Abstract

The velocity profiles of arm movements display a number of regularities, including bell-shaped velocity profiles and an inverse relationship between velocity and curvature in extemporaneous drawing movements (described as a 2/3 power law by Lacquaniti et al., 1983). Here we propose a modification of the minimum-jerk model (Flash and Hogan, 1985) that simultaneously accounts for both regularities. For a given path of the hand in space, our model predicts that the velocity profile will be the one that maximizes smoothness. Analysis reveals that in both two and three dimensions, the 2/3 power law is mathematically equivalent to setting the jerk along the normal vector to zero; the power law generates velocity 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 modified minimum-jerk model provides a very close fit to the observed velocity profiles in all cases. Its performance is uniformly better compared to the 2/3 power law, suggesting that the correlation between velocity and log 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 velocity profile of a complex arm movement is stronger than previously thought, and extends over a wider range of motor tasks. The implications of these findings for motor planning and online control are discussed.

The majority of human arm movements produced in everyday life are aimed at achieving a certain effect on the external environment. Since the same effect can typically be achieved in a number of ways, the exact movement executed at a particular time is not completely determined by the motor task. Consider the example of reaching. All that is specified in this task is the final location of the finger tip at the end of the movement; however, the same finger position can be obtained for a large number of arm configurations, and the hand can take many different paths to the desired location. Yet we can 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 velocity 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), and piecewise-planar $3D$ hand paths (Soechting and Terzuolo, 1987). In this work we focus on regularities in hand paths and velocity profiles in arm movements that are spatially constrained. Previous experimental observations particularly relevant in this context are that reaching movements have straight paths in Cartesian space and characteristic bell-shaped velocity profiles (Morasso, 1981); and that the instantaneous tangential velocity of curved movements is inversely related to the curvature of the hand path, even when subjects attempt to maintain constant velocity (Jack, 1895; Lacquaniti et al., 1983).

A general framework for understanding the regularities observed in underconstrained movements 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. This approach is appealing since it can provide not only a quantitative fit to data, but also suggest a computational mechanism for how the motor system actually arrives at the class of movements that display a particular regularity (i.e. by minimizing a functional, subject to the task constraints). Moreover, the optimization framework can provide hints as to why a class of movements might be preferred. A number of models based on optimization theory have been considered (Nelson, 1983; Uno et al., 1989; Wann et al., 1988; Flash and Hogan, 1985), the cost being energy, force, impulse, peak velocity/acceleration, duration, or torque change. The model which 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), in which the cost is defined as the squared jerk (third derivative of hand position over time), integrated over the entire movement. In reaching tasks the minimum jerk model predicts straight paths and bell-shaped velocity profiles that closely match experimental observation; the model predictions are equally accurate in tasks involving one intermediate (via) point along a curved path. Since this model is purely kinematic, arm dynamics cannot be taken into account, i.e. it predicts kinematic invariance. However, it has been shown that arm dynamics actually affects the hand path in via-point (Uno et al., 1989), as well as obstacle avoidance (Sabes, 1996) tasks. Furthermore, the minimum-jerk model in its current form is restricted to reaching and via-point movements, and it is not clear how to apply it in situations in which the spatial constraints are specified differently.

The class of motor tasks we consider involve movements with hand paths that are constrained spatially. It has been proposed (Lacquaniti et al., 1983) that the velocity $v(t)$ of such movements is related to the curvature $\kappa(t)$ through a power law: $v(t) = \gamma.\kappa(t)^{-1/3}$, where γ is a constant gain factor. The validity of this law is typically demonstrated through a correlation between v and $\kappa^{-1/3}$, or between $\log(v)$ and $\log(\kappa)$. While 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 velocity profiles are captured, apart from the general trend that velocity decreases with increasing curvature. It has been argued, for example, that the actual velocity profiles can be skewed relative to the power law prediction, and the correlation between $\log(v)$ 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 velocity profiles. Finally, in its original form the power law has two limitations: a) it does not apply to paths that have inflection points or straight segments (i.e. for $\kappa = 0$ the predicted velocity is infinite), and it is inaccurate at low curvature; b) the fact that $v(t)$ decreases towards 0 at the endpoints of discrete movements cannot be captured. To fix the first problem, Viviani and Schneider (1991) suggested a modified power law: $v(t) = \gamma.(\kappa(t) + \epsilon)^{\beta}$, which has free parameters that need to be determined.

Despite these limitations, the power law seems to provide a surprisingly good description of the velocity profiles of complex movements, which has motivated 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 velocity, the hand traces a similar ellipse obeying the power law. The problem with this explanation is the same problem that faces the equilibrium-point hypothesis, i.e. the required stiffness is much larger than what has been measured experimentally under dynamic conditions (Bennett et al., 1992; Gomi and Kawato, 1996). Another proposed explanation (Pollick and Sapiro, 1996) is that a point moving according to the power law maintains constant “affine velocity”. While this fact is interesting mathematically, it is not obvious how to relate affine velocity to arm movements, and why it might be advantageous to keep it constant. An alternative suggestion, that we will consider in the current paper, is that maximally smooth (i.e. minimum-jerk) movements have velocity profiles described by the power law (Wann et al., 1988; Viviani and Flash, 1995). While the minimum-jerk model generally predicts lower velocities at high curvature, this relationship has not been clearly demonstrated, or explained.

Here we propose a modification of the minimum-jerk principle that makes it applicable to movements with arbitrary paths. For an observed path of the hand in space, we compute the velocity profile along the path that minimizes jerk. This modified principle, like the power law, can no longer make predictions about the path of the movement, but given the actual path for an individual trial it can generate a velocity profile that can be compared to the observed velocity profile. We analyze the relationship between the new model and the power law, and outline the reasons why the two generate similar velocity predictions for a given path. Next, we present a series of experiments demonstrating that our model predicts the velocity profiles of complex arm movements uniformly better than the power law. Finally, we study the movement characteristics that affect the accuracy of the model, and consider the implications of our results.

A constrained minimum jerk model

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, and the tangential velocity 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 predicts that the temporal profile of the movement is the scalar function $s(t)$ that minimizes:

$$J = \int_0^T \left\| \frac{d^3}{dt^3}\mathbf{r}(s(t)) \right\|^2 dt \tag{1}$$

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 velocity profiles – Flash and Hogan (1985) computed both the path and the velocity given a set of via points. Let us denote the term inside the integral by L:

$$L = \left\| \frac{d^3}{dt^3}\mathbf{r}(s) \right\|^2 = \left\| \frac{d^2}{dt^2}\mathbf{r}'(s)\dot{s} \right\|^2 = \left\| \frac{d}{dt}(\mathbf{r}''(s)\dot{s}^2 + \mathbf{r}'(s)\ddot{s}) \right\|^2$$

$$L = \left\| \mathbf{r}'''(s)\dot{s}^3 + 3\mathbf{r}''(s)\dot{s}\ddot{s} + \mathbf{r}'(s)s^{(3)} \right\|^2 \tag{2}$$

This minimization problem is invariant to rotations and translation 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 (2). However, the derivatives of $\mathbf{r}$ can be expressed as functions of the curvature and torsion. The path $\mathbf{r}$ satisfies the system of differential equations:

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

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 (2):

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

and since the tangent, normal, and binormal are orthogonal,

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

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 (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 (3). We now substitute in (3):

$$L = \left\| \mathbf{n}(\kappa'\dot{s}^3 + 3\kappa\dot{s}\ddot{s}) + \mathbf{t}(s^{(3)} - \kappa^2\dot{s}^3) + \mathbf{b}(\dot{s}^3\kappa\tau) \right\|^2 \tag{5}$$

$$L = \left\| \mathbf{n}(A_s') + \mathbf{t}(s^{(3)} - A_s\kappa) + \mathbf{b}(A_s\tau) \right\|^2$$

$$L = A_s'^2 + (s^{(3)} - A_s\kappa)^2 + A_s^2\tau^2$$

¿From the definition of A_s, we have:

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

The power law states that $A_s^{1/3} = const$, 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 $2D$ curves $\mathbf{r}(t) = (x(t), y(t))$, this can be shown more directly. The power law $v(t) = const.\kappa(t)^{-1/3}$ in $2D$ can be written as:

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

$$x'y'' - y'x'' = const$$

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.

We want to know whether the velocity profile given by the power law is close to the solution of the jerk minimization problem (1). In figure 1 we have plotted three synthetic paths, for which we have computed numerically (see below) the minimum jerk (solid) and the power law (dashed) velocity profiles. 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 velocity 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, however: in example B where the curvature is almost constant the power law predicts a velocity profile which 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 velocity only using the local curvature, comes close to solving the global minimization problem (1)? One intuitive answer is the following: since the total jerk is a sum of two (in $3D$ of three) nonnegative terms, and the power law minimizes one of them (normal jerk is set to zero), it is an efficient approximation to the solution of (1). It seems however that this is not always the complete answer. In panel D of figure 1 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 velocity profile increase both terms, meaning that in this case the optimal velocity 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 velocity 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.

Minimization algorithm

Using (4), we can apply variational calculus and solve for the function $s(t)$ that minimizes the constrained jerk. It turns out that 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 (Press et al., 1992) for solving the resulting nonlinear differential equation fail. The alternative is to minimize (1) directly, by representing $s(t)$ as a linear combination of fixed basis functions, and using gradient descent with respect to the coefficients of the basis functions. We used sixth-order B-splines as basis functions, and a preconditioned conjugate gradient method to find the optimal coefficients (Gill and Murray, 1981). This minimization always converged, but frequently found local minima, and it was necessary to restart it many times before a good solution could be found. An additional problem is that it is not clear what a good solution is.

An alternative 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 velocity 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). To perform this minimization, we used a nonlinear simplex method (in Matlab) to find the passage times through the specified via points. Given those passage times, one can obtain analytically the complete movement, compute its jerk, and pass it as an error function to the minimization algorithm. We found that the resulting path is equivalent to $\mathbf{r}(s)$, and the velocity profiles obtained in this way were indistinguishable from the velocity profiles obtained with the direct method above, when it was run enough times to find a global minimum.

Experiment 1

Methods

To test our model, and compare it to the power law, we studied 4 tasks in which a path was specified and subjects were asked to execute an arm movement along that path. To avoid drawing conclusions from an isolated movement, we used several paths in each task. The same 9 right handed subjects participated in all 4 tasks. Short breaks were given between tasks. The entire experiment lasted about 45 minutes. End-point position was recorded at 100 Hz with an Optotrak 3020 infrared system connected to an IBM PC compatible computer.

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 figure 2 printed on it. Subjects held a pen in their right hand; the task was to position the pen 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 velocity 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 horizontal table that was positioned at shoulder height. We used a VGA projector to display computer generated images on a tilted screen, which the subject 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 an LED attached to it that provided endpoint 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 ten 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 velocity. 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 through 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, and finished at the end point. The configurations we used are shown in figure 3. There were a total of 6 configurations, every other configuration being a 90 degree rotation of the preceding one. Subjects executed 10 trials in each block. As before, movement velocity 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 approximately 10x10x10 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, since 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 velocity prior to movement onset. Each wireframe model was used in 10 trials, there were a total of three models. Movement velocity 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 prespecified 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 the cubic spline representation of the data for each trial, all subsequent steps (i.e. extracting velocity and curvature) were applied directly, without any further filtering.

We then computed a velocity profile prediction from the original power law, the modified power law, and the constrained minimum jerk method (see below). To compute the difference between each prediction and the actual velocity profile, we expressed all velocity profiles as a function of arc length s (which resulted in better alignment), and scaled them so that the average velocity was one and the path lengths were equal. The measure we used was the mean absolute deviation between the scaled velocity profiles, expressed in percent:

$$Err(v, v^*) = 100 \left\langle \left| \frac{v(s)}{\langle v(s) \rangle} - \frac{v^*(s)}{\langle v^*(s) \rangle} \right| \right\rangle$$

where $\langle . \rangle$ signifies averaging over s. The same procedure was repeated over the entire movement, and over the middle 60% of the movement. The latter was done in order to assess the accuracy of the power law away from the end points, where we already know it is inaccurate (see figure 1). We examined a number of velocity profiles to ensure that the 20% of the movement eliminated at each end was enough to cover the parts of the velocity profile that seemed affected by acceleration and deceleration.

The prediction for the original power law was obtained directly from the equation $v(s) = \gamma.\kappa(s)^{-1/3}$. For paths with inflection points, we added a small positive constant to the curvature, that 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 velocity profile. The modified power law $v(s) = \gamma.(\kappa(s) + \epsilon)^{\beta}$ has two free parameters (ϵ, β). We optimized them (using the nonlinear simplex method) separately for each trial, so that the error defined above was minimized.

The constrained minimum jerk model was implemented as discussed in the previous section. In its present form, the minimization requires the velocity and acceleration at the end points (in principle this could be avoided, by minimizing over those parameters as well). When applied to the entire movement, the model has no free parameters: the velocities and accelerations at the end points are zero, which are known a priori, and not extracted from the data we are trying to predict. When applied to the middle 60% of the movement, the model used the actual endpoint velocity and acceleration measured experimentally. Note however that we are predicting rather complex velocity profiles, thus the two end points of a submovement contain relatively very little information. In contrast, the modified power law uses the entire velocity profile to extract the values of its free parameters.

Results and discussion

In the first two tasks, for which the complete path was specified and the movement was superimposed on it, subjects were able to follow the specified paths accurately. Figure 4 shows an example of a subject tracing template 1 in task 2. Note that the trial-to-trial variability in the path is much smaller than it is in the (scaled) velocity profile. Also, the velocity profiles differ significantly in shape, in contrast with reaching tasks where it is possible to align them almost perfectly by scaling.

In the $3D$ task, the majority of subjects were not able to 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 velocity profile given the actual, not the desired path of the hand, so we can still apply them and assess the accuracy of their predictions.

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 five out of the ten trials (dotted) that the subject made, on the original and the rotated configuration of via-points. The solid line is the minimum jerk path, which remains invariant to rotation. Note that the deviations are not just isolated trials, but stable strategies utilized by the subjects; also, the actual paths depended on the orientation of the template, contrary to the model prediction. While some of these deviations may be the result of arm dynamics affecting the path, this factor is unlikely to account for the large differences (i.e. 5A) between the paths that different subjects adopted for the same via-point configuration. A more likely explanation is that multiple via-point tasks 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 have already planned a smooth path that passes through the via-points in the specified order. Whatever the reason for these discrepancies, we would conclude that the original minimum jerk model is accurate on the average, but cannot account for the significant between subject variability on multiple via-point tasks.

We now turn to the comparison of our model and the power law. Figure 6 shows summary error statistics, for all tasks, over the complete (top) and middle 60% of the movement (bottom). 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 7 out of the 8 comparisons (the $3D$ task/complete movements being the exception) all 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 in task 2 (complete movements). Figure 7 shows that the ordering holds for all templates we tested (and not just the ones that include inflection points). It is interesting that the constrained minimum jerk model shows a 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.

A comparison of prediction error over subjects revealed a strong correlation between how rapidly subjects moved on the average, and how well the constrained minimum jerk model predicted their velocity profiles (a much weaker correlation was observed for the modified power law). Figure 8 shows a scatterplot of mean movement duration (over all templates) and mean prediction error for our model in tasks 2 and 3 (the correlations are $r = 0.76, p < 0.05$ and $r = 0.73, p < 0.05$ respectively). There are several possible interpretations of this result. It may be that the "tempo" of arm movements fluctuates over time, thus movements that take longer have less predictable velocity profiles. It is also possible that at high average velocities, movements are for some reason more regular. Since we did not vary the size of the templates systematically, duration was correlated with speed. Another possibility is that the effect is entirely due to between-subject variability, i.e. some subjects naturally produce more predictable velocity 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.

Experiment 2

To eliminate the possibility that between-subject variability is responsible 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 velocity (and size). A new group of 8 subjects was recruited for this experiment.

Methods

The setup was the same as in tasks 2 and 3 above - subjects were asked to make arm movements on a horizontal table, by tracing a continuous curve projected on the table. We used a subset of 5 templates, each shown in 3 blocks of 10 identical trials. In block 1, the template was big (same size as in experiment 1), and desired movement duration was 3 seconds. In block 2, the template was also big, and duration was 1.5 seconds. In block 3, the template was scaled down by a factor of 2, and desired duration was 1.5 seconds. Thus, blocks 1 and 2 were matched for size, blocks 2 and 3 for movement duration, and blocks 1 and 3 for average velocity (assuming of course that subjects produced movements of the specified durations, which they did). 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 duration was shown (in units of 1/100 seconds). 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 was analyzed in exactly the same way as in experiment 1. All subjects traced the paths accurately, and were able to adjust the duration of their movements in the first 2-3 trials in each block. Prediction error statistics, averaged over templates, are shown in figure 9. Again, we found that the constrained minimum jerk method predicts velocity 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 two times better than the modified power law, and four times better than the original power law. The difference is still present when velocity profiles are computed for the middle 60% of the movements.

It can be seen in figure 9 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 only for the constrained minimum jerk predictions; for both complete movements and middle 60%, the prediction error for the large 3 seconds movement was larger than the error for both the large 1.5 seconds, and the small 1.5 seconds movement (all significant at the $p < 0.01$ level). This result indicates that the accuracy of the prediction depends on the duration, not the average velocity or spatial scale of the movement. We can also conclude that the difference in experiment 1 was not entirely due to systematic between-subject variability, since the second experiment yielded the result within subjects. Note that the difference is much smaller for the power law predictions, and a trend can only be seen when the power law is applied to the middle 60% of the movement.

Analysis of velocity profiles

Thus far we have considered statistics based on the average prediction errors. In this section we perform a more detailed analysis of the actual velocity profiles, and the trial-to-trial variability mentioned earlier. We focus on the data from experiment 2, in particular the small 1.5 second movements.

Figure 10 shows velocity profiles of all 8 subjects for 4 of the 5 templates used in experiment 2 (numbers corresponding to the template positions in figure 2). We have scaled the last 5 velocity profiles (complete movements) in each block, and averaged them. Solid lines are the actual velocity profiles, the dashed lines are the constrained minimum jerk predictions, and the dotted lines are the modified power law predictions. Since scaling does not align the velocity 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 prediction for the average trial (which would be meaningless since 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 velocity profiles it predicts in template 2 correspond to a region of constant curvature (as was noted in figure 1). We conclude that the power law captures very well the maxima and minima of the actual velocity profiles, but otherwise it is not a satisfactory model of the details present in the experimental data. While some of the errors are due to skewing (as Wann et al., 1988 argued), the predominant error is in the predicted values at the velocity 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. For example, in template 2, subjects 1, 5, and 6 show the second velocity peak earlier than predicted, while the other subjects perform in agreement with the model. Still, there is some systematic error left. The first velocity maximum on template 1, and the second maximum on template 3 are both slightly underestimated for all subjects. Overall, the predictions of the model are rather close to the experimentally observed velocity 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, since the averaging procedure is likely to obscure some details. For each template, 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 (solid), modified power law (dotted), and constrained minimum jerk (dashed) velocity profiles for the selected trials (middle 60%) are shown in figure 11. 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 figure 10) - 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 panel 2, which correspond to a path with strictly positive curvature.

We now consider the trial-to-trial variability in the movements executed by one subject in one block of 10 trials. The question we ask is, does the constrained minimum jerk model extract any information regarding the velocity profile from the path traced on an individual trial, or is the trial-to-trial variability in path and velocity uncorrelated (at least from the point of view of the model). To answer that question, we have to estimate the variability of velocity profiles, and compare it to the prediction error of the model. To this end we define a new prediction method, which is simply the average velocity profile over a block of 10 trials. Figure 12 shows that the constrained minimum jerk model reduces the trial-to-trial variability by about 50%, therefore the path for an individual trial contains a significant amount of information about the velocity profile on that trial, which will be lost if we look only at the average path. This is particularly interesting since, as shown in figure 4, paths are much less variable than velocity profiles.

Another issue that can be addressed by analyzing these data is movement segmentation. Viviani and Cenzato (1985) suggested that in cases where the gain factor in $v(t) = \gamma(t).\kappa(t)^{-1/3}$ is not constant, it is piecewise constant (this was demonstrated for a pattern of two nested ellipses). They argued that the positions along the movement where the gain factor changes correspond to segmentation boundaries. We computed the gain factor for a number of discrete movements, and did not find any case in which it was convincingly piecewise constant. Figure 13 shows a typical plot—five consecutive trials for one subject on template 2, middle 60%. While the velocity and curvature are strongly correlated, dividing them to compute a gain factor over time does not produce a piecewise constant function.

There have been other attempts to infer segmentation based on the kinematic characteristics of an individual movement (e.g. Morasso and Mussa Ivaldi, 1982). Hogan and Flash (1987) develop an interesting argument against such an approach to segmentation. They argue that if a movement is the solution to a global minimization problem (as in the single via-point task), there is no sense in which it is segmented, because the alleged segments cannot be independent. This argument, however, has a potential flaw: just because a particular movement is a solution to a global optimization problem, it doesn't mean that the same movement cannot result from a local computation that happens to yield an efficient approximation to the global optimization problem. In fact, we saw earlier that for some carefully chosen paths the power law, which is purely local, is a very good approximation to the constrained minimum jerk principle. While it turned out that for most paths the power law does not approximate the minimum jerk solution particularly well, and the actual velocity profiles are much closer to the minimum jerk model than the power law, there still remains the possibility that some other local model can account for all our data as well as the constrained minimum jerk model does.

This issue is very difficult to resolve without having a concrete candidate for such a local rule, and we presently do not have one. One way to approach the problem without having an alternative local rule, is the following: assume that a subject executes a movement which clearly cannot have been planned completely in advance; if we then apply our global optimization method and find that it still works, we would have evidence that some efficient local approximation must have been used. In an attempt to do that, we analyzed reaching movements to targets placed from 20 to 40 cm away from the staring point. Subjects were instructed to make a movement which is as fast as possible, and as accurate as possible. It is well known that the paths of such movements are not straight near the end, and the velocity profiles have secondary peaks; it is generally agreed that these features of reaching movements are the result of visually guided corrections, and thus are not preplanned (Milner, 1992). We want to know whether the curvature at the end of such movements is enough for the constrained minimum jerk model to predict the secondary peaks in velocity profiles (the power law is not applicable since these features occur very close to the end of the movement). We found that this is not the case - examples are shown in figure 14 (the final portions of the path and velocity profiles are plotted). The minimum jerk prediction (dashed) is skewed, but it never displays secondary peaks (much more curvature would be needed for the model to predict secondary velocity peaks). While this result could mean that the model only applies to movements that have been planned as one unit, a more likely interpretation is that the small visually guided corrections near the end of fast and accurate reaching movements are controlled differently, and thus display a different path-velocity relationship.

General discussion

This paper has presented a constrained minimum jerk model of the relationship between hand path and velocity 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 velocity profile of straight reaching movements, and the path-velocity relationship in extemporaneous movements with complex paths. The velocity predictions of the model are equivalent to those of the original minimum-jerk model when the path is straight, and more accurate for multiple via-point tasks. Our model sacrifices the ability to predict the path of the movement, but our results also indicated that in multiple via-point tasks the path is not well predicted by the minimum jerk model. We showed that although 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-velocity relationship previously observed in extemporaneous movements to include the acceleration and deceleration phases of discrete movements, which are much more common.

The high accuracy of the constrained minimum jerk model allowed us to make several important observations:
a) the relationship between path and velocity of arm movements with complex hand paths is significantly stronger than previously thought, and it extends over the entire movement;
b) this relationship is stronger for movements of shorter duration, and is not affected by spatial scale or average velocity (it is also rather uniform across tasks and movement paths);
c) the path of the hand on a single trial contains significant information about the velocity profile on the same trial, to the extent that the velocity profile can be predicted from the path. We found the predictive accuracy to be approximately twice that obtained from the average velocity profile over a block of identical trials.

The last observation rules out the possibility that subjects maintain a central motor plan which is affected by noise acting independently on the path and velocity profile. Whatever the source of trial-to-trial variability, the effects on the hand path and the velocity profile are coupled in such a way so as to maintain the path-velocity relationship. Since the movements we studied are not executed at maximum speed, one likely source of variability is the effect of online (probably visually guided) corrections. Such online corrections, however, have to be incorporated smoothly in the movement, in a manner that is consistent with the overall path-velocity 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).

One possible explanation for the finding that prediction error increases with movement duration (and not average velocity or spatial scale) is the following: there is an open-loop motor plan, which exists prior to movement onset; the actual movement is the result of this open-loop strategy, plus noise/correction effects that are integrated along the movement. While this scenario could explain our result, it predicts that the deviation between the minimum jerk and the observed velocity profiles should increase with time, which does not seem to happen. A more likely explanation is that the "tempo" of the movement is not held constant (as the model assumes), but it fluctuates over time, resulting in bigger prediction errors for movements that take longer.

Finally, we consider the implications of the basic observation that the velocity profile of a complex arm movement is strongly related to the path of the hand in tasks where no temporal constraints are imposed (other than overall duration). It is possible that in such tasks the path is planned first (based, for example, on visual information and dynamic constraints), and then the velocity profile is computed as a function of path. This would imply, however, that the machinery involved in the velocity computation is not utilizable whenever temporal constraints are explicitly present in the task. In fact, the work on the original minimum-jerk model (Flash and Hogan, 1985) strongly suggests that in reaching and simple via-point tasks, the path is also planned using the same optimization principle. The hypothesis that we propose, then, is that the motor system has a general way of optimizing jerk (or a related quantity), and it uses it in combination with whatever set of constraints are implied by the task to "fill in" the details of the movement that are not specified by the desired effect on the external environment. The actual computation could proceed in an iterative fashion, so that both the task-related constraints and the added smoothness constraints are satisfied simultaneously.

Why should the same optimization mechanism be used across tasks if there seem to be many logical choices of a quantity to be optimized (see Nelson, 1983), and it is not obvious why jerk might be the best candidate? Furthermore, if minimizing the use of some particular resource is crucial, avoiding a movement altogether should be the best possible strategy. Why is it then that both humans and other animals make so many movements that it would seem could easily be avoided? We suggest that the answer is not in the actual quantity being optimized (which could be anything that leads to reasonable movements), but in the optimization itself. If the motor system has evolved a general and efficient mechanism to perform one particular optimization under arbitrary combinations of constraints, the same machinery can be used as a general purpose motor planner in a wide range of ill-posed problems, as long as it provides acceptable solutions.

References

Bennett, D.J., Hollerbach, J.M., Xu Y., and Hunter, I.W. (1992) Time-varying stiffness of human elbow joint during cyclic voluntary movement. Exp Brain Research, 88: 433-442.

Bizzi, E., Hogan, N., Mussa-Ivaldi, F.A., and Giszter, S. (1992) Does the nervous system use equilibrium-point control to guide single and multijoint movements. Behavioral and Brain Sciences, 15: 603-613.

De Boor, C. (1978) *A practical guide to splines.* New-York: Springer-Verlag.

Flash, T., and Hogan, N. (1985) The coordination of arm movements: an experimentally confirmed mathematical model. Journal of Neuroscience, 5(7): 1688-1703.

Gill, P.E., Murray, W. (1981) *Practical Optimization.* Academic Press.

Gribble, P., and Ostry, D. (1996) Origins of the power law relationship between movement velocity and curvature: modeling the effects of muscle mechanics and limb dynamics. Journal of Neurophysiology, in press.

Gomi, H., and Kawato, M. (1996) Equilibrium-point control hypothesis examined by measured arm stiffness during multijoint movement. Science, 272: 117-120.

Hogan, N., and Flash, T. (1987) Moving gracefully: quantitative theories of motor coordination. TINS, 10(4): 170-174.

Jack, W.R. (1895) On the analysis of voluntary muscular movements by certain new instruments. Journal of Anatomy and Physiology, 29: 473-478.

Lacquaniti, F., Terzuolo, C., and Viviani, P. (1983) The law relating the kinematic and figural aspects of drawing movements. Acta Psychol. 54: 115-130.

Milner, T.E. (1992) A model for the generation of movements requiring endpoint precision. Neuroscience, 49(2): 487-496.

Morasso, P. (1981) Spatial control of arm movements. Exp Brain Research, 42: 223-227.

Morasso, P., and Mussa Ivaldi, F.A. (1982) Trajectory formation and handwriting: a computational model. Biological Cybernetics, 45: 131-142.

Nelson, W.L. (1983) Physical principles for economies of skilled movements. Biological Cybernetics, 46: 135-147.

Pellison, D., Prablanc, C., Goodale, M.A., and Jeannerod, M. (1986) 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 Research, 62: 303-311.

Pollick, F.E., and Sapiro, G. (1996) Constant affine velocity predicts the 1/3 power law of planar motion perception and generation. Vision Research, in press.

Prakash, N. (1981) *Differential Geometry. An Integrated Approach.* TATA McGraw-Hill Publishing Company Limited.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (1992) *Numerical Recipes in C.* Cambridge University Press.

Sabes, P. (1996) The planning of visually guided arm movements: feedback perturbation and obstacle avoidance studies. Ph.D. Thesis, MIT.

Soechting, J.F., and Terzuolo, C.A. (1986) An algorithm for the generation of curvilinear wrist motion in an arbitrary plane in three dimensional space. Neuroscience, 19: 1393-1405.

Soechting, J.F., and Terzuolo, C.A. (1987) Organization of arm movements in three-dimensional space. Wrist motion is piecewise planar. Neuroscience, 23(1): 53-61.

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

Viviani, P., and Cenzato, M. (1985) Segmentation and coupling in complex movements. Journal of Experimental Psychology: Human Perception and Performance, 11: 828-845.

Viviani, P., and Schneider, R. (1991) A developmental study of the relationship between geometry and kinematics of drawing movements. Journal of Experimental Psychology: Human Perception and Performance, 17: 198-218.

Viviani, P., and Flash, T. (1995) Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. Journal of Experimental Psychology: Human Perception and Performance, 21(3): 2-53.

Wann, J., Nimmo-Smith, I., Wing, A.M. (1988) Relationship between velocity and curvature in movement: equivalence and divergence between a power law and a minimum-jerk model. Journal of Experimental Psychology: Human Perception and Performance, 14: 622-637.

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

Figure captions

Figure 1. A,B,C - Examples of the constrained minimum jerk (solid) and power law (dashed) velocity profile predictions for the synthetic paths on the left. The jerk vector is plotted at several points along the path (for the minimum jerk velocity profile). D - The tangential and normal jerk components for the middle portion of example C. Two small perturbations are applied to the minimum jerk velocity profile (solid), and the corresponding effects on the jerk components are shown.

Figure 2. Templates used in tasks 1 and 2, drawn to scale. The circle marks the starting point on each template. The order of presentation is row-first.

Figure 3. Via-point configurations used in task 3, drawn to scale. The circle marks the starting point, the star marks the end point. Every other configuration (right column) is a 90 degree rotation of the preceding one.

Figure 4. Example of one subject tracing template 1 on task 2. Top plot - the actual paths (dotted) are superimposed on the template (solid). Bottom plot - the corresponding velocity profiles, scaled to unit average velocity.

Figure 5. Examples of paths traced in task 3. Each pair of plots is data from one subject (six different subjects shown). The solid line is the minimum jerk path (which was not shown to the subjects), and the dotted lines are the actual paths on the last 5 trials of the 10 trial block. The filled circle marks the end point.

Figure 6. Summary error statistics for the three models, computed over the entire movement (top) and the middle 60% (bottom), for all tasks. For each subject, the median error score in each block of 10 trials was computed, and the results were averaged over the 8 templates.

Figure 7. Summary error statistics by template - task 2, complete movements. For each template, we computed the median error for each subject, and averaged over subjects.

Figure 8. Scatterplot of movement duration vs. prediction error for the constrained minimum jerk model, tasks 2 and 3. Each data point represents one subject, the results were averaged over all trials the subject made in the corresponding task.

Figure 9. Summary error statistics for experiment 2, complete movements (top) and middle 60% (bottom). The results are computed in the same way as in figure 6.

Figure 10. Average velocity profiles for all subjects, small 1.5 seconds movement in experiment 2 (4 of the 5 templates shown, numbers correspond to template positions in figure 2). Actual velocity profile - solid, constrained minimum jerk prediction - dashed, power law - dotted. All velocity profiles were scaled before averaging.

Figure 11. Typical examples of actual velocity profiles (solid), constrained minimum jerk prediction (dashed) and power law (dotted) on single trials - experiment 2, middle 60% of the movement.

Figure 12. Comparison between the trial to trial variability in velocity profiles (mean velocity profile prediction) and the performance of the three models.

Figure 13. The gain factor is the ratio of $v(t)$ and $\kappa^{-1/3}$ - middle 60% of template 2, experiment 2, small 1.5 seconds movements.

Figure 14. Final portions of hand paths (left) and velocity profiles (right) in rapid and accurate reaching movements. The dashed line is the prediction of the constrained minimum jerk model applied to the portion of the movements shown here.