

Symbolic time series analysis of ultrasonic signals for fatigue damage monitoring in polycrystalline alloys

Shalabh Gupta, Asok Ray and Eric Keller

Mechanical Engineering Department, The Pennsylvania State University, University Park, PA 16802, USA

E-mail: szg107@psu.edu, axr2@psu.edu and eek105@psu.edu

Received 29 December 2005, in final form 21 April 2006

Published 21 June 2006

Online at stacks.iop.org/MST/17/1963

Abstract

The paper presents the concept and experimental validation of an analytical tool for fatigue damage monitoring in polycrystalline alloys. Ultrasonic signals are utilized for early detection of fatigue damage during the crack initiation period. Small microstructural changes occurring inside the material during the initial stages of fatigue damage cause attenuation and distortion of transmitted waves at the receiver end. The anomaly detection algorithm is based on time series analysis of ultrasonic data and is built upon the principles of *symbolic dynamics*, *information theory* and *statistical signal processing*. Experiments have been conducted for both constant amplitude and block loading of 7075-T6 aluminium alloy compact specimens on a special-purpose test apparatus that is equipped with ultrasonics sensors and a travelling optical microscope for fatigue damage monitoring.

Keywords: symbolic time series analysis, anomaly detection, fatigue damage

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Damage due to fatigue phenomenon is one of the most commonly encountered sources of structural degradation in human-engineered complex electromechanical systems. Detection of fatigue damage at an early stage is essential because the accumulated damage could potentially cause catastrophic failures in the system, leading to loss of expensive equipment and human life [1]. Therefore, it is necessary to develop prognosis capabilities for reliable and safe operation of the system and for enhanced availability of its service life. In the current state-of-the-art, direct measurements of fatigue damage at an early stage (e.g., crack initiation) are not feasible due to lack of analytical models and sensing devices. This paper attempts to address this inadequacy by taking advantage of the sensitivity of the ultrasonic impedance to small microstructural changes occurring inside the material during the crack initiation period.

Several model-based approaches have been developed for structural health monitoring and life prediction of mechanical structures [2–7]. Apparently, no existing model, solely based on the fundamental principles of molecular physics [8], can adequately capture the dynamical behaviour of fatigue damage at the grain level. In general, these models are critically dependent on the initial defects in the materials which are difficult to identify and model. Specifically in the short crack region the appearance of many crack nucleation sites can be treated as random events. Moreover, uncertain usage patterns (e.g., random overloads) and fluctuations under environmental conditions (such as temperature and humidity) may adversely affect the performance of mechanical systems leading to unanticipated failures. The random distribution of flaws in identically manufactured structural components leads to different behavioural patterns of fatigue damage evolution [1, 9]. Consequently, the analysis of time series data from available dedicated sensors is essential for monitoring the evolving fatigue damage in real time [10].

Several techniques based on various sensing devices (e.g., ultrasonics, acoustic emission and eddy currents) have been proposed in the recent literature for fatigue crack monitoring [11–13]. Ultrasonic sensing methodology has been effectively utilized for microstructural analysis in polycrystalline alloys to examine the fatigue phenomenon [14, 15]. Impedance of the ultrasonic signals was shown to be sensitive to small microstructural changes occurring during the early stages of fatigue damage [16, 17]. However, the issues of fatigue monitoring during all stages of fatigue life with appropriate signal processing and data analysis tools were not addressed.

Symbolic time series analysis (STSA) is based on conversion of time series data to discrete symbol sequences [18, 19]. Anomaly detection using STSA [20] is a pattern recognition method that has been recently developed [21], and a comparative evaluation of this novel analytical method shows its superior performance relative to other existing pattern recognition tools in terms of early detection of small changes in dynamical systems [22, 23] and robustness to noisy environments [24].

The paper presents and validates the novel concept of STSA for fatigue damage monitoring in polycrystalline alloys. It emphasizes the need for online updating of information derived from sensing devices that are sensitive to small microstructural changes and are capable of issuing early warnings during fatigue damage evolution [25]. To this end, ultrasonic sensors have been adopted for early detection of fatigue damage during the crack initiation period. The STSA method has been experimentally validated on a fatigue damage testing apparatus [23]. The computer-controlled apparatus is equipped with a variety of sensing instruments such as ultrasonic transducers, optical microscope and extensometers [26]. Information obtained by symbolic time series analysis of ultrasonic data is utilized in this paper for real-time monitoring of fatigue damage under both constant and block loading conditions on compact specimens made of 7075-T6 aluminium alloy.

The paper is organized in five sections, including the present section, and an appendix. Section 2 presents the underlying concepts and essential features of symbolic time series analysis for anomaly detection and is supported by the appendix that introduces the concepts of symbolic dynamics and encoding to provide the necessary background. Section 3 describes the experimental apparatus which is designed and constructed for fatigue monitoring using multiple sensing instruments. Section 4 describes the experimental procedure and presents the results and discussion of fatigue monitoring using symbolic time series analysis of ultrasonic data. The paper is concluded in section 5 along with recommendations for future research.

2. Symbolic time series analysis (STSA)

This section presents the underlying concepts and essential features of STSA [20] for anomaly detection in complex dynamical systems [21]. While the details are reported in previous publications [21, 24], the essential concepts of STSA and construction of a finite state machine from the generated symbol sequence are briefly described in the appendix for completeness of this paper.

In STSA, a data sequence is converted to a symbol sequence by partitioning a compact region of the phase space of the dynamical system, over which the trajectory evolves, into finitely many discrete blocks. Each block is labelled as a symbol, where the symbol set Σ is called the *alphabet set* that consists of $|\Sigma|$ different symbols. (Note $|\Sigma| \geq 2$.) As the system evolves in time, it travels through various blocks in its phase space and the corresponding symbol $\sigma \in \Sigma$ is assigned to it, thus converting a data sequence to a symbol sequence (see the appendix).

2.1. Methodology for anomaly detection

Fatigue damage monitoring is formulated as a two time scale problem. The *fast time scale* is related to the response time of machinery operation. Over the span of a given time series data sequence, the structural dynamic behaviour of the system is assumed to remain invariant, i.e. the process has stationary dynamics at the fast time scale. In other words, the variations in the internal dynamics of the system are assumed to be negligible on the fast time scale. The *slow time scale* is related to the time span over which the process may exhibit non-stationary dynamics. Observable non-stationary behaviour can be associated with the anomalies evolving at a slow time scale. In general, a long time span in the fast time scale is a tiny (i.e. several orders of magnitude smaller) interval in the slow time scale. For example, evolution of fatigue damage in structural materials (causing a detectable change in the dynamics of the system) occurs on the slow time scale; the fatigue damage behaviour is essentially invariant on the fast time scale. Nevertheless, the notion of fast and slow time scales is dependent on the specific application, loading conditions and operating environment. From the perspective of fatigue monitoring, sensor data sets are collected on the fast time scale at different slow time epochs separated by regular intervals. Further details are presented in section 4.

Symbolic time series analysis for behavioural pattern identification requires the following steps.

- Time series data acquisition from appropriate sensors at different slow time epochs.
- Transformation of time series data from the continuous domain to the symbolic domain by partitioning the data sequences into finitely many discrete blocks [18, 19] and calculation of respective state probability vectors at different time epochs (see section 2.3).
- Statistical pattern identification based on the deviation of this vector information from the nominal condition.

2.2. Wavelet space (WS) partitioning

A crucial step in the symbolic time series analysis is partitioning of the phase space for symbol sequence generation [20]. Several partitioning techniques have been reported in the literature for symbol generation [10, 27, 28], primarily based on symbolic false neighbours. These techniques rely on partitioning the phase space and may become cumbersome and extremely computation intensive if the dimension of the phase space is large. Moreover, if the time series data are noise-corrupted, then the symbolic false neighbours would rapidly grow in number and require a large symbol alphabet to capture

the pertinent information on the system dynamics. Therefore, symbolic sequences as representations of the system dynamics should be generated by alternative methods because phase-space partitioning might prove to be a difficult task in the case of high dimensions and presence of noise. The wavelet transform [29] largely alleviates these shortcomings and is particularly effective with noisy data from high-dimensional dynamical systems [24].

This paper has adopted a wavelet-based partitioning approach [21, 24] for construction of symbol sequences from the time series data. In this method, the time series data are first converted to wavelet domain, where wavelet coefficients are generated at different time shifts. The wavelet space is then partitioned with alphabet size $|\Sigma|$ into segments of coefficients on the ordinate separated by horizontal lines. The choice of $|\Sigma|$ depends on specific experiments, noise level and also the available computation power. A large alphabet may be noise-sensitive while a small alphabet could miss the details of signal dynamics. The partitioning is done such that the regions with more information are partitioned finer and those with sparse information are partitioned coarser. This is achieved by maximizing the Shannon entropy [30], which is defined as

$$S = - \sum_{i=1}^{|\Sigma|} p_i \log(p_i), \quad (1)$$

where p_i is the probability of the i th state and summation is taken over all possible states. Each partition region is identified as a state $q_j \in Q$ (for further details, see the appendix). Uniform probability distribution of states is a consequence of maximum entropy that makes the partition coarser in regions of low data density and finer in regions of high data density.

2.3. Calculation of the state probability vectors

Once the partitioning is done with the alphabet size $|\Sigma|$ under the nominal condition (time epoch t_0), it is kept constant for all (slow time) epochs $\{t_1, t_2, \dots, t_k, \dots\}$, i.e. the structure of the partition is fixed under the nominal condition. Therefore, the partitioning structure generated under the nominal condition serves as the reference frame for the data analysis at subsequent slow time epochs.

Definition 2.1. The probability of transitions from the state q_j to state q_k belonging to the set Q of states under a transition $\delta : Q \times \Sigma \rightarrow Q$ is defined as

$$\pi_{jk} = P(\sigma \in \Sigma \mid \delta(q_j, \sigma) \rightarrow q_k), \quad \sum_k \pi_{jk} = 1. \quad (2)$$

Thus, the irreducible stochastic matrix $\mathbf{\Pi} \equiv [\pi_{ij}]$ describes all transition probabilities between the states. The left eigenvector \mathbf{p} corresponding to the unit eigenvalue of $\mathbf{\Pi}$ is the state probability vector under the (fast time scale) stationary condition of the dynamical system [21]. The time series data under the nominal condition are set as the reference point. The *state transition matrix* $\mathbf{\Pi}^0$ is generated to obtain the *state probability vector* \mathbf{p}^0 whose elements are the stationary probabilities of the state vector, where \mathbf{p}^0 is the left eigenvector of $\mathbf{\Pi}^0$ corresponding to the (unique) unity eigenvalue. (See the appendix and the two figures therein for more explanation.) Subsequently, state probability vectors $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^k, \dots$ are obtained at slow time epochs $t_1, t_2, \dots, t_k, \dots$ based on the respective time series data.

2.4. Weighted partition

The gradual evolution of the probability vector on the slow time scale represents small increments in fatigue damage. To this end, the elements of the state probability vector are weighted because the states may not be uniformly sensitive to evolving anomalies. Based on the fact that the amplitude of ultrasonic signals is very sensitive to small microstructural changes [16, 17], it is logical to assign larger weights to the higher-energy segments of the partition, which are more sensitive to perturbations during the early stages of crack initiation. Similarly, smaller weights are assigned to lower-energy segments of the partition, which are less sensitive because they bear useful information only at the crack propagation stage when the ultrasonic signal becomes much less effective due to signal attenuation. The remaining partition segments are assigned weights of intermediate magnitude. As such, the state weights are directly related to the partitioning of the signal space (i.e. the space of wavelet coefficients), where the range of coefficients is partitioned at discrete values of $\{\gamma_j : j = 0, 1, \dots, |\Sigma|\}$. Each segment of the partition (see section 2.2) is assigned a corresponding weight w_j as defined below.

Definition 2.2. The weight matrix W , with a positive-definite diagonal structure, is defined as

$$W = \text{Diag} \left\{ w_j = \frac{\vartheta_j}{\sum \vartheta_j}; j = 1, \dots, |\Sigma| \right\} \quad (3)$$

such that

$$0 \leq w_j \leq 1 \quad \text{and} \quad \sum_j w_j = 1,$$

where the segment energy

$$\vartheta_j = \left(\frac{\gamma_j + \gamma_{j-1}}{2} \right)^2.$$

This is still an active area of research based on thermodynamic formalism of dynamical systems [31]. A statistical mechanical analogy has been recently reported, where it is shown that the partition regions can be represented by thermodynamic energy states and the weights are analogous to the corresponding values of energy [32]. The review of the thermodynamic analogy is beyond the scope of this paper and is not discussed any further.

2.5. Damage evolution and pattern identification

Behavioural pattern changes occur in electro-mechanical systems due to accumulation of faults and progression of anomalies. The pattern changes are quantified as deviations from the nominal behaviour (i.e. the probability distribution under the nominal condition). The resulting anomalies (i.e. deviations of the evolving patterns from the nominal pattern) are characterized by a scalar-valued function, called the *anomaly measure* ψ , that is quasi-static in the fast time scale and is monotonically non-decreasing in the slow time scale.

The state probability vector at any time instant corresponds to a singleton point on the unity-radius hypersphere. During fatigue damage evolution, the tip of the probability vector moves along a path on the surface of this hypersphere. The initial starting point of the path

is the probability vector with uniform distribution obtained with maximum entropy partitioning (see section 2.2). As the damage progresses, the probability distribution changes; eventually when a very large crack is formed, complete attenuation of the ultrasonic signal occurs and consequently the tip of the probability vector reaches a point where all states have zero probabilities of occurrence except the one which has a probability 1 (i.e. a delta-distribution); this state corresponds to the partition region where all data points are clustered due to complete attenuation of the signal.

In the context of fatigue damage, the anomaly measure is formulated on the following assumptions.

- *Assumption 1.* The damage evolution is an irreversible process (i.e. with zero probability of self-healing) and implies the following conditions:

$$\psi(t) \geq 0; \quad \psi(t + \delta) - \psi(t) \geq 0 \forall t \geq t_0 \forall \delta > 0. \quad (4)$$

- *Assumption 2.* The damage accumulation between the two time epochs is a path function, i.e. dependent on the path traversed to reach the target state from the initial state.

In the context of fatigue damage in polycrystalline alloys at room temperature, the crack length is traditionally defined by a straight line joining the starting point to the tip of the crack but, in reality, the actual crack follows a complicated path (possibly fractal in ductile materials). In fact, at the initial stages of fatigue damage, there can be multiple short cracks oriented in different directions. Therefore, the crack length alone does not provide complete information on fatigue damage evolution. Since ultrasonic signals are highly sensitive to small microstructural changes, signal distortion is a good index of anomaly growth. The tip of the probability vector, obtained through a symbolic time series analysis, moves along a curved path on the surface of the unity-radius hypersphere between the initial point \mathbf{p}^0 (i.e. uniform distribution obtained under maximum entropy partitioning) and the final point at very large crack formation \mathbf{p}^f (i.e. δ -distribution due to complete attenuation of the signal). Phenomena such as piling up of dislocations, strain hardening or reflections from multiple crack surfaces affect the ultrasonic signals in a variety of ways. An increase of the ultrasonic amplitude is also observed during very early stages of fatigue damage due to hardening of the material. On the other hand, ultrasonic signals attenuate sharply at the crack propagation stage upon development of a large crack. As such, distortion of ultrasonic signals at a single time epoch may not uniquely determine the state of fatigue damage. The rationale is that two signals may exhibit similar characteristics but, in terms of actual incurred damage, the states are entirely different. Consequently, fatigue damage is a path function instead of being a state function. This assessment is consistent with assumption 1 implying that the damage evolution is irreversible. That is, at two different time epochs, the damage cannot be identical unless the net damage increment is zero. Consequently, by assumption 2, the anomaly measure should follow the traversed path of the probability vector, not the straight line joining the end points (i.e. the tips of the probability vectors).

The anomaly measure, based on the path between the nominal state and the completely damaged state, can be different even for identical test samples and under the

same loading conditions because of the stochastic nature of fatigue phenomena. As such, analysis of a stochastic data set collected under identical experimental conditions is essential for identification of variations in different data sets. This problem is still under active investigation and will be reported in a forthcoming publication.

The following distance function is derived between probability vectors at two time epochs:

$$d(\mathbf{p}^k, \mathbf{p}^l) \equiv \sqrt{(\mathbf{p}^k - \mathbf{p}^l)^T W (\mathbf{p}^k - \mathbf{p}^l)}. \quad (5)$$

The algorithm for computation of the anomaly measure ψ compensates for spurious measurement and computation noise in terms of the sup norm $\|\mathbf{e}\|_\infty \equiv \max(|e_1|, \dots, |e_m|)$ of the error in the probability vector (i.e. the maximum error in the elements of the probability vector). The algorithm is presented below.

- (i) $\psi^0 = 0, \delta\psi^1 = 0, \tilde{\mathbf{p}} = \mathbf{p}^0, k = 1;$
- (ii) if $\|\mathbf{p}^k - \tilde{\mathbf{p}}\|_\infty > \epsilon$ then $\delta\psi^k = d(\mathbf{p}^k, \tilde{\mathbf{p}})$ and $\tilde{\mathbf{p}} \leftarrow \mathbf{p}^k;$
- (iii) $\psi^k = \psi^{k-1} + \delta\psi^k;$
- (iv) $k \leftarrow k + 1, \delta\psi^k = 0;$ go to step (ii).

The real positive parameter ϵ , associated with robustness of the anomaly measure for measurement and computation noise, is identified by performing an experiment with a sample with no notch. Since there is no notch there is practically no stress augmentation and relatively no fatigue damage. As such, the parameter ϵ is estimated as

$$\epsilon \approx \max_{l \in \{1, \dots, N\}} (\|\mathbf{p}^{l+1} - \mathbf{p}^l\|_\infty) \quad (6)$$

from N consecutive observations with $N \gg 1$. The algorithm works in the following fashion: the reference point $\tilde{\mathbf{p}}$ is initialized to the starting point \mathbf{p}^0 and the anomaly measure ψ^0 is set to 0. At any slow time epoch t_k , if the state probability vector moves such that the distance travelled in any particular direction (i.e. the sup norm $\|\bullet\|_\infty$) is greater than ϵ as specified in step (ii), then the anomaly measure is incremented by $\delta\psi^k = d(\mathbf{p}^k, \tilde{\mathbf{p}})$ and the reference point is shifted to the current point \mathbf{p}^k . The procedure is repeated at all slow time epochs. As such, the total path travelled by the tip of the probability vector represents the deviation from the nominal condition and the associated damage.

2.6. Summary of STSA anomaly detection

The STSA procedure of anomaly detection is summarized below.

- Time series data acquisition from appropriate sensor(s) at time epoch t_0 , i.e. the nominal condition, when the system is assumed to be in the healthy state (i.e. zero anomaly measure).
- Generation of the wavelet transform coefficients, obtained with an appropriate choice of the wavelet basis.
- Maximum entropy partitioning of the wavelet space under the nominal condition (see section 2.2) and generation of the corresponding symbol sequence. The partitioning is fixed for subsequent time epochs.
- Generation of the state probability vector \mathbf{p}^0 at time epoch t_0 .

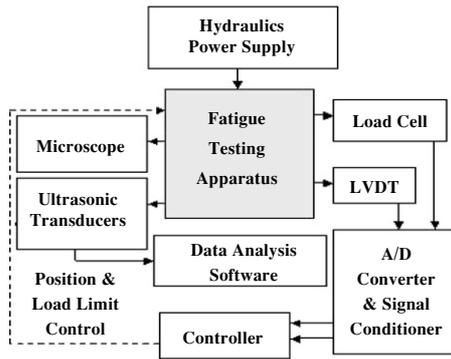


Figure 1. Schematic of the fatigue damage test bed.

- Time series data acquisition at subsequent slow time epochs, $t_1, t_2, \dots, t_k, \dots$ and their conversion to the wavelet domain to generate respective symbolic sequences based on the partitioning at time epoch t_0 .
- Generation of the state probability vectors $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^k, \dots$, at slow time epochs, $t_1, t_2, \dots, t_k, \dots$ from the respective symbolic sequences.
- Computation of scalar *anomaly measures* $\psi^1, \psi^2, \dots, \psi^k, \dots$, at time epochs, $t_1, t_2, \dots, t_k, \dots$.

3. Description of experimental apparatus

This section briefly describes the experimental apparatus which is designed to study the fatigue damage in mechanical structures. Figure 1 shows the schematic of the test bed consisting of the fatigue damage test apparatus embedded with multiple sensing devices along with the software interfaces for control and real-time damage monitoring. The experimental apparatus, shown in figure 2, is a special-purpose uniaxial fatigue testing machine, which is operated under load control or strain control at speeds up to 12.5 Hz; a detailed description of the apparatus and its design specifications are reported in [26]. The test specimens are subjected to tensile–tensile cyclic loading by a hydraulic cylinder under the regulation of computer-controlled electro-hydraulic servo-valves. The feedback signals that are generated from the load cell and the extensometer are processed by signal conditioners that include standard amplifiers and signal processing units. The controller governs the hydraulic servo-valve for operation under specified load and position limits. The damage estimation and life prediction subsystem consists of data analysis algorithm and the associated computer hardware. The process instrumentation and the control module of the fatigue test apparatus are briefly described below.

- *Closed loop servo-hydraulic unit and controller.* The instrumentation and control of the computer-controlled uniaxial fatigue test apparatus includes a load cell, an extensometer, an actuator, the hydraulic system and the controller. The servo-hydraulic unit can excite the system with either random loads or random strains at variable amplitudes. The control module is installed on a computer which is dedicated to machine operation. The controller operates the machine according to a schedule file which contains the specifications of loading profile and the

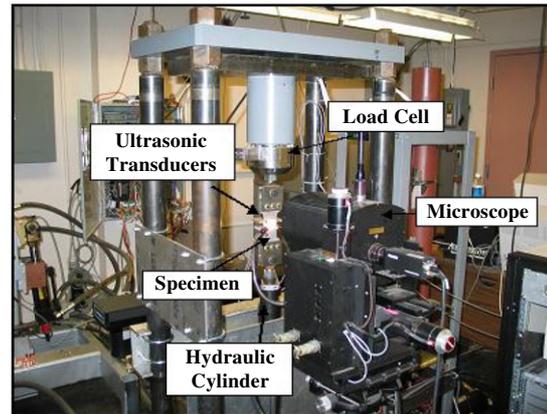


Figure 2. Computer-instrumented apparatus for fatigue testing.

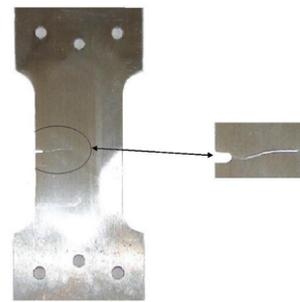


Figure 3. Cracked specimen with a side notch.

number of load cycles. The real-time data from the extensometer and load cell are supplied to the controller for operation under specified position and load limits.

- *Subsystem for data acquisition, signal processing and engineering analysis.* In addition to the computer for controlling the load frame, a second computer is used for real-time image data collection from the microscope to monitor the growth of surface cracks. The instrumentation for ultrasonic flaw detection scheme is connected to a third computer. The real-time ultrasonic data collected on this computer are transferred at regular intervals to a fourth computer on which the data analysis algorithm is installed. The algorithm based on symbolic time series analysis generates the anomaly measures at different slow time epochs, and the plots are displayed on the screen. These laboratory computers are interconnected by a local dedicated network for data acquisition, data communications and control.

Figure 3 shows a compact specimen of the 7075-T6 aluminium alloy used for testing in the fatigue damage test apparatus. The specimens are 3 mm thick and 50 mm wide with a slot on one side of 1.58 mm diameter and 4.57 mm length. The notch is made to increase the stress concentration factor that ensures crack initiation and propagation at the notch end. The test specimens have been subjected to sinusoidal loading under tension–tension mode (i.e. with a constant positive offset) at a frequency of 12.5 Hz. The dc offset is provided in the load cycling to ensure that the specimen is always under tension. Since inclusions and flaws are randomly

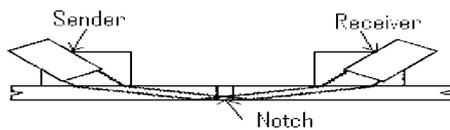


Figure 4. Schematic of ultrasonic sensors on a test specimen.

distributed across the material, small cracks appear at these defects and propagate and join at the machined surface of the notch even before microscopically visual cracks appear on the surface.

The test apparatus is equipped with a variety of dedicated sensors for monitoring the fatigue damage. Two types of sensors that have been primarily used for damage detection are as follows.

(1) *Travelling optical microscope.* The travelling optical microscope, shown as part of the test apparatus in figure 2, provides direct measurements of the visible portion of a crack. The resolution of the optical microscope is about $2\ \mu\text{m}$ at a working distance of 10 to 35 cm. The growth of the crack is monitored continuously by the microscope which takes images of the surface of the specimen at regular intervals. The crack length can be calculated automatically by movement of the microscope from the notch end to the tip of the crack. The data acquisition software also allows for manual operation and image capture at the desired moment.

(2) *Ultrasonic flaw detector.* A piezoelectric transducer is used to inject ultrasonic waves in the specimen, and a receiver transducer is placed on the other side of the notch to measure the transmitted signal, as seen in figure 4. The ultrasonic signals produced are 10 MHz sinusoidal waves and they are triggered during a very short portion at the peak of every load cycle. Ultrasonic measurements are taken at stress levels that exceed the crack opening stress and this causes maximum attenuation of the ultrasonic waves. Note that if crack closure occurs at low loads, then an alternative method would be needed to detect anomalies.

Since material characteristics (e.g., voids, dislocations and short cracks) influence ultrasonic impedance, a small fault in the specimen is likely to change the signature of the signal at the receiver end. Therefore, the signal can be used to capture minute changes during the early stages of fatigue damage [25]. A significant amount of internal damage (e.g., dislocations, short cracks and microstructural defects) occurs before the crack appears on the surface of the specimen when it is observed by the microscope [33]. This internal damage inside the specimen can cause detectable attenuation and/or distortion of the ultrasonic waves [16]. An elaborate description of the properties of ultrasonic waves in solid media is provided by Rose [34]. The crack propagation stage starts when this internal damage eventually develops into a single large crack. Subsequently, the crack growth rate increases rapidly and when the crack becomes sufficiently large, complete attenuation of the ultrasonic signal occurs at the receiver end.

The ultrasonic flaw detection technique is easy to install at the potential damage site and is capable of detection of fatigue damage before the onset of widespread fatigue crack propagation. An optical microscope is only capable

Table 1. Load scheduling for type II loading.

Blocks	Number of cycles	Maximum load
B_0	Until $500\ \mu\text{m}$ crack	81.0 MPa
B_1	10 000 cycles	67.0 MPa
B_2	10 000 cycles	50.0 MPa
B_3	10 000 cycles	39.8 MPa
B_4	Until failure	74.1 MPa

of detecting cracks when they appear on the front surface of the specimen. Therefore, the study in this paper is based on analysing the ultrasonic data for monitoring fatigue damage during both crack initiation and crack propagation stages.

4. Experimental results and discussion

The fatigue tests have been conducted on 7075-T6 aluminium specimens where the aim of the experiments is fatigue damage monitoring using ultrasonic sensing technique before the onset of widespread fatigue leading to fracture. The tests have been performed at 12.5 Hz frequency under two different types of loading conditions:

type I: constant loading

type II: block loading (see table 1).

For type I loading, the specimens (see section 3) are subjected to a sinusoidal load where the maximum and minimum loads are kept constant at 89.3 MPa and 4.85 MPa, respectively. For type II loading, the sample is initially loaded with a sinusoidal loading with maximum amplitude of 81 MPa. As soon as the microscope detects the crack and the crack length reaches $500\ \mu\text{m}$, the load cycling is shifted to block loading which consisted of four blocks from B_1 to B_4 . The minimum load in all regions is kept the same at 4.85 MPa. The loading schedule is summarized in table 1.

Ultrasonic waves at 10 MHz are triggered at the peak of each sinusoidal load cycle where the stress is maximum and the crack is open causing maximum attenuation of the ultrasonic waves. Since the ultrasonic frequency is much higher than the load cycling frequency, data collection is performed for a very short interval in the time scale of load cycling. The slow time epochs have been chosen to be 1000 load cycles (i.e. $\sim 80\ \text{s}$) apart. At the onset of each slow time epoch, the ultrasonic data points are collected on the fast time scale of 50 cycles (i.e. $\sim 4\ \text{s}$), which produced a string of 15 000 data points. It is assumed that during the fast time scale of 50 cycles, the system remains in a stationary condition and no major changes occur in the fatigue crack behaviour. These sets of time series data points collected at different slow time epochs have been analysed using the STSA method to calculate the anomaly measures at those slow time epochs.

The nominal condition at the slow time epoch t_0 is chosen to be ~ 0.5 kilocycles to ensure that the electro-hydraulic system of the test apparatus has come to a steady state, and it is assumed that no significant damage occurs till that point. This nominal condition is chosen as a benchmark where the anomaly measure is chosen to be zero. The anomalies at subsequent slow time epochs, $t_1, t_2, \dots, t_k, \dots$, are then calculated using the procedure summarized in section 2.6.

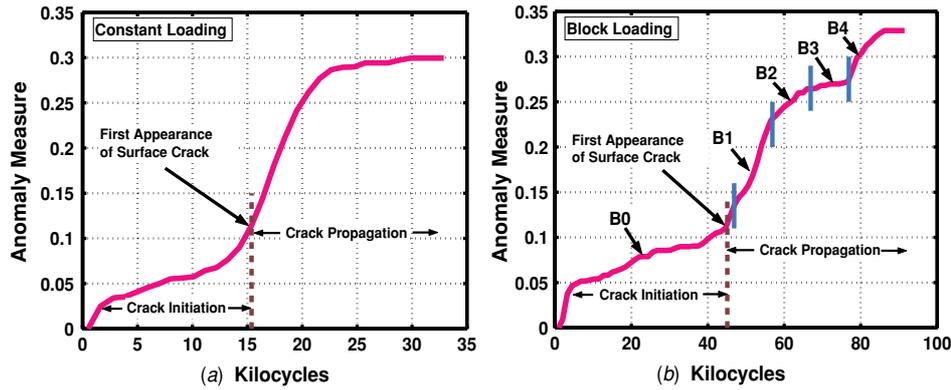


Figure 5. Profiles of anomaly measures exhibiting evolving fatigue damage obtained by symbolic time series analysis of ultrasonic data for two experiments. Left plate: constant loading and right plate: block loading.

Any particular value of the anomaly measure greater than zero indicates that the tip of the probability vector has moved along a path starting from the nominal condition which signifies that changes have occurred inside the specimen. It is emphasized that the anomaly measure is relative to the nominal condition which is fixed in advance and should not be confused with the actual damage at an absolute level. However, inferring fatigue damage from the observed anomaly measure is an inverse problem that is a topic of future research.

The alphabet size for partitioning has been chosen to be $|\Sigma| = 8$ while the wavelet basis is chosen to be ‘gaus2’ [35] (see section 2). Absolute values of the wavelet scale series data have been used to generate the partition because of the symmetry of the data sets about their mean. Increasing the alphabet size $|\Sigma|$ does not improve the results and creates a large number of states, many of them having very small or zero probabilities. This algorithm enables detection of crack initiation with only eight states and is computationally very fast in the sense that the code execution time is several orders of magnitude smaller than the process response time. The wavelet basis, ‘gaus2’, provides better results than the wavelet bases of the Daubechies family [29] because the ‘gaus2’ wavelet basis closely matches the shape of the sinusoidal ultrasonic signals.

Figure 5 exhibits the plots of anomaly measures obtained by symbolic time series analysis of ultrasonic data for two loading conditions: (a) constant loading and (b) block loading. The left plate of figure 5 shows the fatigue damage evolution under constant amplitude loading. The first appearance of a surface crack as observed by the microscope occurred at ~ 15.4 kilocycles which is indicated by the vertical dashed line. The region towards the right of this vertical line is described as the crack propagation phase and the region towards the left is described as the crack initiation phase. During the crack initiation phase, multiple small cracks coalesce together to form a single large crack. It is observed from the plot that the slope of anomaly measure changes sharply during the end of crack initiation phase which indicates the transition from crack initiation to crack propagation. This occurs approximately upon appearance of a surface crack. An abrupt increase in the slope (i.e. a sharp rise in the curvature) of the anomaly measure profile provides a clear insight into a forthcoming failure. In the crack propagation region, the growth of fatigue

damage takes place significantly faster than the crack initiation as shown by the slope of the anomaly measure. After a sufficiently large crack has developed, the ultrasonic signals attenuate completely leading to a complete failure. However, the crack initiation region towards the left of the vertical line is of significant importance because the microscope is unable to show any sign of damage during this phase.

The right plate of figure 5 shows the fatigue damage evolution under block loading. Five different blocks are shown in the figure, each with different amplitude sinusoidal loading as described in table 1. The first appearance of a surface crack as observed by the microscope occurred at ~ 45.1 kilocycles which is indicated by the vertical dashed line. The crack appeared on the surface when it was $\sim 200 \mu\text{m}$ long. The crack length reached $\sim 500 \mu\text{m}$ at ~ 46.9 kilocycles when the loading was shifted from that of block B_0 to block B_1 . Thereafter, each block was shifted after a period of 10 000 cycles. The load is reduced from 81.0 MPa to 67.0 MPa after a 500 μm long crack developed. In the region of block B_1 , the slope of the anomaly measure indicated a small drop for a few cycles, but then it went up high because the specimen is already in the crack propagation phase. As the load is further reduced to 50.0 MPa during block B_2 , the slope of the anomaly measure dropped as compared to that in block B_1 indicating the slowing down of the crack propagation. Further reduction of the load to 39.8 MPa during block B_3 again decreased the anomaly measure slope indicating further slow down of fatigue crack growth. Finally, the load was increased to 74.1 MPa during block B_4 which caused an increase in the anomaly measure and a rapid growth of fatigue crack till failure. These observations indicate that the ultrasonic signals are able to capture different growth rates of fatigue damage under different loading conditions. As similar to a constant amplitude loading case, there is significant rise of the anomaly measure during crack initiation region towards the left of the vertical dashed line. This shows that the subsurface microstructural changes that occur before the onset of widespread fatigue crack propagation are captured by the ultrasonic signals.

The two plates in figure 5 show a relatively large slope of anomaly measure from the start of cyclic loading to ~ 3 – 5 kilocycles. This is the stage where microstructural damage (e.g., due to dislocation movements and accumulation, and

persistent slip band formation) induces hardening of the strained components [33]. The ultrasonic impedance changes sharply due to these deformities causing a sharp rise of the slope of the anomaly measure in the very early cycles. After these initial effects subside, a modest reduction of the slope takes place for the remaining part of the crack initiation phase. Similar phenomena have been reported by Berkovits and Fang [36] in acoustic emission experiments on smooth specimens of Incoloy 901 at room temperature. Further experiments and microstructural analysis are necessary to confirm these findings.

5. Conclusions and future work

This paper validates a novel tool of real-time anomaly detection, which relies on symbolic time series analysis (STSA) of measured variables and is built upon the principles of *symbolic dynamics*, *information theory* and *automata theory*. Efficacy of the anomaly detection tool has been demonstrated by experimental validation on a laboratory apparatus for monitoring fatigue damage in 7075-T6 aluminium alloy compact specimens under constant amplitude and block loading. Time series analysis of ultrasonic sensor signals is capable of detecting early growth of fatigue damage significantly before the onset of large crack propagation, as seen by an optical microscope.

The anomaly measure is computed in real time, which is an indication of microstructural changes occurring inside the material. The exact correlation between actual damage and ultrasonic measurement is not provided in this paper and is an area of future work. Potential scope areas of future research are summarized below.

- Validation of the STSA technique for early detection of fatigue damage under spectral loading.
- Statistical analysis of time series data of fatigue damage, collected under identical loading and environmental conditions, to account for manufacturing and material uncertainties.
- Interpretation of phase changes in the fatigue damage evolution analogous to those in statistical mechanics.
- Development of real-time life extending and damage mitigating control laws based on the damage information generated by STSA.

Acknowledgments

This work has been supported in part by the US Army Research Laboratory and the US Army Research Office under grant no. DAAD19-01-1-0646.

Appendix. Symbolic dynamics, encoding and state machine construction

This appendix briefly describes the concepts of *symbolic dynamics*, encoding nonlinear system dynamics from observed time series data, and state machine construction. It also presents a procedure for online computation of the machine state probability vectors that are representatives of the evolving fatigue damage.

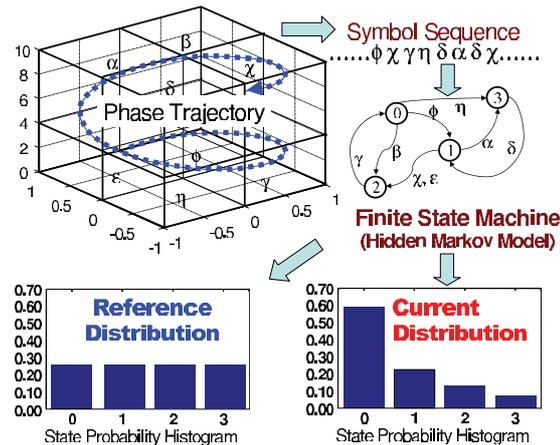


Figure 6. Partitioning, symbol generation and machine construction.

Let a continuously varying physical process be modelled as a finite-dimensional dynamical system in the setting of an initial value problem:

$$\frac{dx(t)}{dt} = f(x(t), \theta(t_s)); \quad x(0) = x_0, \quad (A.1)$$

where $t \in [0, \infty)$ denotes the (fast-scale) time, $x \in \mathbb{R}^n$ is the state vector in the phase space and $\theta \in \mathbb{R}^\ell$ is the (possibly anomalous) parameter vector varying in (slow-scale) time t_s .

A tool for behaviour description of nonlinear dynamical systems is based on the concept of formal languages for transitions from smooth dynamics to a discrete symbolic description [19]. The phase space of the dynamical system in equation (A.1) is partitioned into a finite number of cells, so as to obtain a coordinate grid of the space. A compact (i.e. closed and bounded) region $\Omega \in \mathbb{R}^n$, within which the (stationary) motion under the specific exogenous stimulus is circumscribed, is identified. Encoding of Ω is accomplished by introducing a partition $\Phi \equiv \{\Phi_1, \dots, \Phi_m\}$ consisting of m mutually exclusive (i.e. $\Phi_j \cap \Phi_k = \emptyset \forall j \neq k$) and exhaustive (i.e. $\bigcup_{j=1}^m \Phi_j = \Omega$) cells. The dynamical system describes an orbit by the time series data as $\mathbb{O} \equiv \{x_0, x_1, \dots, x_k, \dots\}$, $x_i \in \Omega$, which passes through or touches the cells of the partition Φ .

Let the cell, visited by the trajectory at a time instant, be denoted as a random variable S that takes a symbol value $s \in \Sigma$. The set Σ of m distinct symbols that label the partition elements is called the *symbol alphabet*. Each initial state $x_0 \in \Omega$ generates a sequence of symbols defined by a mapping from the phase space into the symbol space as

$$x_0 \rightarrow s_i0s_{i1}s_{i2} \dots s_{ik} \dots \quad (A.2)$$

The mapping in equation (A.2) is called *symbolic dynamics* as it attributes a legal (i.e. physically admissible) symbol sequence to the system dynamics starting from an initial state. (Note that a symbol alphabet Σ is called a generating partition of the phase space Ω if every legal symbol sequence uniquely determines a specific initial condition x_0 , i.e. every symbolic orbit uniquely identifies one continuous space orbit.) Figure 6 pictorially elucidates the concepts of partitioning a finite region of the phase space and mapping from the partitioned space into the symbol alphabet. This

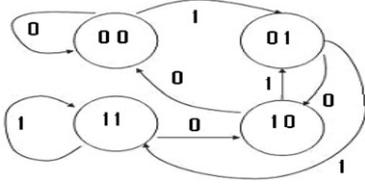


Figure 7. Finite state automaton with $D = 2$ and $\Sigma = \{0, 1\}$.

represents a spatial and temporal discretization of the system dynamics defined by the trajectories. Figure 6 also shows the conversion of the symbol sequence into a finite-state machine as explained in later sections.

Symbolic dynamics can be viewed as coarse graining of the phase space, which is subjected to (possible) loss of information resulting from granular imprecision of partitioning boxes. However, the essential robust features (e.g., periodicity and chaotic behaviour of an orbit) are expected to be preserved in the symbol sequences through an appropriate partitioning of the phase space [19]. Although the theory of phase-space partitioning is well developed for one-dimensional mappings, very few results are known for two- and higher-dimensional systems [31].

Sole usage of the model in equation (A.1) may not always be feasible due to unknown parametric and non-parametric uncertainties and noise. A convenient way of learning the dynamical behaviour is to rely on the additional information provided by (sensor-based) time series data [10, 31].

A.1. State machine construction

The partitioning (see figure 6) is performed at the slow time epoch t_0 of the nominal condition that is chosen to be the healthy state having zero anomaly measure. A finite state machine is then constructed, where the states of the machine are defined corresponding to a given *alphabet* set Σ and window length D . The alphabet size $|\Sigma|$ is the total number of partition segments while the window length D is the length of consecutive symbol words [21], which are chosen as all possible words of length D from the symbol sequence. Each state belongs to an equivalence class of symbol words of length D or more, which is characterized by a word of length D at the leading edge. Therefore, the number n of such equivalence classes (i.e. states) is less than or equal to the total permutations of the alphabet symbols within words of length D . That is, $n \leq |\Sigma|^D$; some of the states may be forbidden with zero probability of occurrence. For example, if $\Sigma = \{0, 1\}$, i.e. $|\Sigma| = 2$ and if $D = 2$, then the number of states is $n \leq |\Sigma|^D = 4$ and the possible states are 00, 01, 10 and 11, as shown in figure 7.

The choice of $|\Sigma|$ and D depends on specific experiments, noise level and also the available computation power. A large *alphabet* may be noise-sensitive while a small alphabet could miss the details of signal dynamics [24]. Similarly, while a larger value of D is more sensitive to signal distortion, it would create a much larger number of states requiring more computation power. In this paper, the window length is set to $D = 1$; consequently, the set of states \mathcal{Q} is equivalent to the symbol alphabet Σ . Therefore, selection

of the parameters $D = 1$ and $|\Sigma| = 8$ leads to a finite state machine with eight states, which is very fast in computation and is also capable of early detection of anomalies. However, other applications, such as two-dimensional image processing, may require larger values of the parameter D . Using the symbol sequence generated from the time series data, the state machine is constructed on the principle of sliding block codes [18]. The window of length D on the symbol sequence $\dots \sigma_{i_1} \sigma_{i_2} \dots \sigma_{i_k} \dots$ is shifted to the right by one symbol, such that it retains the last $(D-1)$ symbols of the previous state and appends it with the new symbol σ_{i_k} at the end. The symbolic permutation in the current window gives rise to a new state. The machine constructed in this fashion is called the D -Markov machine [21] because of its Markov properties.

Definition A.1. A symbolic stationary process is called D -Markov if the probability of the next symbol depends only on the previous D symbols, i.e. $P(\sigma_{i_0} | \sigma_{i_1} \dots \sigma_{i_{-D}} \sigma_{i_{-D-1}} \dots) = P(\sigma_{i_0} | \sigma_{i_1} \dots \sigma_{i_{-D}})$.

The finite state machine constructed above has D -Markov properties because the probability of occurrence of the symbol σ_{i_k} on a particular state depends only on the configuration of that state, i.e. the previous D symbols. Once the alphabet size $|\Sigma|$ and word length D are determined under the nominal condition (i.e. time epoch t_0), they are kept constant for all slow time epochs $\{t_1, t_2, \dots, t_k, \dots\}$ (see section 2.2). That is, the partitioning and the state machine structure generated under the nominal condition serve as the reference frame for data analysis at subsequent slow time epochs.

The states of the machine are marked with the corresponding symbolic word permutation and the edges joining the states indicate the occurrence of a symbol σ_{i_k} . The occurrence of a symbol at a state may keep the machine in the same state or move it to a new state.

On a given symbol sequence $\dots \sigma_{i_1} \sigma_{i_2} \dots \sigma_{i_l} \dots$ generated from the time series data collected at a slow time epoch, a window of length D is moved by keeping a count of occurrences of word sequences $\sigma_{i_1} \dots \sigma_{i_D} \sigma_{i_{D+1}}$ and $\sigma_{i_1} \dots \sigma_{i_D}$ which are respectively denoted by $N(\sigma_{i_1} \dots \sigma_{i_D} \sigma_{i_{D+1}})$ and $N(\sigma_{i_1} \dots \sigma_{i_D})$. Note that if $N(\sigma_{i_1} \dots \sigma_{i_D}) = 0$, then the state $q \equiv \sigma_{i_1} \dots \sigma_{i_D} \in \mathcal{Q}$ has zero probability of occurrence. For $N(\sigma_{i_1} \dots \sigma_{i_D}) \neq 0$, the transitions probabilities are then obtained by these frequency counts as follows:

$$\begin{aligned} \pi_{jk} &\equiv P(q_k | q_j) = \frac{P(q_k, q_j)}{P(q_j)} = \frac{P(\sigma_{i_1} \dots \sigma_{i_D} \sigma)}{P(\sigma_{i_1} \dots \sigma_{i_D})} \\ &\Rightarrow \pi_{jk} \approx \frac{N(\sigma_{i_1} \dots \sigma_{i_D} \sigma)}{N(\sigma_{i_1} \dots \sigma_{i_D})} \end{aligned} \quad (\text{A.3})$$

where the corresponding states are denoted by $q_j \equiv \sigma_{i_1} \sigma_{i_2} \dots \sigma_{i_D}$ and $q_k \equiv \sigma_{i_2} \dots \sigma_{i_D} \sigma$.

A.2. Stopping rule for determining symbol sequence length

This section presents a stopping rule that is necessary to find a lower bound on the length of the symbol sequence required for parameter identification of the stochastic matrix $\mathbf{\Pi}$. The stopping rule [37] is based on the properties of irreducible stochastic matrices [38]. The state transition matrix is constructed at the r th iteration (i.e. from a symbol sequence

of length r) as $\mathbf{\Pi}(r)$ that is an $n \times n$ irreducible stochastic matrix under stationary conditions. Similarly, the state probability vector $\mathbf{p}(r) \equiv [p_1(r)p_2(r) \dots p_n(r)]$ is obtained as

$$p_i(r) = \frac{r_i}{\sum_{j=1}^n r_j}, \quad (\text{A.4})$$

where r_i is the number of symbols in the i th state such that $\sum_{i=1}^n r_i = r$ for a symbol sequence of length r . The stopping rule makes use of the Perron–Frobenius theorem [38] to establish a relation between the vector $\mathbf{p}(r)$ and the matrix $\mathbf{\Pi}(r)$. Since the matrix $\mathbf{\Pi}(r)$ is stochastic and irreducible, there exists a unique eigenvalue $\lambda = 1$ and a corresponding left eigenvector $\mathbf{p}(r)$ (normalized to unity in the sense of absolute sum). The left eigenvector $\mathbf{p}(r)$ represents the state probability vector, provided that the matrix parameters have converged after a sufficiently large number of iterations. That is,

$$\mathbf{p}(r) = \mathbf{p}(r)\mathbf{\Pi}(r) \quad \text{as } r \rightarrow \infty \quad (\text{A.5})$$

Following equation (A.4), the absolute error between successive iterations is obtained such that

$$\|(\mathbf{p}(r) - \mathbf{p}(r+1))\|_\infty = \|\mathbf{p}(r)(\mathbf{I} - \mathbf{\Pi}(r))\|_\infty \leq \frac{1}{r}, \quad (\text{A.6})$$

where $\|\bullet\|_\infty$ is the max norm of the finite-dimensional vector \bullet .

To calculate the stopping point r_{stop} , a tolerance of η ($0 < \eta \ll 1$) is specified for the relative error such that

$$\frac{\|(\mathbf{p}(r) - \mathbf{p}(r+1))\|_\infty}{\|\mathbf{p}(r)\|_\infty} \leq \eta \quad \forall r \geq r_{\text{stop}}. \quad (\text{A.7})$$

The objective is to obtain the least conservative estimate for r_{stop} such that the dominant elements of the probability vector have smaller relative errors than the remaining elements. Since the minimum possible value of $\|\mathbf{p}(r)\|_\infty$ for all r is $\frac{1}{n}$, where n is the dimension of $\mathbf{p}(r)$, the best worst case value of the stopping point is obtained from equations (A.6) and (A.7) as

$$r_{\text{stop}} \equiv \text{int}\left(\frac{n}{\eta}\right), \quad (\text{A.8})$$

where $\text{int}(\bullet)$ is the integer part of the real number \bullet .

References

- [1] Ozekici S 1996 *Reliability and Maintenance of Complex Systems (NATO Advanced Science Institutes (ASI) Series F: Computer and Systems Sciences vol. 154)* (Berlin: Springer)
- [2] Meggiolaro M and Castro J 2004 Statistical evaluation of strain-life fatigue crack initiation predictions *Int. J. Fatigue* **26** 463–76
- [3] Johannesson P, Svensson T and Jacques M 2005 Fatigue life prediction based on variable amplitude tests—methodology *Int. J. Fatigue* **27** 954–65
- [4] Ishihara S and McEvily A 2002 Analysis of short fatigue crack growth in cast aluminium alloys *Int. J. Fatigue* **24** 1169–74
- [5] Bjerken C and Melin S 2003 A tool to model short crack fatigue growth using a discrete dislocation formulation *Int. J. Fatigue* **25** 559–66
- [6] Chapetti M 2003 Fatigue propagation threshold for short cracks under constant amplitude loading *Int. J. Fatigue* **25** 1319–26
- [7] Ramsamooj D 2003 Analytical prediction of short to long fatigue crack growth rate using small- and large-scale yielding fracture mechanics *Int. J. Fatigue* **25** 923–33
- [8] Pathria R 1996 *Statistical Mechanics* (Amsterdam: Elsevier)
- [9] Sobczyk K and Spencer B 1992 *Random Fatigue: Data to Theory* (Boston, MA: Academic)
- [10] Abarbanel H 1996 *The Analysis of Observed Chaotic Data* (New York: Springer)
- [11] Grondel S, Delebarre C, Assaad J, Dupuis J and Reithler L 2002 Fatigue crack monitoring of riveted aluminium strap joints by Lamb wave analysis and acoustic emission measurement techniques *NDT&E Int.* **35** 137–46
- [12] Cook D and Berthelot Y 2001 Detection of small surface-breaking fatigue cracks in steel using scattering of Rayleigh waves *NDT&E Int.* **34** 483–92
- [13] Zilberstein V, Walrath K, Grundy D, Schlicker D, Goldfine N, Abramovici E and Yentzer T 2003 Mwm eddy-current arrays for crack initiation and growth monitoring *Int. J. Fatigue* **25** 1147–55
- [14] Anson L, Chivers R and Puttick K 1995 On the feasibility of detecting pre-cracking fatigue damage in metal matrix composites by ultrasonic techniques *Compos. Sci. Technol.* **55** 63–73
- [15] Vanlanduit S, Guillaume P and Linden G 2003 Online monitoring of fatigue cracks using ultrasonic surface waves *NDT&E Int.* **36** 601–7
- [16] Rokhlin S and Kim J-Y 2003 *In situ* ultrasonic monitoring of surface fatigue crack initiation and growth from surface cavity *Int. J. Fatigue* **25** 41–9
- [17] Kenderian S, Berndt T, Green R and Djordjevic B 2003 Ultrasonic monitoring of dislocations during fatigue of pearlitic rail steel *Mater. Sci. Eng.* **348** 90–9
- [18] Lind D and Marcus M 1995 *An Introduction to Symbolic Dynamics and Coding* (Cambridge: Cambridge University Press)
- [19] Badii R and Politi A 1997 *Complexity Hierarchical Structures and Scaling in Physics* (Cambridge: Cambridge University Press)
- [20] Daw C, Finney C and Tracy E 2003 A review of symbolic analysis of experimental data *Rev. Sci. Instrum.* **74** 915–30
- [21] Ray A 2004 Symbolic dynamic analysis of complex systems for anomaly detection *Signal Process.* **84** 1115–30
- [22] Chin S, Ray A and Rajagopalan V 2005 Symbolic time series analysis for anomaly detection: A comparative evaluation *Signal Process.* **85** 1859–68
- [23] Gupta S, Ray A and Keller E 2006 Symbolic time series analysis of ultrasonic data for early detection of fatigue damage *Mechanical Syst. Signal Process.* at press (Available online at www.sciencedirect.com)
- [24] Rajagopalan V and Ray A 2006 Symbolic time series analysis via wavelet-based partitioning *Signal Process.* at press (Available online at www.sciencedirect.com)
- [25] Keller E and Ray A 2003 Real time health monitoring of mechanical structures *Struct. Health Monit.* **2** 191–203
- [26] Keller E E 2001 Real time sensing of fatigue crack damage for information-based decision and control *PhD Thesis* Department of Mechanical Engineering, Pennsylvania State University, State College, PA
- [27] Davidchack R, Lai Y, Bolt E and Dhamala H 2000 Estimating generating partitions of chaotic systems by unstable periodic orbits *Phys. Rev. E* **61** 1353–6
- [28] Kennel M and Buhl M 2003 Estimating good discrete partitions from observed data: symbolic false nearest neighbors *Phys. Rev. E* **67** 084102
- [29] Mallat S 1998 *A Wavelet Tour of Signal Processing 2/e* (New York: Academic)
- [30] Cover T M and Thomas J A 1991 *Elements of Information Theory* (New York: Wiley)
- [31] Beck C and Schlögl F 1993 *Thermodynamics of Chaotic Systems: An Introduction* (Cambridge: Cambridge University Press)
- [32] Gupta S, Ray A and Srivastav A 2006 Pattern identification in complex systems: a statistical thermodynamic approach *Int.*

-
- Conf. on Manufacturing Science and Engineering: ASME 2006 (Michigan, MI)* under review
- [33] Suresh S 1998 *Fatigue of Materials* (Cambridge: Cambridge University Press)
- [34] Rose J 2004 *Ultrasonic Waves in Solid Media* (Cambridge: Cambridge University Press)
- [35] *Wavelet Toolbox for Use with MATLAB* 1996 (Natick, MA: Mathworks)
- [36] Berkovits A and Fang D 1995 Study of fatigue crack characteristics by acoustic emission *Eng. Fract. Mech.* **51** 401–16
- [37] Ray A 2005 Signed real measure of regular languages for discrete-event supervisory control *Int. J. Control* **78** 949–67
- [38] Bapat R and Raghavan T 1997 *Nonnegative Matrices and Applications* (Cambridge: Cambridge University Press)