Coordinate Transformations in the Genesis of Directed Action

C. R. Gallistel

Department of Psychology

UCLA

Los Angeles, CA 90095-1563

A paradox of psychology is that the analytic difficulty of a problem may be inversely related to the phenomenological difficulty. Things that seem trivially easy to do are often the most difficult to understand analytically; whereas things that seem difficult--things that require "real brains"--have been rather successfully modeled with modest effort. We can program computers to do many things that we find conceptually challenging--inverting matrices, solving partial differential equations--while we are a long way from being able to program machines to do what we find so easy that it requires no thought--for example, reaching out to pick up a pencil.

The deep difficulty in understanding basic aspects of perception is widely appreciated. It has proved extraordinarily difficult to program a computer to segment a visual image into components that correspond to objects in the three-dimensional space from which the image is projected. Thus, the first deep problem that prevents the manufacture of robots that can pick up pencils is getting them to pick out the relevant portion(s) from the image of a scene. Less widely appreciated are the deep difficulties that confront us when we try to understand how the brain generates simple directed actions. Given that it has somehow determined where it wants the end of a limb to go, how can it control the pattern of motor neuron firing in such a way as to make something as mechanically intractable as a vertebrate forelimb move to where it is supposed to go?

I will attempt to bring conceptual order and coherence to some interesting recent findings by arguing that: 1) The problem of controlling the trajectory of a limb is a problem in computing a series of coordinate transformations. 2) There is evidence that anatomically localized stages of the neural circuitry perform computationally distinct coordinate transformations. This suggests that the nervous system itself also treats the problem as if it had the sub-structure of a series of coordinate transformations. 3) The control of the eye's movement is the most intensively studied simple movement for good reasons, namely, moving the eye is mechanically simpler than moving a limb, yet several of the fundamental problems that arise in considering the muscular control of simple directed movements appear to be present even within the simplified mechanical situation provided by the eye. Thus, the study of oculomotor control is of central interest in developing and refining our understanding of the control of directed actions. 4) Finally, the computation of coordinate transformations is the foundation of other important capacities, for example, the capacity to navigate. An understanding of the neurobiological basis of the brain's ability to compute coordinate transformations in the control of simple directed actions may yield principles that have broad application in cognitive neuroscience.

Coordinate Transformations

A coordinate transformation, in the very general sense in which I will use it here, is an operation that maps or relates points specified by coordinates in one 'space' to points specified by coordinates in a different framework. 'Space' is in quotes because it will often refer to something that is a space only in the mathematical sense of a metric space. A metric space is a collection of points on which a

distance relation is defined, so that it is possible to say how far apart points in the space are[1]. The dimensionality of a metric space is the number of variables whose values must be specified in order to specify a point in that space--one for a one-dimensional space (a line), two for a plane, three for ordinary physical space, and n for an n-dimensional space. The dimensions of an abstract space need not and often do not correspond to the dimensions of a physical space. Coordinate transformations take the vector (string of numbers or quantities) that specifies a point in one space and generate the vector for a corresponding point in another space. The only restriction on the transformations I consider under the heading of coordinate transformations is that the transformation carries nearby points in one framework to nearby points in the other framework. Thus, points that lie close to each other in the first space cannot correspond to points that are sprinkled all over the place in the second space.

In directed limb movements, the primary sensory input arises in one framework while the movement is effected in a different framework, which is, generally speaking, of higher dimensionality than the primary sensory space. This is illustrated in Figure 1, which is based on a well known experiment by (Fukson, Berkinblit, & Feldman, 1980) demonstrating that the spinal frog solves a variable coordinate transformation problem when it directs its scratching action of its hind foot toward an irritated spot on the skin of its foreleg. The sensory signal from the irritation specifies the location of the target in a coordinate framework anchored to the forelimb. The axes of this 2-dimensional forelimb-skin space are indicated by the lines labeled d and c in Figure 1, for distance down the limb and position around its circumference. A coordinate system like this, which is defined by (or anchored to) a one or two-dimensional sensory surface such as the skin of a forelimb or the retina or the basilar membrane, I will call a sensory space.

The coordinate system within which the action must be effected is suggested by the angles of the hind limb joint, labeled α, β, γ, δ. Such a system, where points are defined by joint angles (or, in reference to the eye, rotation angles) will hereafter be called a joint space (even when speaking of the eye, which, of course, has no joints, but behaves like the ball part of a ball joint). In the present case, the real joint space is only suggested by the illustration, because it shows each hind limb joint as having only one variable angle. As a result, the joint space portrayed in Figure 1 has only four dimensions, four angles that must be controlled or specified by the motor outflow. The actual joint space of the hind limb would be four-dimensional only if the joints were all hinge joints, that is, only if there were only one possible axis of rotation per joint. In fact however, some of the joints have more degrees of freedom than a hinge. For example, in the figure, it appears that the upper leg can only move forward and backward at the hip joint, but in fact it can also be raised and lowered, and it can be rotated. (In other words, the hip joint is a ball joint with three degrees of freedom.) Thus, during a movement of the real hind limb, three angles must be controlled at the

---

[1] A distance is defined if there is a procedure that, given any two points p, q in the space, specifies a quantity (a scalar) d ≥ 0, such that d(p,q) = d(q,p), d(p,q) + d(q,r) ≥ d(p,r) and d(p,q) = 0 iff p = q.

first joint alone. The diagram of all the variable angles in a multi-jointed limb is too complex for present purposes, so let us pretend that this is a two-dimensional frog, in which case the angles of the four joints of its hind limb define a four-dimensional metric space.[2]

Figure 1. *Tracings from film strips of the wiping motions of the hind limb of a high spinal frog wiping at an irritation of the elbow of its forelimb (small black square). The generation of the movements shown in the lower two panels requires a coordinate transformation that combines a skin surface position signal with signals indicating the positions of the limbs relative to the body. Redrawn from Fukson, et al. (1980) with permission of the author and publisher.*
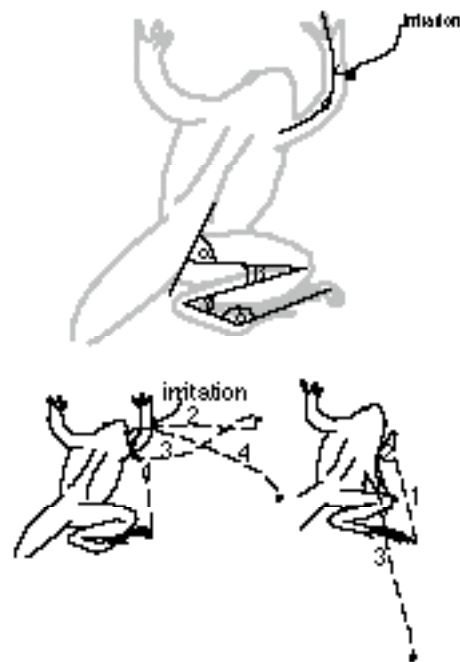
To scratch the right spot, the frog must adjust the angles α, β, γ, and δ so as to place the tip of its hind foot on the irritated patch. In the language of coordinate transforms, its neuromuscular system must effect a transformation that maps the point $<d_i, c_i>$ in sensory space to an appropriate point $<\alpha_i, \beta_i, g_i, \delta_i>$ in joint space. What makes this particularly challenging is that, as indicated in the lower half of Figure 1, the relation between these two coordinate frameworks changes as the position of the forelimb changes. The transformation that maps points in forelimb-skin space into hind limb-joint space must vary as a function of the position of the forelimb.

The experiment of (Fukson et al., 1980) shows that the spinal cord of the frog adjusts the motor output to the hind limb to compensate for the change in the position of the forelimb, that is, it in effect computes this variable coordinate transformation (lower part of Figure 1). The problem is to understand

---

[2] This simplification is the more appropriate in that in some of the work to be discussed later, involving microstimulation of motor centers in the spinal frog, the movements of the leg were (generally but not always) constrained to the horizontal plane.

how it does so. Are the neural processes in the spinal cord organized in terms of a series of coordinate transforms? If so, which coordinate transforms are performed, and where? When the transformations go from a space of lower dimensionality to a space of higher dimensionality, what further constraints does the nervous system impose in order to obtain a unique mapping? What structural aspects of the circuitry and what properties of the cellular level events within that circuitry effect the transformation (the mapping)?

Or, if the nervous system does not organize its operations in terms of a series of coordinate transformations, then by what principles can we understand what the nervous system is doing?

Two Contrasting Conceptions

While it may be natural for someone with a standard mathematical training to think about the frog's problem in terms of coordinate transformations, it does not follow that this conceptualization of what is happening will lead to an understanding of how the nervous system accomplishes what it accomplishes. In fact, theories about how the nervous system solves this problem may be contrasted on the basis of the extent to which they assume that conceptualizing the problem this way leads to valid insights about the underlying processes. This may be illustrated by considering two extreme positions. Although the views here to be contrasted are more extreme than would probably be defended by any contemporary theorist, the first view is roughly the view that an engineer with experience in the design of robot arms might bring to the problem (e.g., (Hollerbach, 1982)), while the second view would be more congenial to a connectionist modeler (e.g., Smolensky, 1988). In fact, of course, the views of individual scientists are apt to evolve over time, so variants of both views may be found in the work of a single influential figure (Robinson, 1975, 1985; Robinson, 1992).

One Transformation After the Other: An Engineer's View

In this conceptualization, there is a separate stage of computation for each of the stages that arise in a conventional physical analysis of the problem, such as would be made by an engineer trying to make a robot that did what the neural tissue in the spinal cord of a frog so readily does.

First, both the primary sensory coordinates for the irritation (in forelimb-skin space) and the primary sensory coordinates that specify the position of the hind foot in joint space are mapped (transformed) into the same three-dimensional space, for example, the Cartesian coordinate framework with the rostro-caudal, medio-lateral, and dorso-ventral axes used by anatomists to describe positions relative to a body. Hereafter, I refer to this coordinate system as <u>body-centered space</u>. Coordinate frameworks like the framework for body-centered space are also called extrinsic coordinates to distinguish them from intrinsic coordinates systems, which are defined by joints or muscles. The transformation of the forelimb-skin point into a point in body-centered space requires, in addition to the signal from the irritated patch of skin, signals that

specify a point in forelimb-joint space, the point defined by the current position of the forelimb. Thus, this sensory transformation combines a point in forelimb-joint space with a point in forelimb-skin space to yield a point in extrinsic space.

In this conception, mapping the two points into a common extrinsic system of coordinates (body-centered space) is a precondition for planning the <u>path</u> of the movement that will bring the tip of the hind foot to the irritated patch of skin. The planning algorithm must also specify the <u>time course</u> of the movement, where it will be along the path at successive moments. The path in body-centered space is a set of points in that three-dimensional space constituting what a geometer would call the 'curve' connecting the starting point to the end point (curves in this usage include straight lines), while the <u>trajectory</u> in body-centered space is a set of points (curve) in 4-dimensional space, the fourth dimension being the temporal dimension. Thus, a path has no temporal dimension, while a trajectory does. The planning of the trajectory of a movement is commonly called <u>kinematic</u> planning, at least in robotics (the engineers' use of the terms kinematic and kinetic differs somewhat from that of physicists).

When a trajectory in body-centered space has been specified, the planned trajectory is realized by means of a series of transformations, each of which may be conceived of as a mapping from points specified in one system of coordinates to points specified in a different system of coordinates, in other words, as a coordinate transformation. First, there is the <u>inverse kinematics</u> transformation of the trajectory. This transformation carries a trajectory in body-centered space, which by definition has three non-temporal dimensions, into a trajectory in the higher-dimensional joint space. This transformation poses a knotty problem, because it carries points into a higher dimensional space. All of the transformations considered so far carried points from a space of higher dimensionality into a space of lower dimensionality. These higher-to-lower transformations were realizable by functions, in the mathematical sense, that is, operations or processes that produce for any given input one and only one output. For a point in hind limb-joint space (that is, for specified values of the variable angles of the joints of the hind limb), there is one and only one point where the tip of the foot can be in body-centered space. Thus, there is a function that carries points in the joint space into points in the body-centered space. (This is the <u>forward kinematics</u> transformation.) The reverse is not true. For a point in body-centered space, there are many corresponding points in hind limb-joint space, that is, there are many different combinations of angles for the hind limb joints all of which place the tip of the foot at the same point in body-centered space. You can verify this by reaching out to touch a point on your desk, then varying the configuration (hence, the joint angles) of your forearm while keeping your finger on that same point.

This lower-to-higher mapping problem, first highlighted by (Bernstein, 1967), is called the <u>degrees of freedom</u> problem. It means that the problem of finding 'the' trajectory in joint space corresponding to a trajectory in body-centered space is not well defined. The nervous system must make it well defined by imposing restrictions on trajectories in joint space, restrictions that reduce the variety of possible trajectories to one actual trajectory. The question

then becomes, what restrictions does the nervous system impose in computing this transformation?

The inverse kinematic transformation specifies the values of the joint angles as a function of time. The final stage in computing the time course of the signals to be sent to the muscles is to solve the <u>inverse dynamics</u> problem, which is: given the time courses of the desired changes in joint angles, find the time course of the force to be exerted by each relevant muscle. The pulling forces exerted by muscles are not pure torques, that is, they are not directed perpendicular to the direction of joint rotation. And, they are related in extremely complicated ways to the torques that do develop (Zajac & Gordon, 1989). Finding the inverse dynamics can also be thought of as a coordinate transformation problem, this time from trajectories in joint space to trajectories in muscle space. The dimensionality of muscle space is the number of muscles that move the limb. Because this number is greater than the dimensionality of joint space--there are more muscles controlling a limb than there are degrees of freedom in its joints--the degrees of freedom problem arises again. The reappearance of the degrees of freedom problem, together with the strong non-linearities in the biomechanics of limb and muscle, makes the inverse dynamics an intimidating problem(Saltzman, 1979). It is impossible to derive analytic expressions for the requisite functions[Hazan, 1991 #2511]. It is at this point that the engineer may begin to wonder how the nervous system could compute this particular, very messy transformation and whether it can in some sense avoid doing so.

One Big Look-Up Table: A Radical Connectionist's View

Because the inverse transformations required in the above conceptualization of the problem are refractory to analytic treatment, one is led to question whether the conceptualization in terms of a series of coordinate transformations conforms at all to what actually occurs in the nervous system. A connectionist modeler might be inclined to reject this kind of computational decomposition of the problem into a sequence of transformations and think instead of a single overall mapping from sensory vectors (the array of signals in the first-order sensory axons) to motor vectors (the array of signals in the motor neurons). On this view, the connections in the network of interneurons (the hidden layer) intervening between the sensory neurons(input layer) and the motor neurons (output layer) have been adjusted by error-correcting feedback processes so that different patterns of input evoke optimized outputs (by some criterion defined by the feedback process). The nervous system may act as a gigantic look-up table, a table that specifies outputs given inputs, but does not do so by means of any analytically describable process.

In this view, the relevant inputs--primary visual afferents, primary somatosensory afferents, primary muscle and joint afferents, etc.--come from many different frameworks, so they cannot be conceived of as together defining points in any one space. Also, one school of connectionism has tended to emphasize the possibility that within the hidden layer (the network of interneurons), there may be no pattern to the activity of individual neurons or of

pools of neurons that would relate in any systematic way to the kinds of coordinate frameworks that an engineer uses to conceptualize the problem [Hazan, 1991 #2511; Kalaska, 1992 #1545; Lehky, 1992 #1698; Lockery, 1993 #1982; Smolensky, 1988 #2506]. Indeed, the pattern of connectivity and of interneuron activity that arises under the influence of error-correcting feedback may vary radically from one network to the next due to variations in the initial conditions of the network. In this case, knowledge of the pattern of intervening neuronal activity in one network that solves the problem might offer no insight into what is going on in another network that solves the same problem, even though both networks developed their structure under the impact of the same experiences and the same error-correcting feedback process.

If the radical connectionist vision gives an accurate conception of how the nervous system solves the problem of directed action, then the attempt to describe what is going on in the nervous system in terms of a series of coordinate transformations is doomed from the outset. Einstein once remarked that the most incomprehensible thing about the universe was that it was comprehensible--by which he meant mathematically describable. Perhaps the nervous system's way of solving difficult problems is not mathematically describable.

There are, of course, compromises between the two extremes just described. One interesting compromise treats connectionist networks as non-analytic function approximators, that is, physical devices that can be tuned to approximate almost any function, including functions for which it is impossible to derive an analytic expression, such as the inverse dynamics function in the control of a multi-joint limb. Networks may even be conceived of as linearizing strongly nonlinear dynamic control problems, so as to present to the higher levels of the nervous system a set of basis functions from which any desired kinetics may be realized by additive composition (about which more later).

Directed Limb Movements

Kinematics

Path Characteristics

The kinematics of the wrist during directed human arm movements have been studied in a variety of experiments. For a movement between two points, there is very little variability in the trajectory, regardless of the speed of the motion or the load (weight) carried by the hand (C.G. Atkeson & Hollerbach, 1985; Flash & Hogan, 1985; Lacquaniti, Soechting, & Terzuolo, 1982, 1986; Morasso, 1981; Soechting & acquaniti, 1981). The trajectories are usually straight (or at least only moderately curved) lines in either body-centered space (Morasso, 1981) or joint space(Hollerbach & Atkeson, 1987). (A straight line in joint space means that the ratios of the angular velocities of the joints are maintained throughout the movement.) In general, a trajectory cannot be straight in joint space if it is straight in body-centered space, and vice versa, although there is an important exception, namely, when the trajectory in body-centered

space lies along a straight line through the shoulder (Hollerbach & Atkeson, 1987).

Because maximum kinematic smoothness is realized by straight paths (Hogan & Flash, 1987) , one might hope to deduce from the straightness or lack of straightness of the trajectories in the two kinematic spaces whether the trajectory is planned in body-centered space or joint space (or neither). The fact that the trajectory can be strongly curved in either one space or the other depending on the work space (the region of body-centered space within which the starting and ending points of a trajectory are found) does not permit an unequivocal decision in favor of either planning space (Hollerbach, et al., 1987) .

The fact that freely chosen trajectories are curved in some parts of either body-centered space or joint space might even be thought to favor the third alternative--that is, there is no planning space, the radical connectionist view. However, it is not clear why on a radical connectionist view, the trajectories should tend toward straightness in most of the work space, nor why they should be so similar between subjects. The answer would presumably lie in something about the criteria that the error-correcting feedback process uses to determine error. One suspects, however, that the specification of this criterion would amount to assuming that the "teacher" (the feedback determining process) has a space in which it evaluates trajectories, that is, the teacher does more than assess whether or not the desired endpoint was reached. There is also a question whether the teacher can teach in the absence of internal models of the dynamics (Atkeson, 1989; Jordan, 1994a). Internal models are the sort of thing that a <u>radical</u> connectionist eschews, but they are the sort of thing that a moderate connectionist might imagine that a neural network provides.

It has also been suggested that the variety of trajectories observed in body-centered and joint space might be a byproduct of optimizing the smoothness of joint torques (Uno, Kawato, & Suzuki, 1989), that is, it may be a byproduct of a dynamic rather than kinematic planning. (Maximizing smoothness means minimizing jerk, which is the third derivative of position as a function of time.) In short, the data on kinematics per se do not point unequivocally to a particular system of coordinates in which the planning of limb trajectories are carried out.

Evidence from End-Point Variability

Recent results by(Gordon, Ghilardi, & Ghez, 1994) on the variability of movement end points suggest a stage in which the trajectory is planned in body-centered space, rather than joint space. Gordon, et al. found that the directional variability in the endpoint (that is, the dispersion of the directions of the endpoints of repeated trajectories around the mean direction relative to the starting point) was constant and independent of the length of the trajectory. This constancy held for patterns involving different joints and for both slow and fast movements. By contrast, end point variability along the axis of the trajectory (that is, variability in the length of a repeated trajectory) increased markedly but non-linearly with distance. Gordon, et al. argue that their results imply that the direction and extent of the movement in body-centered space are independently

computed. In vector terms, the orientation and length of the movement vector in body-centered space are separately computed and make separable contributions to the error in the endpoint.

Trajectory Adaptation Experiments

Strong evidence that trajectories are planned in body-centered space comes from recent adaptation experiments by(Wolpert, Ghahramani, & Jordan, 1994, 1995). They used an ingenious set-up in which the subjects moved a computer mouse on a digitizing tablet. While doing so, they looked at an illusory view of a spot that seemingly marked the position of their hand relative to the target square (the square that marked the end point of the movement they were to make). The target and the hand-spot were projected via a computer-controlled onto a mirror interposed between the subject's head and hand, creating a virtual image in the plane of the hand's movement (an image that appeared to originate from the surface of the digitizing tablet along which the hand was moving). This arrangement enabled them to provide erroneous visual information about the hand's position during the course of the movement without an error at the beginning and end of the movement. That is, they were able to make straight trajectories appear curved and vice versa, while keeping the perception of the location of the beginning and ends of the trajectories constant. They had subjects make repeated back and forth movements that were either transverse (for some subjects) or sagittal (for other subjects)--that is, either perpendicular to or in the sagittal plane of body-centered space. During the first 80 such movements, the trajectory was made to appear more curved. During the final 20 trials, the subjects made the same movement in the absence of visual feedback, so that the experimenter's could assess the aftereffect of the adaptation experience.

        A purely dynamic planning process--a process that computed the time-course of the muscle forces required to bring the hand from its starting position to its ending position without regard to the trajectory through 3-dimensional body-centered space--would not be affected by this artificial alteration in the apparent trajectory of the hand during the adaptation phase. If, however, there is a stage that plans the trajectory in body-centered space, and if that stage takes visual input as a reliable indicator of position in body-centered space, then this stage should register a trajectory error. This trajectory error might then be used to make adjustments in the inverse kinematics transformation so as to offset the error in the trajectory through body-centered space. This was what they in fact found: In the twenty movements made in the absence of visual feedback following the adaptation phase, the subjects' trajectories showed a significant increase in curvature, an increase in the direction that would offset the visual curvature, that is, straighten the apparent trajectory. This is strong evidence for a stage that plans a trajectory in body-centered space.

   In these experiments the pre-adaptation trajectories conditions were gently curved. In companion experiments(Wolpert et al., 1994, 1995), the authors show that this curvature is predicted by errors in the visual perception of the straightness of the trajectory. Subjects perceived a straight trajectory of a moving spot as in fact curved; to get them to perceive the trajectory as straight,

the actual trajectory had to be curved. There was a highly significant correlation between the curvature a subject perceived as straight and the curvature of the subject's actual hand movements.

Not all trajectory curvature can be explained as a consequence of the misperception of straightness in body-centered space. When subjects move their hand from a position in front of them to a position not far above their shoulder, the curvature in the trajectory is much too great to be explained in this way. The pronounced curvature in the trajectory through body-centered space that is observed in this case (arguably a special case) is presumably dictated by the greater simplicity or smoothness of the trajectory in joint space. Thus, it appears, that planning in body-centered space is not obligatory; plans may be constructed in joint space.

Velocity Profiles

The velocity profiles of directed trajectories are bell shaped: the wrist accelerates smoothly along its path to reach a maximal tangential velocity midway in the movement, then decelerates just as smoothly, so that the second half of the velocity profile is nearly the mirror image of the first (Figure 2). These Gaussian velocity profiles minimize the jerk ((N. Hogan & Flash, 1987), which suggests that smoothness considerations play an important role in planning the trajectory. Although, it is not at all intuitively obvious, these smoothness considerations also dictate straight line trajectories in body-centered space (Flash & Hogan, 1985) Smoothness considerations in joint space may also play a fundamental role in specifying unique solutions for the inverse kinematics transformation, that is, the nervous system may pick out the smoothest trajectory in joint space that realizes the already planned trajectory in body-centered space.
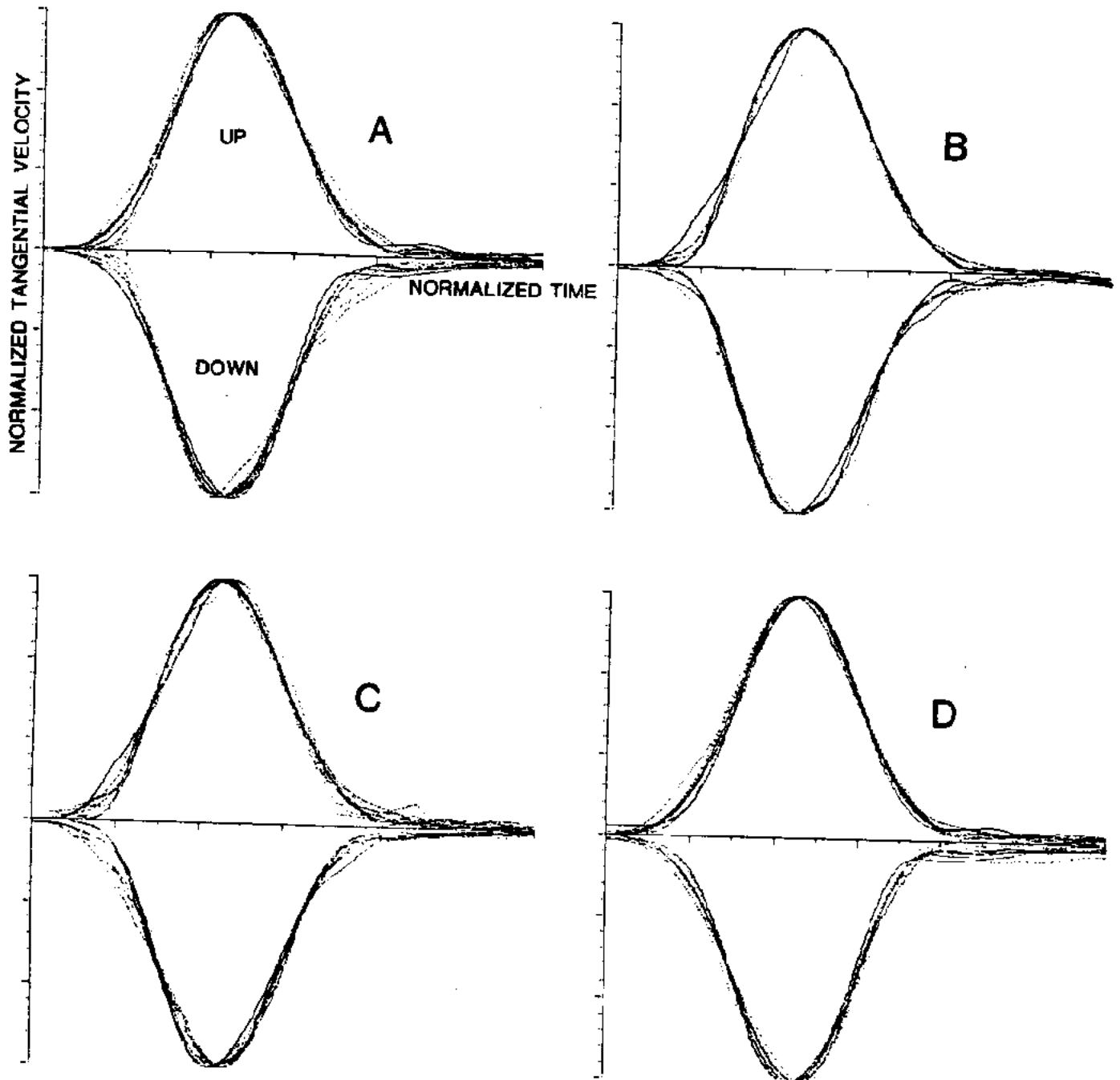
Figure 2. *Tangential velocity profiles normalized for speed and distance and aligned at their peaks. (a) Different speeds. (b) Different loads. (c) Different targets. (d) Different subjects. Reproduced from Hollerbach (1990) with permission of the author and publisher.*

Dynamics

A separable stage of dynamic planning, working to achieve a prespecified kinematics, is suggested by the fact that the velocity profiles are invariant (except

for scaling) in the face of substantial changes in the speed at which the movement is executed and the weight carried by the hand (Figure 2). Changing the speed and the weight carried changes substantially the forces that must be applied to achieve these profiles. Because of the non-linear biomechanics of the arm, the changes in the required forces are presumably not a simple scaling up in the forces required at lower speeds or when carrying less weight. That is, the pattern of required forces--the <u>relative</u> strengths of the forces that must be applied and the relative times at which they must be applied--must be altered in order to maintain the same profile as speed and load vary. This suggests that the kinematics of the movement are specified independently of the dynamics and that the mapping from the kinematic specification to the dynamic specification (the inverse dynamics) is varied so as to take into account the change in load or overall speed.

Can the Inverse Dynamics Problem Be Finessed by Feedback Control?

Because the inverse dynamics are so nonlinear and therefore difficult to model (compute), one is motivated to look for a control scheme that does not require the nervous system to have a model of the dynamics. One way to avoid computing the inverse dynamics is to use high-gain, error-correcting feedback control to develop the necessary muscle forces during the course of the movements, without computing in advance the motor nerve signals required to cause the muscles to develop the requisite forces. This latter, more computationally demanding approach to motor signal programming is called feed-forward or predictive control. (For an introduction to control systems concepts in a motor-control perspective, see(M.I. Jordan, 1994).)

Can the muscle-spindle organ do the job? At first glance, the muscle spindle system would seem structurally suited to provide the requisite feedback control. Muscle spindle organs are embedded in muscles. Each such organ has a specialized muscle fiber of its own, innervated by a distinct class of motor neurons (the gamma efferents). The gamma efferents can command the shortening of the spindle organ's muscle fiber independently of the shortening of the fibers of the muscle itself. The sensory fibers innervating the spindle organ respond both to the difference in length between the spindle fibers and the surrounding muscle fibers and to the rate of change in this difference. Thus, the spindle organ would appear designed to signal the error and the first derivative of the error between a muscle length command carried by the gamma efferents and the actual length of the muscle. This putative error signal is relayed to the spinal cord by the Ia afferents, the fastest conducting sensory fibers in the body, and these afferents make monosynaptic connections on the alpha motor neurons, the neurons whose signaling controls the development of force by the muscle in which the spindle organ is embedded.

This structure seems to cry out that this organ provides the feedback control of muscle force required to avoid computing in advance the inverse dynamics. However, measurements of both the gain and the feedback time in this loop show that it can be at most a minor player in the determination of muscle forces(Hasan & Stuart, 1988; Loeb, 1987). The sluggishness of the

feedback loop prevents its making any contribution to the control of movements that are executed in roughly half a second or less(N. E. Hogan, Bizzi, & Mussa-Ivaldi, 1987).

The moving-equilibrium-point hypothesis. Another suggestion is that increased speed and load compensation is achieved by stiffening the limb during the execution of its trajectory, allowing the visco-elastic properties of the muscles to solve the inverse dynamics problem without neural involvement. The general idea behind this approach to the dynamics problem is that the trajectory is planned as a series of positions in body-centered space. Corresponding to each position in body-centered space, there is a motor command that sets the length-tension curves for the muscles of the limb in such a way that this position is the equilibrium point of the mechanical system. (Thus, in this model, there is also no explicit computation of the inverse kinematics, the requisite trajectory in joint-angle space.) If the limb is not at that point, then the force-field due to the visco-elastic properties of the muscles (the net force moving the limb as a function of the limb's deviation from the equilibrium position) drives the limb toward the equilibrium position. This popular model of limb control was first proposed by(Feldman, 1974), see also (Berkinblit, Gelfand, & Feldman, 1986). It has been elaborated and experimentally tested by Emilio Bizzi and his collaborators. These experiments have demonstrated that the position of the equilibrium point changes smoothly and consistently during a movement and that there are invariants in the force fields surrounding equilibrium points(Bizzi & Mussa-Ivaldi, 1990; Bizzi, Mussa-Ivaldi, & Giszter, 1991; Mussa-Ivaldi, Giszter, & Bizzi, 1994).

The experimental findings show that the moving-equilibrium-point hypothesis captures an important aspect of the neurobiological approach to dealing with the dynamics, namely, that the trajectory and time course of directed limb movements is stabilized by some co-contraction of opposing muscles. This co-contraction of opposing visco-elastic elements (muscles) creates a dynamic force field whose equilibrium point (point where there is no net force) is centered on the momentary planned position of the limb's endpoint. This dynamic force field automatically corrects modest deviations from the planned trajectory.

However, this hypothesis cannot readily explain the fact that the kinematics of reaching movements are invariant under substantial changes in speed and substantial changes in load (the weight in the hand). As the movement is speeded up, the inertial forces that must be overcome, which are modest at low speeds, increase rapidly, and these forces are complex. It is these nonlinear reactive inertial forces that make the inverse dynamics problem so formidable. The force fields seen during low-speed movements are not sufficient to keep the limb from veering widely away from the usual trajectory in reaction to these forces. Increasing the stiffness of the limb (that is, the intensity of the force field at any given deviation from the equilibrium position) by increasing the amount of co-contraction in opposing muscles is energetically inefficient. It uses much of the energy that goes into muscle contraction to oppose the contraction of other

muscles instead of to do work (move loads). Moreover, no plausible amount of stiffening can keep veering within the small limits actually observed.

There seems to be no escaping the fact that the forces that will oppose the movement are to some extent anticipated, so that components of the command signals sent to the muscles in advance function to develop the torques required to offset these reactive forces. This has led to the suggestion that in order to generate those components of the motor signal that reflect the need to offset the reactive forces, the trajectory planning process transiently specifies bogus equilibrium points well beyond the desired end point(N. Hogan & Flash, 1987). This bogus trajectory, which deviates from the actually intended trajectory in a manner that generates the torque components required to offset the reactive forces is called a virtual trajectory. There is no experimental evidence for this kind of bogus virtual trajectory. And, perhaps more importantly, it is not clear how the system could choose an appropriate virtual trajectory without solving the inverse kinematics and inverse dynamics problem, the problems that this model was intended to finesse.

Another problem with the moving-equilibrium-point hypothesis is that it does not support efficient adaptive modification of motor commands during practice, modifications that overcome changes in limb dynamics produced by growth, pathology, and working with implements that have substantial inertia (hammers, stones, spears, poles, tennis rackets, etc.--Atkeson, 1989) . The problem is that because the nervous system has no representation of the relation between the kinematics and the dynamics, it has no way of relating the errors, which are kinematic (deviations from planned trajectories), to the dynamics (the pattern of forces that generated the trajectory). Thus, it cannot use observed trajectory errors to make goal-directed changes in the planning process; it can only proceed by random trial and error.

Adaptation to Artificially Manipulated Dynamics

While roboticists have not yet made computer-controlled robots that can reach and wipe as well as a frog can, they have nonetheless made enough progress to permit highly instructive experimental manipulations of the forces that a human arm encounters during a reach.(Shadmehr & Mussa-Ivaldi, 1994) had subjects make targeted reaching movements while grasping a handle on the end of a low-inertia robot arm. The robot arm had computer-controlled torque motors at its joints, which enabled the computer to generate forces on the hand as it moved. Sensors in the joints gave the computer moment-by-moment information on the angular position and velocity of its joints, from which it could calculate the position and velocity of the subject's hand in body-centered space or the corresponding values in joint-angle space (that is, the angular position and velocity of the subject's joints) by just the sort of coordinate transformations that are the focus of this review. This arrangement enabled the experimenters to program the computer to generate novel velocity-dependent force fields. Because these forces were velocity-dependent (like the forces that oppose the stirring of molasses), they did not exist until the hand began to move.

The subjects made control reaches with the same hand in two different work spaces, one on their right and one on their left, before the velocity-dependent force fields were introduced. Then they made 1,000 reaches in the right work space in the presence (mostly) of the artificial velocity-dependent force field. When the subjects first encountered the novel forces, their trajectories veered far from the straight lines characteristic of normal reaches, and the velocity profiles were multimodal rather than bell shaped. However, as they practiced making reaches in this bizarre dynamic environment, the trajectories again became straight and the velocity profiles bell shaped. During the subjects' adaptation to the artificially imposed force fields, a few trials where inserted on which the field was turned off. On these trials, the subjects trajectories veered away from the normative in ways more or less opposite and equal to the veers induced by the artificial force field when they first encountered it. These veers are the aftereffects of the adaptation. They show that the nervous system is sending signals to the muscles that anticipate and counteract forces that will develop during the course of the planned movement. When the anticipated forces do not materialize the muscle forces intended to counteract those forces cause veers.

The aftereffects prove that muscular forces designed to offset forces that will develop during the movement are programmed in advance into the signals sent to the muscles. In other words, the problem of generating the requisite forces is not solved primarily by feedback control; it is solved by feed forward control. This same conclusion--that trajectory constancy is not obtained by stiffening the limb-- follows from the finding that as subjects adapt to experimentally imposed perturbations in limb movements, limb stiffness declines rather than increases(Milner & Cloutier, 1993; van Emmerik, 1991). The nervous system tries to minimize deviations from programmed trajectories by increasing the gain of the visco-elastic feedback control loop only at first, before it has learned to compensate for the new reactive forces by feed-forward control.

The fact that the nervous system learns to compensate for new reactive forces by feed-forward control strongly suggests that it has a modifiable dynamic model of the limb(M.I. Jordan, 1994; M. I. Jordan, 1994), which enables it to solve the inverse dynamics problem, the problem of calculating the forces required to implement a given trajectory and velocity profile. The model of the limb's dynamics has feedback-settable parameters. When the world changes so as to invalidate the model, the resulting error signals adjust the parameters of the model to make it once again a usable model of the limb's dynamics. The model may even be context-specific; that is, when you pick up a familiar implement like a tennis racket or a hammer, the central nervous system may switch-in a different model of the limb's dynamics, a model that incorporates the contributions of the implement.

Perhaps the most elegant aspect of the Shadmehr and Mussa-Ivaldi experiment is that they trained their subjects with velocity-dependent force fields defined in the two different kinematic spaces-- body-centered space and joint-angle space. For one group, the velocities and the resulting artificial forces were

specified in joint-angle space, while for another group, they were specified in body-centered space. The force fields were chosen so that they were essentially identical when the hand was moving in the right work space where the adaptation training took place. Thus, there was no way for the subject to know which space defined the artificial force field to which he was adapting. That is, there was no way of knowing whether a given force (with a certain magnitude and direction) was evoked by a joint-angle velocity or by the resulting velocity of the hand in body-centered space. However, when the adapted subjects moved their hand over into the left work space, the force-fields produced by the two different programs were no longer identical in body-centered space. In fact, they were nearly orthogonal, that is, the force produced by one program in response to a given velocity of the hand in body-centered space was almost perpendicular to the force produced by the other program in response to the same velocity.

When the subjects shifted their reaches over to the left work space, they were tested with both field-generating programs and also with the imposed force field turned off. This allowed a decisive test of two important questions: (1) Did the adaptation transfer from one work space to another? (2) If so, did the transfer occur in body-centered space or in joint-angle space?

The adaptation did transfer from the right work space to the left one, and the transfer clearly occurred in joint-angle space, not body-centered space. When there was no force-imposing program in the left work space (the space in which they had not practiced), both groups of subjects showed strong and very similar veers (adaptation aftereffects). When the force field that depended on velocities in body-centered space was present in the transfer work space, the subjects showed strong veers, regardless of which program they had adapted to. The adaptation they had learned in the training work space did not enable them to compensate properly for the 'same' forces in the new work space, where 'same' means same if defined in body-centered space. By contrast, when the force field that depended on the joint-angle velocity vector was operative in the new work space, both groups of subjects compensated well. They did not make veers in the new work space when they encountered the 'same' velocity-dependent forces they had encountered in the previous work space, where 'same' now means same if defined in joint-angle space.

The fact that proper compensation transferred only when the force fields were defined with reference to joint-angle space not body-centered space is strong evidence that the inverse dynamics are computed after an inverse kinematics transformation. What the subjects learned were the forces required to implement a given trajectory (and velocity profile) in joint-angle space. How they could learn this if the trajectory in joint-angle space were not represented in the nervous system is unclear. The experiments by Wolpert, et al. (1995) showing adaptation to false-visual curvature give strong evidence for trajectory planning in body-centered space. The Shadmehr and Mussa-Ivaldi results give strong evidence that the dynamics are computed from a trajectory specified in joint space. Together, these experiments argue for the succession of transformations envisioned by the engineering analysis: first a trajectory in body-centered space; then a derived trajectory in joint space (the inverse kinematics computation);

finally, a derived feed-forward trajectory in muscle-force space (the inverse dynamics computation).

The transfer of the adaptation to the new work space, implies that the inverse dynamics are not computed by means of a look-up table. A look-up table gives an output (e.g., a pattern of motor-neuron signals) for each input (e.g., a set of sensory signals indicating limb position and velocity and a set of command signals indicating a desired trajectory). The values of the signals that come into a look-up table do not define a position in a space, at least if the look-up table is what is called a 'dumb' look-up table, one that does not do interpolation. With a dumb look-up table, there is no sense in which one incoming pattern of signals is close to (similar to) another pattern. Because there is no sense in which patterns recognized as distinct inputs are close or distant from one another, the dumb look-up table neither interpolates or extrapolates.

Interpolation implies that inputs may be ordered along various continuous dimensions, that is, that they are points in a space on which a metric may be defined. (Look-up tables that do interpolation are called smart look-up tables.) Extrapolation implies not only that the inputs and outputs to the mapping are points in metric spaces, but also that the mapping between one region of the input space and its corresponding region in the output space defines the mapping for other regions of the input space. This is something akin to analyticity in the theory of functions, and, indeed, neural networks that specify a function over its entire domain are called function-approximators. (A function is analytic if its behavior over any region of its domain (input) defines its behavior over its entire domain.)

The subjects in the Shadmehr and Mussa-Ivaldi experiment extrapolated the adaptation they learned during training in the right work space to the left work space. Moreover, and particularly to be remarked, this extrapolation was on the mark. The extrapolated control signals enabled them to produce straight trajectories in body-centered space in a new region of joint-space (and body-centered space), provided that the force field was programmed in joint space. Anyone who has used a polynomial function to fit nonlinear data should be impressed by this, because the best-fitting polynomial generally deviates wildly from any plausible further data as soon as it gets outside the region for which one already has data. In other words, polynomial functions fitted to observations seldom correctly anticipate what will happen in a new region of the space being experimentally explored; they do not extrapolate correctly. The Shadmehr and Mussa-Ivaldi transfer results imply that in generating the motor commands from the joint-space kinetics, the nervous system operates with something akin to analytic functions, which are sufficiently accurate reflections of the limb's biomechanics that they yield valid extrapolations.

Dynamic Basis Functions

One of the most important ideas in the theory of metric spaces is the concept of a basis function, which may be thought of as a generalization of the notion of the basis for a vector space, one of the fundamental ideas in linear algebra. Recent experimental results suggest that this idea may have relevance to our

understanding of how the nervous system generates the muscle commands that induce a limb to follow a planned trajectory(Giszter, Mussa-Ivaldi, & Bizzi, 1993; Mussa-Ivaldi et al., 1994).

A basis for a vector space is a set of points (that is, vectors) that may be scaled (multiplied by a scaling factor) and combined (by vector addition) to yield any point (any vector) in the space. The traditional bases for any three-dimensional physical space such as the body-centered space are the vectors that specify a point one arbitrarily chosen metric unit along each of the orthogonal axes--the x, y, and z axes by reference to which the positions of points in the space are defined. These orthogonal unit vectors are <1, 0, 0>, <0, 1, 0> and <0, 0, 1>. Scaling one of these vectors means multiplying each of its components by some scaling factor, which can be any real number. Adding scaled vectors means adding (in the conventional arithmetic sense) corresponding components (the first component of one to the first component of the other, the second to the second, and the third to the third). Obviously, by these two operations applied to these orthogonal unit vectors, one can generate any sequence of three real numbers one wants which is to say any point in the space. Thus, for example, the point <3.2, √2, π> is obtained by scaling the first basis vector by 3.2, the second by √2, and the third by π, then adding these scaled orthogonal unit vectors.

A basis function generalizes this idea to the (infinite) sets of points that constitute curves, that is, trajectories. ('Curves' in this usage include straight lines.) Probably the most generally familiar basis for continuous function space are the sine and cosine curves used in the Fourier decomposition and synthesis of a function. Roughly speaking, the Fourier theorem asserts that: (1) one may obtain any continuous curve (thus, any trajectory) by scaling and adding to each other (superimposing) some set of sine and cosine curves; and (2) for any given function, the requisite set and the required scaling factors are unique. Thus, the sine and cosine functions constitute a basis for a trajectory space--you can get any trajectory you want by scaling them and superimposing them, adding them up point by point. (For an illustrated introduction to the superimposing of sine and cosine curves, see (Gallistel, 1980).) Technically, speaking, it requires an infinite set of sine and cosine curves to make a basis for continuous function space, but, practically speaking, a good approximation to almost any trajectory can be achieved by scaling and superimposing a modest fixed set of sine and cosine curves--a set containing on the order of 50 curves. Thus, a finite set of sine and cosine curves can serve as a practical basis for generating any trajectory one wants.

The sine and cosine curves are by no means the only basis for trajectory space. There are many other possibilities, some of which offer appreciable advantages over the Fourier set in some contexts (e.g., wavelets, see(Strang, 1994)). The importance of the Fourier basis set in the present context is only to serve as an illustration of the general idea of a set of basis functions.

This short introduction to the concept of a set of basis functions is by way of preamble to a review of the results obtained by (Giszter et al., 1993) and (Mussa-Ivaldi et al., 1994)., working with the spinal frog. They mapped out

the force-fields produced by brief (0.3 s) trains of low intensity (1-10 µA, 1 ms wide) anodal stimulating pulses delivered (at 40 Hz) through a microelectrode to interneuron pools in the spinal cord. To map the force field induced by the stimulation (the active force field), they anchored the ankle to a 6-axis force transducer on an x-y positioning stage and moved the foot to different points in body-centered space. At each position of the foot, they noted the force vector exerted by the limb before and during stimulation. By subtracting the force vector before stimulation from the force vector that developed during stimulation, they derived the <u>active</u> force vector as a function of time since the onset of stimulation.

The force vectors obtained from a grid-work of points in body-centered space define a force field. The static force field observed in the absence of stimulation and the active field developed during stimulation were generally <u>convergent</u>. That is, there was a point in space at which there was no force--the <u>equilibrium point</u>--and at all other points the force vectors pointed along curves that converged on this equilibrium point. The static equilibrium point was the position to which the unstimulated frog's ankle would return if released from the positioning device, in other words, its resting posture. The active equilibrium point was the point to which the ankle would move during stimulation if it were free to move and if the static, prestimulation force field were absent. Not surprisingly, the strength of the vectors in the active field increased during stimulation.

The total force field acting on the ankle at any moment during the stimulating train was the superimposition (sum) of the static field and of the active field. The active field grew stronger over time since the onset of stimulation. Therefore, the equilibrium point of the summated force field moved during stimulation--to a point that was intermediate between the static and the active equilibrium points. The stronger the active field became relative to the static field, the nearer the equilibrium point of the summated field approached that of the active field. The authors term the trajectory of this moving equilibrium point of the summated field the virtual trajectory, to distinguish it from an actual trajectory the ankle would trace out if it were free to move. When they did allow the ankle to move from its resting position during stimulation, its actual trajectory was generally close to this virtual trajectory.

When they determined the active force fields produced at different depths as they moved the stimulating electrode deeper into the lateral portion of the cord at any given point along the anterior-posterior axis of the cord, they found that the force fields were very similar at different depths--a finding reminiscent of the columnar structure of receptive fields in sensory areas of the cortex. When they varied the strength (stimulating current) or duration of the train of pulses, they found that both of these manipulations served chiefly to change the lengths of the force vectors not their orientation nor the time course of their development. If the stimulating train is itself conceived of as a 'pulse' input whose height is defined by the strength of stimulation (current x pulse frequency) and whose width is defined by train duration, then we can say that varying the energy in the pulse (its area) by varying either its width or its height

scales the active force field. The importance of this is that scaling the basis functions is one of the two basic operations by which other functions are synthesized from basis functions.

More interestingly--from the standpoint of where we are headed, which is toward a concept of a dynamic basis set --they found that the active force fields from many different stimulating sites fell into only about four distinct classes of fields. Stimulation at sites in one class created a force field that moved the tip of the leg forward and out, stimulation at sites in another class moved it back and in, and so on. Most interestingly, when (Mussa-Ivaldi et al., 1994). measured the force fields produced by stimulating simultaneously through two different electrodes that produced different classes of force fields, they found that the resulting active force field was the superimposition (the adding up) of the component active force fields. That is, for any given position of the frog's ankle in space and at any moment in time after stimulation onset, they could predict the active force vector by adding the active force vectors at that point and time obtained when stimulating at each field individually.[3]

Because the force fields superimposed, (Mussa-Ivaldi et al., 1994). could predict the virtual trajectory produced by stimulating the two classes of interneuron pools simultaneously. Note, however, that it was not the virtual trajectory itself that could be obtained by superimposing the two virtual trajectories; rather it was the force field that could be obtained by superimposing the two force fields. Thus, if these different classes of force-field producing interneuron pools are basis functions for the generation of trajectories--which is what (Mussa-Ivaldi et al., 1994) suggest they are--then the space for which they constitute a basis set is a dynamic space not a trajectory space.

These interneuron pools in the spinal cord may be a fundamental part of the neural machinery that computes the inverse dynamics. In fact, they could be the part whose output is modified during adaptation to novel dynamics, although it should be borne in mind that evidence of such adaptations has not been obtained in the frog. The system could use information about how previous trajectories deviated from those specified to alter the parameters of the circuits that create these force fields in ways that compensated for the novel dynamics. This adaptive capacity at the level of the dynamic basis functions would make changes in the dynamics "transparent" to higher planning stages--something that they did not have to take into consideration in carrying out their coordinate

---

[3] A surprising finding in this research was the superimposition (additive combination) of the force fields elicited by direct stimulation of muscles in a limb whose inverse kinematics were not experimentally constrained (i.e., where many different points in joint space could correspond to a single point in anatomical space). This suggests that the biomechanics of limbs are not as intractably non-linear as they are sometimes supposed to be. The muscles appear to insert in such a way that their effects combine additively even though they need not.

transformations. However, as already noted, this kind of adaptation requires that the nervous system have a model of the limb's dynamics.

Dynamic basis functions may simplify the control problem from the standpoint of the higher stages of motor planning in two ways: (1) by reducing the degrees of freedom, the number of variables whose values need to be specified in order to obtain the desired output; and (2) by linearizing the 'apparent' dynamics. They may, in effect, give the next stage up four control levers (assuming four basis functions) to "pull" on--for example, one giving a forward-and-outward force field converging on a stable position at the forwardmost and outermost limit of reach; a second giving a backward-and-outward force also converging on a stable limb position, that is, an equilibrium point; a third giving a backward-and-inward force (likewise convergent); and a fourth giving a forward-and-inward force (again convergent). Because the force fields superimpose when the next stage up pulls on more than one lever at a time, the consequences of various combinations of lever pulls are easily computed or modeled. This is the advantage of having a stage that linearizes the apparent dynamics.

Conclusions

The behavioral evidence--particularly the evidence from adaptation experiments--favors the view that there are separable stages in the planning of a directed limb movement--a stage that plans the trajectory in three-dimensional body-centered space; an inverse kinematics stage, which translates this into a trajectory in joint space; and, finally, an inverse dynamics stage, which translates the joint-space trajectory into a dynamic, convergent force field.

A convergent force field is defined in that sub-space of body-centered space consisting of all the positions that may be occupied by the end point of the limb. The force vectors acting on the end point of the limb at the positions in this space are all directed along curves that converge on a single equilibrium point, a point where the force vector is zero. In a dynamic, convergent force field, the equilibrium point traces out a trajectory, that is, it changes position over time.

Evidence from spinal cord microstimulation in the frog suggests that the inverse dynamics may be implemented by means of a modest set of interneuron pools, with each pool producing a dynamic basis function. These basis functions are the primitives from which all dynamic convergent force fields are synthesized by superimposition--the point-by-point addition of the force vectors. Concurrent activation of these interneuron pools generates the dynamic, convergent force field required to implement a joint-space trajectory. Thus, the problem of computing the inverse dynamics becomes the problem of determining which combinations of these basis functions must be activated and in what temporal sequence in order to produce the joint-space trajectory.

In humans at least, the system that computes the inverse dynamics adapts to an altered dynamic environment. This adaptation is not realized by stiffening the limb, as is required by the moving-equilibrium-point hypothesis. Nor is it realized by computing a virtual trajectory that deviates from the desired

trajectory in such a way as to realize the desired trajectory in the face of the altered dynamics. The adaptation is invariant under translations of the altered dynamics computed in joint space but not under translations computed in body-centered space, which appears to be strong evidence that the joint-space trajectory is explicitly computed in the course of generating a movement. Moreover, the ability of the system to adapt to these perturbations with feed-forward corrections--that is by programming forces that are calculated in advance to offset the altered dynamics--is thought to require at least a crude internal model of the dynamics, called a model of the "plant" (C. G. Atkeson, 1989; M.I. Jordan, 1994). An internal model of the plant is a neural network that generates from a copy of the control signals sent to the muscles the expected pattern of signals from the sensors that report the limb trajectory actually achieved. In other words, the neural model is a stand-in for the limb itself; it mimics the manner in which the limb converts motor signals into sensory signals, signals that indicate the consequences of the muscle forces developed in response to the motor signals. Such a model is thought to be required in order for the system to make appropriate adjustments in the inverse dynamics mapping when the trajectories actually achieved deviate systematically and repeatedly from the trajectories specified by the planning process.

In planning trajectories in body-centered space, the system generally adopts trajectories that maximize smoothness, which is to say trajectories that minimize the jerk (the first derivative of acceleration). Jerk-minimizing trajectories are straight lines. This straightness is not a by-product of some other planning goal (e.g., a purely dynamic goal), because if the trajectories actually produced are made to appear curved when they in fact are not, the subject adapts, that is, he begins to produce trajectories that are in fact curved but appear straight.

In computing the inverse kinematics, the nervous system must impose additional constraints on the solution to make the problem well posed, that is, to insure that there is a unique solution, because there are infinitely many joint-space trajectories that will implement any given trajectory in body-centered space (the degrees of freedom problem). The further constraint that is imposed may again be a smoothness constraint, that is, the inverse kinematics transformation may pick out from this infinitude of joint-space trajectories the one that minimizes the jerk in joint space. It should be noted that this joint-space trajectory will not (necessarily) be the smoothest possible joint-space trajectory; rather, it will be the smoothest joint-space trajectory that implements the required body-centered trajectory. In other words, the system may first plan the smoothest possible body-centered trajectory, then plan the smoothest joint-space trajectory consistent with this body-centered trajectory (sequential constraint satisfaction).

Saccadic Eye Movements

A saccadic eye movement is a high-speed ballistic movement of the eyes from one direction of gaze to another direction of gaze. It may be directed to a punctate visual target; or it may be elicited by an auditory stimulus and directed

toward the computed position of the source; or its direction may be specified by purely internal processes in the absence of any punctate target stimulus. It functions to bring the image of the source position (the position of the distal stimulus in head-centered space) onto the fovea. The movement is ballistic in the sense that it is not influenced by the movement of the visual field across the retina that ordinarily occurs during a saccade. Manipulating this retinal feedback experimentally, by, for example, stabilizing the retinal image does not alter the trajectory of a saccade.

Although there is much that remains to be understood about the control of saccadic eye movements, this is nonetheless the best understood system from a neurobiological standpoint, and the one that provides the most compelling evidence for neuroanatomically distinct stages of coordinate transformation.

Integrator Coordinates: A Neurally Imposed Framework

One of the coordinate transformations that the saccade-generating neural circuitry must compute is necessitated by the peculiarities of rotational kinematics, namely, the non-orthogonality of rotations about orthogonal axes. A ball rotating in a socket has three degrees of rotational freedom: it can rotate horizontally (to the left or to the right about a vertical axis); it can rotate vertically (up or down about a transverse axis in the horizontal plane); and it can rotate torsionally (clockwise or counterclockwise about a sagittal axis in the horizontal plane, an axis that passes through the pupil when the eye looks straight ahead).[4] To appreciate the non-orthogonality of positional changes in a three-dimensional rotational framework, imagine that the eye rotates horizontally 90° away from straight ahead gaze. This rotation brings the pupil of the eye to the point where the transverse axis of rotation enters the eyeball. (Such an extreme rotation is anatomically impossible, but the consequences of this rotation are easy to visualize, and the conclusions that follow apply in intermediate degree for any intermediate rotation.) Now imagine that the eye rotates 45° 'vertically,' that is, about the transverse axis, which now passes

---

[4] The consequences of a sequence of such rotations can be hard to visualize. To follow this exposition, the reader may find it useful to take the ball out of a computer mouse and mark a cross at the "pupil" of this ball, taking care to distinguish the horizontal and vertical bars of this cross. The axis for horizontal rotations passes vertically through the center of the ball. When the ball is held with a finger tip on top and the thumb on the bottom, it rotates about this axis. The axis for vertical rotations passes transversely through the center of the ball. Holding the ball at the points where a transverse axis intersects its surface, allows one to rotate it vertically. To make pure torsional rotations, you have to place the finger tip on the pupil and the thumb diametrically opposite the pupil, so that the axis of rotation lies in the sagittal plane, rather than the transverse plane.

through the pupil of the eye. During this rotation, the pupil remains in the horizontal plane (because it coincides with the axis of the rotation) and a cross drawn on the pupil rotates into an x Imagine that the eye subsequently rotates 90° horizontally back to a straight-ahead gaze. In this sequence, the eye never rotated about our chosen torsional axis, but it has nonetheless undergone a 45° torsional rotation. Thus, a retinal receptive field that had a vertical orientation before we started will now be oriented 45° away from vertical. Two equal and opposite rotations about our vertical axis (90° to one side and then 90° back) with an intervening rotation about our transverse axis yield a change in eye position that could have been produced by a single rotation about the torsional axis and no rotation about the transverse axis. This is weird. It happens, because as soon as the eye rotates horizontally by any amount away from the straight ahead position, then any rotation about the transverse axis (the axis for 'vertical' rotations) becomes to some extent also a rotation about the torsional axis.

Donders law says that torsional changes in the eye's position in the socket do not occur during saccades. That is, at any moment during any saccade the eye occupies a position that could have been reached from the straight-ahead, zero-torsion primary position of the eye by a rotation about an axis that lies in the plane defined by transverse and vertical axes that are perpendicular to the direction of gaze when the eye is in the primary position. These axes define a transverse plane that divides the eye into a front half and a back half. This plane--the plane perpendicular to the primary direction of gaze--is sometimes called Listing's plane. All eye positions reachable from the primary position by a rotation about an axis in this plane have the property that the images of vertical lines with vertical meridians on the retina (great circles intersecting at the vertical axis) and the images of horizontal lines align with horizontal meridians (great circles intersecting at the transverse axis). Thus, the orientations inertial space of the lines to which simple cells in V1 are most sensitive is the same for all such positions of the eye. A receptive field that is vertical on the retina will be maximally stimulated by a vertical line, no matter what the direction of gaze nor how that direction was reached.

The eye's position is specified in a three-dimensional rotational space, all of whose dimensions are angular (degrees of rotation). Often, the positions of the eye during saccades are said to lie in Listing's plane, because Listing's plane may also be conceived of as the zero-torsion plane in rotational space[5]. The dual meaning of the term 'Listing's plane' is confusing, because the two planes referred to are not the same. In fact, they are planes in two different, non-isomorphic spaces. One plane, which I will hereafter call Listing's axes plane, is

---

[5]This is a plane in the mathematical sense not the physical sense. Mathematically, a 'plane' is a two-dimensional set of points in a three-dimensional space, that is a set of three-dimensional position vectors that do not vary along one of the three dimensions after an appopriate rotation of the framework; thus, the set of all positions in rotational space that have zero torsion constitute a plane in the mathematical sense.

defined by a set of axes of rotation in a Cartesian head-centered space, while the other, which I will called Listing's position plane, is defined by a set of eye positions in a rotational space. Listing's coordinates specify these two-dimensional eye positions (that is, these directions of 0-torsion gaze) in terms of the direction in which the gaze is imagined to depart from the primary position (the origin of the system of coordinates) and the magnitude of the rotation made in this direction.

Donders' law (sometimes called Listing's law) has sometimes been thought to be a consequence of the biomechanics of the eye, but the eye occupies positions that are not in Listing's position plane during the smooth portions of vestibular nystagmus (Crawford & Vilis, 1991), during smooth pursuit eye movements(G. Westheimer & McKee, 1973), and during sleep (Nakayama, 1975). Moreover, whenever the eye is not in its primary position at the start of a saccade, then the axis of rotation during that saccade does not lie in Listing's axes plane. When the eye deviates α degrees from the primary position at the start of the saccade, then the axis of rotation for the saccade lies in a plane that is tilted by $\alpha/2°$ away from Listing's axes plane in the direction of α (Villis & Tweed, 1991). In other words, the axes of rotation that maintain the eye in Listing's position plane only lie in Listing's axes plane if the starting point of the saccade is the primary position. For other starting positions, the axis of the rotation lies outside Listing's axes plane. Finally, and most tellingly, Listing's axes plane varies substantially within a subject over time. It varies more widely than can plausibly be attributed to variations in biomechanics (Crawford, 1994; Ferman, Collewijn, & Van den Berg, 1987; Tweed & Villis, 1990).

One often imagines that coordinate frameworks are imposed on the nervous system by our analysis--that the framework is in the mind of the theorist or in the practical necessities of the experimentalist, not in the brain of the subject (Robinson, 1992). In implementing Donders' law, however, the nervous system establishes a coordinate framework of its own, because the positions that the eye assumes during saccadic eye movements have zero torsion only if one correctly identifies the primary position of the eye, and it is the nervous system that determines what that primary position is. Recall that Listing's axes plane is the transverse plane through the center of the eye <u>orthogonal to the direction of gaze when the eye is in its primary position</u>. If the experimenter assumes a primary position of the eye that is not the one the nervous system specifies, then the experimenter's measurements of eye position during saccades will yield points (three-angle eye-position vectors) that do not have zero torsion in the experimenter's system of coordinates[6]. The positions assumed by the eye as it jumps around have zero-torsion only when the experimenter correctly identifies the primary position of the eye. In fact, this is in essence how one determines the primary direction of gaze and hence, the orientation of Listing's position plane:

---

[6]Clinicians refer to torsions that arise from choosing a coordinate system other than the zero-torsion Listing's system as "false torsions."

one rotates the coordinate-framework in which the measured eye positions are expressed until one finds the orientation that minimizes the departures from 0 along the torsional dimension (Tweed, Cardera, & Vilis, 1990). The primary position of the eye varies widely between subjects--by as much as 30°--and it also varies substantially within subjects over time--by as much as 14° (Crawford, 1994; Ferman et al., 1987; Tweed & Villis, 1990).

Crawford (1994) demonstrated that the neural integrators that maintain static eye positions establish an intrinsic coordinate system for eye positions. His experiments exploited two aspects of our extensive knowledge of the neurobiology of oculomotor control. First, the motor signals that govern the eye muscles during and after a saccade are programmed in a pulse-and-step pattern (Robinson, 1975). The pulse is an initial burst of motor neuron firing, which generates the force required to accelerate the eye to the peak velocity that it reaches during the saccade. The greater the magnitude of the saccade, the greater the peak velocity, and the stronger this initial burst of firing. (The size of the pulse also varies as a function of the position of the eye in the orbit at the beginning of the saccade.) The step is the sustained change in firing required to maintain the eye in its new position. Remarkably, the nervous system computes the step by integrating the pulse (Robinson, 1989). Thus, the change in the sustained rate of motor neuron firing is proportional to the area of the pulse. Second, the nervous system decomposes sustained changes in eye position (position steps) into horizontal and vertical-torsional components, which are realized by distinct integrators. The integrator for the horizontal component is in the nucleus prepositus hypoglossi (Cannon & Robinson, 1987; Cheron & Godaux, 1987; Straube, Kurszan, & Büttner, 1991), while the integrators for the vertical-torsional components are in the interstitial nucleus of Cajal (Buttner, Buttner-Ennever, & Henn, 1977; Crawford & Vilis, 1991, 1993; King & Moore, 1991)

It is possible to knock out an integrator and hence the step change in firing that it produces without knocking out the pulse that determines the saccadic trajectory. When the saccadic pulse remains but the step is gone, the eye jumps to its new position in response to the burst of motor neuron firing that constitutes the pulse, but then, because the step change in firing is absent, it drifts back to the origin (0-point) of the integrator's axis, following an exponential time course.

Crawford (1994) reasoned--and confirmed by simulation--that if he knocked out the vertical-torsional integrators but not the horizontal integrator, then the eye would drift back to a resting position on the horizontal axis of Listing's plane after each saccade. Following saccades with different horizontal components, the eye would come to rest at different points along the horizontal axis. Also the post-saccadic drift trajectories would parallel the vertical axis of the intrinsic coordinate system. Put another way, from the static resting positions to which the eye drifted, one could determine the horizontal axis of the integrator-imposed coordinate system. And, from the drift trajectories, one could determine the orientation of the vertical axis. In short, knocking out the vertical-torsional integrators makes the axes of the neurally imposed coordinate framework

manifest in the post-saccadic drifts in eye position. These two drift-determined axes define a plane. If Listing's zero-torsion coordinate framework is imposed by the neural integrators in the intact preparation, then the orientation of the plane determined by drift trajectories after knocking out the vertical-torsional integrators should agree with the orientation determined from the eye positions observed in the intact preparation.

Crawford made repeated experiments on the same monkeys. In each experiment, he first determined Listing's position plane from the positions assumed by the eye during normal saccades and then he temporarily knocked out the vertical-torsional integrators by injecting muscimol into the interstitial nucleus of Cajal (which is a midline structure). The axes established by the post-saccadic drifts observed after temporarily knocking out the vertical-torsional integrators defined a plane that aligned closely with the zero-torsion plane derived from pre-injection saccades. Moreover, the orientations of the planes defined in these two different ways showed strong day-to-day within subject covariation. The primary position of the eyes (the origin of Listing's zero-torsion coordinate system) changed substantially from day to day--by as much as 14.° This change was seen both in the orientation of the zero-torsion plane in the intact subject and in the orientation of the plane defined by the drift trajectories after muscimol injection.

It is important to bear in mind that the axes of the integrator coordinate system are not the axes of Listing's coordinates. The axes of Listing's coordinates are direction (angle with respect to the horizontal plane) and amplitude (angle of rotation from the primary position in the specified direction). There is no vertical axis in the coordinate framework that Listing suggested for describing torsion-free eye positions. By contrast, the axes of the integrator's coordinates are horizontal and vertical. This means that the horizontal and vertical-torsional pulses that cause saccadic eye movements must be computed in such a way as to maintain the eye in the zero-torsion positions demanded by Donders' law. To do this, the circuitry that generates these motor pulses must take into account the position of the eye at the start of the saccade. The burst of motor neuron firing that produces the same <u>change</u> in eye position (e.g., 10° to the right and 5° up) differs depending on the starting position of the eye. In other words, the innervation received by the eye muscles must specify a change from one absolute position of the eye to another, not simply a relative change in position (Nakayama, 1975; G Westheimer, 1981)

By specifying saccadic changes in eye position in a zero-torsion coordinate framework of its own devising, the oculomotor system reduces the degrees of freedom in saccadic eye movements from three to two. This neurally imposed reduction in the degrees of freedom of a joint is not unique to the eye. The orientation of the head during combined head and eye gaze shifts and the orientation of the wrist during pointing are similarly constrained (Hepp & Hepp-Reymond, 1989; Tweed & Vilis, 1990). Thus, it may be possible to discover neurally imposed coordinate systems at other joints with three degrees of rotational freedom.

The Representation of Saccades in the Superior Colliculus

Whereas the brainstem circuits that generate the motor neuron firing that moves the eye must reckon with the position of the eye in the orbit as well as the desired change in position, the superior colliculus simply specifies the desired change in the direction of gaze in a coordinate system whose origin is the current presaccadic direction of gaze. The activity of neurons in the deep layers of the superior colliculus specify the deviation between the direction of gaze and the target, the distal position whose image is to be brought onto the fovea by the saccade. Target positions computed in other coordinate frameworks--the retinal framework in the case of visual inputs, a head-centered coordinate framework in the case of auditory inputs--are mapped into this common coordinate system for the production of saccadic changes in the direction of gaze.

The Mapping of Computed Gaze Error

A fascinating aspect of this coordinate system is that there is a topographic mapping of computed gaze error onto anatomical dimensions of the superior colliculus, so that adjacent neurons (or adjacent columns of neurons) in the superior colliculus represent adjacent positions of the distal target in a gaze-centered coordinate system (D.L. Sparks & Groh, 1995). This mapping of gaze-error is unlike other familiar topographic mappings, such as the mapping of the retina onto V1 or the mapping of the cochlear membrane onto the primary auditory cortex, in that it is not a topographic mapping of a sensory surface. A stimulus falling anywhere on the retina or vibrating any point along the basilar membrane of the cochlea can activate any position in the deep collicular mapping of gaze-error. In fact, because units in this mapping of gaze-error space may be driven by either visual or auditory input, the firing of units in this mapping does not necessarily indicate which sensory epithelium (retinal or cochlear) was stimulated, let alone the position on one of those sensory surfaces excited by the proximal stimulus. In short, this is a mapping of the position of the distal stimulus, not of the position(s) of the proximal stimul(us/i).

The topographic mapping of gaze-error is delineated by the results of microstimulation at different sites in the superior colliculus and by recordings from neurons at those sites. Stimulating a site in the colliculus elicits a saccade that is a systematic function of the position stimulated. The elicited saccade shows only a weak dependence on the position of the eye in the orbit. Thus, for example, if the stimulation elicits a saccade of magnitude 5° directed up and to the right at an angle of 45° from the horizontal when the eye is initially in the primary position, then renewed or continued stimulation will elicit a second such saccade, at the end of which the eye is roughly 10° away from the primary position along the 45° direction line. A roughly but not exactly equivalent statement is that both saccades will have a horizontal and vertical components of about 3.5°.

Figure 3 shows the motor map obtained by (Robinson, 1972) in one such study. The appropriate coordinate framework for describing these results is an interesting question, with computational and neurobiological implications.

Robinson described the saccades he elicited in terms of their direction and magnitude. His map of the effects of stimulation is given in Figure 3B. As he move his stimulating electrode from the rostral margin of the left superior colliculus back to the caudal margin along a line that bisected the collicular surface, he elicited horizontal saccades of increasing magnitude. If he moved his
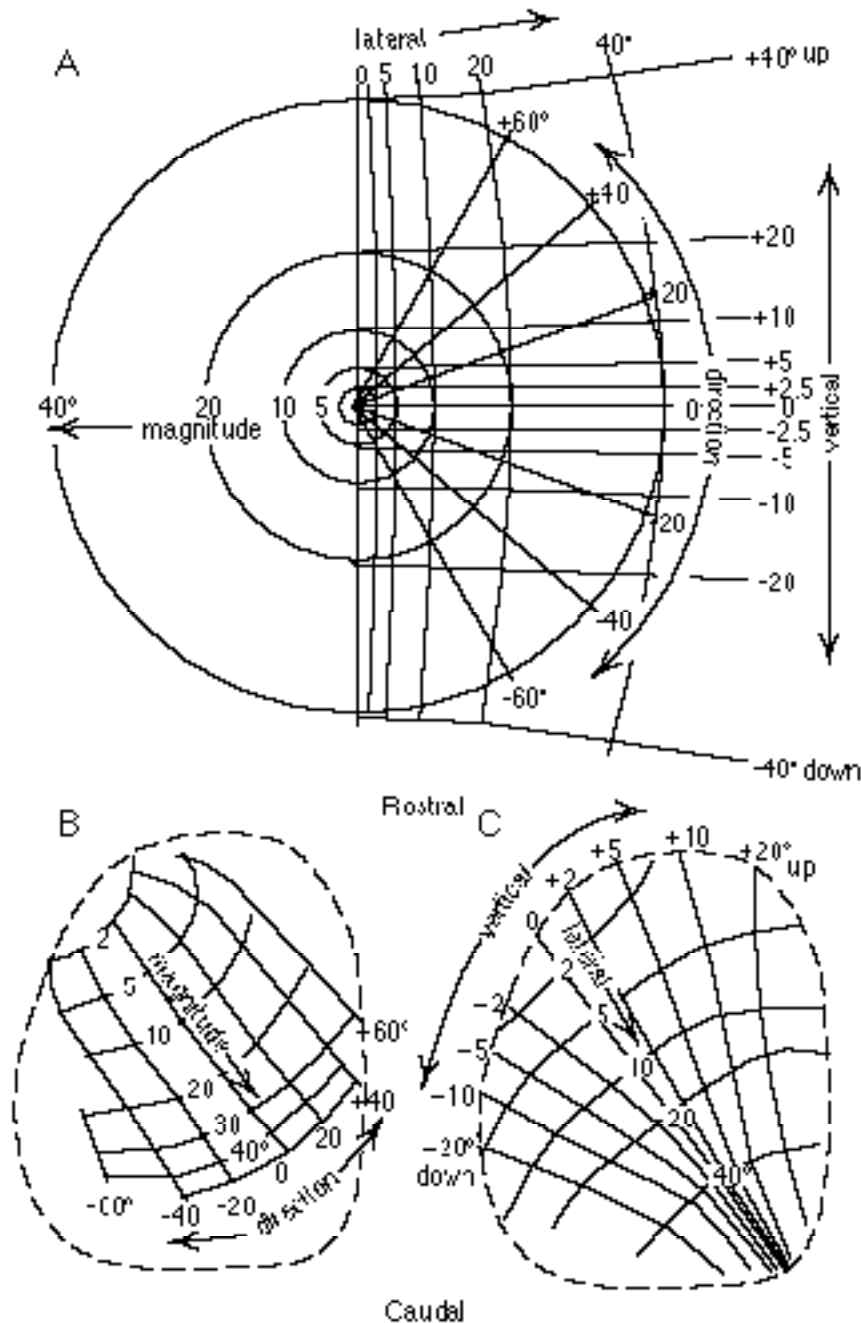
Figure 3. A. *Two coordinate frameworks for representing saccades--the magnitude-direction framework and the horizontal-vertical framework (also called lateral-vertical, or azimuthal-elevational). Formulae for converting m-d coordinates to h-v coordinates frameworks are: v = sin⁻¹[sin(d)sin(m)] and h = tan⁻¹[cos(d)tan(m). Formulae for the*

*reverse conversion are m = cos⁻¹[cos(h)cos(v)] and d = tan⁻¹[tan(v)/sin(h)]. **B**. The results of Robinson's microstimulation experiment represented with the m-d system of coordinates. (Reproduced from Robinson (1972) by permission of the author and publisher. **C.** Approximately how the results would look if they were represented in h-v coordinates.*

stimulating electrode away from this central axis toward the dorsomedial margin of the colliculus, the saccades elicited were directed upward as well as laterally. The closer the electrode was to the dorsomedial margin, the farther upward vertical component of the saccade. Thus, the direction of the elicited saccade (its deviation from horizontal) varied as a function of how far the electrode was from the horizontal axis (the line of stimulating positions that elicited purely horizontal saccades). Another way of describing the same results is to say that the horizontal components of the elicited saccades were determined by the position of the stimulating electrode along a rostro-caudal axis, while the vertical components were determined by its position along a dorsomedial to ventrolateral axis. The coordinate framework for this alternative description is given in Figure 3C.

At this same time, (Schiller & Stryker, 1972) published the results of an experiment in alert monkeys in which they immobilized one eye, so that they could determine the "receptive field" of units at a given position in the superior colliculus. They left the other eye mobile, so that they could determine the saccade elicited by stimulating through the electrode used for recording. They found that when, for example, the units at a given position were sensitive to stimuli located 10° lateral to the fovea and 10° above the horizontal, then stimulating at that site elicited a saccade with a 10° horizontal component and a 10° vertical component, that is, a saccade whose direction (45°) and magnitude (14°) were such as to foveate a stimulus at the distal position to which the units at that site were most sensitive.

The term 'receptive field' is enclosed in quotes above, because subsequent work by (Mays & Sparks, 1980) showed that visually sensitive units in the deeper layers of the superior colliculus do not have retinal receptive fields in the ordinary meaning of the term. There is no fixed area of the retina where a stimulus must fall in order to elicit firing from these deeper units, because the deep layers combine retinal position information with eye position information to map the position of visual sources in a gaze-centered coordinate system. Mays and Sparks used a double-saccade paradigm to dissociate the retinal position of the proximal stimulus from the gaze-error of the distal stimulus source at the time the foveating saccade is made.

In the double-saccade paradigm, there are two brief flashes. The second flash is delivered before the eye initiates the saccade that foveates the first flash. Both flashes are gone before the eye makes any saccade. In response to the two flashes, the monkey makes two saccades--an initial saccade that foveates the position from which the first flash originates and a second saccade that foveates the position from which the second flash originates. Suppose that one records from a site that is ordinarily excited by a stimulus 10° to the right and 10° above

the direction of gaze. Suppose one delivers a first flash that is 10° to the left and 10° below the direction of gaze and a second flash that is in the direction of gaze. Because the second flash is in the direction of gaze, it falls on the fovea, not on the putative receptive field of the units one is recording from. However, the first flash elicits a saccade 10° to the left and 10° down, at the end of which the gaze error for the distal position from which the second flash came is now 10° right and 10° up. A flash from this position would activate the recording site. However, the second flash came and went before the eye made the first saccade, and, as already noted, this second flash fell on the fovea well outside the putative receptive field of the units recorded from. Nonetheless, the units show a response to the second flash, a response that only develops after the eye makes the first saccade. That is, the units respond to the second flash only when the distal position from which that flash originated comes to occupy the position in a gaze-centered framework to which these units are sensitive. Thus, their response to a flash is jointly determined by information from the sensory epithelium indicating where the flash fell on the retina and by a position signal indicating the intervening displacement of the eye. The signal giving the eye's displacement vector and the signal giving the retinal position of the proximal stimulus (the retinal position vector) are combined is such a way as to make the units in the deep layers of the superior colliculus sensitive to the position of the distal stimulus in a coordinate framework centered on the current direction of gaze. The combinatorial operation is equivalent to vector subtraction.

Units in these same deep layers are also sensitive to the distal position (azimuth and elevation) of an auditory stimulus source. This position is computed in head-centered coordinates by a sequence of operations that use small differences in the intensity and time of arrival of sound waves at the two ears. The sequence of computations culminates in a head-centered topographic map of auditory space in the external nucleus of the <u>inferior</u> colliculus (Konishi, 1995).[7] This map is then relayed to the superior colliculus, where it is transformed into a gaze-centered map of auditory space. Because the mapping of auditory space in the superior colliculus is gaze-centered rather than head centered, a coordinate transformation is required. An eye displacement vector must be subtracted from the head-centered auditory position vector to yield a gaze-centered auditory position vector. A consequence of this coordinate transform is that the position in the deep layers of the superior colliculus excited by a sound source at a given head-centered position changes as the eyes change position in the orbit (Jay & Sparks, 1987). Equivalently, the position in head centered coordinates of the <u>auditory</u> stimulus source to which a unit in the deep layers of the superior colliculus is most sensitive changes as the <u>eye</u> changes its position in the orbit.

The transformations demonstrated neurobiologically by the work of Sparks and his collaborators are the sorts of transformations that we inferred

---

[7]At least in the barn owl, where work on the neurobiology of auditory scene analysis is most extensive. An equivalent analysis has yet to be done in a primate.

from the ability of the spinal frog to scratch its elbow with its hind limb regardless of the relative position of the two limbs (Figure 1). In that case, the sensory epithelium was the surface of the forelimb rather than the retina. The displacements that had to be combined with this sensory signal to determine the direction in which the hind limb had to move were the changes in the angular positions of the two limbs. The great importance of the work on the superior colliculus--and similar work in other areas(Bruce, 1990; Bruce & Goldberg, 1990; Gnadt & Andersen, 1988; Russo & Bruce, 1994) --is that it demonstrates that these sorts of coordinate transformations are explicitly carried out in the operation of the nervous system. The nervous system carries out the inferred coordinate transformations and creates electrophysiologically observable topographic mappings of the resulting spaces. It does not 'act as if' it carried out such transformations; it really does them.

The Position or Displacement Signal Comes from Corollary Discharge

The eye position or eye displacement vector that is used in computing the current gaze error does not come from sensory receptors, because animals in which this sensory input has been eliminated nonetheless compensate for changes in eye position (Guthrie, Porter, & Sparks, 1983). This finding has implications similar to the finding that the reaching system can adapt to aritificial dynamic perturbations with feed-forward corrections, forces that are programmed in anticipation of forces that will arise during a movement. The adaptation findings imply that the system has an internal model of the dynamics. The finding that the eye position or displacement vector is based on corollary discharge or efference copy implies that the system has an internal model of the relation between these command discharges and the changes in position that they may be expected to produce. Recall also that the enduring change in the discharge of the motor neurons that holds the eye in its new position after a saccade is derived by integrating the pulse--the signal that specifies the velocity and duration of the saccade. Thus, the displacement signal in these coordinate transformations derives from integrating a velocity signal with respect to time.

Computational Implications of Choosing Different Coordinates

The units in the deep layers of the superior colliculus, like units in the many other topographic mappings that have been discovered in sensory/perceptual areas, are broadly tuned. Each unit has a best direction, a gaze error for which it fires most vigorously, but it also fires to some extent prior to saccades in a cluster of directions centered around its best direction. Thus, prior to any saccade, there is a substantial population of active neurons. Each active neuron may be thought of from a computational perspective as a vector pointing at a particular gaze error, the gaze error represented by its position in the topographic mapping of gaze-error space. Two questions arise; (1) What is the combinatorial operation that determines the resulting saccade from the activity of a population of adjacent active neurons each of which codes for a slightly different gaze error? (2) What are the computationally meaningful coordinates for describing the position of a unit in gaze-error space?

Two simple kinds of combinatorial decision processes might decide the direction: 1) Winner take all--the saccade made is determined by the coordinates of the most active unit. 2) A vector combining operation in which the gaze error that determines the saccade is derived from the coordinates of all the active units, weighting the contribution of each according to how active it is. (Lee, Rohrer, & Sparks, 1988) and (D.L Sparks, Lee, & Rohrer, 1990) report results that point toward the second alternative. In a winner take all scheme, inactivating with lidocaine the center of the active population--that is, temporarily knocking out the units that are firing most vigorously--should alter the direction and magnitude of the resulting saccade, because it should knock out the "winner," shifting the site of greatest activity (hence, the location of the winner) away from the center of the population. Conversely, inactivating units on the margin of the active population--units that are less active than units at the center--should have no effect on the direction and magnitude of the resulting saccade, because the activity level of the winning position is unaffected by such an injection. However, Lee, et al. (1988) showed that inactivating the center of the population did not alter the direction and magnitude of the saccade, although it did greatly reduce its velocity (see also (D.L Sparks et al., 1990)). They further showed that inactivating units on the margin of the active population moved the resulting saccade away from the saccades coded for by the area that was inactivated. Thus, if the center of the active population coded for purely lateral (0-direction) saccades of magnitude 10°, and one inactivated cells coding for saccades lying above and to the left of that in gaze-error space, then the resulting saccade was down and to the right of the saccade that one would have observed in the absence of the inactivation. This shows that the activity of units on the margin of the active population helps to determine the saccade, which implies that the gaze-errors indicated by all the active units are summed or averaged to determine the saccade.

But which is it--summing or averaging? In describing the Lee, et al. (1988) experiment, I mixed two distinct frames of reference that may be used to describe gaze errors and the resulting saccades. I referred to the center of the active population in terms of the direction (0°) and magnitude (10°) of the indicated gaze-error. Direction and magnitude are the coordinates in Figure 3B. However, in the same sentence, I referred to the inactivated cells as lying above and to the left of the gaze error indicated by the center of the active population. The 'lateral' (to the left, to the right) and 'vertical' (above, below) dimensions are the coordinates used in Figure 3C. This highlights the fact that it is not obvious simply from the data what kinds of vectors we should use to represent a unit's contribution to the determination of the saccade vector.

From the standpoint of the computations needed to specify a saccade, it matters which framework we use. If we use a two-dimensional direction-and-magnitude framework, then we have to assume that the combinatorial operation is equivalent to vector averaging. Moreover, we have to imagine a conversion somewhere else from the direction-and-magnitude specification of the saccade vector to a specification in terms of the horizontal and vertical components of the saccade, because we know that there are separate brainstem nuclei for the horizontal and vertical components. Finally, we have to use a different

combinatorial operation to specify saccade velocity from the overall level of activity in the population. On the other hand, if we think of active units as equivalent to three-dimensional vectors, then we have to assume only simple vector summation as the combinatorial operation that specifies the saccade and this same operation gives the velocity of the saccade. Moreover, if we assume that the positions of units in the superior colliculus specify the horizontal and vertical components of saccades (Figure 3C), rather than their direction and magnitude (Figure 3B), then this same operation gives what the next stage needs, namely, the horizontal and vertical components of the required saccade.

If we follow the conventional practice of thinking of vectors as arrows, then in a two-dimensional, direction-and-magnitude framework, an active unit is at the point of an arrow whose tail is at the origin of the coordinate system--the upper left corner of Figure 3B, the point toward which the direction lines are converging. (This point corresponds to the direction of gaze.) The activity of the unit is not a dimension of the vector in this scheme; it serves merely to determine the relative weight that will be given to that vector when the error vectors are averaged. Thinking of the active units as arrows lying on the collicular surface makes it obvious why we cannot combine the vectors for active units by simple summation. Vector summation is equivalent to moving one of the arrows so that its tail coincides with the point of the other arrow. The position of the point of the moved arrow is then the vector sum. Thus, summing two adjacent units, both specifying a saccade of 0° direction and magnitude 10° would yield a gaze error of 0° direction and a magnitude of roughly 30° (as may be verified by carrying out this operation on Figure 3B). That is, the vectors resulting from simple summation are much too long; they specify impossibly big saccades. To make the results of the combinatorial operation plausible, the sum of the weighted error vectors must be normalized by the sum of the weights, that is, the nervous system must compute not the vector sum but the vector average. In computing this average, it must first scale each vector by its firing rate, then sum the scaled vectors, then divide that sum by the sum of all the firing rates.

However, we need not think of the active units as corresponding to two-dimensional vectors. It takes three numbers to describe an active unit--two to specify its position and one to specify its firing rate. This encourages us to think of active units as three-dimensional vectors. To visualize these vectors as arrows, we may think of the surface in Figure 3C as wrapped on a globe. The origin of the coordinate system, where the vector tails are pinned, is not on the surface shown in Figure 3C, it is at the center of the globe. The three-dimensional vector corresponding to an active unit does not lie on the surface of Figure 3C; rather it penetrates that surface. Where it penetrates the surface gives a position in gaze-error space. The third dimension of these vectors--their length--is the firing rate of the unit. The greater the length of one of these vectors, the farther it projects beyond the gaze-error surface. This length may be thought of as the vigor with which an active unit urges the eye to make the move indicated by the unit's position in gaze-error space. This vigor translates into the velocity of the saccade; the more vigorously the unit fires, the faster the saccade it calls for. If these are the neurobiologically meaningful coordinates for describing positions in the superior colliculus's mapping of gaze-error space, then the specification of the horizontal and vertical components of the saccade and its velocity require only

the equivalent of vector summation--the summation of the vectors corresponding to all the units in the active population. The point where the resultant vector penetrates the surface gives the lateral and vertical components of the saccade. The length of the resultant vector specifies its velocity.

The vector-summation model, which replaces vector averaging with vector summation by treating the vectors as three dimensional rather than two dimensional, gives a computationally simpler explanation of the results of the lidocaine experiment.[8] Lidocaine injected into the center of the population will

---

[8]One of the schemes proposed by

(Sparks, & Mays, 1990)   for extracting the horizontal component of a saccade from the activity of a population of units in the superior colliculus implicitly assumes that the horizontal-vertical framework is the appropriate framework (see Figure 1 in
Sparks & Mays, 1990).

not alter the position at which the resultant vector penetrates the surface, because the direction of resultant (where it penetrates the surface) will not be altered by knocking out arrows distributed symmetrically around the resultant. However, by removing the longest contributors to the resultant (the most active units), the injection will reduce the length of the resultant. And, the length of the resultant specifies velocity. Thus, knocking out vectors in the center of the active population should reduce the velocity of the saccade without altering its direction and magnitude. Lidocaine injected to one side of the population center, will shift the point at which the resultant vector penetrates the surface away from the site of injection by removing the contribution of the vectors pointing to the positions where the lidocaine was injected. If the lidocaine is injected on the caudal (more lateral) margin of the population, it will shift the direction of the resultant medially towards gaze errors of smaller magnitude, resulting in hypometric saccades, saccades that are shorter than they should be. If it is injected on the rostral (more medial) margin, it will shift the resultant laterally toward gaze errors of greater magnitude, producing hypermetric saccades, saccades that are bigger than they should be. These are the results that Lee, et al. (1988) in fact obtained.

Concluding Remarks

Behavioral and neurobiological data support the conclusion that the positions of targets, the trajectories of target-directed movements, and the trajectories of forces required to drive those movements are represented in the nervous system in a sequence of different, explicitly computed coordinate frameworks. Thus, an understanding of the circuitry and cellular mechanisms by which coordinate transforms of various kinds may be computed is a fundamental goal of integrative neuroscience. The behavioral analysis of the coordinate transformations that mediate directed actions has a major role to play in the description and elucidation of these mechanisms, because the behavioral analysis tells us what transformations to look for.

In many cases, the computed spaces are realized through the topographic arrangement of units tuned to different positions in the space. Units at a given location in these topographic mappings of abstract spaces become excited when the target or trajectory occupies a small region of that space, the region to which the unit is tuned. The center of this region, the position that corresponds to the maximum firing, is the best position of unit. The mapping of positions in the space to which neurons are tuned is said to be topographic when units that are adjacent neuroanatomically are tuned to adjacent best positions.[9] Adjacent here

---

[9]The term 'topographic,' which has become standard in the neurobiological literature, may have arisen from a confounding of 'topographic' with 'topological'. In mathematics, a topological mapping is one that preserves adjacency but not distance and angle. A topographic map, on the other hand, is a map of (some part of) the earth's surface that encodes elevations as well as horizontal

has a purely mathematical meaning, because some of the spaces that may be topographically mapped may not have a simple physical interpretation. An example of a space without a simple physical interpretation of adjacency would be a trajectory space. "Points" in such a space specify trajectories formed by the superposition of basis functions. The different basis functions constitute the axes of such a space. The dimensionality of the space is equal to the number of basis functions from which trajectories are synthesized. The values along these dimensions give the scaling values for the basis function--how much of this basis function to use when adding up the various contributions to determine the resultant function. Points in such a space are adjacent if the scaling values for corresponding basis functions are all adjacent, that is, if the value of each dimension in one vector is adjacent to the value for the same dimension in the other vector. As psychologists become more sophisticated about coordinate transforms and basis functions, these sorts of highly abstract 'spaces' are likely to play a more prominent role in our analysis of directed actions.

Directed action has been treated here in a circumscribed and literal sense-- to refer to actions in which the optical axis of the eye is to be directed to a given target or a hand or foot is to be moved to a given target. Directed action in a more general sense--action that accomplishes some goal--requires for its analysis a much broader range of principles and mechanisms. Some of those mechanisms- -the different kinds of elementary units of behavior, how they are coordinated into complex units of behavior, the hierarchical structure of the resulting complex units, and the mechanisms for coordination within that hierarchy--have been described and illustrated elsewhere (Gallistel, 1980, 1994).

However, the coordinate transformations that mediate simple directed actions may prove directly relevant to another aspect of directed action-- navigation, moving the whole animal to a target.. Navigation depends on the construction and use of a cognitive map. Both the construction of a map and its use require coordinate transformations strikingly similar to those discussed here (Gallistel, 1990, 1998). To construct a cognitive map, the brain combines a position vector in a body-centered coordinate framework--the position of a terrain feature relative to the animal--with a displacement vector (the animal's displacement in the geocentric coordinate framework). Similarly, the superior colliculus combines the position of a stimulus in a retinal or head-centered coordinate framework with a vector indicating the displacement of that coordinate framework relative to a larger framework. In the construction of a cognitive map, the displacement vector arises from dead reckoning--the integration of the animal's velocity with respect to time to yield the body's displacement vector. Similarly, the signals that specify the eye's displacement vector appear to arise from the integration of eye velocity signals with respect to time.

When the brain adds a vector that specifies a terrain feature in a body-centered coordinate system to the body-displacement vector from the animal's

---

locations (latitudes and longitudes). In neurobiology, however, 'topographic mapping' means a mapping that preserves adjacency.

dead-reckoning mechanism, it maps the position of the terrain feature into a geocentric coordinate framework. By routinely doing this coordinate transformation, the animal builds up a representation of the different terrain features in its environment in a common geocentric coordinate system, even though those features may have been viewed at widely differing times and places.

Using the map involves the inverse transformation, mapping from the position of a goal on the geocentric map to the position of the goal in a body-centered framework. It is this inverse transformation that enables the animal to orient toward goals it cannot currently perceive by reference to its cognitive map and its own position and orientation on that map. Thus, the study of coordinate transformations and their neurobiological implementation is of broad significance for our understanding of the computational processes that underlie higher cognitive function.

References

Atkeson, C. G. (1989). Learning arm kinematics and dynamics. <u>Annual Review of Neuroscience</u>, <u>12</u>, 157-184.

Atkeson, C. G., & Hollerbach, J. M. (1985). Kinematic features of unrestrained vertical movements. <u>Journal of Neuroscience</u> <u>5</u>, 2318-2330.

Berkinblit, M. V., Gelfand, I. M., & Feldman, A. G. (1986). Model of the control of movements of a multijoint limb. <u>Biophysics</u>, <u>31</u>, 142-153.

Bernstein. (1967). <u>The coordination and regulation of movements</u>. Oxford: Pergamon Press.

Bizzi, E., & Mussa-Ivaldi, A. (1995). Toward a neurobiology of coordinate transformations. In M. S. Gazzaniga (Ed.), <u>The cognitive neurosciences</u> (pp. 495-506). Cambridge, MA: MIT Press.

Bizzi, E., & Mussa-Ivaldi, F. A. (1990). Muscle properties and the control of arm movement. In D. O. Osherson, S. M. Kosslyn, & J. M. Hollerbach (Ed.), <u>An invitation to cognitive science. vol 2. Visual cognition and action</u> (pp. 213-242). Cambridge, MA: MIT Press.

Bizzi, E., Mussa-Ivaldi, F. A., & Giszter, S. (1991). Computations underlying the execution of movement: A biological perspective. <u>Science</u>, <u>253</u>, 287-291.

Bruce, C. J. (1990). Integration of sensory and motor signals for saccadic eye movements in the primate frontal eye fields. In <u>Signal and sense: Local and global order in perceptual maps</u> (pp. 261-314). New York: Wiley-Liss.

Bruce, C. J., & Goldberg, M. E. (1990). Primate frontal eye fields: III. Maintenance of a spatially accurate saccade signal. <u>Journal of Neurophysiology</u>, <u>64</u>, 489-508.

Buttner, U., Buttner-Ennever, J. A., & Henn, V. (1977). Vertical eye movement related unit activity in the rostral mesencephalic reticular formation of the alert monkey. <u>Brain Research</u>, <u>130</u>, 239-252.

Cannon, S. C., & Robinson, D. A. (1987). Loss of the neural integrator of the oculomotor system from brainstem lesions in the monkey. <u>Journal of Neurophysiology</u>, <u>57</u>, 1383-1409.

Cheron, G., & Godaux, E. (1987). Disabling of the oculomotor integrator by kainic acid injections in the prepositus-vestibular complex of the cat. <u>Journal of Physiology (London)</u>, <u>394</u>, 267-290.

Crawford, J. D. (1994). The oculomotor neural integrator uses a behavior-related coordinate system. <u>Journal of Neuroscience</u>, <u>14</u>(11, part 2), 6911-6923.

Crawford, J. D., & Vilis, T. (1991). Axes of eye rotation and Listing's law during rotations of the head. Journal of Neurophysiology, 65, 407-423.

Crawford, J. D., & Vilis, T. (1993). Modularity and parallel processing in the oculomotor integrator. Experimental Brain Research, 96, 443-456.

Feldman, A. G. (1974). Change of muscle length due to shift of the equilibrium point of the muscle- load system. Biofizika, 19, 534-538.

Ferman, L., Collewijn, H., & Van den Berg, A. V. (1987). A direct test of Listing's law. I. Human ocular torsion measured in static tertiary positions. Vision Research, 27, 929-938.

Flash, T., & Hogan, N. (1985a). The coordination of arm movements: An experimentally confirmed mathematical model. Journal of Neuroscience, 5, 1688-1703.

Fukson, O. I., Berkinblit, M. B., & Feldman, A. G. (1980). The spinal frog takes into account the scheme of its body during the wiping reflex. Science, 209, 1261-1263.

Gallistel, C. R. (1980). The organization of action: a new synthesis. Hillsdale, NJ: Lawrence Erlbaum.

Gallistel, C. R. (1990). The organization of learning. Cambridge, MA: Bradford Books/MIT Press.

Gallistel, C. R. (1994). Elementary and complex units of behavior. In G. d'Ydewalle, P.Celen, & P. Bertelson (Ed.), Current advances in psychological science: An international perspective (pp. 157-175). Hillsdale, NJ: L. Earlbaum Assoc.

Gallistel, C. R. (1998). Brains as Symbol Processors: The Case of Insect Navigation. In S. Sternberg & D. Scarborough (Vol Ed.), Conceptual and methodological foundations. vol 4 of An invitation to cognitive science. 2nd ed (D. Osherson, Series Ed.). Cambridge, MA: MIT Press. pp 1-51

Giszter, S. F., Mussa-Ivaldi, F. A., & Bizzi, E. (1993). Convergent force fields organized in the frog's spinal cord. Journal of Neuroscience, 13(2), 467-491.

Gnadt, J. W., & Andersen, R. A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. Experimental Brain Research, 70, 216-220.

Gordon, J., Ghilardi, M. F., & Ghez, C. (1994). Accuracy of planar reaching movements: I. Independence of direction and extent variability. Experimental Brain Research, ???(???), ???

Guthrie, B. L., Porter, J. D., & Sparks, D. L. (1983). Corollary discharge provides accurate eye position information to the oculomotor system. Science, 221, 1193-1195.

Hasan, Z., & Stuart, D. G. (1988). Animal solutions to problems of movement control: The role of proprioceptors. Annual Review of Neuroscience, 11, 199-223.

Hasan, Z. (1991). Biomechanics and the study of multijoint movements. In D. R. Humphrey, & H.-J. Freund (Ed.), Motor control: Concepts and issues (pp. 75-84). New York: Wiley.

Hepp, K., & Hepp-Reymond, M.-C. (1989). Donders' and Listing's Law for reaching and grasping arm synergies. Abstracts 15 Annual Meeting of the Society for Neuroscience. 15, p. 604

Hogan, N., & Flash, T. (1987). Moving gracefully: Quantitative theories of motor coordination. Trends in Neurosciences, 10, 170-174.

Hogan, N. E., Bizzi, E., & Mussa-Ivaldi, F. A. (1987). Controlling multijoint motor behavior. Exercise and sport Sciences Review, 15, 153-190.

Hollerbach, J. M. (1982). Computers, brains, and the control of movement. Trends in Neurosciences, 5, 189-192.

Hollerbach, J. M., & Atkeson, C. G. (1987). Deducing planning variables from experimental arm trajectories: Pitfalls and possibilities. Biological Cybernetics, 56, 279-292.

Jay, M. F., & Sparks, D. L. (1987). Sensorimotor integration in the primate superior colliculus. II. Coordinates of auditory signals. Journal of Neurophysiology, 57, 35-55.

Jordan, M. I. (1994a). Computational aspects of motor control and motor learning. In H. Heuer, & S. Keele (Ed.), Handbook of perception and action: motor skills New York: Academic.

Jordan, M. I. (1994b). Computational motor control. In M. S. Gazzaniga (Ed.), The cognitive neurosciences (pp. 597-609). Cambridge, MA: MIT Press.

Kalaska, J. F., & Crammond, D. J. (1992). Cerebral cortical mechanisms of reaching movements. Science, 255, 1517-1523.

King, A. J., & Moore, D. R. (1991). Plasticity of auditory maps in the brain. Trends in Neurosciences, 14(1), 31-36.

Konishi, M. (1995). Neural mechanisms of auditory image formation. In M. S. Gazzaniga (Ed.), The cognitive neurosciences (pp. 269-277). Cambridge, MA: MIT Press.

Lacquaniti, F., Soechting, J. F., & Terzuolo, C. A. (1982). Some factors pertinent to the organization and control of arm movement. Brain Research, 252, 394-397.

Lacquaniti, F., Soechting, J. F., & Terzuolo, C. A. (1986). Path constraints on point-to-point arm movements in three-dimensional space. Neuroscience, 17, 313-324.

Lee, C., Rohrer, W. H., & Sparks, D. L. (1988). Population coding of saccadic eye movements by neurons in the superior colliculus. Nature, 332, 357-360.

Lehky, S. R., Sejnowski, T. J., & Desimone, R. (1992). Predicting responses of nonlinear neurons in monkey striate cortex to complex patterns. Journal of Neuroscience, 12(9), 3568-3581.

Lockery, S. R., & Sejnowski, T. J. (1993). The computational leech. Trends in Neurosciences, 16(7), 283-290.

Loeb, G. E. (1987). Hard lessons in motor control from the mammalian spinal cord. Trends in Neurosciences, 10, 108-113.

Mays, L. E., & Sparks, D. L. (1980). Dissociation of visual and saccade-related responses in superior colliculus neurons. Journal of Neurophysiology, 43, 207-232.

Milner, T. E., & Cloutier, C. (1993). Compensation for mechanically unstable loading in voluntary wrist movements. Experimental Brain Research, 94, 522-532.

Morasso, P. (1981). Spatial control of arm movements. Experimental Brain Research, 42, 223-237.

Mussa-Ivaldi, F. A., Giszter, S. F., & Bizzi, E. (1994). Linear combinations of primitives in vertebrate motor control. Proceedings of the National Academy of Sciences of the United States of America, 91(16), 7534-7538.

Nakayama, K. (1975). Coordination of extra-ocular muscles. In G. Lennerstrand, & P. Bach-y-Rita (Ed.), Basic mechanisms of ocular motility and their clinical implications (pp. 193-208). Oxford: Pergamon.

Robinson, D. A. (1972). Eye movements evoked by collicular stimulation in the alert monkey. Vision Research, 12, 1795-1808.

Robinson, D. A. (1975). Oculomotor control signals. In P. Bach-y-Rita, & G. Lennerstrand (Ed.), Basic mechanisms of ocular motility and their clinical implications (pp. 337-374). Oxford: Pergamon.

Robinson, D. A. (1985). The coordinates of neurons in the vestibulo-ocular reflex. In A. Berthoz, & M. Jones (Ed.), Adaptive mechanisms in gaze control. Facts and theories (pp. 297-311). New York: Elsevier.

Robinson, D. A. (1989). Integrating with neurons. Annual Review of Neuroscience, 12, 33-46.

Robinson, D. A. (1992). Implications of neural networks for how we think about brain function. Behavior and Brain Science, 15, 644-655.

Russo, G. S., & Bruce, C. J. (1994). Frontal eye field activity preceding aurally guided saccades. Journal of Neurophysiology, 71, 1250-1253.

Saltzman, E. (1979). Levels of sensorimotor representation. Journal of Mathematical Psychology, 20, 91-163.

Schiller, P. H., & Stryker, M. (1972). Single-unit recording and stimulation in superior colliculus of the alert rhesus monkey. Journal of Neurophysiology, 35, 915-924.

Shadmehr, R., & Mussa-Ivaldi. (1994). Adaptive representation of dynamics during learning of a motor task. Journal of Neuroscience, 14(5, Part 2), 3208-3224.

Smolensky, P. (1988). On the proper treatment of connectionism. Behavioral & Brain Sciences, 11, 1-74.

Soechting, J. F., & Lacquaniti, F. (1981). Invariant characteristics of a pointing movement in man. Journal of Neuroscience, 1, 710-720.

Sparks, D., & Mays, L. E. (1990). Signal transformations required for the generation of saccadic eye movements. Annual Review of Neuroscience, 13, ???-???

Sparks, D. L., & Groh, J. F. (1995). The superior colliculus: A window for viewing issues in integrative neuroscience. In M. S. Gazzaniga (Ed.), The cognitive neurosciences (pp. 565-584). Cambridge, MA: MIT Press.

Sparks, D. L., Lee, C., & Rohrer, W. H. (1990). Population coding of the direction, amplitude, and velocity of saccadic eye movements by neurons in the superior colliculus. Cold Spring Harbor Symposia on Quantitative Biology, 55, 805-811.

Strang, G. (1994). Wavelets. American Scientist, 82, 250-255.

Straube, A., Kurszan, R., & Büttner, U. (1991). Differential effects of bicuculine and muscimol microinjections into the vestibular nuclei on simian eye movements. Experimental Brain Research, 86, 347-358.

Tweed, D., Cardera, W., & Vilis, T. (1990). Computing three dimensional eye position quaternions and eye velocity from search coil signals. Vision Research, 30, 97-110.

Tweed, D., & Vilis, T. (1990). Listing's law for gaze-directing head movements. In A. Berthoz, W. Graf, & P. P. Vidal (Ed.), The head-neck sensory-motor system Chichester: Wiley.

Tweed, D., & Villis, T. (1990). Geometric relations of eye position and velocity vectors during saccades. Vision Research, 30, 111-127.

Uno, Y., Kawato, M., & Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement. Biological Cybernetics, 61, 89-101.

van Emmerik, R. E. A. (1991). Kinematic adaptations to perturbations as a function of practice in rhythmic drawing movements. Journal of Motor Behavior, 24, 117-131.

Villis, T., & Tweed, D. (1991). What can rotational mechanics tell us about the neural control of eye movements? In D. R. Humphrey, & H.-J. Freund (Ed.), Motor control: Concepts and issues (pp. 85-99). New York: Wiley.

Westheimer, G. (1981). Donders', Listing's, and Hering's laws and their implications. In B. L. Zuber (Ed.), Models of oculomotor behavior and control (pp. 149-160). Boca Raton, Fla: CRC Press.

Westheimer, G., & McKee, S. P. (1973). Failure of Donders' law during smooth pursuit eye movement. Vision Research, 13, 2145-2153.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1994). Perceptual distortion contributes to the curvature of human reaching movements. Experimental Brain Research, 98, 153-156.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). Are arm trajectories planned in kinematic or dynamic coordinates? An adaptation study. Experimental Brain Research, 103, 460-470

Zajac, F. E., & Gordon, M. E. (1989). Determining muscle's force and action in mutli-articular movement. Exercise and Sport Science Reviews, 17, 187-230.