Pergamon

0021-8928(95)00002-X

J. Appl. Maths Mechs, Vol. 59, No. 1, pp. 9–18, 1995

0021-8928/95 \$24.00+0.00

THE APPLICATION OF GALERKIN'S METHOD IN LAGRANGIAN DYNAMICS†

I. I. KOSENKO

Sergiyev Posad, Moscow Region

(*Received* 20 *January* 1994)

A procedure is proposed for approximating solutions of the Cauchy problem in Lagrangian mechanics. The technique is based on a certain version of Galerkin's method. It involves transforming the dynamical system to canonical form. The approximation algorithm is shown to be equivalent to procedures based on variational formulations of the Cauchy problem in Lagrangian dynamics. A procedure for the approximate construction of a solution of Mathieu's equation is considered as an example.

1. REDUCTION TO A HAMILTONIAN SYSTEM

The technique proposed in [1] is equivalent to the widely used algorithms for solving Cauchy problems in holonomic mechanics by using variational principles. Suppose that in a time interval $[t_0, t_1]$ it is required to approximate the motion of a holonomic mechanical system with generalized coordinates $q_1, q_2, \ldots, q_n$ and kinetic energy

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

where T_i $(i = 0, 1, 2)$ are homogeneous functions of the velocities with T_2 positive definite. Let $\mathbf{Q}(t, \mathbf{q}, \dot{\mathbf{q}}) \in \mathbf{R}^n$ be a vector of generalized forces.

In the Cauchy problem one considers the equations of motion

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{\mathbf{q}}} - \frac{\partial T}{\partial \mathbf{q}} = \mathbf{Q} \tag{1.1}$$

together with initial conditions

$$\mathbf{q}(t_0) = \mathbf{q}^0, \quad \dot{\mathbf{q}}(t_0) = \dot{\mathbf{q}}^0 \tag{1.2}$$

Assuming that the functions $T(t, \mathbf{q}, \dot{\mathbf{q}})$, $\mathbf{Q}(t, \mathbf{q}, \dot{\mathbf{q}})$ are sufficiently smooth, we can state that the solution $\mathbf{q}(t)$ belongs to the Sobolev space $H^2([t_0, t_1], \mathbf{R}^n)$ (henceforth denoted simply by H_n^2) of vector-valued functions that are square integrable together with all their derivatives up to and including the second order. By (1.2), the space H_n^2 may be replaced by the smaller affine subspace H_{n0}^2 of functions that satisfy these conditions. The corresponding tangent space (the space of variations) $T_\mathbf{q}H_{n0}^2 = \mathring{H}_n^2$ is the set of all functions $\delta\mathbf{q} \in H_n^2$ that satisfy the initial conditions

$$\delta\mathbf{q}(t_0) = \mathbf{0}, \quad \delta\dot{\mathbf{q}}(t_0) = \mathbf{0}$$

We can now formulate the Cauchy problem for the system of equations (1.1) as follows. Assume that the conditions of the uniqueness theorem for (1.1) hold throughout $[t_0, t_1]$. Then a solution $\mathbf{q} \in H_n^2$ satisfying conditions (1.2) will be a function $\mathbf{q} \in H_{n0}^2$. The elements of H_{n0}^2 may be expressed as $\mathbf{q}(t) = \mathbf{q}^0 + \mathbf{q}^1(t)$, where $\mathbf{q}^1 \in \mathring{H}_n^2$. If $\mathbf{q}(t)$ is a solution of the Cauchy problem, then, by uniqueness, $\mathbf{q}(t) + \delta\mathbf{q}(t)$ $(\delta\mathbf{q} \in \mathring{H}_n^2)$ will also be a solution if and only if $\delta\mathbf{q}(t) \equiv \mathbf{0}$. This condition may be replaced by the integral form of D'Alembert's principle as the equation of work in the interval $[t_0, t_1]$

$$\int_{t_0}^{t_1} (\langle T_{\dot{\mathbf{q}}}, \delta\dot{\mathbf{q}}\rangle + \langle T_\mathbf{q} + \mathbf{Q}, \delta\mathbf{q}\rangle)dt - \langle T_{\dot{\mathbf{q}}}, \delta\mathbf{q}\rangle\big|_{t_1} = 0 \quad (\forall \delta\mathbf{q} \in \mathring{H}_n^2) \tag{1.3}$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in Euclidean space $\mathbf{R}^n$.

†*Prikl. Mat. Mekh.* Vol. 59, No. 1, pp. 10–20, 1995.

Equation (1.3) is interpreted as Hamilton's principle [2–4]. To apply the technique proposed here, we turn to Hamiltonian dynamics, as defined by the variables $\mathbf{q}$, $\mathbf{p} = T_{\dot{\mathbf{q}}}$. If $\mathbf{q} \in H^2_n$, then $\dot{\mathbf{q}} \in H^1{}_n = H^1([t_0, t_1], \mathbf{R}^n)$. Therefore, when we change to the vector of canonical variables $\mathbf{z} = (\mathbf{q}, \mathbf{p})^T$, we must take $H^1_{2n} = H^1([t_0, t_1], \mathbf{R}^{2n})$ as a functional space of trajectories. The Cauchy problem

$$\mathbf{q}(t_0) = \mathbf{q}^0, \quad \dot{\mathbf{p}}(t_0) = \dot{\mathbf{p}}^0 = T_{\dot{\mathbf{q}}}(t, \mathbf{q}^0, \dot{\mathbf{q}}^0)$$

reduces H^1_{2n} to the affine subspace $H^1_{2n,0}$ consisting of all functions $\mathbf{z}(t)$ such that $\mathbf{z}(t_0) = \mathbf{z}^0 = (\mathbf{q}^0, \mathbf{p}^0)^T$. The tangent space to $H^1_{2n,0}$ at a point $\mathbf{z} \in H^1_{2n,0}$ is the linear subspace $\mathring{H}^1_{2n}$ of functions $\delta\mathbf{z}(t)$ such that $\delta\mathbf{z}(t_0) = \mathbf{0}$.

Since

$$T(t, \mathbf{q}, \dot{\mathbf{q}}) = \langle \mathbf{p}(t, \mathbf{q}, \dot{\mathbf{q}}), \dot{\mathbf{q}} \rangle - K(t, \mathbf{q}, \mathbf{p}(t, \mathbf{q}, \dot{\mathbf{q}}))$$

it follows that

$$\delta T = \langle T_{\dot{\mathbf{q}}}, \delta\dot{\mathbf{q}} \rangle + \langle T_{\mathbf{q}}, \delta\mathbf{q} \rangle = \delta\langle \mathbf{p}(t, \mathbf{q}, \dot{\mathbf{q}}), \dot{\mathbf{q}} \rangle - \langle K_{\mathbf{q}}, \delta\mathbf{q} \rangle - \langle K_{\mathbf{p}}, \delta p \rangle$$

Thus

$$\int_{t_0}^{t_1} (\langle \dot{\mathbf{q}} - K_{\mathbf{p}},\ \delta\mathbf{p} \rangle + \langle -\dot{\mathbf{p}} - K_{\mathbf{q}} + \mathbf{Q},\ \delta\mathbf{q} \rangle) = dt = 0 \quad (\forall\ \delta\mathbf{z} = (\delta\mathbf{q},\ \delta\mathbf{p})^{\mathrm{T}} \in \mathring{H}^1_{2n})$$

Let us retain the notation $\langle \cdot, \cdot \rangle$ for the scalar product in $\mathbf{R}^{2n}$. Then Eq. (1.3) may be written in the equivalent form

$$\int_{t_0}^{t_1} \langle -J\dot{\mathbf{z}} - K_{\mathbf{z}} + \mathbf{F}, \delta\mathbf{z} \rangle dt = 0 \quad (\forall \delta\mathbf{z} \in \mathring{H}^1_{2n}) \tag{1.4}$$

where J is a symplectic matrix in $\mathbf{R}^{2n}$, $K_{\mathbf{z}} = (K_{\mathbf{q}}, K_{\mathbf{p}})^{\mathrm{T}}$, $\mathbf{F} = (\mathbf{P}, \mathbf{0})^{\mathrm{T}}$.

Since $\mathbf{z} \in H^1_{2n}$, it follows that $\dot{\mathbf{z}} \in L_2 = L_2([t_0, t_1], \mathbf{R}^{2n})$. Thus the first factor in (1.4) is $-J\dot{\mathbf{z}} - K_{\mathbf{z}} + \mathbf{F} \in L_2$. Next, we know that H^1_{2n} is continuously embeddable in L_2 and dense there. Since $\mathring{H}^1_{2n}$ is dense in the metric of L_2 in H^1_{2n}, it follows that it is also dense in L_2. The scalar product of L_2 is continuous there with respect to either of its arguments. Equation (1.4) holds for any $\delta\mathbf{z} \in L_2$, and so almost everywhere in $[t_0, t_1]$ the system of Hamiltonian equations

$$J\dot{\mathbf{z}} = -K_{\mathbf{z}} + \mathbf{F} \tag{1.5}$$

is satisfied.

Moreover, it is obvious from (1.5) that the function on the right is continuous with respect to $t \in [t_0, t_1]$ (since T and $\mathbf{P}$ are assumed to be sufficiently smooth in their arguments). Hence this equation holds everywhere in $[t_0, t_1]$ for solutions of the Cauchy problem for the Hamiltonian system (1.5). Multiplying both sides of (1.5) by the symplectic matrix $-J$, we obtain a system of equations in Cauchy form

$$\dot{\mathbf{z}} = \mathbf{Z}(t, \mathbf{z}) \quad (\mathbf{Z}(t, \mathbf{z}) = JK_{\mathbf{z}}(t, \mathbf{z}) + (\mathbf{0}, \mathbf{P}(t, \mathbf{z}))^{\mathrm{T}}) \tag{1.6}$$

2. THE PROBLEM OF CHOOSING A BASIS

We can now apply the results of [1, 5]. Fixing a natural number m, we will seek a solution of system (1.6) in the form

$$\mathbf{z}(t) = \sum_{k=1}^{m} \mathbf{z}_k \left[\cos k\pi \frac{t - t_0}{t_1 - t_0} - 1 \right]$$

and then apply the ideas outlined in [5]. However, the reduction carried out in [1] will enable us to use the metric of L_2 in the space of derivatives $\mathbf{y} = \dot{\mathbf{z}}$.

As in [1], we consider the space CA of absolutely continuous functions in $[t_0, t_1]$ with values in $\mathbf{R}^{2n}$.

Now let

$$D\colon \mathrm{CA} \to L_1 = L_1([t_0, t_1], \mathbf{R}^{2n})$$

be the operator of differentiation with respect to t: $(D\mathbf{z})(t) = \dot{\mathbf{z}}(t)$. We know [1] that D is surjective, since the antiderivative of any summable function is absolutely continuous. The space CA can be split into a direct sum $\mathrm{CA} \simeq \mathbf{R}^{2n} + \mathrm{CA}^0$, where CA^0 is the set of all absolutely continuous functions $\mathbf{z}(t)$ such that $\mathbf{z}(t_0) = 0$.

We also know [1] that the restrictions

$$D\colon \mathrm{CA}^0 \to L_1, \quad D\colon \mathbf{z}^0 + \mathrm{CA}^0 \to L_1 \ (\mathbf{z}^0 \in \mathbf{R}^n)$$

are homoeomorphisms. Finally, the Cauchy problem for the differential equation (1.6) may be reduced to a functional equation $\mathbf{y} = T(\mathbf{y})$ in L_1 (or in L_2), where T is the non-linear operator defined by

$$(T(\mathbf{y}))(t) = \mathbf{Z}(t,(D^{-1}\mathbf{y})(t)), \quad (D^{-1}\mathbf{y})(t) = \mathbf{z}^0 + \int_{t_0}^{t_1} \mathbf{y}(\tau)d\tau$$

Then the required approximation may be written in terms of undetermined coefficients as

$$\mathbf{y}(t) = \sum_{k=1}^{m} \mathbf{y}_k \sin k\pi \frac{t-t_0}{t_1-t_0} \tag{2.1}$$

Denoting the corresponding projection operator by P_m, we can derive the system of Galerkin equations of the form

$$\mathbf{y} = P_m T(\mathbf{y}) \quad (\mathbf{y} \in P_m L_2, \ m \in \mathbf{N})$$

and moreover the group of $2n$ equations corresponding to the kth harmonic in (2.1) may be obtained by substituting the function

$$\delta\mathbf{z} = \mathbf{e}_j \sin k\pi \frac{t-t_0}{t_1-t_0} \quad (j = 1,2,\ldots,2n)$$

into (1.4).

Thus, the procedure considered in [2–4] can be transformed correctly to Hamiltonian form, and only then should one apply Galerkin's method.

In fact, the examples of dynamical systems considered in [2–4, 6, 7] are such that the kinematic energy has constant coefficients, independent of $\mathbf{q}$. In such cases the Legendre transformation becomes a trivial operation, which eliminates many problems that arise in the implementation of projection algorithms. The basis functions used in [6, 7] are the Legendre polynomials.

Thus, let us assume that

$$T_2(t,\mathbf{q},\dot{\mathbf{q}}) = \tfrac{1}{2}\langle A\dot{\mathbf{q}},\dot{\mathbf{q}}\rangle, \quad T_1(t,\mathbf{q},\dot{\mathbf{q}}) = \langle \mathbf{a},\dot{\mathbf{q}}\rangle, \quad T_0(t,\mathbf{q}) \equiv 0$$

The constant quantity T_0 maybe omitted. The elements of the positive definite symmetric matrix A and the components of the vector $\mathbf{a}$ are constants.

We shall show that application of the projection method for Eq. (1.1) with condition (1.2) is equivalent to using the projection method for the equation

$$\mathbf{y} = JK_{\mathbf{z}}(\mathbf{y}) + (\mathbf{0},\mathbf{P}(\mathbf{y}))^{\mathrm{T}}, \quad \mathbf{y} = (\dot{\mathbf{q}},\dot{\mathbf{p}})^{\mathrm{T}} \tag{2.2}$$

$$JK_{\mathbf{z}}(\mathbf{y}) = (B\mathbf{p}(\mathbf{y}) + \mathbf{b}, 0)^{\mathrm{T}}, \quad B = A^{-1}, \quad \mathbf{b} = -A^{-1}\mathbf{a}$$

$$\mathbf{P}(\mathbf{y})(t) = \mathbf{P}(t,\mathbf{q}(\mathbf{y})(t), \ \mathbf{p}(\mathbf{y})(t))$$

$$\mathbf{q}(\mathbf{y})(t) = \mathbf{q}^0 + \int_{t_0}^{t} \dot{\mathbf{q}}(\tau)d\tau, \quad \mathbf{p}(\mathbf{y})(t) = \mathbf{p}^0 + \int_{t_0}^{t} \dot{\mathbf{p}}(\tau)d\tau$$

These relations are obtained by using the Legendre transformation $\mathbf{p} = A\dot{\mathbf{q}} + \mathbf{a}$, which yields $\dot{\mathbf{q}} =$

$A^{-1}(\mathbf{p}-\mathbf{a})$. The kinetic energy in the Hamiltonian system is given by the formula

$$K(\mathbf{p}) = \langle \mathbf{p}, \dot{\mathbf{q}} \rangle - T(\dot{\mathbf{q}}) = \tfrac{1}{2}\langle B\mathbf{p}, \mathbf{p} \rangle + \langle \mathbf{b}, \mathbf{p} \rangle$$

By virtue of the initial conditions, application of Galerkin's method to the work equation (1.3) is equivalent to its application to the system of Lagrange equations.

Let us consider the problem of choosing the basis functions. It turns out that the equivalence of the Galerkin equations for the Hamiltonian system, on the one hand, and the Lagrangian system, on the other, hinges on this choice.

In fact, let us first consider the trigonometric functions. The derivative of the phase variable $\mathbf{z} = (\mathbf{q}, \mathbf{p})^{\mathrm{T}}$ may be expanded relative to the basis of L_2 consisting of the functions

$$\left\{ \mathbf{e}_j \sin k\pi \frac{t-t_0}{t_1-t_0} \right\} \quad (j = 1, 2, \ldots, 2n;\ k = 1, 2, \ldots) \tag{2.3}$$

The antiderivative can then be expanded uniquely in terms of the basis functions

$$\left\{ \dot{\mathbf{e}}_j \cos k\pi \frac{t-t_0}{t_1-t_0} \right\} \quad (j = 1, 2, \ldots, 2n;\ k = 0, 1, \ldots) \tag{2.4}$$

Suppose that the system of Galerkin equations (2.2) corresponding to the Cauchy problem for the canonical equations may be written as

$$\dot{\mathbf{z}}(t) = \sum_{j=1}^{2n} \sum_{k=1}^{m} z_{jk} \sin k\pi \frac{t-t_0}{t_1-t_0} \mathbf{e}_j$$

In that case a uniform approximation of the solution of system (1.6) is

$$\mathbf{z}(t) = \sum_{j=1}^{2n} \sum_{k=0}^{m} Z_{jk} \cos k\pi \frac{t-t_0}{t_1-t_0} \mathbf{e}_j,$$

$$Z_{jk} = \frac{t_1-t_0}{\pi k} z_{jk} \quad (k = 1, 2, \ldots, m)$$

$$Z_{j0} = z_j(t_0) - \frac{t_1-t_0}{\pi} \sum_{k=1}^{m} \frac{z_{jk}}{k}$$

If the approximations of the coordinates and momenta are considered separately, we have

$$\mathbf{q}(t) = \sum_{j=1}^{n} \sum_{k=0}^{m} Q_{jk} \cos k\pi \frac{t-t_0}{t_1-t_0} \mathbf{e}_j,$$

$$\mathbf{p}(t) = \sum_{j=1}^{n} \sum_{k=0}^{m} P_{jk} \cos k\pi \frac{t-t_0}{t_1-t_0} \mathbf{e}_j$$

$$Q_{jk} = Z_{jk}, \quad P_{jk} = Z_{n+j,k} \quad (j = 1, 2, \ldots, n)$$

where the basis vectors in $\mathbf{R}^n$ are denoted by the same symbols as those of $\mathbf{R}^{2n}$: $\mathbf{e}_j$.

On the other hand, following the techniques of [2–4, 6, 7], suppose that the Galerkin equations have been set up for the system of Lagrange equations and an approximation has been found for the solution of the Cauchy problem. In accordance with the preceding discussion, we use the basis (2.4). A direct check will then verify that the trigonometric functions do not yield a convenient basis for Galerkin approximation of the solution of the Cauchy problem for the second-order Lagrange equations.

If one uses the trigonometric basis, there may be difficulties in satisfying the initial conditions; secular terms may also appear when the operator D^{-1} is applied. It should nevertheless be noted that the use of technique of [1, 5] produces results in terms of the trigonometric basis, provided that the dynamics of the system are represented in the Hamiltonian formulation (or by the Lagrange equations transformed in some way to Cauchy form).

It turns out that the equivalence of the techniques of [2–4, 6, 7] and that of the present paper may be established in terms of bases that use the classical orthogonal polynomials. Let $\{p_k(\tau)\}_{k=0}^{\infty}$ be one such system, defined in the interval $[-1, 1]$. It is convenient to define a transformation of the independent variable $t \to \tau$, say, by the formula

$$\tau = -2\frac{t-t_0}{t_1-t_0}+1$$

One approach to the use of Chebyshev polynomials in Galerkin's method may be found in [5]. Henceforth, in order to avoid the cumbersome notation due to the above transformation, we shall assume from the start that the dynamical system in question is defined for $t \in [-1, 1]$ and that the initial data are defined at $t = 1$: $\mathbf{q}(1) = \mathbf{q}^0$, $\dot{\mathbf{q}}(1) = \dot{\mathbf{q}}^0$ or $\mathbf{q}(1) = \mathbf{q}^0$, $\mathbf{p}(1) = \mathbf{p}^0$ (for a Hamiltonian system).

3. EQUIVALENCE

Thus, let $\{\varphi_k(t)\}_{k=0}^{\infty}$ be an orthogonal basis in $L_2([-1, 1], \alpha)$, whose elements are square integrable in Lebesgue's sense with weight function $\alpha(t) \geqslant 0$. The corresponding Radon measure $d\mu_1 = \alpha(t)dt$ is finite if the functions $\varphi_k(t)$ are classical orthogonal polynomials. It is known [8] that each of the systems $\psi_k(t) = \dot{\varphi}_{k+1}(t)$, $\chi_k(t) = \dot{\psi}_{k+1}(t)$ $(k = 0, 1, \ldots)$ forms an orthogonal basis in the spaces $L_2([-1, 1], \beta)$, $L_2([-1, 1], \gamma)$, respectively. The weight functions $\beta(t)$, $\gamma(t) \geqslant 0$ for the derivatives may be obtained using standard rules [8]. Thus, the elements of the space H_n^2 will now be functions defined in $[-1, 1]$ whose second derivatives are square integrable with weight γ and their first derivatives are square integrable with weight β. The functions themselves are square integrable with weight α.

It is clear that the bases are transformed into one another by differentiation. Neither do any problems arise when applying the operator D^{-1}. In fact, suppose we have an approximation of generalized accelerations

$$\ddot{\mathbf{q}}(t) = \sum_{k=0}^{m} \ddot{\mathbf{q}}_k \chi_k(t)$$

Then a uniform approximation of the generalized velocities may be obtained by the formula

$$\dot{\mathbf{q}}(t) = \sum_{k=0}^{m+1} \dot{\mathbf{q}}_k \psi_k(t)$$

$$\dot{\mathbf{q}}_k = \ddot{\mathbf{q}}_{k-1} \quad (k = 1, 2, \ldots, m+1), \quad \dot{\mathbf{q}}_0 = \left[\dot{\mathbf{q}}^0 - \sum_{k=0}^{m} \ddot{\mathbf{q}}_k \psi_{k+1}(1)\right]\psi_0^{-1}(1)$$

Note that the zeroth element of any system of classical orthogonal polynomials is a constant. We can now derive formulae for the uniform approximation of the generalized coordinates

$$\mathbf{q}(t) = \sum_{k=0}^{m+2} \mathbf{q}_k \varphi_k(t)$$

$$\mathbf{q}_k = \dot{\mathbf{q}}_{k-1} \quad (k = 1, 2, \ldots, m+2), \quad \mathbf{q}_0 = \left[\mathbf{q}^0 - \sum_{k=0}^{m+1} \dot{\mathbf{q}}_k \psi_{k+1}(1)\right]\psi_0^{-1}(1)$$

We now return to the Hamiltonian system and form an orthogonal basis in the space of derivatives of the phase variables $\mathbf{z}(t)$

$$\{\mathbf{e}_j \psi_k(t),\ \mathbf{e}_{n+j}\chi_l(t)\} \quad (j = 1, 2, \ldots, n; \quad \mathbf{e}_j, \mathbf{e}_{n+j} \in \mathbf{R}^{2n}; \quad k, l = 0, 1, \ldots)$$

We define the finite-dimensional approximation space E_m as the linear span of the basis functions

$$\{\mathbf{e}_j \psi_k(t),\ \mathbf{e}_{n+j}\chi_l(t)\} \quad (j = 1, 2, \ldots, n; \quad k = 0, 1, \ldots, m+1, \quad l = 0, 1, \ldots, m)$$

Hamilton's equations (1.6) are

$$\dot{\mathbf{q}} = B\mathbf{p} + \mathbf{b}, \quad \dot{\mathbf{p}} = \mathbf{P}(t, \mathbf{q}, \mathbf{p})$$

and the corresponding functional notation in the space of derivatives is

$$\dot{\mathbf{q}} = B(D^{-1}\dot{\mathbf{p}}) + \mathbf{b}, \quad \dot{\mathbf{p}} = \mathbf{P}(t, D^{-1}\dot{\mathbf{q}}, D^{-1}\dot{\mathbf{p}}) \tag{3.1}$$

where the antidifferentiation operator is defined as in (2.2). Galerkin's system of projection equations for (3.1) in E_m may be written as

$$\dot{\mathbf{q}} = B(D^{-1}\dot{\mathbf{p}}) + \mathbf{b}, \quad \dot{\mathbf{p}} = [\mathbf{P}(t, D^{-1}\dot{\mathbf{q}}, D^{-1}\dot{\mathbf{p}})]_m \tag{3.2}$$

where the brackets denote expansion of the vector-valued function $\mathbf{P}(t, D^{-1}\dot{\mathbf{q}}, D^{-1}\dot{\mathbf{p}})$ in terms of the basis $\{\mathbf{e}_j\chi_l(t)\}$ ($\mathbf{e}_j \in \mathbf{R}^n$) and omission of all the terms $\mathbf{e}_j\chi_l(t)$ for $l > m$. As for the first subsystem of Eqs (3.1), the operations on the right do not lead to functions outside E_m. In fact, we have $E_m = E_m^q \times E_m^p$, where E_m^q is the linear span of the functions

$$\{\mathbf{e}_j\psi_k(t)\} \quad (\mathbf{e}_j \in \mathbf{R}^n; \quad k = 0, 1,\dots, m+1)$$

and E_m^p that of the functions

$$\{\mathbf{e}_j\chi_l(t)\} \quad (\mathbf{e}_j \in \mathbf{R}^n; \quad l = 0, 1,\dots, m)$$

Therefore, if $(\dot{\mathbf{q}}, \dot{\mathbf{p}})^{\mathrm{T}} \in E_m$, then $\dot{\mathbf{p}} \in E_m^p$. Furthermore, by the relation $D^{-1}\dot{\mathbf{p}} \in E_m^q$ already proved, and since B and $\mathbf{b}$ are constant matrices, it follows that $B(D^{-1}\dot{\mathbf{p}}) + \mathbf{b} \in E_m^q$.

On the other hand, the Lagrangian equations (1.1) are

$$A\ddot{\mathbf{q}} = \mathbf{Q}(t, \mathbf{q}, \dot{\mathbf{q}})$$

whence we obtain the Galerkin system of equations (in the same space E_m)

$$A\ddot{\mathbf{q}} = [\mathbf{Q}(t, \mathbf{q}, \dot{\mathbf{q}})]_m \tag{3.3}$$

where the brackets have the same meaning as before.

Theorem. The solutions

$$\mathbf{q}(t) = \sum_{k=0}^{m+2} \mathbf{q}_k\varphi_k(t) \tag{3.4}$$

of the system of equations (3.3) satisfying the initial conditions

$$\mathbf{q}(1) = \mathbf{q}^0, \quad \dot{\mathbf{q}}(1) = \dot{\mathbf{q}}^0 \tag{3.5}$$

correspond in one-to-one fashion with the solutions

$$\dot{\mathbf{q}}(t) = \sum_{k=0}^{m+1} \dot{\mathbf{q}}_k\psi_k(t), \quad \dot{\mathbf{p}}(t) = \sum_{k=0}^{m} \dot{\mathbf{p}}_k\chi_k(t) \tag{3.6}$$

of the system of equations (3.2). The correspondence is effected by means of the Legendre transformation.

Proof. One can indeed show that systems (3.2) and (3.3) are equivalent, in the sense that each transforms into the other within E_m. Suppose first that a solution (3.4) satisfies system (3.3) with condition (3.5). In fact, we must consider the combined system of Eqs (3.3) and (3.5).

We shall prove that the functions

$$\dot{\mathbf{q}}(t) = \sum_{k=0}^{m+1} \mathbf{q}_{k+1}\psi_k(t), \quad \dot{\mathbf{p}}(t) = \sum_{k=0}^{m} A\mathbf{q}_{k+2}\chi_k(t)$$

satisfy system (3.2). Construct the antiderivatives

$$\begin{gathered} D^{-1}\dot{\mathbf{q}}(t) = \mathbf{q}^0 + \int_1^t \dot{\mathbf{q}}(\tau)d\tau = \mathbf{q}^0 + \sum_{k=0}^{m+1} \mathbf{q}_{k+1}\varphi_{k+1}(t) - \sum_{k=0}^{m+1} \mathbf{q}_{k+1}\varphi_{k+1}(1) \\ D^{-1}\dot{\mathbf{p}}(t) = \mathbf{p}^0 + \int_1^t \dot{\mathbf{p}}(\tau)d\tau = \mathbf{p}^0 + \sum_{k=0}^{m} A\mathbf{q}_{k+2}\psi_{k+1}(t) - \sum_{k=0}^{m} A\mathbf{q}_{k+2}\psi_{k+1}(1), \quad \mathbf{p}^0 = A\dot{\mathbf{q}}^0 + \mathbf{a} \end{gathered} \tag{3.7}$$

Multiply the second relationship in (3.7) by the matrix $B = A^{-1}$, and then add the vector **b** to both sides. Taking the relationship between **b** and **a** (see (2.2)) into account, we have

$$(B(D^{-1}\dot{\mathbf{p}}))(t)+\mathbf{b}=\dot{\mathbf{q}}^0-\sum_{k=0}^{m+1}\mathbf{q}_{k+1}\psi_k(1)+\sum_{k=0}^{m+1}\mathbf{q}_{k+1}\psi_k(t)$$

The function on the right is equal to $\dot{\mathbf{q}}^0$ at $t = 1$. By (3.5)

$$\mathbf{q}_1\psi_0(1)+\sum_{k=1}^{m+1}\mathbf{q}_{k+1}\psi_k(1)=\dot{\mathbf{q}}^0$$

For the classical orthogonal polynomials, $\psi_0(1) \equiv \psi_0(t) \equiv$ const. Therefore

$$(B(D^{-1}\dot{\mathbf{p}}))(t)+\mathbf{b}=\sum_{k=1}^{m+1}\mathbf{q}_{k+1}\psi_k(t)$$

i.e. the first of Eqs (3.2) holds identically.

We shall now show that the second equation of (3.3) is identical to (3.3). Indeed, as A is a constant matrix, $\dot{\mathbf{p}} = A\ddot{\mathbf{q}}$. By conditions (2.2), the functions $D^{-1}\dot{\mathbf{q}}(t)$, $D^{-1}\dot{\mathbf{p}}(t)$ are identical with $\mathbf{q}(t), A\dot{\mathbf{q}}(t) + \mathbf{a}$, respectively. Therefore, by the definition of the function $\mathbf{P}(t, \mathbf{q}, \mathbf{p}) = \mathbf{Q}(t, \mathbf{q}, B\mathbf{p} + \mathbf{b})$, the series expansion of $\mathbf{Q}(t, \mathbf{q}, \dot{\mathbf{q}})$ in terms of polynomials $\{\chi_k(t)\}_{k=0}^{\infty}$ is identical with the same expansion for the function

$$\begin{aligned}&\mathbf{P}(t,D^{-1}\dot{\mathbf{q}}(t),D^{-1}\dot{\mathbf{p}}(t))=\mathbf{P}(t,\mathbf{q}(t),A\dot{\mathbf{q}}(t)+\mathbf{a})=\\&=\mathbf{Q}(t,\mathbf{q}(t),B(A\dot{\mathbf{q}}(t)+\mathbf{a})+\mathbf{b})=\mathbf{Q}(t,\mathbf{q}(t),\dot{\mathbf{q}}(t))\end{aligned} \tag{3.8}$$

The same is true for segments of these series. Consequently, the right-hand sides of Eqs (3.2) (the second group) and (3.3) are identically equal. Thus the second group of equation in system (3.2) is also satisfied.

Conversely, suppose that solution (3.6) satisfies system (3.2). Then the functions

$$\mathbf{q}(t)=D^{-1}\dot{\mathbf{q}}(t)=\mathbf{q}^0-\sum_{k=1}^{m+2}\dot{\mathbf{q}}_{k-1}\varphi_k(t)+\sum_{k=1}^{m+2}\dot{\mathbf{q}}_{k-1}\varphi_k(t)$$

satisfy the initial condition $\mathbf{q}(1) = \mathbf{q}^0$. Since the first group of equations in (3.2) is satisfied, and the function $D^{-1}\dot{\mathbf{p}}(t)$ satisfies the initial condition $D^{-1}\dot{\mathbf{p}}(1) = A\dot{\mathbf{q}}^0 + \mathbf{a}$, it follows that

$$\dot{\mathbf{q}}(1)=B(D^{-1}\dot{\mathbf{p}}(1))+\mathbf{b}=B(A\dot{\mathbf{q}}^0+\mathbf{a})+\mathbf{b}=\dot{\mathbf{q}}^0$$

Thus, all the initial conditions (3.5) are satisfied.

Again by the Legendre transformation or, what is the same, via the first group of Eqs (3.2), we have $A\mathbf{q} + \mathbf{a} = D^{-1}\mathbf{p}$. Hence $A\ddot{\mathbf{q}} = \dot{\mathbf{p}}$, i.e. the left-hand sides of Eqs (3.2) (the second group) and (3.3) are identical. But the same is true of the right-hand sides of these systems of equations, as follows from the fact that the matrix A and vector **a** are constant and from relationships (3.8). Indeed

$$\mathbf{Q}(t,\mathbf{q},\dot{\mathbf{q}})=\mathbf{Q}(t,D^{-1}\dot{\mathbf{q}},B(D^{-1}\dot{\mathbf{p}})+\mathbf{b})=\mathbf{P}(t,D^{-1}\dot{\mathbf{q}},D^{-1}\dot{\mathbf{p}})$$

Hence the segments of the appropriate series are also identical.

Remark. By the results obtained in [1, 5], this theorem implies the legitimacy of the projection approximations of the examples in [6, 7], where the basis functions used were indeed classical orthogonal polynomials (shifted Legendre polynomials).

4. EXAMPLE

To illustrate the technique used in [1] for approximating solutions of the Lagrange equations, let us consider Mathieu's equation

$$\ddot{q} + (\omega^2 + \varepsilon \cos t)q = 0 \tag{4.1}$$

We have a holonomic system with one degree of freedom and Lagrangian

$$L(t, q, \dot{q}) = 1/2\, \dot{q}^2 - 1/2\, (\omega^2 + \varepsilon \cos t)q^2 \tag{4.2}$$

To fix our ideas, suppose we wish to approximate the solution with initial data $q(0) = q_0$, $\dot{q}(0) = \dot{q}_0$ in the interval $t \in [0, T]$. Transform to a new argument $\tau \in [-1, 1]$ by the formula

$$\tau = -2T^{-1}t + 1 \tag{4.3}$$

We obtain an equation of the form

$$q'' + [\Omega^2 + \mu \cos(a\tau + b)]q = 0 \tag{4.4}$$

where the prime denotes differentiation with respect to τ, $\Omega = \omega T/2$, $\mu = \varepsilon T^2/4$, $a = T/2$, $b = -T/2$. Equation (4.4) describes a dynamical system with Lagrangian

$$\Lambda(t, q, q') = 1/2\,(q')^2 - 1/2\,[\Omega^2 + \mu \cos(a\tau + b)]q^2$$

The desired solution of the Cauchy problem must satisfy the initial conditions

$$q(1) = q_0,\quad q'(1) = q_0' = -T\dot{q}_0/2 \tag{4.5}$$

Equation (1.3) for (4.4) may be written as

$$\int_1^{-1} [\Lambda_{q'}\delta q' + \Lambda_q \delta q]d\tau - \Lambda_{q'}\delta q|_{-1} = 0$$

The technique in question dictates the following procedure. The solution will be approximated using Chebyshev polynomials of the first and second kinds. Working in the space $L_2[-1, 1]$, let us approximate the function $q''(\tau)$ by expanding it with undetermined coefficients in terms of Chebyshev polynomials of the second kind $\{U_k(\tau)\}_{k=0}^{\infty}$ and taking a segment of the series

$$q''(\tau) = \sum_{k=0}^{N} z_k U_k(\tau) \tag{4.6}$$

where N is the order of the approximation in the Galerkin scheme. Thus, one can obtain a uniform approximation of the generalized velocity as a segment of the series expansion in terms of Chebyshev polynomials of the first kind

$$q'(\tau) = \sum_{k=0}^{N+1} y_k T_k(\tau) \tag{4.7}$$

where the $(N+2)$-dimensional coefficient vector $\mathbf{y} = (y_0, y_1, \ldots, y_{N+1})^{\mathrm{T}}$ is calculated in terms of the vector $\mathbf{z} = (z_0, z_1, \ldots, z_N)^{\mathrm{T}}$ by the formulae

$$y_0 = q_0' - \sum_{j=0}^{N} (j+1)^{-1} z_j,\quad y_k = k^{-1} z_{k-1}\quad (k = 1, 2, \ldots, N+1) \tag{4.8}$$

The function $q'(\tau)$ thus constructed satisfies the initial condition at $\tau = 1$. The generalized coordinate $q(\tau)$ is also approximated uniformly, by using the expansion

$$q(\tau) = \sum_{k=0}^{N+2} x_k T_k(\tau) \tag{4.9}$$

which is obtained by using standard formulae for the antiderivatives of the polynomials $T_k(\tau)$ [9]. When that is done

$$x_0 = q_0 - y_0 - \frac{y_1}{4} + \sum_{j=0}^{N+1} \frac{y_j}{j^2 - 1},\quad x_1 = y_0 - \frac{y_2}{2},\quad x_2 = \frac{y_1}{4} - \frac{y_3}{4}$$

$$x_k = \frac{y_{k-1} - y_{k+1}}{2k}\quad (k = 3, 4, \ldots, N+2),\quad x_{N+2} = \frac{y_{N+1}}{2(N+2)}$$

Formulae (4.8) were derived by using the standard relation $U_k(\tau) = (k+1)^{-1} T_{k+1}'(\tau)$. In order to set up Galerkin's equations, one must substitute from (4.7) and (4.9) into Eq. (4.4), also carrying out the following additional operations.

Put $A(\tau) = \Omega^2 + \mu \cos(a\tau + b)$. This function admits of the easily verified expansion [9]

$$A(\tau) = \sum_{k=0}^{\infty} A_k T_k(\tau)$$

$A_0 = \Omega^2 + \mu J_0(a)\cos b, A_k = 2\mu J_k(a)\cos(b + k\pi/2)$ $(k = 1, 2, \ldots)$, where $J_k(a)$ are Bessel functions of the first kind.

To complete the derivation of Galerkin's equations, we expand the function $A(\tau)q(\tau)$ first in terms of the polynomials $T_k(\tau)$ and then in terms of $U_k(\tau)$. In so doing we use the well-known property

$$T_m(\tau)T_n(\tau) = (T_{m-n}(\tau) + T_{m+n}(\tau))/2 \tag{4.10}$$

which holds for arbitrary integers m and n (bearing in mind that $T_{-n}(\tau) = T_n(\tau)$).

Let us introduce the following notation for the segment of the Chebyshev series consisting of the first $N + 3$ terms

$$[A(\tau)q(\tau)]_{N+2} = \sum_{k=0}^{N+2} Y_k T_k(\tau) \tag{4.11}$$

where, by (4.10)

$$Y_0 = A_0 x_0 + \frac{1}{2}\sum_{j=0}^{N+2} A_j x_j$$

$$Y_k = \frac{1}{2}\left[\sum_{j=0}^{k} A_{k-j}x_j + \sum_{j=k}^{N+2} A_{j-k}x_j + \sum_{j=0}^{N+2} A_{j+k}x_j\right] \quad (k = 1, 2, \ldots, N+2)$$

Expressing the polynomials of the first kind in (4.11) in terms of those of the second kind, via the standard formula

$$T_k(\tau) = (U_k(\tau) - U_{k-2}(\tau))/2$$

and retaining the first $N + 1$ terms of the expansion, we obtain

$$\left[[A(\tau)q(\tau)]_{N+2}\right]_N = \sum_{k=0}^{N} Z_k U_k(\tau)$$
$$Z_0 = Y_0 - Y_2/2, \quad Z_k = (Y_k - Y_{k+2})/2 \quad (k = 1, 2, \ldots, N) \tag{4.12}$$

Finally, the Galerkin system will consist of $N + 1$ algebraic equations in the $N + 1$ undetermined components of the vector $\mathbf{z}$

$$\mathbf{z} + \mathbf{Z}(\mathbf{z}) = \mathbf{0} \quad (\mathbf{z} \in \mathbf{R}^{N+1}, \mathbf{Z}: \mathbf{R}^{N+1} \to \mathbf{R}^{N+1}) \tag{4.13}$$

The vector-valued function $\mathbf{Z}(\mathbf{z})$ is determined by the components of the expansion (4.12). In the present example the system of equations (4.13) is linear in $\mathbf{z}$. In the general case, however, it is desirable to have an initial approximation for the solution of Eqs (4.13). When ε is small, the number μ will not be too large for moderate values of T. It turns out that in actual computations T may be quite large. It is only important to ensure that the norm of the derivative of the finite-dimensional operator $\mathbf{Z}(\mathbf{z})$ should be less than unity. One can then use Newton's method to find the solution of Eq. (4.13), taking as the initial approximation $\mathbf{z}^0 \in \mathbf{R}^{N+1}$ the coefficients of a segment of the expansion of the function $(q^0)''(\tau)$ in terms of the polynomials $U_k(\tau)$, where $q^0(\tau)$ is the solution of the unperturbed linear problem

$$q'' + \Omega^2 q = 0 \tag{4.14}$$

with the same initial conditions (4.5). The solution of this equation is

$$q^0(\tau) = q_0\cos\Omega\tau + q_0'\Omega^{-1}\sin\Omega\tau$$

Therefore

$$(q^0)''(\tau) = \sum_{k=0}^{\infty} \varsigma_k T_k(\tau) = \sum_{k=0}^{\infty} z_k^0 U_k(\tau), \quad \varsigma_0 = q_0 J_0(\Omega)$$

$$\varsigma_{2k-1} = 2(-1)^{k-1} q_0' \Omega^{-1} J_{2k-1}(\Omega), \quad \varsigma_{2k} = 2(-1)^k q_0 J_{2k}(\Omega) \quad (k = 1, 2, \ldots)$$

We can now compute the vector of the initial approximation for the solution of system (4.13) as

$$z_0^0 = \varsigma_0 - \varsigma_2/2, \quad z_k^0 = (\varsigma_k - \varsigma_{k+2})/2 \quad (k = 1, 2, \ldots, N)$$

REFERENCES

1. KOSENKO I. I., The construction of phase trajectories of a Hamiltonian system in the neighbourhood of an equilibrium position. *Prikl. Mat. Mekh.* **53**, 4, 531–538, 1989.
2. BAILEY C. D., A new look at Hamilton's principle. *Foundat. Phys.* **5**, 3, 433–451, 1975.
3. BAILEY C. D., On a more precise statement of Hamilton's principle. *Foundat. Phys.* **11**, 3/4, 279–296, 1981.
4. HITZL D. L. and LEVINSON D. A., Application of Hamilton's law of varying action to the restricted three-body problem. *Celest. Mech.* **22**, 3, 255–266, 1980.
5. 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.
6. HITZL D. L., Implementing Hamilton's law of varying action with shifted Legendre polynomials. *J. Comput. Phys.* **38**, 2, 185–211, 1980.
7. AGRAWAL O. P. and SAIGAL S., A novel, computationally efficient, approach for Hamilton's law of varying action. *Int. J. Mech. Sci.* **29**, 4, 285–292, 1986.
8. BATEMAN H. and ERDELYI A., *Higher Transcendental Functions*, II. McGraw-Hill, New York, 1953.
9. PASZKOWSKI S., Zastosowania numeryczne wielomianów i szeregów Czebyszewa. Państrv. Wid. Nauk., Warszawa, 1975.

Translated by D.L.