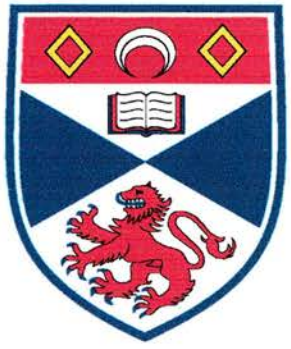
Non-Analytic Shifts In Smooth Goal-Directed Human Behaviour



A thesis to be submitted to the

UNIVERSITY OF ST ANDREWS

for the degree of

DOCTOR OF PHILOSOPHY

by

Adrian Peter Wale

School of Computer Science

University of St Andrews

April, 2006

Declarations

I, Adrian Peter Wale, hereby certify that this thesis, which is approximately 52,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date 28/4/06 *Signature of Candidate*

I was admitted as a research student in April, 2000, and as a candidate for the degree of Ph.D. April, 2001; the higher study for which this is a record was carried out in the University of St Andrews between 2000 and 2005.

Date 28/4/06 *Signature of Candidate*

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 28/4/06 *Signature of Supervisor*

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

Date 28/4/06 *Signature of Candidate*

Abstract

The thesis is aimed at finding the form of explanation and creating the associated computing methodology required to provide an effective computational explanation of smooth goal-directed behaviour. Smooth behaviour has typically been explained using analytic components. It is hypothesised that goal-directed smooth behaviour would benefit from a new hybrid form of explanation involving non-analytic as well as analytic aspects in order to account better for the type of plastic and persistent adaptation seen in natural agent behaviour. The thesis investigates strategies used by animate agents to control the shape of their motor actions in pursuing goals with a view to establishing their components.

The hypothesis that there are non-analytic components in natural smooth goal-directed behaviour is empirically tested in the arena of human hand movement kinematics in a variety of experimental settings. The presence of these components in the behaviour is demonstrated in various ways involving the agent constantly redirecting itself so as to remain projecting non-analytically through the goal.

The demonstrations begin with an investigation of a simplest base case of a behaviour that involves a smooth merge between two parallel linear movements. A further series of experiments generalizes the methodology to provide successful predictions for cases involving different ratios for the central movement, different directions at the ends of the movement, and with smooth external perturbations added to the movement.

Computing and cognitive applications of the methodology are given. It is concluded that the new hybrid form of explanation and methodology is supported by the empirical evidence as being an appropriate one in many cases for providing an effective computational explanation of goal-directed smooth behaviour.

Acknowledgements

Many people have contributed to the completion and correction of the thesis over the years and to them I offer my sincere thanks. Any errors that remain are, of course, my own.

In particular I'd like to thank: my supervisor, Mike Weir, for his tireless efforts throughout; my parents, Mary and Peter Wale, for their love and support; Gordon Milligan, for the use of his shift curve generator program; Jon Lewis and Graeme Bell, of the Intelligent Computation Group, for their help and advice; all of my brave subjects for giving up their time to participate; everyone else at the School of Computer Science who played a part in making this possible.

The thesis was funded by the EPSRC.

Contents

Chapter 1

The Search for An Effective Computational Explanation of Smooth Goal-Directed Behaviour

1.1. Aristotle's Four Forms of Causal Explanation

A primary aim of the thesis as a whole is to discover the form of causal explanation necessary for the computational study of goal-directed behaviour. This chapter provides the historical background to the thesis. It traces the history of causal explanations of goal-directed behaviour from Aristotle to the present day. It is argued that a paradigm shift in the study of motion in general in the Renaissance led to a form of explanation being used in contemporary goal-directed behaviour research that may be inadequate. It is then argued that the reintroduction of the form of explanation lost to dynamics in the Renaissance may be necessary to explain goal-directed behaviour. The thesis investigates the validity of these claims.

To observe the origins of the debate we are required to go back over two thousand years, perhaps somewhat predictably, to ancient Greece. Aristotle, amongst his many other contributions to science and philosophy, provided an influential system of causal explanation. For him, causal explanation concerned the explanation of change in general. His system, which he described in Physics Book II, comprised of four complementary forms of causal explanation: material cause, formal cause, efficient cause and final cause.

The material cause involves explanation by reference to the material substance out of which something is made. As Aristotle (Hankinson, 1998, p.132) puts it, 'the cause is said to be the existing thing out of which something comes to be, e.g. the bronze of the statue, or the silver of the phial, and their genera.'

The formal cause involves explanation by reference to the features that distinguish some entity from any other. It is about the blueprint or pattern that causes something to exist. As Juarrero (1999, p.2) put it, formal cause concerns 'that which makes anything that sort of thing and no other.' In Aristotle's words (Hankinson, 1998, p.132) formal cause concerns the 'form or the template (*paradeigma*)' and says that 'this is the formula (*logos*) of the what-it-is-to-be, and its genera...' He gives the example of the ratio 2:1 being the formal cause of the octave. Although he is not tying the notion of formal cause purely to mathematics. Formal cause is at work everywhere, including in biological agents. As Hankinson (1998, p.135) puts it, 'as far as living things are concerned, it is the dynamic package of internal structural attributes in virtue of which the process of maturation unfolds.'

The efficient cause involves explanation by reference to the significant forces that brought about the existence of some thing or event. As Aristotle (Hankinson, 1998, p.132) said, the efficient cause is '...in general the agent of the thing produced, and the changer of the thing changed.' One example would be that of an apple falling from a tree and hitting the ground due to the force of gravity. Another would be a pool ball entering a pocket because the force imparted upon it by another ball propelled it there.

The final cause (see Hassing, 1997) involves explanation by reference to 'the goal or purpose toward which something aims' (Juarrero, 1999, p.2). As Aristotle (Hankinson, 1998, p.132) described it, 'Moreover, there is the end (*telos*). This is that for the sake of which...' He gives the example of a person walking in order to be healthy. Aristotle believed that final cause and formal cause are related. As Dusek (2001, p.81) says, 'Often the final and formal causes would coincide. The goal of the entity in development would be its fully realized structure.'

According to Aristotle all four causes are at work in any process, whether inanimate or animate (Dusek, 2001, p.81). Woodfield (1976, p.4) provides an example of this complementarity by suggesting four answers to the question 'Why did this building come to be as it is?' The material cause version would be 'Because of the bricks and mortar.' In terms of the formal cause the answer would be 'Because it is a house.' An efficient cause account would say 'Because the builder made it.' Whilst a final cause account would read 'Because it is for living in.'

1.2. The Rise of Final Cause in the Study of Motion

The work of Aristotle was influential during the Hellenistic Period, which is defined by scholars as the time from the death of Alexander the Great in 323BC to the end of the Roman Republic in 31BC, and this influence continued into the time of Imperial Rome (Honderich, 1995, pp. 349-350).

However, Aristotelianism in the empire was to decline during the first millennium AD although some scholars still studied the works of Aristotle and his followers. A particular blow came in 529AD when the Emperor Justinian closed the schools of philosophy in Athens (Honderich, 1995, p.50; Roberts, 1992, p.243). It was not until the 12th Century that Aristotelianism staged a revival in Europe as the work of Arabic scholars such as Avicenna and Averroës, commenting on Aristotle, was translated into Latin and entered Europe via Spain (Honderich, 1995, p.50-51).

The first response of the Christian Church to this revival was somewhat alarmist and reactionary. As Honderich (1995, p.51) puts it, 'Initially, Aristotle was seen as a threat to Christian orthodoxy, and in 1210 the Council of Paris banned the study of his natural philosophy and threatened to excommunicate anyone who studied it.'

Later in the 13th Century many of Aristotle's works, which had previously been lost to the West, were discovered in Constantinople. Two Dominicans, Bruno the Great and Thomas Aquinas, studied these works and attempted to integrate Aristotle's ideas as far as possible with the religious and scientific beliefs of the age. Albert attempted to reinterpret Aristotle's works on the basis of what he believed Aristotle would argue

for if he were an enlightened 13th Century scholar. Aquinas, meanwhile, attempted to provide a synthesis of many aspects of Aristotle's work with the latest in 13th Century Christian and scientific thought. But the synthesis lost much of Aristotle's real meaning. As Honderich (1995, p.761) puts it, 'Aquinas pays respectful attention to pagan philosophers and chiefly to Aristotle, whose works he expounds in detail. But whenever he writes in his own voice, as an ordained teacher of theology, Aquinas systematically transforms every Aristotelian doctrine he touches, often in a direction quite opposed to Aristotle's own intention.'

In terms of the study of motion such reinterpretation of Aristotle led to an emphasis of final cause explanation over the other forms of causal explanation. This was because the 'concept of final causes supplied a convenient foundation for religiously orientated teleology,' as Honderich (1995, p.262) puts it. Through the Christianising of final cause explanation Aristotle's pagan view of a world of objects moving to bring about the ends required by their nature had been transformed into a world of objects with God-given purposes.

Des Chene (1996, p.168) provides an example of the type of explanation that resulted from this version of final cause, in which theorists attempted to make an educated guess about the ends provided for things by the Creator. He quotes the French bishop, Abra de Raconis, who wrote, 'The two most important ends to which the sea is instituted are, first, that it should be the common domicile of fish, and second that in it there should be navigation to provide commerce and necessary goods; but to both ends its saltiness is most fit, since saltiness keeps the sea from putrefying and makes it stronger and denser so as to hold the greater weight of ships.' It is ascriptions of

purpose such as these that Voltaire famously satirised in *Candide*. At one point Candide's tutor Pangloss teaches him that 'Noses, you observe, were made to support spectacles: consequently we have spectacles. Legs, it is plain, were created to wear breeches, and are supplied with them,' (Voltaire, 1759, p.2).

Whilst Voltaire may have been able to popularly ridicule these explanations in the middle of the 18th Century, such distortions of Aristotle's work had dominated the study of motion in the period leading up to the Renaissance. The work of the pagan Aristotle had become a pillar of the Christian orthodoxy.

1.3. The Replacement of Final Cause by Efficient Cause

The Renaissance saw the general hegemony of the Christian Church over science and mathematics increasingly challenged. Christianised Aristotelianism came under attack on both methodological and empirical grounds. The new advances forced the Church to question whether it should ever accept the logic of science over the logic of religious authority (see Blackwell, 1991, for an extended discussion of this issue).

The development of the modern scientific method during the Renaissance was one of the key challenges. Francis Bacon argued that scientific knowledge is obtained through the processes of induction and experimentation. The process of induction involves general inferences being made from specific experimental results. As Peltonen (1996, p.16) writes with regards to Bacon, 'He wanted to replace the Aristotelian method of syllogism with an entirely new scientific method. According to Bacon, a crucial weakness in the old Aristotelian logic was that it jumped from

empirical particulars to first principles (axioms), which formed the premises of deductive reasoning. But Bacon believed that the most general axioms should form the end rather than the beginning of scientific inference, and his own methodology was designed to avoid Aristotle's mistake.'

Syllogism involved propositions such as 'All men are mortal' and 'Greeks are men' leading to the conclusion 'Greeks are mortal' to use the example given by Honderich (1995, p.862). Induction, on the other hand, involved reasoning like '[A]ll objects we know of attract each other with a force inversely proportional to the square of the distance between them, so perhaps they all do so, and always will do so,' to use the example of Blackburn (1994, p.192). Syllogisms provided deductively valid inferences unlike the results of the inductive method but the power of the inductive method to drive scientific progress and develop an increasingly more accurate view of the world was undeniable.

Advances in the science of motion were at the forefront of the empirical attack on the old worldview. The starting point of the Scientific Revolution in the Western World is conventionally considered to be 1543AD (Fernández-Armesto, 1997, 150-151). This is the date of publication of *De Revolutionibus Orbium Coelestium*, which translates as *On The Revolutions of Celestial Orbs*, by Nicholaus Copernicus. It was written in the 1530s and was circulated amongst the academic community, but Copernicus, afraid of reprisals by the Church, would not allow it to be published (Hellemans & Bunch, 1988, p.108). The reason for this fear was that his work contradicted the idea that the Earth was at the centre of the cosmos. This idea, which was developed into a complex system of epicycles by Ptolemy, still formed a part of the Aristotelian pillar

of Christian orthodoxy. For Copernicus, theory and observation told him that the planets, including the Earth, circled the Sun, and he thus took the theoretical step from a geocentric cosmology to a heliocentric cosmology.

The idea that the Earth and the other planets revolved around the sun was slow to catch on, even after publication, but was aided by Kepler's laws of planetary motion and the astronomical observations of Galileo. As Fernández-Armesto (1997, p.151) says, '...Copernicus's system took a long time to command assent. Still, refined and vindicated by Kepler in the early years of the seventeenth century, it gradually remoulded visions of the cosmos...'

The reaction of the Christian Church was to say very little about the subject for some time until a Decree of the Congregation of the Index in 1616 formally condemned Copernicanism. As Blackwell (1991, p.54) says, 'So seventy years of near silence on the question of the relation of Copernicanism to the Bible is followed by a period of a little more than two years (December 1613 to March 1616) in which the issue evolves quickly from a casual after-breakfast topic of conversation to a formal condemnation.' The reason for this sudden shift, argues Blackwell, was the fact that scientists like Kepler and Galileo had begun to question the very authority of the Church and scripture. As Hellemans & Bunch (1988, p.134) say, 'In letters Galileo argued that not everything in the Bible should be taken literally. Galileo further argued that the Church had the burden of proving that the Copernican system is wrong.' Galileo travelled to Rome to defend his views, the result being that the works of Copernicus were themselves withdrawn until corrected to the Church's liking. Galileo, however,

continued to hold Copernican views and was put on trial in 1633, with the Church forcing him to publicly recant his views and putting him under house arrest.

It was Galileo, though, who did more than anyone to overturn the orthodoxy. To start with he considered Aristotle's belief that heavier bodies fall faster than lighter ones and disproved it with a thought experiment. As Hahn (1998, p.76) describes it, 'He contemplated a cannonball in free fall. Next, he considered it to be cut into two pieces, and he thought of both pieces falling side by side...From the point of view of the fall, nothing had changed. Therefore, the cannonballs in the two situations must drop with the same speed. According to Aristotle, however, each of the two separate halves of the ball – being lighter – must drop more slowly than the ball as a whole. Aristotle had to be wrong. Galileo had discovered that all bodies fall in the same way, regardless of their weight, assuming that they are heavy enough for air resistance to be ignored.' In his subsequent studies of motion he performed simple, elegant experiments such as rolling balls down inclined planes and timing them. As Hellemans & Bunch (1988, p.119) describe, such experiments allowed him to deduce three laws:

1) Any body moving on a horizontal plane will continue at the same speed until a force opposes it.

2) In a vacuum, all bodies fall at the same speed, no matter what their weight or constitution.

3) A body falling freely or rolling down an inclined plane undergoes uniform acceleration.

It was these laws that would form the basis for Isaac Newton to bring about the final replacement of the Christianised Aristotelian final cause account by the efficient cause account of modern physics. Central to the project was the conceptualisation of the motion of a body and changes in that motion as being the result of external forces acting on the body. Newton provided three Universal Laws of Motion which built on and surpassed Galileo's laws. As Cohen & Smith (2002, pp. 64-68) report them:

1) Every body perseveres in its state of being at rest or of moving uniformly straight forward [i.e. moving uniformly forward in a straight line] except insofar as it is compelled to change its state by forces impressed.

2) A change in motion is proportional to the motive force impressed and this motion is directed along the straight line in which this force is impressed.

3) To any action there is always an equal and opposite reaction.

Newton was also a major figure in the fundamental advances in the mathematics of the science of motion during the latter stages of the Renaissance. Newtonian calculus and associated tools such as the Taylor series were constructed to provide the new science, based on efficient causal explanation, with a mathematical basis. In dynamics, of particular importance was, as Gullberg (1997, p.676) puts it, 'the clarification and use of the concepts of velocity, acceleration, etc.' by the mathematicians of that time.

By the end of the Renaissance the science of motion had been provided with empirically based theory and mathematical tools that scientists believed could, in principle, explain every motion. This view, taken to its most extreme form, led to the Laplacian idea of a fully state-determined universe. As Dusek (2001, p.81-82) puts it,

'Laplacian determinism, formulated by the physicist Laplace around 1800, generalized efficient causal explanation to claim that if a sufficiently large mind "the Laplacian spirit" knew all positions and motions of particles in the universe it could predict every future event in the universe.' People were left in no doubt that the new science, based on efficient cause, was superior to the old final cause science of Christianised Aristotelianism. Because a successful empirical and mathematical basis had been provided for efficient cause, and no such basis had been provided for any version of final cause explanation, efficient cause became the dominant form of causal explanation for motion in general.

1.4. Causal Explanation in the Present Day Study of Goal-Directed Behaviour

Since the Renaissance efficient cause explanation has come to dominate scientific thought about causation. As Weir (1984, p.42) says, 'Indeed, the very notion of cause itself has become implicitly understood as meaning efficient cause. If the cause of some behaviour is asked for today, the expected answer is the thing that produced the behaviour. For example, the cause of an avalanche may be an initial small disturbance to a part of the snow. The initial disturbance producing the avalanche is prior to the avalanche, hence it is an efficient cause.'

With the rise of efficient cause and the abandonment of final cause, contemporary researchers typically use approaches in which all moving objects, whether inanimate or animate, are subject to the same efficient causal physical laws. As Juarrero (1999, p.21) puts it, 'The basic particles populating this mechanistic world of modern physics are independent of each other and do not interact to achieve a goal or purpose as they

would were Aristotle's final cause still in place. Once Aristotle's final causes were disallowed, Newtonian mechanics left no room for either objects in the external world or end-states to serve as intentional objects of desire and goals of action.'

In the Renaissance, experiments such as those performed by Galileo provided the empirical groundwork for the eventual replacement of Aristotelian dynamics with that of modern dynamics. The experiments that Galileo performed were required in order to supplant the status quo of his age. The empirical evidence supported the truth of the new approach and so the new approach eventually became the status quo. In the present day it would require an analogous demonstration to replace modern dynamics with an account that includes final cause explanation (as complementary to efficient cause explanation) in animate motion. The lack of any such demonstration means that there has never been any substantive technical or scientific distinction demonstrated to warrant to reintroducing final cause explanations in the case of animate motion.

The lack of a demonstration that animate motion is a special case in terms of requiring explanations that involve final causes has left contemporary research into the goal-directed behaviour of animate agents typically using efficient cause explanations alone and typically not using final cause explanations.

1.5. The Possibility of Reintroducing Final Cause Explanation into the Study of Smooth Goal-Directed Behaviour

The paradigm shift (see Kuhn, 1996, for an extended discussion of this concept) in the study of motion during the Renaissance occurred because the empirical and

mathematical basis for efficient cause physics was superior to that of Aristotelian Physics. The work of Copernicus, Galileo, Newton and many others had seen to it that a powerful new research program had built up a head of steam and had left explanations involving final cause trailing way behind.

As Woodfield (1976, pp.8-9) puts it, 'The main reason why Aristotle's doctrine faded out is simply that the new men of science stopped asking teleological questions. To them, final causes were scientifically irrelevant,' and, '[It] was not a reasoned refutation of the doctrine of final causes, but rather a switch of interest away from it.'

The fact that final cause explanation was never refuted means that it is still legitimate today to hypothesise that the study of goal-directed behaviour in particular might benefit technically from some form of final cause explanation.

By virtue of being physical bodies, both animate agents and inanimate entities may be shown to obey physical forces. Consequently, efficient causation influences animate agent motion. There may nevertheless also be a role for final causation in how the agent directs the organisation of physical forces to produce desirable motion. The hypothesis will be that both efficient and final causation may be technically relevant to animate goal-directed behaviour.

However, it must be clearly understood by the reader that the version of final cause being suggested here does not constitute a backwards version of efficient cause where the effect temporally precedes the cause. But rather, the future being referred to is an

intended future state of affairs that the agent has in mind when organising their behaviour in the present.

The goal-directed behaviour chosen for analysis is that of human motion, i.e. human kinematics. It is common for studies of human kinematics to analyse the smooth aspects of the behaviour (for example, Novak et al, 2000) and this approach will be taken here.

To test the hypothesis it would first be necessary to develop a technical approach involving final cause explanation that would allow computational and experimental predictions to be made that are distinct from those made solely on the basis of efficient cause explanation alone.

It would then be necessary to test the hypothesis empirically. Ideally the design should involve an experiment or series of experiments that could analyse the contributions of the types of cause involved.

The empirical aim is to provide an initial example of an experiment where the hybrid approach involving efficient and final cause explanation of smooth behaviour can be shown to provide a more effective and predictive explanation than the purely efficient cause approach and to then generalize any success to several new settings. This would provide a small but significant contribution towards demonstrating whether or not final cause has a useful role to play in the future scientific study of smooth goal-directed behaviour.

1.6. The Thesis

My background as a cognitive scientist has naturally reinforced in me a belief in the power of interdisciplinarity as a research strategy in the study of the intelligent behaviour. It was perhaps predictable then that a thesis on goal-directed behaviour would retain an interdisciplinary approach. What was less predictable was the particular disciplinary nexus that I would find myself, out of necessity, working at in support of the computing science involved. It is a nexus that includes history of science, philosophy of causation, smooth computation, and experimental kinematics.

The central aim of the thesis as a whole is to demonstrate what is necessary to provide an effective computational explanation of smooth goal-directed behaviour with particular reference to human kinematics. (Where kinematics is 'the study of the motion of bodies without reference to mass or force,' as Borowski and Borwein (1999, p.321), define it. See Zatsiorsky (1997) for an introduction to the study of human kinematics and Novak et al. (2000) for a kinematics paper of particular relevance to the present thesis).

In the broad the project will involve: providing the historical context of the thesis; discussing the status of contemporary approaches; defining a theoretical approach that involves final cause explanation that is an alternative to the efficient cause approach; developing new mathematical and computational tools of analysis capable of identifying behaviour associated with both of the competing approaches; describing and analysing human kinematics experiments designed to be as decisive as possible on the issue of which approach is superior for the effective explanation of the smooth

goal-directed tasks investigated in the thesis; discussing the implications of the results and suggesting future research into smooth goal-directed behaviour.

As has been discussed earlier in this chapter, the prevailing paradigm since the Renaissance for the study of motion in general has been that based on efficient cause explanations. Final cause explanations have fallen into disuse. The reason for this lies in the fact that the efficient cause approach has a scientific and computational basis to support it, unlike final cause. It will be shown in the thesis, though, that final cause explanation can also be placed on a hard scientific and computational footing. It then becomes a case of deciding which theory, the purely efficient cause approach or the hybrid approach involving both efficient cause and final cause, is the most effective and predictive in explaining the phenomena observed in the smooth goal-directed tasks investigated in the thesis.

The thesis presents the application of custom-built computation and mathematics for analysing the experimental results. These tools are designed to be simple enough for use by scientists in the everyday study of smooth goal-directed behaviour. The mathematics and computation used is likely to be unfamiliar to the reader at first. However, it is clearly verifiable and, as will be demonstrated, it is also necessary for the analysis of smooth goal-directed behaviour.

In the thesis a series of experiments will be presented that provide empirical data on goal-directed behaviour. The experiments are designed to be as decisive as possible in choosing between an efficient cause approach to the study of goal-directed behaviour and an approach that reintroduces final cause explanation. The experiments discussed

in the thesis all concern simple goal-directed human hand movements that are effectively displayed on a computer screen using a mouse. The initial experiment attempts to look at a simplest base case of a goal-directed human hand movement that involves a smooth merge between two parallel linear movements. This simple design can be thought of as a fundamental experiment in the scientific study of smooth goal-directed behaviour. It can be seen as analogous to Galileo's simple, fundamental experiment of Renaissance Physics involving rolling balls down inclined planes. Further experiments then increase the complexity of the movement under scrutiny step-by-step along a variety of experimental dimensions, thus broadening the empirical base for which the hybrid account can be seen to be preferable.

This chapter has outlined the thesis and explored its historical context. The following chapter looks at contemporary research into goal-directed behaviour. It is argued that the form of explanation used in contemporary approaches may be inadequate because final cause explanation, in the strong sense of the goal of the agent as being a distinctively effective influence on the behaviour of the agent, is not used. However, it is believed that although these approaches presently concentrate on the efficient causal influences of the behaviour they are also all basically compatible with final cause accounts. They only fail to take final cause into account because of the lack of an available empirical, mathematical and computational basis.

The next two chapters develop a theoretical and computational account of smooth goal-directed behaviour that includes final cause as well as efficient cause as a distinctive influence on the behaviour. Chapter 3 develops a theoretical account of smooth goal-directed behaviour that emphasises the role of plastic and persistent

changes in course towards the goal during steering tasks. Smooth goal-directed behaviour is analysed in terms of the concepts of goal, path and agency. Chapter 4 provides a computational account explaining how the characterisation of changes in course as non-analytic forms called shifts can potentially be used in smooth goal-directed behaviour research. It presents examples of the derivative profiles associated with the type of path thought to underlie smooth goal-directed behaviour and describes the fundamental properties of such path types.

Currently there are no well-accepted general non-analytic principles of smooth goal-directed behaviour that make predictions about the kinematics involved. The next two chapters in the thesis are aimed at beginning the process of uncovering such non-analytic principles of animate motion that would be analogous to the laws of inanimate motion discovered by Galileo and Newton. The path type hypothesised to underlie smooth goal-directed behaviour in Chapters 3 and 4 is exploited in Chapters 5 and 6 to suggest general laws for smooth goal-directed agent kinematics which are then tested empirically.

Chapter 5 reports the first empirical evidence that course changes exist in simple human smooth goal-directed behaviour that can be predicted using shift theory. The first experiment provides a simplest base case of a behaviour that involves a smooth merge between two parallel linear movements with equal width and height between the start and end of the motion being analysed. It is shown that a modified form of the hypothesised type of trajectory underlies the smooth goal-directed behaviour investigated here.

At this stage, it is important to be aware of what the techniques used to analyse the experiments described here provide over and above that provided by standard differential kinematics approaches. Standard approaches to differential kinematics analyse the dynamical effect of anatomical connections along the kinematic chain. In terms of temporal causation, they provide an efficient causal account of anatomical dynamics with no reference to the effect of a goal. As such, they can provide a retrospective accounting for the sequence of past pushes in what has happened, or an account of what would happen if the agent moved in a particular way. This type of methodology, however, is not, by itself, able to predict the sequence of values in voluntary direction by the agent in pursuit of goals. This is because the goal may affect the sequence through shaping the behaviour. Consequently, both the goal and how it shapes behaviour need to be predictable. The approach taken in the thesis, on the other hand, allows investigation of voluntary movement, enabling prediction of future sequences of velocities by taking into account the hypothesised goal of the subject.

Chapter 6 shows the broadening of this empirical base so as to demonstrate the potential for generalization of the new approach. It is shown that the path types found in the initial experiments reported in Chapter 5 are still present in cases involving different width to height ratios for the movement, different directions at the two ends of the movement, and also in cases where the subjects are being sent off course during the motion by a disturbance from the steered path.

Chapter 7 describes further generalizations of the new approach that may be possible. It then goes on to describe computational applications of the theory and methodology

to areas such as human-computer interaction, computer animation and robotics, which can be provided with an extra level of smoothness and flexibility in their goal-directed behaviour. It also describes cognitive application including the identification of aiming and firing behaviours in skilled goal-directed actions which may, for example, be involved in sporting activities.

The chapter shows the link between the ability to reliably provide predictable fits for the goal-directed paths being observed (demonstrated in Chapter 5 and generalised to new settings in Chapter 6) and the applications examined. For example, the approach may be used in kinematics studies for the production of hypotheses and the analysis of sports performance, potentially leading to improved performance and decreased injuries. This chapter also describes how the new method may be an improvement on existing methods and provides a description of applications that may be possible as a result of such generalizations by, for example, providing new functions that may be used by artificially animate agents to drive their goal-directed behaviour, e.g. a robot moving its arm to grasp a moving object.

Finally, the Epilogue provides an overview of the core model and the main results and implications of the thesis.

1.7. Summary

Aristotle believed that every process, whether inanimate or animate, required explanation using four forms of causal explanation (material, formal, efficient and final). A Christianised form of final cause explanation was the favoured form for

explaining goal-directed behaviour in the Western World from Aquinas to the Renaissance.

In the Renaissance, because efficient causal explanation had a successful rigorous basis and final cause explanation lacked such a basis, efficient cause became the dominant form of causal explanation for motion in general. As a result, contemporary research into the smooth goal-directed behaviour of animate agents typically uses efficient cause explanation and typically does not use final cause explanation.

The main overall aim of the thesis is to uncover what form of explanation and methodology provides an effective computational explanation of smooth goal-directed behaviour. It will attempt to take the first steps towards providing non-analytic principles of smooth goal-directed behaviour that can then be used to allow the construction of theoretical models capable of empirical testing. With this in place a variety of computational and cognitive applications may be possible.

Chapter 2

Contemporary Scientific Approaches to Goal-Directed Behaviour

2.1. Introduction

In this chapter a variety of scientific approaches to goal-directed behaviour are discussed including evolutionary biology, reinforcement learning, neuroscience, cognitive science, complex systems theory, ecological psychology, kinematics and control systems theory.

Approaches to goal-directed behaviour can be seen to commonly draw upon four different types of distinction in terms of what they see as characteristic of goal-directed behaviour as opposed to other forms of behaviour (see Weir, 1984, pp.59-121). One type of theory focuses on the behaviour itself (for example: Rosenblueth, Wiener, & Bigelow, 1943; Russell, 1945; Braithwaite, 1946; Sommerhoff, 1974). A second type of theory characterises the goal as a desired or realised state of affairs, thus it is the ends of the behaviour that are the important thing (for example:

Braithwaite, 1946; Sommerhoff 1974; Woodfield, 1976). A third type of theory sees a particular type of internal state of the producer of the behaviour as being characteristic of goal-directed behaviour (for example: Taylor, 1966; Woodfield, 1976). The fourth theory type sees the particular form of explanation used as being the important thing (for example: Taylor, 1964; Wright, 1968). The scientific approaches discussed in this chapter overlap across these four types, seeing several characteristics of goal-directed behaviour as being distinctive.

Each of the scientific approaches to goal-directed behaviour discussed below focuses on particular aspects of goal-directed behaviour and therefore none of them on their own provide a complete causal explanation of goal-directed behaviour.

Further, it will be argued that, even when taken together, they do not presently provide a complete explanation of goal-directed behaviour. This is because the causal support present for the approaches is essentially that of efficient cause.

The causal stance adopted in this thesis is relatively rare, i.e. the inclusion of final cause as well as efficient cause. It is an interesting one though, with a variety of applications, and has support for example from Weir (1984) and more recently Juarrero (1999).

The aim of Chapter 2 is to provide an overview concerning the issue of whether this causal stance fundamentally conflicts with the findings and characteristic aims of scientific approaches most relevant to goal-directed behaviour.

It will be put forward that the characteristic results or descriptions already established by such approaches are complemented by, rather than clash with, the extra account that would be provided by the additional use of final causes.

2.2. Evolutionary Biology

Aristotle's theory of animate motion saw the behaviour of animate agents as predominantly concerned with final cause. As Honderich (1995, p.56) explains, 'Aristotle was convinced that teleological explanation was the key to the study of natural organisms. What determined a thing's nature was what counted as its successful operation: its achieving what it is good for it to achieve (as is implicit in his ethical writings). These goals, and being organized so as to achieve them, is what makes the species the one it is' and 'The distinctive goal of each biological kind is what determines its respective essence.' For Aristotle, the defining goal of the human species was to live a life of rational activity and all other goals were subsidiary to that end. (For a full discussion of Aristotle's theory of biology see Lennox, 2001).

As was described in the previous chapter, the Aristotelian version of final cause was Christianised in the time leading up to the Renaissance, and eventually left out of the physical science of motion altogether in favour of purely efficient cause accounts. But in the case of animate motion teleological explanation persisted. Animate functions were considered to be God's design.

When Darwin (1859) proposed that the mechanism underlying biological evolution was natural selection the seeds were sown for the removal of teleology from

biological causal explanations as well. As Honderich (1995, p.94) describes the present attitude towards causality in biology, 'Most contemporary philosophers of biology now hold that functional explanations in biology are in fact disguised [efficient] *causal* explanations, which explain biological traits not by looking forward to *future* beneficial results, but by looking backwards to the *past* evolutionary histories in which such results led to the natural selection of the traits in question.' So animate functions are no longer considered to be due to God's design but are naturally selected according to efficient cause influences.

As Hale, Margham & Saunders, (1995, p.428) explain the current state of the theory, 'Organisms which are better adapted to the environment in which they live produce more viable young, so increasing in proportion in the population; thus particular characteristics are 'selected' and others are lost. Such a mechanism depends on the variability of individuals within the population. Such variability arises through mutation and other genetic events, the beneficial variants being preserved by natural selection.'

Biological theorists still use teleological language to some extent. Dawkins' idea of 'selfish genes' using bodies as vehicles in order to further their own chances of replication is a prime example (Dawkins, 1976; Dawkins, 1982). But such language always comes with the warning that this is just a way of speaking about the phenomena and that the real causes of the phenomena are efficient causes – events that occurred to a creature's ancestors in the past.

It should be noted here that teleology can be split into two common aspects – function and goal. As in the above description by Honderich, evolutionary theory is often called upon to provide functional explanation.

As Woodfield (1976, p.109) defines such explanation, '[Functional explanation] explains the existence or presence of an item by reference to a function that it has. It answers the questions: 'What is X there for?', 'Why does X exist?' or 'Why does X do so and so?'

Evolutionary theory may thus be called upon to explain why birds have wings for example. In particular, evolution explains how apparently teleological function is in fact efficiently caused function that develops through genetic disposition.

It is important not to confuse the two aspects of function and goal (see Woodfield, 1976). Goal can be used in loose teleological statements, such as the goal of the heart being to pump blood around the circulatory system. The latter though is really a functional statement. The sense of goal and goal direction developed in this thesis will be clearer if goal is more tightly circumscribed. In particular, goal is to be linked to a common everyday sense. Goal will be tied to the behaviour as it happens and will be used to explain the direction of motion. "The agent moved to the left in order to reach the stream," is an example of this common teleology. Goal then is a feature that operates within the lifetime of the behaviour and ceases to exist outside it.

This means that evolutionary biologists could deny the need for final cause in their own subject but have no problem with its introduction into goal-directed behaviour in the sense defined above.

2.3. Reinforcement Learning

An approach that will prove to be of relevance to the present thesis encompasses the phenomena of reinforcement learning. The Behaviourist approaches of the 20th Century provide an early account based on this idea. The Behaviourists saw goal-directed behaviour as the result of conditioning. As Gregory (1987, p.159) explains, 'After experiencing a number of pairings of a signal, for example a tone or light, and a reinforcer, in this case food...dogs came to salivate to the signal as much as they did to the food. Similarly...rats...would readily perform some action, such as pressing a lever, that procured food.' The development of these behaviours 'requires the animal to experience a relationship or association between the signal and food in the case of Pavlovian or classical conditioning, and between the action and food in instrumental or operant conditioning.'

For them, behaviours were considered to be responses to stimuli learnt and reinforced on the basis of their past consequences, thus making their explanations a form of efficient cause explanation. As Skinner (see Gregory, 1987, p.74) presents the behaviourist view, 'Although it is the product only of past consequences, behaviour is useful because it may have similar consequences in the future. We refer indirectly to that future when we say that we act with a purpose, intention, or expectation. Such 'states of mind' are not, however, the causes of our behaviour.'

Behaviourism foundered on insisting that behaviour was essentially learnt reflexes without significant internal psychological states. Cognitivism rose instead by placing internal psychological states as primary components of behaviour.

Modern versions of reinforcement learning still look at it as a ubiquitous form of learning that shapes our lives every day. As Sutton & Barto (1998, p.3) explain it, 'The idea that we learn by interacting with our environment is probably the first to occur to us when we think about the nature of learning. When an infant plays, waves its arms, or looks about, it has no explicit teacher, but it does have a direct sensori-motor connection to its environment. Exercising this connection produces a wealth of information about cause and effect, about the consequences of actions, and about what to do in order to achieve goals. Throughout our lives, such interactions are undoubtedly a major source of knowledge about our environment and ourselves. Whether we are learning to drive a car or to hold a conversation, we are all acutely aware of how our environment responds to what we do, and we seek to influence what happens through our behavior. Learning from interaction is a foundational idea underlying nearly all theories of learning and intelligence.'

Modern reinforcement learning though no longer has a need to avoid reference to internal states– this can be seen in approaches like that of Sutton & Barto, for example. In their efforts to develop artificial reinforcement learning systems Sutton & Barto (1998, pp.7-9) identify four components underlying reinforcement learning. The 'policy' of the agent is a mapping from their current perceived state to some action to be performed when the agent is in that state. The 'reward function' maps the

perceived state to a number that is the reward or goal that the agent is seeking. The agent is aiming to maximise this reward in total. The 'value function' is a measure of the long-term desirability of a particular state given the states that are likely to follow it. Finally the 'model' is an optional fourth component that allows the agent to model the environment for planning a future course of action. So for Sutton & Barto reinforcement learning can cover all levels of learning from trial and error learning up to deliberative planning.

The modern approach has no need to avoid reference to goals either. For Sutton & Barto (1998, p.6) even an everyday occurrence such as preparing breakfast involves a plethora of opportunities for goal-directed reinforcement learning. As they put it, 'Closely examined, even this apparently mundane activity reveals itself as a complex web of conditional behavior and interlocking goal-subgoal relationships... Each step is guided by goals, such as grasping a spoon, or getting to the refrigerator, and is in service of other goals, such as having the spoon to eat with once the cereal is prepared and of ultimately obtaining nourishment.'

Be that as it may, the fundamental intuition often presently given is still that behaviour is primarily guided by the mechanism of previous results of actions biasing future actions in favour of those actions that provided the greatest rewards in the past. Such causal explanation remains efficient causal.

However, reinforcement learning aims to account for the development of associations that are not hardwired. It does not aim to account for the universally hardwired properties of plasticity and persistence that enable the goals that arise from the

associations to be pursued. Persistence, first introduced into the characterisation of goal-directed behaviour by Russell (1945), has recently been defined as 'the ability to continue trying for the goal in the event of perturbation throwing the agent off course,' (Wale & Weir, 2002, p.79). Meanwhile, plasticity, first introduced into the characterisation of goal-directed behaviour by Braithwaite (1946), has been defined as 'the ability to behave in a flexible way in pursuit of a goal,' (Wale & Weir, 2002, p.79). The latter account (of plasticity and persistence) is extra to, rather than contradictory of, reinforcement learning.

The reinforcement learning mechanism is only one component of goal-directed behaviour. This is because it is concerned with specific questions about a particular type of adaptation, i.e. learning that goes on between behaviours. So although the reinforcement learning component of causal explanation in goal-directed behaviour is purely efficient causal, there remains the possibility that final cause may be present in some other adaptive aspect such as the first time learning that goes on within a behaviour.

2.4. Neuroscience

Another component of the causal explanation of goal-directed behaviour concerns the neural processes underlying the behaviour (see Carlson (2004) for example). As Horgan (1996, p.159) explains, '[M]odern neuroscientists are interested less in how and why our minds evolved, in a historical sense, than in how they are structured and work right now. The distinction is similar to the one that can be made between cosmology, which seeks to explain the origins and subsequent evolution of matter,

and particle physics, which addresses the structure of matter as we find it here in the present. One discipline is historical and thus necessarily tentative, speculative and open-ended. The other is, by comparison, much more empirical, precise and amenable to resolution and finality.'

Contemporary research now has an increasingly sophisticated assortment of empirical techniques involving PET[1], large scale electrode and SQUID[2] arrays and functional MRI[3] that provide brain data and images at different spatial and temporal scales (Carter, 1998; Posner & Raichle, 1997). The empirical evidence can be used to help construct mathematical models of neural phenomena such as firing rates and receptive field structures as well as the neural basis of learning mechanisms such as reinforcement learning and representational learning (Dayan & Abbott, 2001).

In terms of its attitude to causation the neuroscientific approach sees the processes that occur in the brain to control movement as the causes of the behaviour. These neural correlates of behaviour, being prior to the behaviour, are therefore efficient causes of the behaviour. They see the behaviour as the result of firing patterns in the brain. The question then arises as to what role the goal plays in influencing the firing patterns. There may be a temptation to view the role of the goal as solely that of an internal representation, i.e. to limit the goal's influence to being part of the current firing pattern.

An example of this, which is of relevance to the experiments described later, can be seen in a study by Novak, Miller and Houk (2000, p.430). These authors performed

[1] Positron Emission Tomography.
[2] Superconducting QUantum Interference Device.
[3] Magnetic Resonance Imaging.

experiments involving simple human hand movements where subjects were required to turn a knob in response to an LED target. They found that the behaviour involved a primary movement and corrective submovements. On the basis of this result and the neuroscientific evidence available to them they suggested the possible brain mechanisms that were responsible for the behaviour. In their words, 'We propose that two brain processes are responsible for planning and controlling movements: one to initiate the primary movement and corrective submovements, and one to regulate and tune the parameters. The initiator process might be realized by a circuit between the basal ganglia, thalamus, and motor cortical areas, and the regulator process by a loop between the motor cortical areas, red nucleus, pons, thalamus, and the cerebellum (Houk et al. 1993). Corrective submovements may be generated by evaluating efference copy and afferent feedback. These circuits could form functional modules which interact together to control movements (Houk and Wise 1995).'

The view taken here is that there is the possibility that influence may also lie outside of the current state in the neural process and reside in the plastic and persistent direction of the sequence of firing patterns towards attempting the goal.

As will be seen, the plastic and persistent direction alluded to above will be quite consistent with neuroscientific accounts of how behaviour is caused by brain processes such as the above. It also offers the possibility of an aspect of the goal providing a final cause directing the brain process.

2.5. Cognitive Science

The rise of the digital computer infused cognitive science, the interdisciplinary science of the mind, with the computer metaphor of the mind as its theoretical spirit guide. Since then cognitive science has spawned a host of increasingly more sophisticated approaches based on the principles of information processing.

Cognitive psychology started as a reaction to the excesses of behaviourism in the mid-twentieth century. The behaviourists made little mention of the internal mechanisms of animate agents, concentrating instead on overt behaviour. Cognitive psychologists believed that it was necessary to understand the information processing that went on in the brain but at a more abstract level than that of neuroscience. Central to this view was the representational theory of mind based on the digital computer metaphor.

As well as using more traditional psychological experiments of cognitive psychology, cognitive scientists have also made use of computer simulations to test their theories in the field of Artificial Intelligence (AI). As Blackburn (1994, p.26) explains it, 'Modelling a psychological phenomenon on a computer is a way of showing how the phenomenon is possible in a physical world, and is also a way of bringing out the complexities in apparently simple tasks.' Such approaches have been applied to a variety of problems from the simulation of perceptual systems and natural language understanding to game playing and medical diagnosis (Rich & Knight, 1991, pp.4-5). Computer modelling approaches account for the goal-directed movement of agents in terms of the processes of computing a trajectory, computing the angles of the joints, computing the forces, initiating the movements and receiving feedback to act as a corrective function to be used in to the next round of computation and initiation of movement (Johnson-Laird, 1993, pp.195-214).

The idea here is that of the mind as a formal symbol manipulating system analogous to a digital computer. As Newell & Simon (1976) put it, 'A physical symbol system consists of a set of entities, called symbols, which are physical patterns that can occur as components of another type of entity called an expression (or symbol structure). Thus, a symbol structure is composed of a number of instances (or tokens) of symbols related in some physical way (such as one token being next to another). At any instant of time the system will contain a collection of these symbol structures. Besides these structures, the system also contains a collection of processes that operate on expressions to produce other expressions: processes of creation, modification, reproduction and destruction. A physical symbol system is a machine that produces through time an evolving collection of symbol structures. Such a system exists in a world of objects wider than just these symbolic expressions themselves.'

In terms of goal-directed behaviour the goal of a symbol manipulating system is a part of the current (rather than future) symbolic state of the system. The exact nature of the goal, as Johnson-Laird (1993, p.205) puts it, 'may be a value of some physical quantity, such as body temperature, or a *representation* of a desired state.' The fact that the goal is a part of the state of the system means that the future behaviour of the system is fully determined on the basis of prior states. The form of explanation involved therefore uses purely efficient cause.

A further example of this attitude to goal-directed behaviour in cognitive science is that of Rosenbaum & Krist (in Heuer & Keele, p.3). As they put it, 'In seeking to understand the control of motor behaviour, it is natural to inquire into the antecedents

of action – the events that precede and allow for the execution of voluntary movements. In this chapter, we consider the *representations* formed before movements are carried out and which are vital for the successful realization of the actor's intentions. Such representations have been called *plans* or *motor programs*. The term 'motor program' connotes a representation of forthcoming activity that codes details about movements. The term 'plan' connotes a higher level representation.'

So, for cognitive scientists, although they may use the language of final cause explanation they still believe that the distinctive thing about goal-directed behaviour is the representational antecedents of the behaviour, i.e. the mental or program states. Their accounts are thus purely efficient causal explanations.

Goals though have many aspects, of which internal representation is just one. In particular, internal representation does not *per se* aim to give an account of the plasticity and persistence mechanism of goal-directed behaviour. The cognitive science account focuses on what the goal is perceived to be by the agent and how this affects the symbolic actions rather than the details of their internal plastic and persistent shape. A final causal account of goal-directed plasticity and persistence would again be complementary to, rather than contradictory of, such representationalist approaches.

2.6. Complex Systems Theory

An approach to the study of motion that has had some success in recent decades is that of complex systems theory (Coveney & Highfield, 1995). The researchers behind these approaches have a shared belief that behaviour, whether inanimate or animate, can be usefully studied in terms of complex adaptive systems. Coveney & Highfield (1995, p.425) define this approach as, 'The study of the behavior of macroscopic collections of simple units (e.g. atoms, molecules, bits, neurons) that are endowed with the potential to evolve in time.' At the heart of this view is the notion that global properties of the system emerge from the basic units underlying the behaviour. Its mathematical basis is that of dynamical systems theory (Arrowsmith & Place, 1992).

One important feature of this approach is the demonstration that complex systems, such as the weather, for example, can be highly sensitive to the initial conditions of the system. Small changes to the initial conditions can lead to vastly different future outcomes. This renders prediction very difficult. However, sometimes systems settle down more robustly into stable states known as attractors which provide a way of describing the long-term behaviour of a system (Coveney & Highfield, 1995, pp.150-189).

An example of a brain-related complex systems approach in cognitive science is that of connectionism, which takes a sub-symbolic stance to the issue of representation and involves parallel distributed processing (Rumelhart & McClelland, 1986; McClelland & Rumelhart, 1986; Beale & Jackson, 1990). Artificial neural network models are inspired by aspects of real biological neural networks that allow for the emergence of higher-level competences. As Martel Johnson & Erneling (1997, p.166) describe it, 'A connectionist system involves the three general levels of (1) input units

or "nodes," (2) hidden or processing nodes, and (3) output nodes. Such a system is able to learn to solve problems by a "training-up process," in which it is subjected to a regime of trial and error. The result of each trial is immediately "corrected" on the basis of feedback considerations about whether and how much each of the system's parts – especially the "weight" assigned to each connection between an input or output node and a middle or processing node – has (or has not) contributed toward a correct solution. This allows the system to gradually modify, or "shape" itself, until it finally reaches a state where it is capable of transforming new, untested inputs into desired outputs. For example, given a certain group of sounds, the system can report that the sounds are the voice of *X*; or given a particular set of colored lines and patches, it can correctly identify ("recognize") them as a visual representation of *Y*'s face.'

Connectionism has benefits for cognitive science in terms of areas, as Martel Johnson (1997, pp.165-166) puts it, 'where it is necessary to arrive at a solution by progressive elaboration and refinement of certain initially vague indications, traces, or patterns, rather than by following algorithmic steps to a supposedly inevitable conclusion. Examples of such cases are problems of learning how to recognise a face or distinguish multiple star systems on the basis of fragmentary and fluctuating optical and microwave evidence, as contrasted with problems like dividing one number into another, or determining the first prime number that occurs after 10,000,000,000.'

But in terms of goal-directed behaviour current standard techniques in connectionist models are still state-determined. As Wale & Weir (2002, p.83) put it, 'The standard deterministic methods used in neural network training, by being based on gradient

descent in various forms, are each driven by the fixed shape of the typically smooth error-weight surface in the different paths taken. There are also various other methods that use variable travel surfaces... However, even though the travel surfaces vary, their actions in forming a path are equivalent to a single fixed surface underlying the path.' So given the initial set of values of the system the rest of the behaviour of the system can be fully determined. Only an efficient cause explanation is required, hence connectionists, like Classical AI researchers, may use teleological terms but do not use final cause explanation as identifying a distinctive influence on the behaviour.

Juarrero (1999, pp.126-128), however, believes that present day dynamical systems theory principles such as autocatalysis, self-organisation and attractors can be seen as potentially paving the way towards a scientifically well-founded form of final cause explanation. As she puts it, 'An autocatalytic web functions as an "attractor"... a rudimentary precursor of final cause.' She goes on, 'It would be anthropomorphic to call this vectorial characterisation of autocatalytic structures "goal-intended" or "purposive"; it would be even more absurd to say that this dissipative structures act as they do "for a reason." And yet a precursor of teleology is detectable in the way such structures of process operate. Whether embedded in physical, chemical, biological, psychological, or social processes, homologous, irreversible dynamics appear to be at work in constructing nature's levels of organisation along with the emergent properties characteristic of each... A naturalized account of purposiveness such as I am attempting must identify other biotic and abiotic processes that are rudimentary, primitive – "proto" goal-directed. Autocatalytic and other self-organizing structures are examples of precisely that.'

On this kind of view the goal of an animate agent could be considered as an attractor in a dynamical state space. However, such attractors may be just states that the system ends up in as the result of differential equations that govern the motion. The goal is not a teleological influence on the behaviour. As Cohen & Stewart (1994, p.207) put it, 'The name "attractor" appeals to mathematicians, but it carries the unfortunate suggestion that dynamical systems are goal-oriented – that states end up on the attractor because they know in advance that they have to go there. On the contrary, we only find out what the attractor looks like by watching where the initial states go.'

Juarrero (1999, p.8) argues that the standard covering-law model of explanation, which seeks to explain behaviour in terms of initial conditions and efficient causal laws of nature, is inadequate if animate agents are considered complex adaptive systems. Their embeddedness in time and space means that there are infinitesimally small differences in the conditions at one moment that can lead to very different states at the next moment. As she puts it, 'If human beings and their behavior are complex adaptive phenomena, the precise pathway that their actions will take is simply unpredictable. Covering-law models are therefore clearly inadequate to explain these processes.' Instead she suggests a narrative model of explanation that accounts for the trajectory in terms of the interplay between part and whole at each moment.

Determinate complex systems approaches in general though still retain the intuition that if the existing state of the system were to be fully known at a point in time (together with any parameters) then the future behaviour of the system can be fully determined. Such a system will be referred to as *state-determined.* The intuition contains the implicit notion that it is the past influences that push entities along into

the future even if it is impractical to pin down the initial state precisely or useful on occasion to view the system from the point of view of the attractor. The approach, therefore, strictly speaking is still using effectively efficient causal accounts for supporting mechanism. In this sense its apparent final causal accounts are equivalent to corresponding efficient causal accounts.

However, although efficient causation predominates in present determinate complex systems, the primary concern of complex systems theory is with accounting for complex properties. It would be interesting to consider whether a complex system property such as goal direction could be accounted for with a different supporting determinacy.

2.7. Ecological Psychology

Ecological psychology emphasises the rich and primary nature of the environment in providing enabling and determining information that influences the behaviour (for example: Gibson, 1979, Smith, 2003). This is in contrast to the focus of cognitivism on internal representations providing the key information for processing. It brings together a loose collective of researchers who, as Neisser (see Martel Johnson & Erneling, 1997, p.248) puts it 'all want to know what human beings actually do (or see, or know) in their normal commerce with their various environments.'

For ecological psychologists the environment is awash with 'affordances' in objects and processes that can be perceived by agents and that enable the agent to perform certain behaviours. Affordances could be very obvious properties of objects, like a

flight of stairs affording an agent the possibility of reaching another floor of a building. But there are also much more subtle affordances in the environment. An example of this is the notion of 'flow' in the study of visual perception. As Gregory (1987, p.294) describes this approach, 'Gibson first of all moved away from the traditional experiments with pictures, and what is seen with a single static eye, towards the observer moving around freely and viewing moving objects in natural conditions. He was led to this by considering pilots landing on fields, where the 'flow lines' of motion are important for seeing the landing-point and estimating height and speed. From such considerations of 'visual flow', and texture gradients, he developed what he called 'Ecological Optics'. This almost ignored retinal images, and active brain processes, in favour of regarding perception as 'picking up information from the ambient array of light'... Gibson tried to explain object perception by supposing that some 'higher order' features are invariant with motion and rotation and are 'picked up' with no perceptual computing or processing being required. His search for such invariances, and for just what the visual system uses under various conditions... has proved useful for developing computer vision.'

In terms of goal-directed behaviour the ecological psychology worldview leads to a conception of the animate agent as a detector of, processor of and reactor to, environmental information. The agent has multiple goals that are, at least in part, associated with the changing energy flows of the environment itself. The goal of the agent can change because the agent detects a new goal and the path to a goal can change because of the detection of new affordances.

Such an approach would seem to imbue goals embedded in the environment with the power to influence behaviour as and when they are set. So knowing the current internal state of the agent may not be enough to determine the future behaviour of the agent. The behaviour may also depend on the currently engaged environmental goal of the agent. If framed carefully this might be taken to allow for the reintroduction of final cause so long as the environmental information is a distinct influence on the behaviour.

2.8. Kinematics

In general the study of kinematics is the study of motion without reference to force or mass (Borowski & Borwein, 1999, p.321). The central theme of this approach is that kinematics studies can be used to fruitfully investigate the goal-directed behaviour of animate agents. (For an introduction to human kinematics see Zatsiorksy, 1997).

In this approach movements are observed in controlled experiments using limb monitoring sensors and devices that provide data that can be analysed in terms of displacement over time and in terms of the derivative profiles such as velocity and acceleration (for example Magill (1998, pp.23-26); Morris & Paradiso (2003); Novak et al. (2000)).

Both linear and angular motion can be described. As Magill (1998, p.26) puts it 'In kinematic descriptions of movement, the measures of displacement, velocity and acceleration can refer to either linear or angular motion. Linear motion describes the movement of all parts of a moving object, while angular motion refers to movement

that occurred for some parts of the object but not for other parts.' Novak et al. (2000) used a knob twiddling task based on angular motion whereas the experiments to be analysed in Chapters 5 and 6 of the present thesis look at linear motion.

A further feature is that typically a system under investigation will have more than one dimension to it. Kinematics researchers attempt to take into account all of the necessary dimensions that they require to build up their models of the behaviour. This may involve, for example, placing sensors on many places of a person's limbs to analyse walking behaviour, for example.

In terms of causal explanation, in practice, kinematics researchers impose an efficient causal model on their results by default by using off-the-shelf functions to fit the data that are state-determined. This prevents any possibility of identifying final cause components to the behaviour under scrutiny.

Final cause explanation could be reintroduced relatively painlessly by, instead of imposing state-determined functions on the data, coming up with more general functions capable of reflecting possible greater plasticity and persistence in the data.

2.9. Control Systems Theory

Control systems theory is commonly centred on the notion of feedback to control a system automatically and its history goes at least back to ancient Greece (Dorf, 1986, p.4) where it was applied to a float regulator mechanism that could be used, for example, in water clocks.

More recently control theory lies at the heart of cybernetics, an approach that seeks to explain the behaviour of the whole system, either mechanical or biological, by considering the functions and interactions of its constituent parts and their interactions with the environment. Central to cybernetics is the notion of feedback from the environment and how such feedback can contribute to the overall stability of the system. As Gregory (1987, p.174-177) explains, 'It is a theory of feedback systems, i.e. self-regulating systems, the theory being applicable to machines as well as to living systems... Feedback mechanisms are characterized by the fact that the input is controlled by the output, and thus stabilizes the output, or makes the performance relatively independent from disturbing influences. One can consider the stability of the output as the 'goal' of the system. To turn the argument round, whenever behaviour in biology is encountered that can be described as goal-directed, i.e. teleological, it is very likely that a feedback mechanism is involved.'

In its founding document (Rosenblueth, Wiener & Bigelow, 1943) the modern science of cybernetics was presented as a general science of control and communication. What is of particular interest for present purposes is the form of explanation that they believed was required. As Gardner (1985, p.20) puts it, 'The authors introduced a then-radical notion: that it is legitimate to speak of machines that exhibit feedback as "striving toward goals," as calculating the difference between their goals and their actual performance, and as then working to reduce those differences. Machines were purposeful.'

However, it can be argued that such machines are nevertheless, state-determined and thus goals are not a teleological influence on the behaviour of the system. As Rosenblueth, Wiener & Bigelow put it, 'Teleology has been discredited chiefly because it was defined to imply a cause subsequent in time to a given effect [a final cause]... we suggest that a teleological study is useful if it avoids problems of causality and concerns itself merely with an investigation of purpose.' For them the goal is simply a part of the state (Weir, 1984). Knowing the complete current state would allow the future evolution of the system to be determined. A truly teleolgical goal-directed system could require different goal states to be reachable from the same current state.

So for control theorists, although they may sometimes use teleological language, in fact the present system design is that of a state-determined system and their explanations are purely efficient causal. Out of all the approaches discussed in this chapter, control systems theory is the one in which state-determinacy is most embedded.

Goals take the role of a selective parameter rather than a dynamic developer of the process. The goal parameter selects *a priori* a state-determined phase portrait of a family of trajectories rather than develops the portrait shape dynamically.

Nevertheless, it is at least conceivable that feedback and differential equations may be designed to lead to adaptive behaviour other than through state-determined design. In particular, feedback that uses the goal as a future derivative state rather than as a current parameter of a state-determined system to determine its response to going

away from the goal is a different type of feedback to that of standard control system theory, but is feedback nonetheless. (The notion of a derivative state here refers to the set of values for height, velocity, acceleration and subsequent rates of change for a behavioural trajectory at a particular point).

2.10. Conclusion

All of the approaches analysed in this chapter provide purely efficient causal accounts of goal-directed behaviour by default rather than necessity. However, these approaches concern specific aspects of goal-directed behaviour that may be compatible with a complementary approach in which final cause is reintroduced through a distinct aspect of plasticity and persistence.

Two of the approaches, the ecological psychology and kinematics approaches, seem especially promising in terms of the potential reintroduction of final cause explanation.

In the ecological psychology approach goal information is latent in the environment. The current internal state of the agent may therefore not be enough to predict the future behaviour of the agent. From the same current internal state they may move in very different ways depending upon the goal that they become engaged with in the environment. However, incorporation of the environmental goal information as well as the current state information might allow prediction of a unique path.

In the kinematics, goal information can be characterised, more specifically, as information about the derivative state underlying a particular desirable future outcome. Knowing the current derivative state of an animate agent may not be enough to determine their future kinematic behaviour especially if plastic persistence is required for the task. Multiple paths are possible from the same current derivative state and it is an interesting question as to whether by taking into account the goal state information as well, a unique path can be predicted.

In the next chapter, the theme of plasticity and persistence introduced in the present chapter as a potentially final causal aspect of goal-directed behaviour is developed. A continuous state space approach provides the basis for describing a computational framework for C^{∞} smooth goal-directed behaviour. A problem that will become significant in later chapters, and which is called the Parallel Pipes Problem, is introduced for the first time to illustrate plastic persistence. Plastic persistence is encapsulated by the computational and predictive notion of non-analytic projection.

Chapter 3

A Computational Framework for C^{∞} Smooth Goal-Directed Behaviour

3.1. Introduction

It is acknowledged in computer science that the ability for goal-directed behaviour is a key characteristic of intelligent agents. In their description of agent structure Russell & Norvig (2003, pp.44-55) demonstrate this by showing how agents with goals are an improvement on reflex agents.

Russell & Norvig identify two types of reflex agents. Simple reflex agents respond only to their current percepts. They have a set of rules with conditions that they attempt to match to the current state of the environment (as defined by the percept). If a condition is met then they act according to the rule associated with that condition. This allows for a limited form of intelligent agent because correct decisions can only

be made reliably if the environment is sufficiently observable, i.e. if all the necessary information is present in the current percept.

The second type of reflex agent identified by Russell & Norvig is that of model-based reflex agents. Such agents get over the problem of partial observability of the world by maintaining a current internal model or state that allows the tracking of aspects of the world that are not evident to them in the current percept.

However, knowing the environment is not necessarily enough to make correct decisions. Sometimes the correct decision depends on which of the multiple possible future states, i.e. goals, the agent is trying to get to. Such agents require goal information. Goal-based agents are able to use information about their desired future states as well as information about the current state of the world in their decision-making. At least two important subfields of AI (search and planning) are aimed at providing techniques for agents to find the best action sequences to goals.

A significant modification to goal-based agents is to allow them to assign values to the desirability of future states so that they can maximise the expected reward (utility) from their behaviour. Such utility-based agents are able to choose between action sequences to a particular goal, to choose between multiple goals to pursue and also to choose action sequences that are a trade-off between multiple goals, not all of which can be satisfied (e.g. speed and accuracy).

The ability for exhibiting goal-directed behaviour is also a key feature in distinguishing between animate and inanimate entities. In the thesis goal-directed

behaviour is taken to be the way that animate agents manifest their intentional behaviour in space-time.

In this chapter a computational framework will be developed in which animate agents are seen as those that are capable of acting with behavioural flexibility that can be characterised in terms of plasticity and persistence and how these concepts come together in the notion of continual redirection towards a goal.

In the following discussion of smooth behaviour we will need the definition of a Taylor series. The Taylor series allows prediction of separate values for the function representing a curve based on the derivative state at a particular point on the curve. Following Jeffrey (1996, p.560) this is:

$$f(x) = f(x_0) + f(x-x_0)f'(x_0) + (x-x_0)^2/2! * f''(x_0) + \ldots + (x-x_0)^n/n! * f^{(n)}(x_0) + \ldots$$

Here, f is the function, x_0 is the location of a derivative state, and $(x- x_0)$ is the distance from x_0 used by the Taylor series to form a prediction for the values of f at x, and where the derivative state is $< f(x_0), f'(x_0), \ldots, f^n(x_0) \ldots >$, $n \rightarrow \infty$.

A specific type of goal-directed behaviour will be focussed upon here and in the thesis generally, which involves paths that are C^∞ smooth, i.e. continuously differentiable up to all orders of derivative (see Penrose (2004, pp.107-112) for discussion of C^∞ smoothness). There are two types of C^∞ smooth path – analytic and non-analytic. The term 'analytic' has several senses mathematically (as defined by Borowski & Borwein, 1999, pp.16-17). The sense of interest in the thesis is the second sense listed

by Borowski and Borwein, which they also refer to as 'real analytic'. Analytic in this sense concerns real functions that possess derivatives of all orders and agree with their Taylor series predictions locally. Non-analytic C^{∞} smooth paths are those that do not obey their Taylor series predictions (see Chillingworth, 1976, pp.82-87).

C^{∞} smooth goal-directed behaviour can thus be modelled in three ways. Either using analytic paths, non-analytic paths, or paths that have both analytic and non-analytic parts. In this chapter these three approaches are discussed as candidates for providing the computational framework with a model that best suits the characterisation of animate agents developed here.

3.2. The Characterisation of Goal-Directed Behaviour

The term 'goal-directed behaviour' was originated by Russell (1945). He emphasised the role of *persistence* in goal-directed behaviour. This is the idea that the agent will continue to try for its goal. Recently it has been put forward that persistence is 'the ability to continue trying for the goal in the event of perturbation throwing the agent off course,' (Wale & Weir, 2002, p.79).

Braithwaite (1946) added the concept of *plasticity* to the characterisation of goal-directed behaviour. A recent definition of plasticity suggests that it is 'the ability to behave in a flexible way in pursuit of a goal,' (Wale & Weir, 2002, p.79).

Such plastic persistence in goal-directed behaviour can be seen in several types of real world situation that require the agent to redirect itself in order to maintain on course

for a goal. One type of situation is when agents face environmental perturbations. An example of this is a yachtsman who redirects a yacht back on course when blown off course by a gust of wind.

Plastic persistence is also required in cases where the agent itself lacks skill at a particular task. A learner driver may start to veer to the wrong side of the road, for example. In such a case exhibiting plastic persistence in the behaviour to redirect the car is vital in order to avoid a possible collision.

A third case where plastic persistence is needed is when the agent must traverse a sequence of goals where the trajectory to one goal must evolve into a different trajectory for a subsequent goal. An example of this is a skier traversing a path through a sequence of slalom gates.

In the construction of artificially animate agents the ability to behave with plastic persistence is vital for traversing obstacles and correcting performance errors so as to keep on course for the presently engaged goal, as well as for taking the smoothest path through a series of goals.

However, the notion of plastic persistence is not quite general enough in its coverage of goal-directed behaviour. By defining persistence in terms of continual redirection in the face of perturbation this leaves out cases where an agent simply finds itself off course for a newly engaged goal and has to redirect towards the new goal. To cover these cases as well the term plastic redirection will be used as the general term. The term plastic redirection will be used in the rest of the thesis with the more specific

term plastic persistence being reserved for cases involving continued pursuit of an existing goal.

So, on the basis of the account developed in this section, goal-directed behaviour can be considered as behaviour that involves plasticity and continual redirection towards the currently engaged goal. Animate agents can be distinguished from inanimate objects on the basis that they can perform goal-directed behaviour characterised in the way developed here.

3.3. A State Space Approach

In computational approaches the behaviour of intelligent agents can be investigated using state space trajectories. Such agents are modelled as using some computational procedure that helps them to find their way through a state space to their desired outcomes. State spaces can be either discrete or continuous.

A discrete state space is one that has a finite number of distinct states (Russell & Norvig, 2003, p.42). Examples of task environments involving discrete state spaces include crossword puzzle solving, poker and interactive English tutoring (Russell & Norvig, 2003, p.43).

A continuous state space is where the states come from a continuum of possible states. Examples of task environments involving continuous state spaces include medical diagnosis, image-analysis, and taxi driving (Russell & Norvig, 2003, p.43).

Many real world problems involve continuous state spaces. One such problem, involving human agents using a computer mouse to move a ball between two parallel pipes displayed on a screen, will be investigated in the thesis. This problem will be referred to in the thesis as the Parallel Pipes Problem (PPP).

Of course, it could be pointed out that the mouse movement in such a case is recorded as discrete pixel positions on the screen and discrete timings in milliseconds and therefore the state space used should be a discrete one. However, this input will nevertheless be treated in the thesis as continuous because the discrete data represent

continuous underlying behaviour in spacetime and it is this C^{∞} smooth behaviour that is being investigated. This is analogous to the case of the digital camera described by Russell & Norvig (2003, p.42). Input from a digital camera is, although technically discrete, 'typically treated as representing continuously varying intensities and locations.'

Goal-directed agents are often modelled in research as those that are trying to traverse a trajectory towards a goal or goals that are part of the state space. Russell & Norvig (2003, p.49), use the example of a taxi driver with the goal of reaching the passenger's desired destination. As they put it, 'as well as a current state description the agent needs some sort of goal information, which describes situations that are desirable – for example, being at the passenger's destination. The agent program can combine this with information about the results of possible actions...in order to choose actions that achieve the goal.'

In our terms agents capable of C^{∞} smooth goal-directed behaviour are specifically agents that traverse continuous state space trajectories with plastic redirection and persistence as defined in Section 3.2.

A point that needs to be cleared up at this juncture is the relation between state space approaches and state determinacy. In the thesis state-determined is taken to mean the behaviour of an isolated physical system in which 'the state of the system can be represented as a single-valued function of the initial state and the time co-ordinate,' (Sommerhoff, 1969, p.154). In other words, given the state of the system at one moment there will only ever be one subsequent trajectory possible. As Sommerhoff

(1969, pp.154-155) explains biologists use this definition wherever possible because it has been the only way that they have had of making single-valued predictions. However, this thesis will investigate other possibilities for state evolution and making predictions as well.

In the smooth arena when analytic paths are being taken, state determinacy holds. From any given point, knowing its full derivative state, a single C^{∞} smooth analytic path will result. There is also though a family of non-analytic paths that also project C^{∞} smoothly from the point with the known derivative set. So if C^{∞} smooth goal-directed behaviour involves non-analytic paths then knowing the full derivative state of the system at the given point will not be enough to reliably predict the subsequent behaviour. The behaviour may then not necessarily be state-determined. So using a state space approach does not imply that the behaviour is also necessarily being modelled in a state-determined way. Knowing the full state of the system at a point does not necessarily imply knowing the state of the system at any other point.

The rest of this chapter will develop an account of the type of paths that may be involved in the continuous state space trajectory and that may best allow plastic redirection and persistence.

3.4. The Unsmooth Versus Smooth Distinction

In the natural world whenever a human kicks a football or a hawk swoops down on its prey the agent's action is typically performed, over significant periods, to be as smooth as possible. With practice the action is free from excessive undulations, stops

and jerks. The free kick is taken and followed through in smooth movements and the ball ends up in the back of the net. The talons grasp the mouse as the hawk executes a smooth swoop. In the above, there is a sense of smoothness will be referred to in the thesis as 'undulatory smoothness'. It refers to paths that have minimal changes in gradient overall.

There is also a second sense of smoothness at work in the above examples that will be of use in the search for the path type that is most likely to underlie goal-directed behaviour. C^{∞} smooth paths are those paths that are continuously differentiable up to all orders of derivative over the whole interval. Discontinuity at some order of derivative at some point would mean that the path was unsmooth at the point of discontinuity and thus the overall path would not be C^{∞} smooth, only piecewise smooth. As Borowski & Borwein (1999, p.450) define it, 'a function is piecewise smooth on an interval if it is continuous thereon and is continuously differentiable except at finitely many points where the derivative may have jump discontinuities.'

The C^{∞} smoothness assumption of the world in general is used extensively in science. Newtonian physics, for example, makes the assumption that the motion of all objects in the universe is explicable in terms of C^{∞} smooth paths, amenable to description using laws expressible in terms of differential equations.

It is supposed here that goal-directed behaviour also has distinctive and computationally useful principles that appear in C^{∞} smooth behaviour that is also often undulatory smooth.

3.5. The Analytic Versus Non-Analytic Distinction

As described in Section 3.1, C^{∞} smooth paths can be either analytic or non-analytic. The relevant mathematical definition of analytic being that it involves real-valued functions, has continuous derivatives of all orders and obeys its Taylor series predictions locally (Borowski & Borwein, 1990, p.16). An example of an analytic path type is the familiar function $y = x^2$ shown below in Figure 3.1.

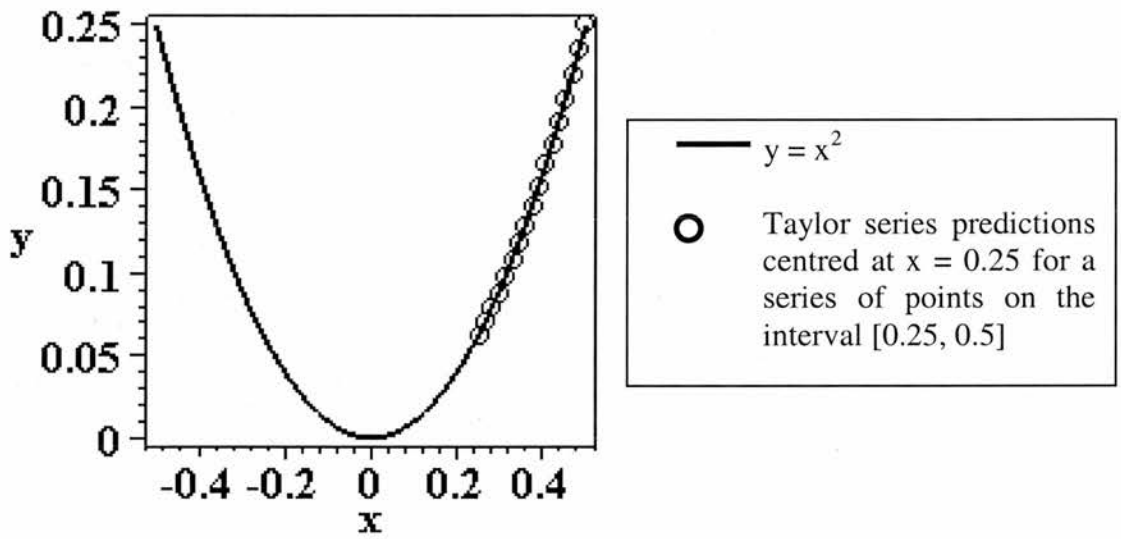


Figure 3.1. Graph of $y = x^2$ with overlaid Taylor series predictions centred at x = 0.25 for a series of points on the interval [0.25, 0.5].

The Taylor series prediction from any point in an analytic curve is a state-determined (as defined in Section 3.3) *projection* of the rest of the curve. If enough terms are used, i.e. up to all non-zero derivative orders, then the prediction will be accurate for the rest of the curve at least locally. This is illustrated in Figure 3.1 for a Taylor series prediction centred at x = 0.25 and continued up to x = 0.5. The dots represent

predictions made using the Taylor series prediction over the interval [0.25, 0.5] with all non-zero derivative orders and precisely follows the actual analytic curve, $y = x^2$, itself.

If a truncated Taylor series is used, i.e. one that does not include all of the non-zero terms of the series, then there will be a discrepancy between the actual curve and the Taylor series prediction from a point. However, the more terms that are used, the more accurate the Taylor series becomes. As Gullberg (1997, p.771) puts it, 'The first term of a converging series is the first approximation, the sum of the first and second terms the next approximation; by adding an increasing number of terms, the degree of accuracy is limited only by our finite capacity of adding terms.'

Analytic paths may be found in the typical paths taken by objects being moved smoothly by purely efficient causal influences such as physical pushes or pulls which are state-determined. In an example adapted from Hahn (1998, p.166) a cannonball dropped from the leaning tower of Pisa, for example, obeys the equation:

$$y(t) = y(0) - (a/2)t^2$$

Where: t is the time elapsed; y(t) is the vertical position after time t; y(0) is the initial position at t = 0; a is the acceleration of the cannonball. (In this example a is equal to the acceleration due to gravity, known as g, which is taken to be $9.8 m/s^2$ close to the Earth's surface).

However, C^{∞} smooth paths do not have to be fully analytic paths. They can also be either partially or fully non-analytic. The Taylor series fails to predict the rest of the curve locally wherever there is non-analyticity. An example of a C^{∞} smooth path that is non-analytic at a certain point is that of the function $y = e^{-1/x^2}$ shown in Figure 3.2 below. This is non-analytic at the origin because a Taylor series prediction[1] centred at this point fails to predict any other point on the curve locally.

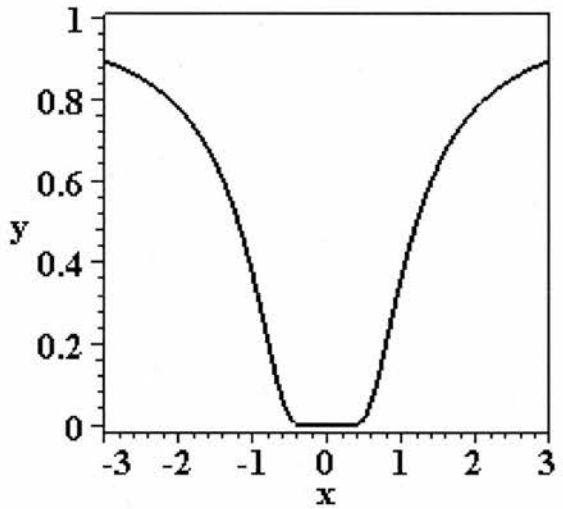


Figure 3.2. Graph of $y = e^{-1/x^2}$.

An example of a C^{∞} smooth path that is analytic nowhere, i.e. non-analytic everywhere, is provided by Merryfield (1992, pp.132-138). This is illustrated in Figure 3.3. The function is:

$$y(x) = \sum_{k=0}^{\infty} 2^{-2^{k/2}} \cos(2^k x)$$

[1] Technically a Taylor series centered at 0 is referred to as a Maclaurin series (Gullberg, 1997, p.769).

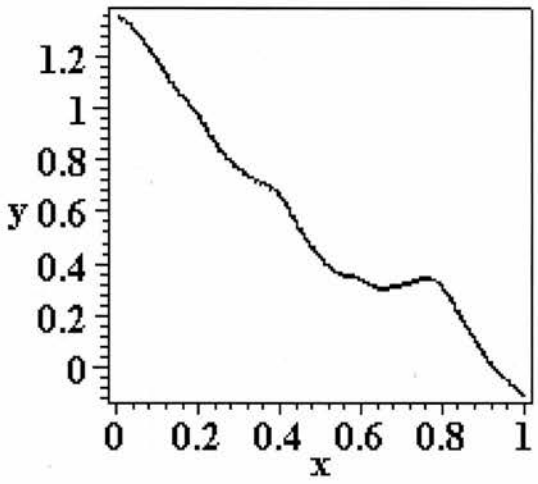


Figure 3.3. A C^{∞} smooth path that is analytic nowhere.

A Taylor series prediction centred at t = 0 in the case of $s = e^{-1/t^2}$ fails to predict the rest of the curve. In the case of Merryfield's curve, the Taylor series fails to predict the rest of the curve from anywhere and this is true no matter how many terms of the Taylor series are used.

Contemporary science predominantly uses analytic paths rather than non-analytic paths and they have numerous uses. However, as Chillingworth (1976, pp.86-87) says:

> 'Despite the enticing advantages of analytic functions, implicitly relied upon in much mathematical modelling in physics, engineering and so on, there are two objections – one practical and one philosophical – to working with analytic functions. The practical objection is that functions do arise in real life that look harmless but are in fact *not* analytic even though they are C^{∞}... The philosophical objection is that there are no grounds for assuming that smooth

functions and maps which model processes should share the many very special properties of analytic functions.'

He uses the function $s = e^{-1/t^2}$ as an example of his practical objection, with its non-analyticity at zero (see Figure 3.2). He illustrates his philosophical objection by pointing out that:

'[A]n analytic function cannot be perturbed a small amount locally without also perturbing it globally by what for distant points may be a very large amount. This kind of rigidity is undesirable for models of the real world, and particularly for dynamical systems... where the idea of stability under small local perturbations plays an important role. Therefore analyticity is not only an unwarranted assumption, but is a constraint that it may be quite inappropriate to impose on a mathematical model.'

Despite objections such as these, analytic paths have been used successfully to describe the motion of both inanimate and animate entities. In the inanimate case, Newton, for example, described the orbit of the moon around the Earth in analytic terms (Hahn, 1998, pp.191-193). In the animate case, kinematics studies, which look at motion without regard to force or mass (Magill, 1998, pp.23-26), also use analytic functions to describe the agent's changing position and associated derivative profiles in analytic terms. This latter case is of direct relevance in modelling the C^{∞} smooth goal-directed behaviour being investigated in the thesis. The fact that such studies use analytic paths but do not use non-analytic paths (even though they could do) is of great significance because non-analyticity may be present in the underlying animate behaviour. This possibility will be developed in the following section.

3.6. What C^{∞} Smooth Paths are Goal-Directed Agents Most Likely to Take?

In Section 3.2 the goal-directed behaviour of animate agents was characterised as showing plasticity, persistence and continual redirection so as to keep on course for a goal, i.e. plastic redirection. A state space approach was then developed in Section 3.3 that is consistent with this characterisation, i.e. a continuous state space approach. The types of C^{∞} smooth path that are available to be taken by animate agents were then discussed in Section 3.4. In this section an assessment of these path types will be made in terms of how well they fit with the earlier characterisation of goal-directed behaviour and thus how plausible it is that they are the path type underlying the behaviour.

Our hypotheses about likely paths are based on the theoretical characterisation of C^{∞} smooth goal-directed behaviour developed in the thesis in terms of plastic redirection. They are designed to be testable according to empirical evidence. In particular, our analysis will be based on what is plausible given what paths are found to exist *in situ* rather than *a priori* likelihoods of existence. So if one type of function (i.e. analytic or non-analytic) were shown to be more common than the other type then this would not be used to prejudge the experimental results within the framework of the statistically more common type of function and to only consider that type of function in the analysis.

There are three possible C^{∞} smooth models of goal-directed behaviour and each one can now be introduced. They are:

Model 1. Analytic.

Model 2. Piecewise analytic, i.e. analytic sections with non-analytic joins.

Model 3. Non-analytic.

Model 1, involving analytic paths, suffers from inbuilt rigidity in that such paths have to obey their Taylor series predictions everywhere. This means that from any particular point there is only one possible trajectory projecting from that path. The limitation imposed on goal-directed behaviour by this property becomes evident when considering cases of goal movement and perturbation on the agent.

Firstly, take the case of goal movement. For example, suppose that an agent does traverse part of the motion using a particular analytic path. But then, at this latest moment, the goal moves in some way that the agent is unable to anticipate or the agent itself instantaneously changes goal. It may be that the agent, locked in to its analytic path will no longer be able to achieve its engaged goal using that path. To achieve the engaged goal the agent would be required to be redirected, i.e. change to a new path. But this is not allowed within the framework of Model 1 because the agent is bound to a single analytic trajectory and can therefore not smoothly redirect.

A second case is that of perturbation. In the thesis the term perturbation will be used to refer to added displacements to an agent's trajectory that are not corrections by the agent. Examples of this include a yacht being blown off course by a gust of wind or a learner driver veering to the wrong side of the road through lack of skill (as discussed

in Section 3.2). In such cases the agent may no longer be able to achieve the goal by continuing on their steered trajectory. However, if this trajectory is a purely analytic one then they will be unable to redirect on to a new path that projects through the goal.

The above examples of goal movement and perturbation demonstrate that modelling redirection using analytic paths would require taking a rethink. One minimal first step is to use an approach that minimises the number of points where analyticity must be sacrificed. In other words, using Model 2, the piecewise analytic approach, rather than the purely analytic approach of Model 1. This means having analytic sections with joins (where instantaneous redirection on to a new analytic path occurs) that are either unsmooth, or, in order to keep the C^{∞} smoothness requirement, are non-analytic.

Thus, for our purposes, Model 2 is an improvement on the pure analytic approach as it can potentially cater for goal movement and perturbation of the agent due to lack of skill or external influences without breaking the C^{∞} smoothness requirement. However, there are two drawbacks with this model.

Firstly, if there is continuous goal movement that the agent is unable to anticipate, or if there is continuous perturbation of the agent, then the agent would ideally want to redirect towards the presently engaged goal at every point. The rigidity of the analytic sections prevents such redirection. This would mean that non-analyticity would be required at every point and there would be no analytic sections. In other words, taking Model 2 and then increasing the number of points where instantaneous redirection is required logically leads you to adopt Model 3, i.e. non-analyticity at all points.

However, it could still be that Model 2 provides a correct description of the way that natural animate agents work in practice if they prefer analytic trajectories and are only capable of redirecting at a finite number of points. The truth is, of course, an empirical matter.

A second drawback of Model 2 is that expecting the agent to continually change path characteristic between analytic and non-analytic sections makes Model 2 overly complex.

Model 3, which involves modelling redirection using fully non-analytic paths, requires no change of path characteristic at any points. It also exhibits no undesirable inflexibility in between.

So although it will contain more unfamiliar elements compared to an analytic approach, the purely non-analytic approach provides a simpler, more unified and flexible model of redirection. This is better suited to the plastic redirection requirements for goal-directed behaviour. The approach has the potential for improved results in modelling C^{∞} smooth paths taken by goal-directed agents. It is this model, i.e. Model 3, that will be developed over the next section and the following chapter.

3.7. Directed Non-Analytic Versus Undirected Non-Analytic Paths

Having shown that Model 3, which uses purely non-analytic paths for points or intervals where there is active goal direction, has the potential to model C^{∞} smoothness in goal-directed paths in a simpler and more unified way than the other two models, this model can now be developed further. In particular it will be important to show how the model can be extended to model the projections associated with goal direction.

3.7.1. Directedness

Very often in goal-directed behaviour changes in goal position and perturbation of the agent due to lack of skill or external influences lead to the agent requiring redirection towards its goal. A current path shape, and hence derivative state, is likely, in a struggle for goal direction, to sometimes be a reflection of what has been needed to get to the present point rather than what is needed in the future. This means that the current path may suffer from goal indirection, i.e. it is unlikely to project through the goal in such conditions.

At other stages however, some paths may be directed. These paths will be hypothesised to project towards the same locus from all points. Projection of C^{∞} smooth goal-directed behaviour needs to be non-analytic to reflect the inherent plastic redirection and hence to be different from the Taylor series projection.

3.7.2. Analytic and Non-Analytic Projections

The first thing to point out is that non-analytic paths have both analytic and non-analytic projections from them at the same point. The analytic projection from a point on a non-analytic path is just the analytic path that the Taylor series, using all derivatives, predicts from that point. This is illustrated in the C^{∞} smooth, non-analytic path shown in Figure 3.4. In this case, with the Taylor series centred at x = 0.5, the path predicted is a linear one, whilst a quadratic prediction is produced with the Taylor series centred at x = 0.25. (N.B. In both of the predictions the *whole* Taylor series is used not a truncated Taylor series). These predictions disagree with the actual rest of the path because of the non-analyticity at the points of projection.

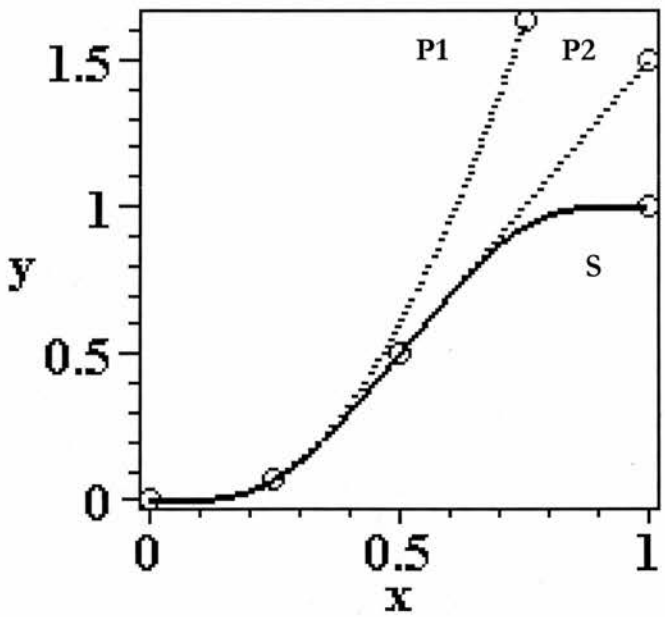


Figure 3.4. A C^{∞} smooth non-analytic s-curve (S) with polynomial analytic projections of P1 and P2. P1: $y = (-9/50) + x + 4(x-1/4)^2$ at x = 1/4; P2: y = (-1/2) + 2x at x = 1/2.

Since the projection needs to be non-analytic, as discussed in Section 3.7.1, a notion of projection different from that of projection from the Taylor series is required. However, there would appear to be no framework or mechanism in the community that provides any such notion of non-analytic projection. There is not, for example, a non-analytic equivalent of the Taylor series available at the moment.

To start the construction of such a framework, observations are made here about various general features of non-analytic projection. A specific mechanism enabling these features is then described in Chapter 4 and later.

A non-analytic projection from a point, rather than an analytic projection, is based on analysis of the derivative state at two points. Two points are used rather than one so as to obtain information as to how deviation from the Taylor series is proceeding.

The two derivative states on the curve will then need to be fed into a mechanism that allows the rest of the path to be interpolated (between those points) and extrapolated (beyond those points). [This is done in chapter 4].

It is not the case that non-analytic paths will always project through a goal. Sometimes non-analytic paths may suffer from goal indirection, i.e. they may not consistently project non-analytically through a goal locus from all points on the path. There may be no goal, or even if the behaviour is goal-directed, the path may be unexpectedly perturbed off course for example.

For the latter case, in Figure 3.5 for example, we see a goal-directed path S that starts at time t_o. At time t_1 the path projects non-analytically through the goal Gt_1. This projection is followed from t_1 to t_2. At t_2 the path still projects non-analytically through the goal at Gt_1 (Projection A). However, if the goal of the agent changes at this moment to Gt_2 then Projection A is no longer appropriate. The agent would then be suffering from indirection. To remedy this a redirection would be required such that the agent projects non-analytically along Projection B, which does project non-analytically through the new goal at Gt_2.

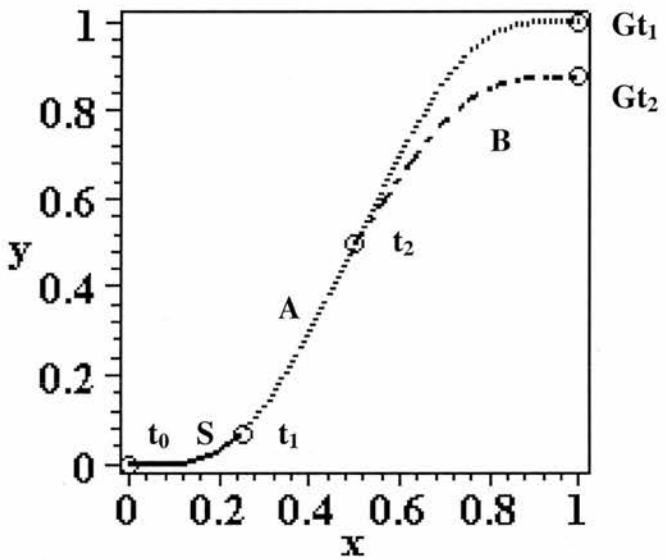


Figure 3.5. An example of redirection in action. Curve S with (A) A non-analytic projection passing from the point at time, $t_1 = 0.25$, through the goal at Gt_1, and (B) A non-analytic projection at time, $t_2 = 0.5$, through the goal at Gt_2.

Of course in goal-directed activity the paths are, according to our earlier characterisation, paths that often do project towards the goal, and indeed some non-analytic paths may be directed. These paths project non-analytically towards the same

locus from all points, e.g. the realised path on the Projection A in Figure 3.5 on the interval t = [0.25, 0.5] projects through the goal Gt_1 from every point on the interval.

Perhaps the most important feature of Model 3 is the fact that the non-analytic projections merge C^∞ smoothly with whatever curve has already been traversed, i.e. they redirect smoothly and pass through through the goal. Redirection becomes a matter of the agent simply taking a new non-analytic trajectory. Model 2, on the other hand required changing path characteristic from analytic to non-analytic at the joins (if maintaining the C^∞ smoothness requirement) in order to instantaneously redirect onto a new analytic trajectory. Non-analytic projection thus avoids the indirection or inflexibility problem that is endemic for the analytic approach and thus better typifies the redirection requirement for goal-directed behaviour.

All in all, given the novel nature of non-analytic projection, the non-analytic approach has the potential for novel and improved results for the analysis of goal direction – provided a predictive mechanism can be established.

3.8. Conclusion

Goal-directed behaviour is considered to be a key characteristic of intelligent and animate agents.

In computational approaches goal-directed behaviour can be analysed in terms of state space trajectories or paths.

The goal-directed behaviour of such agents can be characterised in terms of plastic and persistent action that keeps the path of the agent on course for the goal of the agent. It is supposed in the thesis that the plastic and persistent goal-directed behaviour of animate agents can be fruitfully examined through analysis of its smooth paths.

C^{∞} smooth paths can be analytic or non-analytic or have both analytic and non-analytic sections. Smooth closed form analysis to date predominantly uses analytic path types. However, it is argued that analytic paths have an inbuilt rigidity and indirection that is inappropriate for modelling the plasticity and persistence, respectively, in goal-directed behaviour.

Analytic paths are rigid because they have to obey Taylor series predictions. They are indirect because they are unlikely to continually project through the goal in the face of perturbations (such as lack of skill and external forces) and goal movement. Analytic approaches consequently have to model redirection by allowing for non-analytic points at the joins of non-analytic sections or by breaking the C^{∞} smoothness requirement and thus allowing discontinuities to appear at some derivative order at the joins.

We consider non-analytic paths as an alternative. Non-analytic paths and their sections may or may not be directed by goals. Directed non-analytic paths are those that project non-analytically towards the same locus from all points along the path.

The projection property of directed non-analytic paths means that they are able to model plasticity and persistence in a simpler and more unified manner than analytic paths.

It is therefore hypothesised that the goal-directed behaviour of animate agents may utilise directed non-analytic paths and that the locus that points project towards may be the goal locus of the agent.

In summary, a theory has been constructed that predicts that goal-directed behaviour is characterised by smooth non-analytic path types that repeatedly project through the goal locus rather than by the analytic path types standardly used by contemporary smooth closed form analysis. The non-analytic characterisation provides a simpler and more unified model of redirection than a piecewise analytic approach, with the potential for novel and improved results provided a predictive mechanism can be established.

Chapter 4

A Non-Analytic Computing Methodology For C^{∞} Smooth Goal-Directed Behaviour

4.1. The Tools of Efficient Cause

Chapter 1 presented the historical reasons why efficient causal analysis came to dominate modern science. Chapter 2 then looked at the techniques, and the attitudes towards final causal analysis, that this domination led to in contemporary approaches to goal-directed behaviour. Chapter 3 then developed a computational framework for use in C^{∞} smooth goal-directed behaviour research. In Chapter 4 a non-analytic computing methodology for C^{∞} smooth goal-directed behaviour research is developed.

Efficient causal analysis, as described in Chapter 1, sees behaviour as a result of the past events prior to the motion pushing the entity forward. Such events can result from a variety of sources, e.g. gravity, neural firing patterns and the actions of other agents.

An important addition to the armoury of efficient causal analysis came with the construction of the Newtonian differential calculus. The calculus was built for the calculation of rates of change of a dependent variable, and so the slopes of curves (Borowski & Borwein, 1999, p.64; Hahn, 1998). Applying the idea of a rate of change to successive derivative orders allows for the clarification of kinematics concepts like velocity (the rate of change of displacement with respect to time), acceleration (the rate of change of velocity with respect to time), jerk (the rate of change of acceleration with respect to time) and snap (the rate of change of jerk with respect to time), etc., and for the use of such concepts in a kinematics analysis of motion.

The Newtonian calculus itself is causally neutral. That is, the derivative variation the calculus reveals and measures may be seen in a variety of behaviours caused in a variety of ways, both analytic and non-analytic, both efficient and final causal, i.e. all behaviours with a degree of smoothness.

However, the Newtonian calculus also provides the basis for the important efficient causal computational method, described in Section 3.5, i.e. the Taylor series. The Taylor series is effectively a method that takes into account the influences of all the various derivative orders in an efficient causal manner to predict and analyse curves.

The derivative state in the Taylor series is the smooth computational version of an efficiently caused deterministic system state. By this it is meant that:

(4.1) $Q(t+1) = G(Q(t), S(t))$

where Q(t+1) is the internal state at time t+1, and G is a relation between the internal state Q(t), which is the (efficient) causal summary of the system's history up to time t and the current input to the system (S) at time t determining Q(t+1) (Minsky, 1972, p.17). In the case of a Taylor series, given a displacement, velocity, acceleration, etc. at t = 0, it is expected that the internal derivative state determines the next state at time t = 1, 2, ... , so that there is an efficient causal state determining the behaviour. The system in which the Taylor series is embedded depends in this thesis on the experimental setting. The setting and the system will be described in Chapter 5.

The Newtonian calculus and the Taylor series thus provide an efficient cause based framework for C^{∞} smooth motions with practical and conceptual computational tools for studying analytic behaviour.

4.2. The Need for New Tools of Computational Analysis

In the previous chapter a theory of C^{∞} smooth goal-directed motion was set out that hypothesises the existence of a non-analytic component to the behaviour for enabling *plasticity* and *persistence* in pursuit of a goal. Plasticity was defined in terms of behavioural flexibility in pursuit of a goal and persistence was defined in terms of the

ability to keep trying for the goal when off course (see Section 3.2 for examples of this characterisation).

The presence of a non-analytic component in such C^{∞} smooth goal-directed behaviour has consequences for computational analysis because analysis is usually carried out through analytic functions (which, as we know, have to obey their Taylor series predictions locally). Analysis using such purely analytic tools is inappropriate in the present case though. The Taylor series, for example, can only predict C^{∞} smooth *analytic* motion. The Taylor series cannot predict the course of C^{∞} smooth *non-analytic* motion and neither can any other tool that is analytically based.

Conventionally, when working with Real World data that is smooth to at least some degree, a data modelling assumption is made that presupposes analyticity. For example, a single analytic function such as a polynomial may be used to model the data. Alternatively, if a single analytic function does not provide a good model, a series of analytic functions may be used to build up a partially smooth, piecewise analytic function (see Model 2 described in Section 3.6), as in splines (see Press et al, 1992, pp.105-128). This may be done even though the data may be smoother, e.g. C^{∞} smooth, than the piecewise analytic approximation.

Such an assumption is overly biased when the investigation is concerned with the possibility that its C^{∞} smooth behaviour may contain lawful non-analytic aspects. In short, all trace of non-analyticity is prevented from appearing by the assumption of analyticity. In order to tackle the possibility of non-analytic components in the motion head-on rather than avoiding the issue, new tools, not solely analytically based, are

needed if C^{∞} smooth non-analytic behaviour as well as analytic behaviour is to be revealed, identified and predicted.

4.3. Laws of Goal-Directed Behaviour and Their Relation to Curve Fitting

The primary aim of the thesis is the construction and testing of initial laws or principles of non-analytic C^{∞} smooth goal-directed behaviour in the form of hypotheses. These principles will be a modest start, and will be principles of symmetry for tasks with various degrees of symmetry. They are intended to offer indicators for generalisation to more complex scenarios. Some data modelling will, nevertheless, be attempted on the basis of these new laws or principles and their implied constraints in a specific empirical setting. This section lays out the relation between these two activities.

There are three well-accepted types of data modelling. Firstly, curve fitting without the use of underlying laws. Secondly, modelling on the basis of underlying laws. Thirdly, modelling under known constraints.

As Press et al. (1992, p.656) explain it,

> 'Given a set of observations, one often wants to condense and summarize the data by fitting it to a "model" that depends on adjustable parameters. Sometimes the model is simply a convenient class of functions, such as polynomials or Gaussians, and the fit supplies the appropriate coefficients. Other times, the model's parameters come from some underlying theory that the data are supposed to satisfy; examples are coefficients of rate equations in a complex

network of chemical reactions, or orbital elements of a binary star. Modelling can also be used as a kind of constrained interpolation, where you want to extend a few data points into a continuous function, but with some underlying idea of what the function should look like.'

In the study of C^{∞} smooth goal-directed behaviour what we would like to do, ideally, is to use a theoretical model based on general laws as the basis for modelling of data. However, there are currently no well-accepted general laws or principles of non-analytic goal-directed behaviour that make predictions about the *kinematics* involved, i.e. the way that the motion changes in displacement over time and in the derivatives of displacement such as velocity, acceleration, etc. without consideration of force or mass.

So the thesis is aimed at taking the first steps towards providing such laws or principles in the form of hypotheses that can then be used to allow the construction of theoretical models capable of empirical testing in C^{∞} smooth goal-directed behaviour, and in particular here, its kinematics. Data modelling can then be attempted on the basis of these hypotheses in a specific empirical setting. In other words, the thesis will use characteristics such as C^{∞} and undulatory smoothness, plasticity, persistence and various kinds of symmetry to construct a theoretical model that can then be used to serve as the basis for law-based modelling. This in turn will give some idea of what the underlying function could look like and so allow Press et al's second type of constrained interpolation albeit in a limited form. The thesis is thus more relatable to the last two of Press et al's types of data modelling rather than their first type of fitting to a 'convenient class of functions'.

Because these are the first steps in establishing possible non-analytic principles, the decisions to be made are more qualitative than quantitative, e.g. is the function analytic or non-analytic? The constrained interpolation will be limited in being based on a limited degree of fit to the data – a fit good enough though to give confidence in the qualitative decisions. The intention is to investigate possible laws that may serve as the starting point for future models to provide good fits in the classical sense.

As such, the thesis is a *precursor* to classical data modelling – providing hypotheses which are then tested for their validity – rather than directly being any of the three types of curve fitting (to some 'convenient class of functions', to accepted laws, or to constrained interpolation).

In particular the thesis aims to provide observer prediction of the agent trajectory on the basis of derivative states rather than *a posteriori* curve fitting (or curve shaping). By itself, the latter is a relatively weak way of modelling autonomous behaviour. This is because its principles do not model the generating behaviour that allows the agent to cope with new situations. Demonstrating the ability to predict trajectories in advance in natural animate agents, on the other hand, rests on principles that model how the behaviour will be generated. This also opens up the possibility that the principles enabling the predictions could be applied in artificially animate agents. A particularly useful aim in line with this would be to show that the hypothesised shift away from Taylor series predictions can become a function that artificially animate autonomous agents may use to drive their goal-directed behaviour.

4.4. The Non-Analytic Shift Curve

The comment was made in Section 3.7.2 that there appears to be no mechanism in the community for non-analytic projection, e.g. no non-analytic analogue of the Taylor series. In line with previous comments about the necessarily limited and introductory nature of the thesis as far as non-analytic laws or principles are concerned (see Section 4.3), it is more appropriate to consider, not a general form, but a specific form of non-analytic projection suited to the thesis that may be generalised later.

The specific form argued for is initially motivated by the smooth symmetric motion reported in human kinematic studies. For example, in studying kinematic properties of rapid hand movements, Novak et al. (2000, p.419) report that:

> 'The movements without overlapping submovements on average had a near symmetric bell-shaped velocity profile that was independent of speed.'

In order to see what a symmetric form of non-analytic projection might look like, consider the following problem. Suppose human agents are required to use a mouse to move a ball between two parallel pipes on a computer screen (see Figure 4.1). This task, which will be referred to as the Parallel Pipes Problem, is symmetrical in the start and finish motions and about the midpoint between the two pipes.

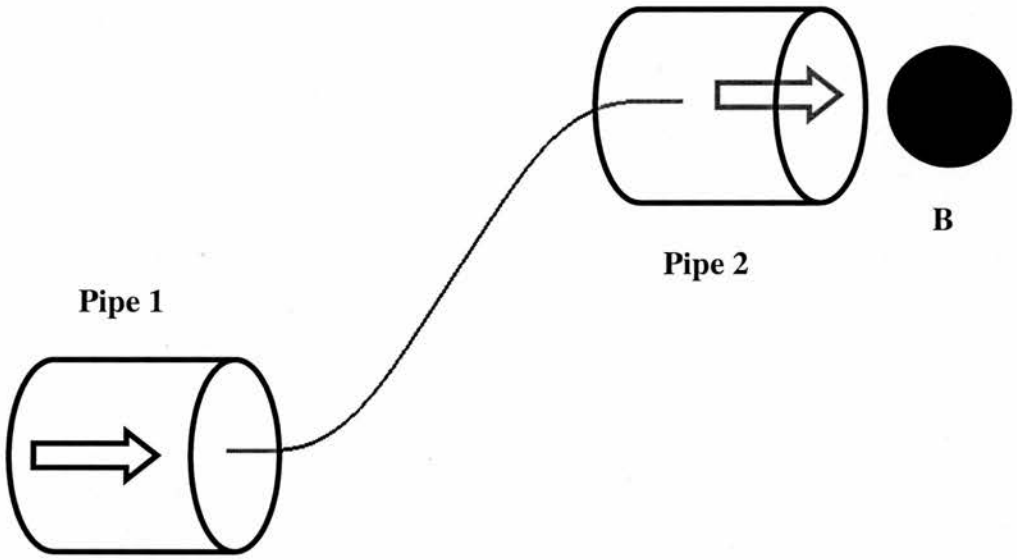


Figure 4.1. The Parallel Pipes Problem. The ball, B, must move through Pipe 1 then up to and through Pipe 2. (N.B. this figure is an illustration of the task, not a screenshot).

The avoidance of unnecessary changes in curvature or its sign suggests an s-shape as a rough initial intuition for achieving a low frequency symmetrical smooth change in displacement between the pipes (see Figure 4.4). But precisely what form does the s-shape take and why?

It may be noticed that motion along the pipes is, in theory, linear and hence analytic. Consequently, the s-shape represents a deviation from the linear analytic projection, which is non-analytic, at least at the inward pipe ends. But what about non-analyticity elsewhere in the motion between the pipes?

The kinematics literature suggests smooth symmetry may be present in agent motion (e.g. Novak et al, 2000). Weir (2005) suggests that, as with other smooth behaviour, symmetrical functions are important for measuring non-analytic behaviour and also may be generated when the problem is simple and symmetrical.

Hence a kinematic analysis of the simple and symmetrical Parallel Pipes Problem stands a good chance of revealing shifts in goal-directed agent behaviour. In particular, the shifts may be responsible for non-analytic motion between the pipes. The establishing of such shifts would provide a contribution firstly to the predictive kinematic sequencing of voluntary human limb movements and secondly to the shape analysis of smooth goal-directed behaviour.

The first contribution comes about because the shifts enable prediction of how the kinematic values vary in their temporal sequence due to the influence of the voluntary goal. This is complementary to existing kinematic prediction which either concerns involuntary behaviour, or is more spatial, i.e. centred on how the derivatives occuring at one bodily location in the kinematic chain affect those in another part of the chain. The second contribution comes about because the shift provides a smooth form for analysing the shape of an agent's goal-directed behaviour.

The next step is to analyse possible answers for the above two questions, (i.e. what symmetries are plausible and what type of curve would be required to implement such symmetries in a smooth movement such as the one introduced in Chapter 3). To do this the symmetry for a particular type of s-shape is explored.

4.4.1 Curve Symmetry

There are many possible s-shapes for the motion between the pipes, but one in particular will be described in detail, as it will provide a basis for non-analytic projection in the thesis. This s-shape will be referred to as a *shift*. A key notion behind the shift curve is the hypothesis that, after taking into account various agent choices, the agent may execute the most symmetric path around the midpoint of the shift interval. Agent choices are task dependent to some extent and are best described in conjunction with the experiments and will be discussed in Chapters 5 and 6 where the experiments are laid out. Here the symmetry itself is explored. The motivation for exploring the symmetry of the shape is that it aids prediction of the projection. But precisely what symmetries are plausible and what type of curve would be required to implement such symmetries in a smooth movement such as the one introduced in Chapter 3?

To begin answering this question, a few basic features may be noted. An s-shaped motion implies monotonically increasing velocity from zero during the interval from the origin to midway. At the midway point velocity peaks to retain one of the symmetries described below which is rotational symmetry (see Figure 4.3). This also implies zero acceleration (see Figure 4.4). The monotonicity then reverses with the curve monotonically decreasing velocity to zero during the second half of the interval (see Figure 4.3). The s-shape discussed in this chapter will standardly have a 1:1 ratio between its overall displacement and its width. Other ratios are discussed in Section 4.5.

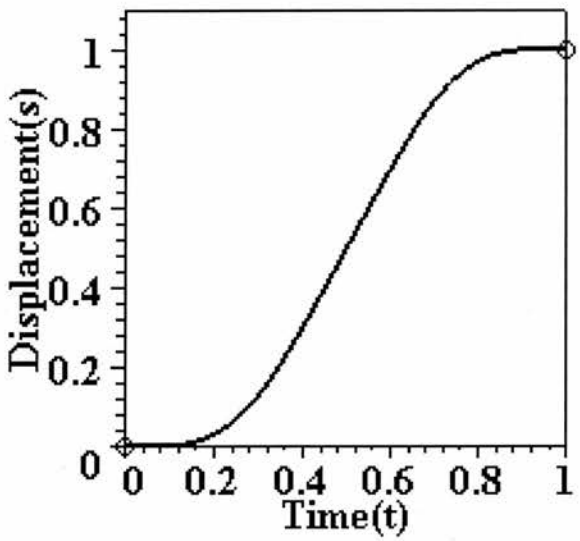


Figure 4.2. The shift curve (zeroth derivative).

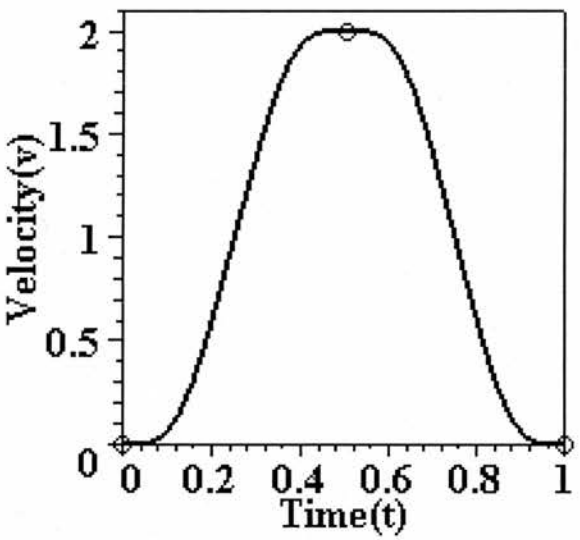


Figure 4.3. The shift curve (first derivative).

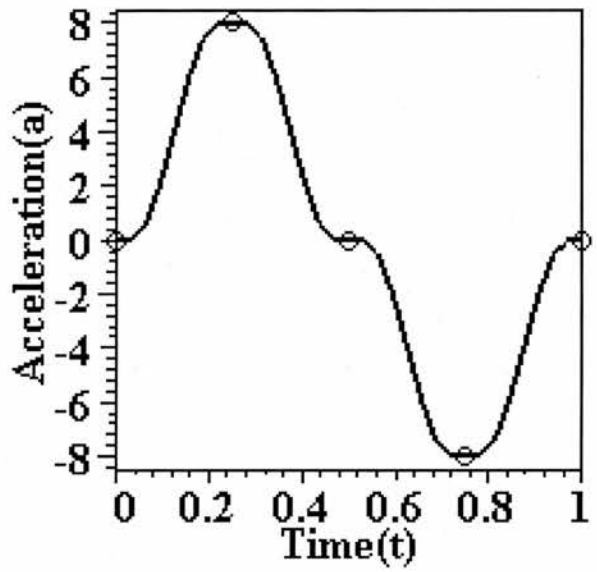


Figure 4.4. The shift curve (second derivative).

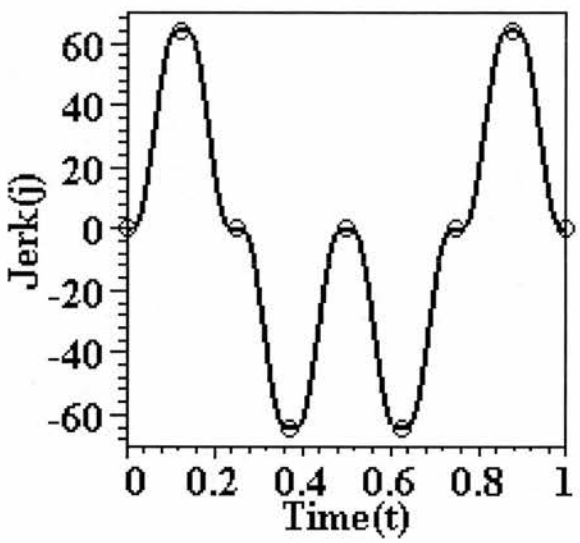


Figure 4.5. The shift curve (third derivative).

The first symmetries considered for the s-shape are those of rotation and reflexion. Rotational and reflexive symmetries around the midpoint of the interval alternating through all derivative orders are symmetries that are plausible on the basis that, all

things being equal, an agent has no reason to do otherwise due to the way the task itself is symmetrical about the midpoint. It will be seen through the examples below that the alternation is the result of alternating mixes of positive and negative rates of change at successive orders of derivative.

The first of these two symmetries is that of rotational symmetry about a point, e.g. a cubic polynomial symmetrical about the origin. An example of this, $y = x^3$ over the interval [-1, 1], is shown in Figure 4.6. In the construction of the shift curve this type of symmetry may be applied around the midpoint of the interval for the zeroth order and even orders of derivative.

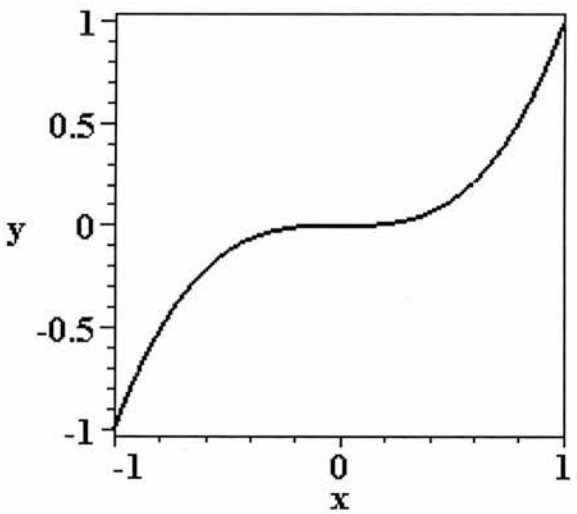


Figure 4.6. Graph of $y = x^3$ over the interval [-1, 1].

The second symmetry for the shift curve is that of reflexive symmetry about a vertical line, e.g. the quadratic which is the first derivative of $y = x^3$ and which is symmetric about the line $x = 0$ (illustrated in Figure 4.7). In construction of the shift curve this

type of symmetry may be applied around the midpoint of the interval for odd orders of derivative.

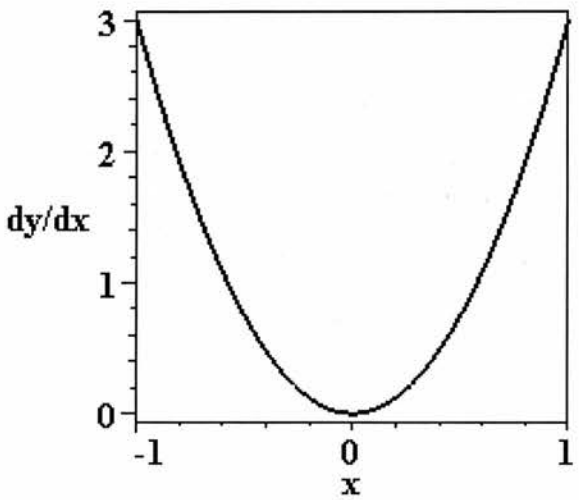
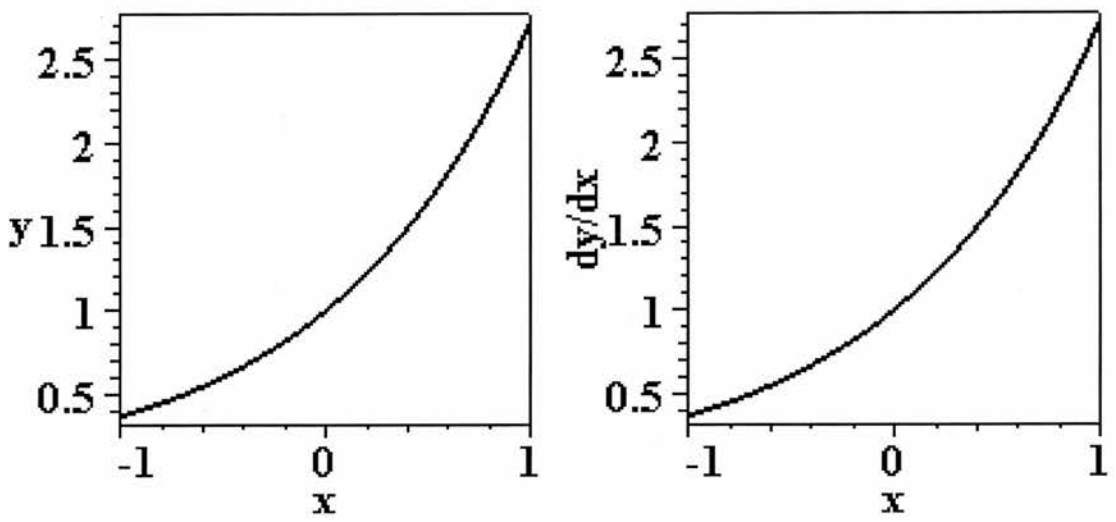


Figure 4.7. Graph of $dy/dx = 3x^2$ over the interval [-1, 1].

A third symmetry may be added in a similar way to the rotational and reflexive symmetries. This third symmetry for the shift curve is that of recursive symmetry through derivative orders.

An example of a curve that demonstrates this symmetry, although not the first and second symmetries, is that of $y = e^x$. This is illustrated in Figure 4.8. The first derivative is just a repeat of the profile at the previous order, and this shape repeats through all derivative orders.



a) b)

Figure 4.8. a) Graph of $y = e^x$ over the interval [-1, 1]. b) Graph of $dy/dx = e^x$ over the interval [-1, 1].

In the construction of the s-shaped shift curve this type of symmetry may be plausibly applied around the midpoint of each interval for all orders of derivative. For example, if you consider the first half of the interval, in which the velocity increases from zero to a peak at halfway as another interval with an s-shape, it appears plausible for the agent to change velocity in a way that also has rotational and reflexive symmetries around the midpoint of this new interval. The same argument can be applied to the change in velocity during the second half of the motion. It can also be applied when considering the changes in acceleration and changes at all higher derivative orders in their version of the s-shape. In other words the symmetrical s-shape at the zeroth order may copy across to all other orders.

In more detail, the recursion entails that the s-shape is copied down to the next order of derivative from the origin to midway (and a back-to-back s-shape from the midway

to the end). This copying down the orders can be seen, for example, by observing the whole zeroth order (see Figure 4.2) then the first half of the first order (see Figure 4.3) then the first quarter of the second order (see Figure 4.4) and the first eighth of the third order (see Figure 4.5). This type of recursive symmetry then implies that there is not only zero acceleration at midway but also all higher orders are zero as well due to the copying of the zero-valued tail end gradients of the s-shape into the higher order profiles.

4.4.2 Analytic Projections

The flat zero-valued nature of the derivative profiles above the 1st order at mid-way results in a linear analytic projection at the midway point (see Projection P2 in Figure 4.9). Note that this linear projection is that of the whole Taylor series, and not a truncated form of it. At the quarterway and three-quarterway points there are quadratic projections due to non-zero velocity and acceleration coupled with zero values for higher order derivatives (see Projections P1 and P3, respectively, on Figure 4.9). These projections are again those of the whole Taylor series. This continues through all derivatives, with the order of the polynomial projection increasing by one at each new bisection of the s-shapes present at the previous order (see Figures 4.10, 4.11 and 4.12 for the projections present at the first, second and third order derivatives, respectively, from the quarterway, halfway and three-quarterway points).

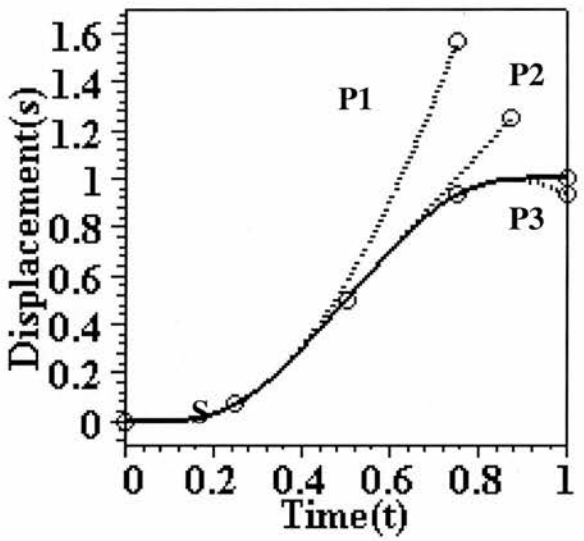


Figure 4.9. Analytic projections from the shift curve zeroth order derivative. P1: the quadratic analytic projection from the quarterway point. P2: the linear analytic projection from the halfway point. P3: the quadratic analytic projection from the three-quarterway point.

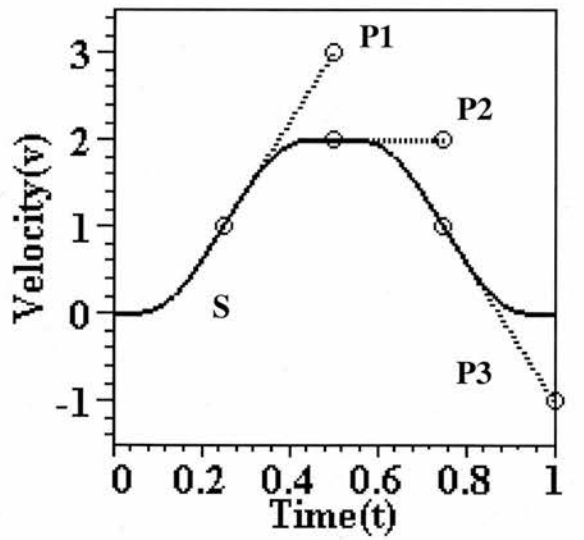


Figure 4.10. Analytic projections from the shift curve (S) first order derivative. P1: the linear analytic projection from the quarterway point. P2: the linear analytic projection from the halfway point. P3: the linear analytic projection from the three-quarterway point.

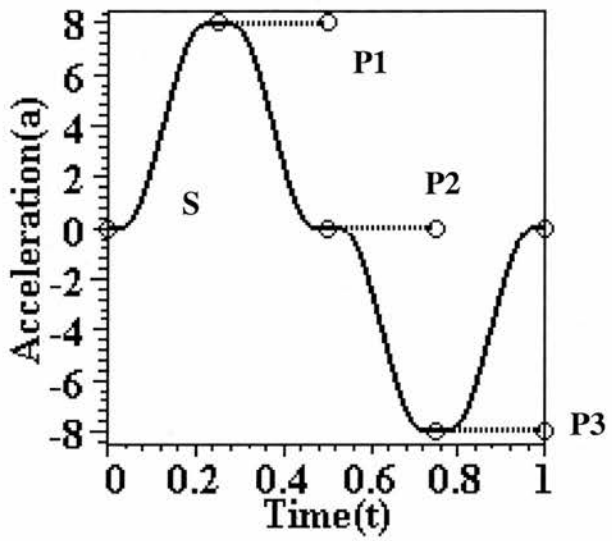


Figure 4.11. Analytic projections from the shift curve (S) second order derivative. P1: the linear analytic projection from the quarterway point. P2: the linear analytic projection from the halfway point. P3: the linear analytic projection from the three-quarterway point.

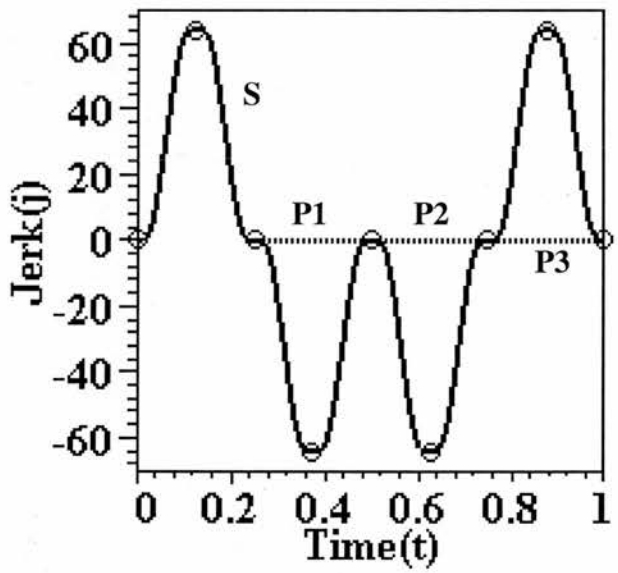


Figure 4.12. Analytic projections from the shift curve (S) third order derivative. P1: the linear analytic projection from the quarterway point. P2: the linear analytic projection from the halfway point. P3: the linear analytic projection from the three-quarterway point.

4.4.3 Non-Analyticity

The change in analytic projection required at each point by the s-shape suggests that no single one of the analytic projection polynomials could ever join the two pipes smoothly to become the rest of the s-shape. Indeed, if any polynomial did join one of the two pipes smoothly at its inward end, it would be a horizontal straight line by virtue of its derivative state at this end – which is impossible.

The polynomial and changing nature of each analytic projection shows that the type of curve you need to implement the s-shape in this strongly symmetric fashion cannot be piecewise analytic (as in Model 2 described in Section 3.6) and so is *non-analytic everywhere* during the shift (as in Model 3 described in Section 3.6).

The C^{∞} smooth non-analytic curve that these symmetries results in, called the shift curve or s-curve, and which is defined through the three symmetries, has been constructed by my supervisor (Weir) prior to the thesis investigation. This curve has the zeroth, first, second and third order derivatives seen in Figures 4.2, 4.3, 4.4 and 4.5, respectively, and also the derivative profiles suggested by extending the established pattern. The curve is defined exactly by the pair of differential equations (Equations 4.2 and 4.3 below):

(4.2) $$s'(t) = 2s(2t) \qquad t \in [0, \tfrac{1}{2}]$$

(4.3) $$s'(t) = 2 - s'(t - \tfrac{1}{2}) \qquad t \in \left]\tfrac{1}{2}, 1\right]$$

coupled with the basic requirements for an s-curve, in particular its known end points and rotational symmetry. See Appendix A for a more detailed and constructive treatment of the shift curve's specification.

In the thesis, the way the shift curve is non-analytic through its recursion will be emphasised more than its specification through the above differential equations. The non-analytic values themselves will be based on a tabular version of the shift curve. As Penrose (2004, p.104) explains this approach to the description of functions, 'Nowadays one prefers to think in terms of 'mappings', whereby some array A of numbers (or of more general entities) called the *domain* of the function is 'mapped' to some other array B, called the target of the function... The essential point of this is that the function would assign a member of the target B to each member of the domain A... This kind of function can be just a 'look-up table'. The actual height and time values of the hypothesised curve described in this chapter can be found in a table in Appendix B.

4.4.4. The Shift and its Qualifiers

The deviation from the Taylor series prediction that can be seen throughout the s-curve will be referred to from now on as a *shift*. A shift can be of *any derivative order* and can be a *full* shift or a *partial* shift in that order of derivative. For example, a full displacement shift is shown in Figure 4.2. If only the interval [0, 1/4] of Figure 4.2 is taken then the resulting profile is a partial displacement shift of one quarter. Shifts can also be *upward* or *downward*. A full upward velocity shift is shown over the interval [0, 1/2] of Figure 4.3 whereas a full downward velocity shift is shown over the interval [1/2, 1] of Figure 4.3. Finally, shifts can also be positive or negative. Figure 4.4 has a positive upward shift over [0, 1/4] followed by a positive downward shift over [1/4, 1/2] followed by a negative downward shift over [1/2, 3/4] then finally a negative upward shift over [3/4, 1]. In other words, there is an increasing acceleration then a decreasing acceleration then an increasing deceleration and finally a decreasing deceleration.

4.4.5 Shifts as a Particular Form of Non-Analytic Shift Projection

Besides being possible as the motion for a particular problem such as the Parallel Pipes Problem, shifts may also be used more generally.

In the case of the Parallel Pipes Problem (see Figure 4.1) the shift is able to provide a C^{∞} smooth deviation from the Taylor series starting at the rightmost end of the Pipe 1 and ending at the leftmost end of Pipe 2. As such, it can provide a symmetric smooth merge from one derivative state to another.

But also, the derivative states between which the merge takes place may differ in other ways than just the zeroth order derivative values as in the Parallel Pipes Problem (i.e. only the vertical displacement value is different at the two temporal ends of the merge, whilst the velocity and all higher derivative values are zero at both ends of the merge). So, for example, another problem may involve ending the merge with a different value for velocity than at the start, as well as a different vertical displacement value.

In the terminology used in the thesis a shift in which only displacement values at the two ends of the merge differ is known as a displacement shift. This is the simplest case of shift, and is the only type of shift necessary to complete the Parallel Pipes Problem task successfully. If the two ends of the merge differ in their velocity as well as their displacement, but not at higher orders, then it is called a velocity shift. The next three higher order shifts are known as acceleration, jerk and snap shifts.

Higher order differences follow from higher order shifts and differences in various combinations of orders follow from various combinations of shifts. Such combining of shifts can allow for asymmetric paths to be built up even though the basic unit is symmetric. This is analogous to the use of linear gradients in a Taylor series building up non-linear curves.

Hence shifts can provide particular symmetric unitary plastic deviations from the Taylor series. The shift forms a symmetric bridge between the current state and the goal state. The state-determinacy in the Taylor series is replaced in the case of a goal-

directed shift by the current and goal states together determining what happens – a shift determinacy.

A key concept of this shift determinacy is that the agent will take the most symmetric curve in merging plastically between derivative states unless there is reason to do otherwise.

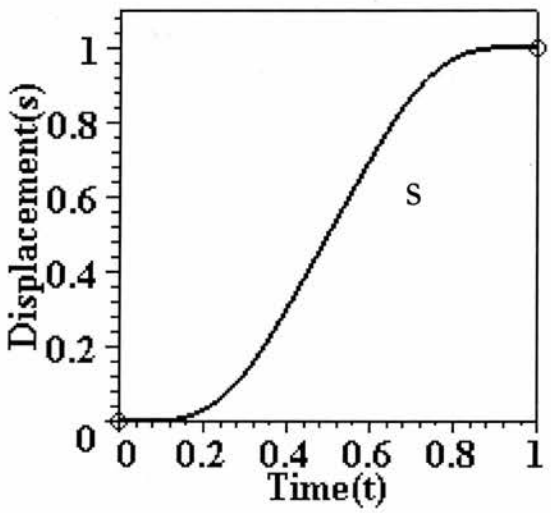


Figure 4.13. A C^{∞} smooth non-analytic s-curve (S) that is a symmetric shift.

It is hypothesised that the path will continue such as to complete a symmetric shift. Knowing the two derivative states on the curve allows the rest of the path to be interpolated (between those points) and extrapolated (beyond those points) to the two ends of the shift. If the two known derivative states are at the end points of the shift (at (0, 0) and (1, 1) in Figure 4.13, for example) then the path between the two is a pure interpolation. For goal-directed behaviour, the completion point of the shift may often be the goal of the agent, e.g. the point (1, 1) in Figure 4.13.

4.5 Kinematic Symmetry

My thesis concerns the exploitation of the shift curve for suggesting general principles for non-analytic goal-directed agent kinematics. In particular the shift will be used as a predictor for the derivative profiles to be expected in various human hand movement experiments.

This approach is based on the idea that a low-order shift is intuitively plausible as an appropriate non-analytic projection for an agent's goal-directed behaviour in simple symmetric tasks that may then be generalised to more complex and less symmetric action. The reasoning behind using the shift being that such projections provide the most symmetric merge of the current path with the goal.

Supposing that the shift function underlies agent kinematics in a variety of tasks or behaviours is to thus suppose that symmetry underlies this variety. So why should symmetry be assumed?

One justification can be found from starting from the counterpart analytic symmetry of linearity and answering a related question, i.e. why should linearity be assumed in a variety of behaviours? In this case the reasoning is that the trajectory will be linear because there is no cause for it not to be linear. Now at first sight, this seems like a rather weak justification. After all, why should other shapes not be justified in the same way?

However, consider the causal element in the reasoning. If another shape were to be given the same justification, it would not seem to be appropriate and would lead people to search for the underlying causes of the distinctive bends in the shape, rather than to presume that the bends just happen.

The reasoning is also supposing that there may well be possible causes of non-linearity but that these are extraneous causes that can be removed. Consider predicting what will happen when a stone is kicked across a pond and hypothesising a linear trajectory for example. Obstacles on the ice and the way a stone is kicked may cause non-linearity unless controlled for. These are extraneous causes of non-linearity that may be removed. *If* this is true and the causes of non-linearity are indeed extraneous rather than inherent, the linearity seems inevitable. The reasonableness of the assumption of linearity comes down to whether undermining causes are extraneous or inherent. The assumption of non-analytic symmetry in the problem at hand also comes down to this same question.

There are also scientific reasons for the assumption of linearity. Firstly, the hypothesis is testable, i.e. the trajectory will or will not be linear once extraneous causes are removed, and the outcome provides useful information as to whether the behaviour is inherently linear or not. This kind of approach was at least acceptable to the Renaissance scientists and worked for them. The same type of reasoning applies when hypothesising non-analytic symmetry.

There are reports already in the literature of smooth symmetry in kinematics (Novak et al., 2000, for example). It thus appears to be both a plausible hypothesis and one

that provides valuable information even if it does not come off. It is in short, a reasonable and worthwhile scientific hypothesis.

Returning now to the projective use of the shift, in restricted settings, i.e. low-order symmetrical experimental set-ups, a low-order shift may be started where the projection through the goal is a continuation, a completion, of the shift. The shift projection follows from the local curve segment (part of the way in) due to the recursive property. This can be contrasted with a different non-analytic curve such as that provided by Merryfield (1992, pp.132-138) and reproduced in Figure 3.3, where the local curve segment does not appear to be readily usable to predict a future projection. Indeed to predict the continuation of Merryfield's curve with any confidence may require the symbolic form of the curve to be known *a priori*. This may well not be available when dealing with Real World problems.

Overall the idea of the approach is to observe the data points and project from the data what happens in between and beyond the data points. In the analytic case attempts to do this use a single derivative state. As will be seen in the following section a non-analytic equivalent to this can be based on a pair of derivative states. This pair determines a shift from which a non-analytic projection follows.

4.6. The Use of a Shift Calculus.

Newtonian differential calculus is based on the notion of a gradient, i.e. the slope of the curve at a point. The gradient is measured as the ratio of the change in displacement to the change in width for an interval along the curve. Shifts can be

thought of as a non-analytic analogue of the gradients in Newtonian differential calculus in that they also have ratios that measure variation, i.e. variation away from the Taylor series prediction. This ratio is also the change in displacement to the change in width along an interval. In the case of the shift, the interval is the interval containing a full shift. This shift ratio enables different shifts (of the same derivative order) to be distinguished (see Figure 4.14 of two different shifts and their ratios).

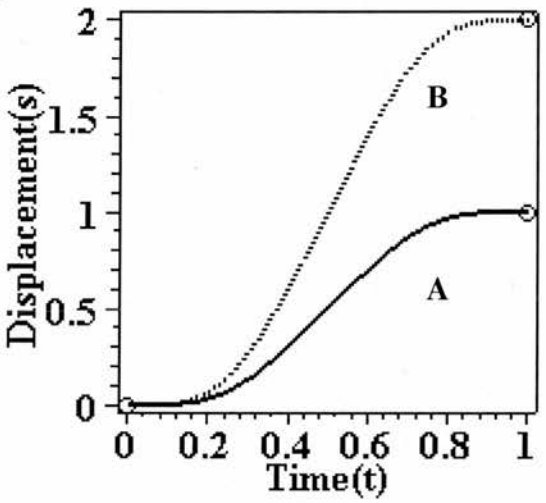


Figure 4.14. (A) a 1:1 full displacement shift and (B) a 2:1 full displacement shift.

In this section a method will be described for identifying whether a non-analytic shift is present in a curve or not and also the type and magnitude of any such shift. An extended (as opposed to infinitesimal) calculus for using shifts will be presented in the thesis.

4.6.1. Taylor Series Filtering: Method

Any C^{∞} smooth curve can be tested for deviation away from the Taylor series prediction. In order to perform this test the first step is to *centre* the Taylor series at the start point for possible deviation. Predictions for the rest of the curve are made on the basis of the derivative state at this point (e.g. the origin in Figure 4.15). The second step is to then *filter* the curve through the Taylor series. By this it is meant that the displacements suggested by the Taylor series prediction are subtracted from the displacements actually reached along the curve at each time, t. The displacements that remain after filtering form a curve that will be called the *residual* (see Figure 4.16). The equation for finding the residual is:

$$R(t) = A(t) - TSP(t) \tag{4.4}$$

Where R(t) is the displacement of the residual curve at time t, A(t) is the displacement of the actual curve at time t and TSP(t) is the Taylor series prediction of the displacement at time t.

By observation of the residual of a curve after Taylor series filtering it is possible to discover the presence or absence of deviation from the Taylor series in a number of ways, as discussed in the next section.

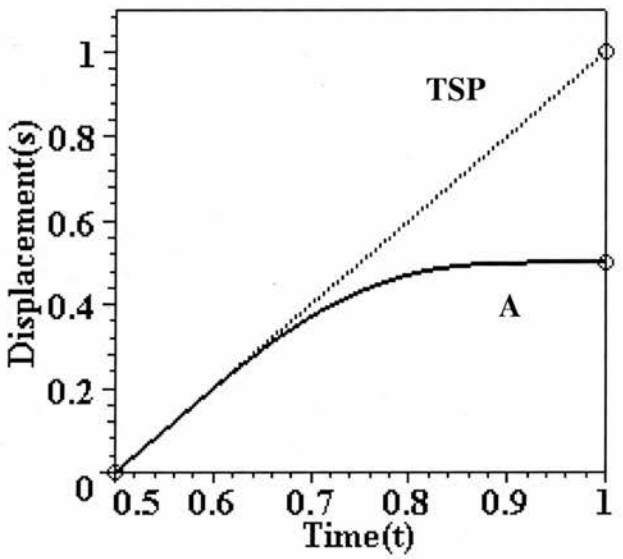


Figure 4.15. Graph of an actual curve (A), which is a translation of the second half of the shift curve to the origin, and the Taylor series prediction centred at t = 0.5 (TSP).

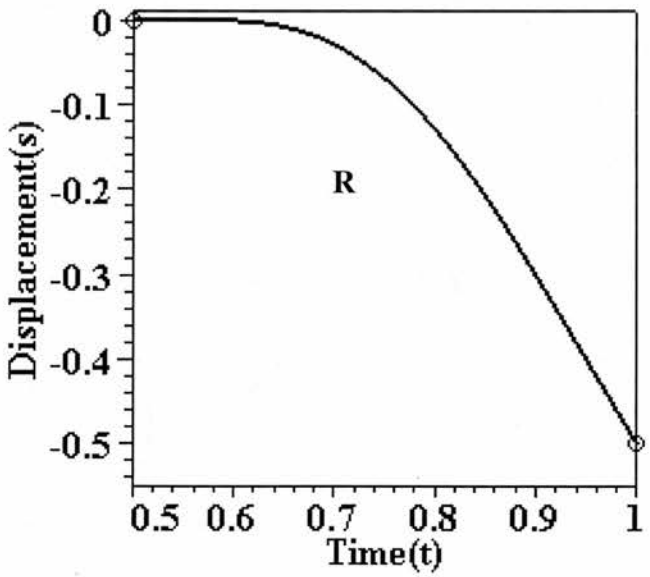


Figure 4.16. Graph of the residual curve (R) after subtracting the displacements of each point on the Taylor Series Prediction (TSP) in Figure 4.13 from the actual displacements, (A), at the same times.

4.6.2. Application of the Full Taylor Series Filter: Conclusions

There are three possible results of Taylor series filtering using a full Taylor series in terms of the type of curve that is present. Firstly, if a zero-valued residual occurs then deviation from Taylor series predictions is absent and therefore the actual curve is analytic. In Figure 3.1, for example, the Taylor series predictions overlap the actual curve. Subtracting the prediction from the actual curve leaves a zero-valued residual – the actual curve has been recognised as analytic (in this example it is the polynomial $y = x^2$).

Secondly, if a non-zero-valued residual occurs then deviation from Taylor series predictions is present so that the actual curve is non-analytic at least at the initial point of prediction. Figure 4.15 shows the actual curve A (which is the second half of the shift curve translated to the origin) and its Taylor Series Prediction (TSP) centred at $t = 0$. Subtracting TSP from A leads to the residual found in the curve (R) in Figure 4.16, showing that significant deviation from the Taylor series is present and thus that the curve is non-analytic (at least at the start point).

Thirdly, if the actual curve is a shift, then there will be an extra property observable in the residual curve. Namely that the residual curve will be a partial copy of the original shift up to a change in sign. The partial copy property for the residual is present because of the recursion in the shift. At every point in the shift a deviation from the Taylor series centred at the point starts up which is a partial copy of the shift (up to a change in sign). This property is exhibited by the residual curve R, shown in Figure

4.16, for example, where the residual is just a negative copy of the first half of the shift curve itself. Thus the actual curve (A), in Figure 4.15, is recognised as a shift. (See Appendix A for more detail on this property).

4.6.3. Application of the Truncated Taylor Series Filter: Conclusions

Of course in situations where Real World data are being analysed, such as in the experiments to be described in Chapters 5 and 6, it may not be possible to use the whole Taylor series in filtering. In such cases a truncated Taylor series using only a finite number of terms is all that is available. So can a truncated Taylor series be of any use in helping to distinguish between analytic curves and the non-analytic shift curve?

The answer lies in considering the number of non-zero terms that are evaluated in the approximation. If sufficient non-zero terms are evaluated to yield a good approximation of the full underlying non-zero Taylor series then a distinction can still be made between an analytic curve and a shift.

A good approximation will result in a significant non-zero subtraction of the truncated Taylor series prediction from the actual curve because for the curve in question to be of interest, it will be non-zero. This will make the residual in the analytic case recognisably different from the original curve. Consider the case of s = sin(t), shown in Graph 4.17, for example. If a truncated Taylor series is used with an increasing number of non-zero terms the prediction more accurately reflects the actual curve, see Figure 4.18. In this case just using one term gives a zero displacement prediction

(TSP1), two terms (TSP2) a straight line of non-zero gradient, four terms (TSP4) starts to give some of the bend required for the true s = sin(t) curve and six terms (TSP6) does even better (N.B. predictions using three and five terms are left out as the third and fifth terms of the Taylor series prediction for s = sin(t) are both zero and so add nothing to the prediction with one fewer term and so are just identical to them).

Subtracting the prediction in each case from the actual curve gives the residual curves shown in Figure 4.19. For one term (R1) the residual is just the curve s = sin(t) itself. For two terms (R2) the residual starts to flatten out towards the start before heading off into the negative. For four terms (R4) the residual flattens out even more before heading off into the positive. With six terms (R6) the residual has almost flattened out completely with just a slight negative dip becoming noticeable towards the end of the interval, meaning that the Taylor series prediction with six terms continues to improve the approximation of the actual curve, s = sin(t). So with two, four or six terms the residual looks very different to the actual curve and is diminishing to zero at least locally as more terms are added.

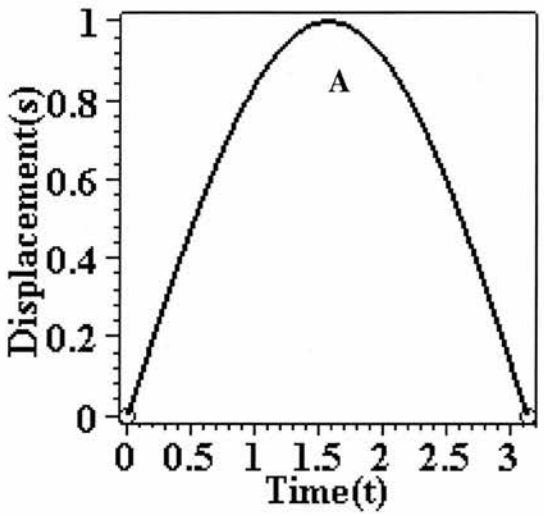


Figure 4.17. Graph of an actual curve (A) s = sin(t) over the interval [0, π].

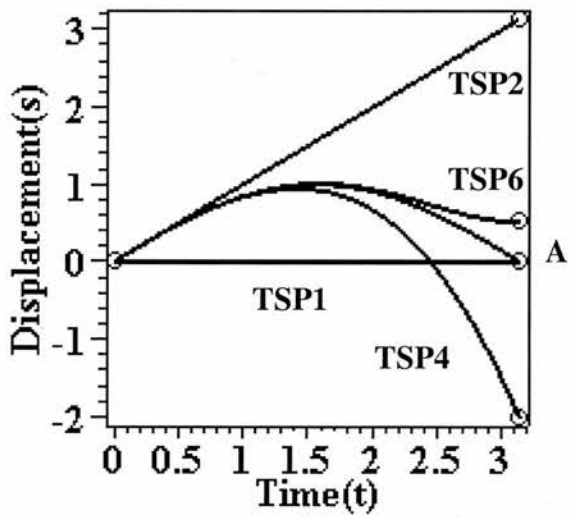


Figure 4.18. Truncated Taylor series predictions from t = 0 on the curve s = sin(t), shown here as A, over the interval [0, π] with one term (TSP1), two terms (TSP2), four terms (TSP4) and six terms (TSP6).

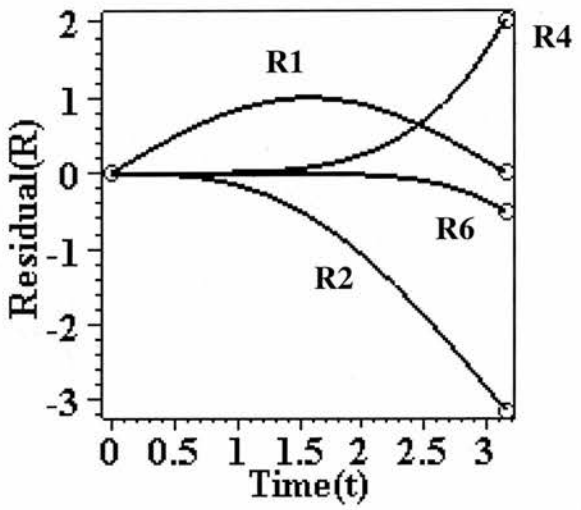


Figure 4.19. Residual of s = sin(t) over the interval [0, π] using a truncated Taylor series with one term (R1), two terms (R2), four terms (R4) and six terms (R6).

But what if the curve is the non-analytic shift curve? If it is the non-analytic shift curve then this can be distinguished from the analytic case. The reason being that a good approximation will make the residual in the non-analytic shift curve case recognisably similar to the original curve despite the significant non-zero subtraction. For example, consider the case discussed earlier in Figure 4.15. Here there is the second half of the shift curve translated to the origin and a full Taylor series prediction made. What happens when a truncated Taylor series is used instead? Figure 4.20 shows the predictions using one, two and three terms of the Taylor series from midway into the shift. With one term (TSP1) there is a zero prediction, with two terms (TSP2, solid line) there is a linear prediction and with three terms (TSP3, dotted line) there is the same linear prediction as with two terms. The latter occurs because the analytic projection is linear. This leads to the residual curves shown in Figure 4.21.

The residual with one term (R1) is just the original curve (A) itself. The residual with two (R2, solid line) and (R3, dotted line) is the same negative copy of the first half of the shift curve.

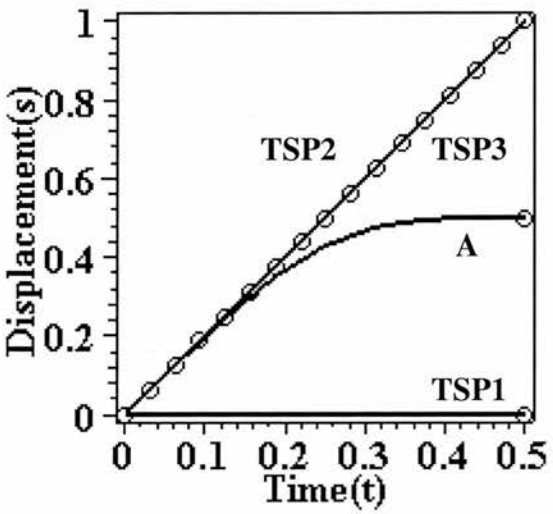


Figure 4.20. Taylor series predictions from t = 0 for the second half of the non-analytic shift curve (A) translated to the origin using a truncated Taylor series with one term (TSP1), two terms (TSP2, solid line) and three terms (TSP3, dotted line).

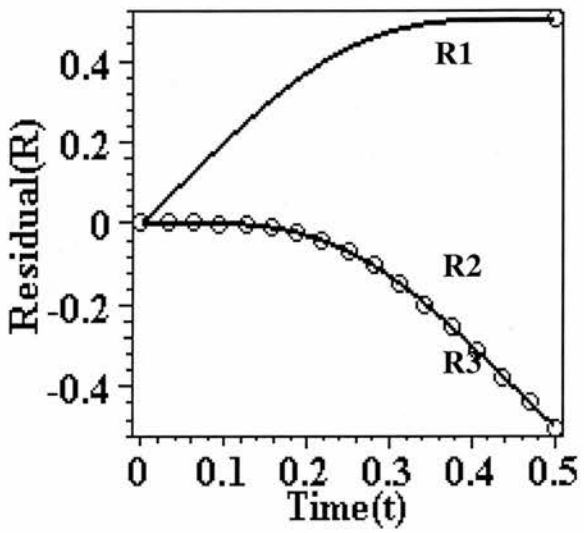


Figure 4.21. Residuals for the second half of the non-analytic shift curve translated to the origin, using a truncated Taylor series with one term (R1), two terms (R2, solid line) and three terms (R3, dotted line).

4.6.4. Non-Analytic Projection Through the Taylor Series Filter: Method

The Taylor series filter can also be used to create a non-analytic projection in the form of the completion of a detected shift. The first step in doing this is to compute the *stage* of a shift. This may have to be done using the measured data rather than *a priori* knowledge of the shift. Computing the stage of the shift may then be done using the ratio of the product of the filtered velocity and the width to the filtered displacement at that point:

(4.5) $T = V_f W / S_f$

Where T is the stage of the shift, V_f is the filtered velocity, W is the width, and S_f is the filtered displacement. In an initial example of this method consider the case where the width of the shift interval is known to be 1 such as in the shift curve shown in Figure 4.2. By examining Figure 4.2 (the zeroth derivative) and Figure 4.3 (the first derivative) at the halfway point and supposing that this is the filtered displacement and velocity at that point then equation 4.5 yields:

(4.6) $T = 2 * 1 / (1/2) = 4$

Doubling the displacement of the shift yields $S_f = 1$ and $V_f = 4$ at the halfway point. This now gives:

(4.7) $T = 4 * 1 / (1) = 4$

So whatever the vertical scaling of the displacement shift the same value for $V_f W / S_f$ results, if the width of the shift is the same. However, this restriction on the width needs to be removed to allow for the computation of the stage in shifts of any width. Suppose that the displacement shift shown in Figure 4.2 now has its width halved. The effect of this is to also double the velocity. This yields the zeroth and first order profiles shown in Figure 4.22 and 4.23, respectively. This now gives:

(4.8) $T = 4 * (1/2) / (1/2) = 4$

So yet again a value of 4 for the ratio is computed. But what do these three cases have in common? It is that they are all found at the halfway point of the shift in question, whether it is scaled horizontally or vertically.

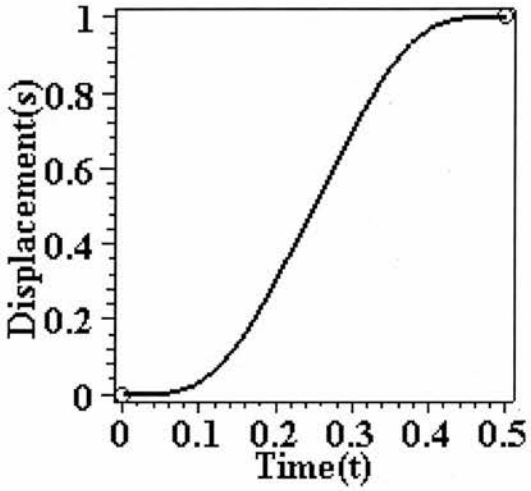


Figure 4.22. The non-analytic shift curve with halved width.

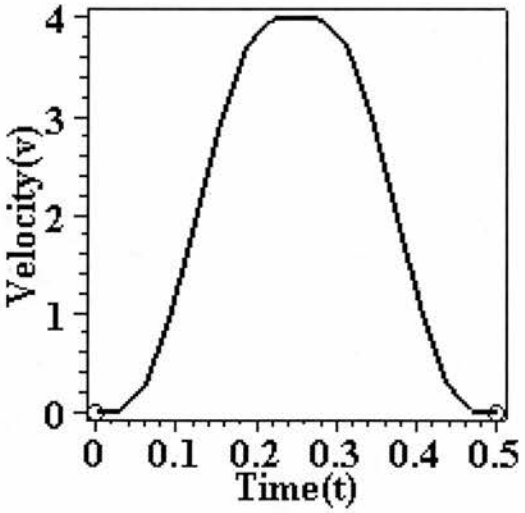


Figure 4.23. The first derivative of the non-analytic shift curve with halved width.

Indeed every stage on the standard shift curve (i.e. curve A in Figure 4.14) has a unique $V_f W / S_f$ ratio associated with it and the ratio is the same at the same stage across all shifts whatever the scaling. This general principle can be seen in Figure 4.24, which shows the $V_f W / S_f$ ratio for all stages (as a fraction of the overall shift) along the shift. This means that the stage that a shift has reached can be identified uniquely by computing the $V_f W / S_f$ ratio.

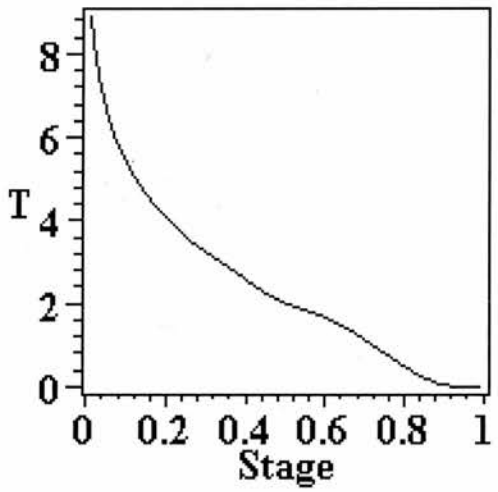


Figure 4.24. The $V_f W/S_f$ ratio (T) for the standard shift with the stage given as the fraction of the actual shift that has been completed.

The second step in creating a non-analytic projection is to compute the *shift ratio*. This is the ratio of the filtered displacement change of the (full) shift to its width. This may be expressed as the ratio of the actual filtered displacement change at the stage computed by the first step to that of a shift with a 1:1 ratio over the same shift interval. Consider the example of a doubled vertical displacement shift, i.e. curve B in Figure 4.14. The stage that has been reached is worked out by computing the $V_f W / S_f$ ratio. Suppose this is 4 (see Equation 4.7). This value is associated with the

halfway stage in any shift. The overall relative shift magnitude (SM) can now be worked out by dividing the actual filtered vertical displacement (A) at the point in question, which is 1 in this case, by the height reached by the shift curve at the same stage (S), which is 1/2 in this case, to yield an overall shift magnitude of 2. In equation form, for the shift magnitude (SM):

(4.9) $SM = A / S = 1 / (1/2) = 2$

To work out the shift ratio (SR) the overall relative shift magnitude must be divided by the overall shift width (SW) achieved at that point. In the current example, the overall shift width is the same as in the standard shift curve (i.e. 1). In equation form:

(4.10) $SR = SM / SW = 2 / 1 = 2$

In other words the shift has been correctly found to have double the shift ratio of the standard 1:1 shift curve using a pair of derivative states extracted from the data.

Computing the stage for a pair of points on a curve and the shift ratio allows interpolation between these two points using the comparable section of the shift curve (with the appropriate scaling on the basis of the shift ratio) fitted between this pair of points. By projecting non-analytically using the fitted shift it is also possible to extrapolate beyond the end points of the interpolated section to make predictions about the continuation of the curve. So suppose, for example, that the stage was found to be the quarterway stage and that the shift ratio was found to be double that of the standard shift for the same interval. This would allow the doubled displacement shift,

i.e. curve B in Figure 4.14, to be interpolated between the origin of the filter and the quarterway point and extrapolated beyond those two points as the rest of the (doubled displacement) shift.

For a worked example using the above methodology that uses experimental data see Appendix C.

4.7. The Empirical Motivation

The above has laid out a non-analytic computing methodology for C^{∞} smooth goal-directed behaviour centred around the shift.

Shifts may be viewed analogously to gradients in the theoretical sense of the transitions in smooth variation being able to be seen as combinations of shifts as well as gradients. This theoretical background has been put in place to facilitate the primary motivation of the thesis, which is an empirical motivation.

The empirical motivation of the thesis is to try and isolate non-analytic shifts in human goal-directed hand movements through the analysis of kinematics profiles. The experiments will be controlled to the extent that such shifts are allowed to emerge naturally if they exist in the movement. This is analogous to isolating laws and shapes of inanimate motion by pushing a stone across ice. In this example, by controlling for the effects of friction, the level of the ice, etc., the law and linear shape of inanimate motion f = ma (where f is the force on the body, m is the mass of the body and a is the acceleration of the body) is allowed to emerge.

In some cases, such as for the experimental settings used later, an s-shape will be appropriate. But it is an empirical matter as to whether this s-shape is the shift curve or some other alternative curve. If it is some alternative curve then it must lack at least some of the symmetry shown by the shift curve and may also be less smooth or less plastic. It is thus an interesting question for the nature of goal direction as to what exactly the human agent does exhibit in terms of symmetry, smoothness and plastic persistence.

4.8. Summary

The efficient cause based contemporary framework for C^{∞} smooth motion has the Newtonian differential calculus and the Taylor series to provide it with a mathematical basis for studying analytic behaviour.

The theory of goal-directed motion developed in the previous chapter needs its own analogous non-analytic tools for the prediction of potentially non-analytic C^{∞} smooth goal-directed behaviour if the experiments in the following chapters are to be analysed on a similar mathematical basis.

The theory also needs new laws of C^{∞} smooth goal-directed behaviour that can be used for modelling. However, as no such laws, that are predictive with regards to non-analytic derivative profiles in the motion, are currently accepted the thesis will be primarily concerned with developing initial laws or principles in the form of hypotheses. These hypotheses involve modest principles of symmetry intended to

form a basis for future models to provide good fits in the classical sense. As such the thesis is a precursor to classical data modelling.

The three symmetries of rotation, reflexion, and recursion are sufficient to define a particular C^{∞} smooth non-analytic curve. Every bisected section of the path of the resulting s-curve represents a shift away from the Taylor series prediction at some order of derivative, so that the curve is nowhere analytic.

This notion of a shift is analogous in many ways to that of a gradient in the Newtonian differential calculus. The shifts can be taken to form the basis of an extended calculus that allows symmetries in future paths to be analysed and predicted in goal-directed behaviour.

In particular, a Taylor series filter can be used that allows the presence, type and magnitude of non-analytic shifts away from the Taylor series prediction to be discovered.

The key properties of the shift curve are: its non-analyticity (introduced in section 3.5) – this allows the hypothesised plasticity and persistence of goal-directed behaviour to be modelled and measured; the presence of back-to-back shift copies at higher orders of derivative (first introduced in section 4.4.1 on curve symmetry) – these provide the non-analyticity everywhere along the shift through their recursive symmetry (see pp. 90-92); the presence of particular polynomial analytic projections from each point on the curve (introduced in section 4.4.2) – these are the analytic predictions of the Taylor series at each point; the presence of partial copied shifts after filtering (first

introduced in section 4.6.2) – these allow non-analyticity in the data to be revealed; the shift ratio (first introduced in section 4.6.4) – this allows a range of shifts to be modelled and measured.

Chapter 5

The Parallel Pipes Problem

5.1. Introduction

In the previous two chapters a computational framework and computing methodology were developed for use in the investigation of C^{∞} smooth goal-directed behaviour. Chapters 5 and 6 present an empirical study in which these novel techniques can be applied.

The experiments in these two chapters involve the study of human hand movement kinematics. They are aimed at investigating C^{∞} smooth goal-directed behaviour in such a way as to decide between the models introduced in the previous two chapters. In particular an attempt is made to isolate non-analytic components of C^{∞} smooth goal-directed behaviour for the first time.

Chapter 5 will look at two experiments involving the Parallel Pipes Problem, in which human subjects are required to use a computer mouse to move a ball within and between two parallel pipes displayed on a screen. The first experiment places no constraints on the agent, other than a time limit, whereas the second experiment constrains the agent to move smoothly in the undulatory sense as well. (For a discussion of smoothness see Section 3.4). Chapter 6 will then report experiments that aim to generalize the empirical base by introducing other height to width ratios for the movement between the two pipes, angled pipes and perturbations during the motion. [The raw data for all five experiments can be found on the CD attached to the thesis. See the 'read me' file on the CD for more details.]

5.1.1. Previous Work

Kinematics in general is the study of the motion of bodies without reference to mass or force (Borowski & Borwein, 1999, p.321). For an overview of specifically the kinematics of human motion see Zatsiorsky, 1998. For an overview of arm and hand movement kinematics see Schaal, 2002.

One aspect of kinematics concerns the kinematic geometry of human motion, i.e. body position, displacement (Zatsiorsky, 1998). This provides a description of the location and orientation of the human body in space. Soechting & Flanders (1992) look at frames of reference, vectors and coordinate systems for the human body that describe the control of eye and limb movements, for example. Engin (1980) studied the relative motion of the arm with respect to the torso, using a matrix based methodology.

Another aspect of kinematics looks at posture of the body (Zatsiorsky, 1998). This provides an account of the overall joint configuration of the body. Cappozzo et al. (1995) and Cappozzo et al. (1996) for example, investigate the position and orientation of bones during movement. Here, anatomic reference frames are used to provide a basis for standardization. Anatomic landmarks are used to define the anatomic frames, to help determine those landmarks that are to be used to define the anatomic frame, and to act as supplementary landmarks.

A further area of kinematics focuses on the geometry and kinematics of the individual joints themselves (Zatsiorsky, 1998). O'Connor & Zavatsky (1990), for example, investigated the kinematics (and mechanics) of the cruciate knee ligaments. Blankevoort, Huiskes & De Lange (1988), investigated knee joint motion, looking, for instance, at the torque-rotation curves of the knee joint under flexion. Such analysis can provide an account of the degrees of freedom and ranges of movement present in the joints and how these features are utilised.

Taken together, these three aspects provide a means for analysing the effect of the anatomical connection between the joints within states of the body.

This anatomical feature of body state is found to be controlled for sufficiently in the thesis by keeping to the same type of whole body posture, i.e. a sitting position holding a mouse with minor movement of the elbow and wrist. Consequently, it does not prove to be a major issue for analysis that affects the outcomes of the experiments presented here.

The analysis of goal-directed behaviour for the thesis, and in particular its plastic persistence, falls instead within the realm of what is known as differential kinematics, i.e. motion governed by derivatives (Zatsiorsky, 1998, Chapter 3). Plamondon et al. (1993) model the velocity profiles of rapid movements, for example. Plamondon & Alimi (1997) provide an overview of the speed versus accuracy trade-off found in rapid-aimed movements. This trade-off, which has been theorised about widely, will be of significance in the experiments described later.

An important part of differential kinematic analysis provides a means for relating motion at one joint to motion at another joint in the same *kinematic chain* (a serial anatomical linkage of rigid bodies).

Major aspects of such analysis include: the direct kinematic problem, in which end point velocity is searched for in terms of the known joint angular velocities; the inverse kinematic problem, where joint angular velocities are searched for given the end point velocity.

In summary, this methodology analyses the dynamical effect of anatomical connections along the kinematic chain. In terms of temporal causation, it provides an efficient causal account of anatomical dynamics. As such, it can provide a retrospective accounting for what has happened, or an account of what would happen if the agent moved in a particular way. The methodology though is not directly relevant to the prediction of the sequential shape of voluntary direction by the agent in pursuit of goals.

This is because neither a retrospective nor a what-if account provides a predictive account of whether an action will occur or what sort it will be. Reference to goals, on the other hand, allows predictive models to get a grip on voluntary direction. Moving a hand up and then down, for example, may be seen to be due to one or more of a number of goals being engaged to influence the direction of the action sequence. For example, the hand may move up and down to direct the winch of a crane. Alternatively, swatting a fly could involve one goal of moving the hand up to swat and a further goal of actually bringing the hand down on the fly. If the fly flies away, the goal is likely to change and with it the action sequence. Further examples of voluntary direction in the pursuit of goals will be seen in the hand movements analysed in the experiments in the thesis.

The thesis investigates strategies used by animate agents to control the shape of their motor actions in pursuing goals. This is currently a very active field of research with researchers keen to find out why the central nervous system prefers some movement patterns over others even though it is not directly using differential measures based on kinematics chains itself (Zatsiorsky, 1998, p.208). Berkenblit et al. (1986), for example, hypothesised simultaneous action of the joints but with independent, individually controlled signals to the joints. Gutman et al. (1990) attempted to explain the standardly observed bell-shaped velocity profiles in reaching movements to targets on the basis of the nonlinearity of subjective time.

It is a contention of the thesis that the directed shifts in voluntary movement can be fruitfully predicted by reference to both the putative goal states and the efficient

causal states as part of an abstract state sequence independent of the particular kinematic chains involved.

In particular, the thesis focuses on the search for predictive fits in time based differential kinematics settings. Consequently, it is the branch of differential kinematics concerned with abstract causal state sequences or trajectories in time, independent of kinematic chains, that will form the experimental framework for the thesis. The thesis seeks the predictive fits in hand movements (see Schaal, 2002 for a survey of arm and hand movements).

A previous study of particular relevance is that of Novak et al. (2000) who looked at multiple (6) agents in an experiment involving turning a knob from rest through a certain angular displacement and then coming to rest again at an LED target. The Novak study and other related studies involve time based differential kinematics settings, the results of which have been analysed from a purely efficient causal perspective using analytic methodology.

The thesis seeks to undertake a similar investigation to the Novak study, but from a final causal as well as an efficient causal perspective.

5.1.2. The Present Study

This chapter will look at the Parallel Pipes Problem without and with movement and undulatory smoothness constraints. The precise details of the constraints used will be provided in the Sections 5.2.2 and 5.3.2.

The experimental design is analogous to that of Novak et al. (2000) but the present study deals with vertical displacement versus time rather than angular displacement versus time. This is because the present study is concerned with the shape of the deviation from one horizontal path to a parallel horizontal path over time. Linear horizontal paths are focussed on because this leads the subject to have the most clearly identifiable derivative states at the inward ends of the pipes (i.e. zero derivatives at all orders in the vertical displacement versus height domain).

Smoothing of the zeroth order vertical displacement versus time profile is performed using a sliding boxcar filter. This filter involves taking a sample under investigation at a particular time point and an equal number of neighbours either side and averaging the vertical displacements for the samples to give the smoothed reading for the central sample. In the three-sample case the vertical displacement at the time point in question and the vertical displacements of one neighbouring point to either side of the sample are included in the average for the smoothed version of the central point. This is repeated for all points in the profile.

From this vertical displacement versus time profile, velocity, acceleration, jerk and snap profiles can be generated by numerical differentiation. The smoothing process in the zeroth order should be iterated just enough times, according to a pre-specified smoothness threshold, to allow smooth snap profiles to appear without removing the phenomena of interest.

The motivation for this smoothing process is that it reduces noise in the data whilst being causally neutral in that such smoothing is not biased to produce an analytic, piecewise analytic, non-analytic shift curve, or some other non-analytic curve outcome. It does not prejudice the Taylor series to be particularly obeyed or disobeyed at any point. It is also not an agent constraint and is not to be confused with undulatory smoothness agent constraints imposed later on.

Some of the statistical methods used for analysis of the kinematic profiles are similar to those used by Novak et al. (2000). These include the K (Nagasaki, 1989) and C (Soechting, 1984) parameters, which measure symmetry and shape of the velocity profile respectively (and which will be explained in detail in Section 5.2.3). Performance measures include trial success rate, movement time, peak velocity, and mean squared jerk. Improvement measures involve the use of t-tests on the various performance measures from beginning to end.

In the present study the results of Taylor series filtering are also presented. This is the most important test, for the purposes of the thesis, because it may provide indicators as to whether the path involved is analytic, piecewise analytic, the non-analytic shift curve or some other non-analytic curve.

The experiment without movement or undulatory smoothness constraints, Experiment 1, shows in what ways *the agent chooses* to override undulatory smoothness and the midway symmetry in the task, i.e. to override an s-shape. It will be argued later (in Section 5.2.4) that this overriding is caused by the agent acting upon perceived needs for overall speed and accuracy into the second pipe.

A minimum number of constraints (that initially focussed on movement but in later versions settled on undulatory smoothness) are then imposed to try and remove these (and only these) agent choices in Experiment 2 so that an underlying s-shape emerges. The constraint ultimately chosen is a local curvature constraint in the height-time domain. As with the smoothing method used, the constraints chosen here are intended to be causally neutral, i.e. to not bias the result in favour of any of the alternative curve types. Instead the bias is towards a temporally even shape (i.e. one that is free from sudden jerks and evenly paced) and hence temporally symmetric shape (using the movement constraint) and a low frequency shape (using an undulatory smoothness constraint). In short, the bias is towards an s-shape without being towards any particular type of s-shape.

The question then investigated is what type of symmetries (out of rotational, reflexive and recursive) and what type of s-shaped curve (out of analytic, piecewise analytic, or non-analytic shift curve) emerges from encouraging temporally even, low frequency shapes.

5.1.3. Simplicity, Inevitability, and Generalizability

Experiment 1 seeks to begin investigation of the basic unit in the empirical investigation of C^{∞} smooth goal-directed behaviour.

This unit should provide an account of a *simple*, smooth, symmetrical movement, making it as clear as possible to show how to choose between the analytic, piecewise analytic, non-analytic shift curve and other non-analytic curve models.

The general type of curve that the experiment with undulatory smoothness constraints produces should be simple but not so simple that a particular curve or model type of curve is *inevitable*. In fact there are many possible curves allowable as solutions within the constraints imposed, including analytic, piecewise analytic and the non-analytic shift curve. A more complex type of curve treated as a unitary curve would make it more difficult to choose between the above alternatives.

The type of curve that the experiment undulatory smoothness constraints produces should be one that is smooth in the undulatory sense – a low frequency shape. Less variation through low frequency simplifies and clarifies the underlying trends – which is a good approach for a base unit – whilst not forcing a particular model to be the only viable one. It also minimises any tendency to jerk.

The type of curve that the experiment with undulatory smoothness constraints produces may be one that maintains the alternating rotational and reflexive symmetries around midway, and the recursive symmetry discussed in Chapter 4, but none of these are forced on the agent. The agent is free to be asymmetrical in all three respects if this is their inherent way of behaving.

Experiment 2, with movement or undulatory smoothness constraints in its basic design, is intended as an improvement on Experiment 1 in attempting to evolve the basic unit in the empirical investigation of C^{∞} smooth goal-directed behaviour.

With this foundation stone in place future generalizations can then be made to understand what may occur in more *general* and complex settings. Experiments 3, 4 and 5 (to be presented in Chapter 6) seek to provide such generalisation in cases involving different height to width ratios between the pipes, different angles of second pipe and perturbations to the steered course, respectively.

5.2. Experiment 1: The Parallel Pipes Problem

5.2.1. Introduction

Experiment 1 involves the subjects moving a ball between two parallel pipes under time constraints (the latter feature imposed to encourage the subjects to focus on the experimental goal rather than develop and vary their own subgoals). There were no movement or smoothness constraints imposed upon them in this experiment. The experimental setup is shown in Figure 5.1, which represents the actual screen displayed to the subject at the start of each trial with arrows added to the figure to illustrate the task of the subject.

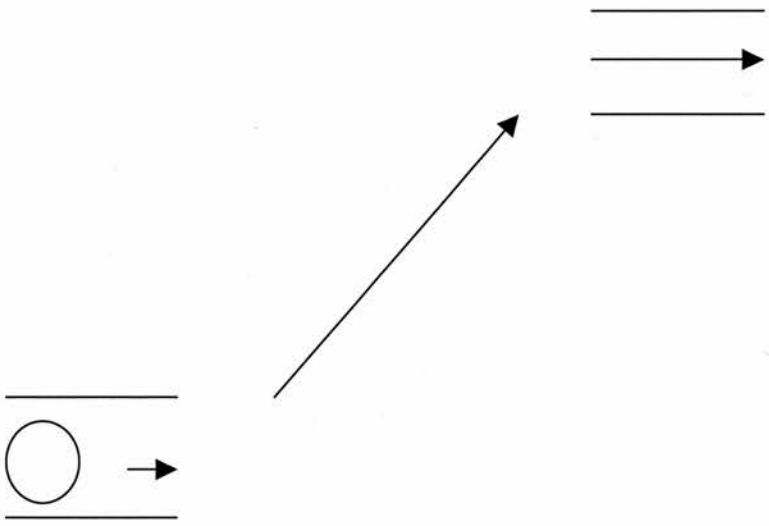


Figure 5.1. Experimental setup. The ball and the two pairs of parallel lines appear on screen. The arrows are added in this figure to illustrate the subject's task.

This setup has been chosen to be the simplest one possible to allow the phenomena of interest to be most clearly seen. Specifically, pipes are used because they lead the subject into a particular spatial path at the two ends of the behaviour. The parallelism of the pipes encourages zero vertical movement at the start and end points of the shift, i.e. the end of the first pipe and start of the second pipe. Potentially this allows for the simplest shift – a full height shift – to be isolated, the motivation for this being that, if confirmed, it could be used as a base unit for C^{∞} smooth goal-directed behaviour. A partial height shift from the end of the first pipe would involve entering the second pipe with a non-zero vertical velocity as well as a change in height and would require an angled second pipe for it to happen. The linearity and parallelism of the pipes encourages zero valued derivatives, in terms of vertical displacement with respect to time, at all derivative orders at the start and end points of the shift (apart from the

zeroth order at the entrance to the second pipe). A ball is used because moving a ball within the pipes was considered to be an intuitive task for the subjects. Ball movement is under the agent's control in both height and width directions so that the agent's hand movement matches the movement of the ball on the screen. This means that the path of the ball provides a copy of the state space trajectory for the subject's hand and goal and also makes the task more intuitive for the subject.

Kinematic analysis using existing methods in the literature is performed to capture aspects of performance and improvement and allow comparisons to be made with other similar studies. The more novel Taylor series filtering is then used to decide between the analytic, piecewise analytic, non-analytic shift curve and other non-analytic alternatives. The shift curve values used for the analysis can be found in Appendix B.

The working hypothesis is that the shape produced could be an underlying non-analytic shift curve similar to that shown in Figure 4.9 that gives rise to the derivative profiles shown in subsequent figures. Suppose first of all that the filter includes sufficient terms to represent the full Taylor series reasonably accurately in its analytic projection. If the working hypothesis is true then a significant non-zero residual profile should result from Taylor series filtering that can be closely and predictively fit by sections of the non-analytic shift curve. If some other non-analytic curve is present then a non-zero residual unlike the shift curve should be present. If, on the other hand, the shape is analytic then zero or (in practice) near-zero residuals should be observed at all points tested. For piecewise analytic shapes, intervals between a pair of consecutive joins should return zero or near-zero residuals and intervals

containing joins should return significant non-zero residuals because the joins are significant points of non-analyticity. If the filter contains insufficient terms to represent the full Taylor series reasonably accurately then a non-zero residual unlike the shift curve should be present.

5.2.2. Method

The experimental setup (as shown in Figure 5.1) involves two parallel pipes and a ball. The two pipes are both 32 pixels (roughly 90mm) vertically by 64 pixels (roughly 180mm) horizontally (using 1152x864 screen resolution). The ball is 22 pixels (roughly 60mm) in diameter. The separation of the pipes is 256x256 pixels (roughly 720x720mm) between the central exit point of the first pipe and the central entrance point to the second pipe.

Six subjects participated voluntarily in the study. They were all healthy, normal adults and were aged between 22 and 50 years old. Five were male and one was female and all were right-handed. So age, sex, and handedness as well as factors such as degrees of freedom of the limbs, hand size and arm length were all potential sources of variations in this (and the later) experiments. It would be of great interest if, despite this variation, similar strategies were employed by the different agents in their kinematic profiles and common shapes and curve types seen.

The apparatus included a desktop computer with 700 MHz Pentium III processor, 384MB RAM and running Windows XP Pro. A Logitech light mouse was used with pointer speed set at 7 (of 11), the PS/2 report rate set at the highest level and with

mouse acceleration, smart move and cursor trails all switched off. A 17-inch Belinea CRT monitor set at 1152x864 pixel resolution and a screen refresh rate of 75Hz was used. A large mouse mat was used to avoid the mouse running over the edge during a trial. A workbench and chair set at a height for performing the task with the arm kept horizontal were also used. In general this equipment and these settings were used because they were found to be the best currently available for allowing the agent to perform the task as they wanted to whilst as much as possible avoiding the influence of extraneous factors such as unsmoothness, uncontrolled perturbation and fatigue.

The subject is seated at the workbench with arm resting horizontally on the workbench with their eyes roughly 25cm from the screen. The first thing they see is a display like that in Figure 5.1 (without the arrows). They are then instructed that they are to move the ball through the first pipe, up to the second pipe and then through the second pipe. To do this they are required to move the mouse pointer on to the ball and left click to 'pick the ball up' and hold the left mouse button down through the whole of the behaviour to continue 'holding the ball'. This is to encourage a regular movement suited to the agent. To show that they are over the ball and can pick it up the mouse pointer changes from an arrow to a crosshair. If they release the left mouse button during the course of the trial before completion then they fail the task instantly, with the ball disappearing (and a fail message appearing). This method is used to keep the hand gripping the mouse in the same way throughout the motion thereby reducing the degrees of freedom. They are told to do the movement as quickly and accurately as they can and that hitting the pipe walls or timing out or releasing the ball will result in failure of that particular trial. They are also told that there is no time limit on the

overall task, a run of 100 trials, and that they are allowed to reposition the mouse on the pad as they wish between trials.

They are then allowed a small number of practice trials to orient themselves to the task requirements. They then perform 4 runs. Dropping the ball does not count as a completed trial and this trial is not used in the analysis, although hitting the pipe and timing out do count as completed trials. This is because the latter conditions still produce complete and analysable shapes (although these are not used in the tables of statistics shown later). Subjects are allowed a break of up to a few minutes between each run so as to minimise the effects of fatigue.

After each trial, success or fail (together with any reason for failure) is shown on screen to reinforce in the subject what is required of him or her. For success the message 'Congratulations!!!' appears. Failure messages include: 'Fail. You dropped the ball!', 'Fail. You hit a pipe!', and 'Fail. Too slow!' (N.B. the message boxes have to be closed by the subject and doing so brings the next ball on to the screen. This gives the subject some leeway to go at their own pace during the experiment, again reducing fatigue effects). These messages serve as rewards or punishments in the process of reinforcement learning leading to a tendency towards producing the shape that best allows success and thus improving success over runs.

5.2.3. Results

Preparation of the results first required smoothing. This is because the raw data inevitably contains jaggedness that obscures the smooth underlying path taken by the

hand. Figure 5.2 shows a single trial in the height-time profile and the first derivative generated on the basis of this profile. Smoothness is significantly lacking much more in the first derivative than the zeroth but the observed jaggedness is different across trials and does not, therefore, represent a significant part of any common strategy of the subjects across trials. Also, the analysis of low frequency s-shapes may treat such jaggedness as high frequency noise.

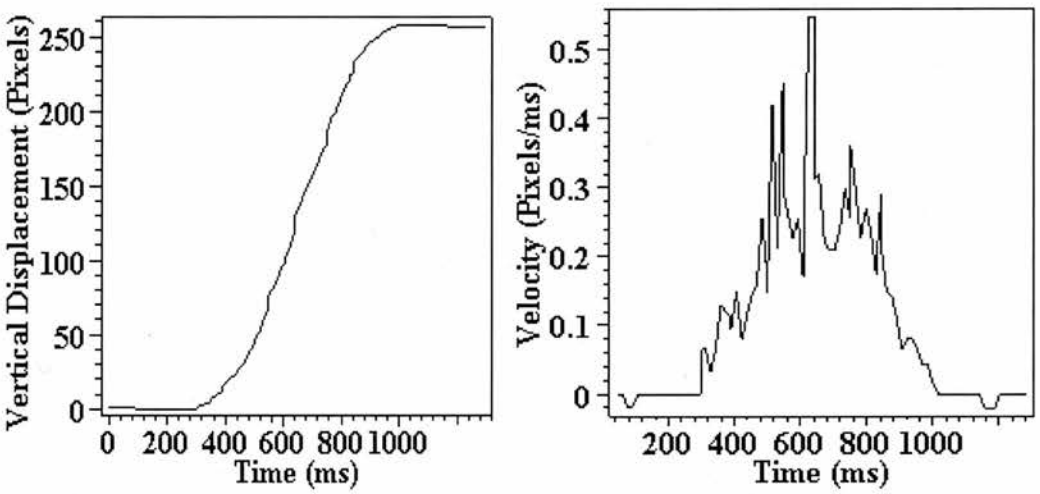


Figure 5.2. a) Raw data in the vertical displacement versus time domain. b) Velocity profile generated from the raw data without smoothing.

Iterative smoothing of the height-time profile using the three sample sliding boxcar filter described in Section 5.1.2 was performed to reduce noise and the next four derivatives (velocity, acceleration, jerk and snap) were then generated.

Standard kinematic analysis was performed on the smoothed profiles (using the methods first introduced in 5.1.2). As in the Novak (2000) study, sometimes shapes for single trials were found that were performed with a single smooth movement whilst in other trials one or more secondary submovements overlapped a primary

movement. These secondary submovements are still relatively low frequency shapes and are thus not smoothed out by the smoothing process (in individual trials) unlike the high frequency shapes. Figure 5.3 (a-e) provides an example of a single smooth movement without any noticeable secondary submovements present. Shown are the vertical displacement, velocity, acceleration, jerk and snap versus time for the central section of the motion, i.e. between the two pipes.

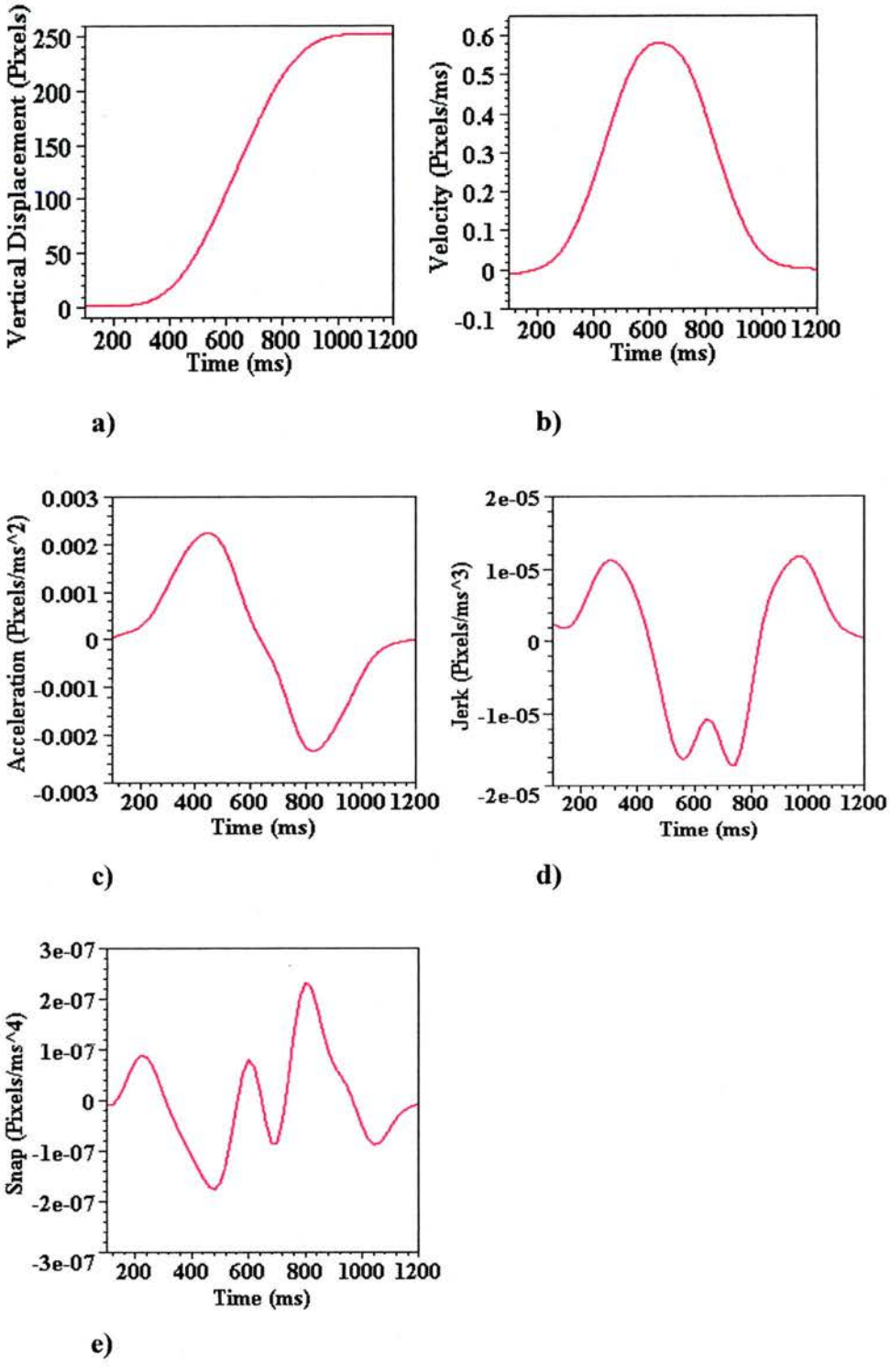
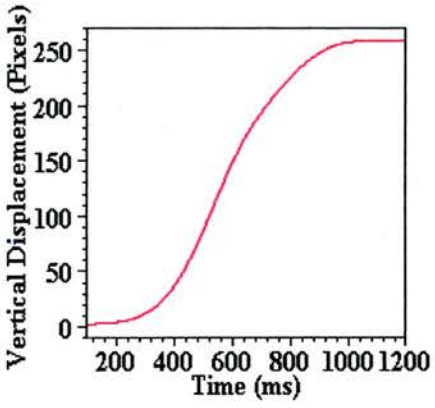
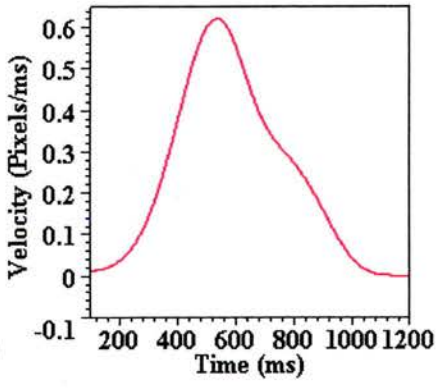


Figure 5.3 (a-e). The kinematic profiles for a single trial without any secondary submovements for vertical displacement, velocity, acceleration, jerk and snap.

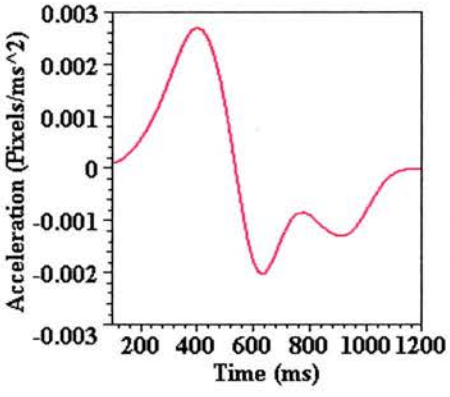
Figure 5.4 (a-e) shows a single trial where a secondary submovement overlaps the primary movement. It is most noticeable as the cause of the bulge in the second half of the velocity profile and the subsequent asymmetries in the higher orders.



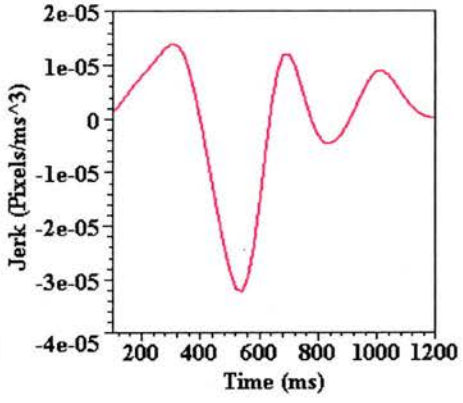
a)



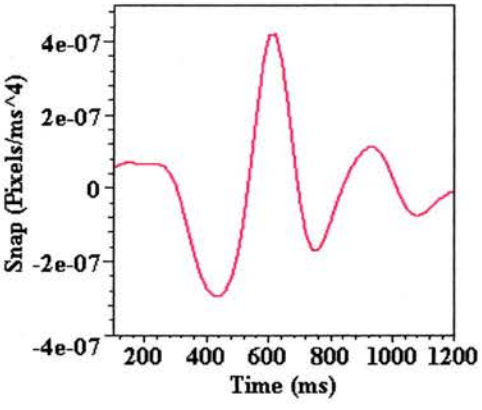
b)



c)



d)



e)

Figure 5.4 (a-e). Single trial kinematic profiles with one secondary submovement for vertical displacement, velocity, acceleration, jerk and snap.

To account conceptually for profiles that include secondary submovements the working hypothesis of a simple shift may be modified to include several shifts overlapping each other. Figure 5.5, for example, shows two components of a possible combined shift for a task of width one and height two. The component S1 is a full height shift (i.e. one with a one to one height to width ratio) whereas component S2 is completed over half the horizontal interval and thus has a two to one height to width ratio. In Figure 5.6 the displacement versus time profile for the combined shift is shown (C). This is the addition of the two components in Figure 5.5, S1 and S2. Figure 5.7 shows the resulting velocity versus time profile (VC) for the combined shift. The example shown here demonstrates that a midway asymmetrical curve can be simulated using shift combinations that are, in their shift components (over the interval where the shift takes place) midway symmetrical. (N.B. this section is aimed only at describing a general strategy that could be used for developing hypotheses about non-analytic curve shapes using shift combinations rather than attempting to closely model the secondary submovement phenomenon).

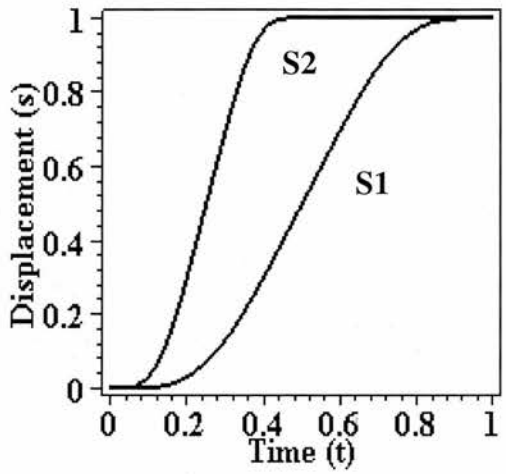


Figure 5.5. The two components of a possible combined shift. S1 is a full height shift component with a height to width ratio of 1:1. S2 is a halved height shift component with a height to width ratio of 2:1 (N.B. the line S2 continues with constant value 1 from t = 1/2 for the remainder of the interval as no further height change due to this component is occurring).

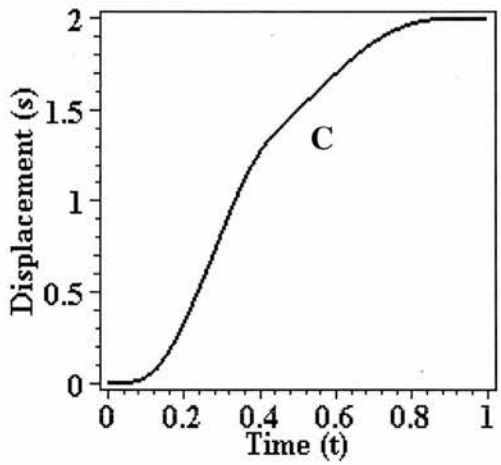


Figure 5.6. The combined height shift (C) resulting from the addition of the two shift components (S1 and S2) shown in Figure 7.1.

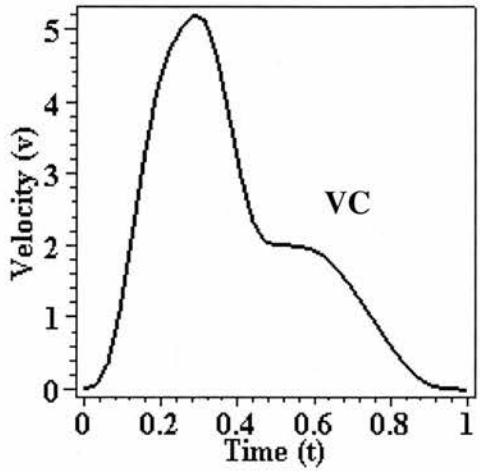


Figure 5.7. The velocity profile (VC) for the combined height shift shown in Figure 5.6.

After smoothing, the next stage in the analysis involved superimposing the profiles between the pipes for all trials on the same unit interval in the height and time domain. This was performed on the data to line up trials taking different times over the same canonical space. This was to allow meaningful superposition of shapes without bias towards any particular one of the alternative models, in order to extract a common strategy up to horizontal scaling. What is of interest in the analysis here is the shape underlying the agent behaviour in the height direction versus time rather than actual timings and heights. For more on scaling derivative profiles, see: Carter & Shapiro (1984); Gielen et al. (1985).

After placing all of the profiles on this new one by one space, averaging of the data could then take place for each run of 100 trials for each agent (where each agent performed 4 runs). These averages were then smoothed in the same manner as the individual trials had been earlier. Particularly of interest was the averaging and

smoothing of the success cases for each run only, because this was where the shape across individual trials was most stable. The average values between trials 75 and 100 varied at most by 5%, for example. The analysis of runs allows the statistical and kinematic properties to be shown that the subjects tended towards in developing a more successful strategy. It is these properties that are required for our isolation of the hypothesised base unit for C^{∞} smooth goal-directed behaviour, i.e. the non-analytic shift curve, because random high frequencies in the fail cases may mask the base unit shape. The fail cases may, for example, have non-zero derivatives at the end of the first pipe or non-zero (beyond the zeroth order) derivatives at the start of the second pipe. In the results presented below only the success cases are included in the measures.

Table 5.1 shows the results of Experiment 1 for percentage success, movement time, peak velocity and mean jerk squared, where the latter three refer to the averages for the success cases only. Significant changes in a metric between Run 1 and Run 4 at the probability, $p < 0.05$ significance level are represented by an underline.

Subject	Run	Trials	Percentage success	Movement time	Peak velocity	Mean Jerk^2
1.1	1	100	60	1341.95	2.325	1928.05
	4	100	62	1275.74	2.417	2402.73
1.2	1	100	42	2141.95	3.830	26403.54
	4	100	27	1934.70	3.903	26238.44
1.3.	1	100	70	1752.50	3.226	11063.14
	4	100	72	1620.82	3.405	13256.14
1.4	1	100	59	1676.92	3.204	14005.69
	4	100	74	1628.26	3.532	16122.47
1.5	1	100	63	1794.78	2.414	4522.29
	4	100	70	1897.50	2.552	4417.12
1.6	1	100	42	2041.21	2.545	4301.50
	4	100	70	1907.01	2.958	6842.77
Average	1	600	56	1791.47	2.924	10370.70
	2	600	59	1764.89	3.067	11771.57
	3	600	61.3	1762.39	3.154	12699.43
	4	600	62.5	1710.67	3.128	11546.61
All		2400	59.7	1757.35	3.068	11597.08

Table 5.1. Table of standard statistics (for movement time, peak velocity and mean jerk squared only the success cases within that run are used). N.B. Underlines represent significant changes in a metric at the 0.05 significance level.

The percentage success for each of the six subjects was intermediate thereby avoiding floor or ceiling effects. This is also important because had the task been too easy the subjects would have been able to develop and vary subgoals and any underlying shape that may be there would have been unlikely to emerge. If the task had been too

difficult then a shape may have been forced on the subjects by the constraints of the task itself.

Overall percentage success averaged over subjects for each run increased monotonically over runs whilst movement time averaged over subjects for each run decreased monotonically over runs, which indicates improved task ability.

Peak velocity was slightly higher in all six subjects in Run 4 than in Run 1, and significantly higher for 5 of the 6 subjects, thus supporting a hypothesis that peak velocity tends to increase over runs. The peak velocity value was typically around the 3 mark, and therefore around 50% higher than the peak velocity of 2 predicted by the non-analytic shift curve model.

Jerk is the rate of change of acceleration with respect to time. As Zatsiorsky (1998, p.207) describes its general significance: 'In general, the change in acceleration is associated with the changes in the magnitude of the acting forces, and thus, the changes in stresses within the moving body. In engineering applications, the jerk is registered when the propagation of the deformation waves is of interest. When movement is performed at only one joint, the jerk is associated with the rate of muscle force development. In biomechanics and motor control, there is one additional stimulus for studying jerk: it has been suggested that a skilled performance is characterized by a decrease in jerk magnitude. According to the minimum-jerk hypothesis, in skilled people the arm moves in a maximally smooth way.'

No particular pattern was apparent for mean squared jerk with the average across agents for runs and altogether around six and a half times higher than the 1774.3 value predicted by the non-analytic shift curve model. This is perhaps indicative of the fact that secondary submovements are adding extra jerk peaks and troughs to some of the movements and these effects are not being totally removed in the smoothing and averaging processes.

The results of further statistical measures are shown in Table 5.2 below. These include LMS symmetry, K, C and the modified K statistics for the acceleration and deceleration phases of the motion. The results shown allow comparison of the first and last runs for each subject.

Subject	Run	Trials	LMS symmetry	K	C	K (acceleration peak)	K (acceleration trough)
1.1	1	100	0.901	0.434	3.637	0.579	0.332
	4	100	1.277	0.436	3.781	0.615	0.314
1.2	1	100	3.999	0.352	5.986	0.732	0.132
	4	100	5.515	0.328	6.100	0.447	0.149
1.3.	1	100	3.786	0.327	5.044	0.665	0.154
	4	100	4.661	0.318	5.322	0.677	0.147
1.4	1	100	4.151	0.302	5.014	0.673	0.163
	4	100	3.805	0.339	5.521	0.885	0.157
1.5	1	100	1.507	0.325	3.780	0.613	0.204
	4	100	1.839	0.353	3.994	0.691	0.223
1.6	1	100	2.901	0.344	3.987	0.558	0.225
	4	100	5.446	0.295	4.634	0.931	0.186
Average	1	600	2.874	0.348	4.575	0.637	0.202
	2	600	3.032	0.353	4.783	0.665	0.196
	3	600	3.653	0.338	4.933	0.687	0.186
	4	600	3.758	0.345	4.892	0.708	0.196
All		2400	3.329	0.346	4.796	0.674	0.195

Table 5.2. Tables of LMS symmetry, K, modified K and C statistics for Runs 1 and 4 for each subject.

A least mean squared (LMS) measure of midway symmetry was used that compares pairs of points at equal distances of the two sides of the midway line of the vertical displacement versus time profile. The results show the LMS symmetry value averaged over subjects for each run increased by over 30% from Run 1 to Run 4, implying a move away from midway rotational symmetry in the zeroth order (and thus away from midway symmetries in higher orders of derivative as well). In all cases it was noticeably above the zero LMS symmetry value predicted by the shift curve model.

The K statistic is a second symmetry measure based on the relative time to peak velocity. This follows Nagasaki (1989) who used this statistic in the analysis of velocity and acceleration asymmetries in human arm movements. $K = 0.5$ refers to a symmetric profile, i.e. the velocity peak appears at the halfway time point in the motion. $K < 0.5$ shows a shorter acceleration period than deceleration period, i.e. the velocity peak appears before the halfway time point in the motion. $K > 0.5$ shows a longer acceleration period than deceleration period, i.e. the velocity peak appears after the halfway time point in the motion.

The results (fifth column from the left) show that K was below 0.5 for all runs of all subjects and thus for all subjects over each run and altogether as well. This implies that a shorter acceleration period than deceleration period is typically used (although analysis of individual trial profiles show that there are atypical cases where the acceleration phase is longer than the deceleration phase). The overall average K measure of 0.346 was considerably below the predicted K value of 0.5 – the value if midway symmetrical movement had been used.

The C statistic is a measure of the squatness of the velocity profile regardless of the amplitude or duration of the motion. This follows Soechting (1984), who used it in the analysis of the effects of target size on the spatial and temporal characteristics of human pointing movements. It is the ratio of peak velocity to the average velocity of the motion, as shown in Equation 5.1 below:

5.1. $C = V(Peak) / V(Ave)$

where V(Peak) is the velocity peak value and V(Ave) is the average velocity value.

For all subjects the C value was higher in Run 4 than in Run 1, implying that the shape of the profile is becoming less squat. In all cases the C value was substantially higher than the C value of 3 predicted by the shift curve model.

A Modified K statistic was also used to examine the midway symmetry present during the acceleration phase of the motion and the deceleration phase of the motion separately. The unmodified version looked over the whole interval for midway symmetry in the velocity profile. The modified version looks first at the acceleration phase as a unit interval in its own right and measures the relative time to peak acceleration during this phase. It then looks at the deceleration phase as a unit interval in its own right and looks at the relative time to peak deceleration during this phase. In each case a value of Modified K = 0.5 would imply midway symmetry within the interval in question. A value of Modified K < 0.5 would imply the acceleration peak or trough being arrived at early (left leaning asymmetry) and Modified K > 0.5 would imply the acceleration peak or trough being arrived at late (right leaning asymmetry).

The symmetric shift curve model predicts that the acceleration peak will appear halfway into the acceleration phase (Modified K = 0.5) and that the deceleration peak will appear halfway into the deceleration phase (Modified K = 0.5). In fact the acceleration peak starts above 0.5 and increases monotonically over runs (see Table 5.2, second column from the right). The deceleration peak, on the other hand, stays at around the Modified K = 0.2 mark throughout (see Table 5.2, first column from the right).

Taylor series filtering (as described in Chapter 4), using all derivative terms up to and including the fourth order, snap, was also performed in an attempt to decide between the analytic, piecewise analytic or non-analytic shift curve alternative models (see Figure 5.8 for an example of a Taylor series prediction from the halfway point of one run for one subject – averaged over the successful trials).

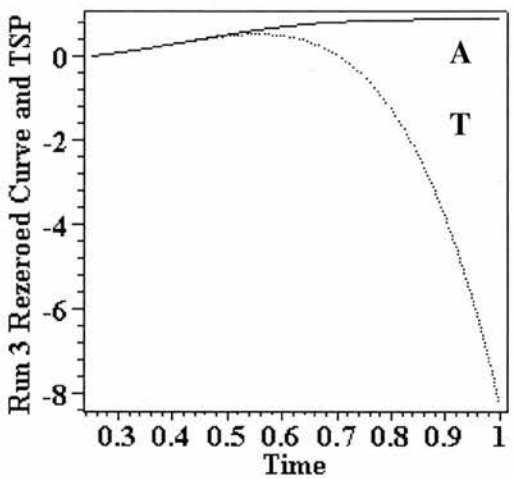


Figure 5.8. An example of a Taylor series prediction (dotted line, T) from the halfway point of the Actual curve (solid line, A) for a single run of a single agent averaged over successful trials.

The results of Taylor series filtering demonstrated significant non-zero residuals at all points tested (see Figure 5.9 for one such example. N.B. the circles shown in this figure and comparable later figures 5.11, 5.12 and 5.13 represent points on the fit curve and are shown like this rather than as a line because of the high degree of overlap between the fit curve and the residual). This means that the presence of the non-analytic shift curve cannot be ruled out (it would have been ruled out had a zero residual been found). However, the significant midway asymmetry in the profiles means that it is not yet appropriate to test conclusively for the presence of the hypothesised base unit.

Be that as it may, Figure 5.9 shows a case where a reasonable, though not exact, fit to a residual can be obtained using a *scaled* and *partial* shift curve section and this is typically the case, with similar results appearing in 63 of 72 points tested. The type, magnitude and sign of the shifts found were variable across agents although less variable within subjects across runs.

The fact that the fitted curves are partial and scaled is important, since these features run counter to the working hypothesis has no scaling within its description. For Figure 5.9, for example, the working hypothesis suggests that the residual should be 32/64 of the shift curve over the half-interval rather than the 15/64 shown. The shift fit to a significant non-zero residual is nevertheless interesting.

At this stage though, enhancing the midway symmetry as a next step may see residuals being produced that are more in line with the hypothesis that a full vertical displacement shift will be used by the subjects.

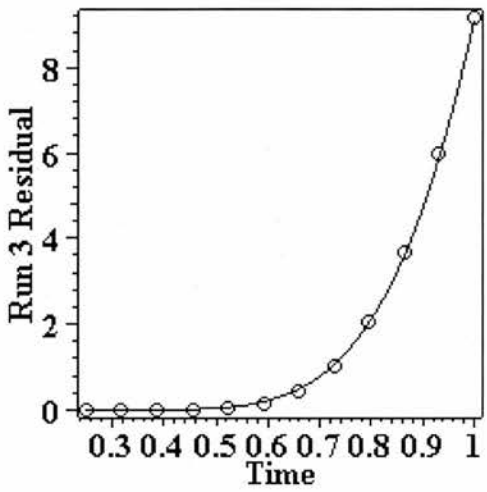


Figure 5.9. The resulting residual (solid lines) when the Taylor series prediction in Figure 5.8 is subtracted from the actual curve in Figure 5.8. A fit using a negative version of the first 15/64 of the shift curve with negative scaling is also shown (circles representing points on the full fit curve).

5.2.4. Discussion

Standard kinematic analysis showed significant improvements in task ability on metrics such as percentage success and performance, with the former monotonically increasing overall and the latter monotonically decreasing overall. The hypothesis that peak velocity would increase over runs (Novak et al., 2000) was supported.

Further significant statistical analysis looked at symmetry. It was found that there was an increasing midway asymmetry in the velocity domain (K statistic). There were also midway asymmetries in the acceleration and deceleration phases of the motion (modified K statistic). The presence of the midway asymmetries are thought to be due to the influence of the accuracy requirement for entry into the second pipe. Subjects may be setting their own goal of pulling up quickly and flattening out level with the second pipe for easy entry (this accords with the reports of some subjects). This is indicative of there being a trade-off between speed and accuracy. For more on this type of trade-off see Plamondon & Alimi (1997). For more on the associated issue of control of approach see Lee et al. (1992) and the Tau hypothesis.

In the crucial test to try and decide between the analytic, piecewise analytic and non-analytic shift curve models it was found that non-zero residual curves that could be reasonably fitted by partial scaled shift curve sections resulted when Taylor series filtering (as described in Chapter 4) was performed. This means that a non-analytic shift curve component may be required in the explanation of the behaviour – zero residuals would have refuted this possibility and non-zero but non-shift curve residuals would have entailed other explanations. The asymmetries in the profiles make it inappropriate for the hypothesised base unit to be decided upon just yet. The addition of constraints that allow midway symmetry to emerge more strongly may lead to something more like a full vertical displacement shift.

5.3. Experiment 2: The Parallel Pipes Problem Under Undulatory Smoothness Constraints

5.3.1. Introduction

The extra work to try to isolate a symmetric shift first involved encouraging temporal evenness through the interval, i.e. even speed, so that general midway symmetry is encouraged without prejudicing which specific curve symmetries are shown.

In early attempts at Experiment 2 movement constraints were added in an attempt to induce midway symmetry without forcing any particular type of midway symmetrical curve.

Several attempts at experiments with such constraints had mixed success. Limiting the horizontal speed or acceleration or requiring the agent to move within a general speed limit (including both horizontal and vertical components of the motion) all seemed to enable the agent to achieve greater midway symmetry than in the unconstrained case. However, this was still limited in the degree of overall symmetry obtained and was at the price of introducing artificially distorted profiles and tasks that were difficult for the subjects to succeed at. This could well have been because they were more preoccupied with finding the best strategy to obey the extra constraints rather than concentrating on the more fundamental feature for the purposes of isolating the basic unit – to complete the task quickly and smoothly.

In the later version of Experiment 2, the results of which are presented in this chapter, it was found that the movement constraints described above could be replaced by a more effective undulatory smoothness constraint. This involved limiting the localised curvature of the motion according to the curvature equation below, Equation 5.1.

5.1. $f(x) = y'' / (1 + (y')^2)^{3/2}$

As Borowski and Borwein (1999) define it, curvature is 'the rate of inclination of the tangent to a curve relative to the length of arc.' In Experiment 2, x is time, y is the vertical distance moved by the hand on the two-dimensional plane of the mouse mat (and thus by the ball on the two-dimensional screen) over time. Therefore y' is the velocity of the hand/ball and y'' is the acceleration of the hand/ball.

During an experimental run this value was calculated at each point by plugging the present values of hand/ball velocity and acceleration into Equation 5.1. The threshold set for this value to return a fail was determined through pre-experimental investigation such that the adherence to this constraint would not diminish the amount of success cases too much whilst still encouraging smoother performance. The result was to encourage undulatory smoothness. In turn, the greater undulatory smoothness may allow greater symmetry to emerge as the asymmetric rush to prepare for slower careful entry into the second pipe is literally smoothed out by the subject. This constraint also does not force a particular type of curve in terms of the four possibilities, i.e. analytic, piecewise analytic, non-analytic shift curve and general non-analytic, and does not force any greater degree of symmetry on the agent.

5.3.2. Method

The experimental setup was the same as that for Experiment 1. Six subjects participated voluntarily in the study. They were all healthy, normal adults and were aged between 19 and 50 years old. Four were male and two were female and all were right-handed. The apparatus used was also the same as Experiment 1.

The procedure was the same as Experiment 1 except that the subjects were also told that to succeed they must also perform the task smoothly over time. If the subjects failed on the basis of smoothness the ball disappeared and the message 'Fail. Not smooth enough!' appeared in a message box. During the short practice session prior to the experimental run subjects were allowed to get a feel for what was required of them in the sense of undulatory smoothness as well as in terms of the pipe and time limit constraints.

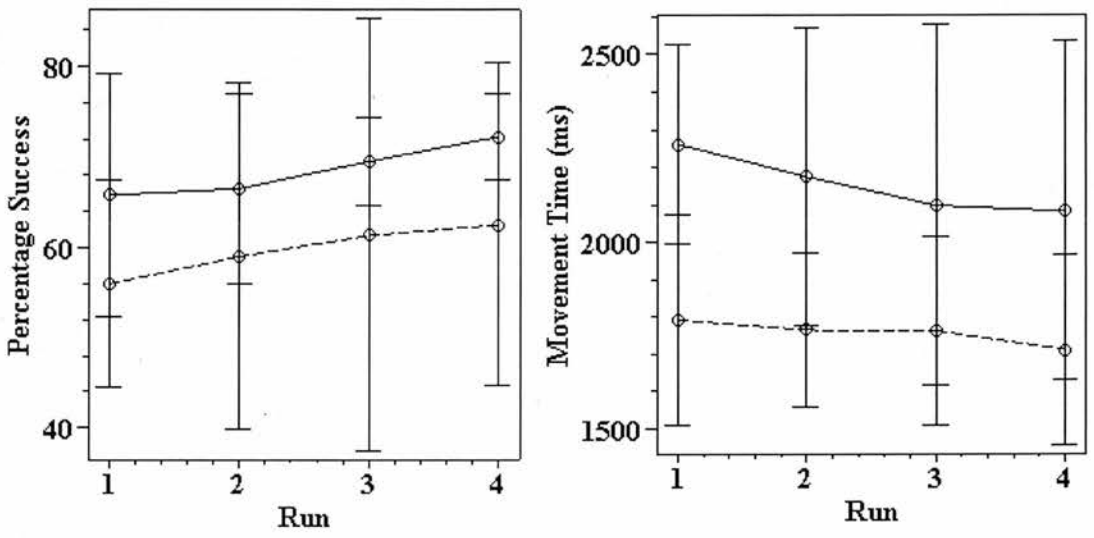
5.3.3. Results

Table 5.3, below, shows the standard statistics for the Experiment 2. Again, significant changes in a metric from beginning to end are indicated with an underline (with $p < 0.05$).

Subject	Run	Trials	Percentage success	Movement time	Peak velocity	Mean jerk ^{2}
2.1	1	100	68	2313.33	2.266	1825.45
	4	100	77	2089.90	2.011	1181.21
2.2	1	100	90	1778.33	2.355	2553.71
	4	100	72	1257.24	2.899	6850.23
2.3.	1	100	58	2228.14	2.208	2359.68
	4	100	68	2347.15	2.644	4549.19
2.4	1	100	57	2481.67	2.597	4429.32
	4	100	75	2433.69	2.537	3370.00
2.5	1	100	69	2247.33	2.471	4118.27
	4	100	76	2438.62	2.898	6302.03
2.6	1	100	53	2518.47	2.494	4454.39
	4	100	65	1924.55	2.535	4510.92
Average	1	400	65.8	2261.21	2.397	3290.14
	2	400	66.5	2174.50	2.530	4069.28
	3	400	69.5	2097.71	2.633	4682.07
	4	400	72.2	2081.86	2.587	4460.60
All		2400	68.5	2153.82	2.537	4125.52

Table 5.3. Table of standard statistics (for movement time, peak velocity and mean jerk squared, only the success cases within that run are used). N.B. Underlines represent significant changes in a metric at the 0.05 significance level.

In general, percentage success was much higher than in Experiment 1 whilst movement time was slower as a result of the undulatory smoothness constraint making subjects behave more carefully (see Figure 5.10 below). As in Experiment 1 the percentage success increased monotonically over runs whilst the movement time also decreased monotonically over runs.



a) b)

Figure 5.10. a) Graph of percentage success over runs comparing Experiment 2 (solid line) and Experiment 1 (dashed line). b) Graph of movement time over runs comparing Experiment 2 (solid line) and Experiment 1 (dashed line).

Peak velocity increased significantly for three of the six subjects although it decreased significantly for one subject. Overall peak velocity was noticeably lower in Experiment 2 than Experiment 1. The lower peak velocity in Experiment 2 can be seen as a direct effect of the introduction of the undulatory smoothness constraint.

The mean jerk squared showed no tendency towards increasing or decreasing averaged over runs. Also, the overall mean jerk squared was far lower for each run and for all runs together in Experiment 2 than in Experiment 1. This may be indicative of the lowering of acceleration changes resulting from the enforcement of the undulatory smoothness constraint.

As in Experiment 1, a further bank of statistical tests was used to look at issues of symmetry and shape by comparing the first and last runs for each subject (Table 5.4, below).

Subject	Run	Trials	LMS symmetry	K	C	K (acceleration peak)	K (acceleration trough)
2.1	1	100	0.939	0.421	3.544	0.530	0.294
	4	100	0.567	0.440	3.149	0.576	0.375
2.2	1	100	1.189	0.384	3.685	0.542	0.240
	4	100	1.751	0.370	4.535	0.534	0.190
2.3.	1	100	2.586	0.317	3.466	0.478	0.233
	4	100	3.615	0.331	4.143	0.667	0.193
2.4	1	100	2.919	0.333	4.062	0.558	0.260
	4	100	2.133	0.369	3.969	0.722	0.238
2.5	1	100	3.087	0.336	3.874	0.517	0.236
	4	100	5.311	0.301	4.540	0.612	0.197
2.6	1	100	2.241	0.341	3.903	0.576	0.261
	4	100	1.899	0.348	3.969	0.762	0.229
Average	1	600	2.160	0.355	3.756	0.534	0.254
	2	600	2.306	0.361	3.961	0.585	0.242
	3	600	2.560	0.358	4.123	0.631	0.238
	4	600	2.546	0.360	4.051	0.646	0.237
All		2400	2.403	0.359	3.973	0.599	0.243

Table 5.4. Tables of LMS symmetry, K, modified K and C statistics for runs 1 and 4 for each subject.

The LMS midway symmetry measure was generally a lot lower than in Experiment 1. In four out of six cases this measure increased significantly over runs though implying a move away from midway symmetry over runs (although a significant decrease can be seen in the case of Subject 2.1).

The K value was slightly higher in each run than in the comparable run in Experiment 1, and overall, although it still fell some way short of the predicted 0.5. The K value also stayed fairly constant over runs.

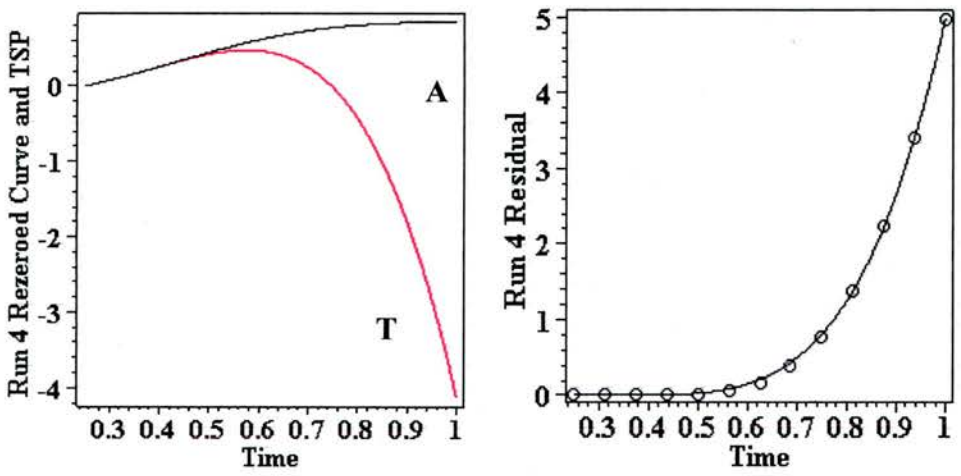
The C value was much lower in each run than in the comparable run in Experiment 1, with the squatter profiles resulting from the lowering of the peak velocity.

The Modified K statistic showed noticeably more symmetric profiles than in Experiment 1. The average K value for the acceleration peak is much closer to symmetric than in Experiment 1 for each run. The K value starts at around 0.5 in Run 1 and increases monotonically to around 0.6 in Run 4, implying a move away from symmetry, however. The average K value for the acceleration trough stays fairly constant, as in Experiment 1, but at a more symmetric value.

As in Experiment 1, Taylor series filtering, using all derivative terms up to and including snap (as described in Chapter 4) was performed in an effort to decide between the alternative models.

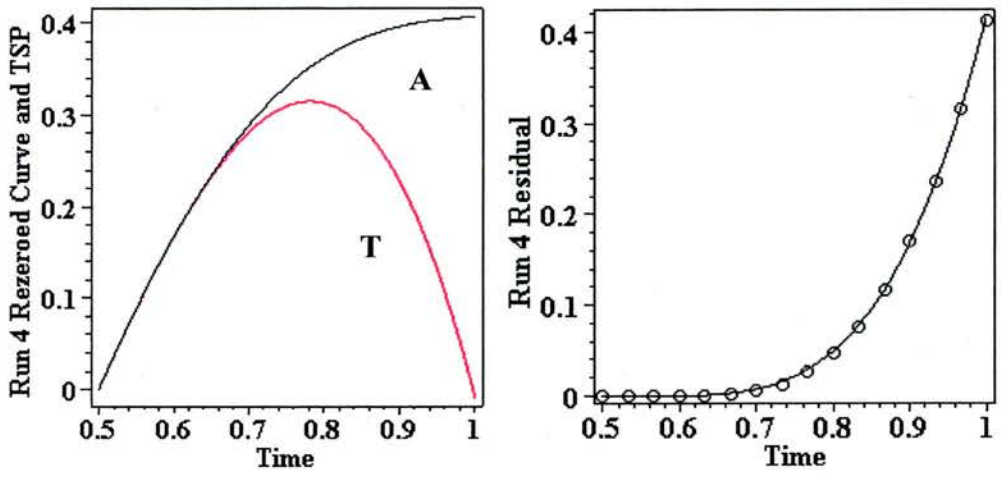
Although the symmetry is improved, the residuals were found to clearly not be full shifts. It is important to note at this point that, due to this fact, it will be partial scaled shifts that are tried for fit in the following analysis and found relative to the full 1:1 height shift.

Figures 5.11, 5.12 and 5.13, below, show examples of Taylor series predictions from the quarterway, halfway and three-quarterway, respectively, alongside partial scaled shift curve section fits to the residuals that result from Taylor series filtering. Each of these graphs is for Run 4 of Subject 2.1. For a worked example that lays out the calculations (using the data from Run 4 of Subject 2.1) see Appendix C.



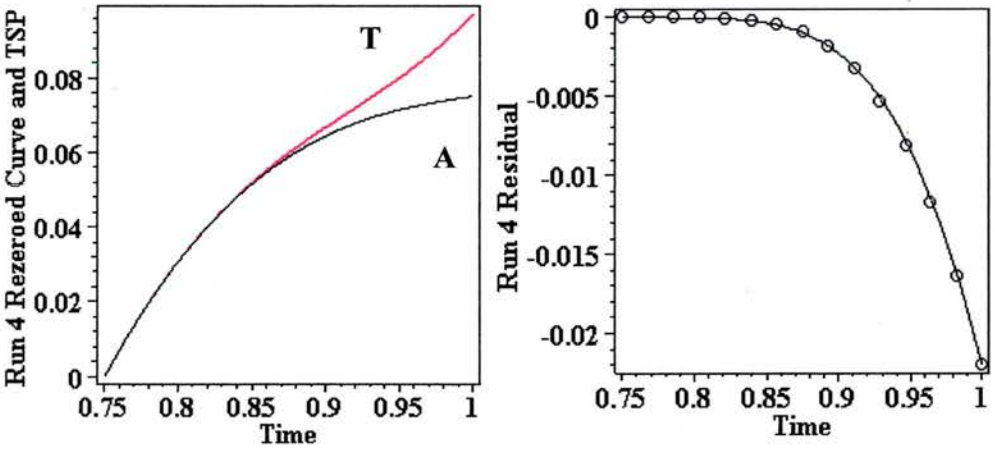
a) b)

Figure 5.11. Quarterway analysis. a) Taylor series prediction (T) from Actual curve (A) for one run of one subject. b) Residual curve (solid line) and closest partial scaled shift curve section fit (circles representing points on the full fit curve) to the residual.



a) b)

Figure 5.12. Halfway analysis. a) Taylor series prediction (T) from Actual curve (A) for one run of one subject. b) Residual curve (solid line) and closest scaled shift curve section fit (circles representing points on the full fit curve) to the residual.



a) b)

Figure 5.13. Three-quarterway analysis. a) Taylor series prediction (T) from Actual curve (A) for one run of one subject. b) Residual curve (solid line) and closest partial scaled shift curve section fit (circles representing points on the full fit curve) to the residual.

The graphs in Figs 5.11, 5.12 and 5.13 show an initially close fit by the truncated Taylor series to the actual curve followed later by more visible deviation from this series. An important question is whether the deviation is due to unmeasured high order terms in an analytic function kicking in in their effect after the horizontal distance travelled is large enough, or is due to a shift.

The fit of a partial scaled shift to the residual indicates the latter to have occurred and further analysis investigates this possibility.

It should be noted that the given shifts providing fit to the residuals are not claimed to be the shift with the best fit. As previously mentioned in Chapter 4, the aim is rather to provide a close and predictive (enough) fit to help decide between the various analytic and non-analytic models.

The given shifts are chosen on the basis of their achieving above a given threshold of fit. A very high degree of fit to scaled shift curve sections is achieved in each case using a 90% similarity threshold. The fits are performed on the basis of the methods described in Chapter 4. The first of these uses the ratio of filtered velocity (V_f) multiplied by the width (W) to the filtered height (S_f), i.e. $V_f\ W\ /\ S_f$, for a pair of points to provide the *stage* T that the shift has reached at the latter point, giving Equation 5.1 below:

[5.1] $$T = V_f W / S_f$$

The ratio of actual height reached (A) to shift curve section height (S) for those points gives the vertical scaling of the shift, giving Equation 5.2 below.

[5.2] $$\text{Vertical Scaling} = A / S$$

With the fitted shift generated in this way the percentage error between the shift and the actual curve can be calculated at each available point in the second half of the interval and the average absolute percentage error taken. The average absolute percentage error is given by the Equation 5.3 below.

[5.3] $$E = \left(\sum_{i=1}^{i=n} \left| (R_i - P_i) / R_i \right| * 100 \right) / n$$

Where E is the average absolute percentage Error, R_i is the height of a Residual at a point, P_i is the shift Prediction using the methods described above and n is the number of points used, i.e. all the available points in the second half of the interval.

Only the available points in the second half of the interval are used because very low values such as those in the first half of the interval are relatively unreliable for the numerical processes being performed. On this measure, fits can be reliably achieved within 10% giving a similarity threshold of 90% for saying that the fit is similar enough to the actual curve. Table 5.4(a-c), below, shows the fitting results for each run of each subject for each of the three points tested. In each case the 'width' measure of the shift refers to the fraction of the shift curve (from its origin) used in the fit, e.g the first eighth or the first quarter of the shift curve. The 'magnitude' measure

of the shift refers to the vertical scaling relative to the standard 1:1 shift curve required to achieve the fit.

Subject	Shift	Run 1	Run 2	Run 3	Run 4
2.1	Width	0.1875	0.171875	0.1875	0.1875
	Magnitude	295	310	340	220
2.2	Width	0.203125	0.171875	0.15625	0.203125
	Magnitude	305	1320	2150	1150
2.3	Width	0.25	0.234375	0.25	0.25
	Magnitude	102	240	208	405
2.4	Width	0.25	0.21875	0.21875	0.21875
	Magnitude	275	418	645	380
2.5	Width	0.265625	0.296875	0.28125	0.265625
	Magnitude	152	34	185	245
2.6	Width	0.21875	0.265625	0.234375	0.234375
	Magnitude	490	230	504	532

Table 5.4(a). Widths and magnitudes of shift curve sections used to fit from quarterway points for each run of each subject.

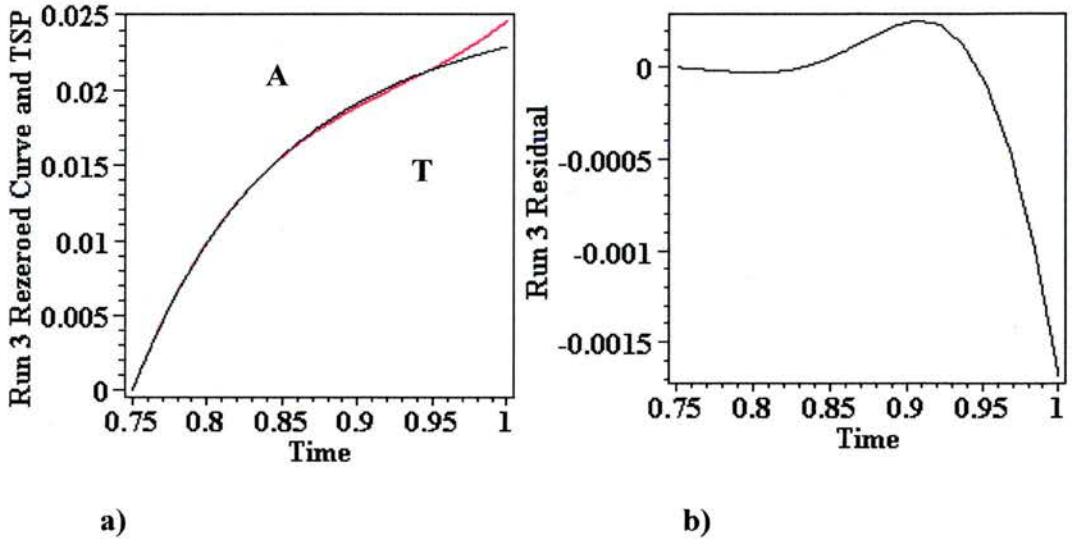
Subject	Shift	Run 1	Run 2	Run 3	Run 4
2.1	Width	No fit	0.203125	0.09375	0.234375
	Magnitude	No fit	-24	-480	7.5
2.2	Width	0.21875	0.234375	0.21875	0.21875
	Magnitude	-40	-32.5	-75	-25
2.3	Width	0.203125	0.21875	0.21875	0.1875
	Magnitude	-7.2	-31	-42	-22
2.4	Width	0.234375	0.21875	0.1875	0.1875
	Magnitude	-16	-29	-64	-77
2.5	Width	0.125	0.25	0.15625	0.359375
	Magnitude	-150	-16	33	-1.87
2.6	Width	0.328125	0.21875	0.21875	0.296875
	Magnitude	3.37	15	-22	-3.55

Table 5.4(b). Widths and magnitudes of shift curve sections used to fit for halfway points for each run of each subject.

Subject	Shift	Run 1	Run 2	Run 3	Run 4
2.1	Width	0.203125	0.140625	No fit	0.21875
	Magnitude	0.61	-2.25	No fit	-0.52
2.2	Width	No fit	0.265625	0.171875	0.171875
	Magnitude	No fit	0.15	-0.76	0.86
2.3	Width	No fit	No fit	0.1875	0.1875
	Magnitude	No fit	No fit	0.34	-0.95
2.4	Width	No fit	0.25	No fit	0.125
	Magnitude	No fit	-0.185	No fit	2.95
2.5	Width	0.203125	No fit	0.203125	0.21875
	Magnitude	-1.26	No fit	0.55	0.49
2.6	Width	0.09375	0.15625	0.234375	0.1875
	Magnitude	22.2	0.67	-0.66	-0.84

Table 5.4(c). Widths and magnitudes of shift curve sections used to fit for three-quarterway points for each run of each subject.

In the following, a 'case' represents one run for one subject. The results show that in 64 of 72 cases a close fit to a simple section of the shift curve can be achieved. Figure 5.14, below, shows an example of 1 of the 8 'no fit' cases which seem to result from cases where the residual and actual curves cross each other.



a) b)

Figure 5.14. Example of an atypical case of residuals that cannot be fit well by a simple shift curve section. a) Subject 2.4, Run 3, three-quarterway Taylor series prediction (T) from Actual curve (A). b) The resulting Residual curve (R).

In the case of the quarterway predictions good fits are achieved for all agents and all runs (see Table 5.4). One 'no fit' case appears for the halfway case and seven 'no fit' cases for the three-quarterway predictions.

One possible explanation for the good fits with a partial scaled low order shift, i.e. a partial scaled height shift, is that the subject may be perturbed in exiting the first pipe and entering the free space between the pipes for a short distance before settling down to a simple behaviour. That is, the perturbation puts the subjects into various non-zero derivative states before the subject is in sufficient control to execute a simple, i.e. low order, shift to gain access to the second pipe.

The perturbation theory suggests that, in line with the derivative states arising from perturbation, subjects will execute the various scaled shifts that are appropriate responses to these varied states. Consequently, the variation in the scaling across runs and subjects may be simply due to the varied nature of the perturbation.

A place where there is constancy is the good fit universally achieved at the early time of $t = 1/4$.

A low order non-analytic shift projection is a good predictor at $t = 1/4$ because there are sufficient terms in the measured set of derivatives to make for a truncated Taylor series close to the full series in its effect as well as the agent executing a low order shift. The reason for this sufficiency may be theorised as due to the perturbation and movement away from zero vertical displacement being small in a key respect as well as varied. The perturbation is sizeable in shaping the behaviour through low order derivatives so that only a partial height shift is then needed to enter the second pipe. However, the perturbation may also be small in generating small average values for the unmeasured high order derivatives. The absence of these derivatives in the truncated Taylor series filter would then be of no great consequence and so lead to the truncated analytic projection being the relatively effective representative of the full Taylor series it seems to be.

The actual shift may also start relatively close to $t = 1/4$ so that the deviation from the Taylor series is relatively small at this stage, and the shift is still very flat. Consequently, taking the fitted shift to start at $t = 1/4$ would then generate only the small errors observed.

The two factors of low order perturbation from a zero-valued derivative state and a nearby origin for the low order shift may thus combine to produce a reasonably accurate residual at t = 1/4 that can be predictively fitted by a low order shift.

The good fits relatively early on, e.g. at t = 1/4, provide an explanation of the poor fits later on, e.g. at t = 3/4. The later filtering is predictably bad at points requiring unmeasured high order derivatives for good prediction. The time t = 3/4 was chosen because it may be expected to yield good shift fits for 1:1 shifts beginning at the exit of the first pipe. Only the first three orders of derivative (zeroth, first and second) are non-zero at t = 3/4 for these shifts, as explained in Chapter 4. Hence projections based on measurements including these orders would be reasonably accurate. However, given that a (partial) scaled shift is actually found at t = 3/4 to be a good fit, t = 3/4 can be predicted in the poor ('no fit') cases to require measurement of unmeasured higher orders of derivatives for accurate projection because the terms involving the latter are significantly non-zero. If a later time of half the actual shift interval is used, i.e. t = 5/8 instead of t = 3/4, a good fit is expected to return due to only measured low orders of derivative being significantly non-zero. And indeed this hypothesis has been confirmed numerically. These points that allow good fits can be thought of as analagous to points on a radio dial where radio stations can be found. As long as you pick the right point you can get a good signal.

This reasoning applies to other points that we look at as well. Any point that is difficult to work with as a filtering origin can be viewed as a point on a path that could be better analysed using a shift curve section from some other point more

amenable to analysis because the measured derivatives are a large enough set at the latter point to adequately represent the Taylor series.

For the initial working hypothesis, the supposition is that the residual could produce a *copy* of the original curve if the latter was a full height shift. This hypothesis has to be revised in the light of partial scaled shifts fitting the residual. Nevertheless, the successful repeated fit by the *same* partial scaled shift at t = 1/4 and t = 5/8 shows the copy property to be present.

It is worth noting that the shift theory can also supply estimated values for unmeasured high order derivative values over the interval in which the shift appears. This is because the partial height shift fitted may be decomposed into its higher order components and so supply estimates for high order values for times following that of the filter origin. For example, Figure 5.15 (below) shows a fitted quarter shift curve stretched over the interval from the quarterway point (at t = 1/4) to the end of the actual shift (at t = 1) together with the extrapolation of the full shift (shown here in the zeroth and first derivative). Hence the residual is fitted by the zeroth derivative, but may also be given estimates for its higher order derivatives through decomposition into higher order derivatives such as the first derivative profile in Figure 5.15. The actual residual should contain later high order values close to these estimates because it is closely fitted by the shift.

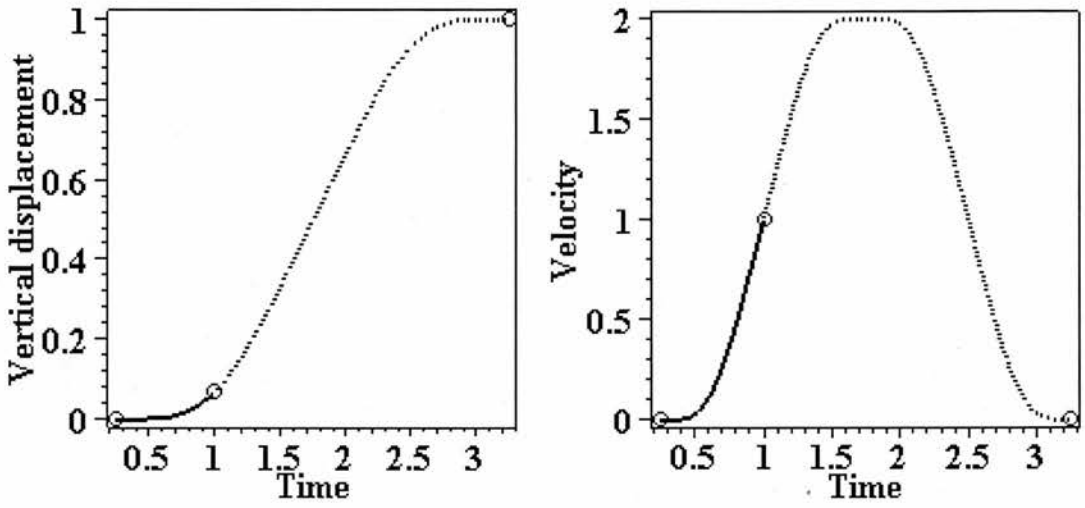


Figure 5.15. A quarter of a standard 1:1 shift (solid lines) stretched over the interval t = 0.25 to t = 1 and the completion of the full shift (dotted lines). a) Zeroth order derivative profile. b) First order derivative profile.

The full set of underlying derivative variations is thereby accounted for.

The universal fit according to the pairing of the observed goal derivative state at the entrance to the second pipe and a state at or near t = 1/4 suggested that the subjects all had a common cognitive strategy of taking aim over roughly the first quarter of the interval followed by a 'firing' phase over the rest of the interval between the pipes consisting of a ballistic zeroth order partial shift.

The only exception to this strategy appeared to be in the secondary submovements. These submovements occur in arbitrary places in the interval between the pipes and are consequently missing from the averaged profiles. Such submovements may be a corrective motion that modifies the aim and fire strategy. That is, they may be part of

an aim, correct, and fire strategy with correction occurring if the aiming was not skilled enough to be sufficiently on target for a final ballistic movement.

For profiles of individual trials we concentrated on the simpler aim and fire strategy by leaving out profiles with secondary submovements. We then tested successful individual trials and again found the same degree and universality of fit from predictions based on the observed goal derivative state at the entrance to the second pipe and one at or near t = 1/4. The individual trials sampled showed the same high degree of fit found for the averaged profiles. An example of such a predictive fit is shown in Figure 5.16 below.

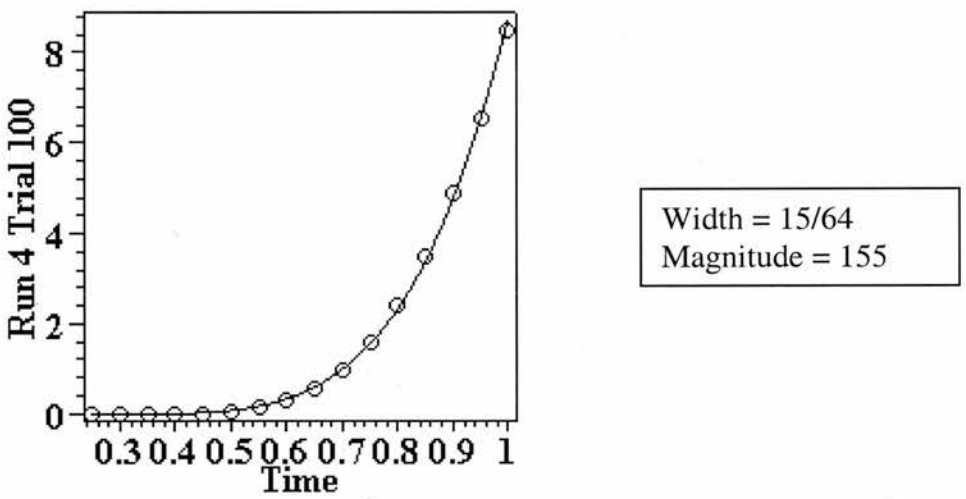


Figure 5.16. Graph of predictive fit for a single individual trial of one subject (Run 4 Trial 100 of Subject 2.1).

5.3.4. Discussion

Experiment 2 imposed an undulatory smoothness constraint on the subjects that was aimed at allowing greater midway symmetries to emerge. This was partially achieved

but task specific factors such as getting up out of the first pipe early to level off for easy entry into the second pipe were still an issue. The subjects did not perform the task using a simple fully symmetric shift.

However, Taylor series filtering did typically produce residuals that could be closely fit by simple partial scaled shift curve sections. The width and magnitude of these sections were fairly consistent within each run for quarterway start points between the pipe ends. However, the width and magnitudes differed somewhat across runs and agents.

It was argued that for filter origins where a good fit was not achieved this was due to not enough terms of the Taylor series being available for the analysis. Using other points as filter origins where enough terms are available allowed predictions involving the troublesome points as a later part of some simple shift prediction. Derivative variation of all orders during the shift is accounted for through the fitted shift and its decomposition. By using shift values, the derivatives in the account are consistent with a low frequency shape between as well as at the discrete data points.

5.4. Conclusion

This chapter presented controlled human hand movement kinematics experiments to investigate C^{∞} smooth goal-directed behaviour and decide between the models introduced earlier, i.e. analytic, piecewise analytic, non-analytic shift curve and general non-analytic. The experiments were designed to investigate a possible base unit of C^{∞} smooth goal-directed behaviour, i.e. the shift.

In both experiments the desired C^{∞} smoothness is not contradicted because no evidence of any gaps or the precursors to gaps (sharp gradient changes) was found in any derivative order used. Indeed, lack of C^{∞} smoothness would mean that jump discontinuities in the zeroth order or its equivalent in higher orders would be occurring. Such discontinuities may be a reasonable way of modelling high frequency phenomena but not the low frequency phenomena investigated in the present study.

The behaviour investigated lies in a 3D vertical displacement versus horizontal displacement versus time cube. In the Experiments presented in this and the following chapters the vertical displacement versus time domain has been chosen for the purposes of kinematic analysis. Behaviour in the other domains, (horizontal displacement versus time (also a kinematic domain) and vertical displacement versus horizontal displacement (a purely spatial domain)) has also been analysed and, whilst interesting in itself, is outside the main focus of the thesis and so is not included here.

In particular, kinematic deviation from the Taylor series – if it exists – could be most expected to be shown by the vertical displacement versus time domain rather than the other two domains. This is because the vertical displacement versus horizontal displacement domain does not involve time and so is not kinematic, and horizontal displacement versus time is a secondary feature of the kinematic task teleology – which is to change vertical displacement from the existing zero to a goal height over time.

Experiment 1 showed that extraneous variables lead to asymmetries in the resulting vertical displacement versus time kinematic profile and its derivatives. The asymmetries are caused by two things. Firstly, subjects pull up early for easy linear entry into the second pipe. Secondly, there are often one or more submovements in individual runs (not over runs) overlapping the primary movement due to lack of skill requiring correction to get back on course for the goal of entering the second pipe.

Experiment 2 showed that using an undulatory smoothness constraint based on limiting changes in local curvature allowed greater symmetry to emerge but, because of remaining task specific factors, the initial hypothesis of a single fully symmetric shift curve being present in the central phase of the motion could not be proven.

The results of Taylor series filtering demonstrated the existence of significant non-zero residuals that could be fit closely by shift curve sections using the V_f W / S_f and A / S (Actual height to Standard shift curve height) ratio methods described in Chapter 4 and exceeding a 90% similarity threshold. This key test would have refuted the possibility that there is an underlying shift curve had zero residuals been found.

A possible explanation for the partial scaled shift is that subjects may have been *perturbed* in coming out of the first pipe but settled down to execute a partial scaled shift over the rest of the interval from around the quarterway point. This also allowed the relative poorness of some fits at the halfway and three-quarterway points to be explained. If these points had been part of a standard 1:1 shift they would have had only low order non-zero derivatives, but actually the shift is scaled and so these points may sometimes have significantly non-zero derivatives present. However, because of

the fitting of a simple scaled shift over the interval from the quarterway point it is possible to identify other intermediate points other than halfway and three-quarterway where only low order derivative terms are required.

The bottom line is that as long as you take a non-analytic projection from a point where enough terms of the Taylor series are available then the projection successfully predicts a simple shift curve section to fit the rest of the residual curve. Any points that you can't project a non-analytic predictive fit from directly can still be seen as forming a part of a single simple shift curve section from some previous point where enough terms were available.

It is now possible to assess the main aim of the chapter – to choose between the alternative models. The analytic model's only support comes from the fact that, by eye, some may argue that a reasonable fit is achieved at early points near to the filter origin. This is not surprising from a non-analytic point of view since any shift would only just be starting at these points. At further distances this fit becomes poor, even by eye. The analytic model is not a good predictor of behaviour anywhere in the interval between the pipes (except, trivially, at the entry to the second pipe).

In the case of the non-analytic shift curve model, the original hypothesis of a 1:1 shift being present is not supported. However, significant non-zero residuals are found that can be fit using a 90% threshold with partial scaled shift curve sections, thus giving predictive ability. The prediction may use the derivative state current at a suitable moment (e.g. t = 1/4 in the experiments) and a future goal derivative state (e.g. t = 1 in the experiments) to predict the intervening behaviour.

The piecewise analytic model is not supported because no mixture of flat or near flat and steeper residuals was found.

So the final decision is that the model with the greatest support is the non-analytic shift curve model, with the modification that partial scaled height shifts are used rather than full 1:1 height shifts. This model allows accurate prediction of behaviour over significant intervals using the measureable derivatives unlike any of its rivals.

Chapter 6

Shift Generalization

6.1. Introduction

The two experiments in Chapter 5 were aimed at isolating a base unit for C^{∞} smooth goal-directed behaviour. This base unit was hypothesised to be the non-analytic shift curve. The Parallel Pipes Problem, in which subjects were required to move through one horizontal pipe on the screen, then up to and through a second horizontal pipe, was used to investigate this hypothesis.

The results showed that the subjects did not perform the task using a full symmetric height shift as originally hypothesised. It was argued that this may have been due to task specific factors such as the subjects moving upwards early on out of the second pipe to make it easier to level off and achieve linear entry into the second pipe.

However, subjects did perform the task in such a way that significant non-zero residuals resulted from Taylor series filtering (see Chapter 4) that could be predictively fit, using fractions of the shift curve from the origin, above a 90% threshold when an average absolute percentage error measure was used. In particular it was found that subjects tend to use roughly the first quarter of the interval to set themselves up for using a simple shift away from the Taylor series prediction during the remainder of the interval.

This led to a revision of the original hypothesis of a full symmetric height shift over the interval to a new hypothesis that predicts that partial low order scaled height shifts will be employed by the subjects after an initial period during which they corrected for the perturbation associated with changing from one phase of motion to the next, i.e. coming out of the constraints of the first pipe and entering free space whilst being required to move through that free space in such a way as to achieve entry into the second pipe.

This chapter investigates three cases that modify the parameters of Experiment 2 to investigate how well the base unit can be used for the purposes of predictive generalization to new experimental settings.

In each new experiment standard statistics are presented that allow comparison with other kinematics studies thus allowing the present work to be linked into the contemporary kinematics community. Psychological explanation is attempted in interpreting this data to give readers a feel for the data. Some of the statistics, specifically those to do with symmetry measures, are particularly relevant to the core

of the thesis as they say things about the shifts present in the data. As well as the standard statistics the key test – involving the use of Taylor series filtering – is also performed in each case.

Experiment 3 investigates the effects of changing the vertical displacement to horizontal displacement ratio for the central phase of the motion, i.e. between the two pipes. Figure 6.1 shows this setup in comparison to the original setup. In Experiments 1 and 2 the ratio of the horizontal displacement to vertical displacement between the central exit point of the first pipe and the central entry point to the second pipe was 1:1. An illustration of the position of the second pipe in Experiment 1 and 2 relative to the first pipe is shown using dashed lines. In Experiment 3, where the subjects must traverse a path of similar ratio to that shown highlighted by the arrows in Figure 6.1, the ratio is now is 2:1. It is hypothesised that significant non-zero residuals will appear that, after vertical scaling, will allow predictive fits above the 90% threshold to be achieved using low order partial scaled shifts from around the quarterway mark.

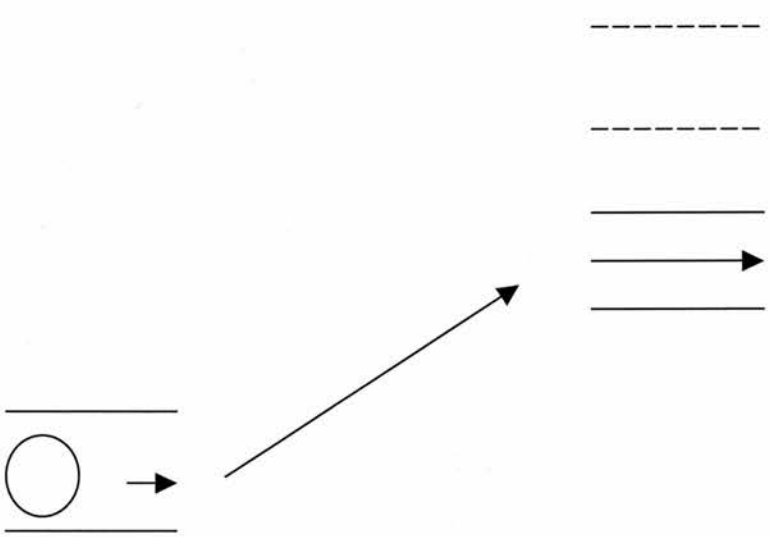


Figure 6.1. An illustration of the halved vertical displacement task used in Experiment 3. The subjects must traverse the path highlighted here by the arrows. The dashed lines represent the position of the second pipe in Experiment 1 and 2.

Experiment 4 investigates the effects of changing the angle of the second (still linear) pipe in relation to the first pipe. This provides evidence of how the hypothesised model generalises to cases where the start and end first order derivative values of the motion differ. Figure 6.2 shows the new 45 degree second pipe setting (solid line) in comparison to the original horizontal second pipe setting (dashed line). Because of the difference in the end derivative state in comparison to Experiment 2, the type of shift used may be different, but the difference may be predictable. One modification to the initial hypothesis suggests that back-to-back velocity shifts of different magnitudes will be present. By taking into consideration the results of Experiment 2 and 3 an alternative hypothesis is that significant non-zero residuals will again appear that can

be predictively fit above the 90% threshold using low order shifts over most of the interval after an initial period of settling in.

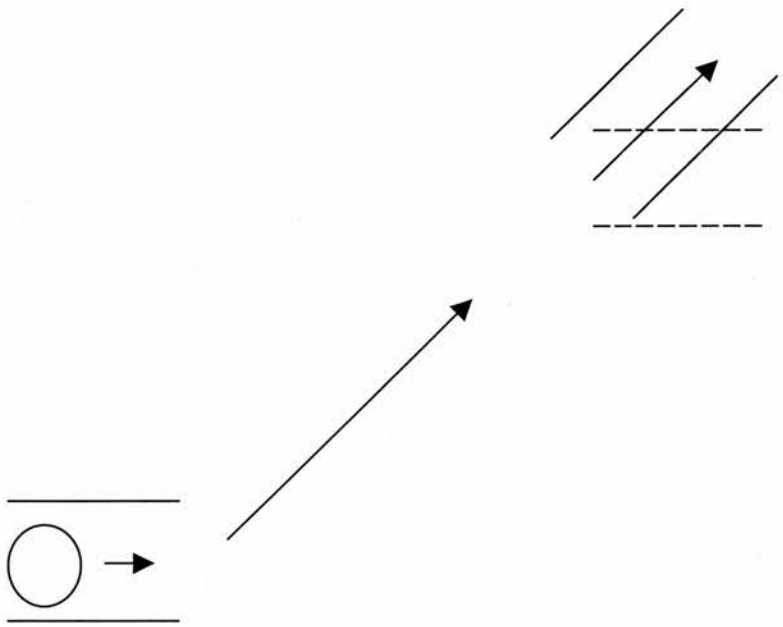


Figure 6.2. An illustration of the 45-degree angled second pipe task used in Experiment 3. The subjects must traverse the path highlighted here by the arrows. The dashed lines represent the position of the second pipe in Experiments 1 and 2.

Experiment 5 investigates the effects of introducing an *external* perturbation over the central phase of the motion, which smoothly pushes the agent off course at every point on the path (N.B. this is distinct from the *internal* perturbation which was referred to earlier when describing the disruption caused to the path by the subject themselves changing from one phase of motion inside the first pipe to a new phase of motion in free space). The key point to appreciate here is that if the subject is projecting a path from a point that goes through the goal then in pushing them away from the goal they will require non-analyticity to move back towards the goal. If you perturb them at every point then they must redirect non-analytically at every point to

remain projecting through the goal. The key questions then are if this projection does indeed occur and, if so, how the non-analytic response is put together.

Experiment 5 uses the same pipes settings as in Experiment 3, i.e. with the same 2:1 horizontal displacement to vertical displacement ratio between the two pipes. A negative perturbation based on the non-analytic shift curve halved in its vertical displacement is added at each point in the horizontal displacement as the agent travels across the screen. Figure 6.3 shows the relationship between the symmetric shift curve (S) and the shape of the perturbation (P) encountered by the subjects in the vertical displacement versus horizontal displacement domain (neither of which, of course, are represented on the screen during the task). So the overall task requires the subjects to vertically move more than the actual distance between the pipes to attain the second pipe. In fact the agent has to move double the vertical distance between the pipes. It is hypothesised that subjects, although finding the task more difficult than in Experiments 2 and 3, may overcome the perturbation by again employing low order shifts that are predictive fits in some way to the residuals above the 90% threshold after the initial settling in period.

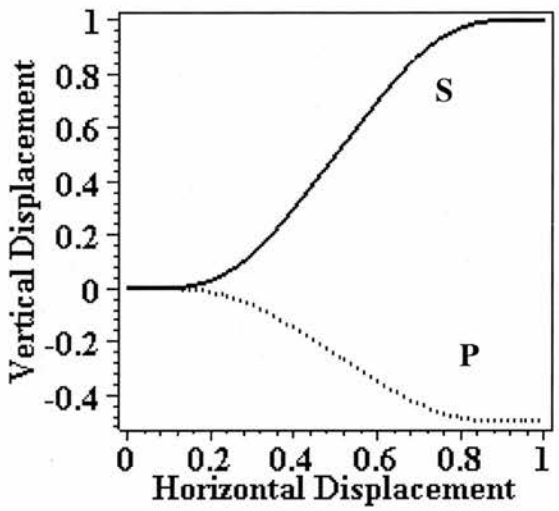


Figure 6.3. The relationship between the symmetric shift curve (S) and the perturbation (P) encountered by the subjects in Experiment 5.

6.2. Experiment 3: The Parallel Pipes Problem With Scaling

6.2.1. Introduction

The first generalisation to tackle investigates the effects of changing the scaling of the central phase of the motion, i.e. the horizontal displacement to vertical displacement ratio between the two pipes. This was changed from 1:1 in Experiment 2 to 2:1 in Experiment 3. This is an important first step because any further ratios may then reasonably be expected to follow similar principles.

In this experiment it is hypothesised that, when the curves are rescaled onto the 1:1 unit interval similar kinematic profiles and statistical test values will appear to those for Experiment 2 (see Chapter 5). Most importantly it is hypothesised that similar

significant non-zero residuals with predictive fits above the 90% threshold using low order partial scaled shifts after an initial period of aiming will appear as in Experiment 2. If this latter prediction holds then it will provide the first demonstration that the hypothesised base unit of C^{∞} smooth goal-directed behaviour, i.e. the non-analytic shift curve, can successfully be used for predictive generalisation over tasks.

6.2.2. Method

The design was the same as that in Experiment 2 except that the vertical displacement between the two pipes in Experiment 3 was half the vertical displacement between the two pipes in Experiment 2 (as shown in Figure 6.1). Again six healthy subjects participated voluntarily, and there were 4 males and 2 females, with an age range of 24-50 and with 5 right-handers and 1 left-hander. The apparatus and experimental procedure were the same as in Experiment 2.

6.2.3. Results

The kinematic profiles averaged for each run for each subject show that the subjects use a similar strategy as in Experiment 2. As can be seen in Figure 6.4 there is again midway asymmetry in the typical profiles that may be due to subjects pulling up early and flattening out for easier linear entry into the second pipe. Certainly this is what some subjects reported doing.

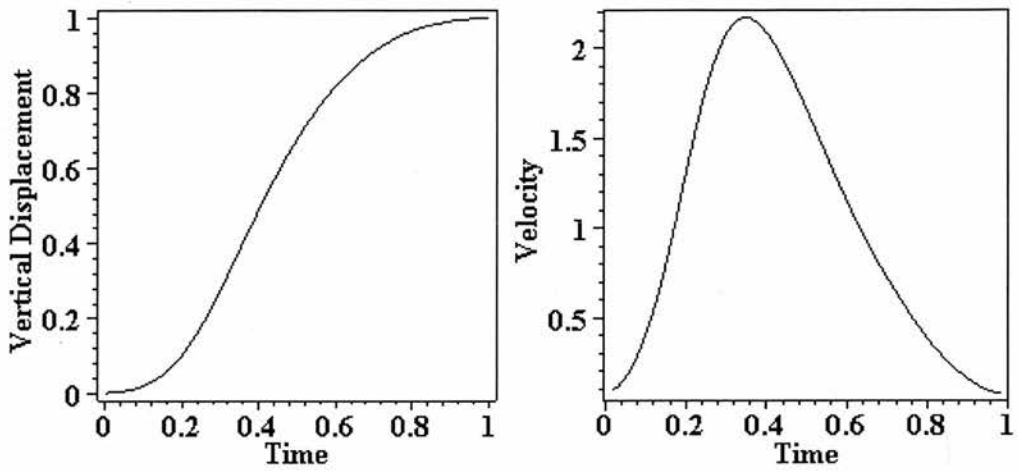


Figure 6.4. Typical kinematic profiles for a single agent in a single run. (a) The vertical displacement versus time. (b) The velocity versus time.

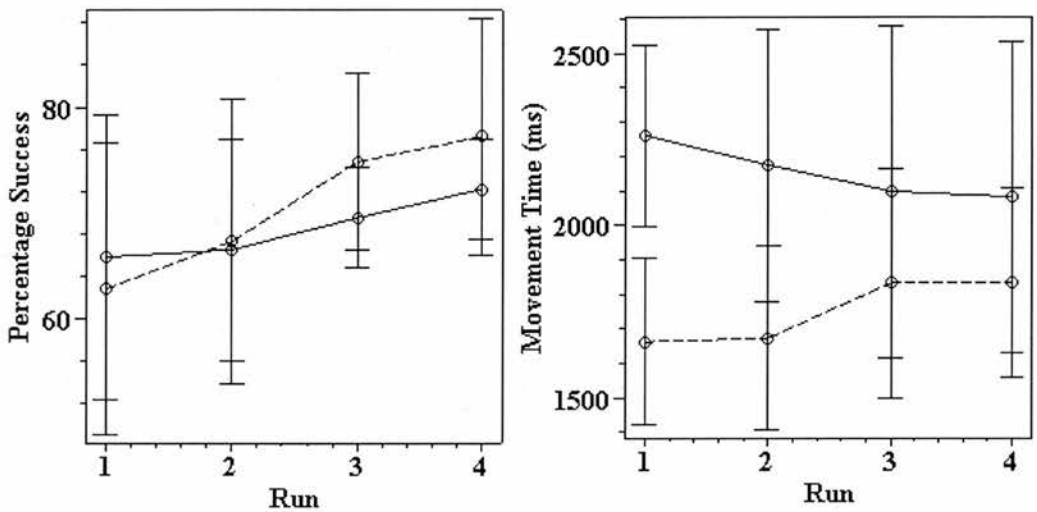
Table 6.1, below, presents the standard statistics for percentage success, movement time, peak velocity, and mean jerk squared. Peak velocity and mean jerk squared values are based on the values after rescaling of the halved vertical displacement curve onto the unit space thus allowing direct comparison of values in this experiment to those in Experiment 2.

Subject	Run	Trials	Percentage success	Movement time	Peak velocity	Mean Jerk^2
3.1	1	100	60	1438.72	2.440	4709.39
	4	100	58	1511.29	2.409	3113.32
3.2	1	100	80	2127.58	2.876	8046.30
	4	100	91	1787.44	3.248	11115.88
3.3.	1	100	55	1593.47	1.994	2454.29
	4	100	74	2169.61	2.027	2851.52
3.4	1	100	42	1606.12	2.589	5483.78
	4	100	75	2041.86	2.738	7403.51
3.5	1	100	65	1663.46	2.201	1995.66
	4	100	82	1518.76	2.203	2129.95
3.6	1	100	75	1545.68	2.667	6544.29
	4	100	83	1970.63	2.689	6018.51
Average	1	600	62.8	1662.51	2.461	4872.29
	2	600	67.3	1672.09	2.548	5306.11
	3	600	74.8	1831.66	2.578	5538.84
	4	600	77.2	1833.27	2.552	5438.78
All		2400	70.5	1749.88	2.535	5289.00

Table 6.1. Table of standard statistics (for movement time, peak velocity and mean jerk squared only the success cases within that run are used). N.B. Underlines represent significant changes in a metric at the 0.05 significance level.

The percentage success averaged over agents for Experiment 3 increased monotonically over runs as it did in Experiment 2 (see Figure 6.5(a)). However, the movement time averaged over agents increased monotonically over runs in contrast to Experiment 2 where it decreased monotonically over runs (see Figure 6.5(b)). This difference may be expected because the vertical displacement that the subjects need to traverse has been halved whilst retaining the same time limit. In Experiment 2

subjects start off near to the time limit and achieve greater success by speeding up whereas in Experiment 3 they start off well below the time limit and achieve greater success by slowing down a bit.



a) b)

Figure 6.5. a) Graph of percentage success over runs comparing Experiment 2 (solid line) and Experiment 3 (dashed line). b) Graph of movement time over runs comparing Experiment 2 (solid line) and Experiment 3 (dashed line). (N.B. In this and other trend graphs below error bars are shown to one standard deviation).

The overall peak velocity in Experiment 2 is almost identical to the overall peak velocity for Experiment 2 (2.537 in Experiment 2 as opposed to 2.535 in Experiment 3). However, a noticeable difference lies in the fact that the mean squared jerk in Experiment 3 is only 78% of that in Experiment 2. This is perhaps a product of the fact that the subjects in Experiment 3 found it easier to smoothly achieve their secondary aim of getting up early out of the first pipe to flatten out for easy entry into

the second pipe than the subjects in Experiment 2 because of the vertical distance traversed in the Experiment 3 task is half that in the Experiment 2 task. This secondary goal requires less fluctuation in regulating acceleration and deceleration for the goal to be achieved, thus lowering the mean squared jerk. As in Experiment 2 there is no evidence that subjects seek to increase or decrease mean squared jerk systematically over runs.

The results of further statistical measures are shown in Table 6.2 below. These include LMS symmetry, K, C and the modified K statistics for the acceleration and deceleration phases of the motion.

Subject	Run	Trials	LMS symmetry	K	C	K (acceleration peak)	K (acceleration trough)
3.1	1	100	3.015	0.366	3.822	0.594	0.297
	4	100	3.031	0.349	3.779	0.555	0.273
3.2	1	100	5.384	0.282	4.511	0.572	0.179
	4	100	7.129	0.265	5.091	0.661	0.159
3.3.	1	100	0.588	0.370	3.130	0.551	0.325
	4	100	1.079	0.331	3.182	0.717	0.314
3.4	1	100	2.747	0.319	4.056	0.612	0.195
	4	100	4.561	0.267	4.292	0.890	0.169
3.5	1	100	0.732	0.427	3.445	0.527	0.313
	4	100	1.345	0.368	3.449	0.685	0.289
3.6	1	100	1.929	0.341	4.175	0.641	0.208
	4	100	1.928	0.351	4.206	0.685	0.193
Average	1	600	2.399	0.351	3.857	0.583	0.253
	2	600	2.430	0.354	3.990	0.645	0.241
	3	600	3.154	0.330	4.040	0.704	0.233
	4	600	3.179	0.322	4.000	0.699	0.233
All		2400	2.790	0.339	3.971	0.658	0.240

Table 6.2. Table of LMS symmetry, K, modified K and C statistics for Runs 1 and 4 for each subject.

The LMS symmetry value increased over runs and was slightly worse in Experiment 3 than it was in Experiment 2. At the same time a drop was also seen in the K value for Experiment 3 in comparison to Experiment 2. These two indicators imply a slight worsening of midway symmetry that ties in with the hypothesis that subjects found it easier to achieve their secondary aim of getting up quickly out of the first pipe for easy linear entry into the second pipe.

The C value in Experiment 3 is almost identical to Experiment 2 showing similarity in this aspect of the kinematic strategies used by the subjects in the two experiments. Given that the peak velocity was also very similar in the two experiments this implies that the average velocity (after scaling) is also similar in the two experiments because the C value is the ratio of the peak velocity to the average velocity of the motion.

The Modified K values, which look at the relative time to peak or trough acceleration within the acceleration and decleration phases of the motion show a tendency towards increased asymmetry over runs and greater asymmetry than in Experiment 2. The acceleration peak (over the acceleration interval) becomes skewed further to the right of centre within that interval over runs. The acceleration trough (over the deceleration interval) becomes skewed further to the left of centre within that interval over runs. These increasing asymmetries may be seen as a side effect of the fact that the secondary aim of getting up early and flattening out early for easier linear entry into the second pipe is being achieved more easily in Experiment 3 than in Experiment 2.

Taylor series filtering (see Chapter 4) was performed from the quarterway, halfway and three-quarterway points for each of the four runs for each of the six subjects. The results of shift curve predictive fitting using the V_f W / S_f and A / S ratios method as described in Chapter 4 and applied in Chapter 5 (to the results of Experiment 2) are shown in Table 6.3(a-c) below for the width and vertical magnitude of the shift found. (See Appendix C for a worked example of the methodology using Run 4 of Subject 2.1 in Experiment 2).

Subject	Shift	Run 1	Run 2	Run 3	Run 4
3.1	Width	0.25	No fit	0.203125	0.25
	Magnitude	142	No fit	198	48
3.2	Width	0.34375	No fit	0.171875	No fit
	Magnitude	72	No fit	-1000	No fit
3.3	Width	0.21875	0.21875	0.25	0.296875
	Magnitude	380	320	210	33
3.4	Width	0.25	0.375	0.4375	No fit
	Magnitude	420	47	21	No fit
3.5	Width	0.203125	0.21875	0.21875	0.25
	Magnitude	215	195	300	175
3.6	Width	0.25	0.21875	0.25	0.234375
	Magnitude	725	700	380	520

Table 6.3(a). Widths and magnitudes of shift curve sections used to fit from quarterway points for each run of each subject.

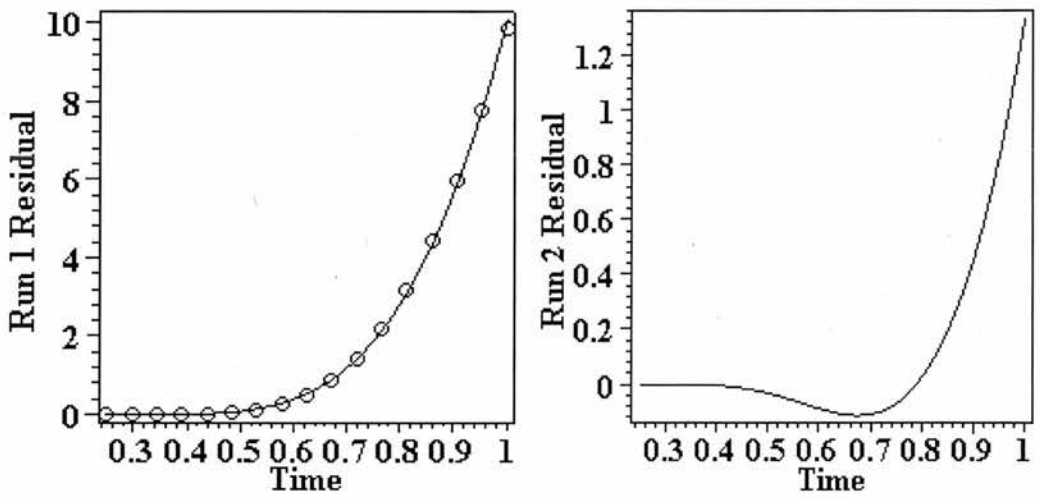
Subject	Shift	Run 1	Run 2	Run 3	Run 4
3.1	Width	0.21875	No fit	0.21875	0.21875
	Magnitude	-22	No fit	-22	-21
3.2	Width	0.1875	0.25	0.203125	No fit
	Magnitude	20	6.8	24.5	No fit
3.3	Width	0.21875	0.28125	0.125	0.3125
	Magnitude	-27	-10.5	-24	2.8
3.4	Width	0.375	0.25	0.265625	No fit
	Magnitude	-0.65	-1.42	14.5	No fit
3.5	Width	0.28125	0.171875	0.125	No fit
	Magnitude	.25	-37	-100	No fit
3.6	Width	0.421875	0.21875	No fit	0.359375
	Magnitude	-0.52	-59	No fit	-1.08

Table 6.3(b). Widths and magnitudes of shift curve sections used to fit for halfway points for each run of each subject.

Subject	Shift	Run 1	Run 2	Run 3	Run 4
3.1	Width	No fit	0.125	0.25	0.25
	Magnitude	No fit	6.5	-0.3	-0.29
3.2	Width	0.21875	0.234375	No fit	0.1875
	Magnitude	0.14	0.125	No fit	-0.18
3.3	Width	0.1875	0.21875	No fit	0.125
	Magnitude	0.49	0.55	No fit	10
3.4	Width	0.1875	0.3125	0.15625	0.125
	Magnitude	1.84	0.031	1.07	3.55
3.5	Width	0.28125	No fit	0.21875	No fit
	Magnitude	0.165	No fit	-0.54	No fit
3.6	Width	No fit	0.203125	0.171875	0.140625
	Magnitude	No fit	-1.07	0.46	2

Table 6.3(c). Widths and magnitudes of shift curve sections used to fit for three-quarterway points for each run of each subject.

As in Experiment 2 the typical curves (found in 57 of 72 cases) are those that can be predictively fit beyond the 90% similarity threshold using a low order partial scaled shift (see Figure 6.6(a)). The atypical cases (see Figure 6.6(b)) dip in the middle meaning that no single section of the non-analytic shift curve can be fitted to them, although combinations of shift sections may be used.



a) b)

Figure 6.6. a) Run 1 quarterway residual (solid) for Subject 3.1 and a predictive fit to the residual (circles representing points on the full fit curve). The shift curve section used is the first quarter of the shift curve translated onto the interval [0.25, 1] and scaled horizontally to fit the residual. b) Run 2 quarterway residual (solid) for Subject 3.1 showing a case where the residual cannot be fitted using a single shift curve section.

From the quarterway point there is typically a good predictive fit using roughly the first quarter of the shift curve stretched over the remainder of the interval, although unlike in Experiment 2 there are 4 cases out of the 24 that lack a good fit (there were no bad fits from the quarterway point in Experiment 2).

Predictions from the halfway and three-quarterway points show a greater number of bad fits and greater inconsistency in terms of width and magnitude than from the quarterway point.

A possible explanation for the atypical cases which result in bad fits involving dips is provided by shift theory. It may be that such cases are caused by the presence of significant unmeasured higher order derivatives. If this is the case then it should still be possible to find nearby points, or 'other nearby low order tunings' according to the radio dial analogy (see Section 5.3.3), from which more measurable residuals and therefore good predictive fits can be achieved.

Investigation of this possibility suggests that this is indeed the case. Figure 6.7 below shows an example where a run that produced a bad fit (Subject 3.2, Run 4) could be reassessed at a nearby point, in this case t =3/8, to provide a good fit using a low order partial scaled shift. This pattern held for the other 'no fit' points at t = ¼ as well. In each case moving further on from t = ¼ rather than nearer to the origin was found to work. This suggests that the subjects employed a slightly extended settling in period prior to executing the low order partial scaled shifts in these cases.

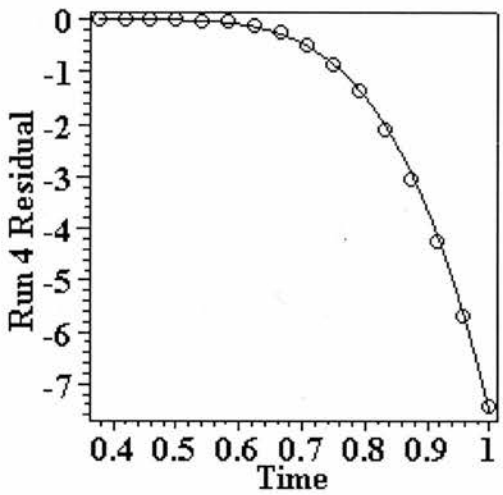


Figure 6.7. Run 4 residual (solid line) for Subject 3.2 from the t = 3/8 point which was a 'no fit' case from the quarterway point. The predictive fit is also shown (circles representing points on the full fit curve).

6.2.4. Discussion

Kinematic profiles display significant similarities to those of Experiment 2 in maintaining peak velocity, average velocity, and therefore C values. However, midway asymmetries have been accentuated slightly because the secondary goal of moving up early and flattening out for easy entry into the second pipe has been made easier for the subjects.

Again, predictive fits beyond a 90% similarity threshold using low order partial scaled shifts were universally achieved in all cases tested. There were more bad fits found in Experiment 3 at the three standard bisection points than in Experiment 2. Nevertheless, as in Experiment 2, the modified non-analytic model using partial

scaled shifts provided a typically effective way of predicting the behaviour from these points in the vast majority of cases. And in fact by looking slightly further into the interval than the t = ¼ point good fits were found in all the runs that had returned bad fits from the quarterway, suggesting that subjects had just used slightly more time for their settling in period. In conclusion, the bottom line is that the base unit, the symmetric non-analytic shift curve, has been successfully used for the first time to provide full generalization to a new experimental setting.

6.3. Experiment 4: The Non-Parallel Pipes Problem

6.3.1. Introduction

In Experiments 1, 2 and 3, parallel linear pipes were used to ensure zero derivatives at all orders at the end of the first pipe and derivatives of 1 in the zeroth order and zero at all higher orders at the start of the second pipe, in an attempt to test for the base unit.

For the purposes of this section non-parallel pipes will be taken to be pairs of linear pipes that are not parallel to each other (as opposed to the linear pipes used in Experiments 1, 2 and 3). Such pipes are used in Experiment 4 to take the first step towards generalising the approach to cases where the start and end derivative states are different beyond the zeroth order.

The aim is to investigate what effect changing the angle of the second pipe has on the results in terms of the kinematic profiles, the various statistical measures and also in

terms of whether Taylor series filtering results in significant non-zero residuals that can be predictively fit (beyond the 90% similarity threshold) using low order partial scaled shifts.

A velocity peak is predicted in Experiment 4 because the subjects still require a substantial acceleration phase to achieve the time limit followed by a deceleration phase to make for easier entry into the second pipe. This velocity peak may appear more centrally (i.e. closer to t = 0.5) in the profile on the basis that there is now less incentive to move quickly upwards early on in the motion to level out for easy linear entry into the second pipe. However, slowing down in the second half for greater accuracy into the second pipe may mean that the velocity peak is still reached earlier than t = 0.5. The question then is what shape the two phases either side of the velocity peak will take.

Also, the acceleration and deceleration phases of the motion will have different magnitudes in terms of velocity change because if they were the same then a full height shift would have been performed and the path would be ending horizontally rather than at the 45 degree angle required. What cannot be predicted in advance are the actual values that these velocity change magnitudes for the acceleration and deceleration phases will be and what the overall velocity profile looks like.

The original hypothesis in the parallel pipes experiment of a full height shift over the interval may be modified in two simple ways for prediction in the non-parallel pipes setting.

A first and minimal modification to the original hypothesis is to use two full back-to-back velocity shifts as per those underlying the full height shift, except that they have different ratios. The hypothetical velocity profile in Figure 6.8, for example, has the velocity change from 0 to 3 during the acceleration phase and from 3 to 1.5 during the deceleration phase, giving a ratio of the two velocity change magnitudes of 2:1 in favour of the acceleration phase. This is only illustrative and should not be taken to represent a firm prediction of the specific values. The actual value of this ratio would be an empirical matter for instance.

Secondly, the original hypothesis may also be modified in another way more in line with the empirical results in the parallel pipes case. This modification anticipates that the subjects will employ low order partial scaled shifts that can be predictively fit to the residuals after an initial period of settling in when moving out of the first pipe and into free space.

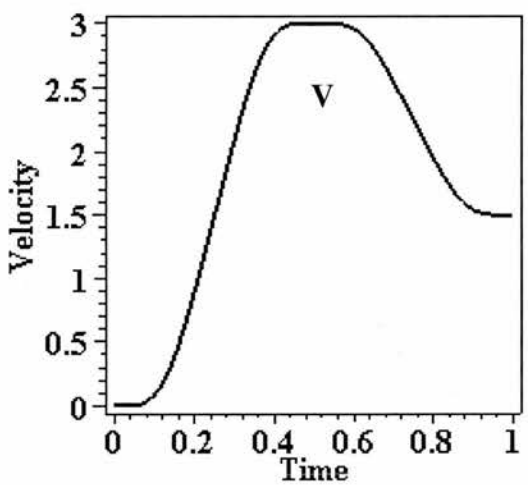


Figure 6.8. Hypothesised velocity profile (V) for Experiment 4.

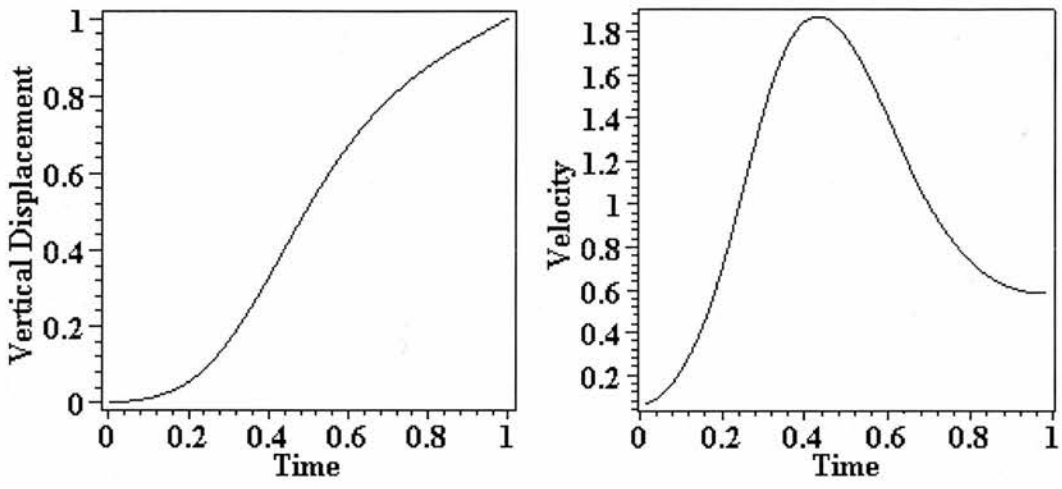
6.3.2. Method

The design was the same as that in Experiment 2 except that the second pipe was at a 45-degree angle rather than horizontal on screen as in Experiment 2 (as shown in Figure 6.2). Again six healthy subjects participated voluntarily, and there were 4 males and 2 females, with an age range of 26-50 and with all 6 subjects right-handed. The apparatus and experimental procedure were the same as in Experiment 2.

6.3.3. Results

Figure 6.9 shows a typical height and velocity profile averaged over Run 4 for one subject. The velocity profile lacks the flattened peak of the profile shown in Figure 6.8. This is counter to the first modification of the original hypothesis, but may be consistent with the second modification to the original hypothesis. This is because the latter supposes that the subjects employ single and partial rather than full and back-to-back velocity shifts over the rest of the interval after the settling in period.

The peak is more central (i.e. closer to t = 0.5) than in Experiment 2, as predicted. The ratio of the velocity change during the acceleration phase and the deceleration phase is roughly 3:2 for the subject and run shown in Figure 6.8 rather than the 2:1 used in Figure 6.8. However, this ratio is found to vary across subjects and across runs for the same subject.



a) b)

Figure 6.9. Typical kinematic profiles for a single agent in Run 4. (a) The vertical displacement versus time. (b) The velocity versus time.

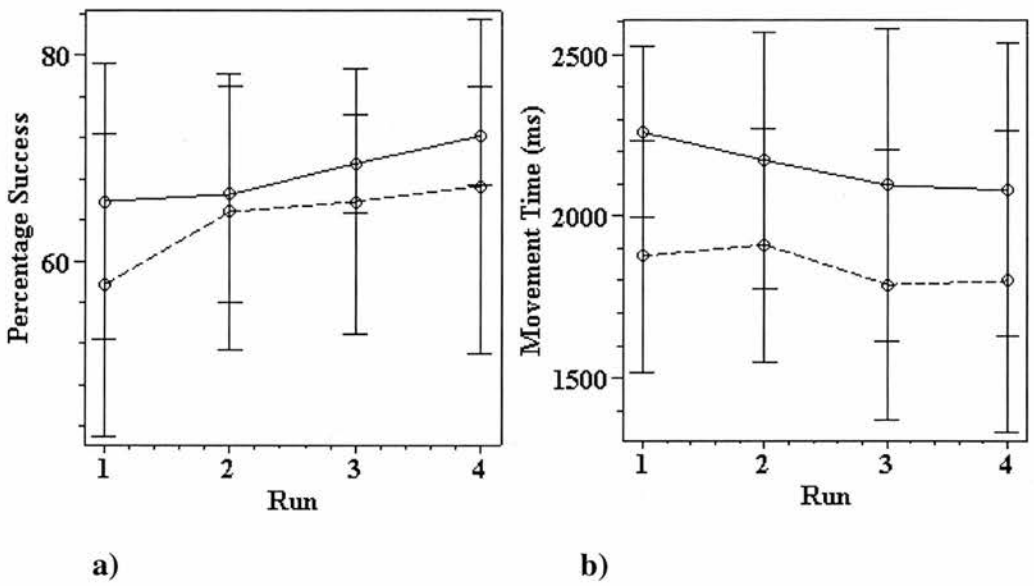
Table 6.4, below, presents the standard statistics for percentage success, movement time, peak velocity, and mean jerk squared. It is predicted that performance will improve over runs, and that peak velocity will increase over runs.

Subject	Run	Trials	Percentage success	Movement time	Peak velocity	Mean Jerk^2
4.1	1	100	42	1490.00	1.698	596.93
	4	100	60	1921.73	1.918	1322.82
4.2	1	100	61	1982.70	2.047	2784.76
	4	100	64	1844.41	2.387	4715.56
4.3.	1	100	71	2085.54	1.848	1563.04
	4	100	87	1599.36	2.089	2464.74
4.4	1	100	48	1465.67	1.651	1010.89
	4	100	42	981.38	1.696	874.74
4.5	1	100	78	2398.45	2.221	3365.99
	4	100	82	2220.52	2.215	2949.54
4.6	1	100	46	1836.78	2.335	4240.54
	4	100	69	2233.57	2.463	5951.50
Average	1	600	57.7	1876.52	1.967	2260.36
	2	600	64.8	1911.89	1.991	2395.36
	3	600	65.8	1788.23	2.082	2785.78
	4	600	67.3	1800.16	2.128	3046.48
All		2400	63.9	1844.20	2.042	2621.99

Table 6.4. Table of standard statistics (for movement time, peak velocity and mean jerk squared, only the success cases within that run are used). N.B. Underlines represent significant changes in a metric at the 0.05 significance level.

Percentage success typically increased over runs as movement times decreased indicating improved performance (see Figure 6.10). Peak velocity increased slightly over runs, as predicted (Novak et al., 2000). Mean jerk squared values were typically lower than in Experiment 4 by around 40%. This may be because the task was simply easier for the subjects in terms of the regulation of acceleration and deceleration required because they were no longer having to steepen and then flatten completely

their trajectory orientation. The 45 degree second pipe could be entered with just the steepening phase.



a) b)

Figure 6.10. a) Graph of percentage success over runs comparing Experiment 2 (solid line) and Experiment 4 (dashed line). b) Graph of movement time over runs comparing Experiment 2 (solid line) and Experiment 4 (dashed line).

The results of further statistical measures are shown in Table 6.5 below. These include LMS symmetry, K, C and the modified K statistics for the acceleration and deceleration phases of the motion. The fact that the pipes are non-parallel makes the exact values of these symmetry and shape measures empirical matters thus meaning that the same values as in the previous experiments cannot be expected to occur in these measures for Experiment 4.

Subject	Run	Trials	LMS symmetry	K	C	K (acceleration peak)	K (acceleration trough)
4.1	1	100	0.641	0.516	2.671	0.555	0.425
	4	100	0.345	0.449	3.012	0.854	0.358
4.2	1	100	1.476	0.378	3.219	0.514	0.173
	4	100	2.152	0.322	3.748	0.611	0.222
4.3.	1	100	0.403	0.358	2.906	0.556	0.254
	4	100	0.649	0.352	3.283	0.736	0.239
4.4	1	100	0.340	0.416	2.600	0.496	0.352
	4	100	0.481	0.350	2.672	0.411	0.290
4.5	1	100	0.935	0.376	3.487	0.559	0.221
	4	100	1.035	0.380	3.477	0.602	0.270
4.6	1	100	2.098	0.292	3.670	0.581	0.178
	4	100	1.862	0.303	3.865	0.882	0.172
Average	1	600	0.982	0.389	3.092	0.544	0.267
	2	600	0.940	0.378	3.129	0.604	0.274
	3	600	1.214	0.379	3.272	0.639	0.271
	4	600	1.087	0.359	3.343	0.683	0.259
All		2400	1.056	0.377	3.209	0.596	0.268

Table 6.5. Tables of LMS symmetry, K, modified K and C statistics for Runs 1 and 4 for each subject.

The Least Mean Squared symmetry measure was somewhat ironically noticeably lower in Experiment 4 than in Experiment 2 implying greater midway symmetry even though the angled second pipe meant that the path could not be symmetrical in the sense that there had to be non-zero velocity in the vertical direction heading into the second pipe (see Figure 6.9).

The K value was closer to the symmetry-indicative 0.5 mark in Experiment 4 than in Experiment 2 in Runs 1-3 but by Run 4 the K values for the two experiments were almost identical. It may be that as movement times decreased and therefore average velocities increased this introduced the need to employ a secondary aim of getting up early out of the first pipe for easier linear entry into the second pipe. This need may have manifested itself in Experiments 1, 2 and 3 in all runs because of the greater total angle the subjects were required to turn through relative to Experiment 4.

The C value was much lower than in Experiment 2. This indicates a squatter velocity profile in Experiment 4 than in Experiment 2 and is linked to the lower peak velocity and higher average velocity in Experiment 4 (where higher average velocity can be inferred from the average movement time being lower in Experiment 4 than in Experiment 2).

The Modified K statistic for the acceleration peak and trough showed similar changes in Experiment 4 as in Experiment 2. The acceleration peak K value started off fairly close to the symmetric 0.5 mark and increased over runs, thus introducing an increased asymmetry in the rightward direction. The acceleration trough value stayed fairly constant over runs but was also asymmetric with the trough leftward asymmetric.

Taylor series filtering (see Chapter 4) was performed from the quarterway, halfway and three-quarterway points for each of the four runs for each of the six subjects. The results of shift curve predictive fitting using the $V_f W / S_f$ and A / S ratios method are shown in Table 6.6(a-c) below for the width and vertical magnitude of the shift found.

Subject	Shift	Run 1	Run 2	Run 3	Run 4
4.1	Width	0.171875	0.125	0.171875	0.1875
	Magnitude	187	625	192	345
4.2	Width	No fit	0.40625	0.28125	0.296875
	Magnitude	No fit	10.5	135	115
4.3	Width	0.234375	0.28125	0.25	0.25
	Magnitude	145	83	185	250
4.4	Width	0.25	0.25	0.1875	0.25
	Magnitude	55	84	190	47
4.5	Width	0.15625	0.234375	No fit	0.1875
	Magnitude	240	-102	No fit	180
4.6	Width	0.359375	0.421875	0.3125	0.34375
	Magnitude	45	8.5	82	65

Table 6.6(a). Widths and magnitudes of shift curve sections used to fit from quarterway points for each run of each subject.

Subject	Shift	Run 1	Run 2	Run 3	Run 4
4.1	Width	0.234375	No fit	No fit	0.171875
	Magnitude	-10.3	No fit	No fit	-23
4.2	Width	No fit	0.25	0.140625	0.21875
	Magnitude	No fit	20.5	120	-16
4.3	Width	0.1875	No fit	0.171875	0.21875
	Magnitude	-23	No fit	-22	-19.5
4.4	Width	0.1875	0.15625	0.265625	No fit
	Magnitude	10.2	-3	4.7	No fit
4.5	Width	0.1875	0.21875	0.171875	0.1875
	Magnitude	-103	-26	-80	-59
4.6	Width	0.171875	No fit	0.25	0.125
	Magnitude	15	No fit	-4.5	-100

Table 6.6(b). Widths and magnitudes of shift curve sections used to fit for halfway points for each run of each subject.

Subject	Shift	Run 1	Run 2	Run 3	Run 4
4.1	Width	0.171875	0.125	0.3125	0.328125
	Magnitude	-1.3	-3.8	-0.056	-0.048
4.2	Width	No fit	0.15625	0.296875	0.203125
	Magnitude	No fit	2.25	-0.084	0.62
4.3	Width	0.203215	0.21875	0.25	0.25
	Magnitude	0.56	0.7	0.2	0.102
4.4	Width	0.1875	0.25	0.15625	0.25
	Magnitude	-1.05	0.26	1.3	-0.125
4.5	Width	0.1875	0.21875	0.25	0.265625
	Magnitude	1.65	0.43	0.46	0.168
4.6	Width	0.25	0.3125	0.109375	0.25
	Magnitude	-0.34	-0.065	1.55	0.45

Table 6.6(c). Widths and magnitudes of shift curve sections used to fit for three-quarterway points for each run of each subject.

Overall the residuals were very much of the typical shape seen before in Experiment 2 and 3 rather than the atypical shape with a dip (see Figure 6.6 for examples of both typical and atypical shapes). In fact only 9 of 72 points tested showed the atypical profile thus allowing the other 63 points to be predictively fit closely (i.e. beyond the 90% similarity threshold) by a low order partial scaled shift curve section (see Figure 6.11 for an example of such a predictive fit for one run of one subject).

As in the previous experiment the two runs from the quarterway point that did not allow predictive fits because of dips in the residuals were able to be predictively fit when the prediction was taken from slightly further into the interval at t = 3/8. So again these are cases where the subjects have taken a little more time over the settling in period before executing a low order partial scaled shift.

In summary, the results once again provide support for the generality of the modified non-analytic shift curve model. This time the generalization is to cases involving different angles between the two pipes.

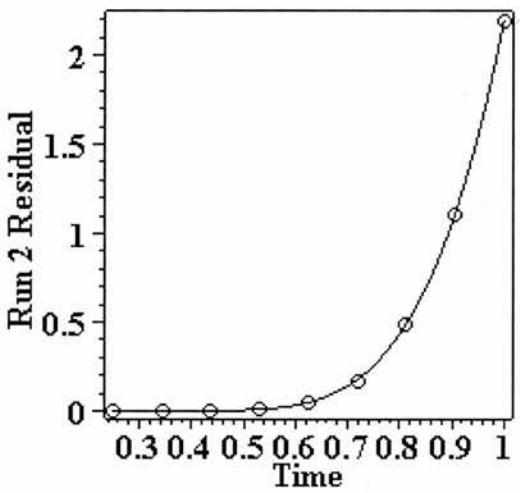


Figure 6.11. Run 2 quarterway residual (solid) for Subject 4.1 and a predictive fit to the residual (circles representing points on the full fit curve).

6.3.4. Discussion

The kinematic profiles found started out more midway symmetric initially than in previous experiments which may be because of the lowered demand for getting up early out of the first pipe for easier entry into the second pipe due to less (unsigned) angle change being required. However, as the movements became faster over runs this secondary aim may have become more influential until by the end the average K value was almost identical to that for Experiment 2.

The shape of the velocity profile did turn out to have an acceleration phase and a deceleration phase as predicted, i.e. there was an identifiable velocity peak present. The empirical fact of the matter in terms of the ratio of the magnitude of the velocity change in the acceleration phase to the magnitude of the velocity change in the deceleration phase was that it was roughly 3:2.

The first modification of the original hypothesis using full back-to-back velocity shifts was unsupported by the results. However, in contrast, the second modification was strongly supported by the results. As in Experiment 2 and 3, residuals that could be predictively fit above the 90% similarity threshold using low order partial scaled shifts were typically observed (for 63 of the 72 cases tested at the three standard bisection points of the task interval). The 9 cases of a poor fit had a good fit returned by retuning to a nearby filter origin. So, the general strategy observed in the previous experiments of subjects using an initial settling in period upon leaving the first pipe and entering free space appears to also hold true when the two pipes are non-parallel. Thus a further degree of generality has been shown to exist for the modified non-analytic shift curve model using low order partial scaled shifts instead of full height shifts.

6.4. Experiment 5: The Parallel Pipes Problem With Perturbation

6.4.1. Introduction

In real world situations, perturbations can be internal or external, i.e. they can be due to errors on the part of the subject or outside forces such as being pushed by

something. They can also be smooth or unsmooth, i.e. they may be the result of additive displacements that do not break the C^{∞} smoothness of the behaviour or they may be very sudden jerks that are perhaps best modelled as breaking with C^{∞} smoothness at some order of derivative. They can also be analytic or non-analytic, i.e. they may or may not obey Taylor series predictions in their path of influence. They can also be of varying magnitude and may take the subject closer to or further away from its goal.

Experiments 1, 2, 3 and 4 involved an environment with no external effects on the course of the subject's path. As described above though, the real world is often not free from such effects. An important generalisation then, is to show how subjects react to the introduction of external perturbations into the part of the task moving the ball between the pipes. The perturbations to be introduced are taken to be C^{∞} smooth displacements to the steered path that are beyond the agent's control and which are added dynamically during the behaviour.

Experiment 5 takes a first step towards generalizing the approach to cases where external perturbation takes place. The perturbation used is C^{∞} smooth and, with learning, can be anticipated by the subject. The smoothness and anticipation are important features in the experimental design. This is because the subjects may be able to anticipate the smooth perturbation and smoothly compensate in a non-analytic projection that continuously goes through the goal. The question and the empirical test is whether or not the subjects will actually do this, and if so, will use low order shifts as non-analytic projections in the face of (external) perturbation.

The type of smooth perturbation used in Experiment 5 is that of a negative perturbation (i.e. one that continually influences the agent such as to push them further from the goal) based on the non-analytic shift curve so as to enable it to be fully smooth. The perturbation used is shown in Figure 6.3 and, as described in Section 6.1, is half the magnitude in vertical displacement of the standard shift curve at each point (in horizontal displacement not time). This halving, combined with the fact that horizontal and vertical distances between the pipes are the same as in Experiment 3, means that the overall vertical distance travelled is the same as in Experiment 2. The fact that the perturbation is continuous means that continual redirection will be required by subjects if they are to continually keep on course for the goal. Using a non-analytic perturbation does not mean the actual path taken has to be a non-analytic outcome. It might be that subjects only redirect occasionally, for example, which would give a piecewise analytic outcome.

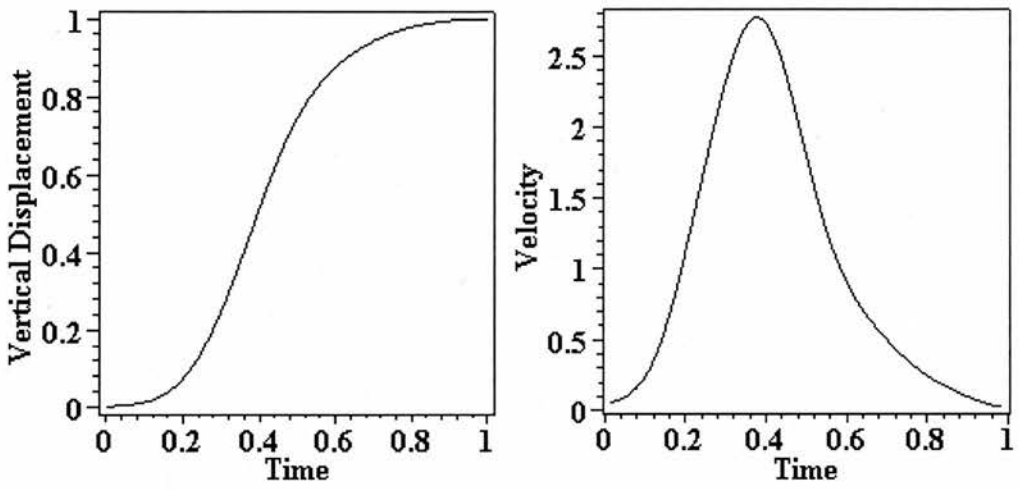
Performance is likely to be poorer in Experiment 5 than in Experiment 2 because of the greater difficulty of the task in being perturbed off course, albeit smoothly. In terms of the kinematic profiles it is by no means obvious what strategy subjects will use to overcome the perturbation. A minimal modification to the original hypothesis is that a symmetric actual path which is a full height shift is possible, given that the perturbation itself is symmetric (see Figure 6.3). A second possibility is that subjects will again tend to use an initial settling in period to get them on course for executing a low order partial scaled shift over the rest of the interval which allows them to ride the perturbation. On this hypothesis statistics similar to those in Experiments 2 and 3 would be predicted. Of course, there are other possible hypotheses as well.

6.4.2. Method

The design of the experiment was the same as that in Experiment 3, which in broad terms is the same as Experiment 1 and 2 but with a halved vertical distance between the two pipes, except that in Experiment 5 a perturbation is introduced that smoothly pushes the subject away from the goal at every point (as shown in Figure 6.2). This meant that over the course of the motion the vertical displacement travelled through was the same as in Experiments 1 and 2. Again six healthy subjects participated voluntarily, and there were 5 males and 1 female, with an age range of 21-32 and with all 6 subjects being right-handed. The apparatus and experimental procedure were the same as in Experiment 2. Subjects were not told about the perturbation order to prevent any attempts at explicitly trying to work out the perturbation shape at any stage.

6.4.3. Results

Figure 6.12 shows a typical height and velocity profile averaged over one run for one subject. Again a midway asymmetry in the velocity domain and a velocity peak can be seen as was found in Experiment 2 and 3. This would be predicted on the basis of a revised hypothesis that subjects will perform low order partial scaled shifts over the latter part of the interval after an initial settling in period as they enter free space, even in the face of external perturbation.



a) b)

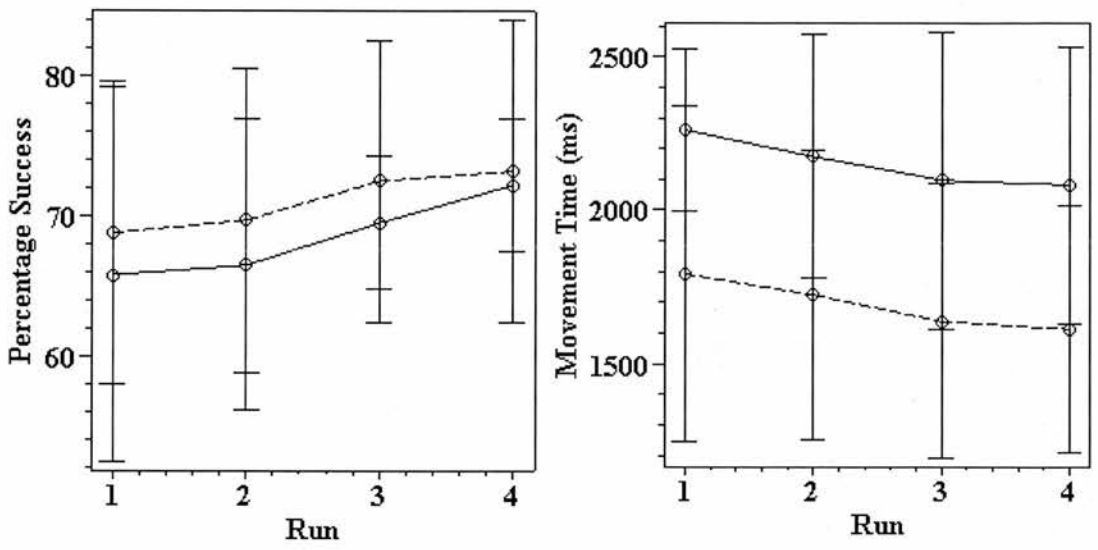
Figure 6.12. Typical kinematic profiles for a single agent in a single run. (a) The vertical displacement versus time. (b) The velocity versus time.

Table 6.7, below, presents the standard statistics for percentage success, movement time, peak velocity, and mean jerk squared.

Subject	Run	Trials	Percentage success	Movement time	Peak velocity	Mean Jerk^2
5.1	1	100	57	1253.54	2.315	2372.09
	4	100	81	1083.70	2.710	4714.97
5.2	1	100	73	1934.81	2.967	10244.62
	4	100	53	1365.47	3.109	10748.06
5.3.	1	100	69	2110.65	2.922	9408.73
	4	100	72	2015.33	3.264	13893.51
5.4	1	100	68	2190.37	2.898	10126.12
	4	100	74	2048.05	3.262	12711.52
5.5	1	100	59	974.25	2.766	5730.55
	4	100	75	1345.07	2.633	4876.39
5.6	1	100	87	2296.80	2.331	3910.11
	4	100	84	1826.63	2.609	5238.90
Average	1	600	68.8	1793.40	2.670	6965.37
	2	600	69.7	1725.02	2.743	7564.15
	3	600	72.5	1637.98	2881	8733.22
	4	600	73.2	1614.04	2.931	8697.23
All		2400	71.0	1692.61	2.814	7989.99

Table 6.7. Table of standard statistics (for movement time, peak velocity and mean jerk squared only the success cases within that run are used). N.B. Underlines represent significant changes in a metric at the 0.05 significance level.

Percentage success increased monotonically over runs whilst movement time decreased monotonically over runs showing improved performance (see Figure 6.13). Percentage success was higher in each run in Experiment 5 than in Experiment 2 and movement time was lower in each run in Experiment 5 than in Experiment 2. This may be because the increased difficulty of the task kept the subjects motivated and alert for longer.



a) b)

Figure 6.13. a) Graph of percentage success over runs comparing Experiment 2 (solid line) and Experiment 5 (dashed line). b) Graph of movement time over runs comparing Experiment 2 (solid line) and Experiment 5 (dashed line).

Peak velocity was higher for each run of Experiment 4 than for Experiment 2. It increased monotonically over runs as predicted.

In Experiment 5 the overall mean jerk squared is roughly double that found in Experiment 2. This may imply that to achieve redirection in the face of perturbation the subjects are actively regulating their accelerations and decelerations far more than in Experiment 2.

The results of further statistical measures are shown in Table 6.8 below. These include LMS symmetry, K, C and the modified K statistics for the acceleration and deceleration phases of the motion.

Subject	Run	Trials	LMS symmetry	K	C	K (acceleration peak)	K (acceleration trough)
5.1	1	100	0.805	0.407	3.621	0.573	0.289
	4	100	1.767	0.379	4.239	0.894	0.197
5.2	1	100	2.295	0.353	4.639	0.639	0.183
	4	100	6.198	0.267	4.873	0.425	0.169
5.3.	1	100	3.827	0.316	4.576	0.610	0.174
	4	100	5.647	0.300	5.110	0.639	0.158
5.4	1	100	1.613	0.389	4.530	0.632	0.259
	4	100	1.796	0.397	5.097	0.752	0.204
5.5	1	100	2.148	0.364	4.324	0.616	0.207
	4	100	1.467	0.386	4.118	0.803	0.242
5.6	1	100	3.484	0.303	3.658	0.485	0.289
	4	100	3.765	0.305	4.090	0.555	0.206
Average	1	600	2.362	0.355	4.225	0.593	0.234
	2	600	2.805	0.350	4.293	0.605	0.237
	3	600	3.112	0.343	4.508	0.656	0.206
	4	600	3.440	0.339	4.588	0.678	0.196
All		2400	2.930	0.347	4.403	0.633	0.218

Table 6.8. Tables of LMS symmetry, K, modified K and C statistics for Runs 1 and 4 for each subject.

The LMS symmetry measure monotonically increased over runs, implying decreased midway symmetry whilst the K value decreased monotonically over runs and away from the K = 0.5, symmetry-indicative, value. So a similar pattern is emerging as in Experiment 2 and 3 with subjects moving away from midway symmetry as they implement their secondary aim of getting up early out of the first pipe to then flatten out early for easy linear entry into the second pipe.

In Experiment 5 the C value increases monotonically over runs as predicted and is higher for each run than in Experiment 2. This implies that, although the peak velocity and the average velocity are both increasing, the peak velocity is increasing more, giving rise to a less squat velocity profile.

The Modified K values for the acceleration and deceleration phases of the motion show even greater asymmetries than in Experiment 2 for each run. In the case of the acceleration peak this is an increasing rightward asymmetry and in the case of the acceleration trough a leftward asymmetry.

Taylor series filtering (see Chapter 4) was again performed from the quarterway, halfway and three-quarterway points for each of the four runs for each of the six subjects. The results of shift curve predictive fitting using the V_f W / S_f and A / S ratios method are shown in Table 6.9(a-c) below for the width and vertical magnitude of the shift found.

Subject	Shift	Run 1	Run 2	Run 3	Run 4
5.1	Width	0.203125	0.171875	0.171875	0.15625
	Magnitude	340	460	620	1020
5.2	Width	0.234375	0.28125	No fit	No fit
	Magnitude	510	245	No fit	No fit
5.3	Width	0.265625	0.328125	0.28125	0.296875
	Magnitude	350	142	300	154
5.4	Width	0.21875	0.203125	0.203125	0.140625
	Magnitude	610	700	780	2750
5.5	Width	0.21875	0.21875	0.21875	0.21875
	Magnitude	670	357	540	530
5.6	Width	0.328125	0.296875	0.265625	0.328125
	Magnitude	37	75	350	36

Table 6.9(a). Widths and magnitudes of shift curve sections used to fit from quarterway points for each run of each subject.

Subject	Shift	Run 1	Run 2	Run 3	Run 4
5.1	Width	0.171875	0.1875	0.1875	0.21875
	Magnitude	-120	-32	-51	-88
5.2	Width	0.21875	No fit	0.203125	No fit
	Magnitude	-60	No fit	43	No fit
5.3	Width	0.234375	0.15625	0.265625	0.4375
	Magnitude	-38	-35	-2.3	1
5.4	Width	0.28125	0.203125	0.1875	0.203125
	Magnitude	-20	-90	-131	-110
5.5	Width	0.21875	0.234375	0.21875	0.21875
	Magnitude	-50	-13	-33	-23
5.6	Width	0.28125	0.140625	0.171875	0.140625
	Magnitude	-0.99	-75	-12	54

Table 6.9(b). Widths and magnitudes of shift curve sections used to fit for halfway points for each run of each subject.

Subject	Shift	Run 1	Run 2	Run 3	Run 4
5.1	Width	0.328125	No fit	No fit	0.21875
	Magnitude	0.03	No fit	No fit	-0.3
5.2	Width	0.0625	No fit	0.34375	0.296875
	Magnitude	138	No fit	0.047	0.065
5.3	Width	0.25	No fit	No fit	0.3125
	Magnitude	0.175	No fit	No fit	-0.024
5.4	Width	0.1875	No fit	0.3125	0.28125
	Magnitude	2.15	No fit	0.092	-0.24
5.5	Width	0.203125	0.25	0.3125	0.28125
	Magnitude	0.6	-0.07	-0.054	-0.11
5.6	Width	0.109375	0.1875	No fit	No fit
	Magnitude	4.75	0.85	No fit	No fit

Table 6.9(c). Widths and magnitudes of shift curve sections used to fit for three-quarterway points for each run of each subject.

Overall the residuals were very much of the typical shape seen before in Experiment 2 and 3 rather than the atypical shape with a dip (see Figure 6.6 for examples of both typical and atypical shapes). In Experiment 5, 12 of the 72 points tested showed the atypical profile whilst the other 60 points could be predictively fit closely (i.e. beyond the 90% similarity threshold) by a low order partial scaled shift curve section (see Figure 6.14 for an example of such a predictive fit for one run of one subject). Eight

of the no fit cases were from the three-quarterway point, which ties in with the findings of Experiment 2 and 3 that predictions from this point are more unreliable because sometimes they lack necessary higher order terms.

As in Experiments 3 and 4 the runs from the quarterway point that did not allow predictive fits (two cases in this experiment) because of dips in the residuals were able to be predictively fit when the prediction was taken from slightly further in to the interval at t = 3/8. So, as before, a little more time has been used for settling in by the subjects before executing a low order partial scaled shift.

Overall, the results once again provide support for the generality of the modified non-analytic shift curve model. In this experiment it has been shown that subjects are able to learn how to smoothly ride smooth external perturbations so that they can employ predictable low order partial scaled shifts.

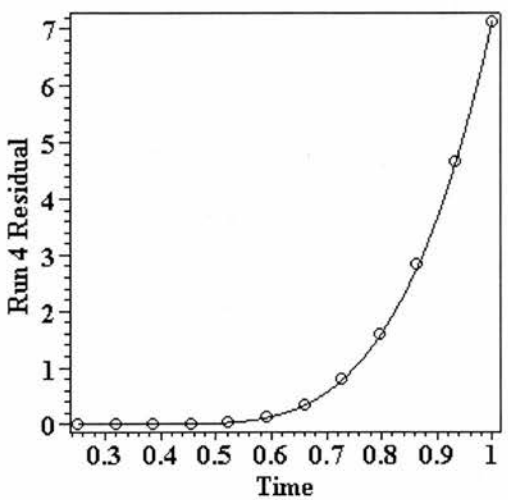


Figure 6.14. Run 4 quarterway residual (solid) for Subject 5.1 and predictive fit to the residual (circles representing points on the full fit curve).

6.4.4. Discussion

Overall the results showed reasonable similarity to those of Experiment 2, although, because the subjects were faced with a more difficult task they may have been motivated and alert for longer. This may have accentuated the effects of their secondary aim for getting up early out of the first pipe and flattening out for easier entry into the second pipe.

The doubling of the mean jerk squared value in comparison to Experiment 2 seems particularly to indicate that subjects may be actively regulating their movement far more often to remain on course for the goal.

As in Experiment 2 and 3 residuals that could be predictively fit above the 90% similarity threshold using low order partial scaled shifts were typically observed (at 60 of 72 points tested). Again the general strategy observed in the earlier experiments of subjects using an initial settling in period upon leaving the first pipe and entering free space before making such a shift appears to be present (although again they occasionally use slightly more than one quarter of the interval for settling in). The recurrence of the strategy is in line with the non-analytic state-goal model as described in Chapter 4. It also means that an extra degree of generality has again been achieved for the modified non-analytic shift curve model using low order partial scaled shifts, in that it can be predictive in tasks where subjects are forced to constantly learn about and adapt to an external perturbation. The answers to the key questions raised in 6.1 about Experiment 5 are that non-analytic projection through

the goal does indeed appear to be continuous even with external (C^{∞} smooth) perturbation, and is put together using similar shifts to those without perturbation.

6.5. Conclusion

The previous chapter looked at experiments aimed at investigating the basic unit of non-analytic shift present in C^{∞} smooth goal-directed behaviour. It was found that the typically low order partial scaled shifts were employed by subjects rather than the full height shifts originally hypothesised. This chapter attempted to generalize the modified version of the model to three new experimental settings.

Experiment 3 investigated scaling of the motion by halving the vertical distance between the central exit point of the first pipe and the central entrance point to the second pipe. Experiment 4 attempted generalization to cases involving non-parallel pipes. Experiment 5 then looked at the effects of adding a C^{∞} smooth non-analytic perturbation over the course of the motion.

In all three cases predictive fits to residuals were typically achieved (beyond the 90% similarity threshold) using low order partial scaled shifts over the remainder of the interval after an initial settling in period of internal perturbation as the movement changed from obeying the constraints inside the first pipe to those of moving up through free space to aim for successful entry into the second pipe.

The main conclusion that follows is that the modified model using partial scaled shifts can be used for predictive generalization to cases involving other scalings of

horizontal and vertical distance between pipes, to cases involving non-parallel pipes and to cases where subjects must rapidly adapt to external perturbations to their path.

Chapter 7

Computing Generalization and Applications

7.1. Introduction

In the previous chapters a computational approach to the study of C^{∞} smooth goal-directed behaviour has been developed. The novel non-analytic theory and competing theories were used to generate hypotheses for the empirical studies presented in chapters 5 and 6. These experiments were then analysed using the new tools of analysis that were necessary for identifying non-analytic components in the behaviour. The empirical studies focussed on deciding between alternative models in C^{∞} smooth human hand movement kinematics experiments using a mouse to steer a ball between two pipes.

The initial experiment was limited to a movement with a 1:1 height to width ratio between the two pipes and therefore generalizations were required to demonstrate the further applicability of the new approach. The beginnings of this generalization

process were described in the experiments presented in Chapter 6, involving generalization to new height to width ratios between the two pipes, new angles between the two pipes and the introduction of external perturbation to the motion.

In Chapter 7 further computational settings that the approach can potentially be generalized to in the near or further future are described. This chapter also describes how the new method may be an improvement on existing methods and provides a description of applications that may be possible as a result of incorporating the general principles behind the predictive ability to provide new functions that can be used by artificially animate agents to drive and generate their goal-directed behaviour, e.g. a robot moving its arm to grasp a moving object.

7.2. Further Generalizations

The approach described in earlier chapters marked the difference between methods that are able to identify non-analyticity at work in C^{∞} smooth goal-directed behaviour and not being able to do so. It also marked the difference between methods allowing fully smooth merges between two derivative states and methods not enabling such merges. The approach was applied to mouse movement in human kinematics.

Whilst such an approach represents a significant and vital first step, common to the development of any novel computational application, there are inevitably limitations to be removed to take things further.

The approach was initially shown to allow for the identification of an s-shaped non-analytic shift curve that would serve as the basic unit for shifts in kinematic behaviour. The approach was limited in practice. At first the shift shape was only established between parallel pipes at a given distance apart (see Chapter 5 for the results of Experiments 1 and 2 to this end). Generalization was then achieved in terms of allowing variable ratios of height to width between the two pipes, for variable angles between the two pipes, and for variable smooth perturbation based on the s-shaped non-analytic shift curve over the full interval of the shift between the two pipes (see Chapter 6 for the results of Experiments 3, 4 and 5 for each of these three generalizations, respectively).

Further generalizations that could readily be made would include allowing all pipe shapes, all perturbation types, dynamic goals, multiple sequential goals, and the removal of pipe and curvature constraints and also to cases beyond mouse movements to other C^{∞} smooth movements.

The ultimate generalization would be to predict everyday real world animate behaviour without experimental constraint. This, of course, would be very difficult at present precisely because the real world is largely unconstrained and uncontrolled with many interactions and many independent variables. The measurement of unfettered real world behaviour is a difficulty for the application of scientific principles generally, rather than shift theory in particular. Such generalization is not expected of our best contemporary physics for inanimate behaviour, for instance. This is not to say that shift theory would not apply to such real world behaviour, rather that it would be very difficult to achieve precise prediction.

In the rest of this chapter each near and further future generalization is, in turn, given more detail. The computing and cognitive applications that may result are then explored.

7.2.1. All Pipe Shapes

In the experimental settings in chapters 5 and 6 only linear pipes are used. This means that the agents always have zero derivative values in all orders with respect to the linear axis of the pipe that they are passing through.

In the pipes experiments this was useful in encouraging zero derivative values in all such orders at the exit from the first pipe and the entrance to the second pipe. This allowed deviation from and to the most measurable analytic condition, the zero derivative state, to be most readily and reliably observed over a broad spectrum of experiments, including angled pipes and external perturbation.

In the near future, a generalization still within the pipes paradigm would be to allow pipes of non-linear shapes such that they allow for non-zero higher order derivatives at the ends of the shift corresponding to those of the pipe shape relative to a given linear axis. For example, a linear first pipe might be paired with a quadratic or a cubic second pipe.

Although the partial scaled shifts observed demonstrate a large variety of derivative state merges, the use of non-linear shapes would demonstrate that the new method

could, in practice, analyse fully smooth and predictable merges of any pair of measurable derivative states (something that is not possible using analytic methods) between pipes without zero derivative states being involved.

7.2.2. Beyond Pipes

Presently the experiments performed so far have all involved movement between two pipes. This was useful in helping to lead the subjects into a particular externally constrained path at the two ends of the behaviour.

However, in the real world, cases of hand/limb movement under such constraints are only sometimes present, e.g. in a gearshift, and in many cases are not present. Thus a future generalization would be to analyse smooth merges between measurable start and end derivative states without pipes constraining the values of those derivative states (again this is not possible using analytic approaches).

Another generalization would be to extend investigation to n-dimensional problems (such as the gait analysis described later) beyond the two dimensions of the pipes problem.

7.2.3. All Perturbation Types

In Experiment 5 (see Chapter 6) the investigation involved an external perturbation to the motion, i.e. a displacement to the motion not under the control of the agent. The

only perturbation type used was a vertical displacement shift of a fixed size based on the non-analytic shift curve (see Chapter 4).

However, perturbations could be of other magnitudes or displacement directions, they could be analytic or of other non-analytic shapes, and they could also be unsmooth. A future species of generalization would be to investigate the ability of agents to cope with these other types of perturbation.

Perturbation is a relatively small issue in well-structured environments, i.e. those environments that Russell & Norvig (2003, pp.40-42) describe as fully observable, deterministic, episodic, static and discrete. This may be because in such environments, the outcome of action closely approximates the intended outcome. In the ill-structured environments often encountered in the real world, on the other hand, such as real world scenarios that are novel for the agent, intended outcome and action are often more loosely coupled. Greater perturbation is the result. This is why shifts without external triggers like pipes are likely to occur.

If the plasticity and persistence in the face of perturbation provided by the new non-analytic theory has extension to various perturbation types then this could provide improved results for successful action by artificial agents in handling perturbation in ill-structured environments.

7.2.4. Dynamic Goals

In the five experiments presented in Chapters 5 and 6 the second pipe is always static with respect to the first pipe during the course of a particular trial. In the real world, static targets are present in many areas of life such as reaching for a glass from a table or hitting a golf ball from the tee. However, there are also many occasions where targets are moving, e.g. when catching a cricket ball or swatting a fly.

A significant near future advance would be to introduce moving second pipes to investigate the effects of dynamic goals on agent behaviour. This would allow the demonstration of the point made in Section 3.6 that the purely analytic Model 1 is unlikely to be viable in cases involving goal movement that cannot be anticipated by the agent prior to the change of goal. This is because the agent would be committed to a unique analytic path early on that is unlikely to pass through later versions of the goal. The success often observed of natural agents attaining moving goals, in practice, suggests that a non-analytic component is present.

Beyond the motivation of better modelling, moving goals are also an important aspect of many real world environments and would thus improve the generality of a successful model. By using the shift approach, which allows for plastic redirection (which cannot be done by existing methods), improved analysis in dynamic goal cases may be possible.

7.2.5. Multiple Sequential Goals

In the experiments described earlier only behaviours with single goals set by the experimenter were considered, i.e. those that could be considered in isolation of other

goals. The subjects make it through the second pipe and the experimental goal is satisfied. This isolation is present in various real world scenarios. Turning the pages of a book, for example, can be considered as a series of independent events because between each page turn there is a rest for reading to take place. Similarly, in the earlier experiments (see chapters 5 and 6) each trial is performed in isolation of the others.

However, in the real world, paths for several goals may be required to be coordinated evolve smoothly into each other. So another near future advance would be to generalize to tasks involving multiple goals where the paths for the consecutive goals are required to evolve smoothly into each other such as a skier traversing a series of slalom gates.

In allowing for fully smooth merges the shift approach provides an account for the way multiple sequential goals are achieved that may enhance existing methods. For example, when an agent takes a particular path to its next goal it may be that, in striving for this goal, the agent finds itself on a poor path for the subsequent goal. Being able to employ C^{∞} smooth non-analytic shifts would allow the agent a way of correcting for the subsequent goal whilst retaining full smoothness, unlike analytic or semi-analytic variation.

7.2.6. State-Space Kinematics In n-Dimensions

The kinematic analysis of the pipes experiments used 2 dimensions, i.e. the vertical displacement of the mouse and the time dimension. Data from a third dimension, the

width, was also available though, thus allowing analysis of a three-dimensional state-space as shown in Figure 7.1.

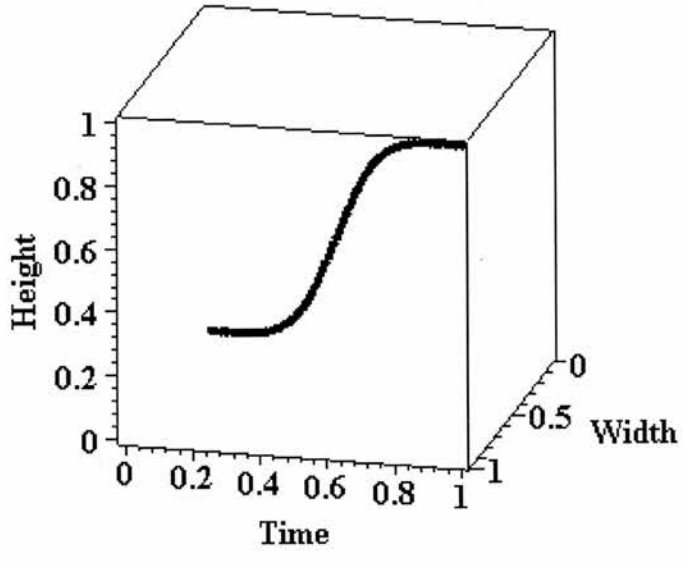


Figure 7.1. A three-dimensional height-width-time portrait of a smoothed actual single trial.

Many interesting movements also occur in systems where more than 1 or 2 dimensions as well as time must be taken into account. In gait analysis, for example, sensors are placed on multiple limb components which are analysed together as a multi-dimensional system changing over time (Magill, 1998; Kocsis et al., 1999; Morris & Paradiso, 2003).

Each independent variable, i.e. each limb component, may have its variation plotted against time and analysed for the presence of shifts. The coordination of actions in the n-D phase portrait may have simultaneously coordinated goals and shifts.

7.3. Potential Applications

7.3.1. Introduction

There are various potential applications that may arise from the novel approach developed in the thesis. This section discusses potential computational applications such as in robotics, computer animation, and human-computer interaction. It then goes on to look at the potential for applying the new methods to cognitive analysis such as the study of golf swing kinematics in sports science, for example. Each area is examined to see if it could benefit from fully smooth rule-following behaviour with low frequency undulation.

7.3.2. Computing Applications

7.3.2.1. Powered Mouse Movement

Present mouse drivers allow for accelerated movement but not in a way that incorporates extended non-analytic prediction of future paths.

A more informed mouse driver would be able to predict the goal location on the screen for the user using the current non-analytic projection and provide more directed acceleration in realising that goal state, e.g. arriving at a button, an icon or a place in a file window.

Powered mouse movement using the non-analytic goal prediction mechanism has the potential to decrease the time and movement required to achieve on-screen goals. This in turn may improve productivity and decrease strain injuries.

7.3.2.2. Computer Animation

Films such as the Lord Of The Rings trilogy use AI to drive the behaviour of artificially animated agents, e.g. orcs, elves, dwarves, hobbits and humans. Shift theory would allow these creatures to move using current state and goal state information to direct themselves fully smoothly using low order non-analytic shifts.

Providing the characters with non-analytic C^{∞} smooth goal-directed components to their behaviour would allow them greater plasticity and persistence in their behaviour in a more automatic and fully smooth manner. This could make for more realistic animation.

7.3.2.3. Artificial Control Theory

Just as human kinematics researchers have been constrained to use analytic functions for their analysis and design toolkit, so too for designers of artificial control systems. Fully smooth behaviour can be designed using analytic functions, but the role of any goals directing the system behaviour is then different from that in the thesis.

In artificial control theory, the goal may be an attractor of some sort such as a command signal (Rosenblueth, Wiener & Bigelow, 1943; Weir, 1984; Brogan, 1991;

Dorf, 1992; Lewis, 1992). In common with the pipes experiments, it is worth describing the role of the goal in present artificial systems whilst a single fixed goal is being attempted. The goal takes the role of an input parameter in an artificial state-determined system with input. As such, the goal has the ability to *select* the phase portrait *a priori*, the family of trajectories that the system will follow, with the goal as an attractor in the portrait.

The trajectories in smooth phase portraits for artificial control systems presently conform to being state-determined and analytic functions by design. The goal is therefore able to influence, *a priori*, what shape of analytic trajectories the system follows, but not their shape during behaviour or the fact that they are analytic.

The role of the goal in the theory developed in the thesis differs in being supposed to induce non-analytic shapes of trajectory in the event during the behaviour of the goal's derivative state not lying on the analytic projection of the current derivative state. The role here is of the goal being more of a dynamic portrait developer rather than a static portrait selector.

The relevance of the difference for application in artificial control theory is that the non-analytic plasticity seen in human control may be usefully transferred to artificial control. This may be done by adopting non-analytic shifts for use in the latter analysis and changing the role of the goal and design toolkit. In particular, if the best efforts of the standard analytic artificial control systems are in practice producing an analytic path not projecting through the goal due to perturbation or otherwise, non-analytic shifts may be applied to project non-analytically through the goal.

This may be of particular benefit to roboticists. Robots have to face from the problem that, in real world scenarios, hardware implementation of software instructions is often imperfect. The outcome of their actions often differs from the intended outcome. In various circumstances, their paths may consequently be perturbed away from the goal.

It could be that cases of action error leading to goal failure could be avoided to begin with or else corrected for on the fly if robots were provided with the extra level of plasticity and persistence offered by non-analytic shift theory.

7.3.3. Cognitive Applications

At the moment there is no methodology for psychologists to identify clearly the non-analytic aspects of C^{∞} smooth goal-directed movements on a day-to-day basis. Developed suitably, the use of the principles and methods provided in the thesis for identifying such behaviour could provide psychologists with a useful tool for enabling both analytic and non-analytic components of C^{∞} smooth goal-directed movements to be analysed.

For instance, shift analysis may be used to affirm or disaffirm cognitive elements in behaviour. In the pipes experiments, for example, the start of the single shift part of the way into the space between the pipes could be taken to suggest two cognitive phases to the behaviour. Specifically, the agent could be hypothesised to take aim for

entry into the second pipe in the early part of the space and then fire, i.e. execute a final shift as a type of ballistic movement.

In the two pipes experiment then, the aiming phase is supposed to occur from the exit of the first pipe, through the settling in period, and up to the quarterway point (or in some cases just after the quarterway point). At the end of the aiming phase the subjects have got themselves into position to project a low order partial scaled shift over the rest of the interval and into the second pipe, which they duly follow.

Another way the new methods could help cognitive analysis is in allowing improved analysis of the influence of reinforcement learning effects in the development of skilled actions. In the pipes experiments, for example, subjects may produce regular aiming and firing phases in their behaviour through the process of reinforcement learning, even in the face of external perturbation. The ability to produce the equivalent of shift combinations may be inbuilt with reinforcement learning then smoothing variation between attempts towards a regularly successful primary low order movement.

These points about aiming, firing, and reinforcement learning could not be detected or made without shift analysis.

By allowing the identification of aiming and firing phases of the behaviour this could be helpful in areas of sports science such as golf kinematics studies, for example. Many researchers are currently interested in the kinematic analysis of golf swings (for example, see: Neal, Abernethy, Moran & Parker, 1990; Dillman & Lange, 1994;

Nesbitt, Cole, Hartzell, Oglesby & Radich, 1994; McLaughlin & Best, 1994; Lee, Erickson & Cherveny, 2002). The shapes of movement and the derivative states found in a golf swing are highly constrained and may therefore be comparable to those found in the pipes experiment. Such research may benefit from the new methods through the generation of new hypotheses about the expected kinematics, such as the non-analytic symmetries, asymmetries and shapes that are likely to be present in the derivative profiles under different conditions. In Experiment 1, for example (see Chapter 5) asymmetry resulted from accuracy and speed influences. There may be other sorts of cognitive trade-off that could be examined through shifts. For example, in different golf shots the player may be concentrating on power or on accuracy or on attempting to achieve some trade-off between the two. A drive down a wide fairway is about generating as much distance through power hitting as possible without worrying too much about the exact positioning of the ball on the fairway whereas a delicate chip on to a green is about holding back on the power and achieving as much accuracy as possible.

Generalization of the novel method would also allow the analysis of other sporting behaviours that involve sequences of movements that evolve into each other (such as in gymnastics) or plastic redirection in response to goal movement (such as remaining on course to make a tackle on a swerving rugby player) or being knocked off course (such as a yachtsman being hit by a gust of wind).

Ultimately the sports players themselves could benefit from such advances in analysis. Firstly, the new tools could be used to analyse performance and compare it to a simulation of the ideal performance. The sports player could then be trained to

produce something more like the ideal performance, bringing his or her golf handicap down, for example. Such analysis may also help in discovering the causes of sporting injuries and allow for the identification of changes to performance that could potentially minimise the risk of injuries to the player.

7.4. Conclusion

The new methods presented in the thesis provide an approach that allows for the identification of non-analytic components in C^{∞} smooth goal-directed behaviour and for smooth merges between derivative states whereas the previously existing methods do not.

With further generalization the methods developed in the earlier chapters may be used in many areas that involve C^{∞} smooth goal-directed behaviour to allow fully smooth rule-following with low frequency undulation.

Specific examples of where this may be useful in the real world are in robotics, computer animation, human-computer interaction, and computational psychology. It may be used in golf kinematics studies, for example, for the generation of hypotheses and the analysis of performance, potentially leading to improved performance and decreased injuries.

Epilogue

One of the central problems in the understanding of natural agents and in the engineering of artificially intelligent agents in their C^{∞} smooth goal-directed behaviour concerns the appropriate level and type of plastic and persistent adaptation in the behaviour of such agents in the pursuit of their goals.

The predominant determinate system is the state-determined system where, after taking input into account, the current state is the single element that determines the behaviour. Such state-determined systems have been able to demonstrate plastic and persistent adaptation up to a point in the sense that they can follow pre-programmed instructions for what to do under given circumstances and use state-based feedback to remain on course for a goal in the event of error. These properties allow such systems to set courses that appear to be good *a priori* and to be able to correct their path dynamically relative to that of a system without feedback. However, if the state-determined path, the analytic path in smooth behaviour, projects away from the goal during the behaviour, there is no further corrective mechanism. In this sense, the system is unable to correct relative to itself when its path is going off course analytically. It was argued in the thesis that plastic and persistent adaptation beyond

that assumed by state-determined systems is conceivable, in particular where non-analytic paths are possible.

Chapter 1 described the historical reasons behind this. In particular it showed that, although the notion of final cause as a distinctive effective influence on the behaviour has been lost from science, it was not lost through falsification (Woodfield, 1976, pp.8-9). It was lost, rather, by default, because of the success of approaches based purely on efficient causal accounts of the forces and mechanisms.

Chapter 2 then went on to describe the present day attitudes to causality in goal-directed behaviour research that have resulted from the falling by the wayside of final causal accounts in science. What is clear is that many areas that currently only use state-determined computing methods could potentially benefit from the reintroduction of final cause methods that provide an extra layer of plastic and persistent adaptation beyond that provided by their efficient cause based methods.

One of the major themes of the thesis was to form and investigate computational principles for such adaptation. To do this it was necessary to develop methods capable of identifying and analysing non-analytic components in C^{∞} smooth goal-directed behaviour if such were present.

Chapter 3 described a computational framework for C^{∞} smooth goal-directed behaviour. This framework involved the use of a continuous state space approach which introduced the idea that smooth inanimate and animate motion can be distinguished by the ability of the latter to involve non-analytic path characteristics in

which (non-analytic) redirection of a path occurs such that the agent of the behaviour remains on course for a goal. An important theme here is that non-analytic projections can merge fully smoothly with whatever curve has already been traversed so as to (non-analytically) redirect through the goal, unlike analytic projections.

Chapter 4 described a non-analytic computing methodology for C^{∞} smooth goal-directed behaviour. It was shown that new tools of analysis were needed in order to identify non-analytic components in the motion because the tools of (solely) efficient cause based science are concerned only with analytic or piecewise analytic components. A main aim of the thesis was to develop initial principles of goal-directed behaviour in the form of hypotheses about symmetry in kinematic profiles. A non-analytic shift curve was introduced that contained three fundamental symmetries and was hypothesised to be the basic unit of C^{∞} smooth shifts in animate behaviour. In computing terms, the shift is a type of smooth switch enabling agents to fully smoothly merge one path with another. A method of analysis using the principle of Taylor series filtering was introduced which was designed to distinguish between analytic, piecewise analytic, and non-analytic shift curve possibilities.

With a computational framework and a non-analytic computing methodology having been constructed for C^{∞} smooth goal-directed behaviour, a second theme of the thesis could then be explored. The extra layer of plastic and persistent adaptation provided by the novel approach was demonstrated in human C^{∞} smooth goal-directed behaviour in hand movement kinematics experiments using a mouse.

Chapter 5 described experiments aimed at isolating non-analyticity in a shift that could be used as a basic unit. These experiments involved the movement of a ball between and through two parallel pipes with a one to one height to width ratio between them. It was initially hypothesised that the basic non-analytic shift curve would be used by the subjects to traverse the path between the two pipes. This hypothesis was found to be incorrect in Experiment 1 (where free movement was allowed between the pipes) and in Experiment 2 (where a localised curvature constraint was introduced). However, although the agents did not use the full shift, the shift and its symmetries nevertheless formed a common basis for adaptation in providing a regular curve. It was found that the subjects used a strategy that may be described as a form of aiming (over roughly the first quarter of the interval) and firing (over the remainder of the interval). The firing phase could be predictively fit using low order partial scaled shift curve sections. The non-analytic shift curve model was thereby found to have the greatest support of the alternative models, with the modification that partial scaled height shifts are used rather than the full 1:1 height shifts originally hypothesised. This model allows accurate prediction of behaviour from the current state to the goal over significant intervals using the measurable derivatives unlike its rivals.

Chapter 6 described experiments aimed at extending the approach to new computational settings involving a different width to height ratio for the central region, a second pipe at an angle to the linear axis of the first pipe, and a smooth external perturbation to the motion. It was found that the presence of a predictive fit to the path during the firing phase of the behaviour, after an initial period of settling in, using a low order partial scaled shift curve was robust in the face of each new

condition. Thus the modified model using partial scaled shifts was shown to be extendable for predictive generalization to cases involving other scalings of horizontal and vertical distance between pipes, to cases involving non-parallel pipes, and to cases where subjects must continually adapt to smooth external perturbations to their path.

Because they are computational, such principles could also potentially be used in scientific and engineering solutions to a wider range of problems of C^{∞} smooth goal-directed behaviour. Chapter 7 described generalizations and possible applications of the new non-analytic method and why they could provide improvements on existing solutions.

As far as generalizations are concerned, near future work is likely to be kept within the realm of the pipes paradigm for the sake of control and providing clear confirmation of hypotheses about principles but with generalization to different pipe shapes, dynamic goals and different types of perturbation. Such work would begin to allow more interesting and complex experimental settings and an increasing number of computational problems to be solved.

Beyond this, behaviour without pipe constraints and with multiple sequential goals would be investigated. A further important generalization would be to go from the 2 dimensions analysed in the pipes experiments to the n dimensions of more complex kinematic settings. It may be that the coordination of actions in the n-D phase portrait has simultaneously coordinated goals and shifts.

In Chapter 7 it was described how the novel approach may be usefully applied in various areas of computing such as robotics, computer animation and human-computer interaction. The novel approach may also have cognitive application such as in the efforts of sports scientists to understand the kinematic structure of C^{∞} smooth goal-directed movements, potentially leading to improved performance and decreased injuries.

The elements used in the determinate modelling of goal-directed behaviour in the thesis are both similar and different to that used in the more standard state-determined system. The similarity is that the current state appears in both types of determinacy. The difference is that a goal state is an additional element in the thesis method relative to the standard system. The thesis has used the current derivative state as the efficient cause, and the goal derivative state as the final cause. The resultant model has demonstrated successful prediction of the path shape between the states using measurable derivatives. In doing this, the thesis overall provides the first steps towards showing that final causal explanation and mechanism can be usefully introduced through computer science as part of a hybrid account combined with efficient causal explanation and mechanism.

Appendix A

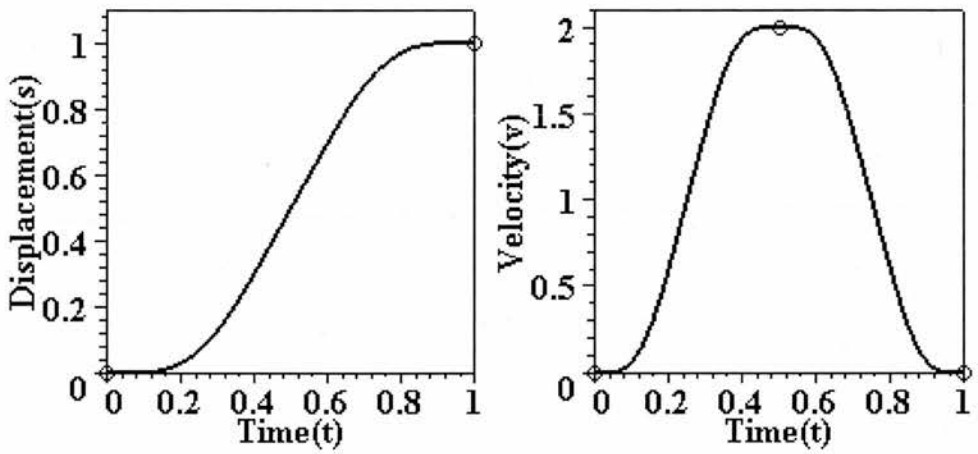
This appendix aims to provide details of how the shift curve used in the thesis is defined exactly.

It does this by first of all giving a graphical account of the inherent symmetries in the curve across the zeroth, first, second and third derivative orders and beyond. These symmetries are incorporated into two key differential equations that define an exact solution in conjunction with the basic requirements of an arbitrary s-shape. The differential equations are converted into difference equations for finite sets of sample points with semi-linear interpolation. An algorithm is developed in two stages to enable the construction of arbitrarily close semi-linear approximations to the underlying unique, continuous, and fully smooth curve that can be used for calculation purposes.

The details that follow are taken from a technical report (Weir, 2005).

The symmetry requirement suggests an s-shape for the shift function (Figure A1(a)). This has rotational symmetry about its mid-point. Furthermore, the rotational

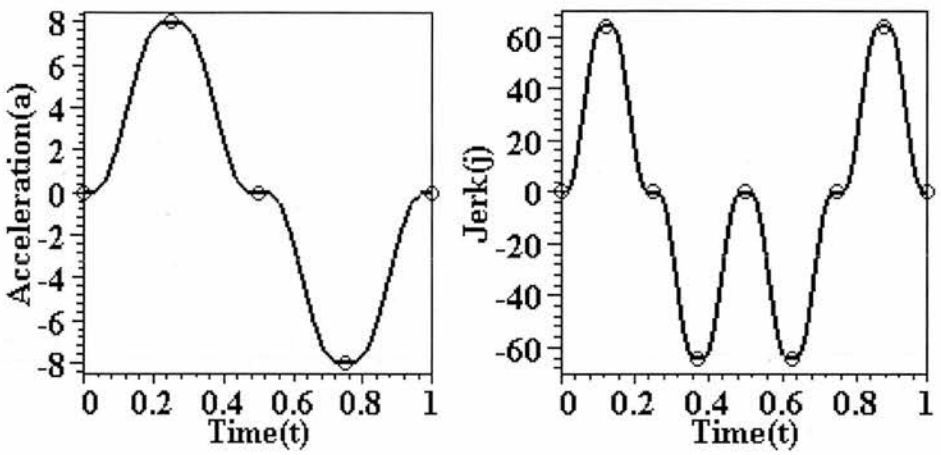
symmetry leads to an additional reflective symmetry between bisected halves of the interval in the odd derivative orders (Figure A1(b)).



a) b)

Figure A1. a) The shift curve (zeroth derivative). b) The shift curve (first derivative).

This leaves the shape of each half of the "s" to be determined. By having each bisection of the shift interval contain a suitably scaled and reflected pair of back-to-back copies of the "s" in the previous order profile, a further symmetry, a recursive symmetry, is induced. This symmetry determines the shape of each half of the "s" so that it copies across all the derivative orders (Figure A2 (a and b) show the copying for the next two orders).



a) **b)**

Figure A2. a) The shift curve (second derivative). b) The shift curve (third derivative).

This determination will be described now through a velocity-height relation.

The back-to-back copying of the heights in Figure A1. (a) by the velocities in Figure A1.(b) may be represented by the pair of differential equations A1 and A2 below:

[A1] $s'(t) = 2s(2t) \qquad t \in [0, \frac{1}{2}]$

[A2] $s'(t) = 2 - s'(t - \frac{1}{2}) \qquad t \in \left]\frac{1}{2}, 1\right]$

where $t \in R$, s is the shift function, and s' is its first order derivative. This pair of equations is sufficient to generate a unique curve from the basic requirements of an arbitrary s-shape as will now be shown.

These requirements provide known height and velocity-height relations a priori simply through the known end points and rotational symmetry. The relations generate $\left(2^n-2\right)$ equations in terms of $\left(2^n-2\right)$ unknown symmetric heights and 2^n unknown symmetric velocities which determine a specific s-shape approximated semi-linearly by $\left(2^n+1\right)$ sample points (where n may be thought of as a degree of resolution of the curve). The semi-linear approximation to the above pair of differential equations by a corresponding pair of finite-difference equations can then be shown to provide a further 2^n linear simultaneous equations to eliminate the unknown velocities from the original $\left(2^n-2\right)$ equations. This leaves the original $\left(2^n-2\right)$ equations in terms of $\left(2^n-2\right)$ unknown symmetric heights. With the symmetries between pairs of heights taken into account, there are $2^{(n-1)}-1$ equations in terms of $2^{(n-1)}-1$ unknown heights that may be solved to generate unique and exact solutions for all the unknowns. As n tends to infinity, a unique and exact C^{∞} smooth solution is converged to.

The velocity-height relations that result from the basic requirements can be demonstrated by extracting them from an algorithmic summary of my shift construction made by Gordon Milligan in collaborative work. The extract is given in the form of Algorithm A1 below. The equations referred to in the algorithm are given immediately after the algorithm.

1. Choose a positive integer n for the curve's resolution, $n \geq 2$.
2. Set the number of discrete sample points, *sample_no*, in $s_n^0(t)$ using Eqn. A.3.
3. Define the width of the trapezoidal areas, *trap_width*, in $s^1(t)$ using Eqn. A.4.
4. Set the values of the 3 a priori known sample heights at $t = 0$, $t = 1/2$ and $t = 1$ according to Eqns. A.5, A.6, and A.7 respectively.
5. Equate the symmetric velocities v_i using Eqn. A.8.
6. Define $2^{(n+1)}$ trapezia, T_i, in $s_n^1(t)$ that cover the curve's velocity profile in terms of the sample velocities v_i according to Eqn. A.9.
7. Equate the $2^{(n-1)} - 1$ unknown heights, h_i, in the left half of the curve $s^0(t)$ with the trapezia T_i expressed in terms of the sample velocities v_i according to Eqn. A.10.
8. Equate the symmetric heights h_i using Eqn. A.11.

where $s_n^0(t)$ and $s_n^1(t)$ are the semi-linear approximations to the zeroth and first order forms showing the heights and velocities of an s-shape of resolution n respectively.

Algorithm A1. The provision of velocity-height relations through a semi-linear s-shape curve construction.

The equations used in the algorithm A1 are given below:

$$sample\_no = 2^n + 1 \qquad [A3]$$

$$trap\_width = \frac{1}{2^{(n+1)}} \qquad [A4]$$

$$h_1 = 0 \qquad [A5]$$

$$h_{\left(2^{(n-1)}+1\right)} = \tfrac{1}{2} \qquad [A6]$$

$$h_{\left(2^n+1\right)} = 1 \qquad [A7]$$

$$v_{\left(2^{(n+1)}+2-i\right)} = v_i \qquad i = 2...2^n \qquad [A8]$$

$$T_i = \frac{1}{2}.trap\_width.\left(v_{i+1} + v_i\right) \qquad i = 1...2^{(n+1)} \qquad [A9]$$

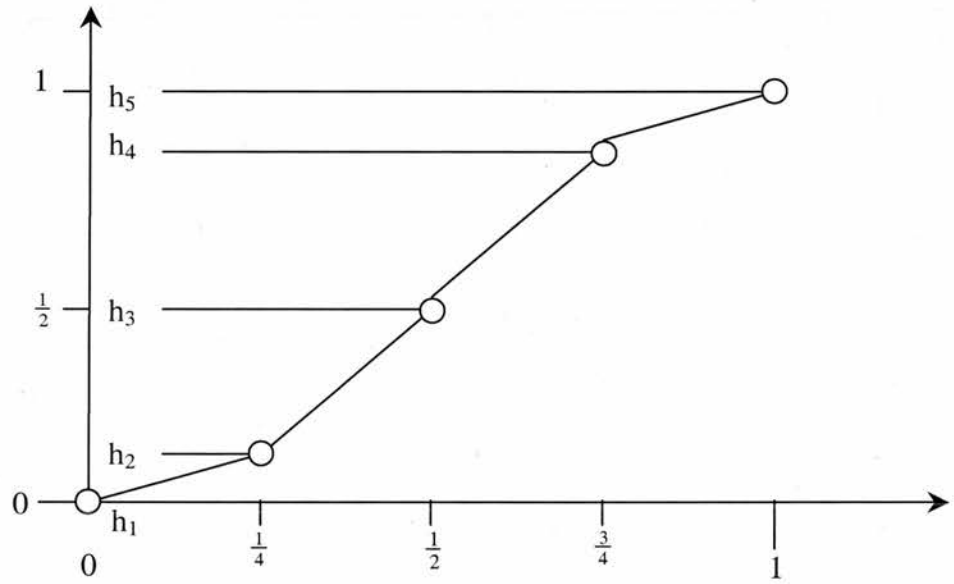
$$h_i = \sum_{p=1}^{2(i-1)} T_p \qquad i = 2...2^{(n-1)} \qquad [A10]$$

$$h_{\left(2^n+2-i\right)} = 1 - h_i \qquad i = 2...2^{(n-1)} \qquad [A11]$$

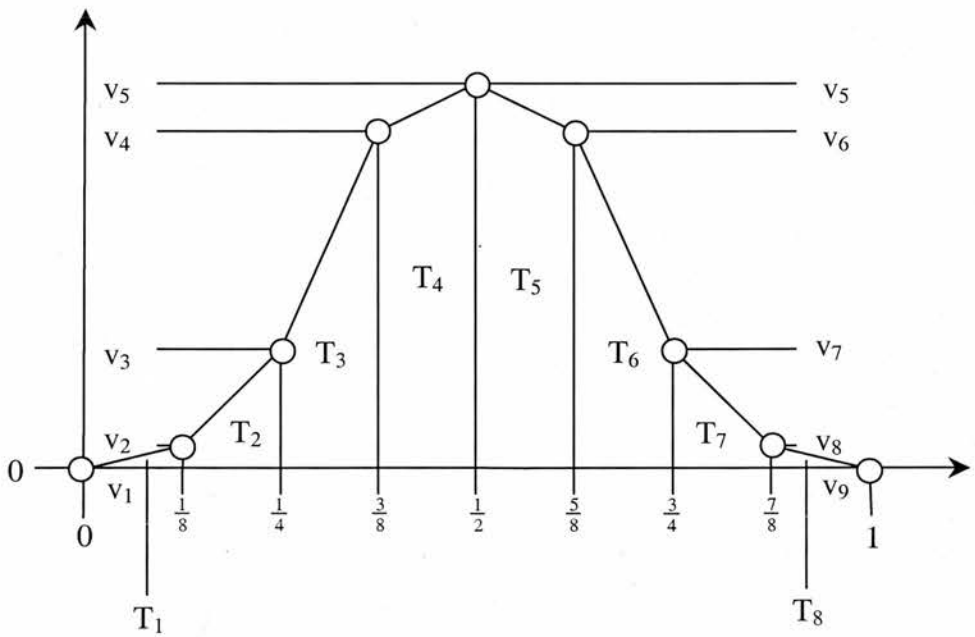
A graphical illustration of the velocity-height relations resulting from algorithm A1 is shown in Figure A1.

Figure A1.

An arbitrary s-shaped curve with a resolution of n = 2 and rotational symmetry about (1/2, 1/2).



(a) The height profile. The value of h_2 (=h_4) is unknown.



(b) The velocity profile. The values of v_2 (= v_8), v_3 (= v_7), v_4 (= v_6), and v_5, are unknown.

The pair of fundamental differential equations may then be shown to provide a unique solution for the velocity-height relation by inserting their finite-difference approximations in step 6 through equations [A1] and [A2], and solving the equations specified in step 7 in a new step 8. We then have the solution algorithm A2 shown below.

1. Choose a positive integer n for the curve's resolution, $n \geq 2$.
2. Set the number of discrete sample points, *sample_no*, in $s_n^0(t)$ using n and Eqn. A.3.
3. Define the width of the trapezoidal areas, *trap_width*, in $s^1(t)$ using n and Eqn. A.4.
4. Set the values of the 3 a priori known sample heights at $t = 0$, $t = 1/2$ and $t = 1$ according to Eqns. A.5, A.6, and A.7.
5. Equate the symmetric velocities v_i using using Eqn. A.8.
6. Define $2^{(n+1)}$ trapezia, T_i, in $s_n^1(t)$ that cover the curve's velocity profile in terms of the sample heights h_i according to according to Eqns. A1., A.2, and Eqn. A.9.
7. Equate the $2^{(n-1)}-1$ unknown heights, h_i, in the left half of the curve $s^0(t)$ with the trapezia T_i expressed in terms of the sample heights h_i according to Eqn. A.10.
8. Solve the $2^{(n-1)}-1$ equations in *step 7* for unique values for the $2^{(n-1)}-1$ unknown heights, h_i.
9. Equate the symmetric heights h_i using Eqn. A.11.

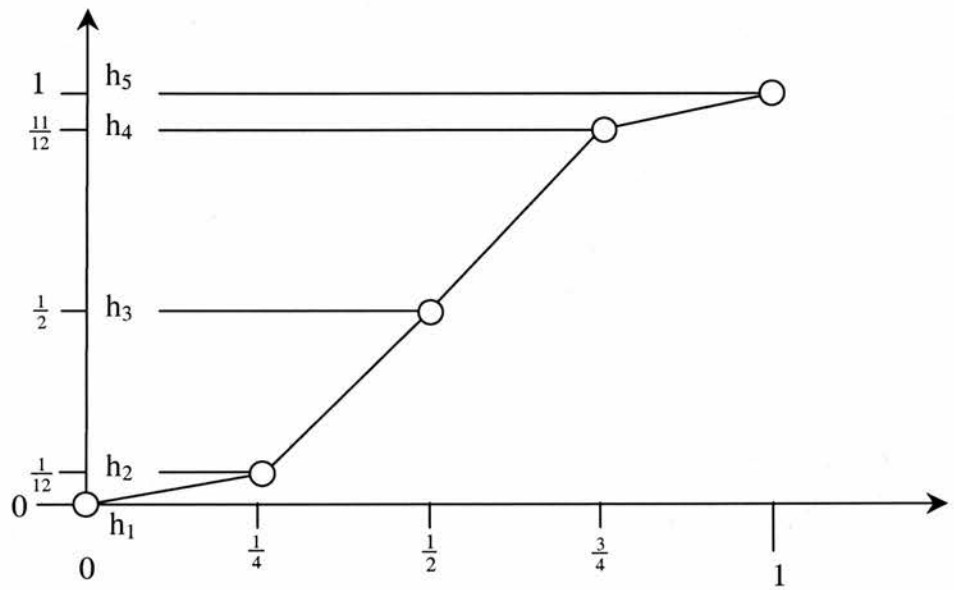
where $s_n^0(t)$ and $s_n^1(t)$ are the semi-linear approximations to the zeroth and first order profiles showing the heights and velocities of an s-shape of resolution n respectively.

Algorithm A2. The realisation of unique solution for shift curve construction.

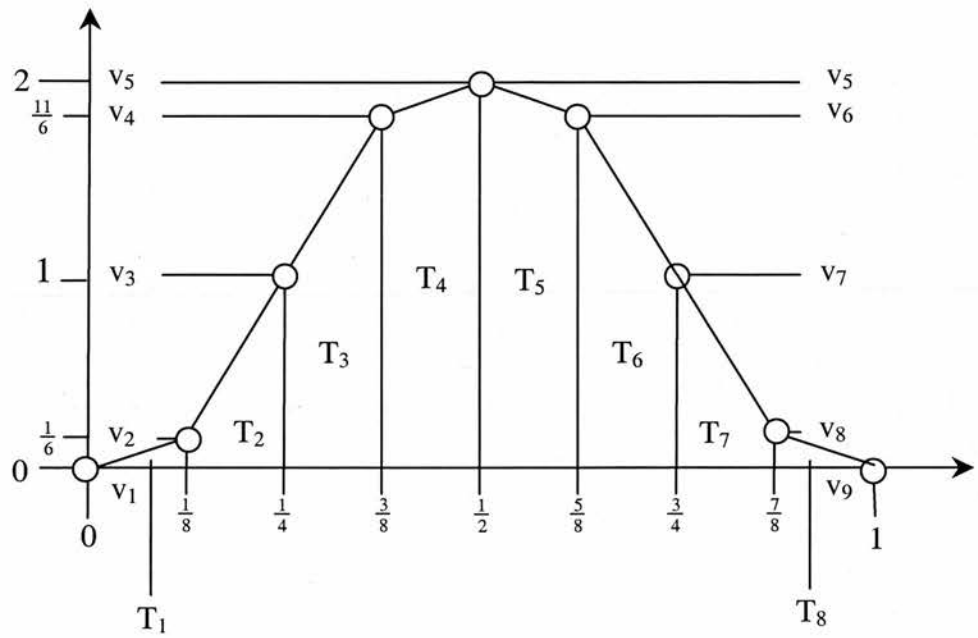
The semilinear graphs of the height and velocity profiles now have complete, unique, and exact detail for the sample points as shown in Figure A2. As n tends to infinity, a unique and exact C^{∞} smooth solution is converged to.

Figure A2.

A semi-linear approximation to the shift curve with a resolution of n = 2.



(a) The height profile. The value of h_2 ($=h_4$) is unique and exact.



(b) The velocity profile. The values of v_2 (= v_8), v_3 (= v_7), v_4 (= v_6), and v_5, are unique and exact.

Appendix B

Time	Height	Time	Height
0	0	0.515625	0.531 249 994 000
0.015625	0.000 000 006 056	0.53125	0.562 499 237 100
0.03125	0.000 000 762 935	0.546875	0.593 738 307 700
0.046875	0.000 011 692 350	0.5625	0.624 927 266 400
0.0625	0.000 072 733 561	0.578125	0.655 973 821 900
0.078125	0.000 276 178 027	0.59375	0.686 726 125 000
0.09375	0.000 773 874 919	0.609375	0.716 979 974 400
0.109375	0.001 770 025 586	0.625	0.746 500 651 000
0.125	0.003 499 348 959	0.640625	0.775 044 765 200
0.140625	0.006 205 234 830	0.65625	0.802 368 927 000
0.15625	0.010 131 107 300	0.671875	0.828 238 417 600
0.171875	0.015 511 582 390	0.6875	0.852 449 035 700
0.1875	0.022 550 964 350	0.703125	0.874 848 931 900
0.203125	0.031 401 068 070	0.71875	0.895 345 815 000
0.21875	0.042 154 184 980	0.734735	0.913 899 745 700
0.234375	0.054 850 254 360	0.75	0.930 501 302 100
0.25	0.069 498 697 910	0.765625	0.945 149 745 700
0.265625	0.086 100 254 360	0.78125	0.957 845 815 100
0.28125	0.104 654 185 000	0.796875	0.968 598 931 900
0.296875	0.125 151 068 100	0.8125	0.977 449 035 700
0.3125	0.147 550 964 300	0.828125	0.984 488 417 600
0.328125	0.171 761 582 400	0.84375	0.989 868 927 000
0.34375	0.197 631 073 000	0.859375	0.993 794 765 200
0.359375	0.224 955 234 800	0.875	0.996 500 651 000
0.375	0.253 499 349 000	0.890625	0.998 229 974 400
0.390625	0.283 020 025 600	0.90625	0.999 226 125 000
0.40625	0.313 273 875 000	0.921875	0.999 723 821 900
0.421875	0.344 026 178 100	0.9375	0.999 927 266 000
0.4375	0.375 072 733 600	0.953125	0.999 988 307 000
0.453125	0.406 261 692 300	0.96875	0.999 999 237 100
0.46875	0.437 500 762 900	0.984375	0.999 999 994 000
0.484375	0.468 750 006 000	1	1
0.5	0.500 000 000 000		

Table B1. Shift curve zeroth order height values for specific points in time (as a fraction of the 1:1 interval being analysed) to 12 decimal places.

Appendix C

The methodology for generating predictive fits is described fully in Section 4.6. Below is a worked example based on this methodology using real data taken from Experiment 2 (see Section 6.3).

Worked Example of Predictive Fit Generation

The aim is to start with the data for one run of one subject in one experiment and finish with a value for the shift ratio so that a predictive fit is achieved.

1. Generating a residual

The aim in this first step is to use predicted values from the Taylor series and compute filtered residual values for variables. Filtering means subtracting the Taylor series predictions from the actual vertical displacements to give filtered displacements. Further filtered derivatives such as the filtered velocity may then be extracted from the filtered displacements as well. To begin, the variation interval is standardly made [0, 1] by scaling for ease of comparison between curves. The residual curve R has a

value at some time t, where time is measured as a fraction of the variation interval [0, 1], that is given by Equation C1, below:

(C1) $$R(t) = A(t) - TSP(t)$$

where A is the actual curve and TSP is the Taylor Series Prediction. An example of the various curves involved with the residual for Subject 2.1, Run 4, from $t_0 = 1/2$ is shown in Figure C1, below.

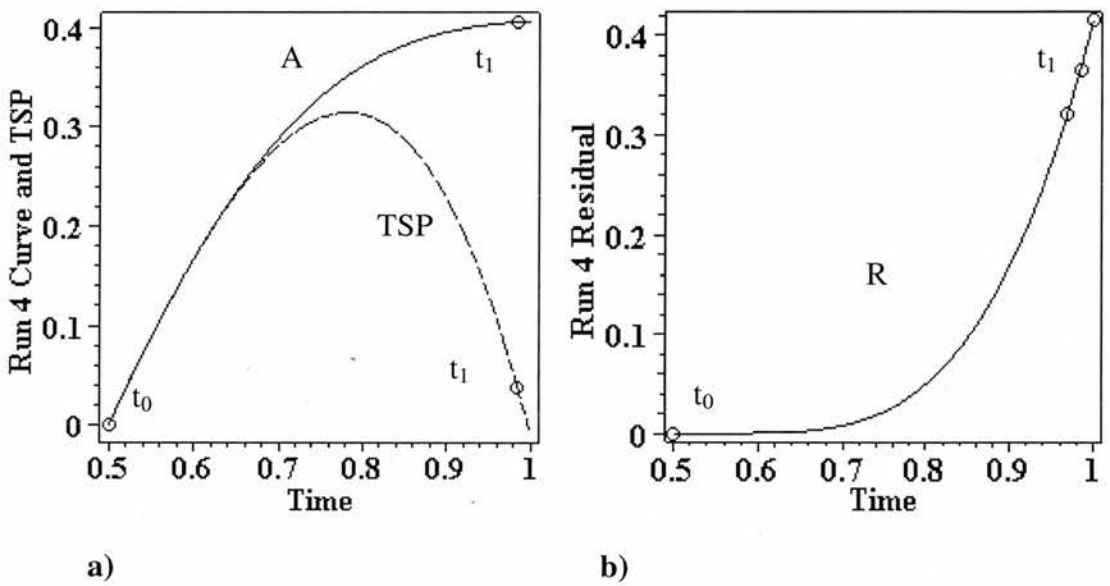


Figure C1. a) Actual curve (A) and Taylor Series Prediction (TSP). b) Residual curve. The times $t_0 = 1/2$, $t_1 = 63/64$, and two extra points a small distance either side of t_1 provide data for Equation C8 in the stage calculation.

2. Calculation of the $V_f W / S_f$ ratio and the stage of the shift.

The aim here is to take data for the residual curve, computed above, and produce the *stage* that the shift has reached by the end of an interval. This production requires the prior computation of a ratio V_fW/S_f. The processing is based on derivative states made up from displacement-velocity pairs.

The first such derivative state at time t_0 is the leftmost one for the variation interval and, in the example shown in Figure C1, is taken from t = 1/2. The filtered vertical displacement and the filtered velocity are always both zero at the leftmost state due to the Taylor series starting with the actual curve derivatives at this state. The second derivative state at time t_1 is the rightmost one for the interval, which, in the example in Figure C1, is taken from t = 63/64.

The first operation in finding the stage that the shift has reached at the rightmost point involves calculating a ratio given by Equation C2, below:

(C2) $$T = V_f W / S_f$$

where V_f is the filtered velocity at the rightmost of the two points being used, W is the width of the time interval between the two points and S_f is the filtered vertical displacement at that rightmost point.

As discussed in Section 4.6.4 each T value picks out a unique stage for the shift regardless of its shift ratio or interval. This relationship is shown in Figure C2. So the

second operation required for calculating the stage is to read off the stage value on the graph for the calculated T value at time t_1, i.e. for the rightmost of the two points.

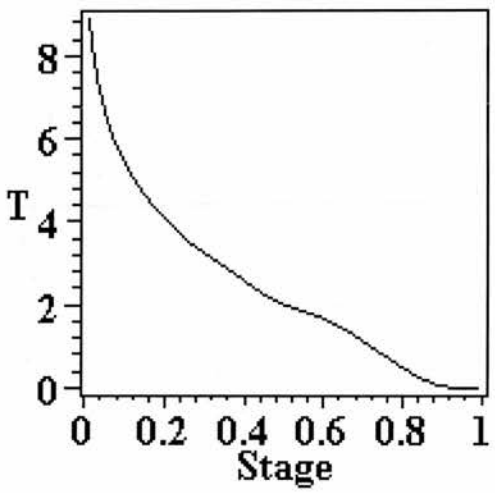


Figure C2. The V_fW/S_f ratio (T) for a shift with the stage given as the fraction of the actual shift that has been completed.

To develop the present example, equations are first required for the filtered velocity, width and filtered displacement. The computations of W, S_f, and V_f are as follows:

(C3) $W = t_1 - t_0$

(C4) $S_f = R(t_1) = A(t_1) - TS(t_1)$

(C5) $V_f = d(S_f) / dt$

So for W:

(C6) $W = (63/64) - (1/2) = 31/64$

Performing the subtraction of the Taylor series prediction from the actual curve at t_1 yields S_f:

(C7) $$S_f = A(t_1) - TS(t_1) = 0.405124 - 0.0389301 = 0.366194$$

Such subtraction is repeated for S_{f+x} and S_{f-x} where x is the magnitude of the small fractional time between consecutive readings. In this case x = 1/64. S_{f+x} is the filtered displacement at t = 1 and S_{f-x} is the filtered displacement at t = 31/32, i.e. 1/64 either side of the time for S_f which is t = 63/64 in the current example. This gives $S_{f+x} = 0.415683$ and $S_{f-x} = 0.321069$.

To generate V_f the gradient of S_f at t_1 is required. This is calculated using the following approximation:

(C8) $$V_f = (S_{f+x} - S_{f-x}) / ((t_1 + x) - (t_1 - x))$$

So we have:

(C9) $$V_f = (0.415683 - 0.321069) / (1 - (31/32)) = 3.02765$$

Hence we can express T in terms of values for W, V_f, and S_f through substitution of C6, C7 and C9 into C2:

(C10) $$T = V_f W / S_f = 3.02765 * (31/64) / 0.366194 = 4.00476$$

The look-up table in Appendix B requires stages given in sixty-fourths. The T value here approximately corresponds to the stage 13/64 for the next step (by reading off the graph shown in Figure C2).

3. Calculation of the shift ratio.

The aim in this final section is to calculate the shift ratio (SR) from the stage, this requires an intermediate computation of the overall relative shift magnitude (SM). First of all, taking 13/64 as the stage prediction, the ratio of the actual height (A) reached over the interval to the height (S) reached by the 1:1 shift curve over the same interval is the overall relative shift magnitude (SM). The value of S at the stage of 13/64 (i.e. S = 0.031 401 068 070) may be taken from the look-up table in Appendix B. In the case of the current example this shift magnitude (SM) is given by Equation C11 below:

(C11) $$SM = A / S = 0.405124 / 0.0314011 = 12.9016$$

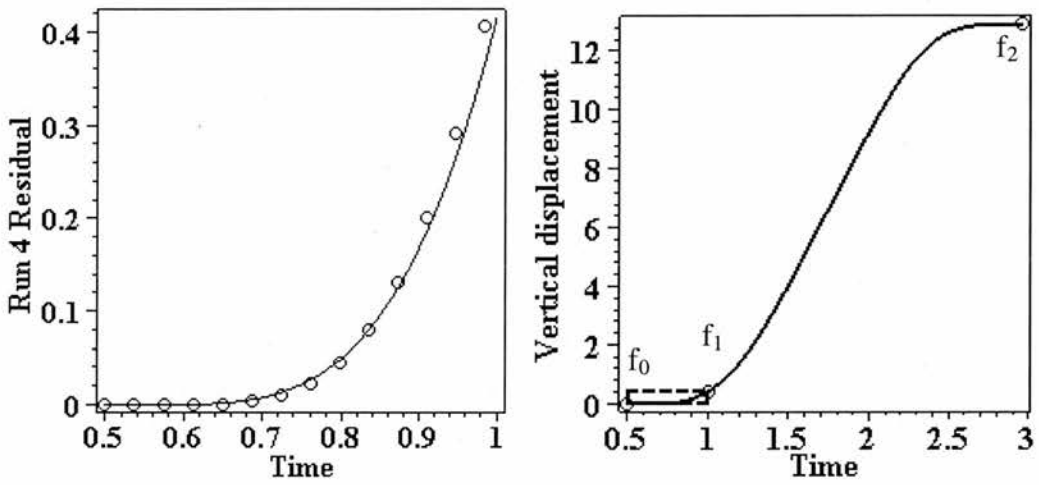
Since the stage prediction of 13/64 corresponds to a time interval of 1 - (31/64) = 33/64, the time interval for the full shift, the overall shift width (SW), is (64/13) * (33/64) = 2.53846. [The time at the end of the full shift is therefore 2.53846 + (1/2) = 3.03846].

The shift ratio (SR) can be worked out by taking the overall filtered vertical displacement, i.e. the overall relative shift magnitude (SM) and dividing it by the

overall horizontal displacement, i.e. the overall shift width (SW) as in Equation C12, below:

(C12) $SR = SM / SW = 12.9016 / 2.53846 = 5.08245$

The resulting predictive fit (represented as dots) to the residual (solid line) is shown in Figure C3, below.



a) b)

Figure C3. a) Predictive fit (represented as dots) to the residual (solid line) for Subject 2.1, Run 4. b) The fit found in Figure C3(a) is the section of a full shift curve between f_0 = (1/2, 0) and $f_1 = (t_1, A(t_1))$ = (63/64, 0.405124). The full shift curve has a shift ratio of SR to 1, i.e. 5.08245:1. The box in the lower left hand corner is an enlargement of Figure C3(a).

References

Aristotle. Physics Book II.

Arrowsmith, D.K. & Place, C.M. 1992. *Dynamical Systems*. Chapman and Hall.

Beale, R. & Jackson, T. 1990. *Neural Networks*. IOP.

Berkenblit, M.B., Gelfand, I.M., & Feldman, A.G. 1986. *A Model for the Control of Multi-Joint Movements.* Biofizika, **31**, 128-138.

Blackburn, S. 1994. *The Oxford Dictionary of Philosophy*. OUP.

Blackwell, R. J. 1991. *Galileo, Bellarmine, and the Bible*. University of Notre Dame Press.

Blankevoort, L., Huiskes, R. & De Lange, A. 1988. *The Envelope of Passive Knee Joint Motion*. Journal of Biomechanics, **21**, 705-720.

Borowski, E. J. & Borwein, J. M. 1999. *Unwin Hyman Dictionary of Mathematics.* Bookmart Ltd.

Braithwaite, R.B. 1946. *Teleological Explanation.* Proc. Aristotelian Society, XLVII, 1946-7.

Brogan, W.L. 1991. *Modern Control Theory*, 3rd ed. Prentice-Hall.

Cappozzo, A., Cattani, F., Della Croce, U. & Leardini, A. 1995. *Position and Orientation of Bones During Movement: Anatomical Frame Definition and Determination.* Clinical Biomechanics, **10**, 171-178.

Cappozzo, A., Cattani, F., Leardini, A., Benedetti, M.G., & Della Croce, U. 1996. *Position and Orientation of Bones During Movement: Experimental Artefacts.* Clinical Biomechanics, **11**, 90-100.

Carlson, N.R. 2004. *Physiology of Behaviour. 8th Edition.* Pearson.

Carter, M.C. & Shapiro, D.C. 1984. *Control of Sequential Movements: Evidence for Generalized Motor Programs.* Journal of Neurophysiology, **52**, 2318-2330.

Carter, R. 1998. *Mapping The Mind.* Weidenfeld and Nicolson.

Chillingworth, D.R.J. 1976. *Differential Topology With A View to Applications.* Research Notes in Mathematics, **Vol. 9.** Pitman Publishing.

Cohen, I. B. & Smith, G. E. 2002. *The Cambridge Companion to Newton*. CUP.

Cohen, J. & Stewart, I. 1994. *Collapse of Chaos*. Penguin.

Coveney, P. & Highfield, R. 1995. *Frontiers of Complexity*. Faber and Faber.

Darwin, C. 1859. *The Origin Of Species*.

Dawkins, R. 1976. *The Selfish Gene*. OUP.

Dawkins, R. 1982. *The Extended Phenotype*. OUP.

Dayan, P. & Abbott, L.F. 2001. *Theoretical Neuroscience*. MIT Press.

Des Chene, D. 1996. *Physiologia. Natural Philosophy in Late Aristotelian and Cartesian Thought*. Cornell University Press.

Dillman, C.J. & Lange, G.W. 1994. *How Has Biomechanics Contributed to the Understanding of the Golf Swing?* In *Science and Golf II: Proceedings of the World Scientific Congress of Golf.* Edited by Cochran, A.J. & Farrally, M.R. Published by E & FN Spon.

Dorf, R.C. 1992. *Modern Control Systems*. Addison-Wesley.

Dusek, V. 2001. *Aristotle's Four Causes and Contemporary "Newtonian" Dynamics.* In 'Aristotle and Contemporary Science, Volume Two,' edited by Sfendoni-Mentzou, D., Hattiagadi, J. & Johnson, D. M. Peter Lang Publishing.

Engin, A.E. 1980. *On The Biomechanics of the Shoulder Complex.* Journal of Biomechanics, **13,** 575-590.

Fernández-Armesto, F. 1997. *Truth. A History and a Guide for the Perplexed.* Black Swan.

Gardner, H. *The Mind's New Science.* Basic Books.

Gibson, J.J. 1979. *The Ecological Approach To Visual Perception.* Lawrence Erlbaum.

Jeffrey, A. 1996. *Mathematics for Engineers and Scientists, 5th Edition.* Nelson Thornes.

Gielen, C.C.A.M., van der Oosten, K., & Pull ter Gunne, F. 1985. *Relation Between EMG Activation Patterns and Kinematic Properties of Aimed Arm Movements.* Journal of Motor Behaviour, **17**, 421-442.

Gleick, J. 1987. *Chaos.* Cardinal.

Gregory, R.L. (ed.) 1997. *The Oxford Companion To The Mind.* OUP.

Gullberg, J. 1997. *Mathematics From The Birth of Numbers.* W. W. Norton & Company.

Gutman, S.R. Gottlieb, G.L., & Corcos, D.M. 1990. *Exponential Model of a Reaching Movement Trajectory with Nonlinear Time.* Comments on Theoretical Biology, **2**, 357-383.

Hale, W.G., Margham, J.P. & Saunders, V.A. 1995. *The Unwin Hyman Dictionary of Biology.* Harper Collins.

Hahn, A. J. 1998. *Basic Calculus From Archimedes to Newton to its Role in Science.* Springer.

Hankinson, R.J. 1998. *Cause and Explanation in Ancient Greek Thought.* Clarendon Press.

Hassing, R.F. (ed.). 1997. *Final Causality in Nature and Human Affairs.* Catholic University of America Press.

Hellemans, A. & Bunch, B. 1988. *The Timetables of Science. A Chronology of the Most Important People and Events in the History of Science.* Simon and Schuster.

Heuer, H. & Keele, S. (eds.). 1996. *Handbook of Perception and Action (II)*. New York: Academic Press.

Honderich, T. (ed.) 1995. *The Oxford Companion to Philosophy*. OUP.

Horgan, J. 1996. *The End Of Science*. Abacus.

Houk, J.C., Kiefer J. & Barto, A.G. 1993. *Distributed Motor Commands In The Limb Premotor Network*. Trends Neurosci 16: 27-33.

Houk, J.C. & Wise, S.P. 1995. *Distributed Modular Architectures Linking Basal Ganglia, Cerebellum, and Cerebral Cortex: Their Role In Planning and Controlling Action*. Cereb Cortex 5: 95-110.

Johnson-Laird, P. 1993. *The Computer and The Mind*. Fontana Press.

Juarrero, A. 1999. *Dynamics In Action. Intentional Behavior As A Complex System.* MIT Press.

Kocsis, L., Kiss, R.M., Knoll, Z, & Jurak, M. 1999. *Bute's Ultrasound-Based Measuring Technique and Model for Gait Analysis*. Physical Education Vol. 1, No. 6, 1999, pp. 1-13.

Kuhn, T. S. 1996. *The Structure of Scientific Revolutions*. The University of Chicago Press.

Lee, D.N., Young, D.S., & Rewt, D. 1992. *How Do Somersaulters Land On Their Feet?* Journal of Experimental Psychology, **18**, 1195-1202.

Lee, N., Erickson, M. & Cherveny, P. 2002. *Measurement of the Behaviour of a Golf Club During the Golf Swing.* In *Science and Golf IV: Proceedings of the World Scientific Congress of Golf.* Edited by Thain, E. Published by Routledge.

Lennox, J.G. 2001. *Aristotle's Philosophy of Biology.* CUP.

Lewis, F.L. 1992. *Applied Optimal Control and Estimation.* Prentice-Hall.

Magill, R.A. 1998. *Motor Learning. Concepts and Applications. 5th Edition.* McGraw-Hill.

Martel Johnson, D. & Erneling, C.E. (eds.). 1997. *The Cognitive Revolution.* OUP.

McClelland, J.L. & Rumelhart, D.E. (eds.). 1986. *PDP Volume II.* MIT Press.

McLaughlin, P.A. & Best, R.J. 1994. *Three-Dimensional Kinematic Analysis of the Golf Swing.* In *Science and Golf II: Proceedings of the World Scientific Congress of Golf.* Edited by Cochran, A.J. & Farrally, M.R. Published by E & FN Spon.

Merryfield, K.G. 1992. *A Nowhere Analytic C^{∞} Function.* Missouri Journal of Mathematical Sciences, Vol. 4, No. 3, (Fall 1992), 132-138.

Minsky, M. 1972. *Computation: Finite And Infinite Machines*. Prentice-Hall.

Morris, S.J. & Paradiso, J.A. 2003. *A Compact Wearable Sensor Package for Clinical Gait Monitoring*. Offspring **Vol. 1**, No. 1, pp. 7-15, January 31, 2003.

Nagasaki, H. 1989. *Asymmetric Velocity and Acceleration Profiles of Human Arm Movements*. Exp Brain Res **74**:319-326.

Neal, R.J., Abernethy, B., Moran, M.J. & Parker, A.W. 1990. *The Influence of Club Length and Shot Distance On the Temporal Characteristics of the Swings of Expert and Novice Golfers*. In *Science and Golf: Proceedings of the First World Scientific Congress of Golf.* Edited by Cochran, A.J. Published by E & FN Spon.

Nesbitt, S.M., Cole, J.S., Hartzell, T.A., Oglesby, K.A. & Radich, A.F. 1994. *Three-Dimensional Kinematic Analysis of the Golf Swing*. In *Science and Golf II: Proceedings of the World Scientific Congress of Golf.* Edited by Cochran, A.J. & Farrally, M.R. Published by E & FN Spon.

Newell, A. & Simon, H.A. 1976. *Computer Science As Empirical Enquiry*. Communications of the ACM, **19(3),** pp.113-126.

Novak, K.E., Miller, L.E. & Houk, J.C. 2000. *Kinematic Properties of Rapid Hand Movements In A Knob Turning Task*. Exp Brain Res (2000) **132**: 419-433.

O'Connor & Zavatsky. 1990. Kinematics and Mechanics of the Cruciate Ligaments of the Knee. In V.C. Mow, V.C. Ratcliffe, & S.L-Y. Woo (Eds.), *Biomechanics of Diarthroidal Joints* (Vol. 2. pp. 197-241). Springer-Verlag.

Peltonen, M., ed. 1996. *The Cambridge Companion to Bacon*. CUP.

Penrose, R. 2004. *The Road to Reality. A Complete Guide to the Laws of the Universe*. Jonathan Cape.

Peterson, D. (ed.). 1996. *Forms of Representation*. Intellect.

Plamondon, R. & Alimi, A.M. 1997. Speed/Accuracy Trade-offs in Target-directed Movements. *Behavioral and Brain Sciences* **20** (2): 279-349.

Plamondon, R., Alimi, A.M., Yergeau, P., & Leclerc, F. 1993. *Modeling Velocity Profiles of Rapid Movements: A Comparative Study*. Biological Cybernetics, **69**, 119-128.

Posner, M.I. & Raichle, M.E. 1997. *Images of Mind*. Scientific American Library.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P. 1992. *Numerical Recipes in C. The Art Of Scientific Computing. Second Edition.* CUP.

Rich, E. & Knight, K. 1991. *Artificial Intelligence*. McGraw-Hill.

Roberts, J. M. 1992. *History of the World*. Helicon.

Rosenblueth, A., Wiener, N. & Bigelow, J. 1943. *Behaviour, Purpose and Teleology*. Philosophy of Science, **Vol. 10**.

Rumelhart, D.E. & McClelland, J.L. (eds.). 1986. *PDP Volume 1*. MIT Press.

Russell, E.S. 1945. *The Directiveness of Organic Activities*. CUP.

Russell, S.J. & Norvig, P. 2003. *Artificial Intelligence. A Modern Approach. Second Edition*. Prentice Hall.

Schaal, S. 2002. *Arm and Hand Movement Control*. In: Arbib MA (ed) The Handbook of Brain Theory and Neural Networks, 2nd Edition. MIT Press, Cambridge MA, pp. 110-113.

Smith, B. 2003. *The Ecological Approach to Information Processing*. In: Nyiri JC (Ed.): Mobile Learning: Essays on Philosophy, Psychology and Education. Vienna: Passagen Verlag 2003; 17-24.

Soechting, J.F. 1984. *Effect of Target Size On Spatial and Temporal Characteristics of a Pointing Movement In Man*. Exp Brain Res **54**:121-132.

Soechting, J.F. & Flanders, M. 1992. *Moving in Three-dimensional Space: Frames of Reference, Vectors, and Coordinate Systems.* Annual Review of Neuroscience **15**, 167-191.

Sommerhoff, G. 1969. *The Abstract Characteristics of Living Systems.* In 'Systems Thinking', Emery, F.E. (ed.). Penguin.

Sommerhoff, G. 1974. *Logic of the Living Brain.* John Wiley & Sons.

Sutton, R.S. & Barto, A.G. 1998. *Reinforcement Learning*. MIT Press.

Taylor, C. 1964. *The Explanation of Behaviour.* Routledge & Keegan Paul.

Taylor, R. 1966. *Action and Purpose*. Prentice-Hall.

Voltaire. 1759. *Candide.*

Wale, A.P. & Weir, M.K. 2002. *Measurement and Design of Goal-Directed Behaviour.* Proceedings of the Seventh Neural Computation and Psychology Workshop, pp. 78-89. World Scientific.

Weir, M.K. 1984. *Goal-Directed Behaviour.* Gordon and Breach Science Publishers.

Weir, M.K. 2005. *Fully Smooth Change and the Shift Function*. Intelligent Computation Group Internal Report. School of Computer Science, University of St. Andrews.

Woodfield, A. 1976. *Teleology*. CUP.

Wright, L. 1968. *The Case Against Teleological Reductionism*. Brit. J. Phil. Science, **19**, 1968-69.

Zatsiorsky, V.M. 1998. *Kinematics of Human Motion*. Human Kinetics Europe Ltd.