Health prognosis based on a novel approach for damage accumulation calculation

Kai-Uwe Dettmann [1], **and Dirk Söffker** [2]
University of Duisburg-Essen, Chair of Dynamics and Control (SRS), Duisburg, NRW, 47057, Germany
[1] *Research Associate (Ph.D. student)*
kai.dettmann@uni-due.de
[2] *Professor (Head of the Chair)*
soeffker@uni-due.de

ABSTRACT

Core aspects of diagnosis and prognosis are based on the knowledge of the actual state-of-damage. Every mechanical system damage increases due to applied stresses. This contribution focuses on systems being affected by mechanical loads, leading to failure if a certain damage level is exceeded. According to the literature, mathematical models are known that describe qualitatively the damage progression based on experimental data. Hence, those models are valid for certain systems under certain operating conditions and depend on the underlying experimental data.

The intention of this contribution is to calculate with a general model the damage progression for different load histories, independent from specific load collective-based experiments. One novel aspect is to conclude from failures of individuals of a considered set of systems used with individual load profiles to the underlying, problem/application-specific damage accumulation relation. This is done by a nonlinear mathematical model, calculating the caused damage due to the applied stress. Costly and/or safety relevant systems are not subject of this investigation but mechanical (e.g., friction, wear, etc.) systems.

1 INTRODUCTION

The central aspects of diagnosis and prognosis concerning the usage of technical systems are based on the knowledge of the actual state-of-damage especially of critical components and/or systems. The supervision of those seems to be useful to ensure the fault free operation. A safe and reliable operation can be disturbed by various hazards as well as by the usage itself. This contribution focuses on systems being stressed/overstressed by mechanical loads. The proposed concept uses a mathematical model to calculate the damage caused by mechanical stresses.

The main idea of using an experimental-based model approach for damage calculation is discussed in the following. A brief introduction to known damage accumulation ideas is given. Hence, the new approach is explained and the results of calculations are shown. This contribution closes with a discussion of the results and an outlook to future work.

2 DAMAGE ACCUMULATION

The calculation of hazard rates and remaining lifetime is usually realized using assumptions of underlying damage accumulation laws. These mathematical relations use the experimental-based knowledge about the S-N-curves (also known as Wöhler curve, (Wöhler, 1870)) of specimens with applied specific loads or load collectives. The S-N-curve describes the relation between the magnitude of a cyclical stress (S) against the cycles to failure (N).

In general, tests with a large sample size of specimen have to be done to examine the lifetime of a component experimentally. But even though (identical operating conditions, etc.) the collective of comparable systems fail at different points in time. This is due to stochastic reasons like material defects, etc. and can be considered by non-deterministic approaches. The well known mathematical descriptions of underlying damage accumulation law e.g., (Palmgren, 1924), and (Miner, 1945) neglect such effects. Expanding approaches e.g., (Henry, 1955), (Marco and Starkey, 1954), and (Hwang and Han, 1986) consider some side effects by adapting the laws to specific problems e.g., experimentally determined material properties, shape of specimen, operating temperature, ratio between creepage and fatigue, etc. Other approaches consider more complex (e.g., nonlinear) material behaviors, maximal tolerable strain borders, etc. Nearly all classical linear approaches calculate the damage based on experimental results (S-N-curve). Those curves were obtained by stressing comparable mechanical systems (uniaxial tensile specimen) under comparable operating conditions. The results of the realization of this idea are only valid for the tested specimen under the described operating conditions. The transferability of those results to real applications (e.g., multi-axial stress, different specimen shape, different operating temperature and so on) is not always given. Even a large number of tests lead to an ambiguous stress vs. load-cycle relation due to stochastic phenomenons and material behavior. The idea itself, operating a set of comparable systems under defined conditions until failure, is indeed transferable to all types of systems. This is shown in this contribution.

In literature the scattering stress values sometimes are neglected and the S-N-curve is handled as a deterministic information, which does not consider the variance of maximal tolerable stress at a certain number of cycles to failure. There are other approaches to modify the S-N-line e.g., for worst- or best-case scenarios see (Wolters, 2008) for details.

The majority of the damage accumulation hypotheses known from literature assume that all stresses above a certain level somehow negatively influence the system damage, independent of the actual system's state-of-damage, load history, and point in time when the stress is applied. Under this restrictions only a vague information about the accumulated damage/remaining lifetime can be stated. Furthermore, the structure and parameters of the used mathematical models are static. This means that the information gathered under real operating conditions, like observed stochastical failures, etc. are not used to adapt/modify the formerly assumed hypotheses. Hence, the statements about the actual state-of-damage do not improve over the system usage. The term "system usage" will be introduced later. For purpose of performance optimization, the feedback of the failure information is necessary.

To overcome some of these drawbacks a new approach is developed in the following. The collective of identical systems will not be operated under identical operating conditions but arbitrary and known ones. Real systems are operated under real varying conditions until they fail at different points in time. Unequal operating conditions (as they appear in reality), the load history and point in time when the stress is applied will be considered and fed back to the damage calculation model. Assuming identical damaging effects as well as identical accumulation the accuracy of the calculation improves over the number of failed systems.

As explained in literature the applied stress mainly influence the damage propagation. Hence, the stress profile, described by peak amplitude, stress over time, etc. contains the information how grave the incremental damage is. Damage due to rusting, weathering, and so on is not considered here. Therefore, the applied stresses over time are measured and analyzed for their incremental damaging content.

Another new aspect considered in the following is that each system might not start working with initial damage level equal to zero. This is another main difference between the classical approaches and the proposed one and widens the practicality of the proposed method.

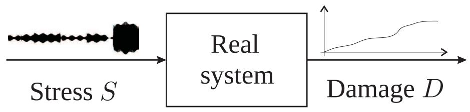

Figure 1: Schematic input/output relation of a system; damage progression due to an affecting stress profile

2.1 Introducing a general mathematical damage model

In the following the core idea of the contribution is illustrated by means of a mechanical system. In Figure 1 the general connection between stress and damage is shown in principle.

To gain a high reproducibility without simplifying the problem significantly, the damaging behavior of the real systems are determined in a first step by simulation with known parameters. Once the applicability of this approach is shown, the method has to be validated in a second step with a real system. Based on the simulated damage accumulation model, the damage progression can be approximated for given (measured) stress signals. The different synthetic stress profiles used in this contribution are based on experimentally determined stress profiles and will be denoted with S^1, S^2, and so on. This means that real stresses are considered within the mathematical approach. Therefore realistic assumptions about the stress are made and the novel model will adapt to e.g., specific operating conditions. In this context one further connection is introduced. The load profile S^1 leads to the damage progression D^1 of the system unambiguously. Accordingly S^i leads to D^i.

In reality the stress profile $S(t = 0...t_e)$ and the point of failure $D_e(t_e)$ are known but the damage propagation from one point to another is not. Even more, the unambiguous assignment of one damage level D to one point in time is not possible by measurement. Additionally when the system starts working at t_0, an initial damage D_0^i might be already present. This fact will be detailed below.

As introduced and shown above, the stress signal and its history is the only input to the system. The direct consequence of this is that the only effect causing the damage is the stress. Hence, the stress bears the damage content.

A system is defined as broken down, once the degradation D exceeds the maximum tolerable damage $D^* = D_e^i$ and assumed to be 1. This occurs at different points in time as the systems underly different time-dependent operating conditions (stress profiles, etc.). As depicted in Figure 2 for three different systems, points in time t_e^i of failure are known. The damage shortly before those points can be approximated. The more it is propagated in the past, the more uncertain the calculation becomes. This is denoted by the disappearing and broadening damage curve runs.

For this consideration shown here it is necessary to eliminate the time dependency by introducing the meta parameter X that denotes that system specific quantities can be used (e.g., driven kilometers, time of usage, useful life, and others). Identical (relating to their

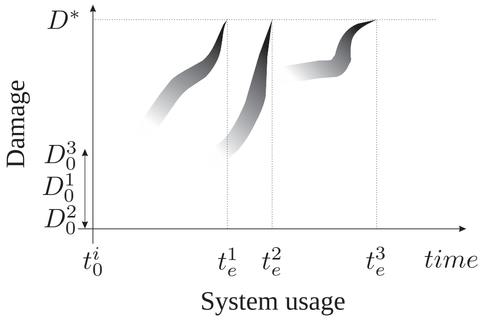

Figure 2: Probable damage progression of systems shortly before failure; gray areas denote the uncertainty of the actual state-of-damage leading to the point of failure D^* at different points in time

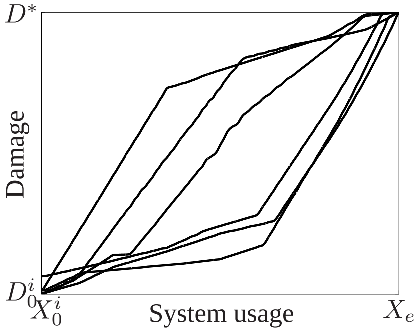

Figure 3: Different damage progressions due to different stress profiles S^i, normalized to unified representation

damaging behavior) systems fail at different points in time as shown in Figure 2.

By decoupling the damage progression for the considered systems from the time, a consistent graphical representation can be depicted. This is shown in Figure 3.

The time is not just substituted but normalized to 100% of possible system usage X. Furthermore it is defined that the system fails once the system usage reaches/exceeds X_e. Hence, all systems have the condition $D^i = D^*$ at $x_e^i = X_e$.

The absolute number X_e (maximum kilometers/lifetime/etc.) is prior not known. This is not a drawback but necessary to adapt the proposed method to arbitrary prevailing systems and conditions. The beginning of the system usage $x_0^i = X_0$ is identical for all i systems whether the initial damage D_0^i is not.

In the following the procedure of concluding from arbitrary stress profiles to each resulting damage progression is developed.

The stress, acting on a system, is measured and processed by several methods. In accordance to the classical damage accumulation approaches, this is done by calculating the incremental damage caused by the stress increment. The size of the increments Δx is problem/system specific.

A stress applied in the beginning of the system usage may damage the system in a different way than the same stress applied in the middle or end of the system usage. The effect of the incremental damage is considered to be depended on the point x the stress is applied. This is done by defining a weighting function that modifies the influence of the incremental damage.

Additionally the equivalent stress is modified by a second consideration. The influence of the height and duration of a stress onto the incremental damage need not necessarily be a linear one. This means that the doubling of the stress amplitude might lead to a triplication of the incremental damage and vice versa. Hence, huge stress amplitudes acting for a long time damage the system disproportionately high, while small short-term stresses can be neglected as they do not influence the overall system damage significantly. So a nonlinear connection between stress and damage is considered by a second weighting function. In the consequence the applied structure of the approach is much more general than classical approaches and includes known ones.

Once the equivalent weighted stress is calculated, the curve run is used to derive the damage progression. As mentioned above the damage is stress-inherent. A model for generating in monotonic increasing curve run is used. This is realized by integrating the absolute value of the stress-equivalent curve run over the whole system usage. The geometrical interpretation of the calculation result is the area beneath the curve. This number bears the meaning of tolerated stress over system usage. After the failure of several systems and applying the above described procedure, different areas (maximum values) are obtained. The value of those areas is subsequently used for normalization purpose. As introduced this area is assumed to be directly corresponding to the available system life. The result is a normalized monotonic increasing curve run, directly resulting from the applied stress.

A sketch of the proposed normalization process is depicted in Figure 4.

The assumption made here is that this curve represents the damage progression of the system and is used as a starting hypothesis. In accordance with the known literature the maximum tolerable damage D^* is defined as a constant value. In reality this value scatters. In contrast to the literature, this fact of uncertainty of this threshold is considered here and explained in the following.

As expected and introduced, systems fail differently even if they are stressed under comparable operating conditions. The main point of this assumption is that all systems fail at D^* at a certain point in system usage ($x_e^i = X_e$). That means that the individual tolerable damage slightly differs and the systems are already damaged as starting their operation at x_0^i. Hence, D_0^i at the start point in system usage x_0^i is greater than zero c.f., Figure 2. By assuming these varying individual initial conditions D_0^i, the scattering D^* has to be converted into a constant value and represents the maximum life for the whole set of systems.

This is done as follows. The system with the maximum tolerated stress is assumed to be the healthiest one. The individual initial damage D_0^i of all other al-

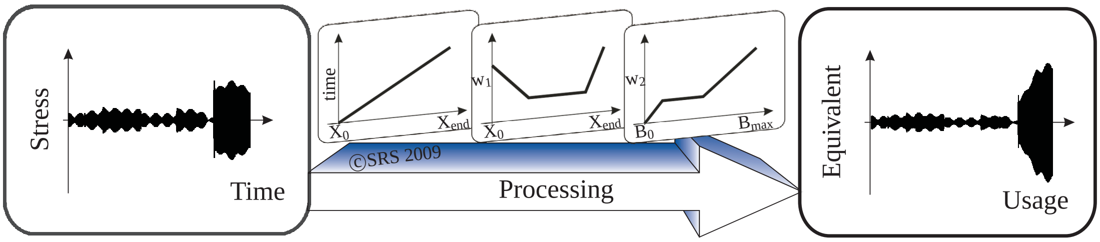

Figure 4: Calculation of unified stress profile for standardization purpose

ready failed systems can therefore be calculated. Only the healthiest system with the dominating stress tolerance starts at $D_0 = 0$. The equivalent damage curve is normalized to the constant value D^* by the value of the area. Here again the iterative and adaptive aspect of the proposed algorithm is shown. Additionally it has to be noted that this method can only calculate the initial damage a posteriori, once the data base of failed systems is sufficient. That means that enough systems must fail with different initial damages, before the method can predict the damage progression reliable. The idea of damage normalization to a deterministic value for D^* is consists of two main aspects. On the one hand the initial damage is considered and not set to zero like in the classical approaches. On the other hand the systems fail at the same point in system usage in analogy to the classical approaches. This has some positive effects on the following mathematical considerations.

The next step is the calculation of the mathematical connection between stress and the derived damage progression. Here, the stress is oscillating with constant frequency and amplitude. A constant monotonic increasing damage progression is obtained (which is valid for non-biological systems) and approximated by a spline. Therefore the system usage-damage-curve is known and the state-of-damage can be calculated. This is done by a NonLinear AutoRegressive model with eXogenous input (nlarx), which is an one layer neural network with an sigmoid activation function.

Once the model is parametrized/trained, it can be used for first damage progression prognosis. By observing other systems, the prognosis of the damage propagation to small usage horizons can be done.

2.2 Damage calculation using a nlarx-model

In general, all mathematical models have to be structured, parameterized, and adapted to the special given problem (here: material, stress profile, and so on). Due to the fact that the problem constraints are not sufficiently known this task can not be solved by look-up tables for material behavior or others. Therefore, a parametric model is assumed which is typically described by difference or differential equations. Hence, a classical parameter identification task is obtained which can be solved with different mathematical approaches. This contribution focuses on time-domain parametric models and methods, since the experimental data are obtained in time-domain (Nelles, 2001).

A one-layer time-delay-neural-network in nonlinear autoregressive structure for single input/output data is

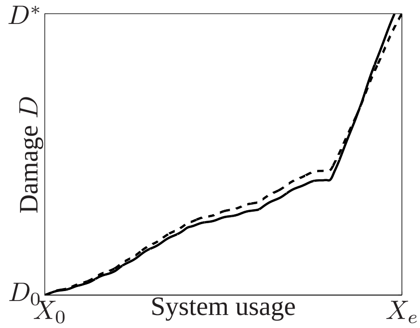

Figure 5: Dashed line: new approach; continuous line: linear approach

chosen for the calculation of an unknown input-output-relation

$$\hat{y}_k = f(y_{k-1}, \ldots, y_{k-na}, u_{k-1}, \ldots, u_{k-nb}). \quad (1)$$

The output $\hat{y}_k$ is calculated in two steps: first, the input and output signals are delayed to different degrees. Second a nonlinear activation function $f(\cdot)$ (here a static neural network) estimates the output. In (Nelles, 2001) a sigmoid function is proposed for the nonlinear activation function, which is used in this context. Other functions for nonlinear dynamic modeling e.g., Hammerstein models, Wiener models, neural or wavelet network are also possible.

This general nlarx model is used to calculate the damage progression by identification.

To check whether the proposed idea meets the classical approaches, like (Palmgren, 1924) and (Miner, 1945), Figure 5 depicts two curve runs. The dashed line shows the damage progression, calculated by the considered new approach. The continuous curve is derived from a classical approach which uses the continuous and classified stresses.

As shown both curve runs fit accurately. Although the proposed approach is not based on experimental results (e.g., physical effects/observations) it meets the classical one.

The proposed idea is tested by simulation with four different systems which are operated under different conditions. The content (amplitude, frequency) of the stress signal for all system is the same but the sequence and overlapping of those oscillation is not. That means that all systems are stressed with the same stress energy resulting in a characteristic damage progression.

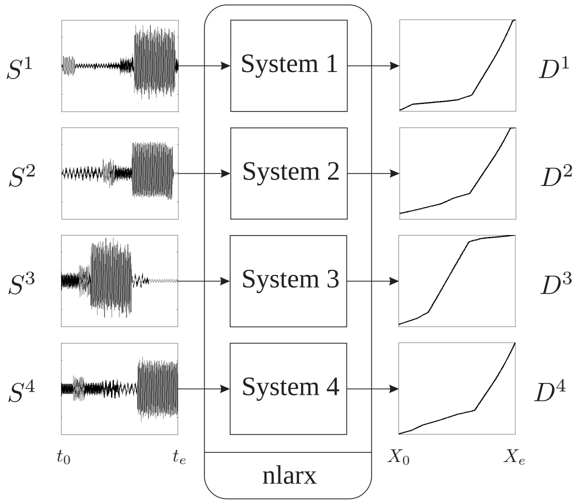

Figure 6: Several different stress profiles acting on identical systems leading to different damage propagation are approximized by one mathematiclal damage accumulation model

As demonstrated in Figure 6, different damage progression are the result although the overall input energy is the same.

The mathematical model has to fit all stress profiles with an already trained neural net. The result is that the model can calculate the incremental damage for a given interval. As known, the prediction of nonlinearities strongly depend on the chosen initial conditions. Here this means that accumulated damage strongly depends on the chosen boundaries (initial damage D_0^i) which is discussed in the following.

2.3 Results

The individual system life is continuously reduced by the applied stress until the complete amount is depleted. Exactly at that moment, it exceeds the overall system damage the maximum tolerable limit D^* and fails. Hence, the result of the calculation depends on the assumption about the initial damage D_0^i. After making this assumption, the average damage progression of the system can be calculated for each increment Δx. The derived model cannot be used to calculate the damage at an arbitrary point x, if the stress history is unknown. Therefore the continuous measurement of the damaging input signal has to be guaranteed.

The application of the proposed method is shown in the following. One system is considered to start working at X_0^1 and is stressed with the stress profile S^1. The damage prediction for the system can be made based on the nlarx-model and the intended stress. The initial damages D_0^1 is unknown until the system fails. For realizing a reliable prediction about the remaining possible system usage of the other systems, an appropriate assumption has to be made. This could be a worst-case assumption, which considers the measured distribution of initial damages. As mentioned above, a sample size of sufficient failures is the base for a statistically firm prediction.

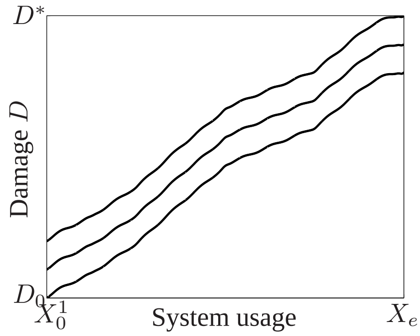

Figure 7: Remaining possible system usage, dependent on chosen initial damage D_0

The results of the damage prediction with different D_0 are depicted in Figure 7. As shown, different assumptions for D_0^1 lead to a different maximum system usages (offset).

The conclusion is that the achieved nlarx model is able to calculate the damage progression due to an applied load profile qualitatively. The most probable state-of-damage and the remaining system usage can be calculated. Finally, real stress profiles are used for adapting the general nlarx algorithm. The quantitative information about the exact remaining system usage is not available due the unknown correct initial damage.

Due to the aspect that the mathematical model for damage calculation is derived directly from simulated input/output signals and not from predefined/assumed physical connections, the method is sensitive to the chosen initial damage and cannot calculate the quantitative damage progression but the qualitative one. Hence, a damage prediction for known stress increments can be done more detailed. On the basis of this considerations further methods can be tested. This includes optimized maintenance strategies, reducing/enlarging inspection intervals, and so on.

3 CONCLUSIONS

The knowledge about the online state-of-damage of a system/component is one central aspect for exact predictive condition monitoring. The direct measurement of this state or correlation to a physical effect is hardly possible. Hence, nowadays strategies use static knowledge and several assumptions about the environmental/operating conditions to at least realize a preventive maintenance. One drawback is that the system is not used up to its maximal possible point of usage.

It is assumed that the damage progression is stress inherent. Hence the stress-damage connection is investigated in this contribution and the applied stress is investigated. In contrast to static linear approaches (e.g., Palmgren-Miner) an adaptive method is proposed here. Additionally no preliminary time consuming test have to be made to examine the material behavior because normal operating conditions differ from standard tests. This implies that arbitrary (within the normal range of operating conditions) stress profiles can be applied and the incremental damage is estimated. The idea of adapting the assumed model to the real operating conditions is explained. Furthermore it is assumed that all systems fail if a certain damage level is exceeded. Due to the individuality of the systems the initial damage need to be different than zero.

While the first estimation of the damage propagation is fairly bad due to a barely existing database, the estimation quality improves over the number of failed systems. After a certain number of failures the most probable mathematical model for calculating the average damage can be estimated. Hence, the damage progression for a given input stress signal can be derived by simulation.

As shown, the proposed approach can reconstruct the damage progression of a system, if the stress profile is measured from the beginning. Dependent on the assumed initial damage, the system can be operated until a predefined safety limit. In contrast to the strategies known from the literature this approach is dynamic. The main aspect is the ability to adapt to special operating conditions, systems, and stress profiles.

Up till now the whole strategy was tested by numerical implementations. For improvement and validation purpose this will be extended with real measurements and real systems. Further enhancements in determination e.g., of the weighting functions are necessary. Once the information about the average state-of-damage of a collective is obtained, strategies of extending the average system usage up to a certain point become possible.

REFERENCES

(Henry, 1955) D. L. Henry. A theory of fatigue-damage accumulation in steel. *Transac. of the ASME*, 77:913–917, 1955.

(Hwang and Han, 1986) W. Hwang and K. S. Han. Cumulative damage models and multi-stress fatigue life prediction. *Composite Materials*, 20:125–153, 1986.

(Marco and Starkey, 1954) S. M. Marco and W. L. Starkey. A concept of fatigue damage. *Transac of the ASME*, 76:626–662, 1954.

(Miner, 1945) M. A. Miner. Cumulative damage in fatigue. *Applied Mechanics*, 12:159–164, 1945.

(Nelles, 2001) O. Nelles. *Nonlinear System Identification*. Springer, Berlin, 2001.

(Palmgren, 1924) A. Palmgren. Die Lebensdauer von Kugellagern. *VDI-Z*, 68:339–341, 1924.

(Wöhler, 1870) A: Wöhler. Über die Festigkeitsversuche mit Eisen und Stahl. *Zeitschrift für Bauwesen*, 20:73–106, 1870.

(Wolters, 2008) K. Wolters. *Formalisms, simulation, and potentials of a dependability concept for optimized system utilization [in German]*. Dr.-Ing. thesis, Chair of Dynamics and Control (SRS), Engineering Faculty, University of Duisburg-Essen, Duisburg, Germany, May 2008.