Pergamon

J. Appl. Maths Mechs, Vol. 59, No. 4, pp. 513–523, 1995

0021–8928/95 $24.00+0.00

0021–8928(95)00060-7

APPROXIMATION OF THE SOLUTIONS IN BOUNDARY-VALUE PROBLEMS IN LAGRANGIAN MECHANICS†

I. I. KOSENKO

Sergiyev Posad

(*Received* 21 *November* 1994)

A projection method is described for constructing motions in Lagrangian mechanics. A Galerkin scheme is constructed for the trajectory in configuration space, using Hamilton's variational principle. The class of admissible paths for which the principle is considered consists of motions with fixed initial and terminal positions.

1. FORMULATION OF THE PROBLEM

Suppose we are given a holonomic mechanical system with n degrees of freedom. The configuration manifold is a domain $D \subset \mathbf{R}^n$. The motion of the system is considered over a finite time interval $[t_0, t_1] \subset \mathbf{R}$. The kinetic energy $T(t, \mathbf{q}, \dot{\mathbf{q}})$ is a sufficiently smooth function of its arguments, defined on a set $\Delta = [t_0, t_1] \times D \times \mathbf{R}^n$. In addition, in the general case it is assumed that

$$T(t,\mathbf{q},\dot{\mathbf{q}}) = T_2(t,\mathbf{q},\dot{\mathbf{q}}) + T_1(t,\mathbf{q},\dot{\mathbf{q}}) + T_0(t,\mathbf{q}) \tag{1.1}$$

$$T_2(t,\mathbf{q},\dot{\mathbf{q}}) = \frac{1}{2}\langle \mathrm{a}(t,\mathbf{q})\dot{\mathbf{q}}, \dot{\mathbf{q}}\rangle, \quad T_1(t,\mathbf{q},\dot{\mathbf{q}}) = \langle \mathbf{b}(t,\mathbf{q}), \dot{\mathbf{q}}\rangle$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean scalar product in $\mathbf{R}^n$ and $a(t, \mathbf{q})$ is the symmetric kinetic energy matrix, which is positive-definite throughout $[t_0, t_1] \times D$. It is assumed that this property of $a(t, \mathbf{q})$ also holds in a certain neighbourhood of $[t_0, t_1] \times D$.

We also assume defined on Δ a vector-valued function of generalized forces $\mathbf{Q}: \Delta \to \mathbf{R}^n$, which is sufficiently smooth with respect to $\mathbf{q}$ and $\dot{\mathbf{q}}$ for almost all $t \in [t_0, t_1]$, and square integrable with respect to t. All this can be summed up in a single condition

$$\mathbf{Q} \in L_2([t_0, t_1], \ C^1(D \times \mathbf{R}^n, \mathbf{R}^n))$$

It is additionally assumed that $\mathbf{Q}$ depends on the velocities to at most the second degree; this takes in a good many applications. The coefficients of the terms of the second degree in $\dot{\mathbf{q}}$ must be sufficiently smooth functions of t. The terms of the first degree in $\dot{\mathbf{q}}$ may have coefficients of which we demand only square integrability with respect to t.

The variational problem is formulated in the standard way: among all sufficiently smooth paths $\mathbf{q}: [t_0, t_1] \to D$ in the configuration space that satisfy the boundary conditions

$$\mathbf{q}(t_0) = \mathbf{q}_0, \quad \mathbf{q}(t_1) = \mathbf{q}_1$$

it is required to find one that satisfies the second-order Lagrange differential equation

$$(T_{\dot{\mathbf{q}}}) - T_{\mathbf{q}} = \mathbf{Q} \tag{1.2}$$

or what is the same, the variational equation

$$\int_{t_0}^{t_1} (\langle T_{\dot{\mathbf{q}}}, \delta\dot{\mathbf{q}}\rangle + \langle T_{\mathbf{q}}, \delta\mathbf{q}\rangle + \langle \mathbf{Q}, \delta\mathbf{q}\rangle) dt = 0 \tag{1.3}$$

†*Prikl. Mat. Mekh.* Vol. 59, No. 4, pp. 540–552, 1995.

for arbitrary sufficiently smooth functions $\delta\mathbf{q}$ satisfying the boundary conditions

$$\delta\mathbf{q}(t_0) = \delta\mathbf{q}(t_1) = \mathbf{0}$$

Throughout, we shall assume that this problem indeed has a solution, and that the solution is unique. We wish to design a uniform approximation procedure for the solution over the interval $[t_0, t_1]$ in the configuration space. In velocity space we confine ourselves to mean-square approximation. In what follows it will be more convenient to work with Eq. (1.3), after a preliminary transformation using integration by parts

$$\int_{t_0}^{t_1} \langle T_{\dot{\mathbf{q}}}(t,\mathbf{q}(t),\dot{\mathbf{q}}(t)) - \int_{t_0}^{t} (T_{\mathbf{q}}(\tau,\mathbf{q}(\tau),\dot{\mathbf{q}}(\tau)) + \mathbf{Q}(\tau,\mathbf{q}(\tau),\dot{\mathbf{q}}(\tau)))d\tau, \delta\dot{\mathbf{q}}\rangle dt = 0 \tag{1.4}$$

2. FUNCTIONAL MODEL

We will use a construction similar to that used in [1, 2], but with several substantial correctives. We first define the Hilbert space in which Eq. (1.4) will be modelled. This will be the space of paths $\mathbf{z}$: $[t_0, t_1] \to R^n$ satisfied by the conditions

$$\mathbf{z}(t_0) = \mathbf{z}(t_1) = \mathbf{0} \tag{2.1}$$

Given points $\mathbf{q}_0$ and $\mathbf{q}_1$ in the domain D, let us connect them in some way by a piecewise-smooth path $\mathbf{q}^0$: $[t_0, t_1] \to D$. If the structure of the set D permits, this may be, for example, a segment of a straight line or a vector-valued function with suitable analytical properties. In the general case the path $\mathbf{q}^0(t)$ may be a polygonal line connecting the points $\mathbf{q}_0, \mathbf{q}_1 \in D$. Yet another construction of $\mathbf{q}^0$, in perturbation problems, involves using trajectories of the unperturbed motion.

We now change to new generalized coordinates $\mathbf{q} \to \mathbf{z}$, by the formula

$$\mathbf{q} = \mathbf{q}^0(t) + \mathbf{z} \tag{2.2}$$

When this is done the Lagrange equations (1.2) and Hamilton's principle (1.3) retain their previous forms. The kinetic energy and generalized forces, however, are modified by the substitution (2.2). A solution of the variational problem

$$\int_{t_0}^{t_1} (\langle T_{\dot{\mathbf{z}}}(t,\mathbf{z}(t),\dot{\mathbf{z}}(t)) - \int_{t_0}^{t} (T_{\mathbf{z}}(\tau,\mathbf{z}(\tau),\dot{\mathbf{z}}(\tau)) + \mathbf{Q}(\tau,\mathbf{z}(\tau),\dot{\mathbf{z}}(\tau)))d\tau, \delta\dot{\mathbf{z}}\rangle dt = 0 \tag{2.3}$$

must now be sought among the functions $\mathbf{z}(t)$ that satisfy (2.1). Note that, after applying (2.2), the representation of the kinetic energy in the form (1.1) is changed. Nevertheless, we will retain the previous notation for the functions $a(t, \mathbf{z})$, $\mathbf{b}(t, \mathbf{z})$, $T_0(t, \mathbf{z})$. This is all the more natural in view of the fact that the matrix $a(t, \mathbf{z})$ will have the same values as before at points corresponding by (2.2).

The Hilbert space $\mathring{H}^1$ of paths $\mathbf{z}(t)$ satisfying condition (2.1) is defined by the scalar product

$$(\mathbf{z}_1, \mathbf{z}_2)^0 = \int_{t_0}^{t_1} \langle \dot{\mathbf{z}}_1(t),\ \dot{\mathbf{z}}_2(t)\rangle dt \tag{2.4}$$

This equality defines a metric in the space

$$\mathring{H}^1 = \mathring{H}^1([t_0, t_1], \mathbf{R}^n)$$

of functions with square-integrable derivatives that satisfy boundary conditions (2.1). We know that the norm in $\mathring{H}^1$, which is defined, according to (2.4), by

$$\|\mathbf{z}\|^0 = \left(\int_{t_0}^{t_1} \|\dot{\mathbf{z}}(t)\|^2\, dt\right)^{1/2} \tag{2.5}$$

is equivalent to the norm

$$\|\mathbf{z}\|^1 = \left(\int_{t_0}^{t_1} \|\mathbf{z}(t)\|^2 \, dt + \int_{t_0}^{t_1} \|\dot{\mathbf{z}}(t)\|^2 \, dt \right)^{1/2} \tag{2.6}$$

of the Sobolev space $H^1 = H^1([t_0, t_1], \mathbf{R}^n)$, if $\mathring{H}^1$ is treated as a subspace of H^1. This means that the use of the boundary conditions (2.1) enables us to pass from the norm (2.6) to the equivalent norm (2.5). In particular, one can speak of $\mathring{H}^1$ as embedded in the space of absolutely continuous functions in $[t_0, t_1]$ with values in $\mathbf{R}^n$. Hence it follows that a small distance in the norm (2.5) guarantees a uniformly small distance in the interval $[t_0, t_1]$.

We shall not need all elements of the space $\mathring{H}^1$, but only those for which the path $\mathbf{q}^0(t) + \mathbf{z}(t)$ belongs to D for all $t \in [t_0, t_1]$. It can then be shown, as in [1, 2], that the set

$$\Omega = \{\mathbf{z} \in \mathring{H}^1 : \mathbf{q}^0(t) + \mathbf{z}(t) \in D \;\forall t \in [t_0, t_1]\}$$

is a domain in $\mathring{H}^1$.

We now consider the left-hand side of Eq. (2.3). It defines a linear form in the tangent space $T_z\Omega = \mathring{H}^1$. For $\delta\mathbf{z} \in \mathring{H}^1$ we denote this form by $I_z(\delta\mathbf{z})$. It turns out that when $\mathbf{z} \in \Omega$ the linear functional L_z: $T_z\Omega \to \mathbf{R}$ is continuous.

To prove this, it will suffice to show that an element $\mathbf{X}(\mathbf{z}) \in \mathring{H}^1$ exists such that the functional L_z may be expressed as a scalar product in $\mathring{H}^1$

$$L_z(\delta\mathbf{z}) = (\mathbf{X}(\mathbf{z}), \delta\mathbf{z})^0 \tag{2.7}$$

which automatically implies the continuity of L_z.

Let us put

$$(\mathbf{l}(\mathbf{z}))(t) = T_{\dot{\mathbf{z}}}(t, \mathbf{z}(t), \dot{\mathbf{z}}(t)) - \int_{t_0}^{t} (T_{\mathbf{z}}(\tau, \mathbf{z}(\tau), \dot{\mathbf{z}}(\tau)) + \mathbf{Q}(\tau, \mathbf{z}(\tau), \dot{\mathbf{z}}(\tau)))d\tau \tag{2.8}$$

also introducing the following linear operators: averaging, $\mathbf{A}$: $L_2 \to L_2$; centring about the mean, $\mathbf{C}$: $L_2 \to L_2$; and primitive, $\mathbf{B}$: $L_2 \to H^1$, where $L_2 = L_2([t_0, t_1], \mathbf{R}^n)$, as follows:

$$(\mathbf{A}\mathbf{x})(t) = (t_1 - t_0)^{-1} \int_{t_0}^{t_1} \mathbf{x}(\tau)d\tau, \quad (\mathbf{B}\mathbf{x})(t) = \int_{t_0}^{t} \mathbf{x}(\tau)d\tau, \quad \mathbf{C}\mathbf{x} = \mathbf{x} - \mathbf{A}\mathbf{x}$$

Then it is obvious that for any $x \in L_2$ we have $\mathbf{BCx} \in \mathring{H}^1$, i.e. $(\mathbf{BCx})(t_0) = (\mathbf{BCx})(t_1) = \mathbf{0}$.

On the other hand, the variational problem (2.3) may be replaced by the equivalent problem

$$\int_{t_0}^{t_1} \langle(\mathbf{Cl}(\mathbf{z}))(t), \; \delta\dot{\mathbf{z}}(t)\rangle dt = 0 \tag{2.9}$$

Indeed, the integral (2.9) differs from (2.3) by the amount

$$\int_{t_0}^{t_1} \langle(\mathbf{Al}(\mathbf{z}))(t), \; \delta\dot{\mathbf{z}}(t)\rangle dt = \langle(\mathbf{Al}(\mathbf{z}))(t), \delta\mathbf{z}(t)\rangle \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \langle(\mathbf{Al}(\mathbf{z}))^{\cdot}(t), \; \delta\mathbf{z}(t)\rangle dt = 0$$

since conditions (2.1) are satisfied and the mean value is independent of $t \in [t_0, t_1]$: $(\mathbf{Al}(\mathbf{z}))(t) \equiv \mathbf{const}$.

Since $(\mathbf{Bx})\,(t) \equiv \mathbf{x}(t)$, the function

$$(\mathbf{X}(\mathbf{z}))(t) = (\mathbf{BCl}(\mathbf{z}))(t) \tag{2.10}$$

satisfies Eq. (2.7) if $\mathbf{l}(\mathbf{z}) \in L_2$.

Indeed, $T_{\dot{z}}$ depends linearly on $\dot{z}$ with sufficiently regular coefficients and a linear function as free term. Therefore $T_{\dot{z}} \in L_2 \subset L_1$, where $L_1 = L_1([t_0, t_1], \mathbf{R}^n)$.

The degree of $T_{\dot{z}}$ with respect to $\dot{z}$ is at most two. The terms of degree zero in T_z are sufficiently regular. Therefore

T_{0z} is integrable with respect to t. Those terms that are linear in $\dot{z}$ also yield a function that is integrable with respect to t (as does $T_{\dot{z}}$). The quadratic terms have the form $T_{2\dot{z}} = \langle a_z\dot{\mathbf{z}}, \dot{\mathbf{z}}\rangle/2$ and may have the upper limit

$$\|\langle a_z\dot{\mathbf{z}},\dot{\mathbf{z}}\rangle/2\| \leq \text{const}\,\|\dot{\mathbf{z}}\|^2$$

Since $\dot{\mathbf{z}} \in L_2$, it follows that $\langle a_z\dot{\mathbf{z}}, \dot{\mathbf{z}}\rangle/2 \in L_1$. Finally, we can state that $T_z \in L_1$.

By our assumption, the vector-valued function **Q** also depends on $\dot{z}$ to degree of at most two, and the coefficients in the quadratic terms must be regular. In that case **Q** will consist of terms involving the derivative of a regular function and at most two functions in L_2. Therefore $\mathbf{Q} \in L_1$.

Thus, we have proved that $T_z + \mathbf{Q} \in L_1$. Therefore $\mathbf{B}(\mathbf{T}_z + \mathbf{Q}) \in \mathrm{CA} \subset L^2$.

Now, recalling Riesz' representation theorem for continuous linear functionals in Hilbert space [3], according to which the element $\mathbf{X}(\mathbf{z})$ in the scalar product in (2.7) is uniquely defined, we conclude that formula (2.10) uniquely defines an operator $\mathbf{X}: \Omega \to \mathring{H}^1$.

The variational equation of Hamilton's principle (2.9) has now been reduced to a functional equation in $\Omega \subset \mathring{H}^1$, of the form

$$\mathbf{X}(\mathbf{z}) = \mathbf{0} \tag{2.11}$$

In what follows we shall use a different representation of this equation

$$\mathbf{z} = \mathbf{Z}(\mathbf{z}) \tag{2.12}$$

To transform (2.11) to this form, let us consider the structure of the operator **X** in more detail. It follows from (2.8) that

$$(\mathbf{l}(\mathbf{z}))\,(t) = a(t, \mathbf{z}(t))\,\dot{\mathbf{z}}(t) + (\boldsymbol{\lambda}(\mathbf{z}))\,(t)$$

$$(\boldsymbol{\lambda}(\mathbf{z}))(t) = \mathbf{b}(t, \mathbf{z}(t)) - \int_{t_0}^{t} (T_z(\tau, \mathbf{z}(\tau),\ \dot{\mathbf{z}}(\tau)) + \mathbf{Q}(\tau, \mathbf{z}(\tau), \dot{\mathbf{z}}(\tau)))d\tau$$

Hence we conclude, integrating the first term by parts and applying the operator **B**, that

$$\begin{aligned}
&(\mathbf{BCl}(\mathbf{z}))(t) = (\mathbf{Bl}(\mathbf{z}))(t) - (\mathbf{BAl}(\mathbf{z}))(t) = a(t, \mathbf{z}(t))\mathbf{z}(t) + (\mathbf{Y}(\mathbf{z}))(t)\\
&(\mathbf{Y}(\mathbf{z}))(t) = -\int_{t_0}^{t} \dot{a}(\tau, \mathbf{z}(\tau), \dot{\mathbf{z}}(\tau))\mathbf{z}(\tau)d\tau + \mathbf{B}\boldsymbol{\lambda}(\mathbf{z}))(t) -\\
&-(t - t_0)\,(\mathbf{Al}(\mathbf{z}))(t),\ \ \dot{a}(t, \mathbf{z}, \dot{\mathbf{z}}) = a_t(t, \mathbf{z}) + a_z(t, \mathbf{z})\dot{\mathbf{z}}
\end{aligned} \tag{2.13}$$

where the evaluation of $\mathbf{Bl}(\mathbf{z})$ also involves integration by parts. We now define the operator **Z**: $\Omega \to \mathring{H}^1$ in Eq. (2.12) by

$$(\mathbf{Z}(\mathbf{z}))(t) = -a^{-1}(t, \mathbf{z}(t))(\mathbf{Y}(\mathbf{z}))(t) \tag{2.14}$$

In the formalism proposed in [1, 2], one can construct in $\mathring{H}^1$ a system of subspaces $E_m \subset \mathring{H}^1$ ($m = 1, 2, \ldots$), exhausting the space, with the help of an orthonormal basis $\{\mathbf{e}_j\chi_k\}$ ($j = 1, 2, \ldots, n$; $k = 1, 2, \ldots$), where the vectors $\{\mathbf{e}_j\}_{j=1}^{n}$ form an orthonormal basis in $\mathbf{R}^n$ and the system $\{\chi_k\}_{k=1}^{\infty}$ a basis in the space of scalar functions that vanish at the endpoints of the interval $[t_0, t_1]$. As the functions χ_k one can take, e.g. the system

$$\chi_k(t) = (2/(t_1 - t_0))^{1/2}(\pi k)^{-1}\sin(k\pi(t_1 - t_0)^{-1}(t - t_0)) \tag{2.15}$$

($k = 1, 2, \ldots$).

To construct a Galerkin scheme, we define projection operators whose ranges are the finite-dimensional subspaces P_m: $\mathring{H}^1 \to E_m$, as projections which are orthogonal in the metric of $\mathring{H}^1$.

3. DIFFERENTIABILITY

The approximation theorem for solutions of Eq. (2.12) requires the operator $\mathbf{Z}\colon \Omega \to \overset{\circ}{H}{}^1$ to have certain properties. In particular, $\mathbf{Z}$ should be fairly regular. Namely, it can be proved that the operator $\mathbf{Z}\colon \Omega \to \overset{\circ}{H}{}^1$ is continuously Fréchet-differentiable.

Indeed, let us use the structure (2.14) of the operator. It is clear from (2.13) that the operator $\mathbf{Y}\colon \Omega \to \overset{\circ}{H}{}^1$ can be defined by

$$\mathbf{Y}(\mathbf{z}) = \mathbf{BC}(\boldsymbol{\lambda}(\mathbf{z}) - \boldsymbol{\alpha}(\mathbf{z})), \quad \boldsymbol{\alpha}(\mathbf{z}) = \dot{a}(t, \mathbf{z}(t), \dot{\mathbf{z}}(t))\mathbf{z}(t)$$

We have used the fact that $\mathbf{A}((a\mathbf{z})^{\cdot}) = \mathbf{0}$.

However, in order to verify that $\mathbf{Z}$ is smooth, it is not enough to verify this for $\mathbf{Y}$. One must also take into consideration that the vector-valued function $(\mathbf{Y}(\mathbf{z}))(t)$ must be premultiplied by the matrix $-a^{-1}(t, \mathbf{z}(t))$. This multiplication defines a linear operator in $\overset{\circ}{H}{}^1$. Since by assumption $a(t, \mathbf{z})$ is a positive-definite symmetric matrix and a sufficiently regular function of its arguments, the same is true of the matrix $a^{-1}(t, \mathbf{z})$.

Let $\mathbf{h} \in \overset{\circ}{H}{}^1$. Denote the operator defined by the formula

$$\mathbf{h}(t) \to a(t, \mathbf{z}(t))\mathbf{h}(t) \tag{3.1}$$

by $\boldsymbol{\Gamma}(\mathbf{z})$. Clearly, for $\mathbf{z} \in \overset{\circ}{H}{}^1$, the action of $\boldsymbol{\Gamma}(\mathbf{z})$ does not take us out of $\overset{\circ}{H}{}^1$, i.e. $\boldsymbol{\Gamma}(\mathbf{z})\colon \overset{\circ}{H}{}^1 \to \overset{\circ}{H}{}^1$. It is easy to show that $\boldsymbol{\Gamma}(\mathbf{z})$ is an element of the algebra $L(\overset{\circ}{H}{}^1, \overset{\circ}{H}{}^1)$ of continuous linear operators in $\overset{\circ}{H}{}^1$.

Now, if $\mathbf{z} \in \Omega$, the matrix $a^{-1}(t, \mathbf{z}(t))$ has the same properties as $a(t, \mathbf{z}(t))$ and is moreover its inverse. Hence, the operator, $\boldsymbol{\Gamma}(\mathbf{z})$, has an inverse $\boldsymbol{\Gamma}^{-1}(\mathbf{z})\colon \overset{\circ}{H}{}^1 \to \overset{\circ}{H}{}^1$, and the latter is bounded and linear: $\boldsymbol{\Gamma}^{-1}(\mathbf{z}) \in L(\overset{\circ}{H}{}^1, \overset{\circ}{H}{}^1)$.

Thus, if $\mathbf{z} \in \Omega$, the operator $\boldsymbol{\Gamma}(\mathbf{z})$ is continuously invertible in $\overset{\circ}{H}{}^1$. Denote its action on an element $\mathbf{h} \in \overset{\circ}{H}{}^1$ by the composition symbol $(\boldsymbol{\Gamma}(\mathbf{z})\mathbf{h})$. Then the operator $\mathbf{Z}(\mathbf{z})$ may be defined as the composition of $-\boldsymbol{\Gamma}^{-1}(\mathbf{z})$ and $\mathbf{Y}(\mathbf{z})$, i.e. $\mathbf{Z}(\mathbf{z}) = (-\boldsymbol{\Gamma}^{-1}(\mathbf{z}) \circ \mathbf{Y}(\mathbf{z}))$.

It is well known [4] that a bilinear mapping

$$(\cdot \circ \cdot)\colon L(\overset{\circ}{H}{}^1, \overset{\circ}{H}{}^1) \times \overset{\circ}{H}{}^1 \to \overset{\circ}{H}{}^1$$

is continuous and of norm 1. It is also well known [4] that a continuous bilinear mapping is continuously differential with respect to both its arguments. Moreover, if $\mathbf{u}_1(\mathbf{z})$ and $\mathbf{u}_2(\mathbf{z})$ are continuously differentiable as functions $\mathbf{u}_1\colon \Omega \to L(\overset{\circ}{H}{}^1, \overset{\circ}{H}{}^1)$, $\mathbf{u}_2\colon \Omega \to \overset{\circ}{H}{}^1$, then the same is true of the function $(\mathbf{u}_1(\mathbf{z}) \circ u_2(z))$, and Leibniz's differentiation formula will hold [4].

The problem has been reduced to proving that the mapping $\boldsymbol{\Gamma}\colon \Omega \to L(\overset{\circ}{H}{}^1, \overset{\circ}{H}{}^1)$, $\mathbf{Y}\colon \Omega \to \overset{\circ}{H}{}^1$ is continuously differentiable. The proof for $\boldsymbol{\Gamma}^{-1}$ is the same as for $\boldsymbol{\Gamma}$.

Since the elements of the range of $\boldsymbol{\Gamma}$ are bounded linear operators, the elements of the range of the Fréchet derivative $\boldsymbol{\Gamma}'$ (if it exists) are continuous bilinear forms.

Noting that the operator $\boldsymbol{\Gamma}(\mathbf{z})\colon \overset{\circ}{H}{}^1 \to \overset{\circ}{H}{}^1$ was defined by formula (3.1), one can prove that its derivative is defined by the formula

$$((\boldsymbol{\Gamma}'(\mathbf{z})\mathbf{h})\mathbf{h}_1)(t) = (a_{\mathbf{z}}(t, \mathbf{z}(t))\mathbf{h}(t))\mathbf{h}_1(t) \quad (\mathbf{h}, \mathbf{h}_1 \in \overset{\circ}{H}{}^1) \tag{3.2}$$

where $a_z(t, \mathbf{z}(t))\mathbf{h}(t)$ is a linear operator in $\mathbf{R}^n$ which depends linearly on the components of the function $\mathbf{h} \in \overset{\circ}{H}{}^1$. Since the kinetic energy matrix is sufficiently smooth, the same will hold for the matrix of the tensor of its partial derivatives. It is therefore easy to verify that formula (3.2) defines a continuous bilinear form

$$\boldsymbol{\Gamma}'(\mathbf{z}) \in L(\overset{\circ}{H}{}^1, L(\overset{\circ}{H}{}^1, \overset{\circ}{H}{}^1))$$

To prove Fréchet differentiability, one estimates the difference $\boldsymbol{\omega}(\mathbf{h})/\|\mathbf{h}\|^0 = (\boldsymbol{\Gamma}(\mathbf{z} + \mathbf{h}) - \boldsymbol{\Gamma}(\mathbf{z}) - \boldsymbol{\Gamma}'(\mathbf{z})\mathbf{h})/\|\mathbf{h}\|^0$ and proves that $\boldsymbol{\omega}(\mathbf{h})/\|\mathbf{h}\|^0 \to 0$ as $\|\mathbf{h}\|^0 \to 0$. Having shown that the function $\boldsymbol{\Gamma}\colon \Omega \to L(\overset{\circ}{H}{}^{11}, \overset{\circ}{H}{}^1)$

is differentiable, one can prove that the derivative $\Gamma': \Omega \to L(\mathring{H}^1, L(\mathring{H}^1, \mathring{H}^1))$ is continuous, again using the smoothness of $a(t, \mathbf{z})$.

We will now analyse the operator $\mathbf{Y}: \Omega \to \mathring{H}^1$. It is the composition of three operators, defined by a commutative diagram

$$\begin{array}{ccc} \Omega & \xrightarrow{\mathbf{Y}} & \mathring{H}^1 \\ \downarrow \mathbf{F} & & \uparrow \mathbf{B} \\ L_2 & \xrightarrow{\mathbf{C}} & L_2 \end{array}$$

where the operator $\mathbf{F}: \Omega \to L_2$ is defined by $\mathbf{F}(\mathbf{z}) = \boldsymbol{\lambda}_{\mathbf{z}}(\mathbf{z}) - \boldsymbol{\alpha}_{\mathbf{z}}(\mathbf{z})$.

It is obvious that the linear operator $\mathbf{B}: L_2 \to \mathring{H}^1$ is bounded, and it is even isometric in the subspace of L_2 of functions with zero mean. Hence $\mathbf{B}$ is continuously differentiable on $\mathbf{C}(L_2)$. The operator $\mathbf{C}: L_2 \to L_2$ is also linear and bounded, since it has the form $\mathbf{C} = \mathbf{I} - \mathbf{A}$, where the averaging operator $\mathbf{A}$ is bounded. Finally, we see that the linear operator $\mathbf{BC}$: $L_2 \to \mathring{H}^1$ is bounded and, therefore, continuously differentiable.

The proof that the operator $\mathbf{F}: \Omega \to L_2$ is differentiable uses standard methods. $\mathbf{F}$ may be evaluated by the formula

$$(\mathbf{F}(\mathbf{z}))(t) = \mathbf{b}(t, \mathbf{z}(t)) - a_t(t, \mathbf{z}(t))\,\mathbf{z}(t) - (a_{\mathbf{z}}(t, \mathbf{z}(t))\,\dot{\mathbf{z}}(t))\,\mathbf{z}(t) - \\ - \int_{t_0}^{t} (T_{\mathbf{z}}(\tau, \mathbf{z}(\tau), \dot{\mathbf{z}}(\tau)) + \mathbf{Q}(\tau, \mathbf{z}(\tau), \dot{\mathbf{z}}(\tau)))\,d\tau$$

One then proves that the Fréchet derivative $\mathbf{F}'(\mathbf{z})$ may be written as

$$(\mathbf{F}'(\mathbf{z})\mathbf{h})(t) = \mathbf{b}_{\mathbf{z}}(t, \mathbf{z}(t))\mathbf{h}(t) - (a_{t\mathbf{z}}(t, \mathbf{z}(t))\mathbf{h}(t))\mathbf{z}(t) - a_t(t, \mathbf{z}(t))\mathbf{h}(t) - \\ ((a_{\mathbf{zz}}(t, \mathbf{z}(t))\mathbf{h}(t))\dot{\mathbf{z}}(t))\mathbf{z}(t) - (a_{\mathbf{z}}(t, \mathbf{z}(t))\dot{\mathbf{h}}(t))\mathbf{z}(t) - (a_{\mathbf{z}}(t, \mathbf{z}(t))\dot{\mathbf{z}}(t))\mathbf{h}(t) - \\ - \int_{t_0}^{t} (\mathbf{S}_{\mathbf{z}}(\tau, \mathbf{z}(\tau), \dot{\mathbf{z}}(\tau))\mathbf{h}(\tau) + \mathbf{S}_{\dot{\mathbf{z}}}(\tau, \mathbf{z}(\tau), \dot{\mathbf{z}}(\tau))\dot{\mathbf{h}}(\tau))d\tau \tag{3.3}$$

We recall that the function $\mathbf{S}(t, \mathbf{z}, \dot{\mathbf{z}}) = T_{\mathbf{z}}(t, \mathbf{z}, \dot{\mathbf{z}}) + \mathbf{Q}(t, \mathbf{z}, \dot{\mathbf{z}})$ is of at most the second order in the generalized velocities; moreover, the coefficients of the second-order terms are assumed to be regular with respect to t and $\mathbf{z}$, non-regularity with respect to t being admissible only in the coefficients of the terms linear in $\dot{\mathbf{z}}$.

In the proof that $\mathbf{F}'(\mathbf{z}) \in L(\mathring{H}^1, L_2)$, each term in $\mathbf{F}'(\mathbf{z})$ is considered separately. It can in fact be verified that the operator $\mathbf{F}'(\mathbf{z})$: $\mathring{H}^1 \to L_2$ is bounded.

The fact that this operator is the Fréchet derivative of $\mathbf{F}: \Omega \to L_2$ may be verified by estimating the norm of the distance $\boldsymbol{\omega}(\mathbf{h}) = \mathbf{F}(\mathbf{z} + \mathbf{h}) - \mathbf{F}(\mathbf{z}) - \mathbf{F}'(\mathbf{z})\mathbf{h}$ in L_2.

To verify the continuity of $\mathbf{F}'$ with respect to $\mathbf{z}_1, \mathbf{z}_2 \in \Omega$, suppose that $\| \mathbf{z}_1 - \mathbf{z}_2 \|^0 \to 0$ and consider the operator norm

$$\|\mathbf{F}'(\mathbf{z}_1) - \mathbf{F}'(\mathbf{z}_2)\| = \sup_{\|\mathbf{h}_1\|^0 = 1} \|(\mathbf{F}'(\mathbf{z}_1) - \mathbf{F}'(\mathbf{z}_2))\mathbf{h}\|_2$$

Using formula (3.3), one can verify the continuity of each term in this norm.

4. CONTINUOUS INVERTIBILITY

Having convinced ourselves that the operator $\mathbf{Z}: \Omega \to \mathring{H}^1$, and hence also the operator $\mathbf{I} - \mathbf{Z}: \Omega \to \mathring{H}^1$, are continuously differentiable everywhere in Ω, we consider the question of whether the derivative

$$\mathbf{I} - \mathbf{Z}'(\mathbf{z}): \mathring{H}^1 \to \mathring{H}^1 \tag{4.1}$$

is continuously invertible in the tangent space to Ω at $\mathbf{Z}$. It will suffice to verify invertibility at a solution of Eq. (2.12). Denote this solution by $\mathbf{y} \in \Omega$.

In Section 3 we derived an expression for the derivative of the operator $\mathbf{Z}$. The complete expression may be written as

$$\mathbf{Z}'(\mathbf{z})\mathbf{h} = -((\mathbf{\Gamma}^{-1})'(\mathbf{z})\mathbf{h} \circ \mathbf{Y}(\mathbf{z})) - (\mathbf{\Gamma}^{-1}(\mathbf{z}) \circ \mathbf{Y}'(\mathbf{z})\mathbf{h})$$

To prove that the operator (4.1) is invertible, we must solve the equation

$$(\mathbf{I} - \mathbf{Z}'(\mathbf{z}))\mathbf{h} = \mathbf{g}$$

for the function $\mathbf{h} \in \overset{\circ}{H}{}^1$, where $\mathbf{g} \in \overset{\circ}{H}{}^1$ is arbitrary. Let us write out this equation in expanded form for $\mathbf{z} = \mathbf{y}$

$$\mathbf{h} + ((\mathbf{\Gamma}^{-1})'(\mathbf{y})\mathbf{h} \circ \mathbf{Y}(\mathbf{y})) + (\mathbf{\Gamma}^{-1}(\mathbf{y}) \circ \mathbf{Y}'(\mathbf{y})\mathbf{h}) = \mathbf{g} \tag{4.2}$$

The Lagrange equations of the second kind for the vector-valued configuration function $\mathbf{z}(t)$ may be written in the following general form

$$\ddot{\mathbf{z}} = \mathbf{f}(t, \mathbf{z}, \dot{\mathbf{z}}) \tag{4.3}$$

where $\mathbf{f}(t, \mathbf{z}, \dot{\mathbf{z}}) = \mathbf{f}_2(t, \mathbf{z}, \dot{\mathbf{z}}) + \mathbf{f}_1(t, \mathbf{z}, \dot{\mathbf{z}}) + \mathbf{f}_0(t, \mathbf{z}, \dot{\mathbf{z}})$, the $\mathbf{f}_i$ being homogeneous forms of degree i in the velocities ($i = 0, 1, 2$). If we assume that $\mathbf{z} \in \overset{\circ}{H}{}^1$, then $\dot{\mathbf{z}} \in L_2$. It is evident from the analysis in Section 2 of the kinetic energy function $T(t, \mathbf{z}, \dot{\mathbf{z}})$ and the function of generalized forces $\mathbf{Q}(t, \mathbf{z}, \dot{\mathbf{z}})$ that substitution of $\mathbf{z} \in \overset{\circ}{H}{}^{11}$ into the right-hand side of (4.3) yields a function $\mathbf{f}(\cdot, \mathbf{z}, \dot{\mathbf{z}}) \in L_1$. System (4.3) satisfies the well-known Carathéodory conditions [5] for equations with not necessarily continuous right-hand sides. Therefore, the Cauchy problem for (4.3) has a solution, which is moreover unique.

Let $\mathbf{y}(t)$ be a solution of the boundary-value problem (2.1) for (4.3). Since $\ddot{\mathbf{y}} \in L_1$, it follows that $\dot{\mathbf{y}} \in \mathrm{CA}$. Consequently, the limit $\lim \dot{\mathbf{y}}(t) = \dot{\mathbf{y}}(t_0)$ as $t \to t_0$ exists and $\mathbf{y}(t)$ is the solution of the Cauchy problem for (4.3) with initial data

$$\mathbf{z}(t_0) = \mathbf{0}, \quad \dot{\mathbf{z}}(t_0) = \dot{\mathbf{y}}_0$$

Introducing the notation

$$a(t, \mathbf{y}(t)) = a_1(t), \quad (a_{\mathbf{z}}(t, \mathbf{y}(t))\mathbf{h}(t))\dot{\mathbf{y}}(t) = a_2(t)\mathbf{h}(t)$$
$$\mathbf{S}_{\mathbf{z}}(t, \mathbf{y}(t), \dot{\mathbf{y}}(t)) = a_4(t), \quad \mathbf{S}_{\dot{\mathbf{z}}}(t, \mathbf{y}(t), \dot{\mathbf{y}}(t)) = a_3(t)$$

in order to simplify what follows, and noting that $\mathbf{y}(t)$ satisfies Eq. (2.12), while the right-hand side of the equation has the form (2.14), we obtain, after multiplying (4.2) by $a_i(t)$ and differentiating with respect to t

$$(\mathbf{A}(\mathbf{h}))(t) - \mathbf{b}_1(\mathbf{h}) = (a_1(t)\mathbf{g}(t)) \tag{4.4}$$
$$(\mathbf{A}(\mathbf{h}))(t) = a_1(t)\dot{\mathbf{h}}(t) + a_2(t)\mathbf{h}(t) - \int_{t_0}^{t} (a_3(\tau)\dot{\mathbf{h}}(\tau) + a_4(\tau)\mathbf{h}(\tau))d\tau$$
$$\mathbf{b}_1(\mathbf{h}) \equiv \mathrm{const} \equiv (\mathbf{AF}'(\mathbf{y})\mathbf{h})(t)$$

For given $\mathbf{g} \in \overset{\circ}{H}{}^1$, Eq. (4.4) is linear inhomogeneous in $\mathbf{h} \in \overset{\circ}{H}{}^1$. It can be shown that if Eq. (4.4) has a solution, that solution is unique.

Let $\mathbf{h}_1, \mathbf{h}_2 \in \overset{\circ}{H}{}^1$ be solutions of Eq. (4.4) for a given $\mathbf{g}$. Their difference $\Delta\mathbf{h} = \mathbf{h}_1 - \mathbf{h}_2$ will satisfy the corresponding homogeneous equation. It is legitimate to differentiate this equation with respect to t, because only $a_2(t)$ contains $\dot{\mathbf{y}}(t)$, which is an element of CA, and therefore $\dot{a}_2(t) \in L_1$; but this is quite sufficient for the Carathéodory conditions to hold. Differentiating, we obtain a system of equations in variations for the original Lagrangian system

$$(a_1\dot{\mathbf{h}} + a_2\mathbf{h})^{\cdot} - a_3\dot{\mathbf{h}} - a_4\mathbf{h} = 0 \tag{4.5}$$

We know that this system has a resolvent $R(t, t_0)$, and so the general solution of system (4.5) may be written in the form

$$\begin{pmatrix} \mathbf{h}(t) \\ \dot{\mathbf{h}}(t) \end{pmatrix} = R(t, t_0) \begin{pmatrix} \mathbf{h}_0 \\ \dot{\mathbf{h}}_0 \end{pmatrix} \tag{4.6}$$

If we assume that $\mathbf{h} \in \overset{\circ}{H}{}^1$, then $\mathbf{h}_0 = \mathbf{h}(t_0) = \mathbf{0}$. Therefore

$$\mathbf{h}(t) = B(t, t_0)\dot{\mathbf{h}}_0 \tag{4.7}$$

where $B(t, t_0)$ is the upper right block in the matrix $R(t, t_0)$.

Under our assumptions (as is obvious from Eq. (4.5)), the solution satisfies the condition $\ddot{\mathbf{h}} \in L_1$, whence it follows that $\dot{\mathbf{h}}$ is continuous, so that the limit $\lim \dot{\mathbf{h}}(t) = \dot{\mathbf{h}}_0$ exists as $t \to t_0$. Therefore Eq. (4.7) is meaningful. In particular, we put $t = t_1$. We obtain an equation for $\dot{\mathbf{h}}_0$

$$B(t_1, t_0)\dot{\mathbf{h}}_0 = \mathbf{h}(t_1) = \mathbf{0} \tag{4.8}$$

We now introduce another essential assumption

$$\det B(t_1, t_0) \neq 0 \tag{4.9}$$

This means that in the mapping along a trajectory the upper right minor of the determinant

$$\partial(\mathbf{h}, \dot{\mathbf{h}}) / \partial(\mathbf{h}_0, \dot{\mathbf{h}}_0)|_{t=t_1}$$

does not vanish.

Note that $B(t_0, t_0) = 0$. If we use a Legendre transformation to change to canonical momenta: $\dot{\mathbf{z}} \to \mathbf{p}$, then condition (4.9) will be equivalent to

$$\partial\delta\mathbf{z}(t_1)/\partial\delta\mathbf{p}(t_0) \neq 0$$

A canonical transformation with this property is said to be free. However, it is well known that not every canonical transformation is free. For example, no point transformation is free. However, an oscillatory motion described by a Hamiltonian system generates a free canonical transformation. This may be verified for the example of normal coordinates for small oscillations of a conservative system about an equilibrium position. The equations in variations then define rotations in the (coordinate, velocity) plane, which generate a free canonical transformation. We may therefore state that this condition is fairly typical in applications.

By (4.9), the system of equations (4.8) has a unique solution $\dot{\mathbf{h}}_0 = \mathbf{0}$. Therefore, the unique solution of system (4.6) satisfying the boundary conditions (2.1) (or $\mathbf{h} \in \overset{\circ}{H}{}^1$) is $\mathbf{h}(t) \equiv \mathbf{0}$, whence it follows that $\Delta\mathbf{h}(t) \equiv \mathbf{0}$, or $\mathbf{h}_1(t) \equiv \mathbf{h}_2(t)$. This means that Eq. (4.) always has a unique solution.

As a trial solution of Eq. (4.4), let us take

$$\mathbf{h}(t) = B(t, t_0)\mathbf{b} + I(t) \tag{4.10}$$

$$I(t) = \int_{t_0}^{t} (A(t, \tau) + B(t, \tau)C_2(\tau))(\mathbf{G}(\mathbf{g}))(\tau)d\tau$$

where $A(t, \tau)$ is the upper left block of the matrix $R(t, t_0)$. As will soon be clear, the function (4.10) is a solution of an equation of the form

$$\dot{\mathbf{h}}(t) = \int_{t_0}^{t} (C_1(\tau)\mathbf{h}(\tau) + C_2(\tau)\dot{\mathbf{h}}(\tau))d\tau + (\mathbf{G}(\mathbf{g}))(t) + \mathbf{b}(\mathbf{g}) \tag{4.11}$$

Here

$$\begin{aligned} C_1(t) &= a_1^{-1}(t)a_4(t) + (a_1^{-1}(t))^{\cdot} a_2(t) - (a_1^{-1}(t)a_2(t))^{\cdot} \\ C_2(t) &= a_1^{-1}(t)a_3(t) + (a_1^{-1}(t))^{\cdot} a_1(t) - a_1^{-1}(t)a_2(t) \\ (\mathbf{G}(\mathbf{g}))(t) &= a_1^{-1}(t)(a_1(t)\mathbf{g}(t)) - \int_{t_0}^{t} (a_1^{-1}(\tau))\,(a_1(\tau)\mathbf{g}(\tau))^{\cdot} d\tau \\ \mathbf{b}(\mathbf{g}) &= B^{-1}(t_1, t_0)I(t_1) \equiv \text{const} \end{aligned} \tag{4.12}$$

It is obvious from (4.10) that, since $B(t_0, t_0) = 0$, we have $\mathbf{h}(t_0) = \mathbf{0}$. The last relation in (4.12) also shows that $\mathbf{h}(t_1) = \mathbf{0}$, i.e. $\mathbf{h} \in \overset{\circ}{H}{}^1$. One can now deduce from the properties of the equation in variations that differentiation of Eq. (4.10) with respect to t gives

$$\dot{\mathbf{h}}(t) = D(t,t_0)\mathbf{b}(\mathbf{g}) + (\mathbf{G}(\mathbf{g}))(t) + \int_{t_0}^{t} (G(t,\tau) + D(t,\tau)C_2(\tau))(\mathbf{G}(\mathbf{g}))(\tau)d\tau \tag{4.13}$$

where $C(t, \tau)$, $D(t, \tau)$ are the lower blocks of the matrix $R(t, t_0)$.

It can be verified that, expanding (4.13), one obtains an identity on the left of (4.11). As done for (4.4), one now verifies that Eq. (4.11) has a unique solution $\mathbf{h} \in \mathring{H}^1$.

It turns out that this function $\mathbf{h}(t)$ is also a solution of Eq. (4.4).

The proof is as follows: By construction, the quantity $\mathbf{b}_1(\mathbf{h})$ in Eq. (4.4), considered for $\mathbf{h} \in \mathring{H}^1$, is the mean of the function $(\Lambda(\mathbf{h}))(t)$, since we have assumed that $\mathbf{g} \in \mathring{H}^1$, so that the mean of $(a_1\mathbf{g})^{\cdot}$ is zero.

We will now try to transform Eq. (4.11) to the form (4.4). We transform the constant vector as follows:

$$-\mathbf{b} = -a_1^{-1}(t)\mathbf{b}_2 + \int_{t_0}^{t} (a_1^{-1}(\tau))^{\cdot}\,\mathbf{b}_2 d\tau \quad (\mathbf{b}_2 = a_1(t_0)\mathbf{b}) \tag{4.14}$$

If we use the representation (4.12) of the functions $C_1(t)$ and $C_2(t)$ and suitably regroup the terms, Eq. (4.11) can always be expressed—after premultiplication by $a_1(t)$, (since $a_1(t)$, the positive-definite symmetric kinetic energy matrix, is always non-singular when evaluated for the solution being approximated)—as a Fredholm equation

$$\mathbf{v}(t) - (\mathbf{K}\mathbf{v})(t) = \mathbf{0} \tag{4.15}$$

$$(\mathbf{K}\mathbf{v})(t) = \int_{t_0}^{t_1} \kappa(t,\tau)\mathbf{v}(\tau)d\tau, \quad \kappa(t,\tau) =: \begin{cases} a_1(t)(a_1^{-1}(\tau))^{\cdot} & (\tau < t) \\ 0 & (\tau \geqslant t) \end{cases}$$

where the unknown function is

$$\mathbf{v}(t) = (\Lambda(\mathbf{h}))(t) - \mathbf{b}_2 - (a_1(t)\mathbf{g}(t))^{\cdot} \tag{4.16}$$

To permit further investigation, we require that the integral operator $\mathbf{I} - \mathbf{K}$: $L_2 \to L_2$ be invertible. This is an operator of Volterra type whose kernel is sufficiently regular everywhere off the straight line $t = \tau$.

As we know, the sufficient condition for the operator $\mathbf{I} - \mathbf{K}$ to be invertible is that the norm of $\mathbf{K}$ should be less than unity. Then the inverse operator exists, is continuous and can be expanded in series: $(\mathbf{I} - \mathbf{K})^{-1} = \mathbf{I} + \mathbf{K} + \mathbf{K}^2 + \dots$.

If the kinetic energy matrix varies weakly along a solution, then the derivative $(a_1^{-1}(t))^{\cdot}$ is sufficiently small and the invertibility condition will hold. In the simplest case one can consider the situation in which $a_1(t) = \text{const}$. Then $\mathbf{K} = \mathbf{0}$ and the invertibility of the identity operator is trivial.

Thus, let $\mathbf{I} - \mathbf{K}$ be invertible. Then the unique solution of Eq. (4.15) in L_2 will be the function $\mathbf{v}(t) \equiv \mathbf{0}$, i.e. in $[t_0, t_1]$. But this means that, for given $\mathbf{g} \in \mathring{H}^1$, with the function $\mathbf{h} \in \mathring{H}^1$ calculated from formula (4.10) and the vector $\mathbf{b}_2$ found by (4.14), Eq. (4.4) is satisfied.

It remains to verify that $\mathbf{b}_2 = \mathbf{b}_1(\mathbf{h}) = \mathbf{AF}'(\mathbf{y})\mathbf{h}$. Indeed, by definition $\mathbf{b}_1$ is the mean of $(\Lambda(\mathbf{h}))(t)$ over the interval $[t_0, t_1]$. On the other hand, taking a given element $\mathbf{g} \in \mathring{H}^1$, we conclude that $(a_1(t)\mathbf{g}(t))^{\cdot}$ has zero mean. Therefore, as the same is true of $\mathbf{v}(t) \equiv \mathbf{0}$, the vector $\mathbf{b}_2$ will also be the mean of the same function $\Lambda(\mathbf{h}))(t)$, where $\mathbf{h}$ is the solution of Eq. (4.4) given by (4.10). Therefore $\mathbf{b}_2 = \mathbf{b}_1$.

Thus, we have finally established that the operator (4.1) is invertible. We shall now prove, using formula (4.10), that the inverse operator is continuous. Since the norm in $\mathring{H}^1$ is the L_2-norm for the derivative, it follows that in fact we need formula (4.11). Estimating the norm for each term in $(\mathbf{I} - \mathbf{Z}'(\mathbf{y}))^{-1}\mathbf{g}$, we obtain the condition for continuity: $\| (\mathbf{I} - \mathbf{Z}'(\mathbf{y}))^{-1}\mathbf{g} \|^0 \leqslant \mathbf{const} \| \mathbf{g} \|^0$.

Lemma. Suppose that the following non-degeneracy condition holds along the solution $\mathbf{y}(t)$ being approximated, at $t = t_1$

$$\frac{\partial \delta \mathbf{z}}{\partial \delta \dot{\mathbf{z}}_0}(t_1) \neq 0 \tag{4.17}$$

where $(\delta\mathbf{z}_0, \delta\dot{\mathbf{z}}_0)^{\mathrm{T}} \to (\delta\mathbf{z}(t), \delta\dot{\mathbf{z}}(t))$ is the tangent mapping along the trajectory. Suppose also that a unique solution $\mathbf{v} \in L_2$ of the following linear integral equation exists

$$\mathbf{v}(t) - \int_{t_0}^{t} a(t,\mathbf{y}(t))(a^{-1}(\tau,\mathbf{y}(\tau)))^{\cdot}\,\mathbf{v}(\tau)d\tau = \mathbf{0} \tag{4.18}$$

Then the Fréchet derivative operator $\mathbf{I} - \mathbf{Z}'(\mathbf{y})\colon \mathring{H}^1 \to \mathring{H}^1$ at the solution $\mathbf{y}(t)$ is continuously invertible.

5. APPROXIMATION CONDITIONS

We will show that the following approximation conditions hold at solutions of Eq. (2.12)

$$\|\mathbf{y} - P_m\mathbf{y}\|^0 \to 0, \quad \|P_m\mathbf{Z}(P_m\mathbf{y}) - \mathbf{Z}(\mathbf{y})\|^0 \to 0, \quad \|P_m\mathbf{Z}'(P_m\mathbf{y}) - \mathbf{Z}'(\mathbf{y})\| \to 0$$

as $m \to \infty$.

These conditions are verified by a method analogous to that described in [1, 2]. The only difference lies in the technique used to prove that the operator $\mathbf{Z}'(\mathbf{z})\colon \mathring{H}^1 \to \mathring{H}^1$ is compact at a solution $\mathbf{z} = \mathbf{y}$.

Let $\{\mathbf{h}_k\}_{k=1}^{\infty}$ be an infinite system of functions in $\mathring{H}^1$ which is uniformly bounded: $\|\mathbf{h}_k\|^0 \leq c$. It can be shown that in the uniform metric of the space $C = C([t_0, t_1], \mathbf{R}^n)$ these functions are uniformly bounded and equicontinuous. By Arzela's theorem, one can extract a subsequence $\{\mathbf{h}_{mk}\}_{k=1}^{\infty}$ which is convergent in the metric of C.

As shown in Section 3, the operator $\mathbf{I} - \mathbf{Z}'(\mathbf{y})\colon \mathbf{h} \to \mathbf{g}$ is defined by Eq. (4.4) relating the functions $\mathbf{h}, \mathbf{g} \in \mathring{H}^1$. Putting $\mathbf{f} = \mathbf{h} - \mathbf{g} \in \mathring{H}^1$, we obtain an integrodifferential equation for $\mathbf{f}$

$$\mathbf{f} = \mathbf{h} - (\mathbf{I} - \mathbf{Z}'(\mathbf{y}))\mathbf{h} = \mathbf{Z}'(\mathbf{y})\mathbf{h} \tag{5.1}$$

That the functions $\mathbf{h}_k$ are equicontinuous has already been verified. An analogous proof establishes the same for the functions $\mathbf{g}_k$. It suffices to consider the different types of terms occurring on the right of (5.1) and to verify that they are equicontinuous. Using (5.1), one can also verify that $\dot{\mathbf{f}}(t)$ is continuous. Thus, by Arzela's theorem, the sequence of continuous functions $\dot{\mathbf{f}}(t)$ contains a convergent subsequence $\dot{\mathbf{f}}_{mk}(t)$. Hence one can conclude that the elements $\mathbf{Z}'(\mathbf{y})\mathbf{h}_{mk} \in \mathring{H}^1$ converge, since

$$\|\mathbf{Z}'(\mathbf{y})\mathbf{h}_{m_k} - \mathbf{Z}'(\mathbf{y})\mathbf{h}_{m_l}\|^0 = \left(\int_{t_0}^{t_1} \left(\|\dot{\mathbf{f}}_{m_k}(t) - \dot{\mathbf{f}}_{m_l}(t)\|^0\right)^2 dt\right)^{1/2} \leq$$
$$\leq (t_1 - t_0)^{1/2} \sup_t \|\dot{\mathbf{f}}_{m_k}(t) - \dot{\mathbf{f}}_{m_l}(t)\|$$

We have thus proved that the operator $\mathbf{Z}'(\mathbf{z})\colon \mathring{H}^1 \to \mathring{H}^1$ is compact. However, in this case, unlike that of [1, 2], $\mathbf{y}$ is not any point of the domain $\Omega \subset \mathring{H}^1$ but the solution of Eq. (2.12) being approximated.

To sum up, we have proved the following theorem.

Theorem. Suppose that the boundary-value problem for Eq. (1.3) is uniquely solvable in the interval $[t_0, t_1]$ and let $\mathbf{q}(t)$ be the solution. Suppose that the kinetic energy function $T(t, \mathbf{q}, \dot{\mathbf{q}})$ and the function of generalized forces $Q(t, \mathbf{q}, \dot{\mathbf{q}})$ satisfy the conditions formulated in Section 1. If $\mathbf{q}^0(t)$ is a path in the configuration space connecting points $\mathbf{q}_0$ and $\mathbf{q}_1$ in time $t_1 - t_0$, then Eq. (1.3) for the function $\mathbf{q}(t)$ is equivalent to Eq. (2.12) for a function $\mathbf{z}(t) = \mathbf{q}(t) - \mathbf{q}^0(t)$ satisfying condition (2.1).

In addition, if the non-degeneracy condition (4.17) holds for the solution $\mathbf{z}(t)$ being approximated, and the integral equation (4.18) is uniquely solvable, then $\varepsilon > 0$ and an integer N exist such that, for any $m > N$, the equation

$$\mathbf{z}_m = P_m\mathbf{Z}(z_m) \quad (\mathbf{z}_m \in E_m) \tag{5.2}$$

has a unique solution $\mathbf{z}_m$ in the sphere $\|\mathbf{z}' - \mathbf{z}\|^0 \leq \varepsilon$, the following estimate holds

$$\|\mathbf{z}_m - \mathbf{z}\|^0 \leq \|\mathbf{z} - P_m\mathbf{z}\|^0 + \|\mathbf{z}_m - P_m\mathbf{z}\|^0 \to 0 \quad (m \to \infty)$$

and there are constants σ_1 and σ_2 such that

$$c_1\Delta_m \leq \|\mathbf{z}_m - P_m\mathbf{z}\|^0 \leq c_2\Delta_m \quad (\Delta_m = \|P_m\mathbf{Z}(P_m\mathbf{y}) - \mathbf{Z}(\mathbf{y})\|^0)$$

Here, again, as in [1, 2], we have used the appropriate result from [6].

6. DISCUSSION OF THE RESULTS

In conclusion, we note that the theorem is applicable in numerous problems of dynamics. It is clear from the assumptions of the theorem that approximations may be constructed even for motions that involve impact interactions. In such cases, as is well known, the generalized velocities may experience discontinuities as functions of time—a situation that is covered by the metric of mean-square approximation (in the tangent space).

We now consider the technology of computing motions that satisfy the boundary conditions in the configuration space. The main problem in making proper use of the equations of motion is to determine initial velocities such that the motion will surely reach the given terminal point. The "shooting" method is standard in such situations. The exact solution is determined by an iterative process in a finite-dimensional space. Each step, however, requires the computation of an entire trajectory for $t \in [t_0, t_1]$ in the conditions for extremality of Hamilton's principle. Here, however, Hamilton's principle is used to formulate the extremality conditions in such a way that Galerkin's method is applicable in a suitable functional path space.

REFERENCES

1. KOSENKO I. I., The use of Chebyshev polynomials to construct the trajectory of perturbed motion in non-linear mechanics. *Prikl. Mat. Mekh.* **55**, 1, 32–38, 1991.
2. KOSENKO I. I., Galerkin's method in non-linear mechanics. *Dokl. Ross. Akad. Nauk* **335**, 5, 586–588, 1994.
3. TRENOGIN V. A., *Functional Analysis*. Nauka, Moscow, 1980.
4. SCHWARTZ L., *Analyse*, Vol. I. Hermann, Paris, 1967.
5. FILLIPOV A. F., *Differential Equations with Discontinuous Right-hand Side*. Nauka, Moscow, 1985.
6. KRASNOSEL'SKII M. A., VAINIKKO G. M., ZABREIKO P. P., RUTITSKII Ya. B. and STETSENKO V. Ya., *The Approximate Solution of Operator Equations*. Nauka, Moscow, 1969.

Translated by D.L.