Smoothness Maximization Along a Predefined Path Accurately Predicts the Speed Profiles of Complex Arm Movements

EMANUEL TODOROV AND MICHAEL I. JORDAN
Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Todorov, Emanuel and Michael I. Jordan. Smoothness maximization along a predefined path accurately predicts the speed profiles of complex arm movements. *J. Neurophysiol.* 80: 696–714, 1998. The speed profiles of arm movements display a number of regularities, including bell-shaped speed profiles in straight reaching movements and an inverse relationship between speed and curvature in extemporaneous drawing movements (described as a 2/3 power law). Here we propose a new model that simultaneously accounts for both regularities by replacing the 2/3 power law with a smoothness constraint. For a given path of the hand in space, our model assumes that the speed profile will be the one that minimizes the third derivative of position (or ''jerk''). Analysis of the mathematical relationship between this smoothness constraint and the 2/3 power law revealed that in both two and three dimensions, the power law is equivalent to setting the jerk along the normal to the path to zero; it generates speed predictions that are similar, but clearly distinguishable from the predictions of our model. We have assessed the accuracy of the model on a number of motor tasks in two and three dimensions, involving discrete movements along arbitrary paths, traced with different limb segments. The new model provides a very close fit to the observed speed profiles in all cases. Its performance is uniformly better compared with all existing versions of the 2/3 power law, suggesting that the correlation between speed and curvature may be a consequence of an underlying motor strategy to produce smooth movements. Our results indicate that the relationship between the path and the speed profile of a complex arm movement is stronger than previously thought, especially within a single trial. The accuracy of the model was quite uniform over movements of different shape, size, and average speed. We did not find evidence for segmentation, yet prediction error increased with movement duration, suggesting a continuous fluctuation of the ''tempo'' of discrete movements. The implications of these findings for motor planning and on-line control are discussed.

INTRODUCTION

The majority of human arm movements produced in everyday life are underconstrained: the desired effect on the external environment can be achieved in a large number of ways, thus the exact movement executed at a particular time is not completely determined by the motor task. We effortlessly plan and execute underconstrained arm movements that achieve the desired effect and are at the same time smooth and graceful. Furthermore, we can produce appropriate movements in tasks that involve almost arbitrary combinations of constraints: spatial locations that the hand should pass through, objects that have to be avoided, forces that have to be exerted with very accurate timing, requirements on the speed at certain points along the path, desired durations of the entire movement or smaller segments of it,

etc. This impressive ability to take into account a diverse set of task constraints and rapidly ''fill in'' the missing details of the movement suggests that the motor system has very general and efficient mechanisms for solving ill-posed problems. Consequently, a fruitful line of research in the field of motor control has been to identify, mostly through observational studies, regularities in biological movement that are not in any way implied by the task. Once identified, such regularities can be used to infer the structure of the underlying motor system that produces them. Examples of this approach include: Listings law, Fitts law, the isochrony principle (Viviani and Schneider 1991), linearly related joint velocities (Soechting and Terzuolo, 1986), piece-wise planar three-dimensional (3D) hand paths (Soechting and Terzuolo 1987), straight paths and bell-shaped speed profiles of reaching movements (Morasso 1981), and the inverse relationship between instantaneous tangential speed and curvature of the hand path (Jack 1895; Lacquaniti et al. 1983).

In this work, we focus on the relationships observed between the hand paths and speed profiles of arm movements that are spatially constrained. We begin with a brief summary of existing versions of the minimum-jerk model (Flash and Hogan 1985) and the 2/3 power law (Lacquaniti et al. 1983) and compare their implications for the biological mechanisms underlying trajectory formation. We then develop a new model that essentially combines the appealing features of these prior models; it simultaneously accounts for both the bell-shaped speed profiles of straight reaching movements and the inverse relationship between curvature and speed of curved hand movements. In a series of experiments, we demonstrate that the new model provides an accurate description of a number of discrete arm movements with arbitrary hand paths in both two and three dimensions. Finally, we discuss the implications of our observations for the generation of complex hand trajectories.

MODELS OF TRAJECTORY FORMATION

In this section, we distinguish three groups of models (Fig. 1) based on the levels of biological processing they parallel. Our classification emphasizes the implications of different models for motor control rather than mathematical similarities among them.

Complete models of trajectory formation

The most ambitious group of models predicts the outcome of the processing in the entire CNS in a restricted set of tasks. Such models provide a complete ''recipe'' for trajectory
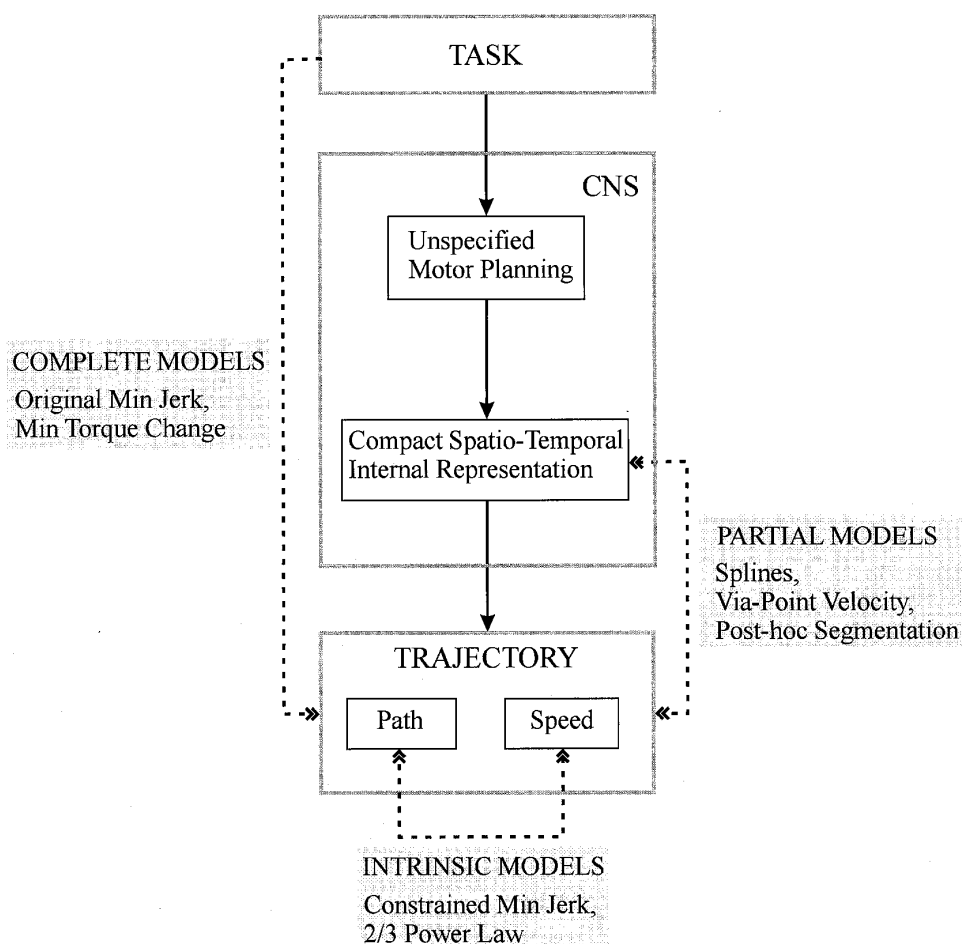
FIG. 1. Schematic representation of our model classification scheme. Grey boxes correspond to groups of models, – – –, predictions these models make. ———, the actual flow of information in the CNS (the internals of the "CNS" box are of course hypothetical). Observed trajectory can be further separated into a priori independent features, in this case speed and path.

generation, starting with the externally specified task description and leading to an observable trajectory. It would be ideal if one could construct a model of this kind that agrees closely with a wide range of experimental data; unfortunately, this has proven to be difficult.

A general method for constructing such models is provided by optimization theory: one defines a "cost" functional and attempts to show that the movements typically produced in a given task are the ones that minimize cost, subject to the constraints imposed by the task. A number of models based on optimization theory have been considered (Flash and Hogan 1985; Nelson 1983; Uno et al. 1989; Wann et al. 1988), the cost being energy, force, impulse, peak speed, peak acceleration, duration, or torque change. The model that seems to account for the largest body of data, while at the same time being attractively simple, is the "minimum-jerk" model (Flash and Hogan 1985). The cost is defined as the squared jerk (3rd derivative of hand position over time), integrated over the entire movement; it is minimized over all trajectories that have the same initial point, final point, and possibly pass through a small set of intermediate (via) points specified in the task itself. In reaching tasks, the minimum-jerk model predicts straight paths and bell-shaped speed profiles that match experimental observations rather well (Flash and Hogan 1985). There are cases, however, where subjects systematically produce nonstraight reaching movements (Atkeson and Holerbach 1985). Be-

cause this model is purely kinematic, it predicts invariance with respect to translation, rotation, and spatial or temporal scaling. It has been shown that planar rotation of the desired movement can significantly affect the hand path in via-point (Uno et al. 1989), as well as obstacle avoidance (Sabes and Jordan 1997) tasks. We show more examples of this phenomenon later.

Models of intrinsic regularities

On the other extreme, we have models that do not explicitly assume anything about the motor task or the flow of information in the CNS. They simply quantify intrinsic relationships between features of the observed trajectories that are not related a priori (in this case, between the path and speed profile). Note that a causal relationship between the two features is not necessarily implied: such models are consistent with any proposed mechanism of trajectory formation that produces the observed relationship either through an explicit computation or as an emergent property of some other, possibly unrelated process.

It has been proposed (Lacquaniti et al. 1983) that for a large class of movements with curved hand paths, the speed $\nu(t)$ is related to the curvature $\kappa(t)$ through a power law: $\nu(t) = \gamma\kappa(t)^{-1/3}$, where γ is a constant gain factor. The validity of this law is typically demonstrated through a correlation between ν and $\kappa^{-1/3}$ or between $\log(\nu)$ and $\log(\kappa)$.

Although the resulting correlation coefficients are rather high (Viviani and Cenzato 1985), it is not immediately obvious from such analyses how much of the details of the speed profiles are captured, apart from the general trend that speed decreases with increasing curvature. It has been argued, for example, that the actual speed profiles can be skewed relative to the power law prediction, and the correlation between $\log(\nu)$ and $\log(\kappa)$ still remains high (Wann et al. 1988). The analysis that could resolve this issue, which thus far has been absent from the literature, is a direct comparison superimposing the actual and predicted speed profiles. In its original form the power law has two limitations: first, it does not apply to paths that have inflection points or straight segments (i.e. for $\kappa = 0$ the predicted speed is infinite), and it is inaccurate at low curvature; and second, the fact that $\nu(t)$ decreases toward 0 at the end points of discrete movements cannot be captured. To fix the first problem, Viviani and Schneider (1991) suggested a modified power law: $\nu(t) = \gamma[\kappa(t) + \epsilon]^{\beta}$ where ϵ depends on the average speed and β is a free parameter estimated from the actual data.

Despite these limitations, the power law seems to provide a surprisingly good description of the speed profiles of complex movements, which has motivated recent attempts to find an explanation of this regularity. Within the framework of equilibrium-point control (Bizzi et al. 1992), Gribble and Ostry (1996) have shown that if the equilibrium point is moved along an ellipse with a constant speed, the hand traces a similar ellipse obeying the power law. A potential problem with this explanation is that the validity of the equilibrium-point hypothesis itself recently has been questioned in studies arguing that the required stiffness is much larger than experimental measurements under dynamic conditions (Bennett et al. 1992; Gomi and Kawato 1996). Another proposed explanation (Pollick and Sapiro 1996) is the mathematically interesting fact that a point moving according to the power law maintains constant ''affine velocity''; it remains to be clarified how affine velocity is related to arm movements and why it might be advantageous to keep it constant. An alternative suggestion, which we will consider in the present paper, is that maximally smooth (i.e. minimum jerk) movements have speed profiles described by the power law (Wann et al. 1988). Viviani and Flash (1995) tested this empirically by applying both the 2/3 power law and their modified minimum-jerk model (see further) to the same experimental data. These authors concluded that the speed-curvature relationship described by the power law is not implicit in the minimum-jerk hypothesis.

Partial models of trajectory formation

Finally, a number of intermediate models have been proposed, attempting to bridge the gap between providing a complete recipe for trajectory formation and simply quantifying intrinsic regularities of observed trajectories. Such models assume that the motor task is transformed (via some unspecified planning process) into a compact spatio-temporal representation, involving a small number of parameters that are sufficient to generate the observed trajectory.

Morasso and Mussa-Ivaldi (1982) suggested that complex hand trajectories are composed of partially overlapping linear ''strokes'' (modeled as B splines with bell-shaped speed profiles) that can be scaled and positioned as to form a desired trajectory. The parameters of the model (timing, scale, and position of each segment) initially were extracted from the experimental data and then adjusted iteratively until a good fit to the observed trajectories was obtained.

Viviani and Flash (1995) proposed a modification of the minimum-jerk model, which assumes that the movement is represented internally as a set of salient points and the velocity (speed direction and magnitude) at these points. The authors extracted the velocities from the observed hand trajectories and used standard optimization methods to solve for the resulting minimum-jerk path and speed profile (the solution is a concatenation of 5th order polynomial splines). This model also has been studied by Yalov (1991), who found that overall it is more accurate than versions of the minimum-jerk model using only the position (and tangent to the path) of a via-point.

Another possible model in this category is ''post hoc segmentation'': start with any model that predicts a local speed profile given a portion of the path, define a segmentation rule that somehow selects segment boundaries given the complete path, and apply the model to each segment. Because the local speed profiles generated by this procedure will have to be scaled appropriately, one has to measure the actual time it takes the subject to traverse each segment of the path and use those times in fitting the model.

Because the hypothetical internal representation is not directly observable and cannot be derived from the task description (otherwise the model would belong to the 1st category), it has to be inferred from the observed trajectory itself. Thus the above models are validated by extracting their parameters from the experimental data and demonstrating that these parameters are sufficient to reconstruct the same data. Strictly speaking, this procedure shows that the family of trajectories produced by the model contains the family of human arm trajectories; in itself it is not a proof that the proposed internal representation actually is being used by the motor system. To support the latter claim and make detailed comparisons to models in the other two categories, we need a way of quantifying how much of the information present in the data is being extracted in fitting the parameters of the model. This may appear to be a problem with the power law as well because it predicts a speed profile given a path extracted from the observed trajectory. The crucial difference is that the path of a trajectory is not a priori related to the speed profile, whereas in the preceding examples, the parameters extracted from the observed trajectory contain information about all aspects of the trajectory (i.e., both spatial and temporal information).

CONSTRAINED MINIMUM-JERK MODEL

The main advantage of the power law is that in principle it can be applied to almost any movement: given an arbitrary path it always predicts a speed profile. The problems with it are related to the particular formula used to predict speed from path. In contrast, all known problems with the original minimum-jerk model are in the path prediction: it can fail to predict the hand path given the configuration of via-points, but there are no known cases where the path was predicted accurately while the speed profile was not.

Our model combines the appealing features of the two models: it preserves and extends the generality of the power law while replacing the particular relationship between curvature and speed with a smoothness constraint. We propose that for an observed path of the hand in space, the speed profile along that path will be the one that minimizes jerk (in general ''smoothness'' is synonymous to ''having small high-order derivatives''; the measure of smoothness we use is one of many possibilities). This constrained principle, like the power law, quantifies an intrinsic relationship between the path and speed of the observed trajectory rather than suggesting a concrete mechanism for trajectory formation (i.e., we do not suggest that speed is being computed explicitly as a function of hand path). Note that we use the path observed on a particular trial to predict the speed profile on that same trial, which is a natural way to model the significant variability in speed profiles in a block of identical trials (see further). In contrast, the original minimum-jerk model predicts a single optimal movement for a given configuration of via-points and thus can only be applied to average data.

Formal definition

Let $\mathbf{r}(s) = [x(s), y(s), z(s)]$ be a 3D curve describing the path of the hand during a particular trial, where s is the curvilinear coordinate (distance along the path), and the tangential speed is $\dot{s}(t)$ ($\dot{s}$ is a time derivative, $\mathbf{r}'$ is the derivative with respect to s, and boldface signifies vector quantities). Our model assumes that the temporal profile of the movement is the scalar function $s(t)$ that minimizes

$$J = \int_0^T \left\| \frac{\mathrm{d}^3}{\mathrm{d}t^3} \mathbf{r}[s(t)] \right\|^2 \mathrm{d}t \qquad (1)$$

This differs from the original formulation of the minimum-jerk hypothesis in that here the path $\mathbf{r}$ is given, and we are only minimizing over the space of speed profiles. Flash and Hogan (1985) computed both the path and the speed given a set of via points. Let us denote the term inside the integral by L

$$L = \left\| \frac{\mathrm{d}^3}{\mathrm{d}t^3} \mathbf{r}(s) \right\|^2 = \left\| \frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathbf{r}'(s)\dot{s} \right\|^2 = \left\| \frac{\mathrm{d}}{\mathrm{d}t} (\mathbf{r}''(s)\dot{s}^2 + \mathbf{r}'(s)\ddot{s}) \right\|^2$$

$$L = \| \mathbf{r}'''(s)\dot{s}^3 + 3\mathbf{r}''(s)\dot{s}\ddot{s} + \mathbf{r}'(s)\dddot{s} \|^2 \qquad (2)$$

This minimization problem is invariant to rotations and translations of the path $\mathbf{r}$. Our goal is to transform the above expression into a form that makes that invariance explicit. We know from differential geometry (e.g., Prakash 1981) that a 3D curve can be defined uniquely, up to translations and rotations, by its curvature $\kappa(s)$ and its torsion $\tau(s)$. It is not possible to simply express $\mathbf{r}$ in terms of κ and τ and substitute in *Eq. 2*. However, the derivatives of $\mathbf{r}$ can be expressed as functions of the curvature and torsion. The path $\mathbf{r}$ satisfies Frenet's formulas

$$\mathbf{t}' = \kappa\mathbf{n}, \quad \mathbf{n}' = \tau\mathbf{b} - \kappa\mathbf{t}, \quad \mathbf{b}' = -\tau\mathbf{n}$$

where $\mathbf{t}(s)$ is the unit tangent, $\mathbf{n}(s)$ is the unit normal, and $\mathbf{b}(s)$ is the unit binormal vector to the curve $\mathbf{r}(s)$ at point s. Using the fact that the derivative of the path with respect to the curvilinear coordinate s is the unit tangent vector, we obtain

$$\mathbf{r}' = \mathbf{t}', \quad \mathbf{r}'' = \kappa\mathbf{n}, \quad \mathbf{r}''' = \kappa'\mathbf{n} + \kappa(\tau\mathbf{b} - \kappa\mathbf{t})$$

We can now substitute in *Eq. 2*

$$L = \| \mathbf{n}(\kappa'\dot{s}^3 + 3\kappa\dot{s}\ddot{s}) + \mathbf{t}(\dddot{s} - \kappa^2\dot{s}^3) + \mathbf{b}(\dot{s}^3\kappa\tau) \|^2 \qquad (3)$$

and because the tangent, normal, and binormal are orthogonal

$$L = (\kappa'\dot{s}^3 + 3\kappa\dot{s}\ddot{s})^2 + (\dddot{s} - \kappa^2\dot{s}^3)^2 + (\dot{s}^3\kappa\tau)^2 \qquad (4)$$

Relation to the power law

Now that we have an explicit relation between curvature and jerk, we analyze the constrained minimum-jerk principle to find how it relates to the power law. We define a function [dependent on $s(t)$] that is equal to the term multiplying the torsion τ in *Eq. 4*

$$A_s = \dot{s}^3\kappa(s)$$

Taking derivatives with respect to time on both sides, we obtain

$$\kappa'(s)\dot{s}^4 + 3\dot{s}^2\ddot{s}\kappa(s) = A_s'\dot{s}$$

$$\kappa'(s)\dot{s}^3 + 3\dot{s}\ddot{s}\kappa(s) = A_s'$$

Note that the term on the left is exactly the term multiplying the normal vector in *Eq. 3*. We now substitute in *Eq. 3*

$$L = \| \mathbf{n}(A_s') + \mathbf{t}(\dddot{s} - A_s\kappa) + \mathbf{b}(A_s\tau) \|^2$$

$$L = A_s'^2 + (\dddot{s} - A_s\kappa)^2 + A_s^2\tau^2 \qquad (5)$$

From the definition of A_s, we have

$$\dot{s}(t) = A_s^{1/3}\kappa^{-1/3}$$

The power law states that $A_s^{1/3} = \text{const}$ or $A_s' = 0$. Thus the power law is equivalent to setting the normal component of the instantaneous jerk to 0, and the binormal component proportional to τ, everywhere along the movement.

For two-dimensional (2D) curves $\mathbf{r}(t) = [x(t), y(t)]$, this can be shown more directly. The power law $v(t) = \text{const}.\kappa(t)^{-1/3}$ in 2D can be written as

$$(x'^2 + y'^2)^{1/2} = \text{const} \left(\frac{\sqrt{(x'y'' - y'x'')^2}}{(x'^2 + y'^2)^{3/2}} \right)^{-1/3}$$

$$x'y'' - y'x'' = \text{const}$$

Taking derivatives, and canceling terms, we get

$$x'/y' = x'''/y''', \quad \mathbf{r}'\|\mathbf{r}'''$$

The jerk vector points along the tangent, which is orthogonal to the normal; therefore the jerk along the normal is zero.

Before analyzing experimental data, we can ask whether the power law and the constrained minimum-jerk model predict similar speed profiles for discrete movements along arbitrary synthetic paths. In Fig. 2, we have plotted three synthetic paths, for which we have computed the speed profiles predicted by our model (——) and the power law (– – –). Also shown is the magnitude and direction of the instantaneous jerk vector at several positions along the path. Note that in most cases the normal jerk is close to zero, and the two speed profiles are rather similar (the deviation near the end points is due to the fact that the power law cannot predict movements that start and stop). There are differences, how-
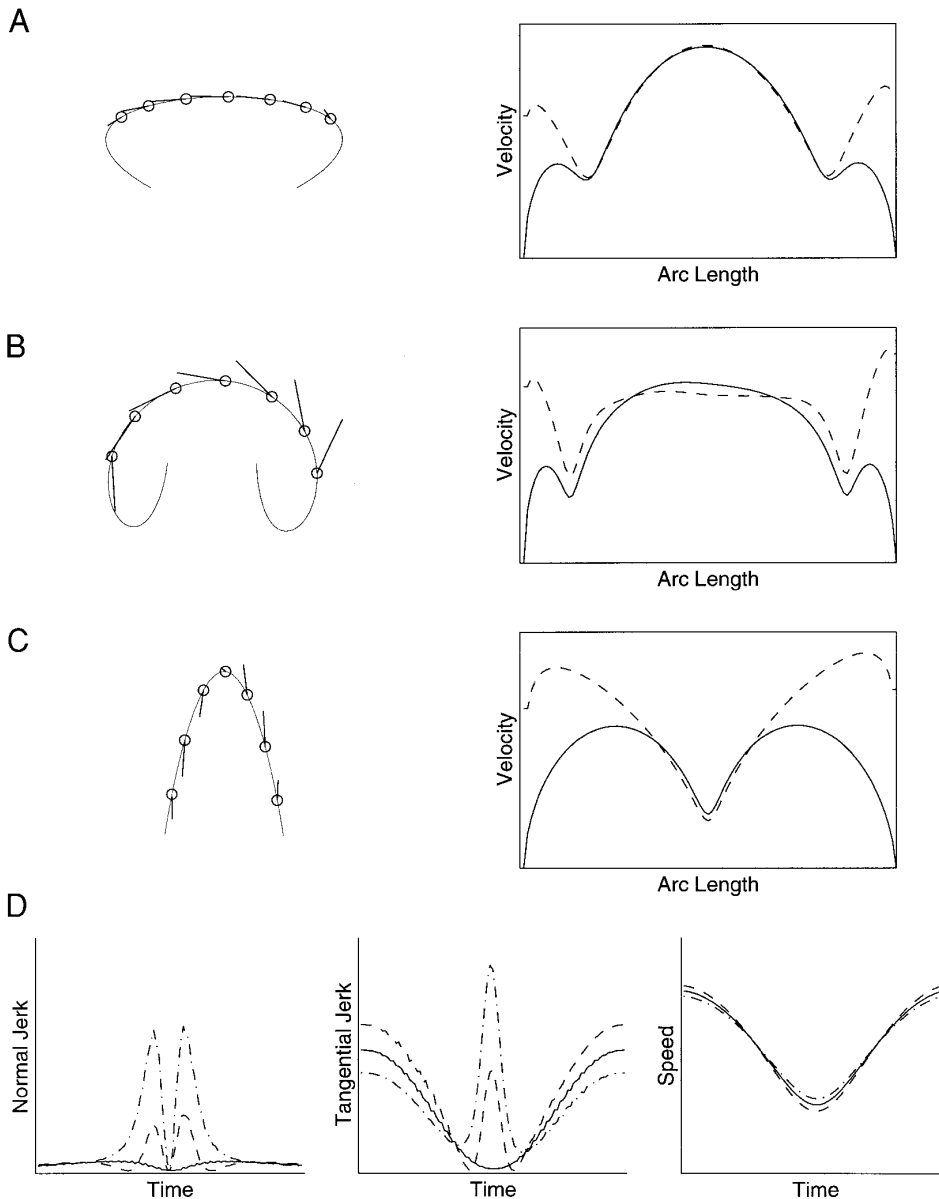
A

B

C

FIG. 2. *A–C*: examples of the constrained minimum-jerk (———) and power law (– – –) speed profile predictions for the synthetic paths on the *left*. Jerk vector is plotted at several points along the path (for the minimum-jerk speed profile). *D*: tangential and normal jerk components for the middle portion of *example C*. Two small perturbations are applied to the minimum-jerk speed profile (———) and the corresponding effects on the jerk components are shown.

D

ever: in *example B* where the curvature is almost constant, the power law predicts a speed profile that is flatter than the minimum-jerk prediction. For most of the paths we have studied, the two predictions are similar in the sense that they go up and down together, but clearly distinguishable (i.e., *example A* is not typical).

Why is it that the power law, which predicts instantaneous speed only using the local curvature, comes close to solving the global minimization problem (*Eq. 1*)? One intuitive answer is the following: because the total jerk is a sum of two (in 3D of 3) nonnegative terms and the power law minimizes one of them (normal jerk is set to 0), it is an efficient approximation to the solution of *Eq. 1* (suggesting that in 3D the difference between the 2 speed predictions will be larger). It seems, however, that this is not always the complete answer. In Fig. 2*D* we have plotted the tangential and normal components of the jerk over the middle portion of *example C*. We can see that small perturbations around the

optimal speed profile increase both terms, implying that in this case the optimal speed profile simultaneously minimizes both terms and not just their sum (this happens around points of high curvature). In summary, the constrained minimum-jerk model and the power law predict similar speed profiles for a given path. There seem to exist a family of paths for which the two are exactly equivalent, and identifying that family may provide further insights, but this is outside the scope of the present paper.

EXPERIMENT 1

Methods

To test our model and compare it with the power law, we studied four tasks in which a path was specified and subjects were asked to execute an arm movement along that path. We used discrete movements of short duration and paths of varying degree of complexity. To avoid drawing conclusions

from an isolated movement, we used several paths in each task. The same eight right-handed subjects participated in all four tasks. Short breaks were given between tasks. The entire experiment lasted ~45 min. End-point position was recorded at 100 Hz with an Optotrak 3020 infrared system.

TASK 1—HAND/FINGER MOVEMENTS. Subjects stood in front of a horizontal table to which we had taped a white sheet of paper (letter format) with the templates shown in Fig. 3 printed on it. Subjects held a pointer in their right hand; the task was to position the pointer near the circle on the current template, wait to hear a beep signaling that the tip of the pen was in the starting area, and then trace the specified path. After 10 trials, a beep of different pitch signaled the subject to move to the next template. No instructions were given regarding the duration or speed of the movement. We observed that subjects used predominantly hand and finger movements in this task.

TASK 2—2D ARM MOVEMENTS. Subjects sat on an adjustable chair and made arm movements on a table positioned slightly below shoulder height so that the movements (made by sliding the arm on the table) occurred in a horizontal plane at shoulder height. We used a VGA projector to display computer generated images on a tilted screen, which the subject

FIG. 3. Templates used in *tasks 1* and *2,* drawn to scale. ○, starting point on each template. Order of presentation is row-1st.

viewed in a semitransparent mirror. The system was calibrated so that the projected images appeared to be on the table (see Wolpert et al. 1995 for details). Subjects held a small cursor (tracked by the infrared cameras), which had a light-emitting diode attached to it that provided end point visual feedback during the movement. The room was dark so subjects could not see their arm. The same eight templates from *task 1,* scaled up by a factor of 8, were projected one at a time. The task was again to position the cursor at the starting point of the template, wait for a beep confirming a correct initial position, and then trace the path. Each path was traced 10 times. After each trial, we displayed a positional error score, which was the maximum deviation (in millimeters) between the desired and actual paths. Again, subjects were allowed to make movements of any speed. This task involved predominantly shoulder and elbow movements.

TASK 3—VIA POINTS. We used the same setup as in *task 2*. The display contained a start point, an end point, and four intermediate points, numbered 1–4. The subject was asked, after moving to the start point and hearing a confirmation beep, to make a movement that passed through the intermediate points in the specified order (without stopping), and finished at the end point. The configurations we used are shown in Fig. 4. There were a total of 6 configurations, every other configuration being a 90° rotation of the preceding one. Subjects executed 10 trials in each block. As before, movement speed was not specified. Note that the configurations used do not correspond to the continuous paths from the previous two tasks.

TASK 4—3D ARM MOVEMENTS. Subjects stood in front of the table, holding the cursor in their right hand and a 3D wireframe model in their left hand. The models were ~10 × 10 × 10 cm and represented smooth 3D curves. The task was to first study the model and then try to make a movement with the right hand that had the same shape (not necessarily the same size). The starting point on the model was marked. There was no specified starting position for the arm because the movement could not be superimposed on the wireframe model. At the beginning of each trial, subjects had to stop moving the cursor and wait for a beep. The experimenter waited for the subject to reposition and then manually activated a procedure that checked for zero speed before movement onset. Each wireframe model was used in 10 trials, there were a total of three models. Movement speed and duration were not specified.

Data analysis

The data from each trial for each subject and task were analyzed separately, i.e., no averaging over trials was performed. The first step was to smooth the raw data. We used cubic spline smoothing (De Boor 1978) applied to $x(t)$, $y(t)$, and $z(t)$. The parameter λ (which determines the amount of smoothing) was set adaptively for each trial, so that the maximum deviation between any smoothed sample point and the corresponding raw data point was within a pre specified radius. That radius was set to 0.002 of the extent of the path—the high accuracy of the Optotrak system obviated the need for substantial smoothing. Once we obtained
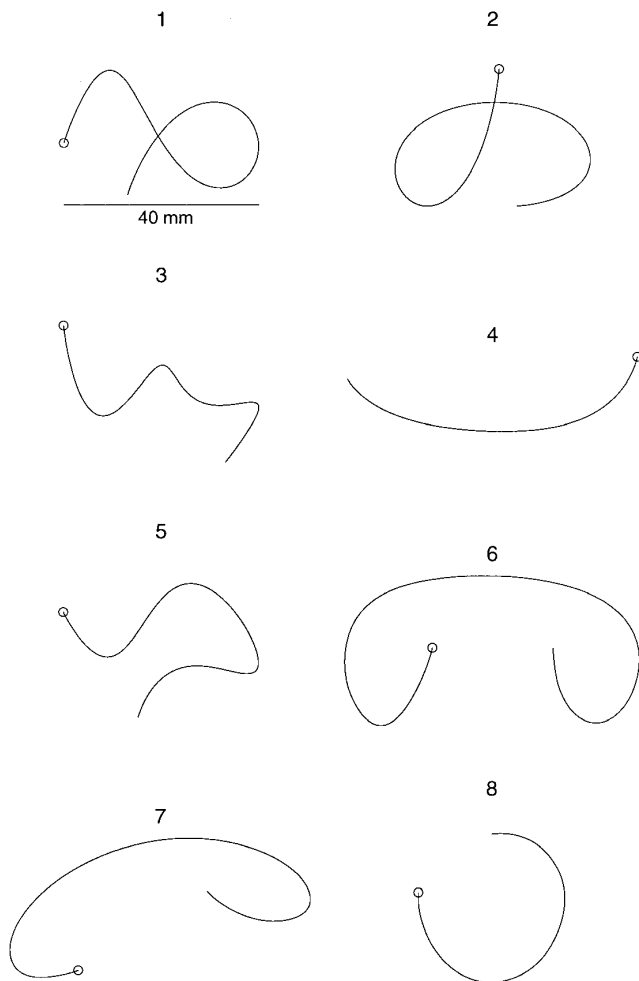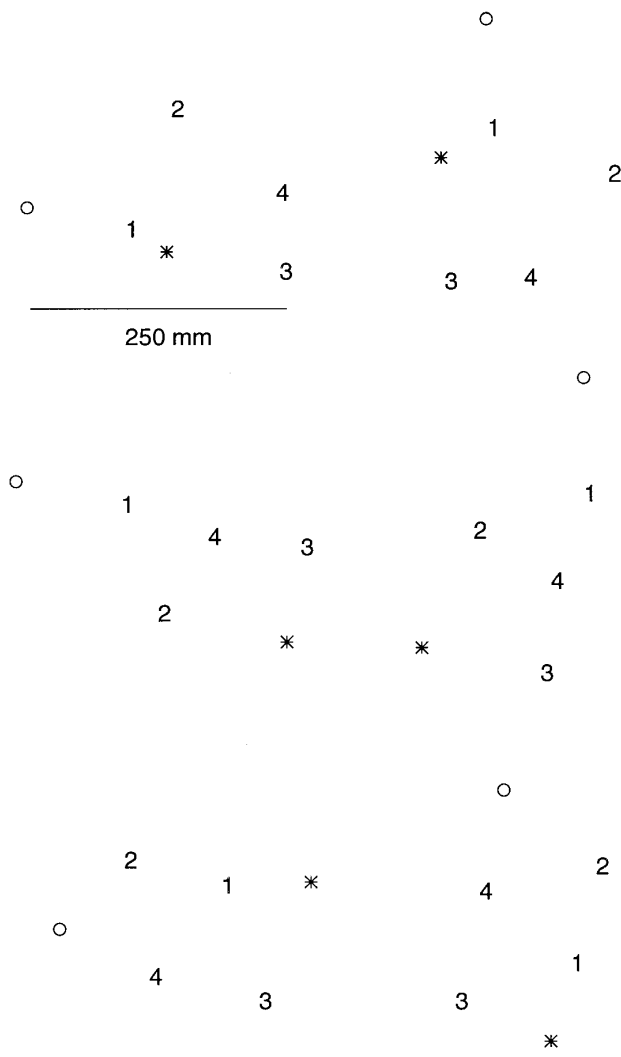
FIG. 4. Via point configurations used in *task 3,* drawn to scale. ○, starting point; *, end point. Every other configuration (*right*) is a 90° rotation of the preceding one.

Maximum deviation	$MAX(\nu1, \nu2) = \max\left[\,|\nu1(s) - \nu2(s)|\,\right]$
Mean absolute deviation	$MAD(\nu1, \nu2) = \langle|\nu1(s) - \nu2(s)|\rangle$
Standard deviation	$STD(\nu1, \nu2) = \sqrt{\langle[\nu1(s) - \nu2(s)]^2\rangle}$
Unexplained variance	$VAR(\nu1, \nu2) = 1 - r[\nu1(s), \nu2(s)]^2$

The first three quantities indicate deviations between instantaneous speed predictions in units relative to the average speed. Thus they are sensitive to the amount of fluctuations in the observed speed profiles (i.e., relatively flat speed profiles will be easier to fit according to these measures). If the speed profiles are not scaled to have unit average speed, then these measures will be sensitive to the average speed. The last measure corresponds to the percent variance (fluctuations around the mean) in the observed speed profile not explained by the predicted speed profile and is affected by difference in shapes rather than scaling factors.

For each model, we computed the predictions over the entire movement and over the middle 60% of the movement (i.e., starting at 20% and ending at 80% along the path). The latter was done to assess the accuracy of the power law away from the end points, where it is likely to be inaccurate because it doesn't have a mechanism for enforcing 0 speed at the initial and final points of discrete movements. We examined a number of speed profiles to ensure that the 20% of the movement eliminated at each end was enough to cover the parts of the speed profile that seemed affected by initial acceleration and final deceleration.

The prediction for the original power law was obtained directly from the equation $\nu(s) = \gamma\kappa(s)^{-1/3}$. For paths with inflection points, we added a small positive constant to the curvature, which was held constant for all trials. Thus the original power law has no free parameters: it takes the path, and the duration of the movement, and predicts a speed profile. The modified power law $\nu(s) = \gamma[\kappa(s) + \epsilon]^{\beta}$ has two free parameters (ϵ, β). We optimized them (with a nonlinear simplex method) separately for each trial and each measure of deviation defined above. Although in the original formulation ϵ depended on average speed, we estimated it from the data to obtain the most accurate fit the modified power law could provide. The possibility of movement segmentation (Viviani and Cenzato 1985) is addressed later.

The prediction of the constrained minimum-jerk model was computed using nonlinear optimization methods for variational problems (see APPENDIX). In its present form, the minimization requires the speed and acceleration at the end points (in principle this could be avoided, by minimizing over those parameters as well). For a discrete movement, both the speed and acceleration at the end points are zero; because this is known a priori, and not extracted from the experimental data, the model has no free parameters when applied to the entire movement. To apply the model to the middle 60% of the movement, we used the actual end-point speed and acceleration measured experimentally. Note, however, that we are predicting rather complex speed profiles, thus the two end points of a submovement contain relatively very little information. In contrast, the modified power law uses the entire speed profile to extract the values of its free parameters.

the cubic spline representation of the data for each trial, all subsequent steps (i.e., extracting speed and curvature) were applied directly, without any further filtering.

We then computed a speed profile prediction from the original power law, the modified power law, and the constrained minimum-jerk method (see further). To compute the difference between each prediction and the actual speed profile, we expressed all speed profiles as a function of are length s and scaled them so that the average speed was one and the path lengths were equal. Scaling of path lengths was necessary because subjects did not follow the specified templates exactly, thus the paths on different trials had slightly different length. Alignment along s does not affect the power law predictions, which are local; it is necessary for our model because it is a global method and prediction errors early in the movement accumulate and cause misalignment towards the end if speed is expressed as a function of time.

We defined four measures of deviation between two speed profiles $\nu1(s)$ and $\nu2(s)$ that are already scaled to have unit average speed; $\langle\cdot\rangle$ signifies averaging over s, $r(.\,,\,.)$ is a correlation coefficient.

Results and discussion

In the first two tasks, for which the complete path was specified and the movement was superimposed on it, subjects could follow the specified paths accurately. In the 3D task, the majority of subjects could not replicate the specified shapes well and did not converge to a very stable movement strategy in 10 trials. After the experiment, subjects reported that they were aware of their poor performance but did not know how to perform the task better. Although this is an important issue to consider, both the power law and the constrained minimum-jerk model predict a speed profile given the actual, not the desired path of the hand, so we still can apply them and assess the accuracy of their predictions. It also was observed that the torsion τ varied rather smoothly instead of being nonzero only at a discrete set of points as suggested by Morasso (1983) and Soechting and Terzuolo (1987). Deviations from piecewise-planar drawing in 3D also have been observed by Gusis (1995); we did not analyze this discrepancy further because it is not the focus of the present study.

In the via-point task, we specified four intermediate points, and subjects were free to choose any path that passed through them in the correct order. Our constrained minimum-jerk method cannot make a prediction about the actual path subjects choose in this task, but the original model (Flash and Hogan 1985) predicts that the path will be the one that minimizes jerk. While this is often the case, we found numerous systematic deviations. Figure 5 shows data from six different subjects; each pair of plots represents the last 5 of the 10 trials ($\cdots$) that the subject made, on the original

and the rotated configuration of via-points. The minimum-jerk path (——) remained invariant to rotation. Note that the deviations are not just isolated trials but stable strategies used by the subjects; also, the actual paths depended on the orientation of the template, contrary to the model prediction. Although some of these deviations may be the result of arm dynamics affecting the path, the large differences (e.g., Fig. 5A) between the paths adopted by different subjects for the same via-point configuration suggest that multiple via-point tasks may have a significant cognitive component, i.e., subjects have to study the configuration and decide what general path they want to take. In fact, we observed that on the initial trials in a new configuration, subjects sometimes stopped and corrected their movement, which is unlikely to happen if they already had planned a smooth path that passes through the via-points in the specified order. Whatever the reason for these discrepancies, we conclude that the original minimum-jerk model cannot account for the significant between subject variability on multiple via-point tasks.

We now turn to a quantitative comparison of our model and the power law. Figure 6 shows summary error statistics, for all tasks and all error measures, over the complete (*left*) and middle 60% of the movement (*right*). For each subject, we computed the median prediction error in each block of 10 trials and then averaged over all templates in a given task. The constrained minimum-jerk model performed uniformly better than the modified power law, which in turn was uniformly better than the original power law (the latter obviously has to be the case). Analysis of variance showed that in the majority of the comparisons (the 3D task/complete movements on the MAD and STD measures being the excep-
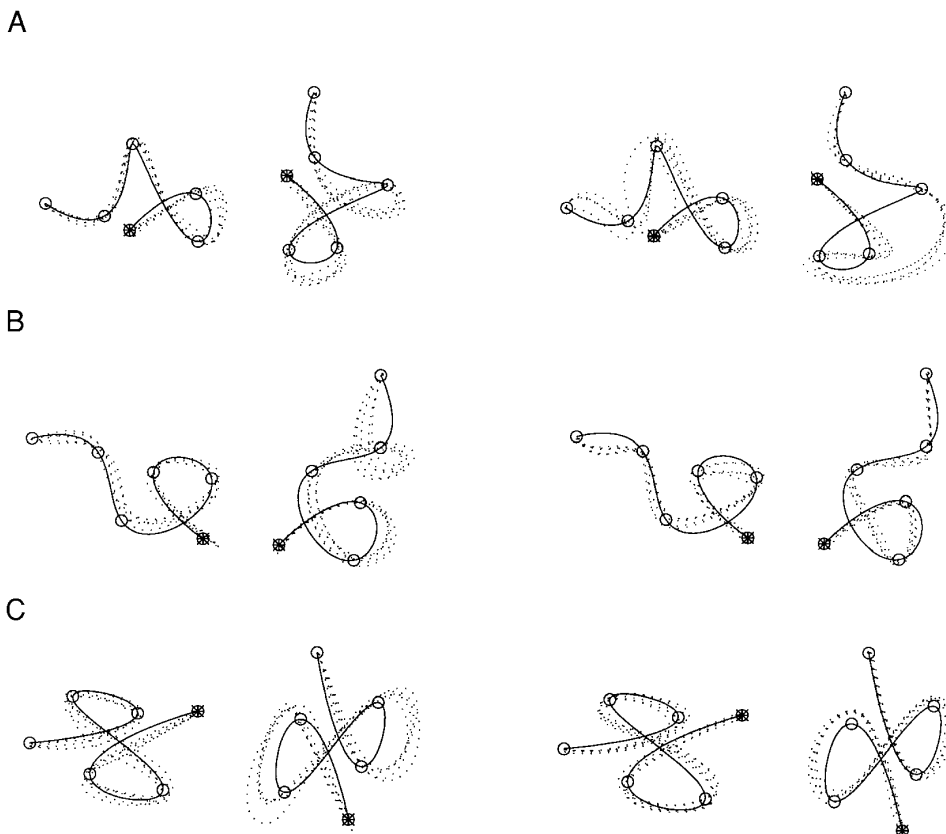
A

B

FIG. 5. Examples of paths traced in *task 3*. Each pair of plots is data from 1 subject (6 different subjects shown). ——, minimum-jerk path (which was not shown to the subjects); $\cdots$, actual paths on the last 5 trials of the 10 trial block. ●, end point.
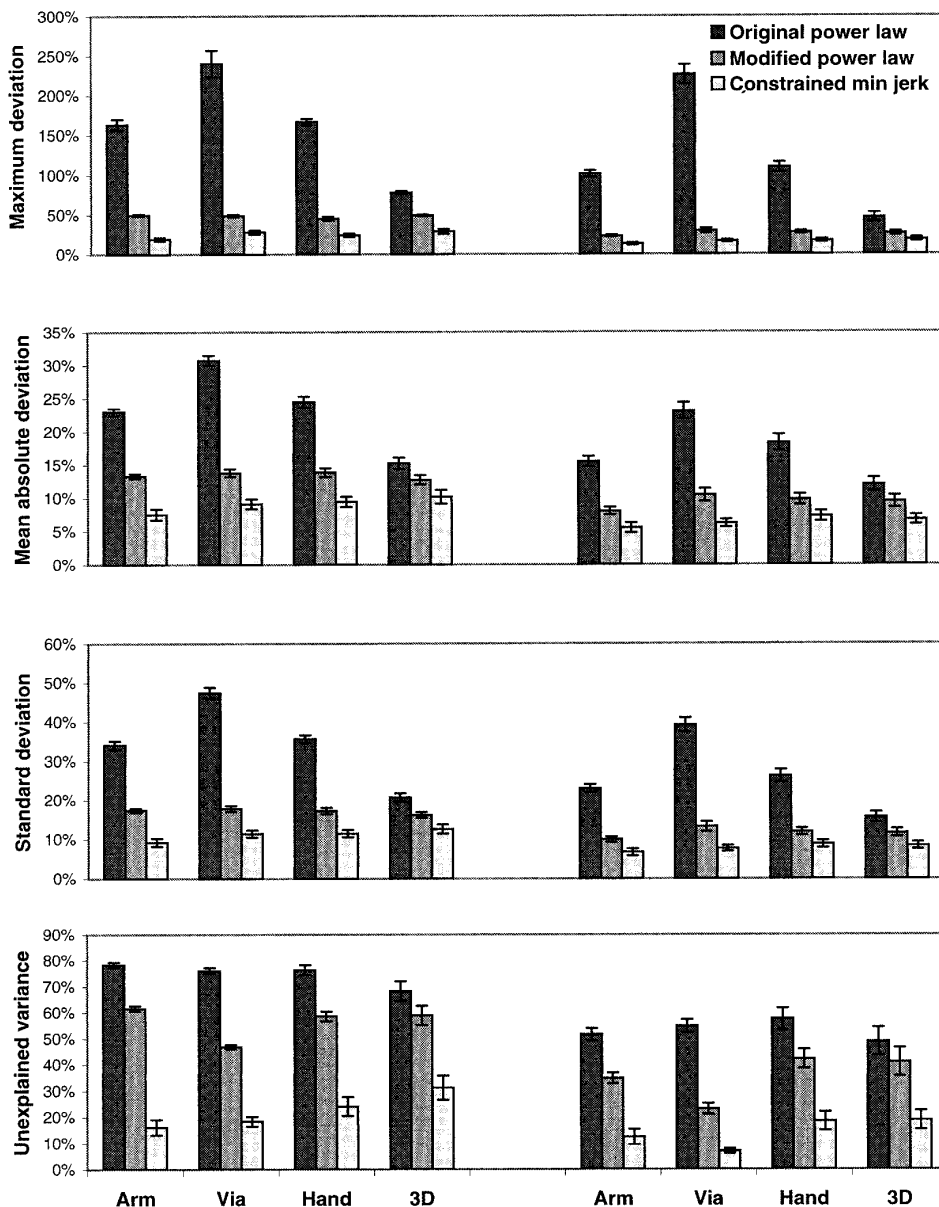
C

FIG. 6. Summary error statistics with standard error bars for the 3 models, computed over the entire movement (*left*) and the middle 60% (*right*) for all tasks and error measures. For each subject, the median error score in each block of 10 trials was computed, and the results were averaged over all templates in the task.

tion), the differences are significant at the $P < 0.05$ level. Furthermore, this ordering holds for every subject we tested. The difference is equally large over the complete movement and the middle 60%, therefore it is not explained by the discrepant predictions of the power law at the end points of discrete movements.

To analyze the dependence of the model performance on the template being traced, we averaged the data over subjects and compared the three prediction methods for each template. Figure 7 (arm task—complete movements) shows that the ordering holds for all templates we tested and not just the ones that include inflection points. This was true for all tasks. It is interesting that the constrained minimum-jerk model shows a rather uniform level of performance, both over different templates and over different tasks. We expected that there might be an effect of trial number, resulting from learning the pattern in the course of the 10 trials, but there was no evidence of that. It is still possible that in more

difficult tasks the performance of the model changes with experience, but there was simply very little learning in our experiments. Indeed the spatial error did not change significantly over trials.

There are some differences between the four error measures we defined. The maximum deviation penalizes the original power law on templates with inflection points because the predicted speed at 0 curvature is very large (it would be infinite if we didn't add a small constant to the curvature). The mean absolute deviation and the standard deviation are very similar up to a scaling factor. The unexplained variance seems to be the most sensitive measure—it is the only one that detects significant differences in the behavior of our model on different tasks and templates. Thus it will be the measure used in the remaining part of the paper, where we study how the model depends on the details of the task. All subsequent analyses also have been performed with the mean absolute deviation measure, yielding very similar results.
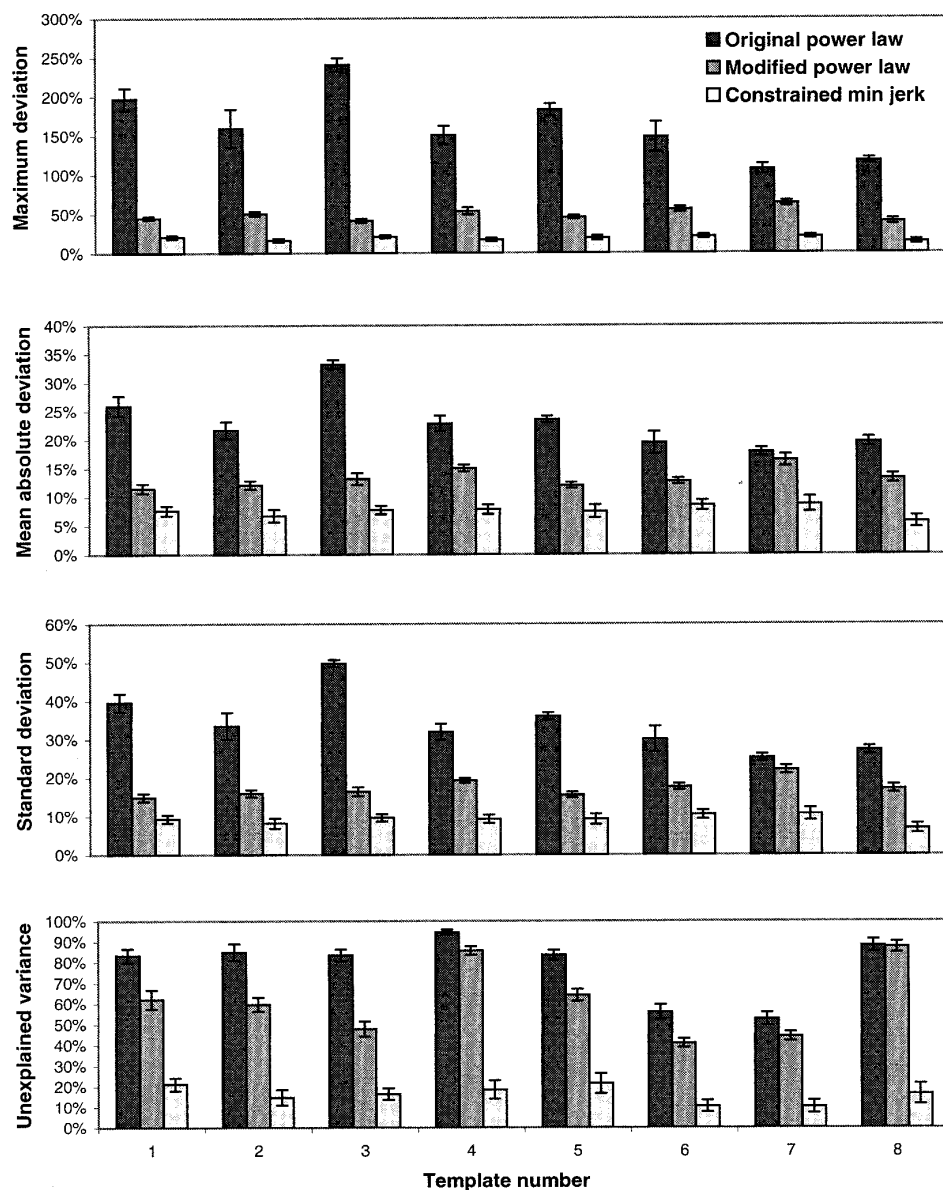
FIG. 7. Prediction errors for all error measures and all templates in *task 2*—complete movements, averaged over subjects.

Analysis of speed profiles

Thus far we have considered statistics based on the average prediction errors. In this section, we perform a more detailed comparison of the actual speed profiles. Figure 8 shows speed profiles of all eight subjects tracing *templates 1, 2, 3,* and *5* from Fig. 3. These data are from the next experiment, where a short movement time (1.5 s) was enforced—these are the conditions under which both our model and the modified power law are most accurate (see further). We have scaled the last five speed profiles (complete movements) in each block for each subject and averaged them; the solid line shows the actual speed profiles; the dashed line shows the constrained minimum-jerk predictions; and the dotted line shows the modified power law predictions. Because scaling does not align the speed profiles perfectly, averaging is not necessarily a meaningful operation, but it helps identify systematic prediction errors. Note that we show averages over the predictions for individual trials not a predic-

tion for the average trial (which would be less meaningful because both prediction methods are nonlinear).

It is evident that the modified power law prediction has significant systematic error, which is not restricted to the end points of the movements. The flat speed profiles it predicts in *template 2* correspond to a region of constant curvature (as was noted in Fig. 2). We conclude that the power law captures very well the maxima and minima of the actual speed profiles but otherwise it is not a satisfactory model of the details present in the experimental data. Although some of the errors are due to skewing (as Wann et al. 1988 argued), the predominant error is in the predicted values at the speed extrema.

The prediction of the constrained minimum-jerk method is much more accurate, and most of the remaining error seems to be due to between-subject variability. Still, there is some systematic error left. The first speed maximum on *template 1,* and the second maximum on *template 3* were both slightly underestimated for all subjects. Overall, the
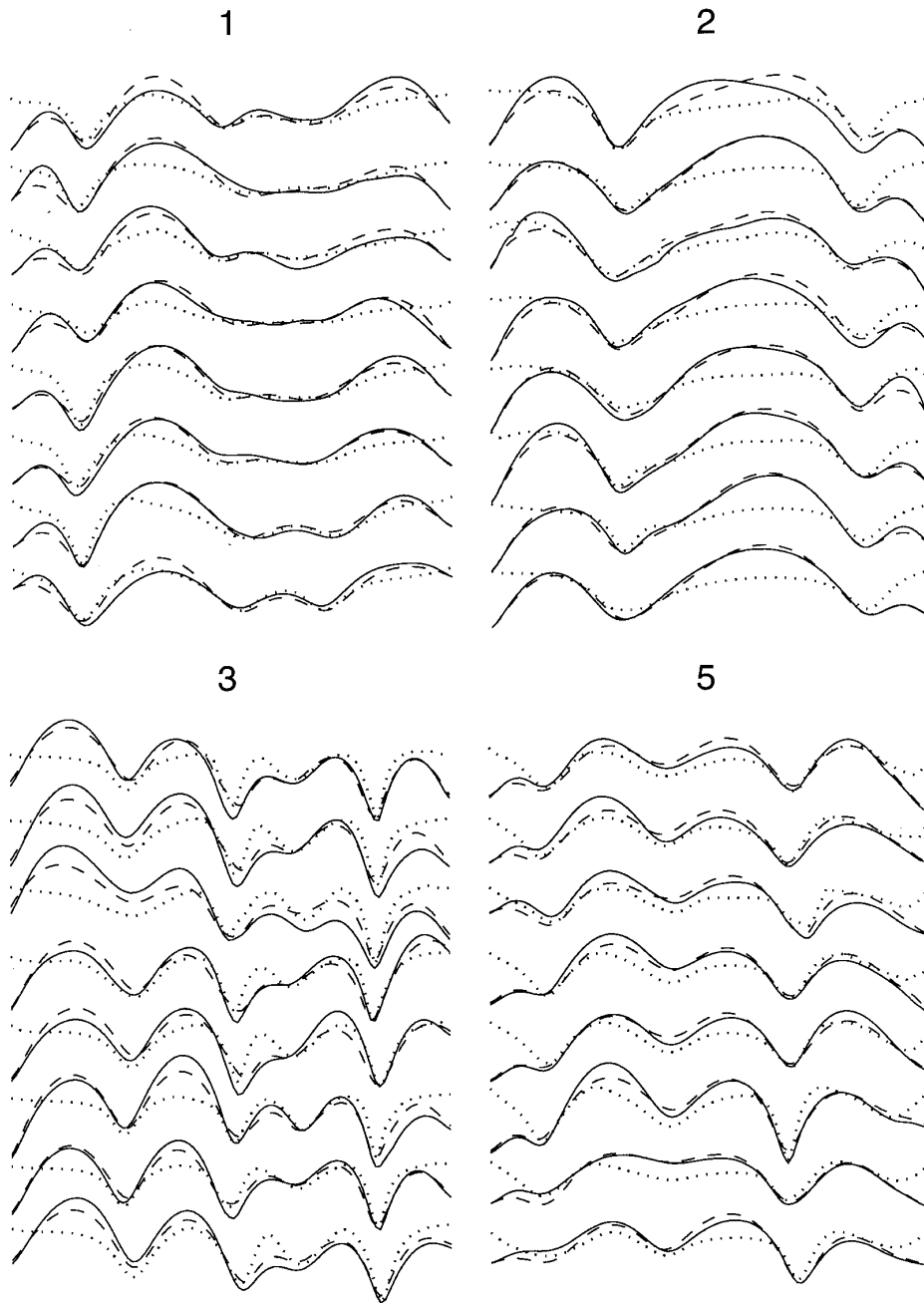
FIG. 8. Average speed profiles for all subjects, small 1.5-s movement in *experiment 2* (4 of the 5 templates shown, numbers correspond to template positions in Fig. 3). Solid line, actual speed profile; dashed line, constrained minimum-jerk prediction; dotted line, power law. All speed profiles were scaled before averaging.

predictions of the model are rather close to the experimentally observed speed profiles, and in some cases (last 2 subjects in *template 2,* for example) it is hard to imagine that any model would be more accurate.

Next we examine data from individual trials because the averaging procedure is likely to obscure some details. For *templates 1, 2, 3, 5,* and *8,* we pooled all trials together and found the one for which both error scores were as close as possible to the corresponding medians. The actual (——), modified power law (· · ·), and constrained minimum-jerk (– – –) speed profiles for the selected trials (middle 60%) are shown in Fig. 9. We can see that there are large errors in the power law prediction that are not related to end point acceleration (deceleration). Note the unnatural sharp peaks, corresponding to inflection points (they were smoothed out

in Fig. 8) — although the modified power law adds a positive constant to the curvature (optimized for the particular trial), this mechanism is apparently not sufficient to deal with inflection points. We might expect that if the paths included straight segments, the performance of the modified power law would be even worse. The inflection points are not the only problem, however, note the deviations in the Fig. 9, *top right*, which correspond to a path with strictly positive curvature.

Tempo fluctuations and segmentation

Both the constrained minimum-jerk model and the modified power law assume that the tempo, or gain, or ''psychological speed'' of the movement is held constant, and the
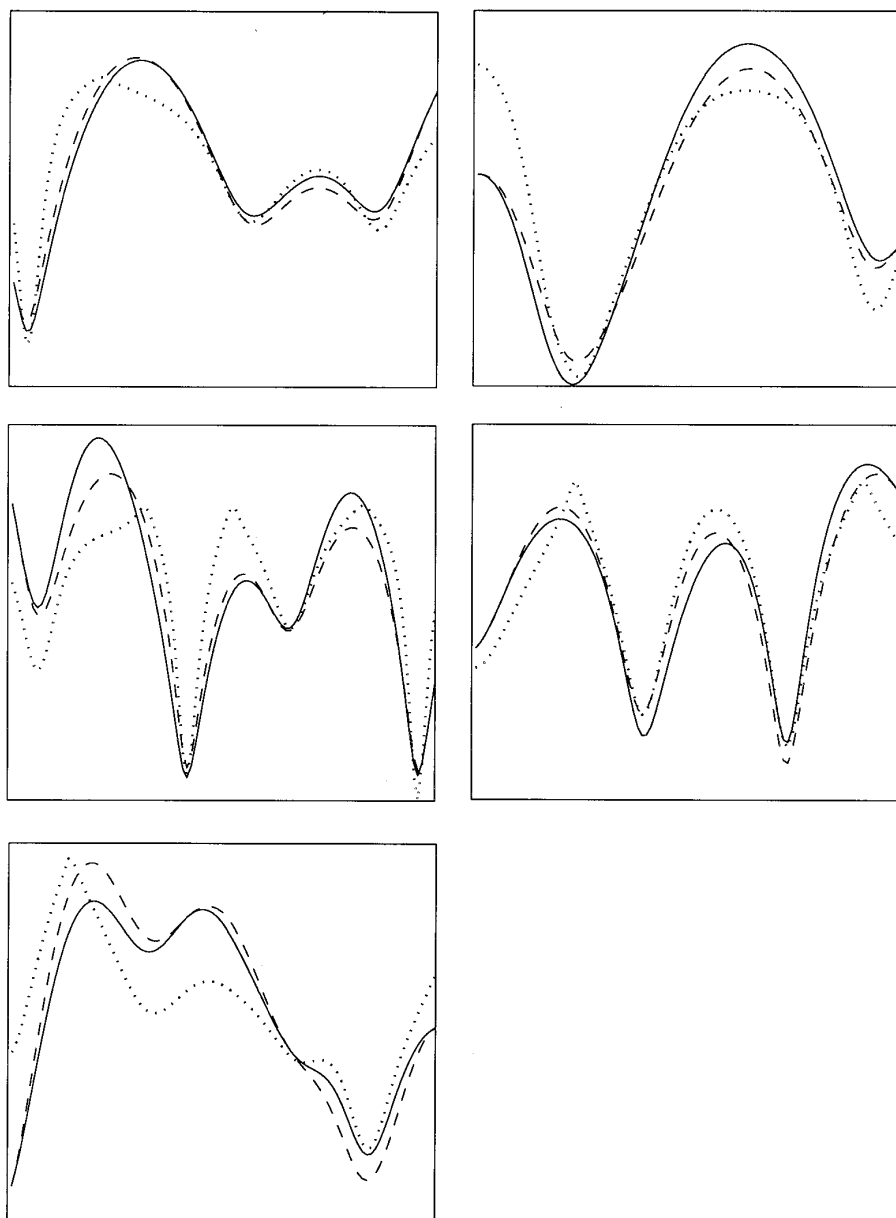
FIG. 9. Typical examples of single trial speed profiles (——), constrained minimum-jerk prediction (– – –), and power law ($\cdot\cdot\cdot$) on single trials—*experiment 2,* middle 60% of the movement.

variation in speed is only related to the path. Clearly that doesn't have to be the case; e.g., we can ask the subjects to speed up toward the end of the movement. It is conceivable that even in the absence of explicit instructions to control speed, the tempo of the movement is variable. For each model, we define the gain (or tempo) as the ratio between the actual and predicted speed profiles at each point in time (the definition is model-dependent by necessity). Viviani and Cenzato (1985) suggested that the gain is a piecewise-constant function of time: in each segment it is proportional to a power function of the corresponding path length: $\gamma = KP^{\alpha}$ (α here corresponds to the parameter γ defined by Viviani and Cenzato).

We computed the gain factor for a number of discrete movements for both models and did not find any case in which it was convincingly piece-wise constant. Figure 10 (*top*) shows a typical plot five consecutive trials for one sub-

ject tracing *template 1.* We then applied the segmented version of the modified power law to the data from all subjects on the arm task—subset of templates used in *experiment 2* (see further). Each movement was segmented at the speed maxima (corresponding to curvature minima). The portions between the end points and the closest segment boundary were eliminated. We applied the modified power law with segmentation with either fixed $\alpha = 0.25$ (which was found to be the optimal fixed value) or α being optimized for each individual trial of each subject (making the comparison to our model somewhat unfair). Figure 10 (*bottom*) shows the speed prediction error, relative to the prediction error of the nonsegmented modified power law, for the two different segmentation methods. The improvement in all cases is much smaller compared with the improvement of the (nonsegmented) constrained minimum-jerk method over the modified power law. This segmentation method also was applied at speed minima (corresponding to
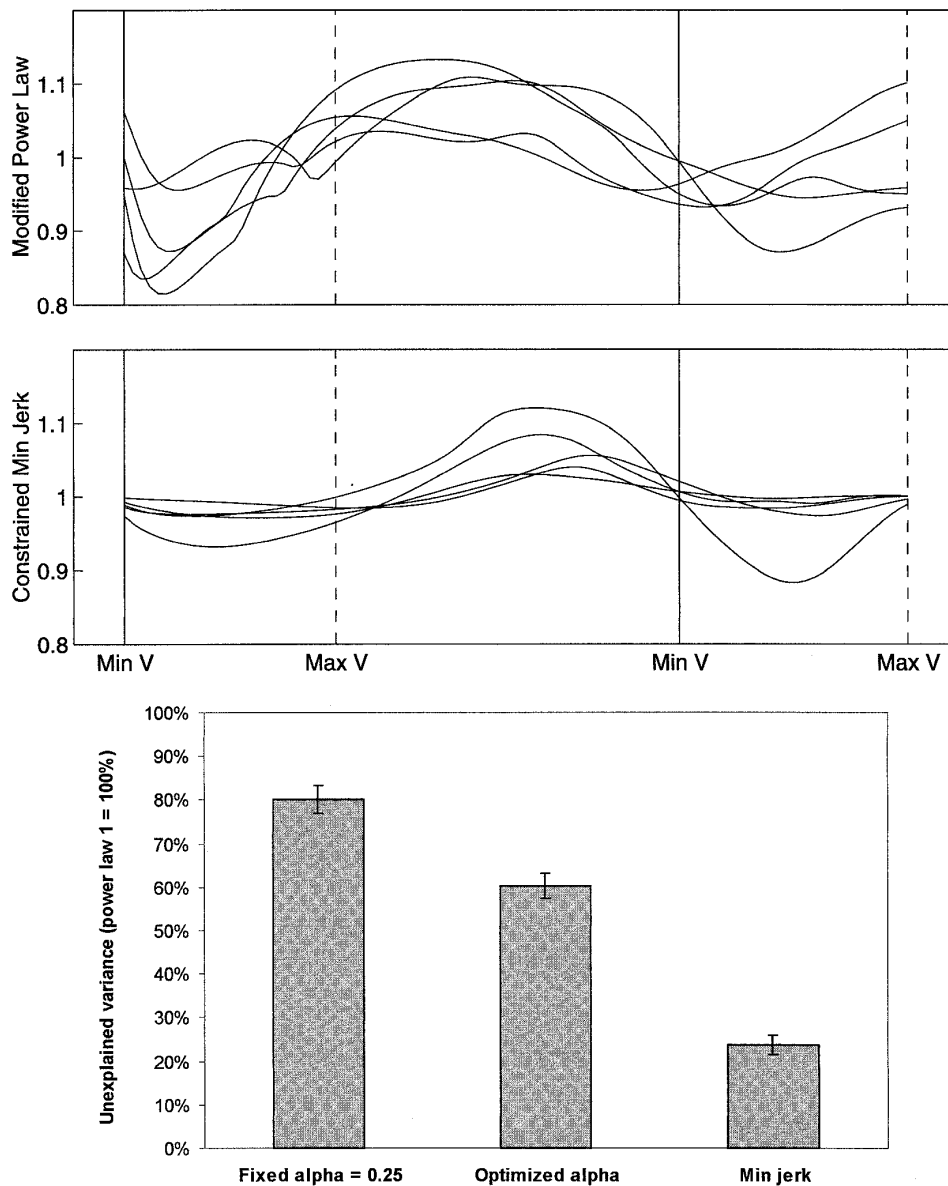
FIG. 10. *Top*: gain fluctuations for the modified power law and our model (consecutive movements of 1 subject tracing *template 1*). Vertical lines correspond to points of maximum and minimum speed. *Bottom*: for each segmentation method (fixed $\alpha = 0.25$ or α optimized separately for each trial) and our model, we show the prediction error relative to the prediction error of the unsegmented modified power law applied over the same portion of the path as the 1 selected by the segmentation scheme.

curvature maxima—as in the model of Morasso and Mussa-Ivaldi 1983) yielding very similar results. Note that the optimal fixed value of $\alpha = 0.25$ is significantly smaller than the value of $\alpha = 0.6$ observed by Viviani and Cenzato (1985). Because $\alpha = 0$ corresponds to a lack of segmentation, this is another indication that segmentation is simply not present in our data, and the tempo fluctuations are continuous and at present unpredictable. The fluctuations in Fig. 10 are systematic because the trials are from the same subject, but they become unpredictable across subjects (as can be inferred from Fig. 8).

Although the details of the gain fluctuations cannot be explained, their magnitude (directly related to prediction error) could vary systematically with some other variable. We already saw that it is not very sensitive to the task or the template being traced. A comparison of prediction error over subjects, however, revealed a strong correlation between how rapidly subjects moved on the average and how well the constrained minimum-jerk model predicted their speed profiles (a weaker correlation was observed for the modified power law). Figure 11 shows a scatter plot of mean movement duration (over all templates) and mean prediction error for our model in all tasks. There are several possible interpretations of this result. It may be that the tempo fluctuations accumulate with movement time, thus movements that take longer have less predictable speed profiles. It is also possible that at high average speed, movements are for some reason more regular. Because we did not vary the size of the templates systematically, duration was correlated with speed [Wann et al. (1988) also observed that the power law holds better for faster movements and attributed the effect to speed instead of duration]. Another possibility is that the effect is a spurious between-subject correlation, i.e., some subjects naturally produce more predictable speed profiles, and the same subjects tend to make faster movements when duration is not specified. The next experiment was designed to distinguish between these possibilities.

Arm Task   r^2=0.78

Via Task   r^2=0.65

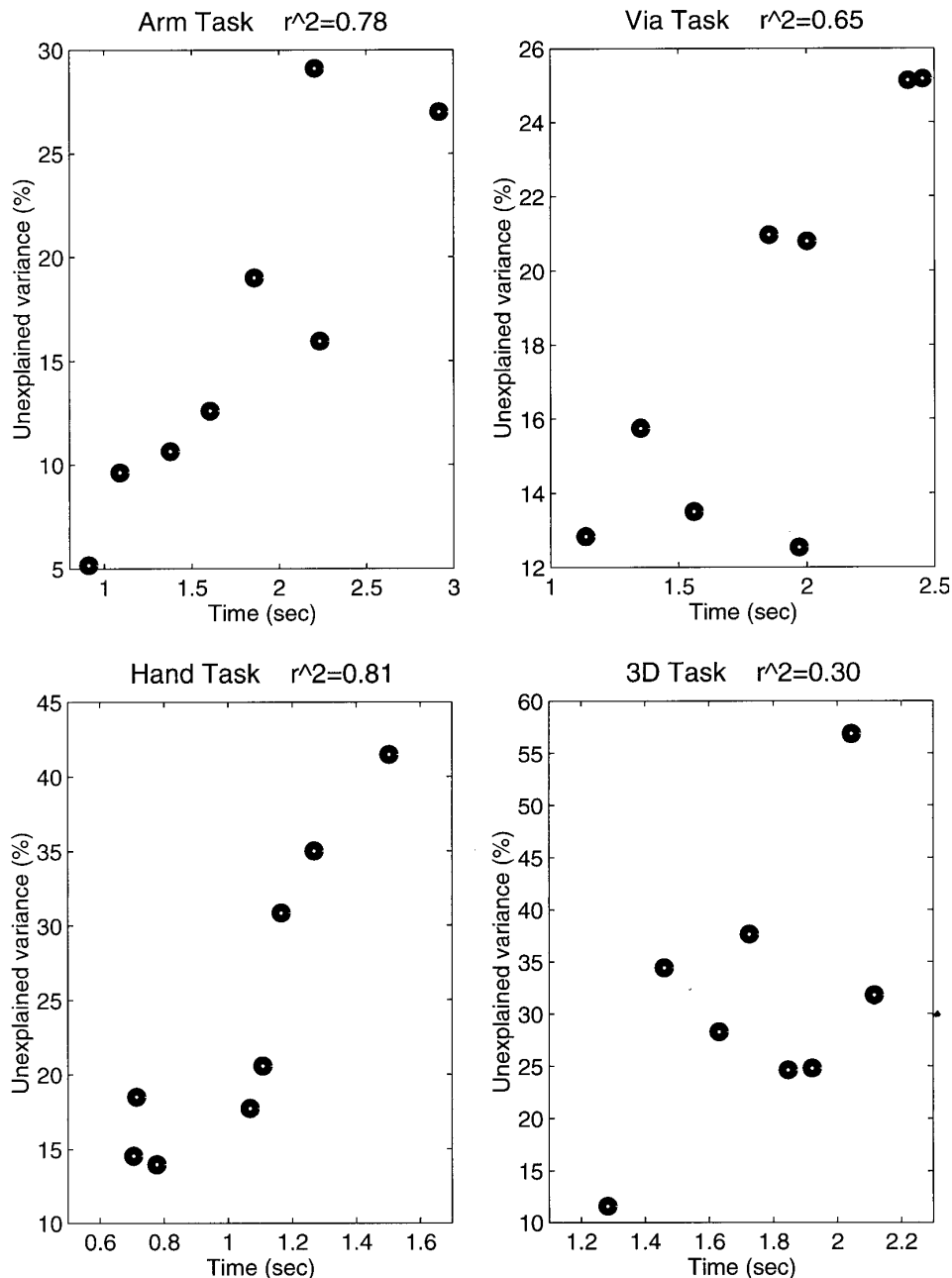## Hand Task   r^2=0.81

3D Task   r^2=0.30

FIG. 11. Scatterplot of movement duration vs. prediction error for the constrained minimum-jerk model, all tasks. Each data point represents 1 subject, the results were averaged over all trials the subject made in the corresponding task.

EXPERIMENT 2

To eliminate the possibility that between-subject variability can account for this correlation, the experiment had a within-subject design. We now specified the duration of the movement and also varied the spatial scale, which allowed us to distinguish between the effects of movement duration and average speed (and size). A new group of eight subjects was recruited for this experiment.

Methods

The setup was the same as in *task 2*—subjects were asked to make arm movements on a horizontal table by tracing a continuous curve projected on the table. To reduce experimental time, we used an arbitrary subset of five templates

(*1, 2, 3, 5,* and *8* in Fig. 3), each presented in three blocks of 10 identical trials. In *block 1,* the template was large (same size as in *experiment 1*), and desired movement duration was 3 s. In *block 2,* the template was also large, and duration was 1.5 s. In *block 3,* the template was scaled down by a factor of 2, and desired duration was 1.5 s. Thus *blocks 1* and *2* were matched for size, *blocks 2* and *3* for movement duration, and *blocks 1* and *3* for average speed (assuming of course that subjects produced movements of the specified durations, which they did) — see Fig. 12, *top.* At the beginning of each trial, the subject positioned the end-point in the starting area and waited for the start circle to disappear, confirming a correct initial position. The computer sounded a beep when the subject started moving and a second beep when the subject stopped moving. After each trial, the actual
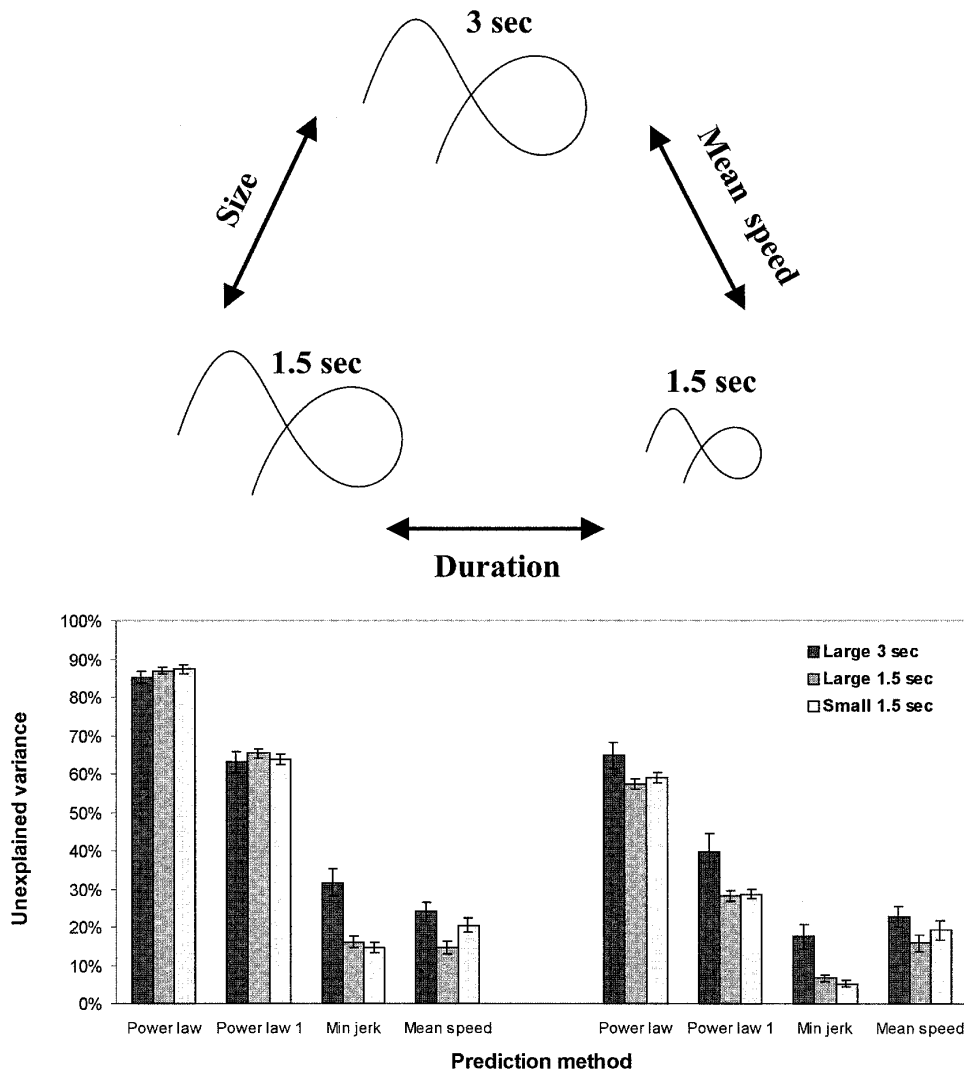
FIG. 12. *Top*: schematic representation of the 3 conditions in the experiment. *Bottom*: summary error statistics for *experiment 2,* complete movements (*left*) and middle 60% (*right*). Results are computed in the same way as in Fig. 6. Last prediction method is the average speed profile over a block of 10 consecutive trials used to predict each individual trial.

duration was shown (in units of 0.01 s). The desired duration was printed on the screen throughout the experiment. Subjects were told what the numbers meant and were instructed to trace the paths for the specified durations.

Results and discussion

The data were analyzed in exactly the same way as in *experiment 1.* All subjects traced the paths accurately and could adjust the duration of their movements in the first two to three trials in each block. Prediction error statistics, averaged over templates, are shown in Fig. 12. Again, we found that the constrained minimum-jerk method predicts speed profiles closer to the ones observed experimentally, compared to both versions of the power law. In this case, the constrained minimum-jerk prediction was about three times better than the modified power law and four times better than the original power law. The difference is still present when speed profiles are computed for the middle 60% of the movements.

It can be seen in Fig. 12 that the prediction errors for the large fast movements are always similar to the small ones and different from the errors for the large slow movements.

Analysis of variance revealed significant differences for the constrained minimum-jerk predictions, and the modified power law applied to the middle 60% of the movements; the prediction error for the large 3-s movement was larger than the error for both the large 1.5-s and the small 1.5-s movements ($P < 0.01$). This result indicates that the accuracy of the prediction depends on the duration not the average speed or spatial scale of the movement. We also can conclude that the correlation in *experiment 1* was not entirely due to systematic between-subject variability because the second experiment yielded the result within subjects. The higher accuracy of our model when applied to the middle 60% instead of the complete movement may be due to the fact that the model is being applied to a movement of shorter duration. However, this also may be caused by using the experimentally observed values of speed and acceleration at the end points of the middle segment.

It is possible that the effect of movement duration on model accuracy is due to the fact that observed speed profiles in short-duration movements are somehow different/simpler, and any model would yield better results. To assess this alternative, we have included in Fig. 12 another prediction method related to the amount of trial-to-trial variability pres-

ent in a block of identical trials. We computed the mean speed profile over each block of 10 trials (separately for each subject), and used it to predict the speed profile for each of the individual trials. The pattern of prediction errors we see in Fig. 12 is very different: it appears that trial-to-trial variability decreases with speed rather than duration and is essentially the same over the entire movement and the middle 60%. Thus the higher model accuracy for shorter durations results from a stronger coupling between path and speed rather than a difference in the speed profiles themselves.

Trial-to-trial variability

It often is assumed that the motor system constructs a plan of the desired trajectory for the current task, and instantiates that plan on each particular trial. Although there is no widespread agreement on what exactly such a plan might contain, it seems reasonable to assume that the actual trajectories produced on consecutive trials on the same task are variations around the same plan, where the variability is due to unspecified sources of noise. Thus the average trajectory over a block of identical trials corresponds to the movement that would result from the plan in the absence of variability. In support of this notion, it has been observed that

the speed profiles of reaching movements of different length and duration can be aligned almost perfectly after straightforward rescaling, and complete models of trajectory formation such as the original minimum-jerk model or the minimum torque-change model (Uno et al. 1988) have been applied to the average trajectory.

Under these assumptions, we can attempt to identify the level of processing where the relationship between path and speed emerges. In particular, the smoothness observed in experimental data may be present already in the plan, in which case, our model should apply better to the average trajectory than to single-trial trajectories (note that averaging by itself does not necessarily produce smoother trajectories, i.e., if we average several optimal trajectories, we will obtain a suboptimal one—because the model is nonlinear). Furthermore, the speed profile predicted from the average path should be more accurate in explaining the single trial speed profiles than the predictions obtained from the single trial paths. In the notation of Fig. 13 [where $MJ(\cdot)$ is the speed predicted by the model for a given path, $\langle\cdot\rangle$ signifies averaging over a block of trials, and : is the error between actual and predicted speed profiles). $MJ(\text{path})$: speed > $MJ(\langle\text{path}\rangle)$: speed > $MJ(\langle\text{path}\rangle)$: $\langle\text{speed}\rangle$. Alternatively, if smoothness is inherent in the (unknown) sources of variability, we should expect the single trial predictions
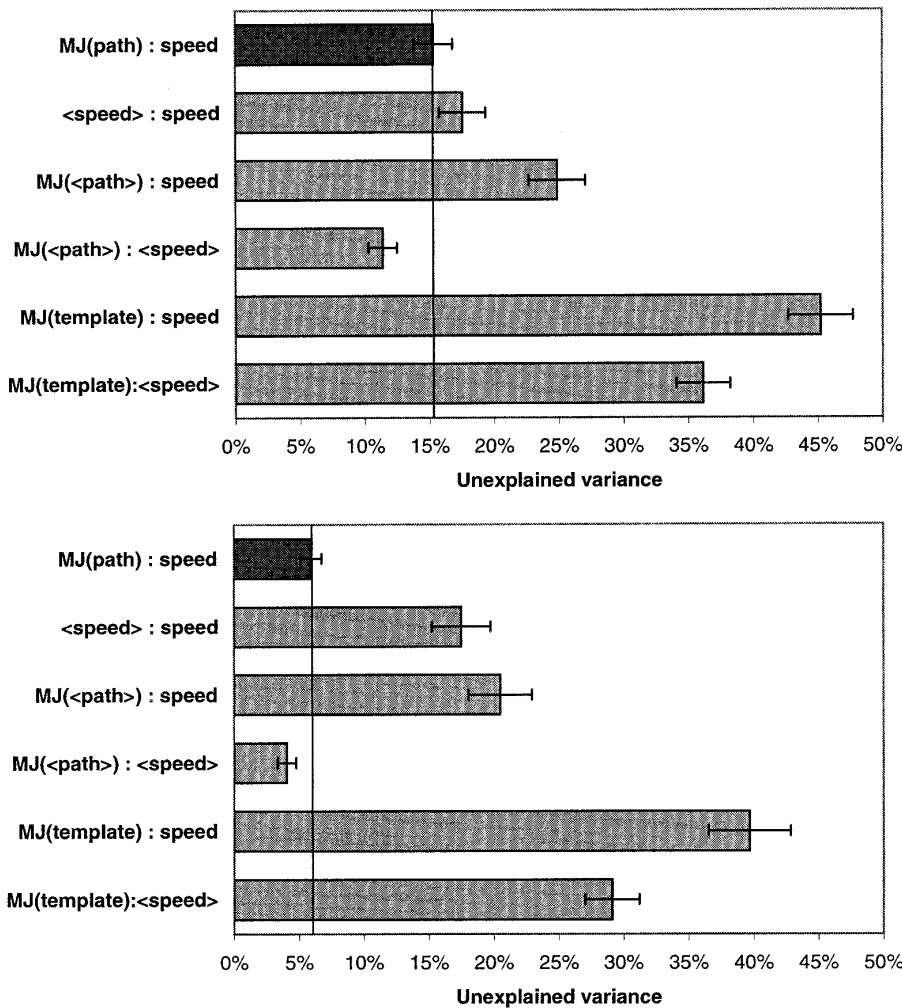

FIG. 13. Comparison of prediction errors of our model (on all 1.5-s movements in *experiment 2*) and possible reconstructions of a "central plan." $MJ(\text{path})$: speed—our model, applied on a trial-by-trial basis (■). $\langle\text{speed}\rangle$: speed—the average speed profile over a block of 10 trials, used to predict each individual speed profile. $MJ(\langle\text{path}\rangle)$: speed—the speed profile generated by our model for the average path and used to predict each individual speed profile. $MJ(\langle\text{path}\rangle)$: $\langle\text{speed}\rangle$—the speed profile generated by our model for the average path and used to predict the average speed profile. $MJ(\text{template})$: $\langle\text{speed}\rangle$—the speed profile generated by our model for the specified template and used to predict the average speed profile. $MJ(\text{template})$: speed—the speed profile generated by our model for the specified template, and used to predict each individual speed profile.

to be more accurate than both average methods: $MJ(\text{path})$: speed $< MJ(\langle\text{path}\rangle)$: speed $< MJ(;\langle\text{path}\rangle)$: $<\text{speed}>$.

Because this analysis is most likely to yield clear results when the model works best, we focused on the data from all 1.5-s movements in *experiment 2*. As Fig. 13 shows, none of the above orderings hold. It is true that the constrained minimum-jerk model describes the average trajectory better than individual trajectories: $MJ(\text{path})$: speed $> MJ(\langle\text{path}\rangle)$: $<\text{speed}>$, therefore smoothness is present in the plan and adding variability decreases the planned smoothness. However, the speed profile predicted from the average path explains the individual speed profiles a lot worse than the single trial predictions: $MJ(\text{path})$: speed $< MJ(\langle\text{path}\rangle)$: speed. In fact, even the average speed profile (which is the best possible constant predictor) performs worse than the single trial prediction method: $MJ(\text{path})$: speed $< <\text{speed}>$: speed. Therefore, the trial-to-trial variability observed in speed profiles is partially predictable given the trial-to-trial variability in the path. Note that speed predictions based on the template rather than the average path are a lot less accurate, suggesting that the spatial deviations from the specified template may be present in the plan.

In summary, the picture that emerges from this analysis is the following: the relationship between speed and path is present in the plan and is stronger than that observed on individual trials. However, the trial-to-trial variability is not at all ''random''; the path and the speed profile represented in the central plan are modified together in a way that preserves the planned smoothness. It is of course possible that a strict separation between planning and execution does not exist, and thus the above analysis yields mixed results.

GENERAL DISCUSSION

This paper has presented a constrained minimum-jerk model of the relationship between hand path and speed profile of complex arm movements. We applied the model to a wide range of spatially constrained motor tasks involving discrete movements of short durations and arbitrary paths. The new model proposed here unifies two sets of experimental observations: the characteristic bell-shaped speed profile of straight reaching movements, and the path-speed relationship in extemporaneous movements with complex paths. The speed predictions of our model are equivalent to those of the original minimum-jerk model when the path is straight and more accurate when the Flash and Hogan (1985) model fails in the path prediction. Although our model uses the entire path (i.e., a continuous curve) to predict speed, it is possible instead to represent the path as a large number of intermediate points and possibly tangents at those points — if such a representation contains sufficient information to reconstruct the continuous path through some interpolation method, this becomes equivalent to our model. In practice, for the family of templates studied here, we have found that the entire path can be reconstructed from ~10 points sampled at equal distances. This does not imply any internal representation consisting of 10 discrete points — it is simply a consequence of the mathematical fact that curves as smooth as the ones used here can be reconstructed easily through interpolation.

We showed that although mathematically the constrained minimum-jerk model yields predictions that are generally similar to the power law, its performance is significantly better than both the original and modified versions of the power law for all tasks and movement templates we studied. Furthermore the model naturally extends the path-speed relationship previously observed in extemporaneous movements to include the acceleration and deceleration phases of discrete movements, which are much more common. Thus it can be argued that the local speed-curvature relationship expressed by the power law, to the extent that it fits experimental data, can be derived from a more global smoothness constraint that relates the path to the speed profile.

In comparing any two models, we have to look not only at their competence in accounting for experimental data, but also at the model complexity (Occam's razor). We emphasize that model complexity has nothing to do with the complexity of the mathematics involved: instead it is related to how many and what free parameters are used in fitting the data. In that sense, the constrained minimum-jerk model and the original power law are equivalent — they both postulate that the speed profile of a movement is a well defined function of hand path, and do not require any extra parameters. The modified power law has higher model complexity because it involves a free parameter (β) extracted from the observed trajectories. Furthermore, all versions of the power law require extra mechanisms for dealing with regions of zero curvature and with zero speed at the end points of a discrete movement. Thus we argue that the statistical complexity of the model presented here, when applied to discrete movements with inflection points or straight segments, is no greater than any version of the 2/3 power law. We also have to compare the difficulties involved in the possible biological implementation of each model. From that point of view, the power law may appear more appealing because it is analytically tractable (i.e., speed can be obtained directly from curvature, without any numerical approximation). If we believe that the biological system actually uses the concrete mathematical formula, this is certainly a big advantage. However, we are dealing with models that only quantify intrinsic relationships between path and speed instead of specifying a recipe for trajectory formation. For example, it has been suggested that a 2/3 power law may emerge out of the complex nonlinear interactions of a large number of direction-tuned neurons in primary motor cortex (Lukashin and Georgopoulos 1993) — such an implementation does not use the analytical tractability of the power law, eliminating any advantage it may have over the model presented here.

Contrary to previous investigations (Viviani and Cenzato 1985) we did not find evidence for segmentation; instead the ''tempo'' seemed to fluctuate rather smoothly and unpredictably. We think this is due to the different nature of the movements studied here. Previous work on segmentation has focused on repetitive movements of long duration, which implied (at least visually) a rather obvious segmentation pattern. Instead we used movement durations of 1–2 s with templates that are not obviously composed of a small number of simple curves concatenated together. It is possible that discrete movements of such short duration are not internally segmented even when they have complex paths. Alternatively, the decomposition into segments may be variable

from trial to trial, preventing us from finding any evidence for it.

The high accuracy of the constrained minimum-jerk model allowed us to make two important observations that have not received previous attention: the relationship between path and speed is stronger for movements of shorter duration and is not affected by spatial scale or average speed (it is also rather uniform across tasks and movement paths) and the path of the hand on a particular trial contains significant information about the speed profile on that same trial; this information is lost if we only study the average path over a block of identical trials.

Presently the model is applied to the complete path, yielding a complete speed profile. It is unlikely that the motor system maintains the path-speed relationship globally, especially over movements of very long duration. Our finding that prediction error increases with movement duration (and not average speed or spatial scale) suggests that the ''sliding window'' over which the model applies best may be rather small, i.e., ~ 1 s. A more local relationship between path and speed is also consistent with observations that the 2/3 power law applies in tracking tasks where little advance planning of maximally smooth movements is possible. Thus it is desirable to modify the procedure for fitting the model so that it can be applied to segments of the path, without requiring end-point speed and acceleration (only duration for that segment). This can be done by minimizing jerk over these parameters as well.

Our examination of trial-to-trial variability suggests that smoothness is centrally planned, and the variability of the hand path and the speed profile are coupled in a way that maintains the planned path-speed relationship. Because the movements we studied are not executed at maximum speed, one likely source of variability is the effect of on-line corrections. Such on-line corrections, however, have to be incorporated smoothly in the movement in a manner that is consistent with the overall path-speed relationship described by the constrained minimum-jerk model. Evidence for such error correction mechanisms is provided by the double-step paradigm, where a reaching movement is smoothly corrected to end at a perturbed target location if subjects are unaware of the perturbation (Pellison et al. 1986). It is also possible that low-level correction/stability mechanisms (e.g., some version of equilibrium point control) combined with a source of noise may give rise to variations around the desired trajectory that preserve its intrinsic properties. The desired trajectory itself does not have to be generated by a servo control mechanism, but only the low-level on-line corrections to it.

Finally, we cannot rule out the possibility that the trial-to-trial variability we studied actually was planned. It is quite difficult to identify the levels of processing on which different regularities emerge based on observational studies alone. The problem is that once a movement is executed, it is impossible to determine whether the subject had planned to produce exactly that movement or the outcome was significantly affected by on-line processes. Perturbative experimental manipulations will be required to address these issues directly.

APPENDIX. CONSTRAINED MINIMIZATION

Minimizing the integral of $L = (\kappa' s^3 + 3\kappa s \ddot{s})^2 + (\dddot{s} - \kappa^2 s^3)^2 + (s^3 \kappa \tau)^2$ with respect to the temporal profile $s(t)$ is a variational problem, which can be approached using a number of techniques. It turns out that the standard techniques for solving variational problems do not work very well in this case.

From the Euler-Lagrange principle, the solution $s(t)$ satisfies the 6th-order nonlinear ordinary differential equation $\partial L / \partial s - d/dt\ \partial L / \partial s' + d^2/dt^2\ \partial L / \partial s'' - d^3/dt^3\ \partial L / \partial s''' = 0$, subject to the constraints on $s(t)$ and its derivatives at the two end points of the movement. This problem is very difficult numerically because the solution depends on high order derivatives of the path, which is recorded experimentally and thus has some noise in it. We found that both shooting and relaxation methods for solving boundary value problems (Press et al. 1992) fail.

Another possibility is to minimize *Eq. 1* directly by representing $s(t)$ as a linear combination of N fixed basis functions $f_i(t)$: $s(t) = \sum_{i=1}^{N} c_i f_i(t)$ and using gradient descent with respect to the coefficients c_i. We used sixth-order B splines as basis functions and a preconditioned conjugate gradient method to find the optimal coefficients (Gill et al. 1981). This minimization always converged but frequently found local minima, and it was necessary to restart it a number of times before a good solution could be found. An additional problem is that it is not clear what a good solution is.

The method we used here is less direct but has the advantage that it is faster and seems to converge to the same minimum regardless of the initial conditions. Rather than using the complete path $\mathbf{r}(s)$, we chose a set of 10 intermediate points, equally spaced along the path (i.e., their positions contain no temporal information). We then found the minimum-jerk movement (in the sense of Flash and Hogan 1985) through those points, given the velocity and acceleration at the two end points. If the resulting minimum-jerk movement has a path very close to $\mathbf{r}(s)$, we are guaranteed that its speed profile is the solution to our original problem (i.e., we are computing a minimum over a given set by minimizing over a superset and ensuring that the solution is in the original set). We describe the minimization procedure in the following text.

It has been shown (Flash and Hogan 1985) that for given passage times *T,* positions *x,* velocities ν, and accelerations *a* at the end points of one segment, the minimum-jerk trajectory is a 5th-order polynomial in *t,* the coefficients of which can be determined easily using the end-point constraints. It is then possible to integrate the squared jerk analytically, and sum it over all segments. The jerk of the optimal trajectory is

$$J_{T,x,\nu,a}(T_{1\ldots N}, x_{1\ldots N}, \nu_{1\ldots N}, a_{1\ldots N}) = \sum_{i=1}^{N-1}$$

$$3\ \frac{3a_i^2 - 2a_i a_{i+1} + 3a_i^2}{(T_{i+1} - T_i)^4} + 24\ \frac{3a_{i+1}\nu_i - 2a_{i+1}\nu_i + 2a_i\nu_{i+1} - 3a_i\nu_{i+1}}{(T_{i+1} - T_i)^3}$$

$$+ 48\ \frac{4\nu_i^2 + 7\nu_i\nu_{i+1} + 4\nu_{i+1}^2}{(T_{i+1} - T_i)^2} + 120\ \frac{(a_i - a_{i+1})(x_i - x_{i+1})}{(T_{i+1} - T_i)^2}$$

$$+ 720\ \frac{(\nu_i + \nu_{i+1})(x_i - x_{i+1})}{T_{i+1} - T_i} + 720\ (x_i - x_{i+1})^2$$

Note that the above expression is quadratic in all ν and *a* and thus can be minimized with respect to those parameters by setting the gradient to 0 and solving (numerically) the resulting system of linear equations. Using this minimization, for given passage times *T* (and intermediate point positions *x*), we can compute the jerk for the optimal trajectory: $J_{T_{1\ldots N}, x_{1\ldots N}} = \sum_{x=X,Y,Z} \min_{\nu_{1\ldots N}, a_{1\ldots N}} J_{T,x,\nu,a}(T_{1\ldots N}, x_{1\ldots N}, \nu_{1\ldots N}, a_{1\ldots N})$. Finally, the function $J_{T_{1\ldots N}, x_{1\ldots N}}$ is minimized over all possible $T_{2\ldots N-1}$ using a nonlinear simplex method. We found that the resulting path is equivalent to $\mathbf{r}(s)$, and the speed profiles obtained in this way were indistinguishable from the speed profiles obtained with the direct method above, when it was run enough times to find a global minimum.

REFERENCES

ATKESON, C. G. AND HOLLERBACH, J. M. Kinematic features of unrestrained vertical arm movements. *J. Neurosci.* 5: 2318–2330, 1985.

BENNETT, D. J., HOLLERBACH, J. M., XU, Y., AND HUNTER, I. W. Time-varying stiffness of human elbow joint during cyclic voluntary movement. *Exp. Brain Res.* 88: 433–442, 1992.

BIZZI, E., HOGAN, N., MUSSA-IVALDI, F. A., AND GISZTER, S. Does the nervous system use equilibrium-point control to guide single and multijoint movements. *Behav. Brain Sci.* 15: 603–613, 1992.

DE BOOR, C. *A Practical Guide to Splines.* New York: Springer-Verlag, 1978.

FLASH, T. AND HOGAN, N. The coordination of arm movements: an experimentally confirmed mathematical model. *J. Neurosci.* 5: 1688–1703, 1985.

GILL, D., MURRAY, W., AND WRIGHT, M. *Practical Optimization.* New York: Academic, 1981.

GOMI, H. AND KAWATO, M. Equilibrium-point control hypothesis examined by measured arm stiffness during multijoint movement. *Science* 272: 117–120, 1996.

GRIBBLE, P. AND OSTRY, D. Origins of the power law relationship between movement velocity and curvature: modeling the effects of muscle mechanics and limb dynamics. *J. Neurophysiol.* 76: 2853–2860, 1996.

GUSIS, A. *Arm Trajectory Planning in 3D Space* (Masters thesis). Rehovot, Israel: Weizmann Institute of Science, 1995.

HOGAN, N. AND FLASH, T. Moving gracefully: quantitative theories of motor coordination. *Trends Neurosci.* 10: 170–174, 1987.

JACK, W. R. On the analysis of voluntary muscular movements by certain new instruments. *J. Anat. Physiol.* 29: 473–478, 1895.

LACQUANITI, F., TERZUOLO, C., AND VIVIANI, P. The law relating the kinematic and figural aspects of drawing movements. *Acta Psychol.* 54: 115–130, 1983.

LUKASHIN, A. V. AND GEORGOPOULOS, A. P. A dynamical neural network model for motor cortical activity during movement: population coding of movement trajectories. *Biol. Cybern.* 69: 517–524, 1993.

MORASSO, P. Spatial control of arm movements. *Exp Brain Res.* 42: 223–227, 1981.

MORASSO, P. Three dimensional arm trajectories. *Biol. Cybern.* 48: 187–194, 1983

MORASSO, P. AND MUSSA IVALDI, F. A. Trajectory formation and handwriting: a computational model. *Biol. Cybern.* 45: 131–142, 1982.

NELSON, W. L. Physical principles for economies of skilled movements. *Biol. Cybern.* 46: 135–147, 1983.

PELLISON, D., PRABLANC, C., GOODALE, M. A., AND JEANNEROD, M. Visual control of reaching movements without vision of the limb. II. Evidence of fast unconscious processes correcting the trajectory of the hand to the final position of a double-step stimulus. *Exp. Brain Res.* 62: 303–311, 1986.

POLLICK, F. E. AND SAPIRO, G. Constant affine velocity predicts the 1/3 power law of planar motion perception and generation. *Vision Res.* 37: 347–353, 1997.

PRAKASH, N. *Differential Geometry. An Integrated Approach.* New York: McGraw-Hill, 1981.

PRESS, W., TEUKOLSKY, S., VETTERLINC, W., AND FLANNERY, B. *Numerical Recipes in C.* New York: Cambridge Univ. Press, 1992.

SABES, P. AND JORDAN, M. Obstacle avoidance and a perturbation sensitivity model for motor planning. *J. Neurosci.* 17: 7119–7128, 1997.

SOECHTING, J. F. AND TERZUOLO, C. A. An algorithm for the generation of curvilinear wrist motion in an arbitrary plane in three dimensional space. *Neuroscience* 19: 1393–1405, 1986.

SOECHTING, J. F. AND TERZUOLO, C. A. Organization of arm movements in three-dimensional space. Wrist motion is piecewise planar. *Neuroscience* 23: 53–61, 1987.

UNO, Y., KAWATO, M., AND SUZUKI, R. Formation and control of optimal trajectory in human multijoint arm movement. *Biol. Cybern.* 61: 89–101, 1989.

VIVIANI, P. AND CENZATO, M. Segmentation and coupling in complex movements. *J. Exp. Psychol. Hum. Percept. Perform.* 11: 828–845, 1985.

VIVIANI, P. AND FLASH, T. Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. *J. Exp. Psychol. Hum. Percept. Perform.* 21: 2–53, 1995.

VIVIANI, P. AND SCHNEIDER, R. A developmental study of the relationship between geometry and kinematics of drawing movements. *J. Exp. Psychol. Hum. Percept. Perform.* 17: 198–218, 1991.

WANN, J., NIMMO-SMITH, I., AND WING, A. M. Relationship between velocity and curvature in movement: equivalence and divergence between a power law and a minimum-jerk model. *J. Exp. Psychol. Hum. Percept. Perform.* 14: 622–637, 1988.

WOLPERT, D., GHAHRAMANI, Z., AND JORDAN, M. I. Are arm trajectories planned in kinematic or dynamic coordinates? An adaptation study. *Exp. Brain Res.* 103: 460–470, 1995.

YALOV, S. *Computational Models of Hand Trajectory Planning in Drawing Movements* (Masters thesis). Rehovot, Israel: Weizmann Institute of Science, 1991.