School of Aerospace Engineering
Atlanta, Georgia 30332

NAG1-939

GRAN
IN-64-CR
217251
30P

A Weak Hamiltonian Finite Element Method for Optimal Control Problems

Dewey H. Hodges[1] and Robert R. Bless[2]

Georgia Institute of Technology, Atlanta, Georgia

Abstract

A temporal finite element method based on a mixed form of the Hamiltonian weak principle is developed for dynamics and optimal control problems. The mixed form of Hamilton's weak principle contains both displacements and momenta as primary variables that are expanded in terms of nodal values and simple polynomial shape functions. Unlike other forms of Hamilton's principle, however, time derivatives of the momenta and displacements do not appear therein; instead, only the virtual momenta and virtual displacements are differentiated with respect to time. Based on the duality that is observed to exist between the mixed form of Hamilton's weak principle and variational principles governing classical optimal control problems, a temporal finite element formulation of the latter can be developed in a rather straightforward manner. Several well-known problems in dynamics and optimal control are illustrated. The example dynamics problem involves a time-marching problem. As optimal control examples, elementary trajectory optimization problems are treated.

Introduction

This paper examines a finite element approach to addressing optimal control problems. Hamilton's principle has traditionally been used in *analytical* mechanics as a method of obtaining the equations of motion for dynamical systems. Bailey [1] followed by several others (see [2 – 4] for example) obtained direct solutions to dynamics problems using a form of Hamilton's principle known as the law of varying action thus opening the door for its use in *computational* mechanics.

More recently, it has been shown that expression of Hamilton's law as a weak form (commonly referred to as Hamilton's weak principle or HWP) provides a powerful alternative to numerical solution of ordinary differential equations in the time domain [5, 6]. The accuracy of the time-marching procedure derived in [5, 6] is competitive with standard ordinary differential equation solvers. However, in order to derive an unconditionally stable algorithm in [5], reduced/selective integration had to be used. Further computational advantages may be obtained in so-called mixed formulations of HWP in which the

[1] Professor, School of Aerospace Engineering. Associate Fellow, AIAA.

[2] Graduate Research Assistant, School of Aerospace Engineering. Student Member, AIAA.

(NASA-CR-185336) A WEAK HAMILTONIAN FINITE ELEMENT METHOD FOR OPTIMAL CONTROL PROBLEMS (Georgia Inst. of Tech.) 30 p CSCL 12A

N89-24873

Unclas
G3/64 0217251

generalized coordinates and momenta appear as independent unknowns [7]. Therein, an unconditionally stable algorithm emerges for the linear oscillator with exact integration. HWP also has shown to be an ideal tool for obtaining periodic solutions for autonomous systems, as well as finding the corresponding transition matrix for perturbations about the periodic solution [8]. These are complex two-point boundary value problems; its utility for these problems and its superior performance in mixed form strongly suggest that it could be used in optimal control problems.

In this paper we show that HWP in mixed form provides a useful parallel to optimal control problems. Although finite elements have been applied to optimal control problems [9], the present formulation is believed to offer advantages over existing formulations. For example, it allows simple choices for shape functions in the finite element formulation [7] which simplifies element quadrature.

We begin by summarizing HWP in both displacement and mixed forms. The mixed form is then applied to initial value problems in dynamics. Then, making use of the well-known analogy between dynamics and optimal control, we develop our Hamiltonian weak form for optimal control problems. Finally, we apply it to two relatively simple optimal control problems to illustrate its power.

Hamilton's Weak Principle for Dynamics

In [1 – 8], one can see the potential of obtaining a direct solution in the time domain very much analogous to obtaining the solution of a beam deflection problem with the beam axial coordinate broken into several segments or finite elements. In the present case, however, it is the time interval which is broken into segments; thus, the term "finite elements in time" has been adopted by several investigators.

Only recently has a mixed formulation of HWP been investigated as a computational tool for finite elements in time [7]. In this section we will formulate the mixed form of HWP and illustrate its application to dynamics problems.

General Development

To this aim, let us consider an arbitrary holonomic mechanical system. The configuration is completely defined by a set of generalized coordinates q. Further, let us denote with $L(q, \dot{q}, t)$ the Lagrangean of the system, Q the set of nonconservative generalized forces applied to the system, and $p = \partial L/\partial \dot{q}$ the set of generalized momenta. Then the following variational equation, known as HWP [5], describes the real motion of the system between the two *known* times t_0 and t_f:

$$\int_{t_0}^{t_f} \delta L \, dt + \int_{t_0}^{t_f} \delta q^T Q \, dt = \delta q^T p \Big|_{t_0}^{t_f} \tag{1}$$

This particular variational equation is said to be in displacement form because it only involves the variation of q. Although this formulation has been shown to be of practical use in

dynamics [5, 6], an even more useful formulation may be derived if independent variations in both displacements and momenta are allowed, resulting in a mixed formulation.

To derive the mixed formulation, we begin by defining the Hamiltonian as

$$H(q,p,t) \equiv p^T\dot{q} - L(q,\dot{q},t) \tag{2}$$

Taking the variation of Eq. (2) and substituting for δL in Eq. (1) results in

$$\int_{t_0}^{t_f} \left(\delta p^T\dot{q} + \delta\dot{q}^T p - \delta H + \delta q^T Q\right)\, dt = \delta q^T p\Big|_{t_0}^{t_f} \tag{3}$$

Integrating the first term of the integrand by parts yields

$$\int_{t_0}^{t_f} \left(\delta\dot{q}^T p - \delta\dot{p}^T q - \delta H + \delta q^T Q\right) dt = \left(\delta q^T p - \delta p^T q\right)\Big|_{t_0}^{t_f} \tag{4}$$

This is called a mixed formulation because it contains independent variations of q and p. It is also a *weak* form in the sense that all boundary conditions are of the natural or weak type.

There are two main advantages of the mixed formulation over the displacement formulation. The first advantage is that the mixed formulation generally provides a more accurate solution for a given level of computational effort than does the displacement formulation. The second advantage is that a simpler choice of shape functions is allowed. Note in Eq. (4) that time derivatives of δq and δp are present. However, no time derivatives of q and p exist. Therefore, it is possible to implement linear shape functions for δq and δp and constant shape functions for q and p within each element.

Let us break the time interval from t_0 to t_f into N equally spaced elements. The nodal values of these elements are t_i for $i = 1, \ldots, N+1$ where $t_0 = t_1$ and $t_f = t_{N+1}$. We define a nondimensional elemental time τ as

$$\tau = \frac{t - t_i}{t_{i+1} - t_i} = \frac{t - t_i}{\Delta t} \tag{5}$$

The linear shape functions for the virtual coordinates and momenta are

$$\begin{aligned} \delta q &= \delta q_i(1-\tau) + \delta q_{i+1}\tau \\ \delta p &= \delta p_i(1-\tau) + \delta p_{i+1}\tau \end{aligned} \tag{6}$$

For the generalized coordinates and momenta

$$q = \begin{cases} \hat{q}_i & \text{if } \tau = 0; \\ \bar{q}_i & \text{if } 0 < \tau < 1; \\ \hat{q}_{i+1} & \text{if } \tau = 1 \end{cases} \tag{7}$$

and

$$p = \begin{cases} \hat{p}_i & \text{if } \tau = 0; \\ \bar{p}_i & \text{if } 0 < \tau < 1; \\ \hat{p}_{i+1} & \text{if } \tau = 1 \end{cases} \tag{8}$$

When these shape functions are substituted into Eq. (4), one can either generate an implicit time-marching procedure for nonlinear problems or apply standard finite element assembly procedures to solve periodic or two-point boundary value problems [8]. When this formulation is applied to the linear oscillator, a time-marching algorithm emerges that is unconditionally stable. Higher-order (so-called p-version) elements could be developed, and they would certainly be attractive for linear problems or for nonlinear problems with nonlinearities of low order. For nonlinear problems in general, use of the crude shape functions allowable with the mixed method would seem to be more efficient than use of higher-order shape functions in a p-version. The reason for this is that, with the exception of the term involving Q, all element quadrature can be done by inspection regardless of the order of the nonlinearities. Detailed comparison of these methods is beyond the scope of the present paper but is being undertaken by the first author at this time.

Example 1: Nonlinear Initial-Value Problems

Applying the shape functions of Eqs. (6) – (8) to Eq. (4) for an initial value problem, we obtain a recursive set of nonlinear algebraic equations of the form

$$f_j\left(\bar{q}_i, \hat{q}_{i+1}, \bar{p}_i, \hat{p}_{i+1}\right) = 0 \qquad j = 1, 2, \ldots, n \tag{9}$$

where n is four times the number of degrees of freedom of the system. Eq. (9) can be solved by a Newton-Raphson method yielding an implicit time-marching procedure. The key advantage of using finite elements and a weak variational approach over numerical integration is that the solution (for linear problems) is stable for *all* time steps. In other words, no matter how large a time step is used, a finite approximation of the solution will be obtained. This unconditional stability is obtained without *ad hoc* procedures such as selective or reduced integration which are necessary in displacement formulations.

We also point out that $2\bar{q}_i = \hat{q}_i + \hat{q}_{i+1}$ and $2\bar{p}_i = \hat{p}_i + \hat{p}_{i+1}$. Thus, it is possible to cut the number of equations and unknowns in half. This can be very useful for a multi-degree of freedom problem.

Consider a simple pendulum composed of a lumped mass m and a weightless bar of length ℓ (see Fig. 1). The single generalized coordinate q is the angular displacement of the bar from the vertical. Denoting the kinetic energy of the system with K and the potential energy with V, then we may define the following:

$$\begin{aligned} K &= \frac{1}{2} m\ell^2 \dot{q}^2 \\ V &= mg\ell(1 - \cos q) \\ L &= K - V \\ p &= \partial L / \partial \dot{q} = m\ell^2 \dot{q} \\ H &= p\dot{q} - L \end{aligned} \tag{10}$$

There are no nonconservative forces Q applied to this system.

Substituting $t = t_i + \tau \Delta t$, along with Eq. (10), and substituting the shape functions defined in Eqs. (6) - (8) into Eq. (4) we obtain

$$\begin{aligned} \Delta t \int_0^1 \Bigg\{ & \left(\frac{\delta q_{i+1} - \delta q_i}{\Delta t} \right) \bar{p}_i - mg\ell \sin \bar{q}_i \left[\delta q_i (1 - \tau) + \delta q_{i+1} \tau \right] \\ & - \left(\frac{\delta p_{i+1} - \delta p_i}{\Delta t} \right) \bar{q}_i - \left(\frac{\bar{p}_i}{m\ell^2} \right) \left[\delta p_i (1 - \tau) + \delta p_{i+1} \tau \right] \Bigg\} d\tau \\ & - \delta q_{i+1} \hat{p}_{i+1} + \delta p_{i+1} \hat{q}_{i+1} + \delta q_i \hat{p}_i - \delta p_i \hat{q}_i = 0 \end{aligned} \tag{11}$$

Carrying out the integration, we obtain the four independent equations

$$\begin{aligned} \hat{p}_i - \bar{p}_i - \frac{mg\ell \Delta t \sin \bar{q}_i}{2} &= 0 \\ \bar{p}_i - \hat{p}_{i+1} - \frac{mg\ell \Delta t \sin \bar{q}_i}{2} &= 0 \\ \bar{q}_i - \hat{q}_i - \frac{\bar{p}_i \Delta t}{2m\ell^2} &= 0 \\ \hat{q}_{i+1} - \bar{q}_i - \frac{\bar{p}_i \Delta t}{2m\ell^2} &= 0 \end{aligned} \tag{12}$$

There are six unknowns; however, for an initial-value problem, we will specify $\hat{q}_i$ and $\hat{p}_i$ and solve for the remaining unknowns. Thus, Eq. (12) is of the form of Eq. (9).

For this simple pendulum example, we will nondimensionalize the variables as follows. If we define $\omega^2 = g/\ell$, then a dimensionless time step $\Delta \bar{t}$ may be defined as $\Delta \bar{t} = \omega \Delta t$. Also, instead of solving directly for p, we will solve for the dimensionless $p/m\ell^2\omega$.

We will start our pendulum at $\hat{q}_1 = 60°$ and $\hat{p}_1 = 0.0$. The equations will be solved for $\Delta \bar{t} = 0.4$, 0.8, and 1.6. Graphs of the solutions are shown in Fig. 2 and Fig. 3 and

$\Delta\bar{t} = 0.4$ gives acceptable results for both displacement and angular velocity. Also, note that even the large 1.6 time step yields a finite approximation of the exact solution.

Weak Principle for Optimal Control

A definite analogy exists between the mixed formulation of HWP in dynamics and the first variation of the performance index in optimal control theory. Specifically, there is a similarity between the generalized coordinates and generalized momenta in dynamics and the states and costates in optimal control theory.

Since the mixed formulation has proven to be so valuable in dynamics, we will formulate a weak variational formulation of the performance index. When deriving this formulation, we will keep two things in mind. First, the resulting formulation *must* satisfy the Euler-Lagrange equations and boundary conditions that have already been established in optimal control theory [10]. Second, whatever terms are necessary will be added to the formulation to transfer all strong boundary conditions to natural or weak boundary conditions.

General Development

We start with a performance index taken from Eq. (2.8.4) of [10]. The first variation of the performance index will be taken in a standard manner, except that states, costates, and controls will have arbitrary variations. Rather than setting the first variation equal to zero, however, it will be set equal to an expression which contains the terms that are necessary to transform all boundary conditions to the natural or "weak" type. The final weak form is then obtained by integration of this equation by parts in such a way that no time derivatives of states or costates appear.

Consider a system defined by a set of n states x and a set of m controls u. Furthermore, let the system be governed by a set of state equations of the form $\dot{x} = f(x, u, t)$. We may denote elements of the performance index, J, with an integrand $L(x, u, t)$ and a discrete function of the states and time at the final time $\phi[x(t_f), t_f]$. In addition, any terminal constraints placed on the states may be placed in the set of q functions $\psi[x(t_f), t_f]$ and adjoined to the performance index by a set of q discrete Lagrange multipliers ν. Finally, we may adjoin the state equations to the performance index with a set of Lagrange multiplier functions $\lambda(t)$ which will be referred to as costates. This yields a performance index of the form

$$J = \int_{t_0}^{t_f} (\lambda^T \dot{x} - L - \lambda^T f)\, dt - \phi\Big|_{t_f} - \nu^T \psi\Big|_{t_f} \tag{13}$$

Taking the first variation of J and setting it equal to an expression chosen so that all boundary conditions are of the weak type, one obtains

$$
\begin{aligned}
\delta J = & \int_{t_0}^{t_f} \left[\delta\lambda^T(\dot{x}-f) + \delta\dot{x}^T\lambda - \delta L - \delta f^T\lambda\right] dt \\
& + \delta t_f \left(\lambda^T\dot{x} - L - \lambda^T f - \frac{\partial\phi}{\partial t} - \nu^T\frac{\partial\psi}{\partial t}\right)\Big|_{t_f} \\
& - \delta x_f^T\hat{\lambda}_f - \delta\nu^T\psi\big|_{t_f} \\
& = \delta t_f \left(\lambda^T\dot{x}\right)\big|_{t_f} - \delta x_0^T\hat{\lambda}_0 - \delta\lambda^T(\hat{x}-x)\Big|_{t_0}^{t_f}
\end{aligned}
\tag{14}
$$

where

$$
\hat{\lambda}_f \equiv \left[\left(\frac{\partial\phi}{\partial x}\right)^T + \left(\frac{\partial\psi}{\partial x}\right)^T \nu\right]\Big|_{t_f} \tag{15}
$$

The right hand side of Eq. (14) contains terms necessary to form all of the proper boundary conditions as natural ones. The quantities $\hat{x}$ and $\hat{\lambda}$ are discrete values of the states and costates at the initial (with subscript 0) and final times (with subscript f). Depending on the problem, these values will either be specified or left as unknowns; they need not coincide with the values of the states or costates taken on within the elements at the beginning and at the end of the time interval. This is yet another indication of the "weakness" of the formulation.

From Eq. (14), we may directly write down a weak formulation. However, as stated earlier, one of the requirements of the formulation is that it satisfies the Euler-Lagrange equations and boundary conditions. To show this, we will integrate the $\delta\dot{x}^T$ term in Eq. (14) by parts, expand the variation of L, substitute Eq. (15), and group terms yielding

$$
\begin{aligned}
& \int_{t_0}^{t_f} \left\{\delta\lambda^T(\dot{x}-f) - \delta u^T\left[\left(\frac{\partial L}{\partial u}\right)^T + \left(\frac{\partial f}{\partial u}\right)^T\lambda\right]\right. \\
& \qquad \left. - \delta x^T\left[\left(\frac{\partial L}{\partial x}\right)^T + \left(\frac{\partial f}{\partial x}\right)^T\lambda + \dot{\lambda}\right]\right\} dt \\
& \qquad - \delta\nu^T\psi\Big|_{t_f} - \delta t_f\left(L + \lambda^T f + \frac{\partial\phi}{\partial t} + \nu^T\frac{\partial\psi}{\partial t}\right)\Bigg|_{t_f} \\
& \qquad + \delta x_f^T\left(\lambda_f - \hat{\lambda}_f\right) - \delta x_0^T\left(\lambda_0 - \hat{\lambda}_0\right) \\
& \qquad + \delta\lambda_f^T(\hat{x}_f - x_f) - \delta\lambda_0^T(\hat{x}_0 - x_0) = 0
\end{aligned}
\tag{16}
$$

where x_0, λ_0, x_f, and λ_f represent the values of those functions at the initial and final times, respectively. The coefficients of $\delta\lambda^T$, δx^T, and δu^T in the integrand are the three

correct Euler-Lagrange equations, Eqs. 2.8.15 – 2.8.17 from [10]. There are also six trailing terms in Eq. (16) from which the boundary conditions can be determined. The first four of these equations and the sixth equation correspond to the correct boundary conditions in [10]. Namely, the requirement for the coefficient of $\delta\nu^T$ to vanish yields Eq. (2.8.21). The requirement for the coefficient of δt_f to vanish is equivalent to Eq. (2.8.20). The requirement for the coefficient of δx_f^T to vanish shows that the final value of λ approaches $\hat{\lambda}_f$ as given in Eq. (15), which corresponds to Eq. (2.8.19). If $\hat{\lambda}_0$ is chosen as zero, the requirement for the coefficient of δx_0^T to vanish requires the initial value of λ to approach zero as the accuracy of the approximation scheme is increased (such as from adding more finite elements); on the other hand, the requirement for the coefficient of $\delta\lambda_0^T$ to vanish requires the initial value of x to approach $\hat{x}_0$, in accordance with Eq. (2.8.18). Finally, the requirement for the coefficient of $\delta\lambda_f^T$ to vanish demands that the final value of x approach the discrete value $\hat{x}_f$; this has no counterpart in [10] since the elements of $\hat{x}_f$ are usually unknown.

Having satisfied our requirement that none of the fundamental equations are altered, we may now derive our weak formulation from Eq. (14). Recall from the mixed formulation for dynamics that we do not want time derivatives of q or p to appear in the formulation. Similarly, we do not wish for the time derivatives of x and λ to appear in the present weak formulation for optimal control. Therefore, we will integrate the $\dot{x}$ term by parts in Eq. (14). The resulting equation is

$$\begin{aligned}
&\int_{t_0}^{t_f} \left\{ \delta\dot{x}^T\lambda - \delta x^T \left[\left(\frac{\partial L}{\partial x}\right)^T + \left(\frac{\partial f}{\partial x}\right)^T \lambda \right] \right. \\
&\left. - \delta\dot{\lambda}^T x - \delta\lambda^T f - \delta u^T \left[\left(\frac{\partial L}{\partial u}\right)^T + \left(\frac{\partial f}{\partial u}\right)^T \lambda \right] \right\} dt \\
&- \delta t_f \left(L + \lambda^T f + \frac{\partial \phi}{\partial t} + \nu^T \frac{\partial \psi}{\partial t} \right) \Bigg|_{t_f} - \delta\nu^T \psi \Big|_{t_f} \\
&- \delta x_f^T \hat{\lambda}_f + \delta x_0^T \hat{\lambda}_0 + \delta\lambda_f^T \hat{x}_f - \delta\lambda_0^T \hat{x}_0 = 0
\end{aligned} \tag{17}$$

This is the governing equation for the weak Hamiltonian method for optimal control problems of the form specified. It will serve as the basis for the finite element discretization described below. It should be noted that normally one will encounter various types of inequality constraints on the states and controls in problems that deal with optimal control. This aspect has not been treated yet and will be the subject of future research.

As in dynamics, we may choose linear shape functions for δx and $\delta\lambda$. We may choose piecewise constant shape functions for x and λ. Thus, we will be working with shape functions similar to those of Eqs. (6) – (8). In addition, note that the time derivatives of u and δu do not appear in the formulation. Thus, we let

$$
\begin{aligned}
u &= \bar{u}_i \\
\delta u &= \delta \bar{u}_i
\end{aligned}
\tag{18}
$$

Plugging in the shape functions described for x, λ, and u, substituting τ for t, and carrying out the simple integration from 0 to 1, we obtain

$$
\begin{aligned}
\sum_{i=1}^{N} \Bigg\langle & \delta x_i^T \left[-\bar{\lambda}_i - \frac{\Delta t}{2} \left(\frac{\partial \bar{f}}{\partial \bar{x}} \right)_i^T \bar{\lambda}_i - \frac{\Delta t}{2} \left(\frac{\partial \bar{L}}{\partial \bar{x}} \right)_i^T \right] + \delta \lambda_i^T \left(\bar{x}_i - \frac{\Delta t}{2} \bar{f}_i \right) \\
& + \delta x_{i+1}^T \left[\bar{\lambda}_i - \frac{\Delta t}{2} \left(\frac{\partial \bar{f}}{\partial \bar{x}} \right)_i^T \bar{\lambda}_i - \frac{\Delta t}{2} \left(\frac{\partial \bar{L}}{\partial \bar{x}} \right)_i^T \right] - \delta \lambda_{i+1}^T \left(\bar{x}_i + \frac{\Delta t}{2} \bar{f}_i \right) \\
& - \delta \bar{u}_i^T \left\{ \Delta t \left[\left(\frac{\partial \bar{L}}{\partial \bar{u}} \right)_i^T + \left(\frac{\partial \bar{f}}{\partial \bar{u}} \right)_i^T \bar{\lambda}_i \right] \right\} \Bigg\rangle \\
& - \delta t_f \left(\hat{L} + \hat{\lambda}^T \hat{f} + \frac{\partial \phi}{\partial t} + \nu^T \frac{\partial \psi}{\partial t} \right) \Bigg|_{t_f} - \delta \nu^T \psi \big|_{t_f} \\
& + \delta x_1^T \hat{\lambda}_0 - \delta \lambda_1^T \hat{x}_0 - \delta x_{N+1}^T \hat{\lambda}_f + \delta \lambda_{N+1}^T \hat{x}_f = 0
\end{aligned}
\tag{19}
$$

This is the general algebraic form of our Hamiltonian weak form for optimal control problems of the form specified. Eq. (19) is a system of algebraic equations. In fact, for N elements, there are $2n(N+1) + mN + q + 1$ equations and $2n(N+2) + mN + q + 1$ unknowns. Therefore, $2n$ of the $4n$ endpoint values for the states and costates ($\hat{x}_0, \hat{\lambda}_0, \hat{x}_f$, and $\hat{\lambda}_f$) must be specified. In general, $\hat{x}_0$ (the initial conditions) is known in accordance with physical constraints. Also, $\hat{\lambda}_f$ can be specified in terms of other unknowns with the use of Eq. (15). Now we have the same number of equations as unknowns.

Normally Eq. (19) can be solved by writing an explicit Jacobian and using a Newton-Raphson solution procedure. For the example problems which follow, the iteration procedure will converge quickly for a small number of elements with a trivial initial guess. Then, the answers obtained for a small number of elements can be used to generate initial guesses for a higher number of elements. Thus, a large number of elements can be solved with a very efficient run-time on the computer.

Unfortunately, for some highly nonlinear problems, trivial initial guesses may not be adequate. This problem may be overcome by noting that the δx_i and δx_{i+1} equations in Eq. (19) are $n(N+1)$ equations that happen to be linear in the $n(N+1)$ unknown costates. Now, we need $n(N+1)$ fewer initial guesses and we can solve for the costates in terms of the other unknowns. This is particularly useful since generating initial guesses for the costates can be difficult. This has proven to be a very useful way of obtaining answers for highly nonlinear equations.

Example 2: A Fixed-Final-Time Problem

As the first optimal control problem, we will examine the transfer of a particle to a rectilinear path (see Fig. 4). This is an example taken from [10], article 2.4. We will let x_1 and x_2 denote the position of the particle at a given time and x_3 and x_4 denote the particle's velocity at a given time. The thrust angle u is the control and the particle has mass m and a constant acceleration a.

The state equations are defined as

$$\dot{x} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} x + \begin{Bmatrix} 0 \\ 0 \\ a\cos u \\ a\sin u \end{Bmatrix} \tag{20}$$

The final time T is fixed and we would like to maximize the final horizontal component of velocity. Thus, $L = 0$ and $\phi = x_3(T)$. There are also two terminal constraints on the states. These are that the particle arrive with a fixed final height (h) and that the final vertical component of velocity be zero. We do not care what the final horizontal component of position is. These constraints can be stated analytically as

$$\psi = \begin{Bmatrix} x_2 - h \\ x_4 \end{Bmatrix} \tag{21}$$

The initial conditions are $x(0) = \hat{x}_0 = \lfloor 0 \quad 0 \quad 0 \quad 0 \rfloor^T$. Finally, we will eliminate the unknown $\hat{\lambda}_f$'s by writing it in terms of other unknowns. In accordance with Eq. (15)

$$\hat{\lambda}_f = \begin{Bmatrix} 0 \\ \nu_1 \\ 1 \\ \nu_2 \end{Bmatrix} \tag{22}$$

With $L = 0$, and t_f fixed, along with all zero initial conditions on the states, then the general formulation of Eq. (19) takes the following form.

$$\begin{aligned} \sum_{i=1}^{N} \Bigg\{ & \delta x_i^T \left[-\bar{\lambda}_i - \frac{\Delta t}{2} \left(\frac{\partial \bar{f}}{\partial \bar{x}} \right)_i^T \bar{\lambda}_i \right] + \delta \lambda_i^T \left(\bar{x}_i - \frac{\Delta t}{2} \bar{f}_i \right) \\ & + \delta x_{i+1}^T \left[\bar{\lambda}_i - \frac{\Delta t}{2} \left(\frac{\partial \bar{f}}{\partial \bar{x}} \right)_i^T \bar{\lambda}_i \right] - \delta \lambda_{i+1}^T \left(\bar{x}_i + \frac{\Delta t}{2} \bar{f}_i \right) \\ & - \delta u_i^T \left[\Delta t \left(\frac{\partial \bar{f}}{\partial \bar{u}} \right)_i^T \bar{\lambda}_i \right] \Bigg\} - \delta \nu^T \psi \big|_{t_f} \\ & + \delta x_1^T \hat{\lambda}_0 - \delta x_{N+1}^T \hat{\lambda}_f + \delta \lambda_{N+1}^T \hat{x}_f = 0 \end{aligned} \tag{23}$$

This produces a system of nonlinear algebraic equations whose size depends on the number of elements N.

These equations are solved by writing down an explicit Jacobian and using a Newton-Raphson algorithm. Trivial initial guesses (that are never changed regardless of input parameters) are used for $N = 2$. These results are then used to obtain the initial guesses for arbitrary N by simple interpolation. In all results obtained to date for this problem, no additional steps are necessary to obtain results as accurate as desired.

Representative numerical results for all four states versus dimensionless time t/T are presented in Figs. 5 – 8. For this example, we have taken $h = 100.0$, $T = 20.0$, and $4h/aT^2 = 0.8897$. The results for 2, 4, and 8 elements are plotted against the exact solution available in [10]. It can easily be seen that $N = 8$ gives acceptable results for all the states.

In Fig. 9, the control angle u versus dimensionless time t/T is presented. Once again, the results are seen to be excellent for $N = 8$.

Three of the four costates are constants for all time and this method yields two of these exactly. The third costate is very close to the exact answer. The fourth costate corresponding to the vertical component of velocity λ_4 is shown in Fig. 10. The results compare nicely with the exact results.

A plot of the relative error of the performance index $J = x_3(T)$ and the endpoint multiplier ν_1 versus the number of elements is shown in Fig. 11. It is seen to be nearly a straight line on a log-log scale. The slope of the line is about -2 which indicates that the error varies inversely with the square of N.

Notice in Fig. 11 that there is a bend in the endpoint multiplier curve. It is a common characteristic of mixed formulations for an error curve not to be monotonically decreasing.

Example 3: A Free-Final-Time Problem

The second optimal control problem is similar to Example 2 except that now the final time is free and we would like to obtain a given horizontal component of velocity (U) in the minimum time (see problem 9, article 2.7 of [10]). Our algorithm from the preceding example is readily modified to fit this problem by noting the following changes. The performance index is now the final time T; so $\phi = 0$ and $L = 1$. Also, there is an additional endpoint constraint on the states; namely that $x_3 = U$. With these changes, we have

$$\hat{\lambda}_f = \begin{Bmatrix} 0 \\ \nu_1 \\ \nu_2 \\ \nu_3 \end{Bmatrix} \tag{24}$$

and

$$\psi = \begin{Bmatrix} x_2 - h \\ x_3 - U \\ x_4 \end{Bmatrix} \tag{25}$$

Along with these changes to our equations, we also pick up the additional δt_f equation. The new system of equations is solved in the same manner described previously. Again, trivial initial guesses are satisfactory for $N = 2$, and these answer are used to obtain initial guesses for arbitrary N.

Representative numerical results for all four states versus dimensionless time t/T are presented in Figs. 12–15 for a case with $ah/U^2 = 0.75$. The results for 2, 4, and 8 elements are plotted against the exact solution available in [10]. It can easily be seen that $N = 8$ gives acceptable results for all the states.

The control angle u versus dimensionless time t/T is presented in Fig. 16. Once again, the results are seen to be excellent for $N = 8$.

As with the fixed time problem, three of the four costates are constants. The costate results are all as accurate or better then the costate depicted in Fig. 10.

A plot of the relative error of the performance index $J = T$ versus the number of elements N is shown in Fig. 17. As before, it is seen to be nearly a straight line on a log-log scale. The slope of the line is about -2 which indicates that the error varies inversely with the square of N.

It should be noted that the computer time on a Cyber 990 is only about 2 seconds for $N = 2$, $N = 4$, and $N = 8$ and 3 seconds for $N = 16$. Thus, the run time is relatively insensitive to N.

Conclusions and Future Work

In this report, a mixed form of Hamilton's Weak Principle has been stated for dynamics problems. Finite elements in time were applied to this formulation and a simple initial-value problem has been used to demonstrate the principles involved. A temporal finite element method based on a mixed form of the Hamiltonian weak principle was then developed for optimal control problems from the dynamics principles. It has been demonstrated that the mixed form allows for a simple choice of shape functions and is essentially self-starting. Two simple optimal control problems have been examined and the results are seen to compare excellently with the exact solutions for even a very few elements. Overall, the method provides very accurate results for the problems investigated to date with only a few elements and for minimal computational effort.

Future research will be done in applying this method to practical problems such as development of on-board trajectory optimization algorithms for launch vehicles [11]. Such applications will require the reliability, efficiency, and self-starting characteristics illustrated in the present approach. However, more work needs to be done to solve the equations as efficiently as possible. One area with great potential for making the method more

efficient is to take advantage of the sparsity of the Jacobian. The method will also have to be extended to allow for inequality constraints on the states and controls.

Acknowledgements

This work was supported in part by NASA Grant NAG-1-939 of which Dr. Daniel D. Moerder is the technical monitor. Technical discussions with Profs. A. J. Calise and M. Borri are also gratefully acknowledged.

References

1. Bailey, C. D., "Application of Hamilton's Law of Varying Action," *AIAA Journal*, Vol. 13, 1975, pp. 433 – 451.

2. Simkins, T. E., "Unconstrained Variational Statements for Initial and Boundary-Value Problems," *AIAA Journal*, Vol. 16, 1978, pp. 559 – 563.

3. Hitzl, D. L., and Levinson, D. A., "Application of Hamilton's Law of Varying Action to the Restricted Three-Body Problem," *Celestial Mechanics*, Vol. 22, 1980, pp. 255 – 266.

4. Baruch, M., and Riff, R., "Hamilton's Principle, Hamilton's Law, 6" Correct Formulations," *AIAA Journal*, Vol. 20, 1982, pp. 687 – 692.

5. Borri, M. et al, "Dynamic Response of Mechanical Systems by a Weak Hamiltonian Formulation," *Computers and Structures*, Vol. 20, No. 1 – 3, 1985, pp. 495 – 508.

6. Peters, David A., and Izadpanah, Amir, "*hp*-Version Finite Elements for the Space-Time Domain," *Computational Mechanics*, Vol. 3, 1988, pp. 73 – 88.

7. Borri, M., *et al.*, "Primal and Mixed Forms of Hamilton's Principle for Constrained and Flexible Dynamical Systems: Numerical Studies," ARO/AFOSR Workshop on Nonlinear Dynamics, Virginia Polytechnic Institute and State University, Blacksburg. Virginia, June 1 – 3, 1988.

8. Borri, M., "Helicopter Rotor Dynamics by Finite Element Time Approximation." *Computers and Mathematics with Applications*, Vol. 12A, No. 1, 1986, pp. 149 – 160.

9. Patten, William Neff, "Near Optimal Feedback Control for Nonlinear Aerodynamic Systems with an Application to the High-Angle-of-Attack Wing Rock Problem," AIAA Paper 88-4052-CP, 1988.

10. Byson, Arthur E. Jr., and Ho, Yu-Chi, *Applied Optimal Control*, Blaisdell Publishing Company, Waltham, Massachusetts, 1969, Chapter 2.

11. Hardtla, J. W., Piehler, M. J., and Bradt, J. E., "Guidance Requirements for Future Launch Vehicles," *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Paper No. 87-2462, Aug. 1987.

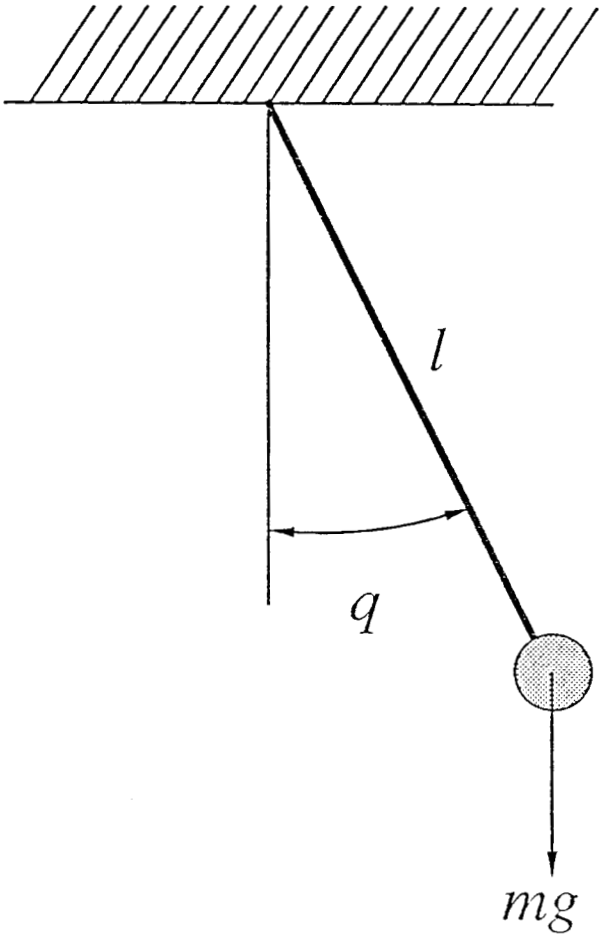

Fig. 1: Nomenclature for Example 1. This is a simple pendulum composed of a lump mass m and a weightless bar of length ℓ.

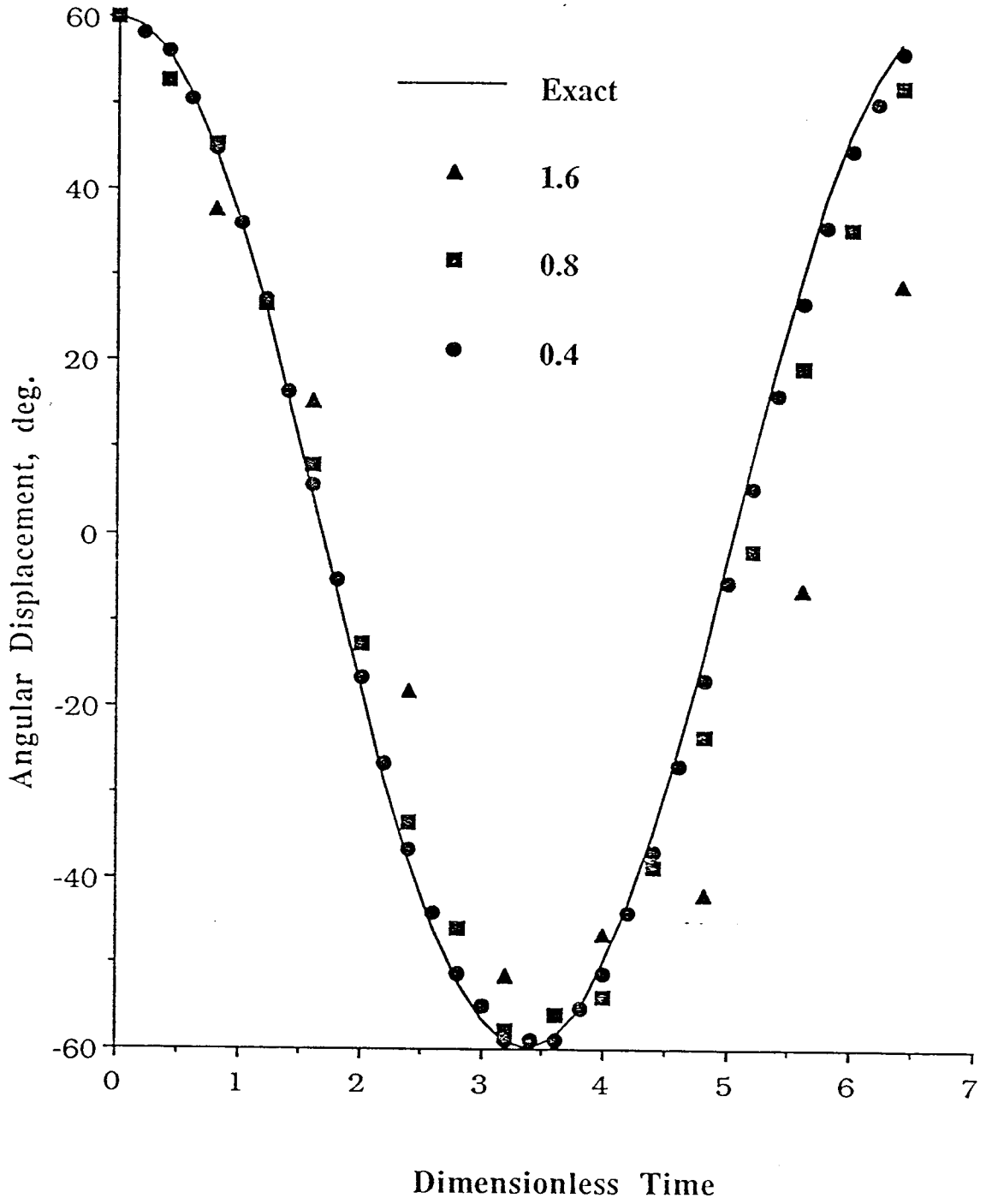

Fig. 2: Angular displacement q versus dimensionless time for three values of the time step $\Delta\bar{t}$. These solutions are being compared to the exact elliptic integral solution. Note that even a large time step yields a stable solution.

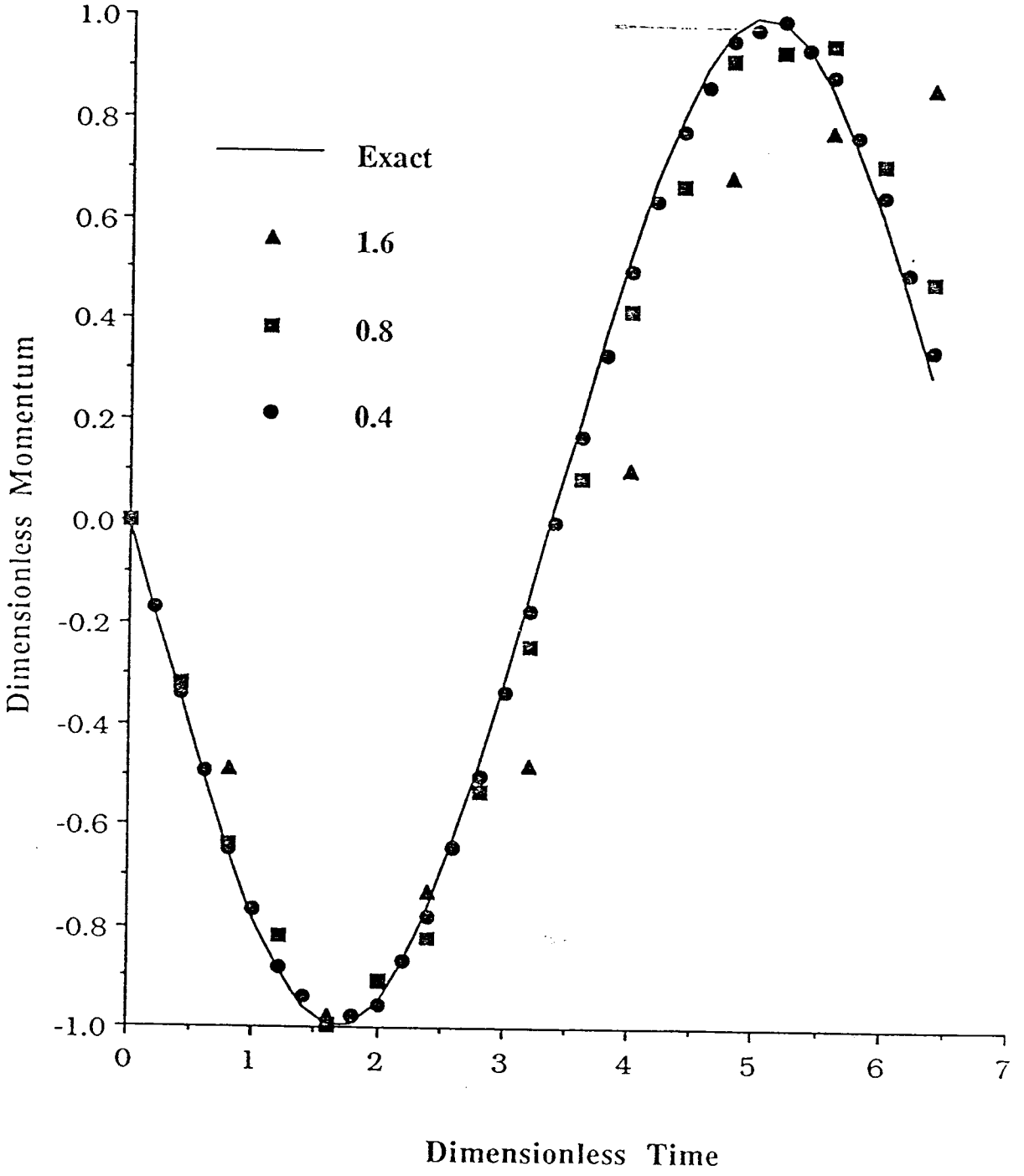

Fig. 3: Dimensionless momentum $p/m\ell^2\omega$ versus dimensionless time for three values of the time step $\Delta\bar{t}$. These solutions are being compared to the exact elliptic integral solution. Note that even a large time step yields a stable solution.

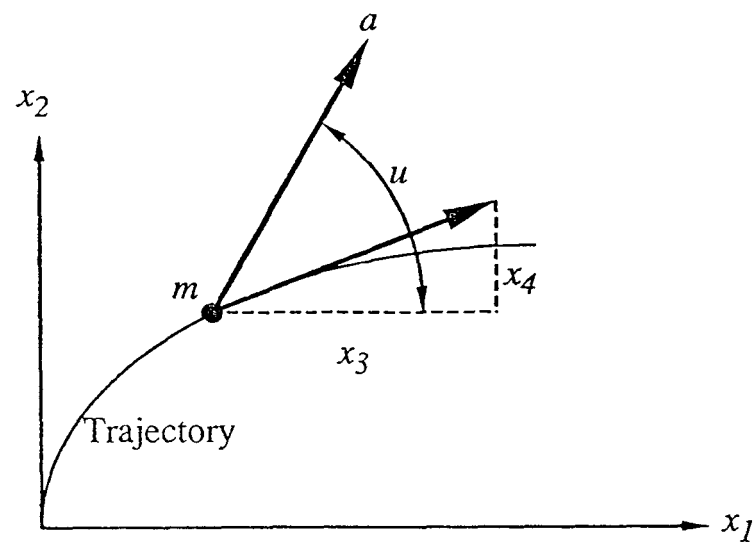

Fig. 4: Nomenclature for Example 2. Transfer of a particle to a rectilinear path with a fixed final time. The problem is to maximize the final horizontal component of velocity subject to specified terminal conditions on the states.

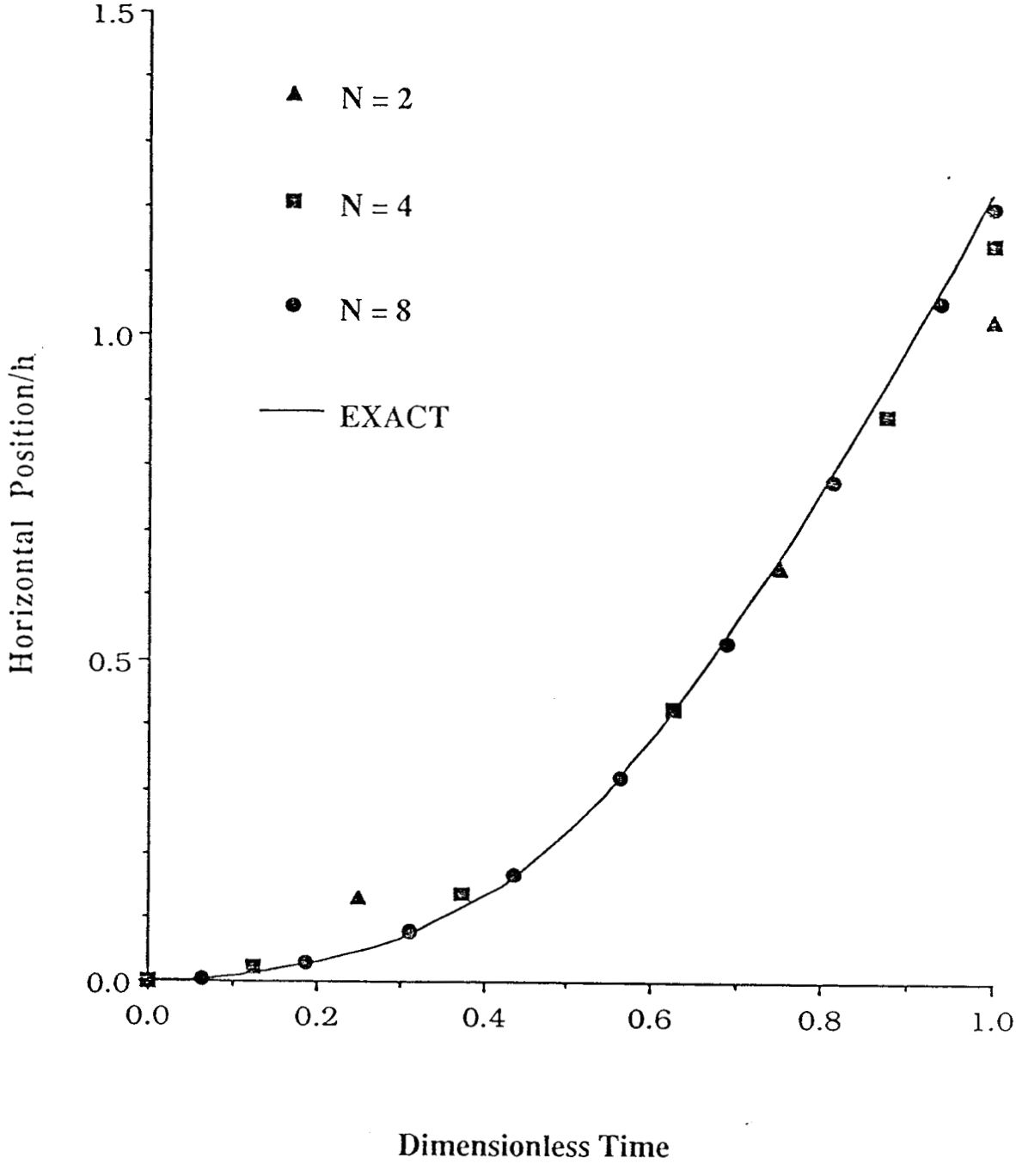

Fig. 5: Dimensionless horizontal position x_1/h versus t/T. Note that the final horizontal component of position is not specified.

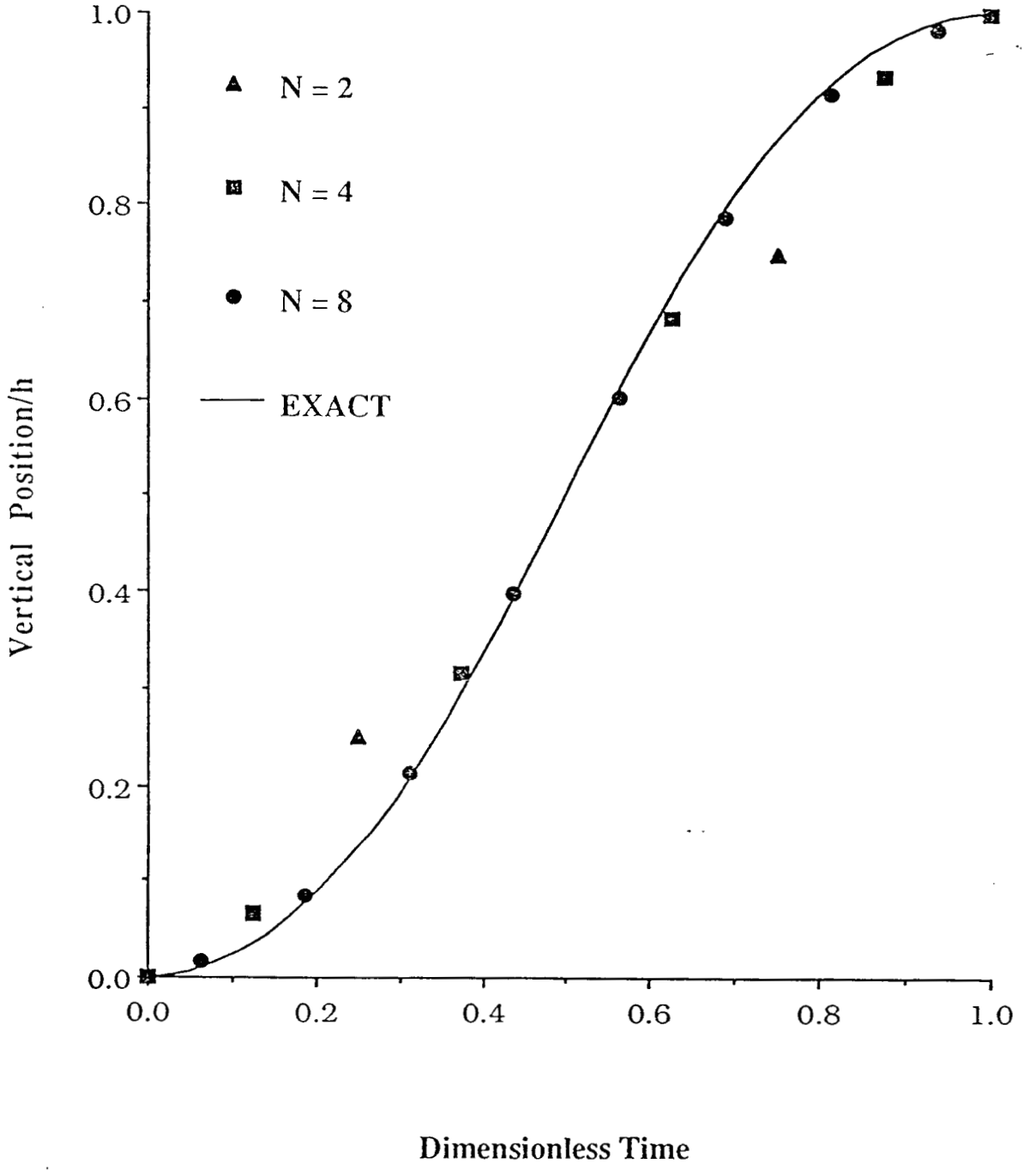

Fig. 6: Dimensionless vertical position x_2/h versus t/T. The final height is constrained to be h at the final time.

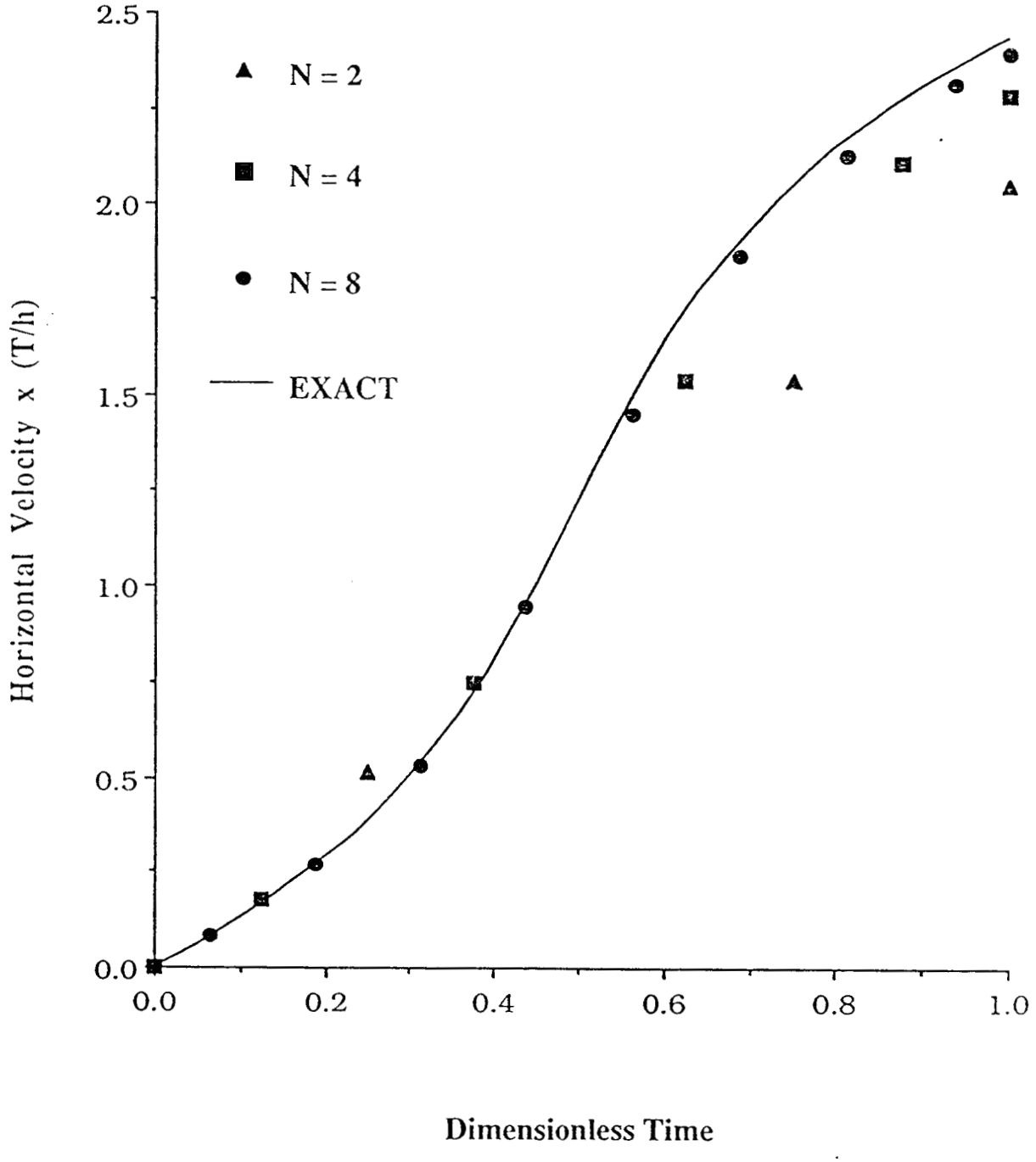

Fig. 7: Dimensionless horizontal velocity $x_3 T/h$ versus t/T. Note that the performance index $J = x_3(T)$.

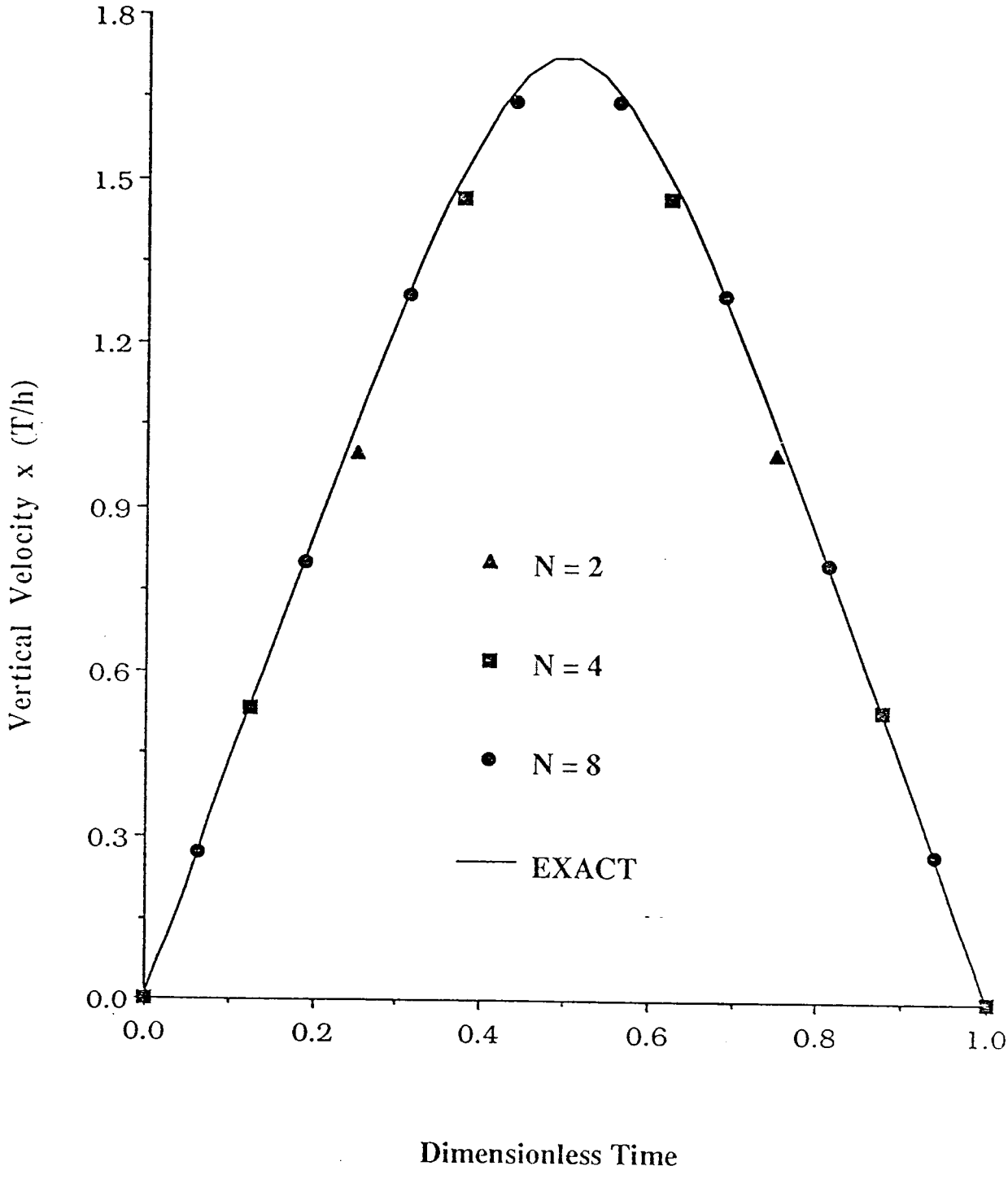

Fig. 8: Dimensionless vertical velocity $x_4 T/h$ versus t/T. The final vertical component of velocity is constrained to be zero at the final time.

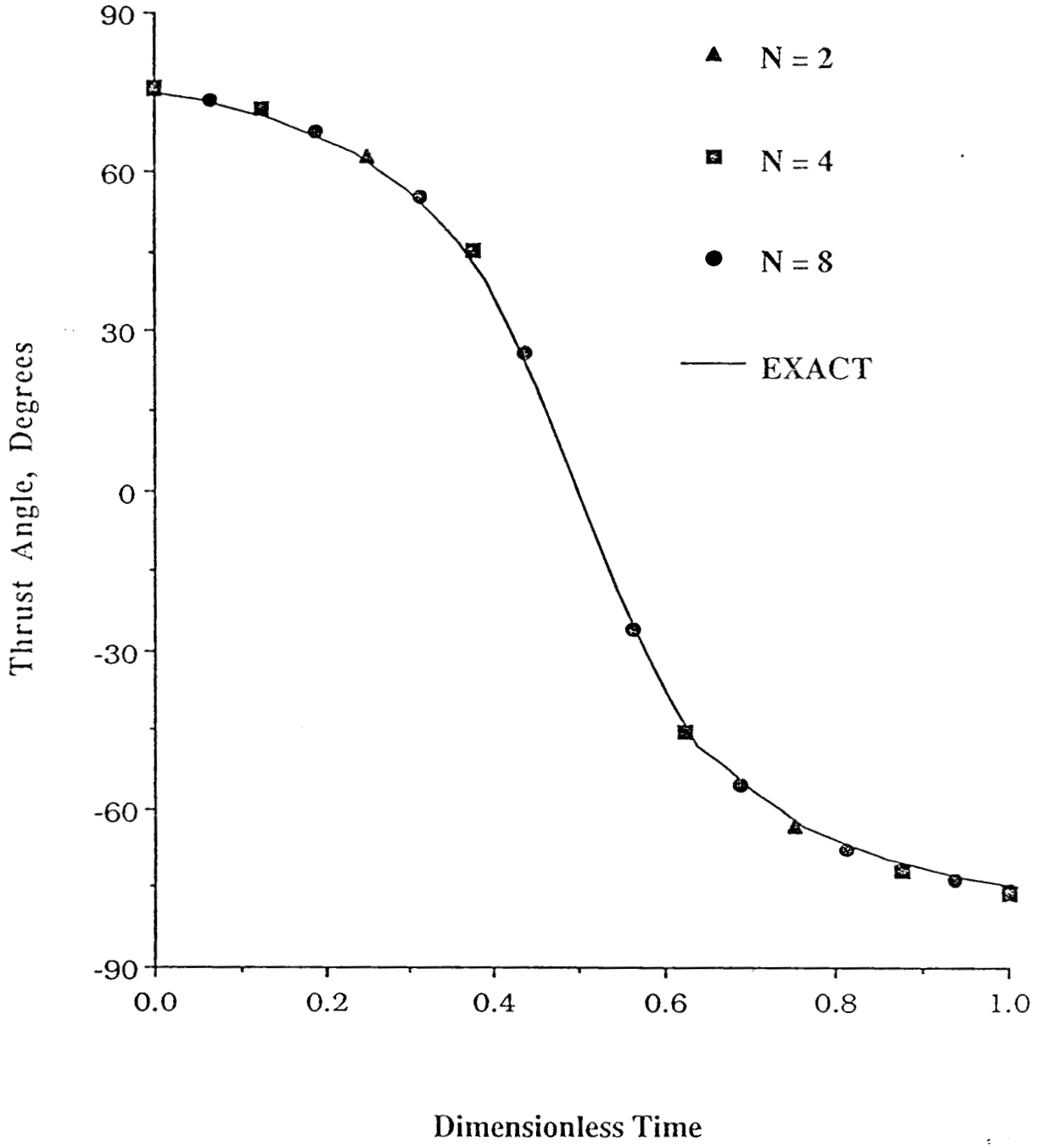

Fig. 9: Control angle u versus t/T. The control is chosen so as to maximize the performance index $J = x_3(T)$.

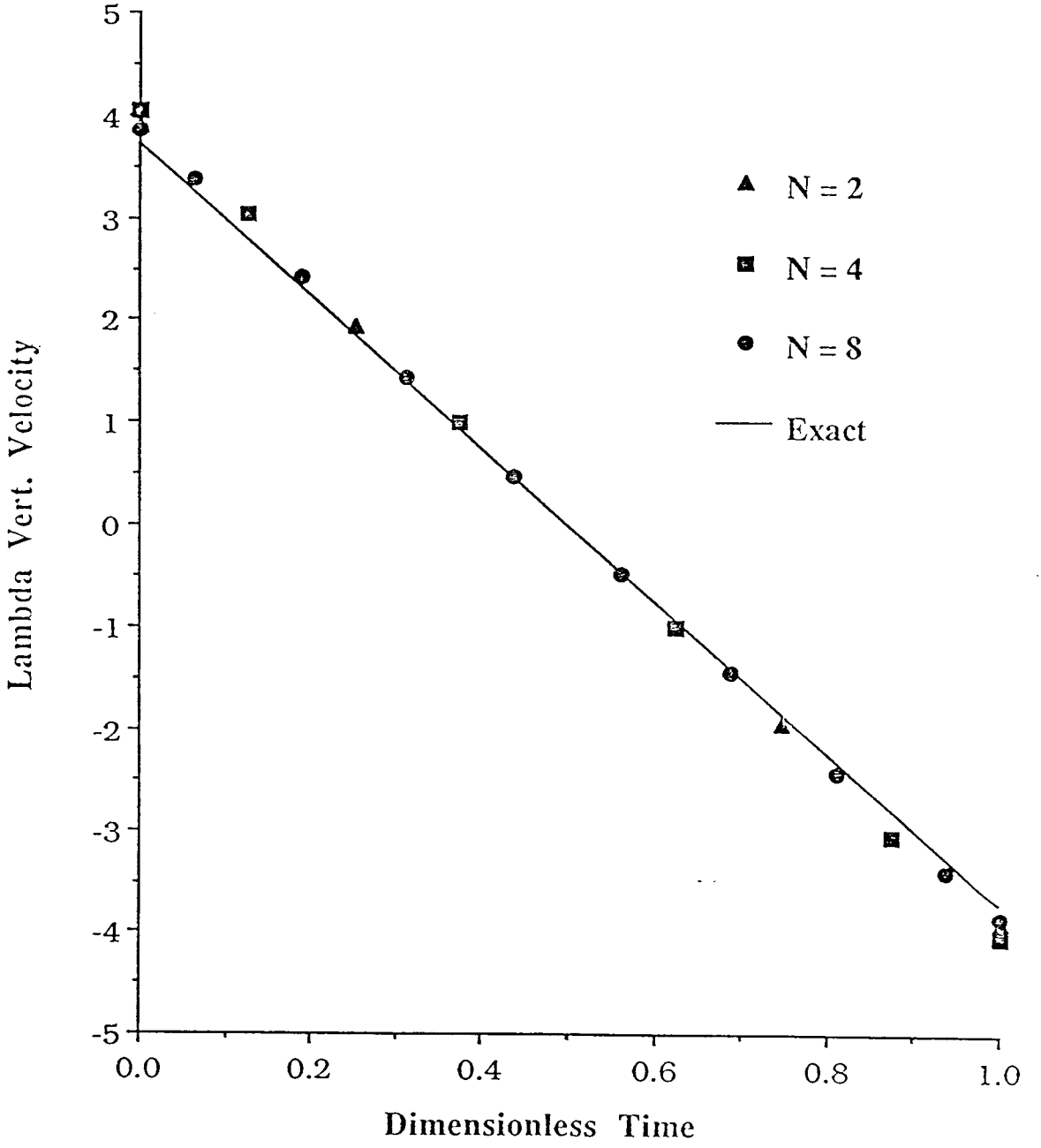

Fig. 10: Vertical velocity costate λ_4 versus t/T. The results for this costate are the least accurate of all the costates.

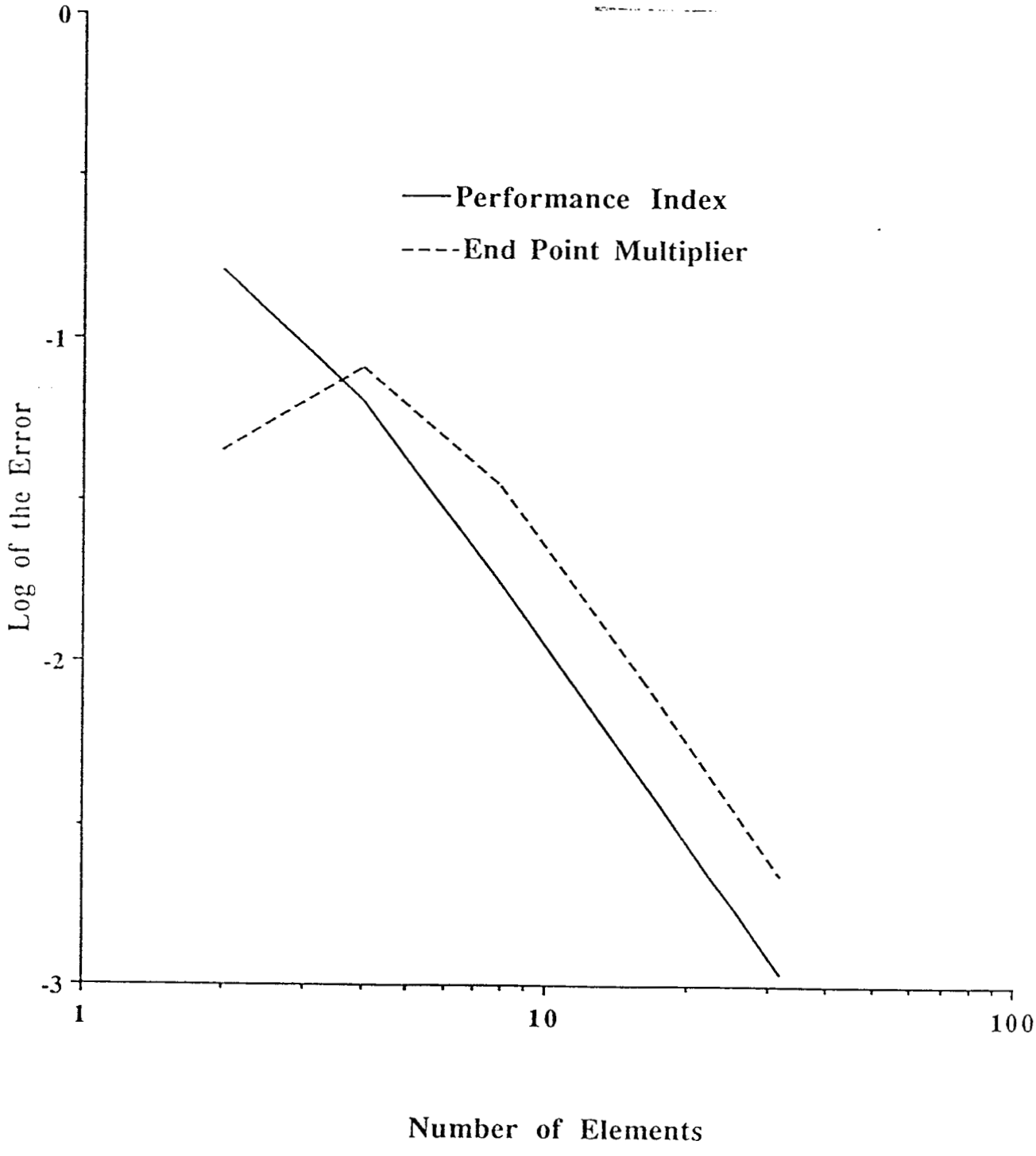

Fig. 11: Relative error of the performance index x_3 and the endpoint multiplier ν_1 versus N. The error varies inversely with the square of N.

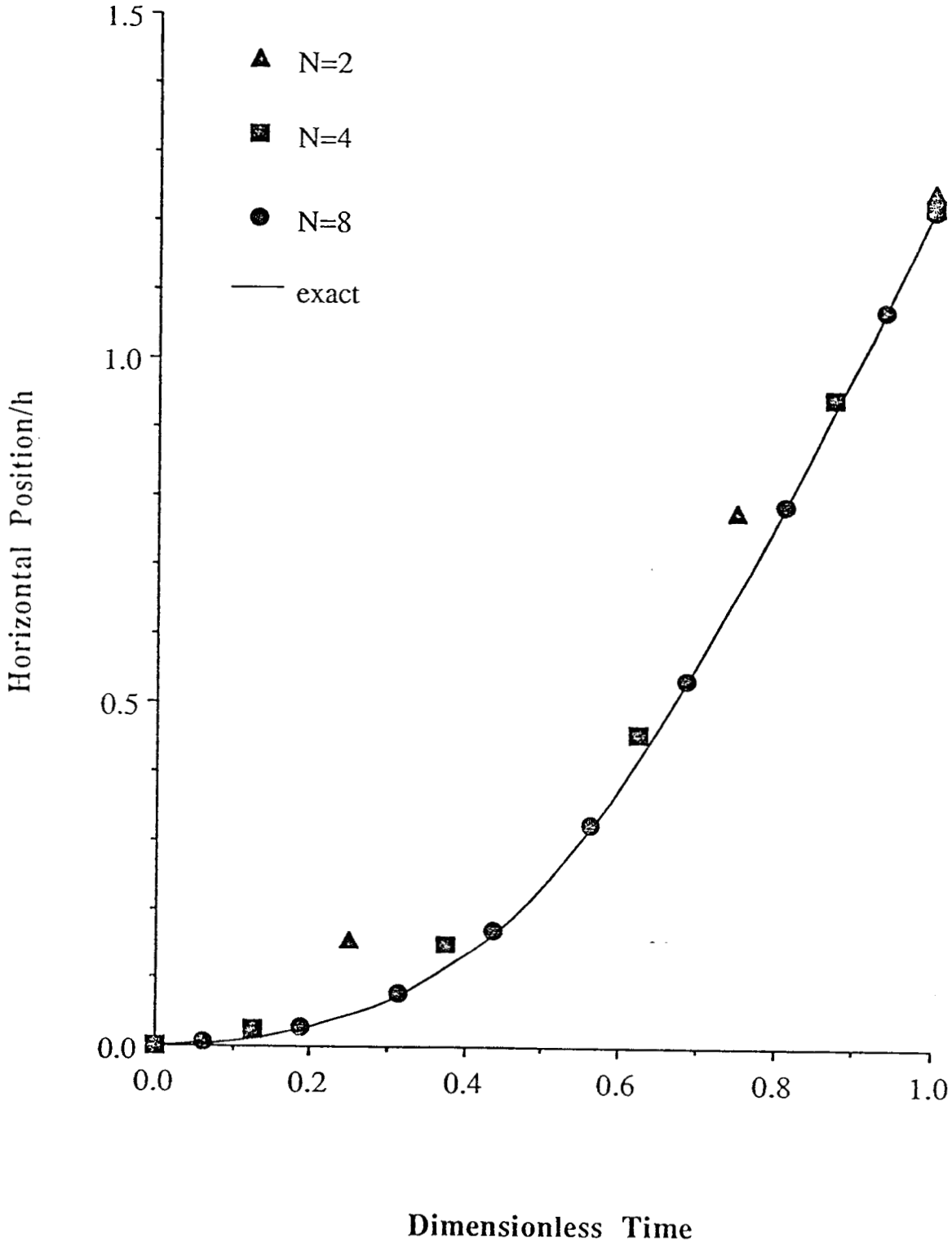

Fig. 12: Dimensionless horizontal position x_1/h versus t/T. Note that the final horizontal component of position is not specified.

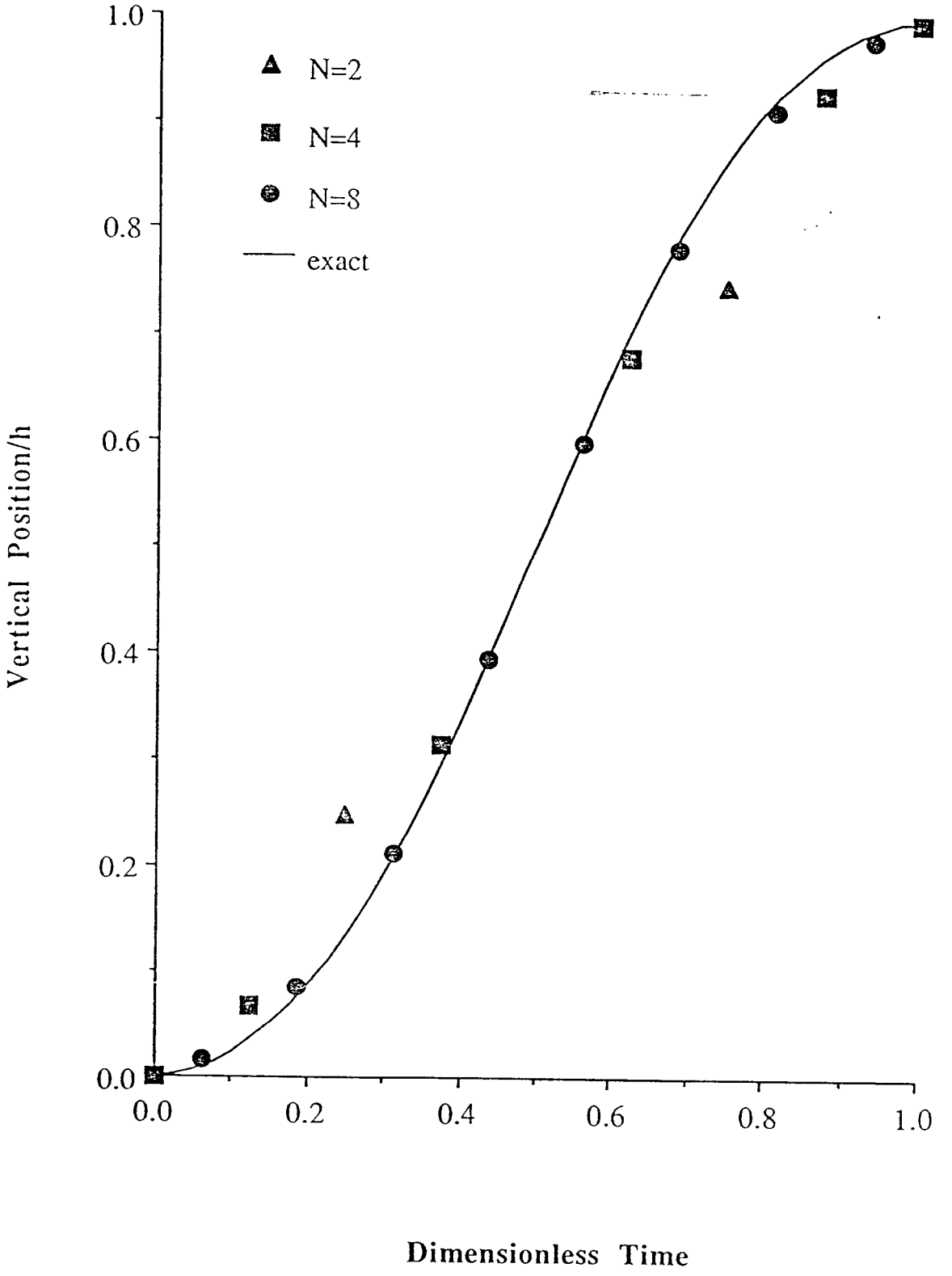

Fig. 13: Dimensionless vertical position x_2/h versus t/T. The final height is constrained to be h at the final time.

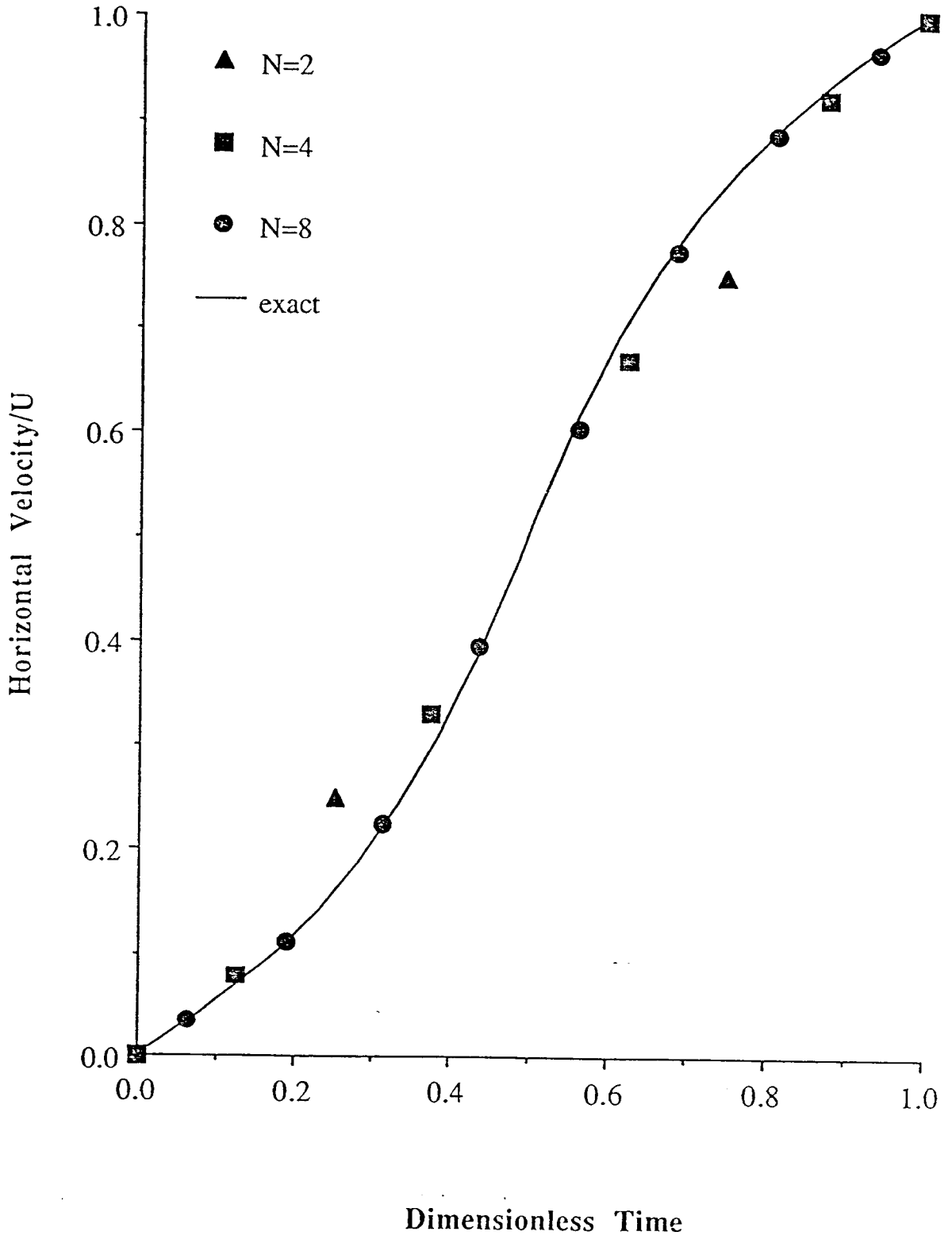

Fig. 14: Dimensionless horizontal velocity x_3/U versus t/T. The final horizontal component of velocity is constrained to be U. The problem is to reach this velocity in the minumum time T.

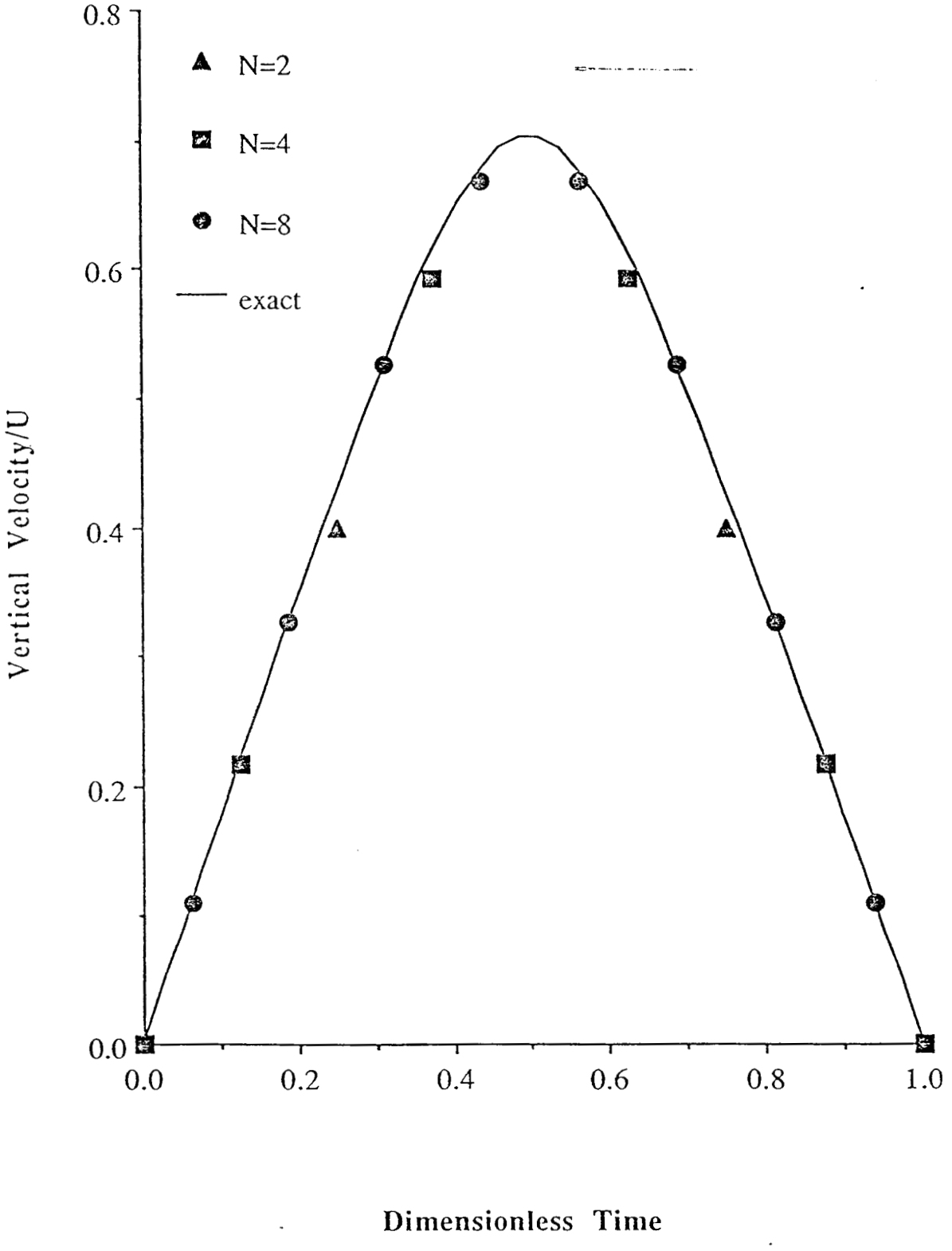

Fig. 15: Dimensionless vertical velocity x_4/U versus t/T. The final vertical component of velocity is constrained to be zero at the final time.

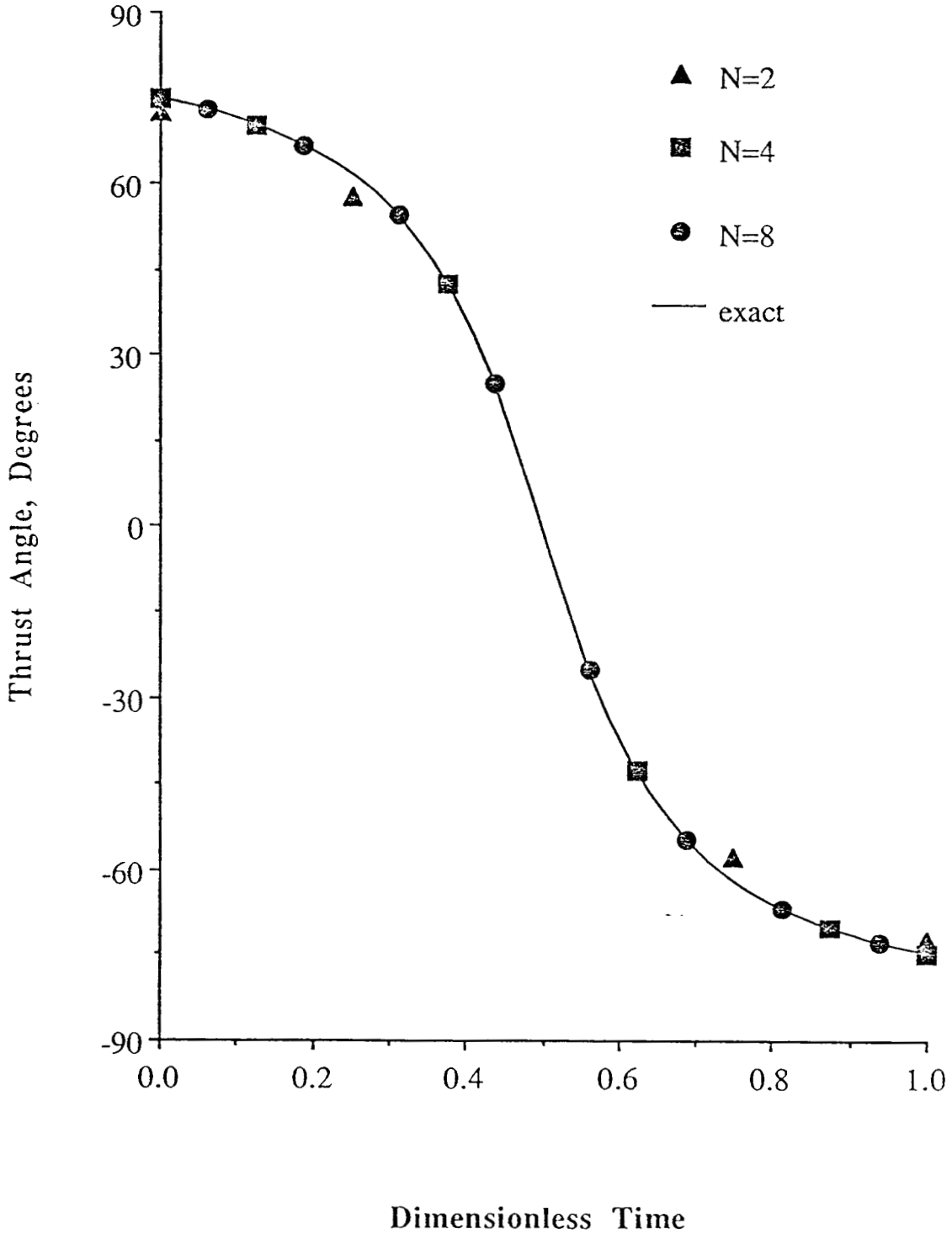

Fig. 16: Control angle u versus t/T. The accuracy of the initial control $u(t_0)$ determines the accuracy of the costates.

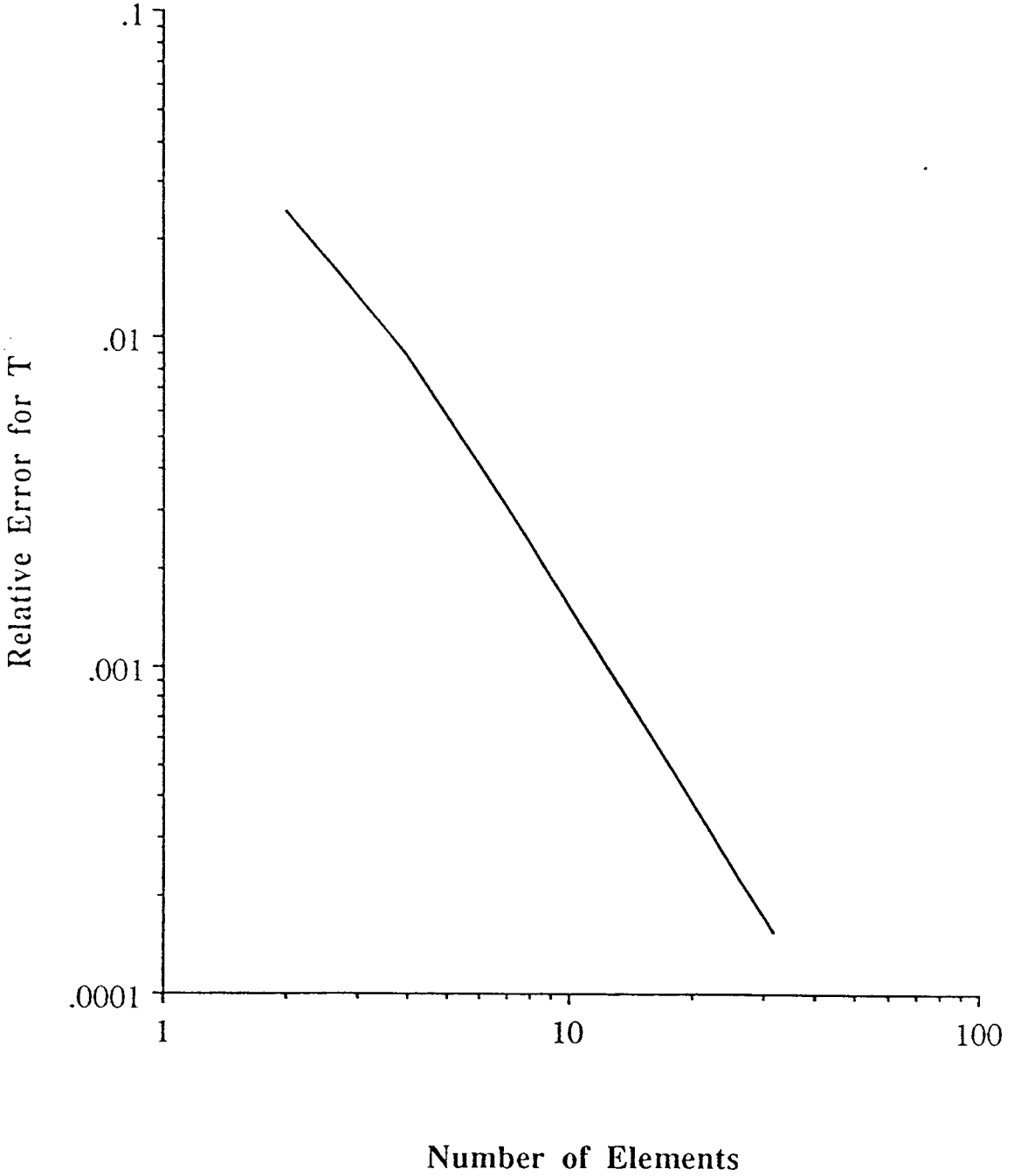

Fig. 17: Relative error of the performance index T versus N. The error varies inversely with the square of N.