

Neural computations underlying the exertion of force: a model

Alexander V. Lukashin, Bagrat R. Amirikian, Apostolos P. Georgopoulos

Brain Sciences Center, Department of Veterans Affairs Medical Center, Minneapolis, USA; and Department of Physiology and Neurology, University of Minnesota Medical School, Minneapolis, Minnesota, USA

Received: 13 September 1995/Accepted in revised form: 9 January 1996

Abstract. We have developed a model that simulates possible mechanisms by which supraspinal neuronal signals coding forces could converge in the spinal cord and provide an ongoing integrated signal to the motoneuronal pools whose activation results in the exertion of force. The model consists of a three-layered neural network connected to a two-joint-six-muscle model of the arm. The network layers represent supraspinal populations, spinal cord interneurons, and motoneuronal pools. We propose an approach to train the network so that, after the synaptic connections between the layers are adjusted, the performance of the model is consistent with experimental data obtained on different organisms using different experimental paradigms: the stiffness characteristics of human arm; the structure of force fields generated by the stimulation of the frog's spinal cord; and a correlation between motor cortical activity and force exerted by monkey against an immovable object. The model predicts a specific pattern of connections between supraspinal populations coding forces and spinal cord interneurons: the weight of connection should be correlated with directional preference of interconnected units. Finally, our simulations demonstrate that the force generated by the sum of neural signals can be nearly equal to the vector sum of forces generated by each signal independently, in spite of the complex nonlinearities intervening between supraspinal commands and forces exerted by the arm in response to these commands.

1 Introduction

Recordings of neuronal activity in the motor cortex of behaving animals have shown that the direction and overall trajectory of arm movements are markedly

related to the activity of motor cortical cells (Georgopoulos et al. 1982, 1986, 1993; Kalaska et al. 1989; Caminiti et al. 1990; Hocherman and Wise 1991; Schwartz 1994). These observations have been extended to the study of forces exerted by the arm against an immovable object (Georgopoulos et al. 1992). A monkey was trained to exert forces on an isometric handle in the presence of a constant force bias. First the monkey was required to exert a postural (static) force P , which compensated a given bias force B . After a holding period, a cue instructed the monkey to exert a force S so that the net force acting on the handle (i.e., the force exerted by the monkey S plus the bias force applied to the handle B) was in a visually specified direction N . Note that the net force is equivalent to the incremental (dynamic) component I of the force exerted by the subject: $I = S - P = S + B = N$ (Fig. 1, left template). Eight net force directions and eight bias force directions were employed. Two principal findings were reported: (i) the activity of single cells showed approximately the same directional tuning properties when the arm exerts a force without moving as when it moves through space (Georgopoulos et al. 1993), and (ii) as a population, the activity of cells reflected the direction of the incremental force I and not the actual force S exerted by the subject. Accordingly, it was hypothesized (Georgopoulos et al. 1992; Georgopoulos 1994) that the postural and incremental signals are controlled by different neural systems, and that these signals would converge in the spinal cord and produce a resulting integrated signal to the motoneuronal pools.

The hypothesis concerning an integration of separate force signals implies a 'linear summation rule'. Indeed, it is seen in Fig. 1 (left template) that the postural signal, which generates the force P , and the incremental signal, which generates the force I , should converge at the level of spinal cord in such a way that the integrated signal would generate the force S that is equal to the vector sum $P + I$. Only in this case will the net force acting on the handle ($S + B$) be equal to the desired net force N . Although an appropriate summation of neuronal signals is

Correspondence to: A.V. Lukashin, Brain Science Center (11 B), VA Medical Center, One Veterans Drive, Minneapolis, MN 55417, USA

conceivable (Redish and Touretzky 1994), the linear summation of *forces* seems unlikely, due to the complex nonlinearities that characterize the mechanical properties of limbs, and the interactions both among neurons and between neurons and muscles (Bizzi et al. 1991; Kalaska and Crammond 1992). However, recent evidence derived by focal microstimulation of the frog's spinal cord (Bizzi et al. 1991; Giszter et al. 1993; Mussa-Ivaldi et al. 1994) has revealed that the simultaneous activation of two distinct spinal sites leads to the vectorial summation of the end-point forces generated by the stimulation of each site separately. Although spinal microstimulation studies have not been performed in primates, it is reasonable to hypothesize a similar plan of spinal organization. Then the synaptic connections between the supraspinal populations and the spinal cord could provide a concomitant activation of a number of the spinal sites associated with different force fields, which are summed in a linear fashion.

The purpose of the present article is to propose and analyze a model that suggests possible mechanisms by which the supraspinal commands could be integrated at the spinal cord level and translated into exertion of a required force. The model consists of a three-layered neural network connected to a two-joint-six-muscle model of the arm. The layers of the network represent supraspinal neuronal populations, spinal cord interneurons, and motoneuronal pools, respectively. The key idea is *to train* the network so that the model reproduces quantitatively experimental data for stiffness characteristics of human arm (Mussa-Ivaldi et al. 1985; Shadmehr et al. 1993; Tsuji et al. 1995), and then *to test* the model by comparing its performance with experimental data obtained from recordings of cell activity in the motor cortex of monkey (Georgopoulos et al. 1992) and by microstimulations of the frog's spinal cord (Bizzi et al. 1991; Giszter et al. 1993; Mussa-Ivaldi et al. 1994). In particular, the model reconciles the nonlinearities that intervene between supraspinal commands and forces exerted by the arm with the hypothesis concerning the linear summation of forces encoded in activities of potentially separated supraspinal neuronal ensembles.

2 Model

2.1 Neural network model

The architecture of the model is shown in Fig. 1 (right). The direction and magnitude of incremental force \mathbf{I} is encoded in activities of directionally tuned motor cortical cells (Georgopoulos et al. 1992). Although the supraspinal representation of postural force \mathbf{P} is less known (see, however, Kalaska et al. 1989; Wise 1993; Georgopoulos 1994), we assume the same distributed representation as for the incremental force. Each population consists of n supraspinal (SS) units of m neurons each. All neurons belonging to the same SS unit possess the same preferred directions \mathbf{C}_i ($i = 1, \dots, n$) with cosine-like tuning function [see (1) below], and the preferred directions \mathbf{C}_i are uniformly distributed in two-dimensional

space (Georgopoulos et al. 1993; Georgopoulos 1994). Therefore, if the activity of the supraspinal population represents a force signal \mathbf{F} ($\mathbf{F} = \mathbf{I}$ or \mathbf{P}), then the activity of the i th SS unit $V_i^{\text{SS}}(\mathbf{F})$ is given by the expression:

$$V_i^{\text{SS}}(\mathbf{F}) = \frac{a_F}{2}(1 + \cos \theta_{FC}) \quad (1)$$

where coefficient a_F is proportional to the number of cells recruited to represent the magnitude of force \mathbf{F} [i.e., the larger the magnitude of the encoded force, the more neurons from a given SS unit are recruited (Evarts et al. 1983; Redish and Touretzky, 1994)], and θ_{FC} is the angle between the direction of force \mathbf{F} and the unit's preferred direction \mathbf{C} .

All units from both supraspinal populations are connected to the four units at the IN layer representing the level of spinal cord interneurons (Fig. 1). To calculate activities of the IN units V_j^{IN} ($j = 1, \dots, 4$) we use the following sigmoid activation function:

$$V_j^{\text{IN}} = \frac{1}{2}[1 + \tanh(T_j + U_j^{\text{SS}})] \quad (2)$$

where T_j is a synaptic input furnished by sources (not shown in Fig. 1) other than the populations coding forces. These signals provide an initial pattern of activity, needed to establish an initial position of the arm (see Sect. 3), before the commands coding forces \mathbf{P} and \mathbf{I} reach the IN layer. The synaptic input U_j^{SS} generated by supraspinal force signals is given by the sum:

$$U_j^{\text{SS}} = \sum_{i=1}^n w_{ji}[V_i^{\text{SS}}(\mathbf{P}) + V_i^{\text{SS}}(\mathbf{I})] \quad (3)$$

where w_{ji} is the weight of connection between the i th unit from the population coding force \mathbf{P} or \mathbf{I} and the j th unit belonging to the IN layer (for the sake of simplicity we assume the same connectivity matrix for both supraspinal populations).

Each of the IN units is connected to all six units at the MN layer representing the level of motoneuronal pools (Fig. 1). Activities of MN units V_k^{MN} ($k = 1, \dots, 6$) are calculated using the relation:

$$V_k^{\text{MN}} = \frac{1}{2}[1 + \tanh(U_k^{\text{IN}})] \quad (4)$$

where U_k^{IN} is a synaptic input generated by IN units:

$$U_k^{\text{IN}} = \sum_{j=1}^4 z_{kj} V_j^{\text{IN}} \quad (5)$$

where z_{kj} is the weight of connection between the j th unit from the IN layer and the k th unit from the MN layer.

2.2 Model of the arm

We use a planar two-joint-six-muscle model utilized previously (Hogan 1985; Flash 1987; Katayama and Kawato 1993). The arm is modeled by two rigid segments ('upper arm' and 'forearm') of equal length (Fig. 1). The upper arm is attached proximally to an immovable 'clavicle' via the shoulder joint and distally to the forearm via the elbow joint. Both single- and double-joint muscles

are included. One end of both flexor and extensor single-joint muscles controlling the shoulder joint is attached to the clavicle at a distance b from the shoulder joint (flexor and extensor are attached at different sides of the joint). The other ends of these muscles are attached to the upper arm segment at a distance L from the shoulder joint (for graphical reasons these distances are not shown in Fig. 1). Single-joint muscles controlling the elbow joint are attached to the upper arm segment at the distances b from the elbow joint and to the forearm segment at a distance L from the elbow joint. Two-joint muscles are attached at distances b from each joint (Fig. 1). Note that the muscle attachment implies that the moment arms of muscles depend upon the joint angles. Below we assume that the distance between the shoulder and elbow joints is equal to L , and use the limit $b \ll L$; specifically, we set $b = 0.01$ m, $L = 0.33$ m.

All muscles are modeled by nonlinear springs with the following length-tension relationship (Feldman 1966; Shadmehr and Arbib 1992):

$$f_k = \alpha \{ \exp[\beta(l_k - l_k^0)] - 1 \} \quad (l_k > l_k^0) \quad (6)$$

$$f_k = 0 \quad (l_k \leq l_k^0)$$

where f_k is a contraction force developed by the k th muscle, l_k is the actual muscle length and l_k^0 is the muscle intrinsic rest length; parameters α and β are constants, and $\alpha = 10$ N, $\beta = 100$ m⁻¹. These specific values of parameters α and β were chosen to fit by (6) the experimental data of Feldman (1966, as presented in fig. 3C of the paper by Shadmehr and Arbib 1992).

2.3 Interactions between neurons and muscles

Each unit from the MN layer innervates one muscle and controls the muscle intrinsic rest length l_k^0 (Feldman 1966; Shadmehr and Arbib 1992). We use the following relation between l_k^0 and activity V_k^{MN} of the motoneuronal unit:

$$l_k^0 = l^{\text{max}} + V_k^{\text{MN}}(l^{\text{min}} - l^{\text{max}}) \quad (7)$$

where l^{min} and l^{max} are the minimal and maximal muscle rest length, respectively, and $l^{\text{min}} = 0.26$ m, $l^{\text{max}} = 0.30$ m. Equation (7) implies that the higher is the activity V_k^{MN} , the shorter is the muscle rest length l_k^0 .

2.4 Equilibrium position and restoring force field

Given a particular set of muscle rest lengths l_k^0 ($k = 1, \dots, 6$), a corresponding equilibrium configuration of the arm can be calculated as follows. Let a configuration of the arm be defined by a pair of shoulder and elbow joint angles (φ_s and φ_e , respectively; $0^\circ \leq \varphi_s \leq 135^\circ$, $0^\circ \leq \varphi_e \leq 180^\circ$). Then the set of actual muscle lengths l_k ($k = 1, \dots, 6$) is uniquely defined by the geometry of the muscle attachments. If muscle rest lengths l_k^0 and muscle actual lengths l_k are known, the forces produced by muscles can be calculated using (6). Let

$(f_1 f_2 f_3 f_4 f_5 f_6)^T$ be the vector of muscle forces (subscripts 1 and 2 refer to the shoulder single-joint flexor and extensor, subscripts 3 and 4 refer to the elbow single-joint flexor and extensor, and subscripts 5 and 6 refer to the double-joint flexor and extensor). The transformation between the vector of muscle forces and the vector of net joint torques $(t_s t_e)^T$, where t_s and t_e are shoulder and elbow torques, respectively, is given by the matrix of muscle moment arms. For the model described above, this relationship is:

$$\begin{pmatrix} t_s \\ t_e \end{pmatrix} = \begin{pmatrix} -r_s & r_s & 0 & 0 & -r_s & r_s \\ 0 & 0 & -r_e & r_e & -r_e & r_e \end{pmatrix} (f_1 f_2 f_3 f_4 f_5 f_6)^T \quad (8)$$

where $r_s = b \sin \varphi_s$, $r_e = b \sin \varphi_e$. The equilibrium configuration of the arm is defined by the condition that both net joint torques, t_s and t_e , are equal to zero. Consequently, the equilibrium joint angles φ_s^{eq} and φ_e^{eq} can be obtained as the solution of the system of two nonlinear equations (8) under conditions $t_s = 0$ and $t_e = 0$.

The force exerted by the end-point of the arm is the result of the difference between current and equilibrium end-point positions. Generally, each equilibrium position is characterized by a field of restoring forces generated by muscles in response to displacements of the end-point from equilibrium position. For any given set of muscle rest lengths, the restoring force field can be calculated as follows. Let a current end-point position be defined by joint angles φ_s and φ_e . Then (8) can be used to calculate the joint torques t_s and t_e [in particular, if the arm configuration is such that $\varphi_s = \varphi_s^{\text{eq}}$ and $\varphi_e = \varphi_e^{\text{eq}}$, (8) gives $t_s = 0$ and $t_e = 0$]. The transformation between net joint torques and x-y components of restoring end-point force (Q_x and Q_y) is given by the inverse Jacobian matrix:

$$\begin{pmatrix} Q_x \\ Q_y \end{pmatrix} = \frac{1}{L \sin \varphi_e} \begin{pmatrix} \cos(\varphi_s + \varphi_e) & -\cos(\varphi_s + \varphi_e) - \cos \varphi_s \\ \sin(\varphi_s + \varphi_e) & -\sin(\varphi_s + \varphi_e) - \sin \varphi_s \end{pmatrix} \begin{pmatrix} t_s \\ t_e \end{pmatrix} \quad (9)$$

The vector field $\{Q_x(\varphi_s, \varphi_e), Q_y(\varphi_s, \varphi_e)\}$ is the restoring force field associated with the given set of muscle rest lengths.

2.5 Performance of the model

All parameters related to the arm construction and elastic properties of muscles will be fixed throughout the rest of the paper. Since the restoring force field and equilibrium end-point position are uniquely determined by activities of the MN units (6–9), the relation between force signals encoded in supraspinal populations (input) and the force exerted by the arm (output) is determined by synaptic connections w_{ji} and z_{kj} between neurons in different layers (Fig. 1, Equations 1–5). Therefore, a desired performance of the model can be attained by training of the network, i.e., by an adjustment of synaptic connections.

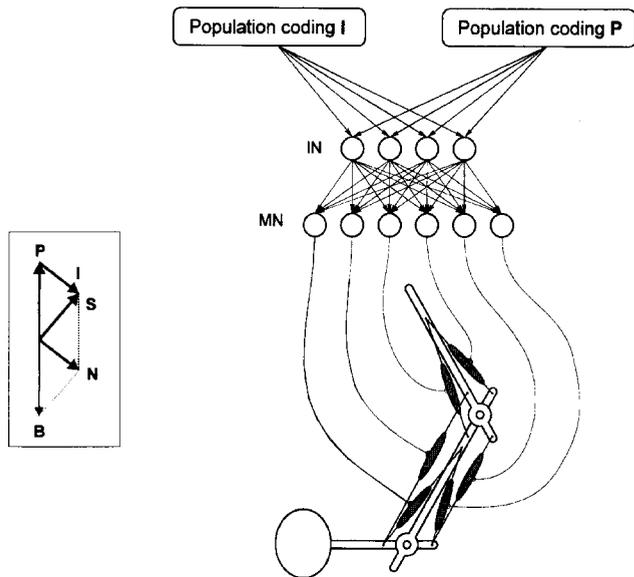


Fig. 1. *Left template:* Forces defined in the text: the force bias B , the postural force P compensating the force bias, the force exerted by the subject S , and the incremental component I of the force S . The net force $S + B$ acting on the handle is equal to the visually instructed force N . *Right:* The model used to simulate the translation of supraspinal commands into a force exerted by the arm. The postural force P and incremental force I are encoded in the activities of two separated populations. Signals coding these forces converge at the level of spinal cord interneurons (the IN layer) and produce an integrated signal to the motoneuronal pools (the MN layer). The pattern of motoneuronal activity determines the force exerted by the arm

3 Training-testing procedure

In this section we describe in detail the method used to adjust the synaptic connections. First the synaptic connections z_{kj} between the IN and MN layers have been adjusted so that any pattern of activity at the IN layer produces a restoring force field whose structure is consistent with experimental data obtained for human arm. In particular, each IN unit is attributed to a specific point in the workspace, which is the equilibrium end-point position produced by the activation of this unit ('directional preference' of the unit given a reference position in the workspace). Then the synaptic connections w_{ji} between supraspinal populations and the IN units have been assigned so that activities of the SS units weighted in accordance with their preferred directions C_i are transformed into activities of the IN units weighted in accordance with their directional preferences in the workspace of the arm (connections z_{kj} between the IN and MN units are kept fixed to retain the structure of force fields). Finally, the overall performance of the model has been tested by independent or simultaneous activations of supraspinal populations.

3.1 Connections between the IN and MN units

Forces exerted by individual muscles depend upon the muscle rest lengths (6), which are controlled by activities of the corresponding MN units (7). Therefore, activities

of different MN units should be correlated to produce muscle synergies that ensure the desired characteristics of restoring force fields. Since any unit from the IN layer is connected to all units at the MN layer (4, 5), correlated patterns of MN activity can be provided by a set of 'synergetic' synaptic connections between the IN and MN units. The problem we address is to find a set of connections z_{kj} such that the restoring force fields generated by the model would be similar to the fields measured for human arm (Mussa-Ivaldi et al. 1985; Shadmehr et al. 1993; Tsuji et al. 1995). In these studies, human arm stiffness characteristics have been determined by measuring the restoring forces for displacements from an equilibrium end-point position. The stiffness was represented graphically by an ellipse, characterized by its size, shape, and orientation. The results indicated that the shape and orientation of the stiffness ellipse are strongly dependent on arm configuration in an orderly fashion (these parameters, however, remain invariant among subjects and over time). To adjust the performance of the model to these experimental data we used the following procedure (Fig. 2).

Four equilibrium positions of the arm were chosen in the workspace (in Fig. 2A these positions are designated as 'T'). For each position, we calculated the set of muscle rest lengths l_k^0 ($k = 1, \dots, 6$) ensuring the size, shape, and orientation of the stiffness ellipse in close agreement with these characteristics measured experimentally (Mussa-Ivaldi et al. 1985) for the human arm at a similar location in the workspace. Note that due to the presence of more muscles than there are joints, different sets of muscle rest lengths may result in the same equilibrium position of the arm. However, it is this redundancy that makes it possible to find a required set of muscle rest lengths imposing the stiffness characteristics as additional conditions. Starting from (8) we have derived explicit relations between muscle rest lengths and the stiffness characteristics of the arm using straightforward algebra, cumbersome as it is (see Appendix).

Each of these equilibrium positions was attributed to one of the IN units: namely, for each of four IN units ($j = 1, \dots, 4$), six parameters z_{kj} ($k = 1, \dots, 6$) were adjusted so that the maximal activity of the j th IN unit (whereas other units are inactive) generates the attributed equilibrium position.

The resulting connectivity matrix z_{kj} ($k = 1, \dots, 6$; $j = 1, \dots, 4$) was fixed and used during the following testing procedure. Different patterns of IN activity were sampled to produce new equilibrium positions, for which stiffness ellipses were calculated. Six of these ellipses are depicted in Fig. 2A. It is seen that the shape and orientation of ellipses strongly depend on arm configuration, reproducing well two main features of the human arm stiffness (Mussa-Ivaldi et al. 1985; Tsuji et al. 1995): (i) the direction of maximum stiffness at any location is approximately oriented along a radial line joining the hand to the shoulder, and (ii) the stiffness ellipse becomes more anisotropic as the hand approaches distal positions.

Since the stiffness describes the relation between force and displacement vectors only in the vicinity of equilibrium, the performance of the model was further tested for

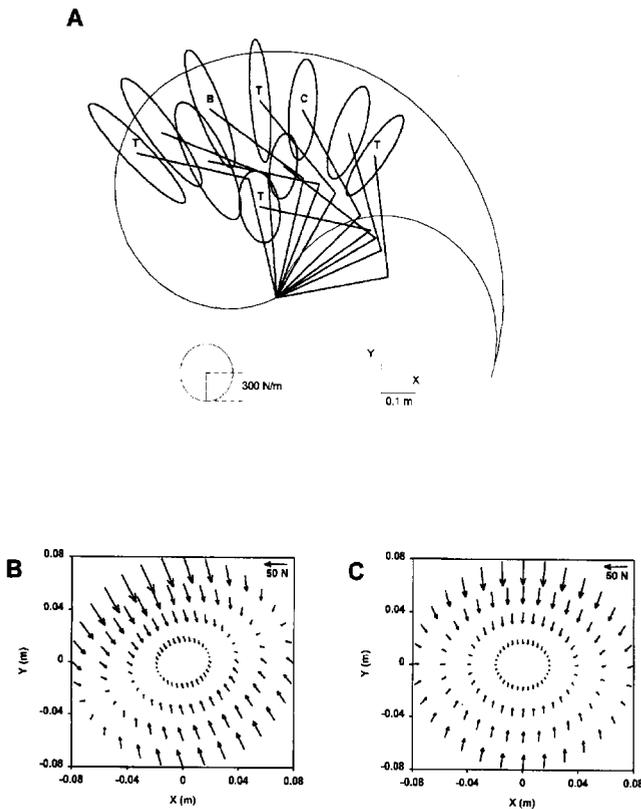


Fig. 2. A Stiffness ellipses at different locations in the workspace of the arm. The two-joint arm is shown schematically by two line segments. The workspace is bounded by a thin curve. The calibration for the stiffness is given by the circle. Four ellipses designated as *T* were used to find weights of connections between the IN and MN layers (see text). The end-point positions in the centers of these ellipses correspond to equilibrium position produced by the activation of one of the IN units, whereas all other units are inactive. Six other ellipses represent examples of the end-point stiffness calculated at different locations to test the performance of the model. B, C Restoring force fields calculated for equilibrium positions designated as *B* and *C* in A. Each force vector is represented by an arrow drawn with its tail at the tip of the end-point displacement vector

relatively large displacements from equilibrium positions. In Fig. 2B and 2C two examples of restoring force fields associated with two different equilibrium positions are shown. The structure of these fields, including a non-linearity with respect to displacement, is similar to that measured experimentally (Shadmehr et al. 1993) for the same range of displacements.

In summary, the use of experimental data for a few equilibrium positions as training ‘examples’ makes it possible to find the connectivity matrix z that provides the required synergy of muscles at other locations in the workspace.

3.2 Connections between the SS and IN units

The connectivity matrix w transforms a force signal encoded in activities of SS units, which are characterized by preferred directions C_i (1), into a signal encoded in activities of IN units, which are attributed to specific

equilibrium positions in the workspace. Consider a point chosen approximately at the center of the workspace. Let D_j be a unit vector pointed from the center of the workspace to the equilibrium end-point position attributed to the j th IN unit (in a sense, D_j is the preferred direction vector of the j th IN unit). We propose the following expression for the connectivity matrix:

$$w_{ji} = \frac{4}{n} \cos \theta_{D_j C_i} \quad (10)$$

where θ_{DC} is the angle between vectors D and C , and $4/n$ is a normalizing coefficient (n is the number of SS units). The idea underlying the use of the above relation is that (10) provides a negative correlation between connection strength w_{ji} and angle between vectors D_j and C_i , i.e., the connection strength is correlated with directional preference of connected units. The cosine function (10) and the normalizing coefficient are chosen to obtain activities of IN units V_j^{IN} (2) in a simple explicit form. Indeed, substituting (1, 3, 10) into (2), one has

$$V_j^{\text{IN}} = \frac{1}{2} [1 + \tanh(T_j + a_p \cos \theta_{PD_j} + a_I \cos \theta_{ID_j})] \quad (11)$$

[to derive (11) we have used the uniformness of distribution of preferred directions C_i ($i = 1, \dots, n$) and straightforward trigonometric relationships].

In summary, the connectivity matrix w (10) transforms supraspinal signals coding forces P and I (1) into a pattern of activity of the IN units (11) in a way that the closer the direction of force to be exerted (either P or I) is to the direction of vector D_j , the higher is the activity of the j th IN unit.

4 Transformation of neuronal signal into exertion of force

After the synaptic connections w_{ji} and z_{kj} were adjusted (see Sect. 3) and fixed, we analyzed the performance of the model as follows (Fig. 3).

First coefficients a_p and a_I characterizing the magnitude of force signals were set to zero, and additional inputs T_j (11) were adjusted to obtain an initial pattern of IN activity that produces an equilibrium end-point position located approximately at the center of the workspace. Using this point as the origin, the preferred direction vectors D_j ($j = 1, \dots, 4$) were defined as unit vectors pointed from the origin to one of four equilibrium positions designated by ‘*T*’ in Fig. 2.

Eight signals coding the postural forces P of equal magnitude ($a_p = 0.3$) and different directions ranging from 0° to 360° were given, and corresponding patterns of the IN activity were calculated in the absence of incremental force signal ($a_I = 0$). Directions of encoded forces are shown by arrows that form the leftmost column in Fig. 3A. Each of these patterns produces a new equilibrium end-point position, which differs from the initial position of the arm. Due to this difference, a restoring force appears. If the arm is free to move, then the restoring force would result in a new position of the arm. Otherwise, the restoring force can be referred to the

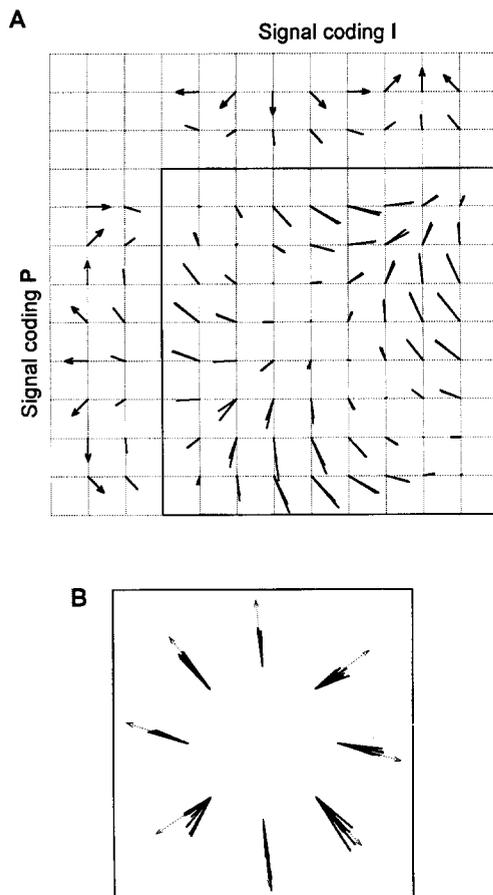


Fig. 3A, B. Transformation of neuronal signal into exertion of force. **A** Outside the square bordered by a thick line: Neuronal signals coding postural forces P and incremental forces I are represented by arrows lined up vertically and horizontally, respectively. The end-point forces exerted in response to these neuronal signals are shown as line segments that form the second column (postural force) and the second row (incremental force). The tail of each force vector is aligned with the grid vertex next right (postural force) or next down (incremental force) from the tail of the corresponding arrow. The calibration for the magnitude of force is given by the length of line segment connecting two adjacent vertices: this value is equal to 50 N. Inside the square: Two forces are drawn from each vertex: (i) the force that is generated in response to the sum of postural and incremental neuronal signals, and (ii) the formal vector sum of two forces generated by each signal separately. **B** The 'net forces' (see text) calculated for eight different directions of incremental signal combined with eight different directions of postural signal shown by arrows in **A**. Each cluster consists of eight net forces (line segments) calculated for the same incremental signal and eight different postural signals. Dashed arrows show directions of net force calculated for eight given incremental signals in the absence of postural signal. The scale for the magnitude of force is magnified by factor 2.5 with respect to the scale used in **A**.

force exerted by the arm against an immovable object measured in the experiment (Georgopoulos et al. 1992). We have calculated end-point forces generated by the eight neuronal signals using (1–11). These forces are shown as line segments that form the second column in Fig. 3A. The tail of each force vector is aligned with the grid vertex next right from the tail of the corresponding arrow. The incremental force signals and forces generated by these signals (in the absence of postural signal)

are shown in Fig. 3A by the upper row of arrows and by the second row of line segments, respectively. These data demonstrate that the model properly transforms neuronal signals of different directions into exertion of force. A difference between directions of force signals and generated forces seen in Fig. 3A can be reduced by an additional adjustment of the connectivity matrix w (data not shown).

To test whether the 'vectorial' summation rule (see Sect. 1) is valid in the framework of the model we calculated restoring end-point forces produced in response to simultaneous activation of both supraspinal populations (i.e. $a_p = 0.3$, $a_i = 0.3$). These data are depicted in Fig. 3A inside the square bordered by a thick line. Two forces are drawn from each of the vertices located inside the square. One force is the result of simultaneous activation of both supraspinal populations (the force signals encoded in the activities of populations are arrows located at the same horizontal/vertical lines crossing at the vertex). Another force is a formal vector sum of two forces generated by independent activation of each population (i.e., the vector sum of forces drawn near the corresponding signals outside the marked square). It is seen (Fig. 3A) that in all 64 cases the sum of neuronal signals is transformed into the force that is nearly equal to the vector sum of forces generated by each signal separately.

The performance of the model can be interpreted in the light of experimental data reported in Georgopoulos et al. (1992; see also Sect. 1 of the present paper for a discussion). To this end, we refer the net force N (Fig. 1, left template) acting on an immovable handle in the presence of bias force B (Georgopoulos et al. 1992) to the force calculated within the model as the difference $S - P$, where S is the force generated by the sum of postural and incremental signals, and P is the force generated in response to the postural signal (we suppose that the postural signal P compensates exactly the bias force B , i.e., $P = -B$). The performance of the model would be consistent with the experimental findings (Georgopoulos et al. 1992) if the direction of 'net force' $S - P$ is close to the direction of the incremental component I , regardless of the direction of postural force, P . In other words, the vectorial summation rule $S = P + I$ should be valid. We have calculated the net forces $S - P$ for eight different directions of incremental signal and eight different directions of postural signal shown in Fig. 3A by arrows. These 64 forces are represented in Fig. 3B by line segments grouped into eight clusters. Each cluster corresponds to the same direction of incremental signal combined with eight different directions of postural signal. Dashed arrows represent net force directions calculated in the absence of postural signal (i.e., directions of incremental force I). It is seen (Fig. 3B) that while the direction of postural force P changes from 0° to 360° within each cluster, the direction of net force $S - P$ remains almost invariant and close to the direction of incremental component I . In fact, this is another interpretation of the vectorial summation rule, which has been demonstrated by the results depicted in Fig. 3A.

5 Summation of restoring force fields

Although parameters of the model have been adjusted to reproduce the results of specific experiments performed on human and monkey (Sects. 2 and 3), it is useful to compare the restoring force fields generated by the model with experiments performed on the spinalized frog (Bizzi et al. 1991; Giszter et al. 1993; Mussa-Ivaldi et al. 1994). The results obtained by microstimulation of the frog's spinal cord suggested that the neural circuits in the spinal cord are organized in a set of control modules that 'store' a few limb postures in the form of convergent force fields acting on the limb's end-point. Moreover, it was demonstrated (Bizzi et al. 1991; Mussa-Ivaldi et al. 1994) that simultaneous stimulation of two distinct spinal sites results in a field of forces proportional to the vector sum of the fields induced by the independent stimulation of each site. These findings have led to a new concept of motor control (Bizzi et al. 1991; Mussa-Ivaldi and Giszter 1992) based on the vector combination of a few convergent fields to produce a vast repertoire of motor behaviour. Below we demonstrate that the proposed model generates restoring force fields whose properties are in close similarity with those reported in (Bizzi et al. 1991; Giszter et al. 1993; Mussa-Ivaldi et al. 1994).

5.1 Justification of the model

To establish a correspondence between the model and experiment (Bizzi et al. 1991; Giszter et al. 1993; Mussa-Ivaldi et al. 1994) we used the following approach. The IN layer of the model (Fig. 1) was disconnected from the supraspinal layers, i.e. the model was 'spinalized'. First, we calculated the force field generated under the condition that all units at the IN layer are inactive (Fig. 4A). Since the MN units are also inactive, the muscle rest lengths in this case are equal to their maximal possible values (7). We refer this field to the 'resting' force field measured in the absence of stimulation (Giszter et al. 1993). The stimulation of a particular spinal site in the experiment was referred to the activation of a particular unit at the IN layer. The activation of distinct units resulted in different force fields converging to different equilibrium points. Two examples are shown in Fig. 4B and C. We referred these fields to the 'total' force fields measured during stimulations of different sites (Bizzi et al. 1991; Giszter et al. 1993). Next, following the experimental procedure described in (Giszter et al. 1993), we calculated the 'active' force fields by subtracting the resting field from the total field. The active fields obtained from the total fields B and C (Fig. 4) are shown in Fig. 4D and E, respectively. Although the scale of forces depicted in Fig. 4 is larger than the scale of forces exerted by the frog's limb (the model has been adjusted to the human's arm), the structure of force fields is similar to that measured experimentally (Bizzi et al. 1991; Giszter et al. 1993; Mussa-Ivaldi et al. 1994).

5.2 Co-activation of distinct sites

Experimental results (Bizzi et al. 1991; Mussa-Ivaldi et al. 1994) suggested that the 'vectorial summation rule' is

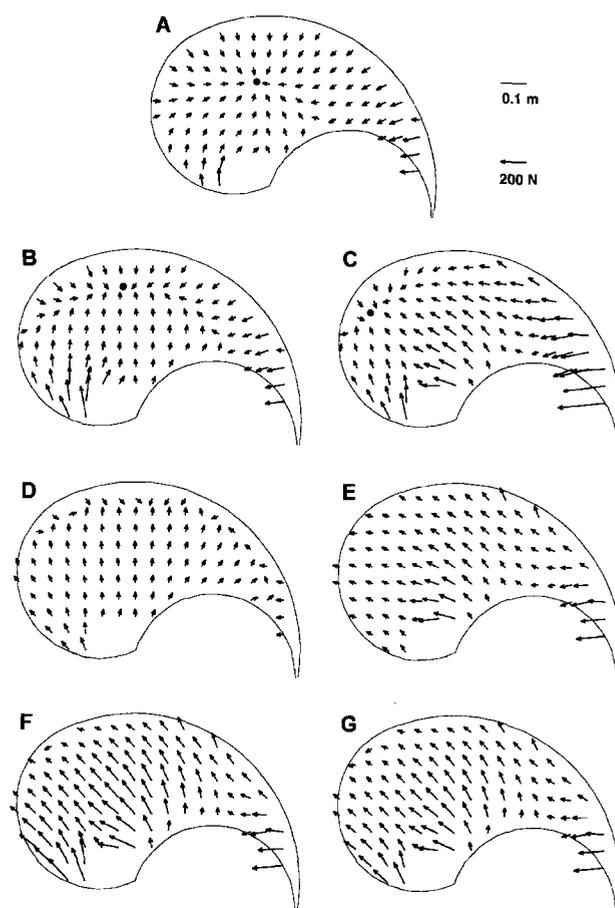


Fig. 4A–G. Vector summation of force fields. Each force vector is represented by an *arrow* drawn with its tail at the tip of the end-point displacement vector. Note that the range of displacements encompasses the whole workspace, which is bounded by a *thin curve*. For graphical reasons the vectors whose magnitude exceed a predefined threshold (50 N) are plotted as forces of 50 N. **A** The 'resting' field generated under the condition that all units at the IN layer are inactive. **B, C** Examples of the 'total' fields generated by separate activation of two different units at the IN layer (the activity of each unit is equal to 0.85). Force fields depicted in **A, B** and **C** converge to equilibrium points located in the workspace (the equilibrium points are indicated by filled circles). **D, E** The 'active' fields obtained by subtracting the resting field **A** from the total fields **B** and **C**, respectively. **F** The active field generated by simultaneous activation of the two units that generate the fields **D** and **E** when activated separately. **G** The field calculated by the vector summation of the active fields **D** and **E**.

valid for the active fields. We compared the active field generated by the simultaneous activation of two IN units (Fig. 3F) with the vector sum of the fields generated by the independent activation of the same units (Fig. 3G). As suggested by Mussa-Ivaldi et al. (1994), to quantify the similarity between the sum and co-activation fields, we calculated the 'cosine' of the angle between these two sampled fields, and used for this measure the value of 0.90 as a threshold for similarity (the maximum value of the similarity is equal to 1). There are four units at the IN layer and, therefore, there are six different pairs of them. We have found that, for all six pairs, the similarity between the sum and co-activation fields exceeded the

threshold value ranging from 0.97 to 0.99 (the similarity between fields depicted in Fig. 4F and G is equal to 0.99).

We extended this analysis comparing the vector sum of fields generated by two random patterns of the IN activity (the activity of each unit is a continuous variable ranging from 0 to 1) with the co-activation field generated by superposition of these patterns. We calculated the sum and co-activation fields for 10 000 pairs of random patterns. The similarity between these fields was dependent on the sampled patterns, ranging from 0.71 to 0.99. Therefore, the co-activation field generated by superposition of two arbitrary patterns may differ dramatically from the sum of the fields generated by each of these patterns independently (e.g., the value of 0.71 corresponds to the average angle between sample vectors equal to 45 degrees). This result was expected because of the nonlinearity of the model. However, the number of cases in which the similarity was below the threshold value 0.90 comprised less than 15% of the total number of trials. Averaged over all trials the similarity between the sum and co-activation fields was 0.96 ± 0.04 , which is consistent with the average similarity across the entire set of experimental data 0.94 ± 0.05 (Mussa-Ivaldi et al. 1994).

5.3 Summary

The force fields exemplified in Fig. 4 have been calculated within the same model that has been adjusted and tested in Sects. 2–4 (Figs. 2, 3). None of the parameters has been changed, neither introduced additionally. Nevertheless, the data shown in Fig. 4 indicate that basic features of the force fields obtained by the spinal stimulation of the frog's limb (Bizzi et al. 1991; Giszter et al. 1993; Mussa-Ivaldi et al. 1994), such as a convergence of total fields to equilibrium points, typical nonlinear dependence of the force direction and magnitude upon the position in the workspace, and the 'vectorial summation rule', are well reproduced within the model.

6 Concluding remarks

The complexity of real biological systems controlling motor behaviors is such that no practical models attempt to describe them in full detail. In this paper we focused on a particular aspect of the problem: the possible mechanisms underlying the *integration* of neuronal commands from different systems to produce a required force. We aimed to develop a model of the 'minimal complexity' that would include those factors of primary importance for the problem, making as few assumptions as possible concerning the details omitted from the model.

Although the degree of isomorphism between the model and motor physiology should not be overstated [for example, the motoneuron activation function (4) does not contain any feedback component], we believe that the model described in this paper is a reasonable compromise between tractability and realism. Indeed, the structure of the model is very simple: it consists of a stan-

dard three-layered feedforward neural network connected to a standard two-joint-six-muscle model of the arm. On the other hand, the model incorporates the following biologically important features: force signals are encoded in the activity of directionally tuned neurons; only a few distinct sites represent the level of spinal cord interneurons; synaptic connections between neural layers result in converging and diverging patterns of influence (in particular, neuronal units from the layer representing spinal cord interneurons make synaptic connections with different pools of motoneurons and activate groups of muscles in a weighted fashion); neuronal units are modeled as elements with nonlinear activation function; the construction of the arm model provides a nonlinear dependence of elastic properties on configuration of the arm, and muscles are modeled by nonlinear springs.

To adjust the performance of the model we used two sets of experimental data. One set (the stiffness characteristics of human arm) was used to train the network, and another set (the correlation between motor cortical activity and the force exerted by monkey, and the structure of restoring force fields generated by stimulation of the frog's spinal cord) was used to test the model's performance. We have shown that the performance of the model with a fixed set of parameters is quantitatively consistent with experimental data obtained on three different organisms using different experimental paradigms, thus suggesting a universal principle of organization of motor control.

The model predicts a specific pattern of connections between supraspinal populations coding forces and spinal cord interneurons. We have shown that a supraspinal signal can be properly translated into a required force if strengths of cortical-spinal connections are weighted in accordance with directional preference of connected units (see also Georgopoulos 1988, 1994, where this type of interactions between cortical and spinal systems was hypothesized). In this paper, the cortical-spinal connectivity matrix is given by an explicit relation (10).

In fact, our investigation originated from experimental findings (Georgopoulos et al. 1992; Mussa-Ivaldi et al. 1994), which raised a question of how an essentially nonlinear system could provide a linear summation of forces coding by different neuronal signals. Obviously, a nonlinear system cannot carry out the linear summation *exactly*, merely by definition of what the nonlinearity is. Indeed, our model, which incorporates nonlinearities at all levels (1–11), never performed the exact vectorial summation of forces (see Figs. 3, 4). However, we have shown that adjustable parameters of the model (synaptic connections) can be chosen in a way that the force generated by the sum of signals is *nearly equal* to the vector sum of forces generated by each signal independently (Figs. 3, 4). From this we conclude that approximately linear summation of forces observed in the experiments (Georgopoulos et al. 1992; Mussa-Ivaldi et al. 1994) cannot be a stringent rule, but may reflect a prevalent tendency, which could be provided by patterns of cortical-spinal and spinal interneuron-motoneuronal connections.

Appendix

For a given arm configuration, specified by the joint angles φ_s and φ_e , the set of muscle lengths $(l_1 \ l_2 \ l_3 \ l_4 \ l_5 \ l_6)$ is defined by the geometry of muscle attachment:

$$\begin{pmatrix} l_1 \\ l_2 \\ l_3 \\ l_4 \\ l_5 \\ l_6 \end{pmatrix} = L \begin{pmatrix} 1 + b/L \cos \varphi_s \\ 1 - b/L \cos \varphi_s \\ 1 + b/L \cos \varphi_e \\ 1 - b/L \cos \varphi_e \\ 1 + b/L(\cos \varphi_s + \cos \varphi_e) \\ 1 - b/L(\cos \varphi_s + \cos \varphi_e) \end{pmatrix} \quad (\text{A1})$$

(the above expression was derived in the limit $b/L \ll 1$). The transformation between the vector of muscle forces $\mathbf{f} = (f_1 f_2 f_3 f_4 f_5 f_6)^T$ and the vector of net joint torques $\mathbf{t} = (t_s \ t_e)^T$ is given by the matrix of muscle moment arms $\mathbf{r} = [\partial l_k / \partial \varphi_a]$ ($k = 1, \dots, 6$; $a = s, e$):

$$\mathbf{t} = -\mathbf{r}^T \mathbf{f} \quad (\text{A2})$$

with

$$\mathbf{r} = \begin{pmatrix} -b \sin \varphi_s & b \sin \varphi_s & 0 & 0 \\ 0 & 0 & -b \sin \varphi_e & b \sin \varphi_e \\ -b \sin \varphi_s & b \sin \varphi_s \\ -b \sin \varphi_e & b \sin \varphi_e \end{pmatrix}^T \quad (\text{A3})$$

[note that (A2) is equivalent to (8)]. The equilibrium configuration of the arm is defined by the condition that both net joint torques are equal to zero: $t_s = 0$, $t_e = 0$. Consequently, from (A2) and (A3) one has

$$\begin{aligned} -f_1 + f_2 - f_5 + f_6 &= 0 \\ -f_3 + f_4 - f_5 + f_6 &= 0 \end{aligned} \quad (\text{A4})$$

The fact that there are only two linear equations (A4) for the six muscle forces reflects the redundancy of the system (the same equilibrium position of the arm can be realized by different sets of muscle forces). Below we use this redundancy to ensure the required relationship between the equilibrium position and the stiffness characteristics of the arm.

Let φ_s^{eq} and φ_e^{eq} be the equilibrium joint angles. Then the joint stiffness matrix is defined as

$$\begin{pmatrix} R_{ss} & R_{se} \\ R_{es} & R_{ee} \end{pmatrix} = - \begin{pmatrix} \frac{\partial t_s}{\partial \varphi_s} & \frac{\partial t_s}{\partial \varphi_e} \\ \frac{\partial t_e}{\partial \varphi_s} & \frac{\partial t_e}{\partial \varphi_e} \end{pmatrix}_{\substack{\varphi_s = \varphi_s^{\text{eq}} \\ \varphi_e = \varphi_e^{\text{eq}}}} \quad (\text{A5})$$

After differentiation using (A2), (A3) one has

$$\begin{pmatrix} R_{ss} & R_{se} \\ R_{es} & R_{ee} \end{pmatrix} = \begin{pmatrix} (f'_1 + f'_2 + f'_5 + f'_6)b^2 \sin^2 \varphi_s^{\text{eq}} \\ (f'_5 + f'_6)b^2 \sin \varphi_s^{\text{eq}} \sin \varphi_e^{\text{eq}} \\ (f'_5 + f'_6)b^2 \sin \varphi_s^{\text{eq}} \sin \varphi_e^{\text{eq}} \\ (f'_3 + f'_4 + f'_5 + f'_6)b^2 \sin^2 \varphi_e^{\text{eq}} \end{pmatrix} \quad (\text{A6})$$

where

$$f'_k = \left. \frac{df_k}{dl_k} \right|_{l_k = l_k(\varphi_s^{\text{eq}}, \varphi_e^{\text{eq}})} = \beta(f_k + \alpha), \quad k = 1, \dots, 6 \quad (\text{A7})$$

where parameters α and β are defined in (8).

The polar orientation of the stiffness ellipse is defined by the relation (Flash and Mussa-Ivaldi 1990):

$$R_{ss} = 2R_{se} = 2R_{es} \quad (\text{A8})$$

which gives one additional equation for the muscle forces. Two other relations can be obtained if the shape parameter, ρ , and the size, σ , of the stiffness ellipse are given. These parameters are expressed in terms of stiffness matrix elements as follows:

$$\begin{aligned} \rho &= (4R_{ee}/R_{ss} - 1) \cot^2(\varphi_e^{\text{eq}}/2) \\ \sigma &= \pi \frac{(R_{ee} - R_{ss}/4)R_{ss}}{L^4 \sin^2 \varphi_e^{\text{eq}}} \end{aligned} \quad (\text{A9})$$

Substitution of R_{ss} , R_{ee} , and R_{se} from (A6), (A7) into (A8), (A9) gives three equations:

$$\begin{aligned} f_1 + f_2 &= 2(\beta^{-1}A(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}}) - \alpha) \\ f_3 + f_4 &= 2(\beta^{-1}B(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}}) - \alpha) \\ f_5 + f_6 &= 2(\beta^{-1}C(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}}) - \alpha) \end{aligned} \quad (\text{A10})$$

where $A(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}})$, $B(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}})$, and $C(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}})$ are explicit functions of the stiffness ellipse shape ρ , size σ , and equilibrium joint angles φ_s^{eq} and φ_e^{eq} (these expressions are too cumbersome to be printed and can be provided on request). To close the system of equations we need one additional equation for which we use the relation

$$f_5 - f_6 = 0 \quad (\text{A11})$$

Then the solution of the system (A4), (A10), and (A11) is given by

$$\begin{aligned} f_1 &= f_2 = \beta^{-1}A(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}}) - \alpha \\ f_3 &= f_4 = \beta^{-1}B(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}}) - \alpha \\ f_5 &= f_6 = \beta^{-1}C(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}}) - \alpha \end{aligned} \quad (\text{A12})$$

Substituting f_k from (A12) and l_k from (A1) into (6), we finally obtain explicit equations for the set of muscle rest lengths $(l_1^0 \ l_2^0 \ l_3^0 \ l_4^0 \ l_5^0 \ l_6^0)$ ensuring for any given equilibrium joint angles φ_s^{eq} and φ_e^{eq} the polar orientation of the stiffness ellipse as well as its specified shape ρ , and size σ :

$$\begin{pmatrix} l_1^0 \\ l_2^0 \\ l_3^0 \\ l_4^0 \\ l_5^0 \\ l_6^0 \end{pmatrix} = L \begin{pmatrix} 1 + b/L \cos \varphi_s + \beta^{-1} \ln(\alpha \beta A^{-1}(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}})) \\ 1 - b/L \cos \varphi_s + \beta^{-1} \ln(\alpha \beta A^{-1}(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}})) \\ 1 + b/L \cos \varphi_e + \beta^{-1} \ln(\alpha \beta B^{-1}(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}})) \\ 1 - b/L \cos \varphi_e + \beta^{-1} \ln(\alpha \beta B^{-1}(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}})) \\ 1 + b/L(\cos \varphi_s + \cos \varphi_e) + \beta^{-1} \ln(\alpha \beta C^{-1}(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}})) \\ 1 - b/L(\cos \varphi_s + \cos \varphi_e) + \beta^{-1} \ln(\alpha \beta C^{-1}(\rho, \sigma, \varphi_s^{\text{eq}}, \varphi_e^{\text{eq}})) \end{pmatrix} \quad (\text{A13})$$

Acknowledgement. This research was supported by contract N00014-94-1-0033 from the Office of Naval Research.

References

- Bizzi E, Mussa-Ivaldi FA, Giszter SF (1991) Computations underlying the execution of movement: a biological perspective. *Science* 253:287–291
- Caminiti R, Johnson PB, Urbano A (1990) Making arm movements within different parts of space: dynamic aspects in the primate motor cortex. *J Neurosci* 10:2039–2058
- Evarts EV, Fromm C, Kroller J, Jennings VA (1983) Motor cortex control of finely graded forces. *J Neurophysiol* 49:1199–1215
- Feldman AG (1966) Functional tuning of the nervous system with control of movement or maintenance of a steady posture. II. Controllable parameters of the muscles. *Biophysics* 11:565–578
- Flash T (1987) The control of hand equilibrium trajectories in multi-joint arm movements. *Biol Cybern* 57:257–274
- Flash T, Mussa-Ivaldi F (1990) Human arm stiffness characteristics during the maintenance of posture. *Exp Brain Res* 82:315–326
- Georgopoulos AP (1988) Neural integration of movement: role of motor cortex in reaching. *FASEB J* 2:2849–2857
- Georgopoulos AP (1994) New concepts in generation of movement. *Neuron* 13:257–268
- Georgopoulos AP, Kalaska JF, Caminiti R, Massey JT (1982) On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J Neurosci* 2:1527–1537
- Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neuronal population coding of movement direction. *Science* 233:1416–1419
- Georgopoulos AP, Ashe J, Smyrnis N, Taira M (1992) The motor cortex and the coding of force. *Science* 256:1692–1695
- Georgopoulos AP, Taira M, Lukashin AV (1993) Cognitive neurophysiology of the motor cortex. *Science* 260:47–52
- Giszter SF, Mussa-Ivaldi FA, Bizzi E (1993) Convergent force fields organized in the frog's spinal cord. *J Neurosci* 13:467–491
- Hocherman S, Wise SP (1991) Effects of hand movement path on motor cortical activity in awake, behaving rhesus monkeys. *Exp Brain Res* 83:285–302
- Hogan N (1985) The mechanics of multi-joint posture and movement control. *Biol Cybern* 52:315–331
- Kalaska JF, Crammond DJ (1992) Cerebral cortical mechanisms of reaching movements. *Science* 259:1517–1523
- Kalaska JF, Cohen DAD, Hyde ML, Prud'homme M (1989) A comparison of movement direction-related versus load direction-related activity in primate motor cortex, using a two-dimensional reaching task. *J Neurosci* 9:2080–2102
- Katayama M, Kawato M (1993) Virtual trajectory and stiffness ellipse during multijoint arm movement predicted by neural inverse models. *Biol Cybern* 69:353–362
- Mussa-Ivaldi FA, Giszter SF (1992) Vector field approximation: a computational paradigm for motor control and learning. *Biol Cybern* 67:491–500
- Mussa-Ivaldi FA, Hogan N, Bizzi E (1985) Neural, mechanical, and geometric factors subserving arm posture in humans. *J Neurosci* 5:2732–2743
- Mussa-Ivaldi FA, Giszter SF, Bizzi E (1994) Linear combinations of primitives in vertebrate motor control. *Proc Natl Acad Sci USA* 91:7534–7538
- Redish AD, Touretzky DS (1994) The reaching task: evidence for vector arithmetic in the motor system? *Biol Cybern* 71:307–317
- Schwartz AB (1994) Direct cortical representation of drawing. *Science* 265:540–542
- Shadmehr R, Arbib MA (1992) A mathematical analysis of the force-stiffness characteristics of muscles in control of a single joint system. *Biol Cybern* 66:463–477
- Shadmehr R, Mussa-Ivaldi FA, Bizzi E (1993) Postural force fields of the human arm and their role in generating multijoint movements. *J Neurosci* 13:45–62
- Tsuji T, Morasso PG, Goto K, Ito K (1995) Human hand impedance characteristics during maintained posture. *Biol Cybern* 72:475–485
- Wise SP (1993) Monkey motor cortex: movements, muscles, motoneurons and metrics. *Trends Neurosci* 16:46–49