DISCRETE–TIME HAMILTONIAN DYNAMICS

by

YOSI SHIBBERU

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

August 1992

ACKNOWLEDGEMENTS

Professor Greenspan's discrete mechanics is the inspiration for the discrete–time theory of Hamiltonian systems proposed in this thesis. I gained valuable insight from Professor Greenspan's lectures on numerical analysis and discrete model theory. The personal interest he has taken in me and all of his students is most appreciated.

July 22, 1992

ABSTRACT

DISCRETE-TIME HAMILTONIAN DYNAMICS

Publication No. ______

Yosi Shibberu, Ph. D.

The University of Texas at Arlington, 1992

Supervising Professor: Donald Greenspan

A discrete–time theory with properties similar to properties of the discrete mechanics of Greenspan and the discrete mechanics of Lee is proposed for Hamiltonian dynamical systems. Equations applicable to arbitrary Hamiltonian systems are derived from a discrete variational principle. The equations completely determine piecewise–linear, continuous trajectories which exactly conserve the Hamiltonian function at the midpoints of each linear segment and exactly conserve, at the vertices, all conserved quadratic functions. For autonomous, positive–definite, linear systems, the equations have solutions identical to solutions obtained by the trapezoidal and midpoint methods. Existence and uniqueness results are presented along with some preliminary work on the coordinate invariance of the theory.

TABLE OF CONTENTS

ILLUSTRATIONS

Figure | Page

TABLES

INTRODUCTION

Hamiltonian systems are used in a wide variety of applications ranging in scope from quantum mechanics to optimal control theory. Computational methods which preserve their special structure are, therefore, of considerable interest.

Newtonian potential systems, a subclass of Hamiltonian systems, can be simulated by using the discrete mechanics equations developed by Greenspan [6]. These equations are invariant with respect to rotation, translation and uniform motion. For the Kepler problem, for example, the equations exactly conserve energy and angular momentum in cartesian coordinates. Labudde [8] has extended the discrete mechanics of Greenspan to include a wide variety of Hamiltonian systems.

More recently, using a Lagrangian formulation, T. D. Lee [10] has developed a discrete mechanics in which trajectories in the configuration space of a system are assumed to be piecewise–linear and continuous. The average value of the energy of a system over each linear segment of the trajectory is conserved at each time step. A distinctive feature of this discrete mechanics is that time plays the role of a dynamic variable.

Symplectic integration schemes have become an increasingly popular way to integrate Hamiltonian systems. Wu [12] has shown that these schemes admit a natural discrete variational principle. In this way, symplectic schemes may be viewed as a type of discrete mechanics.

Discrete mechanics schemes are distinguished from conventional numerical schemes in that they are based on fundamental principles as opposed to approximations of differential equations derived from continuum mechanics. In fact, Lee [10] suggests that discrete mechanics may be even more fundamental than continuum mechanics. The finiteness of physical reality and the dilemmas that the concept of infinity introduces in the continuum theory have been pointed out by Greenspan [6]. These dilemmas do not occur in the discrete theory.

In this thesis, we propose a general discrete–time theory for Hamiltonian systems. The theory is based on a new variational principle which completely determines piecewise–linear, continuous trajectories in the extended phase space of a Hamiltonian system. In the spirit of Hamiltonian dynamics, the theory is completely symmetrical in the way it treats position and momentum. Like the discrete mechanics of Greenspan, the theory exactly conserves energy and conserved quadratic functions such as angular momentum. Like the discrete mechanics of Lee, the theory treats time as a dependent variable. For the simple harmonic oscillator, the theory reduces to the conventional trapezoidal and midpoint schemes commonly used to integrate differential equations.

In Chapter I, we review the variational principles of continuum mechanics and then introduce the basic ideas of the theory for the case of autonomous systems with one degree of freedom. In Chapter II, a discrete variational principle is proposed and used as the basis of the theory. Equations of dynamics are derived and conservation laws are proved. In Chapter III, local existence and uniqueness results are given for general Hamiltonian systems and global results are given for autonomous, positive–definite, linear systems. In Chapter IV, we present numerical results for the Kepler problem (also known as the one body central force problem). Finally, in Chapter V, we discuss some preliminary results on the coordinate invariance of the theory.

CHAPTER I

AUTONOMOUS SYSTEMS WITH ONE DEGREE OF FREEDOM

1.1 Variational Principles of Mechanics

Hamilton's principle is probably the most widely known variational principle of mechanics. This principle states that the integral $\mathcal{I}$ given by (1.1.1) is stationary for the trajectory $q(t)$ of a dynamical system with Lagrangian function $L(q, \dot{q})$ [3].

$$\mathcal{I} = \int_{t_1}^{t_2} L(q(t), \dot{q}(t))\, dt \qquad (1.1.1)$$

The principle of least action is another variational principle of mechanics. The following motivation for the principle of least action is based on [9]. Consider now Legendre's transformation $\dot{q} \to p$ and $L(q, \dot{q}) \to H(q, p)$ where p and $H(q, p)$ are given by:

$$p = \frac{\partial L(q, \dot{q})}{\partial \dot{q}} \qquad (1.1.2)$$

$$H(q, p) = p\,\dot{q}(p) - L(q, \dot{q}(p)) \qquad (1.1.3)$$

For the problems to be considered, $\dot{q}(p)$ in (1.1.3) can be obtained by solving for $\dot{q}$ in (1.1.2). Under Legendre's transformation $\mathcal{I}$ can be expressed as:

$$\mathcal{I} = \int_{t_1}^{t_2} \left[\, p\,\dot{q} - H(q, p) \,\right] dt \qquad (1.1.4)$$

Consider a reparametrization of time given by $t = t(\tau)$. With this reparametrization (1.1.4) becomes:

$$\mathcal{I} = \int_{\tau_1}^{\tau_2} \left[p\,\frac{dq/d\tau}{dt/d\tau} - H(q, p) \right] \frac{dt}{d\tau}\; d\tau$$

Multiplying by $\frac{dt}{d\tau}$ we have:

$$\mathcal{S} \;=\; \int_{\tau_1}^{\tau_2}\left[\,p\frac{dq}{d\tau} - H(q,\,p)\,\frac{dt}{d\tau}\right]d\tau \tag{1.1.5}$$

Define $\wp = -H(q, p)$ and substitute $\wp$ in (1.1.5).

$$\mathcal{S} \;=\; \int_{\tau_1}^{\tau_2}\left[\,p\frac{dq}{d\tau} + \wp\,\frac{dt}{d\tau}\right]d\tau \tag{1.1.6}$$

The integral (1.1.6) is called the action integral of a Hamiltonian dynamical system. The structure of (1.1.6) suggests that just as p is the momentum coordinate corresponding to q, $\wp$ is the momentum coordinate corresponding to t. (It is important to note that the variable t in (1.1.6) is a dependent variable, the independent variable being τ.)

DEFINITION 1.1: (Action Integral)
Assume $p(\cdot)$, $q(\cdot)$, $\wp(\cdot)$, and $t(\cdot)$ are differentiable functions of τ on the interval $[\tau_0,\ \tau_N]$ where q and p are the position and momentum coordinates of a Hamiltonian dynamical system. The action integral of a Hamiltonian dynamical system is defined to be:

$$A(p(\cdot),\, q(\cdot),\, \wp(\cdot),\, t(\cdot)) \;=\; \int_{\tau_0}^{\tau_N}\left[\,p(\tau)\frac{dq(\tau)}{d\tau} + \wp(\tau)\frac{dt(\tau)}{d\tau}\right]d\tau \tag{1.1.7}$$

The trajectory of a Hamiltonian dynamical system can be obtained from the following principle.

DEFINITION 1.2: (Principle of Least Action)
The trajectory of a Hamiltonian dynamical system with Hamiltonian function $H(q, p)$ is given by functions $p(\cdot)$, $q(\cdot)$, $\wp(\cdot)$, and $t(\cdot)$ which cause the action integral to be stationary under the constraint:

$$\wp + H(q,\,p) \;=\; 0 \tag{1.1.8}$$

The endpoints of $q(\cdot)$ and $t(\cdot)$ are to be specified.

The constraint (1.1.8) in Definition 1.2 is necessary because $\wp$ has been defined to be equal to $-H(q, p)$ in (1.1.6).

In the following theorem, the principle of least action is used to derive the dynamical equations of a Hamiltonian system. A discrete version of the principle of least action will be used in Theorem 1.9 to derive equations for discrete–time dynamics.

THEOREM 1.3:
The trajectory of a Hamiltonian dynamical system with Hamiltonian function $H(q, p)$ and initial conditions $q(\tau_o) = q_o$, $p(\tau_o) = p_o$, $t(\tau_o) = 0$, and $\wp(\tau_o) = -H(q_o, p_o)$ is a solution of the following system of differential equations:

$$\frac{dq}{d\tau} = \lambda(\tau)\frac{\partial H(q, p)}{\partial p} \tag{1.1.9}$$

$$\frac{dp}{d\tau} = -\lambda(\tau)\frac{\partial H(q, p)}{\partial q} \tag{1.1.10}$$

$$\frac{dt}{d\tau} = \lambda(\tau) \tag{1.1.11}$$

$$\frac{d\wp}{d\tau} = 0 \tag{1.1.12}$$

The function $\lambda(\tau)$ is an arbitrary function which determines the parametrization of the trajectory.

Proof:

The principle of least action states that the trajectory of a Hamiltonian system causes the action integral (1.1.7) to be stationary when subject to the constraint (1.1.8). Define:

$$g(q, p, \wp) = \wp + H(q, p) \tag{1.1.13}$$

$$f(p(\cdot), q(\cdot), \wp(\cdot), t(\cdot), \lambda(\cdot)) = A(p(\cdot), q(\cdot), \wp(\cdot), t(\cdot)) - \int_{\tau_o}^{\tau_N} \lambda(\tau)\, g\Big(q(\tau), p(\tau), \wp(\tau)\Big)\, d\tau \tag{1.1.14}$$

where $\lambda(\tau)$ is a differentiable function of τ. From (1.1.7), (1.1.13) and (1.1.14)

$$f(p(\cdot), q(\cdot), \wp(\cdot), t(\cdot), \lambda(\cdot)) = \int_{\tau_o}^{\tau_N} L(q(\tau), q'(\tau), p(\tau), t'(\tau), \wp(\tau), \lambda(\tau))\, d\tau \qquad (1.1.15)$$

where $$L(q, q', p, t', \wp, \lambda) = \Big[\, p\, q' + \wp\, t' - \lambda\big(\wp + H(q, p)\big)\Big] \qquad (1.1.16)$$

and where the notation q' and t' has been used for $\dfrac{dq(\tau)}{d\tau}$ and $\dfrac{dt(\tau)}{d\tau}$. The action integral subject to $g(q, p, \wp) = 0$ is stationary when the functional f given by (1.1.15) is stationary. The functional f is stationary when the following Euler–Lagrange equations are satisfied:

$$\frac{d}{d\tau}\Big(\frac{\partial L}{\partial p'}\Big) - \frac{\partial L}{\partial p} = 0 \qquad (1.1.17)$$

$$\frac{d}{d\tau}\Big(\frac{\partial L}{\partial q'}\Big) - \frac{\partial L}{\partial q} = 0 \qquad (1.1.18)$$

$$\frac{d}{d\tau}\Big(\frac{\partial L}{\partial \wp'}\Big) - \frac{\partial L}{\partial \wp} = 0 \qquad (1.1.19)$$

$$\frac{d}{d\tau}\Big(\frac{\partial L}{\partial t'}\Big) - \frac{\partial L}{\partial t} = 0 \qquad (1.1.20)$$

$$\frac{d}{d\tau}\Big(\frac{\partial L}{\partial \lambda'}\Big) - \frac{\partial L}{\partial \lambda} = 0 \qquad (1.1.21)$$

Substituting the Lagrangian function (1.1.16) in equations (1.1.17) – (1.1.20) we arrive at equations (1.1.9) – (1.1.12). Substituting (1.1.16) in (1.1.21) results in the identity:

$$\wp(\tau) + H(q(\tau), p(\tau)) \equiv 0 \qquad (1.1.22)$$

<u>Claim</u>: (1.1.9) – (1.1.12) imply (1.1.22) independently of equations (1.1.21).

Proof of the claim:

$$\frac{dH(q(\tau), p(\tau))}{d\tau} = \frac{\partial H}{\partial q}\frac{dq}{d\tau} + \frac{\partial H}{\partial p}\frac{dp}{d\tau} \qquad (1.1.23)$$

Substituting (1.1.9) and (1.1.10) in (1.1.23) we have that

$$\frac{dH}{d\tau} = \lambda(\tau)\Big(\frac{\partial H}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial p}\frac{\partial H}{\partial q}\Big) \equiv 0$$

Therefore, (1.1.9) and (1.1.10) imply that $H(q(\tau), p(\tau))$ is constant. Equation (1.1.12) implies $\wp(\tau)$ is constant also. We have then that:

$$\wp(\tau) + H[q(\tau), p(\tau)] \equiv \wp(\tau_o) + H[q(\tau_o), p(\tau_o)] = 0$$

as claimed, because from the initial conditions, $q(\tau_0) = q_0$, $p(\tau_0) = p_0$ and $\wp(\tau_o) = -H(q_o, p_o)$. ∎

The claim in the proof of Theorem 1.3 shows that for the Lagrangian (1.1.16) equation (1.1.21) is not independent from equations (1.1.17) – (1.1.20). This dependence is the reason for the arbitrary nature of $\lambda(\tau)$ in equations (1.1.9) – (1.1.12). The situation is very different for the discrete-time theory to be discussed shortly.

1.2 Piecewise-Linear Continuous Functions

Assume the points τ_n, $n = 0, 1, \cdots N$, partition the interval $[\tau_0, \tau_N]$ into N equal intervals of length $\Delta\tau$.

$$\tau_n = \tau_0 + n\Delta\tau \qquad n = 0, 1, \cdots N \tag{1.2.1}$$

where

$$\Delta\tau = \frac{\tau_N - \tau_0}{N} \tag{1.2.2}$$

Let $\hat{x}(\cdot)$ be a piecewise-linear, continuous function of τ as shown in Figure 1.1. Define:

$$x_n = \hat{x}(\tau_n) \qquad n = 0, 1, \cdots N \tag{1.2.3}$$

These x_n's will be called vertices of $\hat{x}(\cdot)$. Clearly, $\hat{x}(\cdot)$ is completely determined by its vertices. Define:

$$\bar{x}_n = \bar{x}_n(x_{n+1}, x_n) = \frac{x_{n+1} + x_n}{2} \qquad n = 0, 1, \cdots N-1 \tag{1.2.4}$$

$$\bar{x}'_n = \bar{x}'_n(x_{n+1}, x_n) = \frac{x_{n+1} - x_n}{\Delta\tau} \qquad n = 0, 1, \cdots N-1 \tag{1.2.5}$$

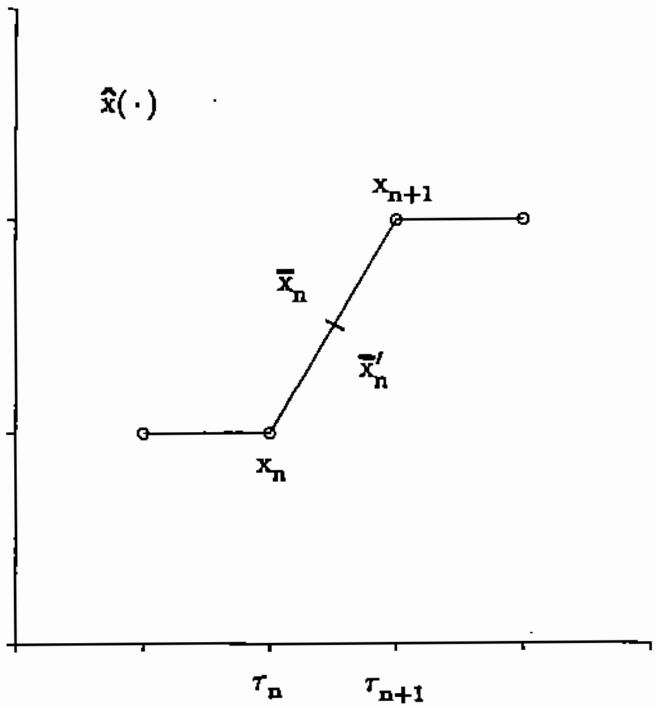


FIGURE 1.1: A piecewise–linear continuous function.

(Note that $\overline{x}'_n$ in (1.2.5) is not the derivative of $\overline{x}_n$ and that both $\overline{x}_n$ and $\overline{x}'_n$ are defined at the midpoints of the partition of $[\tau_o, \tau_N]$.) Since $\hat{x}(\cdot)$ is piecewise–linear, it can be expressed in terms of the values of $\overline{x}_n$ and $\overline{x}'_n$ in the following way:

$$\hat{x}(\tau) = \begin{cases} \overline{x}_n + \overline{x}'_n (\tau - \overline{\tau}_n) & \tau_n \leq \tau < \tau_{n+1} \qquad n = 0, 1, \cdots N-1 \\ x_N & \tau = \tau_N \end{cases} \tag{1.2.6}$$

where
$$\overline{\tau}_n = \frac{\tau_{n+1} + \tau_n}{2}$$

Thus, $\hat{x}(\cdot)$ is completely determined by the values of $\overline{x}_n$ and $\overline{x}'_n$, $n = 0, 1, \cdots N-1$. Since $\hat{x}(\cdot)$ is continuous, $\overline{x}_n$ and $\overline{x}'_n$ must satisfy the following continuity constraint.

<u>LEMMA 1.4</u>: (Continuity Constraint)
A piecewise–linear, continuous function $\hat{x}(\cdot)$ must satisfy the continuity constraint:

$$\frac{\overline{x}_{n+1} - \overline{x}_n}{\Delta\tau} = \frac{\overline{x}'_{n+1} + \overline{x}'_n}{2} \tag{1.2.7}$$

Proof:

Since $\hat{x}(\cdot)$ is continuous,

$$\lim_{\tau \to \tau_{n+1}^-} \hat{x}(\tau) = \hat{x}(\tau_{n+1}) \tag{1.2.8}$$

But from (1.2.6) the left hand side of (1.2.8) is:

$$\lim_{\tau \to \tau_{n+1}^-} \hat{x}(\tau) = \lim_{\tau \to \tau_{n+1}^-} \left[\overline{x}_n + \overline{x}'_n(\tau - \overline{\tau}_n)\right] = \overline{x}_n + \overline{x}'_n\left(\frac{\Delta\tau}{2}\right)$$

while the right hand side of (1.2.8) is:

$$\hat{x}(\tau_{n+1}) = \overline{x}_{n+1} + \overline{x}'_{n+1}\left(\frac{-\Delta\tau}{2}\right)$$

Therefore,

$$\overline{x}_n + \overline{x}'_n\left(\frac{\Delta\tau}{2}\right) = \overline{x}_{n+1} + \overline{x}'_{n+1}\left(\frac{-\Delta\tau}{2}\right)$$

from which it follows that:

$$\frac{\overline{x}_{n+1} - \overline{x}_n}{\Delta\tau} = \frac{\overline{x}'_{n+1} + \overline{x}'_n}{2}$$

∎

Because each of the steps in the proof of Lemma 1.4 is reversible, equation (1.2.7) is also a sufficient condition for the continuity of piecewise–linear functions.

The following lemma will be used in the proof of Theorem 1.9.

LEMMA 1.5:
From (1.2.4) and (1.2.5) it follows that:

$$\frac{\partial\overline{x}_n}{\partial x_n} = \frac{1}{2}, \quad \frac{\partial\overline{x}'_n}{\partial x_n} = -\frac{1}{\Delta\tau}, \quad \frac{\partial\overline{x}_n}{\partial x_{n+1}} = \frac{1}{2}, \quad \frac{\partial\overline{x}'_n}{\partial x_{n+1}} = \frac{1}{\Delta\tau} \qquad n = 0, 1, \cdots N-1$$

1.3 Discrete Principle of Least Action

DEFINITION 1.6: (Discrete Action)
Given a partition of τ from τ_o to τ_N, the discrete action A_N of a Hamiltonian system is defined to be:

$$A_N(p_0 \cdots p_N, q_0 \cdots q_N, \wp_0 \cdots \wp_N, t_0 \cdots t_N) = \sum_{i=0}^{N-1} \left[\overline{p}_i \overline{q}_i' + \overline{\wp}_i \overline{t}_i' \right] \Delta\tau$$

where

$$\overline{p}_i = \overline{p}_i(p_{i+1}, p_i) = \frac{p_{i+1} + p_i}{2}$$

$$\overline{q}_i' = \overline{q}_i'(q_{i+1}, q_i) = \frac{q_{i+1} - q_i}{\Delta\tau}$$

$$\overline{\wp}_i = \overline{\wp}_i(\wp_{i+1}, \wp_i) = \frac{\wp_{i+1} + \wp_i}{2}$$

$$\overline{t}_i' = \overline{t}_i'(t_{i+1}, t_i) = \frac{t_{i+1} - t_i}{\Delta\tau}$$

The following theorem shows that for piecewise–linear, continuous functions, the action integral given by Definition 1.1 is exactly equal to the discrete action.

THEOREM 1.7:
For piecewise–linear, continuous functions, the action integral and the discrete action are equal.

$$A(\hat{p}(\cdot), \hat{q}(\cdot), \hat{\wp}(\cdot), \hat{t}(\cdot)) = A_N(p_0 \cdots p_N, q_0 \cdots q_N, \wp_0 \cdots \wp_N, t_0 \cdots t_N)$$

Proof:

From Definition 1.1

$$A(\hat{p}(\cdot), \hat{q}(\cdot), \hat{\wp}(\cdot), \hat{t}(\cdot)) = \int_{\tau_0}^{\tau_N} \left[\hat{p}(\tau) \frac{d\hat{q}(\tau)}{d\tau} + \hat{\wp}(\tau) \frac{d\hat{t}(\tau)}{d\tau} \right] d\tau$$

Since $\hat{q}(\tau)$ and $\hat{t}(\tau)$ are piecewise linear, from (1.2.6) it follows that $\frac{d\hat{q}(\tau)}{d\tau} = \overline{q}_i'$ and $\frac{d\hat{t}(\tau)}{d\tau} = \overline{t}_i'$ for $\tau_i < \tau < \tau_{i+1}$, $i = 0, 1, \cdots N-1$. Therefore:

$$A[\hat{p}(\cdot), \hat{q}(\cdot), \hat{\wp}(\cdot), \hat{t}(\cdot)] = \sum_{i=0}^{N-1} \left[\overline{q}_i' \int_{\tau_i}^{\tau_{i+1}} \hat{p}(\tau)\, d\tau + \overline{t}_i' \int_{\tau_i}^{\tau_{i+1}} \hat{\wp}(\tau)\, d\tau \right]$$

Since

$$\int_{\tau_i}^{\tau_{i+1}} \hat{p}(\tau)\,d\tau = \bar{p}_i \Delta\tau$$

and

$$\int_{\tau_i}^{\tau_{i+1}} \hat{\wp}(\tau)\,d\tau = \bar{\wp}_i \Delta\tau$$

we have

$$A(\hat{p}(\cdot), \hat{q}(\cdot), \hat{\wp}(\cdot), \hat{t}(\cdot)) = \sum_{i=0}^{N-1} \left[\bar{p}_i \bar{q}_i' + \bar{\wp}_i \bar{t}_i'\right] \Delta\tau$$

The right hand side is the discrete action A_N. Therefore:

$$A(\hat{p}(\cdot), \hat{q}(\cdot), \hat{\wp}(\cdot), \hat{t}(\cdot)) = A_N(p_0 \cdots p_N, q_0 \cdots q_N, \wp_0 \cdots \wp_N, t_0 \cdots t_N) \qquad \blacksquare$$

<u>DEFINITION 1.8</u>: (Discrete Principle of Least Action)
The discrete–time Hamiltonian trajectory of a system with Hamiltonian function $H(q,p)$ is given by piecewise–linear, continuous functions $\hat{p}(\cdot)$, $\hat{q}(\cdot)$, $\hat{\wp}(\cdot)$ and $\hat{t}(\cdot)$ which cause the discrete action to be stationary under the constraint:

$$\bar{\wp}_n + H(\bar{q}_n, \bar{p}_n) = 0 \qquad n = 0, 1, \cdots N-1 \tag{1.3.1}$$

The endpoints q_0, q_N, p_0, p_N, t_0, t_N, $\wp_0$ and $\wp_N$ are to be specified.

The constraint (1.3.1) is less restrictive than (1.1.8) since (1.3.1) is enforced only at discrete points.

<u>THEOREM 1.9</u>: (Main Result)
The discrete–time Hamiltonian trajectory of a system with Hamiltonian function $H(q,p)$ and initial conditions $\hat{q}(\bar{\tau}_o) = \bar{q}_o$, $\hat{p}(\bar{\tau}_o) = \bar{p}_o$, $\hat{t}(\bar{\tau}_o) = \bar{t}_0$, and $\hat{\wp}(\bar{\tau}_o) = -H(\bar{q}_o, \bar{p}_o)$ is a solution of the following system of equations:

$$\frac{\bar{q}_{n+1} - \bar{q}_n}{\Delta\tau} = \frac{1}{2}\left[\lambda_{n+1}\frac{\partial H(\bar{q}_{n+1}, \bar{p}_{n+1})}{\partial \bar{p}_{n+1}} + \lambda_n \frac{\partial H(\bar{q}_n, \bar{p}_n)}{\partial \bar{p}_n}\right] \tag{1.3.2}$$

$$\frac{\bar{p}_{n+1} - \bar{p}_n}{\Delta\tau} = -\frac{1}{2}\left[\lambda_{n+1}\frac{\partial H(\bar{q}_{n+1}, \bar{p}_{n+1})}{\partial \bar{q}_{n+1}} + \lambda_n \frac{\partial H(\bar{q}_n, \bar{p}_n)}{\partial \bar{q}_n}\right] \tag{1.3.3}$$

$$\frac{\overline{t}_{n+1} - \overline{t}_n}{\Delta\tau} = \frac{1}{2}\left[\lambda_{n+1} + \lambda_n \right] \tag{1.3.4}$$

$$\frac{\overline{\wp}_{n+1} - \overline{\wp}_n}{\Delta\tau} = 0 \tag{1.3.5}$$

$$0 = \overline{\wp}_n + H(\overline{q}_n, \overline{p}_n) \tag{1.3.6}$$

where $n = 0, 1, \cdots N-1$.

Proof:

By the discrete principle of least action a discrete-time Hamiltonian trajectory is given by piecewise-linear, continuous functions which cause the discrete action A_N to be stationary under the constraint:

$$\overline{\wp}_n + H(\overline{q}_n, \overline{p}_n) = 0 \qquad n = 0, 1, \cdots N-1$$

where the endpoints $q_0, q_N, p_0, p_N, t_0, t_N, \wp_0$ and $\wp_N$ are fixed. Let:

$$g(\overline{q}_n, \overline{p}_n, \overline{\wp}_n) = \overline{\wp}_n + H(\overline{q}_n, \overline{p}_n) \tag{1.3.7}$$

$$f(p_0 \cdots p_N, q_0 \cdots q_N, \wp_0 \cdots \wp_N, t_0 \cdots t_N, \lambda_0 \cdots \lambda_N) = A_N - \sum_{i=0}^{N-1} \lambda_i \, g(\overline{q}_i, \overline{p}_i, \overline{\wp}_i) \tag{1.3.8}$$

where the λ_i's in (1.3.8) are Lagrange multipliers. We have then that A_N, subject to $g(\overline{q}_n, \overline{p}_n, \overline{\wp}_n) = 0$ for $n = 0, 1, \cdots N\text{-}1$, is stationary when the partial derivatives of f with the possible exception of $\frac{\partial f}{\partial q_o}$, $\frac{\partial f}{\partial q_N}$, $\frac{\partial f}{\partial p_o}$, $\frac{\partial f}{\partial p_N}$, $\frac{\partial f}{\partial t_o}$, $\frac{\partial f}{\partial t_N}$, $\frac{\partial f}{\partial \wp_o}$, and $\frac{\partial f}{\partial \wp_N}$, are equal to zero. The exception is necessary because the endpoints q_0, q_N, p_0, p_N, t_0, t_N, $\wp_0$ and $\wp_N$ are assumed to be specified and therefore partial derivatives with respect to these variables may not be zero. From (1.3.7), (1.3.8) and the definition of A_N:

$$f(p_0\cdots p_N, q_0\cdots q_N, \wp_0\cdots\wp_N, t_0\cdots t_N, \lambda_0\cdots\lambda_N) = \sum_{i=0}^{N-1}\Big[\overline{p}_i\overline{q}_i' + \overline{\wp}_i\overline{t}_i' - \lambda_i\big(\overline{\wp}_i + H(\overline{q}_i,\overline{p}_i)\big)\Big]\Delta\tau \qquad (1.3.9)$$

Equating to zero the partial derivatives $\frac{\partial f}{\partial q_{n+1}}$, $\frac{\partial f}{\partial p_{n+1}}$, $\frac{\partial f}{\partial t_{n+1}}$, and $\frac{\partial f}{\partial \wp_{n+1}}$ for $n = 0, 1, \cdots N-2$ implies equations (1.3.2)–(1.3.5) as follows. From (1.3.9) for $n = 0, 1, \cdots N-2$,

$$\frac{\partial f}{\partial p_{n+1}} = \frac{\partial}{\partial p_{n+1}}\sum_{i=0}^{N-1}\Big[\overline{p}_i\overline{q}_i' + \overline{\wp}_i\overline{t}_i' - \lambda_i\big(\overline{\wp}_i + H(\overline{q}_i, \overline{p}_i)\big)\Big]\Delta\tau \qquad (1.3.10)$$

The terms on the right hand side of (1.3.10) depend on p_{n+1} only for $i = n$ and $i = n+1$

Therefore:

$$\frac{\partial f}{\partial p_{n+1}} = \frac{\partial}{\partial p_{n+1}}\Big\{\overline{p}_n\overline{q}_n' + \overline{\wp}_n\overline{t}_n' - \lambda_n\big(\overline{\wp}_n + H(\overline{q}_n, \overline{p}_n)\big) + \overline{p}_{n+1}\overline{q}_{n+1}' + \overline{\wp}_{n+1}\overline{t}_{n+1}' - \lambda_{n+1}\big(\overline{\wp}_{n+1} + H(\overline{q}_{n+1}, \overline{p}_{n+1})\big)\Big\}\Delta\tau$$

$$= \Big\{\frac{\partial\overline{p}_n}{\partial p_{n+1}}\overline{q}_n' - \lambda_n\frac{\partial H(\overline{q}_n, \overline{p}_n)}{\partial\overline{p}_n}\frac{\partial\overline{p}_n}{\partial p_{n+1}} + \frac{\partial\overline{p}_{n+1}}{\partial p_{n+1}}\overline{q}_{n+1}' - \lambda_{n+1}\frac{\partial H(\overline{q}_{n+1}, \overline{p}_{n+1})}{\partial\overline{p}_{n+1}}\frac{\partial\overline{p}_{n+1}}{\partial p_{n+1}}\Big\}\Delta\tau$$

From Lemma 1.5, $\frac{\partial\overline{p}_n}{\partial p_{n+1}} = \frac{1}{2}$ and $\frac{\partial\overline{p}_{n+1}}{\partial p_{n+1}} = \frac{1}{2}$. Therefore,

$$\frac{\partial f}{\partial p_{n+1}} = \left\{\frac{\overline{q}_{n+1}' + \overline{q}_n'}{2} - \frac{1}{2}\left[\lambda_{n+1}\frac{\partial H(\overline{q}_{n+1}, \overline{p}_{n+1})}{\partial\overline{p}_{n+1}} + \lambda_n\frac{\partial H(\overline{q}_n, \overline{p}_n)}{\partial\overline{p}_n}\right]\right\}\Delta\tau$$

From the continuity constraint on $\hat{q}(\cdot)$

$$\frac{\overline{q}_{n+1}' + \overline{q}_n'}{2} = \frac{\overline{q}_{n+1} - \overline{q}_n}{\Delta\tau}$$

Therefore,

$$\frac{\partial f}{\partial p_{n+1}} = \left\{\frac{\overline{q}_{n+1} - \overline{q}_n}{\Delta\tau} - \frac{1}{2}\left[\lambda_{n+1}\frac{\partial H(\overline{q}_{n+1}, \overline{p}_{n+1})}{\partial\overline{p}_{n+1}} + \lambda_n\frac{\partial H(\overline{q}_n, \overline{p}_n)}{\partial\overline{p}_n}\right]\right\}\Delta\tau$$

Finally, $\frac{\partial f}{\partial p_{n+1}} = 0$ implies equation (1.3.2). Similarly,

$$\frac{\partial f}{\partial q_{n+1}} = \frac{\partial}{\partial q_{n+1}} \sum_{i=0}^{N-1} \left[\overline{p}_i \overline{q}_i' + \overline{\wp}_i \overline{t}_i' - \lambda_i\left(\overline{\wp}_i + H(\overline{q}_i, \overline{p}_i)\right)\right] \Delta\tau$$

$$= \frac{\partial}{\partial q_{n+1}} \Big\{ \overline{p}_n \overline{q}_n' + \overline{\wp}_n \overline{t}_n' - \lambda_n\left(\overline{\wp}_n + H(\overline{q}_n, \overline{p}_n)\right) +$$

$$\overline{p}_{n+1} \overline{q}_{n+1}' + \overline{\wp}_{n+1} \overline{t}_{n+1}' - \lambda_{n+1}\left(\overline{\wp}_{n+1} + H(\overline{q}_{n+1}, \overline{p}_{n+1})\right) \Big\} \Delta\tau$$

$$= \Big\{ \overline{p}_n \frac{\partial \overline{q}_n'}{\partial q_{n+1}} - \lambda_n \frac{\partial H(\overline{q}_n, \overline{p}_n)}{\partial \overline{q}_n} \frac{\partial \overline{q}_n}{\partial q_{n+1}} + \overline{p}_{n+1} \frac{\partial \overline{q}_{n+1}'}{\partial q_{n+1}} -$$

$$\lambda_{n+1} \frac{\partial H(\overline{q}_{n+1}, \overline{p}_{n+1})}{\partial \overline{q}_{n+1}} \frac{\partial \overline{q}_{n+1}}{\partial q_{n+1}} \Big\} \Delta\tau$$

From Lemma 1.5, $\frac{\partial \overline{q}_n'}{\partial q_{n+1}} = \frac{1}{\Delta\tau}$ and $\frac{\partial \overline{q}_{n+1}'}{\partial q_{n+1}} = -\frac{1}{\Delta\tau}$. Therefore,

$$\frac{\partial f}{\partial q_{n+1}} = \left\{ -\frac{\overline{p}_{n+1} - \overline{p}_n}{\Delta\tau} - \frac{1}{2}\left[\lambda_{n+1} \frac{\partial H(\overline{q}_{n+1}, \overline{p}_{n+1})}{\partial \overline{q}_{n+1}} + \lambda_n \frac{\partial H(\overline{q}_n, \overline{p}_n)}{\partial \overline{q}_n} \right] \right\} \Delta\tau$$

$\frac{\partial f}{\partial q_{n+1}} = 0$ implies equation (1.3.3).

$$\frac{\partial f}{\partial \wp_{n+1}} = \frac{\partial}{\partial \wp_{n+1}} \Big\{ \overline{p}_n \overline{q}_n' + \overline{\wp}_n \overline{t}_n' - \lambda_n\left(\overline{\wp}_n + H(\overline{q}_n, \overline{p}_n)\right) +$$

$$\overline{p}_{n+1} \overline{q}_{n+1}' + \overline{\wp}_{n+1} \overline{t}_{n+1}' - \lambda_{n+1}\left(\overline{\wp}_{n+1} + H(\overline{q}_{n+1}, \overline{p}_{n+1})\right) \Big\} \Delta\tau$$

$$= \left\{ \frac{\partial \overline{\wp}_n}{\partial \wp_{n+1}} \overline{t}_n' - \lambda_n \frac{\partial \overline{\wp}_n}{\partial \wp_{n+1}} + \frac{\partial \overline{\wp}_{n+1}}{\partial \wp_{n+1}} \overline{t}_{n+1}' - \lambda_n \frac{\partial \overline{\wp}_{n+1}}{\partial \wp_{n+1}} \right\} \Delta\tau$$

From Lemma 1.5, $\frac{\partial \overline{\wp}_n}{\partial \wp_{n+1}} = \frac{1}{2}$ and $\frac{\partial \overline{\wp}_{n+1}}{\partial \wp_{n+1}} = \frac{1}{2}$. Therefore,

$$\frac{\partial f}{\partial \wp_{n+1}} = \left\{ \frac{\overline{t}_{n+1}' + \overline{t}_n'}{2} - \frac{1}{2}\left[\lambda_{n+1} + \lambda_n \right] \right\} \Delta\tau$$

From the continuity constraint on $\hat{t}(\cdot)$

$$\frac{\overline{t}'_{n+1} + \overline{t}'_n}{2} = \frac{\overline{t}_{n+1} - \overline{t}_n}{\Delta\tau}$$

Therefore,

$$\frac{\partial f}{\partial \wp_{n+1}} = \left\{ \frac{\overline{t}_{n+1} - \overline{t}_n}{\Delta\tau} - \frac{1}{2}\left[\lambda_{n+1} + \lambda_n \right] \right\} \Delta\tau$$

$\frac{\partial f}{\partial \wp_{n+1}} = 0$ implies (1.3.4).

$$\frac{\partial f}{\partial t_{n+1}} = \frac{\partial}{\partial t_{n+1}} \sum_{i=0}^{N-1} \left[\overline{p}_i \overline{q}'_i + \overline{\wp}_i \overline{t}'_i - \lambda_i\Big(\overline{\wp}_i + H(\overline{q}_i, \overline{p}_i) \Big) \right] \Delta\tau$$

$$= \frac{\partial}{\partial t_{n+1}} \Big\{ \overline{p}_n \overline{q}'_n + \overline{\wp}_n \overline{t}'_n - \lambda_n\Big(\overline{\wp}_n + H(\overline{q}_n, \overline{p}_n) \Big) +$$

$$\overline{p}_{n+1} \overline{q}'_{n+1} + \overline{\wp}_{n+1} \overline{t}'_{n+1} - \lambda_{n+1}\Big(\overline{\wp}_{n+1} + H(\overline{q}_{n+1}, \overline{p}_{n+1}) \Big) \Big\} \Delta\tau$$

$$= \left\{ \overline{\wp}_n \frac{\partial \overline{t}'_n}{\partial t_{n+1}} + \overline{\wp}_{n+1} \frac{\partial \overline{t}'_{n+1}}{\partial t_{n+1}} \right\} \Delta\tau$$

From Lemma 1.5, $\frac{\partial \overline{t}'_n}{\partial t_{n+1}} = \frac{1}{\Delta\tau}$ and $\frac{\partial \overline{t}'_{n+1}}{\partial t_{n+1}} = -\frac{1}{\Delta\tau}$. Therefore,

$$\frac{\partial f}{\partial t_{n+1}} = \left\{ - \frac{\overline{\wp}_{n+1} - \overline{\wp}_n}{\Delta\tau} \right\} \Delta\tau$$

$\frac{\partial f}{\partial t_{n+1}} = 0$ implies (1.3.5). Finally, for $n = 0, 1, \cdots N-1$,

$$\frac{\partial f}{\partial \lambda_n} = \frac{\partial}{\partial \lambda_n} \sum_{i=0}^{N-1} \left[\overline{p}_i \overline{q}'_i + \overline{\wp}_i \overline{t}'_i - \lambda_i\Big(\overline{\wp}_i + H(\overline{q}_i, \overline{p}_i) \Big) \right] \Delta\tau$$

$$= \frac{\partial}{\partial \lambda_n} \left[\overline{p}_n \overline{q}'_n + \overline{\wp}_n \overline{t}'_n - \lambda_n\Big(\overline{\wp}_n + H(\overline{q}_n, \overline{p}_n) \Big) \right] \Delta\tau$$

$$= \left(\bar{\wp}_n + H(\bar{q}_n, \bar{p}_n) \right)$$

Equating $\frac{\partial f}{\partial \lambda_n}$ to zero implies equation (1.3.6). $\blacksquare$

The Discrete Principle of Least Action as stated in Definition 1.8 does not completely determine a piecewise–linear, continuous trajectory. As we can see from the equations of Theorem 1.9, the Discrete Principle of Least Action only determines the values of λ_n, $\bar{q}_n$, $\bar{t}_n$, $\bar{p}_n$, and $\bar{\wp}_n$. The values of $\bar{q}_n'$, $\bar{t}_n'$, $\bar{p}_n'$, and $\bar{\wp}_n'$ remain indeterminate. The following corollary shows that if variations in the endpoints of the momentum trajectories are allowed, then the values of $\bar{q}_n'$ and $\bar{t}_n'$ are no longer indeterminate. The values of $\bar{p}_n'$ and $\bar{\wp}_n'$, however, remain indeterminate.

<u>COROLLARY 1.10</u>:
If the endpoints q_0, q_N, t_0, and t_N are specified while the endpoints p_0, p_N, $\wp_0$ and $\wp_N$ are free to vary, then the Discrete Principle of Least Action implies:

$$\bar{q}_n' = \lambda_n \frac{\partial H(\bar{q}_n, \bar{p}_n)}{\partial \bar{p}_n} \tag{1.3.11}$$

$$\bar{t}_n' = \lambda_n \tag{1.3.12}$$

where $n = 0, 1, \cdots N-1$.

Proof:

Since p_0, p_N, $\wp_0$ and $\wp_N$ are free to vary, the function f in (1.3.8) is stationary if in addition to the partial derivatives of Theorem 1.9 the partial derivatives $\frac{\partial f}{\partial p_o}$, $\frac{\partial f}{\partial p_N}$, $\frac{\partial f}{\partial \wp_o}$, and $\frac{\partial f}{\partial \wp_N}$ are equal to zero. Equating $\frac{\partial f}{\partial p_o}$, $\frac{\partial f}{\partial p_N}$, $\frac{\partial f}{\partial \wp_o}$, and $\frac{\partial f}{\partial \wp_N}$ to zero implies equations (1.3.11) – (1.3.12) as follows. From (1.3.9) we have:

$$\frac{\partial f}{\partial p_0} = \frac{\partial}{\partial p_0} \sum_{i=0}^{N-1} \left[\bar{p}_i \bar{q}_i' + \bar{\wp}_i \bar{t}_i' - \lambda_i \left(\bar{\wp}_i + H(\bar{q}_i, \bar{p}_i) \right) \right] \Delta\tau \tag{1.3.13}$$

The terms on the right side of (1.1.13) depend on p_0 only for $i = 0$. Therefore,

$$\frac{\partial f}{\partial p_o} = \frac{\partial}{\partial p_o}\left[\overline{p}_o\overline{q}_o' + \overline{\varphi}_o\overline{t}_o' - \lambda_o\left(\overline{\varphi}_o + H(\overline{q}_o, \overline{p}_o)\right)\right]\Delta\tau$$

$$= \left[\frac{\partial\overline{p}_o}{\partial p_o}\overline{q}_o' - \lambda_o\frac{\partial H(\overline{q}_o, \overline{p}_o)}{\partial\overline{p}_o}\frac{\partial\overline{p}_o}{\partial p_o}\right]\Delta\tau$$

$$= \left[\overline{q}_o' - \lambda_o\frac{\partial H(\overline{q}_o, \overline{p}_o)}{\partial\overline{p}_o}\right]\frac{\partial\overline{p}_o}{\partial p_o}\ \Delta\tau$$

From Lemma 2.2 $\frac{\partial\overline{p}_o}{\partial p_o} = \frac{1}{2} \neq 0$ and $\Delta\tau \neq 0$. Therefore, $\frac{\partial f}{\partial p_0} = 0$ implies:

$$\overline{q}_o' = \lambda_o\frac{\partial H(\overline{q}_o, \overline{p}_o)}{\partial\overline{p}_o} \tag{1.3.14}$$

Similarly,

$$\frac{\partial f}{\partial\varphi_0} = \frac{\partial}{\partial\varphi_0}\sum_{i=0}^{N-1}\left[\overline{p}_i\overline{q}_i' + \overline{\varphi}_i\overline{t}_i' - \lambda_i\left(\overline{\varphi}_i + H(\overline{q}_i, \overline{p}_i)\right)\right]\Delta\tau$$

$$= \frac{\partial}{\partial\varphi_o}\left[\overline{p}_o\overline{q}_o' + \overline{\varphi}_o\overline{t}_o' - \lambda_o\left(\overline{\varphi}_o + H(\overline{q}_o, \overline{p}_o)\right)\right]\Delta\tau$$

$$= \left[\frac{\partial\overline{\varphi}_o}{\partial\varphi_o}\overline{t}_o' - \lambda_o\frac{\partial\overline{\varphi}_o}{\partial\varphi_o}\right]\Delta\tau$$

$$= \left[\overline{t}_o' - \lambda_o\right]\frac{\partial\overline{\varphi}_o}{\partial\varphi_o}\ \Delta\tau$$

From Lemma 2.2 $\frac{\partial\overline{\varphi}_o}{\partial\varphi_o} = \frac{1}{2} \neq 0$ and $\Delta\tau \neq 0$. Therefore, $\frac{\partial f}{\partial\varphi_0} = 0$ implies:

$$\overline{t}_o' = \lambda_o \tag{1.3.15}$$

$$\frac{\partial f}{\partial p_N} = \frac{\partial}{\partial p_N}\sum_{i=0}^{N-1}\left[\overline{p}_i\overline{q}_i' + \overline{\varphi}_i\overline{t}_i' - \lambda_i\left(\overline{\varphi}_i + H(\overline{q}_i, \overline{p}_i)\right)\right]\Delta\tau$$

$$= \frac{\partial}{\partial p_N}\left[\overline{p}_{N-1}\overline{q}_{N-1}' + \overline{\varphi}_{N-1}\overline{t}_{N-1}' - \lambda_{N-1}\left(\overline{\varphi}_{N-1} + H(\overline{q}_{N-1}, \overline{p}_{N-1})\right)\right]\Delta\tau$$

$$= \left[\frac{\partial \overline{p}_{N-1}}{\partial p_N} \overline{q}'_{N-1} - \lambda_{N-1} \frac{\partial H(\overline{q}_{N-1}, \overline{p}_{N-1})}{\partial \overline{p}_{N-1}} \frac{\partial \overline{p}_{N-1}}{\partial p_N} \right] \Delta\tau$$

$$= \left[\overline{q}'_{N-1} - \lambda_{N-1} \frac{\partial H(\overline{q}_{N-1}, \overline{p}_{N-1})}{\partial \overline{p}_{N-1}} \right] \frac{\partial \overline{p}_{N-1}}{\partial p_N} \Delta\tau$$

$$= \left[\overline{q}'_{N-1} - \lambda_{N-1} \frac{\partial H(\overline{q}_{N-1}, \overline{p}_{N-1})}{\partial \overline{p}_{N-1}} \right] \frac{\Delta\tau}{2}$$

$\frac{\partial f}{\partial p_N} = 0$ implies:

$$\overline{q}'_{N-1} = \lambda_{N-1} \tag{1.3.16}$$

$$\frac{\partial f}{\partial \wp_N} = \frac{\partial}{\partial \wp_N} \sum_{i=0}^{N-1} \left[\overline{p}_i \overline{q}'_i + \overline{\wp}_i \overline{t}'_i - \lambda_i \Big(\overline{\wp}_i + H(\overline{q}_i, \overline{p}_i) \Big) \right] \Delta\tau$$

$$= \frac{\partial}{\partial \wp_N} \left[\overline{p}_{N-1} \overline{q}'_{N-1} + \overline{\wp}_{N-1} \overline{t}'_{N-1} - \lambda_{N-1} \Big(\overline{\wp}_{N-1} + H(\overline{q}_{N-1}, \overline{p}_{N-1}) \Big) \right] \Delta\tau$$

$$= \left[\frac{\partial \overline{\wp}_{N-1}}{\partial \wp_N} \overline{t}'_{N-1} - \lambda_{N-1} \frac{\partial \overline{\wp}_{N-1}}{\partial \wp_N} \right] \Delta\tau$$

$$= \left[\overline{t}'_{N-1} - \lambda_{N-1} \right] \frac{\partial \overline{\wp}_{N-1}}{\partial \wp_N} \Delta\tau$$

$$= \left[\overline{t}'_{N-1} - \lambda_{N-1} \right] \frac{\Delta\tau}{2}$$

$\frac{\partial f}{\partial \wp_N} = 0$ implies:

$$\overline{t}'_{N-1} = \lambda_{N-1} \tag{1.3.17}$$

Using the continuity constraint, we can express equations (1.3.2) and (1.3.4) as follows:

$$\frac{\overline{q}'_{n+1} + \overline{q}'_n}{2} = \frac{1}{2} \left[\lambda_{n+1} \frac{\partial H(\overline{q}_{n+1}, \overline{p}_{n+1})}{\partial \overline{p}_{n+1}} + \lambda_n \frac{\partial H(\overline{q}_n, \overline{p}_n)}{\partial \overline{p}_n} \right] \tag{1.3.18}$$

$$\frac{\bar{t}'_{n+1} + \bar{t}'_n}{2} = \frac{1}{2}\left[\lambda_{n+1} + \lambda_n\right] \tag{1.3.19}$$

where $n = 0, 1, \cdots N-1$. Assume for some n, $0 \leq n < N-2$ that:

$$\bar{q}'_n = \lambda_n \frac{\partial H(\bar{q}_n, \bar{p}_n)}{\partial \bar{p}_n} \tag{1.3.20}$$

$$\bar{t}'_n = \lambda_n \tag{1.3.21}$$

Substituting for $\bar{q}'_n$ and $\bar{t}'_n$ in (1.3.18) and (1.3.19) and solving for $\bar{q}'_{n+1}$ and $\bar{t}'_{n+1}$ we have:

$$\bar{q}'_{n+1} = \lambda_{n+1} \frac{\partial H(\bar{q}_{n+1}, \bar{p}_{n+1})}{\partial \bar{p}_{n+1}} \tag{1.3.22}$$

$$\bar{t}'_{n+1} = \lambda_{n+1} \tag{1.3.23}$$

From (1.3.14) and (1.3.15) we see that (1.3.20) and (1.3.21) hold for $n = 0$. By induction, (1.3.20) and (1.3.21) hold for all $n = 0, 1, \cdots N-1$. $\blacksquare$

From the proof of Corollary 1.10, we can see that equations (1.3.16) and (1.3.17) imply (1.3.11) and (1.3.12) independently of equations (1.3.14) and (1.3.15). Thus, allowing variations in the momentum coordinates at $n = N$ yeilds the same equations as the equations obtained by allowing variations in the momentum coordinates at $n = 0$. In Chapter II we present a new variational principle which completely determines piecewise–linear, continuous trajectories for both position and momentum coordinates. The new principle allows variations in the momentum coordinates at $n = 0$ and variations in the position coordinates at $n = N$.

1.4 Preliminary Difference Equations

Equations (1.3.2) – (1.3.6) of Theorem 1.9 can be written in a more compact form as follows:

COROLLARY 1.11:
The discrete–time trajectory of a Hamiltonian dynamical system satisfies the following system of equations for $n = 0, 1, \cdots N-1$:

$$\overline{q}_{n+1} - \overline{q}_n - \frac{\Delta\tau}{2}\left[\lambda_{n+1}\frac{\partial H(\overline{q}_{n+1}, \overline{p}_{n+1})}{\partial \overline{p}_{n+1}} + \lambda_n \frac{\partial H(\overline{q}_n, \overline{p}_n)}{\partial \overline{p}_n}\right] = 0 \quad (1.4.1)$$

$$\overline{p}_{n+1} - \overline{p}_n + \frac{\Delta\tau}{2}\left[\lambda_{n+1}\frac{\partial H(\overline{q}_{n+1}, \overline{p}_{n+1})}{\partial \overline{q}_{n+1}} + \lambda_n \frac{\partial H(\overline{q}_n, \overline{p}_n)}{\partial \overline{q}_n}\right] = 0 \quad (1.4.2)$$

$$H\left(\overline{q}_{n+1}, \overline{p}_{n+1}\right) - H\left(\overline{q}_n, \overline{p}_n\right) = 0 \quad (1.4.3)$$

Proof:

Equations (1.4.1) and (1.4.2) follow directly from (1.3.2) and (1.3.3). From (1.3.6) we have $\overline{\wp}_n = -H(\overline{q}_n, \overline{p}_n)$. Substituting for $\overline{\wp}_n$ in (1.3.5) we obtain (1.4.3). Once λ_{n+1} is obtained from equations (1.4.1) – (1.4.3) t_{n+1} can be obtained explicitly from equation (1.3.4). ∎

We will refer to equations (1.4.1) – (1.4.3) as the preliminary difference equations for autonomous Hamiltonian systems. From equation (1.4.3) it is clear that discrete–time Hamiltonian trajectories exactly conserve the Hamiltonian function at the midpoints of each linear segment. For systems with time dependent Hamiltonians, the right hand side of equation (1.3.5) is not zero, and therefore, the reduction implied by Corollary 1.11 no longer holds true. (A formulation which includes time dependent systems is presented in Chapter II.)

We turn now to some preliminary considerations of existence and uniqueness of solutions to the preliminary difference equations (1.4.1) – (1.4.3). It is obvious from (1.4.1) – (1.4.3) that $\lambda_{n+1} = -\lambda_n$, $\overline{q}_{n+1} = \overline{q}_n$ and $\overline{p}_{n+1} = \overline{p}_n$ is a solution. However, this solution is constant since if $\lambda_{n+1} = -\lambda_n$, equation (1.3.4) implies that $\overline{t}_{n+1} = \overline{t}_n$. Clearly we are interested in solutions for which the time, $\overline{t}_n$, increases with n. Do nonconstant solutions exist? To simplify the discussion of this question, we will focus on only one time step. We will use λ_o, q_o, and p_o, to represent λ_n, $\overline{q}_n$, and $\overline{p}_n$ and we will use q, p, and λ to represent λ_{n+1}, $\overline{q}_{n+1}$, and $\overline{p}_{n+1}$. We will also use h to represent $\Delta\tau$, H_q and H_p to represent $\frac{\partial H(q,p)}{\partial q}$ and $\frac{\partial H(q,p)}{\partial p}$ and we will use H_q^o and H_p^o to represent $\frac{\partial H(q_o,p_o)}{\partial q}$ and $\frac{\partial H(q_o,p_o)}{\partial p}$.

Using the above notation, one step of the preliminary difference equations can be expressed as

$$f(q,p,\lambda) = 0 \tag{1.4.4}$$

where

$$f(q,p,\lambda) = \begin{bmatrix} q - q_o - \frac{h}{2}\left(\lambda H_p + \lambda_o H_p^o\right) \\ p - p_o + \frac{h}{2}\left(\lambda H_q + \lambda_o H_q^o\right) \\ H(q,p) - H(q_o,p_o) \end{bmatrix} \tag{1.4.5}$$

Let Df represent the Jacobian matrix of f. Then by direct computation we have:

$$Df = \begin{bmatrix} 1 - (\frac{h\lambda}{2}) H_{pq} & -(\frac{h\lambda}{2}) H_{pp} & -\frac{h}{2} H_p \\ (\frac{h\lambda}{2}) H_{qq} & 1 + (\frac{h\lambda}{2}) H_{qp} & \frac{h}{2} H_q \\ H_q & H_p & 0 \end{bmatrix} \tag{1.4.6}$$

and

$$\det(Df) = -h^2\left(\frac{\lambda}{4}\right)\Psi(q,p) \tag{1.4.7}$$

where

$$\Psi(q,p) = H_{qq}(H_p)^2 - 2H_{qp}H_qH_p + H_{pp}(H_q)^2 \tag{1.4.8}$$

From (1.4.7) we see that the preliminary difference equations are singular when $h = 0$. Assuming λ_o, q_o, and p_o are given, we will prove in Chapter III that for all sufficiently small nonzero values of h, a sufficient condition for the existence and local uniqueness of solutions to the preliminary difference equations is $\Psi(q_o, p_o) \neq 0$.

CHAPTER II

"DTH" DYNAMICS

In this chapter we present a discrete variational principle for Hamiltonian systems. We show in Section 2.2 that unlike the Discrete Principle of Least Action, this new variational principle, called the DTH Principle of Stationary Action, completely determines piecewise-linear, continuous trajectories which we name DTH trajectories. (DTH is shorthand notation for Discrete-Time Hamiltonian.) The conservation laws of DTH trajectories are described in Section 2.3.

2.1 Symplectic Notation

Let I_{n+1} represent the $(n+1)\times(n+1)$ identity matrix. Let J represent the $(2n+2)\times(2n+2)$ skew-symmetric matrix:

$$J = \begin{bmatrix} 0 & I_{n+1} \\ -I_{n+1} & 0 \end{bmatrix} \qquad (2.1.1)$$

Direct computation shows that

$$J^2 = \begin{bmatrix} -I_{n+1} & 0 \\ 0 & -I_{n+1} \end{bmatrix} \qquad (2.1.1)$$

and

$$(z)^T J\,(z) = 0 \quad \text{for all } z \in \mathbf{R}^{2n+2} \qquad (2.1.3)$$

Let $q = (q_1 \cdots q_n,\ t)^T$ where t represents time and let $p = (p_1 \cdots p_n,\ \wp)^T$ where $\wp$ is defined as in (1.1.6). As in Theorem 1.3, we can show that the Principle of Least Action implies that the trajectory of an n-degree of freedom, possibly time dependent, Hamiltonian system with Hamiltonian function $H(q,p)$ is a solution of the following system of differential equations:

$$\frac{dq}{d\tau} = \lambda(\tau)\,\mathbf{H}_p \tag{2.1.4}$$

$$\frac{dp}{d\tau} = -\lambda(\tau)\,\mathbf{H}_q \tag{2.1.5}$$

where

$$\mathbf{H}_q = \begin{bmatrix} \frac{\partial \mathbf{H}}{\partial q_1} \\ \vdots \\ \frac{\partial \mathbf{H}}{\partial q_{n+1}} \end{bmatrix} \qquad \mathbf{H}_p = \begin{bmatrix} \frac{\partial \mathbf{H}}{\partial p_1} \\ \vdots \\ \frac{\partial \mathbf{H}}{\partial p_{n+1}} \end{bmatrix}$$

and where

$$\mathbf{H}(q,p) = p_{n+1} + H(q,p) \tag{2.1.6}$$

(Note that p and q in equations (2.1.4) – (2.1.5) are vector quantities, not scalar quantities as in equations (1.1.9) – (1.1.10) of Theorem 1.3) Since $t = q_{n+1}$ and $\wp = p_{n+1}$, the equations which correspond to equations (1.1.11) and (1.1.12) of Theorem 1.3 are:

$$\frac{dq_{n+1}}{d\tau} = \lambda(\tau)\,\frac{\partial \mathbf{H}}{\partial p_{n+1}} \tag{2.1.7}$$

$$\frac{dp_{n+1}}{d\tau} = -\lambda(\tau)\,\frac{\partial \mathbf{H}}{\partial q_{n+1}} \tag{2.1.8}$$

Let $z = (q_1 \cdots q_{n+1}, p_1 \cdots p_{n+1})^T$. Using the skew–symmetric matrix J, equations (2.1.4) and (2.1.5) can be combined into one vector equation:

$$\frac{dz}{d\tau} = \lambda(\tau)\,J\,\mathbf{H}_z(z) \tag{2.1.9}$$

where
$$\mathbf{H}_z = \begin{bmatrix} \dfrac{\partial H}{\partial z_1} \\ \vdots \\ \dfrac{\partial H}{\partial z_{2n+2}} \end{bmatrix}$$

Equation (2.1.9) is known as the symplectic form of equations (2.1.4) – (2.1.5).

Next we generalize the results of Section 1.2 to vector valued functions. Consider a vector valued, piecewise–linear, continuous function $\hat{z} : [\tau_o, \tau_N] \rightarrow \mathbf{R}^{2n+2}$. As in Section 1.2, we can express $\hat{z}(\cdot)$ in terms of the values of $\overline{z}^{(k)}$ and $\overline{z}'^{(k)}$.

$$\hat{z}(\tau) = \begin{cases} \overline{z}^{(k)} + \overline{z}'^{(k)}(\tau - \overline{\tau}_k) & \tau_k \leq \tau < \tau_{k+1} \\ z_N & \tau = \tau_N \end{cases} \tag{2.1.10}$$

where

$$\overline{z}^{(k)}\left(z^{(k+1)}, z^{(k)}\right) = \frac{z^{(k+1)} + z^{(k)}}{2}$$

$$\overline{z}'^{(k)}\left(z^{(k+1)}, z^{(k)}\right) = \frac{z^{(k+1)} - z^{(k)}}{\Delta\tau}$$

for
$$k = 0, 1, \cdots N{-}1 .$$

Applying Lemma 1.4 to each component of $\hat{z}(\cdot)$ results in the continuity constraint:

$$\frac{\overline{z}^{(k+1)} - \overline{z}^{(k)}}{\Delta\tau} = \frac{\overline{z}'^{(k+1)} + \overline{z}'^{(k)}}{2} \tag{2.1.11}$$

From the definition of $\overline{z}^{(k)}$ and $\overline{z}'^{(k)}$ we have:

$$\frac{\partial \overline{z}^{(k)}}{\partial z^{(k)}} = \frac{1}{2}\, I \qquad \frac{\partial \overline{z}'^{(k)}}{\partial z^{(k)}} = -\frac{1}{\Delta\tau}\, I$$

$$k = 0, 1, \cdots N{-}1 \tag{2.1.12}$$

$$\frac{\partial \overline{z}^{(k)}}{\partial z^{(k+1)}} = \frac{1}{2}\, I \qquad \frac{\partial \overline{z}'^{(k)}}{\partial z^{(k+1)}} = \frac{1}{\Delta\tau}\, I$$

where I is the $(2n{+}2) \times (2n{+}2)$ identity matrix and where we have used the following notation:

$$\frac{\partial \bar{z}^{(k)}}{\partial z^{(k)}} = \left[\frac{\partial \bar{z}_i^{(k)}}{\partial z_j^{(k)}} \right] \qquad \frac{\partial \bar{z}'^{(k)}}{\partial z^{(k)}} = \left[\frac{\partial \bar{z}_i'^{(k)}}{\partial z_j^{(k)}} \right] \qquad i, j = 1, 2, \cdots 2n+2$$

$$\frac{\partial \bar{z}^{(k)}}{\partial z^{(k+1)}} = \left[\frac{\partial \bar{z}_i^{(k)}}{\partial z_j^{(k+1)}} \right] \qquad \frac{\partial \bar{z}'^{(k)}}{\partial z^{(k+1)}} = \left[\frac{\partial \bar{z}_i'^{(k)}}{\partial z_j^{(k+1)}} \right]$$

2.2 DTH Principle of Stationary Action

The following variational principle is used as the definition of DTH dynamics.

DEFINITION 2.1: (DTH Principle of Stationary Action)
A DTH trajectory is a piecewise–linear, continuous function $\hat{z} : [\tau_o, \tau_N] \rightarrow \mathbf{R}^{2n+2}$ for which the sum:

$$\begin{aligned} \mathcal{A}(\Delta\tau, \lambda_o, \cdots \lambda_{N-1}, \hat{z}(\cdot)) \;=\;& \tfrac{1}{2}(q^{(0)})^T p^{(0)} \\ &+ \sum_{j=0}^{N-1} \left[\tfrac{1}{2}(\bar{z}^{(j)})^T J\, (\bar{z}'^{(j)}) + \lambda_j \mathbf{H}(\bar{z}^{(j)}) \right] \Delta\tau \\ &+ \tfrac{1}{2}(q^{(N)})^T p^{(N)} \end{aligned} \tag{2.2.1}$$

is stationary. The endpoints $q^{(0)}$ and $p^{(N)}$ are to be specified. For a Hamiltonian system with Hamiltonian function $H(z)$, $\mathbf{H}(z)$ is defined to be:

$$\mathbf{H}(z) \;=\; z_{2n+2} + H(z) \tag{2.2.2}$$

Since $z_{2n+2} = p_{n+1}$ and $(z)^T = (q,\ p)^T$ equation (2.2.2) is the same equation as equation (2.1.6).

THEOREM 2.2: (Main Result)
A piecewise–linear, continuous function $\hat{z} : [\tau_o, \tau_N] \rightarrow \mathbf{R}^{2n+2}$ is a DTH trajectory if and only if $\bar{z}^{(k)}$ and $\bar{z}'^{(k)}$ satisfy the following equations:

$$\frac{\bar{z}^{(k+1)} - \bar{z}^{(k)}}{\Delta\tau} = \frac{1}{2} J \left[\lambda_{k+1} \frac{\partial \mathbf{H}(\bar{z}^{(k+1)})}{\partial \bar{z}^{(k+1)}} + \lambda_k \frac{\partial \mathbf{H}(\bar{z}^{(k)})}{\partial \bar{z}^{(k)}} \right] \quad k = 0, 1, \cdots N-2 \tag{2.2.3}$$

$$\bar{z}'^{(k)} = \lambda_k J \frac{\partial H(\bar{z}^{(k)})}{\partial \bar{z}^{(k)}} \quad k = 0, 1, \cdots N-1 \tag{2.2.4}$$

$$H(\bar{z}^{(k)}) = 0 \qquad k = 0, 1, \cdots N-1 \tag{2.2.5}$$

Equations (2.2.3) – (2.2.5) completely determine DTH trajectories. We will refer to these equations as the DTH equations of dynamics. It is important to note that the initial value of $\bar{z}^{(k)}_{2n+2}$, i.e. $\bar{\wp}_o$, must be chosen so that equation (2.2.5) is satisfied at $k = 0$.

Proof:

Assume $\hat{z}(\cdot)$ is a DTH trajectory. From Definition 2.1, we have that for fixed endpoints $q^{(0)}$ and $p^{(N)}$, $\mathcal{A}(\cdot)$ is stationary at $\hat{z}(\cdot)$. Therefore, the following derivatives of $\mathcal{A}(\cdot)$ are equal to zero:

$$\frac{\partial \mathcal{A}}{\partial z^{(k+1)}} = 0 \qquad k = 0, 1, \cdots N-2 \tag{2.2.6}$$

$$\frac{\partial \mathcal{A}}{\partial p^{(0)}} = 0 \tag{2.2.7}$$

$$\frac{\partial \mathcal{A}}{\partial q^{(N)}} = 0 \tag{2.2.8}$$

$$\frac{\partial \mathcal{A}}{\partial \lambda_k} = 0 \qquad k = 0, 1, \cdots N-1 \tag{2.2.9}$$

where we have used the notation:

$$\frac{\partial \mathcal{A}}{\partial z^{(k+1)}} = \begin{bmatrix} \frac{\partial \mathcal{A}}{\partial z_1^{(k+1)}} \\ \vdots \\ \frac{\partial \mathcal{A}}{\partial z_{2n+2}^{(k+1)}} \end{bmatrix} \quad \frac{\partial \mathcal{A}}{\partial p^{(0)}} = \begin{bmatrix} \frac{\partial \mathcal{A}}{\partial p_1^{(0)}} \\ \vdots \\ \frac{\partial \mathcal{A}}{\partial p_{n+1}^{(0)}} \end{bmatrix} \quad \frac{\partial \mathcal{A}}{\partial q^{(N)}} = \begin{bmatrix} \frac{\partial \mathcal{A}}{\partial q_1^{(N)}} \\ \vdots \\ \frac{\partial \mathcal{A}}{\partial q_{n+1}^{(N)}} \end{bmatrix}$$

Equation (2.2.6) implies equation (2.2.3) as follows. For $k = 0, 1, \cdots N-2$ we have from (2.2.1) that:

$$\frac{\partial \mathcal{A}}{\partial z^{(k+1)}} =$$

$$\frac{\partial}{\partial z^{(k+1)}}\left(\tfrac{1}{2}(q^{(0)})^T p^{(0)} + \sum_{j=0}^{N-1}\left[\tfrac{1}{2}(\overline{z}^{(j)})^T J\,(\overline{z}'^{(j)}) + \lambda_j H(\overline{z}^{(j)})\right]\Delta\tau + \tfrac{1}{2}(q^{(N)})^T p^{(N)}\right) =$$

$$\frac{\partial}{\partial z^{(k+1)}}\left(\tfrac{1}{2}(\overline{z}^{(k)})^T J\,(\overline{z}'^{(k)}) + \lambda_k H(\overline{z}^{(k)}) + \tfrac{1}{2}(\overline{z}^{(k+1)})^T J\,(\overline{z}'^{(k+1)}) + \lambda_{k+1} H(\overline{z}^{(k+1)})\right)\Delta\tau =$$

$$\Delta\tau\left(\tfrac{1}{2}\left(\frac{\partial\overline{z}^{(k)}}{\partial z^{(k+1)}}\right)^T J\,(\overline{z}'^{(k)}) + \tfrac{1}{2}\left(\frac{\partial\overline{z}'^{(k)}}{\partial z^{(k+1)}}\right)^T J^T(\overline{z}^{(k)}) + \lambda_k\left(\frac{\partial\overline{z}^{(k)}}{\partial z^{(k+1)}}\right)^T \frac{\partial H(\overline{z}^{(k)})}{\partial\overline{z}^{(k)}} + \right.$$

$$\left.\tfrac{1}{2}\left(\frac{\partial\overline{z}^{(k+1)}}{\partial z^{(k+1)}}\right)^T J\,(\overline{z}'^{(k+1)}) + \tfrac{1}{2}\left(\frac{\partial\overline{z}'^{(k+1)}}{\partial z^{(k+1)}}\right)^T J^T(\overline{z}^{(k+1)}) + \lambda_{k+1}\left(\frac{\partial\overline{z}^{(k+1)}}{\partial z^{(k+1)}}\right)^T \frac{\partial H(\overline{z}^{(k+1)})}{\partial\overline{z}^{(k+1)}}\right) =$$

$$\Delta\tau\left(\tfrac{1}{2}\left(\tfrac{1}{2}I\right)^T J\,\overline{z}'^{(k)} + \tfrac{1}{2}\left(\tfrac{1}{\Delta\tau}I\right)^T(-J)\,\overline{z}^{(k)} + \lambda_k\left(\tfrac{1}{2}I\right)^T \frac{\partial H(\overline{z}^{(k)})}{\partial\overline{z}^{(k)}} + \right.$$

$$\left.\tfrac{1}{2}\left(\tfrac{1}{2}I\right)^T J\,\overline{z}'^{(k+1)} + \tfrac{1}{2}\left(\tfrac{-1}{\Delta\tau}I\right)^T(-J)\,\overline{z}^{(k+1)} + \lambda_{k+1}\left(\tfrac{1}{2}I\right)^T \frac{\partial H(\overline{z}^{(k+1)})}{\partial\overline{z}^{(k+1)}}\right) =$$

$$\Delta\tau\left(\tfrac{1}{2}J\left(\frac{\overline{z}'^{(k+1)} + \overline{z}'^{(k)}}{2}\right) + \tfrac{1}{2}J\left(\frac{\overline{z}^{(k+1)} - \overline{z}^{(k)}}{\Delta\tau}\right) + \tfrac{1}{2}\left(\lambda_{k+1}\frac{\partial H}{\partial\overline{z}^{(k+1)}} + \lambda_k\frac{\partial H}{\partial\overline{z}^{(k)}}\right)\right).$$

From the continuity constraint on $\hat{z}(\cdot)$ we have:

$$\frac{\partial\mathcal{A}}{\partial z^{(k+1)}} = \Delta\tau\left(J\left(\frac{\overline{z}^{(k+1)} - \overline{z}^{(k)}}{\Delta\tau}\right) + \tfrac{1}{2}\left(\lambda_{k+1}\frac{\partial H}{\partial\overline{z}^{(k+1)}} + \lambda_k\frac{\partial H}{\partial\overline{z}^{(k)}}\right)\right)$$

$$= \Delta\tau\, J\left(\left(\frac{\overline{z}^{(k+1)} - \overline{z}^{(k)}}{\Delta\tau}\right) - \tfrac{1}{2}J\left(\lambda_{k+1}\frac{\partial H}{\partial\overline{z}^{(k+1)}} + \lambda_k\frac{\partial H}{\partial\overline{z}^{(k)}}\right)\right)$$

Since J is nonsingular, (2.2.6) implies that:

$$\frac{\overline{z}^{(k+1)} - \overline{z}^{(k)}}{\Delta\tau} = \tfrac{1}{2}J\left(\lambda_{k+1}\frac{\partial H}{\partial\overline{z}^{(k+1)}} + \lambda_k\frac{\partial H}{\partial\overline{z}^{(k)}}\right) \quad k = 0, 1, \cdots N-2$$

Next we show that equations (2.2.7) and (2.2.8) imply (2.2.4).

$$\frac{\partial\mathcal{A}}{\partial p^{(0)}} =$$

$$\frac{\partial}{\partial p^{(0)}}\left(\tfrac{1}{2}(q^{(0)})^T p^{(0)} + \sum_{j=0}^{N-1}\left[\tfrac{1}{2}(\overline{z}^{(j)})^T J\,(\overline{z}\,'^{(j)}) + \lambda_j H(\overline{z}^{(j)})\right]\Delta\tau + \tfrac{1}{2}(q^{(N)})^T p^{(N)}\right) =$$

$$\frac{\partial}{\partial p^{(0)}}\left(\tfrac{1}{2}(q^{(0)})^T p^{(0)} + \frac{\Delta\tau}{2}\,(\overline{z}^{(0)})^T J\,(\overline{z}\,'^{(0)}) + \Delta\tau\,\lambda_0 H(\overline{z}^{(0)})\right) =$$

$$\frac{\partial}{\partial p^{(0)}}\left(\tfrac{1}{2}(q^{(0)})^T p^{(0)} + \frac{\Delta\tau}{2}\left[(\overline{q}^{(0)})^T\,\overline{p}\,'^{(0)} - (\overline{p}^{(0)})^T\,\overline{q}\,'^{(0)}\right] + \Delta\tau\,\lambda_0 H(\,\overline{q}^{(0)}, \overline{p}^{(0)})\right) =$$

$$\frac{1}{2}\,q^{(0)} + \frac{\Delta\tau}{2}\left[\left(\frac{\partial\overline{p}\,'^{(0)}}{\partial p^{(0)}}\right)^T \overline{q}^{(0)} - \left(\frac{\partial\overline{p}^{(0)}}{\partial p^{(0)}}\right)^T \overline{q}\,'^{(0)}\right] + \Delta\tau\lambda_0\left(\frac{\partial\overline{p}^{(0)}}{\partial p^{(0)}}\right)^T \frac{\partial H}{\partial\overline{p}^{(0)}} =$$

$$\frac{1}{2}\,q^{(0)} + \frac{\Delta\tau}{2}\left[\left(-\frac{1}{\Delta\tau}\,I\right)^T \overline{q}^{(0)} - \left(\frac{1}{2}\,I\right)^T \overline{q}\,'^{(0)}\right] + \Delta\tau\lambda_o\left(\frac{1}{2}\,I\right)^T \frac{\partial H}{\partial\overline{p}^{(0)}} =$$

$$\frac{1}{2}\,q^{(0)} - \frac{1}{2}\,\overline{q}^{(0)} - \frac{\Delta\tau}{4}\,\overline{q}\,'^{(0)} + \frac{\Delta\tau}{2}\lambda_0\frac{\partial H}{\partial\overline{p}^{(0)}} =$$

$$\frac{2\,q^{(0)}}{4} - \frac{q^{(1)} + q^{(0)}}{4} - \frac{\Delta\tau}{4}\,\overline{q}\,'^{(0)} + \frac{\Delta\tau}{2}\,\lambda_0\frac{\partial H}{\partial\overline{p}^{(0)}} =$$

$$-\frac{q^{(1)} - q^{(0)}}{4} - \frac{\Delta\tau}{4}\,\overline{q}\,'^{(0)} + \frac{\Delta\tau}{2}\lambda_0\frac{\partial H}{\partial\overline{p}^{(0)}} =$$

$$-\frac{\Delta\tau}{2}\left[\frac{1}{2}\,\frac{q^{(1)} - q^{(0)}}{\Delta\tau} + \frac{1}{2}\,\overline{q}\,'^{(0)} - \lambda_0\frac{\partial H}{\partial\overline{p}^{(0)}}\right] =$$

$$-\frac{\Delta\tau}{2}\left[\overline{q}\,'^{(0)} - \lambda_0\frac{\partial H}{\partial\overline{p}^{(0)}}\right].$$

Thus,

$$\frac{\partial\mathcal{A}}{\partial p^{(0)}} = -\frac{\Delta\tau}{2}\left[\overline{q}\,'^{(0)} - \lambda_0\frac{\partial H}{\partial\overline{p}^{(0)}}\right]$$

Equation (2.2.7) implies

$$\overline{q}'^{(0)} = \lambda_0 \frac{\partial \mathbf{H}}{\partial \overline{p}^{(0)}} \tag{2.2.10}$$

Similarly, from equation (2.2.8) we have:

$$\frac{\partial \mathcal{A}}{\partial q^{(N)}} =$$

$$\frac{\partial}{\partial q^{(N)}}\left(\tfrac{1}{2}(q^{(0)})^T p^{(0)} + \sum_{j=0}^{N-1}\left[\tfrac{1}{2}(\overline{z}^{(j)})^T \mathbf{J}\,(\overline{z}'^{(j)}) + \lambda_j \mathbf{H}(\overline{z}^{(j)})\right]\Delta\tau + \tfrac{1}{2}(q^{(N)})^T p^{(N)}\right) =$$

$$\frac{\partial}{\partial q^{(N)}}\left(\frac{\Delta\tau}{2}\,(\overline{z}^{(N-1)})^T \mathbf{J}\,(\overline{z}'^{(N-1)}) + \Delta\tau\,\lambda_{N-1}\mathbf{H}(\overline{z}^{(N-1)}) + \tfrac{1}{2}(q^{(N)})^T p^{(N)}\right) =$$

$$\frac{\partial}{\partial q^{(N)}}\left(\frac{\Delta\tau}{2}\left[(\overline{q}^{(N-1)})^T\, \overline{p}'^{(N-1)} - (\overline{p}^{(N-1)})^T\, \overline{q}'^{(N-1)}\right] + \right.$$

$$\left. \Delta\tau\,\lambda_{N-1}\mathbf{H}(\overline{q}^{(N-1)}, \overline{p}^{(N-1)}) + \tfrac{1}{2}(q^{(N)})^T p^{(N)}\right) =$$

$$\frac{\Delta\tau}{2}\left[\left(\frac{\partial \overline{q}^{(N-1)}}{\partial q^{(N)}}\right)^T \overline{p}'^{(N-1)} - \left(\frac{\partial \overline{q}'^{(N-1)}}{\partial q^{(N)}}\right)^T \overline{p}^{(N-1)}\right] + \Delta\tau\,\lambda_{N-1}\left(\frac{\partial \overline{q}^{(N-1)}}{\partial q^{(N)}}\right)^T \frac{\partial \mathbf{H}}{\partial \overline{q}^{(N-1)}} + \frac{1}{2}\,p^{(N)} =$$

$$\frac{\Delta\tau}{2}\left[\left(\frac{1}{2}\,\mathrm{I}\right)^T \overline{p}'^{(N-1)} - \left(\frac{1}{\Delta\tau}\,\mathrm{I}\right)^T \overline{p}^{(N-1)}\right] + \Delta\tau\lambda_{N-1}\left(\frac{1}{2}\,\mathrm{I}\right)^T \frac{\partial \mathbf{H}}{\partial \overline{q}^{(N-1)}} + \frac{1}{2}\,p^{(N)} =$$

$$\frac{\Delta\tau}{4}\,\overline{p}'^{(N-1)} - \frac{1}{2}\,\overline{p}^{(N-1)} + \frac{1}{2}\,p^{(N)} + \frac{\Delta\tau}{2}\lambda_{N-1}\frac{\partial \mathbf{H}}{\partial \overline{q}^{(N-1)}} =$$

$$\frac{\Delta\tau}{4}\,\overline{p}'^{(N-1)} - \frac{p^{(N)} + p^{(N-1)}}{4} + \frac{2\,p^{(N)}}{4} + \frac{\Delta\tau}{2}\lambda_{N-1}\frac{\partial \mathbf{H}}{\partial \overline{q}^{(N-1)}} =$$

$$\frac{\Delta\tau}{2}\left[\frac{1}{2}\,\overline{p}'^{(N-1)} + \frac{1}{2}\,\frac{p^{(N)} - p^{(N-1)}}{\Delta\tau} + \lambda_{N-1}\frac{\partial \mathbf{H}}{\partial \overline{q}^{(N-1)}}\right] =$$

$$\frac{\Delta\tau}{2}\left[\,\overline{p}'^{(N-1)} + \lambda_{N-1}\frac{\partial \mathbf{H}}{\partial \overline{q}^{(N-1)}}\right].$$

Thus,

$$\frac{\partial \mathcal{A}}{\partial q^{(N)}} = \frac{\Delta\tau}{2}\left[\overline{p}'^{(N-1)} + \lambda_{N-1}\frac{\partial H}{\partial \overline{q}^{(N-1)}}\right]$$

Equation (2.2.8) implies

$$\overline{p}'^{(N-1)} = -\lambda_{N-1}\frac{\partial H}{\partial \overline{q}^{(N-1)}} \qquad (2.2.11)$$

Now equation (2.2.3) which we have already shown to hold true for $\overline{z}^{(k)}$ $k = 0, 1, \cdots N-1$ can be expressed as two equations.

$$\frac{\overline{q}^{(k+1)} - \overline{q}^{(k)}}{\Delta\tau} = \frac{1}{2}\left[\lambda_{k+1}\frac{\partial H}{\partial \overline{p}^{(k+1)}} + \lambda_k\frac{\partial H}{\partial \overline{p}^{(k)}}\right] \qquad (2.2.12)$$

$$\frac{\overline{p}^{(k+1)} - \overline{p}^{(k)}}{\Delta\tau} = -\frac{1}{2}\left[\lambda_{k+1}\frac{\partial H}{\partial \overline{q}^{(k+1)}} + \lambda_k\frac{\partial H}{\partial \overline{q}^{(k)}}\right] \qquad (2.2.13)$$

Using the continuity constraints

$$\frac{\overline{q}^{(k+1)} - \overline{q}^{(k)}}{\Delta\tau} = \frac{\overline{q}'^{(k+1)} + \overline{q}'^{(k)}}{2}$$

$$\frac{\overline{p}^{(k+1)} - \overline{p}^{(k)}}{\Delta\tau} = \frac{\overline{p}'^{(k+1)} + \overline{p}'^{(k)}}{2}$$

equations (2.2.12) and (2.2.13) become:

$$\frac{\overline{q}'^{(k+1)} + \overline{q}'^{(k)}}{2} = \frac{1}{2}\left[\lambda_{k+1}\frac{\partial H}{\partial \overline{p}^{(k+1)}} + \lambda_k\frac{\partial H}{\partial \overline{p}^{(k)}}\right] \qquad (2.2.14)$$

$$\frac{\overline{p}'^{(k+1)} + \overline{p}'^{(k)}}{2} = -\frac{1}{2}\left[\lambda_{k+1}\frac{\partial H}{\partial \overline{q}^{(k+1)}} + \lambda_k\frac{\partial H}{\partial \overline{q}^{(k)}}\right] \qquad (2.2.15)$$

Assume for some k, $0 \le k \le N-2$ that:

$$\overline{q}'^{(k)} = \lambda_k\frac{\partial H}{\partial \overline{p}^{(k)}} \qquad (2.2.16)$$

Equation (2.2.14) implies:

$$\overline{q}'^{(k+1)} = \lambda_{k+1}\frac{\partial H}{\partial \overline{p}^{(k+1)}} \tag{2.2.17}$$

Equation (2.2.10) implies (2.2.16) holds for $k=0$. Therefore, by induction:

$$\overline{q}'^{(k)} = \lambda_k \frac{\partial H}{\partial \overline{p}^{(k)}} \qquad k = 0,\, 1,\, \cdots\, N{-}1 \tag{2.2.18}$$

Similarly, assume for some k, $0 \le k \le N{-}2$ that:

$$\overline{p}'^{(k+1)} = -\lambda_{k+1}\frac{\partial H}{\partial \overline{q}^{(k+1)}} \tag{2.2.19}$$

Equation (2.2.15) implies:

$$\overline{p}'^{(k)} = -\lambda_k \frac{\partial H}{\partial \overline{q}^{(k)}} \tag{2.2.20}$$

Since by (2.2.11) equation (2.2.19) holds for $k = N{-}2$, we have by induction:

$$\overline{p}'^{(k)} = -\lambda_k \frac{\partial H}{\partial \overline{q}^{(k)}} \qquad k = 0,\, 1,\, \cdots\, N{-}1 \tag{2.2.21}$$

Using symplectic notation, equations (2.2.18) and (2.2.21) can be combined into the following equation, which is equation (2.2.4).

$$\overline{z}'^{(k)} = \lambda_k J \frac{\partial H}{\partial \overline{z}^{(k)}} \qquad k = 0,\, 1,\, \cdots\, N{-}1$$

Equation (2.2.5) follows easily from (2.2.1) and (2.2.9).

$$\frac{\partial \mathcal{A}}{\partial \lambda_k} = H(\overline{z}^{(k)}) = 0 \qquad k = 0,\, 1,\, \cdots\, N{-}1$$

Finally, to conclude the proof, we observe that each step used above is reversible and thus equations (2.2.3) – (2.2.5) not only are necessary, but also sufficient for a piecewise–linear, continuous function to be a DTH trajectory. ∎

2.3 Conservation Laws of DTH Dynamics

The DTH equations for autonomous systems with n degrees of freedom can be reduced in the same way that the equations of autonomous systems with one degree of freedom were reduced in Corollary 1.11. For autonomous systems, the variables $\overline{z}^{(k)}_{n+1} = \overline{t}_k$ and $\overline{z}^{(k)}_{2n+2} = \overline{p}_k$ can be eliminated from equations $(2.2.3)-(2.2.5)$ in the following way. First, we observe that since $\overline{z}^{(k)}_{2n+2} = \overline{p}_k$ and since $\overline{p}_k$ does not appear in $H(\overline{z}^{(k)})$

$$\frac{\partial H(\overline{z}^{(k)})}{\partial \overline{z}^{(k)}_{2n+2}} = 0$$

It follows then, from (2.2.5) that:

$$\frac{\overline{z}^{(k+1)}_{n+1} - \overline{z}^{(k)}_{n+1}}{\Delta\tau} = \frac{1}{2}\left(\lambda_{k+1}\frac{\partial \mathbf{H}}{\partial \overline{z}^{(k+1)}_{2n+2}} + \lambda_k \frac{\partial \mathbf{H}}{\partial \overline{z}^{(k)}_{2n+2}}\right) = \frac{\lambda_{k+1}+\lambda_k}{2} \qquad (2.3.1)$$

Hence, $\overline{z}^{(k)}_{n+1}$ is explicitly dependent on λ_k and λ_{k+1}. Since the system is autonomous, $\overline{z}^{(k)}_{n+1}$ does not appear in the remainder of the equations. Next, observe that equation (2.2.5):

$$\mathbb{H}(\overline{z}^{(k)}) = \overline{z}^{(k)}_{2n+2} + H(\overline{z}^{(k)}) = 0$$

implies

$$\overline{z}^{(k)}_{2n+2} = -H(\overline{z}^{(k)}) \qquad (2.3.2)$$

Since $\dfrac{\partial H(\overline{z}^{(k)})}{\partial \overline{z}^{(k)}_{n+1}} = 0$ for autonomous systems, we have

$$\frac{\overline{z}^{(k+1)}_{2n+2} - \overline{z}^{(k)}_{2n+2}}{\Delta\tau} = \frac{1}{2}\left(\lambda_{k+1}\frac{\partial \mathbf{H}}{\partial \overline{z}^{(k+1)}_{n+1}} + \lambda_k \frac{\partial \mathbf{H}}{\partial \overline{z}^{(k)}_{n+1}}\right) = 0$$

or

$$\overline{z}^{(k+1)}_{2n+2} - \overline{z}^{(k)}_{2n+2} = 0 \qquad (2.3.3)$$

Using (2.3.2) to substitute for $\overline{z}^{(k)}_{2n+2}$ in (2.3.3) we get:

$$H(\overline{z}^{(k+1)}) - H(\overline{z}^{(k)}) = 0 \tag{2.3.4}$$

If we define $\zeta = (q_1 \cdots q_n,\ p_1 \cdots p_n)^T \in \mathbf{R}^{2n}$ (where the variables t and ρ no longer appear) the DTH equations for autonomous systems can be expressed as:

$$\frac{\overline{\zeta}^{(k+1)} - \overline{\zeta}^{(k)}}{\Delta\tau} = \frac{1}{2} J \left[\lambda_{k+1} \frac{\partial H(\overline{\zeta}^{(k+1)})}{\partial \overline{\zeta}^{(k+1)}} + \lambda_k \frac{\partial H(\overline{\zeta}^{(k)})}{\partial \overline{\zeta}^{(k)}} \right] \quad k = 0, 1, \cdots N-2 \tag{2.3.5}$$

$$\overline{\zeta}'^{(k)} = \lambda_k J \frac{\partial H(\overline{\zeta}^{(k)})}{\partial \overline{\zeta}^{(k)}} \quad k = 0, 1, \cdots N-1 \tag{2.3.6}$$

$$H(\overline{\zeta}^{(k+1)}) - H(\overline{\zeta}^{(k)}) = 0 \quad k = 0, 1, \cdots N-1 \tag{2.3.7}$$

Equation (2.3.7) shows that, for autonomous systems, the Hamiltonian function is exactly conserved at the midpoints of DTH trajectories.

Next, we show that all the quadratic conservation laws of a Hamiltonian system are exactly reproduced at the vertices of a DTH trajectory. Let $L(\zeta)$ be the quadratic function

$$L(\zeta) = \frac{1}{2} (\zeta)^T A\, (\zeta) + b^T \zeta + c \tag{2.3.8}$$

where $A = A^T$. We have then that

$$L_\zeta(\zeta) = A\zeta + b \tag{2.3.9}$$

Assume $L(\zeta)$ is conserved for the Hamiltonian system having the Hamiltonian function $H(\zeta)$. Then the Poisson bracket of L and H is identically equal to zero.

$$[L, H] = \left(L_\zeta(\zeta)\right)^T J \left(H_\zeta(\zeta)\right) \equiv 0 \tag{2.3.10}$$

From (2.3.8) and (2.3.9) we have:

$$\frac{L(\zeta^{(k+1)}) - L(\zeta^{(k)})}{\Delta\tau} \quad =$$

$$\frac{1}{\Delta\tau}\left(\frac{1}{2}\,(\zeta^{(k+1)})^T A\;(\zeta^{(k+1)}) \;-\; \frac{1}{2}\,(\zeta^{(k)})^T A\;(\zeta^{(k)}) \;+\; b^T(\zeta^{(k+1)} - \zeta^{(k)})\right) \quad =$$

$$\frac{1}{\Delta\tau}\left(\frac{1}{2}\,(\zeta^{(k+1)} + \zeta^{(k)})^T A\;(\zeta^{(k+1)} - \zeta^{(k)}) \;+\; b^T(\zeta^{(k+1)} - \zeta^{(k)})\right) \quad =$$

$$\left(\frac{\zeta^{(k+1)} + \zeta^{(k)}}{2}\right)^T A\left(\frac{\zeta^{(k+1)} - \zeta^{(k)}}{\Delta\tau}\right) \;+\; b^T\left(\frac{\zeta^{(k+1)} - \zeta^{(k)}}{\Delta\tau}\right) \quad =$$

$$\left(\overline{\zeta}^{(k)}\right)^T A\left(\overline{\zeta}'^{(k)}\right) \;+\; b^T\left(\overline{\zeta}'^{(k)}\right) \quad =$$

$$\left(\left(A\;\overline{\zeta}^{(k)}\right)^T + \; b^T\right)\overline{\zeta}'^{(k)} \quad =$$

$$\left(L_\zeta(\overline{\zeta}^{(k)})\right)^T \overline{\zeta}'^{(k)}. \tag{2.3.11}$$

But from equation (2.3.6) and (2.3.10) we have

$$\begin{aligned}
\left(L_\zeta(\overline{\zeta}^{(k)})\right)^T \overline{\zeta}'^{(k)} \;&=\; \left(L_\zeta(\overline{\zeta}^{(k)})\right)^T \left(\lambda_k\; J\; H_\zeta(\overline{\zeta}^{(k)})\right) \\
&=\; \lambda_k \left(L_\zeta(\overline{\zeta}^{(k)})\right)^T J\left(H_\zeta(\overline{\zeta}^{(k)})\right) \\
&=\; 0
\end{aligned}$$

Therefore, from (2.3.11) we have

$$L(\zeta^{(k+1)}) \;-\; L(\zeta^{(k)}) \;=\; 0 \tag{2.3.12}$$

Clearly, $L(\zeta)$ is exactly conserved at the vertices of a DTH trajectory.

CHAPTER III

EXISTENCE AND UNIQUENESS

3.1 Preliminary Results

In this chapter we present sufficient conditions for the local existence and uniqueness of DTH trajectories. We will prove that, for Hamiltonian functions which satisfy certain conditions, there always exist $\overline{z}^{(k)}$ and $\overline{z}'^{(k)}$ $k = 0, 1, \cdots N{-}1$, which, for sufficiently small values of $\Delta\tau$, determine a piecewise–linear, continuous trajectory and which also satisfy the DTH equations:

$$\frac{\overline{z}^{(k+1)} - \overline{z}^{(k)}}{\Delta\tau} = \frac{1}{2} J \left[\lambda_{k+1} \frac{\partial \mathbb{H}(\overline{z}^{(k+1)})}{\partial \overline{z}^{(k+1)}} + \lambda_k \frac{\partial \mathbb{H}(\overline{z}^{(k)})}{\partial \overline{z}^{(k)}} \right] \quad k = 0, 1, \cdots N{-}2 \tag{3.1.1}$$

$$\overline{z}'^{(k)} = \lambda_k J \frac{\partial \mathbb{H}(\overline{z}^{(k)})}{\partial \overline{z}^{(k)}} \quad k = 0, 1, \cdots N{-}1 \tag{3.1.2}$$

$$\mathbb{H}(\overline{z}^{(k)}) = 0 \quad k = 0, 1, \cdots N{-}1 \tag{3.1.3}$$

Assume λ_o and $\overline{z}^{(0)}$ are chosen so that $\lambda_o > 0$ and $\overline{z}^{(0)}$ satisfies (3.1.3) for $k = 0$. Since $\overline{z}'^{(k)}$ does not appear in equations (3.1.1) and (3.1.3) and since $\overline{z}'^{(k)}$ is explicitly dependent on $\overline{z}^{(k)}$ in equation (3.1.2) to prove the existence of solutions, it is sufficient to prove the existence of λ_{k+1} and $\overline{z}^{(k+1)}$ satisfying equations (3.1.1) and (3.1.3). We will use an induction argument to extend conclusions from $k = 0$ to $k = 1, 2, \cdots N{-}1$. In this chapter, to simplify notation, we will use z^o to represent $\overline{z}^{(k)}$, z for $\overline{z}^{(k+1)}$ and the variable h for $\Delta\tau$. Recall that $\mathbb{H}(z) = z_{2n+2} + H(z)$. Let $\mathbb{H}^o$ represent $\mathbb{H}(z^o)$, $\mathbb{H}^o_z$ represent $\mathbb{H}_z(z^o)$ and $\mathbb{H}^o_{zz}$ represent $\mathbb{H}_{zz}(z^o)$ where $\mathbb{H}_z$ is the gradient and $\mathbb{H}_{zz}$ the Hessian matrix of $\mathbb{H}(z)$. As in Section 2.2, $z \in \mathbf{R}^{2n+2}$. Using the simplified notation, the DTH equations, exclusive of (3.1.2) can be written as follows:

$$z - z^o - \frac{h}{2} J \left[\lambda H_z(z) + \lambda_o H_z^o\right] = 0 \tag{3.1.4}$$

$$H(z) = 0 \tag{3.1.5}$$

As was pointed out in Section 1.4, equations (3.1.4) – (3.1.5) are singular at $h = 0$.

Let us first outline the method of proof which will be given in detail in Sections 3.2 – 3.4. We will use the Newton–Kantorovich Theorem and the Inverse Function Theorem to prove the existence and local uniqueness of a C^2 function $z(h,\lambda)$ which satisfies equation (3.1.4) for bounded values of λ and sufficiently small values of h. We will use $z(h,\lambda)$ to decouple (3.1.5) from (3.1.4) by defining a new function $g(h,\lambda) = H(z(h,\lambda))$. In Section 3.3, we will apply the Newton–Kantorovich Theorem to the decoupled equation $g(h,\lambda) = 0$. We will show that for all sufficiently small $h_o \neq 0$ there exists a $\lambda_1 = \lambda(h_o)$ such that $g(h_o,\lambda_1) = 0$. Finally, in Section 3.4, we will show that a solution to (3.1.4) – (3.1.5) for $\Delta\tau = h_o$, is given by (λ_1, z^1) where $\lambda_1 = \lambda(h_o)$ and $z^1 = z(h_o,\lambda_1)$. We will use the following theorems and lemmas.

<u>THEOREM 3.1</u>: (Newton–Kantorovich Theorem)
Assume:

(i) $D \subset \mathbb{R}^n$, D is convex

(ii) $F : \mathbb{R}^n \to \mathbb{R}^n$

(iii) $F \in C^1(D)$

(iv) $\left\| F_z(z^2) - F_z(z^1) \right\| \leq \gamma \left\| z^2 - z^1 \right\|$ for all $z^1, z^2 \in D$

(v) There exists a $z^o \in D$ such that $\left\| F_z^{-1}(z^o) \right\| \leq \beta$

(vi) $\alpha \leq 0.5$ where $\alpha = \beta\gamma\eta$ and $\eta = \left\| F_z^{-1}(z^o)\, F(z^o) \right\|$

(vii) $\overline{B(z^o, r_-)} \subset D$ where $r_- = \frac{1}{\beta\gamma}\left(1 - \sqrt{1 - 2\alpha}\, \right)$

Define $r_+ = \frac{1}{\beta\gamma}\left(1 + \sqrt{1 - 2\alpha}\, \right)$. Then the sequence $z^{(i)}$ given by:

$$z^{(i+1)} = z^{(i)} - F_z^{-1}(z^{(i)})\, F(z^{(i)}) \qquad i = 0, 1, \cdots$$

where $z^{(0)} = z^o$, is well defined, $z^{(i)} \to z^* \in \overline{B(z^o, r_-)}$, $F(z^*) = 0$ where z^* is the only zero of $F(z)$ in $\overline{B(z^o, r_+)} \cap D$ and:

$$\left\| z^{(1)} - z^* \right\| \leq 2\beta\gamma \left\| z^{(1)} - z^0 \right\|^2$$

Proof: (See [11, p. 155])

THEOREM 3.2: (Implicit Function Theorem)
Assume:

(i) $f : S \subset \mathbb{R}^{k+n} \to \mathbb{R}^n$, S is open

(ii) $f \in C^r(S)$

(iii) There exists a point $(x^1, z^1) \in S$ such that $f(x^1, z^1) = 0$

(iv) $f_z(x^1, z^1)$ is nonsingular

Then there exists an open set $V \subset \mathbb{R}^k$ where $x^1 \in V$ and there exists a unique function $z : V \to \mathbb{R}^n$ such that $z \in C^r(V)$, $z(x^1) = z^1$ and $f(x, z(x)) = 0$ for all $x \in V$.

Proof: (See [1, p. 374] and [2, p. 148])

LEMMA 3.3: (Matrix Perturbation Lemma)
Assume I and E are square matrices where I is the identity matrix. If $\|E\| < 1$, then $(I - E)$ is nonsingular and:

$$\left\|(I - E)^{-1}\right\| \leq \frac{1}{1 - \|E\|}$$

Proof: (See [4, p. 59])

THEOREM 3.4: (Mean Value Theorem)
Assume:

(i) $F : \mathbb{R}^n \to \mathbb{R}^n$

(ii) $D \subset \mathbb{R}^n$ is convex

(iii) $F \in C^1(D)$

Then for all $z^1, z^2 \in D$

$$\left\|F(z^2) - F(z^1)\right\| \leq \sup_{0 \leq \tilde{t} \leq 1} \left\|F_z(z^1 + \tilde{t}\,(z^2 - z^1))\right\| \, \left\|z^2 - z^1\right\|$$

Proof: (See [11, p. 143])

3.2 Existence of a Decoupling Function

THEOREM 3.5:
Assume:

$$f(h,\lambda,z) = z - z^o - \frac{h}{2} J\left[\lambda H_z(z) + \lambda_o H_z^o\right] \tag{3.2.1}$$

where $\lambda_o > 0$, $H \in C^3(U)$, $U \subset \mathbf{R}^{2n+2}$ is open and $z^o \in D_o$ where $H(z^o) = H^o = 0$, D_o is open, bounded and convex and $\overline{D}_o \subset U$. Then there exists a $\delta_o > 0$, an open set $Q_o = Q(\delta_o)$ where $Q(\delta) = (0,\delta) \times (0, 2\lambda_o)$ and there exists a C^2 function $z(h,\lambda)$ which maps Q_o into D_o and for which:

$$f(h,\lambda,z(h,\lambda)) = 0 \quad \text{for all } (h,\lambda) \in Q_o \tag{3.2.2}$$

Before proving Theorem 3.5, we state some facts about $H(z)$ and its derivatives and we prove $f(h,\lambda,z)$ has some special properties. We will use the abbreviation $\sup_{\tilde{\lambda}}$ for the supremum over $\lambda \in (0,\ 2\lambda_o)$ and $\sup_{\tilde{z}}$ for the supremum over $z \in \overline{D}_o$. We will also use the notation H_{zz_i} for $\dfrac{\partial H_z}{\partial z_i}$ and H_{zzz_i} for $\dfrac{\partial H_{zz}}{\partial z_i}$ where $1 \le i \le 2n+2$.

Since $H(z)$ is C^3 on $\overline{D}_o$, which is a compact subset of $\mathbf{R}^{2n+2}$, we can define the following constants:

$$M_1 = \sup_{\tilde{z}} |\, H(\tilde{z}) \,| \tag{3.2.3}$$

$$M_2 = \sup_{\tilde{z}} \left\| H_z(\tilde{z}) \right\| \tag{3.2.4}$$

$$M_3 = \sup_{\tilde{z}} \left\| H_{zz}(\tilde{z}) \right\| \tag{3.2.5}$$

$$M_4 = \sum_{i=1}^{2n+2} \sup_{\tilde{z}} \left\| H_{zzz_i}(\tilde{z}) \right\| \tag{3.2.6}$$

LEMMA 3.6:
Assume the conditions of Theorem 3.5 hold. Then $f(h,\lambda,z)$ has the following properties:

(a) $\displaystyle \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| f(h,\tilde{\lambda},z^o) \right\| = 0$

(b) Given $\delta > 0$ there exists a $\gamma(\delta) > 0$ such that:

$$\left\| f_z(h,\lambda,z^2) - f_z(h,\lambda,z^1) \right\| \le \gamma(\delta) \left\| z^2 - z^1 \right\| \quad \text{for all } z^1, z^2 \in D_o,\ (h,\lambda) \in Q(\delta)$$

(c) For $\beta > 1$ there exists a $\delta_\beta > 0$ such that $f_z(h,\lambda,z)$ is nonsingular and:

$$\left\| f_z^{-1}(h,\lambda,z) \right\| \le \beta \quad \text{for all } z \in D_o \text{ and } (h,\lambda) \in Q(\delta_\beta)$$

(δ_β in this special notation does not denote partial differentiation.)

Proof of (a):

From (3.2.1):

$$\lim_{h\to 0^+} \sup_{\tilde\lambda} \left\| f(h,\tilde\lambda,z^o) \right\| = \lim_{h\to 0^+} \sup_{\tilde\lambda} \left\| -\frac{h(\tilde\lambda+\lambda_o)}{2} \, J \, \mathbf{H}_z^o \right\|$$

$$\leq \lim_{h\to 0^+} h \left(\frac{3\lambda_o}{2}\right) \| J \| \left\| \mathbf{H}_z^o \right\|$$

$$= 0$$

Proof of (b):

By the Mean Value Theorem (Theorem 3.4) with $F = \mathbf{H}_{zz_i}$

$$\left\| \mathbf{H}_{zz_i}(z^2) - \mathbf{H}_{zz_i}(z^1) \right\| \leq \sup_{0 \leq \tilde t \leq 1} \left\| \mathbf{H}_{zzz_i}(z^2 + \tilde t\,(z^2 - z^1) \right\| \, \| z^2 - z^1 \|$$

which implies:

$$\left\| \mathbf{H}_{zz}(z^2) - \mathbf{H}_{zz}(z^1) \right\| \leq \sum_{i=1}^{2n+2} \left\| \mathbf{H}_{zz_i}(z^2) - \mathbf{H}_{zz_i}(z^1) \right\|$$

$$\leq \sum_{i=1}^{2n+2} \sup_{0 \leq \tilde t \leq 1} \left\| \mathbf{H}_{zzz_i}(z^2 + \tilde t\,(z^2 - z^1)) \right\| \, \| z^2 - z^1 \|$$

$$\leq M_4 \| z^2 - z^1 \|$$

Let $\gamma(\delta) = \delta\,\lambda_o \| J \| \, M_4$. Then for $(h,\lambda) \in Q(\delta)$, $z^1, z^2 \in D_o$

$$\left\| f_z(h,\lambda,z^2) - f_z(h,\lambda,z^1) \right\| = \left\| -\left(\frac{h\lambda}{2}\right) J \left[\mathbf{H}_{zz}(z^2) - \mathbf{H}_{zz}(z^1) \right] \right\|$$

$$\leq \left(\frac{h\lambda}{2}\right) \| J \| \left\| \mathbf{H}_{zz}(z^2) - \mathbf{H}_{zz}(z^1) \right\|$$

$$\leq \delta\,\lambda_o \| J \| \, M_4 \| z^2 - z^1 \|$$

$$\leq \gamma(\delta) \| z^2 - z^1 \|$$

Proof of (c):

From (3.2.1)

$$f_z(h,\lambda,z) = I - \left(\frac{h\lambda}{2}\right) J\, \mathbf{H}_{zz}(z)$$

$$= I - E(h,\lambda,z)$$

where $$E(h,\lambda,z) = \left(\frac{h\lambda}{2}\right) J\, \mathbf{H}_{zz}(z)$$

Let $$\delta_\beta = \frac{\beta - 1}{\beta \lambda_o \| J \| M_3}$$

Then for $(h,\lambda) \in Q(\delta_\beta)$ and $z \in D_o$

$$\| E(h,\lambda,z) \| = \left\| \left(\frac{h\lambda}{2}\right) J\, \mathbf{H}_{zz}(z) \right\|$$

$$\leq \left(\frac{h\lambda}{2}\right) \| J \| \, \| \mathbf{H}_{zz}(z) \|$$

$$\leq \delta_\beta \, \lambda_o \| J \| M_3$$

$$\leq \frac{\beta - 1}{\beta}$$

$$< 1$$

By the Matrix Perturbation Lemma (Lemma 3.3) $f_z(h,\lambda,z)$ is nonsingular for all $(h,\lambda) \in Q(\delta_\beta)$ and $z \in D_o$. Moreover,

$$\| f_z^{-1}(h,\lambda,z) \| = \left\| \left(I - E(h,\lambda,z) \right)^{-1} \right\| \leq \frac{1}{1 - \| E(h,\lambda,z) \|} \leq \frac{1}{1 - \frac{\beta - 1}{\beta}} = \beta \qquad \blacksquare$$

Proof of Theorem 3.5:

Consider the family of functions $\{f(h,\lambda,z)\}$ parametrized by $(h,\lambda) \in Q(\delta)$ for some $\delta > 0$. First, using the Newton–Kantorovich Theorem (Theorem 3.1) we will show that for each $(h,\lambda) \in Q(\delta_o)$ where δ_o is sufficiently small, the equation $f(h,\lambda,z) = 0$ has a locally unique zero, $z(h,\lambda)$ to which Newton's method converges quadratically. Then, using the Implicit Function Theorem, we will show that $z(h,\lambda)$ is C^2 on $Q(\delta_o)$.

By assumption, $\mathbf{H} \in C^3(U)$. Since $D_o \subset U$, from (3.2.1) it follows that for each $(h,\lambda) \in Q(\delta)$, $f(h,\lambda,z)$ is a C^2 function which maps the convex set $D_o \subset \mathbf{R}^{2n+2}$ into $\mathbf{R}^{2n+2}$. Thus, for each $(h,\lambda) \in Q(\delta)$ conditions (i) – (iii) of the Newton–Kantorovich Theorem hold true.

Choose $\beta_1 > 1$. By Lemma 3.6(c) there exists a $\delta_{\beta_1} > 0$ such that condition (v) of the Newton–Kantorovich Theorem, with $\beta = \beta_1$, holds true for all $(h,\lambda) \in Q(\delta_{\beta_1})$ and any $z^o \in D_o$. Let $\gamma_1 = \gamma(\delta_{\beta_1})$ where $\gamma(\delta)$ is given by Lemma 3.6(b). Then for all $(h,\lambda) \in Q(\delta_{\beta_1})$ condition (iv) holds true for $\gamma = \gamma_1$. Define $\alpha(h,\lambda) = \beta_1\gamma_1\eta(h,\lambda)$ where $\eta(h,\lambda) = \left\| f_z^{-1}(h,\lambda,z^o)\, f(h,\lambda,z^o) \right\|$. Then

$$
\begin{aligned}
\lim_{h\to 0^+} \sup_{\tilde{\lambda}} \alpha(h,\tilde{\lambda}) &= \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \beta_1\gamma_1\eta(h,\tilde{\lambda}) \\
&\le \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \beta_1\gamma_1 \left\| f_z^{-1}(h,\tilde{\lambda},z^o)) \right\| \left\| f(h,\tilde{\lambda},z^o) \right\| \\
&\le \beta_1^2\, \gamma_1 \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| f(h,\tilde{\lambda},z^o)) \right\| \\
&= 0
\end{aligned}
$$

where we have used Lemma 3.6(a). Thus, there exists a $\delta_{0.5} < \delta_{\beta_1}$ such that for all $(h,\lambda) \in Q(\delta_{0.5})$, condition (vi) of the Newton–Kantorovich Theorem holds true. Since D_o is open, there exists an $r_1 > 0$ such that $\overline{B(z^o, r_1)} \subset D_o$. Define:

$$
r_-(h,\lambda) = \frac{1}{\beta_1\gamma_1}\left(1 - \sqrt{1 - 2\alpha(h,\lambda)}\,\right)
$$

and define $$r_+(h,\lambda) = \frac{1}{\beta_1\gamma_1}\left(1+\sqrt{1-2\alpha(h,\lambda)}\right)$$

We have

$$\begin{aligned}\lim_{h\to 0^+}\sup_{\bar\lambda} r_-(h,\bar\lambda) &= \lim_{h\to 0^+}\sup_{\bar\lambda} \frac{1}{\beta_1\gamma_1}\left(1-\sqrt{1-2\alpha(h,\bar\lambda)}\right) \\ &= \frac{1}{\beta_1\gamma_1}\left(1-\sqrt{1-2\lim_{h\to 0^+}\sup_{\bar\lambda}\alpha(h,\bar\lambda)}\right) \\ &= 0\end{aligned}$$

Thus, there exists a $\delta_{r_1} < \delta_{0.5}$ such that $r_-(h,\lambda) < r_1$ for all $(h,\lambda) \in Q(\delta_{r_1})$. Therefore, $\overline{B(z^o, r_-(h,\lambda))} \subset \overline{B(z^o, r_1)} \subset D_o$ satisfying condition (vii) for all $(h,\lambda) \in Q(\delta_{r_1})$. If we choose $\delta_o < \delta_{r_1}$ all the conditions of the Newton–Kantorovich Theorem are satisfied for each $(h,\lambda) \in Q_o$ where $Q_o = Q(\delta_o)$. This implies there exists a function $z(h,\lambda)$ which maps Q_o into D_o and which has the property that for each $(h,\lambda) \in Q_o$, $z(h,\lambda)$ is the unique zero of $f(h,\lambda,z) = 0$ in $B(z^o, \frac{1}{\beta_1\gamma_1}) \cap D_o$. Uniqueness follows from the fact that, by the Newton–Kantorovich Theorem, $z(h,\lambda)$ is unique in $B(z^o, r_+(h,\lambda)) \cap D_o$ and since $\frac{1}{\beta_1\gamma_1} \le r_+(h,\lambda)$, $B(z^o, \frac{1}{\beta_1\gamma_1}) \subset B(z^o, r_+(h,\lambda))$ for all $(h,\lambda) \in Q_o$.

Finally, we show that $z(h,\lambda)$ is C^2 on Q_o. Let $S = Q_o \times D_o$. $f(h,\lambda,z)$ is C^2 on S and $f_z(h,\lambda,z)$ is nonsingular on S. For $(h_1,\lambda_1) \in Q_o$ and $z^1 = z(h_1,\lambda_1)$, $f(h_1,\lambda_1,z^1) = 0$. By the Implicit Function Theorem (Theorem 3.2) there exists an open set $V \subset Q_o$ containing (h_1,λ_1) and there exists a C^2 function $\tilde z(h,\lambda)$ defined on V such that $\tilde z(h_1,\lambda_1) = z^1$ and $f(h,\lambda,\tilde z(h,\lambda)) = 0$ for all $(h,\lambda) \in V$. If there exists an open set $W \subset V$ containing (h_1,λ_1) such that $z(h,\lambda) = \tilde z(h,\lambda)$ for $(h,\lambda) \in W$, then $z(h,\lambda)$ must also be C^2 on W.

Since, by the Newton–Kantorovich Theorem, $z^1 \in \overline{B(z^o, r_-(h_1,\lambda_1))} \subset B(z^o, \frac{1}{\beta_1\gamma_1})$, by the continuity of $\tilde z(h,\lambda)$ there exists an open set $W \subset V$ containing (h_1,λ_1) such that $\tilde z(W) \subset B(z^o, \frac{1}{\beta_1\gamma_1})$. For each $(h,\lambda) \in W$, the uniqueness of the zeros of $f(h,\lambda,z) = 0$

in $B(z^o, \frac{1}{\beta_1\gamma_1})$ implies that $z(h,\lambda) = \widetilde{z}(h,\lambda)$. Therefore, $z(h,\lambda)$ must also be C^2 on W. Since (h_1,λ_1) is arbitrary in Q_o , $z(h,\lambda)$ must be C^2 on Q_o. ∎

3.3 Hamiltonian Conservation Constraint

In this section we prove that, for all sufficiently small values of $h > 0$, there exist $\lambda(h)$ which satisfy the scalar equation:

$$g(h,\lambda) = 0 \tag{3.3.1}$$

where:

$$g(h,\lambda) = \mathbf{H}(z(h,\lambda)) \tag{3.3.2}$$

Equation (3.3.1) is the decoupled form of the Hamiltonian conservation constraint (equation (3.1.5) of Section 3.1). The "decoupling function" $z(h,\lambda)$ of Section 3.2 is used to decouple (3.1.5) from (3.1.4) resulting in equation (3.3.1). In the proof of Theorem 3.7 below we will use some properties of $g(h,\lambda)$ and its derivatives and therefore we will need to obtain expressions for the derivatives of $z(h,\lambda)$. Obtaining the derivatives of $z(h,\lambda)$ is complicated by the fact that $z(h,\lambda)$ is only given implicitly by equation (3.2.2).

THEOREM 3.7:
Assume the conditions of Theorem 3.5 hold. Let $g(h,\lambda) = \mathbf{H}(z(h,\lambda))$ and let:

$$\Psi(z) = (J\,\mathbf{H}_z(z))^T \mathbf{H}_{zz}(z)\,(J\,\mathbf{H}_z(z))$$

and $\Psi_o = \Psi(z^o)$. If $\Psi_o \neq 0$, then there exists a $\delta_* < \delta_o$ and a function $\lambda(h)$ which maps the interval $(0,\delta_*)$ into $(0, 2\lambda_o)$ and for which:

$$g(h,\lambda(h)) = 0 \quad \text{for all} \quad h \in (0,\delta_*)$$

Before proving Theorem 3.7, we state and prove three lemmas.

<u>LEMMA 3.8</u>:
Assume the conditions of Theorem 3.5 hold. Then as $h \to 0^+$

(a) $z(h,\lambda) \to z^o$

(b) $\mathbf{H}(z(h,\lambda)) \to \mathbf{H}^o \qquad (= \mathbf{H}(z^o))$

(c) $\mathbf{H}_z(z(h,\lambda)) \to \mathbf{H}_z^o \qquad (= \mathbf{H}_z^o(z^o))$

(d) $\mathbf{H}_{zz}(z(h,\lambda)) \to \mathbf{H}_{zz}^o \qquad (= \mathbf{H}_{zz}^o(z^o))$

where the convergence in (a) – (d) is uniform in λ on the interval $(0,\, 2\lambda_o)$.

Proof:

From Theorem 3.5 we have:

$$z(h,\lambda) - z^o = \frac{h}{2} J \left[\lambda \mathbf{H}_z(z(h,\lambda)) + \lambda_o \mathbf{H}_z^o\right]$$

$$\lim_{h \to 0^+} \sup_{\tilde{\lambda}} \left\| z(h,\tilde{\lambda}) - z^o \right\| \leq \lim_{h \to 0^+} \left(\frac{h}{2}\right) \| J \| \left[2\lambda_o M_2 + \lambda_o \mathbf{H}_z^o\right]$$

$$= 0$$

which establishes (a). Using the Mean Value Theorem:

$$| \mathbf{H}(z(h,\lambda)) - \mathbf{H}^o | \leq \sup_{\bar{z}} \left\| \mathbf{H}_z(\bar{z}) \right\| \left\| z(h,\lambda) - z^o \right\| = M_2 \left\| z(h,\lambda) - z^o \right\|$$

Thus
$$\lim_{h \to 0^+} \sup_{\tilde{\lambda}} | \mathbf{H}(z(h,\tilde{\lambda})) - \mathbf{H}^o | \leq M_2 \lim_{h \to 0^+} \sup_{\tilde{\lambda}} \left\| z(h,\tilde{\lambda})) - z^o \right\|$$

$$= 0$$

which establishes (b). Similarly, using the Mean Value Theorem:

$$\lim_{h \to 0^+} \sup_{\tilde{\lambda}} \left\| \mathbf{H}_z - \mathbf{H}_z^o \right\| \leq M_3 \lim_{h \to 0^+} \sup_{\tilde{\lambda}} \left\| z(h,\tilde{\lambda}) - z_o \right\| = 0$$

Thus, (c) is established.

$$\lim_{h\to 0^+} \sup_{\tilde\lambda} \left\| \mathbf{H}_{zz}(z(h,\lambda)) - \mathbf{H}^o_{zz} \right\| \le \lim_{h\to 0^+} \sum_{i=1}^{2n+2} \sup_{\tilde\lambda} \left\| \mathbf{H}_{zz_i}(z(h,\lambda)) - \mathbf{H}^o_{zz_i} \right\|$$

$$\le \lim_{h\to 0^+} \left[\sum_{i=1}^{2n+2} \sup_{\tilde z} \left\| \mathbf{H}_{zzz_i}(\tilde z) \right\| \right] \sup_{\tilde\lambda} \left\| z(h,\tilde\lambda) - z_o \right\|$$

$$= M_4 \lim_{h\to 0^+} \sup_{\tilde\lambda} \left\| z(h,\tilde\lambda) - z_o \right\|$$

$$= 0$$

which establishes (d). ∎

<u>LEMMA 3.9</u>:
Assume the conditions of Theorem 3.5 hold. Then the partial derivatives of $z(h,\lambda)$ are bounded on the set $Q_o = (0, \delta_o) \times (0, 2\lambda_o)$. Moreover, as $h \to 0^+$:

(a) $z_h(h,\lambda) \to \left(\frac{\lambda + \lambda_o}{2}\right) J\, \mathbf{H}^o_z$

(b) $\frac{z_\lambda(h,\lambda)}{h} \to \frac{1}{2} J\, \mathbf{H}^o_z$

(c) $z_{\lambda h}(h,\lambda) \to \frac{1}{2} J\, \mathbf{H}^o_z$

(d) $z_{hh}(h,\lambda) \to \left(\frac{\lambda(\lambda + \lambda_o)}{2}\right) J\, \mathbf{H}^o_{zz}\, J\, \mathbf{H}^o_z$

(e) $\frac{z_{\lambda\lambda}(h,\lambda)}{h^2} \to \frac{1}{2} J\, \mathbf{H}^o_{zz}\, J\, \mathbf{H}^o_z$

where the convergence in (a) – (e) is uniform in λ on the interval $(0, 2\lambda_o)$.

Proof:

Consider the matrix $\left[I - \left(\frac{h\lambda}{2}\right) J\, \mathbf{H}_{zz}\right] = f_z(h,\lambda,z)$. As in the proof of Theorem 3.5,

$$\left\| \left[I - \left(\frac{h\lambda}{2}\right) J\, \mathbf{H}_{zz} \right]^{-1} \right\| = \left\| f_z^{-1}(h,\lambda,z) \right\| < \beta_1 \quad \text{for all } (h,\lambda) \in Q_o,\ z \in D_o \qquad (3.3.3)$$

Proof of (a):

Since Theorem 3.5 implies $z(h,\lambda)$ satisfies equation (3.2.2) we have:

$$z_h - \tfrac{1}{2} J \left[\lambda \mathbf{H}_z + \lambda_o \mathbf{H}^o_z \right] - \left(\frac{h\lambda}{2}\right) J\, \mathbf{H}_{zz}\, z_h = 0 \qquad (3.3.4)$$

$$z_h = \left[I - \left(\frac{h\lambda}{2}\right) J\, H_{zz} \right]^{-1} \tfrac{1}{2} J \left[\lambda H_z + \lambda_o H_z^o \right]$$

$$\|z_h\| \leq \left\| \left[I - \left(\frac{h\lambda}{2}\right) J\, H_{zz} \right]^{-1} \right\| \tfrac{1}{2} \|J\| \left[2\lambda_o \|H_z\| + \lambda_o H_z^o \right]$$

$$\leq \left(\frac{\beta_1 \lambda_o}{2}\right) \|J\| \left[2 M_2 + H_z^o \right]$$

which implies $\|z_h\|$ is bounded on Q_o. From (3.3.4) we have:

$$\lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| z_h - \left(\frac{\lambda + \lambda_o}{2}\right) J\, H_z^o \right\| =$$

$$\lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| \tfrac{1}{2} J \left[\lambda H_z + \lambda_o H_z^o \right] - \left(\frac{\lambda + \lambda_o}{2}\right) J\, H_z^o + \left(\frac{h\lambda}{2}\right) J\, H_{zz}\, z_h \right\| =$$

$$\lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| \left(\frac{\lambda}{2}\right) J\, (H_z - H_z^o) \right\| + \lim_{h\to 0^+} h\, \lambda_o \|J\| M_3 \sup_{\tilde{\lambda}} \|z_h\| =$$

$$\lambda_o \|J\| \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \| H_z - H_z^o \| = 0.$$

which establishes (a)

Proof of (b):

From (3.2.1) we have:

$$z_\lambda - \left(\frac{h}{2}\right) J\, H_z - \left(\frac{h\lambda}{2}\right) J\, H_{zz}\, z_\lambda = 0 \tag{3.3.5}$$

$$z_\lambda = \left[I - \left(\frac{h\lambda}{2}\right) J\, H_{zz} \right]^{-1} \left(\frac{h}{2}\right) J\, H_z$$

$$\|z_\lambda\| \leq h \left(\frac{\beta_1}{2}\right) \|J\| M_2 \leq \left(\frac{\delta_o \beta_1}{2}\right) \|J\| M_2$$

Thus $\|z_\lambda\|$ is bounded on Q_o. Moreover:

$$\lim_{h\to 0^+} \sup_{\tilde{\lambda}} \|z_\lambda\| \leq \lim_{h\to 0^+} h \left(\frac{\beta_1}{2}\right) \|J\| M_2 = 0 \tag{3.3.6}$$

From (3.3.5)

$$\sup_{\tilde{\lambda}} \left\| \frac{z_\lambda}{h} - \frac{1}{2} J H_z^o \right\| = \sup_{\tilde{\lambda}} \left\| \frac{1}{2} J H_z - \frac{1}{2} J H_z^o + \left(\frac{\lambda}{2}\right) J H_{zz} z_\lambda \right\|$$

$$\leq \frac{1}{2} \| J \| \sup_{\tilde{z}} \left\| H_z - H_z^o \right\| + \lambda_o \| J \| M_3 \sup_{\tilde{\lambda}} \| z_\lambda \|$$

Using Lemma 3.8(c) and (3.3.6) we have:

$$\lim_{h \to 0^+} \sup_{\tilde{\lambda}} \left\| \frac{z_\lambda}{h} - \frac{1}{2} J H_z^o \right\| = 0$$

Proof of (c):

From (3.3.5)

$$z_{\lambda h} - \frac{1}{2} J H_z - \frac{h}{2} J H_{zz} z_h - \frac{\lambda}{2} J H_{zz} z_\lambda - \left(\frac{h\lambda}{2}\right) J H_{zz} z_{\lambda h} - \left(\frac{h\lambda}{2}\right) J\ e_1(h,\lambda,z) = 0 \qquad (3.3.7)$$

where:

$$e_1(h,\lambda,z) = \begin{bmatrix} (z_h)^T H_{zzz_1}(z_\lambda) \\ \vdots \\ (z_h)^T H_{zzz_{2n+2}}(z_\lambda) \end{bmatrix}$$

and where:

$$\left\| e_1(h,\lambda,z) \right\| \leq \sum_{i=1}^{2n+2} \left\| (z_h)^T H_{zzz_i}(z_\lambda) \right\|$$

$$\leq \sum_{i=1}^{2n+2} \left\| H_{zzz_i} \right\| \| z_h \| \| z_\lambda \|$$

$$\leq M_4 \| z_h \| \| z_\lambda \|$$

Therefore:

$$z_{\lambda h} = \left[I - \left(\frac{h\lambda}{2}\right) J H_{zz} \right]^{-1} \left[\frac{1}{2} J H_z + \frac{h}{2} J H_{zz} z_h + \frac{\lambda}{2} J H_{zz} z_\lambda + \left(\frac{h\lambda}{2}\right) J\ e_1(h,\lambda,z) \right]$$

$$\| z_{\lambda h} \| \leq \beta_1 \left[\frac{1}{2} \| J \| M_2 + \frac{\delta_o}{2} \| J \| M_3 \| z_h \| + \lambda_o \| J \| M_3 \| z_\lambda \| + \delta_o \lambda_o \| J \| M_4 \| z_h \| \| z_\lambda \| \right]$$

Thus $\| z_{\lambda h} \|$ is bounded on Q_o since $\| z_h \|$ and $\| z_\lambda \|$ are both bounded on Q_o.

$$\lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| z_{\lambda h} - \tfrac{1}{2} J\, H_z^o \right\| \quad =$$

$$\lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| \tfrac{1}{2} J\, H_z - \tfrac{1}{2} J\, H_z^o + \tfrac{h}{2} J\, H_{zz}\, z_h + \tfrac{\lambda}{2} J\, H_{zz}\, z_\lambda + \left(\tfrac{h\lambda}{2}\right) J\, H_{zz}\, z_{\lambda h} + \left(\tfrac{h\lambda}{2}\right) J\, e_1 \right\| \quad =$$

$$\tfrac{1}{2}\|J\| \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| H_z - H_z^o \right\| + \lim_{h\to 0^+} \tfrac{h}{2}\|J\|\, M_3 \sup_{\tilde{\lambda}} \|z_h\| + \lambda_o \|J\|\, M_3 \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \|z_\lambda\| \; +$$
$$\lim_{h\to 0^+} h\, \lambda_o \|J\| M_3 \sup_{\tilde{\lambda}} \|z_{\lambda h}\| + \lim_{h\to 0^+} h\, \lambda_o \|J\|\, M_4 \sup_{\tilde{\lambda}} \|z_h\|\, \|z_\lambda\|$$

$$= \quad 0.$$

where we have used (3.3.6) and the fact that $\|z_h\|$, $\|z_\lambda\|$ and $\|z_{\lambda h}\|$ are bounded on Q_o.

Proof of (d):

From (3.3.4)

$$z_{hh} - \frac{\lambda}{2} J\, H_{zz}\, z_h - \frac{\lambda}{2} J\, H_{zz}\, z_h - \left(\frac{h\lambda}{2}\right) J\, H_{zz}\, z_{hh} \; - \left(\frac{h\lambda}{2}\right) J\; e_2(h,\lambda,z) \;=\; 0$$

where:

$$e_2(h,\lambda,z) \quad = \quad \begin{bmatrix} (z_h)^T H_{zzz_1}(z_h) \\ \vdots \\ (z_h)^T H_{zzz_{2n+2}}(z_h) \end{bmatrix}$$

Using an argument similar to that used for $e_1(h,\lambda,z)$ we can show that:

$$\left\| e_2(h,\lambda,z) \right\| \;\; \le \;\; M_4 \, \|z_h\|^2$$

Solving for z_{hh} we have:

$$z_{hh} \;=\; \left[\, I \; - \; \left(\frac{h\lambda}{2}\right) J\; H_{zz} \right]^{-1} \left[\lambda\, J\, H_{zz}\, z_h + \left(\frac{h\lambda}{2}\right) J\; e_2(h,\lambda,z) \right]$$

$$\|z_{hh}\| \;\le\; \beta \left[2\lambda_o \|J\|\, M_3 \|z_h\| \;+\; \delta_o\, \lambda_o \|J\|\, M_4 \|z_h\|^2 \right]$$

Since $\|z_h\|$ is bounded, $\|z_{hh}\|$ is bounded also. We have:

$$\lim_{h\to 0^+}\sup_{\tilde\lambda}\left\| z_{hh} - \frac{\lambda(\lambda+\lambda_o)}{2} J\, H^o_{zz}\, J\, H^o_z\right\| =$$

$$\lim_{h\to 0^+}\sup_{\tilde\lambda}\left\| \lambda\, J\, H_{zz} z_h - \frac{\lambda(\lambda+\lambda_o)}{2} J\, H^o_{zz}\, J\, H^o_z + \left(\frac{h\lambda}{2}\right) J\, H_{zz}\, z_{hh} + \left(\frac{h\lambda}{2}\right) J\, e_2(h,\lambda,z)\right\| =$$

$$\lim_{h\to 0^+}\sup_{\tilde\lambda}\left\| \lambda\, J\, H_{zz} z_h - \lambda\, J\, H^o_{zz} z_h + \lambda\, J\, H^o_{zz} z_h - \frac{\lambda(\lambda+\lambda_o)}{2} J\, H^o_{zz}\, J\, H^o_z\right\| +$$

$$\lim_{h\to 0^+} h\,\lambda_o \| J \|\, M_3 \sup_{\tilde\lambda} \| z_{hh} \| + \lim_{h\to 0^+} h\,\lambda_o \| J \|\, M_4 \sup_{\tilde\lambda} \| z_h \|^2 \le$$

$$\lambda_o \| J \| \lim_{h\to 0^+}\sup_{\tilde z} \| H_{zz} - H^o_{zz} \| \| z_h \| + \lambda_o \| J \|\, M_3 \lim_{h\to 0^+}\sup_{\tilde\lambda}\left\| z_h - \left(\frac{\lambda+\lambda_o}{2}\right) J\, H^o_z\right\|$$

$$= 0.$$

Thus, (d) is established.

Proof of (e):

From (3.3.5)

$$z_{\lambda\lambda} - \frac{h}{2} J\, H_{zz} z_\lambda - \frac{h}{2} J\, H_{zz} z_\lambda - \left(\frac{h\lambda}{2}\right) J\, H_{zz} z_{\lambda\lambda} - \left(\frac{h\lambda}{2}\right) J\, e_3(h,\lambda,z) = 0 \qquad (3.3.8)$$

where:

$$e_3(h,\lambda,z) = \begin{bmatrix} (z_\lambda)^T H_{zzz_1}(z_\lambda) \\ \vdots \\ (z_\lambda)^T H_{zzz_{2n+2}}(z_\lambda) \end{bmatrix}$$

and where:

$$\| e_3(h,\lambda,z) \| \le M_4 \| z_\lambda \|^2$$

Solving for $z_{\lambda\lambda}$

$$z_{\lambda\lambda} = \left[I - \left(\frac{h\lambda}{2}\right) J\, H_{zz}\right]^{-1} \left[h\, J\, H_{zz} z_\lambda + \left(\frac{h\lambda}{2}\right) J\, e_2(h,\lambda,z)\right]$$

$$\| z_{\lambda\lambda} \| \le \beta_1 \left[h \| J \|\, M_3 \| z_\lambda \| + h\lambda_o \| J \|\, M_4 \| z_\lambda \|^2 \right] \qquad (3.3.9)$$

$$\lim_{h\to 0^+}\sup_{\tilde\lambda}\left\| \frac{z_{\lambda\lambda}}{h} \right\| \le \lim_{h\to 0^+}\sup_{\tilde\lambda} \beta_1 \left[\| J \|\, M_3 \| z_\lambda \| + \lambda_o \| J \|\, M_4 \| z_\lambda \|^2 \right]$$

$$\leq \quad \beta_1 \|J\| M_3 \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \|z_\lambda\| + \beta_1 \lambda_o \|J\| M_4 \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \|z_\lambda\|^2$$

$$= \quad 0$$

Since $h < \delta_o$ in (3.3.9), $\|z_{\lambda\lambda}\|$ is bounded on Q_o and we also have:

$$\lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| \frac{z_{\lambda\lambda}}{h} \right\| = 0$$

From (3.3.8)

$$\lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| \frac{z_{\lambda\lambda}}{h^2} - \frac{1}{2} J\, H^o_{zz}\, J\, H^o_z \right\| =$$

$$\lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| J\, H_{zz} \frac{z_\lambda}{h} - \frac{1}{2} J\, H^o_{zz}\, J\, H^o_z + \frac{\lambda}{2} J\, H_{zz} \frac{z_{\lambda\lambda}}{h} + \frac{\lambda}{2} J \frac{e_3(h,\lambda,z)}{h} \right\| =$$

$$\lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| J\, H_{zz} \frac{z_\lambda}{h} - J\, H^o_{zz} \frac{z_\lambda}{h} \right\| + \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| J\, H^o_{zz} \frac{z_\lambda}{h} - \frac{1}{2} J\, H^o_{zz}\, J\, H^o_z \right\| +$$
$$\lambda_o \|J\| M_4 \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| \frac{z_{\lambda\lambda}}{h} \right\| + \lambda_o \|J\| M_4 \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| \frac{z_\lambda}{h} \right\| \|z_\lambda\| \leq$$

$$\|J\| \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \|H_{zz} - H^o_{zz}\| \left\| \frac{z_\lambda}{h} \right\| + \|J\| M_3 \lim_{h\to 0^+} \sup_{\tilde{\lambda}} \left\| \frac{z_\lambda}{h} - \frac{1}{2} J\, H^o_z \right\|$$

$$= \quad 0.$$

thus establishing (e). ∎

<u>LEMMA 3.10</u>:
Assume the conditions of Theorem 3.5 and Theorem 3.6 hold. Then $g(h,\lambda)$, $g_\lambda(h,\lambda)$ and $g_{\lambda\lambda}(h,\lambda)$ are bounded on Q_o. Moreover, as $h \to 0^+$:

(a) $\dfrac{g(h,\lambda)}{h^2} \to -\left(\dfrac{\lambda^2 - \lambda_o^2}{8}\right)\Psi_o$

(b) $\dfrac{g_\lambda(h,\lambda)}{h^2} \to -\dfrac{\lambda}{4}\Psi_o$

(c) $\dfrac{g_{\lambda\lambda}(h,\lambda)}{h^2} \to -\dfrac{1}{4}\Psi_o$

where the convergence in (a) – (c) is uniform in λ on the interval $(0, 2\lambda_o)$.

Proof:

$$| g(h,\lambda) | = | H(z(h,\lambda)) | \leq M_1$$

Thus $g(h,\lambda)$ is bounded on Q_o.

Proof of (a):

$$\lim_{h\to 0^+} \frac{g(h,\lambda)}{h^2} = \lim_{h\to 0^+} \frac{H(z(h,\lambda))}{h^2} \qquad (3.3.10)$$

Since $\lim_{h\to 0^+} H(z(h,\lambda)) = \mathbf{H}^o = 0$, since $\mathbf{H}^o$ is zero by assumption, we can apply L'Hopital's Rule to (3.3.10)

$$\lim_{h\to 0^+} \frac{g(h,\lambda)}{h^2} = \lim_{h\to 0^+} \frac{(\mathbf{H}_z)^T z_h}{2h} \qquad (3.3.11)$$

But by Lemma 3.8(c) and Lemma 3.9(a)

$$\lim_{h\to 0^+} (\mathbf{H}_z)^T z_h = (\mathbf{H}_z^o)^T \left(\frac{\lambda+\lambda_o}{2}\right) J\, \mathbf{H}_z^o = \left(\frac{\lambda+\lambda_o}{2}\right) (\mathbf{H}_z^o)^T J\, (\mathbf{H}_z^o) = 0$$

where we have used (2.1.3). Therefore, applying L'Hopitals Rule a second time:

$$\lim_{h\to 0^+} \frac{g(h,\lambda)}{h^2} = \lim_{h\to 0^+} \frac{(\mathbf{H}_z)^T z_{hh} + (z_h)^T \mathbf{H}_{zz}(z_h)}{2}$$

Using Lemmas 3.8(c) and (d) and Lemma 3.9(a) and (d) we have:

$$\begin{aligned}
\lim_{h\to 0^+} \frac{g(h,\lambda)}{h^2} &= \lim_{h\to 0^+} \tfrac{1}{2}\Big((\mathbf{H}_z)^T z_{hh} + (z_h)^T \mathbf{H}_{zz}(z_h)\Big) \\
&= \tfrac{1}{2} (\mathbf{H}_z^o)^T \left(\frac{\lambda(\lambda+\lambda_o)}{2}\right) J\, \mathbf{H}_{zz}^o\, J\, \mathbf{H}_z^o + \tfrac{1}{2}\left(\frac{\lambda+\lambda_o}{2} J\, \mathbf{H}_z^o\right)^T \mathbf{H}_{zz}^o \left(\frac{\lambda+\lambda_o}{2} J\, \mathbf{H}_z^o\right) \\
&= \frac{\lambda(\lambda+\lambda_o)}{4} \left(J^T \mathbf{H}_z^o\right)^T \mathbf{H}_{zz}^o\, (J\, \mathbf{H}_z^o) + \frac{(\lambda+\lambda_o)^2}{8} (J\, \mathbf{H}_z^o)^T \mathbf{H}_{zz}^o (J\, \mathbf{H}_z^o) \\
&= -\frac{\lambda(\lambda+\lambda_o)}{4} \Psi_o + \frac{(\lambda+\lambda_o)^2}{8} \Psi_o \\
&= -\frac{(\lambda^2-\lambda_o^2)}{8} \Psi_o
\end{aligned}$$

Thus, (a) is established. Next, since

$$| g_\lambda(h,\lambda) | = | (H_z)^T z_\lambda |$$

$$\leq \sup_{\tilde{z}} \| H_z(\tilde{z}) \| \| z_\lambda \|$$

$$= M_2 \| z_\lambda \|$$

$| g_\lambda(h,\lambda) |$ is bounded on Q_o since by Lemma 3.9 $\| z_\lambda \|$ is bounded on Q_o.

Proof of (b):

$$\lim_{h \to 0^+} \frac{g_\lambda(h,\lambda)}{h^2} = \lim_{h \to 0^+} \frac{(H_z)^T z_\lambda}{h^2} \qquad (3.3.12)$$

From (3.3.5) we have that:

$$z_\lambda = \frac{h}{2} J\, H_z + \left(\frac{h\lambda}{2}\right) J\, H_{zz}\, z_\lambda$$

Substituting for z_λ in (3.3.12) implies:

$$\lim_{h \to 0^+} \frac{g_\lambda(h,\lambda)}{h^2} = \lim_{h \to 0^+} \frac{(H_z)^T \left(\frac{h}{2} J\, H_z + \left(\frac{h\lambda}{2}\right) J\, H_{zz}\, z_\lambda \right)}{h^2}$$

$$= \lim_{h \to 0^+} \frac{1}{2h} (H_z)^T J\, (H_z) + \lim_{h \to 0^+} \frac{\lambda}{2} (H_z)^T J\, H_{zz} \frac{z_\lambda}{h}$$

$$= \frac{\lambda}{2} (H_z^o)^T J\, H_{zz}^o (\tfrac{1}{2} J\, H_z^o)$$

$$= \frac{\lambda}{4} (J^T H_z^o)^T H_{zz}^o (J\, H_z^o)$$

$$= -\frac{\lambda}{4} \Psi_o$$

Next, since $g_\lambda = H_z\, z_\lambda$, we have:

$$g_{\lambda\lambda} = (H_z)^T z_{\lambda\lambda} + (z_\lambda)^T H_{zz}\, (z_\lambda)$$

$$| g_{\lambda\lambda} | \leq \| H_z \| \| z_{\lambda\lambda} \| + \| H_{zz} \| \| z_\lambda \|^2$$

$$\leq \; M_2 \| z_{\lambda\lambda} \| \; + \; M_3 \| z_\lambda \|^2$$

Thus $| g_{\lambda\lambda} |$ is bounded since by Lemma 3.9 both $\| z_\lambda \|$ and $\| z_{\lambda\lambda} \|$ are bounded on Q_o.

Proof of (c)

$$\lim_{h \to 0^+} \frac{g_{\lambda\lambda}(h,\lambda)}{h^2} = \lim_{h \to 0^+} \frac{(H_z)^T z_{\lambda\lambda}}{h^2} + \lim_{h \to 0^+} \left(\frac{z_\lambda}{h}\right)^T H_{zz} \left(\frac{z_\lambda}{h}\right)$$

$$= (H_z^o)^T \tfrac{1}{2} J \, H_{zz}^o \, J \, H_z^o + (\tfrac{1}{2} J \, H_z^o)^T H_{zz}^o (\tfrac{1}{2} J \, H_z^o)$$

$$= -\tfrac{1}{2} \Psi_o + \tfrac{1}{4} \Psi_o$$

$$= -\tfrac{1}{4} \Psi_o$$

Uniform convergence of (a) − (c) follows from the fact that each term which depends on h in the above expressions converges uniformly in λ by Lemma 3.8 and Lemma 3.9. ∎

Proof of Theorem 3.7:

By Theorem 3.5, $f(h,\lambda,z) = 0$ determines a C^2 function $z(h,\lambda)$ which maps $Q_o = (0, \delta_o) \times (0, 2\lambda_o)$ into D_o. Thus, $g(h,\lambda) = H(z(h,\lambda))$ is well defined on Q_o. $\left\{g(h,\lambda)\right\}$ is a family of real–valued functions parametrized by h. We will use the Newton–Kantorovich Theorem to show that there exists a $\delta_* < \delta_o$ such that for all $h \in (0, \delta_*)$, $g(h,\lambda) = 0$ has a locally unique zero, $\lambda(h) \in (0, 2\lambda_o)$ to which Newton's method converges quadratically.

Since $z(h,\lambda) \in C^2(Q_o)$ and $H(z) \in C^3(D_o)$, $g(h,\lambda) \in C^2$. For fixed $h_* \in (0, \delta_o)$ $g(h_*, \lambda)$ is a real valued function define on $(0, 2\lambda_o)$. From the Mean Value Theorem we have:

$$| g_\lambda(h_*,\lambda_2) - g_\lambda(h_*,\lambda_1) | = | g_{\lambda\lambda}(h_*,\tilde{\lambda}) | \, |\lambda_2 - \lambda_1|$$

By Lemma 3.10, $\sup_{\tilde{\lambda}} | g_{\lambda\lambda}(h_*,\tilde{\lambda}) | < \infty$ for all $h_* \in (0,\delta_o)$. Let $\gamma(h) = \sup_{\tilde{\lambda}} | g_{\lambda\lambda}(h,\tilde{\lambda}) |$.

Then:

$$| g_\lambda(h,\lambda_2) - g_\lambda(h,\lambda_1) | \le \gamma(h) |\lambda_2 - \lambda_1| \qquad \text{for } \lambda_1, \lambda_2 \in (0, 2\lambda_o), \; h \in (0,\delta_o)$$

Since by Lemma 3.10, $\frac{g_\lambda(h,\lambda_o)}{h^2} \to -\frac{\lambda_o}{4}\Psi_o \neq 0$ as $h \to 0^+$, there exists a $\delta_1 < \delta_o$ such that $g_\lambda(h,\lambda_o) \neq 0$ for all $h \in (0, \delta_1)$. Therefore, $\beta(h) = \frac{1}{|g_\lambda(h,\lambda_o)|}$, $\eta(h) = \frac{|g(h,\lambda_o)|}{|g_\lambda(h,\lambda_o)|}$ and $\alpha(h) = \beta(h)\gamma(h)\eta(h)$ are well defined for all $h < \delta_1$.

$$\begin{aligned}
\lim_{h\to 0^+} \alpha(h) &= \lim_{h\to 0^+} \beta(h)\gamma(h)\eta(h) \\
&= \lim_{h\to 0^+} \frac{\sup_{\bar\lambda} | g_{\lambda\lambda}(h,\bar\lambda) | \; | g(h,\lambda_o) |}{| g_\lambda(h,\lambda_o) |^2} \\
&= \lim_{h\to 0^+} \frac{\sup_{\bar\lambda} \left| \frac{g_{\lambda\lambda}(h,\bar\lambda)}{h^2} \right| \; \left| \frac{g(h,\lambda_o)}{h^2} \right|}{\left| \frac{g_\lambda(h,\lambda_o)}{h^2} \right|^2} \\
&= \frac{\frac{1}{4}|\Psi_o| \left(\frac{\lambda_o^2 - \lambda_o^2}{8}\right) |\Psi_o|}{\left(\frac{\lambda_o}{4}\right)^2 |\Psi_o|^2} \\
&= 0
\end{aligned}$$

where we have used Lemma 3.10 and the assumption $\Psi_o \neq 0$. Since $\lim_{h\to 0^+} \alpha(h) = 0$, there exists a $\delta_2 < \delta_o$ such that $\alpha(h) < \frac{1}{2}$. By Lemma 3.10(c) there exists a $\delta_3 > 0$ such that $\gamma(h) \neq 0$ for all $h \in (0,\delta_3)$. Thus, there exists a $\delta_4 = \min(\delta_1, \delta_3)$ such that $\beta(h) \neq 0$ and $\gamma(h) \neq 0$ for all $h \in (0, \delta_4)$. Define $r_-(h)$ on $(0,\delta_4)$ as follows:

$$r_-(h) = \frac{1}{\beta(h)\gamma(h)}\left(1 - \sqrt{1 - 2\alpha(h)}\right)$$

Using Lemma 3.10 and the fact that $\alpha(h) \to 0$ as $h \to 0^+$ we have:

$$\lim_{h\to 0^+} r_-(h) = \lim_{h\to 0^+} \frac{| g_\lambda(h,\lambda_o) |}{\sup_{\bar\lambda} | g_{\lambda\lambda}(h,\bar\lambda) |} \left(1 - \sqrt{1 - 2\alpha(h)}\right)$$

$$= \lim_{h\to 0^+} \frac{\left| \frac{g_\lambda(h,\lambda_o)}{h^2} \right|}{\sup_{\tilde\lambda} \left| \frac{g_{\lambda\lambda}(h,\tilde\lambda)}{h^2} \right|} \left(1-\sqrt{1-2\alpha(h)}\,\right)$$

$$\le \lim_{h\to 0^+} \frac{\left| \frac{g_\lambda(h,\lambda_o)}{h^2} \right|}{\left| \frac{g_{\lambda\lambda}(h,\lambda_o)}{h^2} \right|} \left(1-\sqrt{1-2\alpha(h)}\,\right)$$

$$= \frac{\frac{\lambda_o}{4}\,|\Psi_o|\,(\,0\,)}{\frac{1}{4}|\,\Psi_o|}$$

$$= 0$$

Since $\lim_{h\to 0^+} r_-(h) = 0$, there exists a $\delta_5 < \delta_o$ such that

$$\left[\lambda_o - r_-(h),\ \lambda_o + r_-(h)\right] \subset (0,\, 2\lambda_o) \qquad \text{for all } h \in (0,\, \delta_5)$$

All the conditions of the Newton–Kantorovich Theorem are satisfied for each $h \in (0,\delta_*)$ where $\delta_* = \min\{\delta_2,\ \delta_4,\ \delta_5\}$. We conclude that for each $h \in (0,\delta_*)$, Newton's method converges quadratically to a zero $\lambda(h)$ of $g(h,\lambda) = 0$. This zero is unique in the set:

$$\left[\lambda_o - r_+(h),\ \lambda_o + r_+(h)\right] \cap\ (0,\, 2\lambda_o)$$

where:

$$r_+ = \frac{1}{\beta(h)\gamma(h)}\left(1+\sqrt{1-2\alpha(h)}\,\right)$$

∎

3.4 Local Existence and Uniqueness of DTH Trajectories

THEOREM 3.11: (Main Result)
Assume $\mathbf{H} \in C^3(U)$ where $U \subset \mathbf{R}^{2n+2}$ is open. Assume also that $\lambda_o > 0$ and that there exists a $\overline{z}^{(0)} \in U$ such that $\mathbf{H}(\overline{z}^{(0)}) = \mathbf{H}^o = 0$ and $\Psi(\overline{z}^{(0)}) = \Psi_o \neq 0$ where:

$$\Psi(z) \;=\; (J\,\mathbf{H}_z(z))^T \mathbf{H}_{zz}(z)\,(J\,\mathbf{H}_z(z)) \tag{3.4.1}$$

Then, for any positive integer N, there exists a time step $\Delta\tau$ and a locally unique piecewise-linear, continuous trajectory determined by $\overline{z}^{(k)}$ and $\overline{z}'^{(k)}$, where $\overline{z}^{(k)}$ and $\overline{z}'^{(k)}$ satisfy the DTH equations:

$$\frac{\overline{z}^{(k+1)} - \overline{z}^{(k)}}{\Delta\tau} = \frac{1}{2} J \left[\lambda_{k+1} \frac{\partial H(\overline{z}^{(k+1)})}{\partial \overline{z}^{(k+1)}} + \lambda_k \frac{\partial H(\overline{z}^{(k)})}{\partial \overline{z}^{(k)}} \right] \quad k = 0, 1, \cdots N-2 \qquad (3.4.2)$$

$$\overline{z}'^{(k)} = \lambda_k J \frac{\partial H(\overline{z}^{(k)})}{\partial \overline{z}^{(k)}} \quad k = 0, 1, \cdots N-1 \qquad (3.4.3)$$

$$\mathbf{H}(\overline{z}^{(k)}) = 0 \qquad k = 0, 1, \cdots N-1 \qquad (3.4.4)$$

Proof:

Since $\Psi(z)$ is continuous on U and $\Psi(\overline{z}^{(0)}) \neq 0$, there exists an open ball $D_o = B(\overline{z}^{(0)}, r_o)$ such that $\overline{D}_o \subset U$ and $\Psi(z) \neq 0$ for all $z \in D_o$. D_o is open bounded and convex since it is an open ball of $\mathbf{R}^{2n+2}$. Assume $\lambda_k > 0$, $\overline{z}^{(k)} \in D_o$, and $\mathbf{H}(\overline{z}^{(k)}) = 0$. Let:

$$f(h,\lambda,z) = z - \overline{z}^{(k)} - \frac{h}{2} J \left[\lambda \mathbf{H}_z(z) + \lambda_k \mathbf{H}_z(\overline{z}^{(k)}) \right]$$

By Theorem 3.5 there exists a $\delta_o^k > 0$, $Q_k = Q(\delta_o^k)$ and a function $z(h,\lambda)$: $Q_k \rightarrow D_o$ such that $f(h,\lambda,z(h,\lambda)) = 0$ for $(h,\lambda) \in Q_k$. Let $g(h,\lambda) = \mathbf{H}(z(h,\lambda))$. Since $\overline{z}^{(k)} \in D_o$, $\Psi(\overline{z}^{(k)}) \neq 0$. Theorem 3.7 implies there exists a $\delta_*^k < \delta_o^k$ and a function $\lambda(h)$: $(0,\delta_*^k) \rightarrow (0,2\lambda_k)$ such that $g(h,\lambda(h)) = 0$ for all $h \in (0,\delta_*^k)$. Choose any $h_k < \delta_*^k$ and let $\lambda_{k+1} = \lambda(h_k)$ and $\overline{z}^{(k+1)} = z(h_k,\lambda_{k+1})$. Then $\lambda_{k+1} \in (0,2\lambda_k)$ and thus $\lambda_{k+1} > 0$, and $\overline{z}^{(k+1)} \in D_o$ since the range of $z(h,\lambda)$ is in D_o.

$$\mathbf{H}(\overline{z}^{(k+1)}) = \mathbf{H}(z(h_k,\lambda_{k+1})) = g(h_k,\lambda_{k+1}) = 0 \qquad (3.4.5)$$

Since $(h_k, \lambda_{k+1}) \in Q_k$, we also have by Theorem 3.5 that:

$$f(h_k,\lambda_{k+1}, \overline{z}^{(k+1)}) = f(h_k,\lambda_{k+1}, z(h_k,\lambda_{k+1})) = 0 \qquad (3.4.6)$$

Equation (3.4.5) and (3.4.6) imply that given $h_k < \delta_*^k$, $(\lambda_{k+1}, \overline{z}^{(k+1)})$ satisfies the DTH equations (3.4.2) and (3.4.4) for the k^{th} time step. Moreover, there exists an open set in which $(\lambda_{k+1}, \overline{z}^{(k+1)})$ is the only solution to (3.4.2) and (3.4.4).

By assumption, $\lambda_0 > 0$, $\overline{z}^{(0)} \in D_o$, and $\mathbf{H}(\overline{z}^{(0)}) = 0$. Thus, for $k = 0$, $\lambda_k > 0$, $\overline{z}^{(k)} \in D_o$, and $\mathbf{H}(\overline{z}^{(k)}) = 0$. Therefore, by induction, there exists a sequence of values $(h_k, \lambda_{k+1}, \overline{z}^{(k+1)})$, $k = 0, 1, \cdots N-2$, which satisfy the DTH equations (3.4.2) and (3.4.4). Let $\Delta\tau < \min\left\{\delta_*^o, \delta_*^1, \delta_*^2, \cdots \delta_*^{N-1}\right\}$. Then $(\Delta\tau, \lambda_{k+1}, \overline{z}^{(k+1)})$ is also a solution to (3.4.2) and (3.4.4) for $k = 0, 1, \cdots N-2$.

Assume $\overline{z}'^{(k)}$ is given by equation (3.4.3). Then $\overline{z}^{(k)}$ and $\overline{z}'^{(k)}$ determine a piecewise–linear trajectory. Substituting $\overline{z}'^{(k)}$ in equation (3.4.2) implies that $\overline{z}^{(k)}$ and $\overline{z}'^{(k)}$ are related by the equation:

$$\frac{\overline{z}^{(k+1)} - \overline{z}^{(k)}}{\Delta\tau} = \frac{\overline{z}'^{(k)} + \overline{z}'^{(k+1)}}{2} \tag{3.4.7}$$

Equation (3.4.7) shows $\overline{z}'^{(k)}$ satisfys the continuity constraint. Therefore, the trajectory determined by $\overline{z}^{(k)}$ and $\overline{z}'^{(k)}$ is not only piecewise–linear, but also continuous. ∎

The condition $\Psi(\overline{z}^{(0)}) \neq 0$ in Theorem 3.4 has only been shown to be a sufficient condition for local existence and uniqueness of DTH trajectories. For autonomous Hamiltonians with positive–definite Hessian matrices, $\Psi(\overline{z}^{(0)}) = 0$ if and only if $\overline{\zeta}^{(0)}$ is a stationary point of the Hamiltonian vector field of $H(\zeta)$. (Recall that $\zeta \in \mathbf{R}^{2n}$.) In this case, the analytical solution is constant. Typically, $\Psi(z)$ may be zero on a set of isolated points. The condition that $\Psi(z) \equiv 0$ on an open set is much more restrictive. Systems with linear Hamiltonian functions, for example, have $\Psi(z) \equiv 0$. For one degree of freedom systems, $\Psi(z) \equiv 0$ if one of the coordinates is cyclic. Both cases are easily solved analytically.

3.5 Positive–Definite Linear Hamiltonian Systems

In this section we will show that for autonomous, positive–definite, linear Hamiltonian systems, if h_o is sufficiently small, the DTH equations have only two solutions, one of which is constant. For these solutions, $\lambda_{k+1} = \pm \lambda_k$ where the minus sign corresponds to the constant solution.

THEOREM 3.12:
Assume $H(\zeta)$ is an autonomous, quadratic Hamiltonian function given by:

$$H(\zeta) = \frac{1}{2}(\zeta)^T A\,(\zeta) + b^T\zeta + c \tag{3.5.1}$$

where $\zeta \in \mathbb{R}^{2n}$ and A is a symmetric, positive–definite matrix. Assume $\lambda_o \neq 0$ and $\zeta^o \neq 0$. Then for sufficiently small $h_o > 0$, (λ_1, ζ^1) is a solution to the DTH equations if and only if $\lambda_1 = \pm \lambda_o$.

Proof:

First, we show that there is no loss in generality if we assume:

$$H(\zeta) = \frac{1}{2}(\zeta)^T A\,(\zeta) \tag{3.5.2}$$

Given any positive–definite quadratic function:

$$Q(\eta) = \frac{1}{2}(\eta)^T A\,(\eta) + b^T\eta + c \tag{3.5.3}$$

we can use the translation:

$$\eta = \zeta - A^{-1}b \tag{3.5.4}$$

to reduce $Q(\eta)$ to the form given by (3.5.2) plus some constant d. Substituting (3.5.4) in (3.5.3) we have:

$$\begin{aligned} H(\zeta) &= Q(\zeta - A^{-1}b) \\ &= \frac{1}{2}(\zeta - A^{-1}b)^T A\,(\zeta - A^{-1}b) + b^T(\zeta - A^{-1}b) + c \\ &= \frac{1}{2}(\zeta)^T A\,(\zeta) + d \end{aligned} \tag{3.5.5}$$

where $$d = c - \tfrac{1}{2}\, b^T A^{-1} b$$

Since Hamiltonian functions differing by a constant have the same equations, we may assume $d = 0$ in (3.5.5).

For autonomous systems, the DTH equations can be reduced to the form given by equations (2.3.5) – (2.3.7) of Chapter II. For Hamiltonians of the form given by (3.5.2) these equations become:

$$\zeta - \zeta^o = \frac{h_o}{2} J\, A\, (\lambda\zeta + \lambda_o \zeta^o) \tag{3.5.6}$$

$$\tfrac{1}{2}(\zeta)^T A\, (\zeta) - \tfrac{1}{2}(\zeta^o)^T A\, (\zeta^o) = 0 \tag{3.5.7}$$

(Equation (2.3.6) is not used in the proof.) Solving for ζ in (3.5.6) we have:

$$\zeta = \left[I - (\frac{h_o\lambda}{2})\, J\, A\right]^{-1} \left[I + (\frac{h_o\lambda_o}{2})\, J\, A\right] \zeta^o \tag{3.5.8}$$

where by the Matrix Perturbation Lemma (Lemma 3.3) $\left[I - (\frac{h_o\lambda}{2})\, J\, A\right]^{-1}$ exists for sufficiently small values of h_o and for all $\lambda \in (0, 2\lambda_o)$. Assume $\zeta^1 = \zeta(h_o, \lambda_1)$ is a solution to equations (3.5.6) – (3.5.7). Then we have:

$$\begin{aligned}
0 &= \tfrac{1}{2}(\zeta^1)^T A\, (\zeta^1) - \tfrac{1}{2}(\zeta^o)^T A\, (\zeta^o) \\
&= \tfrac{1}{2}(\zeta^1 + \zeta^o)^T A\, (\zeta^1 - \zeta^o) \\
&= \tfrac{1}{2}(\zeta^1 + \zeta^o)^T A \Big(\frac{h_o}{2} J\, A\, (\lambda_1 \zeta + \lambda_o \zeta^o)\Big) \\
&= \frac{h_o}{4}\Big(\lambda_1 (\zeta^1)^T A\, J\, A\, (\zeta^1) + \lambda_1 (\zeta^o)^T A\, J\, A\, (\zeta^1) \\
&\qquad + \lambda_o (\zeta^1)^T A\, J\, A\, (\zeta^o) + \lambda_o (\zeta^o)^T A\, J\, A\, (\zeta^o)\Big) \\
&= \frac{h_o}{4}(\lambda_1 - \lambda_o)(\zeta^o)^T A\, J\, A\, (\zeta^1)
\end{aligned}$$

where we have used the facts that $(\mathrm{A\,J\,A})^T = -(\mathrm{A\,J\,A})$ and:

$$(\zeta)^T \mathrm{A\,J\,A}(\zeta) \;=\; 0 \quad \text{for any } \zeta \in \mathbf{R}^{2n} \tag{3.5.9}$$

Thus either $\lambda_1 = \lambda_o$ as claimed or:

$$(\zeta^o)^T \mathrm{A\,J\,A}(\zeta^1) \;=\; 0 \tag{3.5.10}$$

Using (3.5.6) with $\zeta = \zeta^1$ and $\lambda = \lambda_1$ and using (3.5.9) and (3.5.10) implies:

$$\begin{aligned}(\zeta^o)^T \mathrm{A}(\zeta^1) - (\zeta^o)^T \mathrm{A}(\zeta^o) \;&=\; (\zeta^o)^T \mathrm{A}(\zeta^1 - \zeta^o)\\ &=\; (\tfrac{h_o\lambda_1}{2})(\zeta^o)^T \mathrm{A\,J\,A}(\zeta^1) + (\tfrac{h_o\lambda_o}{2})(\zeta^o)^T \mathrm{A\,J\,A}(\zeta^o)\\ &=\; 0\end{aligned}$$

Therefore

$$(\zeta^o)^T \mathrm{A}(\zeta^1) \;=\; (\zeta^o)^T \mathrm{A}(\zeta^o) \tag{3.5.11}$$

We also have from (3.5.7) that for $\zeta = \zeta^1$

$$(\zeta^1)^T \mathrm{A}(\zeta^1) \;=\; (\zeta^o)^T \mathrm{A}(\zeta^o) \tag{3.5.12}$$

Equations (3.5.11) and (3.5.12) imply that:

$$(\zeta^1 - \zeta^o)^T \mathrm{A}(\zeta^1 - \zeta^o) \;=\; (\zeta^1)^T \mathrm{A}(\zeta^1) - 2(\zeta^1)^T \mathrm{A}(\zeta^o) + (\zeta^o)^T \mathrm{A}(\zeta^o) \;=\; 0$$

Since A is positive–definite, we must have $\zeta^1 = \zeta^o$. Substituting ζ^o for ζ in (3.5.6) we have:

$$0 \;=\; \frac{h_o}{2}\mathrm{J\,A}(\lambda_1\zeta^o + \lambda_o\zeta^o) \;=\; \frac{h_o(\lambda_1 + \lambda_o)}{2}\mathrm{J\,A}\,\zeta^o$$

Since J A is nonsingular and $\zeta^o \neq 0$ we must have $\lambda_1 = -\lambda_o$ as claimed.

Assume $\lambda_1 = \lambda_o$. Then $\zeta^1 = \zeta(h_o,\lambda_o)$ and since ζ^1 is a solution of (3.5.6)

$$\zeta^1 - \zeta^o \;=\; \frac{h_o\lambda_o}{2} J\,A\,(\zeta^1 + \zeta^o) \tag{3.5.13}$$

We use (3.5.13) to show that ζ^1 also satisfies (3.5.7).

$$\begin{aligned}
\tfrac{1}{2}(\zeta)^T A\,(\zeta) \;-\; \tfrac{1}{2}(\zeta^o)^T A\,(\zeta^o) \;&=\; \tfrac{1}{2}(\zeta^1 + \zeta^o)^T A\,(\zeta^1 - \zeta^o) \\
&=\; \tfrac{1}{2}(\zeta^1 + \zeta^o)^T A\Big(\frac{h_o\lambda_o}{2} J\,A\,(\zeta^1 + \zeta^o)\Big) \\
&=\; \frac{h_o\lambda_o}{4}(\zeta^1 + \zeta^o)^T A\,J\,A\,(\zeta^1 + \zeta^o) \\
&=\; 0
\end{aligned}$$

where the last expression is zero by (3.5.9). Thus $\zeta^1 = \zeta(h_o,\lambda_o)$ is a solution to the DTH equations.

Assume $\lambda_1 = -\lambda_o$. Then $\zeta^1 = \zeta(h_o,-\lambda_o)$ and (3.5.8) implies:

$$\zeta^1 \;=\; \zeta(h_o,-\lambda_o) \;=\; \Big[I + (\frac{h_o\lambda_o}{2})\,J\,A\Big]^{-1}\Big[I + (\frac{h_o\lambda_o}{2})\,J\,A\Big]\zeta^o \;=\; \zeta^o$$

Substituting ζ^o for ζ in (3.5.7) we see that $\lambda_1 = -\lambda_o$ also yeilds a solution to the DTH equations, albeit a constant one. ∎

The proof extends in the indicated fashion for any number of time steps. For autonomous, positive–definite, linear systems, Theorem 3.12 states that the only nonconstant solution to the DTH equations is the one for which $\lambda_{k+1} = \lambda_k$. If we choose $\lambda_o = 1$, then by induction, $\lambda_k \equiv 1$ for the nonconstant solution. If $\lambda_k \equiv 1$, equation (2.3.5) reduces to the trapezoidal method and equation (2.3.6) reduces to the midpoint method. We conclude that for autonomous, positive–definite, linear Hamiltonian systems, DTH trajectories can be computed by using the trapezoidal and midpoint rules for integrating differential equations.

CHAPTER IV

NUMERICAL RESULTS

4.1 An Algorithm for Computing DTH Trajectories

Theorem 3.11 provides us with a method for computing DTH trajectories. For fixed $\Delta\tau$, we solve the equation $g(\Delta\tau, \lambda_{k+1}) = 0$ for λ_{k+1} by using Newton's method:

$$\lambda_{k+1}^{(i+1)} = \lambda_{k+1}^{(i)} - \frac{g(\Delta\tau, \lambda_{k+1}^{(i)})}{g_\lambda(\Delta\tau, \lambda_{k+1}^{(i)})} \tag{4.1.1}$$

where $\lambda_{k+1}^{(i)}$ is the i^{th} Newton iterate of $\lambda_{k+1}^{(0)}$ where $\lambda_{k+1}^{(0)} = \lambda_k$. In order to evaluate $g(\Delta\tau, \lambda_{k+1}^{(i)})$ and $g_\lambda(\Delta\tau, \lambda_{k+1}^{(i)})$ in (4.1.1) we need to evaluate $z(\Delta\tau, \lambda_{k+1}^{(i)})$. For fixed values of $\Delta\tau$ and $\lambda_{k+1}^{(i)}$ we can evaluate $z(\Delta\tau, \lambda_{k+1}^{(i)})$ by solving $f(\Delta\tau, \lambda_{k+1}^{(i)}, \overline{z}^{(k+1)}) = 0$ for $\overline{z}^{(k+1)}$ using Newton's method:

$$\overline{z}^{(k+1)}(j+1) = \overline{z}^{(k+1)}(j) - f_{\overline{z}}^{-1}\Big(\Delta\tau, \lambda_{k+1}, \overline{z}^{(k+1)}(j)\Big)\, f\Big(\Delta\tau, \lambda_{k+1}, \overline{z}^{(k+1)}(j)\Big) \tag{4.1.2}$$

where $\overline{z}^{(k+1)}(j)$ is the j^{th} Newton iterate of $\overline{z}^{(k+1)}(0)$ where $\overline{z}^{(k+1)}(0) = \overline{z}^{(k)}$. Thus, we have an " inner " iteration of Newton's method given by (4.1.1) and an " outer " iteration given by (4.1.2). The Newton–Kantorovich Theorem (Theorem 3.1) implies that both iterations converge quadratically. Moreover, for sufficiently small values of $\Delta\tau$, the linear systems which must be solved at each stage are diagonally dominant.

4.2 The Kepler Problem in Cartesian Coordinates

In this section, we will illustrate DTH dynamics by considering the Kepler problem, also known as the one body central force problem. In cartesian coordinates, for appropriately chosen parameters, the Hamiltonian function for the Kepler problem is:

$$H(q_1, q_2, p_1, p_2) = p_1^2 + p_2^2 - (q_1^2 + q_2^2)^{-\frac{1}{2}} \tag{4.2.1}$$

In cartesian coordinates, the angular momentum for this problem is the quadratic function:

$$L(q_1, q_2, p_1, p_2) = q_1 p_2 - q_2 p_1 \tag{4.2.4}$$

$L(q_1, q_2, p_1, p_2)$ is conserved as can be verified by evaluating the Possion bracket $[H, L]$.

The exact values of $q_1(t)$, $q_2(t)$, $p_1(t)$ and $p_2(t)$ for 1 orbit of the Kepler problem for the initial conditions $q_1(0) = 0.5$, $q_2(0) = 0$, $p_1(0) = 0$ and $p_2(0) = 1.2$, are shown in Figures 4.1 – 4.4. Also shown are approximate values obtained by the trapezoidal and midpoint methods, discrete mechanics and DTH dynamics. (Midpoint values were used for the DTH trajectories.) In order to amplify discretization errors, only 25 points per orbit was used. This was achieved by choosing $\Delta t = 0.15$ for the trapezoidal, midpoint and discrete mechanics methods and choosing $\Delta\tau = 0.06277$ for DTH dynamics. (Recall that DTH dynamics has a varying step size.) From Figures 4.1 – 4.4 we see that both the trapezoidal and midpoint methods have significant phase errors – the trapezoidal method has not completed an orbit, while the midpoint method has begun a second orbit.

Table 4.1 illustrates conservation laws obeyed by each method. The midpoint method, discrete mechanics and DTH dynamics conserve angular momentum up to roundoff error while only discrete mechanics and DTH dynamics conserve energy up to roundoff.

TABLE 4.1: Average absolute deviation in energy and angular momentum during 400 orbits of the Kepler problem with 25 points per orbit.

	Energy	Angular Momentum
Trapezoidal Method	1.322×10^{-1}	4.593×10^{-2}
Midpoint Method	5.969×10^{-2}	7.387×10^{-16}
Discrete Mechanics	3.901×10^{-15}	1.738×10^{-15}
DTH Dynamics (midpoints)	1.285×10^{-14}	8.204×10^{-4}
DTH Dynamics (vertices)	1.750×10^{-2}	2.221×10^{-15}

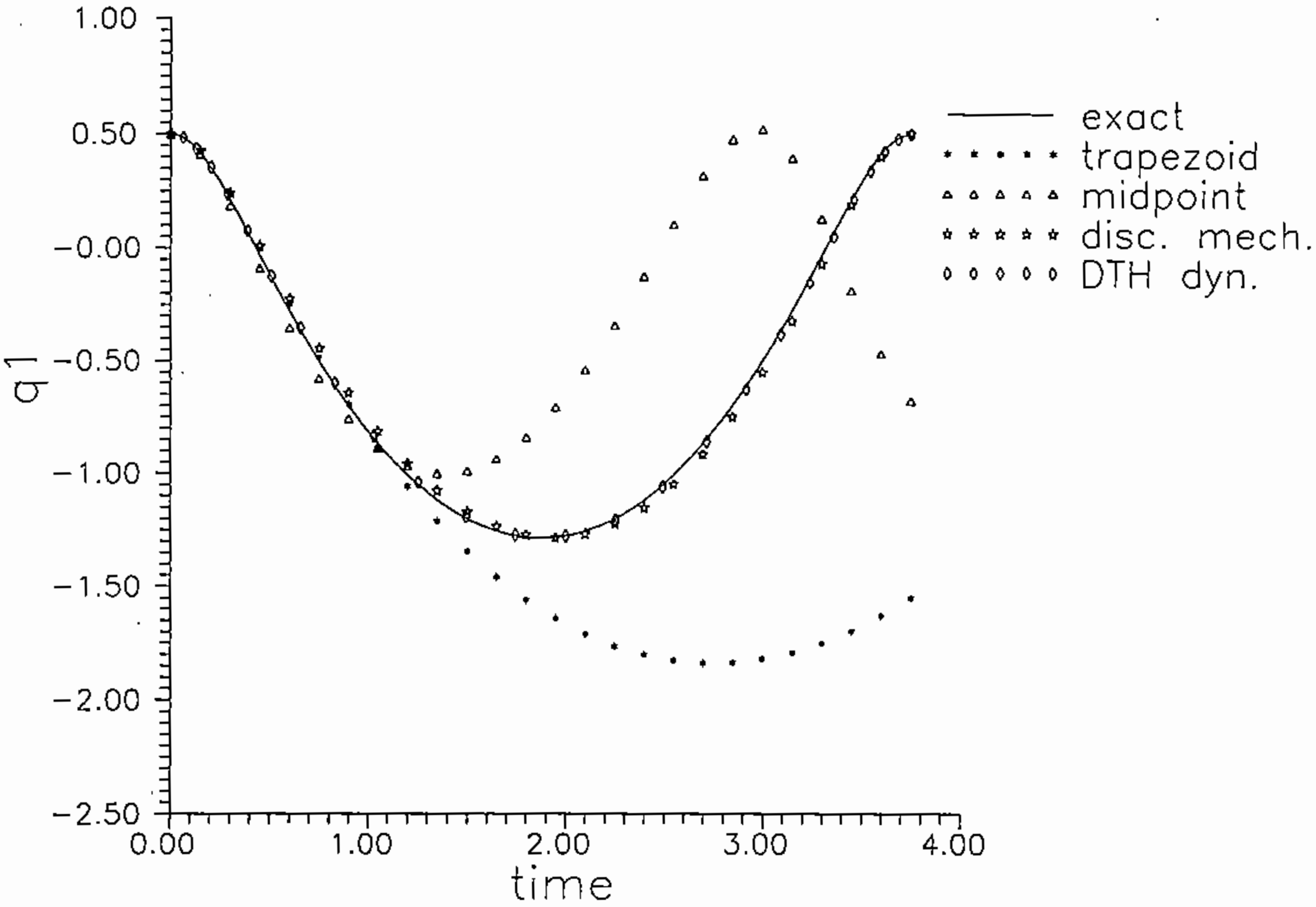


FIGURE 4.1: Kepler problem, 25 points per orbit, 1 orbit.

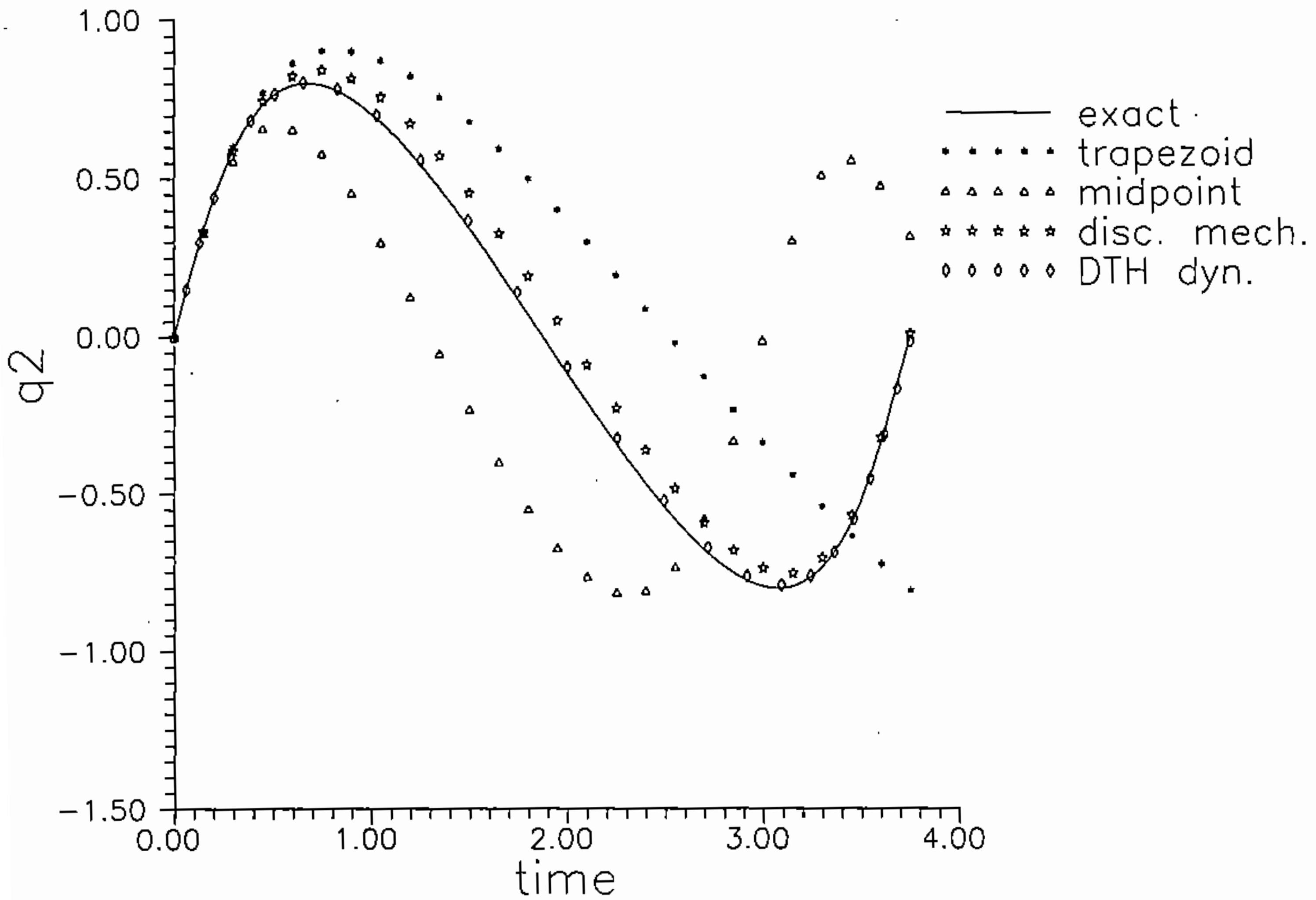


FIGURE 4.2: Kepler problem, 25 points per orbit, 1 orbit.

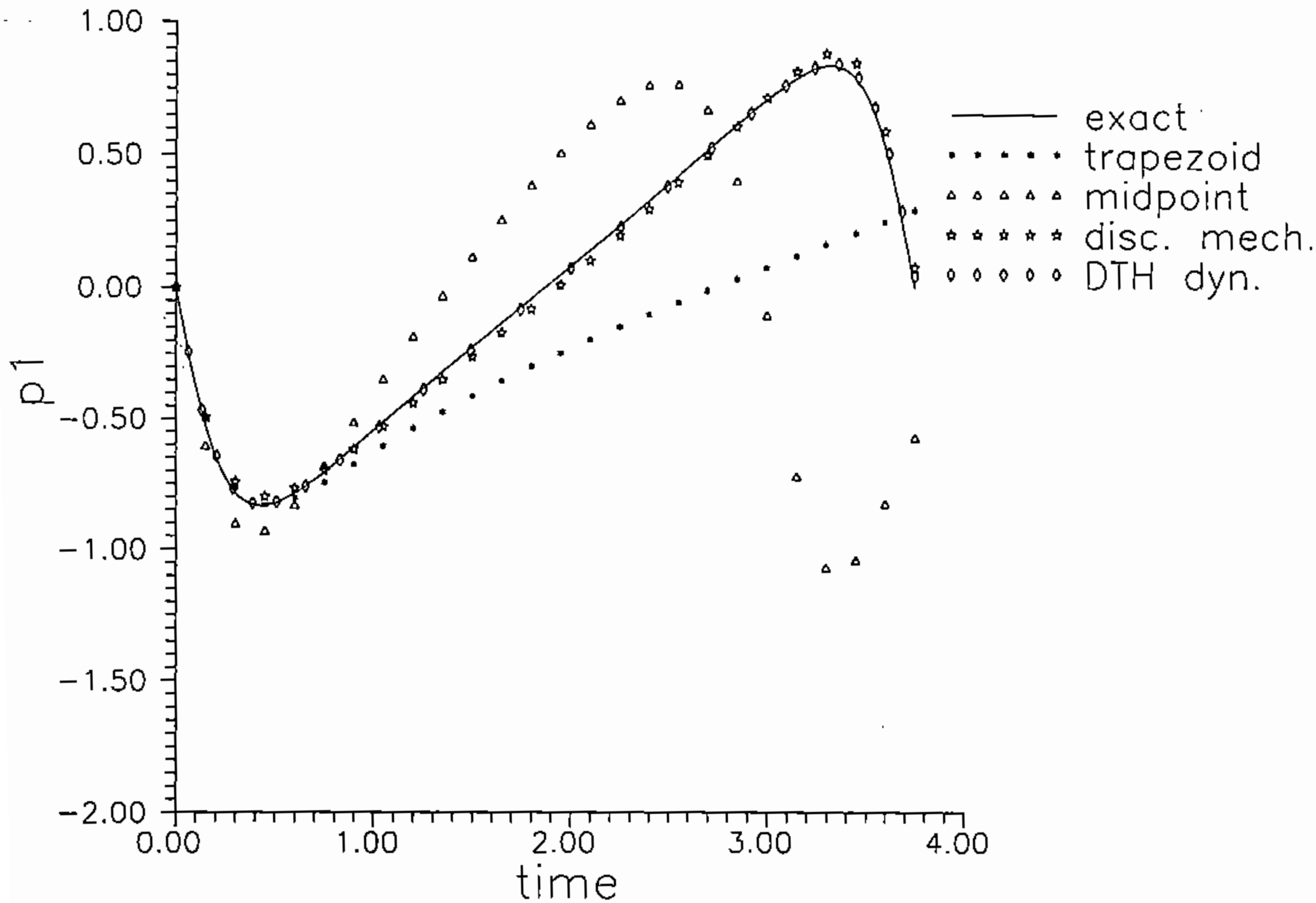


FIGURE 4.3: Kepler problem, 25 points per orbit, 1 orbit.

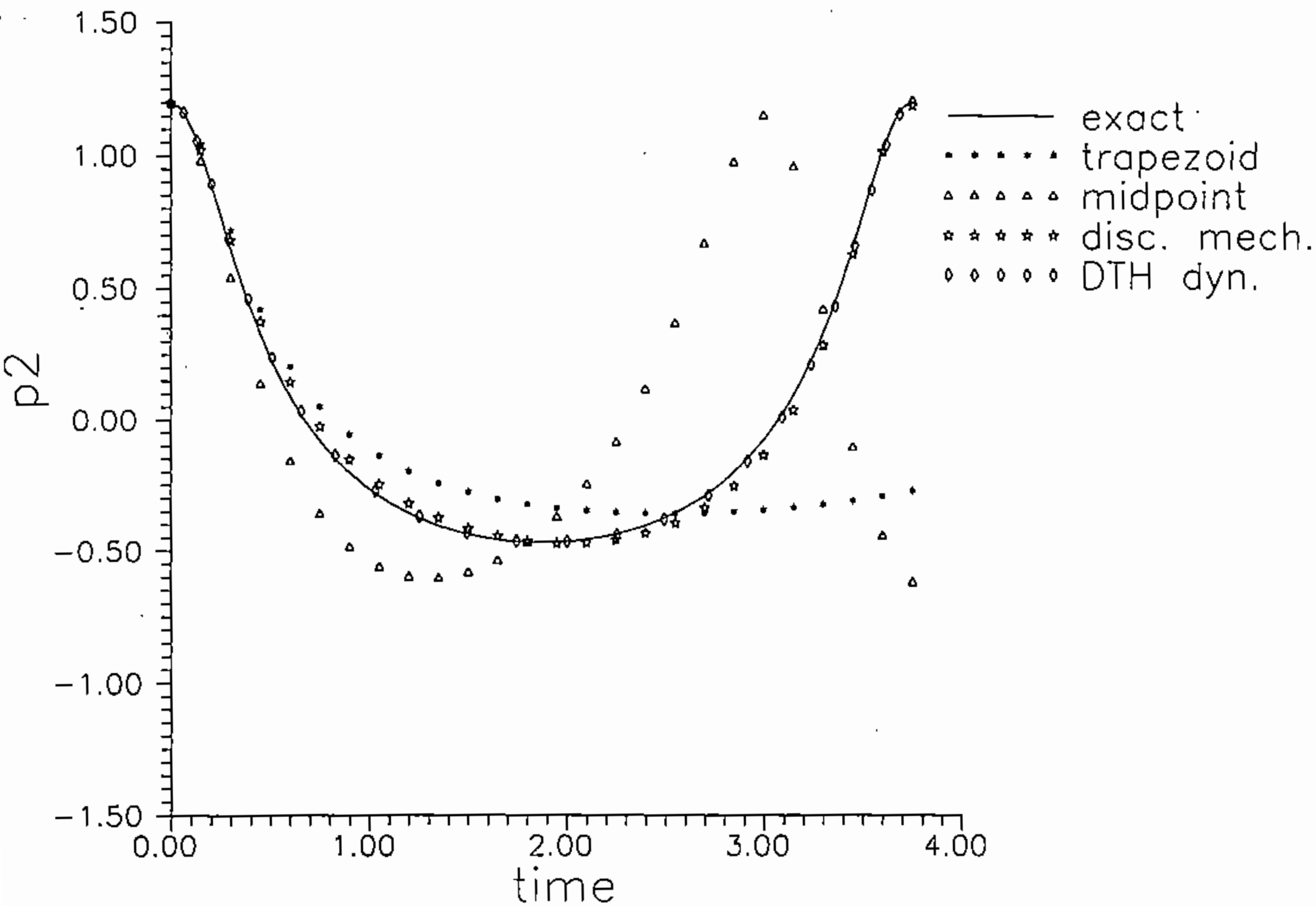
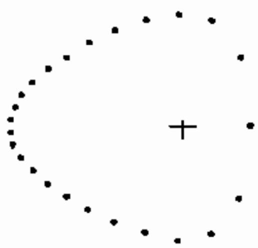
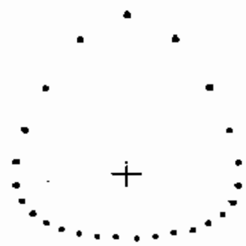


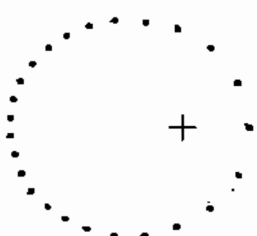
FIGURE 4.4: Kepler problem, 25 points per orbit, 1 orbit.



(a) q_1 vs p_1



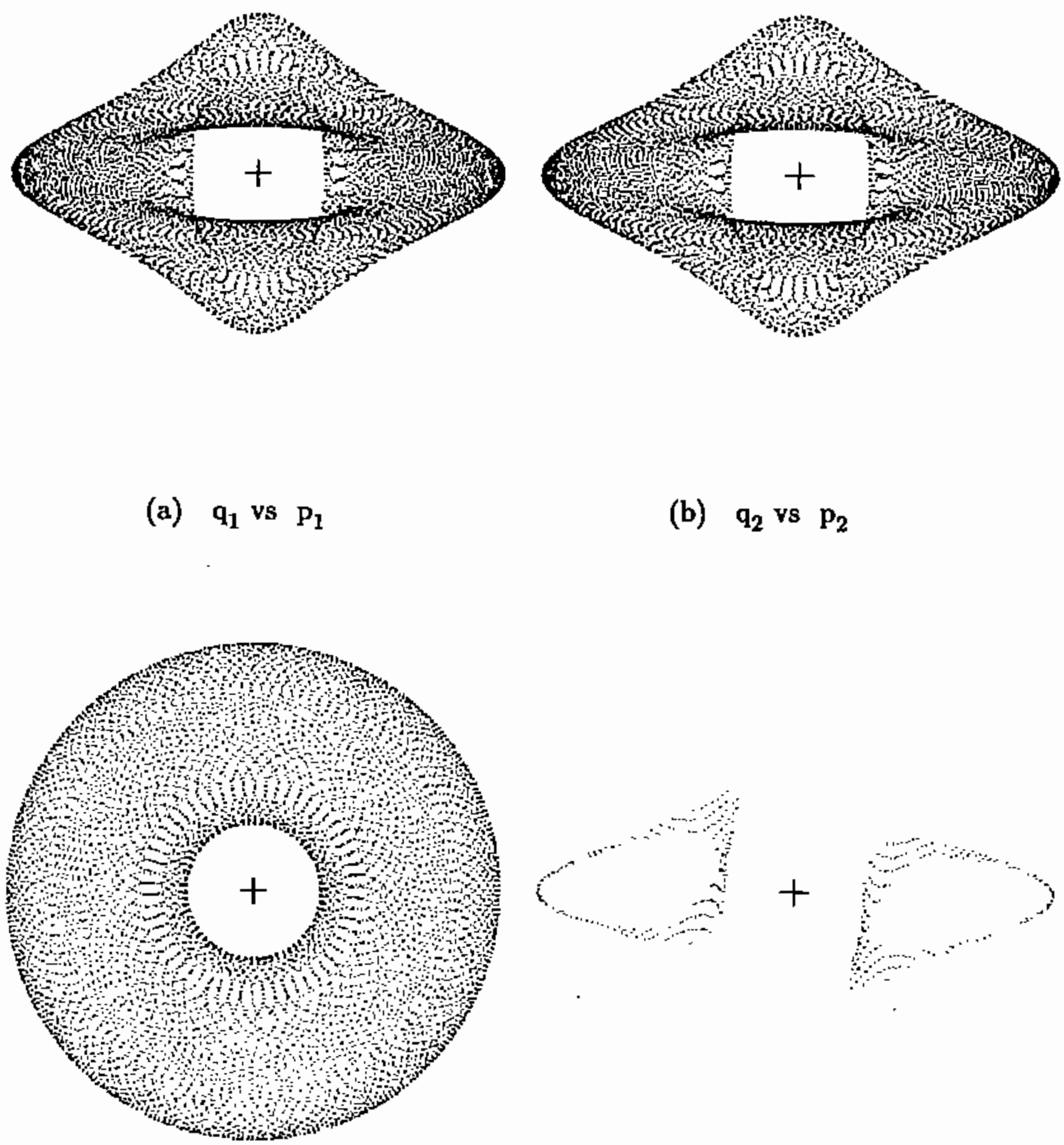
(b) q_2 vs p_2



(c) q_1 vs q_2

(d) q_1 vs p_1, $q_2 = 0$

FIGURE 4.5: Kepler problem, exact orbit, 25 points per orbit, 400 orbits



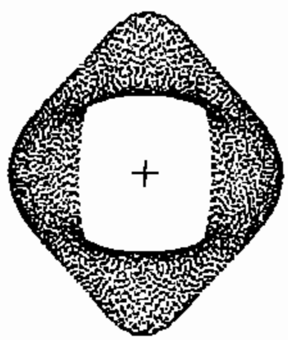
(a) q_1 vs p_1

(b) q_2 vs p_2

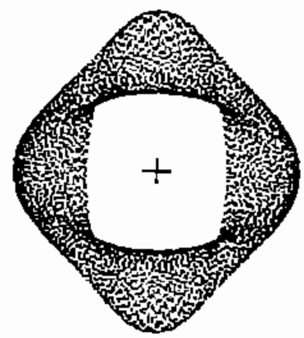
(c) q_1 vs q_2

(d) q_1 vs p_1, $q_2 = 0$

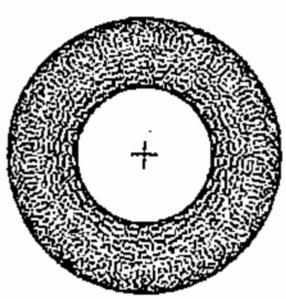
FIGURE 4.6: Kepler problem, trapezoidal method, 25 points per orbit, 400 orbits



(a) q_1 vs p_1



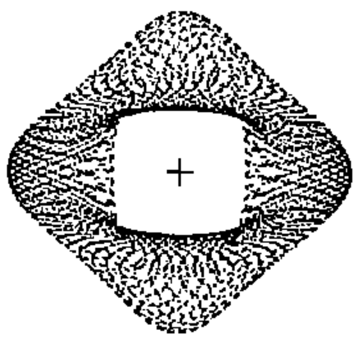
(b) q_2 vs p_2



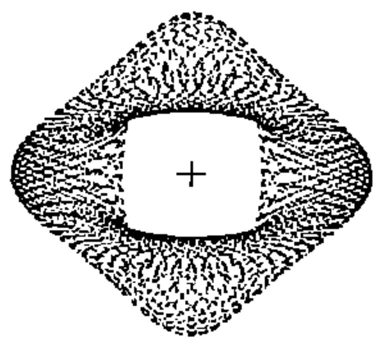
(c) q_1 vs q_2

(d) q_1 vs p_1, $q_2 = 0$

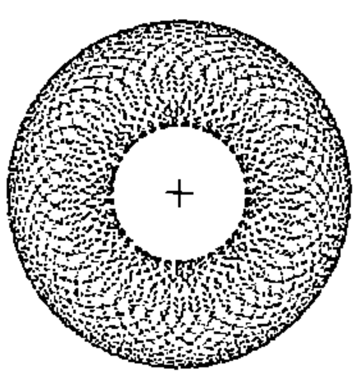
FIGURE 4.7: Kepler problem, midpoint method, 25 points per orbit, 400 orbits



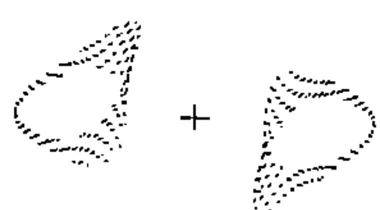
(a) q_1 vs p_1



(b) q_2 vs p_2

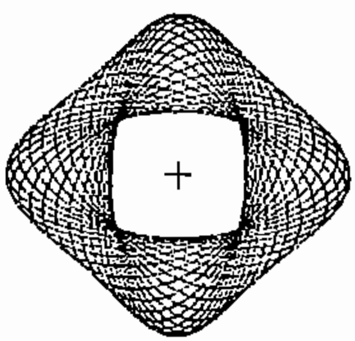


(c) q_1 vs q_2

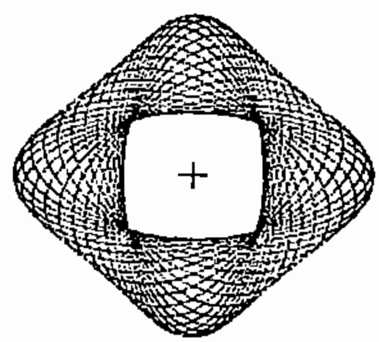


(d) q_1 vs p_1, $q_2 = 0$

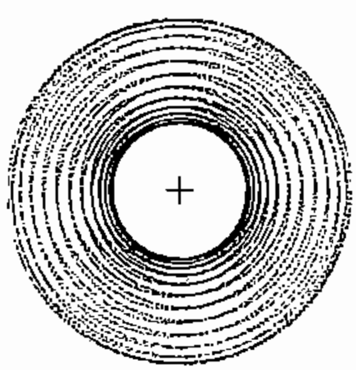
FIGURE 4.8: Kepler problem, discrete mechanics, 25 points per orbit, 400 orbits



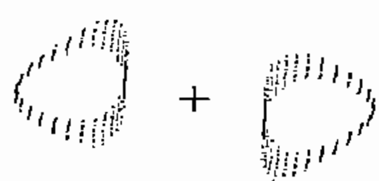
(a) q_1 vs p_1



(b) q_2 vs p_2

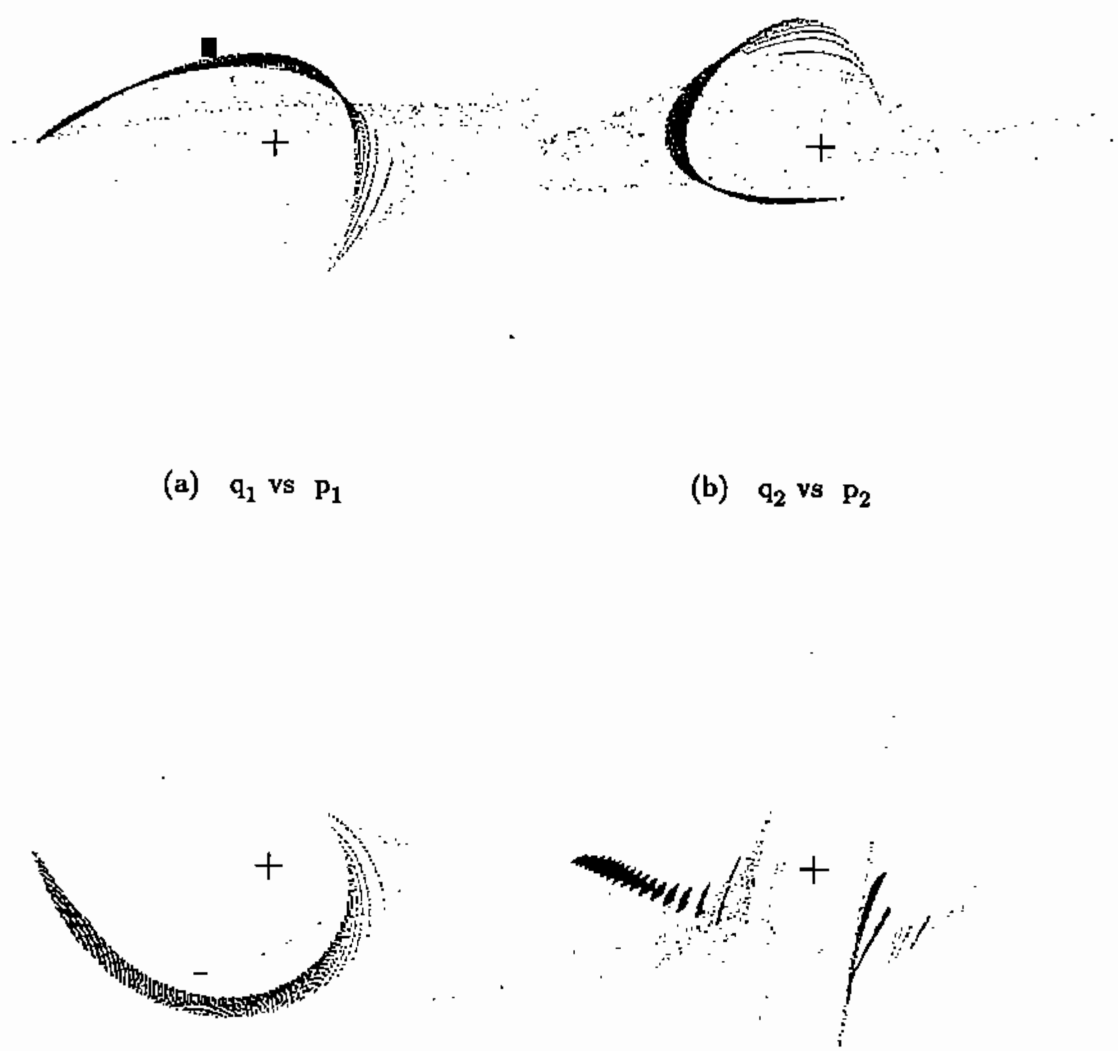


(c) q_1 vs q_2



(d) q_1 vs p_1, $q_2 = 0$

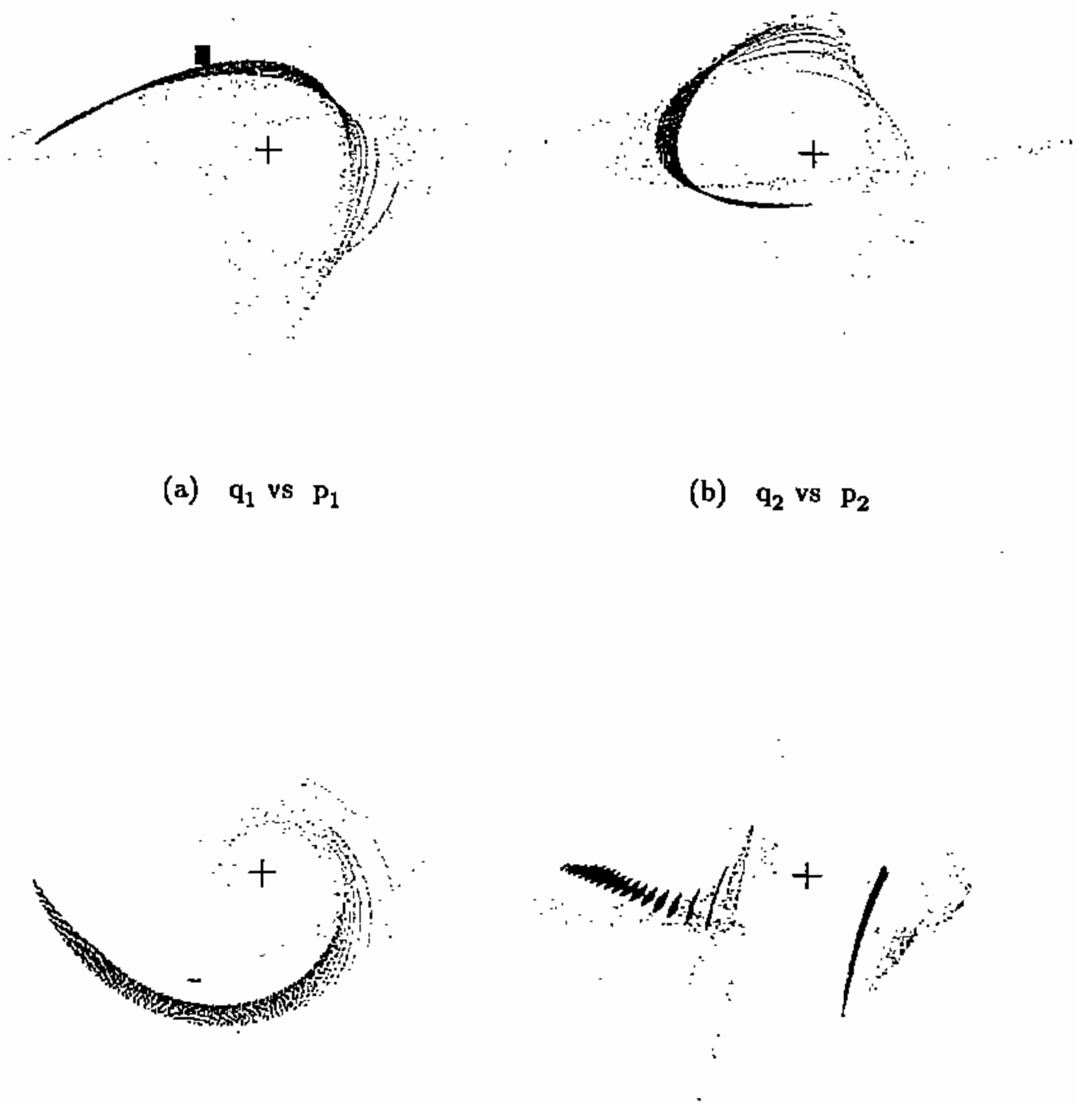
FIGURE 4.9: Kepler problem, DTH dynamics, 25 points per orbit, 400 orbits



(a) q_1 vs p_1 (b) q_2 vs p_2

(c) q_1 vs q_2 (d) q_1 vs p_1, $q_2 = 0$

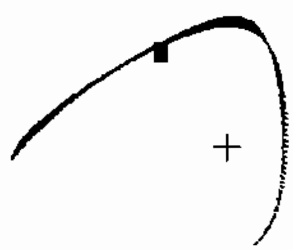
FIGURE 4.10: Kepler problem, trapezoidal method.



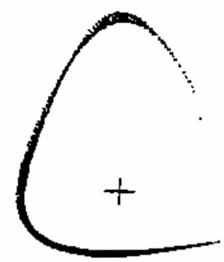
(a) q_1 vs p_1 (b) q_2 vs p_2

(c) q_1 vs q_2 (d) q_1 vs p_1, $q_2 = 0$

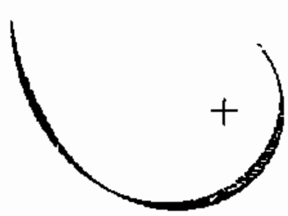
FIGURE 4.11: Kepler problem, midpoint method



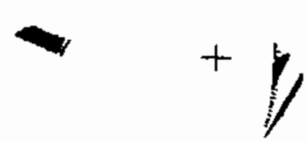
(a) q_1 vs p_1



(b) q_2 vs p_2



(c) q_1 vs q_2



(d) q_1 vs p_1, $q_2 = 0$

FIGURE 4.12: Kepler problem, discrete mechanics

(a) q_1 vs p_1 (b) q_2 vs p_2

(c) q_1 vs q_2 (d) q_1 vs p_1, $q_2 = 0$

FIGURE 4.13: Kepler problem, DTH dynamics.

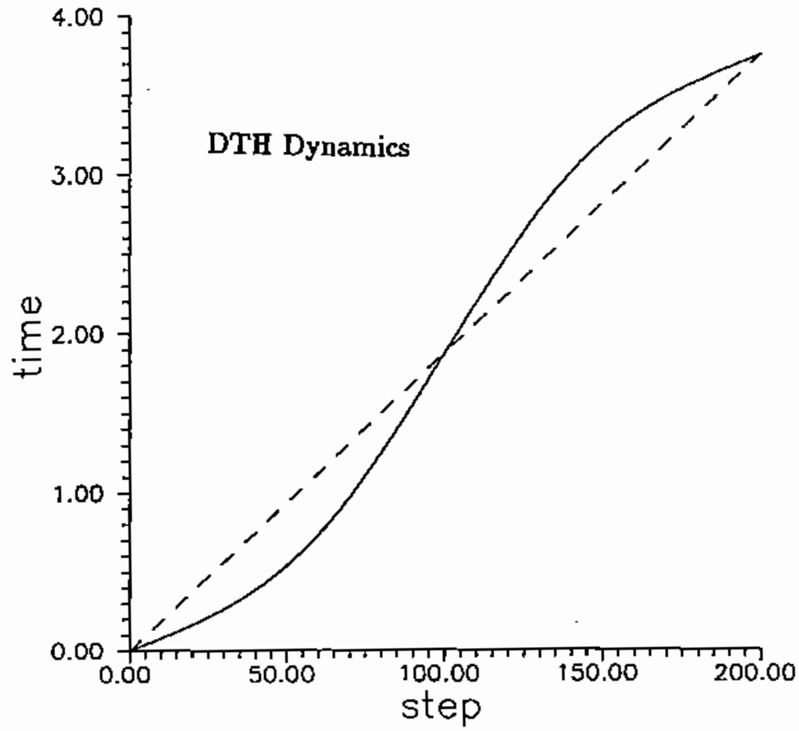


FIGURE 4.14: Kepler problem, 200 points per orbit, 1 orbit.

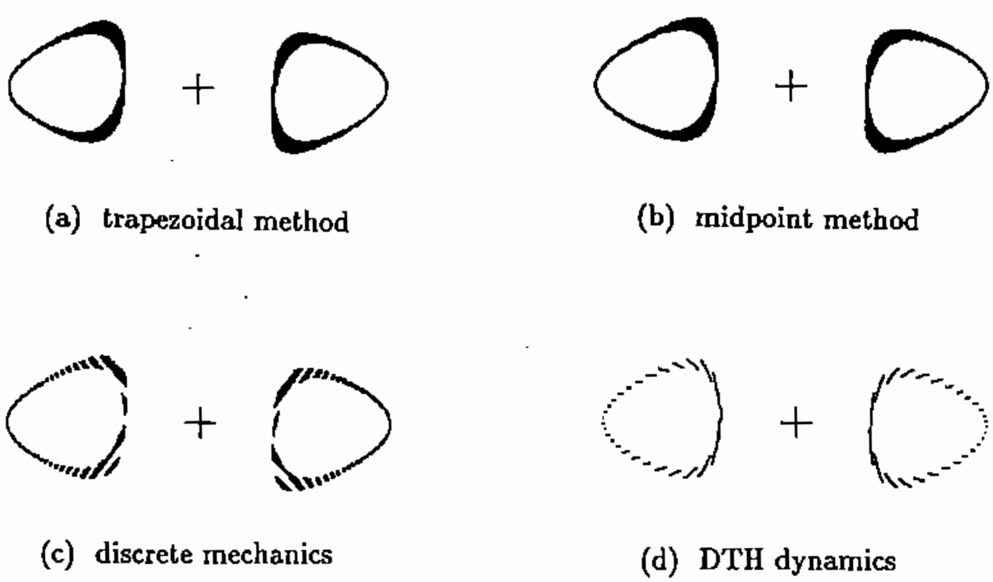


FIGURE 4.15: Kepler problem, Poincare sections, 100 points per orbit, 3,500 orbits.

The complex patterns of the orbits shown in Figures $4.6-4.9$ are due to the precession effects caused by the discretization error of each method. The exact orbit, shown in Figure 4.5, does not precess. The energy conserving properties of discrete mechanics and DTH dynamics is evident from Figure 4.8(c) and 4.9(c) since both annular regions have the same inner and outer radius.

The Poincare sections, Figures $4.5-4.9$ part (d), were computed without interpolating the discrete–time orbits. Instead, the closest point on a given side of the $q_2 = 0$ plane was projected onto this plane resulting in the appearance of short line segments in each figure. These segments are particularly prominent in Figure 4.9(d).

Figures $4.10-4.13$ were obtained by integrating, for 25 time steps, a square set of initial conditions, visible in Figures $4.10-4.13$ part (a), and then replotting these points. Discrete mechanics and DTH dynamics exhibit more regularity than the other methods. DTH dynamics has a markedly different plot because, unlike the other methods, the ending time varies for each set of initial conditions plotted.

Figure 4.14 illustrates that even for a small time step ($\Delta\tau = 0.007846$) the time coordinate of DTH dynamics exhibits nonlinear behavior for the Kepler problem. This suggests that there may exist a limiting trajectory to which the time trajectory converges as $\Delta\tau \to 0$. From Figure 4.15 we see that even for small time steps, DTH dynamics has a significantly slower precession rate than the other methods.

CHAPTER V

COODINATE INVARIANCE

In this chapter, we present arguments which suggest that DTH dynamics is coordinate invariant with respect to a set of symplectic, piecewise–linear, continuous coordinate transformations. In particular, we will establish, under certain assumptions, the equivariance of the DTH equations of dynamics.

Assume (φ, W) are local coordinates of the extended phase space of a Hamiltonian system and assume $\mathbf{H}(z)$ is the Hamiltonian function of the system expressed in terms of (φ, W). Let $\{v^{(1)}, \cdots v^{(N_1)}\}$ represent the vertices of a collection of DTH trajectories which are determined by $\mathbf{H}(z)$ and which lie in $U = \varphi(W) \subset \mathbf{R}^{2n+2}$. Assume there exists a triangulation, $\mathfrak{T}$, of these vertices such that $\mathfrak{T}$ is consistent with the DTH trajectories. (By consistent, we mean that each linear segment of each DTH trajectory is an edge of some simplex, $s^{(j)}$ of $\mathfrak{T}$ where $j = 0, 1, \cdots N_2$.)

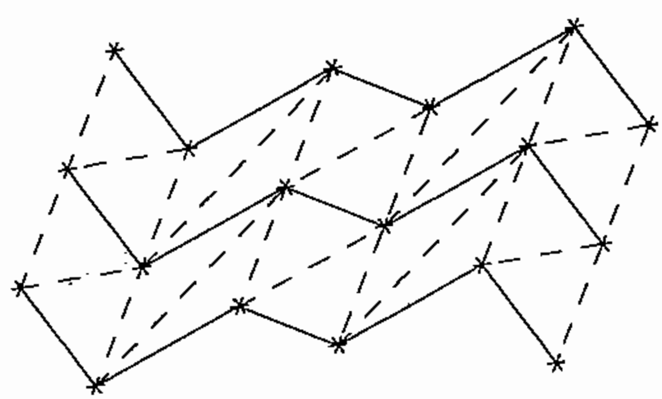


FIGURE 5.1: A triangulation that is consistent with a collection of piecewise–linear, continuous trajectories.

φ^{-1} determines a triangulation of W where $\nu^{(i)} = \varphi^{-1}(v^{(i)})$ are the vertices and $\sigma^{(j)} = \varphi^{-1}(s^{(j)})$ are the simplices of the triangulation. We denote this triangulation by $\mathcal{T}_\varphi$ since it depends on the initial choice of coordinates used to express **H**(z). Assume a second set of vertices $\{V^{(1)}, \cdots V^{(N_1)}\}$ determine an invertible, piecewise–linear, continuous coordinate transformation $T: U \to \mathbb{R}^{2n+2}$ in the following manner. First, define $T(v^{(i)}) = V^{(i)}$, $i = 0, 1, \cdots N_1$. Then, use the values of T at the vertices of the j^{th} simplex of $\mathcal{T}$ to determine the linear transformation:

$$T^{(j)}(z) = A^{(j)}z + b^{(j)} \qquad j = 0, 1, \cdots N_2 \tag{5.1.1}$$

Finally, if $z \in s^{(j)}$, define $T(z) = T^{(j)}(z)$.

We can use the set of all invertible, piecewise–linear, continuous transformations, T, defined in the above manner, to determine a collection of PLC (piecewise–linear, continuous) local coordinates $(\hat{\psi}, W)$ where $\hat{\psi} = T \circ \varphi$. We denote by $\mathbf{PLC}(\mathcal{T}_\varphi)$ the collection of all the coordinate transformations $\hat{\psi}_2 \circ \hat{\psi}_1^{-1}$ where $\hat{\psi}_1$ and $\hat{\psi}_2$ are PLC coordinates defined on W.

Assume $(\hat{\psi}_1, W)$ and $(\hat{\psi}_2, W)$ are PLC coordinates defined on W. Assume $\hat{z}(\cdot) : [\tau_o, \tau_N] \to \hat{\psi}_1(W)$ is a piecewise–linear, continuous trajectory that is consistent with $\mathcal{T}_\varphi$. Then there exists a $T \in \mathbf{PLC}(\mathcal{T}_\varphi)$ such that $\hat{Z}(\cdot) = T\big(\hat{z}(\cdot)\big)$ is a piecewise–linear, continuous trajectory in $\hat{\psi}_2(W)$ where:

$$Z^{(k)} = T(z^{(k)}) \qquad z^{(k)} = T^{-1}(Z^{(k)}) \tag{5.1.2}$$

$$\overline{Z}^{(k)} = T(\overline{z}^{(k)}) \qquad \overline{z}^{(k)} = T^{-1}(\overline{Z}^{(k)}) \tag{5.1.3}$$

$$\overline{Z}'^{(k)} = A^{(j_k)}\overline{z}'^{(k)} \qquad \overline{z}'^{(k)} = \left(A^{(j_k)}\right)^{-1}\overline{Z}'^{(k)} \tag{5.1.4}$$

Equations (5.1.2) – (5.1.4) follow from the piecewise–linearity of T and the consistency of $\hat{z}(\cdot)$ with the triangulation $\mathcal{T}_{\varphi}$.

Consider the subset of $\mathbf{PLC}(\mathcal{T}_{\varphi})$ consisting of all $T \in \mathbf{PLC}(\mathcal{T}_{\varphi})$ that are symplectic — that is, the subset of all transformations having matrices $A^{(j)}$ in (5.1.1) which satisfy the symplectic condition:

$$\left(A^{(j)}\right) J \left(A^{(j)}\right)^{T} = J \tag{5.1.5}$$

We denote this subset by $\mathbf{SPLC}(\mathcal{T}_{\varphi})$.

Consider two different symplectic PLC coordinates, $\hat{\psi}_1$ and $\hat{\psi}_2$. Assume $\hat{\psi}_1$ and $\hat{\psi}_2$ are related by the coordinate transformation $Z = T(z)$ where z are coordinates of $\hat{\psi}_1$, Z are coordinates of $\hat{\psi}_2$ and $T \in \mathbf{SPLC}(\mathcal{T}_{\varphi})$. If $\mathbf{H}(z)$ is a Hamiltonian function for a given Hamiltonian system in the coordinates of $\hat{\psi}_1$, then $\mathbf{K}(Z)$ given by:

$$\mathbf{K}(Z) = (\mathbf{H} \circ T^{-1})(Z) \tag{5.1.6}$$

is a Hamiltonian function for the same system expressed in the coordinates of $\hat{\psi}_2$. DTH dynamics is coordinate invariant with respect $\hat{\psi}_1$ and $\hat{\psi}_2$ if, whenever $\hat{z}(\cdot)$ is a DTH trajectory of $\mathbf{H}(z)$, $\hat{Z}(\cdot) = T\left(\hat{z}(\cdot)\right)$, is the corresponding DTH trajectory of $\mathbf{K}(Z)$. In other words, we have coordinate invariance if the following diagram commutes for all symplectic PLC coordinates $\hat{\psi}_1$ and $\hat{\psi}_2$.

$$\begin{array}{ccc} \underline{\hat{\psi}_1} & & \underline{\hat{\psi}_2} \\ \mathbf{H}(z) & \longrightarrow & \mathbf{K}(Z) \\ \downarrow & & \downarrow \ (\,?\,) \\ \hat{z}(\cdot) & \longrightarrow & \hat{Z}(\cdot) \end{array}$$

For a given set of initial conditions, there may exist more than one DTH trajectory. (An example that occurs in discrete mechanics is given in [5].) In Chapter III, we showed that, for sufficiently small $\Delta\tau$, Newton's method converges to a DTH trajectory that is locally unique. We define this DTH trajectory to be the "principle" DTH trajectory. If $\hat{z}(\cdot)$ is a DTH trajectory of $\mathbf{H}(z)$, then $\overline{z}^{(k)}$ and $\overline{z}'^{(k)}$ must satisfy the DTH equations:

$$\frac{\overline{z}^{(k+1)} - \overline{z}^{(k)}}{\Delta\tau} = \frac{1}{2} J \left[\lambda_{k+1} \frac{\partial \mathbf{H}(\overline{z}^{(k+1)})}{\partial \overline{z}^{(k+1)}} + \lambda_k \frac{\partial \mathbf{H}(\overline{z}^{(k)})}{\partial \overline{z}^{(k)}} \right] \quad k = 0, 1, \cdots N-2 \tag{5.1.7}$$

$$\overline{z}'^{(k)} = \lambda_k J \frac{\partial \mathbf{H}(\overline{z}^{(k)})}{\partial \overline{z}^{(k)}} \quad k = 0, 1, \cdots N-1 \tag{5.1.8}$$

$$\mathbf{H}(\overline{z}^{(k)}) = 0 \quad k = 0, 1, \cdots N-1 \tag{5.1.9}$$

If $\hat{y}(\cdot)$ is a DTH trajectory of $\mathbf{K}(z)$, then $\overline{y}^{(k)}$ and $\overline{y}'^{(k)}$ must satisfy the DTH equations:

$$\frac{\overline{y}^{(k+1)} - \overline{y}^{(k)}}{\Delta\tau} = \frac{1}{2} J \left[\mu_{k+1} \frac{\partial \mathbf{K}(\overline{y}^{(k+1)})}{\partial \overline{y}^{(k+1)}} + \mu_k \frac{\partial \mathbf{K}(\overline{y}^{(k)})}{\partial \overline{y}^{(k)}} \right] \quad k = 0, 1, \cdots N-2 \tag{5.1.10}$$

$$\overline{y}'^{(k)} = \mu_k J \frac{\partial \mathbf{K}(\overline{y}^{(k)})}{\partial \overline{y}^{(k)}} \quad k = 0, 1, \cdots N-1 \tag{5.1.11}$$

$$\mathbf{K}(\overline{y}^{(k)}) = 0 \quad k = 0, 1, \cdots N-1 \tag{5.1.12}$$

Assume $\hat{z}(\cdot)$ is the principle DTH trajectory of $\mathbf{H}(z)$ for the initial conditions λ_o and $\overline{z}^{(0)}$. Assume $\hat{y}(\cdot)$ is the principle DTH trajectory of $\mathbf{K}(z)$ for the initial conditions $\mu_o = \lambda_o$ and $\overline{y}^{(0)} = T(\overline{z}^{(0)})$. To prove equivariance, we need to show that if T maps $\hat{z}(\cdot)$ to $\hat{Z}(\cdot)$, then $\hat{Z}(\cdot) = \hat{y}(\cdot)$, i.e. $\hat{Z}(\cdot)$ is the principle DTH trajectory of $\mathbf{K}(Z)$. What we will actually show is that $\hat{Z}(\cdot)$ is at least a DTH trajectory of $\mathbf{K}(z)$. We do this by showing that $\hat{Z}(\cdot)$ is a solution to the DTH equations (5.1.10) – (5.1.12).

Using (5.1.4), (5.1.8), (5.1.6) and (5.1.5), we show that $\overline{Z}^{(k)}$ and $\overline{Z}'^{(k)}$ satisfy equations (5.1.11) as follows:

$$\begin{aligned}
\overline{Z}'^{(k)} &= A^{(j_k)}\,\overline{z}'^{(k)} \\
&= A^{(j_k)}\,\lambda_k J \frac{\partial H}{\partial \overline{z}^{(k)}} \\
&= \lambda_k \left(A^{(j_k)}\right) J \frac{\partial (K \circ T)}{\partial \overline{z}^{(k)}} \\
&= \lambda_k \left(A^{(j_k)}\right) J \left(\frac{\partial T}{\partial \overline{z}^{(k)}}\right)^T \frac{\partial K}{\partial \overline{Z}^{(k)}} \\
&= \lambda_k \left(A^{(j_k)}\right) J \left(A^{(j_k)}\right)^T \frac{\partial K}{\partial \overline{Z}^{(k)}} \\
&= \lambda_k J \frac{\partial K}{\partial \overline{Z}^{(k)}} \qquad (5.1.13)
\end{aligned}$$

Since λ_k and μ_k in equations (5.1.7) – (5.1.9) and (5.1.10) – (5.1.12) are Lagrange multipliers, they are coordinate invariant quantities. Thus, $\lambda_k \equiv \mu_k$. By substituting μ_k for λ_k in (5.1.13) we have:

$$\overline{Z}'^{(k)} = \mu_k J \frac{\partial K(\overline{Z}^{(k)})}{\partial \overline{Z}^{(k)}} \qquad k = 0,\, 1,\, \cdots N\text{-}1 \qquad (5.1.14)$$

Equation (5.1.14) shows that $\hat{Z}(\cdot)$ satisfies equation (5.1.11). Since $\hat{z}(\cdot)$ and T are both continuous, $\hat{Z}(\cdot) = T\left(\hat{z}(\cdot)\right)$ is continuous. By the continuity constraint on $\hat{Z}(\cdot)$ we have:

$$\begin{aligned}
\frac{\overline{Z}^{(k+1)} - \overline{Z}^{(k)}}{\Delta\tau} &= \frac{\overline{Z}'^{(k+1)} + \overline{Z}'^{(k)}}{2} \\
&= \frac{1}{2} J \left[\mu_{k+1} \frac{\partial K}{\partial \overline{Z}^{(k+1)}} + \mu_k \frac{\partial K}{\partial \overline{Z}^{(k)}}\right] \qquad (5.1.15)
\end{aligned}$$

where we have used (5.1.14). Thus $\hat{Z}(\cdot)$ also satisfies equation (5.1.10). Finally, since

$$\begin{aligned} \mathbf{K}(\overline{Z}^{(k)}) &= (\mathbf{H} \circ T^{-1})(\overline{Z}^{(k)}) \\ &= \mathbf{H}(\overline{z}^{(k)}) \\ &= 0 \end{aligned}$$

$\hat{Z}(\cdot)$ satisfies equation (5.1.12).

In this section, we presented arguments which suggest that the transformations $\mathbf{SPLC}(\mathcal{T}_\varphi)$ map DTH trajectories to DTH trajectories. The question of whether $\mathbf{SPLC}(\mathcal{T}_\varphi)$ maps principle DTH trajectories to principle DTH trajectories still remains. A number of other issues also need to be addressed. One such issue is the question of the existence of triangulations which are consistent with DTH trajectories.

REFERENCES

[1] Tom M. Apostol, *Mathematical Analysis*, Addison–Wesley, Reading, MA, 1974.

[2] Wendell Fleming, *Functions of Several Variables*, Springer–Verlag, New York, 1977.

[3] Herbert Goldstein, *Classical Mechanics*, Addison–Wesley, Reading, MA, 1980.

[4] Gene H. Golub and Charles F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 1989.

[5] Donald Greenspan, *Conservative Numerical Methods for* $\ddot{x} = f(x)$, Journal of Computational Physics, vol. 56, no. 1, October (1984), pp. 28 – 41.

[6] __________ , *Discrete Models*, Addison–Wesley, Reading, MA, 1973.

[7] __________ , *Discrete Numerical Methods in Physics and Engineering*, Academic Press, New York, NY, 1974.

[8] Robert A. Labudde, *Discrete Hamiltonian Mechanics*, International Journal of General Systems, vol. 6, (1980), pp. 3 – 12.

[9] Cornelius Lanczos, *The Variational Principles of Mechanics*, Dover Publications, Mineola, NY, 1970.

[10] T. D. Lee, *Difference Equations and Conservation Laws*, Journal of Statistical Physics, vol. 46, nos. 5/6, (1987), pp. 843 – 860.

[11] James M. Ortega, *Numerical Analysis A Second Course*, Academic Press, New York, NY, 1972.

[12] Yuhua Wu, *The Discrete Variational Approach to the Euler–Lagrange Equations*, Computers and Mathematics with Applications, vol. 20, no. 8, (1990), pp. 61 – 75.