Tooth Flank Fracture – A critical failure mode
Influence of Macro and Micro Geometry

Dr. Stefan Beermann & Dr. Ulrich Kissling, KISSsoft AG

Introduction
Tooth flank fracture is an infrequently seen failure mode observed on cylindrical and bevel gears, especially, but not exclusively, on case-carburized gears. The characteristic of the damage is a crack initiation in the region of the active contact area and inside the material, with a crack propagation to the surface of the loaded flank and in the direction of the root area on the other side of the tooth. Most of the cases mentioned in literature occurred on gears with a larger module.

Since 2014, the ISO/TC 60 workgroup 6 started working on an ISO standard, ISO DTS 19042-1 [1]. This standard defines a calculation method to predict the risk of flank fracture. It is based on research mainly conducted at the FZG in Munich, with the work of Dr. Witzig as the current culmination point. So far, a second draft was issued for comments. The intention of this paper is to provide some results from the calculations according to the defined method to be utilized for the discussion of the standard. For this, a sensitivity study was conducted by systematically varying gear parameters, such as pressure angle, helix angle, normal module, and also modifications (tip relief and lead crowning).

The first part of the paper defines the failure mode investigated and briefly introduces the method set forth in ISO DTS 19042 [1]. It also mentions two points that could be improved for the final standard. In a second part, the results of the parameter variation are presented.

Failure Modes
Gears are loaded with torque, which causes a force wandering over the flank with a mixture of rolling and sliding. This can lead to different failure modes of the teeth. To better define what type of failure mode is treated in this paper and distinguish it from other types of failure, here, the most common ones are briefly described and compared to tooth flank fracture (TFF).

TFF describes a fracture of the tooth in the used area of the flank. The pressure on the flank causes crack initiation below the surface, typically at a preexisting material fault. The crack grows in two directions, into the core of the tooth and towards the surface of the flank. Secondary cracks might occur that are oriented in an angle to the original crack. Typically, as soon as the crack is visible, the tooth breaks.

Fracture of a tooth may occur in the root area. This is either an immediate failure due to high overload, whereas TFF is a fatigue failure, or a crack starts from the surface in the root area and then propagates into the material over many load cycles until the root breaks. The crack is visible before the gear fails and special ink can be used to find cracks before the gear totally fails. TFF is not caused by a crack starting at the surface, but the crack starts inside the material. Also, if TFF occurs, the crack is in the used area of the flank, not in the root area.

TFF is caused by the flank pressure. The same can also cause macropitting, where particles break out of the surface, or micropitting, which is caused by a too-small lubrication film thickness, so that the metal surface gets in direct contact. This causes small pits, which are not directly visible to the naked eye but give the surface a grey appearance. If high pressure is applied and/or high temperature occurs in the contact, the two flank surfaces can be welded together. With the gears continuing to move, the two flanks are then torn apart again, leaving typical scratches on the flank. This is called scuffing. Opposite of all three of these phenomena, TFF is not a damage of the surface but is caused by a crack underneath the surface.

Other damages that can occur on gear flanks are spalling or case crushing, where larger parts of the surface come off. Here, the hardened case is separated from the softer core due to high contact stress. The difference to tooth flank fracture is that the damage is limited to the surface, whereas in the case of tooth flank fracture, a crack grows through the whole tooth.
Several calculation methods are available to assess the risk of occurrence of the above failures. For the tooth root fracture due to fatigue, the most common one is in ISO 6336-3 [2]. For the static tooth root fracture, so far there is no standard; first approaches can be found in the VDI guideline 2736 for plastic gears. The flank pitting is covered by ISO 6336-2 [2], micropitting by ISO TR 15144, and scuffing by ISO TR 13989.

For tooth flank fracture, a new standard is being developed, ISO DTS 19042 [1], so far only available as a draft. It is in the very early stages; however, the outline of the method itself is already quite fixed. It is based on the work from Witzig [3], which itself goes back to Annast [4] and Bruckmeier [5]. Oster [6] was inventing an improved model for the understanding of the mechanism causing damage due to the rolling contact of the two flank surfaces.

Figure 1 – Fracture of a tooth with initial crack on the right side, inside the material

Why Flanks Break

Figure 1 shows a typical fracture image (as published in Witzig [3]) of a tooth that failed due to flank fracture. The origin of the crack is deep in the material and not directly under the surface. In most cases, the crack is initiated by a fault in the metal structure or a non-metallic inclusion, such as manganese sulfide or aluminum oxide. In the beginning, the crack grows in vacuum. The growth into the softer core is faster than into the case. Annast [4] explains this with the residual stresses due to the heat treatment: in the case, the residual stresses are negative (compressive stress). This reduces the effect of the outer stresses, which have to work against the compressive stress keeping the crack closed. Damage in this situation is caused by a combination of mode I (bending) and mode II (shearing) of crack propagation according to Bold [8] (see Figure 2): due to the bending, the crack opens and the friction inside the crack is reduced so that the shearing can widen it. This is the reason why the crack is growing slower in the hard case. In the core, the residual stress is tensile stress (positive), increasing the effect of the outer stresses and causing the crack to grow in mode II, due to shearing.

Figure 2 – Crack propagation modes I, II, and III according to Bold [8]
After a first crack initiation, the crack grows at an angle of about 40–50 degrees to the surface and in the opposite direction, through the core to the back flank. Finally, the tooth breaks with the typical appearance of forced rupture in the crack area towards the back flank.

![Figure 3 – Tooth showing typical flank fracture](image)

**Figure 3 – Tooth showing typical flank fracture**

Bibliography (ISO, 2014)

The typical picture of a crack of the tooth is presented in Figure 3. It is obvious that a large chunk of the tooth is missing. This is due to a secondary crack growing from the surface towards the primary crack. This crack stops when it reaches the primary crack. Witzig showed with FE simulation that the internal primary crack induces an increase of stress close to the surface, underneath the primary crack, and also far below the current contact point (see Figure 4). This stress then leads to the secondary crack.

![Figure 4 – FEM by Witzig to explain the development of secondary cracks](image)

**Figure 4 – FEM by Witzig to explain the development of secondary cracks**

Bibliography

Further investigations by Witzig showed that the development of secondary cracks can be postponed or inhibited by increasing the quality of the flank surface. Witzig applied slide grinding to test gears, thus reducing the surface roughness Ra to below 0.15 μm. However, this improved surface quality could not prevent flank fracture itself, only postpone it. This fits to the model of a crack initiation significantly below the surface.

Summarizing the effect of flank fracture occurs if the material below the surface is exposed to stresses above the permissible limits. The crack is typically initiated at a fault in the material and then grows to the outside until ultimate failure.

**Prediction of Flank Fracture**

To assess the risk of flank fracture, the effective material exposure is calculated and then compared to permissible values. First approaches were concentrating on individual stresses (nominal stress) and yield strength. Witzig and others tried an integrated way, taking all effects causing stresses into account. In the end, however, he also proposed a simplified method for standardization purposes.
The approach used by the ISO method is to calculate a material exposure, taking multiple stresses into account, including the residual stresses. Basically, the local shear stresses are calculated and then divided by the local permissible shear stresses based on the local hardness of the material. So in opposite of, for example, the pitting strength or the root fracture method, this is a highly local method that checks the situation for multiple meshing points and for multiple points into the depth of the material.

For the local hardness, a formula is provided in ISO 19042 to calculate the hardness at depth $y$. The parameters of this formula are the core hardness $HV_{core}$, the surface hardness $HV_{surface}$ and the case-hardened depth (CHD = the depth where the hardness is $HV_{550}$).

The local hardness is calculated according to:

$$HV(y^*) = HV_{core} + (HV_{surface} - HV_{core})(f(y^*))$$

(1)

with a given function $f(y^*)$ modelling the course of the hardness into the depth $y$ of the material, using the dimensionless variable $y^* = y / CHD$.

As a curiosity, this formula needs the CHD value as an input. However, at the depth $y$, the calculated local hardness only depends on $HV_{core}$ and $HV_{surface}$, and in most cases deviates from $HV_{550}$. Since $f(1) = 0.5$, we get for $y = CHD$:

$$HV\left(\frac{CHD}{CHD}\right) = HV(1) = HV_{core} + (HV_{core} - HV_{surface})(f(1)) \approx 0.5(HV_{core} + HV_{surface})$$

(2)

This only gives the value $HV_{550}$ if $HV_{core} + HV_{surface} = HV_{1100}$.

Based on the local hardness, the local permissible shear stress is defined:

$$\tau_{per,CP}(y) = \left(\frac{K_{t,per}}{K_{material}}\right)(HV(y))$$

(3)

Here $K_{t,per}$ is a conversion factor which is fixed to 0.4. The factor $K_{material}$ is the material factor. Witzig was defining this factor to take the different properties of steels into account, caused by the heat treatment. In the ISO standard, this was reduced to the choice between material with at least 800 N/mm² ultimate strength, or at least 900 N/mm² ultimate strength.

As additional parameter influencing the material factor, the tooth thickness is taken into account (see Table 1). Since the material factor is defined constant for intervals of the tooth thickness, this can lead to discontinuities in the permissible shear stress—and therefore also in the results—if the tooth thickness is varied (e.g., by varying the profile shift, the helix angle, or the module). In Figure 5, the normal module was varied. The solutions with normal module between 4 and 6mm and the solutions with normal module larger than 6mm are separated with respect to the safety against flank fracture due to the change in $K_{material}$ when the tooth thickness at the critical point exceeds 10mm. Our proposal here is to amend the standard so that the material factor is interpolated and given as a continuous function over the tooth thickness instead of defining discrete values.

Since the two factors are constant for few discrete values, the permissible shear stress is basically defined by the course of the hardness into the material.

**Table 1 – Definition of $K_{material}$ for case-carburized steel according to ISO 19042, with tensile strength $Rm$ and transverse tooth thickness at point of single tooth contact $st,B-D$**

<table>
<thead>
<tr>
<th>Km</th>
<th>Tooth thickness in mm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3 &lt; st,B-D ≤ 10</td>
</tr>
<tr>
<td>Rm min. 800 N/mm²</td>
<td>1.00</td>
</tr>
<tr>
<td>Rm min. 900 N/mm²</td>
<td>1.13</td>
</tr>
</tbody>
</table>
The local effective stress is calculated by:

\[
\tau_{\text{eff},\text{CP}}(y) = \tau_{\text{eff},\text{L,CP}}(y) - \Delta\tau_{\text{eff},\text{L,RS,CP}}(y) - \tau_{\text{eff,RS}}(y)
\]  

(4)

Where \(\tau_{\text{eff},\text{L,CP}}(y)\) is the local stress without the influence of the residual stresses calculated, based on the so-called dynamic pressure \(p_{\text{dyn}}\). This is the Hertzian pressure calculated according to ISO 15144, multiplied with the square root of the dynamic factor \(K_v\), application factor \(K_A\), and the load distribution factors \(K_{\text{HF}}, K_{\text{HB}}\) from ISO 6336. The value \(\Delta\tau_{\text{eff},\text{L,RS,CP}}(y)\) is the influence of the residual stresses on the total effective stress. And \(\tau_{\text{eff,RS}}(y)\) finally is the quasi-stationary residual stress. So the main influence for \(\tau_{\text{eff},\text{L,CP}}(y)\) is the Hertzian pressure. Further explanations how these stresses are determined are found in [3].

For each contact point investigated, and for a sufficient number of points into the depth of the material, the material exposure:

\[
A_{\text{FF,CP}}(y) = \frac{\tau_{\text{eff,CP}}(y)}{\tau_{\text{per}}(y)} + c_1
\]  

(5)

is calculated. The constant \(c_1\) is the material exposure calibration factor. For case-carburized gears, it is small \((c_1=0.04)\). For other treatment of the gears, the method is not applicable, so there is no definition of \(c_1\) for any other than case-carburized gears. Out of this field of values, the maximum material exposure is extracted:

\[
A_{\text{FF,max}} = \max_{\text{CP},y} A_{\text{FF,CP}}(y)
\]  

(6)
The safety factor against tooth flank fracture is then defined as:

\[ S_{FF} = \frac{1}{\lambda_{FF,max} + c_2} + c_2 \]  

with \( c_2 \) as safety calibration factor. This factor \( c_2 \) is intended to dampen the results so that the safety is not too large when the material exposure is small. The factor is defined as \( c_2 = 0.2 \).

**Variation of Macro Geometry**

The following investigation targets on getting a “feeling” for the behavior of the new method during practical gear design work. It has to be mentioned here that the investigation only covers the results of the method and makes no general statement about real gears.

For the research, the fine sizing module of KISSsoft was used. This software module allows one to vary macro parameters such as module, center distance, pressure angle, or helix angle in a systematic way. The fine sizing varies the parameters as defined by the user, then shows all possible combinations of number of teeth and a variation of the profile shift that leads to gear geometries within the specified boundary conditions. Parameter sets that are not giving correct gear geometry are sorted out automatically. Since this function always gives a variety of solutions which are all practical, tendencies can be identified in an easy and reliable manner.

The first experiment varies the pressure angle of a spur gear with fixed module, center distance, and transmission ratio (\(+/- 4\%\)). The number of teeth and the profile shift are varied here. The range of the pressure angle is 15° to 30°.

The result is presented in Figure 6 and shows a clear trend to higher safety against flank fracture for smaller pressure angles. This is against the trend of the Hertzian pressure, which is reciprocal to the pressure angle. Since the risk of pitting is growing with increased pressure, the risk of flank fracture is hence inverse to the pitting behavior (see below).

![Figure 6 – Influence of the pressure angle on the safety against flank fracture](image)

To investigate a bit further—in a separate series of calculations, the pressure angle was varied from 15 to 30 degrees, without changing anything else. Figure 7 shows the Hertzian pressure as calculated for the pitting safety factor, along with the safety factor itself. The third curve is the product of the two values. This product is more or less constant, thus proving that the safety factor against pitting is inversely proportional to the Hertzian pressure.
In contrast to this result, Figure 8 shows the situation for the safety against flank fracture. Here the product of dynamic pressure and safety factor is changing strongly, so the influence of the flank pressure is not so strong for the flank fracture. Other factors are more important.

This observation fits to results from Witzig [3]. His explanation is that for smaller pressure angle due to the higher curvature, the Hertzian pressure is larger. However, the maximum of the pressure is closer to the surface where the hardness is higher, thus overall reducing the risk of flank fracture.

A second corresponding experiment varies the helix angle instead of the pressure angle (see Figure 9). Again, the trend is just the opposite of the trend from pitting safety: a smaller helix angle reduces the risk of flank fracture. In the cited literature, the influence of the helix angle was not part of the research. So we can currently only accept this fact as resulting from the modeling in the method. A possible explanation is
that the pressure angle is increased by a larger helix angle, so the effect may be indirect due to the influence of the pressure angle as shown above.

Figure 9 – Influence of the helix angle on the safety against flank fracture

A next experiment is to vary the module. Since the center distance is fixed, this automatically means a larger change in number of teeth as well. The first variant presented in Figure 10 shows the results with varying module but fixed hardness depth CHD. There is no clear trend for the flank fracture safety. Although the Hertzian pressure is reduced if the module gets smaller, the safety against flank fracture is not really benefiting from this. The largest achievable safety factor is the same for the full range of the module.

Figure 10 – Safety factors over Hertzian pressure for varying normal module, constant CHD

The situation is completely different if the hardness depth CHD is made dependent on the module—in the manufacture of a gear, the hardness depth CHD is usually defined based on rules, where CHD depends on the module. Therefore, using a different CHD for a different module is realistic. For the experiment, we defined the CHD according to $E_{H \text{opt}}$ in Figure 17 of ISO 6336-5 (see Figure 11). When CHD is depending on the module, then the larger module leads to significantly higher safety (see Figure 11).
The results in Witzig [3] are at least not contradictory: Witzig tried two different modules, 3mm and 5mm, with a hardness depth of 0.5mm and 0.6mm, respectively. So the increase of CHD was not proportional to the increase of the module. Witzig found a strong shift of the measured Wöhlerline to about 30% higher torque. On the other hand, an increase of the hardness depth to 0.7mm on the module 3mm gear did not give an increased load capacity. However, due to the relatively low number of experiments in this field, the picture is not very clear.
Micro Geometry

KISSsoft also allows us to vary the micro geometry (i.e., profile and lead modifications) and perform a contact analysis for each variant. So, it is natural to investigate the influence of micro geometry on flank fracture as well. The term “micro geometry” summarizes the applied flank and profile modifications. These modifications are typically applied to improve the situation for the load distribution (e.g., to avoid or minimize edge contact). Although the design of the micro geometry is a very interesting topic in itself, this paper only concentrates on analyzing the sensitivity of the method investigated to changes in micro geometry.

In a first test, a long linear tip modification was applied to a helical gear set. The amount of the modifications ranged from 0μm to 100μm for module 4.5mm gears. In Figure 13, the smaller safety factor against flank fracture of the two gears is displayed over the amount of tip relief on gear 1 on the x axis and the respective value on gear 2 on the y axis.

![Figure 13](image)

**Figure 13 – Influence of linear tip relief on safety against flank fracture for a helical gear**

Obviously, there is an optimum for the modifications with $C_{a1,\text{opt}}=30 \, \mu\text{m}$ and $C_{a2,\text{opt}}=25 \, \mu\text{m}$. That the modifications have a positive influence is not so clear at all—a tip relief has an effect when two pairs of teeth are in contact. Typically, however, the highest risk for flank fracture is closer to the operating pitch point, usually at one of the points of single tooth contact. In the example above, the course of the Hertzian pressure shows significant higher stress at the beginning and the end of the contact (see Figure 14). The reason is mainly the buttressing effect and the impact shock at the beginning of the contact.
Figure 14 – Hertzian pressure in one cross-section over the meshing for a helical gear set

In Figure 15, the same experiment was carried out for a spur gear set. Here, the missing of the buttressing effect leads to high pressure only in the area of single tooth contact. In Figure 16, the blue curve shows the Hertzian pressure over the meshing without modifications, and the red curve the same, with 100μm modifications on both gears, leading to a high peak at the point of single tooth contact (which in the end means that 100μm is too large a value). The profile modifications have no effect, or a negative one, by shifting the maximum pressure into less advantageous areas of the flank if the value of the modification is chosen too large.

Figure 15 – Influence of linear tip relief on safety against flank fracture for a helical gear
Figure 16 – Hertzian pressure in one cross-section over the meshing for a spur gear set

For lead modification, the situation is more simple: for an ideal gear set without any deviations due to manufacturing and no misalignment of the axes, any modification in the lead direction, such as lead crowning or helix angle modification, will concentrate the load to a smaller fraction of the face width, thus increasing the local stress and therefore reducing all safety factors. Figure 17 shows this situation with crowning on both gears varied from 0 μm to 50 μm. Obviously, the best solution is the one without crowning.

Figure 17 – Influence of crowning on the safety against flank fracture for an ideal gear set

However, since we are not in an ideal world, misalignments and manufacturing tolerances apply. In Figure 18, the same gear set is subject to a misalignment of the shafts of about 25 μm. The variant without crowning has a smaller safety against flank fracture because of the load concentration on one side of the flank. With the crowning applied, the misalignment can be compensated to a certain extent, so the optimal crowning in this case is about 10 μm on both flanks. Still, the safety is slightly lower than in the ideal case, but better than without crowning.
Figure 18 – Influence of crowning on the safety against flank fracture for a misaligned gear set

Flank Fracture versus Pitting

The risk of pitting is mainly dictated by the Hertzian pressure. As seen above, this is not true for the flank fracture. Although dependent on the Hertzian pressure, there are other factors having a stronger influence than the pressure if center distance, transmission ratio, and torque are fixed values—which usually is the case. Figure 19 shows the safety factors against flank fracture and the pitting safety when varying the pressure angle. The trend is clear—a smaller pressure angle favors the safety against flank fracture, whereas a larger pressure angle reduces the risk of pitting. So if the gear set is of the kind that has a risk of flank fracture, a compromise might be necessary to level the two risks.

Figure 19 – Safety against flank fracture over safety against pitting for different pressure angles
Proposed Measures to Improve Safety against Flank Fracture

With the trends shown above, two rules can be derived if the need to increase the safety against flank fracture arises:

1) A smaller pressure angle leads to a higher concentration of the Hertzian pressure. This is negative for the surface (pitting) but leads to a stress concentration in the material areas with high hardness, thus reducing the material exposure further into the depth of the material.

2) A larger module along with an increase of the hardness depth (CHD) reduces the local material exposure by extending the hardened case further to the core. However, a larger module comes along with a reduction in the number of teeth, which is making the situation worse, concerning pitting.

Note that although for both rules there is some explanation from the mechanical theory and measurements conducted by Witzig and others, the basis used for our statements is a calculation method, not the physical world. The future has to show if the method is close enough to reality, so that these two rules are confirmed.

Summary

Tooth flank fracture is an infrequent failure mode. When it occurs, it typically comes as a surprise since optical inspection is not giving any advance warning. The mechanism of the failure seems to be well understood so that the respective standardized calculation method under development is promising to become a reliable tool. Some curiosities like the difference in the hardness from input to output and the discontinuous transition in the material factor should be sorted out before the final version. However, these are only small issues.

If the risk of flank fracture is acute for a gear set, it might be hard to improve the situation without jeopardizing the pitting safety. Since the latter is driven more and more to the limit, we might see an increasing number of flank fractures in the future.

Bibliography


