Journal of Physics: Conference Series **268** (2011) 012001 doi:10.1088/1742-6596/268/1/012001

Distribution of return point memory states for systems with stochastic inputs

A Amann[1], M Brokate[2], D Rachinskii[3] and G Temnov[4]

[1]Tyndall National Institute, University College Cork, Ireland
[2]Zentrum Mathematik - M6, Technische Universität München, Germany
[3]Department of Applied Mathematics, University College Cork, Ireland
[4]Department of Mathematics, University College Cork, Ireland

E-mail: d.rachinskii@ucc.ie

Abstract. We consider the long term effect of stochastic inputs on the state of an open loop system which exhibits the so-called return point memory. An example of such a system is the Preisach model; more generally, systems with the Preisach type input-state relationship, such as in spin-interaction models, are considered. We focus on the characterisation of the expected memory configuration after the system has been effected by the input for sufficiently long period of time. In the case where the input is given by a discrete time random walk process, or the Wiener process, simple closed form expressions for the probability density of the vector of the main input extrema recorded by the memory state, and scaling laws for the dimension of this vector, are derived. If the input is given by a general continuous Markov process, we show that the distribution of previous memory elements can be obtained from a Markov chain scheme which is derived from the solution of an associated one-dimensional escape type problem. Formulas for transition probabilities defining this Markov chain scheme are presented. Moreover, explicit formulas for the conditional probability densities of previous main extrema are obtained for the Ornstein-Uhlenbeck input process. The analytical results are confirmed by numerical experiments.

1. Introduction

1.1. Preisach memory models driven by random input

The response of systems characterised by internal state variables or memory to stochastic inputs, noise and random fluctuations is an area that has received much recent attention in both theoretical and applied studies. Nontrivial effects such as stochastic resonance [1–3] and coherence resonance [4] can be observed in the simplest noise driven bistable or oscillatory systems. In particular, two-state systems such as a non-ideal relay (also known as a non-ideal switch, Schmitt trigger, rectangular loop operator and binary memory element in different applications) or a particle in a double well potential exhibit stochastic resonance in response to stochastic perturbations. This effect underpins the methods of dithering widely used in engineering design, which exploits the constructive role of noise for improving the quality of the output in image processing and signal processing [5].

Various models of systems with memory can be constructed using the formalism of the phenomenological *Preisach model* as a weighted superposition of many two-state non-ideal relays that are individually and independently driven by the same input [6]. For example, the response of the Preisach model to stochastic inputs has been studied in [7, 8] in order to model the *viscosity* or *after-effect* in ferromagnetic materials, which is detrimental for magnetic recording technologies. A similar approach to modelling *creep* in semiconductor materials by means of stochastically driven Preisach memory systems has been adopted in [9]. While in these examples the source of the stochastic component of the input is the internal noise caused by thermal fluctuations, in other areas of engineering and physics the input is stochastic due to its random nature or because it includes an external noise component superimposed on a deterministic input signal. Ubiquitous applications of the Preisach model in magnetizm, superconductivity, piezoelectricity, elastoplasticity, shape memory alloys, economics and finance, modelling capillary hysteresis in flows through porous media such as soils etc. (see, for example, the surveys in [10, 11]) motivated further analysis of the Preisach operator input-output relationship for stochastic inputs of various types including uncorrelated inputs and general diffusion processes [12–17]. This research was focused mainly on finding characteristics of the output time series such as the spectral density and the autocorrelation function. The method was based on the decomposition of the Preisach model into the sum (integral) of two-state non-ideal relays, *i.e.*, independent elementary hysteretic transducers, each producing a binary stochastic output in response to the common input. The decomposition reduced the problem to analysis of the cross-correlation function of the output of a pair of relays, which was then averaged over the set of relays composing the Preisach model. Key computations in this analysis were based on the relation between randomly induced switching of relays and the well studied escape (exit) problems for stochastic processes and, furthermore, partly involved some ingenious techniques such as, for example, the theory of diffusion processes on graphs [18, 19]. This approach resulted, in particular, in analytic and numerical estimates for the spectral density of the output time series, asymptotic formulas for the long-time tales in its autocorrelation, and analysis of $1/f$ noise in the frequency domain.

1.2. Problem statement

Modelling of real world systems is often complicated by the existence of internal degrees of freedom which depend on the history of the system and give rise to hysteretic behaviour. In addition it is often the case that the system of interest has been subject to a random input in the past, but neither this input nor the resulting system evolution are known at a given point in time. Examples of systems with internal memory and random input include magnetization dynamics of magnetic rock in geology, folding dynamics of proteins in biology and modelling of trends, fads and opinions in sociology. Under these circumstances the question arises, *whether it is possible to characterise the internal state or memory structure of the system at a given instant T assuming that this state results from the effect of a random input on the system between some moment in the past and the moment T.* It is the purpose of this paper to investigate this question.

In the following, we will focus on systems with *rate-independent* input-output relationship and *return point memory*, which are characteristic for many real-world systems. These two properties are important idealisations which effectively reduce the relevant memory structure of the system to a specific sequence of shock values (known as *main extrema*) of the input experienced in the past and can generically be modelled by the Preisach operators and their generalisations. Rate-independence of an input-output relationship means that this relationship is invariant with respect to the action of the group of affine transformations of the time scale, implying that shock values of the input have permanent effect on the output in the future and thus are memorised by the system. The magnetization of ferromagnetic material by a fluctuation of external magnetic field which creates a permanent magnet is one example of such memory. In the mathematical theory of hysteretic systems, hysteresis nonlinearities are defined (omitting a few technicalities) as deterministic rate-independent operators with non-local memory [20, 21]. This general definition entails a set of non-trivial properties of hysteresis operators which are sufficient for developing formal concepts with various applications [22–26]. The rate-independence property distinguishes hysteretic memory from other memory types such as in delay differential equations or convolution operators. It is this property that allows one to describe constitutive hysteretic relationships in terms of 2-D input-output diagrams such as the familiar picture showing the nested structure of magnetization curves and hysteresis loops, which present the relationship between the magnetization and the applied magnetic field in ferromagnetic materials.

The *return point memory* (also known as the wiping out memory) property of a rate-independent system ensures that hysteresis loops corresponding to periodic inputs are closed on the input-output diagram (hence there is no effects like ratcheting), which is a reasonable approximation for many real world hysteretic systems. We will use an alternative definition due to [25] whereby the memory state $\omega(T)$ at a given instant T is a vector (with changing dimension), or an infinite sequence, of certain extremum values of the input achieved prior to this instant; the evolution of the state in response to the evolution of the input is defined by a set of simple updating rules dictating which extremum values are included in the vector $\omega(t)$. In particular, for piecewise monotone inputs (or discrete time inputs), each new extremum is included in $\omega(t)$ as the last component at the instant it is achieved; while, whenever the input reaches the value equal to the last component of the vector $\omega(t)$, the last two components are removed from $\omega(t)$. The memory state $\omega(T)$ is a concise record of all those events which happened prior to the moment T and can influence dynamics of the system after this moment in the sense that the state $\omega(T)$ and the input after the moment T completely define the state $\omega(t)$ and the output of the system for $t > T$.

For stochastic inputs which drive the system between moments t_0 and T the memory state $\omega(T)$ is a multidimensional (or infinite-dimensional) random variable. *The goal of this paper is to find, or to characterise, the distribution of the memory state $\omega(T)$ and its parameters at a given moment T for several classes of inputs such as a random walk, Wiener process, Ornstein-Uhlenbeck process and a general stationary diffusion process. An important assumption we make is that the time interval between the moments t_0 and T is large meaning that the initial state $\omega(t_0)$ has negligible effect on the distribution of $\omega(T)$.* More technically, we consider the probability that an input trajectory is confined to a prescribed interval for all times between t_0 and T and assume that the time interval is long enough to make this probability negligibly small.

Thus, our focus is on a *multidimensional memory configuration*, *i.e.*, a detailed description of the system's state at a *single* time moment. This compares to the problem described in the previous subsection focusing on the description of the output of the Preisach model, *i.e.*, a *scalar-valued output time series*. The output value at a given instant can be viewed as an average value of the components of the vector state at this instant. Hence, the two problems are complementary. The most detailed description of system's dynamics would provide information about the multi-dimensional stochastic process with the values in the space of memory states, *i.e.*, description of the stochastic evolution of the state in response to the evolution of a stochastic input. However, the problem in this generality appears to be difficult to approach.

In applications, the state of the Preisach model described in terms of the binary states of all its relay components often models the physical state of the system. The memory vector ω, at a given moment, contains a compact description of the binary function representing the physical state. More precisely, ω can be mapped straightforwardly to this binary function by a procedure proposed in [22] whereby the components of ω encode the staircase boundary separating the domain which represents the relays that are 'on' and the domain which represents the relays that are 'off' on the Preisach plane.

Summarising and extending the discussion of this subsection, let us assume that we control or observe the input of the Preisach model only after a certain instant T. As the output and the state at any instant $t > T$ are defined both by the dynamics of the input between the instants T and t and by the memory state $\omega(T)$, which has recorded the shock values (the main extrema) of the input obtained at times prior to T, information about the state $\omega(T)$ is important for prediction and control of the dynamics of the system after the moment T. In particular, the return point memory property of the Preisach model ensures that the shock values stored as components of the vector $\omega(T)$ will be felt by the systems at the moments $t > T$ when the input reaches again those recorded values. At these moments, the composition of the memory vector $\omega(t)$ changes and, as a result, those moments are characterised by a switch to another input-output branch on the input-output diagram of the model, such as the switch to another magnetization curve on the diagrams depicting the dependence of the magnetization of ferromagnetic materials on the external magnetic field. The jump to another input-output branch is accompanied by a jump of the time-derivative of the output, which is smooth at other times for smooth input time series. Therefore information about the state $\omega(T)$ translates directly to the list of input values for which these jumps (possibly, representing a strong impact on the system) will happen after the moment T, allowing one to predict them.

If the input was not controlled or measured prior to the instant T, then the state $\omega(T)$ is not known or controllable, neither it is typically measurable. However, we make the assumption that the memory state $\omega(T)$ is the result of a random input with known parameters, which has effected the system prior to the instant T, and characterise the distribution and typical parameters of the memory state $\omega(T)$ under this assumption. Because the Preisach nonlinearity has the return point memory property, an equivalent question to ask is what is the distribution of the main extrema of the random process which creates the memory state $\omega(T)$. We approach this question assuming that the random input was effecting the system for a long period of time prior to the moment T.

1.3. Method and results

The two main ingredients of analysis developed in the work, which we refer to in Subsection 1.1, are (a) some form of the time reversion applied to the input; and, (b) reduction of the problem to a one-dimensional escape (exit) problem, or a series of such problems, for a diffusion process. The same two ingredients underpin the approach we adopt here. However, the escape problem component is adapted appropriately as we work with the memory states in the compact form ω of a vector of the past shock values of the input rather than the binary functions representing the 'physical state' of the Preisach model. In particular, we do not resort to the decomposition of the model into the superposition of binary relays.

In order to describe the distribution of the random vector (sequence) $\omega = \omega(T)$, we introduce a Markov chain scheme where the running index n of the component of ω plays the role of the fictitious time. Using the above two ingredients of the analysis, we derive explicit expressions for the transition probabilities of this Markov chain in terms of the SDE defining the input. For the random walk and continuous time Wiener process inputs, along this line of argument, we obtain explicitly the distribution function for the state $\omega(T)$. We then show how to use this result to derive distributions of several random variables naturally associated with $\omega(T)$. In particular, we show that the number q_d of the components of $\omega(T)$ which have values in a given interval I is an exponentially (Poisson) distributed random quantity, and discuss scaling of $q_d = q_d(I)$ with the ends of the interval I, thus obtaining a chrcterisation of the dimension of a typical memory state $\omega(T)$. The random quantity q_d can be viewed as a measure of the amount of information stored in $\omega(T)$ (complexity of $\omega(T)$), or a measure of the effect of the past history of the input on dynamics after the instant T. It is also directly related to the description of the system's state in terms of the input-output diagram. In this description, one distinguishes the input-output curves of different order, such as the primary magnetization curve, secondary magnetization curves and k-th order magnetization curves of any higher order k in the magnetic hysteresis applications. A particular value of q_d places the system to the input-output curve of order $k = q_d$ (exactly, or with controllable accuracy) at the moment T.

Considering general diffusion input processes and using the expressions for the transition probability densities of the subsidiary Markov chain, we derive formulas for the conditional probability density of the n-th component ω^n of the memory vector $\omega = \omega(T)$ conditioned on the values of the two previous components ω^{n-1} and ω^{n-2}. These formulas can be used to obtain the joint probability distribution of those components ω^n whose values are confined to a given interval I. The formulas are specified for the Ornstein - Uhlenbeck processes and compared to their counterparts for the Wiener process with and without drift.

We confirm analytical results by numerical experiments.

Our results are equally applicable to all rate-independent systems with return-point memory as such systems are characterised by the same input-state relationship mapping the input to the memory state $\omega(t)$ (see, [25]). The state-output relationship is model specific. For example, the state-output relationship of the Preisach model ensures that for any periodic input all the hysteresis loops on the input-output diagram are congruent; according to Mayergoyz's theorem this property identifies the Preisach model among all the return-point memory systems, ensuring the decomposition into the sum of relay operators [23]. Another important class of systems with return point memory are spin interaction models stemming from the classical Ising model of the array of coupled spins. The Ising model based approach is quite general and is widely used in many disciplines to study such effects as avalanches, Barkhausen noise, clustering, emergent hysteresis behaviour, and order-disorder phase transitions which are universal to many systems; examples include ferromagnetic materials and earthquake fault systems [27], sand piles [28], capillary effects and percolation in partially saturated porous media [29], phase transitions in solids [30], random networks [31,32], cellular automata [33], and multi-agent models in economics and finance [34]. The roots of this universality, and the universal scaling laws characterising such systems, have been recently revealed in [35,36] (see also the review [37] and references therein). Sethna et al. showed that any system of spins with positive interactions (such as, for example, those in ferromagnetic materials) has the return point memory property [38]. The return point memory state ω recording the shock values of the input can be mapped to the 'physical state' of the model describing the binary state of each spin.

It is worthwhile to note that the so-called *rainflow counting method* which is widely used in damage and fatigue estimates for engineering applications is based on tracking the evolution of a return point memory state $\omega(t)$ driven by a stochastic input [39]. However, it applies to input processes with smooth trajectories, while the focus of this paper is on nonsmooth processes.

This paper, as well as the literature cited in Subsection 1.1, refers to open loop systems. Closed loop stochastic models of simple hydrological systems involving the Preisach operator have been studied in [40] by a combination of analytic and numerical methods. The input of the Preisach operator in these models is yet smooth. Closed loop models involving systems of many switches with jumping thresholds driven by nonsmooth stochastic inputs have been recently proposed and investigated numerically in [41–45] in the context of modelling economics and finance.

The rest of the paper is organised as follows. In the next section, we present the definition of the return point memory states ω and define their evolution in response to the evolution of the input. In Section 3, we discuss systems driven by the random walk input. The purpose of this section is to illustrate the main steps of our approach on the simplest example. Section 4 presents the results for systems driven by continuous time diffusion processes.

2. Memory structure

Consider a continuous scalar-valued input $x(t)$ of the Preisach nonlinearity on the interval $[t_0, T]$. At every time moment t, the memory of the system consists of two arrays of numbers m_k^t and M_k^t $(k = 1, 2, \ldots)$. Here M_1^t, m_1^t are the running global maximum value and global minimum value of x over the interval $[t_0, t]$ defined via

$$M_1^t = \max_{t_0 \le \tau \le t} x(\tau), \quad m_1^t = \min_{t_0 \le \tau \le t} x(\tau). \tag{1}$$

The other entries of the memory at the moment t are defined recursively. In the case where the global maximum is obtained before the global minimum, *i.e.*

$$\tau^t(M_1^t) \le \tau^t(m_1^t), \tag{2}$$

where

$$\tau^t(c) := \max\{\tau \in [t_0, t] : x(\tau) = c\}, \tag{3}$$

we define

$$m_k^t = \min_{\tau^t(M_k^t) \le \tau \le t} x(\tau), \quad M_{k+1}^t = \max_{\tau^t(m_k^t) \le \tau \le t} x(\tau), \quad k = 1, 2, \dots \tag{4}$$

In other words, M_2^t is the maximum value of x on the time interval $[\tau^t(m_1^t), t]$ from the moment $\tau^t(m_1^t)$ when the input achieves the value m_1^t for the last time to the moment t; m_2^t is the minimum value of x on the time interval $[\tau^t(M_2^t), t]$ from the moment $\tau^t(M_1^t)$ when the value M_2^t is achieved for the last time to the moment t, and so on. We will refer to M_k^t and m_k^t as the *main extrema* of x on the interval $[t_0, t]$ or as the *memory elements* of the system, and will call the set of these elements the *memory array* $\omega = \omega(t)$.

Similarly, if the global minimum is obtained before the global maximum, *i.e.*

$$\tau^t(m_1^t) < \tau^t(M_1^t) \tag{5}$$

we define

$$M_k^t = \max_{\tau^t(m_k^t) \le \tau \le t} x(\tau), \quad m_{k+1}^t = \min_{\tau^t(M_k^t) \le \tau \le t} x(\tau), \quad k = 1, 2, \dots \tag{6}$$

Directly from the definition (and under assumption (2)) we observe the following inequality chains

$$\tau^t(M_1^t) \le \tau^t(m_1^t) \le \tau^t(M_2^t) \le \tau^t(m_2^t) \le \cdots \le t \tag{7}$$
$$M_1^t \ge M_2^t \ge \cdots \ge x(t) \tag{8}$$
$$m_1^t \le m_2^t \le \cdots \le x(t) \tag{9}$$
$$M_1^t - m_1^t \ge M_2^t - m_2^t \ge \cdots \ge 0. \tag{10}$$

If equality between any two elements in equations (7)–(9) holds, then these two elements as well as all the elements to the right of them are equal to the final value of the process (or the final element of the chain).

We will also consider discrete time input sequences $x = x(t)$ with integer $t = 0, 1, 2, \dots$ To define the main extrema for them, one considers the piecewise linear extension of $x(t)$ (which is linear between any neighbouring integer times) to a segment of the real line, and applies formulas (2)–(6) to this extension.

For any given moment t, the number of different memory elements can either be finite, for example if x is a discrete time input, or infinite, which is typical for realisations x of a continuous Markov process. The number of different memory elements in the memory array, if finite, is generally variable in time. With increasing time the memory array is updated according to definition (2)–(6) by adding and deleting the appropriate memory elements. Assume for example that we want to obtain the memory array at time $t+\Delta$ starting from a known set of main extrema

M_k^t and m_k^t at time t. Let k_0 be the largest index such that $x(t')$ with $t' \in [t, t+\Delta]$ does not leave the interval of the associated main extrema, *i.e.*

$$k_0 = \max_k \left\{ k : x\left([t, t+\Delta]\right) \subset \left[m_k^t, M_k^t\right] \right\}. \tag{11}$$

Then the memory elements for indices less or equal k_0 are unchanged

$$M_k^{t+\Delta} = M_k^t, \quad m_k^{t+\Delta} = m_k^t \quad \text{for } k = 1, \dots, k_0 \tag{12}$$

and only the memory elements with indices larger than k_0 need to be updated according to (4), (6). These definitions naturally extend the rules for deleting and adding memory elements in response to piecewise monotone inputs to any continuous inputs, see, for example, a detailed discussion in [22]. In particular these definitions also apply if the input is not piecewise monotone, which is typically the case for stochastic inputs.

In this paper we are concerned with stochastic inputs and our aim is to characterise the distribution of the memory array, which becomes a multi-dimensional random quantity (sequence) in this case. In the following section we will first consider a simple classical random walk input process before we turn to more general Markov processes in Section 4.

3. Memory distribution for random walk

3.1. Time reversion and reduction to escape problem

Consider a classical random walk W_t on the integer grid as an input $x(t)$ of the Preisach memory operator, hence the time t is discrete. The point W_t is assumed to move one step left or right with equal probability 1/2, *i.e.*,

$$P[W_{t+1} = W_t + 1] = P[W_{t+1} = W_t - 1] = 1/2.$$

For the purposes of this section, we use the value $x(t)$ as a reference point and characterise the memory at the moment t by the array of values $m_k^t - x(t), M_k^t - x(t)$. For a finite time interval there exist only a finite number of different memory elements

$$m_1^t < m_2^t < \cdots < m_{L_t}^t < x(t) < M_{N_t}^t < \cdots < M_2^t < M_1^t \tag{13}$$

with $|L_t - N_t| \le 1$, while $m_k^t = M_n^t = x(t)$ for all $k > L_t$, $n > N_t$. Moreover, we reorder the entries of the finite memory array (13) to obtain the *memory string*

$$\mathfrak{M}_t = \{\mathfrak{m}_{-L_t}^t, \dots, \mathfrak{m}_{-1}^t, \mathfrak{m}_1^t, \dots, \mathfrak{m}_{N_t}^t\} \tag{14}$$

with

$$\mathfrak{m}_{-L_t}^t < \cdots < \mathfrak{m}_{-2}^t < \mathfrak{m}_{-1}^t < 0 < \mathfrak{m}_1^t < \mathfrak{m}_2^t < \cdots < \mathfrak{m}_{N_t}^t$$

where

$$\mathfrak{m}_{-k}^t = m_{L_t-k+1}^t - x(t), \qquad \mathfrak{m}_n^t = M_{N_t-n+1}^t - x(t)$$

for $k = 1, \dots, L_t$, $n = 1, \dots, N_t$, $|L_t - N_t| \le 1$. Thus, $\mathfrak{M}_t$ is a reordered memory array of the input $y(\tau) = x(\tau) - x(t)$ on the interval $t_0 \le \tau \le t$; $\mathfrak{m}_k^t$ are local maximum values of y for $k > 0$ and local minimum values for $k < 0$. If τ_k^t is the last moment when y reaches the value $\mathfrak{m}_k^t$ before the moment t, *i.e.*, $\tau_k^t = \max\{\tau \in [t_0, t] : y(\tau) = \mathfrak{m}_k^t\}$, then either $t > \tau_1^t > \tau_{-1}^t > \tau_2^t > \tau_{-2}^t > \cdots$ or $t > \tau_{-1}^t > \tau_1^t > \tau_{-2}^t > \tau_2^t > \cdots$, hence the ordering of the moments τ_k^t is defined by the sign of the difference $\tau_1^t - \tau_{-1}^t$. As the ordering is a part of the memory, we will distinguish between the memory strings with the same set of values (14) for $\sigma^t = 1$ and $\sigma^t = -1$ where

$$\sigma^t = \operatorname{sign}(\tau_1^t - \tau_{-1}^t).$$

Note that with this notation the number N_t of positive entries of the memory string and the number L_t of negative entries are related by either $N_t = L_t$ or $N_t = L_t + \sigma^t$.

We will be interested in the distribution of the memory string entries $\mathfrak{m}_k^T$ which lie in a certain interval $-K \le \mathfrak{m}_k^T \le K$. If we continue the input after the moment T, then the other memory string entries do not change until $x(t) - x(T)$ reaches one of the values $\pm K$ at some moment[1] $\tau > T$. Thus, for a vector of integers j_k, σ satisfying

$$-K < j_{-L} < \cdots < j_{-2} < j_{-1} < 0 < j_1 < j_2 < \cdots < j_N < K; \quad |\sigma| = 1$$

with either $L = N$ or $L = N + \sigma$, we are interested in the probability $P_{j_{-L},\dots,j_N,\sigma}$ of the event

$$\mathfrak{m}_{-L}^T = j_{-L}, \dots, \mathfrak{m}_{-1}^T = j_{-1}, \mathfrak{m}_1^T = j_1, \dots, \mathfrak{m}_N^T = j_N; \quad \text{sign}(\tau_1^T - \tau_{-1}^T) = \sigma \tag{15}$$

which implies $L \le L_T, N \le N_T$.

To obtain the desired distribution, we reverse the time and consider a sequence of escape problems for the reversed process.

Consider the reversed process $\hat{W}_t = W_{T-t} - W_T$ for $0 \le t \le T - t_0$. This is also a random walk with the probability 1/2 to go one step left or right starting from zero. Consider the hitting times

$$\hat{\tau}_k = \min\{t \ge 0 : \hat{W}_t = k\} \tag{16}$$

of $\hat{W}_t$ for integer $k \ne 0$ with $|k| \le K$. We assume that $T \gg t_0$, and hence identify $\hat{W}_t$ with a random walk on the infinite interval $t \ge 0$. More precisely, we neglect the probability that $|W_t - W_T|$ never reaches the value K over the time interval $t_0 \le t \le T$. In this approximation, all the hitting times (16) are well-defined and finite, as they are almost surely for the random walk on the infinite time interval $t \ge 0$.

We now observe that under this assumption the event (15) is equivalent to the event

$$\hat{\tau}_{j_1} < \hat{\tau}_{-1} \le \hat{\tau}_{j_{-1}} < \hat{\tau}_{j_1+1} \le \hat{\tau}_{j_2} < \hat{\tau}_{j_{-1}-1} \le \hat{\tau}_{j_{-2}} < \hat{\tau}_{j_2+1} \le \cdots \le \hat{\tau}_{j_N} < \hat{\tau}_{j_{-L}-1} < \hat{\tau}_{j_N+1} \tag{17}$$

for $N = L + 1$, and to the event

$$\hat{\tau}_{j_1} < \hat{\tau}_{-1} \le \hat{\tau}_{j_{-1}} < \hat{\tau}_{j_1+1} \le \hat{\tau}_{j_2} < \hat{\tau}_{j_{-1}-1} \le \hat{\tau}_{j_{-2}} < \hat{\tau}_{j_2+1} \le \cdots \le \hat{\tau}_{j_{-L}} < \hat{\tau}_{j_N+1} < \hat{\tau}_{j_{-L}-1} \tag{18}$$

for $N = L$ if $\sigma = 1$. Similarly, if $\sigma = -1$, then (15) is equivalent to

$$\hat{\tau}_{j_{-1}} < \hat{\tau}_1 \le \hat{\tau}_{j_1} < \hat{\tau}_{j_{-1}-1} \le \hat{\tau}_{j_{-2}} < \hat{\tau}_{j_1+1} \le \hat{\tau}_{j_2} < \hat{\tau}_{j_{-2}-1} \le \cdots \le \hat{\tau}_{j_N} < \hat{\tau}_{j_{-L}-1} < \hat{\tau}_{j_N+1} \tag{19}$$

for $N = L$, and to

$$\hat{\tau}_{j_{-1}} < \hat{\tau}_1 \le \hat{\tau}_{j_1} < \hat{\tau}_{j_{-1}-1} \le \hat{\tau}_{j_{-2}} < \hat{\tau}_{j_1+1} \le \hat{\tau}_{j_2} < \hat{\tau}_{j_{-2}-1} \le \cdots \le \hat{\tau}_{j_{-L}} < \hat{\tau}_{j_N+1} < \hat{\tau}_{j_{-L}-1} \tag{20}$$

for $N = L - 1$. For example, (17) ensures that a realisation of the process $\hat{W}_t$ hits the positive level j_1 before it hits the level -1; then it hits the level $j_{-1} \le -1$ before it hits the level $j_1 + 1$, and so on. Hence the values of such a realisation belong to the interval $[0, j_1] = [\hat{W}_0, \hat{W}_{\hat{\tau}_{j_1}}]$ for $0 \le t \le \hat{\tau}_{j_1}$; then to the interval $[j_1, j_{-1}] = [\hat{W}_{\hat{\tau}_{j_1}}, \hat{W}_{\hat{\tau}_{j_{-1}}}]$ for $\hat{\tau}_{j_1} \le t \le \hat{\tau}_{j_{-1}}$, etc.

As the random walk $\hat{W}_t$ is a time-homogeneous Markov chain, the probability $P_{j_{-L},\dots,j_N,1}$ of the event (17) equals the following product of the probabilities

$$P_{j_{-L},\dots,j_N,1} = P[\hat{\tau}_{j_1}^0 < \hat{\tau}_{-1}^0] P[\hat{\tau}_{j_{-1}}^{j_1} < \hat{\tau}_{j_1+1}^{j_1}] P[\hat{\tau}_{j_2}^{j_{-1}} < \hat{\tau}_{j_{-1}-1}^{j_{-1}}] \cdots P[\hat{\tau}_{j_{-L}-1}^{j_N} < \hat{\tau}_{j_N+1}^{j_N}],$$

[1] In many applications of the Preisach model, the entries $\mathfrak{m}_k^T$ with $|\mathfrak{m}_k^T| > K$ do not affect the dynamics as long as they are constant, *i.e.*, until the moment τ in this case. Moreover, large in absolute value entries $\mathfrak{m}_k^t$ typically have either no or little effect on the future.

where $\hat{\tau}_j^i = \min\{t \geq 0 : \hat{W}_t^i = j\}$ is the hitting time at the level j of a random walk $\hat{W}_t^i$ starting at the level $\hat{W}_0^i = i$. Moreover, for the random walk,

$$P[\hat{\tau}_j^i < \hat{\tau}_k^i] = \frac{k-i}{k-j}, \quad P[\hat{\tau}_j^i > \hat{\tau}_k^i] = \frac{i-j}{k-j} \quad \text{for} \quad j < i < k$$

(see, for example, [46]), hence the probability of the event (15) for $\sigma = 1$ is defined by

$$\frac{1}{P_{j_{-L},\ldots,j_N,1}} = (1+j_1)(1+j_1-j_{-1})(1+j_2-j_{-1})(1+j_2-j_{-2})\cdots(1+j_N-j_{-L})(2+j_N-j_{-L}). \quad (21)$$

Similarly, for $\sigma = -1$,

$$\frac{1}{P_{j_{-L},\ldots,j_N,-1}} = (1-j_{-1})(1+j_1-j_{-1})(1+j_1-j_{-2})\cdots(1+j_N-j_{-L})(2+j_N-j_{-L}). \quad (22)$$

For example, the probability that $\mathfrak{m}_1^T = j, \sigma = 1$ equals $P_{j,1} = (j+1)^{-1}(j+2)^{-1}$ for $j = 1, \ldots, K-1$, and the event $\mathfrak{m}_1^T = -j, \sigma = -1$ has the same probability $P_{-j,-1} = (j+1)^{-1}(j+2)^{-1}$. These probabilities sum up to

$$2\sum_{j=1}^{K-1} \frac{1}{(j+1)(j+2)} = 1 - \frac{2}{K+1}.$$

The complimentary event is that either $\hat{\tau}_K < \hat{\tau}_{-1}$ or $\hat{\tau}_{-K} < \hat{\tau}_1$ with $P[\hat{\tau}_K < \hat{\tau}_{-1}] = P[\hat{\tau}_{-K} < \hat{\tau}_1] = 1/(K+1)$.

A simple numerical check of equation (21) is shown in figure 1, where for $\sigma = 1$ the probabilities $P_{j,k,1}$ as a function of j are plotted for a range of values of k. We observe the expected excellent agreement between the analytical formula (21) and the numerical results.

3.2. Examples

Relations (21), (22) can be used to derive distributions and mean values of parameters of the memory string $\mathfrak{M}_T$. As an illustration, we consider a few examples.

Example 1. Suppose an initial memory state $\mathfrak{M}_T = \mathfrak{M}$ has been created by the random walk input $x(t) = W_t$, $t_0 < t < T$ with $t_0 \ll T$, and we control the input after the moment T. Suppose we increase the input. We ask at which input value the system switches to a new memory branch, that is when the memory string element $\mathfrak{m}_1 = \mathfrak{m}_1^T$ will be deleted. This happens when the increment $\Delta = x(t) - x(T) > 0$ of the input reaches the value $\mathfrak{m}_1$. Thus, we are interested in the probabilities

$$P[\hat{\tau}_1 < \hat{\tau}_{-1}, \mathfrak{m}_1 = m], \quad P[\hat{\tau}_1 > \hat{\tau}_{-1}, \mathfrak{m}_1 = m]$$

for positive integers m. These probabilities are different as the distribution of $\mathfrak{m}_1$ depends on whether $\hat{\tau}_1$ is less or greater than $\hat{\tau}_{-1}$.

According to the previous subsection, the first of these probabilities is $P[\hat{\tau}_1 < \hat{\tau}_{-1}, \mathfrak{m}_1 = m] = P_{m,1} = (m+1)^{-1}(m+2)^{-1}$. Therefore, the corresponding joint probability that $\hat{\tau}_1 < \hat{\tau}_{-1}$ and that the switch happens for $\Delta \geq m$ is

$$P[\hat{\tau}_1 < \hat{\tau}_{-1}, \mathfrak{m}_1 \geq m] = \sum_{k=m}^{\infty} \frac{1}{(k+1)(k+2)} = \frac{1}{m+1}.$$

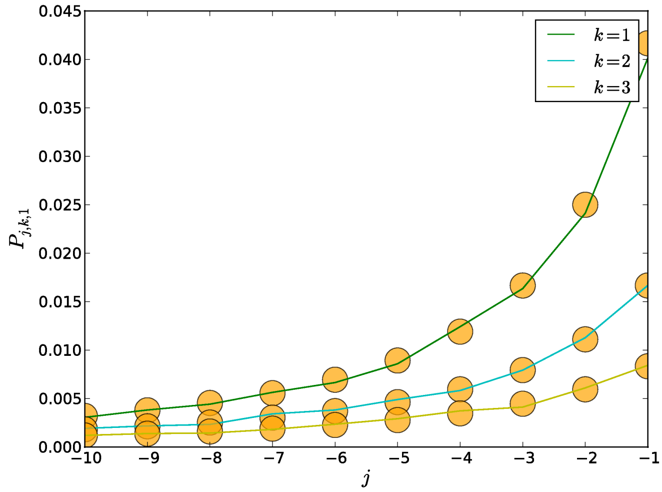

Figure 1. Comparison of analytical results from equation (21) (orange circles) with the averaged outcomes of 10^5 independent simulations of the random walk (solid lines). We plot the joint probability $P_{j,k,1}$ versus the last main minimum j for various values of the last main maximum k.

For $m = 1$, this probability equals $P[\hat{\tau}_1 < \hat{\tau}_{-1}] = 1/2$.

In the case $\hat{\tau}_1 > \hat{\tau}_{-1}$, $P[\hat{\tau}_1 > \hat{\tau}_{-1}, \mathfrak{m}_{-1} = -k, \mathfrak{m}_1 = m] = P_{-k,m,-1} = (k+1)^{-1}(k+m+1)^{-1}(k+m+2)^{-1}$. Summing over k, we obtain

$$P[\hat{\tau}_1 > \hat{\tau}_{-1}, \mathfrak{m}_1 = m] = \sum_{k=1}^{\infty} \frac{1}{(k+1)(k+m+1)(k+m+2)},$$

and for the corresponding cumulative distribution

$$P[\hat{\tau}_1 > \hat{\tau}_{-1}, \mathfrak{m}_1 \geq m] = \sum_{j=m}^{\infty}\sum_{k=1}^{\infty} \frac{1}{(k+1)(k+j+1)(k+j+2)} = \frac{1}{m}\sum_{i=2}^{m+1}\frac{1}{i}.$$

For $m = 1$, this cumulative probability equals $P[\hat{\tau}_1 > \hat{\tau}_{-1}] = 1/2$.

Example 2. The memory string $\mathfrak{M}_t$ changes when the input increment $\Delta = x(t) - x(T)$ reaches the value $\mathfrak{m}_1 = \mathfrak{m}_1^T$ in Example 1 and the system switches to a new memory branch. For instance, if $\hat{\tau}_1 < \hat{\tau}_{-1}$, then the entry $\mathfrak{m}_{-1}^t$ of $\mathfrak{M}_t$ changes from the value $x(T) - x(t)$ to the value $\mathfrak{m}_{-1}^T + x(T) - x(t)$ when Δ reaches the value $\mathfrak{m}_1$. Suppose the jump amplitude $\mathfrak{m}_{-1} = \mathfrak{m}_{-1}^T$ of this entry is used as a measure of the effect of the jump to a new memory branch on the system. Hence, the probability

$$P[\hat{\tau}_1 < \hat{\tau}_{-1}, \mathfrak{m}_1 = m, \mathfrak{m}_{-1} = -k]$$

is of interest, as in our interpretation this is the probability that a jump of strength k happens when Δ reaches the value m. According to the previous subsection, this probability equals $P_{-k,m,1} = (m+1)^{-1}(m+k+1)^{-1}(m+k+2)^{-1}$. Therefore, the probability that $\hat{\tau}_1 < \hat{\tau}_{-1}$ and the strength of the jump is greater or equal than k when Δ reaches the value m is

$$P[\hat{\tau}_1 < \hat{\tau}_{-1}, \mathfrak{m}_1 = m, \mathfrak{m}_{-1} \leq -k] = \sum_{j=k}^{\infty} \frac{1}{(m+1)(m+j+1)(m+j+2)} = \frac{1}{(m+1)(m+k+1)}.$$

For $k = 1$, this is the probability $P_{m,1}$ considered in Example 1.

Further random quantities associated with the random memory string $\mathfrak{M}_T$ such as the number of entries $\mathfrak{m}_k$ in a given interval are considered in the next section.

4. Continuous time stochastic inputs

In this section we consider the situation where the input of the Preisach operator is given by a one-dimensional continuous Markov process $x(t)$, $t_0 \le t \le T$, defined by the Ito stochastic differential equation (SDE)

$$dx = A\,(x)\,dt + \sqrt{B(x)}\,(x)\,dW(t). \tag{23}$$

Here $A\,(x)$ and $B\,(x) \ge 0$ are continuous (not explicitly time dependent) bounded functions of x and $W(t)$ is the standard Wiener process [47]. In addition we assume that there exists a stationary probability distribution whose density $p_s(x)$ fulfills the stationary Fokker-Planck equation,

$$A\,(x)\,p_s\,(x) - \frac{1}{2}\partial_x\Big(B\,(x)\,p_s\,(x)\Big) = 0; \tag{24}$$

here and henceforth the notation ∂_x is used for the ordinary and partial derivatives. We stipulate that the distribution of the initial point $x(t_0)$ follows $p_s(x)$. As a direct consequence, $x(t)$ follows the stationary distribution at each given instant, *i.e.*,

$$p(x,t) = p_s(x) \quad \text{for } t \in [t_0, T] \tag{25}$$

Following Section 2 we assume that the stochastic process $x(t)$ is used as input of the Preisach nonlinearity. To avoid notational clumsiness, let us from now on assume (unless otherwise stated) that condition (2) holds, *i.e.*, the global maximum is assumed to have occurred before the global minimum. Often we will be interested in the memory array at the final time T only, and we introduce the quantities $M_k, m_k, \tau_k^M, \tau_k^m$ to simplify the notation using

$$\mathbf{M}_k = M_k^T, \quad \tau_k^M = \tau^T(M_k^T), \quad \mathbf{m}_k = m_k^T, \quad \tau_k^m = \tau^T(m_k^T), \quad k = 1, 2, \dots \tag{26}$$

In figure 2, a realisation of the Ornstein-Uhlenbeck process is plotted, allowing to illustrate the definitions given above. Since the Markov process defined by (23) is everywhere continuous, but in general not differentiable at points with $B(x) > 0$, it follows that infinitely many different main extrema are possible (generally, the memory array is almost surely infinite). It is our aim to characterise this memory structure.

Here and in the following we will use bold letters for random variables, when the potential for confusion with their corresponding values exists.

In particular we are interested in quantifying the probability distribution of a main extremum (memory element) which was attained *before* a pair of main extrema with known values. For example, we might ask about the probability distribution of a main minimum $\mathbf{m}_{k_0-1}$ under the condition that the immediately following main maximum and minimum have the values $\mathbf{M}_{k_0}$ and $\mathbf{m}_{k_0}$, respectively. This type of question is motivated by the situation, often encountered in practice, that a process is observed for a limited period of time and thus extrema are known up to a certain instant in the past. The problem is then to deduce from those known extrema the probability distribution of previous extreme events for which no recorded data exists. Another related question is about the expected sequence of main extrema in the past. For example if the value of the stochastic process $x(T)$ at time T is known, but not its history, one might be interested to know whether the last main maximum bigger than a certain threshold value $M^0 > x(T)$ has occurred before or after the last main minimum which was below a different threshold value $m^0 < x(T)$.

In order to derive these probability distributions we will first introduce a Fokker-Planck like equation for transition probabilities in reverse time, and then reduce the problem of finding the distribution of previous extrema to a well defined escape problem for the reversed process.

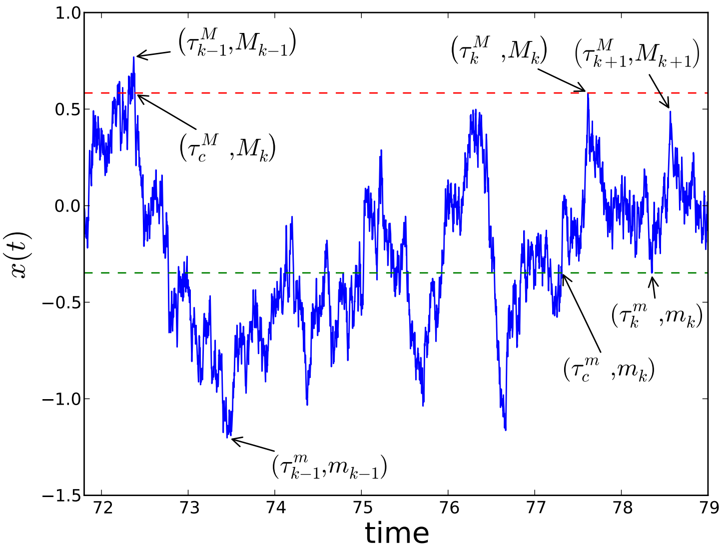

Figure 2. Illustration of the positions of consecutive maxima and minima for a realisation of the Ornstein-Uhlenbeck process with $K = 1$ and $D = 1$. Here M_{k-1} occurring at time τ_{k-1}^M is the global maximum in the shown time interval, and the following main extrema m_{k-1}, M_k, m_k and m_{k+1} are then defined recursively according to (4) and (26). τ_c^M and τ_c^m are defined via $\tau_c^M = \max\{t \in [\tau_{k-1}^M, \tau_{k-1}^m] : x(t) = M_k\}$, $\tau_c^m = \max\{t \in [\tau_{k-1}^m, \tau_k^M] : x(t) = m_k\}$

4.1. Formulation as escape problem

Let us define a reverse transition probability density for $t' < t$ via

$$p_r\left(y, t'|x, t\right) = \frac{p\left(x, t; y, t'\right)}{p_s\left(x\right)} = \frac{p\left(x, t|y, t'\right)p\left(y, t'\right)}{p_s\left(x\right)} = \frac{p\left(x, t|y, t'\right)p_s\left(y\right)}{p_s\left(x\right)}, \tag{27}$$

where equation (25) has been used.

It can be shown that this function fulfills the Fokker-Planck equation in backward time

$$-\partial_y\Big(A(y)p_r(y, t'|x, t)\Big) + \frac{1}{2}\partial_y^2\Big(B\left(y\right)p_r\left(y, t'|x, t\right)\Big) = -\frac{\partial}{\partial t'}p_r\left(y, t'|x, t\right). \tag{28}$$

Note that equation (28) is different from the *backward FPE* for example in the sense of [47]. Instead, it can be regarded as a forward FPE for the reverse transition probability p_r.

Using equation (28) for p_r we can now formulate the problem of finding the distribution of a previous main extremum by using the well developed escape formalism.

Let us consider the case as illustrated in figure 2, and let us assume that some recording of this particular process has started shortly before the instant τ_c^m and thus has established the M_k and m_k as the first known main extrema (realisations of $\mathbf{M}_k$ and $\mathbf{m}_k$). The question is then, what can be assumed about the probability distribution of $\mathbf{m}_{k-1}$. Let us write the probability that $\mathbf{m}_{k-1}$ is below a value $\bar{m}$ under the condition that the next known main extrema are given by M_k and m_k formally as $P(\mathbf{m}_{k-1} < \bar{m}|M_k, m_k)$. This probability can be related to an escape problem in reverse time. Consider the history of the process shown in figure 2 up to τ_c^m. It is clear that we have $\tau_c^M < \tau_{k-1}^m < \tau_c^m$, where $\tau_c^M = \tau^{\tau_c^m}(M_k)$ is the largest time less than τ_c^m at which M_k is obtained. Therefore $\mathbf{m}_{k-1} < \bar{m}$ implies that $\tau^{\tau_c^m}(\bar{m}) > \tau_c^M$. In other words the last time at which the process has entered the interval $[\bar{m}, M_k]$ before reaching m_k at time τ_c^m it has entered this interval from the lower end. On the other hand if $m_{k-1} > \bar{m}$, then the process did last enter the interval $[\bar{m}, M_k]$ at time τ_c^M from the upper end.

In the reverse time formulation as in (28), the problem of last entering an interval is equivalent to the more familiar *first exit* problem from a given interval in forward time [47]. Therefore the probability that m_{k-1} is less than a certain value $\bar{m}$ can be expressed as the probability

$$P(\mathbf{m}_{k-1} < \bar{m}|\mathbf{M}_k = M_k, \mathbf{m}_k = m_k) = \pi(m_k; \bar{m}, M_k) \tag{29}$$

where $\pi(m_k; \bar{m}, M_k)$ is the probability that the reverse time process given by (28) exits the interval $[\bar{m}, M_k]$ through the lower end at $\bar{m}$ when starting at position m_k.

This type of escape problem is well studied in the literature and we here follow the notation of [47] where it is shown that $\pi(m_k; \bar{m}, M_k)$ fulfills the differential equation

$$A(x)\,\partial_x \pi_{\bar{m}}(x) + \frac{1}{2} B(x)\,\partial_x^2 \pi_{\bar{m}}(x) = 0 \tag{30}$$

with the boundary conditions $\pi_{\bar{m}}(\bar{m}) = 1$, $\pi_{\bar{m}}(M_k) = 0$. The explicit solution is given by

$$\pi(m_k; \bar{m}, M_k) = P(\mathbf{m}_{k-1} < \bar{m}|\mathbf{M}_k = M_k, \mathbf{m}_k = m_k) = \frac{I(M_k) - I(m_k)}{I(M_k) - I(\bar{m})} \tag{31}$$

with

$$I(y) \quad = \quad \int_0^y \frac{dx}{B(x)p_s(x)} = \mathcal{N} \int_0^y \exp\left(-\int_0^x \frac{2A(x')}{B(x')}\,dx'\right) dx. \tag{32}$$

Here the normalization constant $\mathcal{N}$ is determined by the normalization condition $\int p_s(x)dx = 1$. Note that $I(y)$ is continuous and monotonic with $I(0) = 0$, but is in general not defined for all values $y \in \mathbb{R}$, since $I(y)$ may diverge to plus or minus infinity at some finite values $y = y_I^M > 0$ or $y = y_I^m < 0$, respectively. If this happens, the domain $\mathcal{O}$ of $I(y)$ is an open interval limited by y_I^M from above or y_I^m from below. Points outside $\mathcal{O}$ cannot be reached by the Markov process $x(t)$ starting at zero, which means that if, for example, we consider two points $a < 0$ and $b > 0$ such that $a \leq y_I^m$ while $I(b)$ is finite, than a process starting at zero will have zero probability of leaving the interval (a, b) at the lower end a. It follows that the function $I(y)$ is invertible and its inverse $I^{-1}(z)$ is well defined for any $z \in \mathbb{R}$.

The probability in equation (31) can also be interpreted as the probability that a (forward time) Wiener process which starts at $I(m_k)$ will first leave the interval $(I(\bar{m}), I(M_k))$ at the lower end boundary at $\bar{m}$, *i.e.*

$$\pi(m_k; \bar{m}, M_k) = \pi^W(I(m_k); I(\bar{m}), I(M_k)) \tag{33}$$

with $\pi^W(x; a, b) = (b - x)/(b - a)$. Therefore the function $I(x)$ relates the characterisation of the main extrema of an arbitrary process with an escape problem of the Wiener process. This connection will be made more explicit in Subsection 4.4 where the memory structure of the Wiener process is studied in greater detail.

To simplify the notation let us also introduce for $x \in \mathcal{O}$

$$\psi(x) = \partial_x I(x) = \frac{1}{B(x)p_s(x)} = \mathcal{N} \exp\left(-\int_0^x \frac{2A(x')}{B(x')}\,dx'\right). \tag{34}$$

Formula (31) for the cumulative conditional probability $P(\mathbf{m}_{k-1} < \bar{m}|\mathbf{M}_k = M_k, \mathbf{m}_k = m_k)$ for m_{k-1} being less than $\bar{m}$ implies that the conditional probability density $p(m_{k-1}|M_k, m_k)$ for the main minimum $\mathbf{m}_{k-1}$ under the condition that the following main extrema are given by M_k and m_k equals

$$p(m_{k-1}|M_k, m_k) = \left.\frac{\partial}{\partial \bar{m}}\right|_{\bar{m}=m_{k-1}} P(\mathbf{m}_{k-1} < \bar{m}|\mathbf{M}_k = M_k, \mathbf{m}_k = m_k) = \frac{I(M_k) - I(m_k)}{[I(M_k) - I(m_{k-1})]^2}\psi(m_{k-1}). \tag{35}$$

We note that there is no explicit k dependence in equation (35), other than via the realisations of the main extrema M_k, m_k and m_{k-1}, which means that formula (35) can be used to calculate the probability density for a main minimum under the condition that the following two main extrema are known independent of k.

Similarly, the conditional density distributions of the previous maximum $\mathbf{M}_{k-1}$ is given (assuming condition (2)) explicitly by

$$p(M_{k-1}|m_{k-1}, M_k) = \frac{I(M_k) - I(m_{k-1})}{\left[I(M_{k-1}) - I(m_{k-1})\right]^2}\psi(M_{k-1}). \tag{36}$$

The generalisation to the case $\tau_k^M > \tau_k^m$ is straightforward.

A combination of equations (35) and (36) can be given in the form of

$$\begin{aligned} p(M_{k-1}, m_{k-1}|M_k, m_k) &= p(M_{k-1}|m_{k-1}, M_k)p(m_{k-1}|M_k, m_k) \\ &= \frac{I(M_k) - I(m_k)}{I(M_k) - I(m_{k-1})}\frac{\psi(m_{k-1})\psi(M_{k-1})}{\left[I(M_{k-1}) - I(m_{k-1})\right]^2} \end{aligned} \tag{37}$$

In this form we can regard $p(M_{k-1}, m_{k-1}|M_k, m_k)$ as the transition probability density from one pair of main extrema (m_k, M_k) to a preceding pair (m_{k-1}, M_{k-1}). This transition probability can be thought to originate from a time discrete two-dimensional master equation on the (m, M) space which proceeds in negative k direction, *i.e.*, $-k$ is the fictitious time. Equation (37) therefore represents the Markov transition chain for this process. This idea is illustrated in figure 3.

The generalization of (37) is straightforward and allows us to write closed form expression for the joint probability of all previous main extrema

$$\begin{aligned} p(M_{k-l}, m_{k-l}, \quad \ldots \quad &, M_{k-1}, m_{k-1}|M_k, m_k) \\ &= p(M_{k-l}, m_{k-l}|M_{k-l+1}, m_{k-l+1})\cdots p(M_{k-1}, m_{k-1}|M_k, m_k) \\ &= \frac{I(M_k) - I(m_k)}{I(M_{k-l}) - I(m_{k-l})}\prod_{i=1}^{l}\frac{\psi(m_{k-i})\psi(M_{k-i})}{\left(I(M_{k-i+1}) - I(m_{k-i})\right)\left(I(M_{k-i}) - I(m_{k-i})\right)}. \end{aligned} \tag{38}$$

which gives formally a complete characterisation of the memory structure before a pair of known main extrema M_k and m_k.

Another related characterisation of past main extrema concerns the sequence in which they have historically occurred. Consider the situation where the input $x(T) = x_0$ of the Preisach nonlinearity is only known at the final time T without any information on the past history. We can then ask the question, whether the last main maximum which is bigger than a certain value $M^0 > x(T)$ has occurred before of after the last main minimum with value less than $m^0 < x(T)$. In other words we want to determine the probability $P\left(\tau^T(M^0) < \tau^T(m^0)|x(T) = x_0\right)$. An escape problem argument as before then leads to the answer that this probability equals the probability of leaving (in reverse time formulation) the interval $[m^0, M^0]$ at the lower end when starting at x_0. Explicitly we find (see equation (31))

$$P\left(\tau^T(M^0) < \tau^T(m^0)|x(T) = x_0\right) = \frac{I(M^0) - I(x_0)}{I(M^0) - I(m^0)}. \tag{39}$$

This equation allows us to predict in which sequence the main extrema are likely to have occurred in the past. The similarity of the expression in (39) with equation (31) allows us to formally write

$$P\left(\tau^T(M^0) < \tau^T(m^0)|x(T) = x_0\right) = P(\mathbf{m}_{k-1} < m^0|\mathbf{M}_k = M^0, \mathbf{m}_k = x_0). \tag{40}$$

This concludes our theoretical investigation on the characterisation of the previous main extrema for a continuous Markov process of the form (23). Next, we discuss a few examples.

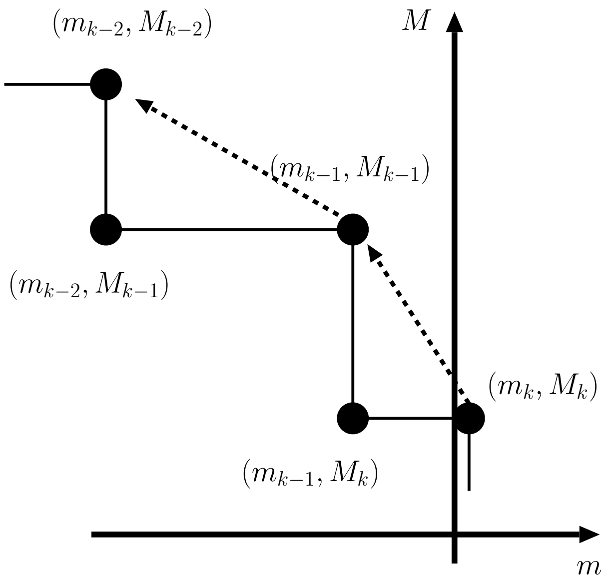

Figure 3. Illustration of the Markov chain between pairs of extrema $(m_k, M_k) \to (m_{k-1}, M_{k-1}) \to (m_{k-2}, M_{k-2}) \to \cdots$ The transition probability density between different pairs of extrema in the (m, M) plane is given by (37). This plane can be interpreted as the Preisach plane, see [22], in which case the staircase line connecting the elements of the chain is the border line between the domain representing the relays which are 'on' and the domain representing the relays which are 'off'.

4.2. Example I: The Ornstein-Uhlenbeck process

Let us consider the case where the input of the Preisach nonlinearity is given by the Ornstein-Uhlenbeck process defined by (23) with $A(x) = -\kappa x$, $B(x) = D$. Then (34) implies

$$\psi(x) = \frac{1}{Dp_s(x)} = \sqrt{\frac{\pi}{\kappa D}} \exp\left(\frac{\kappa x^2}{D}\right) \tag{41}$$

which, according to (31), yields for the probability that m_{k-1} is below $\bar{m}$ for known subsequent main extrema M_k and m_k:

$$P(\mathbf{m}_{k-1} < \bar{m} | \mathbf{M}_k = M_k, \mathbf{m}_k = m_k) = \frac{\int_{m_k}^{M_k} \exp\left(\frac{\kappa}{D} y^2\right)\, dy}{\int_{\bar{m}}^{M_k} \exp\left(\frac{\kappa}{D} y^2\right)\, dy}. \tag{42}$$

The probability density $p(m_{k-1}|M_k, m_k)$ is then again obtained as in equation (35). The probability concerning the sequence of main extrema can then be obtained via equation (38).

As an interesting limit case we now expand the integrand in (42) around m_k via

$$\exp\left(\frac{\kappa}{D} y^2\right) \approx \exp\left(\frac{\kappa}{D} m_k^2\right) \left[1 + (y - m_k)\frac{2\kappa}{D} m_k\right] + \cdots \tag{43}$$

which leads to the simple approximation

$$P(\mathbf{m}_{k-1} < \bar{m} | \mathbf{M}_k = M_k, \mathbf{m}_k = m_k) \quad \approx \quad \frac{M_k - m_k}{M_k - \bar{m}} \left[1 + \frac{\kappa m_k}{D}(m_k - \bar{m})\right]. \tag{44}$$

This approximation is valid if $\sqrt{\frac{\kappa}{D}}(M_k - \bar{m}) \ll 1$ and $\sqrt{\frac{\kappa}{D}} m_k$ stays bounded as $\sqrt{\frac{\kappa}{D}}(M_k - \bar{m})$ decreases.

4.3. Example II: Wiener process with constant drift

If the input for the Preisach nonlinearity is given by a Wiener process with drift, *i.e.*, $A(x) = a$, $B(x) = 1$, then the problem arises that the stationary probability distribution $p_s(x)$ does not exist. The existence of $p_s(x)$ was however a perquisite in the derivation of the Fokker-Planck equation in backward time (28). Nevertheless it turns out that the Wiener process with drift represents an important small scale limit for arbitrary continuous Markov processes. The reason is that for x sufficiently close to m_k we can approximate

$$\psi(x) = \psi(m_k)\exp\left(-\int_{m_k}^{x}\frac{2A(x')}{B(x')}\,dx'\right) \approx \psi(m_k)\exp\left(-\frac{2A(m_k)}{B(m_k)}(x-m_k)\right). \tag{45}$$

This expression is equivalent to the formula for $\psi(x)$ formally obtained from (34) for the Wiener process with drift $a = A(m_k)/B(m_k)$. Note that the Wiener process itself does not possess a stationary distribution, which was required to derive equation (34). Nevertheless we observe that the Wiener process with drift appears naturally in the small scale limit and is therefore a useful abstraction for the dynamics at small scale.

The quantities of interest are not $\psi(x)$ explicitly, but rather the probabilities constructed via $\psi(x)$. For the Wiener process we formally find from (31)

$$\begin{aligned} P(\mathbf{m}_{k-1} < \bar{m}|\mathbf{M}_k = M_k, \mathbf{m}_k = m_k) &= \frac{\int_{m_k}^{M_k}\exp(-2ax)\,dx}{\int_{\bar{m}}^{M_k}\exp(-2ax)\,dx} = \frac{\exp(-2aM_k)-\exp(-2am_k)}{\exp(-2aM_k)-\exp(-2a\bar{m})} \\ &\approx \frac{M_k-m_k}{M_k-\bar{m}}\left[1-a(m_k-\bar{m})\right]. \end{aligned} \tag{46}$$

This last expression is valid up to the first order in $a(M_k - \bar{m})$. We observe that this approximation agrees as expected with the corresponding approximation (44) for the Ornstein-Uhlenbeck process for $a = -\kappa m_k/D$.

In a region around a point x_0 where the ratio $2A(x)/B(x)$ of a given process is about constant, we can characterise the infinitely many main extrema sufficiently close to x_0 by formulas derived from a Wiener process with drift $a = 2A(x_0)/B(x_0)$. At an even smaller scale the drift term can also be neglected and we obtain a characterisation of the main extrema on the basis of the classical Wiener process without drift, *i.e.*, $a = 0$, which we discuss in the next subsection.

4.4. Example III: Wiener Process without drift

In the limit of $a = 0$ the function I defined in (32) becomes formally simply the identity function and we obtain $\psi(x) = 1$. Therefore all previous formulas involving I simplify considerably. For example, formulas (35), (36) for the Wiener process without drift take the form

$$p(m_{k-1}|M_k, m_k) = \frac{M_k - m_k}{(M_k - m_{k-1})^2}, \qquad p(M_{k-1}|M_k, m_k) = \frac{M_k - m_{k-1}}{(M_{k-1} - m_{k-1})^2} \tag{47}$$

and the transition probability (37) in the previously discussed Markov chain is given by

$$p(M_{k-1}, m_{k-1}|M_k, m_k) = \frac{M_k - m_k}{(M_k - m_{k-1})(M_{k-1} - m_{k-1})^2}. \tag{48}$$

This provides the universal characterisation of the main extrema at small scale. Interestingly from equation (48) we can deduce the closed form expression for the joint probability (38) of all previous main extrema

$$\begin{aligned} &p(M_{k-l}, m_{k-l}, \ldots, M_{k-1}, m_{k-1}|M_k, m_k) \\ &\qquad = \frac{M_k - m_k}{(M_k - m_{k-1})(M_{k-1} - m_{k-1})(M_{k-1} - m_{k-2})\cdots(M_{k-l} - m_{k-l})^2}. \end{aligned} \tag{49}$$

Choosing for example $M_k = 1$ and $m_k = 0$ the structure of this formula is directly comparable with the corresponding formula (22) for the discrete random walk process for $\sigma = -1$.

Again, as in the previous section, we remark that there is no stationary probability distribution for the Wiener process and in contrast to our requirement for the derivation of (28) the initial condition of the process cannot be neglected even in the long term limit. Nevertheless, the formulas formally obtained for the Wiener process accurately characterise the memory state of an arbitrary process on a scale at which I is approximately linear.

We can also view the situation from the following perspective. Since the function I is invertible, the main extrema M_k and m_k for a general Markov process $x = x(t)$ are completely characterised if the distribution of the mapped extrema defined as

$$M_k^I = I(M_k), \quad m_k^I = I(m_k) \tag{50}$$

is characterised. Formulas (50) define the main extrema of the process $I(x)$. From (31) it follows that

$$P(\mathbf{m}_{k-1}^I < \bar{m}^I | \mathbf{M}_k^I = M_k^I, \mathbf{m}_k^I = m_k^I) = \frac{M_k^I - m_k^I}{M_k^I - \bar{m}^I}$$

and therefore the conditional probability density for the process $I(x)$ is given by

$$p^I(m_{k-1}^I | M_k^I, m_k^I) = \frac{M_k^I - m_k^I}{\left(M_k^I - m_{k-1}^I\right)^2}$$

which agrees with the corresponding formula (47) for the Wiener process. On the other hand, using

$$p^I(m_{k-1}^I | M_k^I, m_k^I) \partial_y I(y = m_{k-1}) = p(m_{k-1} | M_k, m_k)$$

we directly recover (35). Thus the problem of characterising the distribution of the main extrema of an arbitrary continuous Markov process can be mapped to the problem of characterising the memory structure of the Wiener process without drift by using the invertible function I.

Due to the discussed importance of the Wiener process without drift for the characterisation of main extrema in arbitrary continuous Markov processes let us now consider the differences between main extrema of the Wiener process in more detail. As an alternative equivalent description of the memory array, let us define the strictly monotonically decreasing sequence $\{\mathbf{d}_k\}_{k \geq 1}$ for the differences between the main extrema via

$$\mathbf{d}_k = \begin{cases} \mathbf{M}_{(k+1)/2} - \mathbf{m}_{(k+1)/2} & \text{for } k \text{ odd} \\ \mathbf{M}_{k/2+1} - \mathbf{m}_{k/2} & \text{for } k \text{ even} \end{cases}$$

(cf. (10)) where (2) is assumed. According to formulas (47), in this notation, the conditional probability density for preceding elements in the sequence $\{\mathbf{d}_k\}$ is given via

$$p\left(d_{k-1} | d_k\right) = \frac{d_k}{\left(d_{k-1}\right)^2} \quad \text{for} \quad \mathbf{d}_{k-1} > \mathbf{d}_k; \qquad p\left(d_{k-1} | d_k\right) = 0 \quad \text{for} \quad \mathbf{d}_{k-1} \leq \mathbf{d}_k. \tag{51}$$

We are now interested in the number $\mathbf{q}_d$ of elements of the sequence $\{\mathbf{d}_k\}$ whose value lie in the interval between some maximal value K and some minimal value ϵ. For every $K > \epsilon > 0$, we define the integer valued random variable $\mathbf{q}_d = \mathbf{q}_d(K, \epsilon)$ by

$$\mathbf{q}_m = \min\left\{k : \mathbf{d}_k < \epsilon\right\}, \quad \mathbf{q}_M = \max\left\{k : \mathbf{d}_k < K\right\}, \quad \mathbf{q}_d = \mathbf{q}_m - \mathbf{q}_M.$$

Using the escape problem argument as above we obtain the probability density distribution of the least element $\mathbf{d}_k = \mathbf{d}_{q_m-1}$ satisfying $\mathbf{d}_{q_m-1} \geq \epsilon$ as

$$p\left(d_{q_m-1}\right) = \frac{\epsilon}{\left(d_{q_m-1}\right)^2} \quad \text{for} \quad \mathbf{d}_{q_m-1} > \epsilon\,; \qquad p\left(d_{q_m-1}\right) = 0 \quad \text{for} \quad \mathbf{d}_{q_m-1} \leq \epsilon. \tag{52}$$

A numerical confirmation of formula (52) is shown in figure 4.

We can now interpret d_k as an inverse Markov process with backward transition probabilities given by (51). Combining formulas (51) for k and $k-1$, we obtain $p\left(d_{k-2}|d_k\right) = 0$ for $d_{k-2} \leq d_k$ and

$$\begin{aligned} p\left(d_{k-2}|d_k\right) &= \int p\left(d_{k-2}|d_{k-1}\right) p\left(d_{k-1}|d_k\right)\, \mathrm{d}d_{k-1} \\ &= \int_{d_k}^{d_{k-2}} \frac{d_{k-1}}{\left(d_{k-2}\right)^2} \frac{d_k}{\left(d_{k-1}\right)^2}\, \mathrm{d}d_{k-1} = \frac{d_k}{\left(d_{k-2}\right)^2} \ln\frac{d_{k-2}}{d_k} \quad \text{for} \quad d_{k-2} > d_k. \end{aligned}$$

Similarly, by induction in i, for each $i < k$

$$p\left(d_i|d_k\right) = \int p\left(d_i|d_{i+1}\right) p\left(d_{i+1}|d_k\right)\, \mathrm{d}d_{i+1} = \frac{d_k}{2\left(d_i\right)^2} \left(\ln\frac{d_i}{d_k}\right)^2 \quad \text{for} \quad d_i > d_k$$

with $p\left(d_i|d_k\right) = 0$ for $d_i \leq d_k$.

Relations (52) imply

$$P\left(\mathbf{q}_d = 0\right) = P\left(\mathbf{d}_{q_m-1} > K\right) = \epsilon/K,$$

that is with the probability ϵ/K the memory array does not contain any elements $\mathbf{d}_k$ with values in the range between ϵ and K. Similarly in the case of $\mathbf{q}_d = 1$ there is precisely one $\mathbf{d}_k$, namely $\mathbf{d}_{q_m-1}$, between ϵ and K. With the use of (51), the probability of this event can be expressed as

$$\begin{aligned} P\left(\mathbf{q}_d = 1\right) &= \int_K^\infty \mathrm{d}d_{q_m-2} \int_\epsilon^K p\left(d_{q_m-2}|d_{q_m-1}\right) p\left(d_{q_m-1}|d_{q_m} = \epsilon\right)\, \mathrm{d}d_{q_m-1} \\ &= \int_K^\infty \mathrm{d}d_{q_m-2} \int_\epsilon^K \frac{d_{q_m-1}}{\left(d_{q_m-2}\right)^2} \frac{\epsilon}{\left(d_{q_m-1}\right)^2}\, \mathrm{d}d_{q_m-1} = \frac{\epsilon}{K} \ln\frac{K}{\epsilon}. \end{aligned}$$

Continuing by induction in k, we obtain the following probability distribution of $\mathbf{q}_d$

$$P\left(\mathbf{q}_d = k\right) = \frac{\epsilon}{k!K} \left(\ln\frac{K}{\epsilon}\right)^k \tag{53}$$

which is plotted for $\epsilon/K = 0.1$ in figure 5. That is, for the Wiener process, the number $\mathbf{q}_d$ of the memory elements $\mathbf{d}_k$ between two thresholds ϵ and K has the Poisson distribution with the mean value (intensity)

$$\langle\, \mathbf{q}_d \,\rangle = \ln(K/\epsilon).$$

Acknowledgments

This publication has emanated from research conducted with the financial support of Federal Programme 'Scientists of Innovative Russia', grant 2009-1.5-507-007, Science Foundation Ireland under Grant Number 09/SIRG/I1615, Science Foundation Ireland research grant 07/MI/008 associated with Edgeworth Centre for Financial Mathematics, and Russian Foundation for basic Research, grant 10-01-93112.

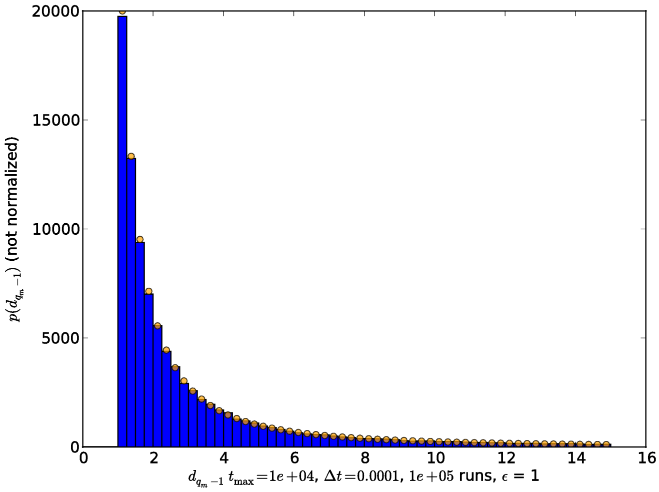

Figure 4. Confirmation of (52) via numerical integration of a Wiener process with time step $\Delta t = 0.0001$ from 0 to $t_{\max} = 10000$. We plot a histogram of the last $\mathbf{d}_m$ which is bigger than $\epsilon = 1$ using a total of 10^5 runs. The circles show the analytical expectation according to (52).

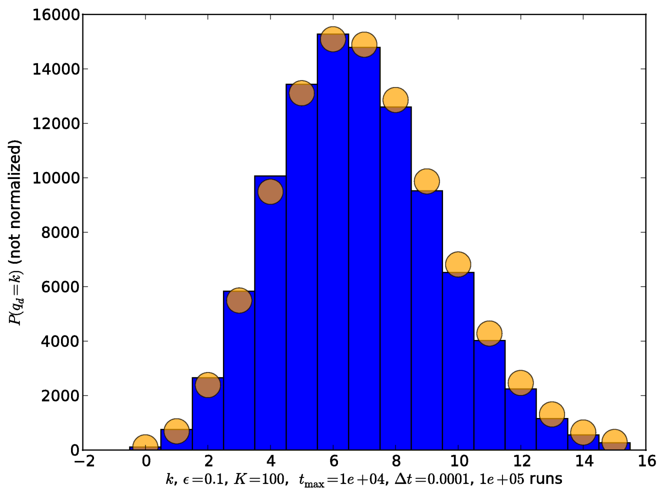

Figure 5. Probability distribution of the number of min-max differences $\mathbf{d}_k$ occurring in the interval $[\epsilon = 0.1, K = 100]$. For the numerical data we integrated a Wiener process with time step $\Delta t = 0.0001$ from 0 to $t_{\max} = 10000$ and recorded the $\mathbf{d}_m$. The circles are the expected values according to (53) with $\epsilon/K = 0.1$.

References

[1] Benzi R, Sutera A and Vulpiani A 1981 *J. Phys. A: Math. Gen.* **14** L453–L457
[2] Benzi R, Parisi G, Sutera A and Vulpiani A 1982 *Tellus* **34** 10-16
[3] Sagues F, Sancho J and García-Ojalvo J 2007 *Rev. Mod. Phys.* **79** 829–882
[4] Lindner B, García-Ojalvo J, Neiman A and Schimansky-Geier L 2004 *Phys. Rep.* **392** 321
[5] Mitaim S and Kosko B 1998 *Proc. IEEE: Special Issue on Intelligent Signal Processing* **86** 2152–2183
[6] Preisach F 1935 *Z. Phys.* **94** 277–302
[7] Mayergoyz I D and Korman C E 1991 *J. Appl. Phys.* **69** 2128–2134
[8] Rugkwamsook P, Korman C, Bertotti G and Pasquale M 1999 *J. Appl. Phys.* **85** 4361–4363
[9] Mayergoyz I D, Adly A, Huang M W and Kraft C 2000 *IEEE Trans. Magn.* **36** 3208
[10] 2006 *The Science of Hysteresis*, three volume set, ed I Mayergoyz and G Bertotti (Elsevier, Academic Press)
[11] 2009 *Hysteresis*, ed R Iyer and X Tan (*IEEE Control Systems Magazine* **1**)
[12] Mayergoyz I D 2006 *The Science of Hysteresis* vol I, ed I Mayergoyz and G Bertotti (Elsevier, Academic Press) pp 529–604
[13] Korman E C 2006 *The Science of Hysteresis* vol I, ed I Mayergoyz and G Bertotti (Elsevier, Academic Press) pp 639–665
[14] Radons G 2008 *Phys. Rev. Lett.* **100** 240602
[15] Radons G 2008 *Phys. Rev. E* **77** 061133
[16] Adedoyin A, Dimian M and Andrei P 2009 *IEEE Trans. Magn.* **45** 3934–3937
[17] Dimian M, Coca E and Popa V 2009 *J. Appl. Phys.* **105** 07D515
[18] Mayergoyz I and Dimian M 2003 *J. Appl. Phys.* **93** 6826–6828
[19] Dimian M and Mayergoyz I D 2004 *Phys. Rev. E* **70** 046124
[20] Visintin A 1994 *Differential Models of Hysteresis* (New York: Springer)
[21] 1997 *Kluwer Encyclopedia of Mathematics* supplement I, ed M Hazewinkel (Kluwer Acad. Publ) pp 310–384
[22] Krasnosel'skii M A and Pokrovskii A V 1989 *Systems with Hysteresis* (Springer)
[23] Mayergoyz I D 1991 *Mathematical Models of Hysteresis* (New York: Springer)
[24] Mayergoyz I D 2003 *Mathematical Models of Hysteresis and Their Applications: Electro-magnetism* (Academic Press)
[25] Brokate M and Sprekels J 1996 *Hysteresis and Phase Transitions* (Springer)
[26] Krejčí P 1996 *Hysteresis, Convexity and Dissipation in Hyperbolic Equations* (Tokyo: Gakkōtosho)
[27] Dahmen K and Ben-Zion Y 2009 *Enciclopedia of Complexity and Systems Science* vol 5, ed C Marchetti and R A Meyers (Springer) pp 5021–5037
[28] Mehta A and Barker G C 2001 *Europhys. Lett.* **56** 626-632
[29] Guyer R A and McCall K R 1996 *Phys. Rev. B.* **54** 18–21
[30] Dahmen K A, Ben-Zion Y and Uhl J T 2009 *Phys. Rev. Lett.* **102** 175501
[31] Gleeson J P 2009 *Phys. Rev. E* **80** 036107
[32] Watts D J 2002 *Proc. of the National Acad. Sci. of USA* **99** 5766–5771
[33] Goicoechea J and Ortin J 1994 *Phys. Rev. Lett.* **72** 2203
[34] R. Cross, M. Grinberg and H. Lamba 2009 *IEEE Control Systems Magazine* **29** 30–43
[35] Sethna J P, Dahmen K and Myers C R 2001 *Nature* **410** 242–250
[36] Dahmen K 2005 *Nature Physics* **1** 13–14
[37] Sethna J P, Dahmen K A and Perkovic O 2005 *The Science of Hysteresis* vol 2, ed G Bertotti and I Mayergoyz (Elsevier) pp 107–168
[38] Sethna J P, Dahmen K, Kartha S, Krumhansl J A, Robetrs B W and Shore J D 1993 *Phys. Rev Lett.* **70** 3347
[39] Rychlik I and Gupta S 2007 *Int. J. of Fatigue* **29** 406–420
[40] Brokate M, McCarthy S, Pimenov A, Pokrovskii A and Rachinskii D 2010 *Environ. Model. Assess.* in press
[41] Lamba H and Seaman T 2008 *Int. J. of Theoretical and Applied Finance* **11** 717–737
[42] Cross R, Grinfeld M, Lamba H and Seaman T 2007 *The European Phys. Journal B* **57** 213–218
[43] Cross R, Grinfeld M, Lamba H and Seaman T 2005 *Physica A* **354** 463–478
[44] Lamba H and Seaman T 2009 arXiv:0709.4242, *Physica A* in press
[45] Lamba H 2009 arXiv:0908.0949v1
[46] Borodin A N and Salminen P 2002 *Handbook of Brownian Motion - Facts and Formulae* (Birkhäuser)
[47] Gardiner C 2009 *Stochastic Methods* (Springer)