Why SHM? A Motivation

Christian Boller
Saarland University, Chair of NDT and Quality Assurance
Fraunhofer IZFP, Campus E 3.1, 66123 Saarbrücken
GERMANY

c.boller@mx.uni-saarland.de

ABSTRACT

A large amount of our engineering infrastructure today is ageing would this be aircraft, ground vehicles, ships or buildings. Damage is a consequence of loads being applied to those engineering structures which has to be tolerated from a design point of view. Maintenance is the action resulting as a consequence and the more a structure ages the more maintenance and specifically inspection is required. Inspection is done by human beings mainly which can become costly the more inspection effort is required. Hence automation of the inspection process becomes an issue of consideration and with this structural health monitoring as well as management, generally found under the acronym of SHM. SHM is the integration of sensing and possibly also actuation into materials and structures such that non-destructive testing (NDT) becomes an integral part of those and inspection is mainly automated. This activity is deeply associated with the damage tolerance design principle which is the major basis for light weight design in aeronautics. Damage tolerance design can however also have a significant impact on operational life extension, which is another dimension in case light weight design may not be a sole design parameter. Within this chapter different reasons for enhanced inspection and hence SHM will be addressed as well as the steps being required to perform structural design in general. These steps do include loads and their consequence with respect to fatigue and fracture. This will then lead to the ways on how to inspect and hence monitor fatigue and fracture in terms of NDT and how this will lead to SHM in the very end including the limitations SHM may have with regard to damage tolerance design.

1.0 THE AGEING INFRASTRUCTURE

Engineering structures today are continuously ageing, would those be in civil engineering, energy generation or aeronautics, to just name a few. Civil engineering buildings such as houses or bridges are not truly designed for a finite life, although one is aware that those will not last forever. In nuclear energy generation, there is currently a significant discussion ongoing with regard to the life extension of existing power plants, which also stems from the fact that many of the power plants have been used less than they were expected to be. This is characteristic with regard to many of the high asset values a society or an operator such as an air force, army or navy has, where any extension of an initially assumed operational life is a potential benefit to the operator.

An area where damage is relevant from a design point of view is aviation. Aircraft are well-designed engineering structures from an engineering point of view. Due to their complexity in structural design as well as in operation, aircraft contain a variety of components as well as loading conditions that can lead to damage during in-service operation. This damage is quantitatively covered during approximately the first half of an aircraft's operational life, specifically in terms of a damage-tolerant design. However, during the second half of the life of those aircraft, a lot of damage does occur. This damage has often not been anticipated in the aircraft's initial design, and thus has to be covered by a variety of additional inspections and possibly even modifications. One of the most challenging aircraft types in that regard is the Boeing B-52 bomber. It was designed and built shortly after World War II and is due to remain in operation until the year 2045. Any modification to be done on this type of aircraft today –and a high number are expected – suffers from a lack of knowledge in the structural design of those old structures as well as in the degree an individual structure has truly aged. The means of identifying such a structure's degree of damage are therefore in dire need.

Aircraft are also very prone to accidental damage, mainly with ground vehicles loading the aircraft. Some of the damage can be clearly recognized such as the one shown in Figure 1. However, many damages are difficult to see. One such critical area is frames around the main cargo door of large aircraft such as the Boeing 747, where heavy loading trucks frequently collide with the lower frame and the locking mechanisms. This has led to accumulated damage and resulted in some cargo door lock failures in the early days of the Boeing 747, which triggered the enhanced inspection of cargo door locks and has since become an issue for continuous monitoring. Aircraft also have to withstand a variety of other types of repeating operational loads. One of them is hard landings where judgement is mainly referred to a pilot's – relatively subjective – judgement. Although landing gears are structures designed safe life and are thus replaced after defined intervals, it may be difficult to judge what effect the hard landing loads may have on a landing gear's adjacent structural components such as fittings, spars or frames. However, aircraft may also be operated to their design limits, which is specifically true with military aircraft. Figure 2 shows three examples a) a Panavia Tornado fighter aircraft, which was initially designed in the late 1960s for the East-West attack in central Europe and is currently used for reconnaissance missions in countries such as Kosovo, Saudi Arabia or Afghanistan, b) load spectra of a F-16 and CF-18, where the true operational spectrum exceeds the design spectrum after a portion of its operational life. If those changes in load spectra are not monitored significant damage will occur to those types of aircraft and the air forces themselves. It has to be considered that a change in the operational conditions such as payloads, manoeuvres or environment will result in a higher level of damage, since the structure, which was initially designed for specific operational conditions, can no longer be changed. In addition to metals, carbon fibre-reinforced polymers (CFRPs) play an increasing role in aircraft structures nowadays, where barely visible impact damage (BVID) is a source of significant concern and thus an option for monitoring.

Figure 1: Accidental damage in aircraft structures resulting from ground vehicles.

Aircraft structures, specifically commercial aircraft, are designed to be damage-tolerant. This allows lighter-weight structures to be designed. Damage tolerance means that the structure considered is able to withstand damages up to a defined size, above which a structure will then fail badly. Allowable damages in aircraft structures can achieve a significant length, ranging from a few millimetres up to a meter or more in size, depending on where a crack is due to initiate. Figure 3 shows an example of a hidden corrosion along a front spar web that starts around the rivets at the inner side and gradually grows in two directions before finally appearing on the surface above the fillet seal. Assuming that a crack of 5 mm or more in length might be reliably detectable by an inspector, the true crack length would already be centimetres or more when finally detected.

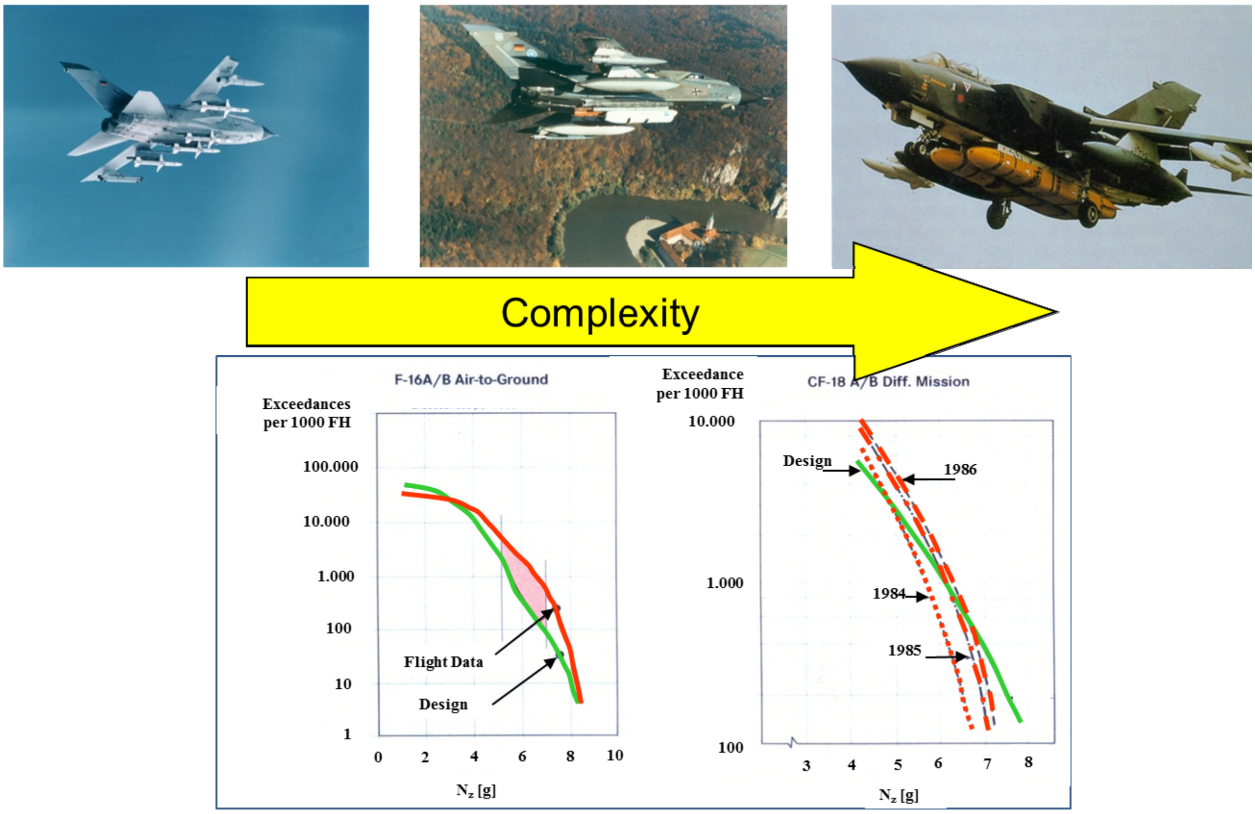


Figure 2: Changes in operational conditions of fighter airplanes of Panavia Tornado (top) and F-16 (bottom left) and CF-18 (bottom right) [1].

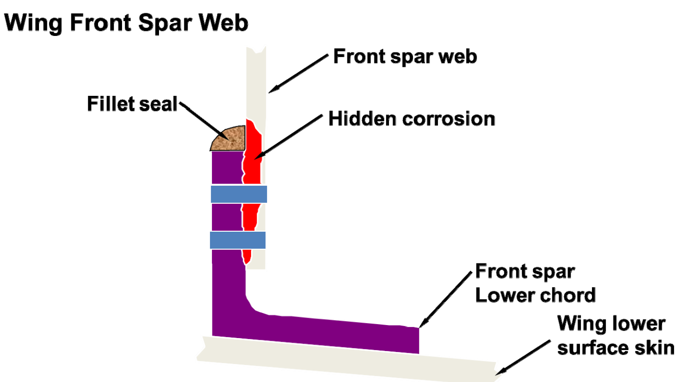


Figure 3: Wing front spar web with hidden corrosion of significant tolerable size.

Once damaged, a structure has to be repaired. Repairs in aircraft structures are standardized up to a certain degree. However, there are many conditions where standardization in repair does not apply such as with accidental damage. In those cases, tailored repair solutions often have to be determined, which might be best monitored continuously if technical conditions allow so. So far, this has unfortunately not happened. In a few cases, some major aircraft have failed catastrophically from repairs, such as with the Japan Airlines B747 in 1985 or the China Airlines B747 in 2001.

Cracking is however not a subject specifically limited to aircraft only. Ground vehicles being overloaded can face the same situation too as well as civil infrastructure. The increased loading in road traffic would this be the increase in number of vehicles or axle loads has led to an increased amount of damage to road infrastructure and resulting infrastructure inspection as shown as an example in Figure 4 below. Serious road infrastructure bridge collapses have been reported during the past years in different even technologically advanced countries where further deterioration is most likely to be expected.

Figure 4: Examples of road infrastructure deterioration (Source: www.strassen.nrw.de).

Damage in aircraft structures can be caused by a variety of factors, which are summarized in Table 1. All of this damage is detected and quantified through non-destructive inspection (NDI). Most of the inspection is done on a visual inspection basis, supported by NDI techniques such as ultrasonic and eddy current testing. It is further supported by specific damage tolerance design criteria and/or loads monitoring if applicable. Inspection however requires a significant amount of effort, specifically when the location to be inspected is located in a hidden place where a significant number of components have to be dismantled and then reassembled. This requires an aircraft to be out of operation for a substantial amount of time. Thus, it would be advantageous if this amount of time could be significantly reduced due to automation of the inspection process. Furthermore, many of the subjective judgements made today by pilots, such as hard landings, could be made more objective through a loads monitoring system.

If SHM is considered to be the integration of sensing and possibly also actuation devices to allow the loading and damaging conditions of a structure to be recorded, analysed, localised and predicted in a way that non-destructive testing becomes an integral part of the structure, then the design as well as life cycle procedure of a structure has to be considered. The design of a structure starts from the applied loads to be assumed which can be static loads mainly in terms of the ultimate load as well as fatigue loads, where the latter are even usually random in terms of their height. Those loads have to be brought into conjunction with a structure's geometry where notches in conjunction with the loading condition (uni- or multi-axial) are the major drivers with respect to the maximum resulting stresses being applied. Those applied stresses have to be smaller than the allowable stresses defined by the material to be used otherwise either the structure's geometry and/or the material used have to be altered. Knowledge of a structure's loads, material and geometry are therefore essential elements to assess a structure's condition in terms of SHM implementation in general. This becomes increasingly essential with a structure's age. The following chapters will therefore specifically elaborate on those with regard to loads, fatigue and fracture before addressing NDT based inspection and the SHM system to be configured in the end.

Table 1: Examples of damage in aircraft structures and current solutions.

Issue	*Solution*	*Comment*
Accident	Visual and/or Non-Destructive Inspection (NDI); Repair	Barely visible damage can be overseen
Loads	Design spectrum; pilot judgement followed by possible NDI	Very subjective judgement
Fatigue & Fracture	Damage Tolerance Design; Major Structural Testing; NDI; Inspection Interval	Can be time consuming and labour intensive
Corrosion	Design; Corrosion Protection Plan; NDI; Inspection Interval	
Multi-Site Damage	NDI; Inspection Interval	
Repair	Quality Assurance; NDI; Inspection Interval	

2.0 LOADS AND OVERLOADS

Loads are the source of any material deterioration, regardless of whether this deterioration happens on a micro or macro scale. Loads are usually associated with mechanical loads, which generate fatigue and fracture as damage. Loads may, however, also be environmental loads such as those generated by temperature, humidity, irradiation or any other type of parameter that might lead to deterioration. Fatigue and resulting fractures are generated in a more complex way when loads are repetitive. Therefore, the time domain signal of a load needs to be known and therefore may have to be monitored. Since mechanical loads are usually the major point of concern in fatigue analysis, considerations in this paper will be limited to mechanical loads only. However, the ideas expressed may be synonymously transferred to other types of loads such as those being generated from the environment. Loads are a fingerprint of a structure's operational conditions. Many of the different structural components therefore do not follow the same pattern of a load sequence, specifically when they are exposed to a stochastic process or combined with additional environmental loads. Only a very limited number of structures operate under a sequence of loads where the height of the load is permanently constant and where the spectrum is called "constant amplitude." Most of the loads sequences have specific spectra. Imagine an aircraft and the load at the wing attachment box, as shown in Figure 5. While the aircraft is on the ground, the load might be slightly compressive or close to zero. However, once the aircraft has taken off, the load increases significantly to a high tensile load that is directly related to the aircraft's flight load. This load cycle is also called the ground-to-air cycle. While in the air, the aircraft will manoeuver and face different gusts. This will result in dynamic loads cycling around the mean load carrying the aircraft. Once the aircraft lands, the tensile load will vanish or even turn into compressive. Different applications will show different time domain loading patterns.

Load sequences are difficult to characterize from a time domain point of view, as can be realised from the example provided in Figure 5. Different load cycle counting procedures have been developed during the 1930s to 1970s where the rainflow cycle counting method [3] has recently turned out to be the most widely accepted. Figure 6 shows a time-domain sequence with its various peaks and troughs, where the shape of the time-domain signal can be imagined as the bottom of a water basin reservoir. Now imagine this reservoir is filled to the top with water (Figure 6 left), and the bottom of the water basin is gradually being shaved from its lowest trough upwards. The first trough that will be hit will be load level -2. Then cut water will flow out from an upper level of +4.5 down to a lower level of -2. This is considered to be the cycle of the largest loading span and has been registered as cycle 3 in the upper right of Figure 6. The next trough to be hit is at load level -1.5, where water will flow out from load level 3.5 down to -1.5. It is considered as cycle 1 in Figure 6. This is followed by the trough at load level -1, where water now only flows out from load level 2.5 down to -1 because the water level has already been reduced from 4.5 to 2.5 due to load cycle 3. Finally, a trough at load level +1 is hit, where water drops from load level 2.5 down to 1 due to the water level reduction already resulting from cycle 1.

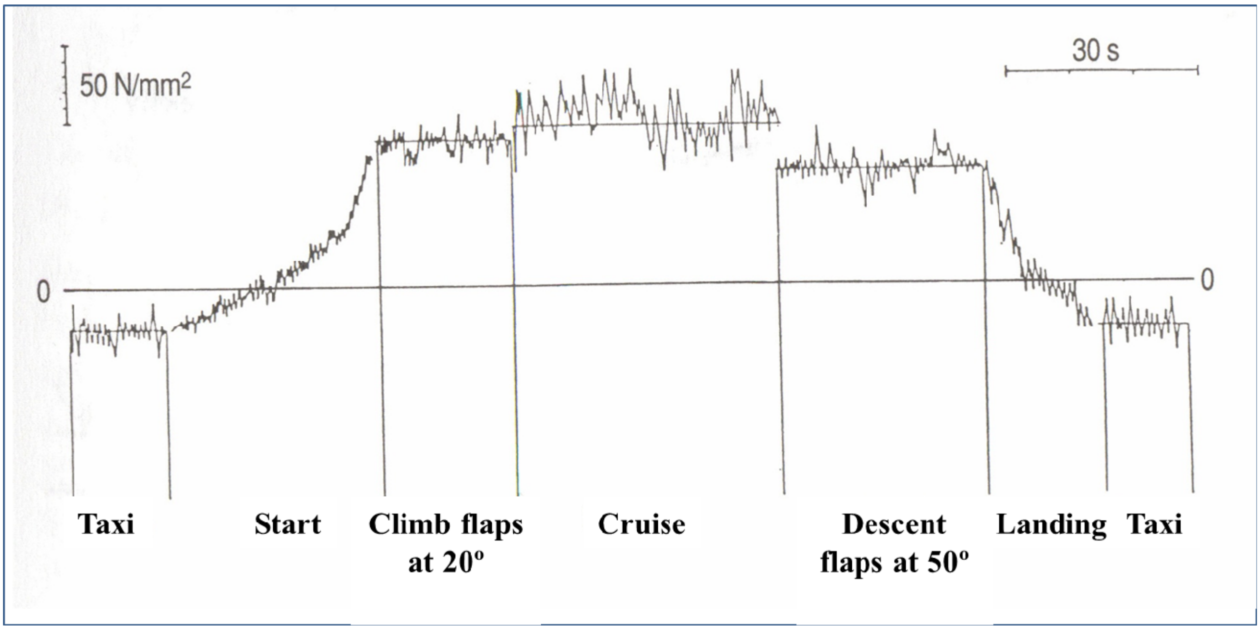


Figure 5: Stress sequence in tension girder of the wing root of a transport aircraft [2].

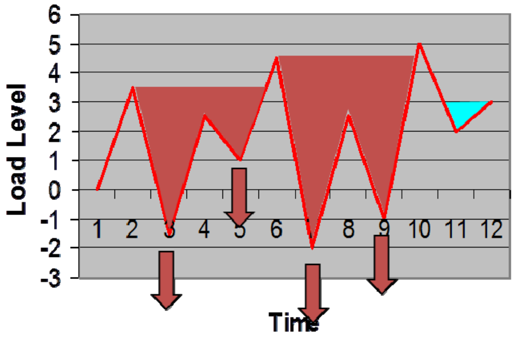
There is a residual remaining with a trough at load level +2, which cannot be considered, since it is unknown where the load sequence will continue. With conventional load sequences and their generally large number of cycles, this residual can be considered to be negligible. Each cycle can now be represented either in terms of the maximum and minimum load/stress or the amplitude and the mean, as shown in Figure 6. The cycles can be further plotted in terms of their loading range as exceedances, as shown on the lower left of Figure 6.

Load cycles determined through the rainflow cycle counting may also be summarized and described in a matrix as the number of load half cycles going from a minimum load level q (trough) to a maximum load level p (peak) within a complete load spectrum. Numbers 14 to 19 indicated in Figure 7 represent arbitrary load levels, which have to be decided in accordance with the specific needs. Figure 7 thus represents a matrix where each of the elements ΔF_{ij} represents the number of cycles going from load class i to load class j.

Going along the horizontal axis and looking at the upper right triangle provides all the peaks, p, with increasing number. Looking at the vertical axis downwards and the matrix's lower left triangle provides all the troughs, q, respectively. Similar considerations can be made with amplitudes a and means m when converting those as follows:

$$q = m - a \qquad p = m + a$$
$$a = (p - q)/2 \qquad m = (p + q)/2$$

Cycle	σ_{max}	σ_{min}	σ_a	σ_m
1	3.5	-1.5	2.5	1.0
2	2.5	1	0.75	1.75
3	4.5	-2	3.25	1.25
4	2.5	-1	1.75	0.75



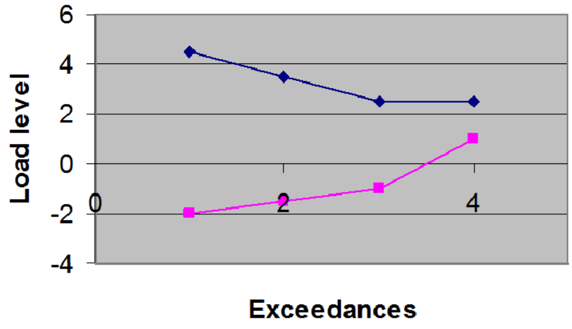


Figure 6: Rainflow cycle counting method [3].

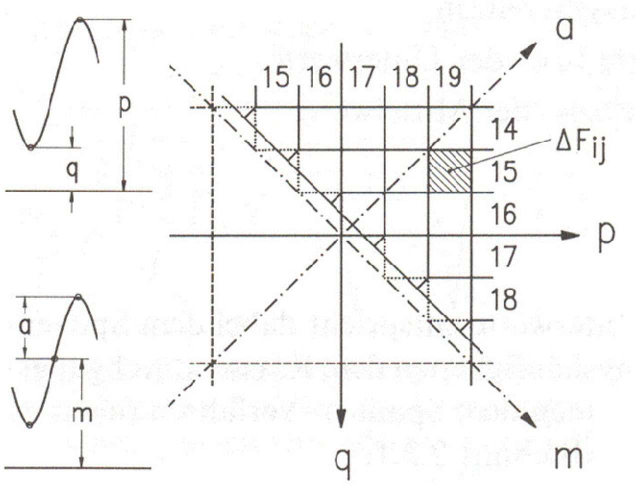


Figure 7: Load matrix descriptions and definitions (figure from [2]).

A rainflow matrix, as shown in Figure 7 and having once been filled, can be pseudo-randomized in accordance with a procedure described in Figure 8. The 32x32 rainflow matrix shown in Figure 8 is filled with numbers a_{ij} . If the randomized service load sequence originates from class α , a term T_α is determined in accordance with $T_\alpha = \sum_{j=\alpha+1}^{n} a_{\alpha j}$ with a random number R_n being determined as:

$$R_n = A \cdot R_{n-1} + (B + 1 + W)(MOD\, M)$$

with $M = 2^r$ and $r = 1 + INT(\ln T_\alpha / \ln 2)$

$$D = M - T_\alpha$$
$$A = (5)MAX(M/2 - 3)$$
$$B = (3)MAX(M/4 - 1)$$

$W = +1$ when turned towards a peak and $W = -1$ when turned towards a trough.

For the very first cycle when loading turns towards a peak, R_{0i} turns out to be:

$$R_{01} = D_i + \sum_{j=i+1}^{16} a_{ij} \quad \text{for } i = 1...15 \text{ and } R_{0i} = D_i \text{ for } R_{0i} = D_i \text{ for } i = 16...32$$

When the first cycle turns towards a trough, R_{0i} is determined as:

$$R_{0i} = D_i \text{ for } i = 1...17 \text{ and } R_{0i} = D_i + \sum_{j=17}^{i=1} a_{ij} \text{ for } i = 18...32$$

Target load level β is achieved when $R_n - \sum_{j=\alpha+1}^{\beta} a_{\alpha j} \leq 0$ and $a_{\alpha\beta}$ is now reduced by 1.

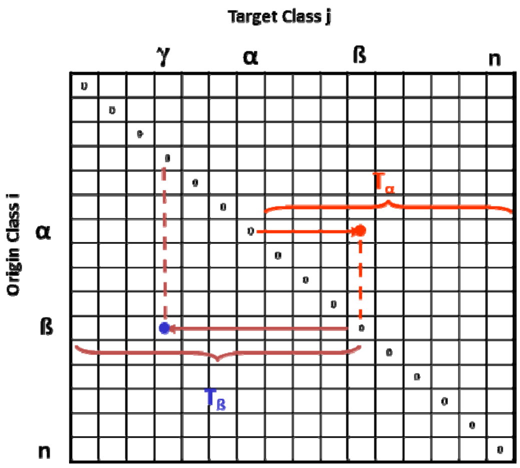


Figure 8: Pseudo-random load sequence generation scheme based on a rainflow matrix [4].

If load level β is a peak (as shown in Fig. 8), the same procedure is determined the opposite way to determine load level γ as shown in Figure 8. This procedure is then continued until all peaks and troughs within the matrix have been absorbed. A resulting load sequence determined might then look like the one shown in Figure 9, where N_0/N_1 is the ratio of the number of cycles of zero crossings versus the complete number of cycles within the load spectrum.

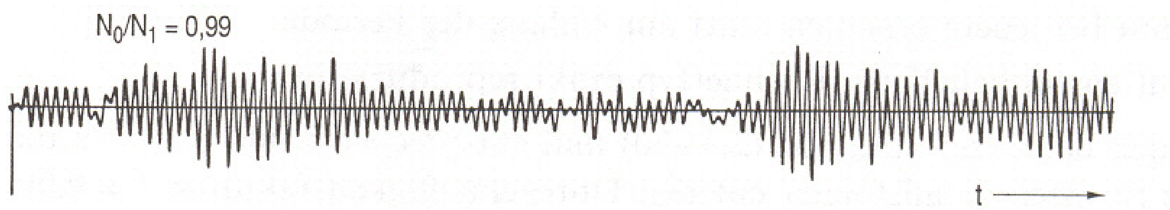


Figure 9: Pseudo-random load sequence generated from a rainflow matrix.

3.0 FATIGUE AND FRACTURE

3.1 Fatigue Life Evaluation

To characterize the fatigue behaviour of an unnotched (material) or a notched component (structure), constant amplitude fatigue tests are performed at different load levels, as shown in Figure 10. The result obtained is a fatigue life curve (S-N curve) for either a crack initiation or complete fracture. Those S-N curves can be linearized when plotted on a log-log scale, where an endurance limit can be seen at fatigue lives of approximately 10^6 cycles. This ultimately results in a bi-linear relationship in a log-log scale. An S-N curve depends on the material, component (geometry/shape), loading and nominal stress definition considered. S-N curves may be determined experimentally or taken from handbooks (i.e. [5]).

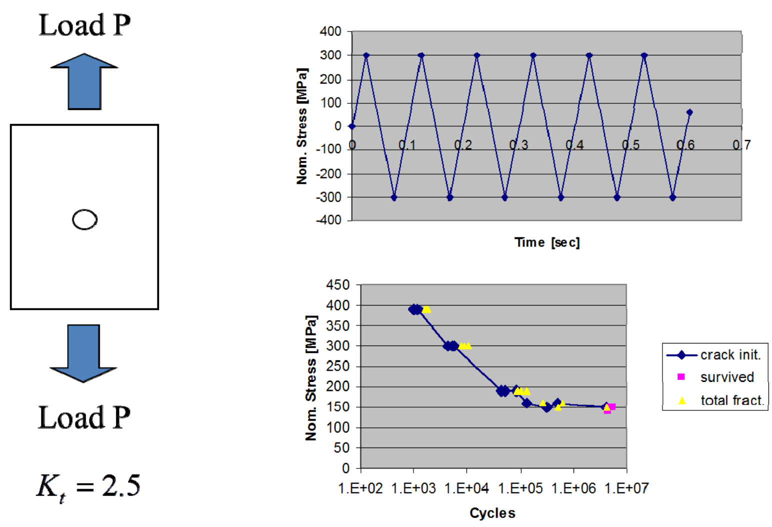


Figure 10: Determination of a constant amplitude S-N curve of a notched specimen.

Once an S-N curve is made available, the fatigue life of a structure can be estimated. The method to conduct an estimation in principle is shown in Figure 11. If a load sequence consists of the following three loading blocks: a) $n_1 = 6$ cycles of a maximum stress of 300 MPa, b) $n_2 = 4$ cycles of a maximum stress of 400 MPa and c) $n_3 = 5$ cycles of a maximum stress of 200 MPa, for the stress range block a) the component would endure approximately $N_1 = 2 \times 10^5$ cycles. This would result in a linearized damage D_i per load cycle to be $D_i = 1/N_1$ or for the complete block $D_1 = n_1/N_1$. The same is applied for blocks b) and c), where N_3 turns out to be infinite in the case of block c). Hence, the damage per loading cycle is zero. Damage is assumed to accumulate linearly in accordance with the Palmgren-Miner rule [6,7], which is defined in Figure 11. Fatigue life is achieved when accumulated damage D turns out to be unity.

With such an approach combined with a FE analysis, accumulated fatigue damage can be virtually calculated for each element, allowing areas of high damage accumulation to be determined. These areas are those where damage monitoring will have to occur first. Therefore, this method of fatigue life evaluation is an essential strategic means for an efficient SHM process.

3.2 Fracture and Crack Propagation

Cracks in general terms are damages that can often be seen on a visual basis considerably. They may range from a few millimetres to centimetres and possibly even metres depending on what the structure considered is able to tolerate without collapse. Cracks can be principally loaded under the three different conditions shown in Figure 12, which are called cracking modes or simply modes. Mode 1 represents an axial tensile loading, mode 2 a bending loading, and mode 3 a twist loading, respectively.

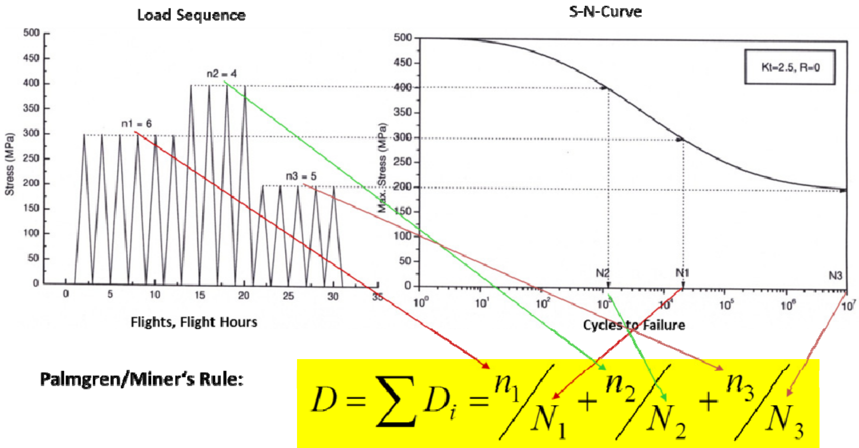


Figure 11: Fatigue life evaluation according to Palmgren Miner's rule.

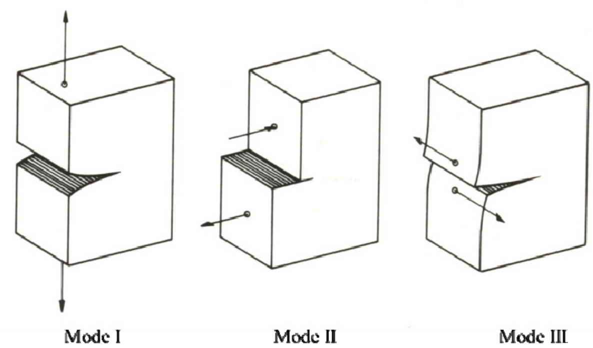


Figure 12: The different fracture modes.

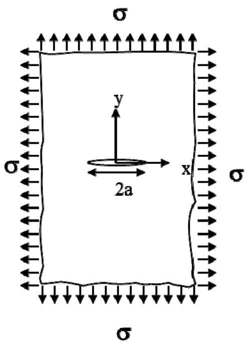
Stresses σ_{kl} around the tip of an arbitrary crack under arbitrary loading can be principally determined on the basis of the following equation:

$$\sigma_{kl} = \frac{1}{\sqrt{2\pi r}}\left[K_I f_{kl}^{I}(\varphi) + K_{II} f_{kl}^{II}(\varphi) + K_{III} f_{kl}^{III}(\varphi)\right] \qquad (3.1)$$

$$k,l = x,y,z \quad \text{Mode 1} \quad \text{Mode II} \quad \text{Mode III}$$

where r is the radius from the centre of the coordinate system, as shown in Figure 13 for a crack in an infinite plate under bi-axial loading. K_I , K_{II} and K_{III} are the stress intensities for the different modes, and $f(\varphi)$ are dimensionless functions, depending on the crack mode. In the case of mode 1, K_I turns out to be:

$$K_I = \sigma\sqrt{\pi a} \quad \left[N/\sqrt{mm^3}\right] \qquad (3.2)$$



$$\sigma_x = \frac{K_I}{\sqrt{2\pi r}} \cos\frac{\varphi}{2}\left(1 - \sin\frac{\varphi}{2}\sin\frac{3\varphi}{2}\right)$$

$$\sigma_y = \frac{K_I}{\sqrt{2\pi r}} \cos\frac{\varphi}{2}\left(1 + \sin\frac{\varphi}{2}\sin\frac{3\varphi}{2}\right)$$

$$\tau_{xy} = \frac{K_I}{\sqrt{2\pi r}} \sin\frac{\varphi}{2}\cos\frac{\varphi}{2}\cos\frac{3\varphi}{2}$$

Figure 13: Stresses under mode 1 around a crack tip for a crack in an infinite plate under bi-axial loading.

Since the geometry of the component has an influence on the stress intensity factor K, a geometry factor Y is additionally introduced, extending Eq. 3.2 to become:

$$K_I = Y\sigma\sqrt{\pi a} \quad \left[N/\sqrt{mm^3}\right] \qquad (3.3)$$

Reference for all geometry factors is a crack in an infinite plate, where Y is unity whereas it will differ for all others. An example of a cracked finite plate is given in Figure 14. Further geometry factors can be found in handbooks such as [8-11] and possibly in other relevant publications, or they may have to be determined by numerical means.

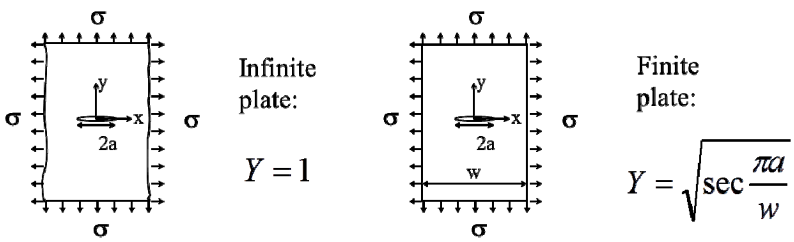


Figure 14: Geometry factors for a cracked infinite and finite plate.

When a component is fatigue loaded, the load alternates between a minimum and maximum load and thus stress. Given that the stress intensity factor K is dependent on stress, it will also alternate and can therefore be differentiated between a minimum and maximum value:

$$\begin{aligned} K_{\min} &= Y\sigma_{\min}\sqrt{\pi a} \\ K_{\max} &= Y\sigma_{\max}\sqrt{\pi a} \end{aligned} \qquad (3.4)$$

and with a range ΔK being:

$$\Delta K = K_{\max} - K_{\min} = Y(\sigma_{\max} - \sigma_{\min})\sqrt{\pi a} \qquad (3.5)$$

K also depends on the crack length, and the influence of the crack length might be less under the minimum load than under the maximum load due to an effect called crack closure. Crack closure is defined as the stress σ_{op} where a crack opens and hence $\sigma_{\min}$ becomes σ_{op}. ΔK is therefore then defined as:

$$\Delta K_{eff} = K_{\max} - K_{\min} = Y(\sigma_{\max} - \sigma_{op})\sqrt{\pi a} \quad (3.6)$$

When a fatigue test is performed on a cracked plate under constant amplitude loading, crack propagation can be observed. A similar test performed at a higher load level will allow the crack to propagate at a higher speed (see Figure 15). When plotting the change in crack propagation over fatigue life da/dN over the range in stress intensity ΔK on a logarithmic scale, a relatively linear relationship is obtained, as seen in Figure 15. This relationship is described as:

$$\frac{da}{dN} = C_P\left(\Delta K_{eff}\right)^{m_P} \quad (3.7)$$

and has been determined by Paris [12]. Further details on how to determine crack opening loads as well as ΔK_{eff} can be found in [13].

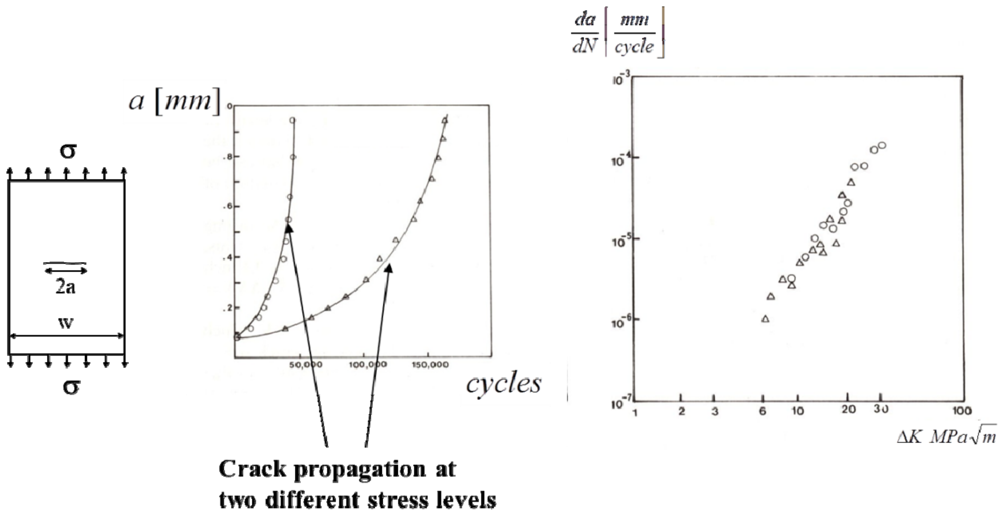


Figure 15: Crack propagation behaviour under two different stress levels.

Crack propagation calculations and thus estimations can now be made loading cycle by loading cycle in accordance with the following steps:

1. Assume an initial crack length a_0 that can be reliably detected by means of non-destructive testing or SHM;
2. Determine the geometry factor Y either from handbooks, literature or numerical analysis;
3. Determine the nominal stress from the maximum or minimum load to be applied within the load cycle considered;
4. Calculate the minimum and maximum stress intensity factors $K_{\min}$ and $K_{\max}$ as well as resulting ΔK_{eff};
5. Calculate crack propagation rate da/dN and thus crack extension Δa for the load cycle considered;
6. Add crack extension Δa to the initial crack length a_0 to determine the new crack length a to be used as the initial crack length in step 1 of the next load cycle;
7. Continue with step 2 accordingly.

With this procedure crack propagation over the fatigue life can be calculated and allows a crack propagation curve to be determined as it is schematically shown on the left hand side of Figure 16. From this curve and calculation an interval ΔN can be determined, that describes the interval between the incident where a crack can be reliably detected by means of NDT and SHM respectively and the incident where a crack becomes critical from a structural integrity point of view. This interval is the interval where a crack can be controlled and hence monitored and since crack propagation is a process that can result in a lot of scatter an inspection is usually recommended to be performed at least after $\Delta N/2$ loading cycles. Another significant parameter is the residual strength in terms of fracture toughness, which is shown on the right hand side of Figure 16 and which describes which minimum crack size must still exist such that the maximum load of an in-service load spectrum can still be withstood. Those two diagrams are the essential ones to perform damage tolerant design and which is one of the major motivations for SHM in general whatever structural type one considers.

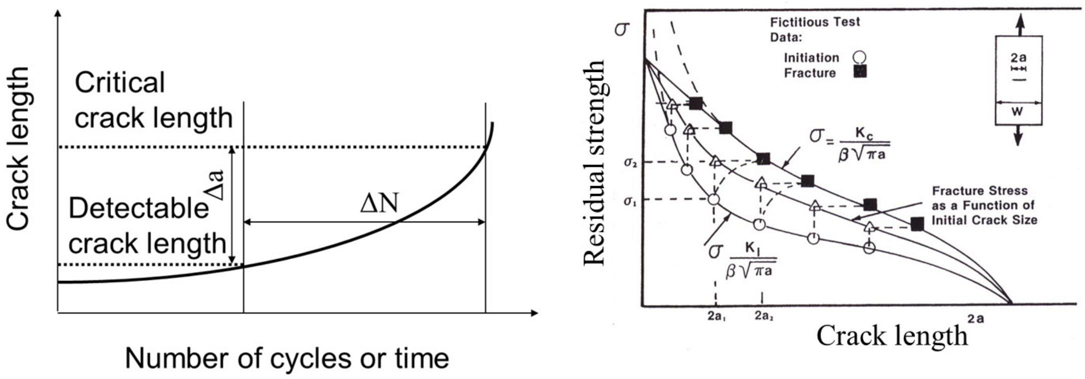


Figure 16: Crack growth (left) and residual strength (right) as key parameters for crack propagation based fatigue life assessment.

4.0 DAMAGE TOLERANCE DESIGN

When being loaded, structures can withstand a considerable number of load cycles until they fracture. Fracture is considered a defined damage condition, for example, a crack of a defined size. When a well-defined component of a material, shape and manufacturing process is fatigue loaded, it will fracture at a specific number of fatigue cycles. If the test is performed again on the next component made of the same material, shape and manufacturing process, it will most likely fail at a different number of fatigue cycles. Further repetitions of the fatigue test will finally result in a scatter band of fatigue lives, as shown in Figure 17. Looking at the resulting S-N curve, this results in a zone where: a) nothing fractures, b) fracture occurs and c) everything has fractured. If a structure is designed so that it will not fracture, it has to be designed so that it falls into zone a). Once the component has achieved the fatigue life defined by the S-N curve of a 0 percent probability of fracture, the component will have to be removed, even though it might still have a substantial amount of remaining fatigue life. This is called *safe life design*. Since the abscissa of an S-N curve is plotted in a logarithmic scale, the scatter band of fatigue lives (zone b) may easily be in the range of a factor of two or higher. Therefore, if the aforementioned component is run after being removed under safe life design conditions until its true fracture life, it may have lasted an additional number of fatigue cycles, possibly the same number as it would have lasted under a safe life design. Allowing a component to be run up to fracture, however, requires understanding the damage mechanisms within the component and inspecting the component at well-defined intervals. This is called *damage-tolerant design* and involves SHM. Damage-tolerant design allows a structure to last longer, or if this is not desired, to apply higher stresses, which results in a lighter-weight design. Aeronautics has specifically taken advantage of the latter, which has now become standard, specifically in commercial aviation.

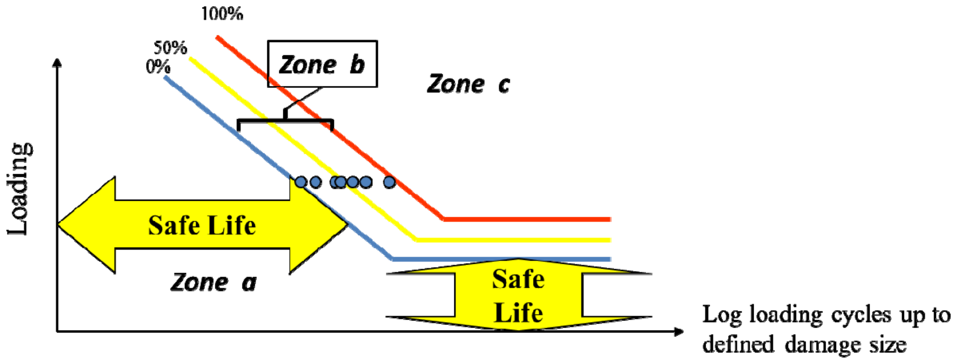


Figure 17: Scatter in S-N curves and the impact on structural design principles.

The idea of damage-tolerant design can also be extended to engineering structures in general. In that case, a component in a structure may fail if neighbouring components are able to take over the loads in the sense of a redundancy. Therefore, the structure with the resulting fracture may even be controllable under these circumstances. This is generally known as *fail safe design*. If microcracking is not considered to impair the structure's integrity, damage tolerance starts from where a crack can be reliably observed and damage progression can be controlled until a tolerable damage constellation within the fail safe area with the size of tolerable damage as defined by structural design is achieved, as shown in Figure 16 left. Since fracture is a stochastic process it is unlikely that a critical crack may be observed after a defined inspection interval of $\Delta N/2$ .It may be much more likely that no detectable crack may be observed after this interval. If this is the case which is considered as »no failure found«, then the assumption of the crack starting at the detectable crack length can be started again and the procedure is repeated as many times until a crack is found that is larger than the defined *detectable crack* but definitely smaller than the *critical crack length* which is shown schematically in Figure 18. Once such a crack (damage) is observed repair has to be done irrespective of how far away the crack is from the critical crack length (or size) wise.

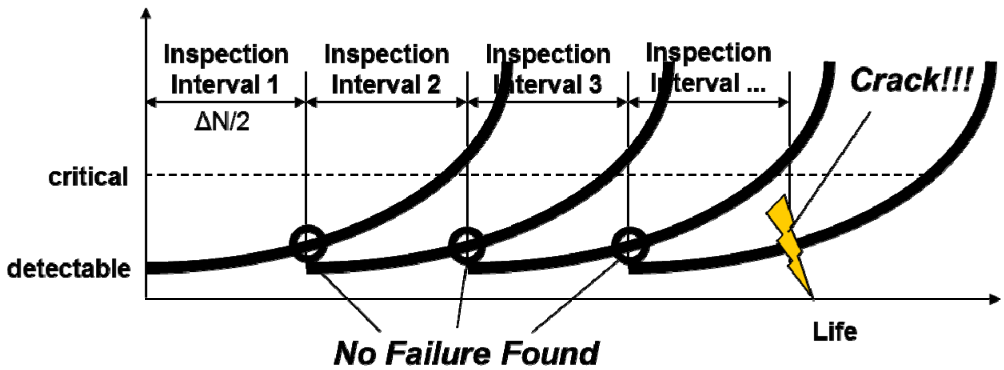


Figure 18: Crack propagation and inspection und damage tolerance principles.

That damage tolerance has its rewards can be seen for the structure shown in Figure 19. Assume an aircraft fuselage component of the type shown in Figure 19 and the condition that frames and stringers might not be inspectable. Hence, the structure can only be inspected from the aircraft's outer side. Under those conditions, the frames and stringers have to be considered broken in case of crack propagation life estimation. Consequently, the inspection interval has to be based on the crack propagation observed only on the fuselage's outer surface. However, if stringers and frames are inspectable, such as with an SHM system, then assumptions for crack propagation could be changed. This may lead to a significant increase in crack propagation life determined or inspection interval defined. If this increase in fatigue life is not of interest for the case considered, then another option is to increase the stresses applied. This would then lead to a decrease in the structural material used and thus result in a lighter-weight design. Again, this is identical to the effect already observed with damage-tolerant design. Therefore, in many cases, implementing SHM merely involves enhancing the idea of damage-tolerant design, or in other words, determining how much potential a structure has in accordance with the damage tolerance principle.

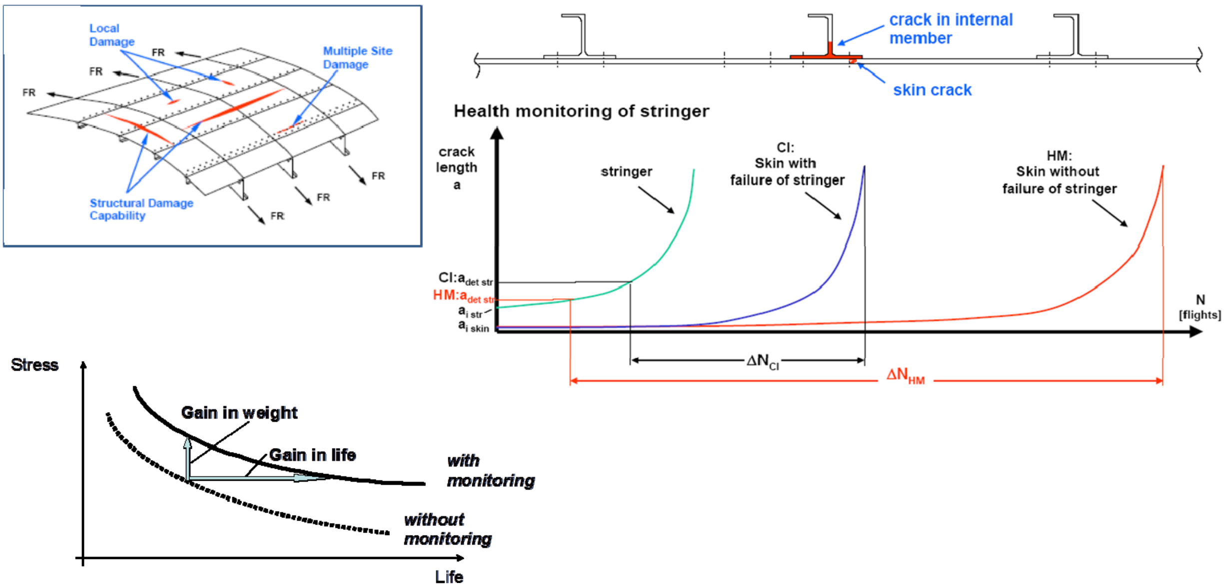


Figure 19: Gain in structural life through health monitoring (HM) of a stringer in an aircraft panel when compared to conventional inspection (CI) (Source: H-J Schmidt, Airbus).

Inspection of engineering structures may often not be easy and pleasant to do. However most of the inspections today are done by humans who have to crawl behind bulkheads or into fuel tanks in the case of an aircraft or onto chimneys, pylons or walls in the case of civil engineering infrastructure. This is laborious, costly and dangerous and triggers the idea of automation being the motivation to get SHM hence developed.

5.0 STRUCTURAL HEALTH MONITORING

As mentioned before, SHM is considered to be the integration of sensing and possibly also actuation devices onto or into materials and structures to allow the loading and damaging conditions of a structure to be recorded, analysed, localised and predicted in a way that NDT becomes an integral part of the structure. Looking at this in a wider context this may not be fully new. In aviation, SHM was recognized after the serious Comet accidents in the early 1950s, although nobody used the expression “SHM” at that time. Acceleration sensors were implemented on fighter airplanes and load spectra were continuously tracked. This resulted in a variety of standard load sequences being used for aircraft qualification [14,15]. Today, a modern fighter aircraft like the Eurofighter Typhoon has a 16-strain gauge operational loads monitoring system. It allows real load sequences to be tracked and damage accumulation to be determined. Similar approaches are made with bridges and other load-critical structures and components. With the development of optical fibre Bragg sensing, a completely new option of sensing has been provided that allows the complexity of sensing to be reduced by an order of magnitude and more. One hundred conventional strain gauges with two wires each can now be easily replaced by a single fibre, giving rise to a multitude of interesting new applications.

A major rule in implementing SHM is: *Maximize a SHM system's efficiency by minimizing the number of sensors to only the most efficient ones.* Sensors are only one part of an SHM system. A major element of an SHM system is simulation, which can be large computational machines doing nothing more than simulating the performance of a structure under various conditions. This simulation requires an input from sensors, and the question is: *Which input will have what impact on an SHM system's performance output?* This is basically an optimization task, which consequently requires minimizing the sensor's input to maximize a structure's performance output.

In the case of SHM, this means starting where the structural design starts, similar to the strength analysis of materials and structures. A strength analysis requires material properties, the structure's geometric shape, and operational loads as inputs. Material properties can be considered to be given, specifically when materials have been qualified through a quality assurance process. Geometric shape can also be considered to be given, where no change is expected over the structure's life. The only major unknown at this stage is operational loads, where assumptions are made prior to a structure's realization, and where safety factors have to be implemented to address any unexpected uncertainties. Reducing those uncertainties is why sensors have to be implemented first, and where major benefits can be expected due to the high safety (or better penalty factors) initially imposed. Determination of the type and sensors' best locations therefore has to come from the structural simulation process itself.

Once an optimum number of load monitoring sensors have been determined, placed on the structure and are operating reliably, the safety factors can be reduced. This results in either a lighter weight design and/or longer endurance, depending on the operational target of the structure considered. The next factor of uncertainty within the structure's operational life is damage, possibly combined with a change in the structural material's condition and the probabilistic nature of damage to occur. Again, simulation might be a very beneficial tool. However, this time, the tools might be used more for the simulation of materials properties on a molecular and possibly even atomistic level. This might generate ideas for sensing other than those traditionally considered in terms of strain monitoring.

With all that information available from the load and material condition monitoring sensors as well as from the simulation algorithms applied, the locations of likely damage occurrence can be determined, even on a prognostic level. Again, this will allow an SHM-based damage monitoring system to be determined on a targeted and thus optimized level. An example for such a concept is shown in Figure 20, where a map indicating the locations ("hot spots") of highest damage accumulation is obtained through prognosis and thus simulation. Simulation has to be performed again for those "hot spots." However, this time it is on the basis of simulating physical waves travelling around the hot spot and the allowable damage to occur, such that the damage tolerated can be reliably detected with the NDT-based transducer configuration to be determined through simulation and then successively realized in hardware.

Although Figure 20 might imply that actuators and sensors or thus transducers might be the key elements of SHM, this is certainly not the case. The SHM system is essential with SHM. An SHM system, as shown in Figure 21, consists of a defined number of transducers that are managed by a signal generator, amplifiers, a multiplexer, a data acquisition unit and a processor. The most important requirement with this SHM system is that the single input and single output interface provides a reliability of information defined by the SHM system user requirements. If the rate of false alarms is set to say 10^{-7} by the SHM system's user, then it is the obligation of the SHM system provider to build a system that will meet that requirement, regardless of how many transducers the system will need, as long as the SHM system meets any other established requirements such as weight and cost.

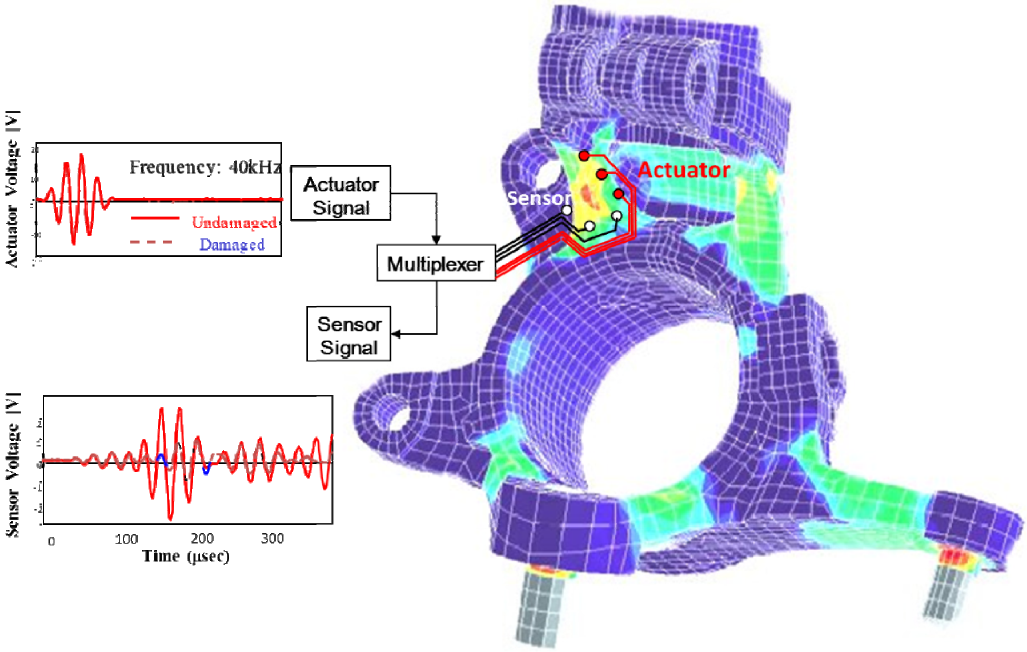


Figure 20: Fatigue life and damage evaluation to determine "hot spots" for damage monitoring.

As to the sensors or generally transducers to be considered there is virtually an absolutely wide range of sensors being developed. However not all of them are used for SHM but rather those which are known from strain measurement of NDT. Those transducers being most popular do include electrical strain gauges and fibre optic sensors from the loads but also modal analysis side or piezoelectric transducers from the magnetics and vibration analysis side. Options do also exist within the electromagnetics side to a limited extent and there are also some special approaches around such as comparative vacuum monitoring (CVM). Many of those techniques are described in more detail in [16] with a selection of those shown in Figure 22 below.

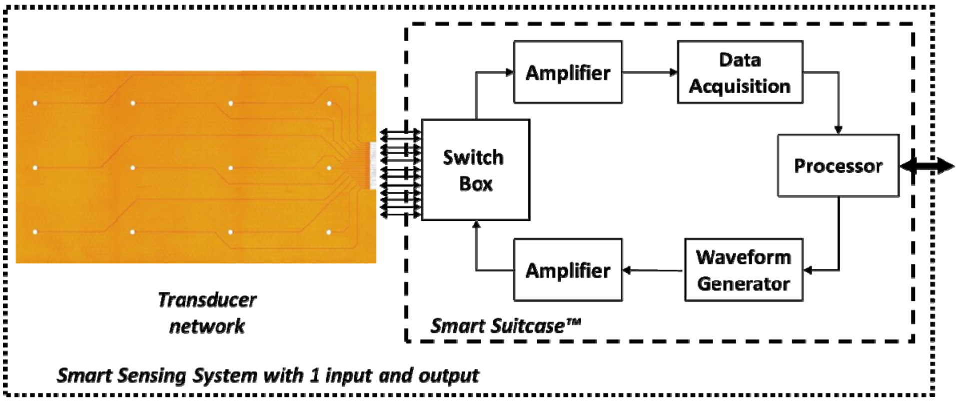


Figure 21: Principle of an SHM sensing system for damage monitoring.

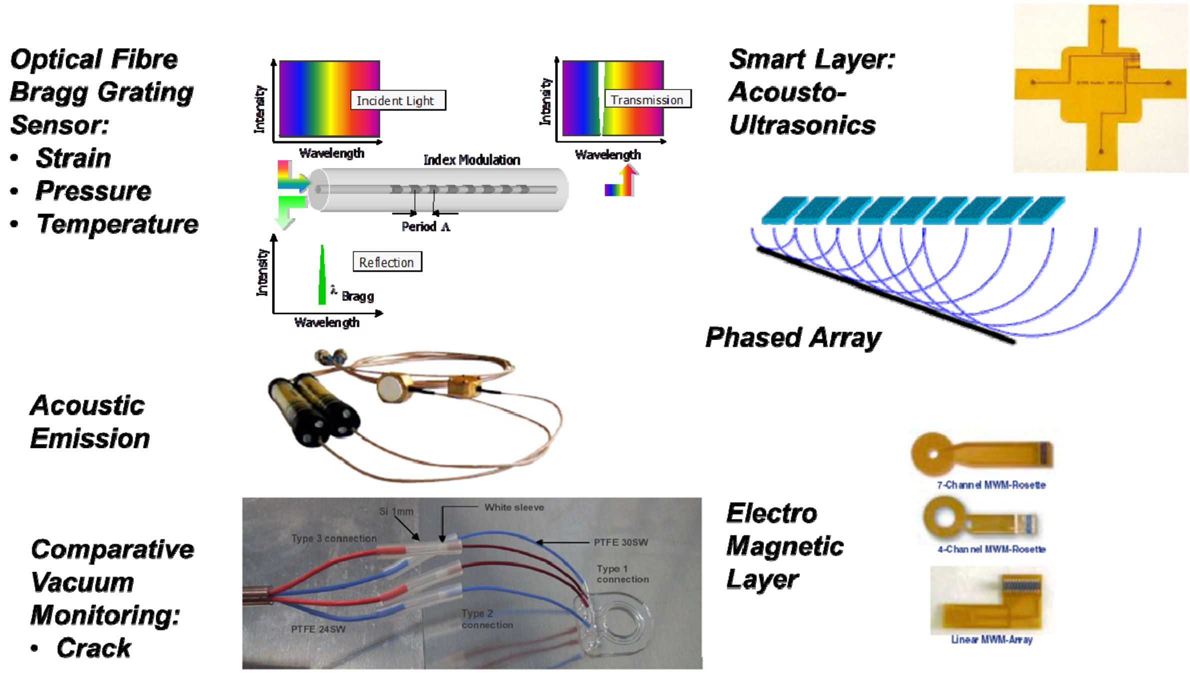


Figure 22: Selection of sensor systems and principles used for SHM.

There is a variety of interesting applications for SHM specifically when a structure is ageing. A major field driving the SHM development is ageing aircraft where the effect of multi-site damage (MSD) along rivet lines has attracted specific interest since monitoring MSD in ageing aircraft may be associated with a lot of additional inspection effort. Therefore, it was explored how well SHM might identify MSD along rivet lines. A substantial overview of this issue is provided in [17], with a major conclusion that cracks of 5 mm in length and longer could be well determined with SHM when compared to conventional NDT. This was confirmed in a follow-up study [18] in which the reasons for limited monitoring sensitivity with SHM were analytically explored and could be explained. Even when not considering the geometric complexity of a rivet hole, the scattered signal resulting from a pure longitudinal Lamb wave generated is definitely not as strong as one would believe. In [18], some simulation results (see Figure 23) have been reported where a longitudinal wave has been sent through a plate with several holes, one of which had a crack of increasing length. Figure 23 shows the intensity of the scattered wave distribution behind the hole for a quarter section only. It can be seen that the maximum distribution follows a hyperbolic distribution that fades away approximately 15 to 20 cm behind the hole. This signal shown in Figure 23 has also been used as a reference for determining the difference to the signal determined for a crack emanating from this hole. It can be observed that the crack must have a substantial size to be detectable at a relatively short distance to the hole and the crack itself. This following is a general guideline: A crack length of 2.5 times the diameter of a hole (rivet) has to be monitored no further away than 12.5 times the size of the hole (rivet) diameter. This shows that damage monitoring is a “hot spot affair” in all regards, and that state-of-the-art SHM may not currently be able to detect damage of a size smaller than traditional NDT can today.

Implementing sensors or generally transducers into a structure very often associates SHM to be equivalent to biology and nature, where millions of sensors are available as a system of nerves allowing a body’s health to be monitored. This association may be true in some regard. However a deeper look into the capabilities and functions of sensors in either biology or engineering provides a rather contradicting image when looking at the criteria summarized in Table 2 below.

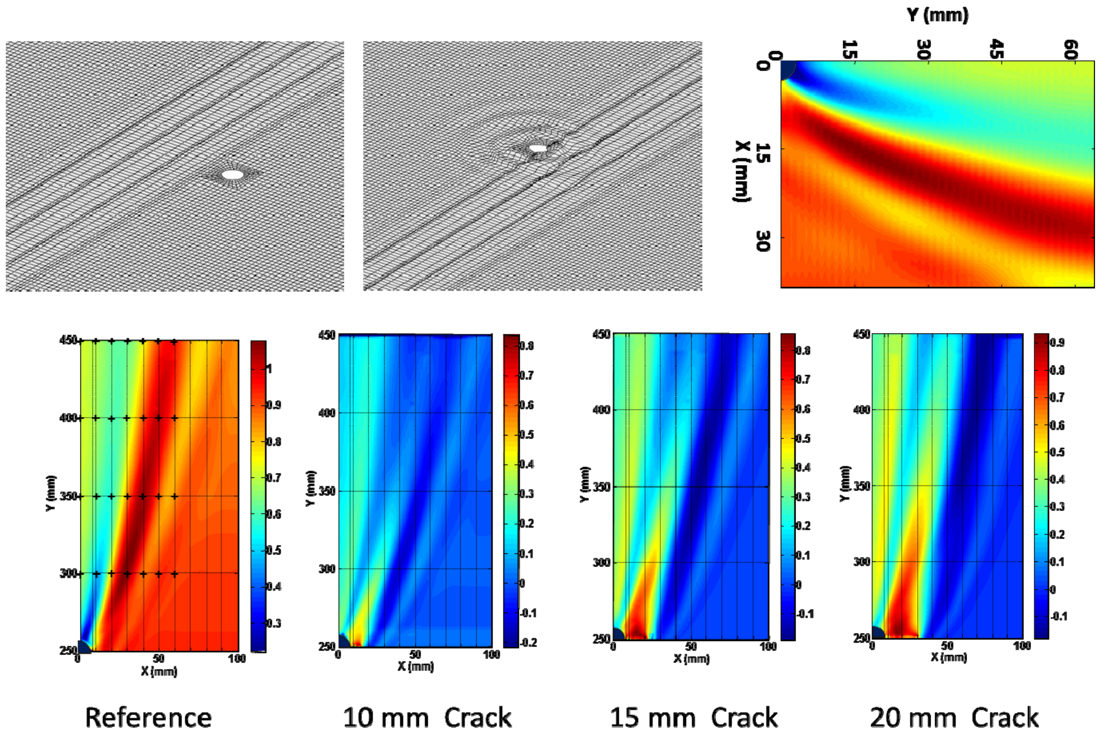


Figure 23: Influence of a hole in an infinite plate on the scattered wave behind the hole (top) and influence of a crack starting from the hole on the differential signal of the scattered wave (bottom).

Table 2: A comparison of sensing functions and capabilities in biology and engineering.

	Biology	***Engineering***
Number of sensors	Maximise	Minimise
Complexity	Low	High
Speed	Low	High
Redundancy	High	Low
Repairability	Self	Mechanical
Replacement	No	Yes
Programming	Self	Man made
Life	Undefined	Defined
Health management	Holistic	Holistic?

What needs to be concluded from here is that nature cannot be simply »copied« in terms of designing SHM into a structure but has rather to be used as a source for inspiration only.

6.0 REFERENCES

[1] Günther G, 1993, DASA (now Airbus Defence & Security) (private communication).

[2] Haibach E, 2006: *Betriebsfestigkeit;* Springer-Verlag, 3rd Edition (in German).

[3] Matsuishi M and T Endo, 1968: Fatigue of metals subjected to varying stress – Fatigue lives under random loading; Preliminary Proc. of JSME Kyushu District Meeting, pp. 37-40 (in Japanese).

[4] Fischer R, M Hück, H-G Köbler and W Schütz, 1970: *Eine dem stationären Gaußprozess verwandte Beanspruchungs-Zeit-Funktion für Betriebsfestigkeitsversuche;* Fortschr.-Ber. VDI-Z Reihe 5, Nr. 30 (in German).

[5] MIL-Handbook 5J: http://femci.gsfc.nasa.gov/links.html (last visited 03/12).

[6] Palmgren A, 1924: Die Lebensdauer von Kugellagern, VDI-Z 58, pp. 339-341 (in German).

[7] Miner M A, 1945: Cumulative damage in fatigue, J. Appl. Mech., 12, pp. 159-164.

[8] Tada H, P C Paris and G R Irwin, 1973: *The stress analysis of cracks handbook*; Del Research.

[9] Sih G C, 1973: *Handbook of stress intensity factors*, Lehigh University Press.

[10] Rooke D P and D.J. Cartwright, 1976: *Compendium of stress intensity factors*, Her Majesty's Stationery Office.

[11] Murakami Y, 1987; *Stress Intensity Factors Handbook*, Japan Soc. Mater. Sci., Pergamon Press, Tokyo/Japan.

[12] Paris P C and F Erdogan, 1963: A critical analysis of crack propagation laws; Trans. ASME, Series D, Vo. 85, pp 528-535.

[13] Elber W., 1971: The significance of fatigue crack closure. Damage tolerance in aircraft structures. ASTM STP 486, pp. 230-242.

[14] de Jonge J B, D Schütz, H Lowak and J Schijve, 1973: A standardized load sequence for flight simulation tests on transport aircraft wing structures; Nat. Aerospace Lab. (NLR) technical report TR 72018 and Fraunhofer LBF report FB-106.

[15] Schütz W, 1989: Standardized stress-time histories: An overview. Development of Fatigue Load Spectra, ASTM STP 1006, pp. 3-16.

[16] Boller C, F-K Chang and Y Fujino (Eds), 2009: *Encyclopedia of Structural Health Monitoring*; 5 Volumes; John Wiley & Sons.

[17] Boller C and W J Staszewski, 2003: *Aircraft Structural Health and Usage Monitoring;* in: Staszewski W J, C Boller and G R Tomlinson, 2003: Health Monitoring of Aerospace Structures; J. Wiley & Sons, pp. 29-73.

[18] Boller C and M R Mofakhami, 2008: *From Structural Mechanics to Inspection Processes: Getting Structural Health Monitoring into Application for Riveted Metallic Structures*; in B Dattaguru, S Gopalakrishnan and V K Aatre (Eds): IUTAM Symposium on Multi-Functional Material Structures and Systems; IUTAM Bookseries Vol 19, Springer, pp. 175-184.