CONSEIL INTERNATIONAL DES MACHINES A COMBUSTION

INTERNATIONAL COUNCIL ON COMBUSTION ENGINES

CO-ORDINATING WORKING GROUP

"CLASSIFICATION SOCIETIES – DIESEL"

(WG2)

Proposal by CIMAC WG4

16.10.2009

IACS UR M53, Appendix IV

Guidance for evaluation of

Fatigue Tests

Index

1. Introduction

Fatigue testing can be divided into two main groups; testing of small specimens and full size crank throws.

For crankshafts without any fillet surface treatment, the fatigue strength can be determined by testing small specimens taken from a full size crank throw. One advantage is the rather high number of specimens which can be then manufactured. Another advantage is that the tests can be made with different stress ratios (R-ratios) and/or different modes e.g. axial, bending and torsion, with or without a notch. This is required for evaluation of the material data to be used with critical plane criteria.

For crankshafts with surface treatment the fatigue strength can only be determined through testing of full size crank throws. For cost reasons this usually means a low number of crank throws. The load can be applied by hydraulic actuators in a 3- or 4-point bending arrangement, or by an exciter in a resonance test rig. The latter is frequently used, although it usually limits the stress ratio to $R = -1$.

Testing can be made using the staircase method or a modified version thereof which is presented in this document. Other statistical evaluation methods may also be applied.

2. Evaluation of test results

2.1. Principles

Prior to fatigue testing the crankshaft must be tested as required by quality control procedures, e.g. for chemical composition, mechanical properties, surface hardness, hardness depth and extension, fillet surface finish, etc.

The test samples should be prepared so as to represent the “lower end” of the acceptance range e.g. for induction hardened crankshafts this means the lower range of acceptable hardness depth, the shortest extension through a fillet, etc. Otherwise the mean value test results should be corrected with a confidence interval: a 90% confidence interval may be used both for the sample mean and the standard deviation.

The test results, when applied in M53, shall be evaluated to represent the mean fatigue strength, with or without taking into consideration the 90% confidence interval as mentioned above. The standard deviation should be considered by taking the 90% confidence into account. Subsequently the result to be used as the fatigue strength is then the mean fatigue strength minus one standard deviation.

If the evaluation aims to find a relationship between (static) mechanical properties and the fatigue strength, the relation must be based on the real (measured) mechanical properties, not on the specified minimum properties.

The calculation technique presented in Chapter 2.4 was developed for the original staircase method. However, since there is no similar method dedicated to the modified staircase method the same is applied for both.

2.2. Staircase method

In the original staircase method, the first specimen is subjected to a stress corresponding to the expected average fatigue strength. If the specimen survives 10^7 cycles, it is discarded and the next specimen is subjected to a stress that is one increment above the previous, i.e. a survivor is always followed by the next using a stress one increment above the previous. The increment should be selected to correspond to the expected level of the standard deviation.

When a specimen fails prior to reaching 10^7 cycles, the obtained number of cycles is noted and the next specimen is subjected to a stress that is one increment below the previous. With this approach the sum of failures and run-outs is equal to the number of specimens.

This original staircase method is only suitable when a high number of specimens are available. Through simulations it has been found that the use of about 25 specimens in a staircase test leads to a sufficient accuracy in the result.

2.3. Modified staircase method

When a limited number of specimens are available, it is advisable to apply the modified staircase method. Here the first specimen is subjected to a stress level that is most likely well below the average fatigue strength. When this specimen has survived 10^7 cycles, this **same** specimen is subjected to a stress level one increment above the previous. The increment should be selected to correspond to the expected level of the standard deviation. This is continued with the same specimen until failure. Then the number of cycles is recorded and the next specimen is subjected to a stress that is at least 2 increments below the level where the previous specimen failed.

With this approach the number of failures usually equals the number of specimens.
The number of run-outs, counted as the highest level where 10^7 cycles were reached, also equals the number of specimens.

The acquired result of a modified staircase method should be used with care, since some results available indicate that testing a runout on a higher test level, especially at high mean stresses, tends to increase the fatigue limit. However, this “training effect” is less pronounced for high strength steels (e.g. UTS > 800 MPa).

If the confidence calculation is desired or necessary, the minimum number of test specimens is 3.

2.4. Calculation of sample mean and standard deviation

A hypothetical example of tests for 5 crank throws is presented further in the subsequent text. When using the modified staircase method and the evaluation method of Dixon and Mood, the number of samples will be 10, meaning 5 run-outs and 5 failures, i.e.:

Number of samples, *n= 10*

Furthermore the method distinguishes between

Less frequent event is failures *C=1*
Less frequent event is run-outs *C=2*

The method uses only the less frequent occurrence in the test results, i.e. if there are more failures than run-outs, then the number of run-outs is used, and vice versa.

In the modified staircase method, the number of run-outs and failures are usually equal. However, the testing can be unsuccessful, e.g. the number of run-outs can be less than the number of failures if a specimen with 2 increments below the previous failure level goes directly to failure. On the other hand, if this unexpected premature failure occurs after a rather high number of cycles, it is possible to define the level below this as a run-out.

Dixon and Mood’s approach, derived from the maximum likelihood theory, which also may be applied here, especially on tests with few samples, presented some simple approximate equations for calculating the sample mean and the standard deviation from the outcome of the staircase test. The sample mean can be calculated as follows:

$$\overline{S_a} = S_{a0} + d \cdot \left(\frac{A}{F} - \frac{1}{2} \right) \qquad \text{when } C=1$$

$$\overline{S_a} = S_{a0} + d \cdot \left(\frac{A}{F} + \frac{1}{2} \right) \qquad \text{when } C=2$$

The standard deviation can be found by

$$s = 1.62 \cdot d \cdot \left(\frac{F \cdot B - A^2}{F^2} + 0.029 \right)$$

where

S_{a0} is the lowest stress level for the less frequent occurrence
d is the stress increment
$F = \Sigma f_i$
$A = \Sigma i \cdot f_i$
$B = \Sigma i^2 \cdot f_i$
i is the stress level numbering
f_i is the number of samples at stress level i

The formula for the standard deviation is an approximation and can be used when

$$\frac{B \cdot F - A^2}{F^2} > 0.3 \qquad \text{and} \qquad 0.5 \cdot s < d < 1.5 \cdot s$$

If any of these two conditions are not fulfilled, a new staircase test should be considered or the standard deviation should be taken quite large in order to be on the safe side.

If increment *d* is greatly higher than the standard deviation *s*, the procedure leads to a lower standard deviation and a slightly higher sample mean, both compared to values calculated when the difference between the increment and the standard deviation is relatively small. Respectively, if increment *d* is much less than the standard deviation *s*, the procedure leads to a higher standard deviation and a slightly lower sample mean.

Example 2.1. Hypothetical test results may look as shown in Figure 2.1. The processing of the results and the evaluation of the sample mean and the standard deviation are shown in Figure 2.2.

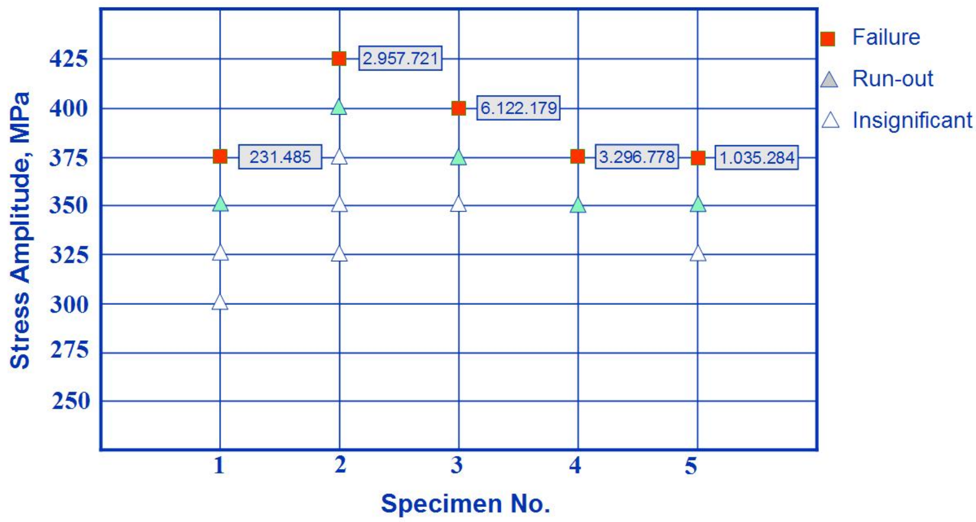
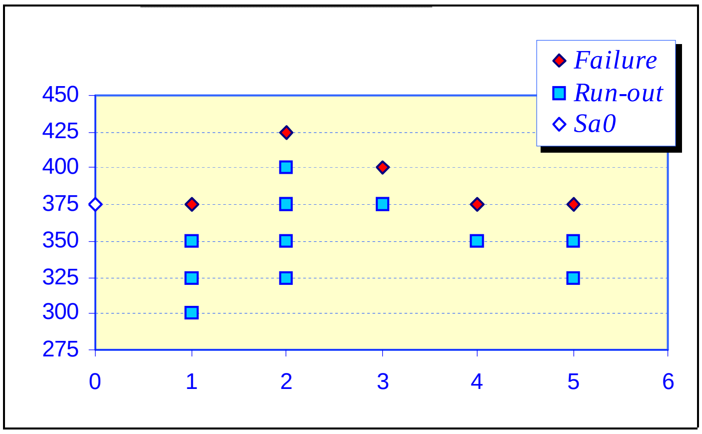


Figure 2.1. Log sheet of a modified staircase test.

Stress level 0, $S_{a0} := 375 \cdot MPa$

Stress increment, $d := 25 \cdot MPa$



i	fi	i*fi	i^2*fi
2	1	2	4
1	1	1	1
0	3	0	0
Σ	5	3	5
	F	***A***	***B***

$F = \sum f_i$

$A = \sum i \cdot f_i$

$B = \sum i^2 \cdot f_i$

i = 0,1,2,... *is the stress level numbering, the numbering ususally starts from zero*

f_i *is number of test specimen at stress level, i*

Level 0 *is the lowest value of the less frequent occurrence in the test results.*

$F := 5$

$A := 3$

$B := 5$

Sample mean

Sample mean,

$$S_a := \begin{cases} S_{a0} + d \cdot \left(\frac{A}{F} - \frac{1}{2}\right) & \text{if } C = 1 \\ S_{a0} + d \cdot \left(\frac{A}{F} + \frac{1}{2}\right) & \text{if } C = 2 \end{cases}$$

$S_a = 377.5\ MPa$

Standard deviation

Sample standard deviation,

$$s := 1.62 \cdot d \cdot \left(\frac{B \cdot F - A^2}{F^2} + 0.029\right)$$

$s = 27.09\ MPa$

Standard deviation ratio,

$$s_r := \frac{s}{S_a}$$

$s_r = 0.072$

Figure 2.2. Processing of the staircase test results.

2.5. Confidence interval for mean fatigue limit

If the staircase fatigue test is repeated, the sample mean and the standard deviation will most likely be different from the previous test. Therefore, it is necessary to assure with a given confidence that the repeated test values will be above the chosen fatigue limit by using a confidence interval for the sample mean.

The confidence interval for the sample mean value with unknown variance is known to be distributed according to the *t*-distribution (also called *student's t-distribution*) which is a distribution symmetric around the average.

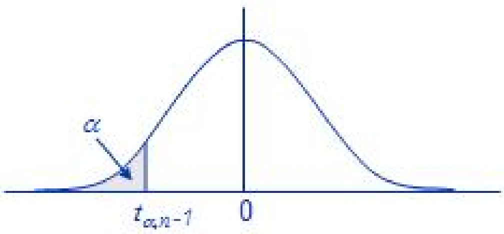


Figure 2.3. Student's t-distribution

The confidence level normally used for the sample mean is 90 %, meaning that 90 % of sample means from repeated tests will be above the value calculated with the chosen confidence level. The figure shows the *t*-value for (1 - α) · 100% confidence interval for the sample mean.

If S_a is the empirical mean and *s* is the empirical standard deviation over a series of *n* samples, in which the variable values are normally distributed with an unknown sample mean and unknown variance, the (1 - α) · 100% confidence interval for the mean is:

$$P\left(S_a - t_{\alpha,n-1} \cdot \frac{s}{\sqrt{n}} < S_{aX\%}\right) = 1 - \alpha$$

The resulting confidence interval is symmetric around the empirical mean of the sample values, and the lower endpoint can be found as;

$$S_{aX\%} = S_a - t_{\alpha,n-1} \cdot \frac{s}{\sqrt{n}}$$

which is the mean fatigue limit (population value) to be used to obtain the reduced fatigue limit where the limits for the probability of failure are taken into consideration.

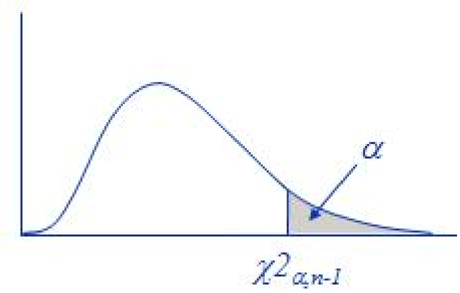
Example 2.2. Applying a 90 % confidence interval (α = 0.1) and n = 10 (5 failures and 5 run-outs) leads to $t_{\alpha,n-1}$ = 1.383, taken from a table for statistical evaluations (E. Dougherty: Probability and Statistics for the Engineering, Computing and Physical Sciences, 1990. Note that $v = n - 1$ in the tables.). Hence:

$$S_{a90\%} = S_a - 1.383 \cdot \frac{s}{\sqrt{n}} = S_a - 0.4373 \cdot s$$

To be conservative, some authors would consider n to be 5, as the physical number of used specimen, then $t_{\alpha,n-1}$ = 1.533.

2.6. Confidence interval for standard deviation

The confidence interval for the variance of a normal random variable is known to possess a chi-square distribution with *n - 1* degrees of freedom.



The confidence level on the standard deviation is used to ensure that the standard deviations for repeated tests are below an upper limit obtained from the fatigue test standard deviation with a confidence level. The figure shows the chi-square for (1 - α) · 100% confidence interval for the variance.

Figure 2.4. Chi-square distribution.

An assumed fatigue test value from *n* samples has a normal random variable with a variance of σ^2 and has an empirical variance s^2. Then a (1 - α) · 100% confidence interval for the variance is:

$$P\left(\frac{(n-1)s^2}{\sigma^2} < \chi^2_{\alpha,n-1}\right) = 1-\alpha$$

A (1 – σ) · 100% confidence interval for the standard deviation is obtained by the square root of the upper limit of the confidence interval for the variance and can be found by

$$s_{X\%} = \sqrt{\frac{n-1}{\chi^2_{\alpha,n-1}}} \cdot s$$

This standard deviation (population value) is to be used to obtain the fatigue limit, where the limits for the probability of failure are taken into consideration.

Example 2.3. Applying a 90% confidence interval (α = 0.1) and n = 10 (5 failures and 5 run-outs) leads to $\chi^2_{\alpha,n-1}$ = 4.168, taken from a table for statistical evaluations (E. Dougherty: Probability and Statistics for the Engineering, Computing and Physical Sciences, 1990). Hence:

$$s_{90\%} = \sqrt{\frac{n-1}{4.168}} \cdot s = 1.47 \cdot s$$

To be conservative, some authors would consider *n* to be 5, as the physical number of the used specimen, then $\chi^2_{\alpha,n-1}$ = 1.064.

3. Small specimen testing

In this connection a small specimen is considered to be one of the specimens taken from a crank throw. Since the specimens shall be representative for the fillet fatigue strength, they should be taken out close to the fillets, as shown in Figure 3.1.

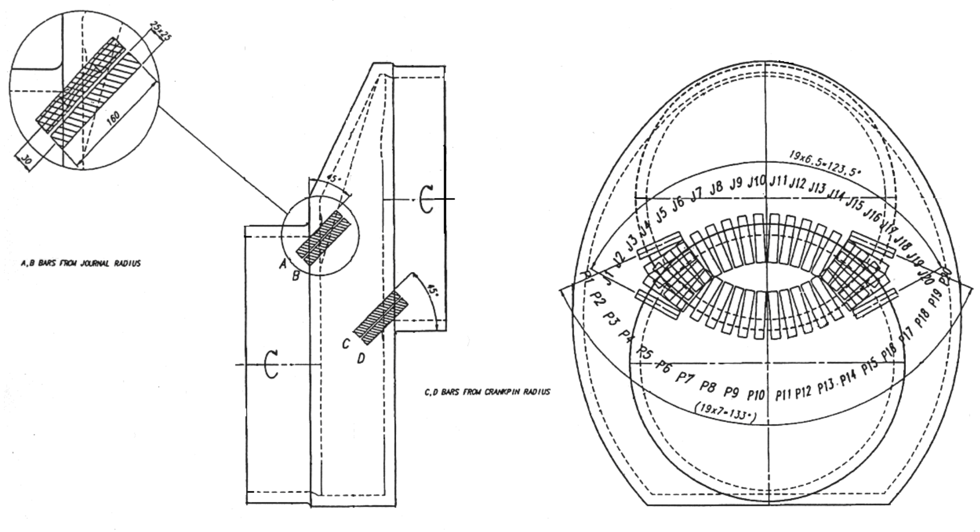


Figure 3.1. Specimen locations in a crank throw.

The (static) mechanical properties are to be determined as stipulated by the quality control procedures.

3.1. Determination of bending fatigue strength

It is advisable to use un-notched specimens in order to avoid uncertainties related to the stress gradient influence. Push-pull testing (stress ratio R = -1) is preferred, but for the purpose of critical plane criteria another stress ratio may be added.

A. If the objective of the testing is to document the influence of high cleanliness, test samples taken from positions approximately 120 degrees in a circumferential direction may be used. See Figure 3.1.
B. If the objective of the testing is to document the influence of continuous grain flow (cgf) forging, the specimens should be restricted to the vicinity of the crank plane.

3.2. Determination of torsional fatigue strength

A. If the specimens are subjected to torsional testing, the selection of samples should follow the same guidelines as for bending above. The stress gradient influence has to be considered in the evaluation.
B. If the specimens are tested in push-pull, the samples should be taken out at an angle of 45 degrees to the crank plane. When taking the specimen at a distance from the (crank) middle plane of the crankshaft along the fillet, this plane rotates around the pin centre point making it possible to resample the fracture direction due to torsion (the results are to be converted into the pertinent torsional values.)

3.3. Other test positions

If the test purpose is to find fatigue properties and the crankshaft is forged in a manner likely to lead to cgf, the specimens may also be taken longitudinally from a prolonged shaft piece where specimens for mechanical testing are usually taken. The condition is that this prolonged shaft piece is heat treated as a part of the crankshaft and that the size is so as to result in a similar quenching rate as the crank throw.

When using test results from a prolonged shaft piece, it must be considered how well the grain flow in that shaft piece is representative for the crank fillets.

3.4. Correlation of test results

When using the bending fatigue properties from tests mentioned in this section, it should be kept in mind that successful continuous grain flow (cgf) forging leading to elevated values compared to other (non cgf) forging (or empirical approaches), will normally not lead to a torsional fatigue strength improvement of the same magnitude. In such cases it is advised to either carry out also torsional testing or to make a conservative assessment of the torsional fatigue strength, e.g. by using no credit for cgf. This approach is applicable when using the Gough Pollard criterion. However, this approach is not recognised when using the von Mises or a multi-axial criterion such as Findley.

If the found ratio between bending and torsion fatigue differs from $\sqrt{3}$, one should consider replacing the use of the von Mises criterion with the Gough Pollard criterion. Also, if critical plane criteria are used, it must be kept in mind that cgf makes the material inhomogeneous in terms of fatigue strength, meaning that the material parameters differ with the directions of the planes.

Any addition of influence factors must be made with caution. If for example a certain addition for super clean steel is documented, it may not necessarily be fully combined with a *K*-factor for cgf. Direct testing of samples from a super clean and cgf forged crank is preferred.

4. Full size testing

4.1. Hydraulic pulsation

A hydraulic test rig can be arranged for testing a crankshaft in 3-point or 4-point bending as well as in torsion. This allows for testing with any *R*-ratio.

Although the applied load should be verified by strain gauge measurements on plain shaft sections for the initiation of the test, it is not necessarily used during the test for controlling load. It is also pertinent to check fillet stresses with strain gauge chains.

Furthermore, it is important that the test rig provides boundary conditions as defined in Appendix III (section 3.1 to 3.3).

The (static) mechanical properties are to be determined as stipulated by the quality control procedures.

4.2. Resonance tester

A rig for bending fatigue normally works with an *R*-ratio of -1. Due to operation close to resonance, the energy consumption is moderate. Moreover, the frequency is usually relatively high, meaning that 10^7 cycles can be reached within some days. Figure 4.1 shows a layout of the testing arrangement.

The applied load should be verified by strain gauge measurements on plain shaft sections. It is also pertinent to check fillet stresses with strain gauge chains.

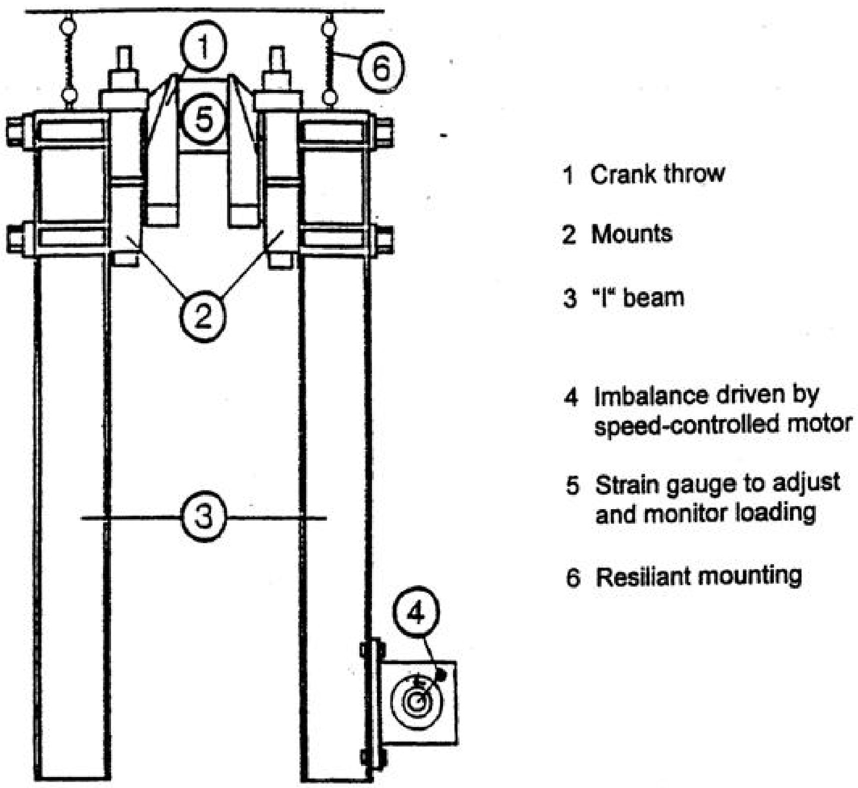


Figure 4.1. A testing arrangement of the resonance tester for bending loading.

Clamping around the journals must be arranged in a way that prevents severe fretting which could lead to a failure under the edges of the clamps. If some distance between the clamps and the journal fillets is provided, the loading is consistent with 4-point bending and thus representative for the journal fillets also.

In an engine the crankpin fillets normally operate with an *R*-ratio slightly above -1 and the journal fillets slightly below -1. If found necessary, it is possible to introduce a mean load (deviate from *R* = -1) by means of a spring preload.

A rig for torsion fatigue can also be arranged as shown in Figure 4.2. When a crank throw is subjected to torsion, the twist of the crankpin makes the journals move sideways. If one single crank throw is tested in a torsion resonance test rig, the journals with their clamped-on weights will vibrate heavily sideways.

This sideway movement of the clamped-on weights can be reduced by having two crank throws, especially if the cranks are almost in the same direction. However, the journal in the middle will move more.

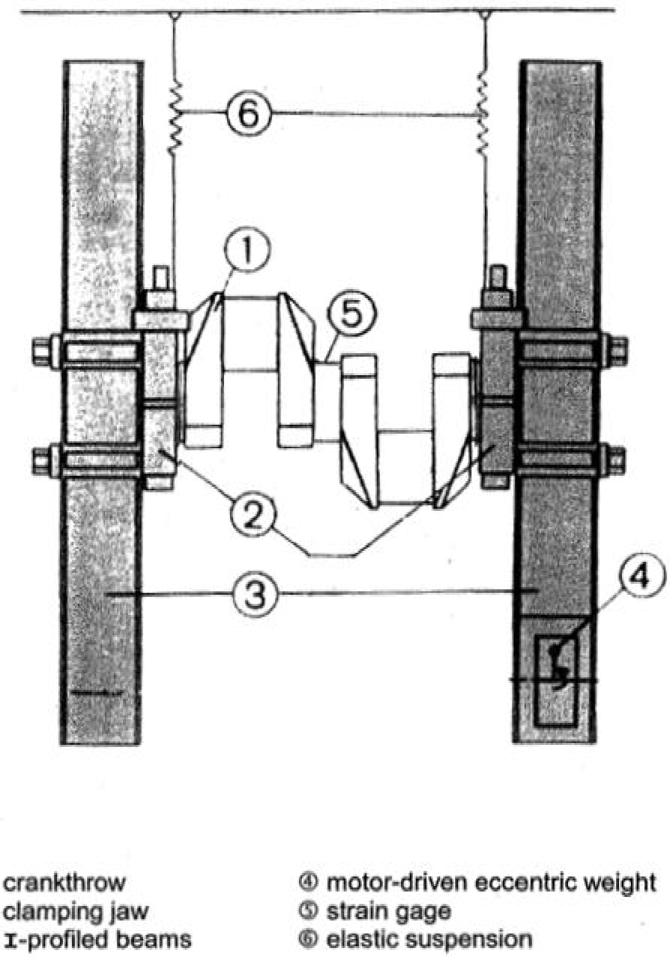


Figure 4.2. A testing arrangement of the resonance tester for torsion loading.

Since sideway movements can cause some bending stresses, the plain portions of the crankpins should also be provided with strain gauges arranged to measure any possible bending that could have an influence on the test results.

Similarly to the bending case the applied load shall be verified by strain gauge measurements on plain shaft sections. It is also pertinent to check fillet stresses with strain gauge chains as well.

5. Use of existing results for similar crankshafts

For fillets or oil bores without surface treatment, the fatigue properties found by testing may be used for similar crankshaft designs providing:

- The same material type
- Cleanliness on the same or better level
- The same mechanical properties can be granted (size versus hardenability)
- Similar manufacturing process

Induction hardened or gas nitrited crankshafts will suffer fatigue either at the surface or at the transition to the core. The surface fatigue strength as determined by fatigue tests of full size cranks, may be used on an equal or similar design as the tested crankshaft when the fatigue initiation occurred at the surface. With the similar design it is meant that the same material type and surface hardness are used and the fillet radius and hardening depth are within approximately ± 30 % of the tested crankshaft.

Fatigue initiation in the transition zone can be either subsurface, i.e. below the hard layer, or at the surface where the hardening ends. The fatigue strength at the transition to the core can be determined by fatigue tests as described above, provided that the fatigue initiation occurred at the transition to the core. Tests made with the core material only will not be representative since the tension residual stresses at the transition are lacking.

For cold rolled or stroke peened crankshafts, the results obtained by one full-size crank test can be applied to another crank size, provided that the base material is of the same type and that the treatment is made in order to obtain the same level of compressive residual stresses at the surface as well as in the depth. This means that both the extension and the depth of rolling or peening must be proportional to the fillet radius.

It has to be noted also what some recent research has shown: The fatigue limit can decrease in the very high cycle domain with subsurface crack initiation due to trapped hydrogen that accumulates through diffusion around some internal defect functioning as an initiation point. In these cases it would be appropriate to reduce the fatigue limit by some percent per decade of cycles beyond 10^7. Based on a publication by Yukitaka Murakami "Metal Fatigue: Effects of Small Defects and Non-metallic Inclusions" the reduction is suggested to be 5 % per decade.