# Lubricated Wear Testing

George Plint MA, CEng, FIMechE Tribology Trust Silver Medal 2017

Phoenix Tribology Ltd

info@phoenix-tribology.com

PLINT Tribology Products from Phoenix Tribology Ltd





When modelling wear and failure mechanisms involving lubricated contacts, we are most usually concerned with modelling real systems such as the contacts in mechanisms, gear-boxes, engines, bearings and machining and forming processes.

The special difficulty with lubricated tests is that in a lubricated contact we have a complex system, the performance of which depends not only on the contacting materials, the contact geometry and conditions of load and relative motion, but also the lubricant additive chemistry. So we are faced simultaneously with issues relating to engineering design, materials and chemistry.



A good starting point is to consider those real processes in which we would not expect wear to occur; we can then consider those process where wear and failure does occur and try to identify the reasons why.

Providing we have a clean lubricant, free of abrasive particles or wear debris, with a contact operating under fully hydrodynamic or elastohydrodynamic conditions and providing complete separation of the material surfaces, we should not expect to generate any wear. If all types of tribological contact operated with hydrodynamic or elastohydrodynamic lubrication, then we would not have any problem with wear; we would need little in terms of lubricant additive chemistry, other than the provision of additives to provide oxidation and shear stability and we could use any moderately viscous fluid. If this were the case, we could lubricate our systems with glycerine instead of oil. This clearly is not the case.

The problem is that in order to generate hydrodynamic separation of the surfaces, any system must fulfil two requirements. Firstly, there must be a mechanism for delivering the lubricant to where it is required and in sufficient quantity to flood the inlet to the contact. Secondly there must be sufficient entrainment velocity to carry the lubricant into the contact in order for it to generate the necessary hydrodynamic film. This latter condition, except in the case of externally pressurized bearings, requires the contacting surfaces to be in motion with at least one surface moving from the direction of lubricant supply and at sufficient speed to generate a thick enough hydrodynamic film.





The slide illustrates what happens in an elasto-hydrodynamic contact between rollers running with the same surface velocity with a fully flooded inlet.

It is important to note that the system is surprisingly very weakly dependent on load, so that once an elasto-hydrodynamic film of sufficient thickness to separate the surfaces is generated, increasing the load has little impact on the level of separation.

In an ideal world, we would like to achieve this kind of lubrication regime, but with many real mechanisms and, more importantly, all mechanisms starting from rest, we are unable to do this.



This slide shows what happens:

- if the surfaces are sliding in opposite directions the zero entrainment condition
- when the surfaces are at rest the start-up condition
- if there is insufficient lubricant to flood the inlet starved lubrication





In the cases where we cannot prevent intimate contact between the mating surfaces we have to rely on lubricant additive chemistry or coatings to limit wear of the surfaces and to provide protection against scuffing, which is the onset of adhesive wear. It follows that, excluding the special case of rolling contact fatigue, all lubricated wear testing is focused on modelling contacts that are operating outside the hydrodynamic or elastohydrodynamic regimes. Under these conditions, lubricant viscosity plays little or no part and the performance of the contact depends critically on the lubricant additive chemistry.

Examples of such contacts include:

- Journal bearings at start-up or during starved lubrication because of failure of supply.
- Elasto-hydrodynamic bearings at start-up or during starved lubrication because of failure of supply.
- Piston ring and liner contacts at bottom and top dead centre, where entrainment velocities fall to zero.
- Piston ring and liner during running because of starved lubrication.
- Gear contacts during running because of negative entrainment.
- Cam and tappet or finger follower contacts during running because of negative entrainment.





It is important to note that every time a tribo-system goes through a start/stop cycles, or indeed a change in operating point, there will be a degree of repeat running-in. The overall wear of the system is thus the sum of multiple contributions, over the life-time of the components, not simply some constant rate wear process.



As an illustration of the issue of negative entrainment consider the case of a simple cam and tappet.





To generate wear in a lubricated contact, there is no point in choosing a test configuration that produces hydrodynamic separation of the surfaces. Tests must be run in boundary or mixed lubrication.

## **Experimental Requirements**

- How much wear do we need to produce?
- What should the correlation criterion be?
- How many tests do we need to perform?
- How do we produce boundary or mixed lubrication regimes, hence wear?
- How do we produce wear and not failure?

#### PLINT Tribology Products

We need to consider the following questions:

- How much wear do we need to produce?
- What should the correlation criterion be?
- How many tests do we need to perform?
- How do we produce a boundary or mixed lubrication regimes, hence wear?
- How do we produce wear and not failure?

A single wear measurement, at the end of the test, tells nothing about the wear rate or wear transitions. We therefore need to supplement post-test wear measurements, with intelligent real-time observations, of which the most commonly used metric is to use changes in friction, electrical contact potential and temperature close to the contact, to indicate possible wear transitions. Hence, in many cases, we do not use real-time wear measurement, but use other methods to identify significant wear transitions.



Fired engine tests typically produce no more than about 10 microns of cylinder wear. Wear occurs on both ring and liner surfaces. Even scuffing failure is preceded by mild wear.

The deepest honing marks are typically about 10 microns deep and are frequently not fully removed during engine life, thus the contact is still running on the original surface and is still influenced by the remains of the original surface topography.

Hence, to model this, we need to be able to generate and then measure a maximum of approximately 10 microns of wear.

Thin hard coatings are typically 2 to 4 microns thick and do not really wear, but fail through fatigue, hence we only really need an on-line mechanism for detecting that failure.

The failure is confined to the coating and there is no point running on the test, once the coating has failed. No one is interested in the friction or wear of the substrate. It is important to ensure test run on the coating, not the substrate.

Thick coatings are typically 20 to 50 microns thick and do wear. To model this, we need to be able to generate and then measure a maximum of approximately 50 microns of wear and this is usually confined to coating.

With polymers running against metals, it is possible to generate significant amounts of wear and once again we potentially need to be able to measure perhaps up to 50 microns of wear.



So how many tests do we need to perform? Wear is a stochastic process. This is obvious with rolling contact fatigue and fretting, but also applies to sliding and sliding-rolling contacts.

The following example, albeit a simple dry sliding contact, data courtesy of Falex Tribology NV, illustrates the point well. Ten samples each, of three different polymers, were run simultaneously, against a steel counter-face, using a crossed-cylinder geometry, on a ten station test machine. The resulting wear scars were measured and the results presented as box plots.

What is of interest here is how the distributions vary significantly, depending on the total number of cycles run. This would suggest that, in this case, the number of tests required to achieve an acceptable confidence level may vary, depending on the chosen number of cycles.



With regard to the question about how many repeat wear tests we need to perform, the answer is that it depends. A box plot allows us to adopt a rational approach to outliers, in other words, measurements that appear to lie an abnormal distance from other values, in a series of repeat measurements. A point beyond an inner "fence" is considered a mild outlier. A point beyond an outer "fence" is considered an extreme outlier. It will be apparent that if a decision is made to discard extreme outliers, there will be a corresponding improvement in the re-calculated confidence level.

A sensible sample size, for a box plot, which is an appropriate method to apply to lubricated sliding wear tests, is perhaps ten repeat tests. In practice, many experimental tribologists hope to get away with just three repeat tests

For tests involving rolling contact fatigue, cycles to failure can be widely different for nominally identical materials, running under nominally identical test conditions. This is because the materials are unlikely to be homogenous and the fatigue process will be influenced by randomly located stress raisers, within the core of the material. Whereas it is possible to impose control on what is going on with the surfaces of our sliding wear test samples, we have no way of controlling the sub-surface conditions that influence rolling contact fatigue life. Depending on the required confidence level, a minimum of ten, but more usually twenty or thirty repeat tests may be required. In the case of rolling contact fatigue of bearings, it is normal to present results in the form of a Weibull plot.





How to produce a boundary or mixed lubrication regime?

In order to generate wear we need to have low surface speeds, to ensure boundary or mixed lubrication.

We cannot sensibly accelerate our experiments simply by increasing the sliding velocity, as that will change our position on the Stribeck curve.

Wear tests are likely to take a long time!



How to produce wear and not failure?

In order to generate wear and not precipitate failure, we need to limit frictional energy input, which requires low sliding speeds.

In order to have some control on lubricant additive action, we need to avoid out of control frictional heating.

We cannot sensibly accelerate our experiments simply by increasing the load.

Wear tests are likely to take a long time!



Before considering how to model boundary lubricated contacts in a bench test, we need first to concentrate a bit on what the lubricant additives work. There are two basic forms of additive protection available to surfaces that are either temporarily or permanently in intimate contact, physisorped and chemically reacted.

Changes in running conditions or test parameters, for example changing load or frequency of events, will usually cause a temporary change in the consistency of these protective layers or films, indicating some kind of limited dynamic stability. These films may be partially destroyed and reformed each cycle, in other words, each specimen pass may knock the top off a number of surface asperities creating new active metal sites. The dwell time between passes allows time for the chemistry to work, so that the film, chemical or physical, can reform before the next specimen pass. There is no way of course of telling how many asperities could be scraped off in this way, although it must be related to the wear.

The rate of formation of the chemically reacted films in particular is considered to be a direct function of the contact temperature; we need temperature for activation and for controlling the rate of reaction, plus a finite time for the chemistry to take effect.





The temperature reached at the surface of the contact (the flash temperature) is strongly influenced by the width of the contact and flash temperature is responsible for many wear and friction effects. Frictional heating within the contact, during relative motion, is generated at the asperities and the resulting flash temperatures can easily rise to several hundred degrees above the bulk temperature of the surrounding material, but last only as long as the individual asperities are in contact.

It follows that a very small distance from the asperities, the temperature distribution is smoothed out to give an interfacial bulk temperature from which we can derive an average contact temperature distributed over the nominal contact area. This average contact temperature will always be higher than the bulk temperature of the material, as measured by some sensor embedded in the bulk material of the specimens some distance from the contacting surfaces.



All wear processes are influenced by temperature, be they the formation of oxides on surfaces, the transformation of microstructure, the formation or break-down of lubricant additive or other tribochemical films, or thermal stress induced failure. To be more specific, wear occurs in conjunction with the dissipation of frictional energy in the contact and this is always accompanied by a rise in temperature.

Different patterns of energy dissipation will give rise to different wear mechanisms.





The friction power intensity (Matveevsky) is simply defined as the amount of energy pumped into the rubbing surfaces as they pass through the contact zone. The temperature achieved in the contact and in the bulk material is directly related to the FPI and the size and thermal characteristics of the materials and their supports.



The "overlap parameter" (Czichos) is the ratio of sliding distance for "body" divided by sliding distance for "counter body". For the thrust washer this is 1, for fretting tests it is close to 1, but for pin on disc tests it is variable, but is typically less than 0.05. The overlap parameter also applies to reciprocating tests; the longer the stroke, the less the linear wear on the plate sample for a given amount of wear on the moving sample.



There are fundamental problems with sliding hertzian point contact test configurations:

- There are no real life engineering applications that involve sliding hertzian point contacts.
- Tests frequently involve a bearing steel ball sliding against a steel flat. There are few, if any, real engineering applications that involve bearing steel sliding in contact with other steels.

If our aim is to model lubricated wear and failure mechanisms in real systems, such as the contacts in mechanisms, gear-boxes, engines, bearings and machining and forming processes, tests involving sliding hertzian point contacts are not going to provide an adequate model. This is the main reason for the oft repeated comment:

"I got no correlation between my bench test and the engine test"



The most important criterion for correlation between model test and actual application is that the test should reproduce the wear and/or failure mechanisms of the real application.

We really need some common sense here! If the wear generated in the bench test looks nothing like the wear in the real system, the model is wrong.

Does any "real" wear process look like this? If we are interested in wear in real systems, we need to understand something about the wear mechanisms involved in both the real system and the corresponding test model.





The following test model example is taken from 2014 inter-laboratory tests as per ASTM D7421 Standard Test Method for Determining Extreme Pressure Properties of Lubricating Oils Using High Frequency, Linear Oscillation Test Machine.

What mechanisms are contributing to the resulting friction and wear response?

What is this test meant to be modelling and what does it mean?



How do we tell the difference between a wear scar and a witness mark?

To illustrate this point, consider the sorts of wear scar produced in HFRR diesel fuel lubricity test.

With the low lubricity (bad) reference fuel, the main wear scar and the stroke end witness marks merge together into one larger wear scar, with much less obvious directionality.

The wear scar with the low lubricity fluid has the appearance of seizure or galling. This would appear to be an example of "junction growth", with the actual area of contact approaching the nominal area of contact.



The difference between the wear scars appears primarily to be a difference of wear mechanism, that being, for the high lubricity fluid, severe adhesive wear, plus something indeterminate on either side of the wear scar, in the direction of motion, and, for the low lubricity fluid, seizure. Why bother measuring the scar dimensions? Why not just report whether the fluid produces seizure or not?

The conventional method of measuring the wear scar on a ball sample is to measure the wear scar diameter in the direction of sliding and transverse to the direction of sliding, but how do we know, in a low wear situation, that what we are measuring is wear and not simply a sort of witness mark, as one would get on the ball in a Brinell hardness test?

One approach is to normalize the nominal scar measurement by dividing it by the calculated initial Hertzian contact area. This way we can determine if the measured wear area is larger than the initial Herztian contact. Although we cannot assume that a normalized wear scar area of 1 indicates no wear, we can assume that a value of 1 indicates a well performing lubricant compared with a lubricant that produces a value well in excess of 1, for the same test.

This approach allows us to determine which lubricants perform well and which ones do not. It is a more rational approach to reporting results than simply relying on an absolute wear scar measurement. In addition to reporting a normalized wear scar measurement, there may also be some benefit in reporting the ellipticity of the wear scar.



It must be remembered that as wear takes place, depending on the specimen configuration, the contact area may change and hence the contact pressure. This is particularly the case with the sliding hertzian point contact.



We have long been aware of the fact that in a sliding hertzian point contact test, the majority of wear (or plastic deformation), occurs at the very beginning of the test and that once the difference in wear between candidate samples has been established, the number of cycles over which the test is run is somewhat arbitrary.

	Average WSD vs Test Cycles Based on ASTM D6079
	75
	Use calleling follower Vite
	100 L
	200
	80
	8
ASTM D6079 diesel	fuel lubricity test run under specified load, stroke, temperature and test time, but

This example shows what happens if we run the ASTM D6079 diesel fuel lubricity test procedure under the specified load, stroke, temperature and test time, but at different frequencies, hence different number of cycles. Tests were run for 75 minutes each at 5, 10, 20, 30, 40 and 50 Hz, the latter being the standard test frequency. The number of cycles per test was therefore 22500, 45000, 90000, 135000, 180000 and 225000 cycles respectively. Two repeat tests were performed at each frequency.

It is apparent that within approximately 50000 cycles the difference in wear between the high and low reference fluids has been established and that not much changes after that.

ASTM D6079 gives a reproducibility figure of 80 microns as the (approximate) 95% confidence level.



It transpires that with the exception of a couple of outliers, tests of 90000 cycles or more, fall within the reproducibility limit.

Plotting the average of each pair of measurements indicates that, perhaps not surprisingly, all the high lubricity measurements fall within the reproducibility limit, indicating that once the initial wear has taken place, further cycles result in very limited additional wear.

With the low lubricity sample, the average measurements for 90000, 135000, 180000 and 225000 all fall within the reproducibility limit.

It would appear that an acceptable result can be achieved within the limits of the standard, running the tests at any frequency from 20 to 50 Hz and that the choice of frequency and number of cycles is pretty much arbitrary.

However, in practice what it demonstrates is a transition from a severe wear process, to a mild wear process, after perhaps as little as 10,000 cycles, after which, the test continues to run on what is an already heavily worn or deformed surface.

This perhaps provides one of the best explanations as to why sliding hertzian point contacts are poor models of the typically mild wear processes associated with real engineering systems.





Sliding hertzian point contact tests are relatively insensitive to increased additive concentration, once a complete, coherent, additive film has been formed. Although the tests may generate a marked difference in wear scar size, between nominally good and nominally bad samples, the test usually lack the necessary sensitivity to distinguish between good candidate samples with differing concentrations of additives. In essence, too small an area of material is being sampled, compared with the available additives in the lubricant sample. Professor Malcolm Fox explains this more scientifically:

"The downward sweep of disparate sigmoidal curves for the friction reduction effects of different additives becomes close to one downward sigmoidal curve on a molar concentration basis. There are only a finite number of absorption sites for additives to latch on to or adhere to, applying Langmuir adsorption theory". (M F Fox).

### Summary of Limitations Hertzian Point Contact Test

Although widely used:

- Not a model of any real engineering contact
- Tests invariably start with potential failure of surface
- · Wear transition at start of test is usually from severe to mild
- Contact pressure changes significantly during test
- Requires wear scars to be measured on two orthogonal axes
- Wear scars frequently somewhat indeterminate
- Not very sensitive to lubricant additive concentration

PLINT Tribology Products

Although it is a widely used test configuration, there are serious limitations with the sliding hertzian point contact test configuration. These can be summarised as follows:

- It is not a model of any real engineering contact.
- Tests invariably start with potential failure of the surface.
- The wear transition at the start of the test is usually from severe to mild.
- Contact pressures change significantly during the test.
- It requires wear scars to be measured on two orthogonal axes.
- Wear scars can frequently be somewhat indeterminate.
- It is not very sensitive to lubricant additive concentration.



So, how does a sliding hertzian line contact compare with a sliding hertzian point contact?

- There are real engineering contacts that involve sliding (but more commonly, sliding-rolling) line contacts.
- Providing sensible running-in sequences are used, the risk of precipitating failure at the very start of the test, can be minimised.
- Contact pressures still change during the test, but much less and with more predictable pressure distribution than with a point contact.
- Wear on the counter-face surface, depending on stroke length, hence overlap parameter, can be more uniformly distributed.
- Although it will take longer to generate a wear scar, simple geometrical scar measurement, on the cylinder specimen, requires just one measurement, in the direction of sliding.
- Wear scar edges are usually readily identifiable.
- The contact geometry is more sensitive to lubricant additive concentration, because a larger amount of surface is being sampled.



The following examples, comparing fuel lubricity tests run with point and line contact geometries, courtesy of SwRI, illustrate the improved sensitivity of the latter over the former.





To illustrate how much more information can be acquired from a sliding line contact test, at sensible overlap parameters, as opposed to a sliding hertzian point contact test, consider the following examples.

These two tests, used the same specimens, running-in process, test load reciprocating frequency and stroke, which was 25 mm; this allows good differentiation between wear patterns at different stroke positions. The first test sequence involved a simple temperature ramp, the nominal aim of which was to precipitate scuffing failure. The second test was run at steady set-point temperature, but followed a stop/start sequence.

Examination of the specimen shows accumulation of compacted fine debris at the stroke end, with debris filling the grinding marks. Away from the stroke end, shiny areas indicate material removal.

Away from stroke end (25% stroke) and under higher magnification, there is a mixture of smearing of material into the grinding marks (on the left) and light abrasion (on the right).



Turning to the moving specimen, what is observed is not a wear scar, but an agglomeration of transferred fine debris material, across the width of the contact. This is more consistent with fine two body abrasive wear than adhesive wear. Mild two-body abrasive wear, as a result of specimen material hardness and roughness, leads to accumulation of material on the moving specimen, as well as at stroke end on the plate specimen.

Although this may eventually lead to seizure between the transferred material on the moving specimen and the source of the transferred material, the fixed plate, this is not an example of an adhesive wear mechanism.

We can summary of Process in this experiment as:

- Mild abrasive wear
- Agglomeration of fine debris at the leading edge
- Adhesion of transferred material
- Like-on-like materials in contact may eventually lead to seizure

In this case, despite the alleged aim of the experiment, we do have a scuffing test.



Instead of running at constant reciprocating frequency, a stop/start cycle is used, with each cycle resulting in a temperature excursion.

During each stop phase, the temperature of the specimens is allowed to cool to a base temperature set-point.

Examination of the samples shows different wear behaviour.




At the stroke end on the plate specimen, we observe material pull-out, with corresponding material transfer to the surface of the moving specimen.



Unlike with the temperature ramp test, where debris accumulates at stroke end, with the stop/start test, we see removal of material at stroke end (white area) and less wear as we move away from stroke end. The stop-start process spreads wear zones further along the contact.





Not surprisingly, adhesion of a pulled out and transferred particle to the moving specimen frequently results in an observable groove in the fixed plate specimen.

So, for the stop/start test, the process is as follows:

- Minimum lubricant entrainment at start of stroke
- Surface propagated fatigue at asperities
- Adhesive pull-out
- The onset of adhesive wear in other words, we have a scuffing test



It is apparent that the two test procedures produce very different wear mechanisms. The temperature ramp test appears to produce what one might term a "false" adhesive wear process. The stop/start test, with the temperature gradient the right way round and cyclic frictional energy input, produces adhesive wear, as typically illustrated in most text books.

If we consider scuffing to be either the onset of adhesive wear, or at very least, some form of adhesive wear process, we should at least use tests that, on examination, actually produce adhesive wear and not some other mechanism.





The key justification for running stop/start type tests is that this is what happens in many real systems; the contact conditions move up and down the Stribeck curve, depending on the system operating point. Furthermore, in real systems, simply changing the operating point can trigger a short term change in wear rate.

This is nicely illustrated in the engine test example from SwRI, in which radioactive tracer technology is used to monitor real-time wear of components.



One of the key issues with a simple line contact specimen arrangement is how to mitigate the geometric stress concentration at either end of a cylindrical specimen, of finite width. One solution is to use a roller with a logarithmic end profile. Another solution is to use a plate specimen with curved edges and a cylinder that is wider than the plate. This then allows the possibility of using plate specimens with non-parallel sides, such that the contact pressure varies with stroke position, for a given load.



It transpires that in terms of r.m.s. and instantaneous friction and contact resistance, with a fully formulated oil and these samples, there is minimal difference between the frictional response of the parallel and wedge specimens: the frictional response appears to be independent of contact width.

Visual inspection of the samples indicates that, with a fully formulated lubricant, the wedge specimens "scuff" at the narrow end. With the parallel sample, a narrow band of adhesive wear is observed, covering the whole stroke length.



With the parallel sample, both reversal positions indicate "severe adhesive wear", characterised by deep holes in surface. The lighter regions outside these areas have cracks in the surface, suggesting "mild adhesive wear". There is a clear distinction between the severe band and the areas subjected to mild adhesive wear.



With the wedge sample, the wide end reversal position indicates "mild adhesive wear" with characteristic directionality. The narrow end is subjected to "severe adhesive wear" with destruction of the oxide or additive protective film and plastic deformation.



Comparison between the different wear regimes is thus as indicated above. Axial surface profilometry shows that with the parallel samples, wear is a maximum at stroke ends and a minimum at mid-stroke.

With the wedge specimen, in comparison with the parallel specimen, the overall wear is much greater and is indeed greater at the narrower end of the wedge, as one would expect.

To summarise:

- Stroke ends: with parallel, wear same at each end. With wedge, wear same at wide end but approximately double at narrow end, where contact pressure is approximately double.
- With parallel, contact pressure is constant with stroke. With wedge, contact pressure is ramping up in one direction and down in the other. Does this action prevent the establishment of a stable lubrication regime, hence contributing to greater wear at the mid-stroke position?

Repeatability with curved edge plate specimens and overlapping pin samples appears to be good, compared with tests where the contact width is less than the plate width. This suggests that in the latter case, variable edge effects may have an important influence on repeatability.

Provided that the wear mechanism is the same at either end of a wedge specimen, the friction is independent of nominal contact area.

Time smoothed friction provides little information with regard to wear transitions under mild regimes.

However, changes in instantaneous contact resistance is a good indicator of wear transitions.

With a curved edge wedge specimen it is possible:

- to generate different wear regimes at either end of the specimen, hence providing more information from a single test run
- to produce mild and severe adhesive wear, while running under conditions of steady load and temperature, in other words, without resorting to the application of ramped loads or temperatures



Now let's move on to sliding-rolling line contacts. I will use as an example, an attempt to model and involute gear with a two roller machine.

One of the most commonly occurring misapprehensions is to assume, wrongly, that the pitch-line velocity of actual gears to be modelled, is the same as the surface sliding speeds of the rollers in a two roller machine test model. The pitch-line velocity determines the contact time for gear tooth pairs. However, the rolling and sliding velocities between the gear tooth pairs depend on the gear tooth profile plus the contact time. It is these velocities that should be modelled in any two roller experiment.





With involute gears we have two contacting surfaces with variable curvature, moving together with a complex combination of rolling and sliding. An added complication is that, away from the pitch point, there is load sharing between overlapping pairs of teeth, adding the uncertainty of dynamic loading, to an already complex system.

This example is based on involute gears, with a 20 degree pressure angle.



Shortly after engagement, the surface of the driving gear is moving with a small velocity relative to the point of contact, whereas the driven gear has a much higher velocity. These are defined as the rolling velocities of the two surfaces.

The sliding velocity of the driven tooth across the surface of the driving gear is in the same direction as the rolling velocities, and is conventionally described as a negative sliding velocity.

At the pitch point, the rolling velocities are equal and there is no sliding in the contact.

As the point of contact nears the end of the contact path, the driving gear is moving faster than the driven gear, relative to the point of contact.

The sliding velocity of the driven tooth, across the surface of the driving tooth is in the opposite direction to the rolling velocities, and this is conventionally described as positive.





Note that in the case of the driving gear, sliding is always away from the pitch point. This imposes a tension in the surface layers and is the reason for the observed greater tendency of the driving gear to pit in the region of the pitch point.

Conditions for the driven gear are the mirror image of those for the driving gear, with sliding always towards the pitch point, imposing compressive forces to the surface layers, thus discouraging pitting.



To take our analysis further, we need to introduce some definitions.

With gears, the relationship between rolling and sliding velocity requires careful definition. Merritt defines the rolling velocity ratio (RVR) as the ratio of the smaller to the larger velocity of the two surfaces relative to the point of contact, taking algebraic sign into account.

For pure rolling, without sliding, RVR = 1.
For pure sliding, RVR = 0.
The sliding velocity ratio, SVR, is defined as shown.
For pure rolling, without sliding, SVR = 0.
For pure sliding, SVR = 1.

	Modelling Invo	lute v Pul	Geai	rs with T	wo Roller Ma	achine			
	Product of FPI and contact transit energy input during its transit of c	time, EP ta ontact zone	ikes into a e, where t	ccount length of tim is transit time in se	e during which material is sul conds	bjected to			
	Energy Pulse:	Ep	=	$\mu$ PV <sub>s</sub> t <sub>t</sub> /A	J/mm <sup>2</sup>				
EP is analogous to Archard Wear Law, however, it uses the friction force rather than applied load, which is perhaps more logical as it takes into account the rubbing conditions (but assumes that the friction coefficient can be measured)									
	Archard Wear Law:	v	=	$k P V_s t_t / A$	mm <sup>3</sup>				
	EP can be regarded as an increme as a measure of total wear	ntal contril	oution to v	vear or surface dam	age in contact. Sum of EPs ca	n be used			
Correct analysis of EP in real contact and subsequent modelling in experimental design significantly enhances probability of achieving a satisfactory emulation									
		P		Phoenix Tribology Ltd					

The Energy Pulse is the product of the Friction Power Intensity (FPI) and the contact transit time. The EP therefore takes into account the length of time during which the material is subjected to energy input during its transit of the contact zone, where  $t_t$  is the transit time in seconds.

The Energy Pulse is analogous to the Archard Wear Law, however, the Energy Pulse equation uses the friction force rather than the applied load. This is perhaps more logical as it takes into account the work done in the contact.

Archard Wear Law:  $\Delta V = k P V_s t_t / A mm^3$ 

Each Energy Pulse can be regarded as an incremental contribution to wear or surface damage in the contact. The sum of the Energy Pulses can be used as a measure of the total wear.



Correct analysis of the EP in the real contact and subsequent modelling in the experimental design significantly enhances the chances of achieving a satisfactory emulation of sliding and combined sliding and rolling contacts.

It is important to note that in many machine components there can be very high FPIs but, because the contact durations are short, the EP is low and hence the incremental damage is low.



We can now examine how the SVR, the RVR and the EP vary at different positions on the gear tooth.

The following schematics are based on two 30 tooth gears with a 20° pressure angle. Note that SVR, which has a negative value, is here plotted as positive, so as to appear above the x-axis.



As the EP is a function of load, it is clear that the EP for a single tooth contact will not only vary with relative sliding velocity, but also as a result of dynamic loading and the sharing of load between successive pairs of teeth.



The EP is zero at the pitch point, despite the load potentially being at a maximum, because the SVR is zero. The EP increases in the direction of the root and the tip, as the SVR increases.





Zero EP at the pitch point provides the machinism for generating micro-pitting. High EP at the tooth tip produces conditions conducive to scuffing.

ller machine	tween dynamically e test geometry	loaded involu	ute gears and conventional	two
	Gear Tooth Pair	Tw	win Rollers	
Contact Geom	Curvature varying with	position on tooth Fix	ixed by disc diameters	
Rolling Veloci	ty Ratio Varying with position or	tooth Fix	ixed with steady state motor speed set-points	
Sliding Velocit	ty Ratio Varying with position or	tooth Fix	ixed with steady state motor speed set-points	
Load	Varying with position or	tooth Fix	ixed with steady state load set-points	
Contact Press	ure Varying with position or	tooth Fix	ixed with steady state load set-points	
Energy Pulse	Varying with position or	tooth Fix	ixed with steady state load and speed set-points	

It will be apparent that there is a significant difference between a pair of dynamically loaded involute gears and a conventional two roller machine test geometry. These can be summarised in the table.

There is, however, one parameter that, for gears, is fully deterministic, but for two roller machines is less certain. With gears, the points of contact between pairs of teeth plus which teeth pairs engage, each cycle, is determined by the design of the gears and the number of teeth on each gear. With a two roller machine, especially where surface speeds are independently controlled, the relationship between corresponding points on each roller is continuously variable.



Parameters affecting performance of both gear and two roller contacts are as follows:

- Contact pressure
- Lubricant film thickness
- Frequency of encounter
- Friction power intensity (FPI)
- Energy pulse (EP)

All these parameters are easy to define for the two roller contact, but less so for the gear tooth contact. However, it is clearly necessary that if we wish to model the complex operating conditions in a gear contact, with a simplified, steady state model, in a two roller machine, we must start by evaluating the conditions in the former.

#### Modelling Involute Gears with Two Roller Machine Parameters affecting performance of both gear & two roller contacts Contact Pressure Relatively straightforward to calculate at gear pitch point Load sharing, combined with varying tooth curvature, renders simple calculation impossible Lubricant Film Thickness Can be estimated using any of established elasto-hydrodynamic film thickness equations (e.g. Dowson and Higginson). Caveat: calculations all assume a fully flooded inlet to contact and minimal side leakage. In practice, most gears run under conditions of starved lubrication Estimating the lubricant film thickness away from the pitch point clearly requires the calculation to be performed taking into account local contact pressure and local entrainment conditions Frequency of Encounter Gear tooth contacts are intermittent, in other words, a given pair of gear teeth is subjected to a brief period of engagement, followed by longer period rotating out of contact, before once again coming into contact · Period out of contact allows time for dissipation of frictional heat and for lubricant additive chemistry to react FPI and EP • FPI and EP clearly have no meaning at the pitch point and thus need to be calculated taking into account local contact pressure, hence local load and sliding velocity Not easy, hence a simple bench-mark estimate is to use hertzian contact pressure and film thickness at pitch point and mean sliding speed across contact. This should provide basic order of magnitude values for initial two roller tests

Because of the general complexity and uncertainty, significant assumptions are inevitable.

Whereas the contact pressure is relatively straightforward to calculate at the gear pitch point, the uncertainty caused by load sharing, combined with varying tooth curvature, renders simple calculation of contact pressure impossible.

The lubricant film thickness at the pitch point can readily be estimated using any of the established elasto-hydrodynamic film thickness calculations, for example, the Dowson and Higginson equation. There is, however, a significant caveat: the calculations all assume a fully flooded inlet to the contact and minimal side leakage. In practice, most gears run under conditions of starved lubrication and there is always side leakage.

Estimating the lubricant film thickness away from the pitch point clearly requires the calculation to be performed taking into account local contact pressure and local entrainment conditions. Once again, that is not easy.

Gear tooth contacts are intermittent, in other words, a given pair of gear teeth is subjected to a brief period of engagement, followed by longer period rotating out of contact, before once again coming into contact. The period out of contact allows time for dissipation of frictional heat and for the lubricant additive chemistry to react. It is well known that in gears running at very high speeds, the frequency of encounter can be too short to allow the chemistry to work, resulting in scuffing.

The FPI and EP clearly have no meaning at the pitch point and thus need to be calculated taking into account local contact pressure, hence local load and local sliding velocity. This is, of course, by no means easy, hence a simple bench-mark estimate is to use the hertzian contact pressure at the pitch point and the mean sliding speed across the contact. This should provide basic order of magnitude values.



Now to some practical choices. It is important to state that there are no established and proven test configurations or test parameters for modelling gear contacts in a two roller machine, hence there is no proven right or wrong answer. However, experiments based on sensible estimates and rational choices are likely to be more meaningful than randomly chosen test parameters.

First let's consider the roller sizes. The contact between involute gears at the pitch point can be modelled as cylinders of the same local contact radius as the gears.

Using rollers of equal radius to the gear radii at the pitch point is a choice made by numerous experimenters. It would seem a rational choice, if one wished to perform experiments modelling conditions at the pitch point. It would seem a somewhat arbitrary choice, if one were intent on modelling conditions away from the pitch point. In practice, however, roller contacts are essentially scalable, so choosing roller diameters that conveniently provide contact radii somewhere within the range of radii of the gear teeth profiles would seem acceptable.



To determine the required two roller machine capacity we need to perform the following calculations, to match the machine capacity to the gear tooth contact speeds and pressures. Having first chosen suitable sized test rollers, we need to perform:

- Speed/RPM Calculations
- Load/Contact Pressure Calculations
- Machine Torque/Power Calculations

	IOPIWO	Rolle	er Experiments		
NPUT DATA				2000 MPa	
Roller 1 - Diameter	70	mm	Friction	189	1 N
Roller 2 - Diameter	70	mm	Torque 1	66.18	5 Nm
Load	18.91	kN	Torque 2	66.18	5 Nm
Traction Coefficient	0.1		Roller 1 - Power	6.9	3 kW
			Roller 2 - Power	17.3	3 kW
Roller 1 - Speed	1000	rpm	Surface Speed 1	3.6	7 m/s
Roller 2 - Speed	2500	rpm	Surface Speed 2	9.1	6 m/s
			Sliding Velocity	5.5	0 m/s
			Friction Power	10.4	0 kW
			Rolling Velocity	6.4	1 m/s
			Slide-Roll Ratio	85.7	0 %

Typical machine torque and power calculations are as shown in the table, this is for 70 mm diameter by 10 mm wide contacts at 2000 MPa contact pressure, with moderately realistic surface speeds.

If you perform these sorts of calculation, you will discover how easy it is to end up specifying an absolutely enormous, high power capacity, two roller machine.

Having calculated the potential machine capacities, it is sensible to:

- Review FPI to confirm that it is sensible
- Calculate the nominal lubricant film thickness



The normal practice with two roller tests is to jet test lubricant into the in-running side of the roller contacts. To model starved lubrication, jetting lubricant against the out-running side of the contact may be worth considering.



Micro-pitting tests should be run at high contact pressures equivalent to those at or near the gear pitch point, but with low sliding velocities, hence low frictional energy input. Note that two roller tests have shown that negative sliding is more conducive to pitting than sliding in a positive direction. The level of asperity engagement can be varied by:

- Varying the lubricant entrainment velocity
- Varying the lubricant inlet temperature, hence viscosity



Scuffing tests should be run at lower contact pressures equivalent to those at or near the gear tip, but with higher sliding velocities, hence high frictional energy input. Typical sliding speeds are between 5 and 20 ms<sup>-1</sup>.

As scuffing is a wear transition (the onset of adhesive wear), tests sensibly involve increasing the severity of conditions within the contact, with the aim of precipitating the transition, but preferably not causing catastrophic failure.

There are various mechanisms for precipitating scuffing in a two roller machine:

Firstly, what about progressively increasing the load:

As EHD film thickness is only weakly dependent on load, the main effect of increasing load is thus to increase the frictional energy input, hence contact temperature.

We can progressively reduce the lubricant film thickness:

There are two methods for achieving this. Firstly, by increasing the lubricant inlet temperature, hence reducing the lubricant viscosity. Secondly, by reducing the lubricant entrainment velocity.

Finally, we can progressively increase the frictional energy input:

This is best achieved by increasing the sliding velocity in the contact, while limiting the entrainment velocity.

# Modelling Involute Gears with Two Roller Machine Two Roller Experiments - Test Procedures

## Running-in

The need for satisfactory running-in of gears is well understood. There is a similar requirement to run-in test rollers. This is best performed at modest loads and low sliding velocities.

Running-in performs two functions, firstly, generating plastic shakedown, which is the process of initial plastic deformation of the sub-surface, and secondly, flattening the peaks of the surface asperities. Shakedown imparts residual stresses to the sub-surface material, after which the contact should be elastic. This is analogous to a controlled work hardening process. The tips of surface asperities are flattened by a combination of plastic deformation and mild wear. Increasing the sliding velocity during running-in alters the shakedown behaviour and increases the risk of scuffing at the asperity tips

The need for satisfactory running-in of gears is well understood. There is a similar requirement to runin test rollers. This is best performed at modest loads and low sliding velocities. Running-in performs two functions, firstly, generating plastic shakedown, which is the process of initial plastic deformation of the sub-surface, and secondly, flattening the peaks of the surface asperities. Shakedown imparts residual stresses to the sub-surface material, after which the contact should be elastic. This is analogous to a controlled work hardening process. The tips of surface asperities are flattened by a combination of plastic deformation and mild wear. Increasing the sliding velocity during running-in alters the shakedown behaviour and increases the risk of scuffing at the asperity tips.



An alternative to the steady state slide-roll ratio behaviour achievable with a two roller machine, a variable slide-roll ratio contact can be generated by imposing a degree of rotation on a test roller, as it is reciprocated against a flat plate.

In this device, a crowned or flat roller is reciprocated, with a linkage connected to the side of the roller opposite the tribo-contact. As the roller is reciprocated, the RVR changes with stroke position, with RVR = 1 (pure rolling) at mid-stroke, and RVR <1 (rolling and sliding) away from the mid-stroke position. Hence, the point of contact moves on both the surface of the roller and the surface of the plate, with a motion similar to a pair of gear teeth rolling backwards and forwards, on either side of the pitch point.

By changing the linkage position, RVR range can be changed.

The device has not been extensively used for micro-pitting tests, but has been successfully used as a scuffing screening test, for gear oils. In this case, scuffing is precipitated in a controlled way, not by increasing the frictional energy input, but by increasing the contact temperature by electrically heating the plate specimen.



Another important sliding-rolling contact is that between a cam and a follower.

I have covered the simplified analysis of a cam-follower contact in another lecture on lubricated friction measurement. The key wear issue with cams and followers is the contact at the cam nose, where a combination of poor entrainment conditions and high peak loads can give rise to scuffing, in other words, the onset of adhesive wear. Under these conditions, additive protection is essential.

In this example, we have a quite complicated "wedge on sphere" geometry, designed to promote rotation on the bucket follower, thus producing a circular wear track. For our bench mark calculations, we use a simplified geometry, comprising a "cylinder on flat" with a 10 mm wide contact width.

Only by doing the appropriate analysis can we establish the lubrication regime under which our component is running, hence determine test conditions, in an appropriate model system, in which to generate relevant and meaningful wear or failure data.



# Now let's move on to lubricated area contacts.

With simple area contacts, lubricant entrainment conditions are poorly defined. Furthermore, in the absence of some form of converging geometry, it is impossible to generate hydrodynamic separation of the surfaces. It follows, that, in real applications, components usually have to be designed to enable lubricant entrainment.



This frequently leads to a reliance on component based tests, particularly in the case for wet clutches, their materials and associated ATFs, face seals and lubricated journal and plain thrust bearings. In these cases, the tribological response depends not just on lubricant and material properties, but importantly, on the design of the components.



In many other lubricated area contacts, the nature of the application is so specific and the contact conditions so complicated, that simple tribometer tests are of little practical use. In order to achieve correlation between wear or failure in a laboratory model and that in the real system, it is usually necessary for the model to be very much application specific and in many cases to use components from the real system.

The examples shown here are perhaps all too specific to be included in a general talk on lubricated wear testing and perhaps need to be covered elsewhere, as topics in their own right.
## <section-header> Conclusions Determinent ensible lubricated wear tests, we must choose systems and procedures that allow tests. to be performed under mixed or boundary lubrication that model the required wear and failure mechanism of the real system that generate wear and not simply precipitate failure that do not precipitate unwanted wear transitions that are at sensible contact pressures with contact temperatures, whether self-generated or externally applied, appropriate for additive chemistry that have cyclic energy input (stop/start) if necessary that do not attempt to accelerate the wear simply by increasing: the do not attempt to accelerate the wear simply by increasing: the soliding webcit.

In order to perform sensible lubricated wear tests, we must choose systems and procedures that allow tests:

- to be performed under mixed or boundary lubrication
- that model the required wear or failure mechanism of the real system
- that generate wear and do not simply precipitate failure
- that do not precipitate unwanted wear transitions
- that are at sensible contact pressures with contact temperatures, whether self-generated or externally applied, appropriate for the additive chemistry
- that limit or control the frictional energy input
- that have cyclic energy input (including stop/start) if necessary
- that do not attempt to accelerate the wear simply by increasing:
  - $\circ$  the sliding velocity
  - o the load

Slide 74



On the subject of sliding hertzian point contact tests:

- Essentially pass/fail test, suitable for quality control applications
- Repeatability depends on having highly finished sample surfaces in order to maximise the number of asperities supporting the applied load at test start
- Nearly impossible to avoid plastic flow at the start of a test, either on or below the surface, and obliteration of the surfaces
- Because of small contact scale, number of available active metal sites is very limited, making tests rather insensitive to additive concentration
- Rate and range of fall in contact pressure at the beginning of the test produces an almost "digital" response
- Most standard tests use material pairs that are unrepresentative of material pairs in real sliding application contacts
- Do not model anything in the real world
- Wasting time trying to:
  - Improve sensitivity
  - Achieve correlation with real applications

Sliding hertzian line contact tests:

• With care, a potential model for certain real systems

Sliding/Rolling hertzian point contact tests:

• A comprehensive model of certain real systems

Lubricated Area Contacts:

- Frequently involve actual application components
- To achieve correlation, require careful duplication of the application contact conditions

Slide 75

Conclusion

If you are going to run a wear test, you should at least know what wear mechanism you are modelling and what wear mechanism or mechanisms you are producing in your test

PLINT Tribology Products

So, why is it so important to get these things right? You are probably familiar with the classification of different test systems given in DIN 50320, showing potential steps between laboratory test and real application, and vice versa. We can usually manage to characterise frictional response relatively quickly, but it usually takes a long time to generate wear or failure, in real systems, hence sensible wear tests are likely to take a long time!

Slide 76



It follows that if we can miss out any of the steps between real system and simple laboratory test, the potential for saving both time and money is enormous. As tribologists, we all dream of achieving convincing correlation between a simple bench test and the real-life application, if possible, missing out all the steps in the middle!