Pascal Hartelman and Peter Molenaar

Detection of developmental transitions

Abstract

Catastrophe theory can be very useful in research concerning developmental transitions. A very powerful and reliable technique for fitting catastrophe models is the method of Cobb. Applying catastrophe theory without restrictions means that coordinate transformations should be implemented in the method of Cobb. However, this will cause over-parametrisation and invariance problems. In this paper it will be shown that these problems can be solved by using non-parametric smoothing techniques and so-called level crossing characteristics of a stochastic process.

Introduction

Since the introduction, by Piaget, of his epigenetic model of development, a lively discussion emerged on the plausibility of this model (Piaget, 1960; Flavell, 1971; Brainerd, 1978; Fisher & Silvern, 1985) and stage-wise developmental models in general. In a stage-wise developmental model one distinguishes stages that correspond to qualitatively different behaviour patterns. Furthermore, development is accompanied by abrupt transitions between these stages. According to Piaget, conservation acquisition necessitates learning new logical rules. A particular developmental stage of a subject can be characterised by the set of logical rules that are used. A transition between stages leads to a sudden change in cognitive capacities. The discussion on the plausibility of stage-wise developmental models, has for a long time been obstructed because of a lack of consensus about the formal definition of both a stage and a transition. Often a transition is described as a 'large and rapid change'. With such a definition it is difficult to construct empirical criteria, necessary for the detection of transitions in experimental research. Recently it has been shown (van der Maas & Molenaar, 1992) that catastrophe theory can be used to develop a formal transition theory and an acceptable methodology. Such a methodology consists of applicable definitions of stages and transitions, as well as statistical techniques (Hartelman, 1997a) for detecting transitions and modelling stage-wise developmental models.

The aim of catastrophe theory is to describe processes that exhibit transitions as well as classifying processes with respect to their qualitative, transitional properties. For the application of catastrophe theory we distinguish three approaches: catastrophe detection, catastrophe modelling and catastrophe analysis. In catastrophe analysis one investigates the known 'equations of motion' of the process by means of the mathematical tools provided by catastrophe theory. In the social sciences this approach is often beyond our reach because the processes under study are most often ill-defined. In catastrophe detection one uses so-called flags (Gilmore, 1981), certain characteristics of the behavioural variable in a catastrophic process, which can be used as indicators of a transition. The application of these flags requires a minimum of information about the system. The application of flags for developmental processes in general, and conservation acquisition in particular, has been discussed thoroughly by van der Maas (1992). This paper deals with catastrophe modelling, which consists of fitting and testing catastrophe models. A very reliable technique for fitting catastrophe models is the method of Cobb. However, in this technique it is not possible to implement smooth co-ordinate transformations, which are crucial in catastrophe theory. This will lead to over-parametrisation and invariance problems. Solution to these problems will be presented in this contribution, which consists of using kernel density estimates and so-called level crossing characteristics of a stochastic process. We will start with a small introduction to catastrophe theory. A description of the method of Cobb will be given in §3. Finally, in §4 the transformation problems and proposed solutions will be discussed. In §5 a statistical technique for testing the presence of a transition will be applied to conservation acquisition data of van der Maas (1992).

Catastrophe Theory

Catastrophe theory deals with gradient systems, which are specific dynamical systems, the behaviour of which can be described by one (smooth) function. These functions, which are often called potential functions, depend on so-called state (behavioural) variables $y_1, \ldots, y_n$ and control variables (model parameters) $c_1, \ldots, c_m$. The state variables describe the state of a system, and the control variables determine the change of the state in time. In catastrophe theory gradient systems are classified with respect to their qualitative properties. This classification is induced by co-ordinate transformations. Two systems are equivalent if their corresponding potential functions can be transformed into one other by a change in the co-ordinate system. This means that in catastrophe theory one is not interested in the quantitative properties of a system but in the qualitative properties of a system, i.e. those properties that do not depend on the choice of the measurement scales. The classification leads to the partition of the class of gradient systems into equivalence classes. Two members of the same equivalence class exhibit the same qualitative behaviour and the corresponding potential functions are the same up to a change in the co-ordinate systems. By using an appropriate co-ordinate transformation $\varphi: (y_1, \ldots, y_n) \rightarrow (x_1, \ldots, x_n)$ one can transform a potential function into a canonical, most simple polynomial form, which is the same for all members of the same equivalence class. This canonical form can thus be used to label the equivalence class. The classification of gradient systems in catastrophe theory is illustrated by the classification of smooth potential functions, using a tree structure, in Figure 1.

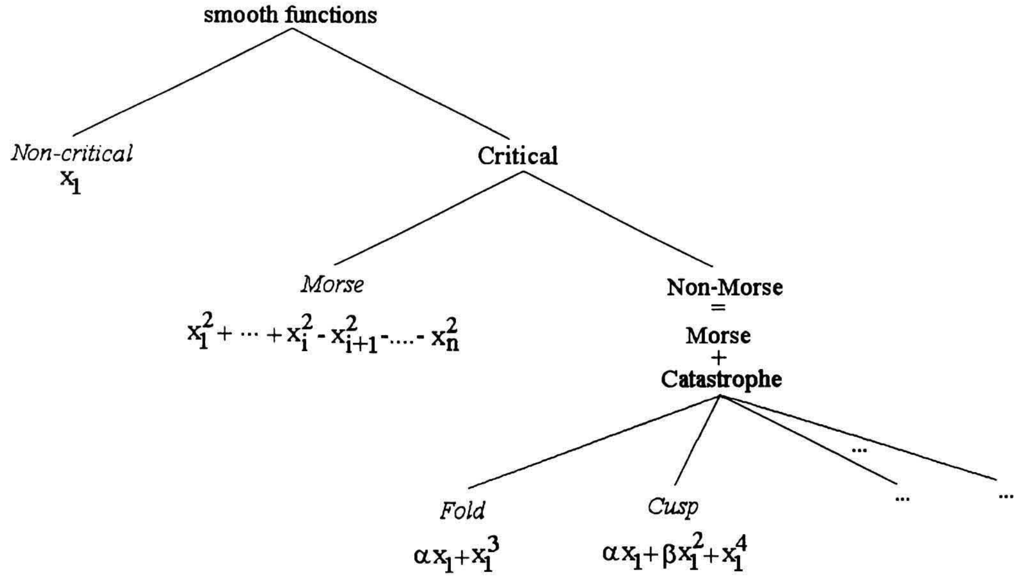

Fig.1. Classification scheme in catastrophe theory. The class of smooth functions can be partitioned into a number of equivalence class (italic). The functions are the corresponding canonical potential functions.

From Figure 1 we see that gradient systems can be partitioned into an equivalence class with so-called non-critical potential functions and a group of systems with critical potential functions. The canonical form of the potential function of non-critical systems is linear in one canonical variable x_1 (The terms critical, non critical, Morse, Non-Morse,... in Figure 1 are connected to functions. However, in the case of gradient systems there is one-to-one relation between gradient systems and the corresponding potential functions. Therefore, we can also use these terms to label the gradient systems). The group of critical systems can be partitioned into an equivalence class of Morse systems and a group of Non-Morse systems. The canonical potential function for Morse systems is a sum of quadratic functions. Finally, the group of Non-Morse systems can be partitioned into a number of equivalence classes, so-called catastrophe systems (Fold, Cusp,...). The canonical potential function of such an equivalence class is the sum of a Morse-function and a so-called catastrophe function. As an example, consider a gradient system with n state variables $y_1,\ldots,y_n$ and m control variables $c_1,\ldots,c_m$. If this system is equivalent to a cusp catastrophe, then the canonical potential function is

$$\left(\alpha x_1 + \beta x_1^2 + x_1^4\right) + \left(x_2^2 + \cdots + x_i^2 - x_{i+1}^2 - \cdots - x_n^2\right) \tag{1}$$

for certain i. The parameters α,β depend on the control variables $c_1,\ldots,c_m$. The first function between brackets is the catastrophe function, the second function is the Morse-function. Only catastrophe systems exhibit transitions. In the case of the cusp

catastrophe potential function in (1), the occurrence of a transition depends only on the form of the catastrophe function, which depends on the parameters α and β. An important advantage of catastrophe theory is that the number of equivalence classes is limited. There are only 7 equivalence classes, i.e. qualitatively different transition models, if the number of state variables does not exceed 2 and if the number of control variables does not exceed 4. Therefore, catastrophe theory can be very useful in those fields of research where model selection is a difficult task.
How a transition can occur will be explained by a discussion of the behaviour of unidimensional gradient systems (with one control variable), the dynamics of which can be described by

$$\frac{dx}{dt} = -\frac{dV\,(x;\,c)}{dx} \tag{2}$$

x is a behavioural variable, that describes the state of the system, c is a control variable and t denotes time. V is the potential function. According to (2) the change in x over time is equal to minus the first derivative of V with respect to the state variable x. If the state of the system x is such that $dV(x;c)/dx$ is zero, then the state x will not change. Such a state is called an equilibrium state. We distinguish two types of equilibrium states: stable and unstable. A small perturbation to a system, which actual state is a stable equilibrium state, leads to a movement back to the stable equilibrium state. In contrast, for an unstable equilibrium state, a small perturbation will lead to a movement away from the unstable equilibrium state, towards a stable equilibrium state. The qualitative form of the potential function determines the qualitative behaviour of the system, which is related to the configuration of the stable and unstable equilibrium states of the system (equilibrium surface). Consider two gradient systems with the following potential functions:

$$V_1\,(x;\,c) = (x - c)^2, \qquad V_2\,(x;\,c) = cx - \tfrac{3}{2}\,x^2 + \tfrac{1}{4}\,x^4 \tag{3}$$

V_1 is a Morse function, V_2 is a special instance of a cusp catastrophe function. The equilibrium surfaces for these systems are depicted in Figure 2.

The equilibrium surface of the first system consists only of stable equilibria. However, for the second system the equilibrium surface has stable and unstable equilibria. At certain c-values (bifurcation points), a stable equilibrium and an unstable equilibrium coalesce and disappear, after which the system is forced to move towards another stable equilibrium. The result is a sudden change in the state of the system: a transition. This example illustrates the difference between Morse-systems and catastrophe systems. The Morse- and non-critical models do not exhibit transitions, whereas catastrophe systems do.

In this short introduction only the most important points in catastrophe theory are discussed. For a more detail description, see Poston & Stewart (1978), Gilmore (1981) or Castrigiano & Hayes (1993).

The most successful applications of catastrophe theory deal with systems for which the equations of motion are known. However, in experimental research we

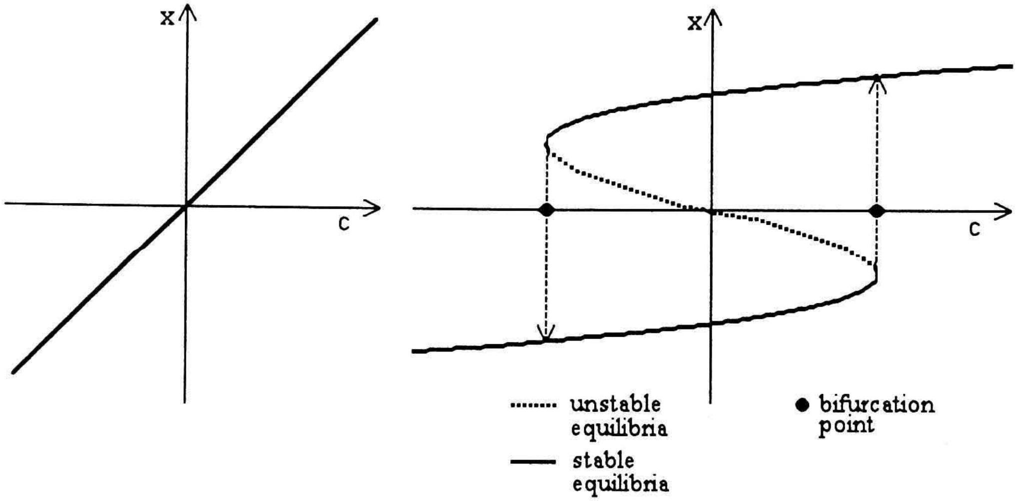

Fig. 2. Equilibrium surfaces of a Morse system with potential function $(x-c)^2$ and a special instance of a catastrophe system with potential function.

often only have limited information, often in the form of a finite number of discrete observations. Moreover, real-life systems are often perturbed by random influences. This raises the questions: how can we apply catastrophe theory to stochastic systems in experimental research. That is, how do we define a stochastic attractor and bifurcation and can we apply the classification scheme in catastrophe theory to stochastic systems? How can we fit catastrophe models to data?

Method of Cobb

A most promising approach to a stochastic catastrophe theory has been developed by Cobb (1978; 1980; 1981). Cobb considers stochastic processes that can be described by a stochastic differential equation. He adds to the gradient system (2), a random process in the form of the derivative of a Wiener process.

$$\frac{dx}{dt} = -\frac{dV(x;c)}{dx} + B(x;c)\frac{dW(t)}{dt} \tag{2}$$

The increments of a Wiener process, i.e. $W(t+\Delta t)-W(t)$, are normally distributed with variance Δt. The function $B(x;c)$ determines the size of the variance of the noise and is called the diffusion function. Cobb often sets this diffusion function to a constant s. The probability density function (PDF) of the state x will ultimately converge to a limiting stationary probability density function (SPDF). This SPDF, which is C $\exp(-V(x;c))$ where C is an integration constant, is used to characterise the behaviour of the stochastic system. Therefore, there is a unique relation between the potential function and the SPDF. The stable (unstable) equilibria correspond to modes (antimodes) of the SPDF. A stochastic bifurcation occurs when the number of modes/antimodes changes as the control variable is varied. The qualitative form of the deterministic potential V is equivalent to the corresponding SPDF. In order to obtain information about the qualitative form of the potential function, one can apply catastrophe theory, and in particular the classification scheme, to the SPDF.

A catastrophe model is fitted by substituting the representative potential function of the class of models that is considered, in the expression for the SPDF. For instance, the SPDF for the cusp catastrophe, with one state variable x and two control variables c_1 and c_2, is

$$p_s(x; \alpha, \beta) = C \exp(\alpha x + \tfrac{1}{2}\beta x^2 - \tfrac{1}{4}x^4) \tag{5}$$

In order to be able to vary the scale and location of the model, Cobb implements a linear scale transformation $x \to (x - \lambda) / \sigma$, where λ is a location and σ a scale parameter, determining respectively the location and scale of the cusp model with respect to the state variable axis. Furthermore, the canonical parameters a,b in the model depend on the measured control variables c_1 and c_2. This relation is in the method of Cobb assumed to be linear

$$a = a_0 + \alpha_1 c_1 + \alpha_2 c_2 \qquad \beta = \beta_0 + \beta_1 c_1 + \beta_2 c_2 \tag{6}$$

The cusp catastrophe model is now fitted, by estimating the parameters λ, σ, α_0, α_1, α_2, β_0, β_1, β_2 by means of maximum likelihood estimation.

Transformation problems of the method of Cobb

Simulation studies (Hartelman, 1997a) show that the method of Cobb is a powerful, robust and reliable tool for estimating the parameters of a catastrophe model. However, there are two important problems associated with this technique. Both problems

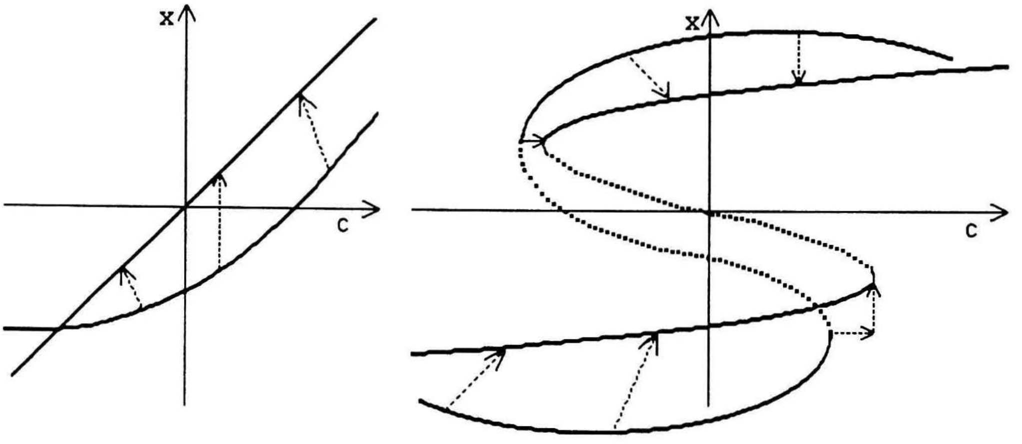

Fig 3. Equivalence of equilibrium surfaces. Two systems are equivalent if there are (smooth) co-ordinate transformations such that the equilibrium surfaces can be transformed into one another.

concern the application of co-ordinate transformations. In catastrophe theory we are interested only in the qualitative behaviour of systems. This, for instance, enables one to partitioning the class of gradient systems into a number of disjunct equivalence classes. The behaviour of two systems in an equivalent class is qualitatively the same. The qualitative properties of interest are those properties that are left unchanged under smooth co-ordinate transformations. Without going into the details, we consider the equilibrium surfaces of the systems (3). In Figure 3 we have drawn alternative equivalent models for both systems.

These are equivalent because there are co-ordinate transformations such that the equilibrium surface of one can be transformed into the other. Notice that the equilibrium surfaces of the Morse system and the cusp system are not equivalent.

The use of co-ordinate transformations is of utmost importance in catastrophe theory, because application of these lead to the celebrated classification scheme. However, in the method of Cobb only linear transformations are used. That is, the state

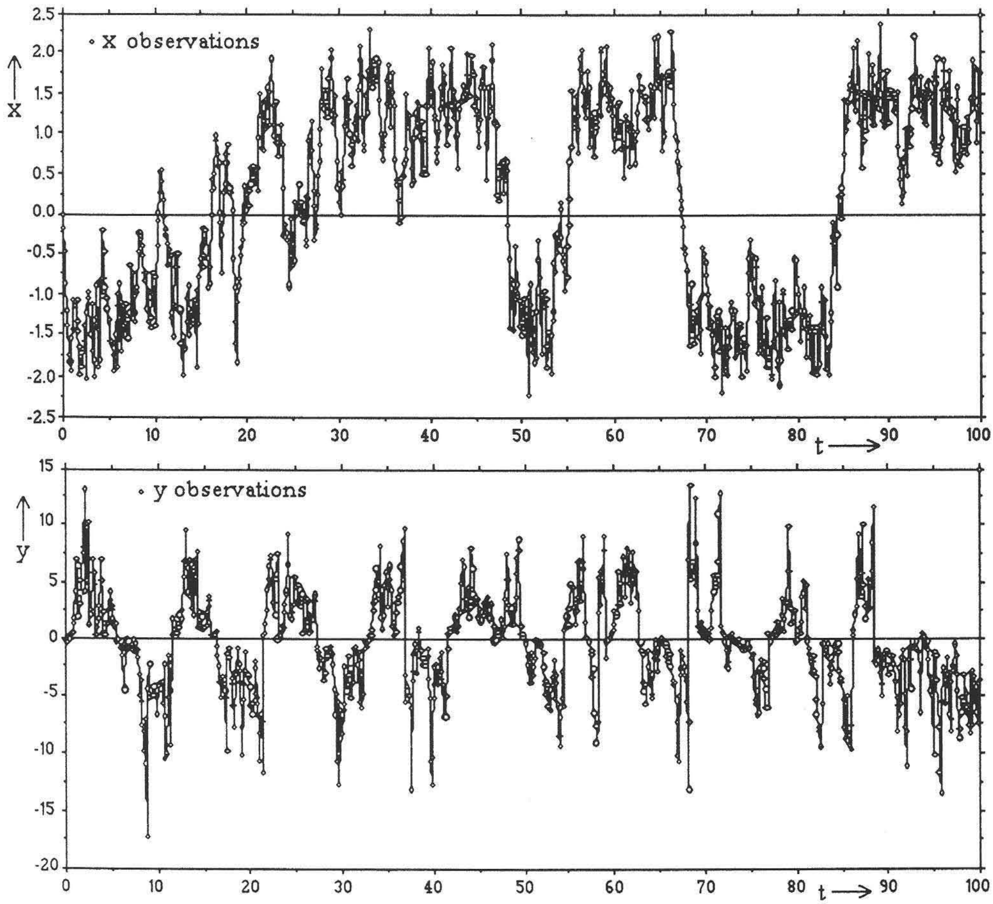

Fig. 4. Time series of two equivalent stochastic systems. The second system is obtained by transforming the state variable x of the first system by means of $y = x + x^3$.

variable is scaled linearly and the canonical parameters α, β depend linearly on the control variables c_1 and c_2 according to (6). If we want to make full use of the benefits of catastrophe theory, then the complete class of smooth co-ordinate transformations should be implemented in the method of Cobb. However, this leads to invariance and over-parametrisation problems.

Invariance

As an example of the invariance problem, consider the time series in Figure 4.

The second time series is sampled from a process that is obtained by transforming the state variable x of the first system by means of the transformation $y = x + x^3$. The control variables are fixed for both processes. According to the approach of Cobb, the SPDF gives information about the number of stable equilibrium states (or attractors). For this purpose we have depicted in Figure 5 the experimental frequency histograms of the processes in Figure 4.

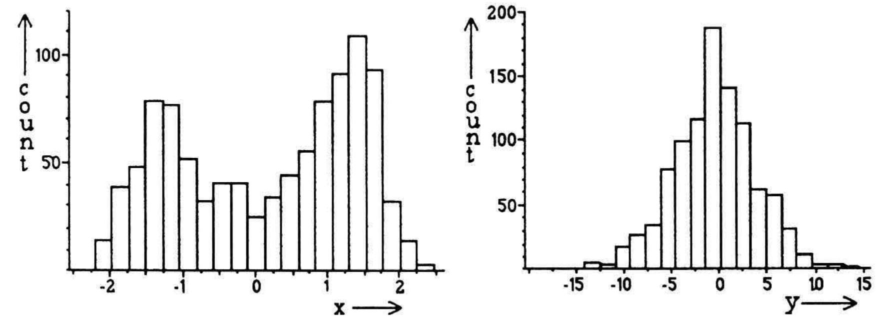

Fig. 5. Frequency histograms of time series in Figure 4.

From Figure 5 we see that the number of modes is not invariant with respect to scale transformations. This means that the stochastic equilibrium states and stochastic bifurcations, as defined by Cobb, are not invariant. The occurrence of a bifurcation thus depends on the measurement scales that are used. The invariance problem is caused by the transformation rule for PDF's, which is not in accordance with the transformation rule for potential functions. Our solution to this problem uses an alternative characterising function of a stochastic process, that transforms properly. This alternative function, which is called the level crossing function, is equal to the SPDF multiplied by the diffusion function, i.e. $\Theta(x) = p_s(x) B(x)$ (for convenience we omit here the explicit dependence on the control variables). It is easy to show that the number of modes of this function is invariant with respect to co-ordinate transformations. If the diffusion function is constant, then the level crossing function is proportional to the SPDF. Therefore, applying the level crossing function can be considered as a generalisation of the method of Cobb. Stochastic stable equilibrium states are defined as modes of the level crossing function. A stochastic bifurcation occurs when the number of these modes changes.

This newly defined characterising function is called the level crossing function, because it is directly related to certain level crossing probabilities. Let a level crossing of a level x be a pair of time series observations that lie on opposite sides of level x (see Figure 6), and let the probability of the occurrence of this event, with time step Δ, be denoted by $p_\Delta(x)$. Then it can be shown (Hartelman, 1997a) that the level crossing probability for a level x divided by $\sqrt{(\Delta/2\pi)}$ is, up to a negligible term if the time step is small, equal to the level crossing function, i.e. $p_\Delta(x)/\sqrt{(\Delta/2\pi)} = \Theta(x) + O(\sqrt{\Delta})$.

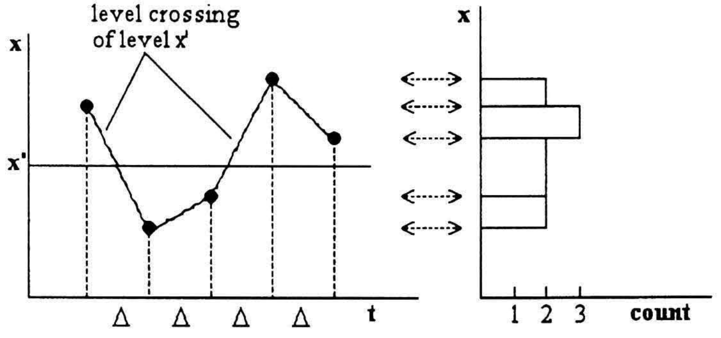

Fig. 6. Level crossing probability and construction of level crossing function.

Furthermore, the number of experimental level crossings can be used to estimate the level crossing function. Let $N_\Delta(x)$ be the experimental number of level crossings of level x from a time series with N observations. It is shown (Hartelman, 1997a) that the fraction of level crossings divided by $(\sqrt{\Delta/2\pi})$ converges almost sure, as N goes to infinity, to the level crossing function up to a negligible term of order $O(\sqrt{\Delta})$. That is

$$\Theta_{N,\Delta}(x) = N_\Delta(x) / N\sqrt{\Delta/2\pi} \qquad \Theta_{N,\Delta}(x) \xrightarrow{\text{almost sure}} \Theta(x) + O(\Delta) \tag{7}$$

The experimental level crossing plots for the time series in Figure 4 are depicted in Figure 7.

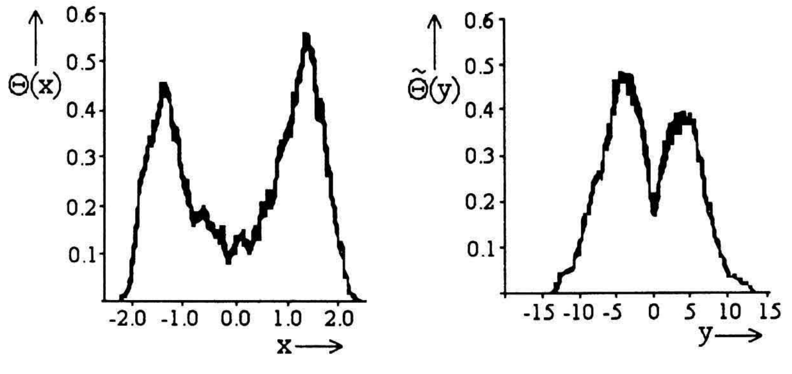

Fig. 7. Experimental level crossing plots for time series in Figure 4.

From Figure 7 we see that, according to our new definitions, and in accordance with the equivalence of the systems bistability is preserved.

Over-parametrisation

The models that are fitted by the method of Cobb are canonical catastrophe models, up to linear scale transformations. However, it is very unlikely that real systems under study are already in this canonical form. Furthermore, testing a catastrophe model versus alternative non-catastrophe models is carried out by comparing the catastrophe model with a linear model by means of likelihood statistics. However, in catastrophe theory one is only interested in the qualitative behaviour of systems, which means that in a proper hypotheses testing procedure one should compare complete equivalence classes. In contrast, in the method of Cobb, the class of catastrophe models is compared to the class of non-catastrophe models, by picking from both classes a model that is supposed to be good representative model. The linear model is in most cases a very poor representative model for the class of non-catastrophe models. As an example, consider the data in Figure 8, which are clearly sampled from a model that exhibits a rapid acceleration, which is not a transition.

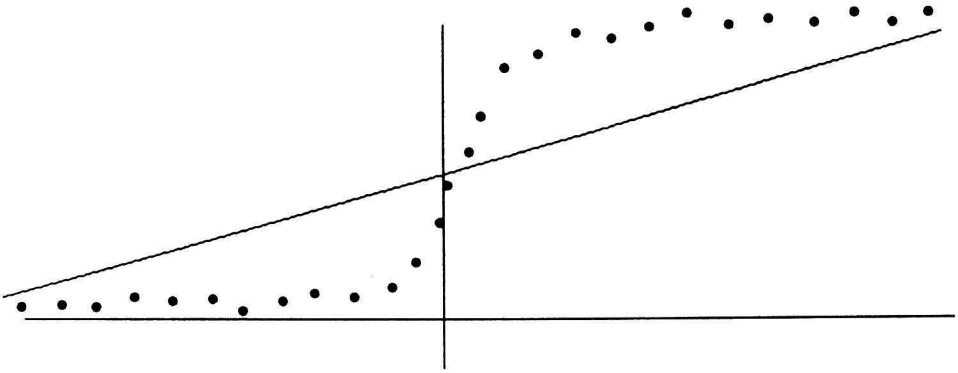

Fig. 8. Logistic growth curve data. Because a linear model does not fit to the data very well, it will be a poor representative model of the class of non-bifurcation models, in case of a bifurcation testing technique.

For these data the linear model is clearly a poor representative of the class of non-catastrophe models. A logistic growth curve would be a more appropriate model.
In order to overcome these problems one should implement the complete class of smooth co-ordinate transformations in the method of Cobb. However, this will lead to over-parametrisation problems. In catastrophe theory, one is allowed to apply a family of co-ordinate transformations ϕ (x, c), i.e. for every fixed c the measurement scale x is transformed into a measurement scale y according to y = ϕ (x, c). The only requirement is that this family of transformation functions is smooth in c. However, applying such a family of co-ordinate transformations to a data set consisting of a finite number of observations, such as in Figure 9, means that every x-value of an observation can arbitrarily be transformed, independent of the other observations.

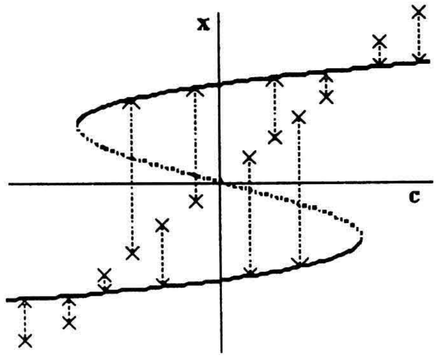

Fig. 9. Any finite set of observations fits perfectly to a cusp if the complete class of co-ordinate transformations is used.

Therefore, the observations in Figure 9, which are clearly lying on a straight line, can be transformed in such a way that they lie on a cusp equilibrium surface. In this way it is possible to fit a catastrophe model perfectly to every data set. To overcome this problem, one either has to restrict the class of allowed transformations (Cobb makes a very drastic restriction to linear transformations), or one has to make smoothness assumptions. In the latter case, smoothing the data by using kernel density estimation is an appropriate and reliable technique.

A kernel density estimate of the density of a random variable is obtained by adding (Gaussian) noise to the observations. Consider a random variable Z, with unknown density f, and observations $z_1,\ldots,z_n$. Let Z^* be a random variable which is obtained by picking randomly one of the observations z_i, and adding Gaussian noise with zero mean and standard deviation h. The kernel density estimate of f is the density of Z^*, which is

$$f_N(z) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}\,h} e^{-\frac{1}{2}\left(\frac{z - z_i}{h}\right)^2} \tag{8}$$

The 'parameter' h controls the amount of smoothing. If h is very small then the kernel density estimate is a sum of large and narrow peaks located at the observations. If h is large then the data is over-smoothed, resulting in a gaussian density with large deviation and located at the mean of the observations. Under certain conditions, with respect to h (Silverman, 1986), the kernel density estimate converges to the true density. In the same way, kernel density estimates can be used to estimate multivariate densities. The kernel method is non-parametric, in the sense that only the 'smoothing parameter' h has to be estimated. h can reliably be determined, for instance, by least-squares cross validation (see Silverman, 1986).

In contrast to the parametric technique of Cobb, one can estimate the density model by using a (multivariate) kernel density estimate of the joint probability density function. The kernel density estimates are smooth and, therefore, catastrophe theory can be applied to these estimates unrestrictedly. Mathematically analysing the kernel density estimate leads to the appropriate canonical model and the necessary co-ordinate transformations to put the model into a canonical form. In this way, the complete class of co-ordinate transformations is applied indirectly by smoothing the data.

Kernel density estimation can also be used to construct an algorithm for testing the presence of a transition. For this purpose, we make use of the mode testing algorithm of Silverman (1981; 1983). Consider the data in Figure 10 sampled from a stochastic catastrophe system with one state variable x and one control variable c.

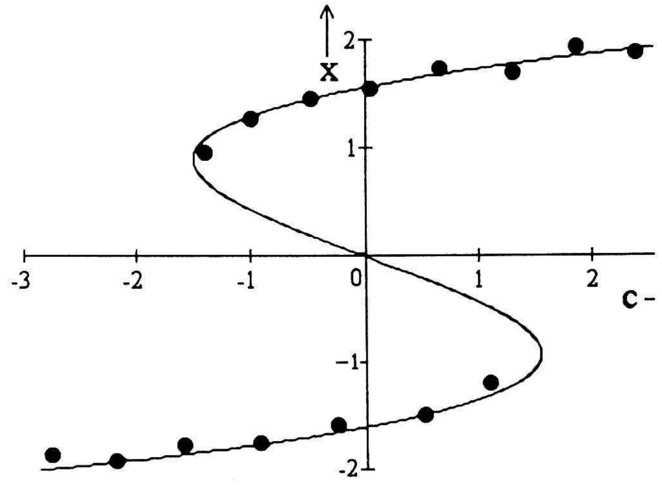

Fig. 10. Example data set and corresponding configuration of modes, with respect to x-direction, of the kernel density estimate.

For this data set we can determine the kernel density estimate of the joint probability density function of x and c. For this purpose we have to estimate, for instance by cross-validation, two smoothing parameters h_x and h_c for respectively the x- and c direction. For every c we can determine the location of the modes of the kernel density estimate in the x-direction. Similar as in the method of Cobb, these correspond to the stable equilibrium states of the system. Assume that the configuration of the modes is as depicted in Figure 10. From Figure 10 we clearly see that the system possesses two bifurcation points. It can be shown (Hartelman, 1997a) that the number of bifurcation points decreases as the amount of smoothing in the x-direction is increased, i.e. if h_x is increased or, in other words, if the standard deviation in the x-direction of the noise on the data is increased. This behaviour is illustrated by Figure 11. Increasing the standard deviation of the noise in the x-direction will eventually lead to the disappearance of the bifurcations.

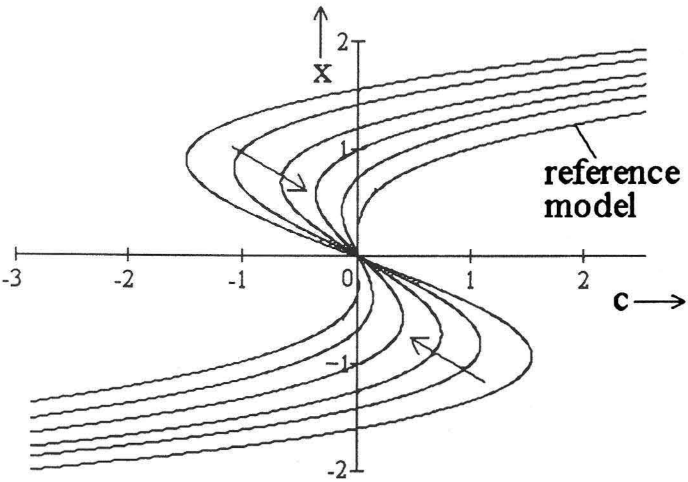

Fig.11. Kernel smoothing of an equilibrium surface. If the amount of smoothing in the x-direction (h_x) is increased, then a bifurcation will eventually disappear.

Therefore, there is, at fixed h_c, a critical amount of smoothing, such that for $h_x > h_{crit}$, the estimated equilibrium surface has no bifurcation points, whereas for $h_x < h_{crit}$ the equilibrium surface does possess a bifurcation point. Intuitively it is clear that h_{crit} will be large for models that do possess a bifurcation, whereas for non-bifurcation model h_{crit} will be small. It has been shown (Hartelman, 1997a) that one can actually use the size of h_{crit} as a statistic to test the occurrence of a bifurcation. To test the hypothesis that a bifurcation is present, one has to compare h_{crit} of the data with h_{crit} of a reference model that does not possess a bifurcation. This reference model should be chosen in such a way that the class of non-bifurcation models is given a 'fair' chance of explaining the data. An appropriate candidate reference model is the kernel density estimate with $h_x=h_{crit}$. As an example, the reference model for the data in Figure 10 is indicated in Figure 11. This model is 'on the edge' between the class of bifurcation models and the class of non-bifurcation models. Furthermore, we stay as close to the data as possible

The following algorithm is proposed to test the hypothesis that a bifurcation is present

a. Determine h_c by cross-validation with respect to the marginal kernel density of c.
b. Determine h_{crit}. The reference model is now completely known.
c. Sample data sets from the reference model.
d. Determine critical window width for sampled data sets.
e. The p-value (probability that a bifurcation is present) is the number of times that a sampled critical window width is smaller than the original h_{crit}.

In this hypotheses testing procedure one does not have to pick two appropriate models and compare them. In contrast, the reference models are data-driven. Moreover the complete class of co-ordinate transformations are implemented in an indirect way. A small simulation study (Hartelman, 1997a) shows that this technique is reliable.

Application of kernel bifurcation testing algorithm to conservation acquisition data

As an example of the use of the bifurcation testing algorithm in the previous section it has been applied to data from a conservation experiment by van der Maas (1992). The aim of the experiment was to find evidence that conservation acquisition is a stage-wise developmental process, accompagnied by transitions. The data that are discussed here are collected by using the item depicted in Figure 12, which is part of a larger computer test that was used in a longitudinal study with 101 children, ranging from 6.2 to 10.6 years of age.

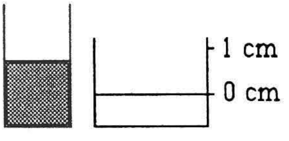

Fig 12. Test item in conservation experiment of van der Maas (1992)

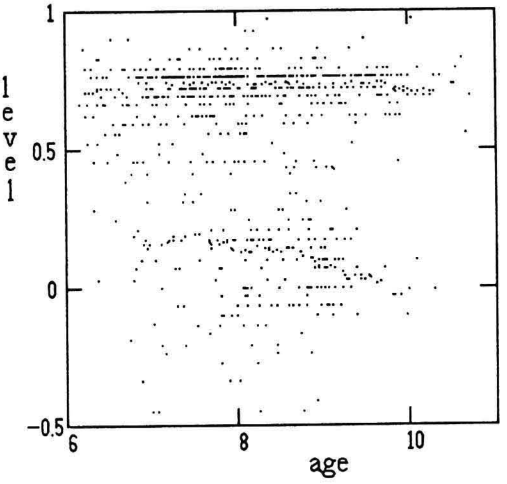

Fig. 13. Data from conservation experiment (van der Maas, 1992), for item in Figure 12. The independent and behavioural variable are respectively age (years) and water level (cm).

The item consists of two glasses of different widths. One of the glasses is filled with water, and a subject has to determine the height of the water level in the second glass, if the water was poured into this glass. The study did not include measures for control variables. For illustrative purposes we have analysed the data with age as control variable. We have omitted observations for which the level response is near the top or bottom of the glass. A margin of 0.25 cm is used. This leads (taking into account missing data) to 721 observations. It is expected that older children give the correct response, i.e. level 0. In contrast, younger children are not capable of taken into account the difference in glass widths and will probably choose a water level that is the same as in the first glass. This is supported by the data, which are plotted in Figure 13. Therefore, there is a change in response, from level 0.75 to level 0, as childen grow older. However, this does not necessarily imply that a transition is present. To investigate whether a transition is present we first determined the kernel density estimate for the data. This is plotted in Figure 14.

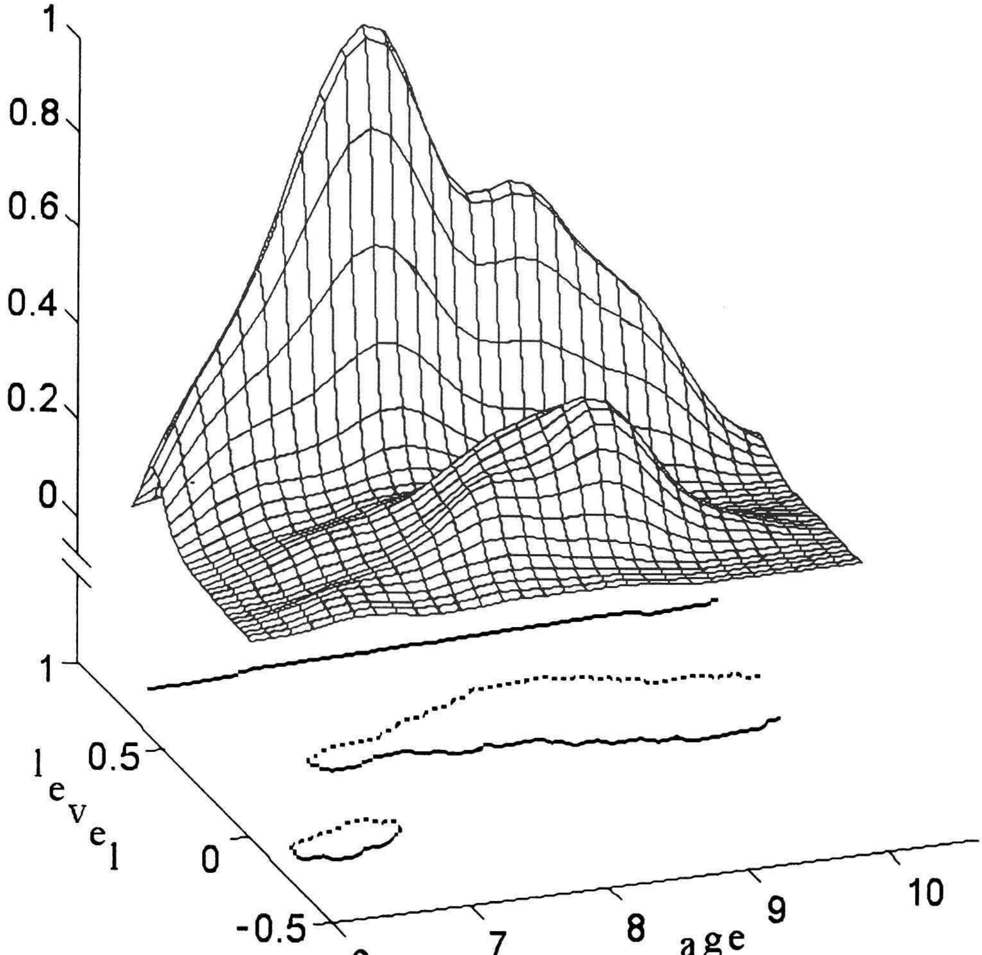

Fig. 14. Kernel density estimate for the conservation data in Figure 12.

Figure 14 also contains, in the horizontal plane, the configuration of modes of the density with respect to the level variable, which corresponds to the configuration of stable and unstable equilibrium states. Figure 14 clearly shows that there are two equilibrium states, one at the level 0.75 and one at the level 0.0. In the first case subjects respond incorrectly by giving the same height as the water level in the first glass. In the second case subjects respond correctly. From Figure 14, and in particular the height of the kernel density estimate at the modes, we also see that the probability that a level response of 0.75 is given decreases as age increases, wheras the probability that a correct response is given increases (the fact that the probability of a correct response decreases for subjects over age 10 is clearly caused by the small number of observations in this region). Figure 14 gives us a strong indication that there is a bifurcation point at ±7 years (we ignore the small circle at the bottom of the Figure 14, because of the small number of observations in the neighbourhood of this circle). In order to ascertain that there is indeed a bifurcation we applied the kernel estimate algorithm in the previous section. This resulted in a p-value, i.e. the probability that a bifurcation is present, of $p=0.87$.

Discussion and conclusions

With catastrophe theory it is possible to develop a methodology for testing the presence of transitions in experimental research. It delivers relevant and applicable definitions of a stage and a transition. Furthermore, the classification scheme can be very useful if one does not have clear knowledge about the appropriate model.

Since the development of catastrophe theory (Thom, 1975) in the early seventies there has only been constructed a few techniques for estimating catastrophe models. Besides the method of Cobb, we mention the regression technique of Guastello (1988), and the least squares technique GEMCAT (Oliva, Desarbo, Day & Jedidi, 1987). With the method of Guastello one is not capable of distinguishing between data from a catastrophe model and completely noisy data (Alexander, Herbert, Deshon & Hanges, 1992). With GEMCAT one is not capable of distinguishing between a stable equilibrium state and an unstable equilibrium state (Hartelman, 1997a).

The method of Cobb is a reliable technique for estimating a catastrophe model using cross-sectional data. However, the domain of application of this method is rather restricted. It can only be used if the noise is homogeneous, i.e. constant for every state variable value, and if linear transformations are sufficient to obtain a canonical form. Furthermore, the statistics which are used by Cobb are not sufficient to reliably test hypotheses with respect to the presence of a transition.

If the stochastic noise term is not homogeneous, then, instead of using the probability density function, one should apply an alternative function, such as the level crossing function, to characterise the stochastic system. A stochastic equilibrium state and transition, defined with respect to the qualitative form of the level crossing function, correspond to qualitative properties of the system. This means that the occurrence of a transition does not depend on the measurement scales. The level crossing function can easily be estimated by using time level crossing properties of time series data. However, experience has shown that, in general, the number of observations should be large and the size of the time step should be small in order to obtain reliable estimates.

If the noise is homogeneous, then the over-parametrisation problem in the method of Cobb can be solved by smoothing the (cross-sectional) data using kernel density estimation. In this way the complete class of co-ordinate transformations can be used. As a result it is possible to develop a transition testing algorithm that is in complete correspondence with catastrophe theory. That is, it is possible to test models with respect to their qualitative properties. In contrast to the method of Cobb, that consists of a comparison of a canonical cusp model and a linear model, the technique in §4.2, compares the class of transition models with the class of non-transition models.

The application of the kernel hypotheses testing technique to the conservation data of §5 is merely discussed as a first example of the statistical technique. However, the p-value as well as the form of the kernel density estimate in Figure 14, strongly indicate that a transition is present. The actual conservation test of van der Maas (1992) consisted of 4 items. The kernel technique has also been applied to the remaining items. The results for two of those items are similar as the one discussed in this paper (Hartelman et al., 1998).

References

Alexander, R. A., G. R. Herbert, R. P. Deshon and P. J. Hanges (1992). An examination of least squares regression modeling of catastrophe theory. *Psychological Bulletin,* **111**(2), 366-374.

Brainerd, C. J. (1978). The stage question in cognitive-developmental theory. *Behavioral and Brain Sciences*, **2**, 173-213.

Castrigiano, D. P. L. and S. A. Hayes (1993). *Catastrophe theory, Reading*, Massachusets. Addison-Wesley.

Cobb, L. (1978). Stochastic catastrophe models and multimodal distributions. *Behavioral Science*, **23**, 360-373.

Cobb, L. (1980). Estimation theory for the cusp catastrophe model. *Proceedings of the Section on Survey Research Methods*, (pp. 772-776). Washington, DC: American Statistical Association.

Cobb, L. (1981). Parameter estimation for the cusp catastrophe model. *Behavioral Science*, **26**, 75-78.

Fisher, K. W. and L. Silvern, (1985). Stages and individual differences in cognitive development. *Annual Review of Psychology*, **36**, 613-648.

Flavell, J. H. (1971). Stage-related properties of cognitive development. *Cognitive Psychology*, **2**, 421-453.

Gilmore, R. (1981). *Catastrophe theory for scientists and engineers.* New York: Wiley.

Guastello, S. J. (1988). Catastrophe modeling of the accident process: Organizational subunit size. *Psychological Bulletin,* **103**, 246-255

Hartelman, P. A. I. (1997a). *Stochastic Catastrophe Theory.* Ph. D dissertation, University of Amsterdam, Dept. of Psychology.

Hartelman, P. A. I., H. L. J. van der Maas, P. C. M. Molenaar (1998). Detecting and modelling Developmental Transitions. *British Journal of Developmental Psychology*, 16, 97-122.

Maas, van der, H. L. J. and P. C. M. Molenaar (1992). Stagewise cognitive development: An application of catastrophe theory. *Psychological Review*, **99**, 395-417.

Oliva, T. A., W. S. Desarbo, D. L. Day and K. Jedidi (1987). GEMCAT: A General Multivariate Methodology for estimating Catastrophe Models. *Behavioral Science,* **32**, 121-137.

Poston, T. and I. Stewart (1978). *Catastrophe theory and its applications.* London: Pitman

Piaget, J. (1960). The general problems of the psychobiological development of the child. In J. M. Tanner and B. Inhelder (Eds.), *Discussions on child development*, Vol. 4, (pp. 3-28), London: Tavistock.

Silverman, B. W. (1981). Using kernel density estimates to investigate multimodality. *J. Roy. Statist. Soc. B,* **43**, 97-99.

Silverman, B. W. (1983). Some properties of a test for multimodality based on kernel demsity estimates. In Kingman, J. F. C. and Reuter, G. E. H. (eds), *Probability, Statistics and Analysis*, (pp. 248-259). Cambridge: Cambridge University Press.
Silverman, B. W. (1986). Density estimation for Statistics and Data analysis. *Monographs on Statistics and Applied Probability*, 26. London: Chapman and Hall.
Thom, R. (1975). *Structural stability and morphogenesis*. Reading, MA.: Benjamin.

Faculty of Psychology, University of Amsterdam