

Dynamics systems vs. optimal control —a unifying view

Stefan Schaal

Progress in brain research

Cite this paper

Downloaded from [Academia.edu](#) 

[Get the citation in MLA, APA, or Chicago styles](#)

Related papers

[Download a PDF Pack](#) of the best related papers 



[Learning movement primitives](#)

Stefan Schaal

[Dynamic movement primitives-a framework for motor control in humans and humanoid robotics](#)

Stefan Schaal

[Movement planning and imitation by shaping nonlinear attractors](#)

Stefan Schaal

CHAPTER 27

Dynamics systems vs. optimal control — a unifying view

Stefan Schaal^{1,2,*}, Peyman Mohajerin¹ and Auke Ijspeert^{1,3}

¹Computer Science & Neuroscience, University of Southern California, Los Angeles, CA 90089-2905, USA

²ATR Computational Neuroscience Laboratory, 2-2-2 Hikaridai Seika-cho Sorako-gun, Kyoto 619-02, Japan

³School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Station 14, CH-1015 Lausanne, Switzerland

Abstract: In the past, computational motor control has been approached from at least two major frameworks: the dynamic systems approach and the viewpoint of optimal control. The dynamic system approach emphasizes motor control as a process of self-organization between an animal and its environment. Non-linear differential equations that can model entrainment and synchronization behavior are among the most favorable tools of dynamic systems modelers. In contrast, optimal control approaches view motor control as the evolutionary or development result of a nervous system that tries to optimize rather general organizational principles, e.g., energy consumption or accurate task achievement. Optimal control theory is usually employed to develop appropriate theories. Interestingly, there is rather little interaction between dynamic systems and optimal control modelers as the two approaches follow rather different philosophies and are often viewed as diametrically opposing. In this paper, we develop a computational approach to motor control that offers a unifying modeling framework for both dynamic systems and optimal control approaches. In discussions of several behavioral experiments and some theoretical and robotics studies, we demonstrate how our computational ideas allow both the representation of self-organizing processes and the optimization of movement based on reward criteria. Our modeling framework is rather simple and general, and opens opportunities to revisit many previous modeling results from this novel unifying view.

Keywords: discrete movement; rhythmic movement; movement primitives; dynamic systems; optimization; computational motor control

Introduction

Before entering a more detailed discussion on computational approaches to motor control, it is useful to start at a rather abstract level of modeling that can serve as a general basis for many theories. Following the classical control literature

from around the 1950s and 1960s (Bellman, 1957; Dyer and McReynolds, 1970), the goal of motor control and motor learning can generally be formalized in terms of finding a task-specific control policy:

$$\mathbf{u} = \pi(\mathbf{x}, t, \alpha) \quad (1)$$

that maps the continuous state vector \mathbf{x} of a control system and its environment, possibly in a time t dependent way, to a continuous control vector \mathbf{u} .

*Corresponding author. Tel.: +1 213 740 9418;
Fax: +1 213 740 1510; E-mail: sschaal@usc.edu

The parameter vector α denotes the problem-specific adjustable parameters in the policy π , e.g., the weights in neural network or a generic statistical function approximator.¹ In simple words, all motor commands for all actuators (e.g., muscles or torque motors) at every moment of time depend (potentially) on all sensory and perceptual information available at this moment of time, and possibly even past information. We can think of different motor behaviors as different control policies π_i , such that motor control can be conceived of as a library of such control policies that are used in isolation, but potentially also in sequence and superposition in order to create more complex sensory-motor behaviors.

From a computational viewpoint, one can now examine how such control policies can be represented and acquired. Optimization theory offers one possible approach. Given some cost criterion $r(\mathbf{x}, \mathbf{u}, t)$ that can evaluate the quality of an action \mathbf{u} in a particular state x (in a potentially time t dependent way), dynamic programming (DP), and especially its modern relative, reinforcement learning (RL), provide a well-founded set of algorithms of how to compute the policy π for complex nonlinear control problems. In essence, both RL and DP derive an optimal policy by optimizing the accumulated reward (in statistical expectation $E\{\}$) over a (potentially $\gamma \in [0, 1]$ -discounted²) long-term horizon (Sutton and Barto, 1998):

$$J = E \left\{ \sum_{i=0}^T \gamma^i r(\mathbf{x}, \mathbf{u}, t) \right\} \quad (2)$$

Unfortunately, as already noted in Bellman's original work (Bellman, 1957), learning of π becomes computationally intractable for even moderately high-dimensional state-action spaces, e.g., starting from ~ 6 to 10 continuous dimensions, as the search space for an optimal policy becomes too

¹Note that different parameters may actually have different functionality in the policy: some may be more low level and just store a learned pattern, while others may be higher level, e.g., as the position of a goal, that may change every time the policy is used. See, for instance, Barto and Mahadevan (2003) or the following sections of this paper.

²The discount factor causes rewards far in the future to be weighted down, as can be verified when expanding Eq. (2) over a few terms.

large or too nonlinear to explore empirically. Although recent developments in RL increased the range of complexity that can be dealt with (e.g., Tesauro, 1992; Bertsekas and Tsitsiklis, 1996; Sutton and Barto, 1998), it still seems that there is a long way to go before general policy learning can be applied to complex control problems like human movement.

In many theories of biological motor control and most robotics applications, the full complexity of learning a control policy is strongly reduced by assuming prior information about the policy. The most common priors are that the control policy can be reduced to a desired trajectory, $[\mathbf{x}_d(t), \dot{\mathbf{x}}_d(t)]$. Optimal control or RL approaches for trajectory learning are computationally significantly more tractable (Kawato and Wolpert, 1998; Peters et al., 2005). For instance, by using a tracking error-driven feedback controller (e.g., proportional-derivative, PD), a (explicitly time dependent) control policy can be written as:

$$\begin{aligned} \mathbf{u} &= \pi(\mathbf{x}, \alpha(t), t) = \pi(\mathbf{x}, [\mathbf{x}_d(t), \dot{\mathbf{x}}_d(t)], t) \\ &= \mathbf{K}_x(\mathbf{x}_d(t) - \mathbf{x}) + \mathbf{K}_{\dot{x}}(\dot{\mathbf{x}}_d(t) - \dot{\mathbf{x}}) \end{aligned} \quad (3)$$

For problems in which the desired trajectory is easily generated and in which the environment is static or fully predictable, such a shortcut through the problem of policy generation is highly successful. However, since policies like those in Eq. (3) are usually valid only in a local vicinity of the time course of the desired trajectory $(\mathbf{x}_d(t), \dot{\mathbf{x}}_d(t))$, they are not very flexible. A typical toy example for this problem is the tracking of the surface of a ball with the fingertip. Assume the fingertip movement was planned as a desired trajectory that moves every second 1 cm forward in tracing the surface. Now imagine that someone comes and holds the fingertip for 10 s, i.e., no movement can take place. In these 10 s, however, the trajectory plan has progressed 10 cm, and upon the release of your finger, the error-driven control law in Eq. (3) would create a strong motor command to catch up. The bad part, however, is that Eq. (3) will try to take the shortest path to catch up with the desired trajectory, which, due to the concave surface in our example, will actually try to traverse through the inside of the ball. Obviously, this behavior is

inappropriate and would hurt the human and potentially destroy the ball. Many daily life motor behaviors have similar properties. Thus, when dealing with a dynamically changing environment in which substantial and reactive modifications of control commands are required, one needs to adjust desired trajectories appropriately, or even generate entirely new trajectories by generalizing from previously learned knowledge. In certain cases, it is possible to apply scaling laws in time and space to desired trajectories (Hollerbach, 1984; Kawamura and Fukao, 1994), but those can provide only limited flexibility. For the time being, the “desired trajectory” approach seems to be too restricted for general-purpose motor control and planning in dynamically changing environments, as needed in every biological motor system, and some biological evidence has been accumulated that completely preplanned desired trajectories may not exist in human behavior³ (Desmurget and Grafton, 2000).

Given that the concept of time-indexed desired trajectories has its problems, both from a computational and a biological plausibility point of view, one might want to look for other ways to generate control policies. From a behavioral point of view, a control policy is supposed to take the motor system from an arbitrary start point to the desired behavior. In most biological studies of arm movements, the desired behavior is simply a goal for pointing or grasping. But there is also the large class of cyclic movements, like walking, swimming, chewing, etc. Both behavioral classes can be thought of as attractor dynamics, i.e., either a point attractor as in reaching and pointing, or a limit cycle attractor as in periodic movement. Systems with attractor dynamics have been studied extensively in the nonlinear dynamic systems literature (Guckenheimer and Holmes, 1983; Strogatz, 1994). A dynamic system can generally be written as a differential equation:

$$\dot{\mathbf{x}} = f(\mathbf{x}, \alpha, t) \quad (4)$$

³It should be noted, however, that some approaches exist that can create time indexed desired trajectories in a reactive fashion (Hoff and Arbib, 1993), but these approaches only apply to a very restricted class of analytically tractable trajectories, e.g., polynomial trajectories (Flash and Hogan, 1985).

which is almost identical to Eq. (1), except that the left-hand side denotes a change of state, not a motor command. Such a *kinematic* formulation is, however, quite suitable for motor control if we conceive of this dynamic system as a kinematic policy that creates kinematic target values (e.g., positions, velocities, accelerations), which subsequently are converted to motor commands by an appropriate controller (Wolpert, 1997). Planning in kinematic space is often more suitable for motor control since kinematic plans generalize over a large part of the workspace — nonlinearities due to gravity and inertial forces are taken care of by the controller at the motor execution stage (cf. Fig. 1). Kinematic plans can also theoretically be cleanly superimposed to form more complex behaviors, which is not possible if policies code motor commands directly. It should be noted, however, that a kinematic representation of movement is not necessarily independent of the dynamic properties of the limb. Proprioceptive feedback can be used on-line to modify the attractor landscape of the policy in the same way as perceptual information (Rizzi and Koditschek, 1994; Schaal and Sternad, 1998; Williamson, 1998; Nakanishi et al., 2004). Figure 1 indicates this property with the “perceptual coupling” arrow.

Most dynamic systems approaches also emphasize removing the explicit time dependency of π , such that the control policies become “autonomous dynamic systems”:

$$\dot{\mathbf{x}} = f(\mathbf{x}, \alpha) \quad (5)$$

Explicit timing is cumbersome, as it requires maintaining a clocking signal, e.g., a time counter that increments at very small time steps (as typically done in robotics). Besides that it is disputed whether biological system have access to such clocks (e.g., Keating and Thach, 1997; Roberts and Bell, 2000; Ivry et al., 2002), there is an additional level of complexity needed for aborting, halting, or resetting the clock when unforeseen disturbances happen during movement execution, as mentioned in the ball-tracing example above. The power of modeling motor control with autonomous nonlinear dynamic systems is further enhanced, as it is now theoretically rather easy to modulate the control policy by additional,

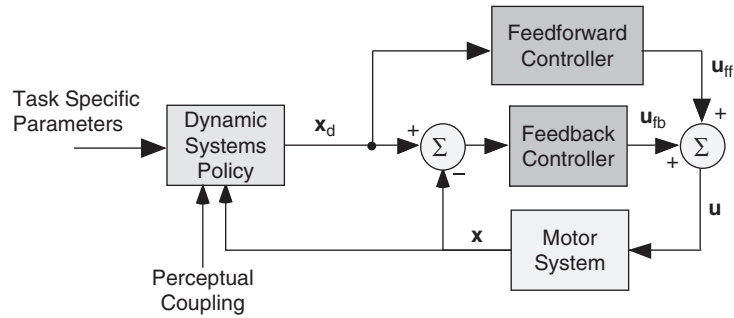


Fig. 1. Sketch of a control diagram with a dynamic systems kinematic policy, in particular how the policy is inserted into a controller with feedback (i.e., error-driven) and feedforward (i.e., anticipatory or model-based) components.

e.g., sensory or perceptual, variables, summarized in the coupling term C :

$$\dot{\mathbf{x}} = f(\mathbf{x}, \alpha) + C \quad (6)$$

We will return to such coupling ideas later in the paper.

Adopting the framework of dynamics systems theory for policy generation connects to a large body of previous work. For invertebrates and lower vertebrates, research on central pattern generators (Selverston, 1980; Getting, 1985; Kopell and Ermentrout, 1988; Marder, 2000; Righetti and Ijspeert, 2006; Ijspeert et al., 2007) has a long tradition of using coupled oscillator theories for modeling. From a behavioral point of view, many publications in the literature deal with coupled oscillator theories to explain perception–action coupling and other behavioral phenomena (Kugler et al., 1982; Turvey, 1990; Kelso, 1995). Thus, at the first glance, one might expect a straightforward and experimentally well-established framework to approach control policies as nonlinear dynamic systems. Unfortunately, this is not the case. First, modeling with nonlinear dynamics systems is mathematically quite difficult and requires usually very good intuition and deep knowledge in nonlinear systems theory — optimization approaches are often much easier to handle with well-established software tools. Second, with very few exceptions (Bullock and Grossberg, 1988; Schöner, 1990), dynamic systems approaches have only focused on periodic behavior, essentially assuming that discrete behavior is just an aborted limit cycle. In contrast, optimization approaches to motor control primarily have focused on

discrete movement like reaching and pointing (e.g., Shadmehr and Wise, 2005), and rhythmic movement was frequently conceived of as cyclically concatenated discrete movements.

The goal of this paper is to demonstrate that a dynamic systems approach can offer a simple and powerful approach for both discrete and rhythmic movement phenomena, and that it can smoothly be combined with optimization approaches to address a large range of motor phenomena that have been observed in human behavior. For this purpose, first, we will review some behavioral and imaging studies that accumulated evidence that the distinction of discrete and rhythmic movement, as suggested by point and limit cycle attractors in dynamic systems theory, actually is also useful for classifying human movement. Second, we will suggest a modeling framework that can address both discrete and rhythmic movement in a simple and coherent dynamic systems framework. In contrast to any other dynamic systems approaches to motor control in the past, the suggested modeling approaches can easily be used from the viewpoint of optimization theory, too, and bridges thus the gap between dynamic systems and optimization approaches to motor control. We will demonstrate the power of our modeling approach in several synthetic and robotic studies.

Discrete and rhythmic movement — are they the same?

Since Morasso’s and his coworkers’ seminal work in the early 1980s (Morasso, 1981, 1983; Abend

et al., 1982), a large amount of focus has been given to stroke-based trajectory formation. In the wake of this research, periodic movement was often regarded as a special case of discrete (i.e., stroke-based) movement generation, where two or more strokes are cyclically connected. In the following sections, we will review some of our own and other people's research that tried to emphasize periodic movement as an independent and equally important function of the nervous system, similar as point attractors and limit cycle attractors in dynamic systems theory require quite different treatment.

Dynamic manipulation as coupled dynamic systems

From the viewpoint of motor psychophysics, the task of bouncing a ball on a racket constitutes an interesting test bed to study trajectory planning and visuomotor coordination in humans. The bouncing ball has a strong stochastic component in its behavior and requires a continuous change of motor planning in response to the partially unpredictable behavior of the ball. In previous work (Schaal et al., 1996), we examined which principles were employed by human subjects to accomplish stable ball bouncing. Three alternative movement strategies were postulated. First, the point of impact could be planned with the goal of intersecting the ball with a well-chosen movement velocity such as to restore the correct amount of energy to accomplish a steady bouncing height (Aboaf et al., 1989); such a strategy is characterized by a constant velocity of the racket movement in the vicinity of the point of racket-ball impact. An alternative strategy was suggested by work in robotics: the racket movement was assumed to mirror the movement of the ball, thus impacting the ball within increasing velocity profile, i.e., positive acceleration (Rizzi and Koditschek, 1994). Both of these strategies are essentially stroke-based: a special trajectory is planned to hit the ball in its downward fall, and after the ball is hit, the movement is reset to redo this trajectory plan. A dynamic systems approach allows yet another way of accomplishing the ball bouncing task: an oscillatory racket movement creates a dynamically stable

basin of attraction for ball bouncing, thus allowing even open-loop stable ball bouncing, i.e., ball bouncing with one's eyes closed. This movement strategy is characterized by a negative acceleration of the racket when impacting the ball (Schaal and Atkeson, 1993) — a quite nonintuitive solution: why would one break the movement before hitting the ball?

Examining the behavior of six subjects revealed the surprising result that dynamic systems captured the human behavior the best: all subjects reliably hit the ball with a negative acceleration at impact, as illustrated in Fig. 2 (note that some subjects, like Subject 5, displayed a learning process where early trials had positive acceleration at impact, but later trials switched to negative acceleration). Manipulations of bouncing amplitude also showed that the way the subjects accomplished such changes could easily be captured by a simple reparameterization of the oscillatory component of the movement, a principle that we will incorporate in our modeling approach below. Importantly, it was hard to imagine how the subjects could have achieved their behavioral characteristics with stroke-based movement generation scheme.

Apparent movement segmentation does not indicate segmented control

Invariants of human movement have been an important area of research for more than two decades. Here we will focus on two such invariants, the 2/3-power law and piecewise-planar movement segmentation, and how a parsimonious explanation of those effects can be obtained without the need of stroke-based movement planning.

Studying handwriting and 2D drawing movements, Viviani and Terzuolo (1980) were the first to identify a systematic relationship between angular velocity and curvature of the end-effector traces of human movement, an observation that was subsequently formalized in the “2/3-power law” (Lacquaniti et al., 1983):

$$a(t) = kc(t)^{2/3} \quad (7)$$

$a(t)$ denotes the angular velocity of the endpoint trajectory, and $c(t)$ the corresponding curvature;

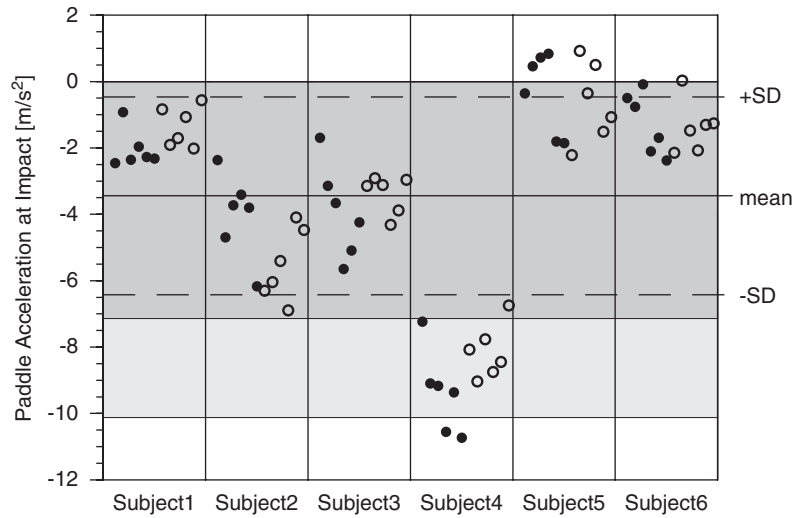


Fig. 2. Trial means of acceleration values at impact, $\ddot{x}_{P,n}$, for all six experimental conditions grouped by subject. The symbols differentiate the data for the two gravity conditions G . The dark shading covers the range of maximal local stability for G_{reduced} the light shading the range of maximal stability for G_{normal} . The overall mean and its standard deviation refers to the mean across all subjects and all conditions.

this relation can be equivalently expressed by a 1/3 power-law relating tangential velocity $v(t)$ with radius of curvature $r(t)$:

$$v(t) = kr(t)^{1/3} \quad (8)$$

Since there is no physical necessity for movement systems to satisfy this relation between kinematic and geometric properties, and since the relation has been reproduced in numerous experiments (for an overview, see [Viviani and Flash, 1995](#)), the 2/3-power law has been interpreted as an expression of a fundamental constraint of the CNS, although biomechanical properties may significantly contribute ([Gribble and Ostry, 1996](#)). Additionally, [Viviani and Cenzato \(1985\)](#) and [Viviani \(1986\)](#) investigated the role of the proportionality constant k as a means to reveal movement segmentation: as k is approximately constant during extended parts of the movement and only shifts abruptly at certain points of the trajectory, it was interpreted as an indicator for segmented control. Since the magnitude of k also appears to correlate with the average movement velocity in a movement segment, k was termed the “velocity gain factor.” [Viviani and Cenzato \(1985\)](#) found that planar elliptical drawing patterns are characterized by a single k

and, therefore, consist of one unit of action. However, in a fine-grained analysis of elliptic patterns of different eccentricities, [Wann et al., 1988](#) demonstrated consistent deviations from this result. Such departures were detected from an increasing variability in the log- v to log- r -regressions for estimating k and the exponent β of Eq. (2), and ascribed to several movement segment each of which has a different velocity gain factor k .

The second movement segmentation hypothesis we want to address partially arose from research on the power law. [Soechting and Terzuolo \(1987a, b\)](#) provided qualitative demonstrations that 3D rhythmic endpoint trajectories are piecewise planar. Using a curvature criterion as the basis for segmentation, they confirmed and extended [Morasso's \(1983\)](#) results that rhythmic movements are segmented into piecewise planar strokes. After [Pellizzer et al. \(1992\)](#) demonstrated piecewise planarity even in an isometric task, movement segmentation into piecewise planar strokes has largely been accepted as one of the features of human and primate arm control.

We repeated some of the experiments that led to the derivation of the power law, movement segmentation based on the power law, and movement

segmentation based on piecewise planarity. We tested six human subjects when drawing elliptical patterns and figure-8 patterns in 3D space freely in front of their bodies. Additionally, we used an anthropomorphic robot arm, a Sarcos Dexterous Arm, to create similar patterns as those performed by the subjects. Importantly, the robot generated the elliptical and figure-8 patterns solely out of joint-space oscillations, i.e., a nonsegmented movement control strategy. For both humans and the robot, we recorded the 3D position of

the fingertip and the seven joint angles of the performing arm.

Figure 3 illustrates data traces of one human subject and the robot subject for elliptical drawing patterns of different sizes and different orientations. For every trajectory in this graph, we computed the tangential velocity of the fingertip of the arm and plotted it versus the radius of curvature raised to the power 1/3. If the power law were obeyed, all data points should lie on a straight line through the origin. Figure 3a, b clearly

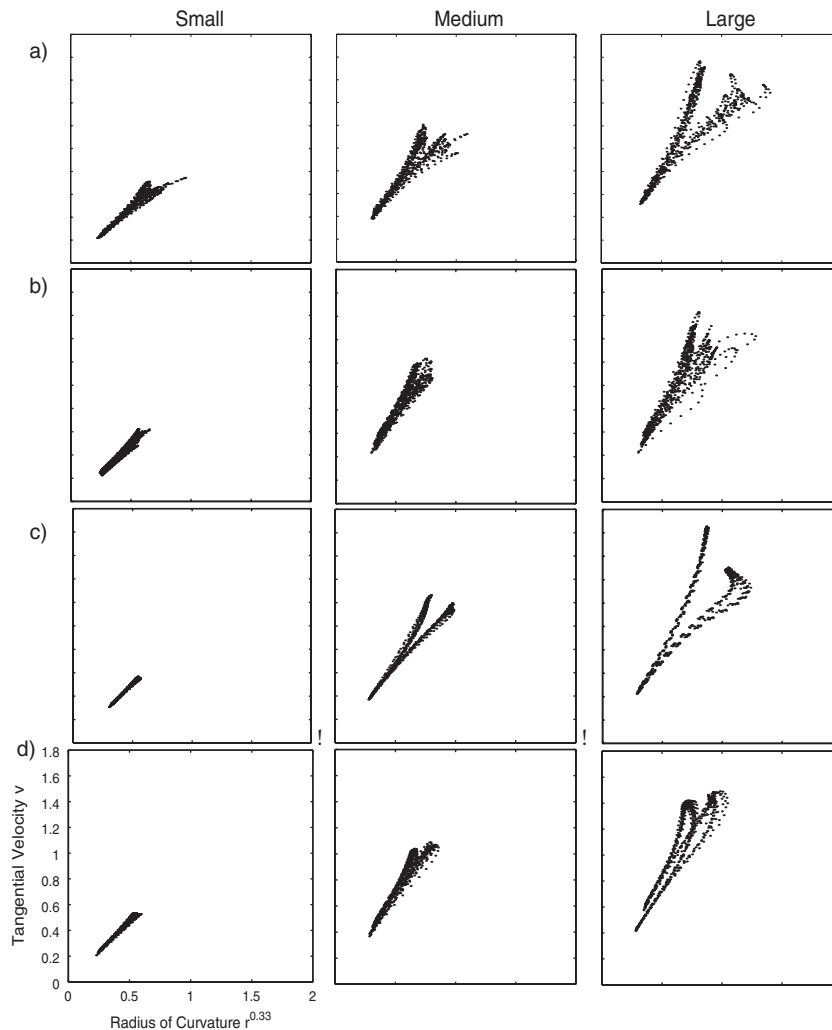


Fig. 3. Tangential velocity versus radius of curvature to the power 1/3 for ellipses of small, medium, and large size for elliptical pattern orientations in the frontal and oblique workspace plane: (a) human frontal; (b) human oblique; (c) robot frontal; (d) robot oblique.

demonstrates that for large size patterns, this is not the case, indicating that the power seems to be violated for large size patterns. However, the development of two branches for large elliptical patterns in Fig. 3a, b could be interpreted that large elliptical movement patterns are actually composed of two segments, each of which obeys the power law. The rejection of the latter point comes from the robot data in Fig. 3c, d. The robot produced strikingly similar features in the trajectory realizations as the human subjects.

However, the robot simply used oscillatory joint space movement to create these patterns, i.e., there was no segmented movement generation strategy. Some mathematical analysis of the power law and the kinematic structure of human arms could finally establish that the power law can be interpreted as an epiphenomenon of oscillatory movement generation: as long as movement patterns are small enough, the power law holds, while for large size patterns the law breaks down (Sternad and Schaal, 1999; Schaal and Sternad, 2001). Using figure-8 patterns instead of elliptical patterns, we were also able to illuminate the reason for apparent piecewise-planar movement segmentation in rhythmic drawing patterns. Figure 4 shows figure-8 patterns performed by human and robot subjects in a planar projection when looking at the figure-8 from the side. If realized with an appropriate width-to-height ratio, figure-8 patterns look indeed like piecewise planar trajectories

in this projection and invite the hypothesis of movement segmentation at the node of the figure-8. However, as in the previous experiment, the robot subject produced the same features of movement segmentation despite the fact that it used solely joint space oscillations to create the patterns, i.e., no movement segmentation. Again, it was possible to explain the apparent piecewise planarity from a mathematical analysis of the kinematics of the human arm, rendering piecewise planarity to be an epiphenomenon of oscillatory joint space trajectories and the nonlinear kinematics of the human arm (Sternad and Schaal, 1999).

Superposition of discrete and rhythmic movement

In another experiment, we addressed the hypothesis that discrete and rhythmic movements are two separate movement regimes that can be used in superposition, sequence, or isolation. Subjects performed oscillatory movements around a given point in the workspace with one joint of the arm, and shifted the mean position of another joint of the same (or the other arm) at an auditory signal to another point. In previous work (Adamovich et al., 1994), it was argued that such a discrete shift terminates the oscillatory movement (generated by two cyclically connected movement strokes) and restarts it after the shift, i.e., the entire system of

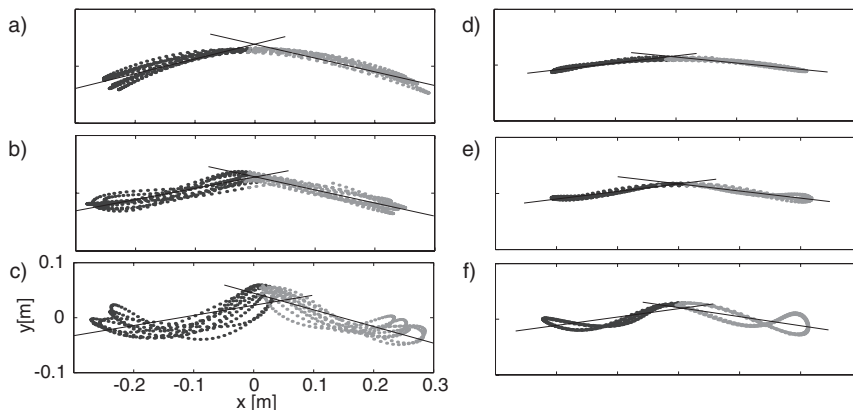


Fig. 4. Planar projection of one subject's figure-8 patterns of small, medium, and large width/height ratio: (a–c) human data; (d–f) corresponding robot data. The data on the left side of each plot belong to one lobe of the figure-8, and the data on the right side to the other figure-8 lobe.

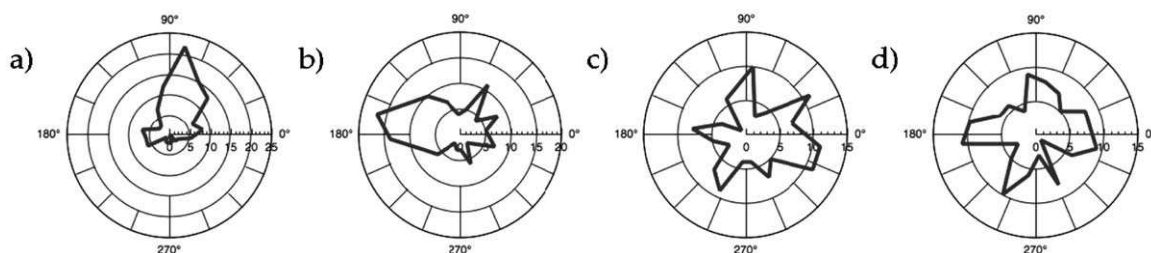


Fig. 5. Polar histograms of the phase of the discrete movement onset in various experimental conditions, averaged over six participants: (a) a rhythmic elbow movement is superimposed with a discrete elbow flexion; (b) a rhythmic elbow movement is superimposed with discrete wrist supination; (c) a rhythmic wrist flexion–extension movement with superimposed discrete shoulder flexion; (d) a right elbow flexion–extension movement superimposed with a discrete left elbow flexion movement. In (a) and (b), the onset of the discrete movement is confined to a phase window of the on-going rhythmic movement. In (c) and (d), no such phase window was found.

rhythmic and discrete movement was assumed to be generated by a sequence of discrete strokes.

Among the most interesting features of this experiment was that the initiation of the discrete movement superimposed onto ongoing rhythmic movement was constrained to a particular phase window of the ongoing rhythmic movement when both discrete and rhythmic movement used the same joint (Adamovich et al., 1994; Sternad et al., 2000, 2002; De Rugy and Sternad, 2003) (Fig. 5a) and even when the discrete and rhythmic movement used different joints (Fig. 5b) (Sternad and Dean, 2003). Furthermore, in both types of experiments the ongoing rhythmic movement was disrupted during the discrete initiation and showed phase resetting. Interestingly, in a bimanual task (Wei et al., 2003), where subjects performed rhythmic movement with their dominant arm and initiated a second discrete movement with their nondominant arm, there was no evidence of a preferred phase window for the discrete movement onset (Fig. 5d).

In Mohajerian et al. (2004), we repeated this experimental paradigm over a systematic set of combinations of discrete and rhythmic movement of different joints of the same arm, and also joints from the dominant and nondominant arm — some of the results are shown in Fig. 5. All observed phenomena of phase windows of the discrete movement onset and phase resetting of the rhythmic movement could be explained by superimposed rhythmic and discrete movement components and spinal reflexes. While the CNS executes

the rhythmic movement, the discrete movement is triggered according to the auditory cue as a superimposed signal. If the rhythmic movement uses a muscle that is also needed for the discrete movement, and if this muscle is currently inhibited by the spinal interneuronal circuits due to reciprocal inhibition, the discrete movement onset is delayed. Such a superposition also leads to phase resetting of the rhythmic movement. Whenever the rhythmic and discrete joint did not share muscles for execution, no phase windows and phase resetting was observed (Fig. 5c, d). One more time, the hypothesis of independent circuits for discrete and rhythmic movement offered an elegant and simple explanation for observed behavioral phenomena.

Brain activation in discrete and rhythmic movement

Among the most compelling evidence in favor of the idea that discrete and rhythmic movement are independent functional circuits in the brain is a recent fMRI study that demonstrated that rhythmic and discrete movement activate different brain areas. Figure 6 illustrates the summary results from this experiment, where subjects performed either periodic wrist flexion–extension oscillations, or discrete flexion-to-extension or extension-to-flexion point-to-point movements with the same wrist. The major findings were that while rhythmic movement activated only a small number of unilateral primary motor areas (M1, S1, PMdc, SMA,

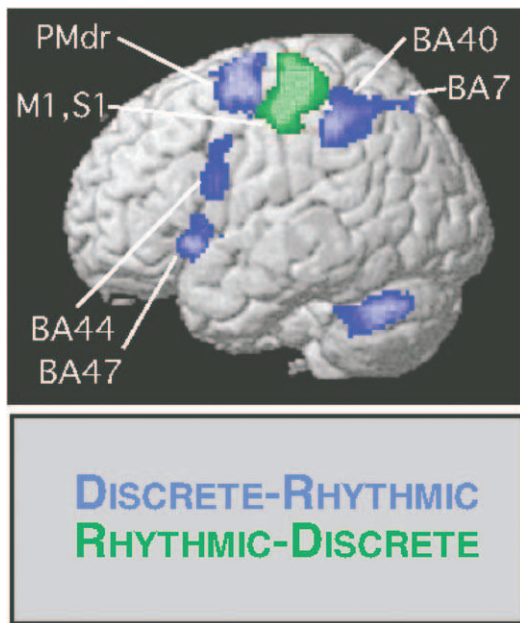


Fig. 6. Differences in brain activation between discrete and rhythmic wrist movements. Abbreviations are (Picard and Strick, 2001): CCZ: caudal cingulate zone; RCZ: rostral cingulate zone, divided in an anterior (RCZa) and posterior (RCZp) part; SMA: caudal portion of the supplementary motor area, corresponding to SMA proper; pre-SMA: rostral portion of the supplementary motor area; M1: primary motor cortex; S1: primary sensory cortex; PMdr: rostral part of the dorsal premotor cortex; PMdc: caudal part of the dorsal premotor cortex; BA: Brodman area; BA7: precuneus in parietal cortex; BA8: middle frontal gyrus; BA 9: middle frontal gyrus; BA10: anterior frontal lobe; BA47: inferior frontal gyrus; BA40: inferior parietal cortex; BA44: Broca's area.

pre-SMA, CCZ, RCZp, cerebellum), discrete movement activated a variety of additional contralateral nonprimary motor areas (BA7, BA40, BA44, BA47, PMdr, RCZa) and, moreover, showed very strong bilateral activity in both the cerebrum and cerebellum (Schaal et al., 2004). Figure 6 shows some of these results in as much as they can be visualized on the surface of the left hemisphere: most important are the Discrete-Rhythmic (blue) areas, which were unique to discrete movement. The Rhythmic-Discrete (green) area is actually active in both rhythmic and discrete movements, just to larger extend in rhythmic movement, which can be explained by the overall larger amount of movement in rhythmic trials.

Control experiments examined whether such unbalanced amounts of movement in rhythmic movement, and, in discrete movement, the much more frequent movement initiation and termination and the associated cognitive effort could account for the observed differences. Only BA40, BA44, RCZa, and the cerebellum were potentially involved in such issues, leaving BA7, BA47, and PMdr as well as a large amount of bilateral activation a unique feature in discrete movement. Since rhythmic movement activates significantly fewer brain areas than discrete movement, it was concluded that it does not seem to be warranted to claim that rhythmic movement is generated on top of a discrete movement system, i.e., rhythmic arm movement is *not* composed of discrete strokes. The independence of discrete and rhythmic movement systems in the brain seemed to be the most plausible explanation of the imaging data, which is in concert with various other studies that demonstrated different behavioral phenomena in discrete and rhythmic movement (e.g., Smits-Engelsman et al., 2002; Buchanan et al., 2003; Spencer et al., 2003).

Discrete and rhythmic movement: a computational model

The previous section tried to establish that a large number of behavioral experiments support the idea that discrete and rhythmic movement should be treated as separate movement systems, and in particular, that there is strong evidence against the hypothesis that rhythmic movement is generated from discrete strokes. We will now turn to a unifying modeling framework for discrete and rhythmic movement, with the special focus to bridge dynamic systems approaches and optimization approaches to motor control. A useful start is to establish a list of properties that such a modeling framework should exhibit. In particular, we wish to model:

- point-to-point and periodic movements,
- multijoint movement that requires phase locking and arbitrary phase offsets between individual joints (e.g., as in biped locomotion),

- discrete and rhythmic movement that have rather complex trajectories (e.g., joint reversals, curved movement, a tennis forehand, etc.),
- learning and optimization of movement,
- coupling phenomena, in particular bimanual coupling phenomena and perception–action coupling,
- timing (without requiring an explicit time representation),
- generalization of learned movement to similar movement tasks,
- robustness of movements to disturbances and interactions with the environment.

As a starting point, we will use a dynamic systems model, as this approach seems to be the best suited for creating autonomous control policies that can accommodate coupling phenomena. Given that the modeling approach suggested below will be able to represent a library of different movements in the language of dynamic systems theory, we conceive of every member of this library as a movement primitive, and call our approach Dynamic Movement Primitives (DMPs) (Ijspeert et al., 2001, 2002a, b, 2003).

We assume that the variables of a DMP represent the desired kinematic state of a limb, i.e., desired positions, velocities, and accelerations for each joint. Alternatively, the DMP could also be defined in task space, and we would use appropriate task variables (e.g., the distance of the hand from an object to be grasped) as variables for the DMP — for the discussions in this paper, this distinction is, however, of subordinate importance, and, for the ease of presentation, we will focus on formulations in joint space. As shown in Fig. 1, kinematic variables are converted to motor commands through a feedforward controller — usually by employing an inverse dynamics model — and stabilized by low gain⁴ feedback control. The example of Fig. 1 corresponds to a classical computed torque controller (Craig, 1986), which has

⁴The emphasis of low gain feedback control is motivated by the desire to have a movement system that is compliant when interacting with external objects or unforeseen perturbation, which is a hallmark of human motor control, but quite unlike traditional high gain control in most robotics applications.

also been suggested for biological motor control (Kawato, 1999), but any other control scheme could be inserted here. Thus, the motor execution of DMPs can incorporate any control technique that takes as input kinematic trajectory plans, and in particular, it is compatible with current theories of model-based control in computational motor control.

Motor planning with DMPs

In order to accommodate discrete and rhythmic movement plans, two kinds of DMPs are needed: point attractive systems and limit-cycle systems. The key question of DMPs is how to formalize nonlinear dynamic equations such that they can be flexibly adjusted to represent complex motor behaviors without the need for manual parameter tuning and the danger of instability of the equations. We will sketch our approach in the example of a discrete dynamic system for reaching movements — an analogous development holds for rhythmic systems.

Assume we have a basic point attractive system, instantiated by the second order dynamics

$$\tau \ddot{z} = \alpha_z(\beta_z(g - y) - \dot{z}) + \dot{f}, \quad \tau \dot{y} = z \quad (9)$$

where g is a known goal state, α_z and β_z time constants, τ a temporal scaling factor (see below) and y, \dot{y} correspond to the desired position and velocity generated by Eq. (9), interpreted as a movement plan as used in Fig. 1. For instance, y, \dot{y} could be the desired states for a one degree-of-freedom motor system, e.g., the elbow flexion–extension. Without the function f , Eq. (9) is nothing but the first-order formulation of a linear spring-damper, and, after some reformulation, the time constants α_z and β_z have an interpretation in terms of spring stiffness and damping. For appropriate parameter settings and $f = 0$, these equations form a globally stable linear dynamic system with g as a unique point attractor, which means that for any start position the limb would reach g after a transient, just like a stretched spring, upon release, will return to its equilibrium point. Our key goal, however, is to instantiate the nonlinear function f in Eq. (9) to change the rather trivial

exponential and monotonic convergence of y towards g to allow trajectories that are more complex on the way to the goal. As such a change of Eq. (9) enters the domain of nonlinear dynamics, an arbitrary complexity of the resulting equations might be expected. To the best of our knowledge, this problem has prevented research from employing nonlinear dynamic systems models on a larger scale so far. We will address this problem by first introducing a bit more formalism, and then by analyzing the resulting system equations.

The easiest way to force Eq. (9) to become more complex would be to create a function f as an explicit function of time. For instance, $f(t) = \sin(\omega t)$ would create an oscillating trajectory y , or $f(t) = \exp(-t)$ would create a speed up of the initial part of the trajectory y — such functions are called forcing functions in dynamic systems theory (Strogatz, 1994), and, after some reformulation, Eq. (9) could also be interpreted as PD controller that tracks a complex desired trajectory, expressed with the help f . But, as mentioned before, we would like to avoid explicit time dependencies. To achieve this goal, we need an additional dynamic system

$$\tau \dot{x} = -\alpha_x x \quad (10)$$

and the nonlinear function f in form of

$$f(x, g, y_0) = \frac{\sum_{i=1}^N \psi_i w_i x}{\sum_{i=1}^N \psi_i} (g - y_0),$$

where $\psi_i = \exp(-h_i(x - c_i)^2)$ (11)

Equation (10) is a simple first order “leaky-integrator” equation as used in many models of neural dynamics (e.g., Hodgkin and Huxley, 1952) — we will call this equation the *canonical* system from now on, as it is among the most basic dynamic systems available to create a point attractor. From any initial conditions, Eq. (10) can be guaranteed to converge monotonically to zero. This monotonic convergence of x becomes a substitute for time: all what time does is that it monotonically increases, similar to the time course of x . Of course, x behaves also a little bit different

from time: it monotonically decreases (which, mathematically, is just a technically irrelevant detail), and it saturates exponentially at the value “0”, which is appropriate as we expect that at this time the movement terminates. Equation (11) is a standard representation of a nonlinear function in terms of basis functions, as commonly employed in modeling population coding in the primate brain (e.g., Mussa-Ivaldi, 1988; Georgopoulos, 1991). Let us assume that the movement system is in an initial state $y = g = y_0$, $z = 0$, and $x = 0$. To trigger a movement, we change the goal g to a desired value and set $x = 1$ (where the value “1” is arbitrary and just chosen for convenience), similar as done with the “go” value in (Bullock and Grossberg, 1988). The duration of the movement is determined by the time constant τ . The value of x will now monotonically converge back to zero. Such a variable is called a “phase” variable as one can read out from its value in which phase of the movement we are, where “1” is the start, and “0” is the end. The nonlinear function f is generated by anchoring its Gaussian basis functions ψ_i (characterized by a center c_i and bandwidth h_i) in terms of the phase variable x . The phase x appears also multiplicative in Eq. (11) such that the influence of f vanishes at the end of the movement when x has converged to zero (see below). It can be shown that the combined system in Eqs. (9)–(11) asymptotically converge to the unique point attractor g .

The example in Fig. 7 clarifies the ingredients of the discrete DMP. The top row of Fig. 7 illustrates the position, velocity, and acceleration trajectories that serve as desired inputs to the motor command generation stage (cf. Fig. 1) — acceleration is equivalent to the time derivative of z , $\ddot{y} = \dot{z}$. In this example, the trajectories realize a minimum jerk trajectory (Hogan, 1984), a smooth trajectory as typically observed in human behavior (the ideal minimum jerk trajectory, which minimizes the integral of the squared jerk along the trajectory, is superimposed to the top three plots of Fig. 7, but the difference to the DMP output is hardly visible). The remaining plots of Fig. 7 show the time course of all internal variables of the DMP, as given by Eqs. (9)–(11). Note that the trajectory of x is just a strictly monotonically decreasing

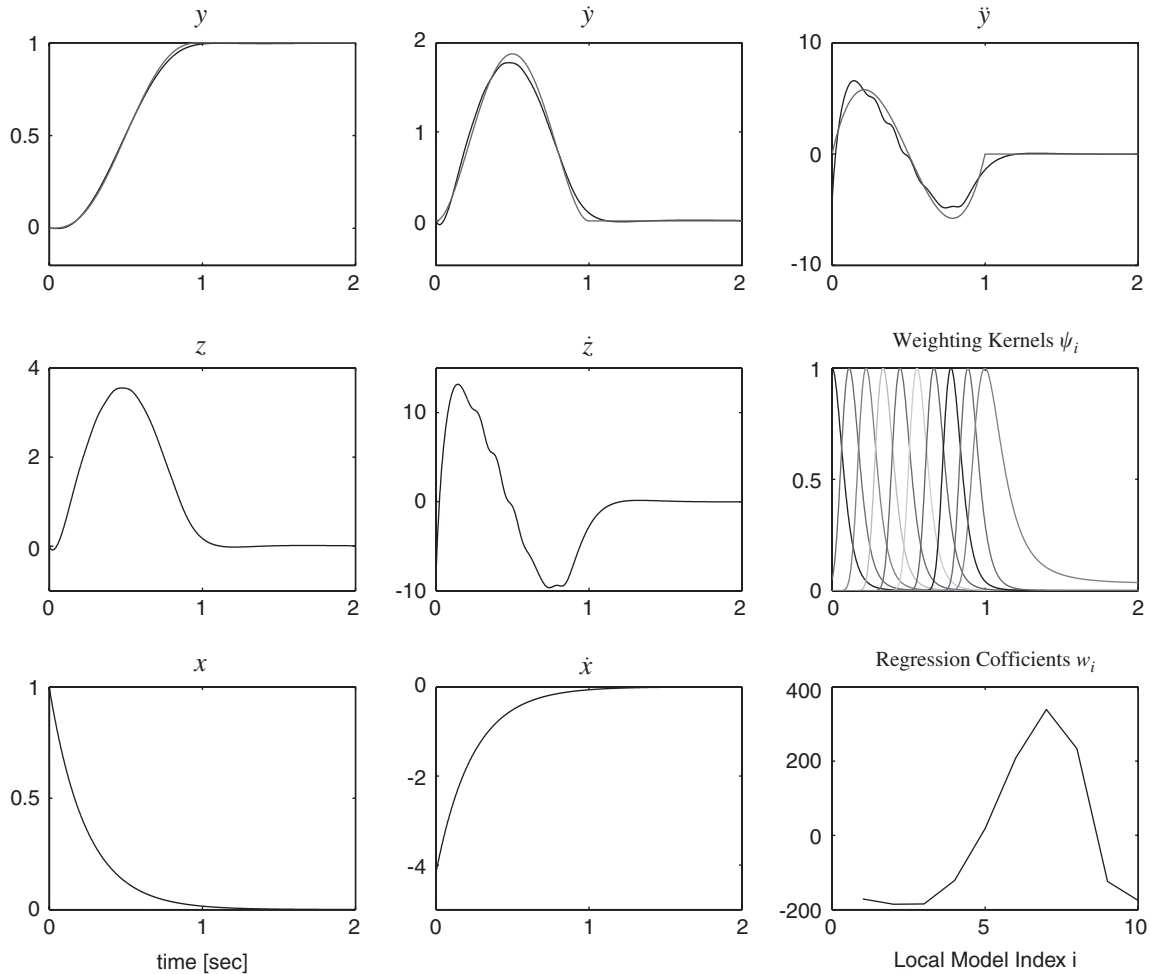


Fig. 7. Example of all variables of a discrete movement dynamic primitive as realized in a minimum jerk movement from zero initial conditions to goal state $g = 1$.

curve. As x multiplies the nonlinearity in Eq. (11), the nonlinearity only acts in a transient way, one of the main reasons that these nonlinear differential equations remain relatively easy to analyze. The basis function activations (ψ_i) are graphed as a function of time, and demonstrate how they essentially partition time into shorter intervals in which the function value of f can vary.

It is not the particular instantiation in Eqs. (9)–(11) that is the most important idea of DMPs, but rather it is the design principle that matters. A DMP consists of two sets of differential equations:

a *canonical* system

$$\tau \dot{\mathbf{x}} = \mathbf{h}(\mathbf{x}, \theta) \quad (12)$$

and an *output* system

$$\tau \dot{\mathbf{y}} = \mathbf{g}(\mathbf{y}, f, \theta) \quad (13)$$

where we just inserted θ as a placeholder for all parameters of these systems, like goal, time constants, etc. The canonical system needs to generate the phase variable x and is a substitute for time for anchoring our spatially localized basis functions Eq. (11). The appealing property of using a phase variable instead of an explicit time

representation is that we can now manipulate the time evolution of phase, e.g., by speeding up or slowing down a movement as appropriate by means of additive coupling terms or phase resetting techniques (Nakanishi et al., 2004) — in contrast, an explicit time representation cannot be manipulated as easily. For instance, Eq. (10) could be augmented to be

$$\tau\dot{x} = \alpha_x x \frac{1}{1 + \alpha_c (y_{\text{actual}} - y)^2} \quad (14)$$

The term $(y_{\text{actual}} - y)$ is the tracking error of the motor system, if this error is large, the time development of the canonical system comes to a stop, until the error is reduced — this is exactly what one would want if a motor act got suddenly perturbed.

An especially useful feature of this general formalism is that it can be applied to rhythmic movements as well, simply by replacing the point attractor in the canonical system with a limit cycle oscillator (Ijspeert et al., 2003). Among the simplest oscillators is a phase representation, i.e., constant phase speed:

$$\tau\dot{\phi} = 1 \quad (15)$$

where r is the amplitude of the oscillator, A the desired amplitude, and ϕ its phase. For this case, Eq. (11) is modified to

$$f(\phi, A) = \frac{\sum_{i=1}^N \psi_i w_i}{\sum_{i=1}^N \psi_i} A,$$

$$\text{where } \psi_i = \exp(h_i(\cos(\phi - c_i) - 1)) \quad (16)$$

with A being the amplitude of the desired oscillation. The changes in Eq. (16) are motivated by the need to make the function f a function that lives on a circle, i.e., ψ_i are computed from a Gaussian function that lives on a circle (called von Mises function). The output system in Eq. (9) remains the same, except that we now identify the goal state g with a setpoint around which the oscillation takes place. Thus, by means of A , τ , and g , we can control amplitude, frequency, and setpoint of an oscillation independently.

DMPs for multidimensional motor systems

The previous section addressed only a one-dimensional motor system. If multiple dimensions are to be coordinated, e.g., as in the seven major degrees of freedom (DOFs) of a human arm, all that is required is to create a separate output system for every DOF (i.e., Eqs. (9) and (13)). The canonical system is shared across all DOFs. Thus, every DOF will have its own goal g (or amplitude A) and nonlinear function f . As all DOFs reference the same phase variable through the canonical system, it can be guaranteed that the DOFs remain properly coordinated throughout a movement, and in rhythmic movement, it is possible to create very complex stable phase relationship between the individual DOFs, e.g., as needed for biped locomotion. In comparison to previous work on modeling multidimensional oscillator systems for movement generation that required complex oscillator tuning to achieve phase locking and synchronization (e.g., Taga et al., 1991), our approach offers a drastic reduction of complexity.

Learning and optimization with DMPs

We can now address how the open parameters of DMPs are instantiated. We assume that goal g (or amplitude A) as well as the timing parameter τ is provided by some external behavioral constraints. Thus, all that is needed is to find the weights w_i in the nonlinear function f . Both supervised and reinforcement/optimization approaches are possible.

Supervised learning with DMPs

Given that f is a normalized basis function representation, linear in the coefficients of interest (i.e., w_i) (e.g., Bishop, 1995), a variety of learning algorithms exist to find w_i . In supervised learning scenario, we can suppose that we are given a sample trajectory $y_{\text{demo}}(t)$, $\dot{y}_{\text{demo}}(t)$, $\ddot{y}_{\text{demo}}(t)$ with duration T , for instance, from the demonstration of a teacher. Based on this information, a supervised learning problem results with the following target for f :

$$f_{\text{target}} = \tau\ddot{y}_{\text{demo}} - \alpha_z(\beta_z(g - y_{\text{demo}}) - \tau\dot{y}_{\text{demo}}) \quad (17)$$

In order to obtain a matching input for f_{target} , the canonical system needs to be integrated. For this purpose, in Eq. (10), the initial state of the canonical system is set to $x = 1$ before integration. An analogous procedure is performed for the rhythmic DMPs. The time constant τ is chosen such that the DMP with $f = 0$ achieves 95% convergence at $t = T$. With this procedure, a clean supervised learning problem is obtained over the time course of the movement to be approximated with training samples (x, f_{target}) .

For solving the function approximation problem, we chose a nonparametric regression technique from locally weighted learning (LWPR) (Vijayakumar and Schaal, 2000). This method allows us to determine the necessary number of basis functions N , their centers c_i , and bandwidth h_i automatically. In essence, every basis function ψ_i defines a small region in input space x , and point falling into this region are used to perform a linear regression analysis, which can be formalized as weighted regression (Atkeson et al., 1997). Predictions for a query point are generated by ψ_i -weighted average of the predictions of all local models. In simple words, we create a piecewise linear approximation of f_{target} , where each linear function piece belongs to one of the basis functions.

As evaluations of the suggested approach to movement primitives, in Ijspeert et al. (2002b), we demonstrated how a complex tennis forehand and tennis backhand swing can be learned from a human teacher, whose movements were captured at the joint level with an exoskeleton. Figure 8 illustrates imitation learning for a rhythmic trajectory using the phase oscillator DMP from Eqs. (15) and (16). The images in the top of Fig. 8 show four frames of the motion capture of a figure-8 pattern and its repetition on the humanoid robot after imitation learning of the trajectory. The plots in Fig. 9 demonstrate the motion captured and fitted trajectory of a bimanual drumming pattern, using 6 DOFs per arm. Note that rather complex phase relationships between the individual DOFs can be realized. For one joint angle, the right elbow joint (R_EB), Fig. 10 exemplifies the effect of various changes of parameter settings of the DMP (cf. also figure caption in Fig. 8). Here it is noteworthy how quickly the pattern converges to the new limit cycle attractor, and

that parameter changes do not change the movement pattern qualitatively, an effect that can be predicted theoretically (Schaal et al., 2003). The nonlinear function of each DMP employed 15 basis functions.

Optimization of DMPs

The linear parameterization of DMPs allows any form of parameter optimization, not just supervised learning. As an illustrative example, we considered a 1 DOF movement linear movement system

$$m\ddot{y} + b\dot{y} + ky = u \quad (18)$$

with mass m , damping b , and spring stiffness k . For point-to-point movement, we optimized the following criteria:

$$\left. \begin{array}{ll} \text{Minimum Jerk} & J = \int_0^T \dot{y}^2 dt \\ \text{Minimum Torque Change} & J = \int_0^T \ddot{u}^2 dt \\ \text{Minimum Endpoint Variance} & \\ \text{with signal dependent noise} & \\ J = \text{var}(y(T) - g) + \text{var}(\dot{y}(T)) & \end{array} \right\} \quad (19)$$

where in the case of the minimum-endpoint-variance criterion, we assumed signal dependent noise $u_{\text{noisy}} = (1 + \varepsilon)u$ and $\varepsilon \sim \text{Normal}(0, 0.04)$.

The results of these optimizations, using the Matlab optimization toolbox, are shown in Fig. 11. As a comparison, we superimposed the results of a minimum jerk trajectory in every plot. The velocity profiles obtained from the DMPs after optimization nicely coincide with what has been obtained in the original literature suggesting these optimization criteria (Flash and Hogan, 1985; Uno et al., 1989; Harris and Wolpert, 1998). What is the most important, however, was that it was essentially trivial to apply various optimization approaches to our dynamic systems representation of movement generation. Thus, we believe that these results are among the first that successfully combined dynamic systems representations to motor control and optimization approaches.

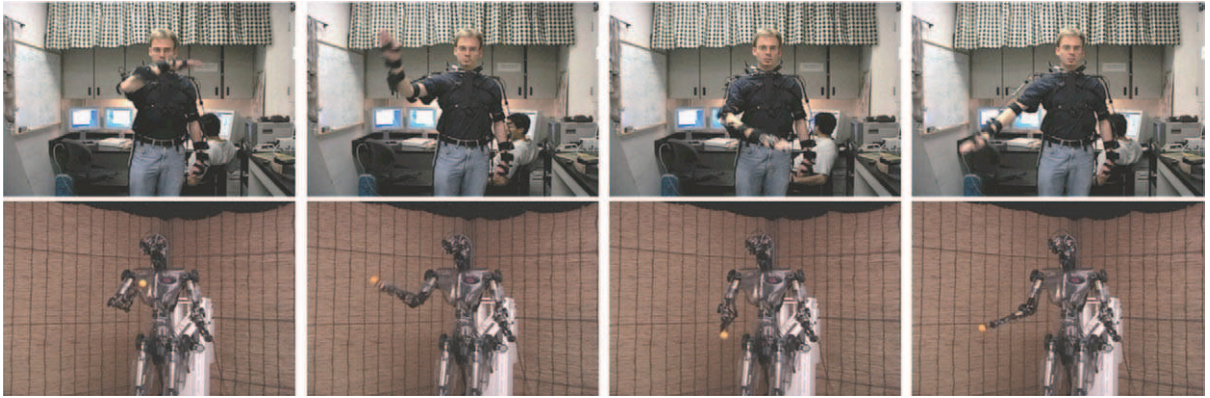


Fig. 8. Humanoid robot learning a figure-8 movement from a human demonstration.

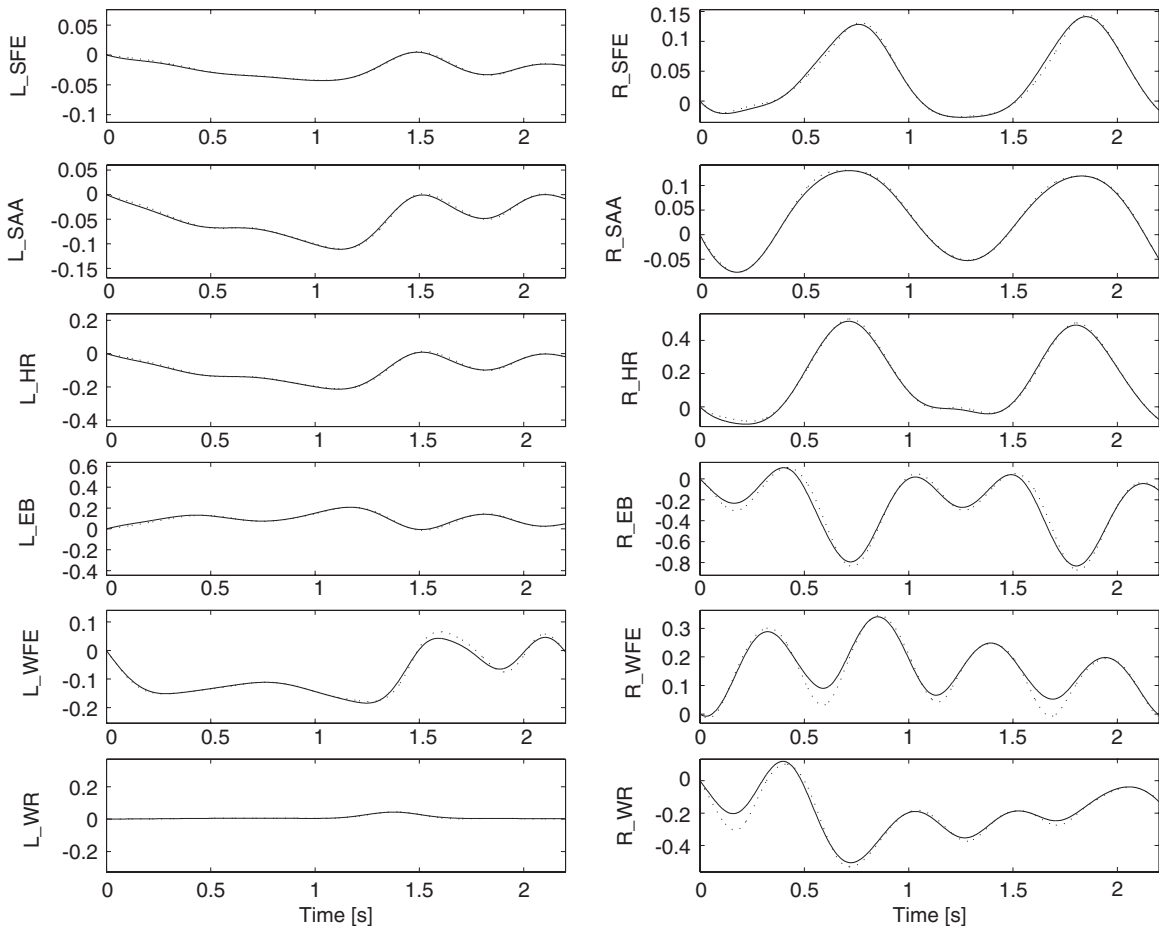


Fig. 9. Recorded drumming movement performed with both arms (6 DOFs per arm). The dotted lines and continuous lines correspond to one period of the demonstrated and learned trajectories, respectively — due to rather precise overlap, they are hardly distinguishable.

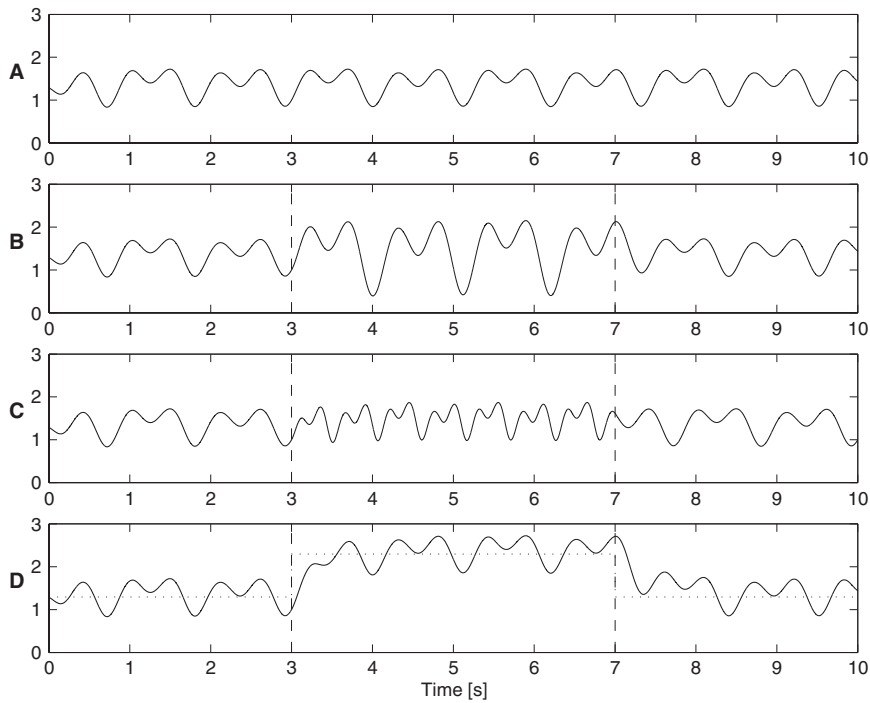


Fig. 10. Modification of the learned rhythmic drumming pattern (flexion/extension of the right elbow, R_EB). (A) Trajectory learned by the rhythmic DMP; (B) temporary modification with $A \leftarrow 2A$ in Eq. (16); (C): temporary modification with $\tau \leftarrow \tau/2$ in Eqs. (9) and (15); (D): temporary modification with $g \leftarrow g + 1$ in Eq. (9) (dotted line). Modified parameters were applied between $t = 3$ s and $t = 7$ s. Note that in all modifications, the movement patterns do not change qualitatively, and convergence to the new attractor under changed parameters is very rapid.

Discussion

This paper addressed a computational model for movement generation in the framework of dynamic systems approaches, but with a novel formulation that also allows applying optimization and learning approaches to motor control. We started by reviewing some of our own work that established evidence that periodic and point-to-point movements need to be investigated as separate functionalities of motor control, similar to the fact that point attractors and limit cycle attractors require different theoretical treatment in dynamic systems theory. We also emphasized that models of movement generation should not have explicit time dependency, similar to autonomous dynamic systems, in order to accommodate coupling and perturbation effects in an easy way. While these requirements favor a dynamic systems formulation of motor control, there has been no

acceptable computational framework so far that combines both the properties of dynamic systems approaches to motor control and the ease of applying learning and optimization approaches, which have played a dominant role in computational motor control over that last years (e.g., Shadmehr and Wise, 2005).

Our formulation of Dynamic Motor Primitives (DMP) offers a viable solution. Essentially DMPs are motivated by the VITE model of Bullock and Grossberg (1988) and other approaches that emphasized that movement should be driven by a difference vector between the current and the desired goal of a movement (for a review, see Shadmehr and Wise, 2005). DMPs create desired trajectories for a movement system out of the temporal evolution of autonomous nonlinear differential equations, i.e., the desired trajectory is created in real-time together with movement execution, and not as a preplanned entity. This

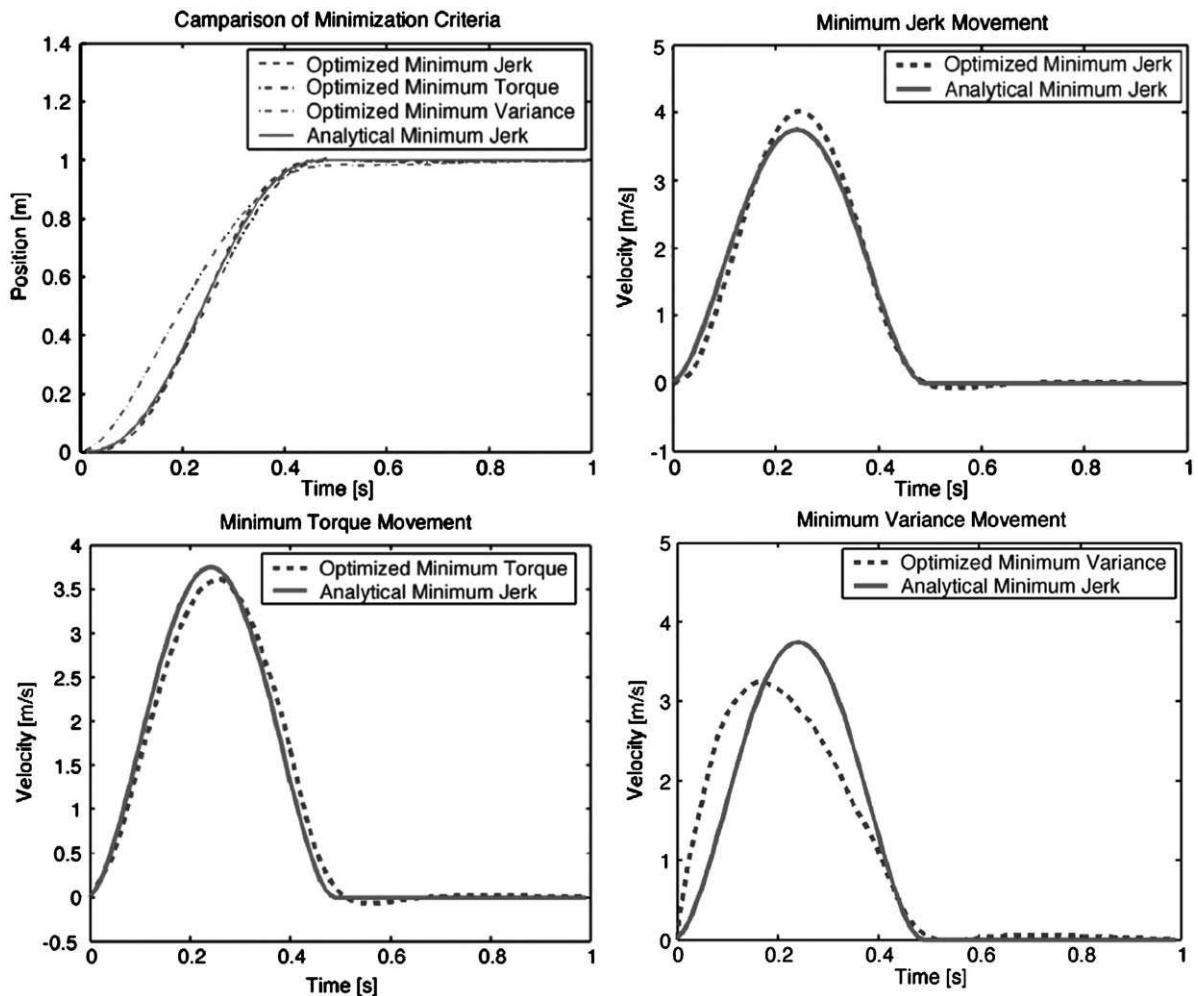


Fig. 11. Optimization results for DMPs for various criteria — see text for explanations.

real-time generation allows also real-time modification of the desired trajectory, a topic that we did not expand on in this paper, but which has been examined in previous work (Ijspeert et al., 2003). Such real-time modification is essential if one wishes to account for perception–action coupling or the reaction to perturbations during movement. Unlike other models of movement generation in the past, DMPs can represent rather complex movements in one simple coherent framework, e.g., a complete tennis forehand can be cast into one DMP. The complexity of a DMP is only limited by the number of basis functions that is provided to its core nonlinearity, a population-code

basis function approximator that could be generated by many areas of the primate brain. This line of modeling opens the interesting question of where and when a complex movement needs to be segmented into smaller pieces, i.e., how complex a movement primitive can be in biology. Another point worth highlighting is that DMPs can represent both discrete and rhythmic movement. Complex multi-DOF periodic patterns can be generated, where all contributing DOFs are easily synchronized and phase locked in arbitrary relationships. This property is unlike traditional coupled-oscillator models for multi-DOF movement generation, which usually have major difficulties

in modeling anything but synchronized in-phase and out-of-phase movement relationships. As a last point, DMPs can be scaled in time and space without losing the qualitative trajectory appearance that was originally coded in a DMP. For instance, a DMP coding a tennis forehand swing can easily create a very small and slow swing and a rather large and fast swing out of the exactly the same equations. We believe that this approach to modeling of movement could be a promising complement in many theories developed for human and primate motor control, and offers to revisit many previous movement models in one simple coherent framework.

Acknowledgments

This research was supported in part by National Science Foundation grants ECS-0325383, IIS-0312802, IIS-0082995, ECS-0326095, ANI-0224419, the DARPA program on Learning Locomotion, a NASA grant AC#98-516, an AFOSR grant on Intelligent Control, the ERATO Kawato Dynamic Brain Project funded by the Japanese Science and Technology Agency, and the ATR Computational Neuroscience Laboratories. We are very grateful for the insightful and thorough comments of the editors of this volume, which helped improving this article significantly.

References

- Abend, W., Bizzi, E. and Morasso, P. (1982) Human arm trajectory formation. *Brain*, 105: 331–348.
- Aboaf, E.W., Drucker, S.M. and Atkeson, C.G. (1989) Task-level robot learning: juggling a tennis ball more accurately. In: *Proceedings of IEEE International Conference on Robotics and Automation*. IEEE, Piscataway, NJ, May 14–19, Scottsdale, AZ, pp. 331–348.
- Adamovich, S.V., Levin, M.F. and Feldman, A.G. (1994) Merging different motor patterns: coordination between rhythmical and discrete single-joint. *Exp. Brain Res.*, 99: 325–337.
- Atkeson, C.G., Moore, A.W. and Schaal, S. (1997) Locally weighted learning. *Artif. Intell. Rev.*, 11: 11–73.
- Barto, A.G. and Mahadevan, S. (2003) Recent advances in hierarchical reinforcement learning. *Discrete Event Dyn. Syst.*, 13: 341–379.
- Bellman, R. (1957) *Dynamic Programming*. Princeton University Press, Princeton, NJ.
- Bertsekas, D.P. and Tsitsiklis, J.N. (1996) *Neuro-Dynamic Programming*. Athena Scientific, Belmont, MA.
- Bishop, C.M. (1995) *Neural Networks for Pattern Recognition*. Oxford University Press, New York.
- Buchanan, J.J., Park, J.H., Ryu, Y.U. and Shea, C.H. (2003) Discrete and cyclical units of action in a mixed target pair aiming task. *Exp. Brain Res.*, 150: 473–489.
- Bullock, D. and Grossberg, S. (1988) Neural dynamics of planned arm movements: emergent invariants and speed-accuracy properties during trajectory formation. *Psychol. Rev.*, 95: 49–90.
- Craig, J.J. (1986) *Introduction to Robotics*. Addison-Wesley, Reading, MA.
- De Rugy, A. and Sternad, D. (2003) Interaction between discrete and rhythmic movements: reaction time and phase of discrete movement initiation against oscillatory movement. *Brain Res.*
- Desmurget, M. and Grafton, S. (2000) Forward modeling allows feedback control for fast reaching movements. *Trends Cogn. Sci.*, 4: 423–431.
- Dyer, P. and McReynolds, S.R. (1970) *The Computation and Theory of Optimal Control*. Academic Press, New York.
- Flash, T. and Hogan, N. (1985) The coordination of arm movements: an experimentally confirmed mathematical model. *J. Neurosci.*, 5: 1688–1703.
- Georgopoulos, A.P. (1991) Higher order motor control. *Annu. Rev. Neurosci.*, 14: 361–377.
- Getting, P.A. (1985) Understanding central pattern generators: insights gained from the study of invertebrate systems. In: *Neurobiology of Vertebrate Locomotion*, Stockholm, pp. 361–377.
- Gribble, P.L. and Ostry, D.J. (1996) Origins of the power law relation between movement velocity and curvature: modeling the effects of muscle mechanics and limb dynamics. *J. Neurophysiol.*, 76: 2853–2860.
- Guckenheimer, J. and Holmes, P. (1983) *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, New York.
- Harris, C.M. and Wolpert, D.M. (1998) Signal-dependent noise determines motor planning. *Nature*, 394: 780–784.
- Hodgkin, A.L. and Huxley, A.F. (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117: 500–544.
- Hoff, B. and Arbib, M.A. (1993) Models of trajectory formation and temporal interaction of reach and grasp. *J. Mot. Behav.*, 25: 175–192.
- Hogan, N. (1984) An organizing principle for a class of voluntary movements. *J. Neurosci.*, 4: 2745–2754.
- Hollerbach, J.M. (1984) Dynamic scaling of manipulator trajectories. *Trans. ASME*, 106: 139–156.
- Ijspeert, A., Nakanishi, J. and Schaal, S. (2001) Trajectory formation for imitation with nonlinear dynamical systems. In: *IEEE International Conference on Intelligent Robots and Systems (IROS 2001)*. Weilea, HI, Oct. 29–Nov. 3, pp. 752–757.

- Ijspeert, A., Nakanishi, J. and Schaal, S. (2003) Learning attractor landscapes for learning motor primitives. In: Becker S., Thrun S. and Obermayer K. (Eds.), *Advances in Neural Information Processing Systems 15*. MIT Press, Cambridge, MA, pp. 1547–1554.
- Ijspeert, A.J., Crespi, A., Ryczko, D. and Cabelguen, J.M. (2007) From swimming to walking with a salamander robot driven by a spinal cord model. *Science*, 315: 1416–1420.
- Ijspeert, J.A., Nakanishi, J. and Schaal, S. (2002a) Learning rhythmic movements by demonstration using nonlinear oscillators. In: *IEEE International Conference on Intelligent Robots and Systems (IROS 2002)*. IEEE, Lausanne, Piscataway, NJ, Sept. 30–Oct. 4, pp. 958–963.
- Ijspeert, J.A., Nakanishi, J. and Schaal, S. (2002b) Movement imitation with nonlinear dynamical systems in humanoid robots. In: *International Conference on Robotics and Automation (ICRA2002)*. Washington, May 11–15.
- Ivry, R.B., Spencer, R.M., Zelaznik, H.N. and Diedrichsen, J. (2002) The cerebellum and event timing. *Ann. N.Y. Acad. Sci.*, 978: 302–317.
- Kawamura, S. and Fukao, N. (1994) Interpolation for input torque patterns obtained through learning control. In: *International Conference on Automation, Robotics and Computer Vision (ICARCV'94)*. Singapore, Nov. 8–11, pp. 183–191.
- Kawato, M. (1999) Internal models for motor control and trajectory planning. *Curr. Opin. Neurobiol.*, 9: 718–727.
- Kawato, M. and Wolpert, D. (1998) Internal models for motor control. *Novartis Found Symp.*, 218: 291–304.
- Keating, J.G. and Thach, W.T. (1997) No clock signal in the discharge of neurons in the deep cerebellar nuclei. *J. Neurophysiol.*, 77: 2232–2234.
- Kelso, J.A.S. (1995) *Dynamic Patterns: The Self-Organization of Brain and Behavior*. MIT Press, Cambridge, MA.
- Kopell, N. and Ermentrout, G.B. (1988) Coupled oscillators and the design of central pattern generators. *Math. Biosci.*, 90: 87–109.
- Kugler, P.N., Kelso, J.A.S. and Turvey, M.T. (1982) On control and co-ordination of naturally developing systems. In: Kelso J.A.S. and Clark J.E. (Eds.), *The Development of Movement Control and Coordination*. Wiley, New York, pp. 5–78.
- Lacquaniti, F., Terzuolo, C. and Viviani, P. (1983) The law relating the kinematic and figural aspects of drawing movements. *Acta Psychol.*, 54: 115–130.
- Marder, E. (2000) Motor pattern generation. *Curr. Opin. Neurobiol.*, 10: 691–698.
- Mohajerian, P., Mistry, M. and Schaal, S. (2004) Neuronal or spinal level interaction between rhythmic and discrete motion during multi-joint arm movement. In: *Abstracts of the 34th Meeting of the Society of Neuroscience*. San Diego, CA, Oct. 23–27.
- Morasso, P. (1981) Spatial control of arm movements. *Exp. Brain Res.*, 42: 223–227.
- Morasso, P. (1983) Three dimensional arm trajectories. *Biol. Cybern.*, 48: 187–194.
- Mussa-Ivaldi, F.A. (1988) Do neurons in the motor cortex encode movement direction? An alternative hypothesis. *Neurosci. Lett.*, 91: 106–111.
- Nakanishi, J., Morimoto, J., Endo, G., Cheng, G., Schaal, S. and Kawato, M. (2004) Learning from demonstration and adaptation of biped locomotion. *Robot. Auton. Syst.*, 47: 79–91.
- Pellizzer, G., Massey, J.T., Lurito, J.T. and Georgopoulos, A.P. (1992) Three-dimensional drawings in isometric conditions: planar segmentation of force trajectory. *Exp. Brain Res.*, 92: 326–337.
- Peters, J., Vijayakumar, S. and Schaal, S. (2005) Natural actor-critic. In: Gama J., Camacho R., Brazdil P., Jorge A. and Torgo L. (Eds.), *Proceedings of the 16th European Conference on Machine Learning (ECML 2005)*, 3720. Springer, Porto, Portugal, pp. 280–291 Oct. 3–7.
- Picard, N. and Strick, P.L. (2001) Imaging the premotor areas. *Curr. Opin. Neurobiol.*, 11: 663–672.
- Righetti, L. and Ijspeert, A. (2006) Design methodologies for central pattern generators: an application to crawling humanoids. In: *Proceedings of Robotics: Science and Systems*. MIT Press, Philadelphia, PA.
- Rizzi, A.A. and Koditschek, D.E. (1994) Further progress in robot juggling: solvable mirror laws. In: *IEEE International Conference on Robotics and Automation*, Vol. 4. San Diego, CA, May 8–13, pp. 2935–2940.
- Roberts, P.D. and Bell, C.C. (2000) Computational consequences of temporally asymmetric learning rules: II. Sensory image cancellation. *J. Comput. Neurosci.*, 9: 67–83.
- Schaal, S. and Atkeson, C.G. (1993) Open loop stable control strategies for robot juggling. In: *IEEE International Conference on Robotics and Automation*, Vol. 3. IEEE, Piscataway, NJ; Atlanta, GA, May 2–6, pp. 913–918.
- Schaal, S., Peters, J., Nakanishi, J. and Ijspeert, A. (2003) Control, planning, learning, and imitation with dynamic movement primitives. In: *Workshop on Bilateral Paradigms on Humans and Humanoids*. IEEE International Conference on Intelligent Robots and Systems (IROS 2003). Las Vegas, NV, Oct. 27–31.
- Schaal, S. and Sternad, D. (1998) Programmable pattern generators. In: *3rd International Conference on Computational Intelligence in Neuroscience*. Research Triangle Park, NC, Oct. 24–28, pp. 48–51.
- Schaal, S. and Sternad, D. (2001) Origins and violations of the 2/3 power law in rhythmic 3D movements. *Exp. Brain Res.*, 136: 60–72.
- Schaal, S., Sternad, D. and Atkeson, C.G. (1996) One-handed juggling: a dynamical approach to a rhythmic movement task. *J. Mot. Behav.*, 28: 165–183.
- Schaal, S., Sternad, D., Osu, R. and Kawato, M. (2004) Rhythmic movement is not discrete. *Nat. Neurosci.*, 7: 1137–1144.
- Schöner, G. (1990) A dynamic theory of coordination of discrete movement. *Biol. Cybern.*, 63: 257–270.
- Selverston, A.I. (1980) Are central pattern generators understandable? *Behav. Brain Sci.*, 3: 555–571.

- Shadmehr, R. and Wise, S.P. (2005) The computational neurobiology of reaching and pointing: a foundation for motor learning. MIT Press, Cambridge, MA.
- Smits-Engelsman, B.C., Van Galen, G.P. and Duysens, J. (2002) The breakdown of Fitts' law in rapid, reciprocal aiming movements. *Exp. Brain Res.*, 145: 222–230.
- Soechting, J.F. and Terzuolo, C.A. (1987a) Organization of arm movements in three dimensional space. Wrist motion is piecewise planar. *Neuroscience*, 23: 53–61.
- Soechting, J.F. and Terzuolo, C.A. (1987b) Organization of arm movements. Motion is segmented. *Neuroscience*, 23: 39–51.
- Spencer, R.M., Zelaznik, H.N., Diedrichsen, J. and Ivry, R.B. (2003) Disrupted timing of discontinuous but not continuous movements by cerebellar lesions. *Science*, 300: 1437–1439.
- Sternad, D., De Rugy, A., Pataky, T. and Dean, W.J. (2002) Interaction of discrete and rhythmic movements over a wide range of periods. *Exp. Brain Res.*, 147: 162–174.
- Sternad, D. and Dean, W.J. (2003) Rhythmic and discrete elements in multi-joint coordination. *Brain Res.*
- Sternad, D., Dean, W.J. and Schaal, S. (2000) Interaction of rhythmic and discrete pattern generators in single joint movements. *Hum. Mov. Sci.*, 19: 627–665.
- Sternad, D. and Schaal, D. (1999) Segmentation of endpoint trajectories does not imply segmented control. *Exp. Brain Res.*, 124: 118–136.
- Strogatz, S.H. (1994) *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Addison-Wesley, Reading, MA.
- Sutton, R.S. and Barto, A.G. (1998) *Reinforcement Learning: An Introduction*. MIT Press, Cambridge.
- Taga, G., Yamaguchi, Y. and Shimizu, H. (1991) Self-organized control of bipedal locomotion by neural oscillators in unpredictable environment. *Biol. Cybern.*, 65: 147–159.
- Tesauro, G. (1992) Temporal difference learning of backgammon strategy. In: Sleeman D. and Edwards P. (Eds.), *Proceedings of the Ninth International Workshop on Machine Learning*. Morgan Kaufmann, Aberdeen, Scotland, UK, July 1–3, pp. 451–457.
- Turvey, M.T. (1990) The challenge of a physical account of action: A personal view. In: Whiting, H.T.A., Meijer, O.G. and van Wieringen, P.C.W. (Eds.), *The Natural Physical Approach to Movement Control*. Amsterdam: Free University Press, Amsterdam, pp. 57–94.
- Vijayakumar, S. and Schaal, S. (2000) Locally weighted projection regression: an $O(n)$ algorithm for incremental real time learning in high dimensional spaces. In: *Proceedings of the 17th International Conference on Machine Learning (ICML 2000)*, Vol. 1. Stanford, CA, pp. 288–293.
- Viviani, P. (1986) Do units of motor action really exist? In: *Experimental Brain Research Series 15*. Springer, Berlin, pp. 828–845.
- Viviani, P. and Cenzato, M. (1985) Segmentation and coupling in complex movements. *J. Exp. Psychol. Hum. Percept. Perform.*, 11: 828–845.
- Viviani, P. and Flash, T. (1995) Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. *J. Exp. Psychol. Hum. Percept. Perform.*, 21: 32–53.
- Viviani, P. and Terzuolo, C. (1980) Space-time invariance in learned motor skills. In: Stelmach G.E. and Requin J. (Eds.), *Tutorials in Motor Behavior*. North-Holland, Amsterdam, pp. 525–533.
- Wann, J., Nimmo-Smith, I. and Wing, A.M. (1988) Relation 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.
- Wei, K., Wertman, G. and Sternad, D. (2003) Interactions between rhythmic and discrete components in a bimanual task. *Motor Control*, 7: 134–155.
- Williamson, M. (1998) Neural control of rhythmic arm movements. *Neural Netw.*, 11: 1379–1394.
- Wolpert, D.M. (1997) Computational approaches to motor control. *Trends Cogn. Sci.*, 1: 209–216.