A NEW APPROACH IN THE USE OF VARIOUS KINDS OF GEOMETRY AS AI TOOLS IN ROBOT CONTROL: THE IDEA OF MINIMUM OPERATION TRANSFORMATIONS

József K. Tar*, J.F. Bitó**, I.J. Rudas*, H.M. Amin Shamshudin***

*Institute of Mathematical and Computational Sciences,
**Centre of Robotics and Automation,
John von Neumann Faculty of Informatics, Budapest Polytechnic, Budapest, Hungary
***Centre for Artificial Intelligence and Robotics, Faculty of Electrical Engineering,
Universiti Teknologi Malaysia, Johor, Malaysia

Abstract: A new branch of Soft Computing (SC) designed for the adaptive control of a special class of non-linear coupled multivariable systems is reported. In contrast to traditional SC its uniform structures are obtained from certain abstract geometry-related Lie groups. Adavantages are: a priori known and reduced structure size, increase in lucidity; simple, short, and explicit algebraic procedure instead of intricate learning. Disadvantage is: limited circle of applicability. Convergence consideratons for the new approach are discussed. Simulations are presented for the control of the inverted pendulum using the generalized Lorentz Group. It is concluded that the method is promising and probably imposes acceptable convergence requirements in many practical cases.

Keywords: Adaptive control, Learning control, Model reduction, Identification algorithms, Abstract geometries, Lie groups.

1. INTRODUCTION

The basic components of SC was almost completely developed by the sixties. In our days SC roughly means a combination of neural networks and fuzzy controllers equipped with several deterministic, stochastic or combined parameter-tuning or "learning" methods. It is becoming more and more popular since evades the development of intricate analytical system models. Instead of that typical problem classes has been identified for which typical uniform SC architectures can be aplied (e.g. multilayer perceptron, Kohonen-network, Hopfield-network, etc.). Fuzzy systems usually use membership functions of typical (e.g. trapezoidal, triangular or step-like, etc.) shapes, and the fuzzy relations can also be utilized in a standardized way.
The "first phase" of using these methods, that is identification of the problem class and finding the appropriate stucture, normally is relatively easy. The following phase, i.e. determining the necessary structure-size and fitting its is far less easy. For neural networks certain solutions starts from a quite big initial network and apply dynamic pruning for getting rid of the "dead" nodes (Reed, 1993). An alternative method starts with small network, and the number of nodes is increased step by step (see e.g. in Fahlmann *et al.*, 1990, and Nabhan *et al.*, 1994). Due to the possible existence of "local optima" in "backpropagation training" inadequacy of a given number of neurons cannot be concluded simply. Improved "learning methods", also including stochastic elements, were seruiously improved in the last decade and released this problem (see e.g. in Magoulas *et al.*, 1997, Chen *et al.*, 1996, Kinnenbrock, 1994, Kanarachos *et al.*, 1998).

In spite of this development for strongly coupled non-linear multivariable systems SC still has considerable drawbacks. The number of the necessary fuzzy rules, as well as that of the necessary neurons in a neural network strongly increases with the degree of freedom and the intricacy of the problem. External dynamic interactions on which no satisfactory information is available for the controller influences the system's behavior in dynamic manner. Both the big structure-size and huge number of tuneable parameters, as well as the time-varying "goal" still mean serious problem.

As the cause of the above problems "generality" and "uniformity" of the "traditional SC structures" can be identified. This makes the idea rise that several "simplified" branches of SC could be developed for narrower problem classes if more specific problem-or task-features could be identified and taken into account in the uniform structures applied. The first steps in this direction were made in the field of Classical Mechanical Systems (CMSs) (Tar *et al.*, 1995), while further refinements were published by Tar *et al.*, 1997, 1999, 2000, on the basis of principles detailed e.g. by Arnold, 1985. This approach used the internal symmetry of CMSs, the ymplectic Group (SG) of Symplectic Geometry. The "result" of the "situation-dependent system identification" was a symplectic matrix correcting the effects of the inaccuracy of the rough dynamic modell used as well as the external dynamic interactions not modeled by the controller.

From purely mathematical point of view all the essential steps used in that control can be realized by other mathematical means than symplectic matrices. SG can be replaced by other Lie groups defined by some "fundamental quadratic expression". For proceeding in this line the convergence properties of the method must be investigated in the case when the particular Lie group used in the control does not describe any internal physical symmetry of the system to be controlled. These considerations are discussed in the sequel the results of which are illustrated by simulation for the control of the inverted pendulum by the use of the "Generalized Lorentz Group" (GLG), that is a potential Lie group seeming to be promising for this purpose.

2. THE CONTROL PROBLEM IN GENERAL

The control task can be formulated as follows: there is given some *imperfect model of the system* on the basis of which some *excitation* is calculated for a desired input $\mathbf{i}^d$ as $\mathbf{e}=\varphi(\mathbf{i}^d)$. The system has its *inverse dynamics* described by the *unknown function* $\mathbf{i}^r=\psi(\varphi(\mathbf{i}^d))=f(\mathbf{i}^d)$ and resulting in a realized $\mathbf{i}^r$ instead of the desired one, $\mathbf{i}^d$. (In Classical Mechanics these values are the *desired* and the *realized joint accelerations*, while the external free forces and the joint velocities serve as the parameters of this time-varying function.) Normally we can obtain information via observation only on the "net" function $f()$, which is the most cases explicitly varies in time. In general no practical tools are available for "manipulating" the nature of this function directly: we can *deform* only its actual input $\mathbf{i}^{d*}$ in comparison with the *desired one*. The aim is to achive and maintain the $\mathbf{i}^d=f(\mathbf{i}^{d*})$ state. We can directly manipulate only the nature of the *model function* $\varphi()$, too.

In analogy with the idea of "renormalization", for solving the above task let us consider a sequence of certain quadratic matrices $\{\mathbf{S}_n\}$ and a series of step-by-step deformed inputs as

$$\mathbf{i}_0;\ \mathbf{S}_1\mathbf{f}(\mathbf{i}_0)=\mathbf{i}_0;\ \mathbf{i}_1=\mathbf{S}_1\mathbf{i}_0;\ \ldots\ ;\ \mathbf{S}_n\mathbf{f}(\mathbf{i}_{n-1})=\mathbf{i}_0;\quad \mathbf{i}_{n+1}=\mathbf{S}_{n+1}\mathbf{i}_n. \tag{1}$$

If the $\mathbf{S}_n\to\mathbf{I}$ convergence is realized a solution is obtained for the control problem. The difficulty in making (1) definite is that it does not have unique solution for the appropriate matrices, and that in general the calculation of the inverse of matrices is very inefficient from computational point of view. It can be conveniently evaded if special restrictions are imposed on $\{\mathbf{S}_n\}$: *a) let its elements be the members of some special Lie group outlined in the next paragraph; this immediately makes the calculation of the inverse matrices efficient; b) this till doesnot resolve ambiguity; to resolve it further, simply realizable restrictions must be imposed on the allowed matrices by bringing them as close to the identity operator as possible, from a special point of view.* For this purpose forst the potential geometries and Lie groups are considered in the next paragraph. Following this the question of convergence will be discussed.

3. POTENTIAL LIE GROUPS

Let $\mathbf{G}$ be a nonsingular quadratic, otherwise arbitrary constant matrix. Let the set $\{\mathbf{v}^{(i)}\}$ a linearly independent full set of vectors corresponding to the dimensions of $\mathbf{G}$. Let this set called "*special according to* $\mathbf{G}$" if it satisfies the restrictions

$$\mathbf{v}^{(i)T}\mathbf{G}\mathbf{v}^{(j)}=G_{ij} \tag{2}$$

The elements of this set can form the columns of a special matrix $\mathbf{V}$ satisfying the equation

$$\mathbf{V}^T\mathbf{G}\mathbf{V}=\mathbf{G}\ \Rightarrow\ \mathbf{V}^{-1}=\mathbf{G}^{-1}\mathbf{V}^T\mathbf{G}, \tag{3}$$

that is the calculation of the inverse of such matrices in general is very easy and computationally cost-efficient. These matrices trivially form a group, consequently their inverse and products also are matrices "of the same kind", that is the member of the same group. They may have the determinant ±1 only. By the use of the unimodular sub-group built up from the generators **H**

$$\mathbf{G}\mathbf{H}+\mathbf{H}^T\mathbf{G}=\mathbf{0} \tag{4}$$

special Lie-groups can be constructed for which the concept of "being as close to the identical transformation as possible" gains definite meaning.

(If **G** corresponds to **I**, $\mathfrak{I}=\left[\begin{array}{c|c}\mathbf{0} & \mathbf{I}\\ \hline -\mathbf{I} & \mathbf{0}\end{array}\right]$, and $\mathbf{g}=\langle 1,1,1,-c^2\rangle$, the *Orthogonal, the Symplectic,* and the *Lorentz Group,* can be obtained, respectively, as the most frequently used groups in Physics --"*c*" is the velocity of light--). The appropriate special sets are the *orthonormal,* the *symplectic,* and the *Lorentzian sets.* In these examples **G** is either symmetric (**I**, **g**) or skew-symmetric ($\mathfrak{I}$), consequently **H** can be constructed of skew-symmetric or symmetric **J** matrices as $\mathbf{H}=\mathbf{G}^{-1}\mathbf{J}$. All the considerations formerly used for constructing a mapping between the *observed* and the *desired* behavior of the controlled system (e.g. Tar *et al.* 1997, 2000) can trivially be repeated in the case of such a group, supposing, that at least one element of the special sets in (2) can be an arbitrary non-zero vector. For using the "*Generalized Lorentz Group*" first a "fictitious dimension" can be "added" to the DOF dimensional problem. So for **G** the diagonal matrix $\mathbf{g}=\langle 1,\ldots,1,-c^2\rangle$ can stand. Now let the DOF dimensional vector **f** stand for the desired/observed joint coordinate acceleration, and let us start with the columns of the DOFxDOF dimensional unit matrix. In the first step let this set be rigidly so rotated that its first vector becomes parallel with **f**. It is easy to so construct the rotation operators that the orthogonal sub-space of the initial and the goal vectors remains unchanged. Let $\mathbf{e}^{(1)}=\mathbf{f}/\sqrt{\mathbf{f}^T\mathbf{f}}=\mathbf{f}/f$. It is trivial that the columns of the following matrix form a generalized Lorentzian set:

$$\left[\begin{array}{c|c|c|c|c}\mathbf{e}^{(1)}\sqrt{f^2/c^2+1} & \mathbf{e}^{(2)} & \ldots & \mathbf{e}^{(DOF)} & \mathbf{f}\\ \hline f/c^2 & 0 & \ldots & 0 & \sqrt{f^2/c^2+1}\end{array}\right] \tag{5}$$

In this solution the *physically interpreted vector* **f** is accomplished with a fictitious (DOF+1)$^{\text{th}}$ component, and it is placed into the last column of a generalized Lorentzian. On the basis of the group properties the proper "Lorentzian" transforming the observed acceleration into the desired one can be calculated as

$$\mathbf{L}=\begin{bmatrix}\ldots & \mathbf{f}^D\\ \ldots & \sqrt{f^{D2}/c^2+1}\end{bmatrix}\mathbf{g}^{-1}\begin{bmatrix}\ldots & \mathbf{f}^R\\ \ldots & \sqrt{f^{R2}/c^2+1}\end{bmatrix}^T\mathbf{g} \tag{6}$$

For control-technical purposes "*c*" may be an arbitrary positive constant. This construction is "as close to the unit matrix as possible" in the sense that its components contain stretch/shrink only in the one-dimensional space for which information is available, and leave the orthogonal sub-spaces unchanged.

4. STABILITY CONSIDERATIONS

The $\mathbf{f}(\mathbf{i}_n)\to\mathbf{i}_0$ requirement can be expressed in more or less restricted forms. For instance, assume, that there exists $0<K<1$ for which

$$\left\|\mathbf{f}(\mathbf{i}_n)-\mathbf{i}_0\right\|\le K\left\|\mathbf{f}(\mathbf{i}_{n-1})-\mathbf{i}_0\right\|\le\ldots\le K^n\left\|\mathbf{f}(\mathbf{i}_0)-\mathbf{i}_0\right\| \tag{7}$$

This requirement trivially guarantees the desired convergence. The question is whether it can be satisfactory (that is not too rigorous) from practical point of view. Let the **S** matrices be written in the form of $\mathbf{S}=\mathbf{I}+\sigma$. The desired convergence now means that $\sigma_n\to\mathbf{0}$. Consider a differentiable function **f**(), and three points in the space: **f**(**x**), **f**(**Sx**), and $\mathbf{S}\mathbf{f}(\mathbf{x})=\mathbf{i}_0$. In the iteration the *present error,* and the *next error* can be expressed as

$$\begin{gathered}\mathbf{H}^P=\mathbf{f}(\mathbf{x})-\mathbf{i}_0=\mathbf{f}(\mathbf{x})-\left[\mathbf{f}(\mathbf{x})+\sigma\mathbf{f}(\mathbf{x})\right]=-\sigma\mathbf{f}(\mathbf{x})\\ \mathbf{H}^n=\mathbf{f}(\mathbf{x}+\sigma\mathbf{x})-\mathbf{i}_0=\\ =\mathbf{f}(\mathbf{x})+\mathbf{D}\sigma\mathbf{x}-\mathbf{i}_0=\mathbf{D}\sigma\mathbf{x}-\sigma\mathbf{f}(\mathbf{x})\end{gathered} \tag{8}$$

In (8) **D** correponds to the derivative of **f**() for near zero σ matrices. For these

$$\begin{gathered}\frac{\left\|\mathbf{H}^n\right\|^2}{\left\|\mathbf{H}^P\right\|^2}=\\ =\frac{\mathbf{x}^T\sigma^T\mathbf{D}^T\left(\mathbf{D}\sigma\mathbf{x}-2\sigma\mathbf{f}(\mathbf{x})\right)+\mathbf{f}(\mathbf{x})^T\sigma^T\sigma\mathbf{f}(\mathbf{x})}{\mathbf{f}(\mathbf{x})^T\sigma^T\sigma\mathbf{f}(\mathbf{x})}\le\\ \le K^2<1\end{gathered} \tag{9}$$

is needed. Now suppose that **f**(**x**) is in the same order of magnitude as **x**, and that the function is "very flat" in the sense that its derivative **D** is very small. In this case the first term in the numerator of (15) can be negligible in comparison with the second one and the additional requirement $\mathbf{x}^T\sigma^T\mathbf{D}^T\sigma\mathbf{f}(\mathbf{x})>0$ may lead to the desired convergence. To show that this does not seem to be too irrealistic in many physical cases, consider e.g. CMSs in which **x** corresponds to the joint acceleration and

$$\mathbf{f}(\mathbf{x})=\mathbf{M}^{-1}\mathbf{M}\mathbf{x}+\mathbf{M}^{-1}(\mathbf{b}-\mathbf{b})\equiv\mathbf{A}\mathbf{x}+\mathbf{c} \tag{10}$$

in which **M** and **b** correspond to the *model inertia* and the Coriolis plus gravitational terms, respectively, while their counterparts denoted by tilde correspond to the real data. With small **M**, that is with under-estimated model inertia, and a model term **b** which can be arbitrarily set, the fulfillment of this requirement is not hopeless. For instance, if $\mathbf{M}=\mathcal{M}\mathbf{I}$,

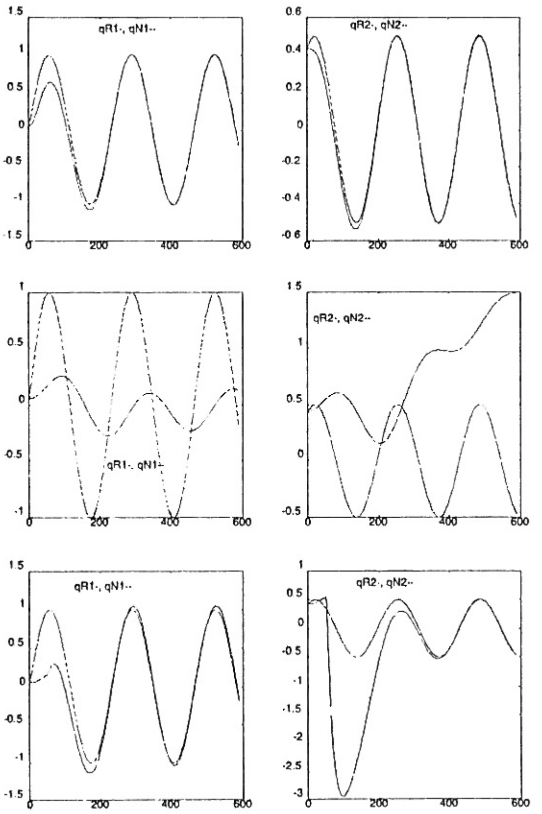

Fig. 1. Comparison of the exact CTC, the non-adaptive, and the adaptive control with Lorenzians

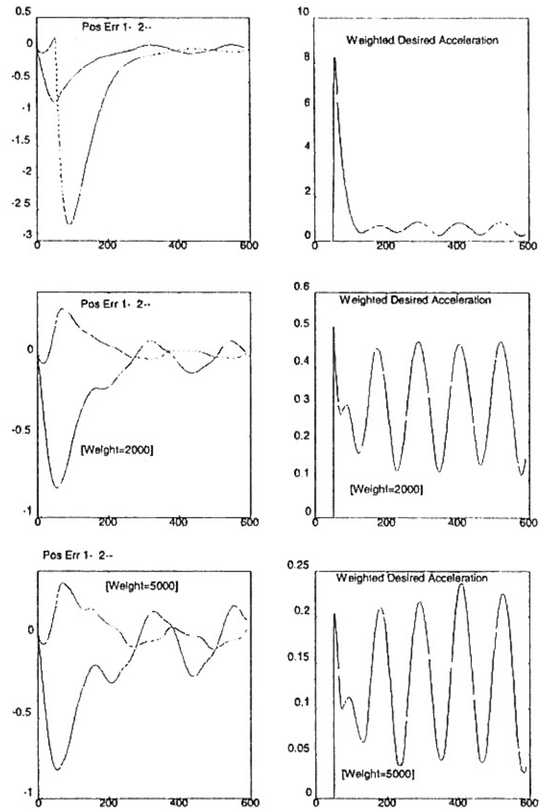

Fig. 2. The effect of the weight factor on the quality of the adaptive control joint coordinate errors and weighted desired accelerations

where $\mathcal{M}$ is a positive scalar, the appropriate term in (9) originating from **b** has the form of

$$\alpha\mathcal{M}\left(\mathbf{M}^{-1}\sigma\mathbf{x}\right)^{T}\sigma\left(\mathbf{M}^{-1}\sigma\mathbf{x}\right) \text{ if } \mathbf{b}=\alpha\sigma\mathbf{x} \qquad (11)$$

where α is real number. It is known only that the real invrse inertia matrix is positive definite (its concrete value is unknown), and σ is known in the control. By changing the sign of α the effect on the convergence can be monitored, while its absolute value can also be manipulated. In the next paragraph simulation examples are presented for the most "popular" paradigm, the control of the inverted pendulum.

5. SIMULATION RESULTS

The inverted pendulum has the usual structure with one linear and one rotational degree of freedom. First three kind of control are compared: a kinematically prescribed PD control with the *exact dynamic model*, the same kinematical control with the rough dynamic model *without and with adaptation*. (c=1 was chosen in the simulations; instead of varying it, a *weighting factor* w=1000 was chosen for "scaling" the joint accelerations (via division) before putting them into the Lorentzians. The proper part of the resulting Lorentzian was multiplied by this factor before using it in the control.) In Fig. 1 typical results are given regarding trajectory reproduction. It is evident that considerable adaptivity was achieved and that the adaptive control well approaches the exact CTC contro. Fig. 2 reveals that the quality of adaptivity considerably depends on the weighting factor and that it is advantageous if the "physically interpreted part" in the genberalized Lorenztians is comparable in norm with the value c=1. This observation is confirmed for slow required motion, too, according to Fig. 3.

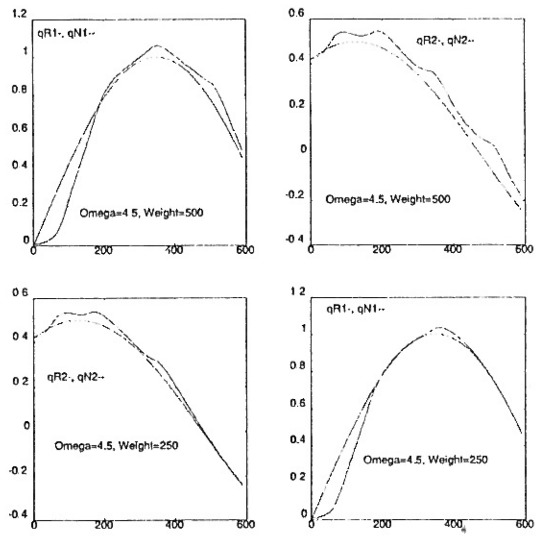

Fig. 3. The effect of the weighting factor for very slow motion along the same path as on Fig.1-2.

To reveal technical details in Fig. 4 the generalized forces to be exerted by the pendulum's drives, the square root of the convergence ratio in (9), the norm of the appropriate generalized Lorentzians are displayed for the fast motion and adaptive control. The norms of the Lorentzians are not constant but they are very close to the norm of the identity transformation of the 3×3 matrix (≅1,732) revealing that in the iteration the **S** matrices were really close to the identity transformation. The figures reveal that the $K<1$ restriction is too rigorous in the practice. In spite of the control sections in which the joint acceleration error is not monotone decreasing the system still is kept at bay. It can be interpreted as follows: though the joint accelerations not always are exactly the same as it is prescribed by the "*kinematic PD control*", the actually achieved accelerations makes the system follow the nominal trajectory with good accuracy. This supposition is proved via Fig. 5 for a slower motion along the same trajectory. The order of magnitude of the desired joint coordinate accelerations is between 1000 and 4000 units, while their error is between 0 and 60 units.

6. CONCLUSIONS

In this paper the possibility of developing a particular special branch of Soft Computing was investigated in which the uniform structures to be used may originate from different abstract Lie-groups. The new approach has two essential advantages in comparison with the "traditional means" of Soft Computing:

- The structure-size and the number of the free parameters is uniquely determined by the degree of freedom of the system to be controlled and the particular group chosen;
- Machine learning can be realized via simple, deterministic, definite algebraic steps limited in number; it is void of the problem of local optima.

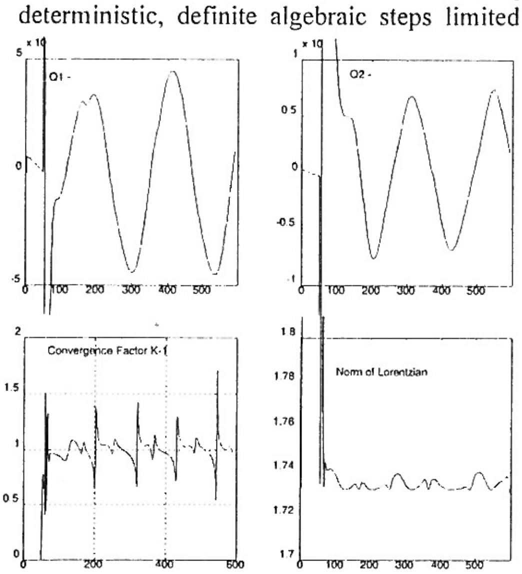

Fig. 4. The joint generalized forces (10^4 N and Nm), the square root of the convergence factor as defined in (9) and the norm of the generalzied Lorenztians for the motion in Figs. 1-2.

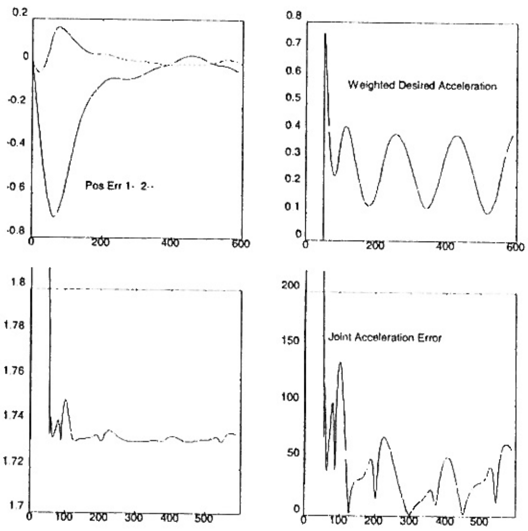

Fig. 5. Trajectory reproduction error, weighted joint coordinate accelerations (w=1000), norm of the Lorenztians, and the joint acceleration errors for the motion of medium velocity along the same trajectory.

The convergence properties of the learning algorithm were investigated in general. A possible particular criterion was suggested which was found to be met in practical cases as classical mechanical systems' control.

The applicability of the approach was demonstrated via simulation in the case of the inverted pendulum, controlled by the use of the Generalized Lorentz Group. It was found that the criterion suggested is even too rigorous for the real needs of adaptivity and it may be released in the future.

It can be expected that the here presented considerations can be extended to a wider class than the control of mechanical systems. In general seeking for different convergence criteria for the learning seems to be expedient.

7. ACKNOWLEDGMENT

The authors gratefully acknowledge the aid by the Malaysian-Hungarian Bilateral Technology Co-operation Project (2000-2001) supporting their work.

8. REFERENCES

Arnold, V.I. (1985): *Mathematical Methods of Classical Mechanics* (in Hungarian) "Mûszaki Könyvkiadó" Budapest, Hungary 1985.

Chen, C., Chang, W. (1996): A Feedforward Neural Network with Function Shape Autotuning. In *Neural Networks*, **9**, pp. 627-641, 1996.

Fahlmann, S., Lebiere, C. (1990): The Cascade-Correlation Learning Architecture. In: *Advances in Neural Information Processing Systems*, **2**, pp. 524-532, 1990.

Kanarachos, A., Geramanis, K. (1998): Semi-Stochastic Complex Neural Networks. In: *IFAC-CAEA '98 Control Applications and Ergonomics in Agriculture*, pp. 47-52, 1998.

Kinnenbrock, W. (1994): Accelerating the Standard Backpropagation Method Using a Genetic Approach. In: *Neurocomputing*, **6**, pp. 583-588, 1994.

Magoulas, G., Vrahatis, N., Androulakis, G. (1997): Effective Backpropagation Training with Variable Stepsize. In: *Neural Networks*, **10**, pp. 69-82, 1997.

Nabhan, T., Zomaya, A. (1994): Toward Generating Neural Network Structures for Function Approximation. In: *Neural Networks*, 7, pp. 89-9, 1994.

Reed, R. (1993): Pruning Algorithms - A Survey. In: *IEEE Transactions on Neural Networks*, **4**., pp. 740-747, 1993.

Tar, .J.K., Kaynak, O.M., Bitó, J.F., Rudas, I.J. (1995): Generalised Canonical Equations of Classical Mechanics: A New Approach towards the Adaptive Treatment of Unmodelled Dynamics. In Proc. of the *International Conference on Recent Advances in Mechatronics (ICRAM'95)*, 14-16 August, 1995, Istanbul, Turkey, pp. 1092-1097.

Tar, J.K., Rudas, I.J., Bitó, J.F. (1997): Group Theoretical Approach in Using Canonical Transformations and Symplectic Geometry in the Control of Approximately Modeled Mechanical Systems Interacting with Unmodelled Environment. In: *Robotica*, **Vol. 15**, pp. 163-179, 1997.

Tar, J.K., Rudas, I.J., Bitó, J.F., Kaynak, O.M. (1999): A New Utilization of the Hamiltonian Formalism in the Adaptive Control of Mechanical Systems Under External Perturbation. In: Intelligent *Automation and Soft Computing*, **Vol. 5**, Issue:4, 1999, pp. 303-312.

Tar, J.K., Rudas, I.J., Balladji, Touré (2000): Minimum Operation Algebraic Adaptive Control for Robot Based on a Novel Branch of Soft Computing. In *Proc. of EUREL '2000, Vol. 2, Session 3, (European Advancded Robotics Systems - Masterclass and Conference - Robotics 2000)*, 12-14 April 2000, The University of Salford, Manchester, UK.