Optimization Principles In Motor Control

Tamar Flash[1], Neville Hogan[2] and Magnus J. E. Richardson[1]

1. Department of Applied Mathematics and Computer Science,
   The Weizmann Institute of Science, Rehovot 76100, Israel.
   tamar@wisdom.weizmann.ac.il, Magnus.Richardson@lps.ens.fr

2. Departments of Mechanical Engineering and Brain and Cognitive Sciences,
   Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

Short title: Optimization Principles In Motor Control

Contact details of corresponding author.

Magnus Richardson
Laboratoire de Physique Statistique, Ecole Normale Superieure
24 rue Lhomond 75231, Paris Cedex 05, France
Magnus.Richardson@lps.ens.fr
Fax: +33 1 44 32 34 33
Tel: +33 1 44 32 38 00

Road map: Biological Motor Control
Related Reading: Equilibrium Point Hypothesis; Eye-Hand Coordination in Reaching Movements; Geometrical Principles in Motor Control; Human Movement: A System-Level Approach; Limb Geometry: Neural Control

INTRODUCTION

Optimization theory has become an important research tool in our attempts to discover organizing principles that guide the generation of goal-directed motor behavior. It provides a convenient way to formulate a coarse-grained model of the underlying neural computation, without requiring specific details of the way those computations are carried out. Generally speaking, this application of optimization theory consists of defining an objective function that quantifies what is to be regarded as optimum (i.e., best) performance and then applying the tools of variational calculus to identify the specific behavior that achieves that optimum. This forces us to make explicit, quantitative hypotheses about the goals of motor actions and allows us to articulate how those goals relate to observable behavior. Not all motor behaviors are necessarily optimal but attempts to identify optimization principles can be useful for developing a taxonomy of motor behavior and gaining insight into the neural processes that produce motor behavior.

Ill-Posed Problems in Motor Behavior

Many optimization-based models in the literature have been developed to address the "excess degrees-of-freedom" problem. How does the motor system select the behavior it uses from the infinite number of possibilities open to it? In mathematical parlance, this is an "ill-posed" problem in the sense that many solutions are possible. For example, most limb segments are moved by a larger number of muscles than appear to be necessary. To reach for a cup of coffee the hand may move along many different paths. The same figural form (e.g., the letter Z or an ellipse) may be drawn using a wide variety of time profiles for the pen's position. The central question is how the nervous system chooses values for the large number of parameters that can be controlled. One appealing possibility is that the nervous system has evolved to select solutions that maximize the organism's fitness, i.e., that are optimal in some sense. More specifically, the hypothesis is that in performing a motor task the brain produces coordinated actions that minimize some measure of performance (such as effort, smoothness, etc.). In this article we review several studies in which the validity of such ideas was examined in the context of planar upper limb movements. Similar ideas have been explored in the context of other effector systems and motor actions such as whole body posture, gait and various sporting activities, but they will not be considered here. The interested reader is referred to Winters and Crago (2000) for further information.

Arm Trajectory formation

Our first topic is the kinematic aspects of movement. Kinematics refers to the time course of limb position, velocity, etc., while dynamics refers to variables such as forces and torques. In principle, even a single-degree-of-freedom movement (e.g., elbow rotation) can be performed in many different ways. Thus, while the hand path is constrained to follow a circular arc, its speed along the path may follow many different time profiles. One way to gain insight into the processes responsible for the selection of specific limb trajectories is to experimentally observe human movements. Patterns or invariances in the observed behavior suggest hypotheses about the way these movements are organized. Optimization theory provides a mathematical tool for concisely formulating and testing these hypotheses. The key step is the identification of an objective function which defines a measure of performance by assigning a cost to each member of the class of possible behaviors under study (e.g., arm trajectories). One member of this class (e.g., one trajectory) will then be selected to maximize or minimize that function. How the objective function is defined determines what aspects of the motor behavior are considered important.

Kinematic vs. Dynamic Objective Functions

In this article we will consider two different types of objective functions that have been proposed (out of the multitude of possibilities) as they reflect two major competing theories of how motor computations are organized. The first type of objective function is based solely on kinematic variables (e.g., limb position and its time derivatives). If a kinematic objective function can be found that leads to optimal trajectories that accurately reproduce the patterns of observed behavior, it implies that the brain ignores non-kinematic factors in selecting and producing that behavior. This would be consistent with a theory that, to produce movement, neural computations are organized hierarchically and executed by proceeding from the abstract (i.e. move to that light over there) to the particular (i.e. activate that set of motoneurons in this manner). The most compelling evidence supporting this idea is the observation that similar kinematic patterns are observed even when widely different musculo-skeletal systems are involved in producing motor behavior. One's signature on paper is equally as recognizable and distinctive as one's signature on a blackboard, despite the enormous differences in the mechanics and physiology of the body parts used to produce it. Nevertheless, a troubling aspect of this theory is that it seems to imply that, at least at the higher levels of the postulated hierarchy, the brain does not take *any* dynamic considerations into account such as the energy required, the loads on the limb segments or the force and fatigue limitations of its peripheral neuromuscular system.

To circumvent this problem within the framework of optimization theory, a second type of objective function may be formulated based on dynamic variables (e.g., joint torques, muscle forces, etc., and their time derivatives). If a dynamic objective function can be found that leads to optimal trajectories that accurately reproduce the patterns of observed behavior, it implies that the brain considers dynamic factors in selecting and producing that behavior. It is also consistent with a theory that neural computations to produce movement are executed in parallel, taking all relevant factors (e.g., dynamics as well as kinematics) into account simultaneously.

Single-Joint Movements.

As has been frequently observed, single-joint movements are characterized by single-peaked, bell-shaped speed profiles. This finding and the tendency of natural movements to be characteristically smooth and graceful, led Hogan (1984) to suggest that motor coordination can be mathematically modeled by postulating that voluntary movements are made, at least in the absence of any other overriding concerns, to be as smooth as possible. For mathematical convenience (there are many other plausible measures of smoothness) maximizing smoothness was expressed as minimizing mean-squared average jerk, the third time-derivative of position. In the single-joint case:

$$C = \int_{t_0}^{t_f} (\frac{d^3\theta}{dt^3})^2 dt \tag{1}$$

where $\theta(t)$ is the joint angle, and t_0 and t_f are the initial and final movement times, respectively. Using variational calculus, the unique time history of joint positions that minimizes this performance measure may be derived analytically. It is described by the following quintic polynomial in time:

$$\theta(t) \quad = \quad c_0 + c_1 t + c_2 t^2 + c_3 t^3 + c_4 t^4 + c_5 t^5 \tag{2}$$

where $c_i, i = 0, ..., 5$ are unspecified coefficients whose values are determined by the conditions at the beginning and end of the movement (boundary conditions). Originally, Hogan (1984) analyzed movements that start and end at rest and therefore assumed zero initial and final velocities and accelerations. Consequently, the predicted trajectories were characterized by symmetric bell-shaped speed profiles. For movements of different amplitudes and durations, the ratio of peak speed to average speed was invariant at 1.88. For a repetitive sequence of movements, speed profiles were again symmetric and this ratio was again invariant, but with a value of 1.57. These predictions appear to be in good agreement with observation. A constant ratio of peak speed to average speed has been reported by several researchers, with values between 1.60 and 1.90, depending on the conditions of measurement. However, a distinctive feature of these minimum-jerk movements is their symmetric speed profile and that is not always observed experimentally. For example, when enhanced accuracy of target acquisition is demanded, an asymmetric speed profile is typically observed, with the peak speed occurring earlier in the movement. This indicates that the simple minimum-jerk theory may need to be modified. One possible way to account for this asymmetry is by adding (to the objective function) a term to minimize hand-to-target error integrated across the movement. An alternative is to modify the boundary conditions.

Another alternative is to use a dynamic objective function. This requires formulation of a model of neuromuscular and skeletal mechanics to relate dynamic variables (e.g., forces) to kinematics. Hasan (1986) proposed a minimum-effort theory of single-joint movement generation based on a model that described neuromuscular behavior as equivalent to a "spring-like" element driving the limb towards a neurally-defined "equilibrium position", determined by simultaneous activation of agonist and antagonist muscle groups. Minimization of effort was expressed as follows:

$$C \quad = \quad \int_{t_0}^{t_f} (\sigma(t)(\frac{d\beta}{dt})^2)dt \tag{3}$$

where σ is the joint stiffness (describing the rate of change of the restoring force generated by the "spring-like" element with its displacement from equilibrium) and $\frac{d\beta}{dt}$ is the time derivative of the equilibrium position. Thus for single-joint movements, optimization theories using both kinematic and dynamic objective functions have been applied with success. A more telling test of these theories is found in multi-joint movements.

MULTI-JOINT MOVEMENTS AND THE QUESTION OF COORDINATES

The kinematics of multi-joint arm movements may be represented in a number of different ways, e.g., as a series of hand positions, joint angles or muscle lengths, etc. Each of these may be considered as an alternative "coordinate frame" for describing the movement. The neural computations underlying multi-joint arm movements may make use of any one (or even several) of these representations. Experimental observations of unconstrained human reaching movements are characterized by approximately straight hand paths and symmetric bell-shaped speed profiles which remain nearly invariant despite changes in movement direction, speed, and starting position. Because these features are only evident in the motions of the hand and not in the movements of individual limb segments, it was proposed that the neural computations underlying movement production take place in terms of hand motion through extracorporeal space and not in terms of joint rotations.

Flash and Hogan (1985) showed that the maximum-smoothness theory reproduced all of these features, provided the objective function was expressed in terms of the Cartesian coordinates of the hand as follows:

$$C = \int_{t_0}^{t_f} \left(\left(\frac{d^3x}{dt^3}\right)^2 + \left(\frac{d^3y}{dt^3}\right)^2 \right) dt \tag{4}$$

where $x(t), y(t)$ describe the hand position coordinates and t_f is the movement duration.

Minimizing this objective function yielded analytic expressions for the hand trajectories. For unrestrained point-to-point movements starting and ending at rest, the model predictions agreed closely with experimental data and successfully accounted for the invariance of hand trajectories under translation, rotation, amplitude and speed scaling.

In more complex curved movements, patterns were again evident in hand kinematics but not in joint kinematics. When subjects were instructed to generate curved or obstacle-avoidance movements, although the hand paths appeared smooth, movement curvature was not uniform; the trajectories displayed two or more curvature maxima. The hand speed profiles also had two or more maxima and the minima between adjacent peaks temporally corresponded to the peaks in curvature

To describe curved and obstacle-avoidance movements, the maximum smoothness model was extended by assuming that a small number of points along the path through which the hand should pass are specified (Flash and Hogan, 1985). The time of passage through those "via" points and the hand velocity at that time were not specified a priori but were predicted by the model. For the simplest case of one via point between the initial and final positions, the theory yielded explicit mathematical expressions for the hand motion (Flash and Hogan, 1985) that reproduced all the features of the experimental observations: distinct maxima in the speed profile with a minimum between them which coincided temporally with a curvature maximum; trajectory shape invariant under translation, rotation, amplitude and time scaling; and nearly equal durations of movement from the initial position to the via point, and from the via point to the final position. The latter observation was referred to as *the isochrony principle* (Viviani and Terzuolo 1982)—the phenomenon that movement durations of large and small segments of a trajectory are roughly equal.

Minimum Torque-Change Models

In contrast to the maximum-smoothness model, Uno, Kawato, and Suzuki (1989) postulated that movement selection optimizes the rate of change of actuator efforts, e.g., joint torques. Although minimizing jerk and minimizing the rate of change of joint torques appear conceptually similar (in a single-joint system with predominantly inertial dynamics they are proportional to one another) there are important differences. First, the objective function is based on dynamic variables: the rate of change of torque. Therefore the predicted motion depends sensitively on the modeled dynamic behavior of the musculoskeletal system. Secondly, the objective function was formulated in terms of joint torques rather than functions of the hand's coordinates as is the case for minimum jerk. This implies that motor computations are based on a joint-space representation of behavior. Although, as outlined above, kinematic patterns are most evident in hand motions in extra-corporeal space, approaches based on either joint or muscle spaces have the advantage that they can generate solutions to important aspects of the ill-posed motor-control problems, such as kinematic redundancy (the apparent excess degrees of freedom) or actuator redundancy (the apparent excess of muscles). The maximum-smoothness model expressed in hand coordinates does not address these issues.

Initially, Uno et al (1989) reported that the performance of the minimum torque-change model surpassed that of the maximum smoothness model. It appeared to account for the small but systematic curvature seen in point-to-point movements and also for the larger curvature seen in movements that pass from the side to the front of the body. However, an independent study (Flash 1990) and a later reappraisal co-authored by some of the original proponents of the minimum torque-change model (Nakano et al 1999) invalidated these conclusions: it was shown that a combination of too-large an inertia and too-small a viscosity contingently led to predictions compatible with experimental results.

However, in a later paper (Nakano, et al 1999) a variant of the minimum torque-change model, the minimum *commanded* torque-change model, was introduced. In this model the commanded torque includes non-zero viscous terms that arise from biochemical and mechanical reaction processes within the muscles, and in this way both the muscles and the link dynamics are considered as controlled objects. Using more realistic, measured physical parameters this second model was again able to reproduce the experimentally verified effects of curvature.

Motor Adaptation Studies

The most critical comparison of these two models arises from their fundamental differences. According to kinematically-based optimization models, neural computations specify intended motions independently of movement dynamics or external load conditions. In contrast, dynamically- based optimization models imply that external loads profoundly influence intended motions. For example, according to the minimum torque-change models, movements in the presence of elastic loads should be more curved than unloaded movements, whereas the maximum-smoothness model predicts no effect.

Investigating motor adaptation to elastic loads, Uno et al. (1989) concluded that the behavior in the presence of the load was different from the unloaded case. Completely different results, however, were obtained in another study in which static elastic loads were unexpectedly introduced during human reaching toward visual targets (Flash and Gurevich, 1997). In the first few trials following load application movements were found to be misdirected and to miss the target but following a small number of practice trials (5-7), the loaded movements tended to follow straight hand paths with symmetric velocity profiles. In another study (Shadmehr and Mussa-Ivaldi, 1994), velocity-dependent force fields were used to perturb the motion and the perturbed trajectories performed in the presence of the new force fields were again found to converge toward the ones seen in the unloaded case. In a third related study, Wolpert et al. (1994) used altered visual feedback conditions that caused an increase in the perceived curvature of aiming movements. This led to significant corrective adaptation of the movements actually produced; the hand movements became curved, thereby reducing the visually perceived curvature. These results support the notion that arm trajectories follow a kinematic plan formulated in extrinsic visual space independent of movement dynamics or external force conditions. They are incompatible with the assumptions of dynamically-based optimization models formulated in terms of intrinsic coordinates.

Furthermore, it should be noted that small deviations from straight line movements do not necessarily imply planning in joint coordinates. Such phenomena are compatible with planning in kinematic space, but with perturbations due to the dynamics of the arm and neuronally controlled muscles at the implementation stage (Flash, 1990). Conclusions with respect to the sensorimotor mapping that associate desired trajectories to motor commands were drawn based on motor adaptation studies. Shadmehr and Mussa-Ivaldi (1994) have analyzed the aftereffects observed when following training in one region of the work space, subjects were asked to perform reaching movements at a nearby space. The patterns of aftereffects suggested that generalization from learning was in terms of intrinsic joint-based coordinates.

RELATION TO PHYSIOLOGY

The kinematic and dynamic objective functions discussed above are based on measures of smoothness in different coordinate frames. Both of these models have in common that they are *phenomenological* approaches. The controller (nervous system) and plant (arm) are treated as a *black box* with the input the experimental task and the output the goal-fulfilling movement. The success (or otherwise) of phenomenological theories in fitting experiment gives insight into which variables the central nervous system might consider important in the movement planning process. Results have been presented above that support the idea that the high-level planning processes in the CNS might be in the coordinates of the hand's position.

The fact that movements are smooth, whether in hand or joint coordinates has been interpreted as being compatible with increasing the predictability of the trajectory or reducing the amount of information needed to internally represent motion plans (Hogan, 1984; Flash and Hogan, 1985). Smoothness maximization and the superposition of elemental movements to generate more complicated arm trajectories are also closely related to regularization-based approaches to learning from examples. Those approaches view learning as equivalent to identifying a function from sparse and noisy data. The tradeoff between accurate data reproduction and "well-behaved-ness" of the mapping is achieved by maximizing the smoothness of the function.

Work has also been done on how the CNS might implement an optimization procedure such as minimum jerk. For example, it has been shown that a minimum-jerk movement planner can be directly implemented by a Radial Basis Function *RBF* network. Another implementation scheme was described by Hoff and Arbib (1992) by showing how the minimum jerk principle could be converted into a real-time controller, in which delays and noise effects could explain a number of experimental observations beyond the fitting of simple point-to-point trajectories. However, looking for neural circuits that can reproduce explicitly the calculations inherent in the phenomenological theories of minimum-jerk and minimum-torque might be a too-literal interpretation of the success of such theories in reproducing experiment. The kinematic and psychophysical observations reported to date do not sufficiently constrain the possible movement-generating algorithms to distinguish the finer details of neural implementation. Nevertheless, these phenomenological theories serve as background, coarse-grained descriptors to which deeper, more biologically-detailed theories must conform.

Recently, some effort has been made in grounding the optimization approach to motor control in a neuro-biological context. It was noted that biological systems are corrupted by noise, the variance of which increases with the size of the signal (Harris and Wolpert 1998). Hence, any pre-planned movement is likely to be off-target when the motor program is run through the noisy neuro-muscular system. As each goal-directed movement has some characteristic level of error this suggests a natural optimization criterion: the CNS chooses movements that minimize the final error in the achievement of the motor task.

Harris and Wolpert (1998) analyzed the predictions of this hypothesis in the context of saccadic eye movements. The error in the final eye position was functionally minimized with respect to the control signal (using a linear model of the plant). It was found that small final error was achieved by low-bandwidth neuronal signals, corresponding to the smooth velocity profiles seen experimentally. This approach was also extended to arm movements in the particular case of two-joint motions in the plane (Harris and Wolpert, 1998). A large range of experimental results were successfully reproduced including the small curvature seen in point-to-point movements. Furthermore, it was claimed that the predicted trajectory of the hand was, to a large degree, independent of the specifics of the model of the plant: the controlling neuronal signal adapts to produce similar output.

The role of noise in the coordination of movement has been further examined in the context of *optimal feedback control* (Todorov and Jordan, 2001). It was noted that the variability and redundancy inherent in, for example, the control of the human arm, are often treated as problems to be overcome in the planning process. In their work on optimal feedback control, it is shown that increased accuracy in the goal-specific parameters of movement can be obtained by allowing the variance to increase in the redundant variables. In fact, their model does not enforce a desired trajectory but corrects only those deviations that interfere with the task, a principle of *minimum intervention*. Despite this minimal formulation, experimentally observed features such as simplifying rules, control parameters and synergies emerge as epiphenomena of the control process. The theory is supported by a number of exemplary experimental results, and provides a satisfying interpretation of the role of variability and the so-called degrees-of-freedom problem.

MOTION PLANNING FOR THREE-DIMENSIONAL MOVEMENTS

For completeness, we briefly mention recent work that has been done on motion in three dimensions, i.e. not confined to a plane. Compared with the success of optimization techniques in two dimensions, the use of cost-function analysis is still in the investigative phase for this more general class of motion. It is known that point-to-point motions in three dimensions are considerably more curved than in the plane. Nevertheless, there has been some success in predicting this more complex behavior using the techniques of the optimization approach. Hypothesized cost functions have included minimum kinetic energy (Soechting, et al 1995), in which it was also shown that a simple Donder's Law rule which expresses a kinematic constraint on eye orientation does not apply to arm motions. Other models that incorporate a description of muscle dynamics and hypothesize the minimization of a metabolic energy cost or consider the effect of final posture of the arm have also been developed, representing attempts to deal with the acute degrees-of-freedom problem found in three-dimensional movements.

DISCUSSION

One of the exciting challenges of brain theory is the need to deal with reality at the level of whole, functioning systems. Traditionally, scientific endeavor has advanced our state of knowledge by delving into finer and finer details of isolated pieces of reality – the essence of the reductionist approach. However, due to the limited amount known of these fine details, and the difficulties involved in studying complex systems of many neurons, it is evident that this "bottom-up" approach is currently severely limited in its ability to describe system-wide behavior: large-scale, strongly-interacting systems exhibit characteristics that emerge primarily from interactions among their parts. To understand them, a "top-down" approach is far more effective, beginning at a coarse-grained macroscopic level, proceeding to finer levels of detail as their structure is discerned. Optimization theory provides a powerful set of mathematical tools that lend themselves well to a top-down approach to studying the brain. As we have reviewed above, optimization theory facilitates a rigorous approach, based on macroscopic observations of psychophysical behavior, to some fundamental and far-reaching questions about the structure of neural computations.

References

[1] *Flash, T. and Hogan, N. 1985. The coordination of arm movements: An experimentally confirmed mathematical model, J. Neurosci., 5:1688-1703.

[2] Flash, T. 1990. The organization of human arm trajectory control. In Multiple Muscle Systems: Biomechanics and Movement Organization. Eds. J. Winters and S. Woo, Springer-Verlag, pp. 282-301

[3] Flash T, Gurevich I. 1997. Arm trajectory generation and stiffness control during motor adaptation to external loads. In Self-Organization, Computational Maps and Motor Control. Eds. P.G. Morasso and V. Sanguinetti Amsterdam: Elsevier; pp. 423-482.

[4] Hasan, Z. 1986. Optimized movement trajectories and joint stiffness in unperturbed inertially loaded movements, Biol. Cybern., 53:373-382.

[5] *Harris C. M. Wolpert D. M. (1998) Signal-dependent noise determines motor planning, Nature, 394:780-784.

[6] Hoff, B., and Arbib, M.A., 1992. A model of the effects of speed, accuracy, and perturbation on visually guided reaching. In Control of Arm Movement in Space: Neurophysiological and Computational Approaches Eds. Caminiti, P. B. Johnson, and Y. Burnod, Experimental Brain Research Series 22, pp. 285-306.

[7] *Hogan, N. 1984. An organizing principle for a class of voluntary movements, J. Neurosci., 4:2745-2754.

[8] Nakano, E., Imamizu, H., Osu, R., Uno, Y., Gomi, H., Yoshioka, T. and Kawato, M. 1999. Quantitative Examinations of Internal Representations for Arm Trajectory Planning: Minimum Commanded Torque Change Model, J. Neurophysiol., 81:2140-2155.

[9] *Shadmehr, R. and Mussa-Ivaldi, F. A. 1994. Adaptive representation of dynamics during learning of a motor task, J. Neurosci., 14:3208-3224.

[10] *Soechting J. F., Buneo C. A., Hermann U. and Flanders M. 1995. Moving Effortlessly in Three Dimensions: Does Donder's Law Apply to Arm Movements?, J. Neurosci., 15:6271-6280.

[11] *Todorov, E. and Jordan, M. I. 2001. Optimal feedback control as a theory of motor coordination. http://www-rcf.usc.edu/ etodorov/

[12] *Uno, Y., Kawato, M. and Suzuki, R. 1989. Formation and control of optimal trajectory in human multijoint arm movement: Minimum torque-change model, Biol Cybern., 61:89-101.

[13] Viviani, P. and Terzuolo, C. 1982. Trajectory determines movement dynamics, Neuroscience, 7:431-437.

[14] *Winters, J. M. and Crago, 2000. Biomechanics and neural control of posture and movement, New York: Springer-Verlag.

[15] *Wolpert, D. M., Ghahramani, Z. and Jordan, M. I. 1995. Are arm trajectories planned in kinematic or dynamic coordinates? An adaptation study, Exp. Brain Res. 103:460-470.