e.,

$$X_1^T X_2 - X_2^T X_1 = 0 \qquad \text{or} \qquad X_1^T X_2 \in \mathrm{Sm}(2n).$$

We call such $4n \times 2n$ matrix $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, according to Siegel [8], a symmetric pair. If, moreover, $|X_2| \neq 0$, then $X_1 X_2^{-1} = N \in \mathrm{Sm}(2n)$ and $\begin{Bmatrix} X_1 \\ X_2 \end{Bmatrix} = \begin{Bmatrix} N \\ I \end{Bmatrix}$.

Similarly, a $2n$-dim subspace $\{Y\} = \begin{Bmatrix} Y_1 \\ Y_2 \end{Bmatrix}$ is $\tilde{J}_{4n}$-Lagrangian if

$$Y^T \tilde{J}_{4n} Y = 0,$$

i.e.,

$$Y_1^T J_{2n} Y_1 = Y_2^T J_{2n} Y_2.$$

The $4n \times 2n$ matrix $Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ is called a symplectic pair. $|Y_2| \neq 0$ implies $Y_1 Y_2^{-1} = M \in \mathrm{Sp}(2n)$ and $\begin{Bmatrix} Y_1 \\ Y_2 \end{Bmatrix} = \begin{Bmatrix} M \\ I \end{Bmatrix}$.

Theorem 7. $\alpha = \begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix} \in \mathrm{GL}(4n)$ *carries every $\tilde{J}_{4n}$-Lagrangian subspace into a J_{4n}-Lagrangian subspace if and only if* $\alpha \in \mathrm{CSp}(\tilde{J}_{4n}, J_{4n})$, *i.e.,*

$$\alpha^T J_{4n}\alpha = \mu \tilde{J}_{4n}, \qquad \textit{for some } \mu = \mu(\alpha) \neq 0. \tag{18}$$

Proof. "if" part is obvious, we need only to prove "only if" part.
Taking $\alpha_0 \in \mathbf{Sp}(\tilde{J}_{4n}, J_{4n})$(it always exists), by Prop. 2 we have

$$\mathbf{CSp}(\tilde{J}_{4n}, J_{4n}) = \mathbf{CSp}(4n) \cdot \alpha_0.$$

Therefore it suffices to show that if α carries every J_{4n}- Lagrangian subspace into J_{4n}-Lagrangian subspace then $\alpha \in \mathbf{CSp}(4n)$, i.e.,

$$\alpha^T J_{4n}\alpha = \mu J_{4n}, \qquad \text{for some } \mu \neq 0.$$

1^0 Take the symmetric pair $X = \begin{bmatrix} I_{2n} \\ 0_{2n} \end{bmatrix}$. By assumption,

$$\alpha X = \begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix} = \begin{bmatrix} A_\alpha \\ C_\alpha \end{bmatrix}$$

is also a symmetric pair, i.e., $A_\alpha^T C_\alpha - C_\alpha^T A_\alpha = 0$. Similarly, $B_\alpha^T D_\alpha - D_\alpha^T B_\alpha = 0$.

2^0 Take the symmetric pairs $X = \begin{bmatrix} S \\ I \end{bmatrix}$, $S \in \mathbf{Sm}(2n)$. Then

$$\alpha X = \begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix} \begin{bmatrix} S \\ I \end{bmatrix} = \begin{bmatrix} A_\alpha S + B_\alpha \\ C_\alpha S + D_\alpha \end{bmatrix}$$

are also symmetric pairs, i.e.,

$$\begin{aligned} 0 &= (\alpha X)^T J_{4n}(\alpha X) \\ &= \left(S^T A_\alpha^T + B_\alpha^T, S^T C_\alpha^T + D_\alpha^T\right) \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} A_\alpha S + B_\alpha \\ C_\alpha S + D_\alpha \end{bmatrix} \\ &= S(A_\alpha^T C_\alpha - C_\alpha^T A_\alpha)S + S(A_\alpha^T D_\alpha - C_\alpha^T B_\alpha) \\ &\quad -(D_\alpha^T A_\alpha - B_\alpha^T C_\alpha)S + B_\alpha^T D_\alpha - D_\alpha^T B_\alpha \\ &= S(A_\alpha^T D_\alpha - C_\alpha^T B_\alpha) - (A_\alpha^T D_\alpha - C_\alpha^T B_\alpha)^T S, \qquad \forall S \in \mathbf{Sm}(2n). \end{aligned}$$

Set $P = A_\alpha^T D_\alpha - C_\alpha^T B_\alpha$. Then the above equation becomes

$$SP = P^T S, \qquad \forall S \in \mathbf{Sm}(2n).$$

It follows that $P = \mu I$, i.e.,

$$A_\alpha^T D_\alpha - C_\alpha^T B_\alpha = \mu I.$$

So

$$\begin{aligned}\alpha^T J_{4n}\alpha &= \begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix}^T \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix} \\ &= \begin{bmatrix} A_\alpha^T C_\alpha - C_\alpha^T A_\alpha & A_\alpha^T D_\alpha - C_\alpha^T B_\alpha \\ B_\alpha^T C_\alpha - D_\alpha^T A_\alpha & B_\alpha^T D_\alpha - D_\alpha^T B_\alpha \end{bmatrix} \\ &= \mu \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} = \mu J_{4n}.\end{aligned}$$

$\alpha \in \mathbf{GL}(4n)$ implies $\mu \neq 0$. Q.E.D.

Remark 8. The inverse matrix of α is denoted by $\alpha^{-1} = \begin{bmatrix} A^\alpha & B^\alpha \\ C^\alpha & D^\alpha \end{bmatrix}$. Then by (18), we have

$$A^\alpha = \mu^{-1} J C_\alpha^T, \qquad B^\alpha = -\mu^{-1} J A_\alpha^T,$$
$$C^\alpha = -\mu^{-1} J D_\alpha^T, \qquad D^\alpha = \mu^{-1} J B_\alpha^T.$$

Definition 9. *A linear map $z \to \hat{z} = g(z) = Mz$ resp. $w \to \hat{w} = f(z) = Nw : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ is called a canonical resp. gradient map if $M \in \mathrm{Sp}(2n)$ resp. $N \in \mathrm{Sm}(2n)$.*

For a linear map $z \longrightarrow \hat{z} = g(z) = Mz : \mathbf{R}^m \to \mathbf{R}^m$, the graph of g,

$$\Gamma_g = \left\{ \begin{bmatrix} \hat{z} \\ z \end{bmatrix} \in \mathbf{R}^{2m} \middle| \quad \hat{z} = g(z) = Mz \right\}$$

is an m-dim subspace of $\mathbf{R}^{2m}$ and $\Gamma = \left\{ \begin{matrix} M \\ I \end{matrix} \right\}$, i.e., it is spanned by the column vectors of $\begin{bmatrix} M \\ I \end{bmatrix}$.

Proposition 10. The graph Γ_g of a linear canonical map $g(z)$ is a $\tilde{J}_{4n}$-Lagrangian subspace. The graph Γ_f of a linear gradient map f is a J_{4n}-Lagrangian subspace.

Let $\alpha = \begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix} \in \mathrm{CSp}(\tilde{J}_{4n}, J_{4n})$. It defines linear transformations

$$\begin{bmatrix} \hat{w} \\ w \end{bmatrix} = \alpha \begin{bmatrix} \hat{z} \\ z \end{bmatrix}, \qquad \begin{bmatrix} \hat{z} \\ z \end{bmatrix} = \alpha^{-1} \begin{bmatrix} \hat{w} \\ w \end{bmatrix}, \tag{20}$$

i.e.,

$$\hat{w} = A_\alpha \hat{z} + B_\alpha z, \qquad \hat{z} = A^\alpha \hat{w} + B^\alpha w, \tag{21}$$
$$w = C_\alpha \hat{z} + D_\alpha z, \qquad z = C^\alpha \hat{w} + D^\alpha w.$$

Theorem 11. *Let* $\alpha = \begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix} \in \mathbf{CSp}(\tilde{J}_{4n}, J_{4n})$. *Then* $M \in \mathbf{Sp}(2n)$ *and M satisfies (17.1) iff* $N = \sigma_\alpha(M) \in \mathbf{Sm}(2n)$ *and N satisfies (17.3). That is, the linear fractional transformation* $\sigma_\alpha : \{M \in \mathbf{Sp}(2n) \mid |C_\alpha M + D_\alpha| \neq 0\} \longrightarrow \{N \in \mathbf{Sm}(2n) \mid |C^\alpha N + D^\alpha| \neq 0\}$ *is one to one and onto.*

Proof. By (17), we know $|C_\alpha M + D_\alpha| \neq 0$ iff $|C^\alpha N + D^\alpha| \neq 0$.

"only if" part. The map $z \to \hat{z} = g(z) = Mz : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ is a linear canonical map. So its graph Γ_g is a $\tilde{J}_{4n}$-Lagrangian subspace. Since $\alpha \in \mathbf{CSp}(\tilde{J}_{4n}, J_{4n})$, by Theorem 7, $\alpha(\Gamma_g)$ is J_{4n}-Lagrangian. Notice that $\Gamma_g = \begin{Bmatrix} M \\ I \end{Bmatrix}$.

$$\alpha(\Gamma_g) = \left\{ \begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix} \begin{bmatrix} M \\ I \end{bmatrix} \right\} = \begin{Bmatrix} A_\alpha M + B_\alpha \\ C_\alpha M + D_\alpha \end{Bmatrix}.$$

By assumption, $|C_\alpha M + D_\alpha| \neq 0$. Therefore

$$\alpha(\Gamma_g) = \begin{Bmatrix} (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1} \\ I \end{Bmatrix} = \begin{Bmatrix} \sigma_\alpha(M) \\ I \end{Bmatrix}$$

and it is J_{4n}-Lagrangian. That is, $N = \sigma_\alpha(M)$ is symmetric.

Substituting α^{-1} for α and noting $\alpha^{-1} \in \mathbf{CSp}(J_{4n}, \tilde{J}_{4n})$, we can get similarly the "if" part. Q.E.D.

Theorem 12. Let $\alpha \in \mathbf{CSp}(\tilde{J}_{4n}, J_{4n})$. Let $z \to \hat{z} = g(z) = Mz$ be a linear canonical map and M satisfy (17.1). Then there exists a linear gradient map $w \to \hat{w} = f(w) = Nw$ and a quadratic function–generating function–$\phi(w) = \frac{1}{2} w^T N w$(depending on α and g) such that

1. $f(w) = \nabla\phi(w)$; (22)
2. $A_\alpha g(z) + B_\alpha z = f(C_\alpha g(z) + D_\alpha z) = \nabla\phi(C_\alpha g(z) + D_\alpha z)$, identically in z; (23)
3. $N = \sigma_\alpha(M) = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1}$; $M = \sigma_{\alpha^{-1}}(N) = (A^\alpha N + B^\alpha)(C^\alpha N + D^\alpha)^{-1}$; (24)
4. $\Gamma_f = \alpha(\Gamma_g)$, $\Gamma_g = \alpha^{-1}(\Gamma_f)$, (25)

where $\nabla\phi = (\phi_{w_1}, \cdots, \phi_{w_{2n}})^T$, $\phi_w = (\phi_{w_1}, \cdots, \phi_{w_{2n}})$, so $\nabla\phi = (\phi_w)^T$.

Proof. The image of Γ_g under α is

$$\alpha(\Gamma_g) = \left\{ \begin{bmatrix} \hat{w} \\ w \end{bmatrix} \in \mathbf{R}^{4n} \mid \hat{w} = A_\alpha Mz + B_\alpha z, w = C_\alpha Mz + D_\alpha z \right\}.$$

By assumption, $|C_\alpha M + D_\alpha| \neq 0$, so $w = (C_\alpha M + D_\alpha)z$ is invertible and its inverse is $z = (C_\alpha M + D_\alpha)^{-1}w$. Set $\hat{w} = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1}w = Nw = f(w)$. By Theorem 11, $f(w)$ is a linear gradient map. Obviously, $\phi(w) = \frac{1}{2}w^T Nw$ satisfies (22).

From the equation

$$N = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1},$$

it follows that

$$A_\alpha M + B_\alpha = N(C_\alpha M + D_\alpha).$$

So

$$A_\alpha Mz + B_\alpha z = N(C_\alpha Mz + D_\alpha z), \qquad \forall z \in \mathbf{R}^{2n},$$

i.e.,

$$A_\alpha g(z) + B_\alpha z = f(C_\alpha g(z) + D_\alpha z), \qquad \text{identically in } z.$$

Since $\Gamma_g = \left\{ \begin{matrix} M \\ I \end{matrix} \right\}$,

$$\alpha(\Gamma_g) = \left\{ \begin{bmatrix} A_\alpha & B_\alpha \\ C_\alpha & D_\alpha \end{bmatrix} \begin{bmatrix} M \\ I \end{bmatrix} \right\} = \left\{ \begin{matrix} A_\alpha M + B_\alpha \\ C_\alpha M + D_\alpha \end{matrix} \right\}.$$

$|C_\alpha M + D_\alpha| \neq 0$ implies $\alpha(\Gamma_g) = \left\{ \begin{matrix} N \\ I \end{matrix} \right\} = \Gamma_f$. Q.E.D.

Theorem 13. *Let $\alpha \in \mathrm{CSp}(\tilde{J}_{4n}, J_{4n})$. Let $\phi(w) = \frac{1}{2}w^T Nw$, $N \in \mathrm{Sm}(2n)$ be a quadratic function, $f(w) = \nabla\phi = Nw$ its induced linear gradient map and N satisfy (17.3). Then there exists a linear canonical map $z \to \hat{z} = g(z) = Mz$ such that*

1. $A^\alpha f(w) + B^\alpha w = g(C^\alpha f(w) + D^\alpha w)$, *identically in w*;
2. $M = \sigma_{\alpha^{-1}}(N) = (A^\alpha N + B^\alpha)(C^\alpha N + D^\alpha)^{-1}$;
   $N = \sigma_\alpha(M) = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1}$;
3. $\Gamma_g = \alpha^{-1}(\Gamma_f)$ $\Gamma_f = \alpha(\Gamma_g)$.

The proof is similar to the one of Theorem 12 and omitted here.

§4. General Canonical and Gradient Maps and Generating Functions

Definition 14. *A map $z \to \hat{z} = g(z) : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ is called a canonical map if its Jacobian $M(z) = g_z(z) \in \mathrm{Sp}(2n)$ everywhere. A map $w \to \hat{w} = f(w) : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ is called a gradient map if its Jacobian $N(w) = f_w(w) \in \mathrm{Sm}(2n)$ everywhere.*

Definition 15. *An m-dim submanifold U of $(\mathbf{R}^{2m}, K)$ is a K-Lagrangian submanifold if its tangent plane T_zU at z is a K-Lagrangian subspace of the tangent space $T_z\mathbf{R}^{2m}$ to $\mathbf{R}^{2m}$ at z for any $z \in U$.*

Proposition 16. The graph Γ_g of a canonical map $z \to \hat{z} = g(z)$ is a $\tilde{J}_{4n}$-Lagrangian submanifold of $\tilde{\mathbf{R}}^{4n}$. The graph Γ_f of a gradient map $w \to \hat{w} = f(w)$ is a J_{4n}-Lagrangian submanifold of $\mathbf{R}^{4n}$.

Theorem 17. *Let $\alpha \in \mathrm{CSp}(\tilde{J}_{4n}, J_{4n})$. Let $z \to \hat{z} = g(z) : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ be a canonical map with Jacobian $M(z) = g_z(z) \in \mathrm{Sp}(2n)$ satisfying (17.1) in (some neighborhood of) $\mathbf{R}^{2n}$. Then there exists a gradient map $w \to \hat{w} = f(w)$ in (some neighborhood of) $\mathbf{R}^{2n}$ with Jacobian $N(w) = f_w(w) \in \mathrm{Sm}(2n)$ satisfying (17.3) and a scalar function -generating function-$\phi(w)$(depending on α and g) such that*

1. $$f(w) = \nabla\phi(w); \tag{26}$$
2. $$\begin{aligned} A_\alpha g(z) + B_\alpha z &= f(C_\alpha g(z) + D_\alpha z); \\ &= \nabla\phi(C_\alpha g(z) + D_\alpha z), \quad \textit{identically in } z \end{aligned} \tag{27}$$
3. $$\begin{aligned} N &= \sigma_\alpha(M) = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1}; \\ M &= \sigma_{\alpha^{-1}}(N) = (A^\alpha N + B^\alpha)(C^\alpha N + D^\alpha)^{-1}; \end{aligned} \tag{28}$$
4. $$\Gamma_f = \alpha(\Gamma_g), \ \Gamma_g = \alpha^{-1}(\Gamma_f). \tag{29}$$

Proof. Under the linear transformation α, the image of Γ_g is

$$\alpha(\Gamma_g) = \left\{ \begin{bmatrix} \hat{w} \\ w \end{bmatrix} \in \mathbf{R}^{4n} \middle| \ \hat{w} = A_\alpha g(z) + B_\alpha z, \quad w = C_\alpha g(z) + D_\alpha z \right\}.$$

Since Γ_g is a $\tilde{J}_{4n}$-Lagrangian submanifold and $\alpha \in \mathrm{CSp}(\tilde{J}_{4n}, J_{4n})$, the tangent plane of $\alpha(\Gamma_g)$

$$\left\{ \begin{matrix} A_\alpha M(z) + B_\alpha \\ C_\alpha M(z) + D_\alpha \end{matrix} \right\}$$

is a J_{4n}-Lagrangian subspace. So $\alpha(\Gamma_g)$ is a J_{4n}-Lagrangian submanifold. By assumption, $|C_\alpha M + D_\alpha| \neq 0$, so, by the Implicit Function Theorem,

$$w = C_\alpha g(z) + D_\alpha z \tag{30}$$

is invertible and its inverse is denoted by $z = z(w)$. Set

$$\hat{w} = f(w) = (A_\alpha g(z) + B_\alpha z)|_{z=z(w)} = A_\alpha g(z(w)) + B_\alpha z(w).$$

Obviously, such $f(w)$ satisfies the identity

$$A_\alpha g(z) + B_\alpha z \equiv f(C_\alpha g(z) + D_\alpha z).$$

The Jacobian of f is

$$\begin{aligned} N(w) &= f_w(w) = \frac{\partial \hat{w}}{\partial w} = \frac{\partial \hat{w}}{\partial z}\frac{\partial z}{\partial w} = \frac{\partial \hat{w}}{\partial z}\left(\frac{\partial w}{\partial z}\right)^{-1} \\ &= (A_\alpha M(z) + B_\alpha)(C_\alpha M(z) + D_\alpha)^{-1} = \sigma_\alpha(M(z)). \end{aligned}$$

By Theorem 11, it is symmetric. So $f(w)$ is a gradient map. By Poincare Lemma, there exists a scalar function $\phi(w)$, such that

$$f(w) = \nabla\phi(w).$$

In addition,

$$\Gamma_f = \left\{ \begin{bmatrix} \hat{w} \\ w \end{bmatrix} \in \mathbf{R}^{4n} \mid \hat{w} = f(w) = A_\alpha g(z(w)) + B_\alpha z(w) \right\} = \alpha(\Gamma_g).$$

Q.E.D.

Theorem 18. *Let $\alpha \in \mathbf{CSp}(\tilde{J}_{4n}, J_{4n})$. Let $\phi(w)$ be a scalar function and $w \to \hat{w} = f(w) = \nabla\phi(w)$ be its induced gradient map and $N(w) = f_w(w) = \phi_{ww}(w)$, the Hessian matrix of $\phi(w)$, satisfy (17.3) in (some neighborhood of) $\mathbf{R}^{2n}$. Then there exists a canonical map $z \to \hat{z} = g(z)$ with Jacobian $M(z) = g_z(z)$ satisfying (17.1) such that*

1. $A^\alpha f(w) + B^\alpha w = g(C^\alpha f(w) + D^\alpha w)$, *identically in w;*
2. $M = \sigma_{\alpha^{-1}}(N) = (A^\alpha N + B^\alpha)(C^\alpha N + D^\alpha)^{-1}$;
   $N = \sigma_\alpha(M) = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1}$;
3. $\Gamma_g = \alpha^{-1}(\Gamma_f)$, $\quad \Gamma_f = \alpha(\Gamma_g)$.

§5. Generating Functions for the Phase Flow of Hamiltonian Systems

Consider the Hamiltonian system

$$\frac{dz}{dt} = J_{2n}^{-1}\nabla H(z), \qquad z \in \mathbf{R}^{2n}, \tag{31}$$

where $H(z)$ is a Hamiltonian function. Its phase flow is denoted as $g^t(z) = g(z,t) = g_H(z,t)$, being a one-parameter group of canonical maps [2,3], i.e.,

$$g^0 = \text{identity}, \qquad g^{t_1+t_2} = g^{t_1} \circ g^{t_2}$$

and if z_0 is taken as an initial condition, then $z(t) = g^t(z_0)$ is the solution of (31) with the initial value z_0.

Theorem 19. *Let $\alpha \in \mathrm{CSp}(\tilde{J}_{4n}, J_{4n})$. Let $z \to \hat{z} = g(z,t)$ be the phase flow of the Hamiltonian system (31) and $M_0 \in \mathrm{Sp}(2n)$. Set $G(z,t) = g(M_0 z, t)$ with Jacobian $M(z,t) = G_z(z,t)$. It is a time-dependent canonical map. If M_0 satisfies the transversality condition (17.1), i.e.,*

$$|C_\alpha M_0 + D_\alpha| \neq 0, \tag{32}$$

then there exists, for sufficiently small $|t|$ and in (some neighborhood of) $\mathbf{R}^{2n}$, a time-dependent gradient map $w \to \hat{w} = f(w,t)$ with Jacobian $N(w,t) = f_w(w,t) \in \mathrm{Sm}(2n)$ satisfying the transversality condition (17.3) and a time-dependent generating function $\phi_{\alpha,H}(w,t) = \phi(w,t)$ such that

1. $f(w,t) = \nabla\phi(w,t);$ (33)
2. $\frac{\partial}{\partial t}\phi = -\mu H(A^\alpha \nabla\phi(w,t) + B^\alpha w),$ (34)
3. $$\begin{aligned} A_\alpha G(z,t) + B_\alpha z &\equiv f(C_\alpha G(z,t) + D_\alpha z, t) \\ &\equiv \nabla\phi(C_\alpha G(z,t) + D_\alpha z, t), \end{aligned} \tag{35}$$

$$\begin{aligned} N &= \sigma_\alpha(M) = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1}, \\ M &= \sigma_{\alpha^{-1}}(N) = (A^\alpha N + B^\alpha)(C^\alpha N + D^\alpha)^{-1}. \end{aligned} \tag{36}$$

(34) is the most general Hamilton–Jacobi equation for the Hamiltonian system (31) and the linear transformation α.

Proof. Since $g(z,t)$ is differentiable with respect to z and t, so is $G(z,t)$. Condition (32) implies that for sufficientlly small $|t|$ and in some neighborhood of $\mathbf{R}^{2n}$,

$$|C_\alpha M(z,t) + D_\alpha| \neq 0. \tag{37}$$

Thus by Theorem 17 there exists a time-dependent gradient map $\hat{w} = f(w,t)$ such that it satisfies (35) and (36).

Set

$$\begin{aligned} \overline{H}(w,t) &= -\mu H(\hat{z})|_{\hat{z}=A^\alpha \hat{w}(w,t)+B^\alpha w} \\ &= -\mu H(A^\alpha \hat{w}(w,t) + B^\alpha w). \end{aligned} \tag{38}$$

Consider the differetial 1–form

$$\omega^1 = \sum_{i=1}^{2n} \widehat{w}_i dw_i + \overline{H}(w,t)dt. \tag{39}$$

$$\begin{aligned} d\omega^1 &= \sum_{i,j=1}^{2n} \frac{\partial \widehat{w}_i}{\partial w_j} dw_j \wedge dw_i + \sum_{i=1}^{2n} \frac{\partial \widehat{w}_i}{\partial t} dt \wedge dw_i + \sum_{i=1}^{2n} \frac{\partial \overline{H}}{\partial w_i} dw_i \wedge dt \\ &= \sum_{i<j} \left(\frac{\partial \widehat{w}_i}{\partial w_j} - \frac{\partial \widehat{w}_j}{\partial w_i} \right) dw_j \wedge dw_i + \sum_{i=1}^{2n} \left(\frac{\partial \widehat{w}_i}{\partial t} - \frac{\partial \overline{H}}{\partial w_i} \right) dt \wedge dw_i. \end{aligned} \tag{40}$$

Since $N(w,t) = f_w(w,t) = \partial \widehat{w}/\partial w$ is symmetric, the first term of (40) is zero.

Notice that $\widehat{z} = G(z,t) = g(M_0 z, t)$,

$$\frac{d\widehat{z}}{dt} = \frac{dg(M_0 z, t)}{dt} = J^{-1} \nabla H(G(z,t)). \tag{41}$$

So $G(z,t)$ is the solution of the following initial–value problem

$$\begin{cases} \frac{d\widehat{z}}{dt} = J^{-1} \nabla H(\widehat{z}), \\ \widehat{z}(0) = M_0 z. \end{cases}$$

Therefore from the equations

$$\widehat{w} = A_\alpha G(z,t) + B_\alpha z, \qquad w = C_\alpha G(z,t) + D_\alpha z,$$

it follows that

$$\frac{d\widehat{w}}{dt} = A_\alpha J^{-1} \nabla H(\widehat{z}), \qquad \frac{dw}{dt} = C_\alpha J^{-1} \nabla H(\widehat{z}).$$

Since

$$\begin{aligned} \frac{d\widehat{w}}{dt} &= \frac{\partial \widehat{w}}{\partial w} \frac{dw}{dt} + \frac{\partial \widehat{w}}{\partial t}, \\ \frac{\partial \widehat{w}}{\partial t} &= (A_\alpha - \frac{\partial \widehat{w}}{\partial w} C_\alpha) J^{-1} \nabla H(\widehat{z}). \end{aligned}$$

On the other hand,

$$\begin{aligned} \nabla_w \overline{H}(w,t) &= \left(\overline{H}_w(w,t) \right)^T = \mu \left(-H_{\widehat{z}} \cdot (A^\alpha \frac{\partial \widehat{w}}{\partial w} + B^\alpha) \right)^T \\ &= -\mu \left(B^{\alpha T} + (\frac{\partial \widehat{w}}{\partial w})^T A^{\alpha T} \right) \nabla H(\widehat{z}) \\ &= (A_\alpha J^{-1} - \frac{\partial \widehat{w}}{\partial w} C_\alpha J^{-1}) \nabla H(\widehat{z}) \qquad \text{(by (19) and } N \in \mathrm{Sm}(2n)) \\ &= \frac{\partial \widehat{w}}{\partial t}. \end{aligned}$$

So $d\omega^1 = 0$. By Poincare Lemma there exists, in some neighborhood of $\mathbf{R}^{2n+1}$, a scalar function $\phi_{\alpha,H}(w,t)$, such that

$$\omega^1 = \widehat{w} dw + \overline{H} dt = d\phi_{\alpha,H}(w,t),$$

i.e.,

$$f(w,t) = \nabla_w \phi_{\alpha,H}(w,t),$$
$$\frac{\partial}{\partial t}\phi = -\mu H(A^\alpha \nabla_w \phi_{\alpha,H}(w,t) + B^\alpha w). \qquad \text{Q.E.D.}$$

Examples of generating functions:

$$\text{(I)} \quad \alpha = \begin{bmatrix} 0 & 0 & -I_n & 0 \\ I_n & 0 & 0 & 0 \\ 0 & 0 & 0 & I_n \\ 0 & I_n & 0 & 0 \end{bmatrix}, \mu = 1, \quad M_0 = J, \quad |C_\alpha M_0 + D_\alpha| \neq 0;$$
$$w = \begin{bmatrix} q \\ \widehat{q} \end{bmatrix}, \quad \phi = \phi(q, \widehat{q}, t);$$
$$\widehat{w} = \begin{bmatrix} -p \\ \widehat{p} \end{bmatrix} = \begin{bmatrix} \phi_q \\ \phi_{\widehat{q}} \end{bmatrix}, \quad \phi_t = -H(\phi_{\widehat{q}}, \widehat{q}).$$

This is the generating function and H.J. equation of the first kind [3].

$$\text{(II)} \quad \alpha = \begin{bmatrix} 0 & 0 & -I_n & 0 \\ 0 & -I_n & 0 & 0 \\ 0 & 0 & 0 & I_n \\ I_n & 0 & 0 & 0 \end{bmatrix}, \mu = 1, \quad M_0 = I, \quad |C_\alpha M_0 + D_\alpha| \neq 0;$$
$$w = \begin{bmatrix} q \\ \widehat{p} \end{bmatrix}, \quad \phi = \phi(q, \widehat{p}, t);$$
$$\widehat{w} = -\begin{bmatrix} p \\ \widehat{q} \end{bmatrix} = \begin{bmatrix} \phi_q \\ \phi_{\widehat{p}} \end{bmatrix}, \quad \phi_t = -H(\widehat{p}, -\phi_{\widehat{p}}).$$

This is the generating function and H.J. equation of the second kind [3].

$$\text{(III)} \quad \alpha = \begin{bmatrix} -J_{2n} & J_{2n} \\ \frac{1}{2}I_{2n} & \frac{1}{2}I_{2n} \end{bmatrix}, \mu = 1, M_0 = I, \quad |C_\alpha M_0 + D_\alpha| \neq 0;$$
$$w = \frac{1}{2}(z + \widehat{z}), \quad \phi = \phi(w,t);$$
$$\widehat{w} = J(z - \widehat{z}) = \nabla\phi, \quad \phi_t = -H(w - \tfrac{1}{2}J^{-1}\nabla\phi).$$

This is the Poincare's generating function and H.J. equation.

If the Hamiltonian function $H(z)$ depends analytically on z then we can give through recursions the explicit expression of the corresponding generating functions.

Theorem 20. *Let $H(z)$ depend analytically on z. Then $\phi_{\alpha,H}(w,t)$ is expressible as a convergent power series in t for sufficiently small $|t|$, with recursively determined coefficients:*

$$\phi(w,t) = \sum_{k=0}^{\infty} \phi^{(k)}(w)t^k, \tag{42}$$

$$\phi^{(0)}(w) = \frac{1}{2}w^T N_0 w, \quad N_0 = (A_\alpha M_0 + B_\alpha)(C_\alpha M_0 + D_\alpha)^{-1}, \tag{43}$$

$$\phi^{(1)}(w) = -\mu(\alpha)H(E_0 w), \quad E_0 = A^\alpha N_0 + B^\alpha = M_0(C_\alpha M_0 + D_\alpha)^{-1}, \tag{44}$$

$$\begin{aligned} k \geq 1, \ \phi^{(k+1)}(w) = & -\frac{\mu(\alpha)}{k+1} \sum_{m=1}^{k} \frac{1}{m!} \sum_{i_1,\cdots,i_m=1}^{2n} H_{z_{i_1},\cdots,z_{i_m}}(E_0 w) \sum_{\substack{j_1+\cdots+j_m=k \\ j_l \geq 1}} \\ & \times (A^\alpha \nabla \phi^{(j_1)})_{i_1} \cdots (A^\alpha \nabla \phi^{(j_m)})_{i_m}, \end{aligned} \tag{45}$$

where $H_{z_{i_1},\cdots,z_{i_m}}(E_0 w)$ is the m-th partial derivative of $H(z)$ w.r.t. $z_{i_1},\cdots,z_{i_m}$, evaluated at $z = E_0 w$ and $(A^\alpha \nabla \phi^{(j_l)}(w))_{i_l}$ is the i_l-th component of the column vector $A^\alpha \nabla \phi^{(j_l)}(w)$.

Proof. Under our assumption, the generating function $\phi_{\alpha,H}(w,t)$ depends analytically on w and t in some neighborhood of $\mathbf{R}^{2n}$ and for small $|t|$. Expand it as a power series as follows

$$\phi(w,t) = \sum_{k=0}^{\infty} \phi^{(k)}(w)t^k.$$

Differentiating it with respect to w and t, we get

$$\nabla \phi(w,t) = \sum_{k=0}^{\infty} \nabla \phi^{(k)}(w)t^k \tag{46}$$

$$\frac{\partial}{\partial t}\phi = \phi_t(w,t) = \sum_{k=0}^{\infty}(k+1)t^k \phi^{(k+1)}(w). \tag{47}$$

By (33),

$$\nabla \phi^{(0)}(w) = \nabla \phi(w,0) = f(w,0) = N_0 w.$$

So we can take $\phi^{(0)}(w) = \frac{1}{2}w^T N_0 w$. Denote $E_0 = A^\alpha N_0 + B^\alpha$. Then

$$A^\alpha \nabla \phi(w,t) + B^\alpha w = E_0 w + \sum_{k=1}^{\infty} A^\alpha \nabla \phi^{(k)}(w)t^k.$$

Expanding $H(z)$ at $z = E_0 w$, we get

$$H(A^\alpha \nabla \phi(w,t) + B^\alpha w) = H(E_0 w + \sum_{k=1}^{\infty} A^\alpha \nabla \phi^{(k)}(w)t^k)$$

$$= H(E_0w) + \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{i_1,\cdots,i_m=1}^{2n} \sum_{j_1,\cdots,j_m=1}^{\infty} H_{z_{i_1},\cdots,z_{i_m}}(E_0w)t^{j_1+\cdots+j_m}.$$
$$\times(A^\alpha\nabla\phi^{(j_1)}(w))_{i_1}\cdots(A^\alpha\nabla\phi^{(j_m)}(w))_{i_m}$$
$$= H(E_0w) + \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{i_1,\cdots,i_m=1}^{2n} \sum_{k\geq m} t^k \sum_{\substack{j_1+\cdots+j_m=k \\ j_l\geq 1}} H_{z_{i_1},\cdots,z_{i_m}}(E_0w).$$
$$\times(A^\alpha\nabla\phi^{(j_1)}(w))_{i_1}\cdots(A^\alpha\nabla\phi^{(j_m)}(w))_{i_m}$$
$$= H(E_0w) + \sum_{k=1}^{\infty} t^k \sum_{m=1}^{k} \frac{1}{m!} \sum_{i_1,\cdots,i_m=1}^{2n} \sum_{\substack{j_1+\cdots+j_m=k \\ j_l\geq 1}} H_{z_{i_1},\cdots,z_{i_m}}(E_0w).$$
$$\times(A^\alpha\nabla\phi^{(j_1)})_{i_1}\cdots(A^\alpha\nabla\phi^{(j_m)})_{i_m}.$$

Substituting this into the R.H.S. of (34) and (47) into the L.H.S. of (34), then comparing the coefficients of t^k on both sides, we get the recursions (44) and (45). Q.E.D.

In the next section when we use generating functions $\phi_{\alpha,H}$ to construct difference schemes we always assume $M_0 = I$. For the sake of convenience, we restate Theorem 19 and Theorem 20 as follows.

Theorem 21. *Let $\alpha \in \mathrm{CSp}(\tilde{J}_{4n}, J_{4n})$. Let $z \to \hat{z} = g(z,t)$ be the phase flow of the Hamiltonian system (31) with Jacobian $M(z,t) = g_z(z,t)$. If*

$$|C_\alpha + D_\alpha| \neq 0, \tag{32'}$$

then there exists, for sufficiently small $|t|$ and in (some neighborhood of) $\mathbf{R}^{2n}$, a time-dependent gradient map $w \to \hat{w} = f(w,t)$ with Jacobian $N(w,t) = f_w(w,t) \in \mathrm{Sm}(2n)$ satisfying the transversality condition (17.3) and a time-dependent generating function $\phi_{\alpha,H}(w,t) = \phi(w,t)$ such that

1. $f(w,t) = \nabla\phi(w,t)$;
2. $\dfrac{\partial\phi}{\partial t} = -\mu H(A^\alpha\nabla\phi(w,t) + B^\alpha w)$,
3. $A_\alpha g(z,t) + B_\alpha z \equiv f(C_\alpha g(z,t) + D_\alpha z, t) \equiv \nabla\phi(C_\alpha g(z,t) + D_\alpha z, t)$,
4. $N = \sigma_\alpha(M) = (A_\alpha M + B_\alpha)(C_\alpha M + D_\alpha)^{-1}$

   $M = \sigma_{\alpha^{-1}}(N) = (A^\alpha N + B^\alpha)(C^\alpha N + D^\alpha)^{-1}$.

Theorem 22. *Let $H(z)$ depend analytically on z. Then $\phi_{\alpha,H}(w,t)$ is expressible as a convergent power series in t for sufficiently small $|t|$, with the recursively* *determined coefficients:*

$$\begin{aligned}
\phi(w,t) &= \sum_{k=0}^{\infty} \phi^{(k)}(w)t^k, \\
\phi^{(0)}(w) &= \frac{1}{2} w^T N_0 w, \qquad N_0 = (A_\alpha + B_\alpha)(C_\alpha + D_\alpha)^{-1}, \\
\phi^{(1)}(w) &= -\mu(\alpha) H(E_0 w), \qquad E_0 = (C_\alpha + D_\alpha)^{-1}, \\
k \geq 1,\ \phi^{(k+1)}(w) &= -\frac{\mu(\alpha)}{k+1} \sum_{m=1}^{k} \frac{1}{m!} \sum_{i_1,\cdots,i_m=1}^{2n} H_{z_{i_1},\cdots,z_{i_m}}(E_0 w) \sum_{\substack{j_1+\cdots+j_m=k \\ j_l \geq 1}} \\
&\quad \times (A^\alpha \nabla \phi^{(j_1)})_{i_1} \cdots (A^\alpha \nabla \phi^{(j_m)})_{i_m}.
\end{aligned}$$

§6. Construction of Canonical Difference Schemes

In this section we consider the construction of canonical difference schemes for the Hamiltonian system (31). By Theorem 18, for a given time-dependent scalar function $\psi(w,t) : \mathbf{R}^{2n} \times \mathbf{R} \to \mathbf{R}$, we can get a time-dependent canonical map $\tilde{g}(z,t)$. If $\psi(w,t)$ approximates some generating function $\phi_{\alpha,H}(w,t)$ of the Hamiltonian system (31) then $\tilde{g}(z,t)$ approximates the phase flow $g(z,t)$. Then fixing t as a time step, we can get a difference scheme –the canonical difference scheme–whose transition from one time–step to the next is canonical. By Theorem 22, generating functions $\phi(w,t)$ can be expressed as a power series. So a natural way to approximate $\phi(w,t)$ is to take the truncation of the series. More precisely, we have

Theorem 23. *Using Theorems 21 and 22, for sufficiently small $\tau > 0$ as the time-step, define*

$$\psi^{(m)}(w,\tau) = \sum_{i=0}^{m} \phi^{(i)}(w)\tau^i, \qquad m = 1,2,\cdots. \tag{48}$$

Then the gradient map

$$w \to \hat{w} = \tilde{f}(w,\tau) = \nabla \psi^{(m)}(w,\tau) \tag{49}$$

defines an implicit canonical difference scheme $z = z^k \to z^{k+1} = \hat{z}$,

$$A_\alpha z^{k+1} + B_\alpha z^k = \nabla \psi^{(m)}(C_\alpha z^{k+1} + D_\alpha z^k, \tau) \tag{50}$$

of m-th order of accuracy.

Proof. Since $\psi^{(m)}(w,0) = \phi(w,0)$, $\psi^{(m)}_{ww}(w,0) = \phi_{ww}(w,0) = f_w(w,0) = N(w,0)$ satisfies the transversality condition (17.3), i.e., $|C^\alpha N(w,0) + D^\alpha| \neq 0$. Thus for sufficiently small τ and in some neighborhood of $\mathbf{R}^{2n}$, $N^{(m)}(w,\tau) = \psi^{(m)}_{ww}(w,\tau)$ satisfies the transversality condition (17.3), i.e., $|C^\alpha N^{(m)}(w,\tau) + D^\alpha| \neq 0$. By

Theorem 18, the gradient map $w \to \hat{w} = \tilde{f}(w,\tau) = \nabla\psi^{(m)}(w,\tau)$ defines implicitly a time-dependent canonical map $z \to \hat{z} = \tilde{g}(z,\tau)$ by the equation

$$A_\alpha \hat{z} + B_\alpha z = \nabla\psi^{(m)}(C_\alpha \hat{z} + D_\alpha z, \tau).$$

That is, the equation

$$A_\alpha z^{k+1} + B_\alpha z^k = \nabla\psi^{(m)}(C_\alpha z^{k+1} + D_\alpha z^k, \tau)$$

is an implicit canonical difference scheme.

Since $\psi^{(m)}(w,\tau)$ is the m-th approximant to $\phi(w,\tau)$, so is $\tilde{f}(w,\tau) = \nabla\psi^{(m)}(w,\tau)$ to $f(w,\tau)$. It follows that the canonical difference scheme given by (50) is of m-th order of accuracy. Q.E.D.

It is not difficult to show that the generating function $\phi(w,t)$ of type (III) is odd in t. Hence Theorem 23 leads to a family of canonical difference schemes of arbitary even order accuracy.

Theorem 24. *Let* $\alpha = \begin{bmatrix} -J_{2n} & J_{2n} \\ \frac{1}{2}I_{2n} & \frac{1}{2}I_{2n} \end{bmatrix}$. *For sufficiently small $\tau > 0$ as the time-step, define*

$$\psi^{(2m)}(w,\tau) = \sum_{i=1}^{m} \phi^{(2i-1)}(w)\tau^{2i-1}, \quad m = 1, 2, \cdots. \tag{51}$$

Then the gradient map

$$w \to \hat{w} = \tilde{f}(w,\tau) = \nabla\psi^{(2m)}(w,\tau)$$

defines implicitly canonical difference schemes $z = z^k \to z^{k+1} = \hat{z}$,

$$z^{k+1} = z^k - J^{-1}\nabla\psi^{(2m)}(\tfrac{1}{2}(z^{k+1} + z^k), \tau) \tag{52}$$

of 2m-th order of accuracy. The case $m = 1$ is the Euler centered scheme.

For the linear Hamiltonian system (31) with the quadratic Hamiltonian $H(z) = \frac{1}{2}z^T S z$, $S \in \mathrm{Sm}(2n)$, the generating function of type (III) is the quadratic form

$$\begin{aligned} \phi(w,\tau) &= -\tfrac{1}{2}w^T(2J\tanh\tfrac{\tau}{2}L)w \\ &= -\sum_{i=1}^{\infty} a_{2i-1} w^T J(\frac{\tau}{2}L)^{2i-1} w, \qquad L = J^{-1}S, \end{aligned} \tag{52}$$

where

$$\tanh\lambda = \lambda - \tfrac{1}{3}\lambda^3 + \tfrac{2}{15}\lambda^5 - \tfrac{17}{312}\lambda^7 + \cdots = \sum_{i=1}^{\infty} a_{2i-1}\lambda^{2i-1},$$

$$a_{2i-1} = 2^{2i}(2^{2i} - 1)B_{2i}/(2i)!, \quad B_{2i}\text{—Bernoulli numbers},$$

$$J\tanh\tfrac{\tau}{2}L \in \mathrm{Sm}(2n).$$

We can easily get (53) by a simple way.

We know, the phase flow of the linear Hamiltonian system

$$\frac{dz}{dt} = J^{-1}Sz = Lz, \qquad L = J^{-1}S,$$

is e^{tL}. Set $z = z(0), \hat{z} = e^{\tau L}z(0) = e^{\tau L}z$. Then

$$\hat{w} = J(z - \hat{z}) = J(I - e^{\tau L})z, \quad w = \frac{1}{2}(z + \hat{z}) = \frac{1}{2}(I + e^{\tau L})z.$$

Hence

$$\begin{aligned} \hat{w} &= 2J(I - e^{\tau L})(I + e^{\tau L})^{-1}w = 2J\frac{e^{-\frac{\tau}{2}L} - e^{\frac{\tau}{2}L}}{e^{-\frac{\tau}{2}L} + e^{\frac{\tau}{2}L}}w \\ &= -2J\tanh(\tfrac{\tau}{2}L)w. \end{aligned}$$

Taking the truncation of (53),

$$\psi^{(2m)}(w,\tau) = -\sum_{i=1}^{m} a_{2i-1}w^T J(\frac{\tau}{2}L)^{2i-1}w,$$

we get the gradient map

$$\hat{w} = \nabla_w \psi^{(2m)} = -2J\sum_{i=1}^{m} a_{2i-1}(\frac{\tau}{2}L)^{2i-1}w.$$

Noting that $\hat{w} = J(z - \hat{z}), w = \frac{1}{2}(z + \hat{z})$, we have

$$J(z - \hat{z}) = -J\sum_{2i-1}^{m} a_{2i-1}(\frac{\tau}{2}L)^{2i-1}(z + \hat{z}).$$

So the $2m$-th order difference scheme is

$$z^{k+1} - z^k = \sum_{i=1}^{m} a_{2i-1}(\frac{\tau}{2}L)^{2i-1}(z^{k+1} + z^k).$$

Set $m = 1$, we get the Euler centered scheme.

If we take the diagonal Pade approximants to $\tanh\lambda$,

$$\frac{R_m(\lambda)}{Q_m(\lambda)} - \tanh\lambda = 0(|\lambda|^{2m+1}),$$

where

$$\begin{aligned} R_0(\lambda) &= 0, \\ R_1(\lambda) &= \lambda, \\ R_m(\lambda) &= (2m-1)R_{m-1}(\lambda) + \lambda^2 R_{m-2}(\lambda), \quad m = 2, 3, \cdots, \\ Q_0(\lambda) &= 1, \\ Q_1(\lambda) &= 1, \\ Q_m(\lambda) &= (2m-1)Q_{m-1}(\lambda) + \lambda^2 Q_{m-2}(\lambda), \quad m = 2, 3, \cdots, \end{aligned}$$

then we get another type symplectic difference schemes

$$z^{k+1}-z^k=\frac{R_m(\frac{\tau}{2}L)}{Q_m(\frac{\tau}{2}L)}(z^{k+1}+z^k),$$

i.e.,

$$z^{k+1}=\frac{Q_m(\frac{\tau}{2}L)+R_m(\frac{\tau}{2}L)}{Q_m(\frac{\tau}{2}L)-R_m(\frac{\tau}{2}L)}z^k. \tag{54}$$

Suppose that $P_m(\lambda)/P_m(-\lambda)$ is the diagonal Pade approximant to e^λ where

$$\begin{aligned}&P_0(\lambda)=1,\\&P_1(\lambda)=2+\lambda,\\&P_m(\lambda)=2(2m-1)P_{m-1}(\lambda)+\lambda^2P_{m-2}(\lambda),\quad m=2,3,\cdots.\end{aligned}$$

Then because

$$\tanh\ \lambda=\frac{e^\lambda-e^{-\lambda}}{e^\lambda+e^{-\lambda}}=\frac{-1+e^{2\lambda}}{1+e^{2\lambda}},$$

the diagonal Pade approximant $R_m(\lambda)/Q_m(\lambda)$ to $\tanh\lambda$ is

$$\frac{R_m(\lambda)}{Q_m(\lambda)}=\frac{-1+P_m(2\lambda)/P_m(-2\lambda)}{1+P_m(2\lambda)/P_m(-2\lambda)}=\frac{P_m(2\lambda)-P_m(-2\lambda)}{P_m(2\lambda)+P_m(-2\lambda)}.$$

Hence

$$R_m(\lambda)=c_m(P_m(2\lambda)-P_m(-2\lambda)),\qquad Q_m(\lambda)=c_m(P_m(2\lambda)+P_m(-2\lambda)),$$

where $c_m=const\neq 0$. It follows that

$$P_m(2\lambda)=\frac{1}{2c_m}(Q_m(\lambda)+R_m(\lambda)),\qquad P_m(-2\lambda)=\frac{1}{2c_m}(Q_m(\lambda)-R_m(\lambda)).$$

In fact $c_m=2^{-m}$. So (54) becomes

$$z^{k+1}=\frac{P_m(\tau L)}{P_m(-\tau L)}z^k.$$

Examples of canonical difference schemes.

$$\begin{cases} p_i^{k+1} = p_i^k - \tau H_{q_i}(p^{k+1}, q^k), \\ q_i^{k+1} = q_i^k + \tau H_{p_i}(p^{k+1}, q^k). \end{cases} \quad i = 1, \cdots, n. \tag{55}$$

When H is separable, $H = U(p) + V(q)$, $H_{q_i}(p^{k+1}, q^k) = V_{q_i}(q^k)$, $H_{p_i}(p^{k+1}, q^k) = U_{p_i}(p^{k+1})$. At this time, (55) becomes

$$\begin{cases} p_i^{k+1} = p_i^k - \tau V_{q_i}(q^k), \\ q_i^{k+1} = q_i^k + \tau U_{p_i}(p^{k+1}), \end{cases} \quad i = 1, \cdots, n. \tag{56}$$

Evidently, (56) is an explicit difference scheme of 1–st order of accuracy. If we set $q's$ at half–integer times $t = (k + \frac{1}{2})\tau$, then (56) becomes

$$\begin{cases} p_i^{k+1} = p_i^k - \tau V_{q_i}(q^{k+\frac{1}{2}}), \\ q_i^{k+\frac{1}{2}+1} = q_i^{k+\frac{1}{2}} + \tau U_{p_i}(p^{k+1}). \end{cases} \quad i = 1, \cdots, n. \tag{57}$$

(57) is a staggered explicit scheme of 2–nd order of accuracy.

b. Second order scheme.

$$\psi^{(2)}(w, \tau) = \psi^{(1)}(w) + \tau^2 \phi^{(2)}(w).$$

The induced gradient map is

$$\hat{w} = \nabla_w \psi^{(2)} = -\begin{bmatrix} \hat{p} \\ q \end{bmatrix} - \tau \begin{bmatrix} \nabla_q H \\ \nabla_p H \end{bmatrix} - \frac{\tau^2}{2} \begin{bmatrix} \nabla_q \sum\limits_{i=1}^n H_{q_i} H_{p_i} \\ \nabla_p \sum\limits_{i=1}^n H_{q_i} H_{p_i} \end{bmatrix}.$$

So the second order scheme is

$$\begin{cases} p_i^{k+1} = p_i^k - \tau H_{q_i}(p^{k+1}, q^k) - \frac{\tau^2}{2} \Big(\sum\limits_{j=1}^n H_{q_j} H_{p_j}\Big)_{q_i}(p^{k+1}, q^k) \\ q_i^{k+1} = q_i^k + \tau H_{p_i}(p^{k+1}, q^k) + \frac{\tau^2}{2} \Big(\sum\limits_{j=1}^n H_{q_j} H_{p_j}\Big)_{p_i}(p^{k+1}, q^k) \end{cases} \quad i = 1, \cdots, n.$$

This scheme is already implicit even when $H(z)$ is separable.

c. The third order scheme is, for $i = 1, \cdots, n$,

$$\begin{aligned} p_i^{k+1} = \ & p_i^k - \tau H_{q_i}(p^{k+1}, q^k) - \frac{\tau^2}{2} \sum_{j=1}^n (H_{q_j} H_{p_j})_{q_i}(p^{k+1}, q^k) \\ & - \frac{\tau^3}{6} \sum_{l,j=1}^n (H_{p_l p_j} H_{q_l} H_{q_j} + H_{q_l q_j} H_{p_l} H_{p_j} + H_{p_l q_j} H_{q_l} H_{p_j})_{q_i}(p^{k+1}, q^k), \\ q_i^{k+1} = \ & q_i^k + \tau H_{p_i}(p^{k+1}, q^k) + \frac{\tau^2}{2} \sum_{j=1}^n (H_{q_j} H_{p_j})_{p_i}(p^{k+1}, q^k) \\ & + \frac{\tau^3}{6} \sum_{l,j=1}^n (H_{p_l p_j} H_{q_l} H_{q_j} + H_{q_l q_j} H_{p_l} H_{p_j} + H_{p_l q_j} H_{q_l} H_{p_j})_{p_i}(p^{k+1}, q^k). \end{aligned}$$

By Theorem 22, as $\mu = 1$,

$$\begin{aligned}
\phi^{(0)}(w) &= \tfrac{1}{2}w^T N_0 w, \qquad N_0 = (A_\alpha + B_\alpha)(C_\alpha + D_\alpha)^{-1},\\
\phi^{(1)}(w) &= -H(E_0 w), \qquad E_0 = (C_\alpha + D_\alpha)^{-1},\\
\phi^{(2)}(w) &= \tfrac{1}{2}(\nabla H)^T A^\alpha E_0^T (\nabla H)(E_0 w),\\
\phi^{(3)}(w) &= -\tfrac{1}{3}(\nabla H)^T A^\alpha \nabla_w \phi^{(2)} - \tfrac{1}{6}(A^\alpha \nabla \phi^{(1)})^T H_{zz}(A^\alpha \nabla \phi^{(1)})\\
&= -\tfrac{1}{6}(\nabla H)^T A^\alpha (E_0^T H_{zz} A^\alpha E_0^T \nabla H + E_0^T H_{zz} E_0 A^{\alpha T} \nabla H)\\
&\quad -\tfrac{1}{6}(\nabla H)^T E_0 A^{\alpha T} H_{zz} A^\alpha E_0^T \nabla H\\
&= -\tfrac{1}{6}\{(\nabla H)^T A^\alpha E_0^T H_{zz}(A^\alpha E_0^T + E_0 A^{\alpha T})\nabla H\\
&\quad +(\nabla H)^T E_0 A^{\alpha T} H_{zz} A^\alpha E_0^T \nabla H\}.
\end{aligned}$$

Here we use, instead of the component notation in Theorem 22, the matrix notation, H_{zz} denotes the Hessian matrix of H; all derivatives of H are evaluated at $z = E_0 w$.

Type (II) .

$$\alpha = \begin{bmatrix} 0 & 0 & -I_n & 0 \\ 0 & -I_n & 0 & 0 \\ 0 & 0 & 0 & I_n \\ I_n & 0 & 0 & 0 \end{bmatrix}, \qquad \alpha^T = \alpha^{-1} = \begin{bmatrix} 0 & 0 & 0 & I_n \\ 0 & -I_n & 0 & 0 \\ -I_n & 0 & 0 & 0 \\ 0 & 0 & I_n & 0 \end{bmatrix};$$

$$w = \begin{bmatrix} q \\ \widehat{p} \end{bmatrix}, \qquad \widehat{w} = -\begin{bmatrix} p \\ \widehat{q} \end{bmatrix}.$$

$$N_0 = -\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}, \qquad E_0 = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}, \qquad A^\alpha E_0^T = -\begin{bmatrix} 0 & 0 \\ I & 0 \end{bmatrix},$$

$$\phi^{(1)}(w) = -H(\widehat{p}, q),$$

$$\phi^{(2)}(w) = -\frac{1}{2}\sum_{i=1}^{n}(H_{q_i} H_{p_i})(\widehat{p}, q),$$

$$\phi^{(3)}(w) = -\frac{1}{6}\sum_{i,j=1}^{n}(H_{p_i p_j} H_{q_i} H_{q_j} + H_{q_i q_j} H_{p_i} H_{p_j} + H_{q_i p_j} H_{p_i} H_{q_j})$$

where $H(z) = H(p_1, \cdots, p_n, q_1, \cdots, q_n)$, $H_{z_i} = \partial H / \partial z_i$.

a. First order scheme.

$$\psi^{(1)}(w, \tau) = \phi^{(0)}(w) + \tau \phi^{(1)}(w).$$

The equation $\widehat{w} = \nabla \psi^{(1)}(w, \tau)$ defines a first order canonical difference scheme

Type (III)

$$\alpha = \begin{bmatrix} -J_{2n} & J_{2n} \\ \frac{1}{2}I_{2n} & \frac{1}{2}I_{2n} \end{bmatrix}, \qquad \alpha^{-1} = \begin{bmatrix} \frac{1}{2}J_{2n} & I_{2n} \\ -\frac{1}{2}J_{2n} & I_{2n} \end{bmatrix}.$$

$$w = \frac{1}{2}(\hat{z} + z), \qquad \hat{w} = J(z - \hat{z}).$$

$$N_0 = 0, \qquad E_0 = I, \qquad A^{\alpha}E_0^T + E_0A^{\alpha T} = 0.$$

$$\phi^{(0)} = \phi^{(2)} = \phi^{(4)} = 0,$$

$$\phi^{(1)}(w) = -H(\frac{1}{2}(\hat{z} + z)),$$

$$\phi^{(3)}(w) = \frac{1}{24}(\nabla H)^T J H_{zz} J \nabla H,$$

$$\psi^{(2)}(w,\tau) = -\tau H,$$

$$\psi^{(4)}(w,\tau) = -\tau H + \frac{\tau^3}{24}(\nabla H)^T J H_{zz} J \nabla H.$$

By Theorem 24 the second order scheme is

$$J(z - \hat{z}) = \hat{\omega} = \nabla_\omega \psi^{(2)}(\omega,\tau) = -\tau \nabla H(\frac{1}{2}(z + \hat{z})),$$

ie.,

$$z^{k+1} = z^k + \tau J^{-1}\nabla H(\frac{1}{2}(z^{k+1} + z^k)).$$

The 4-th order scheme is

$$J(z - \hat{z}) = \hat{\omega} = \nabla_\omega \psi^{(4)}(\omega,\tau) = -\tau \nabla H(\frac{1}{2}(z + \hat{z})) + \frac{\tau^2}{24}\nabla_z((\nabla H)^T J H_{zz} J \nabla H)$$

i.e.,

$$z^{k+1} = z^k + \tau J^{-1}\nabla H(\frac{1}{2}(z^{k+1} + z^k)) - \frac{\tau^3}{24}J^{-1}\nabla_z((\nabla H)^T J H_{zz} J \nabla H)(\frac{1}{2}(z^{k+1} + z^k)).$$

References

[1] Feng Kang, On Difference schemes and symplectic geometry, Proceedings of the 1984 Beijing Symposium on Differential Geometry and Differential Equations—Computation of partial differential equations, Ed. Feng Kang, Science Press, Beijing, 1985, 42–58.

[2] R. Abraham, J. Marsden, Foundations of Mechanics, Addison-Wesley, Reading, Mass, 1978.

[3] V. I. Arnold, Mathematical Methods of Classical Mechanics, New York, 1978.

[4] E. Artin, Geometrical Algebra, New York, 1957.

[5] G. A. Baker, Jr., P. Graves-Morris, Pade Approximants, Part I: Basic Theory, Addison-Wesley, London, 1981.

[6] Feng Kang, Wu Hua-mo, Qin Meng-zhao, Symplectic difference schemes for the linear Hamiltonian canonical systems, to appear.

[7] L. K. Hua, On the theory of automorphic functions of a matrix variable I, II, *Amer. J. Math.*, 66(1944), 470–488, 531–563.

[8] C. K. Siegel, Symplectic geometry, *Amer. J. Math.*, 65(1943), 1–86.