Robotica (1997) volume 15, pp 163–179.

Group theoretical approach in using canonical transformations and symplectic geometry in the control of approximately modelled mechanical systems interacting with an unmodelled environment

J.K. Tar*, I.J. Rudas*, and J.F. Bitó*

E-mail: jktar@zeus.banki.hu, rudas@zeus.banki.hu, jbito@zeus.banki.hu

SUMMARY

In spite of its simpler structure than that of the Euler-Lagrange equations-based model, the Hamiltonian formulation of Classical Mechanics (CM) gained only limited application in the Computed Torque Control (CTC) of the rather conventional robots. A possible reason for this situation may be, that while the independent variables of the Lagrangian model are directly measurable by common industrial sensors and encoders, the Hamiltonian canonical coordinates are not measurable and can also not be computed in the lack of detailed information on the dynamics of the system under control. As a consequence, transparent and lucid mathematical methods bound to the Hamiltonian model utilizing the special properties of such concepts as Canonical Transformations, Symplectic Geometry, Symplectic Group, Symplectizing Algorithm, etc. remain out of the reach of Dynamic Control approaches based on the Lagrangian model. In this paper the preliminary results of certain recent investigations aiming at the introduction of these methods in dynamic control are summarized and illustrated by simulation results. The proposed application of the Hamiltonian model makes it possible to achieve a rigorous deductive analytical treatment up to a well defined point exactly valid for a quite wide range of many different mechanical systems. From this point on it reveals such an ample assortment of possible non-deductive, intuitive developments and approaches even within the investigations aiming at a particular paradigm that publication of these very preliminary and early results seems to have definite reason, too.

KEYWORDS: Canonical transformations; Symplectic geometry; Mechanical systems control; Unmodelled environment.

1. INTRODUCTION

In order to gain a precise quantitative description of the physical processes different physical quantities and concepts must be provided with some real numbers or certain groups of real numbers. In general, the process of this provision is realized by the aid of different measurements. Due to the objective nature of the *measuring process,* especially in the field of technical applications, the arc illusion that *fully objective meaning* can be attributed to the *numbers* being results of the measurements frequently arises. Though this attitude is also supported by practical considerations concerning the direct measurability of certain quantities, it is misleading in the sense that this "provision" (that is the measurement) has many arbitrary possibilities and that physical concepts can be modelled mathematically in a higher level of abstraction.

A particular field of significant practical interest is Classical Mechanics, since in our daily life we meet many equipment items for the behavior of which partly the laws of CM are responsible. Typical examples are industrial robots as non-linear, strongly coupled multiple variable systems for the fault-tolerant control on which many recent efforts were made.[1,2]

In the field of CM, for instance, the basic concept is *the set of possible physical states of the system* forming *a differentiable manifold.* For gaining a quantitative description, differentiable manifolds can be described by the use of *atlases* consisting of *contradiction-free maps* mapping some subsets of the manifold to some *open regions* of $\Re^n$. The coordinates of these maps are not necessarily the direct results of certain measurements: they may—and must—be related to the measurements in indirect ways. By introducing topology-conserving coordinate transformations defined over the coordinates of a given map, *new maps* can be introduced for dealing with the *same physical reality*.

It is evident, that the mathematical form and complexity of the equations describing the same physical process may considerably depend on the properties of the given map, which from this point on, will be referred to as a *particular representation.*

In the control of mechanical devices such as industrial robots and manipulators the practical need for direct measurability of the coordinates to be controlled results in a strong insistence on the Lagrangian model's generalized coordinates. The use of the Lagrangian model may command the high price that the appropriate equations of motion form a second-order non-linear set

* Department of Information Technology, **Centre of Robotics and Automation, Bánki Donát Polytechnic, H-1081 Budapest, Népszínház u. 8., (Hungary).

of differential equations. The dynamic constants (parameters) of the robot as mechanical system in this model are "hidden" in the elements of a positive definite symmetric quadratic matrix referred to as the inertia matrix of the robot. The same parameters also are present in the quadratic expressions regarding the first time derivatives of the generalized coordinates, as well as in the terms expressing the effect of the gravitation. This second order nature, at least within the frames of the Lagrangian model excludes the possibility of using simple geometric concepts in describing the propagation of the state of the system in time. For this purpose *first order equations* are needed in which the first time derivatives of the coordinates describing the state of the system can be interpreted as the elements of the *tangent space of the differentiable manifold of the states.*

As is well known, within the frames of CM this step first was made by Hamilton in the 19th century by introducing the so called *canonical coordinates* as the results of a possible Legendre-transformation. Regarding the mathematical structure of CM this step had far reaching consequences. The theory gained the possibility of having a *pure local geometric interpretation* leading to the concept of *Symplectic Geometry* defined in the tangent space of the states. Symplectic Geometry has considerable formal analogies with the properties of the Euclidean Geometry more familiar in our everyday-life. Both concepts are based on a basic quadratic structure referred to as the *scalar product* and the *symplectic structure,* respectively. On the basis of these concepts the sets of the *orthonormal* and *symplectic sets* of *linearly independent basis vectors* can be introduced.

In Hamiltonian Mechanics the use of symplectic sets of basis vectors instead of orthonormal ones has definite reason: the state propagation of the mechanical systems transforms symplectic sets into symplectic ones in the tangent space of the physical states. Therefore, the description of the mechanical systems by symplectic sets has a kind of transparent "symmetry" which is not "apparent" in the case of other representations.

As in the case of the Euclidean Geometry in which an orthonormal set of basis vectors can be chosen in many arbitrary ways, in Symplectic Geometry also many arbitrary possibilities are available for choosing some symplectic bases. From the point of view of *algorithmic considerations,* in both cases an appropriate number of arbitrary but linearly independent vectors can be chosen in the first step. The free parameters of the arbitrary possible choices for the orthonormal (symplectic) basis are "hidden" in these "initial" vectors. By the use of simple and easily programmable numerical algorithms (the Gram-Schmidt and the Symplectising one) appropriate orthonormal (symplectic) bases can be gained from the initial vectors.

Group Theory-based analysis of the free parameters in the appropriate cases leads to the concepts of the *orthogonal and the symplectic groups,* respectively. From a physical point of view, as the orthogonal group describes an *inner symmetry* of Newton's "absolute space" observable by our senses, the symplectic group expresses an "abstract", non-trivial inner symmetry of the conservative mechanical systems. Both groups consist of *unimodular matrices* which can be inverted by simple matrix multiplications requiring a very limited number of numerical operations. Furthermore, both groups are Lie-groups and can be parametrized in many arbitrary ways with continuous parameters making it possible to use closed analytical formulas for describing the appropriate elements of the groups. In this description the linearly independent vectors of the tangent space of the groups near the vicinity of the unity element (the so-called generators) play a key role. By using the Lie-algebras of the appropriate groups, besides the geometric ones, considerable *algebraic analogies* can be utilized, too. In the Hamiltonian model the propagation of the state of the physical system simply can be related to the gradient of a *scalar function,* the Hamiltonian of the conservative mechanical system.

In spite of the above listed *formal advantages,* the Hamiltonian model has a serious drawback: its canonical coordinates cannot be directly measured. They can be *calculated* from the directly measurable coordinates only in the case when the *dynamical parameters of the system* normally not available in the case of a given robot are known exactly. Besides being complicated and time-consuming processes, these calculations cannot be done without identifying the unknown parameters of the system. As it was shown by Lantos[3] in 1993, normally only different combinations of the dynamical parameters of the robots can be identified via long lasting off-line calculations. (The appropriate "groups" depend on the kinematic structure of the robot arms, too.) The results of these calculations also are influenced by the dynamic interaction between the robot arem and its environment.

On the basis of the above calculations it is clear, that in order to utilize the formal advantages of Hamiltonian Mechanics, an "additional idea" is necessary. The "progenitor" of this idea first was introduced by Jánossy[4] in a quite different context. In his cited work Jánossy made an attempt to generalize Einstein's Special Theory of Relativity via introducing a so called "Deformation Principle". The essence of this idea is to use two different interpretations of the Lorentz Group*. In the "conventional interpretation" a Lorentz transformation yields the description of the *same physical system* by the use of the coordinates of a "system of coordinates" or frame different to the original one. Jánossy's interpretation runs as follows. By using the coordinates of a well defined frame, by transforming the coordinates of a given physical system with a Lorentz transformation the coordinates of a *different possible*

* It is worth noting that the Special Theory of Relativity also has quite strict formal analogies with both the Symplectic and the Euclidean Geometries. The basic concept is a quadratic term describing the propagation of light signals. This basic concept leads to the so-called Minkowski Geometry and to the Lorentz Group in quite a similar manner as the Orthogonal Group is related to the Euclidean, and the Symplectic Group is pertaining to the Symplectic Geometry, respectively.

https://doi.org/10.1017/S0263574797000192 Published online by Cambridge University Press

physical system, the deformed one can be obtained. For maintaining the possibility of this interpretation the introduction of certain restrictions regarding the allowable transformations was necessary. Based on this analogy, a similar "deformation principle" was proposed by Tar *et al.*[5] for the interpretation of the elements of the Symplectic Group strictly related to the Canonical Transformations of the Hamiltonian Mechanics. The main idea was based on a very rough dynamic model of the robot to be controlled. In this model the appropriate local symplectic deformations were interpreted as *adaptive modifications* of the initial rough model. Due to giving up the demand for a *full identification* of the system, the appropriate procedure also was referred to as "*partial identification*".

By using a symple paradigm for numerical simulations some results of preliminary investigations considering the operation of the outlined method were announced in different international conferences. These publications utilized only certain parts of the possibilities mentioned in this introduction. Recent investigations revealed, that on the basis of this "deformation principle", many different, adaptive control ideas can be developed. Besides being strictly related to the mathematical structure of the Hamiltonian Mechanics, these methods also show strong similarities to modern computing technologies such as Soft Computing. In both cases the identification and use of the *exact dynamic model* of the system to be controlled is abandoned. Either the Fuzzy Systems (FS), or the Artificial Neural Networks (ANN) have well defined mathematical structures in which plenty of free parameters are "hidden". The learning process of ANNs essentially consists of some tuning of these free parameters. In the case of supervised training (either purely causal, stochastic or combined) simple rules can be applied for realizing this learning. Fuzzy systems are also based on a plenty of free parameters hidden in the shape of the "membership functions" and "fuzzy relations". There exist tunable "adaptive" fuzzy systems, too. Both solutions may have a strongly parallel operation. It will be shown that considerable parallelism can also be realized in the case of calculations with the Hamiltonian model. Both solutions are based on simple "uniform" structures and modes of operation independent of the particular details of the concrete problem to be solved (min-max operations in the case of a fuzzy system, and learning/operation of the ANN). The combination of the simple algorithms mentioned above, as well as the use of the closed analytical formulas also independent of the particular details of the problem to be solved, also gives similar advantages to the use of the Hamiltonian Model. Furthermore, tuning the parameters within the Hamiltonian Model also gives possibilities for combining it with different Soft Computing methods.

The aim of this paper is to fully summarize the mathematical background of the proposed method in details, and to give illustrative examples regarding its operation via numerical simulations. It also is a goal to introduce different ideas needing further investigations. In *Section* 2 the connection between the Lagrangian and the Hamiltonian models is briefly summarized since this step is the only link which connects the abstract Hamiltonian model to the phenomenologically well established and interpreted Lagrangian description, i.e. to the realm of industrial sensors. In *Section* 3 the advantages of the Hamiltonian model are summarized in a succinct way listing the basic analogies between the *Euclidean* and the *Symplectic Geometry* defined in the *tangential space* of the set of possible physical states of the system in the form of a table. Also in this section those main properties of the *Lie Groups* will be summarized which are systematically utilized in this paper in two different particular cases: in the case of the Orthogonal and the Symplectic Groups. This Section contains a table, in which certain common, most useful properties of *Lie groups* are summarized. *Section* 4 is devoted to the *Deformation Principle* on which the proposed control method is based. *Section* 5 is dedicated to certain ideas already mentioned in connection with the application of the Hamiltonian Model in control technology. In *Section* 6 certain particular considerations pertaining to a particular paradigm and the appropriate simulation results are presented to demonstrate the possibilities "hidden" in the Hamiltonian Model. Conclusions are drawn in *Section* 7*,* while the remaining sections contain those parts which are obligatory components of scientific publications.

2. THE CONNECTION BETWEEN THE LAGRANGIAN AND THE HAMILTONIAN MODEL

As it is well known, phenomenological foundations of CM were established by Galileo in the 16th century by realizing the role of *time as an independent variable* in describing the behavior of mechanical systems[6] and by Newton in the 17th Century by introducing the concept of the *Inertial Systems of Coordinates* with respect to which the behavior of a mass-point can be described in the simplest mathematical form. This description uses *directly measurable physical quantities* as coordinate-vectors, velocities, accelerations and forces for describing the equations of motion for a single mass-point. It is a simple mathematical consequence that by the use of an inertial frame the *kinetic* and *potential energy* of the system can be constructed for a set of mass points interacting with each other and with an external environment, and that the Newtonian equations of motion can be deduced from the energy function via simple mathematical operations, too. From this point of view *rigid bodies* can be considered as special objects for the full description of the motion of which the use of a few independent coordinates, the *generalized coordinates* in this paper consistently denoted by letter "$\mathbf{q}$", can be used. On this basis it became possible to express the Newtonian equations of motion as a simple consequence of a variational principle for conservative systems. This principle is called the *Hamilton Principle* running as

https://doi.org/10.1017/S0263574797000192 Published online by Cambridge University Press

follows:[7] by using the system's *Lagrangian* as

$$L(\mathbf{q}, \dot{\mathbf{q}}) = T - V(\mathbf{q}),$$
$$T = \frac{1}{2} M_{ij}(\mathbf{q})\dot{q}_i\dot{q}_j, \quad (1)$$
$$V(\mathbf{q}) = \text{potential energy}$$

in which T denotes the *kinetic energy*, from the set of the system's prospective trajectories, starting from point $\mathbf{q}_0$ in time t_0 and ending in point $\mathbf{q}_1$ in time t_1, nature chooses the *realized one* for which the integral

$$\int_{t_0}^{t_1} (L + Q_u^{\text{Free}} q_u)dt = \text{Extremum} \quad (2)$$

In equation (2) there is a summation over the operative index denoted by subscript "u". The term Q_u^{Free} denotes the *generalized forces* defined as

$$Q_i^{Free} \equiv \frac{\partial}{\partial q_i} \sum_s F_i^s x_i^s(\mathbf{q}) \quad (3)$$

by using the Cartesian coordinates of the appropriate (sth) mass-points of the body with respect to the inertial frame x_i^s. It can be shown, that in equation (3) for a conventional robot the components of $\mathbf{Q}^{\text{Free}}$ can be interpreted as *the projection on the appropriate joint axis of the external forces/torques* acting on the given arm section or link. At least in principle, the phenomenological basis of the Lagrangian model is contained in this statement. By applying the usual method of partial integration the Euler-Lagrange equations of motion can be derived from the condition in equation (2) as

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = Q_i^{\text{Free}}. \quad (4)$$

Via applying the usual Legendre transformation the so called canonical momentum "$\mathbf{p}$" and the Hamilton function (Hamiltonian) $H(\mathbf{p}, \mathbf{q})$ can be introduced as

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i} = M_{ij}(\mathbf{q})\dot{q}_j, \quad H(\mathbf{q}, \mathbf{p}) \equiv p_s\dot{q}_s - L. \quad (5)$$

Via introducing L into equation (2) the variational principle yields the equation of motion in the terms of the *canonical coordinates* as

$$\dot{p}_s = -\frac{\partial H}{\partial q_s} + Q_s^{\text{Free}}, \quad \dot{q}_s = \frac{\partial H}{\partial p_s}. \quad (6)$$

From practical point of view, it is worth noting that the canonical momentum "$\mathbf{p}$" normally does not have *directly measurable components*. It is related to the directly measurable ones by the elements of the *inertia matrix* M_{ij} which contain the inertia parameters of the robot arm and the gripped work-piece manipulated by the robot. *In practice, normally no accurate information is available on the values of these parameters.*

Though the same parameters also are present in the Lagrangian, *at least the independent variables* of the system within the Lagrangian model *are directly measurable in principle**. This fact may explain why the Hamiltonian formalism is almost completely "neglected" in connection with the control of conventional robots in the present literature.

Though from phenomenological aspect the Hamiltonian model seems to be quite disadvantageous, regarding its *mathematical structure,* it leads to appropriate equations of motion of far simpler structure than that of the Lagrangian model. The main advantage of this model is that both the generalized coordinates and the canonical momentums have "equal rights" within *the set of the first order differential equations* characteristic to the Hamiltonian concept. This simplicity has profound mathematical consequences still poorly utilized in robot control. These consequences are considered in the next section.

3. THE FORMAL ADVANTAGES OF THE HAMILTONIAN MODEL

By "putting together" the components of "$\mathbf{q}$" and "$\mathbf{p}$" in a $2 \times \text{DOF}$ dimensional array defined as $\mathbf{x}^T = [\mathbf{q}^T, \mathbf{p}^T]$, the full energy of the system can be expressed as a simple scalar function $H(\mathbf{x})$. By introducing the *constant skew symmetric matrix of unit determinant*

$$\mathfrak{I} \equiv \begin{bmatrix} \mathbf{0} & \boldsymbol{I} \\ -\boldsymbol{I} & \mathbf{0} \end{bmatrix} \quad (7)$$

and the "$2 \times \text{DOF}$" dimensional array $\tilde{\mathbf{Q}}^{\text{Free}^T} \equiv [\mathbf{0}^T, \boldsymbol{Q}^{\text{Free}^T}]$ for the external generalized forces, the equations of motion have the form of

$$\dot{x}_i = \Im_{ij} \frac{\partial H(x)}{\partial x_j} + \tilde{Q}_i^{\text{Free}}. \quad (8)$$

Equation (8) gives rise to a very simple *geometric interpretation.* Let us suppose, that the appropriate "$\mathbf{x}$" coordinates can be the elements of some open region Ω. In each point "$\mathbf{x}$" of Ω the *tangential space of the states* can be defined as the linear space of the $\left\{\frac{\mathbf{x} - \mathbf{x}^{(n)}}{\|\mathbf{x} - \mathbf{x}^{(n)}\|}\right\}$ vectors in which $\mathbf{x}^{(n)}$ runs over the set of the neighboring points within the infinitesimally close vicinity of $\mathbf{x}$. $\dot{\mathbf{x}}$ evidently is an element of this set. The possible states of the system can be considered as a differentiable manifold for the mathematical description of which atlases consisting of consistent (contradiction-free) maps are used. A particular map realises a mapping of a sub-set of the manifold onto an open sub-set of $\Re^{2 \times DOF}$. *Due to the considerations related to the existence of the inertial frames it can be stated that Nature distinguishes a special map the coordinates of which correspond to those*

* It has to be noted that in practice the situation is not so clear even in the case of the Lagrangian model. In the definition given in equation (3) the F_i^s components should be summed over each "elementary" mass point of the robot arm as a rigid body. However, this summation could be done only if the whole surface of the robot would be covered with local force sensors. It is only a special supposition that the point of the action of the external forces is located in the gripper and that this external interaction can completely be identified via force and momentum sensors.

https://doi.org/10.1017/S0263574797000192 Published online by Cambridge University Press

canonical variables which can be directly introduced from the Lagrangian generalized coordinates.

From this point on a higher level of abstraction can be achieved in describing CM by turning from this *phenomenologically well substantiated "Lagrangian Map"* to other maps. In the tangential space of the system's states this immediately leads to *algebraic* and *geometric* analogies as it is shown below.

3.1. Analogies between the Euclidean and the Symplectic Geometries

From a purely mathematical aspect other maps can be introduced by an arbitrary differentiable and invertible time-independent coordinate transformation $\mathbf{x}'(\mathbf{x})$ treating the Hamiltonian as an invariant scalar that is changing its form as function as $H'(\mathbf{x}'(\mathbf{x})) \equiv H(x)$. The equations of motion according to this new map can be deduced from the original ones as

$$\begin{aligned}\dot{x}'_i &= \frac{\partial x'_i}{\partial x_s}\dot{x}_s = \frac{\partial x'_i}{\partial x_s}\left(\Im_{st}\frac{\partial H}{\partial x_t} + \tilde{Q}_s^{\text{Free}}\right)\\ &= \frac{\partial x'_i}{\partial x_s}\left(\Im_{st}\frac{\partial H'}{\partial x'_u}\frac{\partial x'_u}{\partial x_t} + \tilde{Q}_s^{\text{Free}}\right)\\ &= A_{iu}(\mathbf{x}')\frac{\partial H'}{\partial x'_u} + \tilde{Q}_i'^{\text{Free}} \qquad (9)\end{aligned}$$

It is clear that the structure of equation (9) is very similar to that of equation (8). In this case a *skew symmetric nonsingular and non-constant matrix* $\mathbf{A}(\mathbf{x})$ stands for $\mathfrak{I}$, and a similar, $2 \times DOF$ array stands for the external generalized forces. Both of them obeys well defined *transformation rules* as given in equation (9).

The situation is quite similar to the transformation law of the matrix elements of the *metric tensor* in the case of an Euclidean Geometry when turning to new coordinates defined with curved surfaces. *In this case the coordinate dependence in* "$\mathbf{A}$" *does not convey any essential information on the phyiscal system for the description of which it is used. It is rather characteristic of the more or less arbitrary way according to which the coordinates of a new map can be chosen.* As in the case of the Euclidean Geometry, in which the *possibility for introducing special coordinates leading to the constant unit matrix as the representation of the metric tensor* distinguishes these systems as particular ones, yielding the simplest description, in CM the possibility for choosing special maps on which $\mathbf{A}(\mathbf{x}) \equiv \mathfrak{I} \equiv$ const. also distinguishes these maps leading to the possible simplest form of the equations of motion. Therefore, as in the case of Euclidean Geometry of particular interest are those transformations, which leave the form of the metric tensor unchanged, in CM those coordinate transformations of the form of $\mathbf{x}'(\mathbf{x})$ for which

$$\frac{\partial x'_i}{\partial x_s}\Im_{st}\frac{\partial x'_u}{\partial x_t} = \mathfrak{T}_{Iau} \text{ or in matrix form } \mathbf{S}\mathfrak{I}\mathbf{S}^T = \mathfrak{I} \qquad (10)$$

also have special significance. (Here the matrix "$\mathbf{S}$" stands for the Jacobi matrix of the coordinate transformation). From equation (10) it is clear, that $\det \mathbf{S} = \pm 1$. With the restriction of $\det \mathbf{S} = 1$ these transformations are referred to as *Canonical Transformations,* and the appropriate maps are called *Canonical Maps.* The appropriate Jacobians $\mathbf{S}(\mathbf{x})$ are referred to as *symplectic matrices.*

From this point on it is easy to summarize the main formal analogies between the Euclidean and the symplectic geometries in Table I.

3.2. Other advantages of the Hamiltonian formalism

As normally in different fields of *Classical Physics* the basic laws of nature can be expressed in a *tensorial form* based on the structure of the scalar product, within the frames of *Classical Mechanics* the symplectic structure has a similar distinguished significance. Any measurable physical quantity characteristic of the system must be an *unique function* of the canonical coordinates unambiguously describing its physical state. The evolution of such a quantity $f(\mathbf{x})$ for an autonomous system can be described as a "Poisson Bracket" defined on the basis of the symplectic structure as

$$\dot{f}(\mathbf{x}) = \frac{\partial f}{\partial x_i}\dot{x}_i = \frac{\partial f}{\partial x_i}\Im_{ij}\frac{\partial H}{\partial x_j} = \{f, H\}. \qquad (11)$$

It is clear, that for two arbitrary *scalar functions* $f(\boldsymbol{x})$ and $g(\mathbf{x})$ the $\mathbf{x}'(\mathbf{x})$ canonical transformations lead to the representation $f(\mathbf{x}'(\mathbf{x})) \equiv f(\mathbf{x})$, etc. *leaving both the numerical value and the form of equation* (11) *unchanged.*

From an algebraic point of view the *linear space* of the arbitrary many times continuously differentiable functions can be transformed into an *algebra* by considering the Poisson brackets as a *multiplication.* The skew symmetry of the matrix $\mathfrak{I}$ has two significant consequences which can be utilized in the practice. They are as follows:

(a) If for a function $f(\mathbf{x})\{f, H\} \equiv 0$, it follows that $\{H, f\} \equiv -\{f, H\} = 0$. This fact can be interpreted as a *symmetry principle*; if the *phase current* generated by $f(\mathbf{x})$ as $\dot{x}_i = \Im_{ij}\frac{\partial f}{\partial x_j}$ leaves the Hamiltonian of the system unchanged (that is it is *a symmetry of the system*), the evolution of the system's state defined by the phase current $\dot{x}_j = \Im_{js}\frac{\partial H}{\partial x_s}$ leaves the numerical value of "f" constant. That is, to *each symmetry of the system pertains a characteristic constant.*
(b) For three arbitrary, infinitely many times continuously differentiable function $f(\mathbf{x})$, $g(\mathbf{x})$ and $h(\mathbf{x})$ this algebra has the properties of a *Lie algebra,* i.e. the *Jacobi identity* is satisfied by them: $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} \equiv 0$. If we put the Hamiltonian of the system into the place of $h(\mathbf{x})$, and $f(\mathbf{x})$ and $g(\mathbf{x})$ pertain to some symmetries of the system, their Poisson bracket $\{f, g\}$ will also be a *symmetry* of the same system.

Since the motion of the system is determined by the $2DOF$ constants determining the initial conditions,

https://doi.org/10.1017/S0263574797000192 Published online by Cambridge University Press

Table I. The formal analogies between the Euclidean and the Symplectic Geometries.

Euclidean Geometry (n dimension)	Symplectic Geometry (2n dimension)
Basic concept: Scalar product of vectors: $a_i\delta_{ij}b_j$	**Basic concept:** Symplectic structure: $a_i\Im_{ij}b_j$
Orthogonality: The vectors **a** and **b** are *orthogonal to each other* if $a_i\delta_{ij}b_j=0$. **Linear sub-spaces:** The set of those vectors which are *orthogonal* to a given vector a form a *linear space.*	**Anti-orthogonality:** The vectors **a** and **b** are *anti-orthogonal to each other* if $a_i\Im_{ij}b_j=0$. **Linear sub-spaces:** The set of those vectors which are *anti-orthogonal* to a given vector a form a *linear space.*
Orthonormal basis: The set of "n" linearly independent vectors $\{\mathbf{e}^{(i)} \mid i=1,\ldots,n\}$ which satisfies the restrictions $e_s^{(i)}\delta_{st}e_t^{(j)}=\delta_{ij}$ is referred to as an *orthonormed basis.*	**Symplectic basis:** The set of "$2n$" linearly independent vectors $\{\mathbf{f}^{(i)} \mid i=1,\ldots,2n\}$ which satisfies the restrictions $f_s^{(i)}\Im_{st}f_t^{(j)}=\Im_{ij}$ is referred to as a *symplectic basis.*
Orthogonal matrices: Those matrices, which transform an *orthonormed basis* into another orthonormed one via linear combination of the elements of the original set as $e_k'^{(j)}=O_{ji}\mathbf{e}_k^{(i)}$ are referred to as *orthogonal matrices.* This definition immediately leads to the restriction of $$e_k'^{(j)}e_1'^{(m)}\delta_{kl}=O_{ji}O_{mu}e_k^{(i)}e_1^{(u)}\delta_{kl} = O_{ji}O_{mu}\delta_{iu}=\delta_{jm}$$ or in a more succint style $\mathbf{OIO}^T=\mathbf{I}$.	**Symplectic matrices:** Those matrices, which transform a *symplectic basis* into another symplectic one via linear combination of the elements of the original set as $f_k'^{(j)}=\mathbf{S}_{ji}\mathbf{f}_k^{(i)}$ are referred to as *symplectic matrices.* This definition immediately leads to the restriction of $$f_k'^{(j)}f_1'^{(m)}\Im_{kl}=S_{ji}S_{mu}f_k^{(i)}f_1^{(u)}\Im_{kl} = S_{ji}S_{mu}\Im_{iu}=\Im_{jm}$$ or in a more succinct style $\mathbf{S}\mathfrak{J}\mathbf{S}^T=\mathfrak{J}$.
The Orthogonal Group: From the above definition it immediately follows that the orthogonal matrices form a group, the *orthogonal group,* and that det $\mathbf{O}=\pm1$.	**The Symplectic Group:** From the above definition it immediately follows that the symplectic matrices form a group, the *symplectic group,* and that det $\mathbf{S}=\pm1$.
The Orthogonal Group as a Lie Group: With the restriction of det $\mathbf{O}=1$ the orthogonal matrices form a *Lie group.*	**The Symplectic Group as a Lie Group:** With the restriction of det $\mathbf{S}=1$ the symplectic matrices form a *Lie group.*
The inverse of an Orthogonal Matrix: From the definition it comes, that the inverse of such a matrix can be computed without numerical procedures, since $\mathbf{O}^{-1}=\mathbf{O}^T$.	**The inverse of a Symplectic Matrix:** From the definition it comes, that the inverse of such a matrix can be computed with minimal numerical procedures, since $$\mathbf{S}^{-1}=\mathfrak{J}\mathbf{S}^T\mathfrak{J}^{-1}=\mathfrak{J}\mathbf{S}^T\mathfrak{J}^T.$$
Special orthogonal matrices: The unit matrix I defining the basic structure also is an orthogonal matrix. Furthermore, if **O** is orthogonal matrix, $\mathbf{O}^T$ also is orthogonal matrix.	**Special symplectic matrices:** The matrix $\Im$ defining the basic structure also is symplectic, since $\mathfrak{J}\mathfrak{J}\mathfrak{J}^T=\mathfrak{J}$. Furthermore, if **S** is symplectic, $\mathbf{S}^T$ also is symplectic.

finding some symmetries will help us to solve the *equations of motion.* By the systematic use of the Poisson bracket new constant quantities can be constructed from the known ones.

3.3. Common aspects of the Orthogonal and the Symplectic Groups utilized

Lie groups are special groups the elements of which can be "parametrized" by continuous parameters in the form of $g(\xi)$ and $g(\zeta)$ in a way that the product of the elements $g(\psi)=g(\xi)g(\zeta)$ is a unique function of the parameters $\psi(\xi,\zeta)$ and it is continuously differentiable in *infinite times.* (ξ, ζ, ψ are the elements of $\Re^N$, in which N denotes the dimension of the parameter-space of the group. If $\zeta=0$ corresponds to the *unit element of the group,* in an arbitrary composite scalar function of the variable "t" $g(\zeta(t))$ for which $\zeta(0)=0$ the

$$\frac{d}{dt}g(\zeta(t))=G \tag{12}$$

quantities can be conisdered as the *elements of the tangential space of the group drawn at the unit element.* For the *generators* of a Lie group simple considerations can be done leading to important consequences as is given in Table II. It is evident, that for an arbitrary generator G the function defined by the power series of the matrix exponential $g(t)=\exp(tG)$ represents a *single-parameter sub-group* generated by G. From the finite dimension of the linear space of the generators it

https://doi.org/10.1017/S0263574797000192 Published online by Cambridge University Press

Table II. Certain common properties of Lie groups utilized in the adaptive control.

Special Construction	Consequences
For the given $g(\zeta)$ and $h(\xi)$ functions $g(\zeta(\alpha t))$, $g(\zeta(t))h(\xi(t))$ can be considered with the restriction of $\zeta(0)=0,\ \xi(0)=0\ (\alpha\in\Re)$:	***The generators form a linear space.***
For an arbitrary constant group element "g" and for the above function $h(\xi)$ $gh(\xi)g^{-1}$.	***If H is a generator, then for an arbitrary group element "g" gHg^{-1} also is a generator.***
For an arbitrary constant generator "H" and the above defined function $g(\zeta(t))$ the derivative given below also is a generator: $$\frac{d}{dt}g(\zeta(t)Hg(\zeta(t))^{-1}$$	***If "G" and "H" are generators, then their commutator $[G,H]\equiv GH-HG$ also is a generator.***
Consideration of the *cyclic permutation* of the matrix products $[A,[B,C]]$:	***The generators of a Lie group satisfy the Jacobi equation:*** $$[A,[B,C]]+[B,[C,A]+[C,[A,B]]\equiv 0.$$

immediately can be concluded that by using the appropriate number of linearly independent parameters $G^{(i)}$, the matrix product

$$g(p_1,\dots,p_N)\equiv \exp(p_1G^{(1)})\exp(p_2G^{(2)})\cdots\exp(p_NG^{(N)}) \quad (13)$$

yields *a special continuous parametrization of the group.*

Normally by using special linearly independent generators the power series of the matrix exponentials can easily be expressed in a simple closed analytical form. It also is worth noting that if the vector $\mathbf{v}$ is an *eigenvalue* of the generator "G", then for an arbitrary element of the group "g" the $g\mathbf{v}$ vector will be the eigenvalue of the generator gGg^{-1}, since if

$$G\mathbf{v}=\lambda\mathbf{v},\quad \text{then}\quad gGg^{-1}g\mathbf{v}=\lambda g\mathbf{v}. \quad (13)$$

a special case is $\lambda=0$ to which the vectors left unchanged by the exponential in $\exp(tG)$ pertain. It is trivial that, for a given G the matrix exponential has a closed form, then the matrix exponential for an arbitrary group element "g" can also be expressed in a closed form, since

$$\exp(tgGg^{-1})=g\exp(tG)g^{-1}. \quad (13)$$

For gaining the *generators* of a Lie group the near unity transformations can be considered which must have the form of $\mathbf{T}=\mathbf{I}+\varepsilon\mathbf{A}$, in which ε is a small real number. This can be substituted into the quadratic equation defining the group. By prescribing the fulfillment of the definition up to the first power of ε the appropriate restrictions for the structure of the possible generators can be gained. In the case of the orthogonal group the generators must be *skew symmetric matrices.* For the symplectic group they must have the form of

$$\mathbf{A}=-\left[\begin{array}{c|c}\mathbf{G}^{a,b} & \mathbf{G}^{b,b}\\ \hline \mathbf{G}^{a,a} & -\mathbf{G}^{a,b^T}\end{array}\right] \quad (14)$$

in which $\mathbf{G}^{a,a}=\mathbf{G}^{a,a^T}$, $\mathbf{G}^{b,b}=\mathbf{G}^{b,b^T}$ otherwise are arbitrary matrices of $DOF\times DOF$ dimensions and $\mathbf{G}^{a,b}$ is arbitrary. It is trivial that $\mathbf{A}^T$ has similar structure, i.e. if $\mathbf{A}$ is generator, then $\mathbf{A}^T$ is also a generator. Since the generators form a linear space, *symmetric and skew symmetric generators* can be introduced for the symplectic group as

$$\begin{bmatrix}\mathbf{K} & \mathbf{H}^-\\ \mathbf{H}^- & -\mathbf{K}\end{bmatrix},\quad \begin{bmatrix}\mathbf{J} & \mathbf{H}^+\\ -\mathbf{H}^+ & \mathbf{J}\end{bmatrix} \quad (15)$$

where $\mathbf{H}^+$, $\mathbf{H}^-$ and $\mathbf{K}$ are symmetric, and $\mathbf{J}$ is skew-symmetric. By using these block matrices as linearly independent generators the closed analytical formulas for the exponentials are given below.

$$\begin{aligned}
\exp\left(t\begin{bmatrix}-\boldsymbol{K} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{K}\end{bmatrix}\right)&=\begin{bmatrix}\exp(-t\mathbf{K}) & \mathbf{0}\\ \mathbf{0} & \exp(t\mathbf{K})\end{bmatrix}\\
\exp\left(t\begin{bmatrix}-\mathbf{J} & \mathbf{0}\\ \mathbf{0} & -\mathbf{J}\end{bmatrix}\right)&=\begin{bmatrix}\exp(-t\mathbf{J}) & \mathbf{0}\\ \mathbf{0} & \exp(-t\mathbf{J})\end{bmatrix}\\
\exp\left(t\begin{bmatrix}\mathbf{0} & \mathbf{H}^-\\ \mathbf{H}^- & \mathbf{0}\end{bmatrix}\right)&=\begin{bmatrix}ch(t\mathbf{H}^-) & sh(t\mathbf{H}^-))\\ sh(t\mathbf{H}^-) & ch(t\mathbf{H}^-)\end{bmatrix}\\
\exp\left(t\begin{bmatrix}\mathbf{0} & -\mathbf{H}^+\\ \mathbf{H}^+ & \mathbf{0}\end{bmatrix}\right)&=\begin{bmatrix}\cos(t\mathbf{H}^+) & -\sin(t\mathbf{H}^+)\\ \sin(t\mathbf{H}^+) & \cos(t\mathbf{H}^+)\end{bmatrix}
\end{aligned} \quad (16)$$

For our control technical aims the construction of symplectic group generators leaving certain vectors $\mathbf{u}$ unchanged will be necessary. Such generators can easily be found by decomposing $\mathbf{u}$ into two blocks as $\mathbf{u}=[\mathbf{a}^T,\mathbf{b}^T]^T$, and by prescribing the restrictions for the symmetric and the skew symmetric blocks of the generators as

$$\begin{aligned}\mathbf{Ka}+\mathbf{H}^-\mathbf{b}=\mathbf{0},\quad \mathbf{H}^-\mathbf{a}-\mathbf{Kb}=\mathbf{0},\\ \mathbf{Ja}+\mathbf{H}^+\mathbf{b}=\mathbf{0},\quad -\mathbf{H}^+\mathbf{a}+\mathbf{Jb}=\mathbf{0}.\end{aligned} \quad (17)$$

These restrictions can automatically be satisfied if the matrices $\mathbf{J}$, $\mathbf{H}^-$, and $\mathbf{H}^+$ are constructed of arbitrary vectors taken from the *common orthogonal sub-space* of $\mathbf{a}$ and $\mathbf{b}$. Since the "$\mathbf{a}$ parallel with $\mathbf{b}$ case" may occur very seldom, generally it may be supposed that this sub-space is of the dimension of (DOF-2) spanned by the linearly independent orthonormed basis vectors $\{\mathbf{c}^{(i)} \mid i = 3, \ldots, DOF\}$. The appropriate components in the generators are

$$K_{ij}^{(uv)} = H_{ij}^{(uv)^-} = H_{ij}^{(uv)^+} = \frac{1}{2}(c_i^{(u)}c_j^{(v)} + c_j^{(u)}c_i^{(v)}), \quad J_{ij}^{uv} = \frac{1}{2}(c_i^{(u)}c_j^{(v)} - c_j^{(u)}c_i^{(v)}) \tag{18}$$

Such a set can easily be created by the use of the Gram-Schmidt algorithm started with an initially orthonormed set where its first two elements are replaced by **a** and **b**. The columns of these unit vectors form an *orthogonal matrix* **C** which can be utilized in the control for assigning continuous parameters to certain symplectic matrices for the purpose of parameter-tuning.

4. THE "DEFORMATION PRINCIPLE"

In strict analogy with the idea invented by L. Jánossy, in the field of the Hamiltonian Mechanics the introduction of the following principle can be attempted. Within the frames of the conventional theory the canonical $\mathbf{x}'(\mathbf{x})$ transformations allowed only one way for interpretation: "*the same physical system is described by the coordinates of a new canonical map*". The novel interpretation of *certain* canonical transformations: "*the* $\mathbf{x}'(\mathbf{x})$ *coordinates correspond to another, hypothetical physical system* (*the "deformed one"*) *described by the use of the coordinates of the original map*".

In order to use this principle for control-technical purposes it is necessary to use the directly measurable quantity $\mathbf{y} = [\mathbf{q}, \dot{\mathbf{q}}]$ instead of the canonical $\mathbf{x}$. The question is how these quantities can be related to the canonical formalism of Classical Mechanics. By the use of the functions $H'(\mathbf{y}) \equiv H(\mathbf{x}(\mathbf{y}))$, $\mathbf{L} \equiv \frac{\partial \mathbf{x}}{\partial \mathbf{y}}$, $\tilde{\mathbf{Q}}'^{\text{Free}} \equiv \mathbf{L}\tilde{\mathbf{Q}}^{\text{Free}}$ a skew-symmetric, non-singular matrix $\mathbf{A}(\mathbf{y})$ can be introduced instead of $\Im$ as $\mathbf{A}(\mathbf{y}) = \mathbf{L}^{-1}\Im\mathbf{L}^{-1^T}$. The equations of motion gain a form more or less similar to the canonical ones: $\mathbf{y}_i = A_{ij}(\mathbf{y})\frac{\partial H'(\mathbf{y})}{\partial y_j} + \tilde{\mathbf{Q}}'^{\text{Free}}$. Instead the original canonical transformations modified ones of the form of $\mathbf{z}(\mathbf{y})$, for which $\mathbf{U} \equiv \frac{\partial \mathbf{y}}{\partial \mathbf{z}}$ leaves $\mathbf{A}(\mathbf{y})$ locally unchanged can be considered: $\mathbf{A}(\mathbf{z}(\mathbf{y})) = \mathbf{U}^{-1}\mathbf{A}(\mathbf{y})\mathbf{U}^{-1^T} = \mathbf{A}(\mathbf{y})$. If we apply a *very rough approximate dynamic model* for the system, the matrix **L** can be constant, consequently the matrix **A** can also be constant. It is evident that the above defined **U** matrices form a group, too, and this group is in quite close relation with the symplectic group and has very similar properties: the $\mathbf{L}\mathbf{U}^{-1}\mathbf{L}^1$ matrices are *symplectic*. The applied control method is based on this idea. The appropriate restrictions to be imposed for the purposes of the deformation principle are as follows: In the canonical map directly deduced from the Lagrangian model the *generalized force vectors* have only *DOF* non-zero components. Since a general canonical transformation can "mix together" all the $2DOF$ components of the transformed generalized force vectors, a considerable part of the canonical transformations cannot be applied for deformation purposes. Only those solutions can be accepted, for which the necessarily "truncated", phenomenologically non-interpretable components of the generalized force vector are negligible in comparison with the interpretable parts.

5. POSSIBLE APPLICATIONS OF THE DEFORMATION PRINCIPLE IN ADAPTIVE CONTROL

The essence of the deformation consists in the difference in the phase currents generated by H' and H in the same point of the differentiable manifold. The idea of a *partial system identification* is related to this interpretation: starting with a very rough initial model of constant **M** in L and with a constant $\partial V/\partial \mathbf{q}$ the model establishes a connection between the exerted local generalised forces and the propagation of the state-vector $\dot{\mathbf{y}}^{\text{Mod}}$. In reality, the encoders measure a different propagation $\dot{\mathbf{y}}^{\text{Real}} \neq \dot{\mathbf{y}}^{\text{Mod}}$. It is expected, that the difference can be eliminated by some deformation of the initial model in the form of $H'(\mathbf{z}) = H(\mathbf{y}(\mathbf{z}))$ locally represented by the appropriate matrix **U**.

According to the original canonical formulation, an appropriate symplectic matrix is to be found for which $\mathbf{a} = \dot{\mathbf{x}}^{\text{Real}} = \mathbf{S}\dot{\mathbf{x}}^{\text{Mod}} = \mathbf{S}\mathbf{b}$ is valid. This can be done e.g. in the following way: By making two quadratic matrices of the column vectors **a** and **b** (**A** and **B**) via "putting near them" further linearly independent vectors, the matrix relation $\mathbf{A} = \mathbf{S}\mathbf{B}$ can be prescribed. Due to the group properties of the symplectic matrices this can be satisfied if both **A** and **B** are symplectic and their first column is equal to **a** and **b** respectively. The solution is simply $\mathbf{S} = \mathbf{A}\mathbf{B}^{-1} = \mathbf{A}\Im\mathbf{B}^T\Im^T$.

This situation can be achieved by the *symplectising algorithm,* a simple and easily programmable procedure quite similar to the Gram-Schmidt orthogonalization method frequently used in a Quantum Mechanics for obtaining *orthonormed* basis vectors. The two algorithms can be treated in a strictly "parallel" way, as it is summarized in Table III.

Due to the group properties of the symplectic matrices in each computing cycle of the controller the symplectic deformation applied in step "t" $\mathbf{S}(t)$ can be so "corrected" by the actually computed symplectic correction $\mathbf{T}(t)$ that $\mathbf{S}(t+1) = \mathbf{T}(t)\mathbf{S}(t)$, etc. It can be expected that this modified model will yield a better solution than the original, "rough" dynamic model without any corrections.

An alternative possibility is to abandon the idea of *cumulative corrections* and apply the symplectic deformation in each steps starting directly from the initial rough dynamical model in each step.

It is evident, that in both cases the symplectic model based approach contains a considerable number of "unconstrained parameters" hidden in the columns of matrices **A**, **B**.

Though the symplectising algorithm decreases the

https://doi.org/10.1017/S0263574797000192 Published online by Cambridge University Press

Table III. Formal analogies between the Gram-Schmidt and the Symplectizing algorithms

The Gram-Schmidt Algorithm	**The Symplectizing Algorithm**
Let $\{\mathbf{a}^{(i)} \mid i=1,\ldots,n\}$ a linearly independent set of basis vectors.	Let $\{\mathbf{b}^{(i)} \mid i=1,\ldots,2n\}$ a linearly independent set of basis vectors.
Since $\mathbf{a}^{(1)} \neq 0$, it can be *normed* for forming the *first element* of the orthonormed set $\mathbf{e}^{(i)}$: $$\mathbf{e}^{(1)} = \frac{\mathbf{a}^{(1)}}{\|\mathbf{a}^{(1)}\|}$$	Since $\Im$ is non-singular, none of the $\Im\mathbf{b}^{(j)}$ $(j=1,\ldots,2n)$ vectors can be zero. Due to its skew-symmetry $\mathbf{b}^{(1)T}\Im\mathbf{b}^{(1)}=0$, therefore the remaining set must contain at least one vector, $\mathbf{c}$, for which $\mathbf{b}^{(1)T}\Im\mathbf{c}\neq 0$. Via permutation of the remaining vectors let the index "$n+1$" assigned to it. Via the normalization $\mathbf{b}'^{(n+1)} = \frac{\mathbf{b}^{(n+1)}}{\mathbf{b}^{(1)T}\Im\mathbf{b}^{(n+1)}}$ the vector $\mathbf{b}'^{(n+1)}$ is the *symplectic "mate"* of $\mathbf{b}^{(1)}$.
Those vectors $\mathbf{a}^{(j)}$ of the remaining set which are not orthogonal to $\mathbf{e}^{(1)}$ can be made orthogonal to it by the transformation $$\mathbf{a}'^{(j)} = \mathbf{a}^{(j)} - \mathbf{e}^{(1)}[\mathbf{e}^{(1)T}\mathbf{a}^{(j)}] \neq 0.$$	Those vectors $\mathbf{b}^{(j)}$ of the remaining set which are not anti-orthogonal to the pair $\mathbf{b}^{(1)}$ and $\mathbf{b}'^{(n+1)}$ can be made anti-orthogonal to them by the transformation $$\mathbf{b}'^{(j)} = \mathbf{b}^{(j)} + \mathbf{b}^{(1)}[\mathbf{b}^{(n+1)T}\Im\mathbf{b}^{(j)}] - \mathbf{b}^{(n+1)}[\mathbf{b}^{(1)T}\Im\mathbf{b}^{(j)}] \neq 0.$$
Due to the completeness and linear independence of the original set of vectors the transformed remaining set must consist of $(n-1)$ linearly independent non-zero vectors each of which is orthogonal to $\mathbf{e}^{(1)}$,	Due to the completeness and linear independence of the original set the remaining set must consist of $(2n-2)$ non-zero, linearly independent vectors each of which is anti-orthogonal to the pair $\mathbf{b}^{(1)}$ and $\mathbf{b}'^{(n+1)}$.
The above steps can be repeated within the linear sub-space orthogonal to $\mathbf{e}^{(1)}$.	The above steps can be repeated within the linear sub-space anti-orthogonal to the pair $\mathbf{b}^{(1)}$ and $\mathbf{b}'^{(n+1)}$.
The final result is an orthonormal set of basis vectors.	The final result is a symplectic set of basis vectors.

number of these parameters, within this process the "story" of these parameters cannot be traced in a lucid way. Furthermore, though these parameters do not concern the control task in the given step, the appropriate prediction made on the basis of this estimation influences the behavior of the controlled system in the next step.

In order to deal with the free parameters in a more flexible way, the introduction of the continuous Lie parameters is expedient. It may be done in the following way: Instead of using it a slight modification can be introduced in the form as

$$\mathbf{P}^{*}\mathbf{S}^{\text{Desired}} = \mathbf{S}\mathbf{P}\mathbf{S}^{\text{Measured}} \tag{19}$$

in which $\mathbf{P}^{*}$ and $\mathbf{P}$ are independent matrices containing the continuous parameters of the symplectic group, and they are so chosen, that $\mathbf{P}^{*}\mathbf{u}^{\text{Desired}} = \mathbf{u}^{\text{Desired}}$, and $\mathbf{P}\mathbf{u}^{\text{Measured}} = \mathbf{u}^{\text{Measured}}$. According to the procedure based on the Gram-Schmidt algorithm and equation (18) an orthogonal matrix $\mathbf{C}$ can be constructed in each control step. Due to the group properties of the orthogonal matrices it can be considered as a rotation of an "initial set" forming the columns of the unit matrix as

$$\mathbf{c}^{(i)} = \mathbf{C}\mathbf{e}^{(i)}, \quad e_j^{(i)} = \delta_{ij}. \tag{19}$$

By applying equations (13–14) for the orthogonal group, the appropriate blocks $\mathbf{K}$, $\mathbf{J}$, $\mathbf{H}^{-}$, $\mathbf{H}^{+}$ constructed of the $\{\mathbf{c}^{(i)}\}$ vectors and their analytical functions can simply be obtained from their counterparts constructed of the set $\{\mathbf{e}^{(i)}\}$: $\mathbf{K} = \mathbf{C}\mathbf{K}'\mathbf{C}^{T}$, $\mathbf{J} = \mathbf{C}\mathbf{J}'\mathbf{C}^{T}$, $\mathbf{H}^{+-} = \mathbf{C}\mathbf{H}^{+-\prime}\mathbf{C}^{T}$. These expressions can be applied in the blocks of equation (16). Since the components of the $\{\mathbf{e}^{(i)}\}$ set have a very simple structure, their appropriate matrix exponentials for a $3DOF$ system can be expressed in a closed analytical form as

$$\mathbf{S}^{K33} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \exp(t) & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \exp(-t) \end{bmatrix}$$

$$\mathbf{S}^{H^{-}33} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & ch(t) & 0 & 0 & sh(t) \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & sh(t) & 0 & 0 & ch(t) \end{bmatrix}$$

$$\mathbf{S}^{H^{+}33} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \cos(t) & 0 & 0 & \sin(t) \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -\sin(t) & 0 & 0 & \cos(t) \end{bmatrix}$$

(In the case of 3-DOF systems for the skew-symmetric $\mathbf{J}$ only the trivial zero solution can be obtained.) The matrix denoted by the superscript "**K**" expresses a stretch/shrink along the third axis, "accompanied" by a simultaneous contraction/stretch in the third component of the canonical momentum vector. The matrix that has the superscript H^{-} describes a hyperbolic rotation in the space of the canonical state vectors influencing only the third components of the state vectors. Similar statement

https://doi.org/10.1017/S0263574797000192 Published online by Cambridge University Press

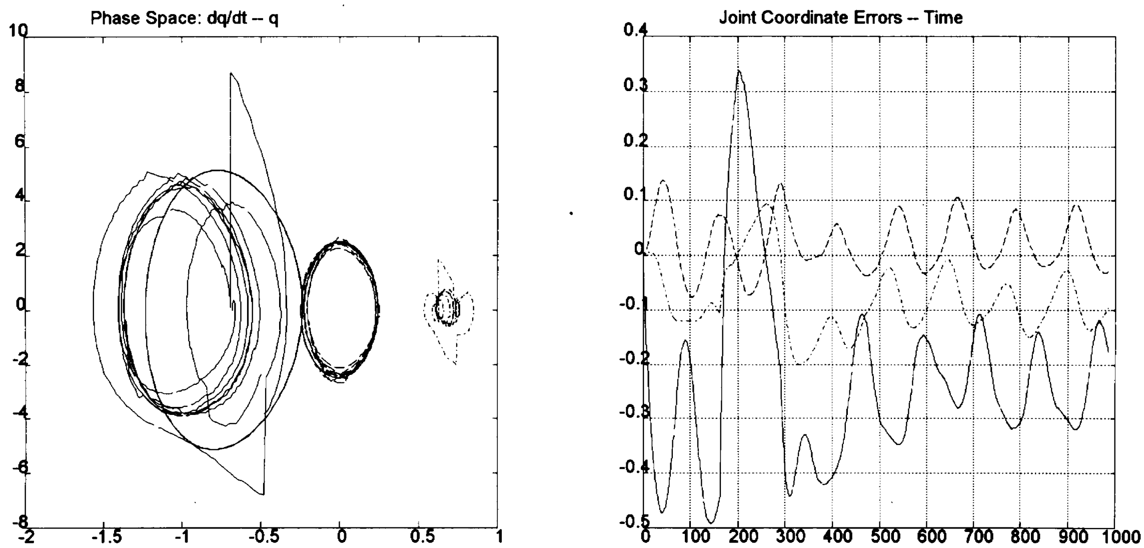

Fig. 1. Typical result of the non-cumulative solution without tuning of adaptive parameters: projections in the planes of the phase-space and the joint coordinate errors.

holds for the matrix denoted by H^+ describing "common rotation". Via associating continuous parameters α_1, α_2, α_3 to $\mathbf{K}$, $\mathbf{H}^-$ and $\mathbf{H}^+$ in the symplectic matrix $\mathbf{P}$, and assigning their counterparts β_1, β_2, and β_3 to $\mathbf{P}^*$, the control of a 3-*DOF* system development of parameter-tuning strategies can be initiated. In the next section the behaviors of certain strategies are presented on the basis of computer simulation.

6. SIMULATION RESULTS

For simulation purposes exactly the same 3-*DOF* robot arm structure was used that in a previous investigation.[8] It consisted of a vertical rod of 5 kg moving up and down (q_1 in m), rotating around itself as a vertical axis (q_2 in *rad*), and a second rod joined to it by a wrist tilting around a horizontal axis (q_3 in *rad*). This latter joint was also translated by q_1 and rotated by q_2. The second rod had negligible mass but carried a point-like small body of variable mass. It also had variable, but constant length (R_0 in m). The three axes were controlled by drives exerting force on q_1 and torque on q_2 and q_3 prescribed by the control strategy. In each cases considered the end-point of the robot arm was desired to be moved with circular frequency $\Omega\, s^{-1}$ along a circle of $0.5\, m$ radius laying in a vertical plane at a distance of $2\, m$ from the vertical axis. In each case, the "*initial rough estimation*" of the dynamic model consisted of a non-singular, constant inertia matrix and a constant gravitational term. No quadratic velocity coupling was taken into account. To make all the further corrections was the task of the symplectising algorithm and the "predictive control". In the space of the *joint coordinates* a linear feedback was applied as

$$\ddot{\mathbf{q}}^D + \ddot{\mathbf{q}}^N - b(\dot{\mathbf{q}}^B - \dot{\mathbf{q}}^N) - c(\mathbf{q}^R - \mathbf{q}^N) \qquad (19)$$

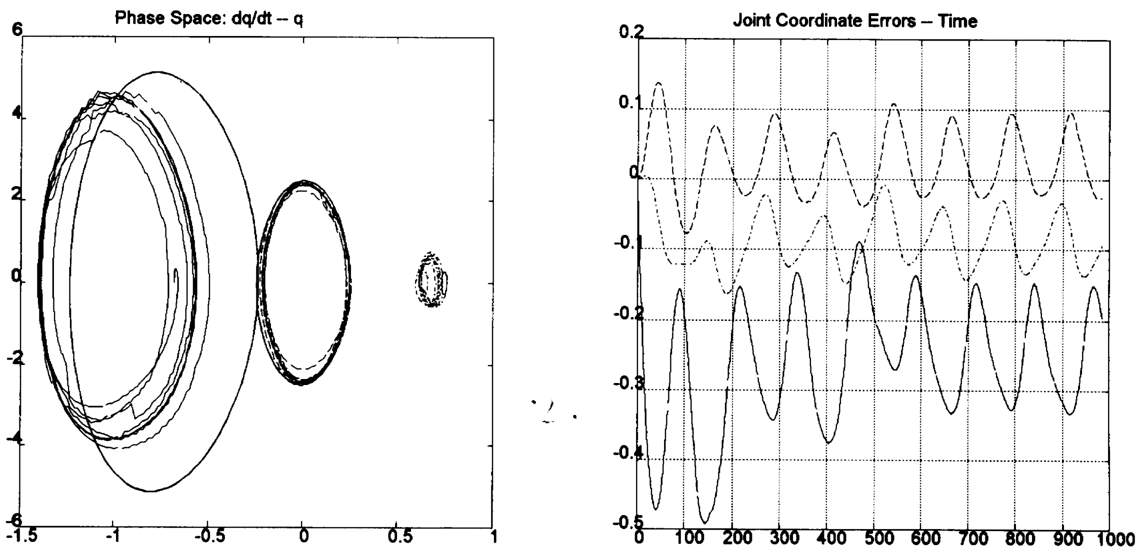

Fig. 2. Typical result of the non-cumulative solution with tuning α_1: projections in the planes of the phase-space and the joint coordinate errors.

https://doi.org/10.1017/S0263574797000192 Published online by Cambridge University Press

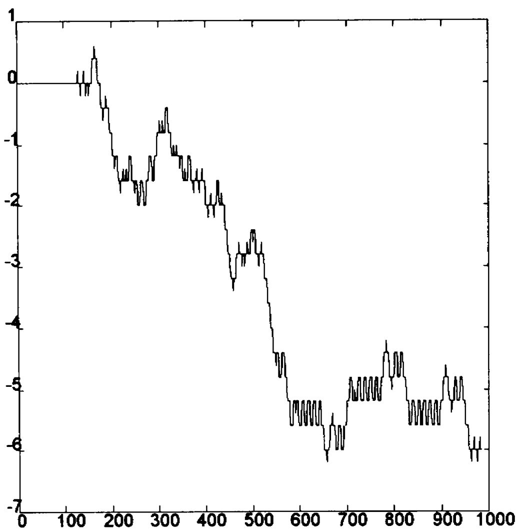

Fig. 3. Variation of α_1 for parameter tuning.

in which the superscript D, R and N correspond to the "desired", the "realized" and the "nominal" values, respectively. (According to the above concepts, there is no "desired" trajectory. The nominal trajectory pertains to the motion to be executed. In the case of a Computed Torque Control (CTC) we have only desired accelerations in each control cycle.) The unmodelled environmental interaction was represented by a spring having viscous friction ("dashpot"). The values $b = 15\,s^{-1}$ and $c = 56\,s^{-2}$ were so chosen that equation (19) approximately corresponds to a single damping constant without any oscillation. In the lack of the exact dynamical model this strategy cannot be precisely implemented. The expected role of the identification method was continuous correction of the initial dynamic model in order to really implement this strategy. In the graphs the *solid,* the *dashed* and the *dashdot* lines correspond to $q_1(Q_1)$, $q_2(Q_3)$, and $q_3(Q_3)$, respectively. (In the figures

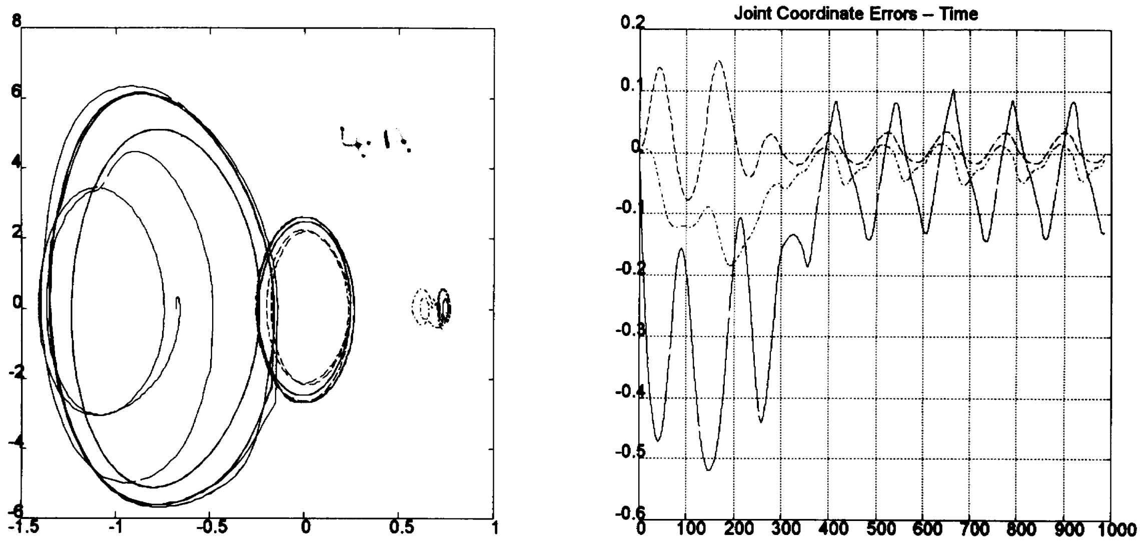

Fig. 4a. Cumulative control without parameter tuning for small viscosity: phase-space, joint coordinate errors.

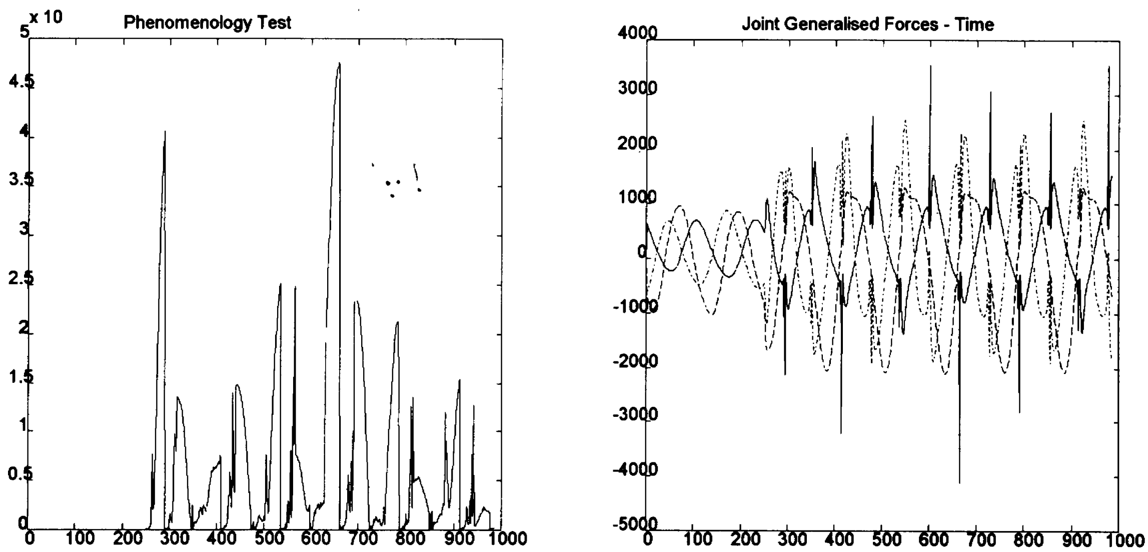

Fig. 4b. Cumulative control without parameter tuning for a small viscosity: phenomenology test (the norm of the truncated, non-interpretable part of the generalized forces) and the generalized forces.

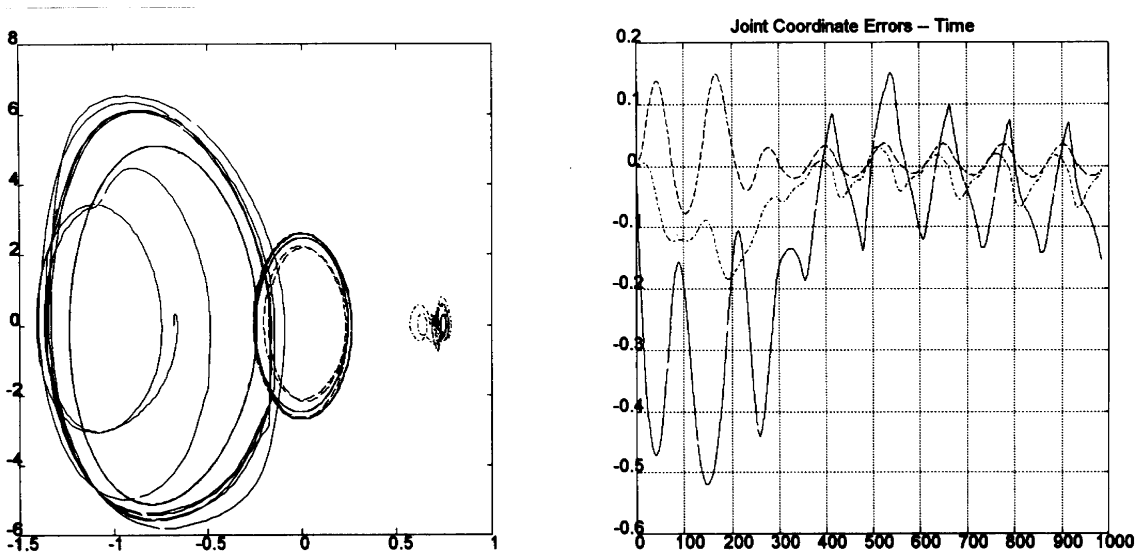

Fig. 5a. Behavior of the cumulative control when tuning α_1 and the viscosity is small: phases-space, joint coordinate errors (dimless).

describing functions *versus* time time is given in δt units, i.e. the duration of one computational cycle. In the simulations it was 5 ms.)

In the first quarter of the duration of the motion only a simple linear feedback based on the rough dynamical model was applied. The symplectic model deformation was in effect only from this point on. Regarding the tuning of the continuous Lie parameters, in each case a very simple strategy was applied; the adaptive parameters were kept moving. If the change in the adaptive parameter coincided with the decrease in the prediction error in the state propagation, this tendency was maintained. Otherwise it was reversed.

The application of a non-cumulative approach, in which the input of the symplectising algorithm were the columns of the symplectic matric, $\mathfrak{I}^T$ did not give satisfactory results as it is given in Figure 1.

Its quality was not better than that of the purely linear approach. Tuning of the adaptive parameter α_1 (the Lie-parameter of the stretch/contraction along $\mathbf{e}^{(3)}$) did not give essential improvement as it is shown in Figures 2–3.

Increasing or decreasing the finite steps in the adaptive parameter did not give essential modification of the above structure. Consequently, further investigations were concentrated on the behavior of the *cumulative approach.*

Regarding the cumulative approach, the symplectizing algorithm with the same inputs showed better results, but it was very sensitive to the viscosity present in the "dashpot" as in the counterpart in an unmodelled external interaction.

A similar solution using the columns of the unit matrix as the input of the symplectizing algorithm resulted in a better quality of motion even without extra parameter tuning. The results are described in Figures 4a–4b.

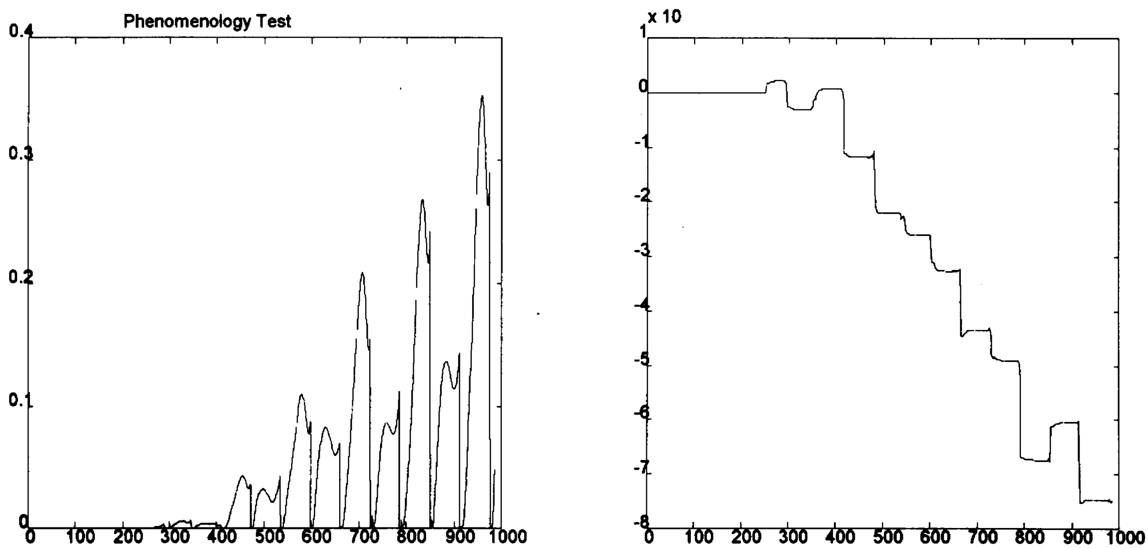

Fig. 5b. Behavior of the cumulative control when tuning α_1 and the viscosity is small: phenomenology test and α_1 in the range of 10^{-3}.

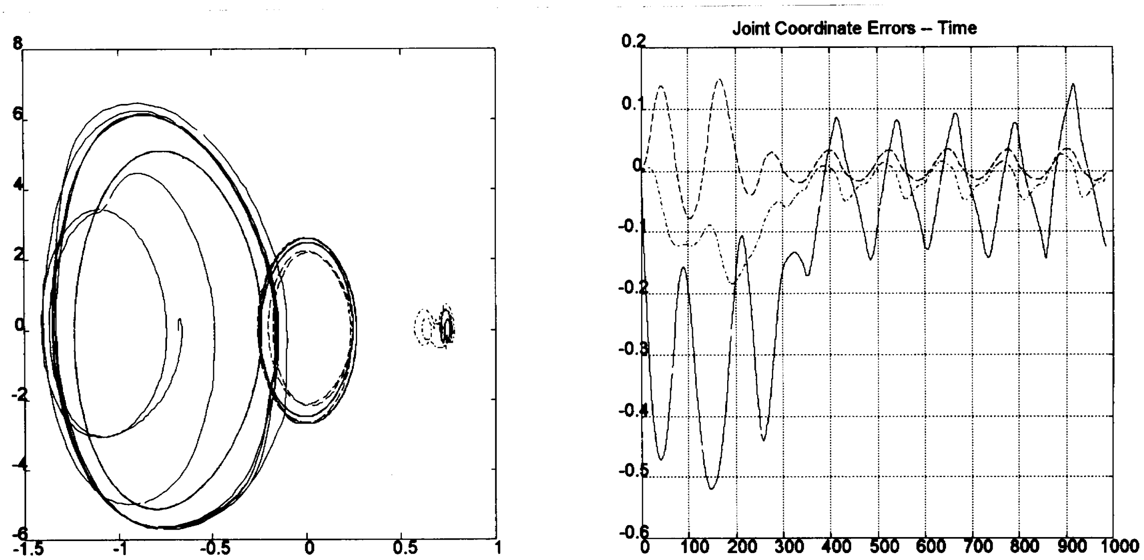

Fig. 6a. Behavior of the cumulative control when tuning α_2 and the viscosity is small: phase-space, joint coordinate errors.

Figures 4a–b reveals, that the cumulative nature of the control by itself means a kind of adaptivity. The applied transformations satisfy the deformation principle since the truncated part of the generalized forces is negligible in this case.

Due to its cumulative nature, this approach is much more sensitive to the variation of the continuous parameters, as it is shown in Figures 5–7: The modification is mostly apparent in the figures describing the joint coordinate errors.

The system is far more sensitive to the variation of α_2 and α_3 "mixing" the phenomenologically interpretable and non-interpretable components.

In Figs 8a–8b each of the six continuous parameters are tuned during 3-3 consecutive steps in the case of low viscosity. In the case of applied tuning the interdependence of these parameters shows a kind of stabilizing effect: none of them can meander far from the zero value pertaining to the identical transformation.

In Figure 9 the effect of a high viscosity coefficient can be traced. It can be stated, that the given strategy is not too sensitive to the viscosity of the external system (dashpot).

The effect of the increased viscosity can well be observed in this phase space, in the shift of the joint coordinate errors and in the change in the shape of the

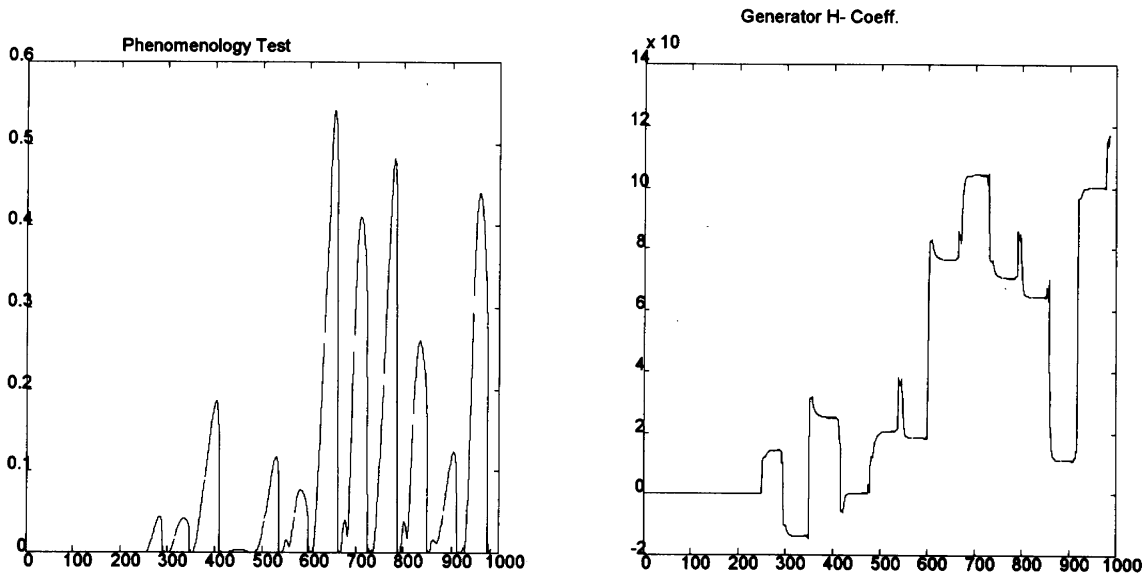

Fig. 6b. Behavior of the cumulative control when tuning α_2 and the viscosity is small: phenomenology test, and α_2 in the range of 10^{-6}.

https://doi.org/10.1017/S0263574797000192 Published online by Cambridge University Press

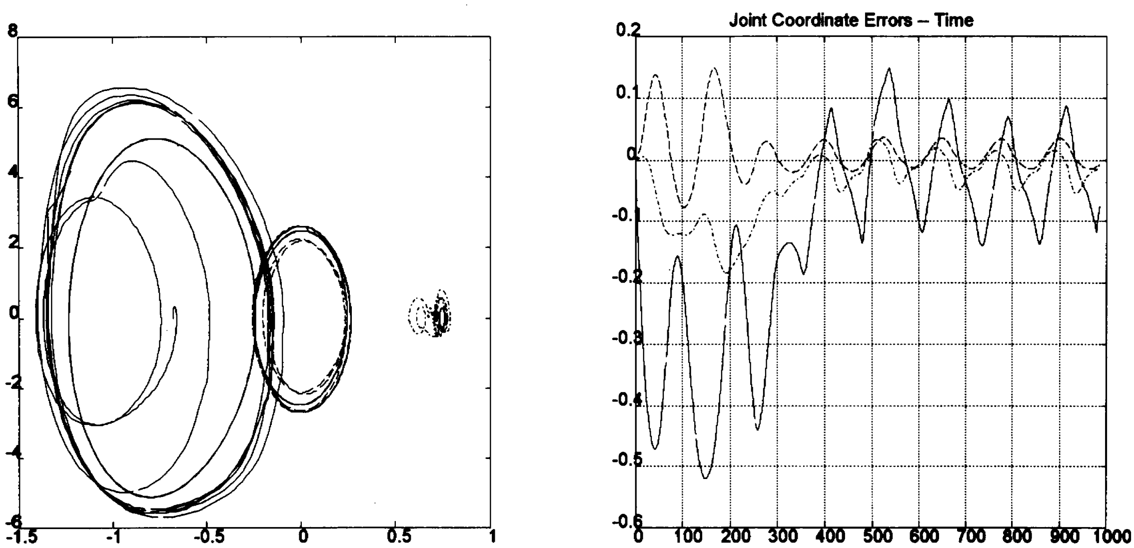

Fig. 7a. Behavior of the cumulative control when tuning α_3 and the viscosity is small: phase space, joint coordinate errors.

curves describing the generalized forces. The data described in Fig. 10a–10b pertain to very high environmental viscosity. The above mentioned tendencies are far more easily observable in these graphs.

7. CONCLUSIONS

In this paper a concise theory has been developed with the aim of making it possible to utilize certain simple concepts inherent in the Hamiltonian Mechanics for control technical purposes. The theoretical approach here presented was strongly based on *simple formal analogies* between otherwise different mathematical concepts and procedures. It was found that on the basis of an idea called "*deformation principle*", the transparent concepts as canonical transformation, Euclidean and the Symplectic Geometry, Orthogonal and Symplectic Groups as Lie groups and their generators, orthonormed and symplectic sets of basis vectors, the Gram-Schmidt and Symplectizing algorithms can be used for inventing robust and adaptive control for mechanical systems in dynamic interaction with an unmodelled environment via tuning of continuous free parameters within closed-form analytical expressions. The number of the independent continuous, tunable parameters strongly increases with the degree of freedom of the mechanical system to be controlled. Neither the complexity nor the structure of the computational operations depend on the particular features of the mechanical system to be controlled. This structure has a kind of "uniformity" and universality like certain ANNs and fuzzy controllers. The proposed algorithms can be run in a strongly parallel way on an appropriate, multiple-processor hardware; for both sides of the control equation computation of the change in the Lie-parameters, the Gram-Schmidt and the Symplectizing algorithms can be run simultaneously in a parallel way.

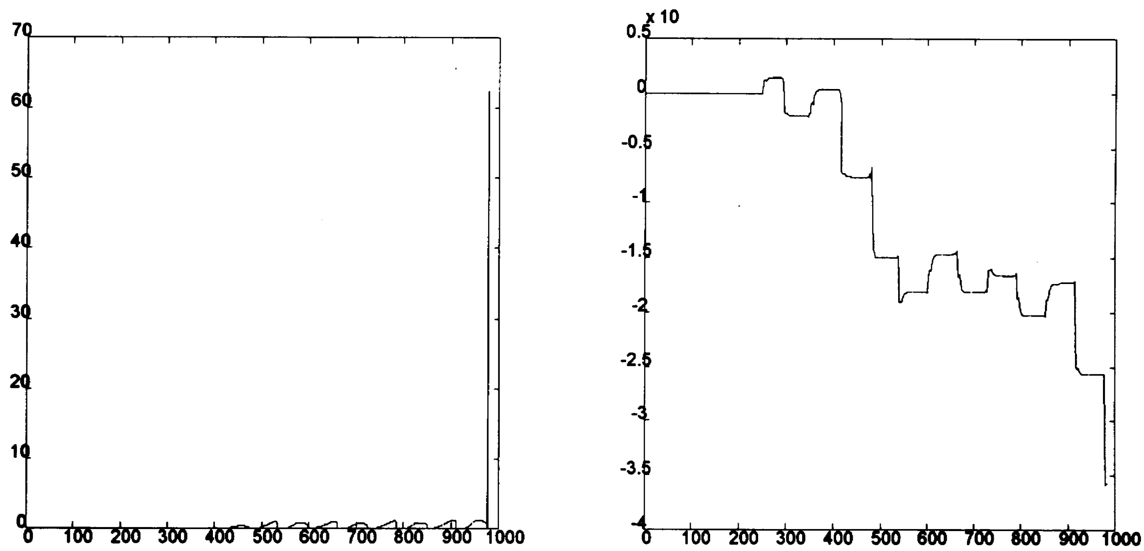

Fig. 7b. Behavior of the cumulative control when tuning α_3 and the viscosity is small: phenomenology test, and α_3 in the range of 10^{-5}.

https://doi.org/10.1017/S0263574797000192 Published online by Cambridge University Press

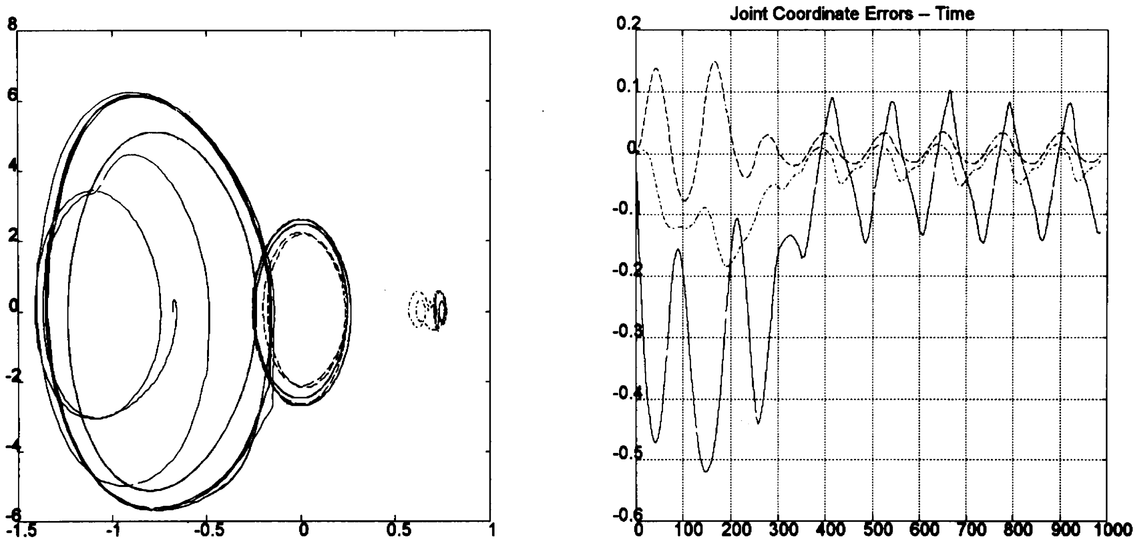

Fig. 8a. Tuning each of the continuous parameters in the case of small viscosity: phase-space, joint coordinate errors.

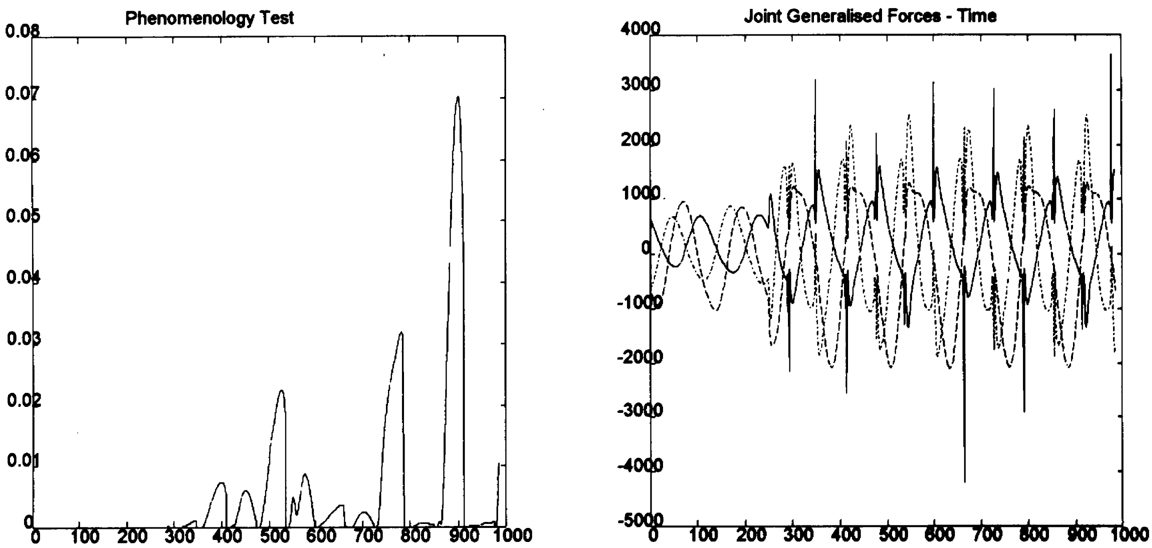

Fig. 8b. Tuning each of the continuous parameters in the case of small viscosity: phenomenology test, generalized forces to be exerted by the drives.

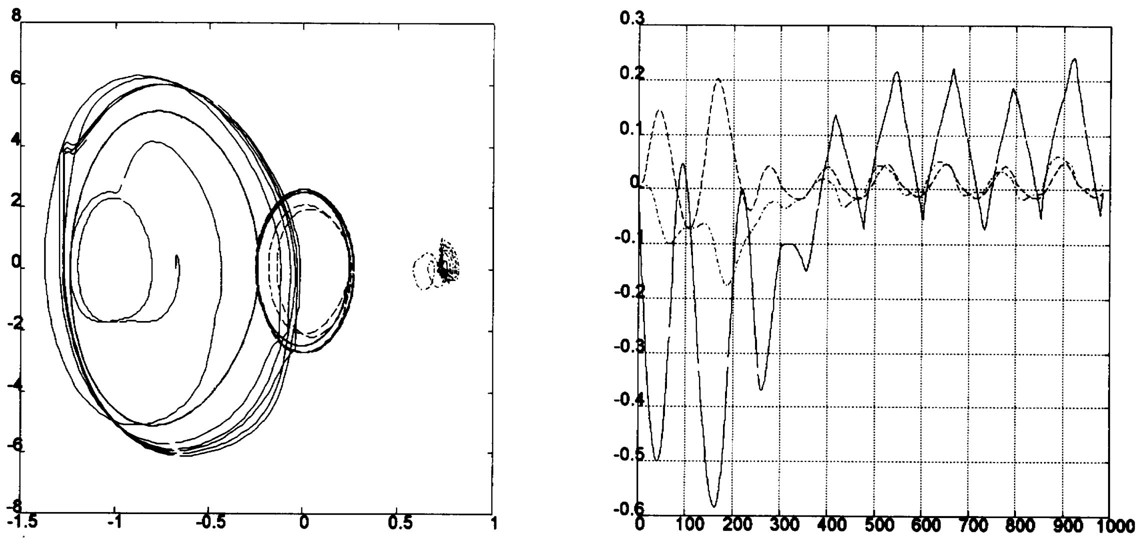

Fig. 9. The effect of a greater viscosity in the environmental coupled system: the phase-space, joint coordinate errors.

https://doi.org/10.1017/S0263574797000192 Published online by Cambridge University Press

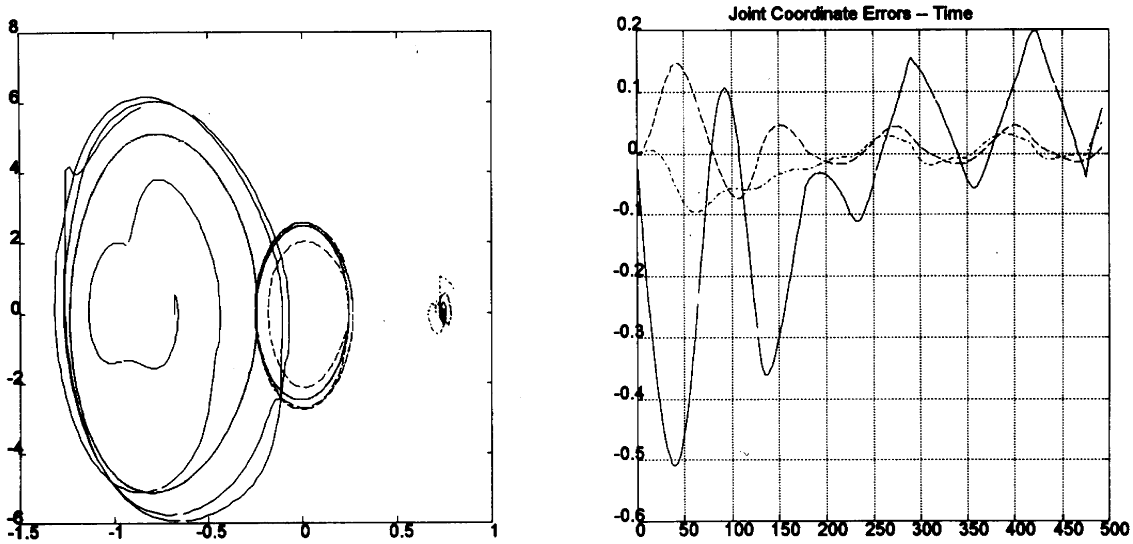

Fig. 10a. The effect of very high viscosity in the environmental coupled system: the phase-space, joint coordinate errors.

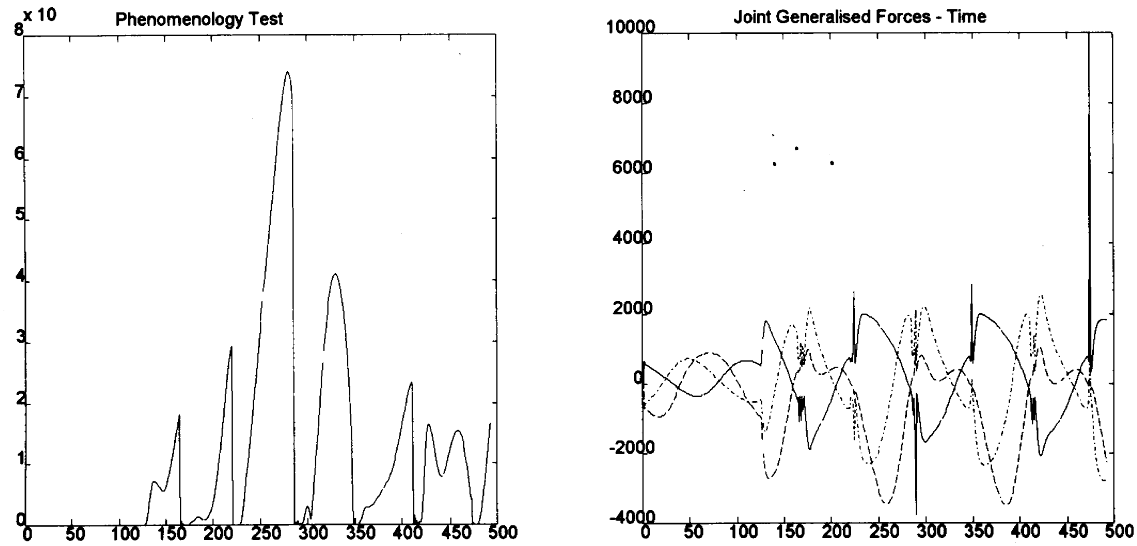

Fig. 10b. The effect of very high viscosity in the environmental coupled system: phenomenology test and the generalized forces.

It was also found, that from a well defined point a great variety of possible tuning strategies can be developed. By the use of a particular paradigm and computer simulations two typical versions were investigated: the *non-cumulative* and the *cumulative* approaches. A comparison of the results revealed that the *cumulative approach* seems to be far more effective that the *non-cumulative* one. The applied simple heuristic tuning strategy based on the consecutive tuning of the independent continuous parameters, according to the results of simple correlation-investigations concerning the accuracy of the prediction of the motion, shows stability near the unit transformation. This strategy was found to be less sensitive to the viscous interactions than its more heuristic progenitors starting the symplectizing algorithm from $\Im^T$ instead of **I**.

It is worthy of note, too, that the control strategy applied also contains important parameters which do not form part of the symplectic model. The possible effects of these parameters were not investigated in this paper. It is likely that further investigations in connection with different paradigms and parameters will be reasonable.

Acknowledgment

The authors gratefully acknowledge the grant provided by the "Fund for the Hungarian Higher Education" (AMFK'95) and that of the "National Scientific Research Fund" (OTKA T019032).

References

1. S. Tosunoglu, "Dynamic Modeling of Compliant Robots for Modular Environment and Fault Tolerant Operations" *Proc. of Dynamics and Stability,* International Conference ob Spatial Mechanisms and High Class Mechanism, October 4–6, 1994, Alma Ata, Kazakhstan (1994) **Vol. 3,** pp. 27–34.
2. S. Tosunoglu, "Fault Tolerant Control of Mechanical Systems" *Proc. of IEEE 21st International Conference on Industrial Electronics* (*IECON'95*), 6–10 November 1995, Orlando, Florida (1995) **Vol. 1,** pp. 110–115.

3. B. Lantos, ‘‘Identification and adaptive control of robots’’ *Int. J. Mechatronics* **2,** No. 3, 149–166 (1993).
4. L. Jánossy, *Theory of Relativity based on Physical Reality* (Akadémiai Kiadó, Budapest, Hungary, 1971).
5. J.K. Tar, O.M. Kaynak, J.F. Bitó, I.J. Rudas & D. Mester, ‘‘A New Method for Modelling the Dynamic Robot-Environment Interaction Based on the Generalization of the Canonical Formalism of Classical Mechanics’’ (Accepted for publication in the ‘‘First International ECPD Conference’’, Athens, Greece, July, 1995).
6. G. Szanosi, ‘‘Polyphonic Music and Classical Physics, the Origin of the Newtonian Concept of Time’’ *History of Science* **28,** 175–191 (1990).
7. V.I. Arnold, ‘‘Mathematical Methods of Classical Mechanics’’ (original issue in Russian by ‘‘*Nauka*’’) Hungarian translation issued by *Müszaki Könyvkiadó Budapest,* Hungary 1985).
8. I.J. Rudas, J.F. Bitó & J.K. Tar, ‘‘An Advanced Robot Control Scheme Using ANN and Fuzzy Theory Based Solutions’’ *Robotica* **14,** part 2, 189–198 (1996).