Determination of energy dissipation during cyclic loading and its use to predict fatigue life of metal alloys

Kassapoglou, Christos

DOI
10.1111/ffe.14096

Publication date
2023

Document Version
Final published version

Published in
Fatigue and Fracture of Engineering Materials and Structures

Received: 10 April 2023 | Revised: 19 June 2023 | Accepted: 21 June 2023

DOI: 10.1111/ffe.14096

ORIGINAL ARTICLE

FFEMS Fatigue & Fracture of Engineering Materials & Structures WILEY

Determination of energy dissipation during cyclic loading and its use to predict fatigue life of metal alloys

Christos Kassapoglou

Aerospace Structures and Materials Section, Aerospace Engineering Department, Delft University of Technology, Delft, Netherlands

Correspondence
Christos Kassapoglou, Aerospace Structures and Materials Section, Aerospace Engineering Department, Delft University of Technology, Delft, Netherlands.
Email: c.kassapoglou@tudelft.nl

Abstract

An approach to determine the specific energy dissipated during cyclic loading of metal alloys with load ratios between 0 and 1 is developed. The dissipated energy per cycle is determined by using a limiting procedure to obtain the area enclosed between successive loading and unloading curves and can be used for low-, medium-, and high-cycle fatigue. Comparisons with published data from two aluminum, one titanium, and one steel alloy show that the predictions capture the test results very well. A universal curve relating cycles to failure to two nondimensional parameters is derived. It is shown, in some limiting cases, that the method leads to power law equations with the parameters in these equations determined directly from the complete stress–strain curve of the material with no need to fit experimental data. Furthermore, if certain conditions are satisfied, the method is consistent with the existence of an endurance limit and Miner's rule.

KEYWORDS

cycles to failure, dissipated energy, loading and unloading curves, stabilized cycle

Highlights

- Model describing the shape of loading and unloading curves during fatigue is developed.
- A closed-form expression for the fatigue cycles to failure initiation is obtained.
- A single master curve for fatigue is created in terms of two variables.
- Comparisons with test results for four alloys show good to excellent agreement.

1 | INTRODUCTION

The concept that fatigue failure occurs when the total energy dissipated during cycling equals the available energy per unit volume, that is, the area under the stress–strain curve, has been around for almost 100 years. As early as 1927, Inglis[1] was among the first to propose this approach. He was followed by other investigators[2,3] who attempted to develop the concept further. Manson[4] and Coffin[5] related the cycles to failure to the width of the hysteresis cycle, and Halford and Morrow[6] proposed a closed-form expression for the energy dissipated per cycle applicable to low cycle fatigue. Their approach was extended to medium and high cycle fatigue by Chang et al.[7]

Examining the relationship between hysteresis and the rate at which damage accumulates during cyclic loading, Erber et al.[8] established phenomenological relations between energy dissipated during hysteresis and cycles to failure. These were linked to earlier work mentioned above but also established a framework for further experimental and simulation work. Within this framework, Tchankov and Vesselinov[9] used an incremental approach to calculate the total hysteresis energy for random applied loading and applied it to predict fatigue failure on a steel alloy.

Constantinescu et al.[10] provided an analysis that is equally applicable to low- and high-cycle fatigue by providing an interpretation of elastic shakedown in terms of dissipated energy. Charkaluk et al.[11] established a damage indicator based on the energy dissipated per cycle using a viscoplastic stress–strain law and three-dimensional finite element simulations.

An attempt to relate energy dissipated per cycle to the creation and propagation of random microcracks in the material was made by Troshchenko.[12] To further understand energy dissipation, Meneghetti[13] developed a model to relate the heat loss per cycle per unit volume to cycles to failure and applied it to a steel alloy. Korsunsky et al.[14] derived an expression for the dissipated energy per cycle under fully reversed loading for a Ramberg–Osgood-type alloy, which is very similar to the expression derived by Halford.[15] They also compared the accuracy of the energy dissipation criterion to two other methods and showed that it is more reliable.

Scott-Emuakpor et al.[16] observed that dividing the monotonic strain energy to failure by the average strain energy dissipated per cycle may not be sufficient if it is necessary to account for the experimentally observed increase of the strain energy dissipated per cycle near the end of life. They carried out experiments on an aluminum alloy and showed an exponential increase of the strain energy dissipated per cycle near final failure. Park et al.[17] studied the effect of twinning and de-twinning during low cycle fatigue loading of a magnesium alloy and found that the plastic energy dissipated per cycle correlates well with the cycles to failure.

A numerical approach to calculate the instantaneous energy dissipated per cycle under thermomechanical fatigue loading was introduced by Gosar and Nagode.[18] One of the advantages of this approach is that it can calculate energy dissipated at any point during a cycle and not after each cycle is completed as most other methods do. Liakat and Khonsari[19] used a finite element approach to calculate the entropy generated and plastic energy dissipated during tension–compression fatigue loading and compared to test results. They showed that the energy generated due to heat conduction was negligible compared to the plastic strain energy.

Fan et al.[20] examined the effects of creep and stress relaxation on fatigue life. They found a power law relation between generalized hysteresis energy and fatigue life of a 316L(N) alloy. Within the same context of fatigue creep interaction, Wang et al.[21] proposed a creep-fatigue model using a modified model based on the dissipated strain energy per unit volume and applied it to a number of steel alloys with satisfactory results. An evaluation of the dissipated energy approach for AISI 316L alloy under multiaxial fatigue was carried out by Feng et al.[22] and showed that, overall, it was superior to models based on the critical plane approach.

All previous works have demonstrated that the dissipated strain energy per unit volume, or energy density, accumulates during cycling as long as there are plastic strains present. And when the total energy accumulated equals the static total energy, the area under the stress–strain curve, fatigue failure occurs. Several issues remain such as the fact that the total plastic energy may change over time and with loading type. In general, most methods rely on a combination of experimental results and simulation to determine the energy dissipated during each cycle and are mostly applicable to low cycle fatigue where the plastic strains are most pronounced. The fact that the dissipated energy changes as a function of cycles and the difficulty to determine its stabilized or limiting value adds to the difficulty of correlating the dissipated energy density with fatigue life.

In this work, an approach to estimate the energy dissipated per cycle is proposed using some basic phenomenological properties of the loading and unloading cycles. The resulting closed-form expression for the dissipated energy is shown to have a finite limit as the number of cycles increase. This limit is divided into the total static strain energy density to obtain accurate fatigue life predictions for low, medium and high cycle fatigue.

2 | APPROACH

2.1 | Tension–tension fatigue with $0 \le R < 1$

We start with the case $\sigma_{min}/\sigma_{max} = R = 0$ because it is easier to introduce the assumptions and the basic equations. After that introduction, we will generalize to any positive R value. The case of $R = 0$ will then follow from the derivation as a special case.

Consider a material with stress–strain curve as shown in Figure 1, loaded to a stress level σ with corresponding strain ε. The failure point is defined by (ε_f, σ_f) and the total area under the curve, the dissipated energy density at failure, is denoted by U_{df}. The material is assumed to follow a Ramberg–Osgood-type of constitutive relation:

$$\varepsilon = \frac{\sigma}{E_o} + K\left(\frac{\sigma}{E_o}\right)^n \tag{1}$$

with E_o the initial slope (Young's modulus) and K and n material constants.

In addition to the static stress–strain curve to failure, the first two unloading–reloading cycles for $R = \sigma_{min}/\sigma_{max} = 0$ are also shown in Figure 1 with residual plastic strains ε_{m1} and ε_{m2}, respectively. The area enclosed in each unloading–reloading cycle is the specific energy dissipated in that cycle and will be determined first. To proceed, the following assumptions are made (see right portion of Figure 1):

1. Each loading cycle is described by a corresponding Ramberg–Osgood stress–strain law as in Equation (1) with, in general, different values for the starting slope E_i, coefficient K_i and exponent n_i. In what follows, it will be assumed that the starting slope of the loading cycle E_i is the same as E_o.
2. The loading curve AHB has the same slope as the original stress–strain curve at point B with coordinates (ε, σ).
3. The area between the straight line AB connecting (ε_{mi}, 0) with (ε, σ) and the unloading curve BIA is the same as the area between line AB and the loading curve AHB. That is, the straight line AB divides the unloading–reloading cycle in two approximately equal regions.

One more assumption is needed, and for this, we resort to Figure 2.

4. It is assumed that the areas of the delineated regions obey the relation:

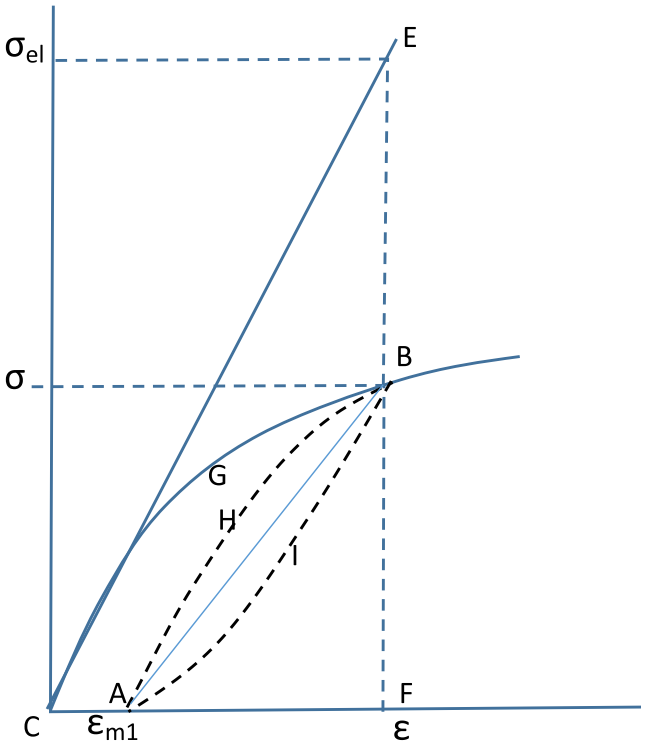

FIGURE 2 Linear versus nonlinear material behaviors. [Colour figure can be viewed at wileyonlinelibrary.com]

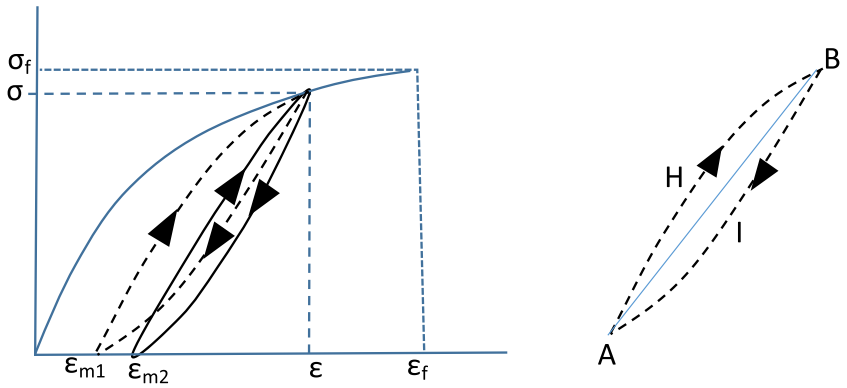

FIGURE 1 First two loading–unloading cycles in stress–strain space for $R = 0$ (left of figure) with first cycle isolated on the right for clarity. [Colour figure can be viewed at wileyonlinelibrary.com]

$$\frac{CEBGC}{CEF} \approx \frac{CGBHA}{CGBFC} \tag{2}$$

where the equality sign used is meant to imply an approximate relation. This equation is somewhat arbitrary but its implication can be understood by considering loading up to strain ε and then unloading. It relates the following quantities: CEBGC is the "unavailable" energy between a purely linear material with stress–strain curve CE and the actual inelastic system with stress–strain curve CGB. CEF is the total available energy if the material were linear. CGBHA is the "unavailable" energy between the actual nonlinear material and the loading curve CHB. Finally, CGBFC is the total energy for the nonlinear system. Then, Equation (2) simply states that the ratio of "unavailable" to available energy for a linear material equals the corresponding ratio for the actual nonlinear material. The validity of this assumption has been checked for epoxy resins[23] where the residual shear strain in a composite laminate after unloading was calculated and was found to be in good agreement with test results. The validity of this assumption for an alloy will be put to test below when fatigue life predictions are compared to test results.

The area under the static curve up to a point (ε, σ) is given by

$$U_d = \sigma\varepsilon - \int_0^\sigma \varepsilon d\sigma \tag{3}$$

with the subscript "d" referring to energy density. Using Equation (1),

$$\begin{aligned} U_d &= \sigma\varepsilon - \int_0^\sigma \left(\frac{\sigma}{E_o} + K\left(\frac{\sigma}{E_o}\right)^n\right) d\sigma \\ &= \sigma\varepsilon - \frac{\sigma^2}{2E_o} - \frac{K}{n+1}\sigma\left(\frac{\sigma}{E_o}\right)^n \end{aligned} \tag{4}$$

and using Equation (1) again,

$$U_d = \frac{\sigma^2}{2E_o} + \frac{nK}{n+1}\sigma\left(\frac{\sigma}{E_o}\right)^n \tag{5}$$

Then, it can be shown that

$$CEBGC = \frac{1}{n+1}K\sigma\left(\frac{\sigma}{E_o}\right)^n + \frac{K^2E_o}{2}\left(\frac{\sigma}{E_o}\right)^{2n} \tag{6}$$

$$CEF = \frac{\sigma^2}{2E_o} + K\sigma\left(\frac{\sigma}{E_o}\right)^n + \frac{K^2E_o}{2}\left(\frac{\sigma}{E_o}\right)^{2n} \tag{7}$$

$$CGBHA = \frac{n}{n+1}K\sigma\left(\frac{\sigma}{E_o}\right)^n - \frac{n_1}{n_1+1}K_1\sigma\left(\frac{\sigma}{E_o}\right)^{n_1} \tag{8}$$

$$CGBFC = \frac{\sigma^2}{2E_o} + \frac{n}{n+1}K\sigma\left(\frac{\sigma}{E_o}\right)^n \tag{9}$$

Equation (1) applied to the loading curve AHB gives

$$\varepsilon - \varepsilon_{m1} = \frac{\sigma}{E_o} + K_1\left(\frac{\sigma}{E_o}\right)^{n_1} \tag{10}$$

where K_1 and n_1 are material constants in the Ramberg–Osgood constitutive relation for the first loading cycle. The quantities $K\left(\frac{\sigma}{E_o}\right)^n$ and $K_1\left(\frac{\sigma}{E_o}\right)^{n_1}$ can be calculated from Equations (1) and (10), respectively, giving

$$K\left(\frac{\sigma}{E_o}\right)^n = \varepsilon - \frac{\sigma}{E_o} \tag{11}$$

and

$$K_1\left(\frac{\sigma}{E_o}\right)^{n_1} = \varepsilon - \varepsilon_{m1} - \frac{\sigma}{E_o} \tag{12}$$

These can be substituted in Equations (6)–(9), which, in turn, are substituted in Equation (2), which is rearranged to give

$$\begin{aligned} &\frac{(E_o\varepsilon - \sigma)}{E_o(n+1)}\left[\frac{(n_1+1)(2E_on\varepsilon - \sigma(n-1))(E_o(n+1)\varepsilon - \sigma(n-1))}{2{E_o}^2(n+1)\varepsilon^2}\right. \\ &\left. + n(n-n_1)\right] \end{aligned} \tag{13}$$

This equation relates the new exponent n_1 to the previous exponent n. Rearranging and solving for n_1,

$$n_1 = -1 - \frac{\frac{n(n+1)^2}{n-1}\frac{2E_o\varepsilon}{\sigma}}{(n-1)\frac{\sigma}{E_o\varepsilon} - (3n+1)} \tag{14}$$

The slope of the stress–strain curve at (ε, σ) is obtained as follows:

$$\frac{d\sigma}{d\varepsilon} = \frac{\sigma}{\frac{\sigma}{E_o} + nK\left(\frac{\sigma}{E_o}\right)^n} \tag{15}$$

with an analogous expression for the slope of AHB at (ε, σ) where n and K are replaced by n_1 and K_1. Then the equality of slopes at (ε, σ), from Assumption 2, gives

$$\frac{\sigma}{\frac{\sigma}{E_o}+nK\left(\frac{\sigma}{E_o}\right)^n}=\frac{\sigma}{\frac{\sigma}{E_o}+n_1K_1\left(\frac{\sigma}{E_o}\right)^{n_1}} \tag{16}$$

Using (11) and (12) to eliminate K and K_1 and simplifying

$$\varepsilon_{m1}=\left(\varepsilon-\frac{\sigma}{E_o}\right)\left(1-\frac{n}{n_1}\right) \tag{17}$$

with n_1 given by Equation (14). For the second unloading–reloading cycle, the equation analogous to Equation (12) reads:

$$K_2\left(\frac{\sigma}{E_o}\right)^{n_1}=\varepsilon-\varepsilon_{m2}-\frac{\sigma}{E_o} \tag{18}$$

Following the same procedure as for ε_{m1}, it can be shown that

$$\varepsilon_{m2}=\left(\varepsilon-\frac{\sigma}{E_o}\right)\left(1-\frac{n}{n_2}\right) \tag{19}$$

We now generalize to any positive R value less than 1. This is shown in Figure 3.

The first two cycles with $0 < R < 1$ are highlighted in black dashed lines. As cycling continues, the cycles will converge to a constant cycle. Two such cycles very near convergence are isolated for clarity and shown on the right of Figure 3 with ε_{mi-1} converging to ε_{mi}. The goal is to determine the area enclosed in the "converged" cycle delineated by curve A_iH_iB for the loading part and BI_iA_i for the unloading part. Consider now the straight segment $A_{i-1}B$. It divides the $i-1$ cycle in two portions, which, from Assumption 3, have equal areas. Similarly, segment A_iB divides cycle i in two equal areas. The triangle $A_{i-1}BA_i$ is made up from the lower half of cycle $i-1$, the upper half of cycle i plus a small portion $A_{i-1}MA_i$ at the bottom. But as the unloading–reloading cycles converge to $A_iH_iBI_iA_i$, A_{i-1} converges to A_i, and thus, in the limit, the area of $A_{i-1}MA_i$ is zero. Therefore, in the limit, the area of triangle $A_{i-1}BA_i$ equals the energy dissipated during the stabilized unloading–reloading cycle. That area for the first two cycles is

$$\begin{aligned}T\Delta A &= \text{Total dissipated energy in first two cycles}\\ &=\frac{(\sigma-R\sigma)}{2}(\varepsilon_{R2}-\varepsilon_{R1})\end{aligned} \tag{20}$$

Calculating the strains ε_{R1} and ε_{R2} and computing the limit of the right hand side of Equation (20) will give the sought for dissipated energy per cycle. We make one additional assumption for the complete unloading–reloading curves corresponding to $R = 0$:

5. The shape of the unloading curve for $R = 0$ is the anti-symmetric of the shape of the loading curve for the same cycle with center of (anti-)symmetry the mid-point of A_iB.

This is an approximation that eliminates the need to calculate the Ramberg–Osgood constants K and n for each cycle when $0 \le R < 1$. It also is consistent with and does not violate Assumption 3 made earlier. Considering the first unloading–reloading cycle of Figure 3, we can, therefore, write

$$\varepsilon-\varepsilon_{R1}=\varepsilon_{Q1}-\varepsilon_{m1} \tag{21}$$

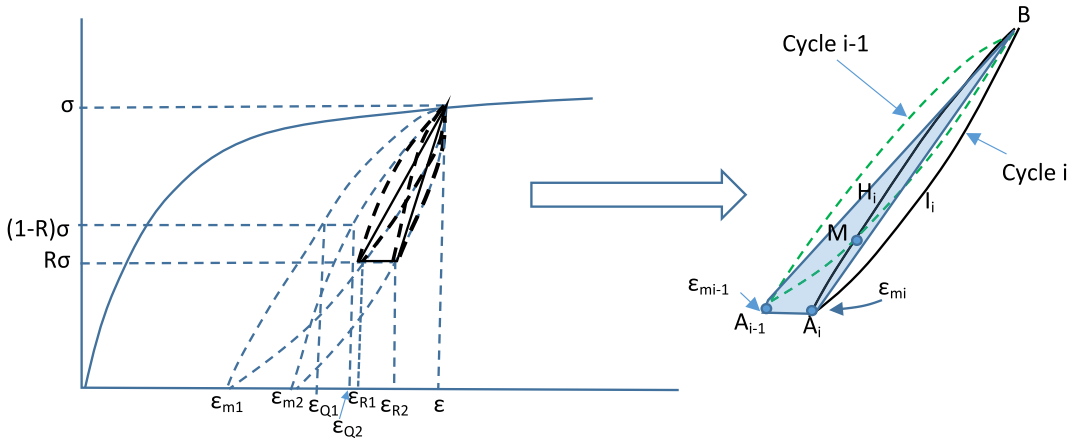

FIGURE 3 Fatigue cycles for $0 \le R < 1$. [Colour figure can be viewed at wileyonlinelibrary.com]

From the stress–strain equation, and since ε_R is on the loading part of the first unloading/reloading cycle, we can write analogous to Equation (10):

$$\varepsilon_{Q1}-\varepsilon_{m1}=\frac{(1-R)\sigma}{E_o}+K_1\left(\frac{(1-R)\sigma}{E_o}\right)^{n_1} \tag{22}$$

and using Equation (21),

$$\varepsilon-\varepsilon_{R1}=\frac{(1-R)\sigma}{E_o}+K_1\left(\frac{(1-R)\sigma}{E_o}\right)^{n_1} \tag{23}$$

In an analogous fashion, looking at Figure 3, and focusing on the second cycle,

$$\varepsilon-\varepsilon_{R2}=\varepsilon_{Q2}-\varepsilon_{m2} \tag{24}$$

Also, in analogy to Equation (22),

$$\varepsilon_{Q2}-\varepsilon_{m2}=\frac{(1-R)\sigma}{E_o}+K_2\left(\frac{(1-R)\sigma}{E_o}\right)^{n_2} \tag{25}$$

Combining (24) and (25),

$$\varepsilon-\varepsilon_{R2}=\frac{(1-R)\sigma}{E_o}+K_2\left(\frac{(1-R)\sigma}{E_o}\right)^{n_2} \tag{26}$$

Subtracting Equation (26) from (23) leads to

$$\varepsilon_{R2}-\varepsilon_{R1}=K_1\left(\frac{(1-R)\sigma}{E_o}\right)^{n_1}-K_2\left(\frac{(1-R)\sigma}{E_o}\right)^{n_2} \tag{27}$$

The first term on the right hand side of Equation (27) can be determined using Equation (12). For the second term, the equivalent of Equation (12) applied to the second cycle reads:

$$K_2\left(\frac{\sigma}{E_o}\right)^{n_2}=\varepsilon-\varepsilon_{m2}-\frac{\sigma}{E_o} \tag{28}$$

Substituting in Equation (27),

$$\varepsilon_{R2}-\varepsilon_{R1}=(1-R)^{n_1}\left(\varepsilon-\varepsilon_{m1}-\frac{\sigma}{E_o}\right)-(1-R)^{n_2}\left(\varepsilon-\varepsilon_{m2}-\frac{\sigma}{E_o}\right) \tag{29}$$

Equations (17) and (19) are used to substitute for ε_{m1} and ε_{m2} in Equation (29). Rearranging gives

$$\varepsilon_{R2}-\varepsilon_{R1}=(1-R)^{n_1}\left(\varepsilon-\frac{\sigma}{E_o}\right)\frac{n}{n_1}-(1-R)^{n_2}\left(\varepsilon-\frac{\sigma}{E_o}\right)\frac{n}{n_2} \tag{30}$$

Using this result to substitute in Equation (20) gives

$$T\Delta A=\frac{\sigma}{2}\left(\varepsilon-\frac{\sigma}{E_o}\right)(1-R)^{n_1+1}\frac{n}{n_1}\left(1-(1-R)^{n_2-n_1}\frac{n_1}{n_2}\right) \tag{31}$$

Equation (31) is valid for two successive cycles. Equation (14) can now be used to determine the ratio n_{i-1}/n_i after appropriate substitution of indices. It can be shown thatwhere

$$\frac{n_{i-1}}{n_i}=\frac{-3an_{i-1}{}^3+n_{i-1}{}^3+2an_{i-1}{}^2+an_{i-1}-2n_{i-1}{}^2+n_{i-1}}{-2a^2n_{i-1}{}^3-4a^2n_{i-1}{}^2-2a^2n_{i-1}+3an_{i-1}{}^2-2an_{i-1}-a-n_{i-1}{}^2+2n_{i-1}-1} \tag{32}$$

$$a=\frac{E_o\varepsilon}{\sigma} \tag{33}$$

Equation (32) is a recursion relation, which would allow calculation of $T\Delta A$ for any cycle i starting with the static cycle value of n. However, as will be discussed below, there is great sensitivity in the value of the hardening exponent n and very small variations in that value accumulate during recursion to huge changes in the prediction for cycles to failure. It is, therefore, preferred to work with the limiting case where the quantity n_{i-1}/n_i has converged, and the energy dissipation cycle has stabilized. It can be shown from Equation (14) that n_1 is greater than n and, as more cycles are applied, n_i keeps increasing. Thus, for a large number of cycles, Equation (32) gives

$$\lim_{n_i\to\infty}\left(\frac{n_{i-1}}{n_i}\right)=\frac{1}{2}\frac{\sigma}{E_o\varepsilon}\left(3-\frac{\sigma}{E_o\varepsilon}\right) \tag{34}$$

This allows replacing n/n_1 and n_1/n_2 in Equation (31) with the right hand side of Equation (34). In addition, for sufficiently large number of cycles, n_1 and n_2 become large and close to each other in magnitude (but not necessarily equal). It is assumed that

$$n_2-n_1\approx 0=>(1-R)^{n_2-n_1}=1 \tag{35}$$

Then, Equation (31) is rewritten:

$$T\Delta A=\frac{\sigma}{2}\left(\varepsilon-\frac{\sigma}{E_o}\right)(1-R)^{n_1+1}\frac{n}{n_1}\left(1-\frac{n_1}{n_2}\right) \tag{36}$$

which, using Equation (34) and after some manipulation, leads to the dissipated energy per cycle when it has stabilized:

$$T\Delta A=\frac{\sigma\varepsilon}{8}\frac{\sigma}{E_o\varepsilon}\left(1-\frac{\sigma}{E_o\varepsilon}\right)^2\left(2-\frac{\sigma}{E_o\varepsilon}\right)\left(3-\frac{\sigma}{E_o\varepsilon}\right)(1-R)^{n_1+1} \tag{37}$$

The cycles to failure or more precisely to crack initiation, N_f, can then be approximated by dividing the static specific energy to failure U_{df} by the energy dissipated per cycle:

$$N_f=\frac{U_{df}}{T\Delta A}=\frac{8U_{df}}{\sigma\varepsilon\frac{\sigma}{E_o\varepsilon}\left(1-\frac{\sigma}{E_o\varepsilon}\right)^2\left(2-\frac{\sigma}{E_o\varepsilon}\right)\left(3-\frac{\sigma}{E_o\varepsilon}\right)(1-R)^{n_1+1}},\ R\neq 0 \tag{38}$$

valid for $0\le R<1$. If $R=0$, Equation (38) simplifies to

$$N_f=\frac{8U_{df}}{\sigma\varepsilon\frac{\sigma}{E_o\varepsilon}\left(1-\frac{\sigma}{E_o\varepsilon}\right)^2\left(2-\frac{\sigma}{E_o\varepsilon}\right)\left(3-\frac{\sigma}{E_o\varepsilon}\right)},\ R=0 \tag{39}$$

It is interesting to note that the cycles to failure for $R\neq 0$ can be obtained from the cycles to failure for $R=0$ simply by multiplying by $1/(1-R)^{n1+1}$. Equation (39), or (38), can be rearranged to give a master curve for fatigue life predictions as a function of two non-dimensional parameters, $U_{df}/(\sigma\varepsilon)$ and $\sigma/E_o\varepsilon$. For $R=0$, Equation (39) yields

$$\frac{N_f}{\frac{8U_{df}}{\sigma\varepsilon}}=\frac{1}{\frac{\sigma}{E_o\varepsilon}\left(1-\frac{\sigma}{E_o\varepsilon}\right)^2\left(2-\frac{\sigma}{E_o\varepsilon}\right)\left(3-\frac{\sigma}{E_o\varepsilon}\right)} \tag{40}$$

Since $\sigma/E_o\varepsilon$ can only vary between 0 (no applied load) to 1 (material response is perfectly elastic), Equation (40) can be plotted in Figure 4. It can be seen from Equation (40) and Figure 4 that if there is a portion of the stress–strain curve, which is perfectly linear, then, for that portion, $\sigma=E_o\varepsilon$, and the fatigue life is infinite. The highest value of σ for which $\sigma=E_o\varepsilon$ would then be the endurance limit for this material. This transition from perfectly linear to nonlinear behavior emphasizes the importance of very accurate strain measurements during a stress–strain test as deviations from linearity that are a small fraction of a percentage point could make the difference between infinite and finite fatigue lives. This will be addressed below when physical interpretations of $\sigma/(E_o\varepsilon)$ are discussed.

Another interesting implication of Equation (40) is the existence of a minimum of the right hand side. It occurs when $\sigma/(E_o\varepsilon)=0.2711$ and leads to a minimum value of the normalized cycles to failure (left hand side of Equation (30)), which is 1.4716. Such values of $\sigma/(E_o\varepsilon)$ are, typically, deep into the plastic range and result in low cycle fatigue.

According to Equation (40), for a given material which fixes the value of U_{df} (and E_o), the cycles to failure depend on two variables: $\sigma\varepsilon$ and $\sigma/(E_o\varepsilon)$. The first is twice the energy density stored in a linear elastic material system when loaded to the point (ε, σ) or the area of triangle CBF in Figure 2. The second can be rewritten in terms of energy densities:

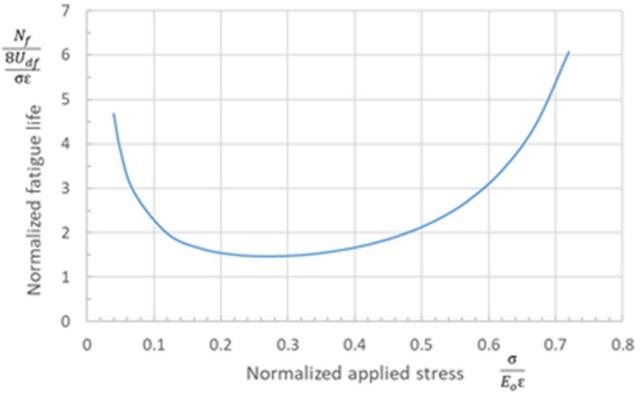

FIGURE 4 Master curve for predicting fatigue life ($R=0$). [Colour figure can be viewed at wileyonlinelibrary.com]

$$\frac{\sigma}{E_o\varepsilon}=\frac{\frac{1}{2}\sigma\varepsilon}{\frac{1}{2}E_o\varepsilon^2} \tag{41}$$

This is the ratio of the energy density of a linear material loaded to (ε, σ) to the energy density (area of triangle CEF in Figure 2) that would be stored in the actual material if it behaved linearly with Young's modulus E_o when strain ε were applied. This is related to Assumption 4 made earlier and emphasizes that the dissipated energy and, through it, fatigue life depend on ratios of available and "unavailable" energies with the word "unavailable" referring to fictitious behavior that the material might exhibit if it were linear or had no plastic residual strains. It also shows that, according to Equation (40), the life would be infinite if there were no plastic strains, that is, if $\sigma = E_o\varepsilon$. This has been postulated and observed in the past with Coffin,[5] among the first to suggest that plastic strains are necessary for finite fatigue lives and that such plastic strains imply that an endurance limit does not exist.

Another way to view this quantity, $\sigma/(E_o\varepsilon)$, is as the ratio of elastic strain σ/E_o to the total strain ε. This becomes particularly interesting when the term $(1-\sigma/(E_o\varepsilon))^2$ at the denominator of Equation (40) is considered. If the elastic strain σ/E_o is close to the total strain ε, the quantity $1-\sigma/(E_o\varepsilon)$ is close to zero, and its square becomes very small. In such situations, this term in the denominator of Equation (40) dominates, and very long lives are predicted. Because of this strong nonlinearity of the cycles to failure, very small changes in the value of ε lead to very large changes in the fatigue life. For example, if $\sigma/(E_o\varepsilon)$ equals 0.95, and the value of ε changes by about 4% to 0.99, the resulting fatigue life increases by a factor of 25. This makes very accurate knowledge of the applied total strain ε extremely important as, for high cycle fatigue, especially close to the endurance limit, differences in that value of even 1%, when $\sigma/(E_o\varepsilon)$ is close to 1, have a large effect on the fatigue life. This sensitivity is also responsible for the large fatigue scatter. On the one hand, nominally identical specimens will easily have variations in the stress–strain curve, which, near the linear region, will easily exceed 1% since the coefficient of variation for Young's modulus of most alloys is in the range 4%–6%. On the other, applying the exact strain (or stress) in an experiment can be a challenge if this kind of accuracy is to be maintained. Small variations will lead to large changes in fatigue life. Based on this discussion, one would expect the fatigue scatter to be higher at lower applied loads and this is borne out by test results.

This sensitivity to the term $(1-\sigma/(E_o\varepsilon))^2$ brings up an interesting trade-off. For materials with high plastic strain capabilities loaded with high plastic strain, this term is close to 1, but the area under the stress–strain curve U_{df} can be quite high. On the other hand, for materials which are nearly linear, this term is very close to 0, which would increase the cycles to failure, but U_{df} is very low. It is the "right" combination of area under the curve U_{df} and material nonlinearity, which will lead to the longest lives.

Some limiting cases deduced from Equation (40) are of interest. If the Ramberg–Osgood-type relation can describe the complete static stress–strain curve to failure, Equation (5) implies:

$$U_{df}=\frac{{\sigma_f}^2}{2E_o}+\frac{nK}{n+1}\sigma_f\left(\frac{\sigma_f}{E_o}\right)^n \tag{42}$$

Using Equations (42) and (11) to substitute in Equation (40),

$$N_f=\frac{8\left(\frac{{\sigma_f}^2}{2E_o}+\frac{nK}{n+1}\sigma_f\left(\frac{\sigma_f}{E_o}\right)^n\right)}{\frac{\sigma^2}{E_o\varepsilon^2}K^2\left(\frac{\sigma}{E_o}\right)^{2n}\left(2-\frac{\sigma}{E_o\varepsilon}\right)\left(3-\frac{\sigma}{E_o\varepsilon}\right)} \tag{43}$$

Now assume that the material in question has a large plastic component, that is,

$$K\left(\frac{\sigma_f}{E_o}\right)^n \gg \frac{\sigma_f}{E_o} \tag{44}$$

Then, the first term in parentheses in the numerator of the right hand side of Equation (43) can be neglected. The denominator can be approximated according to whether the applied stress and strain are low and thus close to the linear region of the stress–strain curve (high cycle fatigue) or relatively high with strain into the plastic region so the elastic strain can be neglected (low cycle fatigue). In the first case,

$$\varepsilon\approx\frac{\sigma}{E_o},\left(2-\frac{\sigma}{E_o\varepsilon}\right)\approx 1,\left(3-\frac{\sigma}{E_o\varepsilon}\right)\approx 2 \tag{45}$$

Then, Equation (43) reads:

$$N_f=\frac{8\left(\frac{nK}{n+1}\sigma_f\left(\frac{\sigma_f}{E_o}\right)^n\right)}{2E_oK^2\left(\frac{\sigma}{E_o}\right)^{2n}}=4\frac{n}{n+1}\frac{1}{K\left(\frac{\sigma_f}{E_o}\right)^{n-1}}\left(\frac{\sigma_f}{\sigma}\right)^{2n} \tag{46}$$

In the second case,

$$\varepsilon\approx K\left(\frac{\sigma}{E_o}\right)^n,\frac{\sigma}{E_o\varepsilon}\ll 2\text{ and}\ll 3 \tag{47}$$

Then, Equation (43) reads:

$$N_f = \frac{8\left(\frac{nK}{n+1}\sigma_f\left(\frac{\sigma_f}{E_o}\right)^n\right)}{6\frac{\sigma^2}{E_o}} = \frac{4}{3}\frac{n}{n+1}\left(\frac{\sigma_f}{E_o}\right)^{n-1}\left(\frac{\sigma_f}{\sigma}\right)^2 \quad (48)$$

If the material stress–strain curve to failure is nearly linear, instead of Equation (44), we have

$$K\left(\frac{\sigma_f}{E_o}\right)^n \ll \frac{\sigma_f}{E_o} \quad (49)$$

Also, because the material is close to linear elastic,

$$\varepsilon \approx \frac{\sigma}{E_o}, \left(2-\frac{\sigma}{E_o\varepsilon}\right)\approx 1, \left(3-\frac{\sigma}{E_o\varepsilon}\right)\approx 2 \quad (50)$$

Substituting then into Equation (43) gives

$$N_f = \frac{8\left(\frac{\sigma_f{}^2}{2E_o}\right)}{2\frac{\sigma^2}{E_o\varepsilon^2}K^2\left(\frac{\sigma}{E_o}\right)^{2n}} = \frac{2}{K^2\left(\frac{\sigma_f}{E_o}\right)^{2n-2}}\left(\frac{\sigma_f}{\sigma}\right)^{2n} \quad (51)$$

Equations (46), (48), and (51) all have the form:

$$N_f = \frac{B}{\sigma^\beta} \quad (52)$$

with B and β constants. This is the classic form for S–N curves first proposed by Basquin[24] where the cycles to failure are assumed to have a power law relationship with the applied maximum stress. The difference here is that, at least for the three limiting cases discussed, the constants B and β are derived from the general formulation and can be determined from material constants without the need of fitting experimental fatigue data. What is also interesting is that these constants are not necessarily the same for low and high cycle fatigue suggesting that a single curve of the form of Equation (52) might not be sufficient.

Finally, it is easy to show that Miner's rule follows naturally from the formulation. If different stress levels are applied for different numbers of cycles, each would dissipate a certain amount of energy $m_i U_{df}/(T\Delta A)_i$, following Equation (38), with m_i the number of cycles with per cycle dissipated energy $(T\Delta A)_i$. If these amounts add up to U_{df}, there is failure. Implicit in this is that the stress–strain curve and, more importantly, the amount of energy dissipated per cycle do not change or any changes have negligible effect on the dissipated energy in the stabilized cycle given by Equation (37). It is recognized here that the equations derived so far, are for stabilized fatigue cycles and neglect the cycles and energy dissipated up to stabilization. Tests have shown that Miner's rule does not always hold and improving the energy dissipation calculation to account for changes during fatigue loading would be a needed refinement.

It should also be noted that the approach can be extended relatively easily to apply to cases with $R < 0$ provided compressive stress–strain curves of sufficient accuracy are available.

3 | COMPARISON WITH TEST RESULTS

As was already discussed, the proposed method requires non-zero plastic strains. As such, it would not work in the linear elastic region of a material unless it is assumed that no material has a perfectly elastic region and there are tiny deviations from the linear behavior, which may not be easily measurable but may be sufficient to create plastic strains. This means that very accurate measurements of the complete stress–strain curves of a given material would be needed. It was already mentioned that small errors in strain measurements lead to large changes in predicted fatigue lives. Thus, the accuracy of extensometers and strain gages, even the thickness of lines used to plot stress–strain curves will have a significant effect on the predictions. As a result, an attempt was made to create a consistent approach to translate published stress–strain curves to data usable in the present approach.

First, the 0.2% strain offset method is used to determine a yield stress σ_y and a yield strain ε_y by drawing a line with slope equal to the initial modulus E_o of a given stress–strain curve and going through strain 0.02% when the stress is zero. Then, a stress–strain curve with initial slope E_o that goes through the $(\varepsilon_y, \sigma_y)$ point is constructed with the form:

$$\sigma = E_o\varepsilon + A\varepsilon^\alpha \quad (53)$$

The advantage of this equation is that it acts as complementary to the Ramberg–Osgood relation, which has the strain on the left hand side. It is recognized here that there are published values for the Ramberg–Osgood constants K and n; however, these are meant to go deep into the plastic regime and may not be as accurate in the nearly linear region where the present method is most sensitive.

Denoting the local slope of a given stress–strain curve at $(\varepsilon_y, \sigma_y)$ by S_y, it can be shown that Equation (53) goes through the point $(\varepsilon_y, \sigma_y)$ and has local slope S_y if

$$\alpha = \frac{(S_y - E_o)\varepsilon_y}{\sigma_y - E_o\varepsilon_y} \quad (54)$$

$$A = \frac{\sigma_y - E_o \varepsilon_y}{{\varepsilon_y}^{\alpha}} \tag{55}$$

It should be noted that A from Equation (55) is always negative. Then, for a number of strain values ε_i (at least four) between 0 and ε_y, the corresponding stress values, σ_i, can be obtained from Equation (53), and Equation (1) is best fitted through the pairs (ε_i, σ_i) to obtain K and n for strains in this range. An example of this is shown in Figure 5 for 6061-T6 aluminum with the stress–strain curve taken from fig. 3.6.2.2.6(m) in literature.[25] The portion with low applied strains up to ε_y is also shown (zoomed in) on the right of Figure 5 showing that the difference between Equation (53) and experiment is very small and only in the vicinity of ε_y. If the applied strain is greater than ε_y, it can be read directly off the stress–strain plot for the material and used in Equations (38) and (39) to predict fatigue life. As this strain is significantly greater than the elastic strain, any errors in its exact value, which would be an issue for lower applied strains, lead to small changes in the predicted fatigue lives.

Predictions of Equations (38) and (39) were generated for four alloys, 7075-T6 aluminum, 6061-T6 aluminum, Ti 6-4 titanium, and 4340 steel, and were compared to test results from literature.[25] Stress–strain curves given in literature[25] were used to obtain E_o, σ_y, and ε_y. U_{df} was obtained by numerical integration of these stress–strain curves. K and n were obtained using the procedure described above via Equations (53) to (55). The values of these variables are summarized in Table 1. Note that, for $R \neq 0$, n_1 in Equation (38) was obtained from Equation (32) using n from Table 1.

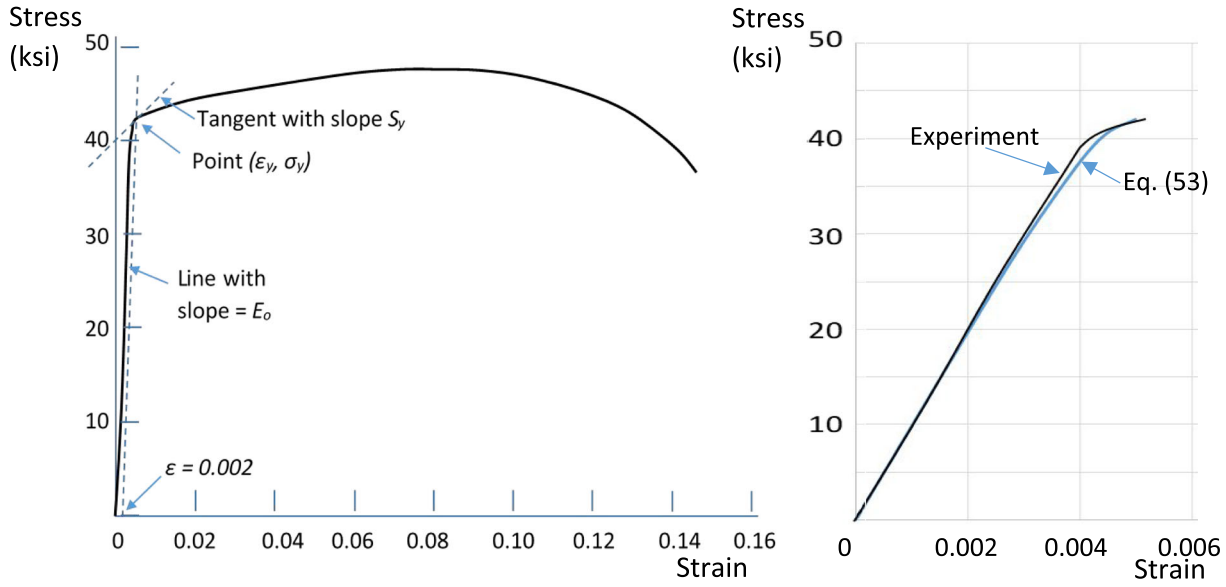

FIGURE 5 Complete stress–strain curve for 6061-T62 aluminum (left) and elastic portion (right). [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Material data used in fatigue life predictions.

Property/material	7075-T6 Al	6061-T6 Al	Ti 6-4	4340 steel
E_o (GPa)	71.60	67.80	116.5	201.1
U_{df} (MPa)	65.01	45.48	100.05	193.7
ε_y	0.0097	0.0052	0.01025	0.0065
σ_y (MPa)	579.1	294.4	947.9	1206
S_y (GPa)	4.387	4.136	6.204	147.7
A (Pa)	-2.716×10^{19}	-5.719×10^{20}	-3.338×10^{17}	-3.512×10^{15}
α	5.648	5.688	4.59	3.449
K	1.564×10^{9}	5.470×10^{12}	3.414×10^{7}	9.705×10^{4}
n	5.90	5.76	5.03	3.74

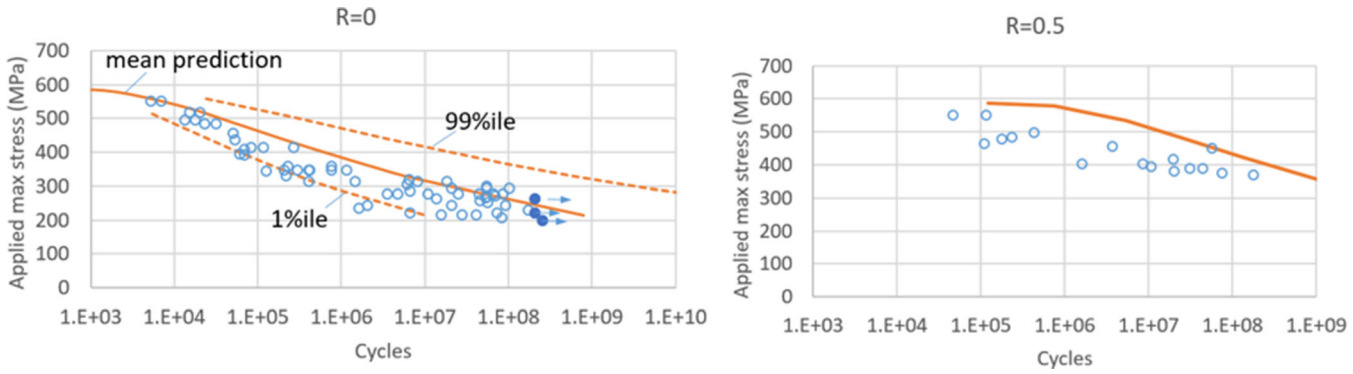

FIGURE 6 Predictions compared to test results (tests denoted by circles) for 7075-T6, $R = 0$ and $R = 0.5$. [Colour figure can be viewed at wileyonlinelibrary.com]

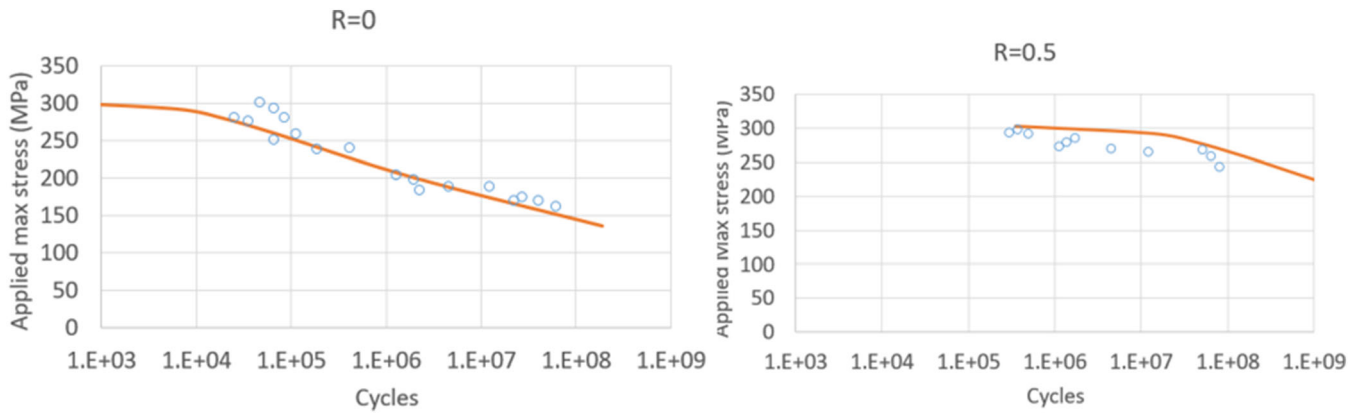

FIGURE 7 Predictions compared to test results (tests denoted by circles) for 6061-T6, $R = 0$ and $R = 0.5$. [Colour figure can be viewed at wileyonlinelibrary.com]

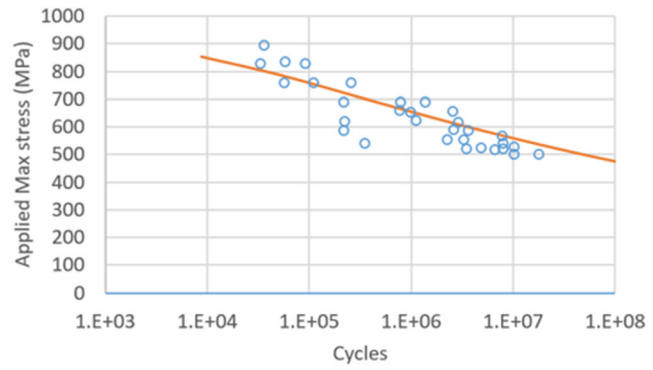

FIGURE 8 Predictions compared to test results (tests denoted by circles) for titanium 6-4, for $R = 0.01$. [Colour figure can be viewed at wileyonlinelibrary.com]

The predictions by Equations (38) and (39) are compared to test results from literature[25] for 7075-T6 aluminum for $R = 0$ and $R = 0.5$ in Figure 6. For the $R = 0$ case, the test data include enough points to give an indication of the scatter. For this case, the predictions of the present method for the 1 and 99 percentile curves are also shown. These are obtained by treating ε_y as a normally distributed variable and using statistical data from literature,[25] namely, the A-Basis and B-Basis values, to backcalculate a coefficient of variation. The mean curve slightly overpredicts and the predicted scatter band contains most of the data. For $R = 0.5$, the predictions are higher than test results by at most a decade on the cycles axis. The predictions for 6061-T6 aluminum for $R = 0$ and $R = 0.5$ are shown in Figure 7. Very good agreement with test results is observed for $R = 0$. For $R = 0.5$, the trend is captured very well but the predictions are higher than the tests again by at most a decade on the cycles axis. The predictions for titanium 6-4 for $R = 0.01$ are shown in Figure 8. There is very good agreement with the test results. Finally, The predictions for 4340 steel for $R = 0.2$ and $R = 0.54$ are shown in Figure 9. In both cases, the predictions are in good agreement with test results.

Overall, the predictions follow the trends from the experiments very well for all four alloys and for $R = 0$ are in very good agreement with test results. For the case

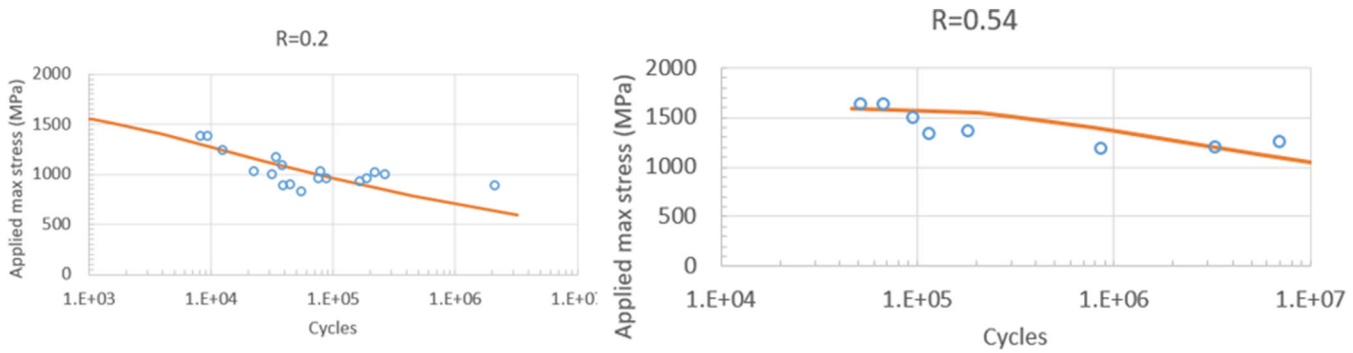

FIGURE 9 Predictions compared to tests (tests denoted by circles) for steel 4340 for $R = 0.2$ and $R = 0.54$. [Colour figure can be viewed at wileyonlinelibrary.com]

where there were sufficient test data, the scatter band predicted by the present method also is in reasonable agreement with the scatter observed in the test results.

4 | SUMMARY AND CONCLUSIONS

A method was developed to determine the energy density dissipated during each fatigue cycle in a metal alloy and was used to predict fatigue life up to damage initiation for cases with $0 \leq R \leq 1$. The resulting expressions for cycles to failure are in closed form and make no use of parameters fitted to experimental results. Only accurate static stress–strain data are needed. The method was shown to degenerate to standard power law equations with the coefficient and exponents in those equations determined directly from the complete static stress–strain curve. It was shown that nonlinear strains must be present for finite fatigue lives, and if a portion of the stress–strain curve is perfectly linear, the fatigue life is infinite for that portion suggesting ways to determine an endurance limit (if the material has one). Comparison of predicted fatigue lives with test results for two aluminum, one titanium, and one steel alloy showed good to very good agreement with tests.

The proposed method is very promising, but it relies on a number of assumptions the validity of which needs to be further examined. (1) Almost the entire stress–strain curve must be nonlinear. For low applied strains, there is great sensitivity to the value of the quantity $\sigma/(E_o\varepsilon)$ and very small errors in that value lead to great changes in the predictions. This on the one hand emphasizes the need to measuring strains to an accuracy better than $\pm 50\,\mu s$ and on the other points to an inherent source of fatigue scatter: Nominally identical specimens will have very different fatigue lives if they have small differences in this value. (2) The formulation is based on a Ramberg–Osgood-type material. As such, the stress–strain curve must be monotonically increasing. This means that materials with significant necking would not be captured by the model if fatigue lives for applied strains past the maximum of the stress–strain curve are needed. All four alloys examined here exhibit some amount of necking. For this reason, the predictions were kept to strains below the point where necking starts and the results are still accurate. (3) It is assumed that the static stress–strain curve does not change with fatigue cycling. This is an approximation as localized extended plasticity may cause load redistribution not captured by a single stress–strain curve meant to represent an entire specimen throughout its fatigue life.

When applied to load ratios $R \neq 0$, the method has the tendency to overpredict test results. This requires further investigation. A single "master" curve for predicting cycles to failure can be constructed and shows that, within the framework of the assumption made, two parameters, $\sigma/(E_o\varepsilon)$ and $U_{df}/(\sigma\varepsilon)$, are sufficient to describe the behavior.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in Department of Defense Handbook, Metallic Materials and Elem at https://automaterials.files.wordpress.com/2018/09/military-handbooks-a-collection.pdf. These data were derived from the following resources available in the public domain: Department of Defense Handbook, Metallic Materials and Elem, https://automaterials.files.wordpress.com/2018/09/military-handbooks-a-collection.pdf.

REFERENCES

1. Inglis NP. 'Hysteresis and fatigue of Wohler rotating cantilever specimen. *Metallurgist*. 1927;1:23-27.
2. Hanstock RF. Damping capacity, strain hardening and fatigue. *Proc Phys Soc*. 1947;59(2):275-287.

3. Feltner CE, Morrow JD. 'Microplastics strain hysteresis energy as a criterion for fatigue fracture. *J Basic Eng*. 1961;83(1):15-22.
4. Manson SS. Behavior of materials under conditions of thermal stress. *NACA TN*. 1953;2933:1.
5. Coffin LF. A study of the effects of cyclic thermal stresses on a ductile metal. *Trans ASME*. 1954;76(6):931-949.
6. Halford GR, Morrow JD. 'Low-cycle fatigue in torsion. *Proc ASTM*. 1962;62:695-709.
7. Chang CS, Pimbley WT, Conway HD. An analysis of metal fatigue based on hysteresis energy. *Exp Mech*. 1968;8(3): 133-137.
8. Erber T, Guralnick SA, Michels SC. Hysteresis and fatigue. *Ann Phys*. 1993;224(2):157-192.
9. Tchankov DS, Vesselinov KV. Fatigue life prediction under random loading using Total hysteresis energy. *Int J Pres Vessels*. 1998;75(13):955-960.
10. Constantinescu A, Dang Van K, Maitournam H. A unified approach for high and low cycle fatigue based on shakedown concept. *Fatigue Fract Eng Mater Struct*. 2003;26(6):561-568.
11. Charkaluk E, Bignonnet A, Constantinescu A, Dang Van K. Fatigue design of structures under thermomechanical loadings. *Fatigue Fract Eng Mater Struct*. 2003;25(12):1199-1206.
12. Troshchenko VT. Nonlocalized fatigue damage of metals and alloys. Part 1. Inelasticity, investigation of methods and results. *Strength Mater*. 2005;37(4):337-356.
13. Meneghetti G. Analysis of the fatigue strength of a stainless steel based on the energy dissipation. *Int J Fatigue*. 2007;29(1): 81-94.
14. Korsunsky AM, Dini D, Dunne FPE, Walsh MJ. Comparative assessment of dissipated energy and other fatigue criteria. *Int J Fatigue*. 2007;29(9-11):1990-1995.
15. Halford GR. The energy required for fatigue. *J Mater ASTM*. 1966;1:3-18.
16. Scott-Emuakpor O, George T, Cross C, Herman Shen MH. Hysteresis-loop representation for strain energy calculation and fatigue assessment. *J Strain Anal Eng Des*. 2010;45(4):275-282.
17. Park SH, Hong SG, Lee BH, Bang W, Lee CS. Low-cycle fatigue characteristics of rolled Mg-3Al-1Zn alloy. *Int J Fatigue*. 2010; 32(11):1835-1842.
18. Gosar A, Nagode M. Energy dissipation under thermomechanical fatigue loading. *Int J Fatigue*. 2012;43:160-167.
19. Liakat M, Khonsari MM. Rapid estimation of fatigue entropy and toughness in metals. *Mater Des*. 2014;62:149-157.
20. Fan Y-N, Shi H-J, Tokuda K. A generalized hysteresis energy method for fatigue and creep-fatigue life prediction of 316L(N). *Mater Sci Eng A Struct Mater*. 2015;625:205-212.
21. Wang RZ, Zhang XC, Tu ST, Zhu SP, Zhang CC. A modified strain energy density exhaustion model for creep-fatigue life prediction. *Int J Fatigue*. 2016;90:12-22.
22. Feng E, Wang X, Jiang C. Multiaxial fatigue evaluation of type 316L stainless steel based on critical plane and energy dissipation. *Fatigue Fract Eng Mater Struct*. 2022;45(12):3486-3499.
23. Socci CA, Kassapoglou C. Prediction of matrix crack initiation and evolution and their effect on the stiffness of laminates with off-axis plies under in-plane loading. *Compos Sci Technol*. 2020; 200:108427.
24. Basquin OH. The exponential law of endurance test. *Proc ASTM*. 1910;10:625-630.
25. Department of Defense Handbook, "Metallic materials and elements for aerospace structures", MIL-Hdbk-5H, Knovel electronic edition, 2003.

How to cite this article: Kassapoglou C. Determination of energy dissipation during cyclic loading and its use to predict fatigue life of metal alloys. *Fatigue Fract Eng Mater Struct*. 2023;1-13. doi:10.1111/ffe.14096