ELSEVIER

Applied Mathematical Modelling 25 (2001) 979–994

APPLIED MATHEMATICAL MODELLING

www.elsevier.com/locate/apm

Fatigue crack growth under variable-amplitude loading: Part I – Model formulation in state-space setting ☆

Asok Ray *, Ravindra Patankar

Mechanical Engineering Department, The Pennsylvania State University, 137 Reber Building, University Park, PA 16802-1412, USA

Received 13 December 1999; received in revised form 5 January 2001; accepted 12 February 2001

Abstract

This two-part paper presents formulation and validation of a non-linear dynamical model of fatigue crack growth in ductile alloys under variable-amplitude loading including single-cycle overloads, irregular sequences, and random loads. The model is formulated in the state-space setting based on the crack closure concept and captures the effects of stress overload and reverse plastic flow. The state variables of the model are crack length and crack opening stress. This paper, which is the first part, presents formulation of the state-space model that can be restructured as an autoregressive moving average (ARMA) model for real-time applications such as health monitoring and life extending control. The second part is the companion paper that is dedicated to model validation with fatigue test data under different types of variable-amplitude and spectrum loading. © 2001 Published by Elsevier Science Inc.

Keywords: Fatigue and fracture; State-space modeling; Crack retardation; Sequence effects

1. Introduction

Models of fatigue crack growth under variable-amplitude loading (e.g., [6,17]) usually rely on a memory-dependent physical variable (e.g., crack opening stress, or reference stress) that requires storage of information on the load history. For example, the crack-opening stress in the FASTRAN model [17] is assumed to depend on the load history over an interval of about 300 cycles. Another example is the strain-life model in which the reference stress obtained by the rainflow method relies on cycle counting that, in turn, depends on the loading history [3,21]. In the current state of the art of fatigue crack growth modeling, the finite interval over which the load history is considered to be relevant may vary with the type of loading as well as with the rules employed for cycle counting. The model predictions, in general, become more accurate if the load history is considered over a longer period. In some instances, however, a short recent history of the applied load might be adequate for crack growth modeling. An extreme example is constant-amplitude

☆ The research work reported in this paper has been supported in part by National Science Foundation under grant no. CMS-9819074; National Academy of Sciences under a research fellowship award to the first author; ARINC Corporation under NASA Langley Research Center Cooperative Agreement No. NCC1-333.

* Corresponding author. Tel.: +1-814- 865-6377; fax: +1-814-863-4848.
E-mail address: axr2@psu.edu (A. Ray).

0307-904X/01/$ - see front matter © 2001 Published by Elsevier Science Inc.
PII: S0307-904X(01)00026-9

Nomenclature

A_k^j	parameter in the empirical equation of S_k^{oss} for $j = 1, 2, 3, 4$
a	crack length
E	Young's modulus
$F(\cdot, \cdot)$	crack length dependent geometry factor
$h(\cdot)$	crack growth function in crack growth equation
ℓ	crack growth retardation delay in cycles
k	current cycle of applied stress
R	stress ratio of minimum stress to maximum stress
S^{flow}	flow stress
$S^{\max}$	maximum stress within a cycle
$S^{\min}$	minimum stress within a cycle
S^{o}	crack opening stress
S^{oss}	crack opening stress under constant amplitude load given by empirical equation
S^{ult}	ultimate tensile strength
S^{y}	yield stress
t	specimen thickness
$U(\cdot)$	the Heaviside function
w	half-width of center-cracked specimen or width of compact specimen
α	constraint factor for plane stress/strain
$\alpha_{\max}$	maximum value of α
$\alpha_{\min}$	minimum value of α
$\Delta a^{\max}$	crack increment above which $\alpha = \alpha^{\min}$
$\Delta a^{\min}$	crack increment below which $\alpha = \alpha^{\max}$
Δa_k	crack increment $= (a_k - a_{k-1})$
ΔK^{eff}	effective stress intensity factor range
ε^{thr}	positive lower bound for absolute value of maximum stress $\{S_k^{\max}, k \geqslant 0\}$
η	decay rate for S^{o}
$\mathbb{R}$	the set of real numbers $(-\infty, \infty)$
τ	time instant
$\mathfrak{I}$	time interval of a cycle

cyclic loading where storage of the load history over the previous cycles may not be necessary. In essence, it is not precisely known to what extent information storage is necessary for calculating the memory-dependent variable in a fatigue crack growth model under a priori unknown variable-amplitude (e.g., single-cycle, block, spectrum, or random) loading. Nevertheless, this memory-dependent variable can be modeled in a finite-dimensional state-space setting by an ordinary difference (or differential) equation. The state at the current cycle is realized as a combination of the state and the input (i.e., cyclic stress) excitation at finitely many previous cycles. Equivalently, the state becomes a function of the fading memory of the input excitation, similar to that of an infinite-impulse response (IIR) discrete-time filter [27]. Holm et al. [7] have proposed an autoregressive (AR) model that generates the (cycle-dependent) crack opening stress as an output in response to the cyclic stress excitation. This concept can be generalized to an autoregressive moving average (ARMA) model [12] that is equivalent to a state-space model.

This two-part paper presents a non-linear model of fatigue crack growth under variable-amplitude loading in ductile alloys following the state-space approach. The proposed model,

hereafter referred to as the *state-space model*, is formulated based on the crack closure concept where the state variables are the crack length a and the crack-opening stress S^o. The crack growth equation in the state-space model is structurally similar to Paris equation [18] modified for crack closure, which has been used in the FASTRAN model [17]. However, the state-space model and FASTRAN use entirely different algorithms for calculating the crack opening stress S^o. As such, the crack length computed by these two models could be different under a given variable amplitude loading although the results are essentially identical under the same constant-amplitude loading. Unlike the existing crack growth models (e.g., [14,20]), the state-space model does not require a long history of stress excitation to calculate the crack-opening stress. Therefore, savings in the computation time and memory requirement are significant.

This paper is the first of a two-part sequence and is organized in five sections and one appendix. Section 2 establishes the rationale for selecting crack opening stress as a state variable. Appendix A summarizes the crack closure concept that is essential for understanding the physics of crack opening. Section 3 formulates the model equations in the state-space setting. Section 4 delineates the features of the state-space model including its characteristics in response to overload and underload excitation. Section 5 summarizes and concludes the first part of the two-part sequence. Part II, which is the companion paper [31], reports model validation with fatigue test data under different types of variable-amplitude and spectrum loading.

2. State variable representation of crack-opening stress

The results of experimental research in fatigue crack growth over the past several decades provide a knowledge base to explore the possibility of modeling the crack-opening stress S^o as a state variable. Schijve [24] collected fatigue crack data for specimens made of 2024-T3 aluminum alloy sheets in uniaxial tension for a constant amplitude load with $S^{max} = 147$ MPa and $S^{min} = 98$ MPa. The experiments were repeated with the same constant amplitude load with the exception of a single overload cycle with $S^{max} = 196$ MPa and $S^{min} = 98$ MPa when the crack length reached 15 mm. Two curves in each of the three plates of Fig. 1 shows the respective profiles of crack length, crack-opening stress, and crack growth rate per cycle, generated from the Schijve data with and without the overload effect. The extents of the shaded regions in two plates of Fig. 1 are qualitative due to inexact information on S^o recorded during the experiments (for example, see [24, p. 12]). Recently, Yisheng and Schijve [30] have observed from experimental data that, upon application of an overload, there is an immediate decrease in S^o followed by a rapid increase and a subsequent slow decrease. Similar results were reported earlier by Newman [15] based on analysis only. This transient behavior of S^o is, to some extent, similar to that of a (linear) non-minimum phase system [10]. Apparently, the initial sharp decrease and the follow-on abrupt increase of S^o that may occur only for a few cycles have no significant bearing on the overall crack growth. Generally, in response to an overload excitation, the crack growth rate exhibits an initial sharp increase and then monotonically increases starting from a lower value, as seen in the bottom plate of Fig. 1. This experimental observation is in agreement with the Paris equation modified for crack-closure, as explained below.

The net effect of a single-cycle overload is an abrupt increase in $(S^{max} - S^o)$, resulting in an increase in the crack growth increment in the present cycle. As S^{max} returns to its pre-overload value, the consequence of increased S^o is a sharp decrease in $(S^{max} - S^o)$, which causes the crack growth rate to diminish (or arrested if $(S^{max} - S^o)$ is non-positive). Shortly after expiration of the overload, S^o starts decreasing slowly from its increased value as seen in the middle plate of Fig. 1. As S^{max} returns to its pre-overload value, The result is a decrease in $(S^{max} - S^o)$, which causes the

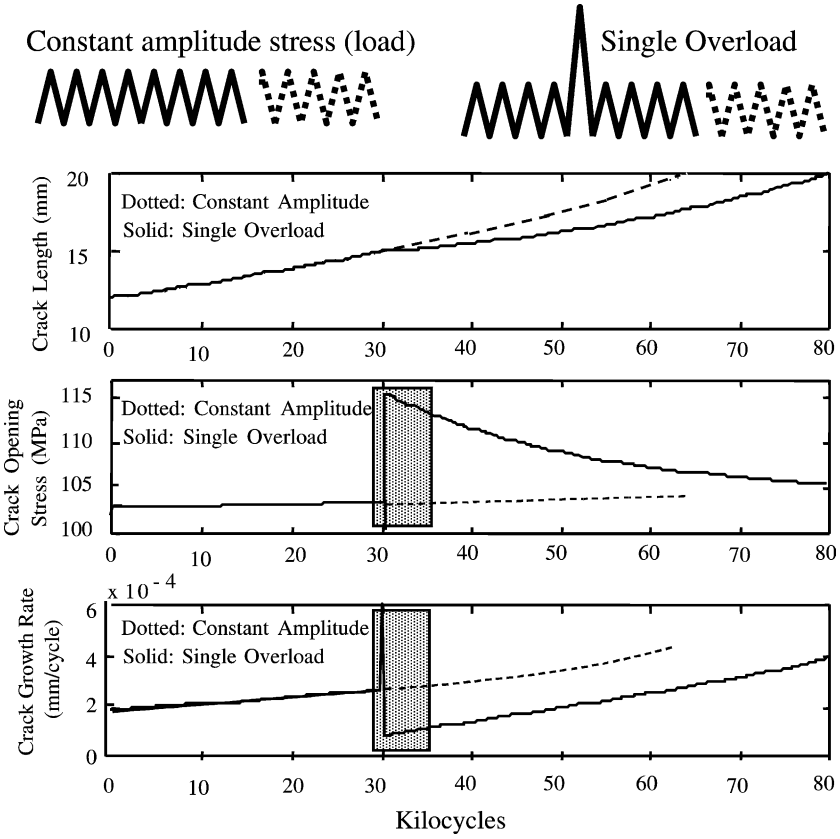

Fig. 1. Overload response of a typical crack growth process.

crack growth rate to diminish. Subsequently, under constant amplitude load, as S^o slowly relaxes back to its original value, crack growth rate also reaches its original higher value as seen in the bottom plate of Fig. 1. The crack growth is therefore retarded due to the fast rise and subsequent slow decay of S^o that can be attributed to the crack-closure effect. Some models assume that crack-closure effects are responsible for the retardation phenomenon [17], while others consider the plastic zone in front of the crack tip [29] to be the root cause. A physical explanation of crack retardation due to enlarged plastic zone is presented below.

Fig. 2 shows the plastic zone radius at the crack tip at different stages. Under constant amplitude loading, the plastic zone size is relatively small as indicated by a black oval. The plastic zone can be perceived as the material's resistance to crack growth. When a single overload is applied as shown by gray color in Fig. 2, the resulting plastic zone becomes larger. When a constant amplitude stress is resumed after the overload, the crack has to propagate through the larger plastic zone shown in gray color. Crack growth through this larger plastic zone is severely retarded because of material's increased resistance to crack growth. Once the crack grows out of the overload plastic zone, a normal crack growth rate prevails upon reaching the plastic zone of original size indicated by the black oval in Fig. 2.

Newman [14] assumed the plastic zone radius ρ to be dependent on the maximum stress in the cycle but independent of the minimum stress. Although such a model for plastic zone can capture the effects of a single-cycle overload effect, it may not be able to adequately explain the well-known sequence effects. For example, Jacoby et al. [9] demonstrated that the crack growth retardation due to crack-closure could possibly be effective beyond the overload plastic zone.

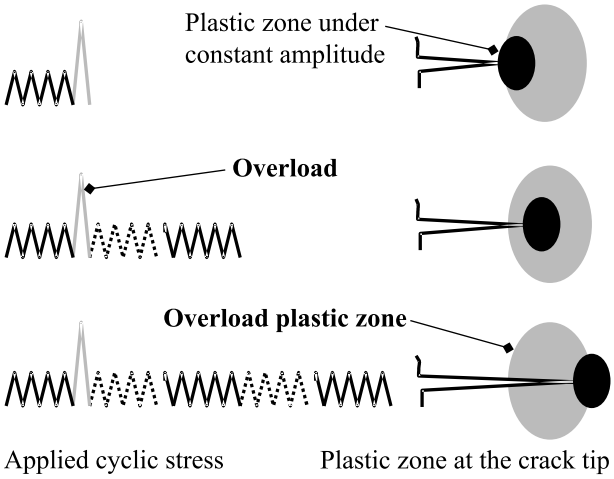

Fig. 2. Impact of fatigue overload on plastic zone size.

Therefore, selection of S^o as a state variable is preferable to that of the plastic zone radius ρ. Schijve [24, p. 19] made the following conclusion after studying various data sets of crack growth:

> The problem of predicting crack growth rates in service cannot be solved without a thorough knowledge of load-time history occurring in service $\cdots$ and knowledge of load sequences is essential.

Treating S^o as a state variable supports the above observation.

Currently available fatigue crack growth models under variable-amplitude loading calculate the crack opening stress, based on a finite history of cyclic stress excitation. These empirical procedures are often cumbersome and computationally intensive. If S^o is selected as a state variable, it can be obtained as the recursive solution of a low-order difference equation. Consequently, depending on the order of the difference equation, the current value of S^o is a function of its finitely many previous values that contain sufficient information on the input stress history required to obtain the crack-opening stress in the next cycle. This state-space model not only simplifies the computation of S^o to a great extent but also represents the physics of crack growth much better with a constitutive equation than with various empirical procedures involving the history of cyclic stress excitation. Dependence on the history of stress excitation can be captured by assuming the existence of at least one independent state variable in addition to the crack length in the crack growth equation. A state-space model for fatigue crack growth with crack-opening stress S^o as an additional state variable is therefore warranted.

3. Formulation of the crack growth model in state-space setting

The state-space model of crack growth is formulated based on mechanistic principles of the crack-closure concept which is briefly described in Appendix A, and is supported by fatigue test data for variable-amplitude cyclic loading (e.g., [13,19,24]). The following definition of a fatigue cycle is adopted for model development in the sequel:

Definition 3.1. The kth fatigue cycle is defined on the time interval:

$$\mathfrak{I}_k = \{\tau \in \mathbb{R} : \underline{\tau}_{k-1} < \tau \leqslant \underline{\tau}_k\} \quad \text{with } \underline{\tau}_{k-1} < \bar{\tau}_k < \underline{\tau}_k,$$

where $\underline{\tau}_k$ and $\bar{\tau}_k$ are the instants of occurrence of the minimum stress $S_k^{\min}$ and the maximum stress $S_k^{\max}$, respectively. The kth fatigue cycle is denoted as the ordered pair $(S_k^{\max}, S_k^{\min})$.

Remark 3.1. A stress cycle is determined by the maximum stress $S^{\max}$ and the following minimum stress $S^{\min}$. The frequency and the shape of a stress cycle are not relevant for crack growth in ductile alloys at room temperature [1]. The load dependence of crack growth is assumed to be completely characterized by peaks and valleys of applied stress at temperatures significantly below one-third of the melting point (e.g. aluminum and ferrous alloys at room temperature). It follows from the above definition that $S_k^{\max} > \max(S_{k-1}^{\min}, S_k^{\min})$.

Before proceeding to develop the fatigue crack growth model, pertinent observations that are critical for model formulation and validation are summarized below:

1. An overload may introduce significant crack growth retardation. Up to certain limits, the tenure of crack retardation effects is increased by:
   - larger magnitudes of the overload excitation;
   - periodic repetition of the overload during the crack propagation life; and
   - application of short blocks of overload instead of isolated single-cycle overloads.
2. Crack retardation may not always immediately follow the application of an overload. There could be a short delay before the crack growth rate starts decreasing. Under some circumstances, a small initial acceleration in crack growth has been observed. The delayed retardation in crack growth due to overload was clearly verified by observation of striation spacing [22].
3. The instantaneous crack growth caused by an overload itself is larger than that expected from a constant-amplitude load equal to the amplitude of the overload. This observation has been confirmed by fractography [13]. The rationale is that the crack opening stress S^{o} picks up in magnitude a few cycles after application of the overload whereas, for constant amplitude load, S^{o} is already at its steady state value equal to S^{oss}. Therefore, the crack growth rate while S^{o} is increasing due to a large $S^{\max}$ is higher than the rate when S^{o} has the steady-state value S^{oss}.
4. An underload has smaller effects on crack growth than an overload of the same magnitude [19]. However, an underload applied immediately after an overload may significantly compensate for the effects of crack growth retardation due to the overload [9,19,23]. If the underload precedes the overload, the compensation is much smaller due to a sequence effect of the overload cycles.
5. In step loading, a high-low sequence produces qualitatively similar results as overload cycles including delayed retardation [22]. Interaction effects after a high-low sequence are barely detectable in the macroscopic sense. However, more accurate measurements and striations do reveal existence of locally accelerated crack growth according to [13].
6. Duration of crack growth retardation depends upon ductility of the material. If ductility of an alloy is modified by heat treatment, a lower (higher) yield strength corresponds to a longer (shorter) retardation period. Moreover, the specimen geometry also affects the retardation period. Schijve [24] tested specimens of different thickness under equivalent single-cycle overload conditions. A reduction in retardation period was observed with increase in thickness.
7. Rest periods at zero stress following a tensile peak overload have no significant influence on subsequent fatigue crack retardation for ductile alloys at room temperature [25].
8. The approximate non-minimum phase behavior of crack opening stress, observed by Yisheng and Schijve [30] and Newman [15] as discussed earlier in Section 2, is explained as follows. Upon application of an overload, S^{o} decreases sharply and then rapidly undergoes an overshoot followed by a slow decay. Similarly, an underload would cause a sharp increase in S^{o} before an undershoot is observed. Dabayeh and Topper [2] measured crack-opening stress on

2024-T351 aluminum alloy specimens using 900 power short focal length optical microscope at 1, 5, 10, 50, 100 cycles immediately after application of an overload. The non-minimum phase behavior of S^{o} was not observed in any one of those specimens. Therefore, existence of the non-minimum phase behavior in the transient response of S^{o} is debatable at this moment. Since the transients having the non-minimum phase behavior, if they exist, are fast, their contributions to overall crack growth are considered to be insignificant relative to the total fatigue life.

Most fatigue crack growth models reported in technical literature are based on modifications of the Paris equation [18] in which the inputs are $S_k^{\max}$ and $S_k^{\min}$ in the kth cycle and the output is the crack length increment Δa_k. It is customary in the fracture mechanics community [1,26] to express the dynamical behavior of fatigue crack growth as a derivative $\mathrm{d}a/\mathrm{d}N$ with respect to the number of cycles, which is essentially Δa_k in the kth cycle as delineated below:

$$\left.\begin{aligned} &\Delta a_k \equiv a_k - a_{k-1} = h\left(\Delta K_k^{\mathrm{eff}}\right) \text{ with } h(0) = 0 \\ &\Delta K_k^{\mathrm{eff}} \equiv \sqrt{\pi a_{k-1}} F(a_{k-1}, w)\left(S_k^{\max} - \max(S_k^{\min}, S_{k-1}^{o}) U\left(S_k^{\max} - S_{k-1}^{o}\right)\right) \end{aligned}\right\} \text{ for } k \geqslant 1 \text{ and } a_0 > 0, \tag{1}$$

where a_{k-1} and S_{k-1}^{o} are the crack-length and the crack-opening stress, respectively, during the kth cycle and change to a_k and S_k^{o} at the expiry of the kth cycle; $F(\cdot,\cdot)$ is a crack-length-dependent correction factor compensating for finite geometry of the specimen with the width parameter w; the non-negative monotonically increasing function $h(\cdot)$ can be represented either by a closed form algebraic equation or by table lookup [17]; and

$$U(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geqslant 0 \end{cases}$$

is the Heaviside unit step function..

Eq. (1) is a first-order non-linear difference equation excited by $S_k^{\max}$ and S_{k-1}^{o} in the kth cycle. Apparently, the crack length a_k can be treated as a state variable in Eq. (1). However, since S^{o} is dependent on the stress history (i.e., the ensemble of peaks $S^{\max}$ and valleys $S^{\min}$ in the preceding cycles), Eq. (1) cannot be readily represented in the state-space setting in its current form. The task is now to make a state-variable representation of the evolution of S^{o} under variable-amplitude cyclic stress excitation, and then augment the crack growth model in Eq. (1) with S^{o} as an additional state variable. It is postulated that a state-space model of crack growth is observable [28]. In other words, the state variables in any given cycle can be determined from the history of measured variables over a finite number of cycles. The crack length a_k is assumed to be measurable. The other state variable, the crack-opening stress S^{o}, can be determined from a finite history of the input (i.e., peaks and valleys of stress excitation) and the output (i.e., crack length measurements), starting from a particular cycle in the past onwards to the current cycle. This concept is analogous to the methods used in the existing crack growth models where either the crack-opening stress or the reference stress is obtained based on the history of cyclic stress excitation.

It is observed from experimental data that S^{o} requires a short period of cycles to rise to a peak value after the application of a single-cycle overload. If a first-order difference equation is postulated to model the transient behavior of S^{o}, then S^{o} can be depicted to have an instantaneous rise, which is a good approximation for most ductile alloys. The application of an overload should generate a positive pulse to excite an appropriate state-space equation. Moreover, once this overload pulse reaches its peak, decay of S^{o} should be very slow. Hence, upon application of a large positive overload, the peak of S^{o} may be significantly larger than its steady-state value. Upon application of another small overload when S^{o} is still larger than its steady-state value, the

smaller overload should not have any significant effect. In other words, a small overload following a large overload should not generate a pulse input to the state-space equation. This implies non-linearity of the forcing function that can be captured by a Heaviside function. As the non-linearity is dependent upon the current value of S^o, a low-order non-linear difference equation can provide a viable model for describing the transient behavior of S^o under overload conditions.

The plastic zone size is largest during a load cycle when S^{max} is applied. As the applied stress is decreased from S^{max}, there is a reverse plastic flow at the crack tip [11]. The reverse plastic flow is at its maximum when the minimum stress S^{min}, even if positive in value, is applied. This reverse plastic flow depletes the large plastic zone caused by S^{max}. If the crack growth leads into a large overload-plastic zone and if an underload is applied next, then depletion of the plastic zone is higher than the one that would be caused by a regular (i.e., higher) S^{min}. This effect reduces the protection against crack growth, which can be stated in other terms as a decrease in S^o. Lardner [11] modeled an elastoplastic shear crack in which the crack was replaced by a linear array of freely slipping dislocations and the plastic zones by coplanar arrays moving against a frictional resistance. Fig. 3 depicts the dislocation density curves by the dotted line $f(x)$ at the peak stress S^{max} and the solid line $g(x)$ at S^{min} following S^{max} where the crack spans the range $\{x \in \mathbb{R} : |x| \leqslant c\}$ and the plastic zone spans the range $\{x \in \mathbb{R} : c \leqslant |x| \leqslant a\}$. The dislocation densities $f(x)$ and $g(x)$ are identical except for the region $\{x \in \mathbb{R} : c \leqslant |x| \leqslant d\}$ where reverse plastic flow occurs. It was observed from the analytical solution that the plastic zone is depleted with a smaller S^{min} and that it is increased with a larger S^{max}.

Rainflow cycle counting [21] has been used in variable-amplitude fatigue models to generate the reference stress, which is analogous to the crack-opening stress to some extent. The rainflow technique remembers the stress history back to the occurrence of least minimum stress. If the new minimum stress is lower than the previous minimum stress, then cycles are counted according to a rule between these two minimum stresses and the stored stress profile is updated starting from the new minimum stress. This is analogous to encountering a new S^{min} that is lower than its past values. This new S^{min} causes a large reverse plastic displacement leading to severe depletion of the plastic zone, wherefrom it has to be built up again by continued application of the stress profile. When the plastic zone is severely depleted, the memory of the previous plastic zone is destroyed and a new memory is built up as the load is applied further on. To accurately predict the crack growth, the state-space model must be able to account for the entire reverse plastic flow.

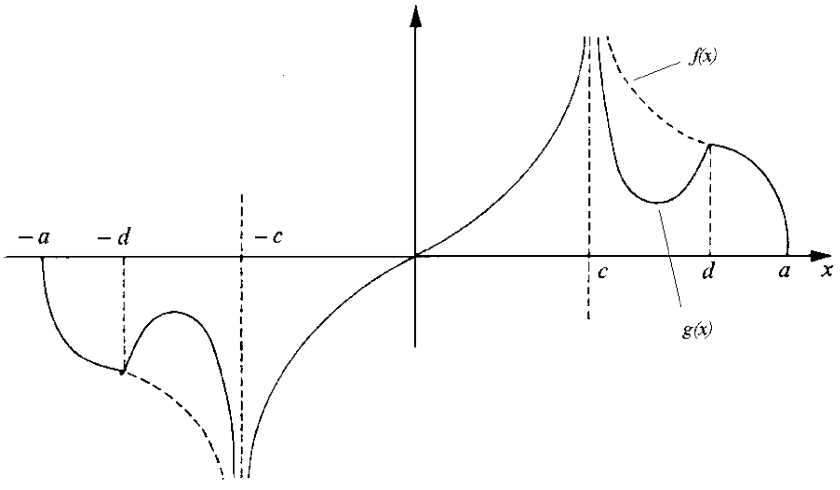

Fig. 3. Dislocation density solution [11].

We now proceed to determine the structure of the difference equation that is excited by the cyclic stress input to generate the crack opening stress. To this end, we first consider the steady-state solution of the difference equation under constant amplitude load. This issue has been addressed by several investigators including Newman [16] and Ibrahim et al. [8]. The steady-state crack-opening stress S^{oss} under a constant amplitude cyclic load is a function of the minimum stress $S^{\min}$, the maximum stress $S^{\max}$, the constraint factor $S^{\min}$ (which is 1 for plane stress and 3 for plane strain), the specimen geometry, and the flow stress S^{flow} (which is the average of the yield strength S^{y} and the ultimate strength S^{ult}). These relationships are shown to be good for most ductile alloys by Newman [16]. One such empirical relation has been used in the FASTRAN model [17].

The objective is to construct the difference equation for (cycle-dependent and non-negative) crack opening stress S_k^{o} such that, under different levels of constant-amplitude load, the forcing function S_k^{oss} at the kth cycle matches the crack-opening stress derived from the following empirical relation [16] that is valid for non-zero peak stress:

$$S_k^{\text{oss}} = S^{\text{oss}}(S_k^{\max}, S_k^{\min}, \alpha, F) = \left(A_k^0 + A_k^1 R_k + A_k^2 (R_k)^2 + A_k^3 (R_k)^3\right) S_k^{\max}, \tag{2}$$

where

$$R_k = \frac{S_k^{\min}}{S_k^{\max}} U\left(S_k^{\max}\right), \quad \text{for all } k \geqslant 0, \tag{3}$$

$$A_k^0 = \left(0.825 - 0.34\alpha_k + 0.05(\alpha_k)^2\right)\left[\cos\left(\frac{\pi}{2}\frac{S_k^{\max}}{S^{\text{flow}}} F(a_{k-1}, w)\right)\right]^{1/\alpha_k}, \tag{4}$$

$$A_k^1 = (0.415 - 0.071\alpha_k)\left(\frac{S_k^{\max}}{S^{\text{flow}}} F(a_{k-1}, w)\right), \tag{5}$$

$$A_k^2 = \left(1 - A_k^0 - A_k^1 - A_k^3\right) U(R_k), \tag{6}$$

$$A_k^3 = \left(2A_k^0 + A_k^1 - 1\right) U(R_k). \tag{7}$$

The constraint factor α_k used in Eqs. (4) and (5) is obtained as a function of the crack length increment Δa_k in Eq. (1). A procedure for evaluation of α_k is presented in Section 2 of the second part. Since α_k does not significantly change over cycles, it can be approximated as piecewise constant for limited ranges of crack length.

Remark 3.2. The inequality in the Heaviside function $U(S_k^{\max})$ of Eq. (3) should be realized by setting $S_k^{\max} \geqslant \varepsilon^{\text{thr}} > 0$ to avoid the singular region around $S_k^{\max} = 0$. The parameter ε^{thr} is selected for code development in Section 2 of the second part. This modification is not necessary for applications where the peak stress is sufficiently tensile.

The following constitutive relation in the form of a non-linear first-order difference equation is proposed for recursive computation of the crack-opening stress S_k^{o} upon the completion of the kth cycle:

$$\begin{aligned} S_k^{\text{o}} = {} & \left(\frac{1}{1+\eta}\right) S_{k-1}^{\text{o}} + \left(\frac{\eta}{1+\eta}\right) S_k^{\text{oss}} + \left(\frac{1}{1+\eta}\right)\left(S_k^{\text{oss}} - S_{k-1}^{\text{o}}\right) U\left(S_k^{\text{oss}} - S_{k-1}^{\text{o}}\right) \\ & + \left(\frac{1}{1+\eta}\right)\left[S_k^{\text{oss}} - S_k^{\text{oss\_old}}\right] U\left(S_{k-1}^{\min} - S_k^{\min}\right)\left[1 - U\left(S_k^{\text{oss}} - S_{k-1}^{\text{o}}\right)\right], \end{aligned} \tag{8}$$

$$\eta = \frac{tS^{y}}{2wE} \text{ for center-cracked specimens of finite width,} \tag{9}$$

where the forcing function S_k^{oss} in Eq. (8) is calculated from Eq. (2) as if a constant amplitude stress cycles (S_k^{max}, S_k^{min}) is applied; similarly, $S_k^{oss\_old}$ is given by Eq. (2) as if a constant amplitude stress cycle $(S_k^{max}, S_{k-1}^{min})$ is applied. For constant-amplitude loading, S^{oss} is the steady-state solution of S^{o}. In general, the inputs S_k^{oss} and $S_k^{oss\_old}$ to Eq. (8) are different from the instantaneous crack-opening stress S_k^{o} under variable-amplitude loading. The Heaviside function $U(S_k^{oss} - S_{k-1}^{o})$ in the third term on the right-hand side of Eq. (8) allows fast rise and slow decay of S^{o}. The last term on the right-hand side of Eq. (8) accounts for the effects of reverse plastic flow. Depletion of the normal plastic zone occurs when the minimum stress S_k^{min} decreases below its value S_{k-1}^{min} in the previous cycle, which is incorporated via the Heaviside function $U(S_{k-1}^{min} - S_k^{min})$. Note that the overload excitation and reverse plastic flow are mutually exclusive.

The dimensionless parameter η in Eq. (9) depends on the specimen thickness t, half-width w, yield strength S^{y}, and Young's modulus E. Following an overload cycle, the duration of crack retardation is controlled by the transients of S_k^{o} in the state-space model, and hence determined by the stress-independent parameter η in Eqs. (8) and (9). Physically, this duration depends on the ductility of the material that is dependent on many factors including the heat treatment of specimens [24]. Smaller yield strength produces a smaller value of η, resulting in longer duration of the overload effect. Smaller specimen thickness has a similar effect [24]. Although a precise relationship for η is not known at this time due to the lack of adequate test data for different materials and different geometry, η could be estimated from the experimental data of a single overload on an identical specimen made of the same material. In the absence of such data, the relationship in Eq. (9) could be used to generate an estimate of η for center-cracked flat specimens of finite width. For different geometrical configurations, η needs to be identified from experimental data under variable-amplitude load excitation.

The model equations (1)–(9) are summarized and re-arranged in Section 2 of the second part to provide adequate information for generating a fatigue crack growth simulation code. Note that the function $h(\Delta K^{eff})$ in Eq. (1) can be represented either by a closed form algebraic equation or by table lookup. These functional relationships and numerical values of the model parameters are available in the FASTRAN manual for different materials [17].

Next we address the issue of (possibly) additional delays associated with the transient response of crack opening stress S_k^{o}, which might be prevalent in some materials. In order to include the effects of delay ℓ (in cycles) in the response of S_k^{oss}, the right-hand side of Eq. (1) can be modified by altering ΔK_k^{eff} as:

$$\Delta a_k \equiv a_k - a_{k-1} = \left(\left(S_k^{max} - S_{k-\ell-1}^{o}\right)\sqrt{\pi a_{k-1}}F(a_{k-1})\right) \quad \text{with } S_{k-\ell-1}^{o} = S_k^{min} \text{ for } \ell \geqslant 0. \tag{10}$$

Since the experimental data may not exactly show the transients of S^{o} during and immediately after a variation of S^{max} or S^{min}, the model may not accurately depict S^{o} in this range. Nevertheless, this (possible) modeling inaccuracy has hardly any effect on overall crack growth. Starting with a higher order difference equation, the order (i.e., the number of state variables) of the present model is reduced to 2 by singular perturbation [28] based on the experimental data of 7075-T6 [19] and 2024-T3 [13] aluminum alloys. The possibility of a higher order model to represent non-minimum phase behavior or delayed response of S_k^{o} is not precluded for other materials.

Remark 3.3. Eq. (10) is identical to Eq. (1) for $\ell = 0$. In that case, the transient response of crack growth is subjected to a built-in delay of two cycles after the application of an overload pulse as seen by examination of Eqs. (1) and (8). For $\ell > 0$, the corresponding delay is $(\ell + 2)$ cycles.

4. Features of the state-space model

The most important feature of the state-space model, formulated in Section 3, is recursive computation of the crack opening stress without the need for information storage of stress excitation except for the minimum stress in the previous cycle. This is evident from the governing Eqs. (1) and (8) for a_k and S_k^o, respectively, that the two-dimensional state-space model of fatigue crack growth has the structure of an ARMA model [12]. In other words, the crack growth equations can be represented by a second-order non-linear difference equation that recursively updates the state variables, a_k and S_k^o, with S_k^{max}, S_k^{min} and S_{k-1}^{min} as inputs and the immediate past information on a_{k-1} and S_{k-1}^o; storage of no other information is required. This implies that the crack length and crack-opening stress in the present cycle are obtained as simple algebraic functions of the maximum and minimum stress in the present cycle as well as the minimum stress, crack length, and crack-opening stress in the immediately preceding cycle. While details of model validation with fatigue test data are presented in the second part, we present qualitative explanations of the model response in Figs. 4–6 to elucidate the role of the Heaviside functions in the constitutive equation of crack opening stress.

Fig. 4 presents qualitative features of the model response to a single-cycle overload excitation based on the first-order difference equation (8) in which S^o attains a peak value in the cycle following the application of a single-cycle overload. The positive edge of this resulting pulse is effective whereas, unlike a linear system, the negative edge is rendered ineffective by the Heaviside function $U(S_k^{oss} - S_{k-1}^o)$ and $U(S_{k-1}^{min} - S_k^o)$ as seen in the third and fourth terms on the right-hand side of Eq. (8). When $U(S_k^{oss} - S_{k-1}^o)$ is zero, S^o decreases at a rate determined by the dimensionless parameter η. Under constant-amplitude loading, the terms of Eq. (8) containing the Heaviside functions are ineffective. Upon application of a single-cycle overload, the amplitude of the input pulse S_k^{oss} on the right-hand side of Eq. (8) depends on the amount of overload which leads to retarded crack growth under the constant amplitude load that follows the overload.

In contrast to a single-cycle overload, a single-cycle underload makes the Heaviside function $U(S_k^{oss} - S_{k-1}^o)$ ineffective while the fourth term on the right-hand side of Eq. (8) becomes effective due to the Heaviside function $U(S_{k-1}^{min} - S_k^{min})$ that accounts for the reverse plastic flow and

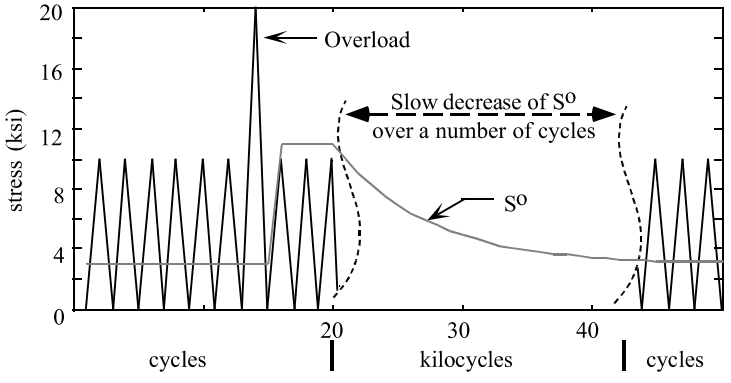

Fig. 4. Overload response of crack opening stress.

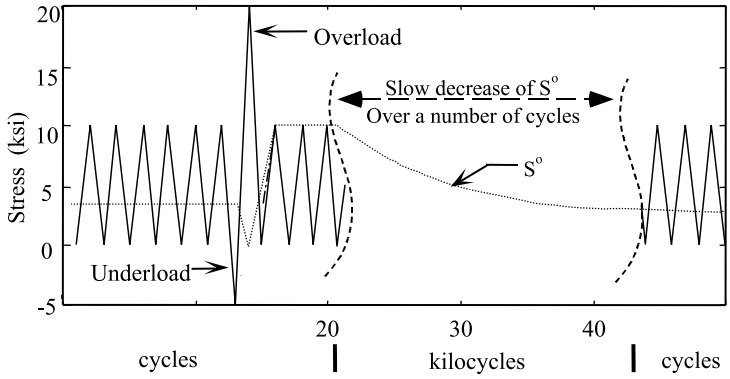

Fig. 5. Underload–overload response of crack opening stress.

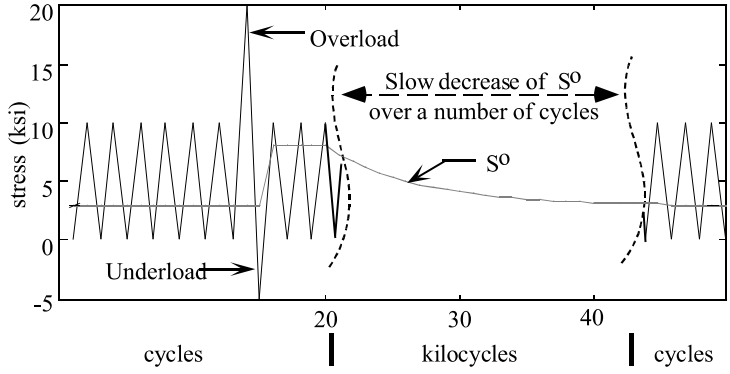

Fig. 6. Overload–underload response of crack opening stress.

depletion of the resulting plastic zone. When the load returns to its normal range from an underload, the Heaviside function $U(S_k^{\text{oss}} - S_{k-1}^{\text{o}})$ again becomes effective and the fourth term on the right-hand side of Eq. (8) becomes ineffective. This brings S^{o} back to its normal value. Thus S^{o} is low only for one cycle during a single-cycle underload that has very little bearing on the overall crack growth rate. Fig. 5 presents qualitative features of the model response to an underload followed by an overload. The major difference between the responses due to a single-cycle overload and an underload–overload is that the preceding underload causes S^{o} to be abnormally low. This reduces the protection of the crack from growing during the follow-on overload cycle and consequently the crack increment becomes larger. The model response after the overload is similar to that for a single-cycle overload.

Fig. 6 presents a qualitative view of how S^{o} is affected by an overload immediately followed by an underload. In the overload–underload cycle, $S_k^{\max}$ is identical to that for a single-cycle overload in Fig. 4 but the corresponding $S_k^{\min}$ is smaller and consequently S_k^{oss} is also smaller. In effect, the forcing function in the third term on the right-hand side of Eq. (8), which is multiplied by the Heaviside function $U(S_k^{\text{oss}} - S_{k-1}^{\text{o}})$, assumes a smaller value for overload–underload than that for a single-cycle overload while the fourth term is ineffective. A single-cycle overload thus retards crack growth more effectively than a similar overload immediately followed by an underload. In essence, the benefits of an overload monotonically diminish with increase in the magnitude of the following underload.

Next we compare the state-space model (which is ARMA) with the AR model proposed by Holm et al. [7]. While both models are piecewise linear and treat the crack-opening stress S^{o} as a state variable, there are several differences in the structures of their governing equations. Specifically, the state-space model uses mechanistic principles and takes advantage of fatigue test data while the AR model is largely empirical. The phenomenon of crack growth retardation requires only one constant parameter η in the state-space model. The AR model uses different constant parameters over the two halves of a cycle to represent the increase and decrease of S_k^{o}. A major drawback of having two different constants is that when frequent overloads (or underloads) are applied, S_k^{o} rises with each successive application and becomes unbounded. Consequently, the AR model is not capable of capturing the effects of a single overload, irregular load sequences, and random loads with the same set of constants. This problem does not arise in the state-space model as the excitation S_k^{oss} applied due to an overload is automatically adjusted by subtracting the current value of S_k^{o} as seen in the third term on the right-hand side of Eq. (8). The effects of an abrupt reduction in S^{min} during the crack retardation period are realized in the fourth term.

5. Summary and conclusions

This paper presents the physical concept and formulation of a fatigue crack growth model in the state-space setting to capture the effects of variable-amplitude cyclic loading in ductile alloys. The state-space model is formulated based on the crack closure concept and captures the effects of stress overload and reverse plastic flow. The model state variables are crack length and crack opening stress that are recursively generated by simple algebraic functions of the maximum and minimum stress in the present cycle as well as the minimum stress, crack length, and crack-opening stress in the immediately preceding cycle. The model is validated with different types of fatigue test data in the second part that is a companion paper.

The crack growth equation in the state-space model is structurally similar to that in the FASTRAN model [17]. However, these two models use entirely different algorithms for calculating the crack opening stress. As such, the crack length computed by the models could be different under a given variable amplitude loading but the results are essentially identical under the same constant-amplitude loading. Unlike the existing crack growth models that require a long history of stress excitation to calculate the crack-opening stress under variable-amplitude loading, the state-space model only needs the information on peak stresses in the present and past cycles. Therefore, savings in the computation time and memory requirement are significant.

Acknowledgements

The authors acknowledge beneficial technical discussions with Dr. James C. Newman, Jr. of NASA Langley Research Center.

Appendix A. The crack-closure concept

Experiments on metallic materials, conducted by Elber [5] demonstrated that fatigue cracks remain closed during a part of every load cycle under both constant-amplitude and variable-amplitude loading. Newman [14] thereafter has proposed an analytical crack-closure model based on the Dugdale model [4]. Plane stress and plane strain conditions are simulated by incorporating a constraint factor on tensile yielding to account for three-dimensional effects. In order to

calculate the contact stresses, also known as closure stresses, and crack-opening stress during crack propagation, the elastic-plastic solution for stresses and displacements in a cracked body must be known. The crack surface displacements, which are used to calculate the contact stresses during unloading, are influenced by plastic yielding at the crack tip and residual deformations left in the wake of an advancing crack. Upon reloading, the applied stress at which the crack surfaces become fully open (i.e., with no surface contact) is directly related to the contact stresses. This applied stress is called the crack-opening stress S^{o}. Measurements of S^{o} are difficult and have been made on only a few materials under a limited number of loading conditions.

Fig. 7 shows a schematic of the crack closure model [14] at the maximum and minimum applied stress. The model spans three regions:

Region 1: Linear elastic region with a (measurable) physical crack of half-length $(a + \rho)$;

Region 2: Plastic region of radius ρ at the crack tip; and

Region 3: Residual crack deformation region along the crack surfaces.

Region 1 is treated as an elastic continuum, and crack surface displacements can be found using available equations. Regions 2 and 3 are composed of rigid perfectly plastic (constant stress) bar elements. The shaded regions in both plates of Fig. 7 indicate the material that is in plastic state.

At any applied stress, the bar elements are either intact (in Region 2) or broken (in Region 3). The broken elements carry only compressive loads provided that the cracked surfaces are in contact. Elements that are not in contact do not affect the calculation of crack-opening displacements. The plastic zone size is determined by requiring that the stress intensity factor at the tip of the plastic zone is zero. The maximum radius ρ^{max} of the plastic zone is calculated based on the largest applied stress. The length of the bar elements in the plastic zone is calculated from a crack surface displacement equation [14].

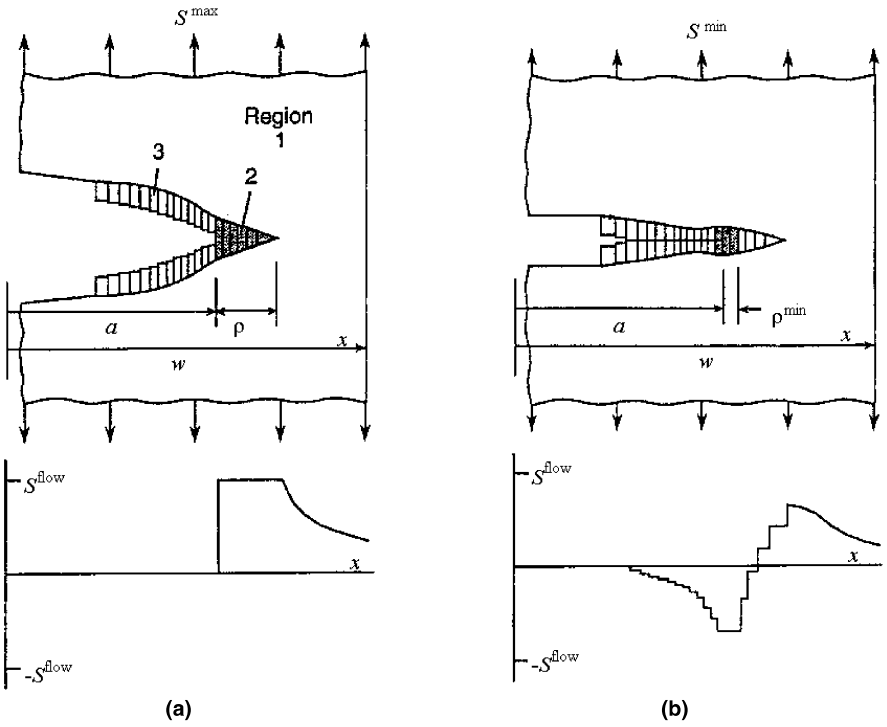

Fig. 7. Crack surface displacements and stress distribution along crack line [14]: (a) maximum stress; (b) minimun stress.

References

[1] T.L. Anderson, Fracture Mechanics, CRC Press, Boca Raton, FL, 1995.
[2] A.A. Dabayeh, T.H. Topper, Changes in crack-opening stress after underloads and overloads in aluminium alloy, International Journal of Fatigue 17 (4) (1995) 261–269.
[3] N.E. Dowling, Fatigue life prediction for complex load versus time histories, Journal of Engineering Materials and Technology, ASME Trans. 105 (1983) 206–214.
[4] D.S. Dugdale, Yielding in steel sheets containing slits, Journal of Mechanics and Physics of Solids 8 (2) (1960) 100–104.
[5] W. Elber, Fatigue crack closure under cyclic tension, Engineering Fracture Mechanics 2 (1970) 37–45.
[6] J.A. Harter, AFGROW Users' Guide and Technical Manual, Report No. AFRL-VA-WP-1999-3016, Air Force Research Laboratory, WPAFB, OH 45433-7542, 1999.
[7] S. Holm, B.L. Josefson, J. deMare, T. Svensson, Prediction fatigue life based on level crossings and a state variable, Fatigue and Fracture of Engineering Materials 18 (10) (1995) 1089–1100.
[8] F.K. Ibrahim, J.C. Thompson, T.H. Topper, A study of effect of mechanical variables on fatigue crack closure and propagation, International Journal of Fatigue 8 (3) (1986) 135–142.
[9] G.H. Jacoby, H.T.M. Van Lipzig, H. Nowack, Experimental results and a hypothesis for fatigue crack propagation under variable amplitude loading, fatigue crack growth under spectrum loads, ASTM STP 595 (1976) 172–183.
[10] T. Kailath, Linear Systems, Prentice-Hall, Englewood Cliffs, NJ, 1980.
[11] R.W. Lardner, Mathematical Theory of Dislocations and Fracture, University of Toronto Press, Toronto, Canada, 1974.
[12] L. Ljung, System Identification Theory for the User, second ed., Prentice-Hall, Englewood Cliffs, NJ, 1999.
[13] J.C. McMillan, R.M.N. Pelloux, Fatigue crack propagation under program and random loads, fatigue crack propagation, ASTM STP 415 (1967) 505–532 (Also Boeing Space Research Laboratory (BSRL) Document D1-82-0558, 1966).
[14] J.C. Newman Jr., A Crack-closure model for predicting fatigue crack growth under aircraft loading, methods and models for predicting fatigue crack growth under random loading, ASTM STP 748 (1981) 53–84.
[15] J.C. Newman Jr., Prediction of fatigue crack growth under variable-amplitude and spectrum loading using a closure model design of fatigue and fracture resistant structures, ASTM STP 761 (1982) 255–277.
[16] J.C. Newman Jr., A crack opening stress equation for fatigue crack growth, International Journal of Fracture 24 (1984) R131–R135.
[17] J.C. Newman Jr., FASTRAN-II - A Fatigue Crack Growth Structural Analysis Program, NASA Technical Memorandum 104159, Langley Research Center, Hampton, VA 23665, USA, 1992.
[18] P.C. Paris, F. Erdogan, A critical analysis of crack propagation laws, Journal of Basic Engineering, ASME Trans. 85 (1960) 528–534.
[19] T.R. Porter, Method of analysis and prediction for variable amplitude fatigue crack growth, Engineering Fracture Mechanics 4 (1972) 717–736.
[20] A. Ray, M.K. Wu, M. Carpino, C.F. Lorenzo, Damage-mitigating control of mechanical systems: Parts I and II, Journal of Dynamic Systems Measurement and Control, ASME Trans. 116 (3) (1994) 437–455.
[21] I. Rychlik, Note on cycle counts in irregular loads, Fatigue and Fracture of Engineering Materials and Structures 16 (4) (1993) 377–390.
[22] J. Schijve, F.A. Jacobs, P.J. Tromp, The effect of load sequence under fatigue crack propagation under random loading and program loading, NLR TR 71014 U, National Aerospace Laboratory NLR, The Netherlands, 1971.
[23] J. Schijve, Fatigue damage accumulation and incompatible crack front orientation, Engineering Fracture Mechanics 6 (1974) 245–252.
[24] J. Schijve, Observations on the prediction of fatigue crack growth propagation under variable-amplitude loading, fatigue crack growth under spectrum loads, ASTM STP 595 (1976) 3–23.
[25] W.N. Sharpe Jr., D.M. Corbly, A.F. Grandt Jr., Effects of rest time on fatigue crack retardation and observation of crack closure, fatigue crack growth under spectrum loads, ASTM STP 595 (1976) 61–77.
[26] S. Suresh, Fatigue of Materials, Cambridge University Press, Cambridge, UK, 1991.
[27] P.P. Vaidyanathan, Multirate Systems and Filter Banks, Prentice-Hall PTR, Englewood Cliffs, NJ, 1993.
[28] M. Vidyasagar, Nonlinear Systems Analysis, second ed., Prentice-Hall, Englewood Cliffs, NJ, 1992.
[29] O.E. Wheeler, Spectrum loading and crack growth, Journal of Basic Engineering, ASME Trans. 94 (1972) 181–186.

[30] W. Yisheng, J. Schijve, Fatigue crack closure measurements on 2024-T3 sheet specimens, Fatigue and Fracture of Engineering Materials and Structures 18 (9) (1995) 917–921.
[31] A. Ray, R. Patankar, Appl. Math. Modelling 25 (2001) 995–1013, Part II.