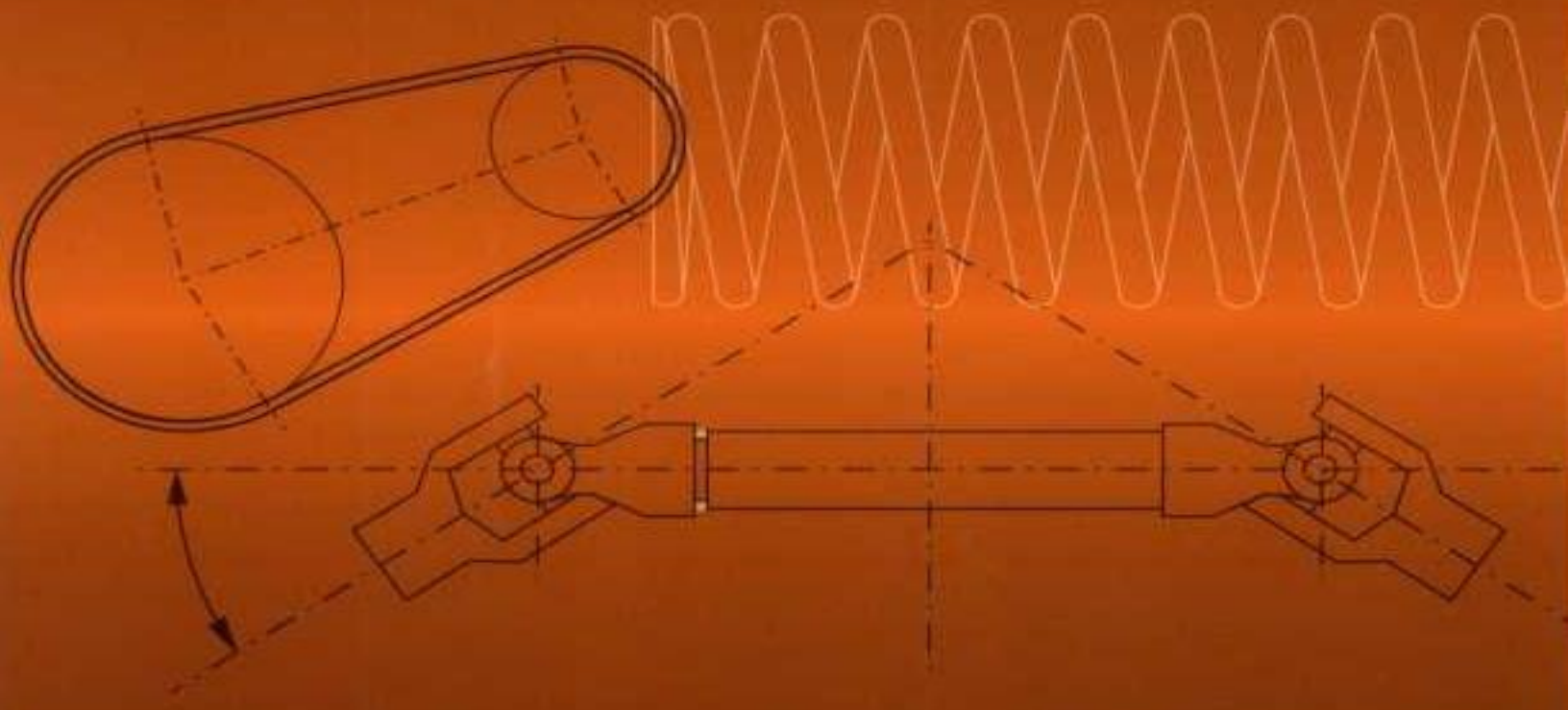
Peder Klit & Niels L. Pedersen

MACHINE ELEMENTS

ANALYSIS AND DESIGN



Polyteknisk
Forlag

Machine Elements • Analysis and Design

This textbook provides undergraduate students with a basic understanding of machine element theory, and introduces tools and techniques to facilitate design calculations for a number of frequently encountered mechanical elements.

The material in the book is appropriate for one or two courses in Machine Elements and/or Mechanical Engineering Design. The material is intended for students who have passed first and second year basic courses in engineering physics, engineering mechanics and engineering materials science.

The book is organized into 13 separate chapters, which in principle can be read independently. The covered subjects are: Tolerances, springs, bearings, shafts, shaft-hub connections, threaded fasteners (bolts), 2D Joint Kinematics, couplings, clutches, brakes, belt drives, gear geometry and strength of gears.

About the authors

Peder Klit and Niels L. Pedersen are both professors in machine elements at the Department of Mechanical Engineering, the Technical University of Denmark, DTU.

Peder Klit & Niels L. Pedersen

MACHINE ELEMENTS

ANALYSIS AND DESIGN

Polyteknisk Forlag

Machine Elements

Analysis and Design

By Peder Klit and Niels L. Pedersen

2nd edition, 2014

ISBN 978-87-502-1068-9

Cover design by Anne Bjørlie

Printed by Livonia

Printed in Latvia 2014

Polyteknisk Forlag

Anker Engelundsvej 1

DK-2800 Lyngby

Phone: +45 7742 4328

Fax: +45 7742 4354

E-mail: forlag@polyteknisk.dk

www.polyteknisk.dk

Preface to the second edition

This book is intended to provide undergraduate students with basic understanding of machine element theory, and to introduce tools and techniques facilitating design calculations for a number of frequently encountered mechanical elements. The material in the book is appropriate for a course in Machine Elements and/or Mechanical Engineering Design for students who have passed first and second year basic courses in engineering physics, engineering mechanics and engineering materials science.

At the end of each chapter in the book, references, which may be useful for further studies of specific subjects or for verification, are given. Students who wish to go deeper into the general theory of machine elements may find the following textbooks inspiring:

- Norton, R. L., "Machine Design, an integrated approach", Prentice-Hall, 2014.
- Shigley, J. E. and Mischke, C. R., "Mechanical Engineering Design", McGraw-Hill, 2004.

Students are encouraged to find supplement information from other sources such as International and National Standards, Internet Catalogues and information provided by companies (online or paper based). Those who are in command of the German language will find numerous German textbooks of very high standard. Outstanding in quality is the textbooks by Niemann and co-authors.

- Niemann, G., Winter,H., Hohn, B. "Maschinenelemente", Springer Verlag, Band I, 2005.
- Niemann, G., Winter,H., "Maschinenelemente", Springer Verlag, Band II, 2003.
- Niemann, G., Winter,H., "Maschinenelemente", Springer Verlag, Band III, 1983.
- Decker, K., "Maschinenelemente, Funktion, Gestaltung und Berechnung", Carl Hanser Verlag, 2011

and an overall mechanical engineering reference book can be recommended as helpful during the study, and afterwards in your professional engineering life as well:

- DUBBEL: Taschenbuch fiir den Maschinenbau, Springer Verlag, 2014.

In this second edition of the book the misprints in the first edition have been corrected and some chapters have been extended. A new chapter on 2D joint kinematics has also been added to the book.

Copenhagen, June 2014

Peder Klit and Niels L. Pedersen

Contents

Chapter 1 Limits, fits and surface properties

1.1 Introduction

The technical and technological development continuously provokes manufacturing of new and more attractive products. Products that "a short time ago" were good and competitive are now outdated or too expensive to produce. It is an ongoing challenge to design for easy and economical production and at the same time fulfil the functional demands.

Product costs mainly originate in design and the designer has a prime responsibility to ensure that the product gives optimum value for money. Cost is just as much an attribute of the design specification as is performance, appearance, reliability, life, safety etc., and is an essential factor to be satisfied by the optimum design solution. A design that fails to meet its cost specification is no better than one that fails to satisfy the performance requirements. Furthermore, when all other factors are equal, the decision by the customer whether or not to buy a product is largely determined by its cost.

Total product cost is the addition of manufacturing cost and selling cost and is shown graphically in Figure 1.1.

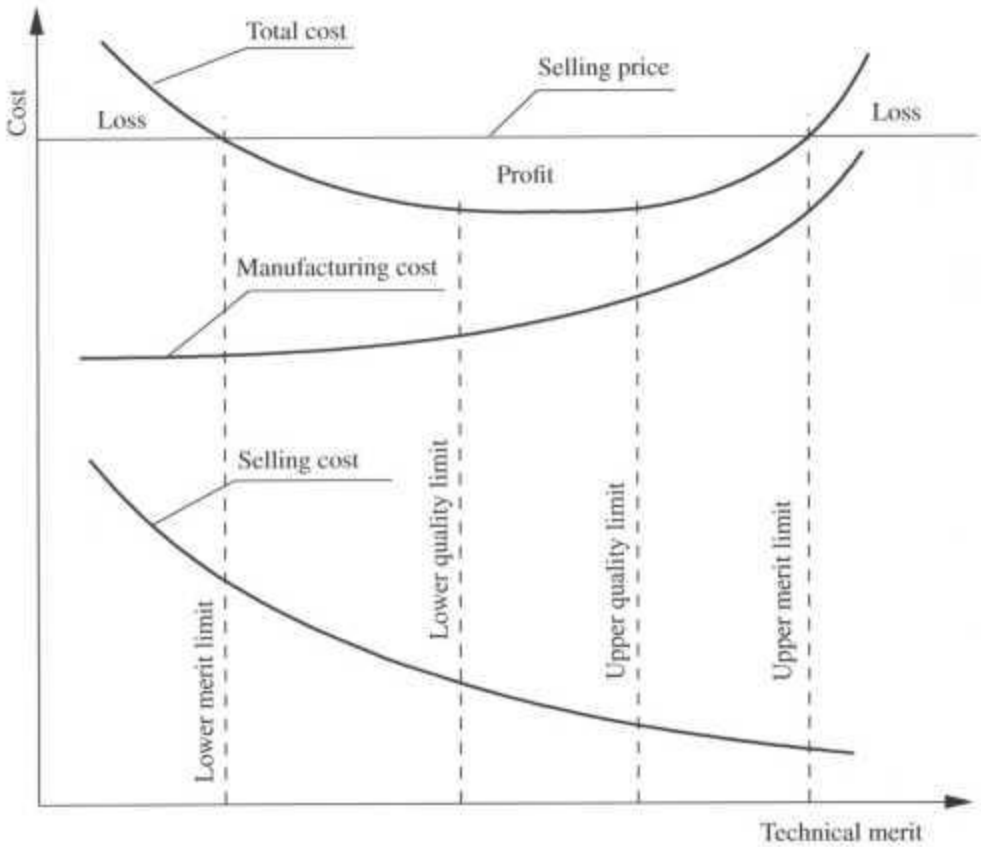
The essence of good design must be the provision of optimum value within the product specification. Any excursion beyond the upper and lower limit of technical merit will rapidly turn a situation of profit into one of loss. Within these limits, there must be upper and lower quality limits set to maximize the profit from the product [1].

Specifications related to manufacturing

Looking at the quality demands for the manufacturing of a product it is important to notice that there is a lower as well as an upper limit to respect. When designing a product it is, next to basic functional demands, important to analyze which productional demands are to be stated for the single components. This includes everything from choice of material, specification of surface characteristics (form and surface texture) and to dimensions (tolerances on lengths, diameters, angles).

1.2 Geometrical tolerances

It maybe required to specify that the faces of a component are flat, parallel, perpendicular to others etc. This is done on drawings by specifying a geometrical tolerance. For instance, the cylinder head on a piston compressor does need to be flat, where it interfaces with the crankcase, which of course also needs to be flat. It does on the other hand not need to have very accurate size tolerances. Cylindrical components may also need geometrical tolerances. Again using the piston compressor as an example, the crankshaft will almost certainly need geometrical tolerances. Several bearing surfaces need to be concentric with each other. The only way to guarantee concentricity is to use one surface as a datum and



[billedtekst start]**Figure 1.1:** Graph illustrating total product cost.[billedtekst slut]

use geometrical tolerances, in order to ensure that the other surfaces do not deviate outside of the limit specified.

The scope of specifying geometrical tolerances on technical drawings is to limit deviations of form, orientation, location and run-out for technical products to be produced. A detailed description is given in [2],

Geometrical tolerances shall be specified only where they are essential for the function.

Indicating geometrical tolerances does not necessarily imply any particular methods of production, measurement or gauging.

A geometrical tolerance applied to a feature defines the tolerance zone within which the feature (surface, axis, or median plane) is to be contained.

1.2.1 Specifying geometrical tolerances

The tolerance requirements are shown in a regular frame that is divided into two or more boxes. These boxes contain from left to right:

- the symbol for the characteristic to be tolerance

- the tolerance value

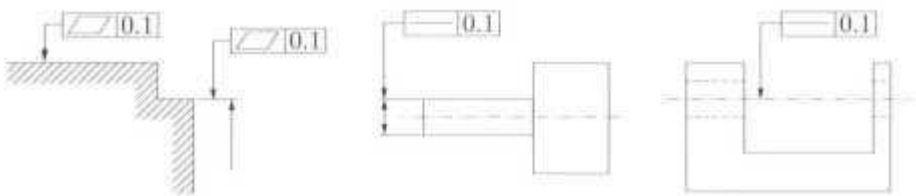
- if appropriate, the letter or letters identifying the datum feature or features.

[billedtekst start]**Figure 1.2:** Examples of geometrical tolerance specifications.[billedtekst slut]

1.2.2 Toleranced features

The tolerance frame is connected to the toleranced feature by a leader line terminating with an arrow in the following way:

- on the outline of the feature
- as an extension of a dimension line, when the tolerance refer to the axis defined by the feature so dimensioned, or
- on the axis when the tolerance refers to the axis.



[billedtekst start]**Figure 1.3:** Examples of tolerance frames connected to features.[billedtekst slut]

Features & tolerances		Toleranced characteristics	Symbols
Single features	Form tolerances	Straightness	—
		Flatness	▱
		Circularity	○
		Cylindricity	⌭
Related features	Orientation tolerances	Parallism	//
		Perpendicularity	⊥
		Angularity	∠
	Location tolerances	Position	⌖
		Concentricity & coaxiality	◎

		Symmetry	
	Run-out tolerance	Circular run-out	

Figure 1.4: Examples of symbols for toleranced characteristics.

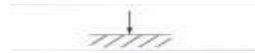
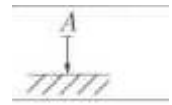
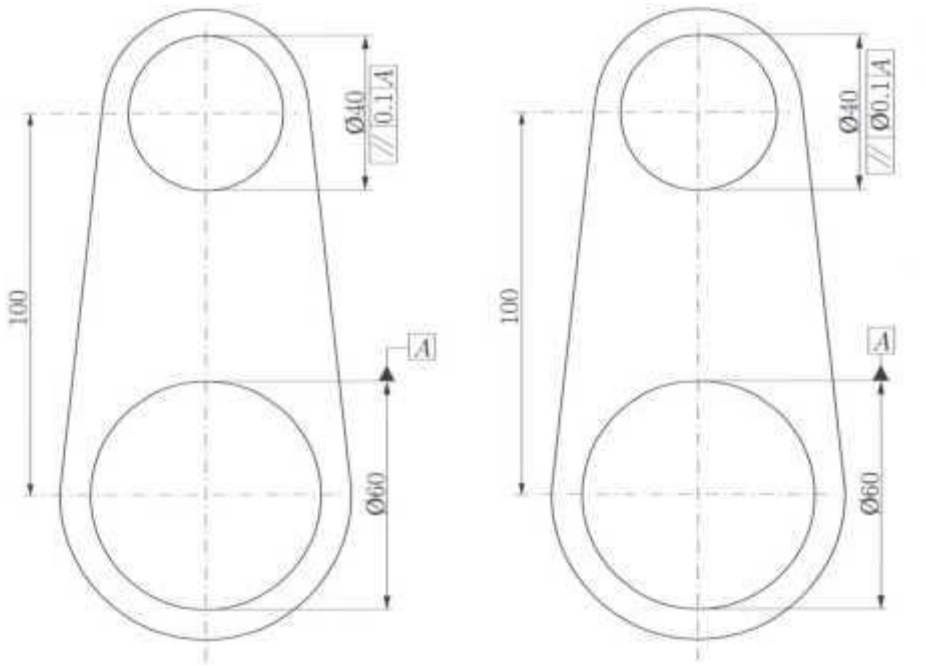
Descriptions		Symbols
Toleranced feature indications	direct	
	by letter	
Datum indications	direct	
	by letter	

Figure 1.5: Examples of identifying datum features.



[billedtekst start]**Figure 1.6:** The width of the tolerance zone is in the direction of the arrow of the leader line joining the tolerance frame to the feature, unless the tolerance zone is preceded by the sign Ø.[billedtekst slut]

1.3 Surface texture

Every surface has some form of texture that consists of a series of peaks and valleys distributed over the surface. These peaks and valleys vary in height and spacing, and have properties that are a result of the way the surface was produced. For example, surfaces produced by cutting

tools tend to have uniform spacing with defined cutting directions, whilst those produced by grinding have random spacing.

The ability of a manufacturing operation to produce a specific roughness depends on many factors. For example, in end mill cutting, the final surface depends on the rotational speed of the end mill cutter,

the velocity of the traverse, the feed rate, the amount and type of lubrication at the point of cutting, and the mechanical properties of the piece being machined. A small change in any of the above factors can have a significant effect on the surface produced.

Measuring surface finish. In the past the evaluation of surface texture was done by comparing the surface to be measured with standard surfaces. A modern surface measuring instrument consists of a stylus with a small diamond tip, transducer, a traverse datum and a processor. The surface is measured by moving the stylus across the surface. As the stylus moves up and down along the surface, the transducer converts this movement into a signal which is then exported to a processor that converts it into a number and usually a visual profile.

1.3.1 Surface Texture Parameters

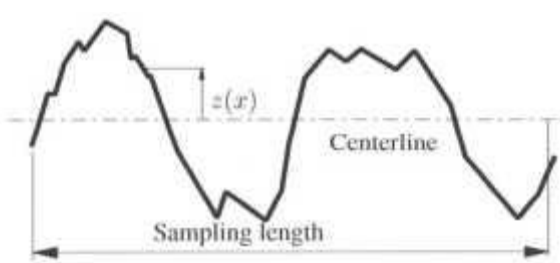
The identification of the surface texture uses a number of parameters. These are different depending on the standard used and on the issue of the relevant standard.

1.3.2 Surface Texture Parameters

The identification of the surface texture uses a number of parameters. These are different depending on the standard used and on the issue of the relevant standard.

Ra - **Arithmetical mean deviation.** Graphically, the average roughness is the area between the roughness profile and its center line divided by the evaluation length (normally five sampling lengths equals one evaluation length).

$$\mathrm{Ra} = \frac{1}{l}\int_0^l |z(x)|dx \tag{1.1}$$



[billedtekst start]**Figure 1.7:** Sketch showing definition of Ra.[billedtekst slut]

or, if the surface profile is measured in equidistant discrete points

$$\mathrm{Ra} = \frac{1}{N}\sum_{i=1}^{N} |z_i| \tag{1.2}$$

Rq - **Root mean square (rms).** This roughness specification is often used in the US.

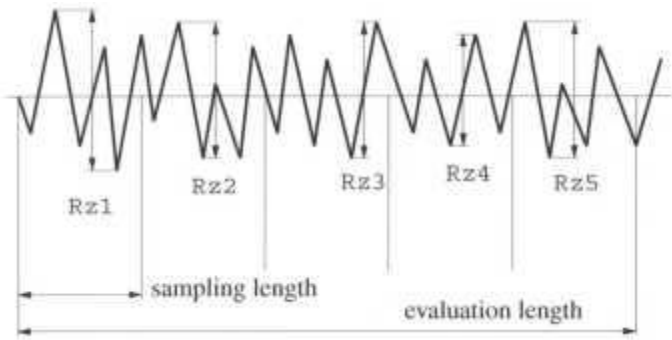
$$\mathrm{Rq} = \left(\frac{1}{l}\int_0^l z(x)^2 dx\right)^{1/2} \tag{1.3}$$

or, if the surface profile is measured in equidistant discrete points

$$Rq = \left(\frac{1}{N}\sum_{i=1}^{N} z_i^2\right)^{1/2} \tag{1.4}$$

Rz - **Mean peak-to-valley profile roughness.** The average peak-to-valley profile roughness is based on one peak and one valley per sampling length. The single largest deviation is found in five sampling lengths and then averaged, see Figure 1.8.

$$Rz = \frac{1}{5}(Rz1 + Rz2 + Rz3 + Rz4 + Rz5) \tag{1.5}$$



[billedtekst start]**Figure 1.8:** Sketch showing definition of Rz (ISO).[billedtekst slut]

The Rz-specification is slightly better than the Ra-specification to ensure good functional characteristics for the surface, but in fact none of the two secures a good bearing surface with good resistance to wear.

Rt - Maximum peak-to-valley height.

$$Rt = max(z) - mm(z) \tag{1.6}$$

For most roughness distributions we find that

$$Ra < Rq < Rz < Rt \tag{1.7}$$

and further the estimates

$$\frac{Rq}{Ra} \simeq 1.1 \tag{1.8}$$

and

$$Ra \cong 0.lRz \tag{1.9}$$

As seen in Figure 1.9 two surfaces with very different bearing characteristics could have the same Ra and Rz values. Especially for surfaces sliding on each other as in bearings, efforts have been made to develop a new method of specifying demands to the surface texture.

ISO 13565-2:1996 [6] defines a number of roughness parameters that may be used to characterize a surface in a more functional way than the classical parameters as for example Ra.

The parameters Rpk, Rk, Rvk, M_r**1**, and M_{r2} (see Figure 1.10) are all derived from the bearing ratio curve based on the ISO 13565-2:1996 standard. The bearing area curve is a measure of the relative cross-sectional area of a plane, passing through the measured surface, from the highest peak to the lowest valley.

		Rz	Ra
Parabola curve		1	0.26
Parabola curve		1	0.26

Figure 1.9: Sketch showing lack of functionality in Rz and Ra.

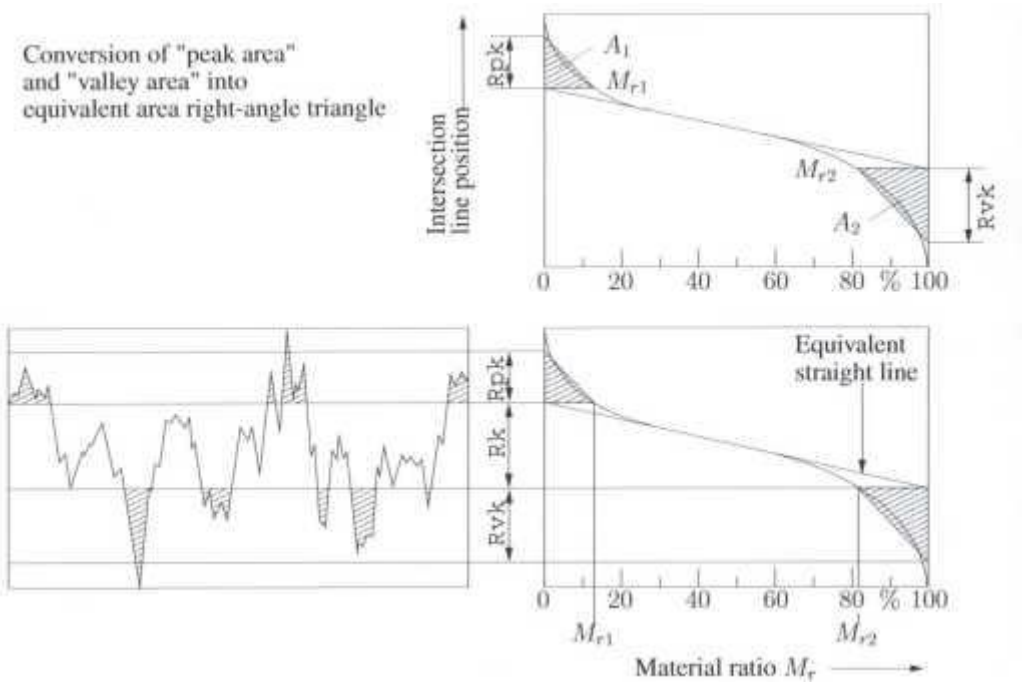
- Rpk, the reduced peak height is a measure of the peak height above the nominal/core roughness. These peaks will be the areas of most rapid wear when the machine is running.
- Rk, the core roughness depth is a measure of the nominal or "core" roughness (peak-to-valley) of the surface with the predominant peaks and valleys removed. It is the long term running surface which will influence the performance and life of the surface. (Also the load bearing area of the surface).
- Rvk, the reduced valley depth, is a measure of the valley depth below the nominal /core roughness. It is a measure of the oil retaining capability of the valleys of the surface produced during the machining process (for example plateau honing).
- M_{r1}, the peak material portion, indicates the percentage of material that comprises the peak structures associate with Rpk. Where the Rpk and Rk depths meet on the material ratio curve.
- M_{r2}, the valley material portion, relates to the percentage of the measurement area that comprises the deeper valley structures given by 100% M_r2. Where the Rvk and Rk depths meet on the material ratio curve.
- A_1, is the 'peak area' of the material ratio curve. It is calculated as the area of a right angled triangle of base length 0% to Mrl and height Rpk.
- A_2, is the 'valley area' of the material ratio curve. It is calculated as the area of a right angled triangle of base length Mr2 to 100% and height Rvk.

A high Rpk implies a surface composed of high peaks providing small initial contact area and thus high areas of contact stress (force/area), when the surface is contacted. Thus Rpk may represent the nominal height of the material that may be removed during a running-in operation. Consistent with Rpk, M_r1 represents the percentage of the surface that may be removed during running-in. Rk represents the core roughness of the surface over which a load may be distributed, once the surface has been run-in. Rvk, is a measure of the valley depths below the core roughness and may be related to lubricant retention and debris entrapment. Rk is a measure of the nominal roughness (peak to valley) and may be used to replace parameters such as Ra, Rt or Rz, when anomalous peaks or valleys may adversely affect the repeatability of

these (i.e. Ra, Rt and Rz) parameters.

The ratios of the various bearing ratio parameters Rpk/Rk (the reduced peak to core ratio), Rvk/Rk (the reduced valley to core ratio), and Rpk/Rvk (the reduced peak to reduced valley ratio) may be helpful in further understanding the nature of a particular surface texture. In some instances, two surfaces with indistinguishable roughness average (Ra) may be easily distinguished by a ratio such as Rpk/Rk. For example a surface with high peaks as opposed to a surface with deep valleys may have the same Ra, but with vastly different Rpk/Rk and Rvk/Rk values.

By considering the ratios such as Rpk/Rk, Rvk/Rk and Rpk/Rvk one may determine quantitatively the dominance of peak structures relative to valley structures. In typical tribological applications such as seals and brakes, these ratio may be useful in differentiating surfaces that have similar surface roughness as measured by Ra.



[billedtekst start]**Figure 1.10:** Definition of Rk, Rpk and Rvk from ISO 13565-2:1996 [6].[billedtekst slut]

Specifying surface texture requirements on drawings Examples are given in Figure 1.11, where the following definitions apply

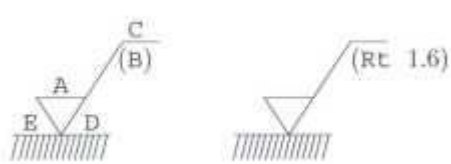
A = Surface texture requirements 1, Ra in micrometers

B = Surface texture requirements 2, Rz or Rt in micrometers

C = Manufacturing process - Turned, ground, plated . . .

D = Surface lay and orientation

E = Machining allowance



[billedtekst start]**Figure 1.11:** The surface texture symbols.[billedtekst slut]

Only quote surface texture where needed. Drilled through holes for bolts need normally no requirements for the surface texture. Bored holes for tight fits on the other hand require a surface quality corresponding to the tolerance specifications.

Guidelines for selection of a suitable surface finish When specifying a surface finish one should first pay attention to the function, and secondly to the manufacturing possibilities and price. Outer

surfaces without specific mechanical function often have to be specified according to "look", "feel" or "clean" conditions. (Ex. instruments for surgery or food processing machines). Surfaces with specific requirements for assembly conditions can be specified according to the following table.

Table 1.1: Guidelines for Arithmetical Average Roughness Ra.

Guidelines for Average Roughness Ra[μm]	0.05	0.1	0.2	0.4	0.8	1.6	3.2	6.3
Gauges for dimension control. Parts for roller bearings. High speed journal bearings. Surfaces for high capacity journal bearings.	-xX	XX	XX	XX-				
Normal bearings and guidance surfaces. Translating and rotating parts against sealingŠs. Surfaces for coating to mirror blank purposes.			-xx	XX	XX	XX-		
Normal for high stressed shafts. Static surfaces with contact to rubber sealingŠs. Surfaces for coating. Seats for ball and roller bearings.				-xx	XX	XX	XX-	
Plain surfaces to be sealed without a gasket. Contacting surfaces for accurate details. Flanks on splines, threads and similar details.					-xx	XX	Xx-	
Plain surfaces to be sealed with a gasket. Normal contacting surfaces in assemblies. Flanks on splines, threads and similar details.						-xx	XX	Xx-
Surfaces without specific mechanical function demands.							-xx	XX→

Another guideline for permissible surface roughness can be derived from functional demands specified by tight tolerances. It is obvious that fine tolerances in a fit are of no relevance if combined with a "rough surface".

Empiric formulas link permissible roughness value to specified tolerance grade IT (See definition later in this chapter).

$$\begin{aligned} Ra &\approx \frac{IT}{12} \quad \text{to} \quad \frac{IT}{9} \\ Rz &\approx \frac{IT}{4} \quad \text{to} \quad \frac{IT}{3} \end{aligned} \tag{1.10}$$

Example:

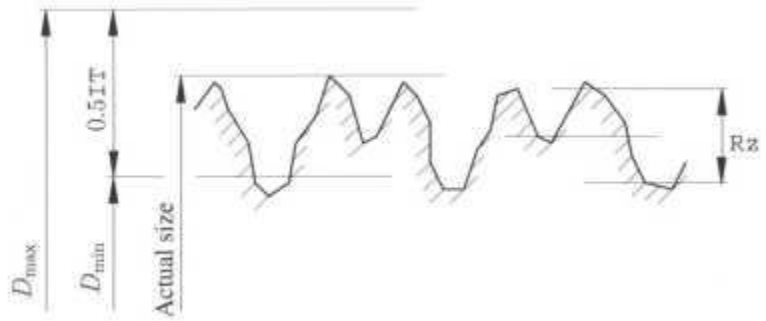
A shaft end has to be machined to 03Ok6 (prepared for mounting a coupling) The average roughness to be specified is:

$$\mathrm{IT6}_{\varnothing 30} = 13\mu m \quad \Rightarrow \quad \mathrm{Ra}_{max} \approx \frac{13}{9} = 1.44\mu m \tag{1.11}$$

closest standardized value to be chosen is Ra = 1.6μm

1.4 Tolerances on lengths, diameters, angles.

Appropriate manufacturing of components require that the dimensions specified on drawings, need to show the acceptable upper and lower limits of size. Within reason, these limits should be as generous as possible in order to keep down manufacturing costs. Obviously, there are situations where it is necessary



[billedtekst start]**Figure 1.12:** Sketch showing surface profile, and limits of size for cylindrical object.[billedtekst slut]

to quote very tight limits in order to provide a particular fit. For instance, it may be necessary for two components that have the same nominal dimensions to be assembled with a transition fit, or the components may need to be pressed together. Clearly the limits on the size of the components, will dictate the kind of fit obtainable. Generally, the designer cannot leave any dimension without a size tolerance. Most dimensions can be covered by an overall drawing tolerance, but areas where particular fits are necessary, need to be identified and given appropriate tolerances. Interchangeability of components is one of the major reasons for using tolerances. It is impossible to guarantee that components will fit together without using comprehensive tolerance specification on interfaces. It is necessary to be able to produce components in batches at any time, in different locations and still be able to guarantee the fit. Imagine you are going to buy a spare part for your car or motor cycle and find that it will not be able to fit, because the size is incorrect. The manufacturer must be able to rely on the tolerance specified on the drawing in order to be able to produce fully interchangeable components with the correct interfacing condition. On the other hand, one off or prototype components and assemblies do not necessarily need comprehensive tolerances. It is often sufficient to allow one component to be machined to fairly relaxed tolerances whilst specifying the mating component to be machined to a specific fit, quoting only the clearance or interference required.

1.4.1 Dimensions and tolerances

When dimensioning components it is appropriate to distinguish between functional dimensions and dimensions of less importance for the function, but necessary for the manufacturing. A third class of dimensions is for general information, but of minor importance for function and manufacturing. All functional dimensions are to be limited with tolerances such that adequate functions will be achieved. These limits are used to define the lower and the upper limit of size. The difference between these two dimensions is called the tolerance. The tolerance is the workspace for the production. It is important to realize that the tolerance must not be regarded as an uncertainty in the production.

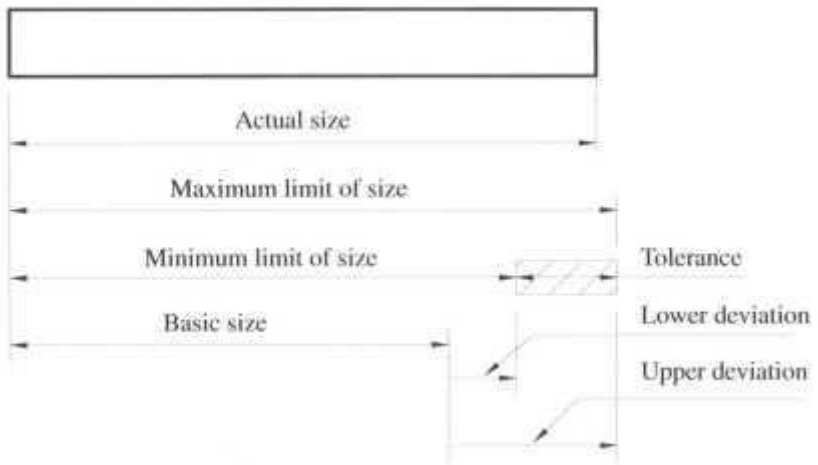
Tolerance definitions:

Actual size (of a part): The size of a part as obtained by measurement.

Maximum limit of size: The greater of the two limits of size.

Minimum limit of size: The smaller of the two limits of size.

Basic size; nominal size: The size from which the limits of size are derived by the application of the upper and lower deviations. See Figure 1.13. The basic size can be a whole number or a decimal number, e.g. 32; 8.75; 0.5; etc.



[billedtekst start]**Figure 1.13:** Tolerance definitions.[billedtekst slut]

Actual deviation: The algebraically difference between the actual size and the corresponding basic size.

Lower deviation: The algebraically difference between the minimum limit of size and the corresponding basic size.

Upper deviation: The algebraically difference between the maximum limit of size and the corresponding basic size.

Tolerance: The difference between the maximum limit of size and the minimum limit of size, or (in other words) the algebraic difference between the upper deviation and the lower deviation. The tolerance is an absolute value without sign.

1.4.2 Fits

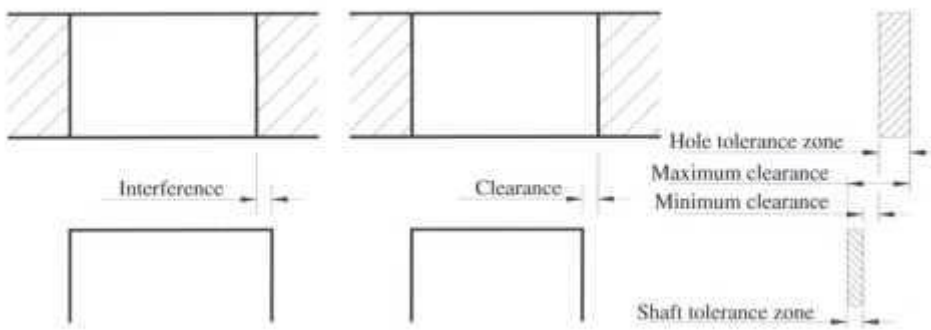
Fits are the (before) assembly relations between two or more parts, all with their own tolerances. The ordinary fit calculations are in one plane and very often only in one direction. A fit calculation result in either a clearance or an interference. Interferences are normally only acceptable for shaft and hub connections to make a shrink fit or a pressure fit.

Fits definitions

Clearance: The positive difference between the size of the "hole" and the "shaft" before assembly, when the dimension of the shaft is smaller than the dimension of the hole. See Figure 1.14.

Interference: The negative difference between the size of the "hole" and the "shaft" before assembly, when the dimension of the shaft is larger than the dimension of the hole.

Fit: The relationship resulting from the difference before assembly, between the two sizesof the two parts that are to be assembled.



[billedtekst start]**Figure 1.14:** Fits definitions.[billedtekst slut]

Clearance fit: A fit that always provides a clearance between the "hole" and the "shaft" when assembled, i.e. the minimum size of the "hole" is either greater than or equal to the maximum size of the "shaft".

Interference fit: A fit that always provides an interference between the "hole" and the "shaft" when assembled, i.e. the maximum size of the "hole" is either smaller than or equal to the minimum size of the "shaft".

Transition fit: A fit that may provide either a clearance or an interference. (The tolerance zone of the "hole" and the "shaft" overlap.)

1.4.3 The quality function deployment

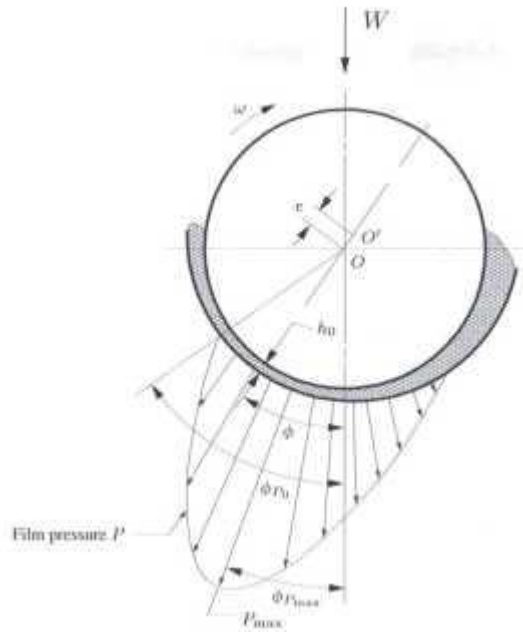
The importance of "choosing" a proper clearance fit for a journal bearing is obvious since in fact it incorporates the load carrying capacity of a journal bearing is strongly dependent on the minimum oil film thickness h_0. For given running conditions (speed, oil viscosity, etc.) it is possible to calculate an optimum oil film thickness h_0 related to the clearance between the journal and the bearing. Because of the manufacturing process, the journal and the bearing must be limited with such tolerances that the required load carrying capacity for the bearing will be achieved for all bearings with those tolerances. An overall high load carrying capacity for all bearings requires fine tolerances. Sometimes it can be reasonable to use the quality function deployment concept in the discussion between the design department and the manufacturing department. A quality function Q can be defined as the fraction between the achieved function and the specified or required function.

$$Q = \frac{\text{achieved function}}{\text{specified function}} \tag{1.12}$$

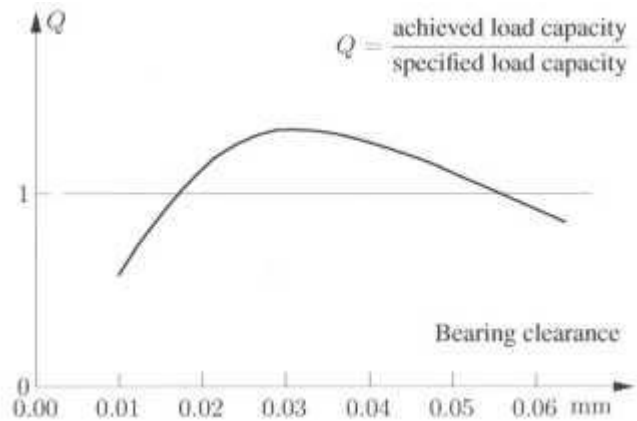
In the above example with the journal bearing the quality function can be expressed as a function of the clearance in the bearing. See Figure 1.15.

1.4.4 Functional dimensioning

The functional dimensioning should be expressed directly on the drawing. The application of this principle will result in the selection of datum features based on the function of the product. The method of locating it is to look into the assembly of which it may form a part. If any datum feature other than one based on the function of the product is used, finer tolerances will be necessary. Products, which would satisfy the functional requirements may be rejected because they exceed these finer tolerances.



[billedtekst start]**Figure 1.15:** Journal bearing.[billedtekst slut]

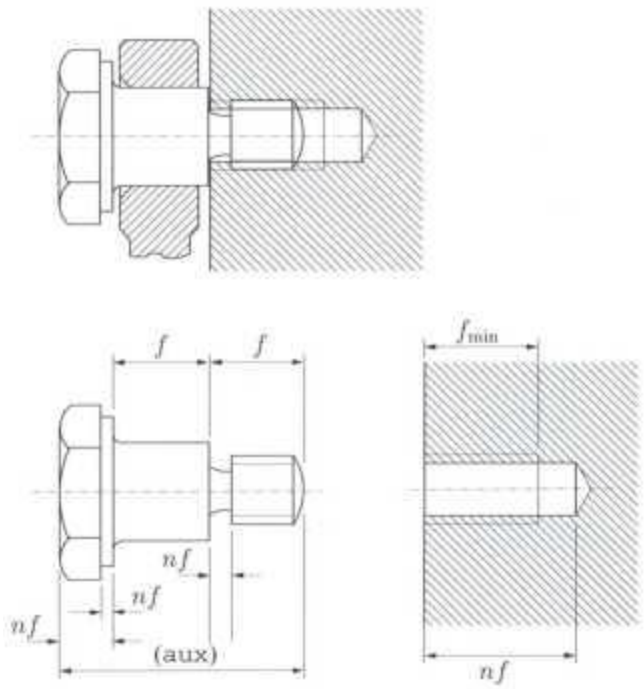


[billedtekst start]**Figure 1.16:** Quality function for a specific journal bearing.[billedtekst slut]

This does not preclude the preparation of special drawings dimensioned from a common datum point, to suit particular numerical controlled machining systems, where it is known that the overall accuracy of the system to be used will be adequate to ensure the finer tolerance arising from dimensioning other than directly from functional datum features, are met.

A dimension is not complete without a tolerance although the tolerance may not always appear with the dimension.

Functional dimensions shall always be shown with explicitly specified tolerances, whereas nonfunctional dimensions may be left without tolerances. In that case information to the manufacturer is



[billedtekst start]**Figure 1.17:** Functional dimensioning, f: functional dimension, nf: non-functional dimension, aux: auxiliary dimension without tolerances (for information only).[billedtekst slut]

to be given elsewhere on the drawing as a reference to standards as ISO 2768-1: General tolerances for linear and angular dimensions without individual tolerance indications [7],

Table 1.2: Permissible variations for linear dimensions.

Nominal dimensions		0.5 up to 3	Over 3 up to 6	Over 6 up to 30	Over 30 up to 120	Over 120 up to 315	Over 315 up to 1000	Over 1000 up to 2000
Permissible	Fine series	±0.05	±0.05	±0.1	±0.15	±0.2	±0.3	±0.5
variations	Medium series	±0.1	±0.1	±0.2	±0.3	±0.5	±0.8	±1.2
	Coarse series	–	±0.2	±0.5	±0.8	±1.2	±2	±3

Dimensioning a shaft to a gearbox will show the idea of functional dimensioning. The shaft is located in axial direction in the gearbox by the two ball bearings. Besides the requirement that there shall be full tooth contact all the time it is required that no interference during assembly operation or during running conditions occurs.

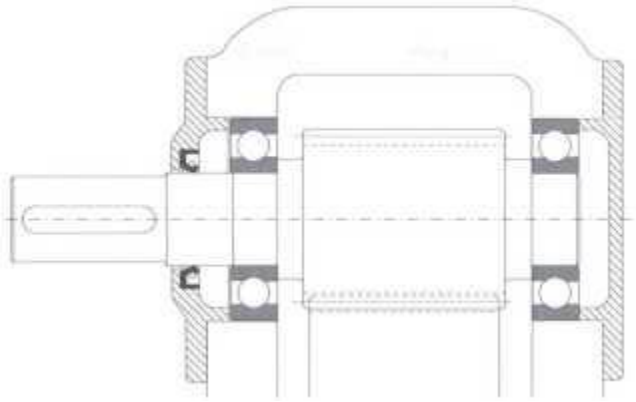
An often-used praxis has been to dimension with basis in the ends of the shaft to make it

"easier" for the manufacturing department in the sense that the dimensions directly reflect the manufacturing process by giving the dimensions that the lathe tool has to move during cutting.

The problem by using this method of dimensioning is that the distance of importance for the axial clearance is given indirectly. Working with absolute tolerances this (functional important) distance can be calculated:

The maximum is: 212.2 – 97.85 – 17.9 = 96.45

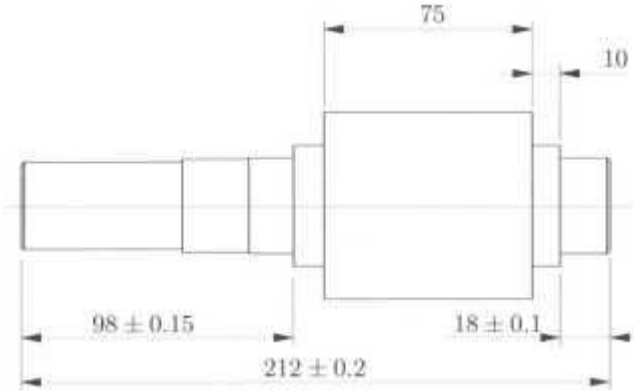
The minimum is: 211.8 – 98.15 – 18.1 = 95.55



[billedtekst start]**Figure 1.18:** Part of gear assembly drawing.[billedtekst slut]

The absolute tolerance on this resulting distance is 0.9 mm which is the same as the sum of the tolerances for the dimensions of relevance for the distance. (0.4 + 0.3 + 0.2 = 0.9).

1.4.5 Dimension chains

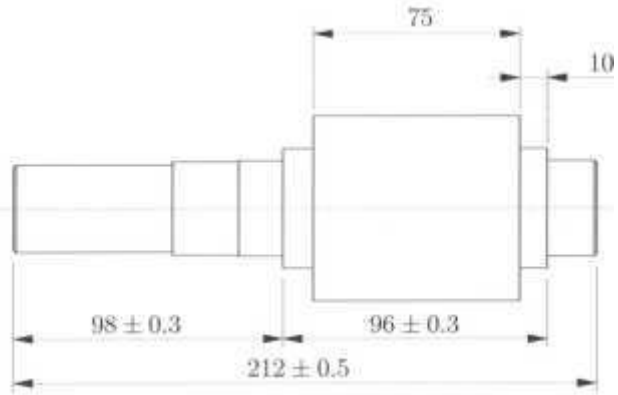


[billedtekst start]**Figure 1.19:** Gear shaft with manufacturing related dimensions.[billedtekst slut]

Because of the tolerance stacking problem it is important to dimension with as few dimensions in the chain as possible (and feasible). Changing the method of dimensioning the shaft for it to include the functional dimension may be advantageous. The dimension chain of importance for the axial clearance will be shorter. The tolerances can often be chosen bigger. The non-functional dimensions can be given even bigger tolerances resulting in lower manufacturing costs. See Figure 1.20.

Calculation of clearance.

Related to the gear assembly drawing (Figure 1.18) we have the information of interest in Figure 1.21 for the calculation of the clearance between the left end cover and the outer ring of the ball bearing.



[billedtekst start]**Figure 1.20:** Gear shaft, more functionally dimensioned.[billedtekst slut]

[billedtekst start]**Figure 1.21:** Dimension chain for the shaft mounted in the gearbox.[billedtekst slut]

First simplifying by looking at the actual dimensions and calculate the actual gab ***w.*** Consider the dimensions in Figure 1.21 exchanged with actual dimension vectors x_1, x_2, ... for all right tending and y_1,y_2,y_3, ... for all left tending vectors, starting in the upper left corner and working clockwise as shown in Figure 1.22. The working direction 8clockwise or counterclockwise) is determined from the location of the clearance in the chain. To get the right sign on the clearance, it needs to be located on "the way back" to the chain origo.

[billedtekst start]**Figure 1.22:** Dimensions in chain regarded as vectors.[billedtekst slut]

The actual clearance can now be calculated to

$$w = (x_1 + x_2 + ...) - (y_1 + y_2 + y_3 + ...) = \Sigma_x - \Sigma_y \qquad (1.13)$$

Calculating the *worst case* max. and min. value of *w*

$$w_{max} = \Sigma x_{max} - \Sigma y_{max} \quad (1.14)$$

$$w_{min} = \Sigma x_{min} - \Sigma y_{min} \quad (1.15)$$

Back to the gear assembly example the clearance can be calculated to

$$w_{max} = 148.3 - (8.8 + 16.88 + 95.7 + 16.88 + 7.8) = 2.24 \quad (1.16)$$

$$w_{min} = 147.7 - (9.2 + 17 + 96.3 + 17 + 8.2) = 0 \quad (1.17)$$

The above calculations are valid lor an *absolute tolerance system.* This is one in which no instances of the tolerances are outside the range. A *worst case* calculation presumes that all the tolerances in the chain are on their most dangerous limit. That is of course not the case in practical life. With a "number of dimensions" in the chain the clearance value will normally be concentrated symmetrically around the mean value of w $(\bar{w})$.

1.5 The ISO-tolerance system

1.5.1 Introduction

The need for limits and fits for machined work pieces was brought about mainly by the inherent inaccuracy of manufacturing methods, coupled with the fact that "exactness" of size was found to be unnecessary for most work pieces. In order for the function to be satisfied, it was found sufficient to manufacture a given work piece so that its size lay within two permissible limits i.e. a tolerance, this being the variation in size acceptable in manufacture.

Similarly, where a specific fit condition is required between mating work pieces, it is necessary to ascribe an allowance either positive or negative to the basic size in order to achieve the clearance or interference required, i.e. a "deviation".

With developments in industry and international trade it became necessary to develop formal systems of limits and fits. ISO 286 [8] describes the internationally accepted system of limits and fits.

1.5.2 Field of application

The ISO system of limits and fits provides a system of tolerances and deviations primary for cylindrical work pieces. However, the tolerances and deviations apply to work pieces of other than circular section as well.

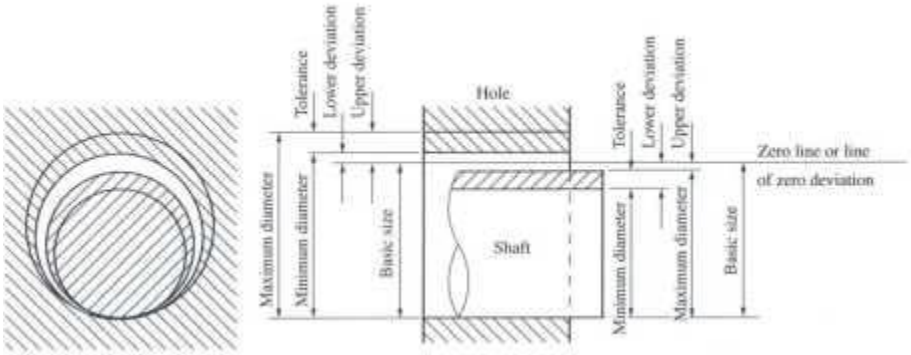
The general term "hole" or "shaft" can be taken as referring to the space contained by (or containing) the two parallel faces (or tangent planes) of any work piece such as the width of a slot or the thickness of a key.

The system provides fits between mating cylindrical holes and shafts or fits between work pieces having features with parallel faces such as the fit between a key and keyway.

1.5.3 Terms and definitions

The following terms and definitions extend the previously given ones:

Standard tolerance (IT): Any tolerance belonging to ISO system. The letters of the symbol IT stand for "International Tolerance" grade.



[billedtekst start]**Figure 1.23:** Basic size and limits of size.[billedtekst slut]

Standard tolerance grades: A group of tolerances (e.g. IT7) is considered as corresponding to the same level of accuracy for all basic sizes.

Tolerance zone: In a graphical representation of tolerances -the zone- contained between two lines representing the maximum and minimum limits of size. It is defined by the magnitude of the tolerance and its position relative to the zero line.

Tolerance class: The term used for a combination of fundamental deviation and a tolerance grade, e.g. h9, D13.

Standard tolerance factor *i* **and** *I*: A factor which is a function of the basic size, and which is used as a basis for the determination of the standard tolerances of the system. The standard tolerance factor ***i*** is applied to basic sizes less than or equal to 500 mm, whereas *I* is applied to basic sizes greater than 500 mm.

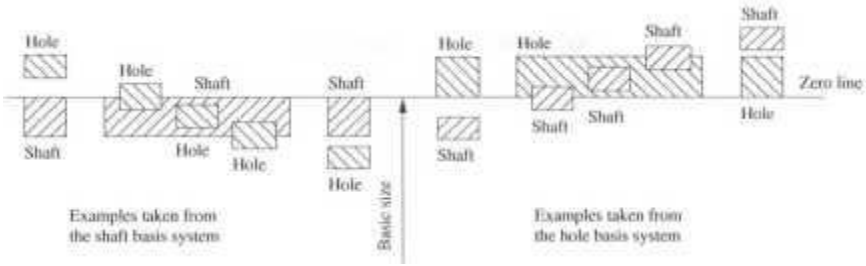
Fit system: A system of fits comprising shafts and holes belonging to a limit system.

Shaft-basis system of fits: A system of fits in which the required clearances or interferences are obtained by associating holes of various tolerance classes with shafts of a single tolerance class. The maximum limit of size of the shaft is identical to the basic size i.e. the upper deviation is zero see Figure 1.26.

Hole-basis system of fits: A system of fits in which the required clearances or interferences are obtained by associating shafts of various tolerance classes with holes of a single tolerance class. The minimum limit of size of the hole is identical to the basic size, i.e. the lower deviation is zero see Figure 1.24.

Maximum material limit (MML): The designation applied to that of the two limits of size which corresponds to the maximum material size for the feature i.e. the upper limit of size for a shaft and the lower limit of size for a hole.

Least material limit (LMU): The designation applied to that of the two limits of size which corresponds to the minimum material size for the feature i.e. the lower limit of size for a shaft and the upper limit of size for a hole.



[billedtekst start]**Figure 1.24:** Examples of fits.[billedtekst slut]

1.5.4 Tolerances and deviations

Basically, the tolerances for holes and shafts can be written as for other functional dimensions, as indicated in Figure 1.25.

In the ISO-tolerance system the same dimension can be written as:

Ø80G7

where

80 = Nominal size.

G = Information about the tolerance zone deviation.

In this example given by the lower deviation.

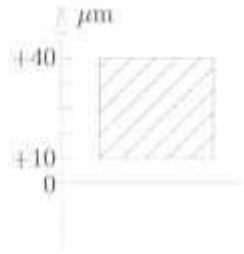
Generally the deviation closest to the zero-line is used.

Upper case (capital) letters: holes.

Lower case (small) letters: shafts.

7 = A code for the size of the tolerance. (Tolerance grade)

The tolerance value is a function of the nominal size.



[billedtekst start]**Figure 1.25:** Example: Hole $Ø80^{+0.040}_{+0.010}$ Nominal size: 80 Upper deviation: +0.040 Lower deviation: +0.010 Tolerance: 0.030mm or 30μm[billedtekst slut]

Standard tolerance grades

The standard tolerance grades are designated by the letters IT followed by a number, e.g. IT7. When the tolerance grade is associated with a letter representing a fundamental deviation to

form a tolerance specification the letters IT are omitted, e.g. h7.

Deviations

The position of the tolerance zone with respect to the zero line, is designated by upper case letters for holes (A .. , ZC) or lower case letters for shafts (a . . . zc), see Figure 1.26.

Designation

A tolerance class is designated by the letter representing the fundamental deviation (location) followed by the number representing the standard tolerance grade.

Examples: H7 (holes); h7 (shafts)

A toleranced size shall be designated by the basic size followed by the designation of the required tolerance class or the explicit deviations.

Examples: $32H7;\ 80js15;\ 100g6;\ 100^{-0.012}_{-0.034}$

A fit requirement between mating features shall be designated by the common basic size, the tolerance class symbol for the hole and the tolerance class symbol for the shaft.

Examples: $52H7/g6$ or $52\frac{H7}{g6}$

Tolerance tables can be found in Appendix A.

1.5.5 Preferred numbers

The French army engineer Col. C. Renard proposed in the late 19th century a set of preferred numbers for use with the metric system.

The system of numbers divides the interval from 1 to 10 into 5, 10, 20, or 40 steps respectively. The factor between two consecutive numbers in a Renard series is constant. The constant is the 5th., 10th., 20th., or 40th. root of 10.

The basic equation for the sequences therefore is

$$q = 10^{(j/b)},\ j = 0..n \tag{1.18}$$

where j is a integer number between $n = 0$ and $n = b - 1$. The integer b is the base number, i.e., $b = 5$, $b = 10$, $b = 20$ or $b = 40$.

The most basic R5 ($b = 5$) series therefore is

$$q_5 = 10^{(j/5)},\ j = 0..4 \tag{1.19}$$

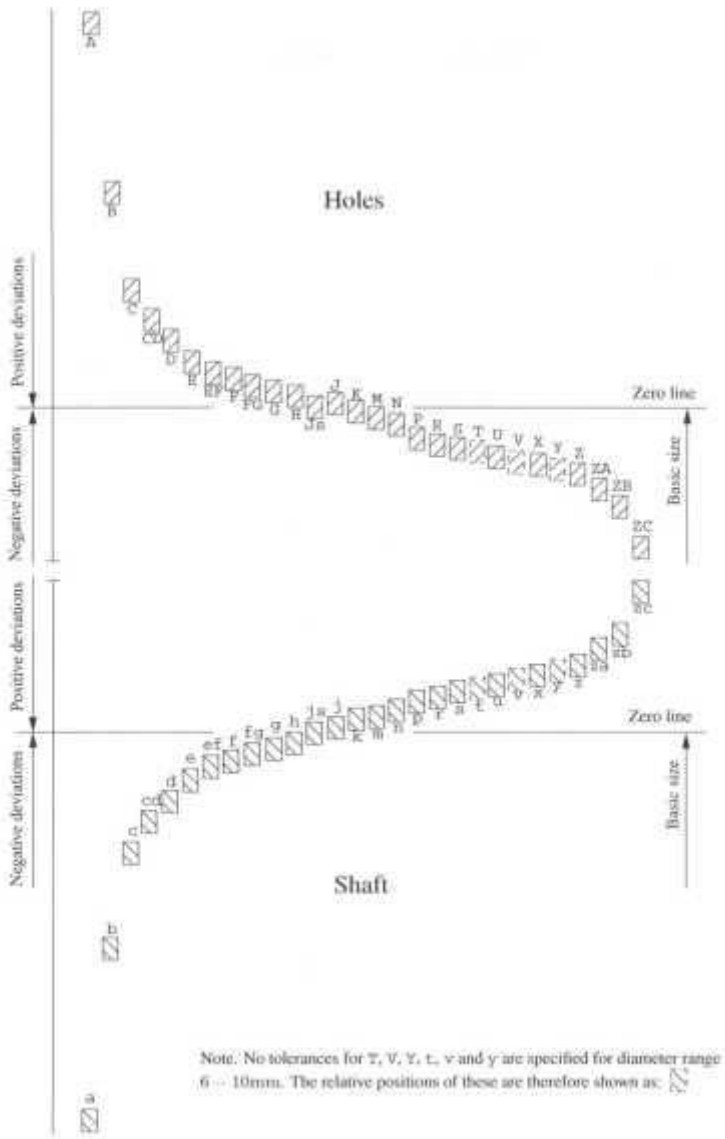
if a finer resolution is needed, we use the R10 ($b = 10$) series

$$q_{10} = 10^{j/10},\ j = 0..9 \tag{1.20}$$

where an even finer grading is needed, the R20 or R40 series can be applied. The values that result from the Renard series need to be rounded in order not to specify a too fine tolerance. The rounded values are defined in ISO 3.

The rounded version of Renard series R5, R10, R20 and R40 are given in Table 1.3

together with the % deviation of the Renard value to the ISO value.



[billedtekst start]**Figure 1.26:** Position of tolerance zones.[billedtekst slut]

1.5.6 Standard tolerance grades IT1 to IT16

The values for standard tolerances ITl to IT16 for basic sizes up to 500mm are determined as a function of the standard tolerance factor ***i.*** The standard tolerance factor i in [μm] is calculated from (1.21)

$$i = (0.45\sqrt[3]{\frac{D}{\text{mm}}} + 0.001\frac{D}{\text{mm}})\mu\text{m} \tag{1.21}$$

where D is the geometric mean of the basic size step in [mm]. If the extreme values of a step is D_1 and $D2$ then the geometric mean is defined as

Table 1.3: Preferred numbers from ISO 3. The R5, R10, R20 and R40 series.

R5	R10	R20	R40	% deviation from (1.18)
1.00	1.00	1.00	1.00	0
-	-	-	1.06	+0.07
-	-	1.12	1.12	-0.18
-	-	-	1.18	-0.71
-	1.25	1.25	1.25	-0.71
-	-	-	1.32	-1.01
-	-	1.40	1.40	-0.88
-	-	-	1.50	+0.25
1.60	1.60	1.60	1.60	+0.95
-	-	-	1.70	+1.26
-	-	1.80	1.80	+ 1.22
-	-	-	1.90	+0.87
-	2.00	2.00	2.00	+0.24
-	-	-	2.12	+0.31
-	-	2.24	2.24	+0.06
-	-	-	2.36	-0.48
2.50	2.50	2.50	2.50	-0.47
-	-	-	2.65	-0.40
-	-	2.80	2.80	-0.65
-	-	-	3.00	+0.49
-	3.15	3.15	3.15	-0.39

-	-	-	3.35	+0.01
-	-	3.55	3.55	+0.05
-	-	-	3.75	-0.22
4.00	4.00	4.00	4.00	+0.47
-	-	-	4.25	+0.78
-	-	4.50	4.50	+0.74
-	-	-	4.75	+0.39
-	5.00	5.00	5.00	-0.24
-	-	-	5.30	-0.17
-	-	5.60	5.60	-0.42
-	-	-	6.00	+0.73
6.30	6.30	6.30	6.30	-0.15
-	-	-	6.70	+0.25
-	-	7.10	7.10	+0.29
-	-	-	7.50	+0.01
-	8.00	8.00	8.00	+0.71
-	-	-	8.50	+1.02
-	-	9.00	9.00	+0.98
-	-	-	9.50	+0.63
10.0	10.0	10.0	10.0	0

$$D = \sqrt{D_1 D_2} \tag{1.22}$$

1.5.7 Formula for standard tolerances in grades IT5 to IT16

The size of the standard tolerance grades follow the numbers given in Table 1.4. It can be seen that the numbers follows the R5 series although the value 6.3 for some reason have been

interchanged with either the value 7.0 or the value 6.4.

Example

Table 1.4: Formula for standard tolerances.

	IT 5	IT 6	IT 7	IT 8	IT 9	IT 10	IT 11	IT 12	IT 13	IT 14	IT 15	IT 16
Value	7 *i*	10*i*	16*i*	25*i*	40*i*	64i	100*i*	160*i*	250*i*	400*i*	640*i*	1000*i*

Calculate the standard tolerance IT7 for the diameter range $D \in]50\text{mm}: 80\text{mm}]$.

$$D = \sqrt{50 \cdot 80}\text{mm} = 63.246\text{mm}$$
$$i = (0.45\sqrt[3]{63.246} + 0.001 \cdot 63.246)\mu\text{m} = 1.856\mu\text{m} \tag{1.23}$$

$IT7_{50\text{–}80} = 16i = 29.7\mu$m rounded up to 30μm compare with Table A. 1 in Appendix A.

1.6 Nomenclature

aux	mm	Auxiliary dimension
b	–	Base number
i	jum	Standard tolerance value
f	mm	Functional dimension
nf	mm	Non functional dimension
q	–	A prefered number
w	mm	Gab (clearance)
x_1, x_2, ...	mm	Dimensions
y_1,y_2,...	mm	Dimensions
D	mm	Geometric mean
Q	–	Quality function (number)
Ra	μm	Surface roughness (Arithmetic mean)
Rk	μm	Surface roughness parameter
Rpk	μm	Surface roughness parameter
Rq	μm	Surface roughness parameter

Rt	μm	Surface roughness parameter
Rvk	μm	Surface roughness parameter
Rz	μm	Surface roughness parameter

1.7 References

[1] BS 4500. ISO *limits and fits.*

[2] ISO 1101. Technical drawings - geometrical tolerancing - tolerancing of form, orientation, location and run-out - generalities, definitions, symbols, indications on drawings.

[3] ISO 129-1. Technical drawings - indication of dimensions and tolerances - part 1: General principles.

[4] ISO 129-2. Geometrical product specifications (gps) - indication of dimensions and tolerances - part 2: Mechanical engineering drawings.

[5] ISO 1302:2002. Geometrical product specifications (gps) - indication of surface texture in technical product documentation.

[6] ISO 13565-2:1996. Geometrical product specifications (gps) - surface texture: Profile method; surfaces having stratified functional properties - part 2: Height characterization using the linear material ratio curve.

[7] ISO 2768-1. General tolerances for linear and angular dimensions without individual tolerance indications.

[8] ISO 286. ISO system of limits and fits.

Chapter 2 Springs

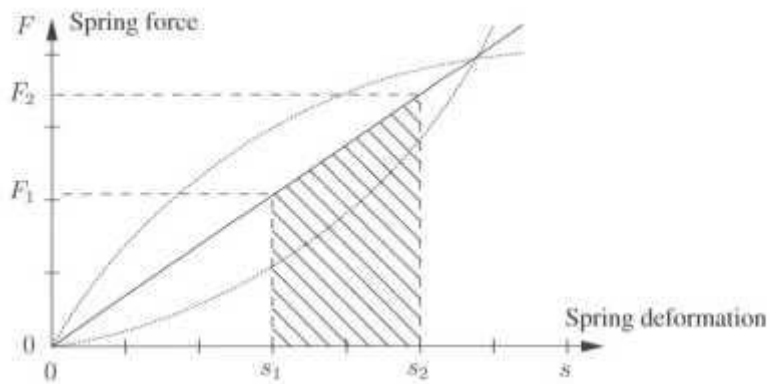
2.1 Introduction

The purpose of using springs can fall into a wide range between the two following extremities:

1. As accumulators for energy. An example is the balancing spring in the hood of an automobile or the balancing spring in a watch. In both cases the energy loss are to be minimized.
2. As energy absorbers. An example is the buffer springs between goods wagons. The springs are in this case designed to absorb most of the energy caused by the collision when shunting and when released transform most of the energy into friction heat.

But for normal springs (at least metal springs) the characteristic functionality will be closest to the first extremity.

In the design situation for a suitable spring the starting point will normally be the working diagram for the spring.

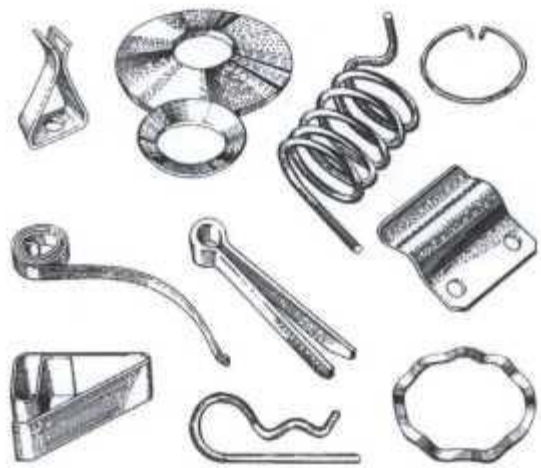


[billedtekst start]**Figure 2.1:** Working diagram for a spring. Shown are both a linear and two nonlinear spring characteristics.[billedtekst slut]

2.2 The design situation

Often the goal is to design a spring with the lowest possible weight which are able to fit properly into the actual application. This is of course of special interest when designing springs for vehicles as automobiles

and airplanes. In other cases the design criteria is of a more complex nature as the spring examples in Figure 2.2 could indicate.



[billedtekst start]**Figure** 2.2: A variety of spring elements.[billedtekst slut]

User specific parameters will normally be related to the spring rate. As is seen from Figures 2.3 and 2.4, relations between the parameters can be shown graphically. The following expressions can be set up.

2.3 Helical springs

The spring rate/stiffness is given as

$$K(s) = \frac{dF}{ds} \tag{2.1}$$

For the helical spring it can be assumed that the spring rate is constant. From this follows that

$$F = Ks \tag{2.2}$$

under the assumption that the displacement ***s*** is measured relative to the unloaded state, i.e., where $F = 0$. The rate can in this case be given as

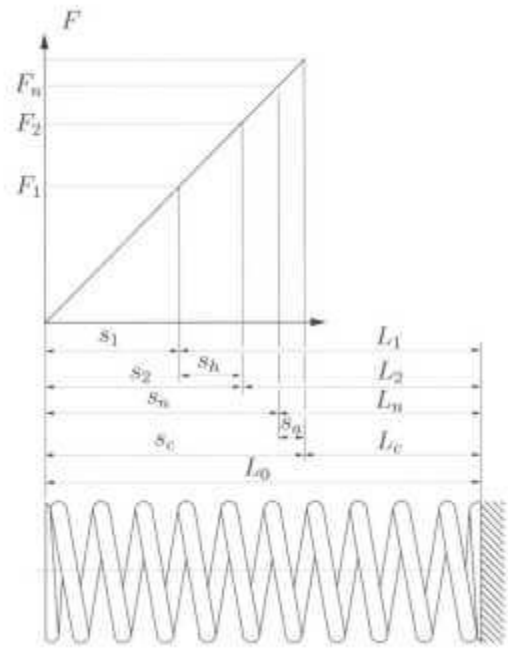
$$K = \frac{F_2 - F_1}{s_2 - s_1} = \frac{F_2 - F_1}{s_h} \tag{2.3}$$

where s_h is a working deflection, see Figure 2.3. The work done by spring deformation is given as

$$W = \int_{s_1}^{s_2} F(s)ds = \int_{s_1}^{s_2} K(s)sds \tag{2.4}$$

In this case with a constant spring rate this simplifies to

$$W = \int_{s_1}^{s_2} Ksds = \frac{Ks_2^2 - Ks_1^2}{2} = \frac{F_2 s_2 - F_1 s_1}{2} \tag{2.5}$$



[billedtekst.start]**Figure 2.3:**Spring diagram for compression spring.[billedtekst slut]

It is normally not all the parameters that are of interest for the designer. Often it is only F_1 F_2 and s_h. There may be requirements to the geometric data of a spring, so that for example its diameter is inside certain limits. Typically, specifications are given for D_e, D_i, L_0, L_1, L_2, L_c and s_a. L_c and s_a are of course only specified for compression springs. Extension springs will very often be manufactured with initial tension. This means that the spring will not deflect until this initial tension is overcome. Hot-formed springs cannot be manufactured with initial tension.

Often the user will set up requirements for the tolerance of the spring. The tolerance may be specified for as well geometrical data as for spring force and the spring rate. Statically loaded springs will be designed to a certain maximum relaxation.

Additionally, requirements can be set up for the natural frequency, ability to absorb shock loads, corrosion resistance, electrical conductivity, operational temperature etc.

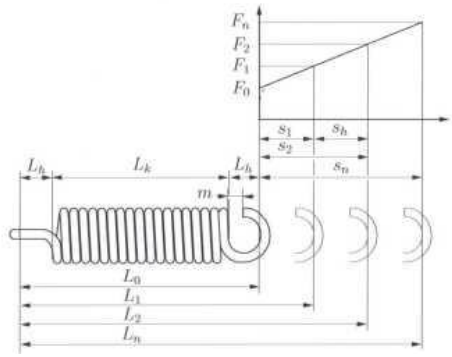
2.3.1 Formulas for helical springs

When a helical spring is loaded with an external load F it will deflect a distance s, see Figure 2.1. The associated external work is

$$W = \frac{Fs}{2} \tag{2.6}$$

As the helical spring deflects the wire will be twisted an angle α due to the torsional moment T, where α is measured at the total length of the wire. The wire twisting can be seen as

the internal work in the wire



[billedtekst.start]]**Figure 2.4:** Spring diagram for extension spring.[billedteskst.slut]

$$W = \frac{T\alpha}{2} \tag{2.7}$$

giving

$$s = \frac{T\alpha}{F} \tag{2.8}$$

When a prismatic beam is loaded by a torsional moment T the shear angle (shear strain) becomes

$$\gamma = \frac{\tau_t}{G} \tag{2.9}$$

where

$$\tau_t = \frac{T}{W_p} = \frac{T}{I_p} r = \frac{T}{\frac{\pi}{32} d^4} r = \frac{T}{\frac{\pi}{16} d^3} \tag{2.10}$$

Finally, the relation between twist angle α and the shear angle γ is

$$\alpha = \frac{\gamma l}{r} = \frac{Tl}{GI_p} \tag{2.11}$$

Combining (2.6) to (2.11) the fundamental formulas for springs may be derived Load on spring

$$F = \frac{Gd^4 s}{8D^3 n} \tag{2.12}$$

Spring deflection

$$s = \frac{8D^3 nF}{Gd^4} \tag{2.13}$$

Spring rate

$$K = \frac{Gd^4}{8D^3n} \tag{2.14}$$

Shear stress

$$\tau_t = \frac{8DF}{\pi d^3} \tag{2.15}$$

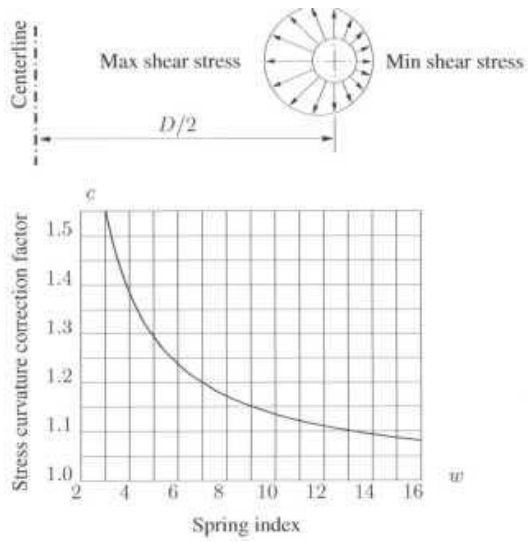
Wire diameter

$$d = \left(\frac{8FD}{\pi \tau_t}\right)^{1/3} \tag{2.16}$$

Active coils

$$n = \frac{Gd^4 s}{8D^3 F} \tag{2.17}$$

2.3.2 Stress curvature correction factor



[billedtekst.start]**Figure 2.5:** Stress curvature correction factor as a function of the spring index.[billedteskst.slut]

The stresses in a spring are unevenly distributed in the wire cross section. The stress is greater on the inside of the coil than on the outside. The curvature correction factor c takes this effect into account. The stress curvature correction factor depends on the spring index $w = \frac{D}{d}$.

$$c = \frac{w + 0.5}{w - 0.75} \tag{2.18}$$

2.3.3 Material properties

Material properties for springs may be found in numerous places. Standards classifies the materials into groups, see [3], [4], [5], [6] and [7]. Material vendors normally distribute updated material data on request.

Table 2.1: Examples of commonly used materials for springs.

Material	*E*	*G*	*ρ*
-	N/mm²	N/mm²	kg/m³
Cold drawn unalloyed steel spring wire, DIN 17223 part 1	206000	81500	7850
/EN10270-1			
Oil quenched and tempered spring steel wire of unalloyed and alloyed steels, DIN 17223 part 2 /EN10270-2	206000	79500	7850
Hot rolled steels for quenched and tempered springs, DIN 17221	206000	78500	7850
Stainless steel after DIN 17224/EN10088: X 12 CrNi 17 7	185000	70000	7900
X 7 CrNiAl 17 7	195000	73000	7900
X 5 CrNiMo 18 10	180000	68000	7900
Tin-bronze CuSn6 F95 after DIN 17682/EN1654 hard-drawn	115000	42000	8730
Cobber-Zinc-Alloy CuZn36 F70 after DIN 17682 hard-drawn	110000	39000	8400
Cobber-Beryllium-Alloy CuBe2 after DIN 17682	120000	47000	8800
Cobber-Cobolt-Beryllium-Alloy CuCoBe after DIN 17682	130000	48000	8800

2.3.4 Relaxation

The relaxation of a spring depends on time, temperature, material and internal tension. There are no formulas for this, therefore, the relaxation must be found from experimental work. From [9] an example is given in Figure 2.6. When the temperature is raised the parameters of the spring changes, especially the shear modulus.

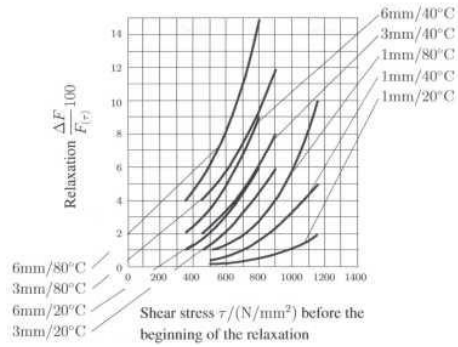
2.3.5 Types of load

DIN 2089, part 1 [8] separate the loads in three categories.

Static load: This is a load situation, where the load is constant over a longer period of time or where the number of load cycles is less than 10^4 in the lifetime of the spring.

Quasi-static load: In this load situation the load varies with time, but the variation is less than 10% of the design torsional stress with static load.

Dynamic load: This concerns load situations where the number of load cycles exceeds 10^4 over the springs lifetime and the stress variation exceeds 10% of the design stress level for static load.



[billedtekst.start]**Figure 2.6:** Relaxation after 48 hours of cold-formed compression springs made of wire type C (see [4] for details).[billedteskst.slut]

2.3.6 Dynamic loading

A spring is considered as dynamically loaded if the number of load cycles exceeds 10^4 during the lifetime of the spring. If the spring is loaded dynamically the stress level should be reduced to minimize the risk of fatigue failure.

The dynamic load is characterized by the following parameters:

The mean value of the load

$$F_{m,d} = \frac{F_{\max,d} + F_{\min,d}}{2} \tag{2.19}$$

and the amplitude of the load

$$F_{a,d} = \frac{F_{\max,d} - F_{\min,d}}{2} \tag{2.20}$$

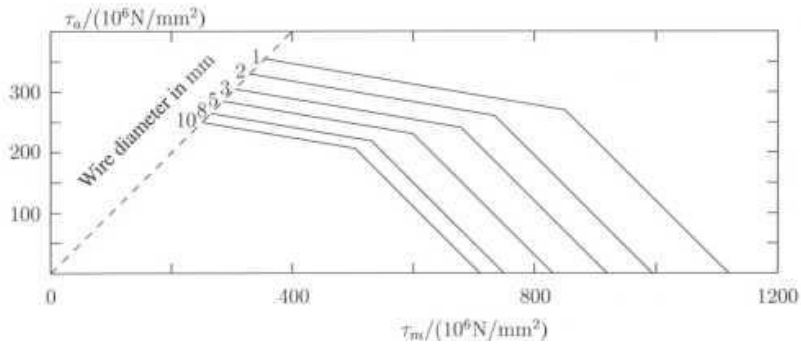
The shear stresses corresponding to these load parameters can be expressed as

$$\tau_m = \frac{8DF_{m,d}}{\pi d^3}c \tag{2.21}$$

and

$$\tau_a = \frac{8DF_{a,d}}{\pi d^3}c \tag{2.22}$$

Generally only compression springs should be used for dynamic loads, since the maximum actual stress in extension springs is very sensitive to the layout of the eyes. An example of material maximum values of shear stress is given by the specific modified Goodman diagram in Figure 2.7. For specific materials limits see e.g. [8]. More details about the Goodman diagram is given in Chapter 4.



[billedtekst.start]**Figure 2.7:** Specific modified Goodman diagram (for the case where $\tau_m > \tau_a$) for spring wire shear strength for 10^6 load cycles, the graph is modified relative to the original one found in [8] and is for wire type C and D (see [8] for details). The load point (τ_m, τ_a) should be below the full line for the specific wire diameter to withstand 10^6 load cycles.[billedteskst.slut]

2.3.7 Optimization

If a number of springs made in the same material are all able to fulfill the requirements set up by the designer, he will often choose the one with the smallest mass, since this normally will be the cheapest one.

The mass of a spring is proportional to the volume of the wire and can be expressed by

$$V_w = \frac{\pi d^2}{4}\pi D n = \frac{n\pi^2 d^2 D}{4} \tag{2.23}$$

the work done by the spring deformation can be expressed as

$$W = \frac{Fs}{2} \tag{2.24}$$

where

$$F = \frac{\pi d^3 \tau}{8D} \tag{2.25}$$

and

$$s = \frac{\pi D^2 n \tau}{Gd} \tag{2.26}$$

and

$$W = \frac{\tau^2}{4G} V_w \tag{2.27}$$

For a given amount of work the shear stress r must be maximized to minimize the volume. The most optimal spring is therefore the one with the highest shear stress level.

The required spring diameter for a given maximum load F_2 can be found

$$D = \frac{\tau_{all} d^3 \pi}{8 F_2} \tag{2.28}$$

where F_2 is the force determining the springs dimensions. All springs with a specified τ_{all} will be equally well suited with regard to work.

The following expressions can be used to guide the search for an optimum spring

$$s = \frac{\pi D^2 n \tau}{Gd} \tag{2.29}$$

$$s = f(D^2 n) \tag{2.30}$$

$$V_w = \frac{\pi^2 d^2 D n}{4} \tag{2.31}$$

$$V_w = f(Dn) \tag{2.32}$$

To achieve the smallest possible volume of wire, the mean diameter should be as large as possible for constant s.

The optimization process should first find the smallest acceptable wire diameter within the solution domain. As the maximum allowable torsional stress is known, the corresponding mean diameter of the coil can be found.

The mean diameter can be found from:

$$D = \frac{\pi \tau_{all}}{8F} d^3 \tag{2.33}$$

If the mean diameter D found is inside the limits of the solution domain, the calculation may continue with a check of the other boundary conditions.

If the mean diameter found is greater than the maximum allowable D_{max}, D should be put equal to D_{max}.

If the calculated mean diameter is smaller than the minimum acceptable mean diameter, the wire diameter should be rejected.

The spring index is expressed as

$$\omega = \frac{D}{d} \tag{2.34}$$

If the calculated spring index is inside the solution domain, the calculation may proceed with check of other boundary conditions.

If the calculated spring index is larger than accepted in the solution domain, then

$$D = d\omega \tag{2.35}$$

If the calculated spring index is smaller than the smallest acceptable in the solution domain, the wire is rejected and the calculation starts over again with a larger wire diameter.

The maximum deflection of one coil is found from

$$s_{1c} = \frac{8D^3}{Gd^4} F_2 \tag{2.36}$$

Now, the necessary number of coils can be found

$$n = \frac{s_2}{s_{1c}} \tag{2.37}$$

If the necessary number of coils is inside the specified limits, the unloaded spring length L_0 is found. In case of compression springs the maximum workload F_2 should not correspond to a solid compressed spring, since the spring rate is not linear at this deflection.

The maximum deflection until solid length is

$$s_c = s_2 + s_a \tag{2.38}$$

The stress in the spring at solid length is:

$$\tau_{ctheo} = \frac{G d s_c}{\pi n D^2} \tag{2.39}$$

This stress level should be lower than the maximum allowable stress level in the spring. If this is not the case the mean diameter of the spring is reduced and the number of coils is increased, which gives a smaller deflection in each coil and thereby a lower stress level.

Dynamically loaded springs must be checked in the Goodman diagram (see Chapter 4). This is done by calculating the stress at F_1 and the amplitude stress for $F_2 - F_1$.

2.3.8 Compression springs

Figure 2.3 shows the spring diagram for a compression spring. Two phenomena are of special interest for compression springs. One is the growing spring diameter when the spring is compressed and the other the natural frequency of the spring.

2.3.9 Growing mean diameter of helix

When a compression spring is loaded so that the coils approach each other, the diameter of the helix increases. The increase can be expressed as

$$\Delta D = 0.1 \frac{m^2 - 0.8md - 0.2d^2}{D} \tag{2.40}$$

where $m = (L_0 - d)/n$ for springs with closed and ground ends, and $m = (L_0 - 2.5d)/n$ for springs with plain ends.

2.3.10 Natural frequency

Springs used in machinery with high speeds can be exposed to resonance phenomena. If a spring is clamped in both ends the natural frequency is

$$f = \frac{d}{2\pi\sqrt{2}nD^2}\sqrt{\frac{G}{\rho}} \tag{2.41}$$

2.3.11 Buckling of spring

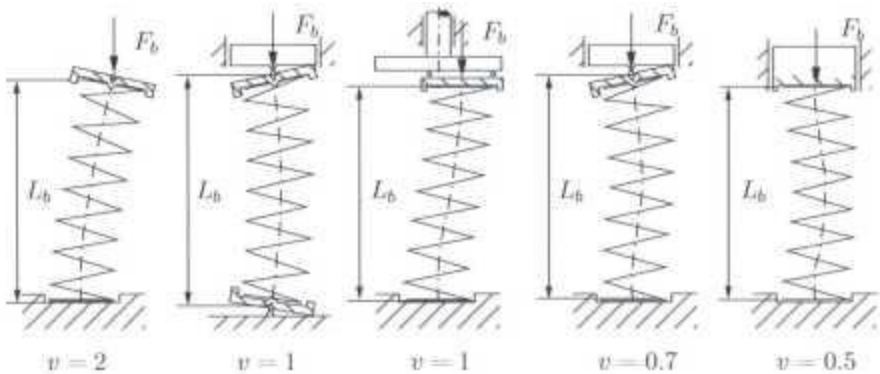
It is important to ensure that the spring height is limited for the buckling or column action to be avoided. If the spring buckles the length corresponding to the buckling load it is called L_b and the corresponding compression of the spring is called s_b. The buckling load for a spring depends upon the boundary conditions. In Figure 2.8 different types of boundary conditions are specified. Corresponding to these boundaries is a factor that influences the following expression

$$\frac{s_b}{L_0} = \frac{1}{2(1-\frac{G}{E})}\left(1-\sqrt{1-2\pi^2\frac{1-\frac{G}{E}}{1+2\frac{G}{E}}(\frac{D}{vL_0})^2}\right) \tag{2.42}$$

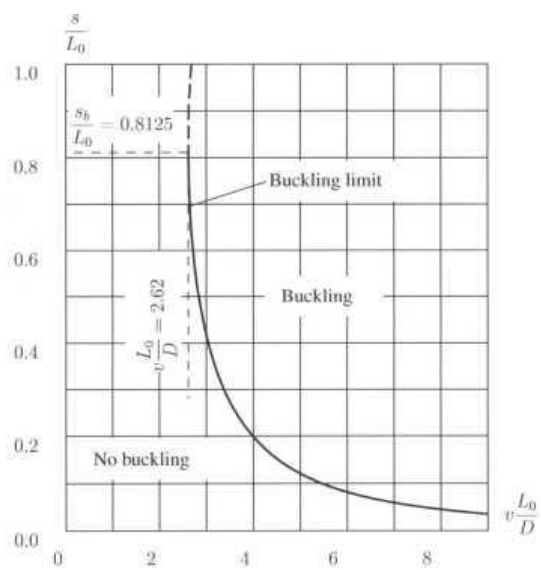
Buckling is avoided if

$$\frac{s_b}{s} > 1 \tag{2.43}$$

The buckling limit can also be seen on Figure 2.9.



[billedtekst.start]**Figure 2.8:** Different types of boundaries for springs .[billedteskst.slut]



[billedtekst.start]**Figure 2.9:** Theoretical buckling limit for helical compression spring (Poisson's ratio is assumed to be v =0.3).[billedteskst.slut]

2.3.12 Statically loaded cold-formed compression spring

The following conditions should be fulfilled

$$d < 7\text{mm},\ D < 200\text{mm},\ L_0 < 630\text{mm},\ n > 2 \text{ and } 4 < \omega < 20$$

All compression springs must be designed so they can be compressed until solid length, without this leading to overload.

The highest allowable torsional stress is $\tau_{all} < 0.56\tau_u$, where τ_u is the lowest ultimate stress level for the material. In addition the following formulas apply:

$$n_t = n + 2 \tag{2.44}$$

$$s_a = (0.0015\frac{D^2}{d} + 0.1d)n \tag{2.45}$$

with ends ground the solid length is

$$L_{sl} = n_t d \tag{2.46}$$

with plain ends

$$L_{sl} = (n_t + 1.5\ )d \tag{2.47}$$

2.3.13 Statically loaded hot-formed compression spring

The following conditions should be fulfilled:

$$8\text{mm} < d < 60\text{mm},\ D_e < 460\text{mm},\ L_0 < 800\text{mm},\ n > 3 \text{ and } 3 < \omega < 12$$

All compression springs must be designed so they can be compressed until solid length without this leading to overload. The maximum allowable stress when compressed to solid length can be seen from DIN 2089, part 1 [8]. In addition the following formulas apply

$$n_t = n + 1.5 \tag{2.48}$$

$$s_a = 0.02\ (D + d)n \tag{2.49}$$

with ends ground the solid length is

$$L_{sl} = (nt - 0.3)d \tag{2.50}$$

with plain ends

$$L_{sl} = (n_t + 1.1)d \tag{2.51}$$

2.3.14 Dynamically loaded cold-formed compression spring

The same geometrical limitations as for a statically loaded springs apply for the dynamically loaded spring. In addition there are some requirements concerning fatigue load.

The calculated τ_e must not exceed the value given in Goodman diagram. When calculating the lower and upper shear stress amplitude values, τ_{el} and τ_{eu}, the stress curvature correction factor c, see Figure 2.5, must be taken into account.

$$\tau_{el} = c\tau_1 \tag{2.52}$$

$$\tau_{eu} = c\tau_2 \tag{2.53}$$

$$\tau_a = \tau_{eu} - \tau_{el} = c(\tau_2 - \tau_1) \tag{2.54}$$

Goodman diagrams for a number of materials are given in DIN 2089, part 1 [81. It should be remembered that these values are only representative if the spring is not exposed to corrosion or frictional wear. All dynamically loaded springs should be shot peened.

In addition the following relations apply

$$n_t = n + 2 \tag{2.55}$$

$$L_{sl} = n_t d \tag{2.56}$$

$$s_a = (0.0015\frac{D^2}{d} + 0.1d)n \cdot 1.5 \tag{2.57}$$

2.3.15 Dynamically loaded hot-formed compression spring

The same limitations as for cold-formed statically loaded compression springs apply. In addition the following special relations apply

$$n_t = n + 1.5 \tag{2.58}$$

$$L_{sl} = (n_t - 0.3)d \tag{2.59}$$

$$s_a = 0.02(D + d)n \cdot 2 \tag{2.60}$$

As described in Subsection 2.3.2 the stresses should be modified with c.

2.3.16 Extension springs

For extension springs the specification is defined according to Figure 2.4.

2.3.17 Initial tension

Most cold-formed extension springs are produced with initial tension. The spring can, however, be calculated as if it were not. Afterwards the influence of the initial tension can be taken into account as described in DIN 2089, part 2 [9]. The stress level depends on the spring index ω. Since stresses are only given for ω = 4 and ω = 12 linear interpolation may be used to find intermediate values.

2.3.18 Statically loaded cold-formed extension springs

The springs must fulfill the following conditions:

$$d < 17\text{mm},\ D < 160\text{mm},\ L_0 < 1500\text{mm},\ n > 3 \text{ and } 4 < \omega < 20$$

At the largest possible spring deflection the highest allowable torsional stress is $\tau_{all} < 0.45\tau_u$, where τ_u is the minimal ultimate stress for the material.

The following special relations apply

$$n_t = n \tag{2.61}$$

$$L_{sl} = (n_t + 1)d \tag{2.62}$$

$$L_0 = L_{sl} + 2L_e \tag{2.63}$$

2.3.19 Statically loaded hot-formed extension springs

The following geometrical limitations apply

$$10\text{mm} < d < 60\text{mm},\ D_e < 460\text{mm},\ L_0 < 1500\text{mm},\ n > .3 \text{ and } 3 < \omega < 12$$

$$n_t = n \tag{2.64}$$

$$L_{sl} = (n_t + 1\ )d \tag{2.65}$$

$$L_0 = L_{sl} + 2L_e \tag{2.66}$$

2.3.20 Dynamically loaded cold-formed extension springs

The following geometrical limitations apply

$$d < 17\text{mm},\ D < 160\text{mm},\ L_0 < 1500\text{mm},\ n > 3 \text{ and } 4 < \omega < 20$$

$$n_t = n \tag{2.67}$$

$$L_{sl} = (n_t + 1)d \tag{2.68}$$

$$L_0 = L_{sl} + 2L_e \tag{2.69}$$

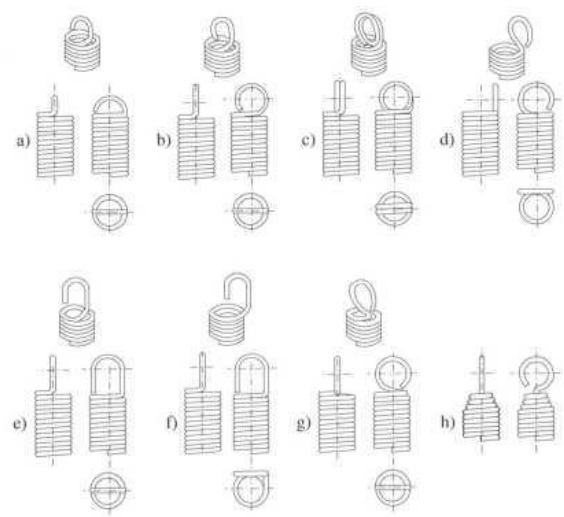
Extension springs are normally not suited for dynamic loading. This is caused by the normal production methods for the spring ends. By bending the hook, a sharp edge is "added" to the bending zone that causes stress concentration.

2.3.21 Dynamically loaded hot-formed extension springs

There are no specifications for this type of springs in the standards, since they should be avoided wherever possible.

2.3.22 Ends of extension springs

A number of different ends can be chosen when an extension spring is specified, see examples in Figure 2.10.



[billedtekst.start]**Figure 2.10:** Different types of ends.[billedteskst.slut]

The ends have the following names:

a Machine half loop

b Full twisted loop

c Double twisted loop

d Side loop

e Raised hook

f Side hook

g English loop

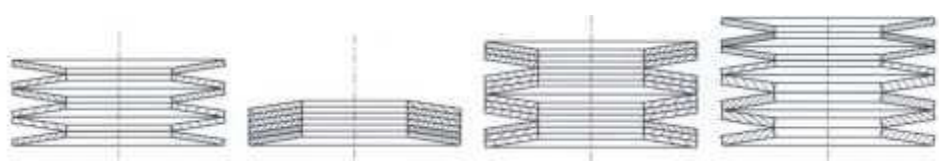
h Conical spring end with loop

The most common of these ends is the full twisted loop (b). Additionally, it is the cheapest one and at the same time relatively strong. The weakest point in an extension spring is in the end and generally it is recommended that the load should be at least 20% lower than in a compression spring.

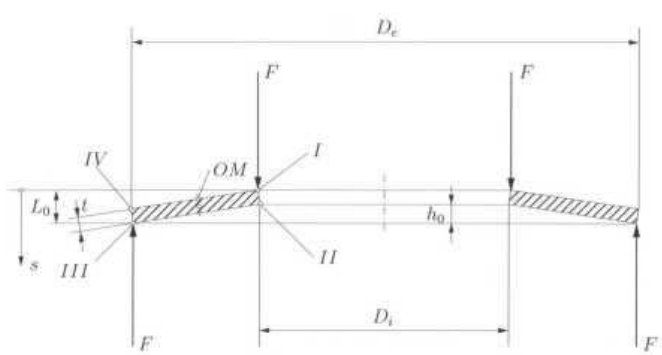
2.4 Belleville springs or coned-disk springs

Belleville springs are made from circular disks that have been squeezed to a conical shape.

The major advantage of Belleville springs is that it is possible to obtain very high spring rates and forces in limited space.



[billedtekst.start]**Figure 2.11:** Belleville springs.[billedteskst.slut]



[billedtekst.start]**Figure 2.12:** Geometry data and location of critical stresses in Belleville springs.[billedteskst.slut]

2.4.1 Formulas for Belleville springs

The fundamental formulas for Belleville springs originate from rather complicated analysis.

Belleville springs have according to the standard [10] two standard designs; one without contact areas as see in Figure 2.12, and one with contact areas.

The flattening force (the force required to flatten the spring completely) is

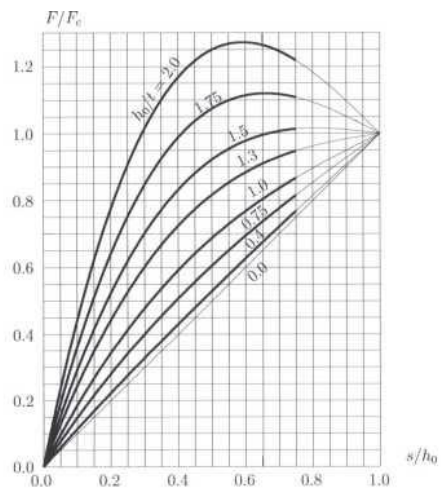
$$F_c = \frac{4E}{1-\nu^2}\frac{t^3 h_0}{C_1 D_e^2} C_4^2 \tag{2.70}$$

Introducing the material constant c_m

$$c_m = \frac{4E}{1-\nu^2} \tag{2.71}$$

Force-deflection relationship

$$F = c_m \frac{ts}{C_1 D_e^2} C_4^2 \left(C_4^2 (h_0 - s)(h_0 - 0.5s) + t^2\right) \tag{2.72}$$



[billedtekst.start]**Figure 2.13:** Deflection-force diagram for Belleville springs.[billedteskst.slut]

Spring rate

$$K = c_m \frac{t}{C_1 D_e^2} C_4^2 \left(C_4^2 (h_0^2 - 3h_0 s + 1.5s^2) + t^2\right) \tag{2.73}$$

Work done by spring deflection

$$W = c_m \frac{ts^2}{2C_1 D_e^2} C_4^2 \left(C_4^2 (h_0 - 0.5s)^2 + t^2\right) \tag{2.74}$$

The stress at the point defined in Figure 2.12 are

$$\sigma_{OM} = -c_m \frac{ts}{C_1 D_e^2} \frac{3}{\pi} C_4 \tag{2.75}$$

$$\sigma_I = -c_m \frac{s}{C_1 D_e^2} C_4 \left(C_3 t + C_4 C_2 (h_0 - 0.5s)\right) \tag{2.76}$$

$$\sigma_{II} = c_m \frac{s}{C_1 D_e^2} C_4 \left(C_3 t - C_4 C_2 (h_0 - 0.5s)\right) \tag{2.77}$$

$$\sigma_{III} = c_m \frac{s}{C_1 D_e^2 \delta} C_4 \left(C_3 t + C_4 C_5 (h_0 - 0.5s)\right) \tag{2.78}$$

$$\sigma_{IV} = -c_m \frac{s}{C_1 D_e^2 \delta} C_4 \left(C_3 t - C_4 C_5 (h_0 - 0.5s)\right) \tag{2.79}$$

where $\delta = D_e/D_i$ is the diameter ratio for the spring. Positive stress is tensile stress and negative stress is compressive stress. The defined parameters C_1 to C_5 depend on δ. For the standard spring design, in Figure 2.12, we can use $C_4 = 1$ and estimate the remaining parameters from Table 2.2. Alternatively one may use

$$C_1 = \frac{1}{\pi} \cdot \frac{\left(\frac{\delta - 1}{\delta}\right)^2}{\frac{\delta + 1}{\delta - 1} - \frac{2}{\ln \delta}} \tag{2.80}$$

$$C_2 = \frac{6}{\pi} \cdot \frac{\frac{\delta - 1}{\ln \delta} - 1}{\ln \delta} \tag{2.81}$$

$$C_3 = \frac{3}{\pi} \cdot \frac{\delta - 1}{\ln \delta} \tag{2.82}$$

$$C_4 = 1 \tag{2.83}$$

$$C_5 = 2C_3 - C_2 \tag{2.84}$$

For a spring with contact areas C_4 depends on the design see [10] for specific value.

2.5 Helical torsion springs

2.5.1 Methods of loading

Helical torsion springs have essentially the same shape as helical compression or tension springs except that the ends are formed in such a way that the spring may be loaded by a torque about the coil axis. Because of the mode of stressing of such springs, the primary stress is flexural in contrast to the helical compression or tension spring where the primary stress is torsional.

Torsion springs are made with a variety of shapes. The design of the spring end is made primarily for the purpose of transmitting external torque to the spring. Such springs are used in a wide variety of applications. A typical method of loading a torsion spring is indicated in Figure 2.14.

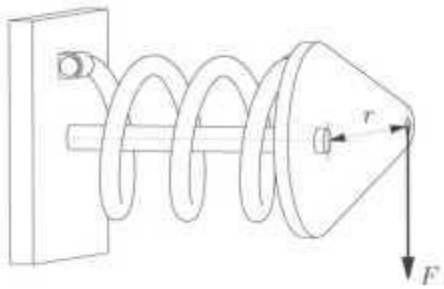
The spring is supposed to be wound around a rod, one end of it being fastened to the rod while the other end has a straight portion projecting radially. If the spring is loaded by a force F at a radius r from the axis to wind the spring the moment tending to twist the spring will be Fr, as indicated in the figure. Because of friction between the spring and guiding rod, the

actual moment will decrease along making it difficult to predict the spring rate.

Since most torsion springs are formed cold, it is advisable to load them in such a way that the spring tends to wind up as the load is applied. The reason for this is that the residual stresses set up as a consequence of the cold winding are in such a direction as to subtract from the peak stress due

Table 2.2: Parameters for Belleville springs.

δ	C_1	C_2	C_3	C_5
1.2	0.29	1.02	1.05	1.08
1.4	0.46	1.07	1.14	1.20
1.6	0.57	1.12	1.22	1.31
1.8	0.65	1.17	1.30	1.43
2.0	0.69	1.22	1.38	1.54
2.2	0.73	1.26	1.45	1.64
2.4	0.75	1.31	1.53	1.75
2.6	0.77	1.35	1.60	1.85
2.8	0.78	1.39	1.67	1.95
3.0	0.79	1.43	1.74	2.05
3.2	0.79	1.46	1.81	2.15
3.4	0.80	1.50	1.87	2.25
3.6	0.80	1.54	1.94	2.34
3.8	0.80	1.57	2.00	2.44
4.0	0.80	1.60	2.07	2.53
4.2	0.80	1.64	2.13	2.62
4.4	0.80	1.67	2.19	2.71
4.6	0.80	1.70	2.25	2.80
4.8	0.79	1.73	2.31	2.89
5.0	0.79	1.76	2.37	2.98



[billedtekst.start]**Figure 2.14:** Helical torsion spring.[billedteskst.slut]

to the loading, provided that the load is in the same direction as that in which the spring was wound. If the direction of loading is such as to unwind the spring, it is advisable to heat-treat by means of a stress-relieving treatment in order to remove residual stresses.

2.5.2 Binding effects

Because a torsion spring (for usual applications) tends to wind up with load, its diameter decreases. If the spring is operated around a rod, it is important that sufficient clearance is being allowed between the rod diameter and the inner diameter of the spring. If this is not done, the spring may bind or wrap around the rod and high stresses may be set up. The clearance necessary may be estimated from the calculated deflection of the ends of the spring that will depend on the end design as seen in Subsection 2.3.22. Thus, if the spring end deflects 90° or one-quarter of a turn and the spring has eight turns, the diameter will

change about 3%. This can be allowed for in design. If the spring fits inside a tube and loaded to unwind sufficient clearance must be allowed between the outside diameter of the spring and the inside diameter of the tube.

2.5.3 Formulas for helical torsion springs

The fundamental formulas derived from basic physics are presented below.

Bending moment load and spring torsional moment

$$M = Fr \tag{2.85}$$

Bending stress in wire

$$\sigma = \frac{M}{W_b} c_t = \frac{32Fr}{\pi d^3} c_t \tag{2.86}$$

where the stress curvature correction factor c_t depends on the spring index (2.34) and is given as

$$c_t = \frac{4w^2 - w - 1}{4w(w-1)} \tag{2.87}$$

Spring deflection angle

$$\alpha = \frac{Ml}{EI} = \frac{64FRl}{E\pi d^4} \tag{2.88}$$

Torsional spring rate

$$K_t = \frac{M}{\alpha} = \frac{E\pi d^4}{64l} \tag{2.89}$$

Length of wire in the spring

$$l = D\pi n \tag{2.90}$$

The deflection work for a helical torsion spring is given by

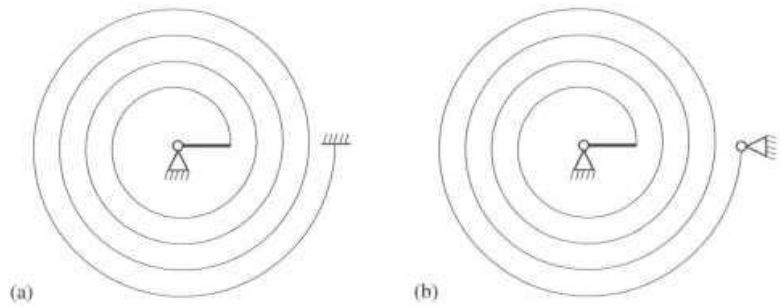
$$W = \int_{\alpha_1}^{\alpha_2} M d\alpha \tag{2.91}$$

using from (2.88) that $M = \alpha EI/l$ in (2.91) gives

$$W = \left[\frac{EI}{l}\frac{\alpha^2}{2}\right]_{\alpha_1}^{\alpha_2} = \frac{EI}{2l}(\alpha_2^2 - \alpha_1^2) \tag{2.92}$$

and finally the work is given by

$$W = \frac{EI}{2l}\frac{l^2}{E^2I^2}(M_2^2 - M_1^2) = \frac{l}{2EI}(M_2^2 - M_1^2) \tag{2.93}$$



[billedtekst.start]**Figure 2.15:** Spiral springs. (a) Clamped outer end. (b) Simply supported outer end.[billedteskst.slut]

2.6 Spiral springs

Spiral springs (Figure 2.15) are flat moment springs that are used in many different applications, e.g. watches or as the rewinding mechanism for the starter cord of small engines (lawnmower engines). The cross section of the spring wire is in most cases rectangular.

In this section we will derive the spring stiffness for spiral springs. The derivation is based on a number of assumptions, these are

- The number of turns is large.
- The thickness h of the wire relative to the radius r allows the theory for straight beams to be used, i.e. $h/r < 1/2$.
- There is no contact between the turns.
- Deformation only due to bending moment.

With the assumption that there is no contact between the turns, we may express a spring in the shape of a spiral as an Archimedes spiral given as

$$x = \alpha\theta\cos(\theta) \tag{2.94}$$

$$y = \alpha\theta \sin(\theta) \tag{2.95}$$

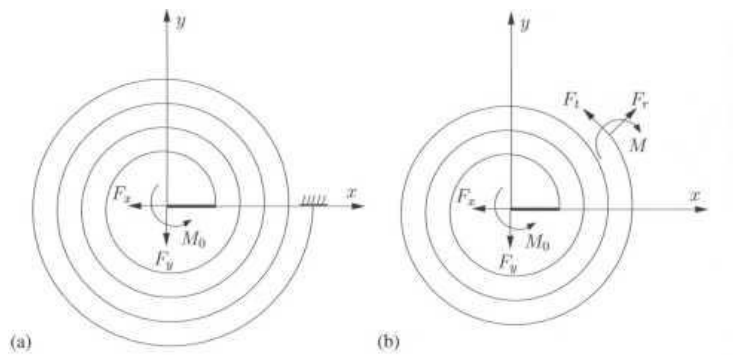
where a is the distance between the turns and the angle lies between two endpoints $\theta_1 \leqq \theta \geqq \theta_2$.

The derivation separates into two cases; one where the outer end is clamped and the second where the outer end is simply supported.

2.6.1 Clamped outer end

The inner part of the spiral spring is assumed to be attached to the center by an infinitely stiff beam. With the outer end clamped the structure is two times statically indeterminate and we must add two redundant generalized forces. The chosen redundant forces F_x and F_y are illustrated in Figure 2.16(a) together with the external applied moment M_0.

The bending moment is given by



[billedtekst.start]**Figure 2.16:** (a) Clamped spiral spring with external applied moment M_0 and redundant forces F_x and F_y. (b) Spiral spring cut at a specific point indicating the internal forces F_r, F_t and the bending moment M.[billedteskst.slut]

$$M = M_0 + F_y\, x - F_x\, y \tag{2.96}$$

With the assumption that the height h of the wire relative to the radius r in the spiral spring fulfills that $h/r < 1/2$ we can express the complementary elastic energy due to the bending moment in the spring as

$$U^c = \int_0^\ell \frac{M^2}{2EI} ds \tag{2.97}$$

where ℓ is the length of the spring, E modulus of elasticity and I moment of inertia of cross section.

From Castigliano's 2nd theorem we know that we may find the displacement by differentiating the energy with respect to the corresponding load. The two displacements due to the two redundant forces must be zero and therefore we find that

$$\frac{\partial U^c}{\partial F_x} = \int_0^\ell \frac{M}{EI}\frac{\partial M}{\partial F_x} ds = 0 \tag{2.98}$$

$$\frac{\partial U^c}{\partial F_y} = \int_0^\ell \frac{M}{EI}\frac{\partial M}{\partial F_y} ds = 0 \tag{2.99}$$

By using (2.96) in (2.98) and (2.99) we find that

$$-M_0 \int_0^\ell y ds - F_y \int_0^\ell xy ds + F_x \int_0^\ell y^2 ds = 0 \tag{2.100}$$

$$M_0 \int_0^\ell x ds + F_y \int_0^\ell x^2 ds - F_x \int_0^\ell xy ds = 0 \tag{2.101}$$

Under the assumption that there are many turns we may show that the following equations are fulfilled

$$\frac{\int_0^\ell y ds}{\int_0^\ell y^2 ds} \approx 0 \qquad \frac{\int_0^\ell xy ds}{\int_0^\ell y^2 ds} \approx 0 \qquad \frac{\int_0^\ell x ds}{\int_0^\ell x^2 ds} \approx 0 \qquad \frac{\int_0^\ell xy ds}{\int_0^\ell x^2 ds} \approx 0 \tag{2.102}$$

From these equations follow directly that

$$F_x \approx 0 \qquad\qquad F_y \approx 0 \tag{2.103}$$

We now find the angle of rotation θ_0 corresponding to the external load M_0 by applying Castigliano's 2nd theorem again. We find

$$\theta_0 = \frac{\partial U^c}{\partial M_0} \approx \frac{M_0}{EI}\int_0^\ell ds = \frac{M_0 \ell}{EI} \tag{2.104}$$

From this equation we directly have the stiffness of a spiral spring that is clamped at the outer end

$$K_c \approx \frac{EI}{\ell} \tag{2.105}$$

The maximum stress in the spring is found directly from the moment and is given by

$$\sigma_{\max} \approx \frac{M_0 h}{2I} \tag{2.106}$$

2.6.2 Simply supported outer end

With the assumption of many turns we can also specify that $\theta_2 = 2\pi n$ where n is an integer number. The inner part of the spiral spring is assumed to be attached to the center by a infinitely stiff beam. With the outer end simply supported, the structure is one time statically indeterminate and we must add one redundant generalized force. In Figure 2.17(a) we show the chosen redundant forces F_x together with the external applied moment M_0 and the reaction force R_y.

Using that the moment is zero at the outer end we may express R_y as a function of M_0 as shown on Figure 2.17(b). The bending moment is then given by

$$M = M_0(1 - \frac{x}{r_1}) - F_x y \tag{2.107}$$

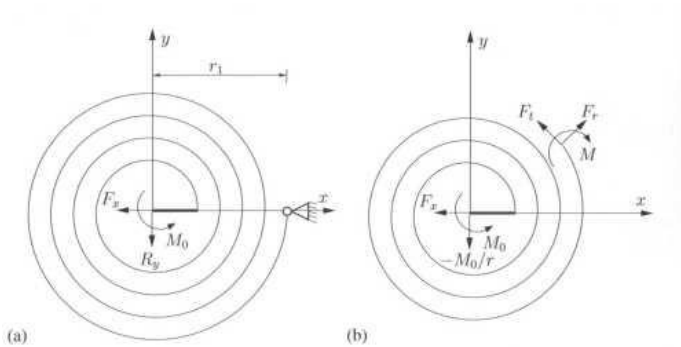
We apply Castigliano's 2nd theorem to find F_x under the assumption that the corresponding displacement is zero.

$$\frac{\partial U^c}{\partial F_x} = \int_0^\ell \frac{M}{EI}\frac{\partial M}{\partial F_x} ds = 0 \tag{2.108}$$

By using (2.107) in (2.108) we find that

$$-M_0 \int_0^\ell y ds + \frac{M_0}{r_1}\int_0^\ell xy ds + F_x \int_0^\ell y^2 ds = 0 \tag{2.109}$$

Using (2.102) we find that



[billedtekst.start]**Figure 2.17:** (a) Simply supported spiral spring with redundant force F_x,the external applied moment M_0 and the reaction force R_y. (b) Spiral spring cut at a specific point indicating the internal forces which includes the bending moment *M.[billedteskst.slut]*

$$F_x \approx 0 \tag{2.110}$$

We now find the angle of rotation θ_0 corresponding to the external load M_0 by applying Castigliano's 2nd theorem again. We find

$$\theta_0 = \frac{\partial U^c}{\partial M_0} = \frac{M_0}{EI}\int_0^\ell (1-\frac{x}{r_1})^2 ds \approx \frac{M_0}{EI}\int_0^\ell 1+(\frac{x}{r_1})^2 ds \tag{2.111}$$

The last approximation in (2.111) is due to the use of (2.102). Finally, we may for a large number of turns show that

$$\frac{1}{r_1^2}\int_0^\ell x^2 ds \approx \frac{\ell}{4} \tag{2.112}$$

The rotation θ_0 is therefore given by

$$\theta_0 \approx \frac{5}{4}\frac{M_0\ell}{EI} \tag{2.113}$$

From this equation we directly find the stiffness of the spiral spring that is simply supported at the outer end

$$K_s \approx \frac{4}{5}\frac{EI}{\ell} \tag{2.114}$$

The maximum stress in the spring is found directly from the maximum value of the moment, which in the case of many turns becomes $M_{max} \approx 2M_0$.

$$\sigma_{max} \approx \frac{M_0 h}{I} \tag{2.115}$$

We notice that the maximum stress for the case of a simply supported outer end is twice the size of the maximum stress for the case of a clamped outer end.

2.7 Supplementary literature

Springs are treated in most machine element textbooks and the following books are recommended for further reading [14], [9] and [1],

2.8 Nomenclature

c	–	Stress curvature correction factor
c_m	N/mm²	Material constant
c_t	–	Stress curvature correction factor
d	mm	Wire diameter
f	rad/s	Natural frequency
h_0	mm	Strike height of Belleville spring
t	mm	Thickness of Belleville spring
l	mm	Length of wire
l	mm	Length of spring
n	–	Number of active coils
n_t	–	Total number of coils
s	mm	Deflection of spring
s	mm	Curve coordinate
s_1, s_2, s_3	mm	Deflection of spring corresponding to loads F_1, F_2 and F_3
s_{1c}	mm	Deflection of one coil
s_a	mm	Sum of minimum distances between coils

s_b	mm	Deflection of spring corresponding to buckling load
s_h	mm	Working deflection
s_n	mm	Deflection corresponding to the smallest allowable spring length
s_c	mm	Deflection to solid length
t	mm	Thickness of Belleville spring
v	–	Buckling sensitivity factor
w	–	Spring index
x	mm	Position
y	mm	Position
C_1, C_2, C_3, C_4, C_5	–	Constants for Belleville springs
D	mm	Mean diameter of helix
D_e	mm	External diameter of helix
D_e	mm	External diameter of Belleville spring
D_i	mm	Internal diameter of helix
D_i	mm	Internal diameter of Belleville spring
D_{max}	mm	Maximum value of mean diameter of helix
E	N/mm^2	Modulus of elasticity
F	N	Load on spring

F_1	N	Minimum working load
F_2	N	Maximum working load
$F_{a,d}$	N	Load amplitude for dynamic loaded spring
F_b	N	Buckling load
F_c	N	Flattening force for Belleville spring
F_{max}	N	Maximum allowable load
$F_{max,d}$	N	Maximum load for dynamic loaded spring
$F_{min,d}$	N	Minimum load for dynamic loaded spring
$F_{m.d}$	N	Mean load for dynamic loaded spring
F_n	N	Load corresponding to the smallest allowable spring length
G	N/mm^2	Shear modulus
I	mm^4	Cross sectional moment of inertia
I_p	mm^4	Cross sectional polar moment of inertia
K	N/mm	Spring rate/stiffness
K_c	Nmm	Stiffness of clamped spiral spring
K_s	Nmm	Stiffness of simply supported spiral spring
K_t	Nmm	Torsional spring rate/stiffness
L_0	mm	Length of spring unloaded
L_1	mm	Length of spring when loaded with F_1
L_2	mm	Length of spring when loaded with $F2$
L_b	mm	Buckling length
L_e	mm	Length of end on an extension spring

L_n	mm	Test length of spring
L_{sl}	mm	Solid length of spring (when coils touch each other)
M	Nmm	Bending moment
s_d	mm	Standard deviation of wire diameter
S_D	mm	Standard deviation of spring diameter
SF_{max}	N	Standard deviation of spring strength
$S_{\tau all}$	N/mm²	Standard deviation of spring shear strength
T	Nmm	Torsional moment
u^c	Nmm	Complementary elastic energy
V_w	mm³	Volume of the wire in a spring
W	Nm	Work done by spring deformation
W_b	mm³	Bending resistance
w_p	mm³	Torsional resistance
α	rad	Angle of deflection
γ	–	Shear strain (angle for torsionally loaded wire)
δ	–	Diameter ratio for Belleville springs
ϱ	kg/m³	Density of spring material
σ	N/mm²	Normal stress
τ_u	N/mm²	Ultimate stress
τ_a	N/mm²	Amplitude of shear stress
τ_{all}	N/mm²	Allowable shear stress
τ_e	N/mm²	Endurance limit in shear
τ_{el}	N/mm²	Endurance limit in shear, lower value in Goodman-diagram

τ_{eu}	N/mm²	Endurance limit in shear, upper value in Goodman-diagram
τ_{m}	N/mm²	Mean shear stress
τ_{ctheo}	N/mm²	Shearing-stress is a spring compressed solid
V	–	Poisson ratio

2.9 References

[1] A. D. S. Carter. *Mechanical Reliability.* MacMillan, 1986.

[2] K. Decker. Maschinenelemente, Funktion, Gestaltung und Berechnung, 18. aktualisierta Auflage. Carl Hanser Verlag, München, Wien, 2011.

[3] DIN 17221. Hot rolled steels for quenched and tempered springs, technical delivery conditions.

[4] DIN 17223 part 1. Round spring steel wire, patented cold drawn unalloyed steel spring wire, technical delivery conditions.

[5] DIN 17223 part 2. Round spring steel wire, oil quenched and tempered spring steel wire of unalloyed and alloyed steels, technical delivery conditions.

[6] DIN 17224. Wire and strip of stainless steels for springs, technical terms of delivery.

[7] DIN 17682. Round spring wire of wrought copper alloys, mechanical properties technical conditions of delivery.

[8] DIN 2089 part 1. Helical compression springs out of round wire and rod, calculation and design.

[9] DIN 2089 part 2. Helical springs made from round wire and rod, calculation and design of extension springs.

[10] DIN 2092. Tellerfedern; Berechnung (German standard), 1992.

[11] DIN 2096 part 2. Helical compression springs made of round rod, quality requirements for mass production.

[12] G. Niemann, Winter H,, and Höhn G. *Maschinenelemente Band I.* Springer-Verlag, Berlin, Heildelberg, New York, 2005.

[13] M. F. Spotts. *Design of Machine Elements.* Prentice-Hall, 1978.

[14] W. Steinhilper and R. Röper. *Maschinen-und Konstruktionselemente.* Springer-Verlag, 1991.

Chapter 3 Rolling element bearings

3.1 Introduction

Rolling element bearings are very important standard elements in modern machinery. They are manufactured in an incomprehensible large number and are therefore relatively cheap and of a very high and uniform quality.

The fundamental bearing theory used to analyze rolling element bearings is called "Elastohydrodynamic Lubrication" and this may be found in advanced textbooks on lubrication theory. See for example [4].

The typical situation for the engineering designer is that he has to select or find an appropriate bearing for his machinery. This chapter will focus on the selection procedure and the associated required information.

Rolling element bearing geometry is ISO-standardized, whereas the bearing selection procedure is different from manufacturer to manufacturer.

The content of this chapter is in broad outline based on the selection procedure from the manufacturer SKF. In the "SKF General Catalogue" (the catalogue with the bearings manufactured by the SKF Company) the selection procedure and other facts are described in detail. Further it is possible to obtain much more information from the SKF Company WEB-site.

3.2 Bearing types

Each type of bearing displays characteristic properties that depend on its design. For example, deep groove ball bearings can accommodate radial loads as well as axial loads.

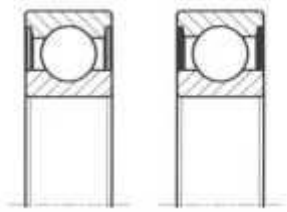
Spherical roller bearings can carry heavier loads and are self-aligning. These properties make them attractive in heavy machinery.

Several application factors have to be considered and weighed against each other when selecting bearing type. These include load carrying capacity and life, friction, permissible speeds, bearing internal clearance or preload, lubrication and sealing.

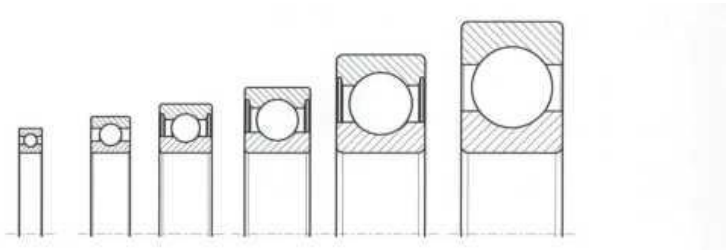
3.2.1 Available space

For small diameter shafts all types of ball bearings can be used, the most popular being deep groove ball bearings, see Figure 3.1. Needle roller bearings are also suitable. For large diameter shafts, cylindrical, spherical and taper roller bearings are available as well as deep groove ball bearings, see Figures 3.1 to 3.9.

When radial space is limited bearings with a small cross section, particularly those with a low cross- sectional height should be chosen. Needle roller bearings without (or with) inner ring are very appropriate



[billedtekst.start]**Figure 3.1:** A deep groove ball bearing (6210) with and without seals.[billedteskst.slut]



[billedtekst.start]**Figure 3.2:** Size comparison of deep groove ball bearings with the same bore diameter. From left to right: 61810, 61910, 6010, 6210, 6310 and 6410.[billedteskst.slut]

as are certain series of deep groove and angular contact ball bearings, cylindrical and spherical roller bearings.

When space is limited in the axial direction, certain series of single row cylindrical roller bearings and deep groove ball bearings can be used for radial and combined loads as well as the various types of combined needle roller bearings.

For purely axial loads, needle roller and cage thrust assemblies (with or without washers) as well as certain series of thrust ball bearings and cylindrical roller thrust bearings can be used. In many cases, one of the principal dimensions of the bearing, generally the bore diameter, is predetermined by the machine design.

3.2.2 Loads

The magnitude of the load is the factor that usually determines the size of bearing used.

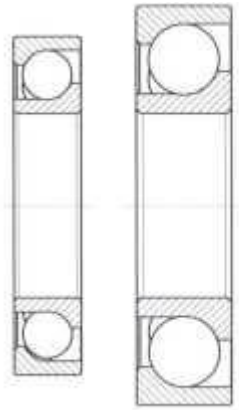
Generally, roller bearings are able to support heavier loads than ball bearings having the same overall dimensions and bearings having a full complement of rolling elements can carry heavier loads than the corresponding caged bearings. Ball bearings are mostly used where loads are light or moderate. For heavy loads and where shaft diameters are large, roller bearings are usually more appropriate.

Purely radial loads can be supported by cylindrical roller bearings having one ring without flanges (NU and N types), radial needle roller bearings and CARB bearings. All other radial bearings can carry some axial load in addition to radial loads.

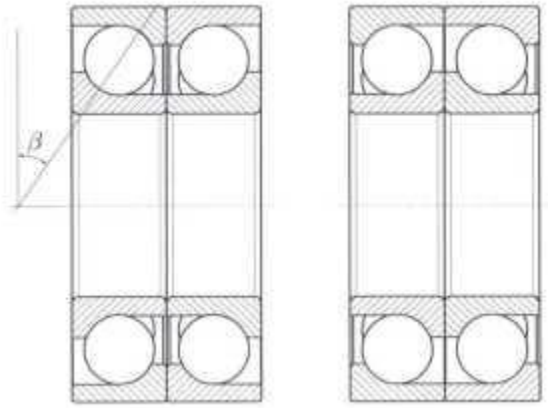
Thrust bearings can be grouped as single direction bearings and double direction

bearings respectively. Single direction thrust ball bearings can only accommodate loads acting in one direction. For loads acting in both directions, double direction bearings are needed. Thrust ball bearings and four-point contact ball bearings are the most suitable types for light or moderate loads that are purely axial.

Angular contact thrust ball bearings can support moderate axial loads at high speeds, see Figure 3.3. The single direction bearings can also accommodate simultaneously acting radial loads, whilst the double direction bearings are normally used only for purely axial loads.



[billedtekst.start]**Figure 3.3:** Two angular contact ball bearings (7212 and 7312).[billedteskst.slut]



[billedtekst.start]**Figure 3.4:** Pairs of angular contact ball bearings. To the left: back-to-back arrangement. To the right: face-to-face arrangement.[billedteskst.slut]

For moderate and heavy axial loads acting in one direction, needle roller bearings, single direction cylindrical and taper roller thrust bearings are suitable, as are spherical roller thrust bearings. Spherical roller thrust bearings can also accommodate simultaneously acting radial loads.

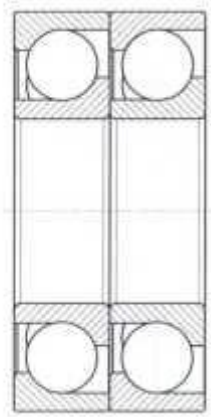
For heavy alternating axial loads, two cylindrical roller thrust bearings or two spherical roller thrust bearings can be mounted adjacent to each other.

3.2.3 Combined load

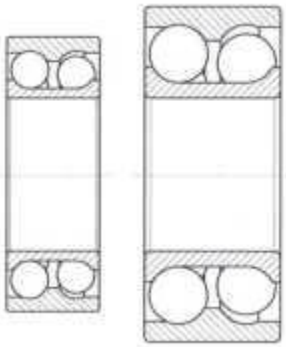
A combined load comprises a radial and an axial load acting simultaneously.

The ability of a bearing to carry axial load is determined by the angle of contact, β, the

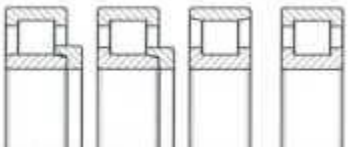
greater the angle, the more suitable the bearing for axial loads.



[billedtekst.start]**Figure 3.5:** Tandem pairs of angular contact ball bearings.[billedteskst.slut]



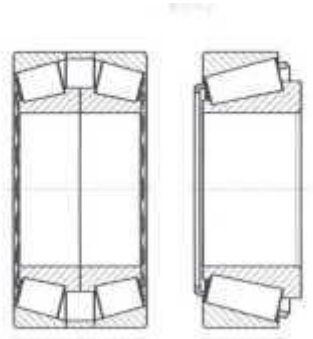
[billedtekst.start]**Figure** 3.6: Double row angular contact ball bearings (3210 and 3310).[billedteskst.slut]



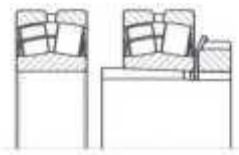
[billedtekst.start]**Figure 3.7:** Cylindrical roller bearings with different layout of rings. From left to right: NU210, NJ210, N210 and NUP210.[billedteskst.slut]

The axial load carrying capacity of deep groove ball bearings depends on the internal clearance in the bearing.

For combined loads, single and double row angular contact ball bearings and single row taper roller bearings are most commonly used. Although deep groove ball bearings and spherical roller bearings are also suitable.

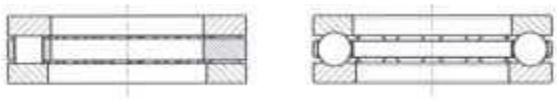


[billedtekst.start]**Figure 3.8:** A double row tapered roller bearing (33210) and a single row tapered roller bearing (30210).[billedteskst.slut]

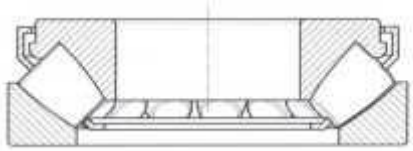


[billedtekst.start]**Figure 3.9:** Double row spherical roller bearings. To the left a bearing with standard bore, to the right a bearing with tapered bore, a tapered seating and a lock nut.[billedteskst.slut]

Single row angular contact ball bearings, taper roller bearings and spherical roller thrust bearings can only accommodate axial loads acting in one direction. For axial loads of alternating direction these bearings must be combined with a second bearing.



[billedtekst.start]**Figure 3.10:** A Thrust roller bearing and a thrust ball bearing.[billedteskst.slut]



[billedtekst.start]**Figure 3.11:** Spherical roller thrust bearing.[billedteskst.slut]

When the axial component of combined loads is large, it may be supported independently from the radial load by a separate bearing. In addition to the thrust bearings some radial bearings, e.g. deep groove ball bearings or four-point contact ball bearings are suitable for this task. To make sure that the bearing is only subjected to the axial load in such cases, the bearing must be mounted with radial clearance for the outer ring.

When the load acts eccentrically on the bearing, tilting moments will arise. Double row bearings, e.g. deep groove or angular contact ball bearings can take up tilting moments. However paired single row angular contact ball bearings or taper roller bearings arranged face-to-face or better still back-to-back, are more suitable.

3.2.4 Misalignment

Angular misalignments between shaft and housing occur for example, when the shaft bends under the operating load, when the bearing seats in the housing are not machined at a single setting or when shafts are supported by bearings in separate housings that are far apart.

Generally, the "rigid bearings" cannot accommodate any misalignment or can only tolerate very minor misalignments.

Self-aligning bearings can in opposition to rigid bearings (i.e. self-aligning ball bearings, CARB bearings, spherical roller bearings and spherical roller thrust bearings) on the other hand accommodate misalignments produced under operating loads and can also compensate for errors of alignment.

3.2.5 Speed

The speed at which rolling element bearings can be operated is limited by the permissible operating temperature. Therefore bearing types with low friction and correspondingly low heat generation in the bearing itself are the most suitable for high-speed operation.

The highest speeds can be achieved with deep groove ball bearings when loads are purely radial and with angular contact ball bearings for combined loads. This is particularly true for the high precision bearings with special cages.

Because of their design, thrust bearings cannot operate at such high speeds as radial bearings.

3.2.6 Stiffness

The stiffness of a rolling element bearing is characterized by the magnitude of the elastic deformation (resilience) in the bearing under load.

Because of the contact conditions between rolling elements and raceways roller bearings (e.g. cylindrical or taper roller bearings) have higher stiffness than ball bearings.

Bearing stiffness can be further enhanced by applying a preload. See [4].

3.2.7 Axial displacement

A shaft is generally supported in a locating and a non-locating bearing, see Figure 3.12.

Locating bearings provide axial location for the machine component in both directions. The most suitable bearings for this task are those that can accommodate combined loads.

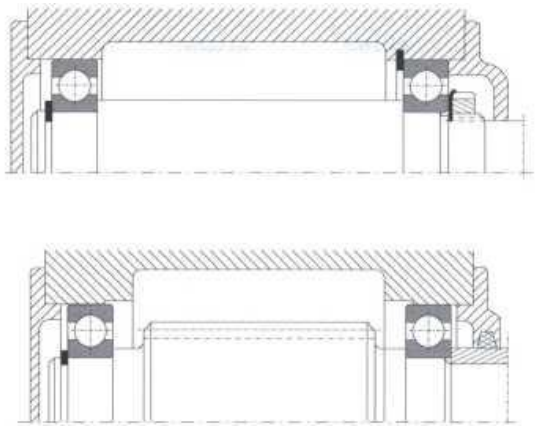
Non-locating bearings must permit movement in the axial direction for the bearings not to be additionally stressed when for example, thermal expansion of the shaft takes place. The most suitable bearings include needle roller bearings and cylindrical roller bearings which have one ring without flanges of the NU and N designs. Cylindrical roller bearings of the NJ design

and some full complement designs can also be used.

If the required axial displacement is relatively large and at the same time the shaft may be misaligned, the CARB type is the ideal non-locating bearing.

All these bearings permit axial displacement of the rollers with respect to one of the raceways so both the inner and the outer rings can be mounted with interference fits. Values for the permissible axial displacement within the bearing are given in the relevant product tables.

If deep groove ball bearings or spherical roller bearings are used as non-locating bearings, one of the bearing rings must have a loose fit, see Figure 3.12. Whether this loose fit is on the inner or outer ring depends on the load situation.



[billedtekst.start]**Figure 3.12:** Two examples of bearing arrangements for a shaft with a locating and a non-locating bearing.[billedteskst.slut]

3.3 Load carrying capacity and life

The size of a bearing to be used for an application is initially selected on the basis of its load carrying capacity in relation to the loads to be carried and the requirements regarding life and reliability. Numerical values termed basic load ratings are used in the calculations to express load carrying capacity. Values for the basic dynamic load rating C and the basic static load rating C_0 are quoted in the bearing tables.

3.3.1 Basic load ratings

The basic dynamic load rating C is used for calculations involving dynamically loaded bearings, i.e. when selecting a bearing which is to rotate under load. It expresses the bearing load which will give an "ISO basic rating life" (defined below) of 10^6 revolutions.

The basic dynamic load ratings for bearings have been determined in accordance with the methods prescribed by ISO 281:1990. They apply to loads that are constant in both magnitude and direction, for radial bearings radial loads and for thrust bearings axial loads that act centrically.

The basic static load rating C_0 is used in calculations when the bearings are lo rotate at very slow speeds, are to be subjected to very slow oscillating movements or are to be stationary under load during certain periods. It must also be taken into account when heavy shock loads of short duration act on a rotating (dynamically stressed) bearing.

The basic static load rating is defined in accordance with ISO 76:1987 as the static load that corresponds to a calculated contact stress at the center of the most heavily loaded rolling element/raceway contact of

- 4600MPa for self aligning ball bearings
- 4200MPa for all other ball bearings

- 4000MPa for all roller bearings

This stress produces a total permanent deformation of rolling element and raceway which is approximately $0.0001D_w$, where D_w is the rolling element diameter, for a ball of 10mm in diameter this

corresponds to a plastic deformation of the two bodies in contact of 1μm. The loads are purely radial for radial bearings and centrically acting axial loads for thrust bearings.

3.3.2 Life

The life of a rolling bearing is defined as the number of revolutions (or the number of operating hours at a given constant speed) which the bearing is capable of taking, before the first sign of fatigue (flaking, spalling) occurs on one of its rings or rolling elements.

Laboratory tests as well as practical experience indicate that seemingly identical bearings operating under identical conditions have different lives. A clear definition of the term "life" is therefore essential for the calculation of bearing size.

All information presented on dynamic load ratings is based on the life that 90% of a sufficiently large group of apparently identical bearings can be expected to attain or exceed. This is called the basic rating life and agrees with the ISO definition. The median life is approximately five times the calculated basic rating life.

There are several other bearing "lives". One of these is the "service life", which is the actual life achieved by a specific bearing before it fails. Failure is not generally by fatigue in the first instance, but by wear, corrosion, seal failure etc.

Bearing life can be calculated with various degrees of sophistication, depending on the accuracy with which the operating conditions can be defined.

3.3.3 Basic rating life equation

The most simple method of life calculation is to use the ISO [5] equation for basic rating life which is

$$L_{10} = \left(\frac{C}{P}\right)^p \tag{3.1}$$

or

$$\frac{C}{P} = (L_{10})^{1/p} \tag{3.2}$$

Where

L_{10} basic rating life (at 90% reliability), millions of revolutions.]

C basic dynamic load rating, kN.

P equivalent dynamic bearing load, kN.

p exponent of the life equation, (p = 3 for ball bearings, p = 10/3 for roller bearings).

For bearings operating at constant speed it may be more convenient to deal with a basic rating life expressed in operating hours using the equation

$$L_{10h} = \frac{10^6}{60n}\left(\frac{C}{P}\right)^p \tag{3.3}$$

or

$$L_{10h} = \frac{10^6}{60n}L_{10} \tag{3.4}$$

where

L_{10h} basic rating life (at 90% reliability), operating hours.

n[rpm] rotational speed

Basic rating life for oscillating bearings

If a bearing does not rotate, but oscillates from a central position through an amplitude angle of 7, then:

$$L_{10osc} = \frac{\pi}{2\gamma} L_{10} \tag{3.5}$$

where

L_{10osc} basic rating life, millions of cycles

γ oscillation amplitude (angle of maximum deviation from center position)

It is not meaningful to calculate a basic rating life if the amplitude of oscillation 7 is very small.

3.3.4 Requisite basic rating life

When determining the bearing size it is general practice to base the calculations on the basic rating life (L_{10}). Therefore it is essential that the required basic rating life for the application under consideration is specified. It usually depends on the type of machine and the requirements regarding duration of service and operational reliability.

Influence of operating temperature on bearing material

At elevated temperatures the dynamic load carrying capacity is reduced. The reduction in dynamic load carrying capacity at different temperatures is taken into account by multiplying the basic dynamic load rating C by a temperature factor obtained from the following table

Table 3.1: Values of temperature factor.

Bearing temperature, °C	150	200	250	300
Temperature factor	1.00	0.90	0.75	0.60

A satisfying operation of bearings at elevated temperatures also depends on whether the bearings have adequate dimensional stability for the operating temperature, if the chosen lubricant will retain its lubricating properties and if the materials of the seals, cage etc. are suitable.

3.3.5 Adjusted rating life equation

In the classic life equation only the influence of bearing load on the life of a given bearing is considered. Where bearings are used in conventional applications, a calculation of the basic rating life L_{10} is normally adequate, since the recommendations regarding requisite life are

based on experience and in fact, therefore, consider factors such as lubrication.

It may be desirable to consider other factors influencing bearing life in more detail. ISO [**6**] introduces an adjusted rating life equation

$$L_{na} = a_1 a_2 a_3 L_{10} \tag{3.6}$$

where

L_{na} adjusted rating life in millions of revolutions (the index n represents the difference between the required reliability and 100%. The term reliability refers to the probability that a bearing will attain or exceed a specified life.

a_1 life adjustment factor for reliability.

a_2 life adjustment factor for material.

a_3 life adjustment factor for operating conditions.

A calculation of the adjusted rating life requires that the operating conditions are well defined and that the bearing loads can be accurately determined, i.e. the calculation should consider the load spectrum, shaft deflection etc.

For the generally accepted reliability of 90% and for bearing materials to which the dynamic load carrying capacity, C, values correspond, and for normal operating conditions, $a_1 = a_2 = a_3 = 1$ and the equations for the basic and adjusted rating lives become identical.

Life adjustment factor a_1

The a_1 factor for reliability is used to determine lives other than the L_{10} life, i.e. lives with a greater probability than 90%. Values of a_1 are given in Table 3.2.

Table 3.2: Values of life adjustment factor a_1 [4].

Reliability %	L_{na}	a_1 [–]
90	L_{10a}	1
95	L_{5a}	0.62
96	L_{4a}	0.53
97	L_{3a}	0.44
98	L_{2a}	0.33
99	L_{1a}	0.21

Life adjustment factor a_2

The standard steels used by bearing manufacturers have better life properties than the steels on which the ISO 281:1990 standard is based. Therefore, the use of $a_2 = 1$ includes a considerable safety margin.

Life adjustment factor a_3

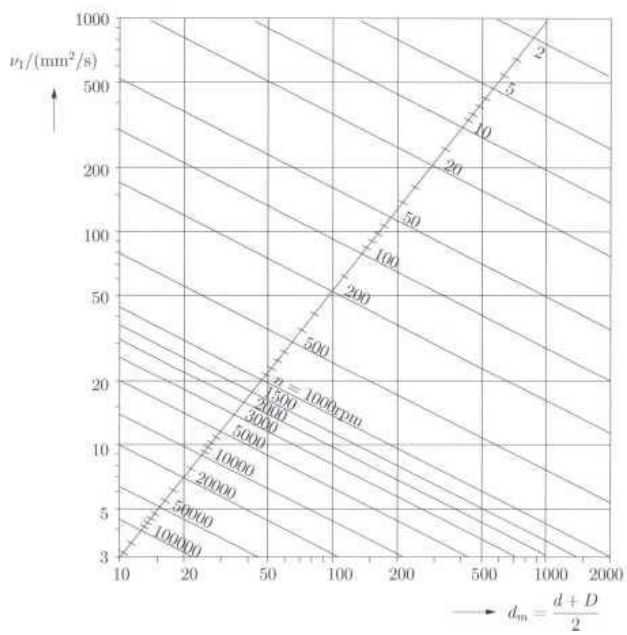
The operating conditions factor a_3 is essentially determined by bearing lubrication if the bearing

operating temperatures are not excessive. Changes in material properties at elevated temperatures are accounted for by reducing the basic dynamic load ratings, see under "Influence of operating temperature". The efficiency of lubrication is primarily determined by the degree of surface separation in the rolling contacts of the bearing. If an adequate load-carrying lubricant film is to be formed, the lubricant must have a given minimum viscosity at the operating temperature, i.e. the temperature of the bearing in operation. Under the cleanliness conditions normally prevailing in an adequately sealed bearing arrangement, the a_3 factor is based on the viscosity ratio κ. This is defined as the ratio of the actual viscosity ν to the viscosity ν_1 required for adequate lubrication, both values being at the operating temperature.

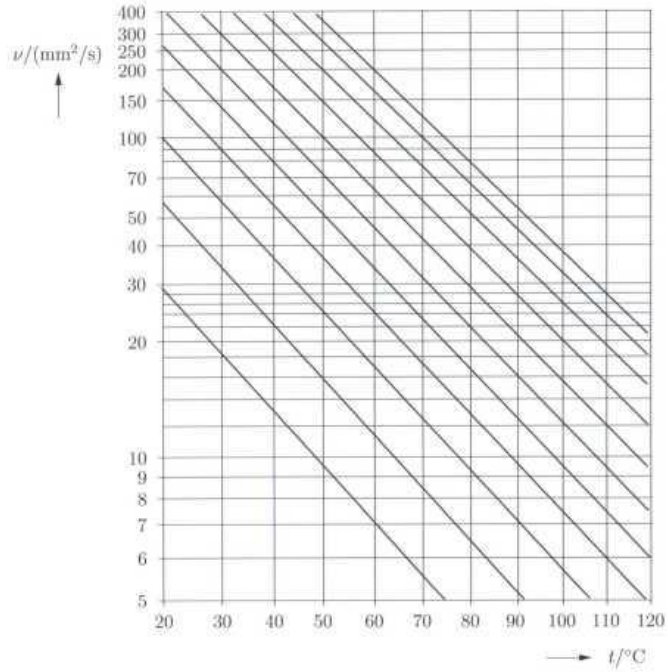
The selection of an oil is primarily based on the viscosity required to provide adequate lubrication for the bearing at the operating temperature.

In order for a sufficiently thick film of oil to be formed in the contact area between rolling elements and raceways, the oil must retain a minimum viscosity at the operating temperature. The kinematic viscosity v_1 required at the operating temperature to ensure adequate lubrication can be determined from Figure 3.13 provided a mineral oil is used. When the operating temperature is known from experience or otherwise determined, the corresponding viscosity at the internationally standardized reference temperature of 40°C can be obtained from Figure 3.14.

Table 3.4 lists the ISO viscosity classes showing the range of viscosity for each class at 40°C. Certain bearing types, e.g. spherical roller bearings, tapered roller bearings and spherical roller thrust bearings, normally have a higher operating temperature than deep groove ball bearings and cylindrical roller bearings under comparable operating conditions.



[billedtekst.start]**Figure 3.13:** Recommended lubricant viscosity at operating temperature [4].[billedteskst.slut]



[billedtekst.start]**Figure 3.14:** Temperature-viscosity relationship [4].[billedteskst.slut]

3.3.6 Combination of life adjustment factors a2 and a3

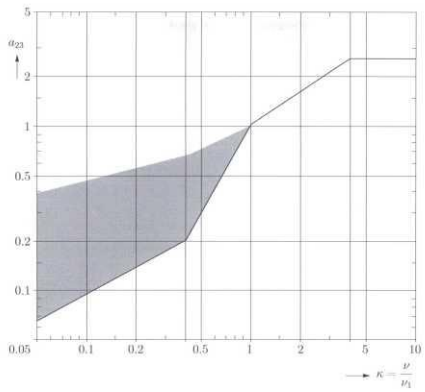
The factors a_2 and a_3 are interdependent as explained above and the manufacturer SKF has replaced them in the adjusted rating life equation by the combined factor a_{23} for material and lubrication so the equation becomes

$$L_{na} = a_1 a_{23} L_{10} \tag{3.7}$$

Provided cleanliness is normal, values of a_{23} can be obtained from Figure 3.15 as a function of the viscosity ratio $\kappa = \frac{\nu}{\nu_1}$. If lubricants containing additives of the EP type are used, higher values may be obtained when $\kappa < 1$ (shaded area).

3.3.7 SKF Life Theory

The classic L_{10} life equation standardized by ISO has been expanded to take the fatigue load limit and several other factors related to lubrication and contamination into account.



[billedtekst.start]**Figure 3.15:** The relationship between κ and a_{23} [4].[billedteskst.slut]

The fatigue load limit P_u represents the load below which fatigue will not occur in the bearing. Values of P_u will be found in the product tables.

To give an idea of the significance of the life theory, a simplified equation illustrating the relationship with the two ISO rating life equations has been derived

$$L_{naa} = a_1 a_{SKF} \left(\frac{C}{P}\right)^p \tag{3.8}$$

or simply

$$L_{naa} = a_1 a_{SKF} L_{10} \tag{3.9}$$

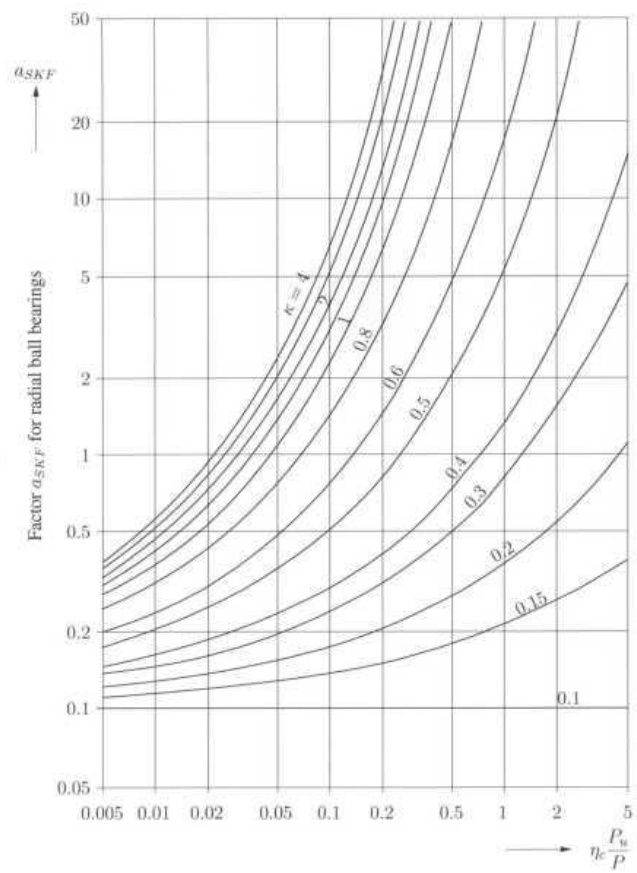
Where

L_{naa} rating life to the SKF Life Theory, millions of revolutions

a_1 life adjustment factor for reliability, see Table 3.2

a_{SKF} life adjustment factor based on the SKF Life Theory

Life adjustment factor a_{SKF}. This factor represents a verycomplex relationship of several factors including lubrication conditions and is related to the viscosity ratio κ, see Figure 3.15. Values of a_{SKF} are given as a function of η_c (P_u/P) for different values of κ in the Figure 3.16.



[billedtekst.start]**Figure 3.16:** The a_{SKF} factor for ball bearings [4], If $\kappa > 4$ use $\kappa = 4$ curve. As the value of $\eta_c \cdot (P_u/P)$ tends to zero, a_{SKF} tends to 0.1 for all values of κ.[billedteskst.slut]

The figures are drawn for typical values of a general safety factor of the type normally associated with fatigue load limits for other mechanical components. The value of this factor depends on the bearing type.

Adjustment factor η_c for contamination. This factor has been introduced to take contamination into account. The influence of contamination on bearing fatigue life depends on a number of parameters including bearing size, relative lubricant film thickness, size and distribution of solid contaminant particles

as well as types of contaminant (soft, hard etc.) The influence of these parameters on bearing life is complex and many of the parameters are difficult to quantify. It is therefore difficult to allocate precise values to η_c which would have general validity. However, some guidance will be found from Table 3.3.

Table 3.3: Values of adjustment factor η_c for different degrees of contamination [4],

Contamination condition	**Adjustment factor** η_c
Very clean Debris size of the order of the lubricant film thickness	1
Clean Conditions typical of bearings greased for life and sealed	0.8
Normal Conditions typical of bearings greased for life and shielded	0.5
Contaminated Conditions typical of bearings without integral seals; coarse lubricant filters and/or particle ingress from surroundings	0.5-0.1
Heavily contaminated	0

An indication of the strong effect of contamination on fatigue life can be obtained from the following example. Deep groove ball bearings 6305 with and without seals were tested in a strongly contaminated environment (gearbox with considerable artificially introduced debris). No failures of the sealed bearings occurred and the tests were stopped for practical reasons after the sealed bearings had run for periods that were at least 30 times longer than the experimental lives of the unsealed bearings. The unsealed bearing lives equalled 10% of the calculated L_{10} life, which corresponds to a η_c factor of 0 as indicated in Table 3.3.

The Figure 3.16 indicates the great importance of cleanliness in lubrication by the rapid reduction in the a_{sKF} factor with diminishing η_c. When bearings with integral seals are used, contamination of the bearing can be kept to a minimum, but the lives of lubricant and seals must also be taken into consideration.

3.4 Calculation example

A deep groove ball bearing 6309 made of standard steel is to operate at a speed of 5000rpm under a constant radial load F_r = 8000N . Oil lubrication is to be used, the oil having a kinematic

viscosity v = 20mm²/s at the operating temperature. The desired reliability is 90% and it is assumed that the operating conditions are ultra-clean. What will be the L_{10}, L_{na} and L_{naa} lives?

a) **Basic rating life** L_{10} (for 90% reliability) From the product tables, the basic dynamic load rating for bearing 6309 is C = 52700N. Since the load is purely radial, $P = F_r$ = 8000N (see "Equivalent dynamic bearing load") and therefore

$$L_{10} = \left(\frac{52700}{8000}\right)^3 = 286 \text{ million revolutions} \tag{3.10}$$

b) **Adjusted rating life** L_{na}

$$L_{na} = a_1 a_{23} L_{10} \tag{3.11}$$

Table 3.4: Specification of kinematic viscosity [4].

Viscosity class according to ISO	**Kinematic viscosity at** 40° C [mm²/s]		
	mean	min	max
ISO VG 2	2.2	1.98	2.42
ISO VG 3	3.2	2.88	3.52
ISO VG 5	4.6	4.14	5.06
ISO VG 7	6.8	6.12	7.48
ISO VG 10	10	9.00	11.0
ISO VG 15	15	13.5	16.5
ISO VG 22	22	19.8	24.2
ISO VG 32	32	28.8	35.2
ISO VG 46	46	41.4	50.6
ISO VG 68	68	61.2	74.8
ISO VG 100	100	90.0	110
ISO VG 150	150	135	165
ISO VG 220	220	198	242
ISO VG 320	320	288	352
ISO VG 460	460	414	506
ISO VG 680	680	612	748
ISO VG 1000	1000	900	1100
ISO VG 1500	1500	1350	1650

Since a reliability of 90% is required, the L_{10a}life is to be calculated and a_1 = 1, see Table 3.2. The a_{23} factor is found in the following way: for bearing 6309 using d and D from the product tables,

d_m = 72.5mm and the requisite oil viscosity at the operating temperature for a speed of 5000rpm is v_1 = 7mm²/s, $k = v/v_1$ = 2.7 and the value of a_{23} = 1.92.

$$L_{10a} = 1 \cdot 1.92 \cdot 286 = 550 \text{ million revolutions} \quad (3.12)$$

c) Rating life to SKF Life Theory

$$L_{naa} = a_1 a_{SKF} L_{10} \quad (3.13)$$

As the desired reliability is 90%, the L_{10aa} life is calculated and a_1 = 1. From the product tables P_u = 1.340kN and P_u/P = 1.34/8 = 0.17. As the conditions are ultra-clean η_c = 1 and therefore for K = 2.7 the value of $CISKF$ is a_{sKF} = 14 so that according to the SKF Life Theory

$$L_{10aa} = 1 \cdot 14 \cdot 286 = 4000 \text{ million revolutions} \quad (3.14)$$

To obtain the corresponding lives in operating hours we multiply by $(1 \cdot 10^6)/(60n)$ where n = 5000rpm. The different lives are then L_{10h} = 950 operating hours, L_{10ah} = 1800 operating hours, L_{10aah} = 13300 operating hours.

If the example were to be calculated for contaminated conditions such that η_c = 0.2, then a_{sKF} = 0.3 and

$$L_{10aa} = 1 \cdot 0.3 \cdot 286 = 86 \text{ million revolutions} \quad (3.15)$$

3.5 Calculation of dynamic bearing loads

The loads acting on the bearing can be calculated if the external forces are known. When calculating the load components for a single bearing, the shaft is considered as being a beam supported on rigid, moment-free supports. Elastic deformations in the bearing, the housing or the machine frame are ignored and so are the moments produced in the bearing as a result of shaft deflection. These simplifications are necessary if a bearing arrangement is to be calculated by hand. The standardized methods for calculating basic load ratings and equivalent bearing loads are based on similar assumptions.

3.5.1 Gear trains

With a gear train, the theoretical tooth forces can be calculated from the power transmitted and the design characteristics of the gear teeth. Additional dynamic forces may be present, produced either in the gear itself or by the input drive or power take-off. Other dynamic forces in gears result from errors of shape of the teeth and unbalance of the rotating components. Because of the requirements for quiet running, gears are made to accommodate high standards of accuracy to ensure that the forces are generally so small that they can be neglected when making bearing calculations. Additional forces arising from the type and mode of operation of the machines coupled to the gear can only be determined when the operating conditions are known. Their influence on the rating lives of the bearings is considered using an "operation" factor.

3.5.2 Belt drives

For belt drives it is necessary to take into account the effective belt pull (circumferential force) which is dependent on the transmitted torque, when calculating bearing loads. The belt pull must be multiplied by a factor that is dependent on the type of belt, its preload, belt tension and any additional dynamic forces. Values are usually published by belt manufacturers.

3.5.3 Equivalent dynamic bearing load

If the calculated bearing load F obtained when using the above information is found to fulfil the requirements for the basic dynamic load rating C, i.e. the load is constant in magnitude and direction and acts radially on a radial bearing or axially and centrically on a thrust bearing, then $P\ F$ and the load may be inserted directly in the life equations.

In all other cases it is necessary to calculate the equivalent dynamic bearing load. This is defined as that equivalent load, constant in magnitude and direction, acting radially on radial bearings or axially and centrically on a thrust bearing which, if applied, would have the same influence on bearing life as the actual loads to which the bearing is subjected.

3.5.4 Constant bearing load

Radial bearings are often subjected to simultaneously acting radial and axial loads. If the resultant load is constant in magnitude and direction, the equivalent dynamic bearing load P can be obtained from the general equation

$$P = XF_r + YF_a \tag{3.16}$$

where P is the equivalent dynamic bearing load in N, F_r is the actual radial bearing load in N, F_a is the actual axial bearing load in N, X is the radial load factor for the bearing and Y is the axial load factor for the bearing.

An additional axial load only influences the equivalent dynamic load P for a single row radial bearing if the ratio F_a/F_r exceeds a certain limiting factor e. With double row bearings even light axial loads are generally significant.

The same general equation is also applied for thrust bearings that can take both axial and radial loads, e.g. spherical roller thrust bearings. For thrust bearings that can carry only purely axial loads, i.e. thrust ball bearings and cylindrical, needle and taper roller thrust bearings, the equation can be simplified provided the load acts centrically.

Notice that the detailed procedure for finding the equivalent load differs from type to type of the bearings.

3.5.5 Fluctuating bearing load

In many cases the magnitude of the load fluctuates. If the load can be divided into a number of forces which are constant for a given number of revolutions, but which are different in magnitude from each other, we use "Miners rule" to determine the lifetime. F_1, F_2... are the constant loads during U_1 U_2, life fraction intervals. The sum of all life fraction intervals is

$$U = U_1 + U_2 + ... \leq 1 \tag{3.17}$$

See Figure 3.17. Denoting the number of revolutions required under load F_1 by N_1, under load F_2 by N_2 etc., and the total number of revolutions required by N we can write Miners rule as

$$\frac{N_1}{L_{10m1}} + \frac{N_2}{L_{10m2}} + \frac{N_3}{L_{10m3}} + \cdots = \frac{N}{L_{10m}} \leqq 1 \tag{3.18}$$

where L_{10m1} L_{10m2}, L_{10m3},… are the rating lives under loads F_1,F_2, ... and L_{10m} is the total rating life in the combined load situation. Rearranging gives

$$L_{10m} = \frac{N}{\frac{N_1}{L_{10m1}} + \frac{N_2}{L_{10m2}} + \frac{N_3}{L_{10m3}} + \cdots} \tag{3.19}$$

and

$$L_{10m} = \frac{1}{\frac{U_1}{L_{10m1}} + \frac{U_2}{L_{10m2}} + \frac{U_3}{L_{10m3}} + \cdots} \tag{3.20}$$

If bearing speed is constant and the bearing load direction is constant, but the magnitude of the load constantly fluctuates between a minimum value F_{min} and a maximum value F_{max} , see Figure 3.17, the mean load can be obtained from

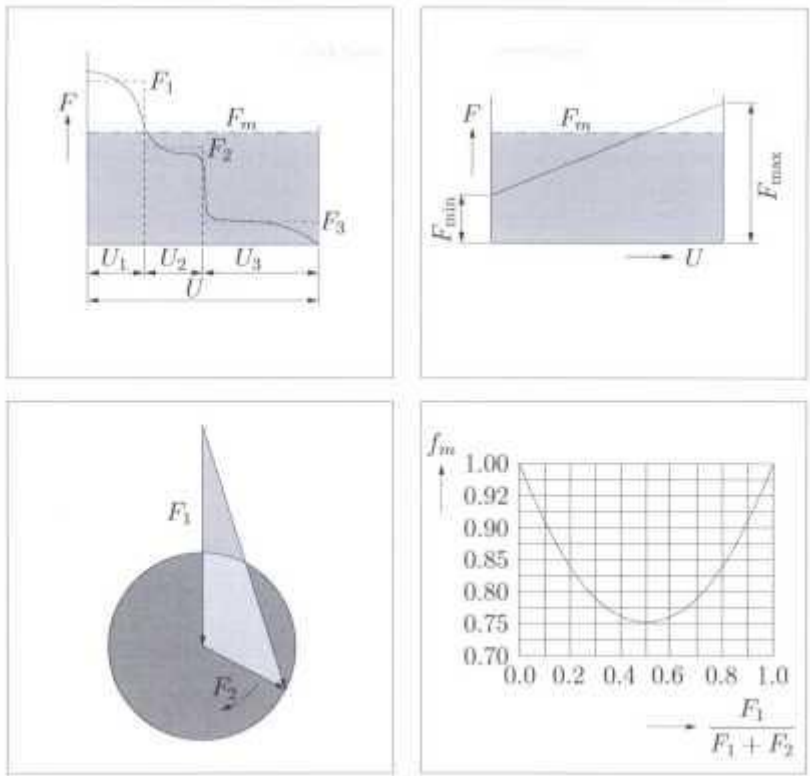
$$F_m = \frac{F_{min} + 2F_{max}}{3} \tag{3.21}$$

If, as illustrated in Figure 3.17, the load on the bearing consists of a load F_1 which is constant in magnitude and direction (e.g. the weight of a rotor) and a rotating constant load F_2 (e.g. an unbalance load), the mean load can be obtained from

$$F_m = f_m (F_1 + F_2) \tag{3.22}$$

values for the factor f_m can be obtained from Figure 3.17.

If the fluctuating load acts in a purely radial direction for radial bearings and in a purely axial direction for thrust bearings, then the equivalent dynamic bearing load $P = F_m$. However if the load acts in any other direction, the general equation for the equivalent dynamic bearing load must be used and F_r and F_a are replaced by the radial and axial components of the mean load F_m respectively.



[billedtekst start]**Figure 3.17:** Types of load combinations [4],[billedtekst slut]

3.5.6 Requisite minimum load

If a rolling bearing is to operate satisfactorily it must be subjected to a given minimum load.

As a general "rule of thumb" a load corresponding to 0.02*C* should be imposed on roller bearings and a load corresponding to 0.01*C* on ball bearings. The importance of applying this minimum load increases where accelerations in the bearing are high, and where speeds are higher than 75% of the speed ratings given in the product tables.

More detailed recommendations for calculating the requisite minimum loads for the different bearing types may be given by the bearing manufacturer.

3.6 Selecting bearing size using the static load carrying capacity

Bearing size should be selected on the basis of the static load ratings C_0 instead of on bearing life under one of the following conditions:

- The bearing is stationary and is subjected to continuous or intermittent (shock) loads.
- The bearing rotates under load at very slow speed and is only required to have a

short life (the life equation in this case, for a given equivalent load P would give such a low requisite basic dynamic load rating C that the bearing selected on a life basis would be seriously overloaded in service).

- The bearing makes slow oscillating or alignment movements under load.
- The bearing rotates and in addition to the normal operating loads, it has to sustain heavy shock loads that act during a fraction of a revolution.

In these cases, the permissible load for a bearing is determined not by material fatigue, but by the permanent deformation at the rolling element/raceway contacts caused by the load.

3.6.1 Stationary bearing

Loads acting on a stationary bearing or one which is slowly oscillating, as well as shock loads on a rotating bearing which act for only part of a revolution, produce flattened areas on the rolling elements and indentations in the raceways. The indentations may be irregularly spaced around the raceway or may be evenly spaced at positions corresponding to the spacing of the rolling elements. If the load acts for several revolutions the deformation will be evenly distributed over the whole raceway.

Permanent deformations in the bearing can lead to vibration in the bearing, noisy operation and increased friction. It is also possible that the internal clearance will increase or the character of the fits may be changed.

The extent to which these changes are detrimental to bearing performance depends on the demands placed on the bearing in a particular application. It is therefore necessary to ensure that permanent deformations cannot occur or occur to a very limited extent only. This is done by selecting a bearing with sufficiently high static load carrying capacity, and at least one of the following demands has to be satisfied:

- Quiet running (e.g. for electric motors).
- Vibration-free operation (e.g. for machine tools).
- Constant bearing friction torque (e.g. for measuring apparatus and test equipment).
- Low starling friction under load (e.g. for cranes).

3.6.2 Static load rating

When determining bearing size based on static load carrying capacity a given safety factor s_0 is used to calculate the requisite basic static load rating. This factor represents the relationship between the basic static load rating C_0 and the equivalent static bearing load P_0.

Static loads comprising radial and axial components must be converted into an equivalent static bearing load. This is defined as that load (radial for radial bearings and axial for thrust bearings) which, if applied, would cause the same permanent deformation in the bearing as the actual load. It is obtained from the general equation

$$P_0 = X_0 F_r + Y_0 F_a \tag{3.23}$$

where P_0 is the equivalent static bearing load in N, F_r is the actual radial bearing load in N, F_a is the actual axial bearing load in N, X_0 is the radial load factor for the bearing and Y_0 is the axial load factor for the bearing.

When calculating P_0, the maximum load which can occur should be used and its radial and axial components inserted in the equation above. If a static load acts in different directions on a bearing, the magnitude of these components will change. In such cases, the components of the load giving the largest value of the equivalent static bearing load P_0 should be used.

3.6.3 Requisite basic static load rating

The requisite basic static load rating C_0 can be determined from

$$C_0 = s_0 P_0 \tag{3.24}$$

where

C_0[N] basic static load rating

P_0[N] equivalent static bearing load

s_0 static safety factor

3.7 Radial location of bearings – Selection of fit

When selecting a fit between inner bearing ring and shaft and between outer bearing ring and housing, the load situation should be considered.

Load conditions

There are three different load situations, characterized as "rotating load", "stationary load" and "direction of load indeterminate".

"**Rotating load**" pertains if the bearing ring rotates and the load is stationary, or if the ring is stationary and the load rotates so that all points on the raceway are subjected to load in the course of one revolution.

A bearing ring subjected to a rotating load will tend to turn on its seating if mounted with a clearance fit, and may result in wearing (fretting corrosion) of the surfaces involved. To prevent this, an interference fit must be specified. The interference required is dictated by the operating conditions.

"**Stationary load**" pertains if the bearing ring is stationary and the load is also stationary, or if the ring and the load rotate at the same speed, so that the load is always directed towards the same position on the raceway.

"**Direction of load indeterminate**" represents variable external loads, shock loads, vibrations and unbalance loads in high-speed machines. These give rise to changes in the direction of load, which cannot be accurately described. When the direction of load is indeterminate and particularly where heavy loads are involved, it is desirable that both rings have an interference fit. For the inner ring the recommended fit for a rotating load is normally used. If the outer ring must be free to move axially in the housing, a somewhat looser fit than that recommended for a rotating load may be used.

Operating and load conditions for interference fit between inner ring and shaft (condition 1)

- Rotating inner ring
- Stationary outer ring

- Constant load direction

and

- Rotating load on inner ring

- Stationary load on outer ring

Interference fit is required between inner ring and shaft as the load is varying in direction relative to the inner ring.

Clearance fit can be used between outer ring and housing as the load has a constant direction relative to the outer ring.

An example of this load situation is belt-driven shafts.

Operating and load conditions for interference fit between inner ring and shaft (condition 2)

- Stationary inner ring
- Rotating outer ring
- Load rotates with outer ring

or

- Rotating load on inner ring
- Stationary load on outer ring

Interference fit is required between inner ring and shaft as the load is varying in direction relative to the inner ring.

Clearance fit can be used between outer ring and housing as the load has a constant direction relative to the outer ring.

An example of this load situation is merry-go-round drives.

Operating and load conditions for interference fit between outer ring and housing (condition 1)

- Stationary inner ring
- Rotating outer ring
- Constant load direction

and

- Stationary load on inner ring
- Rotating load on outer ring

Interference fit is required between outer ring and housing as the load is varying in direction relative to the outer ring.

Clearance fit can be used between inner ring and shaft as the load has a constant direction relative to the inner ring.

An example of this load situation is conveyor idlers and car wheel hub bearings.

Operating and load conditions for interference fit between outer ring and housing (condition

2)

- Rotating inner ring
- Stationary outer ring
- Load rotates with inner ring

or

- Stationary load on inner ring
- Rotating load on outer ring

Interference fit is required between outer ring and housing as the load is varying in direction relative to the outer ring.

Clearance fit can be used between inner ring and shaft as the load has a constant direction relative to the inner ring. Examples of this load situation are vibratory applications and vibrating motors.

Magnitude of the load

The interference fit of a bearing inner ring on its seating will be loosened with increasing load as the ring deforms. Under the influence of rotating load the ring may begin to creep. The degree of interference must be related to the magnitude of the load; the heavier the load, the greater the interference fit required.

Bearing internal clearance

An interference fit of a bearing on a shaft or in a housing means that the ring is elastically deformed (expanded or compressed) and the bearing internal clearance is reduced. A certain minimum clearance should remain.

Temperature conditions

In many applications the outer ring has a lower temperature in operation than the inner ring. This might lead to reduced internal clearance.

In service, bearing rings normally reach a temperature that is higher than the one of the components to which they are fitted. This can result in a reduction of the fit of the inner ring on its seating, while an outer ring expansion may prevent the desired axial displacement of the ring in its housing. Temperature differentials and the direction of heat flow in the bearing arrangement must therefore be carefully considered.

Running accuracy requirements

To reduce resilience and vibration, clearance fits should generally not be used for bearings where high demands are placed on running accuracy. Bearing seats on the shaft and in the housing should be made to tolerance grade 5 for the shaft and to tolerance grade 6 for the housing. Tight tolerances should also be applied to the cylindricity.

Displacement of the non-locating bearing

If non-separable bearings are used as non-locating bearings it is imperative that one of the bearing rings is free to move axially at all limes during operation. Adopting a clearance fit for the ring that carries a stationary load will provide this.

If cylindrical roller bearings have one ring without flanges, needle roller bearings or

CARB toroidal roller bearings are used, both bearing rings may be mounted with an interference fit because axial displacement will take place within the bearing.

3.8 Bearing lubrication

The selection of an oil is primarily based on the viscosity required to provide adequate lubrication for the bearing at the operating temperature.

The viscosity of an oil is temperature dependent, so the viscosity becomes lower as the temperature rises.

In order to obtain a sufficiently thick film of oil in the contact area between rolling elements and raceways, the oil must retain a minimum viscosity at the operating temperature. The kinematic viscosity v_1 required at the operating temperature to ensure adequate lubrication can be determined from Figure 3.14, provided a mineral oil is used.

When the operating temperature is known from experience or otherwise determined, the corresponding viscosity at the internationally standardized reference temperature of 40°C or other test temperatures (e.g. 20°C) can be obtained from Figure 3.15, which is compiled for a viscosity index of 85 or can be calculated.

Table 3.4 lists the ISO viscosity classes showing the range of viscosity for each class at 40°C. Certain bearing types, e.g. spherical roller bearings, taper roller bearings and spherical roller thrust bearings, normally have a higher operating temperature than other bearing types, e.g. deep groove ball bearings and cylindrical roller bearings under comparable operating conditions.

When selecting the oil the following aspects should be considered.

- Bearing life may be extended by selecting an oil whose viscosity v at the operating temperature is somewhat higher than $v\backslash$. However, since increased viscosity raises the bearing operating temperature there is frequently a practical limit to the life enhancement obtained by this means.
- If the viscosity ratio $\kappa = v/v_1$ is less than 1 an oil containing EP additives is recommended and if κ is less than 0.4 an oil with such additives must be used. An oil with EP additives may also enhance operational reliability in cases where $\kappa > 1$ and medium and large size roller bearings are in operation.

Table 3.5: Fits between bearing inner ring and shaft [4],

Fits for solid steel shafts					
Radial bearings with cylindrical bore					
Conditions	**Examples**	**Shaft diameter, mm**			**Tolerance**
		Ball bearings	Cylindrical needle and taper roller bearings	Spherical roller roller	
Rotating inner ring load or direction of load indeterminate					
Light and variable loads ($P \leqq 0.06C$)	Conveyor, lightly loaded gearbox bearings	(18)to 100	≦ 40	–	j6
		(100) to 140	(40) to 100	–	k6
Normal and heavy loads ($P > 0.06C$)	Bearing applications generally, electric motors, turbines, pumps internal combustion engines, gearing woodworking machines	≦18	–	–	j5
		(18) to 100	≦ 40	≦ 40	k5(k6)
		(100) to 140	(40) to 100	(40) to 65	m5(m6)
		(140) to 200	(100) to 140	(65) to 100	m6
		(200) to 280	(140) to 200	(100) to 140	n6
		–	(200) to 400	(140) to 280	p6
		–	–	(280) to 500	r6
		–	–	> 500	r7
Very heavy loads and shock loads with difficult working conditions ($P > 0.12C$)	Axleboxes for heavy railway vehicles, traction motors, rolling mills	–	(50) to 140	(50) to 100	n6
		–	(140) to 200	(100) to 140	p6
		–	> 200	> 140	r6

High demands on running accuracy with light loads ($P \leqq 0.06C$)	Machine tools	≦ 18	–	–	h5
		(18) to 100	≦ 40	–	j5
		(100) to 200	(40) to 140	–	k5
		–	(140) to 200	–	m5
Stationary inner ring load					
Easy axial displacement of inner ring on shaft desirable	Wheels on non–rotating axis				g6
Easy axial displacement of inner ring on shaft unnecessary	Tension pulleys, rope sheaves				h6
Axial loads only					
	Bearing applications of all kinds	≦ 250	≦ 250	≦ 250	j6
		> 250	> 250	> 250	j s6

Table 3.6: Fits between bearing outer ring and housing [4].

Fits for cast iron and steel housings			
Radial bearings - solid housing			
Conditions	**Examples**	**Tolerance**	**Displacement of outer ring**
Rotating outer ring load			
Heavy loads on bearings in thin-walled housings, heavy shock loads (P > 0.12*C)*	Roller bearing wheel hubs, big-end bearings	P7	Cannot be displaced
Normal and heavy loads *(P* > 0.06C)	Ball bearing wheel hubs big-end bearings, crane traveling wheels	N7	Cannot be displaced
Light and variable loads (P ≦ 0.06C)	Conveyor rollers, rope sheaves, belt tension pulleys	M7	Cannot be displaced
Direction of load indeterminate			
Heavy shock loads	Electric traction motors	M7	Cannot be displaced
Normal and heavy loads (P > 0.06*C*), axial displacement of outer ring unnecessary	Electric Motors, pumps crankshaft bearing	K7	Cannot be displaced as a role
Accurate or silent running			
	Small electric motors	J6	Can be displaced

3.9 Nomenclature

a_x	-	Life adjustment factor for reliability
asKF	-	Life adjustment factor - SKF life theory

d	mm	Bearing diameter
dm	mm	Bearing mean diameter
FM	-	Load factor
N	rpm	Rotational speed
P	-	Exponent in dynamic bearing lifetime expression
t	°C	Temperature
C	N	Dynamic load carrying capacity
C_0	N	Static load carrying capacity
D	mm	External diameter of bearing
Dw	mm	Rolling element diameter
F_a	N	Axial load on bearing
F_m	N	Mean load (axial or radial) on bearing

F_r	N	Radial load on bearing
L_{10}	mill.rev	Bearing lifetime
L_{10h}	hours	Bearing lifetime
N	–	Number of revolutions
P	N	Actual dynamic load on bearing
P_0	N	Actual static load on bearing
P_u	–	Fatigue load limit
U	-	Lifetime
X	–	Radial load factor
Y	–	Axial load factor
β	–	The contact angle
γ	rad	Oscillating amplitude angle
K	–	Viscosity ratio
η_c	–	Contamination factor
V	mm^2/s	Kinematic viscosity of lubricant

3.10 References

[1] A. D. S. Carter. *Mechanical Reliability.* MacMillan, 1986.

[2] B. J. Hamrock, B. Jacobson, and S. R. Schmid. *Fundamentals of machine elements.* McGraw Hill, 1999.

[3] T. A. Harris. *Rolling Bearing Analysis.* John Wiley and Sons, 1984.

[4] SKF. *General catalogue.* Catalogue 4000/IV E.

[5] SKF and ISO 281:1990.

[6] SKF and ISO 281:1990/Amd 2:2000.

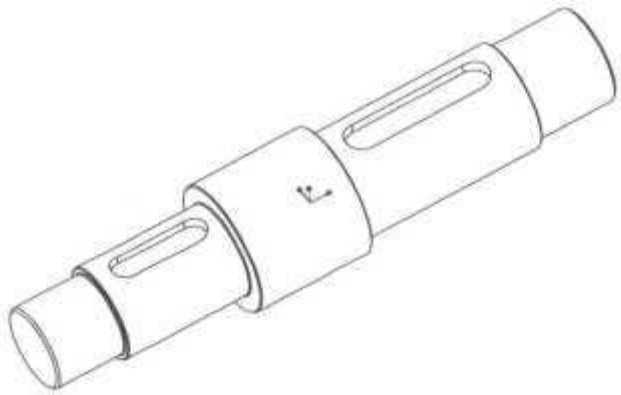
[7] W. Steinhilper and R. Roper. *Maschinen- und Konstruktionselemente.* Springer-Verlag, 1991.

Chapter 4 Shafts

4.1 Introduction

A shaft or axle is usually of circular cross section and mounted with machine elements such as bearings, gears, pulleys, sprockets and other machine elements. The analysis and corresponding dimensioning of the shaft or axle depend strongly on whether it is rotating or stationary.

Shafts and axles are normally designed for the specific application unlike bearings, belts and chain drives. Each shaft and axle should be designed specifically for the application, considering the mounted elements, operating loads and other operating conditions. An example of a shaft is shown in Figure 4.1.



[billedtekst start]**Figure 4.1:** Example of a shaft.[billedtekst slut]

4.1.1 Terminology

Before going further into the analysis it might be useful to clarify the difference between shafts and axles based on their distinct ways of operation.

Axles. The main purpose of an axle is to support different types of machine elements and it is only loaded with shear forces and bending moments. The axle may be fixed to a frame or support or it may rotate. Typical examples of non-rotating axles are the axle in the wire pulley in a lifting block or an axle carrying the wheels of a railway passenger car. The pulleys, wheels and so on are mounted on bearings

that in turn are mounted on the axle. Rotating axles are used in other applications. Examples of rotating axles are seen in support rollers in belt conveyors or the axles in goods wagons.

Shafts. Shafts are normally rotating and the main purpose is to transmit power and thereby torque. A shaft is mounted with machine elements which can lead the torque into and out from the shaft. Typical elements mounted on shafts are gears, belt pulleys, chain sprockets, couplings, brakes and so on. Typical examples of shafts are the input -, output - and intermediate shaft in a gear transmission. A shaft is supported in at least two bearings. Long and slender shafts, as for example propeller shafts in ships may be supported by more than two bearings to prevent transversal vibrations.

In this chapter we will focus on shafts since they represent the most general load situation, but the equations are fully applicable to axles. It is through out the chapter assumed that the shaft is made from a ductile material.

4.2 Types of load

Before a strength (stress) based dimensioning of a shaft can be done the load situation must be clarified. This means that, i.e., bearing reactions and distribution of axial force, bending moment and torque must be found as a first step.

Shear forces Shear force loading is especially note-worthy for short shafts. The slenderness ratio is defined as

$$\Gamma = L/\sqrt{I/A} \tag{4.1}$$

where L is the effective length of the shaft (here we make the assumption that L is the distance between supporting bearings), I cross sectional moment of inertia and A the cross sectional area. From e.g. [6] we know that if $\Gamma < 20$ we should include the influence from the shear force on the deflection, i.e., we cannot use simple Bemoulli-Euler beam theory. For circular cross sections $\Gamma < 20$ corresponds to $L/D < 5$ where D is the diameter of the shaft.

Shafts in gearboxes are normally very short in order to prevent bending deflection of the shaft and here the maximum shear deformation is of the same order of magnitude as the bending deformation. Generally, the shear stress is quite small in short shafts in gearboxes.

Bending moment Bending moment loading on a shaft is caused by the forces acting on the machine elements mounted on the shaft. Power transmitting components such as gear wheels and belt pulleys give rise to forces and moments giving the bending moment loading. The maximum bending stress appears at the outer surface of the shaft.

Torque In power transmitting components a torsional moment (a torque) is found in all or part of the shaft.

Shafts may be subjected to a combination of loads coming from axial forces, bending and torsional moments. The loads may be combined of stationary and time varying components, depending on the specific application. Shafts used in gears for transmission of power will be

subjected to an almost constant torque, together with a reversed bending moment(see later in this chapter).

4.3 Shaft design considerations

In normal shaft design the shaft is usually designed specifically for a certain purpose. This means that a shaft is normally not just a straight shaft with no changes in the cross sectional dimension. We have shoulders/steps to accommodate elements such as bearings, sprockets and gears etc. Generally, there will also be any given number of shaft-hub connections including, i.e., keys, snap rings and cross pins. This is described in details in Chapter 5. Most of these connections require grooves or holes through the shaft. Each of these changes in geometry will cause stress concentrations. By different means we may reduce the size of the stress concentrations, but they will be present and we must deal with them. Also important in shaft design is that the support and the connection of the shaft should be such that it will allow for elongation due to temperature in a way that does not result in stresses.

The shaft designer needs to make decisions on a number of issues (in mentioned order):

– Size and spacing of components being supported
– Bearing positions, to support the shaft
– Appropriate shaft design considering assembly possibilities
– Attachment method for transmission elements to the shaft (pins, keys, splines, press or shrink fit)
– Appropriate shaft diameters and shaft material

4.3.1 Possible modes of failure

The shaft geometry may be determined by considering a number of parameters:

- Maximum allowable stress (yield criterion)
- Fatigue failure
- Maximum allowable deflections
- Critical speeds

hopefully it should be clear that the design of shafts involves many steps and iterations in the design process are probably needed.

4.4 Static loading

Stresses developed in a shaft during operation should never exceed the yield stress of the shaft material. As previously mentioned a shaft is normally subject to a combination of loads, which in turn means that the shaft experiences a combination of stresses such as bending stress, torsional shear stress, shear stress due to shear force etc. To compare the combined stress situation to the materials yield strength a reference stress must be determined based on one of the known stress criteria such as the "Distortion Energy Theory" (von Mises) or the "Maximum

Shear Stress Criterion" (Tresca). In this chapter we consider the criterion based on von Mises.

To find the stresses in the shaft it is necessary to find the bending moment distribution and the distribution of the shear stress. Shafts are inherently three dimensional so static equilibrium in 3D is needed.

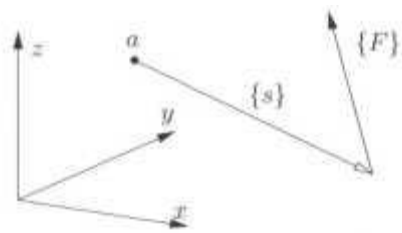
Static equilibrium in 3D

The present subsection is included to show a simple method of describing equilibrium in 3D, especially in relation to the moment equilibrium. The general definition of static equilibrium is

$$\sum\{F\} = \{0,0,0\}^T \tag{4.2}$$

$$\sum\{M\} = \{0,0,0\}^T \tag{4.3}$$

i.e., the sum of forces is zero and the sum of moments is zero. The shaft is not accelerated, but is rotating at constant speed so we should include the gyroscopic forces due to the rotation in the moment equilibrium, see e.g. [5], If we assume that the axis of rotation is balanced, i.e., the rotation axes is a principal axes of inertia then the gyroscopic forces are zero and we are left with (4.3).



[billedtekst start]**Figure 4.2:** General force in 3D.[billedtekst slut]

The moment of a force around a point, a, as shown in Figure 4.2 is by definition given as a cross product

$$\{M\} = \{S\} \times \{F\} \tag{4.4}$$

with distance vector $\{s\} = \{s_1, S_2, S_3\}^T$ and force vector $\{F\} = \{F_1\ F_2, F_3\}^T$. The definition given in (4.4) is directly related to Figure 4.2 and the definition of the distance vector $\{S\}$ (with other definitions the sign changes). It follows that in order to perform successfully moment equilibrium, we must be able to perform the cross product of two vectors.

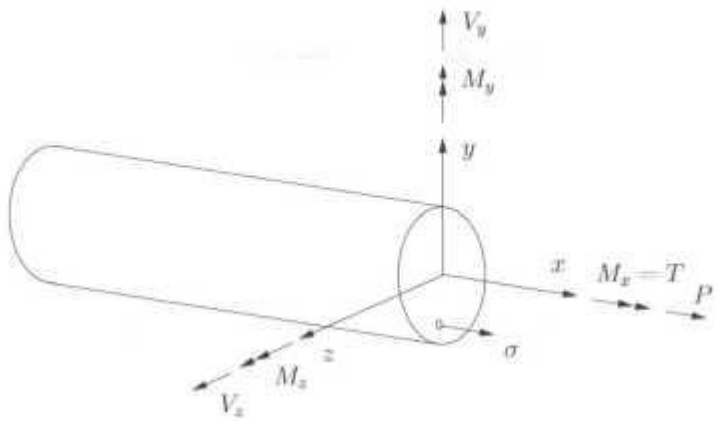
It can be shown that

$$\{M\} = \{s\} \times \{F\} = [\tilde{s}]\{F\} \tag{4.5}$$

where the skew-matrix $[\tilde{s}]$ is defined as

$$[\tilde{s}] = \begin{bmatrix} 0 & -s_3 & s_2 \\ s_3 & 0 & -s_1 \\ -s_2 & s_1 & 0 \end{bmatrix} \tag{4.6}$$

The matrix $[\tilde{s}]$ is a skew symmetric matrix, i.e., $[\tilde{s}]^T = -[\tilde{s}]$ (therefore the name). Many mathematical simplifications can be made using the skew-matrix as seen in [3], The primary advantage of the definition of the skew-matrix is that it transforms the cross product into a form where all standard vector matrix operations apply.



[billedtekst start]**Figure 4.3:** Definition of forces, moments and normal stress on beam cross section.[billedtekst slut]

Bending moment, torsion and axial loading

Assuming that the reaction forces are known, we can find the moment distribution in the shaft. In 3D we use the definition given in Figure 4.3.

Due to the symmetry of the shaft we have that the normal strain in the original straight beam is

$$\epsilon(y,z) \simeq \epsilon_0 - \frac{y}{\rho_y} - \frac{z}{\rho_z} \tag{4.7}$$

where ϵ_0 is the strain at the neutral axes, ρ_z and ρ_y is radius of curvature of the beam in the plane x–y and $z - x$ respectively. The normal stress is given as

$$\sigma(y,z) = E\epsilon(y,z) \tag{4.8}$$

where E is the modulus of elasticity. By force equilibrium we have

$$P = E\epsilon_0 \int_A dA \tag{4.9}$$

$$M_y = -\frac{E}{\rho_z}\int_A z^2 dA = -\frac{E}{\rho_z} I_y \tag{4.10}$$

$$M_z = \frac{E}{\rho_y}\int_A y^2 dA = \frac{E}{\rho_y} I_z \tag{4.11}$$

where A is the cross sectional area and the definition of the cross sectional moment of inertia is used for I_y and I_z.

For a circular shaft of diameter d we have that $I_y = I_z$ and the maximum stress is therefore given by

$$|\sigma_{\max}| = \left|\frac{P}{A}\right| + \frac{\sqrt{M_y^2 + M_z^2} \cdot \frac{d}{2}}{I_y} \tag{4.12}$$

i.e., the maximum bending moment is given by

$$M_{\max} = \sqrt{M_y^2 + M_z^2} \tag{4.13}$$

Both the shear forces V_x and V_y and the torsion T result in shear stress. We may define the resulting shear force as

$$V = \sqrt{V_x^2 + V_y^2} \tag{4.14}$$

For a circular cross section we then know that the maximum shear stress due to the shear force is

$$\tau_{\max} = \frac{16}{3}\frac{V}{\pi d^2} \tag{4.15}$$

If we assume that the shear forces and bending moments are the result of the same external forces then they are perpendicular to each other. From this follows that the normal stress due to the bending moment is maximum, when the shear stress due to the shear force is zero and opposite. This is however not the case for the shear stress from torsion, which is given by

$$\tau = \frac{Tr}{K} \tag{4.16}$$

where $T = M_x$ is the torsional moment, r is the distance from the center and K is the cross sectional torsional stiffness factor. For circular cross sections we know that K is identical to the polar cross sectional moment $K = I_p = 2I_x = 2I_y$.

It follows that the stress is multi-axial and we can calculate a reference stress (non-physical but related to the work of the deviatoric forces) defined by von Mises, in this case given by

$$\sigma_{ref} = \sqrt{\sigma^2 + 3\tau^2} \tag{4.17}$$

It is important to remember that stresses are defined in points, i.e., the normal stress σ and the shear stress τ must be at the same point to be used in (4.17). For a circular shaft this leads to

$$I_y = I_z = \frac{\pi}{64}d^4 \quad I_p = \frac{\pi}{32}d^4 \quad A = \frac{\pi}{4}d^2 \quad \Rightarrow$$

$$\sigma_{\max} = \frac{4P}{\pi d^2} + \frac{32M_{\max}}{\pi d^3} \tag{4.18}$$

$$\tau_{\max} = \frac{16T}{\pi d^3} \tag{4.19}$$

You should notice that the shear stress specified in (4.19) is only due to the torsion. We have

$$\sqrt{\sigma_{\max}^2 + 3\tau_{\max}^2} = \sqrt{(\frac{4P}{\pi d^2} + \frac{32M_{\max}}{\pi d^3})^2 + 3(\frac{16T}{\pi d^3})^2}$$
$$= \frac{4}{\pi d^3}\sqrt{(Pd + 8M_{\max})^2 + 48T^2} \leqq \frac{\sigma_y}{n_s} \tag{4.20}$$

where σ_y is the yield stress and n_s is the safety factor. From (4.20) we see that if there is no axial loading *P,* and the required safety factor n_s is know the diameter can be explicitly given

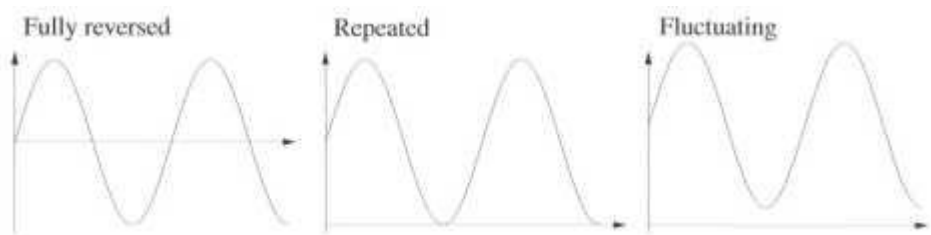
$$d = \left(\frac{4n_s}{\pi\sigma_y}\sqrt{64M_{\max}^2 + 48T^2}\right)^{1/3} \tag{4.21}$$

The stresses calculated in this section are nominal stresses. In practical designs of shafts there is a need for connecting with hubs. This requires design changes different from the simple straight shaft. These design changes normally result in stress concentrations. It is assumed that the shafts are made from ductile material and for this reason these stress concentrations are of no interest when designing with respect to static loading. The shaft will yield locally, and it is only if the whole cross section yields that we will have a problem. However we can not neglect the stress concentrations when designing with respect to fatigue as done in Section 4.5.

4.5 Design for fatigue (cyclic load/dynamic load)

Under conditions where a component is subjected to repeated application of a load, it is possible that this component will experience failure, though the applied stress is well below the yield stress. This failure mode is known as fatigue. In dynamic loading the main design criteria with respect to stresses is therefore fatigue, and we cannot compare the stress level with the yield stress. Instead we must find the material fatigue stress or endurance stress.

Assuming that the time dependent loading is sinusoidal we generally split the time dependent loading into three cases.



[billedtekst start]**Figure 4.4:** Different types of sinusoidal loading (the graphs show stress as a function of time).[billedtekst slut]

If we load a rotating shaft with a constant transverse force the bending stress will be fully reversed although the loading (force) is constant. Shafts must therefore be designed with respect to fatigue, because of the rotation.

The important stress components in relation to fatigue are the amplitude of the stress variation σ_a and the mean value σ_m, these are given by

$$\sigma_a = \frac{\sigma_{\max} - \sigma_{\min}}{2} \tag{4.22}$$

$$\sigma_m = \frac{\sigma_{\max} + \sigma_{\min}}{2} \tag{4.23}$$

in the case of sinusoidal loading as shown in Figure 10.2.

4.5.1 Stress concentration

In dynamic loading the failure mode is fracture so we cannot neglect stress concentrations as done in static loading of ductile materials. Stress concentration are normally defined as

$$\sigma_{max} = Kt\sigma_{nom} \tag{4.24}$$

$$\tau_{max} = K_{ts}\tau_{nom} \tag{4.25}$$

where K_t and K_{ts} are theoretical stress concentration factors. These factors can be found in many books, e.g., [7], [2] or [4], Examples taken from [2] is given in Appendix B. It is important that the nominal stress (σ_{nom} or τ_{nom}) is clearly defined as seen in the graphs in Appendix B.

The value of the stress concentration factor depends only on geometry. It is a theoretical factor, based on usual assumptions of elasticity. All materials do not react in the same way with respect to stress concentration, this is expressed through the notch sensitivity q. We define a new set of concentration factors termed fatigue stress concentration factors as

$$K_f = 1 + q(K_t - 1) \tag{4.26}$$

$$K_{fs} = 1 + q(K_{ts} - 1) \tag{4.27}$$

It is noted that a notch sensitivity of zero, $q = 0$, means that the material acts as if there is no stress concentration, i.e., it is insensitive to stress concentrations whereas $q = 1$ means that the material is very sensitive and will "feel" the whole stress concentration. The notch sensitivity can be approximated from

$$q \cong \frac{1}{1 + \frac{\sqrt{a}}{\sqrt{r}}} \tag{4.28}$$

where r is the notch radius and a is Neuber's constant which is material dependent, i.e., it depends on the ultimate tensile stress of the material and on the specific material. In the Tables 4.1-4.3 Neuber's constant for steel and aluminium is given.

It should be noted that the notch sensitivity depends both on material through the Neuber's constant and also on the geometry through the notch radius. In the next subsection it is shown how we may find the material endurance level that is a material parameter. The question is now if the notch sensitivity should be put on the endurance level because it depends on the material or if the notch sensitivity should be put on the stress since it depends on the geometry. In the present book (as most others) the notch sensitivity is put on the stress, but this means that the stresses calculated from

$$\sigma = K_f\sigma_{nom} \tag{4.29}$$

$$\tau = K_{fs}\tau_{nom} \tag{4.30}$$

are the theoretical values since it describes both the rise in the stress due to geometric changes and also the material sensitivity to these.

Table 4.1: Neuber's constant for steel (after [4])

σ_{ut} [MPa]	345	379	414	483	552	621	689	758

$\sqrt{a}$ [$\sqrt{\text{mm}}$]	0.655	0.595	0.544	0.469	0.403	0.353	0.312	0.277
σ_{ut} [MPa]	827	896	965	1103	1241	1379	1517	1655
$\sqrt{a}$ [$\sqrt{\text{mm}}$]	0.247	0.222	0.197	0.156	0.121	0.091	0.066	0.045

Table 4.2: Neuber's constant for hardened aluminium (after [4])

σ_{ut} [MPa]	103	138	207	276	345	414	483	552
$\sqrt{a}$ [$\sqrt{mm}$]	2.268	1.915	1.401	1.104	0.937	0.816	0.726	0.660

Table 4.3: Neuber's constant for annealed aluminium (after [4])

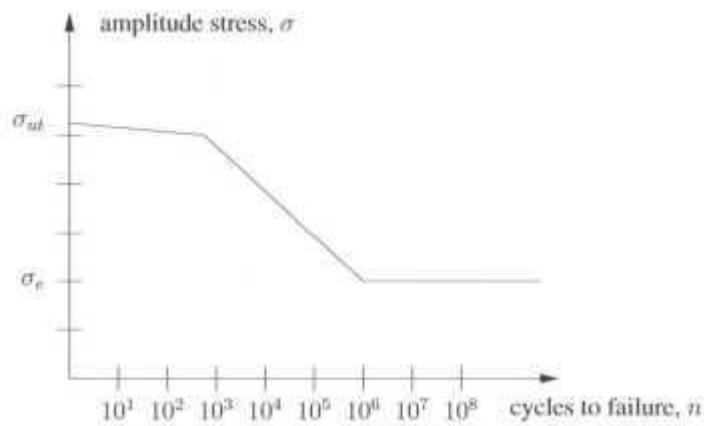
σ_{ut} [MPa]	69	103	138	172	207	241	276	310
$\sqrt{a}$ [$\sqrt{mm}$]	2.520	1.719	1.331	1.094	0.907	0.766	0.635	0.559

4.5.2 S-N curve or Wohler curve

With the stresses known we must find the material data that control fatigue. The area of fatigue is based highly on experiments. A common experiment is a rotating shaft loaded with a transverse constant load, this leads to a fully reversed loading condition

$$\sigma_m = 0 \text{ and } \sigma_a \neq 0 \tag{4.31}$$

By experiments a S-N curve is found (S for stress and N for number of loading cycles). A typical example is shown in Figure 4.5.



[billedtekst start]**Figure 4.5:** Schematic S-N curve or Wohler curve for steel, showing the ultimate tensile stress σ_{ut} and the endurance stress σ_e.[billedtekst slut]

The curve in Figure 4.5 is for steel. It gives the endurance stress for a specific number of cycles. An important point here is that for steel we have an endurance limit σ_e. i.e., a stress level at which no fatigue is possible independent on the number of cycles. This is not the case for e.g. aluminium where no lower limit exists. Therefore, aluminium cannot be designed to prevent fatigue failure. The S-N curve is a curve-fit to experiments and so it is made with a 50% failure likelihood, i.e., 50% of the experiments will lie below and 50% above. The uncertainty with respect to the S-N curve is so large that different researchers do not agree on the stress axes

being a logarithmic scale or a linear scale.

The endurance limit is found from numerous experiments and can be found in books, see e.g. [2]. Alternatively, the producers of the steel will supply these values. As indicated the S-N curve and consequently

the endurance limit is found by experiments with rotational bending of shafts. The experiments are performed under certain constraints.

- Loading condition of shaft
- Size of shaft
- Surface of shaft
- Temperature

A specific endurance level for a machine element should incorporate these items and also the needed level of reliability, remembering that the S-N curve is made for 50% reliability.

4.5.3 Estimation of endurance level

If an experimental or manufacturer given endurance level cannot be found we may estimate the endurance level as shown in [4] or [8]. The estimate for steel is shown below (taken from [4]).

An uncorrected endurance level for steel is

$$\begin{aligned} \sigma'_e &\cong 0.5\sigma_{ut} && \text{for } \sigma_{ut} < 1400\text{MPa} \\ \sigma'_e &\cong 700\text{MPa} && \text{for } \sigma_{ut} \geqq 1400\text{MPa} \end{aligned} \tag{4.32}$$

The corrected (with respect to the specific machine element) endurance level is then given as

$$\sigma_e = C_{load} C_{size} C_{surf} C_{temp} C_{reliab}\, \sigma'_e \tag{4.33}$$

The individual correction factors are:

Loading effect

bending	$C_{load} = 1.0$
axial loading	$C_{load} = 0.7$

Size effect

for $d \leqq 8$mm	$C_{size} = 1.0$
for 8mm $\leqq d \leqq$ 250mm	$C_{size} = 1.189(d/\text{mm})^{-0.097}$
for 250mm $\leqq d$	$C_{size} = 0.6$

Surface effect

$C_{surf} = B(\sigma_{ut}/\text{MPa})^b$ if $C_{sur} > 1.0$, set $C_{Surf} = 1.0$

where the constants B and b are taken from the Table 4.4, and σ_{ut} is expressed in MPa.

Temperature effect

for t $\leqq$ 450°C Ctemp = 1.0

$$\text{for } 450°\text{C} \leqq t \leqq 550°\text{C} \qquad C_{\text{temp}} = 1 - 0.0058(\underline{t}/°\text{C} - 450) \tag{4.34}$$

At higher elevated temperatures (50% of the material absolute melting temperature) creep becomes an important issue and the approach applied here is no longer valid.

Table 4.4: Coefficients for surface factor equation.

Surface finish	*B*	*b*
Ground	1.58	–0.085
Machined or cold-rolled	4.51	–0.265
Hot-rolled	57.5	–0.718
As-forged	272	–0.995

Reliability effect The reliability coefficient is given in Table 4.5.

Table 4.5: Reliability factor for standard deviation of 8% of mean value.

Reliability/ %	50	90	99	99.9	99.99	99.999
C_{reliab}	1.000	0.897	0.814	0.753	0.702	0.659

All of the corrections factors and also the uncorrected value of the endurance level are specified for steel and should not be used for other metals, see [4] for further information.

With the S-N curve and the endurance level we can deal with the case of zero mean value of stress. In the general case this is however not applicable and other methods must be used as discussed in next subsection.

4.5.4 Fluctuating load

In the general case we have

$$\sigma_m \neq 0 \text{ and } \sigma_a \neq 0 \tag{4.35}$$

Many different methods have been proposed based on experimental data, one of these is the modified Goodman diagram (modified compared to the original proposed one). This diagram is shown in Figure 4.6. In the figure is also shown some specific values σ_m^* and σ_a^*. The maximum stress $(\sigma_m^* + \sigma_a^*)$ and the minimum stress $(\sigma_m^* - \sigma_a^*)$ should lie within the max stress line and the min stress line.

An alternative to plotting the stress as a function of the mean stress is to plot the alternating stress as a function of the mean stress. This is done in Figure 4.7. Using the Goodman line and the yield line the working point (σ_m^*, σ_a^*) should lie below the fat line.

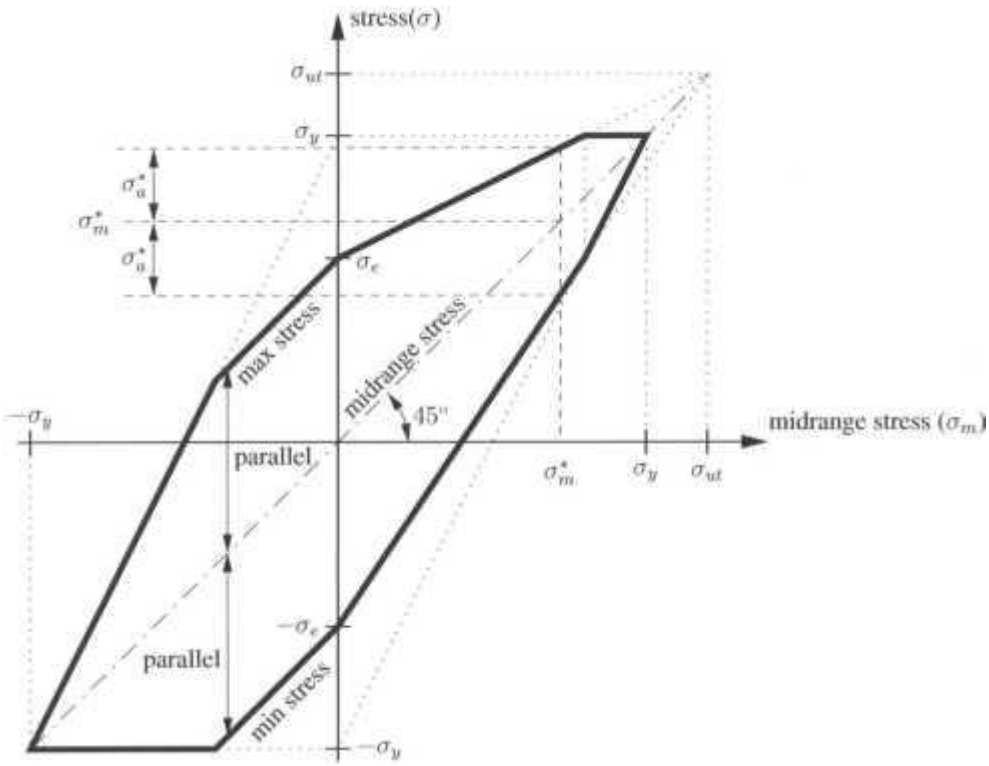
Other proposed fatigue failure envelopes are also presented in Figure 4.7.

Soderberg line:

$$\frac{\sigma_a}{\sigma_e} + \frac{\sigma_m}{\sigma_y} = 1 \tag{4.36}$$

ASME elliptic expression:

$$\left(\frac{\sigma_a}{\sigma_e}\right)^2 + \left(\frac{\sigma_m}{\sigma_y}\right)^2 = 1 \tag{4.37}$$



[billedtekst start]**Figure 4.6:** The modified Goodman diagram, σ_m^* and σ_a^* specifies specific values.[billedtekst slut]

Modified Goodman line (in Germany the Schmidt expression):

$$\frac{\sigma_a}{\sigma_e} + \frac{\sigma_m}{\sigma_{ut}} = 1 \tag{4.38}$$

Gerber parabola:

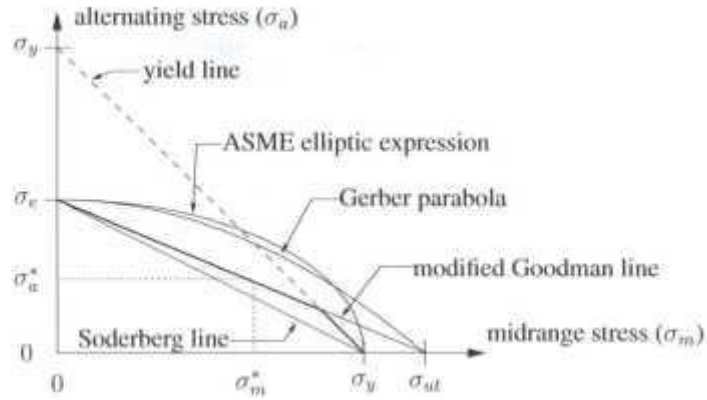
$$\frac{\sigma_a}{\sigma_e} + \left(\frac{\sigma_m}{\sigma_{ut}}\right)^2 = 1 \tag{4.39}$$

Depending on the relation between the mean and alternating stress, different safety factors can be found. For the Goodman assumption we know that the safety factors n_s can be calculated. Under the assumption that it is the Goodman line and not the yield line that is the limiting line, the safety factors are

$$\frac{n_s\sigma_a^*}{\sigma_e} + \frac{\sigma_m^*}{\sigma_{ut}} = 1 \quad \text{assumption of constant mean stress} \tag{4.40}$$

$$\frac{\sigma_a^*}{\sigma_e} + \frac{n_s\sigma_m^*}{\sigma_{ut}} = 1 \quad \text{assumption of constant alternating stress} \tag{4.41}$$

$$\frac{n_s\sigma_a^*}{\sigma_e} + \frac{n_s\sigma_m^*}{\sigma_{ut}} = 1 \quad \text{assumption of constant ratio between mean and alternating stress} \tag{4.42}$$



[billedtekst start]**Figure 4.7:** Different fatigue models. σ_m^* and σ_a^* specifies specific values.[billedtekst slut]

Other possibilities exist and similar equations can be found from the other failure lines.

As for static stresses a reference stress criterion must be applied to be able to compare the combined fatigue stress loading to a uni-axial stress situation. To calculate the mean and alternating stress we have to include the fatigue stress concentration factors. If the fatigue stress concentration factors also should be used on the mean stress is still discussed. Following [8] and [4] the full fatigue stress concentration factors should be put on the mean stress if no yielding occurs. If yielding occurs, we may reduce the concentration factor on the mean stresses, it is however a conservative assumption to apply the full concentration factor. The alternating and mean stress are given by

$$\sigma_a^* = \sqrt{\sigma_{xa}^2 + 3\tau_{xya}^2} = \frac{16}{\pi d^3} B_1 \tag{4.43}$$

$$\sigma_m^* = \sqrt{\sigma_{xm}^2 + 3\tau_{xym}^2} = \frac{16}{\pi d^3} B_2 \tag{4.44}$$

where

$$B_1 = \sqrt{4(K_f M_a)^2 + 3(K_{fs} T_a)^2} \tag{4.45}$$

$$B_2 = \sqrt{4(K_f M_m)^2 + 3(K_{fs} T_m)^2} \tag{4.46}$$

where M_a and M_m is the alternating and mean values of M_{max} (4.13) and T_a and T_m is the alternating and mean part of the torque T. We may now express the safety factor directly in these terms for the Goodman line, under the assumption that the ratio between the bending moment and the torque is constant as defined in (4.42).

$$\frac{1}{n_s} = \frac{32}{\pi d^3} \left(\frac{\sqrt{K_f^2 M_a^2 + \frac{3}{4} K_{fs}^2 T_a^2}}{\sigma_e} + \frac{\sqrt{K_f^2 M_m^2 + \frac{3}{4} K_{fs}^2 T_m^2}}{\sigma_{ut}} \right) \tag{4.47}$$

the same expression for the Soderberg line is found simply by replacing the ultimate stress with

the yield stress ($\sigma_{ut} \rightarrow \sigma_y$).

4.6 Design for shaft deflections

In many applications it is required that the lateral and angular deflections of a shaft should be kept within specified limits to ensure satisfactory operation of the elements located on the shaft.

This is especially true for gears where even small deflections may degrade gear performance and cause noise and vibration.

As a general rule of thumb the stresses in a gear shaft are satisfactorily low if the lateral and angular deflection is within the limits required for good gear operation and performance.

We therefore need to calculate the maximum deflection and rotation of the shaft. In doing this we have to take care of two things:

- The shaft might change diameter → $A(x)$. $I(x)$,$K(x)$
- The shaft is short $L/D < 5$

To overcome the issues it is proposed to use Castigliano's 2nd theorem which states that

$$v_e = \frac{\partial U^c}{\partial P_e} \tag{4.48}$$

i.e., the deflection v_e corresponding to the external load P_e equals the derivative of the complementary energy with respect to the load. This also applies if the load is interchanged with the external moment M_e, then the deflection is the rotation angle θ_e at the point where the moment acts.

$$\theta_e = \frac{\partial U^c}{\partial M_e} \tag{4.49}$$

where θ_e and M_e are corresponding angle and moment.

We therefore need an expression for the complementary energy. This can be found in e.g. [2] for straight beams.

$$U^c = \int_{l_1}^{l_2} \left(\frac{P^2}{2EA} + \beta \frac{V^2}{2GA} + \frac{M^2}{2EI} + \frac{T^2}{2GK} \right) dx \tag{4.50}$$

where the forces are

P = normal force

V = shear force

M = bending moment

T = torsional moment

material data

E = modulus of elasticity

$$G = \text{shear modulus} = \frac{E}{2(1+\nu)}$$

v = Poisson's ratio

cross sectional (geometric) quantities

A = area

I = moment of inertia

K = torsional stiffness factor

β = factor from shear stress

all of the quantities might be a function of the position x. The β factor can be found in books and also in [6] and [2]. The maximum shear stress due to the shear load is

$$\tau_{max} = \mu \frac{V}{A} \tag{4.51}$$

where μ depend on the cross section. The shear stress from the shear force then becomes

$$\tau = \tau_{max} f(y) \tag{4.52}$$

We can now define β as

$$\beta = \frac{\mu^2 \int_A f(y)^2 dA}{A} \tag{4.53}$$

for a circular cross section we find that β = 10/9. The normal force P and the torsional moment T can be neglected in the calculation of the deflection, because these will not contribute, so we may simplify to

$$U^c = \int_{l_1}^{l_2} \left(\beta \frac{V^2}{2GA} + \frac{M^2}{2EI} \right) dx \tag{4.54}$$

For a thorough introduction to energy principles the reader is referred to [6],

4.7 Design for critical shaft speeds

All physical systems that are elastic have eigenfrequencies, i.e., frequencies at which the structure can vibrate without force input (assuming zero damping), and the amplitude of vibration is not limited. These vibrations are termed free vibrations.

In the real application the amplitude of vibration is reduced and free vibration will eventually be removed because of damping. To illustrate a simple ID system a mass spring system is shown in Figure 4.8.

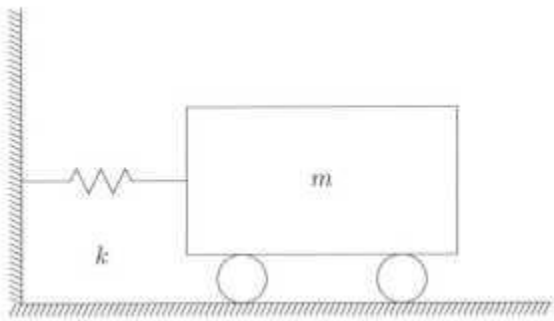
It can be shown that this ID system has one eigenfrequency given by

$$\omega = \sqrt{\frac{k}{m}} \tag{4.55}$$

where k is the stiffness of the spring and m is the mass.

When a system is loaded with external forces at specific frequencies we have the possibility of resonance, when the forcing frequency becomes close to or identical to the eigenfrequency. This shall always be avoided because it leads to large vibrations and resulting stresses.

The main point is that a forcing frequency should be as far away from eigenfrequencies as possible. In continuous structures we do not have a single eigenfrequency, but infinitely many. Eigenfrequencies

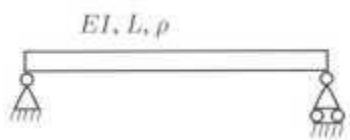


[billedtekst start]**Figure 4.8:** Simple ID mass spring system.[billedtekst slut]

can be calculated using a FE-program, but here a simple analytical estimate is given. The presented method is Dunkerleys method and no background information is given - the reader is referred to books on vibration. The eigenfrequency found is the lowest one relating to transverse vibration.

First we need the lowest eigenfrequencies of the shaft without added hubs. For the simple straight shaft in Figure 4.9 we have

$$\omega_a = \left(\frac{\pi}{L}\right)^2 \sqrt{\frac{EI}{\rho A}} \tag{4.56}$$



[billedtekst start]**Figure 4.9:** Simply supported straight beam (shaft).[billedtekst slut]

Dunkerleys method gives an estimate for the lowest eigenfrequency. The estimate is given by

$$\frac{1}{\omega^2} \approx \frac{1}{\omega_a^2} + \frac{1}{\omega_1^2} + \frac{1}{\omega_2^2} + \frac{1}{\omega_3^2} + \ldots \tag{4.57}$$

where w_i is the eigenfrequency, if we only add hub number i and neglect the mass of the shaft. To find these frequencies the stiffness of the beam at the position of the hub is found and then the frequency is given by this stiffness and the mass of the hub using (4.55).

It is important to note that Dunkerleys method is an estimate and it gives a lower estimate. In real structures not only eigenfrequencies related to transverse vibration exist. A full calculation of the eigenfrequency spectrum of the system might be needed in order to ensure that none of the external loadings have frequencies close to the eigenfrequencies of the system.

4.8 Suggested design procedure, based on shaft yielding

1. Determine the shaft rotational speed, and power or torque to be transmitted by the shaft.
2. Determine the power transmission elements to be mounted on the shaft, their dimensions, and locations.
3. Specify the locations of bearings to support the shaft.
4. Determine the forces and torques exerted on the shaft.
5. Determine the reactions required at the bearings to support the shaft.
6. Produce Shear Force Diagrams, Bending Moment Diagrams, and Torque Diagrams for the shaft.
7. Identify critical points on the shaft, experiencing the largest stresses (due to bending, torsion, and stress concentrations).
8. Determine the required diameter at different sections of the shaft, based upon stresses at the critical points.

4.9 Nomenclature

a	mm	Neuber's constant
d	mm	Shaft diameter
k	N/mm	Stiffness
m	kg	Mass
n	–	Number of cycles
n_s	–	Safety factor
{s}	mm	General geometric vector in 3D
q	–	Notch sensitivity
t	°C	Temperature
v_e	mm	Deflection due to external load
A	mm^2	Cross sectional area
c	–	Correction constants for endurance limit

E	N/mm^2	Modulus of elasticity
$\{F\}$	N	General force vector in 3D
G	N/mm^2	Shear modulus
I	m^4	Cross sectional moment of inertia
I_u	mm^4	Cross sectional moment of inertia
I_z	mm^4	Cross sectional moment of inertia
I_p	mm^4	Cross sectional polar moment of inertia
K	mm^4	Cross sectional torsional stiffness factor
K_t	–	Theoretical stress concentration factor
K_{ts}	–	Theoretical shear stress concentration factor
K_f	–	fatigue stress concentration factor
K_{fs}	–	fatigue shear stress concentration factor
L	mm	Length of beam
M	Nmm	Bending moment
M_a	Nmm	Amplitude bending moment
M_e	Nmm	External bending moment
M_m	Nmm	Mean bending moment
$\{M\}$	Nmm	General moment vector in 3D

p	N	Axial load
P_e	N	External load
T	Nmm	Torque (torsional moment)
T_a	Nmm	Amplitude torque (torsional moment)
T_m	Nmm	Mean torque (torsional moment)
V	N	Shear force
U^c	Nmm	Complementary energy
Γ	–	Slenderness ratio
ϵ	–	Strain
w	rad/s	Eigenfrequency
ρ	kg/mm^3	Density
ρ_u	mm	Radius of curvature
ρ_z	mm	Radius of curvature
σ_0	N/mm^2	Amplitude stress
σ_e	N/mm^2	Endurance limit
σ'_e	N/mm^2	Uncorrected endurance limit
σ_m	N/mm^2	Mean stress
σ_{nom}	N/mm^2	Nominal stress
Σ_{REF}	N/mm^2	von Mises reference stress
τ_{nom}	N/mm^2	Yield stress
Σ_{UT}	N/mm^2	Ultimate tensile stress
τ_{nom}	N/mm^2	Nominal shear stress

θ_e	rad	Rotation due to external load

4.10 References

[1] B. J. Hamrock, B. Jacobson, and S. R. Schmid. *Fundamentals of machine elements.* McGraw Hill, 1999.

[2] G. Hedner. *Formelsamling i Hallfasthetslara.* KTH, Stockholm, 1986.

[3] P. E. Nikravesh. *Computer-Aided Analysis of Mechanical Systems.* Prentice-Hall International, Inc., Englewood Cliffs, Nj 07632, 1988.

[4] R. L. Norton. *Machine design, an integrated approach, fifth edition.* Prentice-Hall Inc., Upper Saddle River, N.J. 07458, 2014.

[5] N. L. Pedersen. *Analysis and synthesis of complex mechanical systems.* Solid Mechanics, DTU, 1998. Ph.D. thesis.

[6] *P. Pedersen. Elasticity - Anisotropy - Laminates with Matrix Formulation, Finite Element and Index to Matrices. Solid Mechanics, DTU, 1998.*

[7] R. E. Peterson. *Stress concentration design factors.* John Wiley & Son, Inc., New York, USA, 1953.

[8] J. E. Shigley and C. R. Michke. *Mechanical Engineering Design 7th ed.* McGraw Hill, Singapore, 2004.

Chapter 5 Shaft-hub Connections

5.1 Introduction

Power transmission elements such as gears, pulleys and sprockets should be mounted on shafts securely locating them in a circumferential as well as an axial position, and making them able to transmit torque between the transmission element and the shaft.

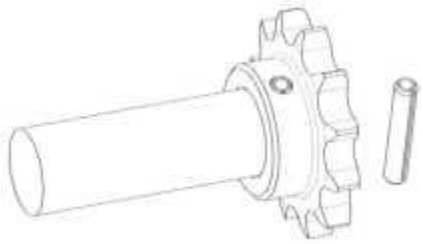
The various types of shaft-hub connections have different characteristics such as

- different ability to transfer torque
- difference in easiness of assembly (and/or disassembly)
- different level of running accuracy

Because of difference in the way of functioning we distinguish between shaft-hub connections based on the principle of "positive connections", where the power is transmitted through shape or geometry of the elements, and shaft-hub connections based on "transmission by friction" where the transmitted power depends on the frictional forces between mating surfaces.

5.2 Positive connections

5.2.1 Pinned and taper-pinned joints

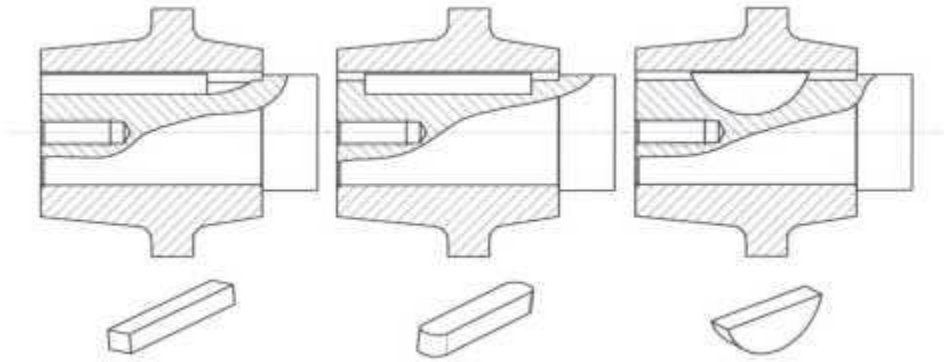


[billedtekst start]**Figure 5.1**: An often used pin connection is a connection with a split tubular spring pin (Available in light, medium and heavy duty series).[billedtekst slut]

The purpose of the pin is to prevent rotational motion between shaft and the element mounted when a torque is transmitted. Pins are suitable for transmitting low to medium torque. They are cheap and easy to assemble and disassemble. Pins are common standardized elements. There are a great variety of pins.

In some designs where the pin is going through the shaft, the pin is seen as a safety component. If a dramatic increase in the torque should happen, the pin will shear before other more expensive machine elements fail. It is then of great importance to choose a pin made of significantly weaker material than the hub and shaft, often from mild steel.

5.2.2 Parallel keys and Woodruff Keys



[billedtekst start]**Figure 5.2:** Two types of ordinary parallel keys and a Woodruff key.[billedtekst slut]

Keys are only suitable for torsional loads in one direction. The dimension criterion for keys is usually the pressure load on the key sides. Although the pressure distribution along the key is very uneven, the normal calculation method is based on even pressure load distribution as indicated previously. In fact there are at least two parameters to have in mind. Firstly, because of the torsional deflection of the shaft the pressure load is highest at the end of the key, where the torsional loads comes in. In practice it is useless to design keys with carrying length of more than 1.5 - 2 times the shaft diameter. Secondly, the pressure varies over the key height. Applying a pressure distribution that will prevent the key from tilting is difficult to sketch. The allowable pressure value is dependent of the material combination of shaft and hub. In [5] the following values are seen: Cast iron hub: $p_{allowable} \leq 50 N/mm^2$ for L_{eff} = 1.6d to 2.1 d. Steel hub: $p_{allowabie} \leq 90 N/mm^2$ for L_{eff} = 1.1 d to 1.4d. In special cases $p = 200 N/mm^2$ is allowable for infrequent high extra loads. The key itself is normally made of "key-steel" which is a medium strength steel (C 45 K).

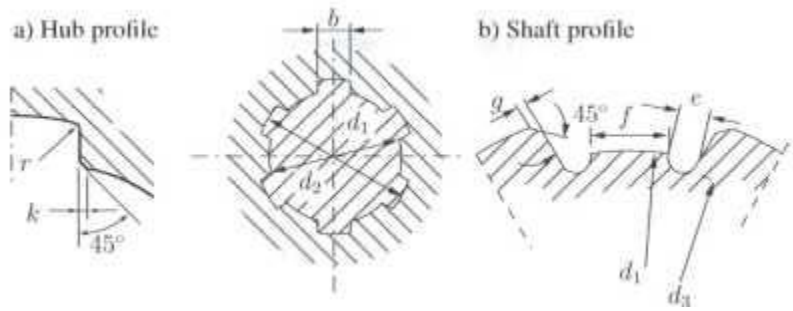
5.2.3 Splined joints

To lower the pressure load on the key sides it is possible to use more than one key on the circumference, but in practice it is very difficult to obtain even load share between the keys.

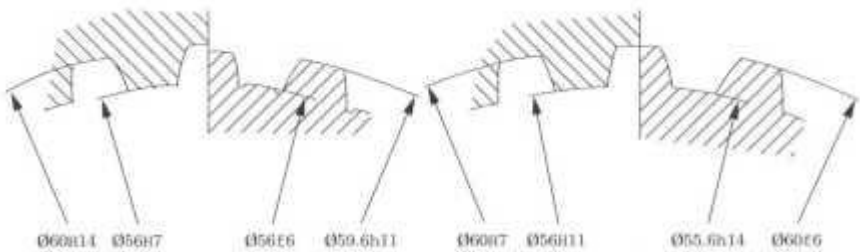
Instead multiple "keys" can be cut directly in the shaft and are then called splines. The key ways in the hub are then manufactured by reaming. See Figures 5.3 and 5.4.

5.2.4 Prestressed shaft-hub connections

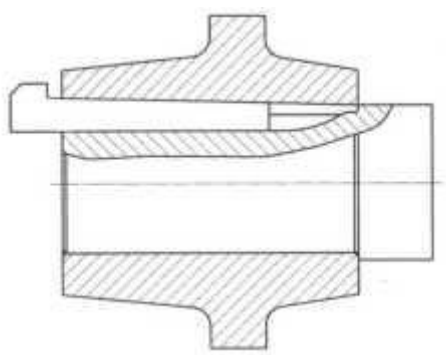
When dynamic loads are applied to the shaft-hub connection and especially dynamic loads in both directions, it is necessary to use prestressed connection. See Figure 5.5.



[billedtekst start]**Figure 5.3:** Straight-sided splines are available in light, medium and heavy series.[billedtekst slut]



[billedtekst start]**Figure 5.4:** Involute splines. Centered on bottom (left), top (right) or flank.[billedtekst slut]



[billedtekst start]**Figure 5.5:** Prestressed connection with Gib-head key. Normally not to be recommended.[billedtekst slut]

5.2.5 Failure of positive connections

The positive connection can all in general fail in two modes

- Shear failure
- Bearing failure

Shear failure is when e.g. a key is loaded such that the key is sheared across its width

$$\tau = \frac{F}{A_{shear}} \tag{5.1}$$

where F is the applied load and A_{sflcar} is the shear area being cut. For a key this is equal to the length times the width. The force on the key can be found from the quotient of the shaft torque and shaft radius. If the shaft torque is constant with time the force will also be and the safety factor can be found by comparing the shear force to the yield stress

$$n_s = \frac{\sigma_y}{\sqrt{3}\tau} \tag{5.2}$$

If the torque is varying with time, then a fatigue failure of the key is possible. In this case we have to use methods as shown in Chapter 4 e.g. the Goodman diagram.

The second mode of failure is bearing failure. We may define an average bearing stress by

$$\sigma = \frac{F}{A_{bearing}} \tag{5.3}$$

where F is the applied load and $A_{bearing}$ is the bearing area, i.e., e.g. the area of contact between the key side and the shaft or the hub. For a square key this equals half the height of the key times the length. A Woodruff key has a different bearing area in the hub than in the shaft. The hub's Woodruff bearing area is smaller and will typically fail first.

The safety factor is in these cases given by

$$n_s = \frac{\sigma_y}{\sigma} \tag{5.4}$$

If one key is insufficient to carry the load splined joints can be used. It is difficult to estimate the number of "teeth" that is in contact and howhard they shear the load from the torque. A usual assumption is that only 75%of the "teeth" are in contact and then calculate the strength of the connection using this assumption.

5.3 Connection with force (Transmission by friction)

5.3.1 Cone interference fit

In a conical shaft hub connection the pressure is established by pressing the conical shaft geometry into the conical hole in the hub. When assembled, it almost behaves as a press or shrink fit (see later in this chapter). The force required can in some cases be obtained by the bolt and nut connection as seen on Figure 5.6. Very often the necessary axial force has to be applied by a hydraulic piston. Afterwards the conical parts can be held in place by a nut.

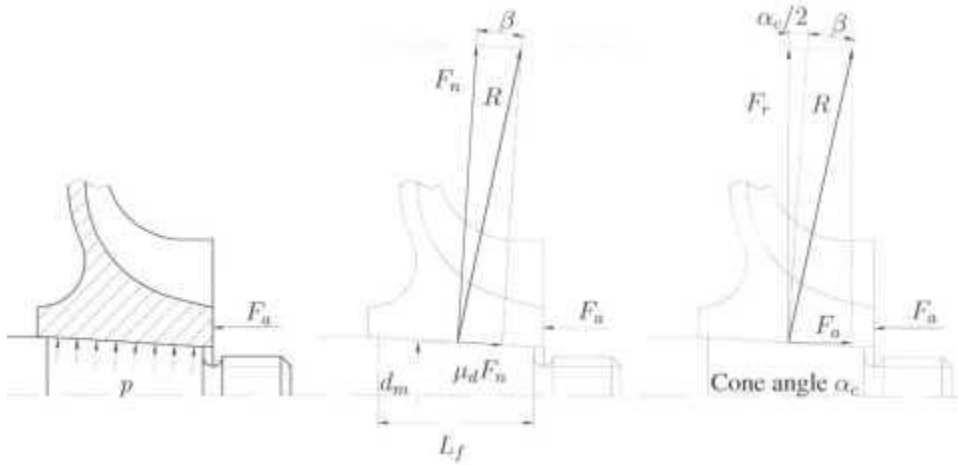
The friction force is

$$F_f \leqq \leq \mu_s F_n \qquad \text{if fixed}$$

$$F_f = \mu_d F_n \qquad \text{if sliding}$$

where F_n is the normal load, μ_s is the static coefficient of friction and μ_d is the dynamic coefficient of friction. The surface area of the frustum of contact is

$$A_{frustum} = \pi \frac{d_1 + d_2}{2} L = \pi d_m L \tag{5.5}$$



[billedtekst start]**Figure 5.6**: A conical shaft-hub connection.[billedtekst slut]

where d_1 and d_2 are the two diameters of the conical shaft at the start and end of contact with the hub, L is the length of contact and d_m is the mean value of d_1 and d_2. By force equilibrium (assuming sliding) we have

$$F_a - p\pi d_m L \sin\frac{\alpha_c}{2} - \mu_d p\pi d_m L \cos\frac{\alpha_c}{2} = 0$$

from which the force F_a can be found

$$F_a = \pi d_m L\, p(\sin\frac{\alpha_c}{2} + \mu_d \cos\frac{\alpha_c}{2}) \Rightarrow$$
$$F_a = \pi d_m L_f\, p(\tan\frac{\alpha_c}{2} + \mu_d) \tag{5.6}$$

where the length L_f is defined in Figure 5.6. If the angle α_c is small then $\cos(\alpha_c/2) \approx$ and we may use the approximation that $L = L_f$ which leads to

$$F_a \approx \pi d_m L_f\, p(\sin\frac{\alpha_c}{2} + \mu_d \cos\frac{\alpha_c}{2}) \tag{5.7}$$

this approximation can be found in many books. In (5.6) and (5.7) we used the assumption of sliding and therefore μ_d if no sliding then μ_d should be interchanged with μ_s and then (5.6) and (5.7) gives the limiting load at which sliding begins.

The torque limit is

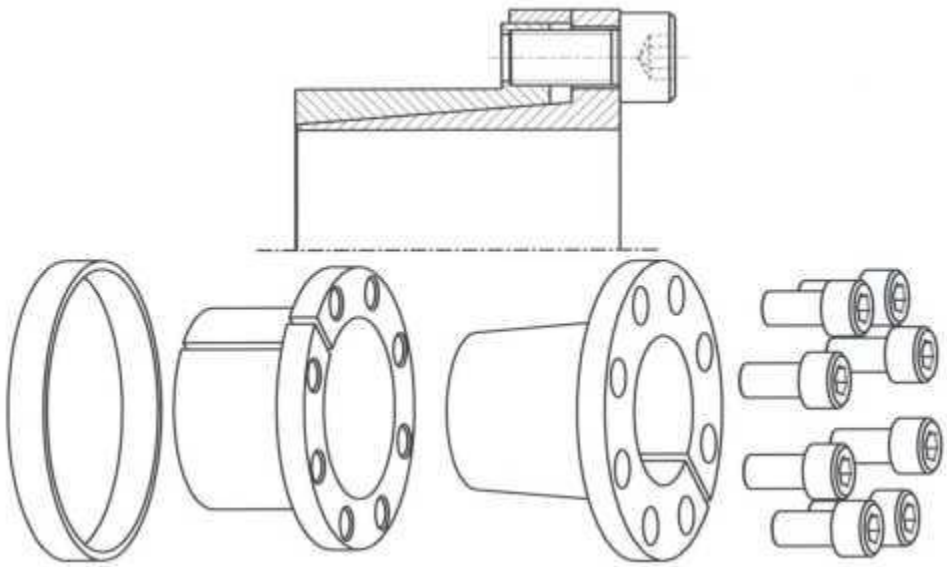
$$T = \mu_s \pi \frac{L_f}{\cos\frac{\alpha_c}{2}} \cdot \frac{d_2^2 + d_1^2 + d_1 d_2}{6} p \approx \mu_s \pi L_f \frac{d_m^2}{2} p \tag{5.8}$$

for the approximation the assumption $d_1 \approx d_2 \approx d_m$ is used.

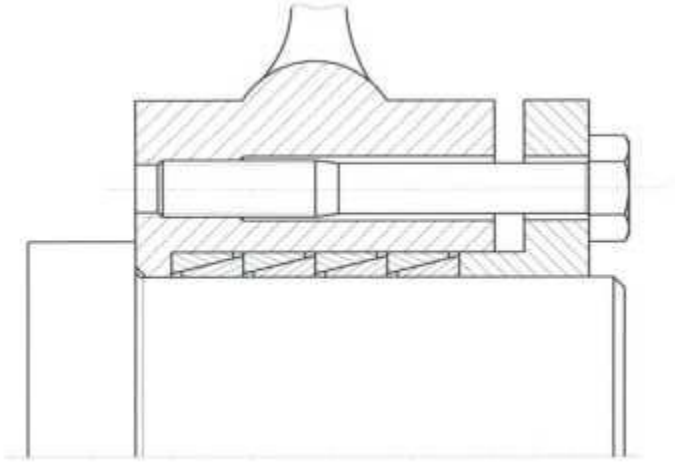
5.3.2 Interference fit with spacers

When easy assembly and disassembly have a high priority combined with medium torque transmission capabilities special spacers have been developed. One of these is the ringfeder-

element as shown in Figure 5.8.



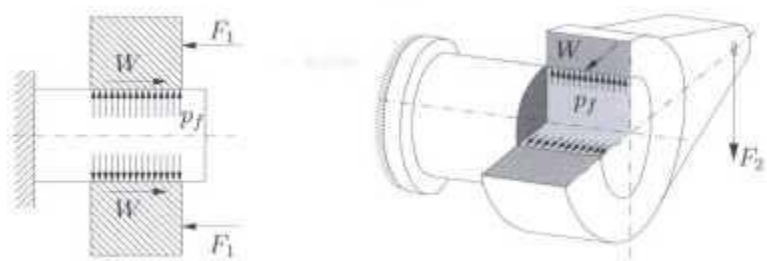
[billedtekst start]**Figure 5.7:** An example of cone clamping elements with radially cut through tapered rings. Self-centering. Ringspann Gmbh. KTR Kupplungstechnik Gmbh.[billedtekst slut]



[billedtekst start]**Figure 5.8:** Shaft hub connection with tapered straining rings. Not self-centering. Ringspann Gmbh.[billedtekst slut]

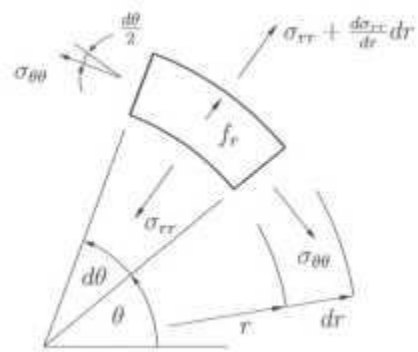
5.3.3 Interference fit (press and shrink fits)

In a press or shrink fit the pressure between shaft and hub is caused by interference. The pressure increases the diameter of the hole in the hub and reduces the diameter of the shaft. Chapter 1 describes the tolerances on shaft and hub that result in specific fits (intervals). Figure 5.9 left shows an axially loaded shrink fit. The right side picture shows a normal torsionally loaded shaft-hub connection. In the following we will describe the stresses and strains that appear due to the fit.



[billedtekst start]**Figure 5.9:** A friction based connection. *(W* = friction force).[billedtekst slut]

Because of symmetry it is practical to use cylindrical coordinates (r, *θ, z),* where *z* is in the axial direction. Assuming axial symmetry, i.e., symmetric with respect to the *z*-axis the force equilibrium becomes independent of *θ.*



[billedtekst start]**Figure 5.10:** Free body diagram of infinitesimal piece of shaft or hub in cylindrical coordinates. Assuming rotational symmetry.[billedtekst slut]

The cut-out in Figure 5.10 shows the stresses under the axial symmetry assumption, which also allow for radial body force f_r. Force equilibrium in the radial direction gives

$$(\sigma_{rr} + \frac{d\sigma_{rr}}{dr}dr)d\theta(r + dr)dz - \sigma_{rr}d\theta rdz - 2\sigma_{\theta\theta}\sin\frac{d\theta}{2}drdz + f_r d\theta rdrdz = 0 \qquad (5.9)$$

removing higher order terms we have

$$\frac{d\sigma_{rr}}{dr} + \frac{\sigma_{rr} - \sigma_{\theta\theta}}{r} + f_r = 0 \qquad (5.10)$$

Because of symmetric with respect to the *z*-axis the corresponding strains are (can be found in e.g. [6])

$$\epsilon_{rr} = \frac{dv_r}{dr} \qquad (5.11)$$

$$\epsilon_{\theta\theta} = \frac{v_r}{r} \qquad (5.12)$$

$$\epsilon_{r\theta} = 0 \qquad (5.13)$$

where v_r is the displacement in the radial direction.

On the sides of the hub we know that there is plane stress and in the center of the hub we must have plane strain due to symmetry. In the following it is assumed that the thickness of the hub is such that we may assume plane stress throughout the hub, i.e.,

$$\sigma_{ZZ} = 0 \tag{5.14}$$

The constitutive equations (see [6]) under the assumption of plane stress are

$$\epsilon_{rr} = \frac{1}{E}(\sigma_{rr} - \nu\sigma_{\theta\theta}) \tag{5.15}$$

$$\epsilon_{\theta\theta} = \frac{1}{E}(\sigma_{\theta\theta} - \nu\sigma_{rr}) \tag{5.16}$$

$$\epsilon_{zz} = -\frac{\nu}{E}(\sigma_{rr} + \sigma_{\theta\theta}) \tag{5.17}$$

$$\epsilon_{r\theta} = \epsilon_{rz} = \epsilon_{\theta z} = 0 \tag{5.18}$$

$$\sigma_{rr} = \frac{E}{1-\nu^2}(\epsilon_{rr} + \nu\epsilon_{\theta\theta}) \tag{5.19}$$

$$\sigma_{\theta\theta} = \frac{E}{1-\nu^2}(\epsilon_{\theta\theta} + \nu\epsilon_{rr}) \tag{5.20}$$

$$\sigma_{zz} = \sigma_{r\theta} = \sigma_{rz} = \sigma_{\theta z} = 0 \tag{5.21}$$

where E is the modulus of elasticity and v is Poisson's ratio. Using (5.11) and (5.12) in (5.19) and (5.20) yields the differential equations

$$\sigma_{rr} = \frac{E}{1-\nu^2}\left(\frac{dv_r}{dr} + \nu\frac{v_r}{r}\right) \tag{5.22}$$

$$\sigma_{\theta\theta} = \frac{E}{1-\nu^2}\left(\frac{v_r}{r} + \nu\frac{dv_r}{dr}\right) \tag{5.23}$$

Using (5.22) and (5.23) in (5.10) yields (after some manipulations)

$$r\frac{d^2v_r}{dr^2} + \frac{dv_r}{dr} - \frac{v_r}{r} = -\frac{1-\nu^2}{E}f_r r \tag{5.24}$$

The differential equation (5.24) is Euler's differential equation. If we assume that the radial body force is given by

$$f_r = \rho r w^2 \tag{5.25}$$

where ϱ is the density and ω is angular speed, then the solution to (5.24) is

$$v_r = c_1 r + c_2\frac{1}{r} - \frac{1-\nu^2}{8E}f_r r^2 \tag{5.26}$$

where c_1 and c_2 are constants that are to be determined from the boundary conditions. Using (5.26) we may express the stress and strain as

$$\sigma_{rr} = c_3 - c_4\frac{1}{r^2} - \frac{3+\nu}{8}f_r r \tag{5.27}$$

$$\sigma_{\theta\theta} = c_3 + c_4\frac{1}{r^2} - \frac{1+3\nu}{8}f_r r \tag{5.28}$$

$$\epsilon_{rr} = \frac{1}{E}\left((1-\nu)c_3 - (1+\nu)c_4\frac{1}{r^2} - \frac{3(1-\nu^2)}{8}f_r r\right) \tag{5.29}$$

$$\epsilon_{\theta\theta} = \frac{1}{E}\left((1-\nu)c_3 + (1+\nu)c_4\frac{1}{r^2} - \frac{1-\nu^2}{8}f_r r\right) \tag{5.30}$$

$$\epsilon_{zz} = -\frac{\nu}{E}\left(2c_3 - \frac{1+\nu}{2}f_r r\right) \tag{5.31}$$

where the new constants $c3$ and c_4 are related to c_1 and c_2 through

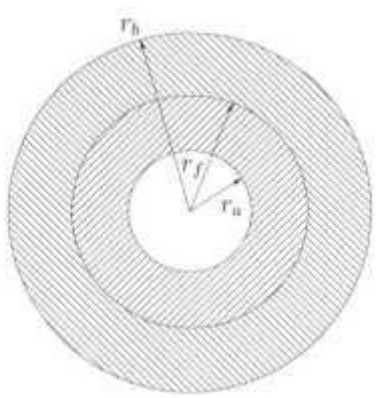
$$c_3 = \frac{E}{1-\nu}c_1 \tag{5.32}$$

$$c_4 = \frac{E}{1+\nu}c_2 \tag{5.33}$$

Expressing the radial displacement in c_3 and c_4 gives

$$u_r = c_3\frac{1-\nu}{E}r + c_4\frac{1+\nu}{E}\frac{1}{r} - \frac{1-\nu^2}{8E}f_r r^2 \tag{5.34}$$

We now have all the needed equations ready to express the relation between the interference and the corresponding pressure in an interference or press fit. First we make the assumption that $f_r = 0$ and let the shaft be a tube, as shown in Figure 5.11. In the figure the inner radius r_a is shown (in case of a solid shaft we use $r_a = 0$), r_b, is the outer radius of the hub and r_f is the fit radius.



[billedtekst start]**Figure 5.11:** Definition of inner, outer and fit radii.[billedtekst slut]

Outer part (hub) First we look at the hub. The boundary condition is that the pressure is zero at the outer radius of the hub and at the fit radius we have the fit pressure p_f. This gives

$$\sigma_{rr}(r = r_b) = c_3 - c_4 \frac{1}{r_b^2} = 0 \tag{5.35}$$

$$\sigma_{rr}(r = r_f) = c_3 - c_4 \frac{1}{r_f^2} = -p_f \tag{5.36}$$

that yields

$$c_3 = -\frac{r_f^2}{r_f^2 - r_b^2} p_f \tag{5.37}$$

$$c_4 = -\frac{r_f^2 r_b^2}{r_f^2 - r_b^2} p_f \tag{5.38}$$

The radial deformation at the fit radius of the hub is

$$v_{r_h}(r = r_f) = \frac{1}{E_h}\left(c_3(1 - \nu_h) r_f + c_4(1 + \nu_h)\frac{1}{r_f}\right) \tag{5.39}$$

By inserting the values of c_3 and c_4

$$v_{r_h}(r = r_f) = \frac{r_f p_f}{E_h}\left(\frac{r_b^2 + r_f^2}{r_b^2 - r_f^2} + \nu_h\right) \tag{5.40}$$

Inner part (shaft) The boundary condition is that the pressure is zero at the inner radius of the shaft and at the fit radius we have the fit pressure p_f. This gives

$$\sigma_{rr}(r = r_a) = c_3 - c_4 \frac{1}{r_a^2} = 0 \tag{5.41}$$

$$\sigma_{rr}(r = r_f) = c_3 - c_4 \frac{1}{r_f^2} = -p_f \tag{5.42}$$

$$\tag{5.43}$$

that yields

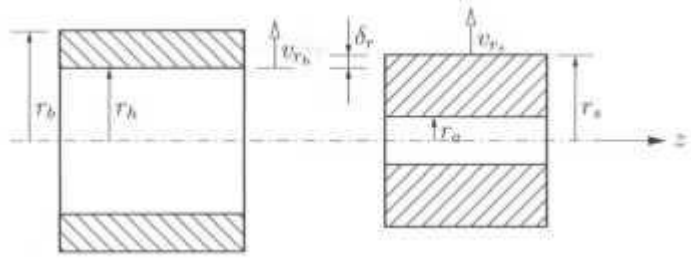
$$c_3 = -\frac{r_f^2}{r_f^2 - r_a^2} p_f \tag{5.44}$$

$$c_4 = -\frac{r_f^2 r_a^2}{r_f^2 - r_a^2} p_f \tag{5.45}$$

The radial deformation at the fit radius of the shaft is

$$v_{r_s}(r = r_f) = \frac{1}{E_s}\left(c_3(1 - \nu_s) r_f + c_4(1 + \nu_s)\frac{1}{r_f}\right) \tag{5.46}$$

By inserting the values of c_3 and c_4



[billedtekst start]**Figure 5.12:** Cut through shaft and hub before assembly.[billedtekst slut]

$$v_{r_s}(r = r_f) = \frac{r_f p_f}{E_s}\left(\frac{r_a^2 + r_f^2}{r_a^2 - r_f^2} + \nu_s\right) \tag{5.47}$$

In Figure 5.12 we show a cut through the shaft and the hub before assembly.

The radius before assembly of the inner radius of the hub is r_h and the outer radius of the shaft before assembly is r_s. After assembly it is known that both of these radii must be displaced to the fit radius r_f, this yields

$$r_h + v_{rh} = r_f \; r_s + v_{rs} = r_f \Rightarrow$$

$$r_s - r_h = v_{rh} - v_{rs} \tag{5.48}$$

The radial and diametral fit is defined as

$$\delta_r = r_s - r_h = v_{rh} - v_{rs} \tag{5.49}$$

$$\delta_d = 2\delta_r \tag{5.50}$$

the diametral fit corresponds to the fit specified in Chapter 1 and can be given as

$$\delta_d = 2(v_{r_h} - v_{r_s}) = 2r_f p_f \left(\frac{r_b^2 + r_f^2}{E_h(r_b^2 - r_f^2)} + \frac{\nu_h}{E_h} - \frac{r_a^2 + r_f^2}{E_s(r_a^2 - r_f^2)} - \frac{\nu_s}{E_s}\right) \tag{5.51}$$

The radial volume force is neglected in the calculations that result in (5.51). The reason for including f_r is the possibility to include radial forces due to the rotational speed. In this case we can use the radial force as defined in (5.25). In many cases with low angular speed the influence from this term can be neglected. For completeness the diametral fit including the influence from the rotational speed is given. The calculations are identical to the ones that lead to (5.51).

$$\begin{aligned}\delta_d = {} & 2r_f p_f \left(\frac{r_b^2 + r_f^2}{E_h(r_b^2 - r_f^2)} + \frac{\nu_h}{E_h} - \frac{r_a^2 + r_f^2}{E_s(r_a^2 - r_f^2)} - \frac{\nu_s}{E_s}\right) \\ & + \frac{1}{2} r_f \omega^2 \left(\frac{\rho_h}{E_h}\left((3 + \nu_h)r_b^2 + (1 - \nu_h)r_f^2\right) - \frac{\rho_s}{E_s}\left((3 + \nu_s)r_a^2 + (1 - \nu_s)r_f^2\right)\right)\end{aligned} \tag{5.52}$$

it is noted that in the case of $w = 0$ then (5.52) becomes identical to (5.51). It should also be noted that the rotational speed will reduce the pressure in the fit, i.e., if a given fit pressure is needed, then we must for $w \neq 0$ specify a higher fit δ_d to achieve this.

Assembly force and transmitted torque

For a press fit the assembly force, F_a, depends on the thickness of the outer member, the difference in diameters of the shaft and hub and of the coefficient of friction. The maximum shear stress is

$$\tau_{\max} = \mu_s p_f = \frac{F_{a,\max}}{A} = \frac{F_{a,\max}}{\pi d_f l} \tag{5.53}$$

where μ_s is the static coefficient of friction. The maximal torque that can be transmitted is

$$T_{\max} = F_{a,\max}\frac{d_f}{2} = \frac{\mu_s \pi\, d_f^2\, l\, p_f}{2} \tag{5.54}$$

where l is the length of the contact zone between the shaft and hub.

Shrink fits

To assemble shaft and hub with a shrink fit a temperature difference between shaft and hub must be established. The easiest way of doing this is by heating the hub. However, it is important to ensure that a temperature increase of the hub will not harm its heat treatment. Gears are especially sensitive to this.

Assuming that there is a linear relationship between thermal strain and temperature, the following expression applies

$$\Delta t_m = \frac{\delta_d}{\alpha d_f} \tag{5.55}$$

Where α is the thermal expansion coefficient. The temperature difference Δt_m may of course be established also by cooling the shaft or by a combination of heating the hub and cooling the shaft.

Applying the temperature difference calculated from (5.55) will exactly allow for assembly. A small decrease in this temperature difference (e.g. by the two parts touching each other) will however make assembly impossible. It is therefore typical to add a handling deflection to the fit. The size of this handling deflection is typical

$$\delta_d^h = 0.001 d_f \tag{5.56}$$

Smoothing out of surfaces

Especially when assembling shaft and hub by pressing it is important to take the associated smoothing out of surfaces into account. Due to the sliding of the surfaces on each other, the surfaces smoothen out and the interference is reduced. The reduction is based on experience and for press fits it amounts to

$$\delta_d^s \approx 8(\mathrm{Ra}_{shaft} + \mathrm{Ra}_{hub}) \tag{5.57}$$

where Ra is the arithmetic mean roughness value (DIN 4768, DIN 4662, ISO 4287/1). For shrink fits the reduction is smaller and can be estimated to

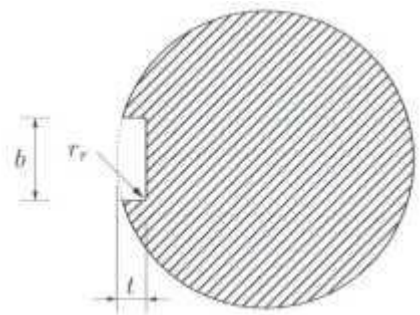
$$\delta_d^* \approx 4(\mathrm{Ra}_{shaft} + \mathrm{Ra}_{hub}) \tag{5.58}$$

5.4 Design modification/optimization

The design of many shaft hub connections, e.g. key and spline designs, is defined in different standard. The reason being that then component form different manufacturing companies can be used together. The designs are often restricted to straight lines and the circular shape. From shape optimization with

respect to minimizing stress concentrations it is known that the circular shape is seldom the optimal choice and therefore there is a potential for improvement. The different standards regulating the designs are based on many years of experience, so in most cases the designs are not very far from an optimal one still although based on the circular shape.

Keyway design



[billedtekst start]**Figure 5.13:** Cross section of prismatic part of parallel keyway. The relative dimensions corresponds to a 100mm shaft diameter according to [2], (t = 10mm, b = 28mm, 0.4mm ≦ r_r ≦ 0.6mm).[billedtekst slut]

For the keyway connection one of the commonly used standards is DIN 6885, [2], In 5.13 a cross section of the prismatic part of a keyway is shown. The standard specifies a range for the fillet radius in the key way. In Figure 5.14 the theoretical shear stress concentration factor, K_{ts} defines as

$$K_{ts} = \frac{\tau_{max}}{\tau_{nom}} \tag{5.59}$$

is shown for the prismatic part of a keyway in pure torsion. The results are taken from [7].

In [7] a curve fitted result for the stress concentration factor is also given.

$$K_{ts} = (1.4786\frac{t}{d} + 0.6326)(\frac{r_r}{d})^{(0.869(\frac{t}{d})^2 - 0.4392\frac{t}{d} - 0.2369)}, \qquad d \in [6:38]\text{mm} \tag{5.60}$$

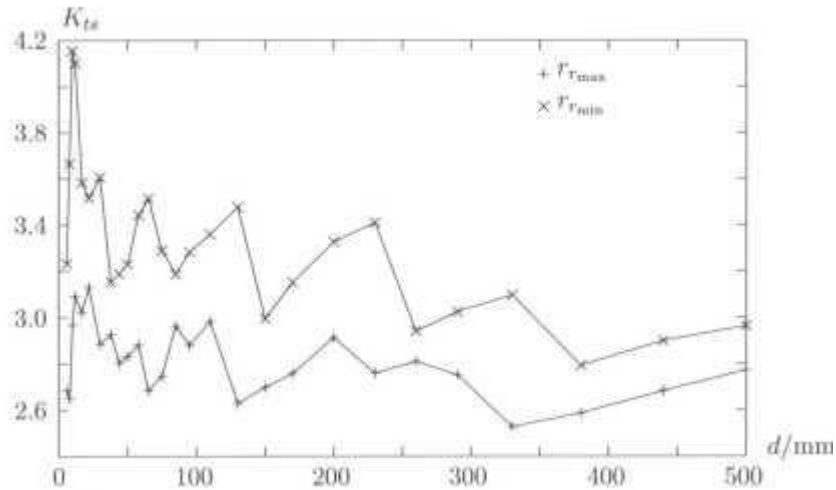
$$K_{ts} = (1.0428\frac{t}{d} + 0.5355)(\frac{r_r}{d})^{(2.8074(\frac{t}{d})^2 - 0.8091\frac{t}{d} - 2476)}, \qquad d \in [38:500]\text{mm} \tag{5.61}$$

where d is the shaft diameter, t the depth of the keyway, see Figure 5.15, and r_r is the fillet radius. With the two curve fits an easy stress concentration factor estimation for the keyways prismatic part in pure torsion for designs that follows DIN 6885 are given. In the case of a specific design that does not follow the standard DIN 6885 the full numerical simulation specified in e.g. [7] is needed.

An improved design is suggested in [7]. The parametrization chosen here is to use the super ellipse due to the simple parametrization and due to previous results obtained with this shape in relation to stress concentrations. The design domain is shown in Figure 5.15, where the

elliptical shape can be seen for the fillet.

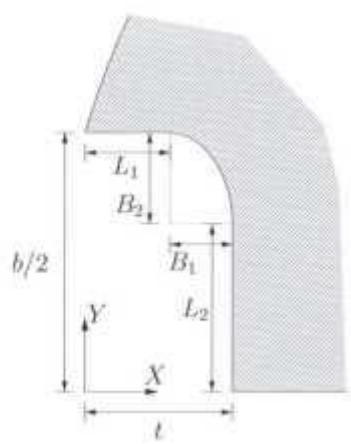
The super ellipse (with super elliptical power η) is in parametric form given by



[billedtekst start]**Figure 5.14:** The stress concentration factor for the prismatic part of a keyway in pure torsion as a function of the diameter. The different design variables are controlled by DIN 6885. The point corresponds to numerical simulations, these are connected by straight lines.[7][billedtekst slut]

$$X = L_1 + B_1 \cos(\theta)^{(2/\eta)}, \qquad \theta \in [0 : \tfrac{\pi}{2}] \qquad (5.62)$$

$$Y = L_2 + B_2 \sin(\theta)^{(2/\eta)}, \qquad \theta \in [0 : \tfrac{\pi}{2}] \qquad (5.63)$$



[billedtekst start]**Figure 5.15:** The design domain: half a keyway where the fillet is a super ellipse with semi-major axes B_1 and B_2.[billedtekst slut]

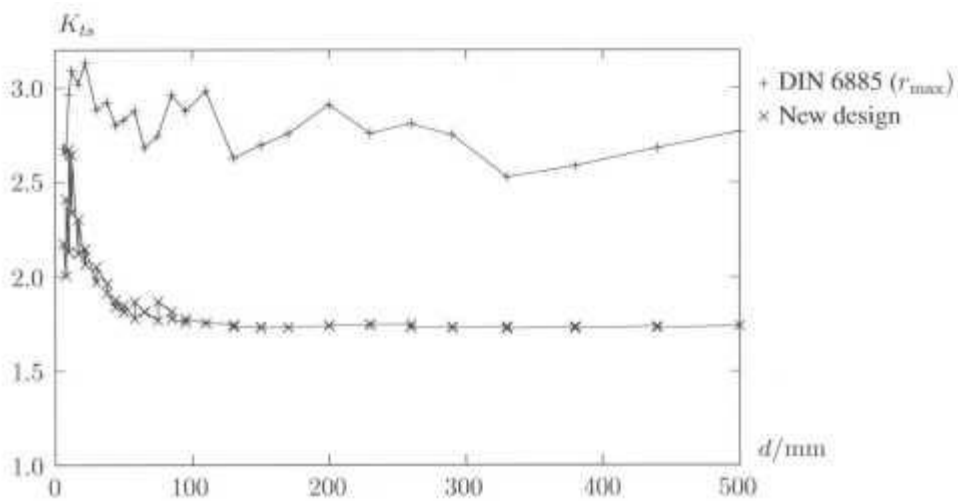
The new design proposal is

- L_1 =minimum allowable shoulder length according to DIN 6885

- $L_2 = 0$
- η=2
- b DIN 6885 standard
- t = 1.4L_1 for 6mm ≦ d ≦ 38mm and t = 1.5L_1 for 38mm ≦ d ≦ 500mm

The keyway design is according to Figure 5.15 fully controlled by five design parameters, width b, depth t, length L_1 and $L_2$and super ellipse power η.

The K_{ts} factors for the diameter range 6mm ≦ d ≦ 500mm is shown in Figure 5.16 together with the minimum obtainable stress concentration using the standard. It can be seen that for most of the diameter range the K_{ts} factor is almost constant. The design is best for the larger diameter range but always better than that given by the standard. The smallest difference is achieved for d = 8mm where the minimum stress concentration specified by DIN 6885 is K_{ts} = 2.65 where the new keyway design has K_{ts} = 2.41, i.e. a 9% reduction in the stress. For most of the diameter range the improvement is much larger with a reduction in the maximum stress of about 35%.



[billedtekst start]Figure 5.16: The stress concentration factor for the prismatic part of a key way in pure torsion as a function of the diameter. The top curve is for the DIN 6885 design with the maximum allowable fillet radius. The bottom curve is for the new proposed keyway design. The point corresponds to numerical simulations, these are connected by straight lines. [7][billedtekst slut]

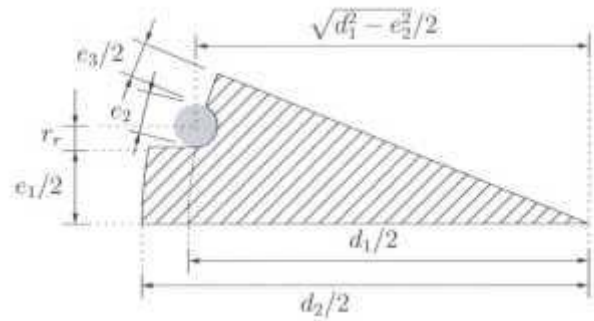
5.4.1 Spline design

Splines in shafts are found in many different types and designs. The possibilities for design modifications are here described relative to straight sided splines as it is found in [8],

The design of straight-sided splines is regulated by standards. Here we refer to the international ISO standard (DIN ISO 14) or/and the German DIN 5464 norm. According to these

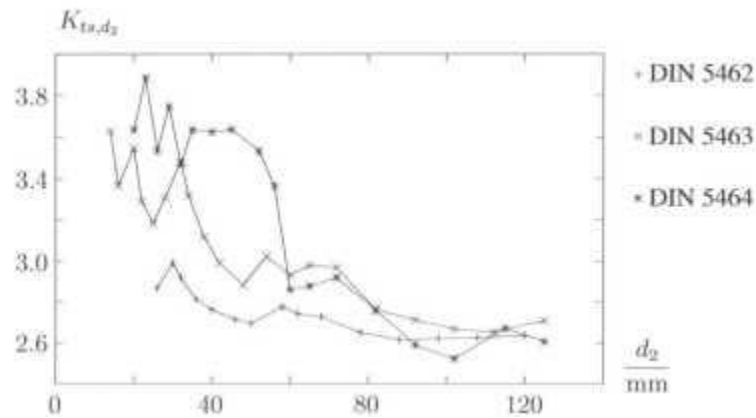
standards the splined shafts fall in three series; light, middle and heavy, depending on the teeth height. The geometric definitions are given in Figure 5.17

The stress concentration factors found are shown in Figure 5.18, the results are taken from [8], and includes the outdated standards DIN 5462 and DIN 5463. It should be noted that the standards do not



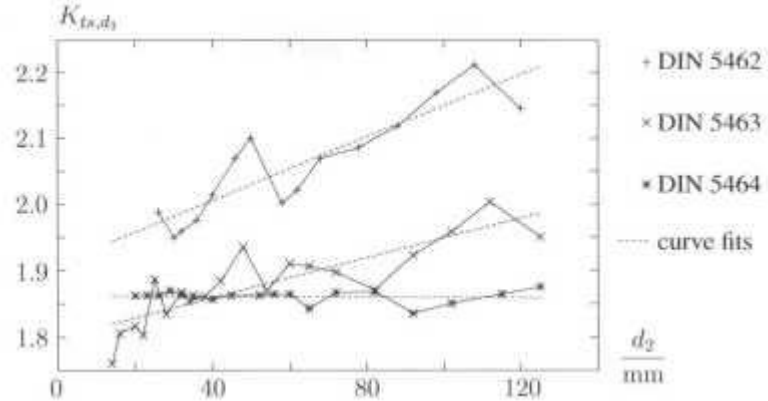
[billedtekst start]**Figure 5.17:** The geometric definition of half a tooth. Definitions are given as specified in DIN 5462-5463.[billedtekst slut]

define the root so the presented stress concentration factors corresponds to a lower limit, i.e. the largest possible r_r is selected.



[billedtekst start]**Figure 5.18:** The stress concentration factor of a splined shaft as a function of the outer diameter d_2 The nominal stress is defined relative to the outer diameter, i.e. $\tau_{nom} = \frac{16M_t}{\pi d_2^3}$. [8] [billedtekst slut]

If the nominal stress is defined relative to the inner diameter d_1 instead of relative to the outer diameter d_2 the stress concentration has a small and more intuitive value. This is because the shear stress will never be able to utilize fully the spline teeth. It will neither be possible to reach a value of 1 since there is material missing compared to a shaft with the outer diameter. Theoretically it is known that the stress concentration factor for a shaft with a semicircular cut in the limit is 2. It is therefore not surprising that the stress concentration defined relative to the inner diameter gives similar values. The stress concentration defined with the nominal stress defined relative to the inner diameter is shown in Figure 5.19.



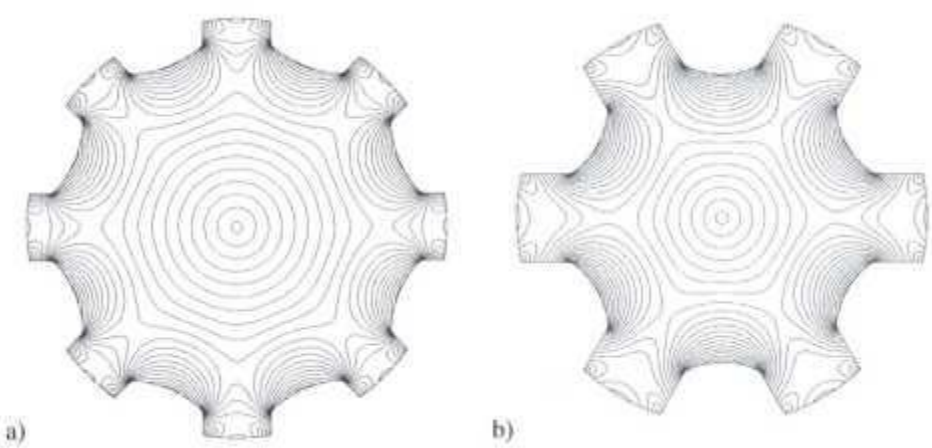
[billedtekst start]**Figure 5.19:** The stress concentration factor of a splined shaft as a function of the outer diameter d_2. The nominal stress is defined relative to the inner diameter, i.e. $\tau_{nom} = \frac{16M_t}{\pi d_1^3}$. Also shown are the linear curve fits to the results. [8][billedtekst slut]

From Figure 5.19 it is noticed that the values are relatively constant facilitating a simple linear curve fit. Linear curve fits to the results are also plotted in the figure. Specific expressions are

$$\text{DIN5462}: \quad K_{ts,d_2} = 1.9095 + 2.4 \cdot 10^{-3}\left(\frac{d_2}{\text{mm}}\right) \tag{5.64}$$

$$\text{DIN5463}: \quad K_{ts,d_2} = 1.7991 + 1.5 \cdot 10^{-3}\left(\frac{d_2}{\text{mm}}\right) \tag{5.65}$$

$$\text{DIN5464}: \quad K_{ts,d_2} = 1.9095 + 3.0 \cdot 10^{-5}\left(\frac{d_2}{\text{mm}}\right) \tag{5.66}$$



[billedtekst start]**Figure 5.20:** Cross sectional design of optimized splines with the resulting stress level iso lines. The iso lines indicate how constant the stress is along the design boundary. Please note that the two splines are scaled to the same size. a) The optimized design for DIN 5362 8×56×62. b) The optimized design for DIN 5363 6× 11 × 14. [8][billedtekst slut]

These curve fitted results should be used with care since they are defined relative to the maximum allowable radius at the tooth root. There are many possibilities for design improvements, see [8], To exemplify Figure 5.20 show two optimized splines scaled to the same size, in order to see the smaller spline iso lines. The design modification is made such that the splines use flange centering. The improvement in the stress concentration factor compared to the DIN norm is 27.1% and 15.2% respectively for the two designs in Figure 5.20.

5.5 Nomenclature

b	mm	Keyway width
d	mm	Shaft diameter
d_1	mm	Diameter of centering shaft surface on splined shaft
d_2	mm	Outer diameter of spline shaft
d_f	mm	Fit diameter
d_m	mm	Cone mean diameter
e_1	mm	Width of spline key
e_2	mm	Fillet diameter of fillet root in spline
e_3	mm	Width of shaft centering surface in spline
f_r	N/mm^3	Radial volume force
l	mm	Length of contact zone in interference fit
K_{ts}		Theoretical shear stress concentration factor
n_s	–	Safety factor
P	N/mm^2	Contact pressure
P_f	N/mm^2	Contact pressure between shaft and hub
r	mm	Radial coordinate
r_a	mm	Internal radius of shaft (if hollow)
r_b	mm	Outer radius of hub

r_f	mm	Fit radius
r_h	mm	Internal radius of hub before deformation
r_r	mm	Root radius
r_s	mm	Outer radius of shaft before deformation
t	mm	Key way depth
v_r	mm	Radial displacement
v_{rh}	mm	Radial displacement of hub fit surface
v_{rs}	mm	Radial displacement of shaft fit surface
A	mm^2	Contact area
$A_{bearing}$	mm^2	Bearing contact area
$A_{frustum}$	mm^2	Bearing contact area of frustum
A_{shear}	mm^2	Shear area
B_1, B_2	mm	Semi-axes of fillet super ellipse of spline root
E	N/mm^2	Modulus of elasticity
E_h	N/mm^2	Modulus of elasticity of hub
E_s	N/mm^2	Modulus of elasticity of shaft
F	N	Force
F_a	N	Assembly force
$F_{a,\max}$	N	Maximum assembly force for press fit
F_n	N	Force normal to surface
F_r	N	Radial force

K_{ts}	–	Theoretical shear stress concentration factor
L_1, L_2	mm	Length of key way planar surfaces
L_{eff}	mm	Active length of key, depending on the shape of the parallel key (straight or rounded ends)
L_f	mm	Length of conical contact
R	N	Resultant force
T	Nmm	Torque (torsional moment)
Ra	mm	Surface roughness (Arithmetic mean)
α	1/°C	Thermal expansion coefficient
α_c	rad	Cone angle
δ_r	mm	Radial fit
δ_d	mm	Diametral fit
Δt_m	°C	Temperature difference between shaft and hub
η	–	Super elliptical power
ϵ_{rr}	–	Radial strain
ϵ_{zz}		Axial stress
$\epsilon_{\theta\theta}$	–	Circumferential strain
μ_d	–	Dynamic coefficient of friction
μ_s	–	Static coefficient of friction
V	–	Poisson ratio
V_h	–	Poisson ratio for hub
V_s	–	Poisson ratio for shaft

ω	rad/s	Angular speed
ρ	kg/mm^3	Density
ρ_h	kg/mm^3	Density of hub
ρ_s	kg/mm^3	Density of shaft
σ_{rr}	N/mm^2	Radial stress
$\sigma_{\theta\theta}$	N/mm^2	Circumferential stress
σ_y	N/mm^2	Yield stress
τ_{max}	N/mm^2	Max shear stress

5.6 References

[1] F. Bodenstein. *Konstniktionselemente des maschinenbaues*. Springer Verlag, 1979.

[2] DIN 6885-1. PaBfedern nuten (in German), 1968.

[3] B. J. Hamrock, B. Jacobson, and S. R. Schmid. *Fundamentals of machine elements.* McGraw Hill, 1999.

[4] A. Leyer. *Maschinenkonstruktionslehre, heft 3. Spezielle gestaltungslehre*, 1. teil. Birkhauser Verlag, Switzerland, 1966.

[5] G. Niemann, Winter H., and Höhn G. *Maschinenelemente Band I.* Springer-Verlag, Berlin, Heildelberg. New York, 2005.

[6] N. Olhoff and A. G. Nielsen. *Noter til styrkelære II (in Danish).* Inst, for Mekanik energi og konstruktion, Technical University of Denmark (DTU), 2001.

[7] N. L. Pedersen. Stress concentrations in keyways and optimization of keyway design. *Journal of Strain Analysis for Engineering Design,* 45(8):593–604, 2010.

[8] N. L. Pedersen. Optimization of straight-sided spline design. *Archive of Applied Mechanics*, 81(10): 1393–1407, 2011.

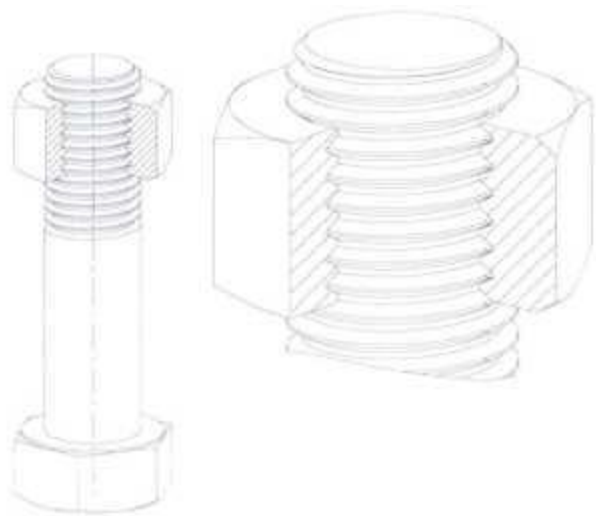
[9] J. E. Shigley and C. R. Michke. *Mechanical Engineering Design 7th ed.* McGraw Hill, Singapore, 2004.

Chapter 6 Threaded Fasteners

6.1 Introduction

A bolted connection is an assembly of components fixed to each other by one or more threaded fasteners. The fastener transmits the static and dynamic forces acting on the components. A threaded fastener without a nut is called a screw. A screw is called a bolt when used together with a nut.

Millions of screws and bolts are applied to mechanical design every day without being carefully dimensioned. However, in many applications careful calculation is needed, such as in cars, airplanes, cranes, steam and gas turbines etc., where mechanical reliability and human safety aspects are vital.

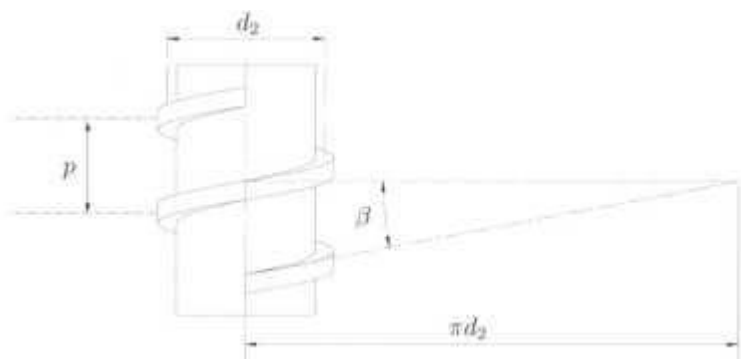


[billedtekst start]**Figure 6.1:** Bolt and nut.[billedtekst slut]

6.2 Characteristics of screw motion

When tightening or loosening a bolt, a screw motion is made along the screw axis. When one complete turn of the screw is made, a relative axial displacement is created along the screw axis which corresponds to the lead p_h. The pitch (or (lank pitch) p shown in Figure 6.2 is the distance from a point on one thread

to the same point on an adjacent thread. For a single threaded screw $P_h = P$ and for a double threaded screw $P_h = 2p$. In the following only single threaded screws are considered. However, it is relatively simple to modify the expressions so that double or triple threaded screws can be analyzed.



[billedtekst start]**Figure 6.2:** Spindle with flat thread.[billedtekst slut]

Unfolding a helix laying on a cylinder with the radius

$$r_2 = \frac{d_2}{2} \tag{6.1}$$

results in an inclined straight line with the pitch angle β where

$$\tan\beta = \frac{p}{\pi d_2} \tag{6.2}$$

6.3 Types of thread

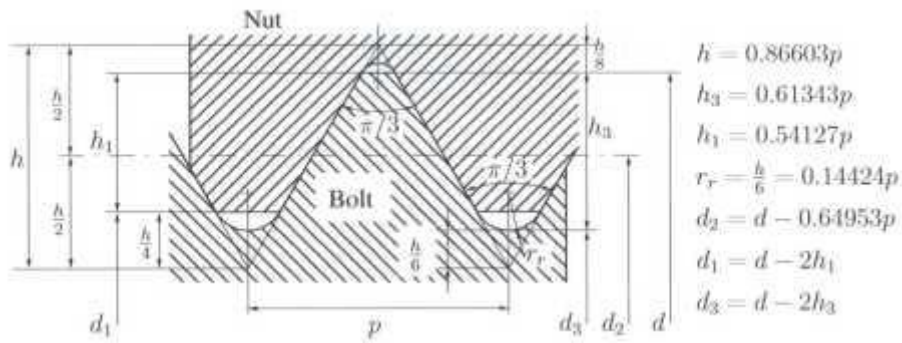
The thread profile is the outline of a thread seen in an axial cross section. The thread flanks are the parts of the thread profile that are in contact with the threaded counterpart. The thread flanks create an angle α called the flank angle. For ISO-metric threads the flank angle is $\alpha = \pi/3$.

V-Thread for fastener-type bolts. The metric ISO thread is the most commonly used in the EU. The profile of the bolt and nut thread can be seen in Figure 6.3. The major diameter d of the bolt thread is equal to the major diameter of the nut thread. This is also referred to as the nominal diameter. The minor diameter d_3 of the bolt is used to calculate the cross-section area of the core.

$$A_3 = \frac{\pi d_3^2}{4} \tag{6.3}$$

The groove and ridge of the thread has the same width along the axis on the pitch diameter d_2 of the bolt and nut.

The theoretical height h of the sharp V-profile with a pitch p, and a flank angle $\alpha = \pi/3$ is seen in Figure 6.3. Referring to Figure 6.3 for metric ISO thread, the following expressions can be deduced



[billedtekst start]**Figure 6.3:** Metric ISO thread.[billedtekst slut]

$$h = \frac{1}{2\tan\frac{\alpha}{2}}p \tag{6.4}$$

$$d_2 = d - \frac{3}{8\tan\frac{\alpha}{2}}p \tag{6.5}$$

$$d_3 = d - \frac{17}{24\tan\frac{\alpha}{2}}p \tag{6.6}$$

$$h_1 = \frac{5}{8}h \tag{6.7}$$

The fillet radius in the thread root

$$r_r = \frac{h}{6} \tag{6.8}$$

The flank overlap h_1 is also referred to as the depth of thread engagement. The fillet radius r_r in the thread root of the boll is specified in the standards. The stressed cross-section of the screw is approximated by

$$A_s = \frac{\pi}{4}\left(\frac{d_2 + d_3}{2}\right)^2 \tag{6.9}$$

A_s is the reference cross-section for strength calculations. In Tables 6.1 and 6.2 are listed the nominal diameter d, pitch p, diameter d_3 and stressed cross-section A_s for a selected series of metric (SO coarse and fine pitch threads.

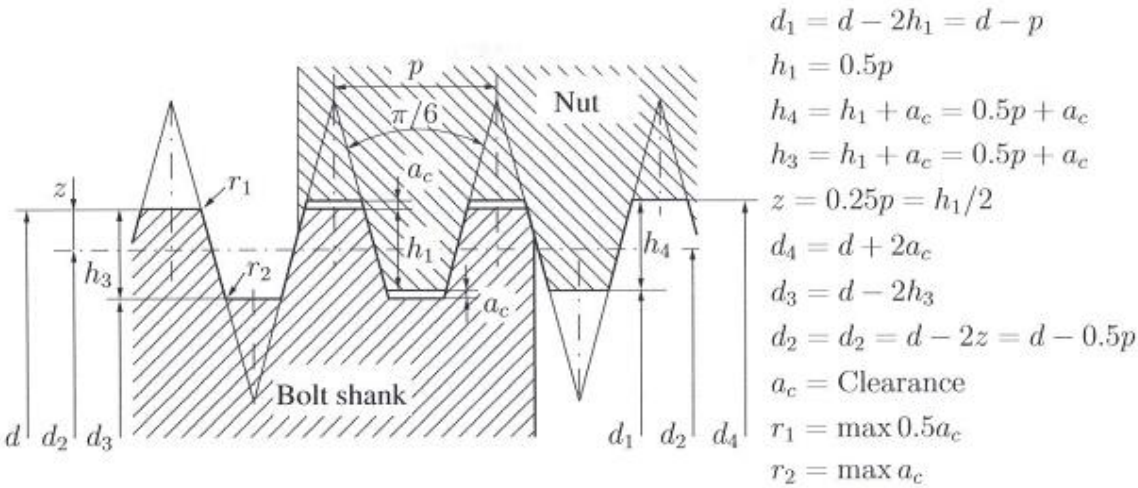
Flat thread. Flat thread for Translation-Type screws produces less friction between nut and bolt than the V-thread. For nominal profiles of the nut and bolt of a metric thread, with a clearance in the major and minor diameter but without flank clearance can be seen in Figure 6.4. The thread is side fitting (as the ISO metric thread) and because of the small flank angle it should therefore be loaded only by axial forces. Table 6.3 contains nominal sizes for trapezoidal threads.

Table 6.1: Extract of metric ISO threads - coarse series.

Nom. dia. d mm	**Pitch** p mm	**Pitch diam.** d_2 mm	**Core diam.** d_3 mm	**Stressed section** A_s mm²
4	0.7	3.545	3.141	8.78
5	0.8	4.480	4.019	14.2
6	1.0	5.350	4.773	20.1
8	1.25	7.188	6.466	36.6
10	1.5	9.026	8.160	58.0
12	1.75	10.863	9.853	84.3
16	2	14.701	13.546	157
20	2.5	18.376	16.933	245
24	3	22.051	20.319	353
30	3.5	27.727	25.706	561

Table 6.2: Extract of metric ISO threads - fine series.

Nominal dia. d mm	**Pitch** P mm	**Pitch diameter** d_2 mm	**Core diameter** d_3 mm	**Stressed section** A_s mm²
8	1	7.35	6.773	39.2
12	1.25	11.188	10.466	92.1
16	1.5	15.026	14.160	167
20	1.5	19.026	18.160	272
24	2	22.701	21.546	384
30	2	28.701	27.546	621



[billedtekst start]**Figure 6.4:** Translation type thread (ISO-profile).[billedtekst slut]

Table 6.3: Extract of metric ISO trapezoidal threads

Nominal dia. d mm	**Pitch** p mm	**Flank dia.** d_2 mm	**Root dia.** d_3 mm
10	2	9.0	7.5
12	3	10.5	8.5
16	4	14.0	11.5
20	4	18.0	15.5
24	5	21.5	18.5
28	5	25.5	22.5
32	6	27.0	23.0
36	6	29.0	25.0

Parametric representation of the thread surface

Any curve on a thread surface with constant distance to the bolt axis is a helix, and a helix has a simple mathematical form. The thread surface can therefore also be expressed in a rather simple

parametric form.

A 3D curve can generally be expressed in parametric form as

$$\{c_1(u)\} = \gamma_x(u)\{i\} + \gamma_y(u)\{j\} + \gamma_z(u)\{k\} \tag{6.10}$$

where $\{i\}$, $\{j\}$ and $\{k\}$ are the unit vectors along the x, y and z coordinate axes respectively. The independent variable u describes the position along the curve.

If we assume that the bolt axis is aligned with the z axis and look at the thread in the x,z plane it is a straight line that can be represented by

$$\{c_2(u)\} = \left(\frac{d_2}{2} + u\right)\{i\} - u \tan\frac{\alpha}{2}\{k\} \tag{6.11}$$

with a further assumption that the bolt is rotated such that the surface of the thread includes the point $\{d_2/2,0,0\}^T$. If we rotate this curve around the z axis we find the surface

$$\{s_1(u,v)\} = (\frac{d_2}{2} + u)\cos v\{i\} + (\frac{d_2}{2} + u)\sin v\{j\} - u\tan\frac{\alpha}{2}\{k\} \tag{6.12}$$

i.e. v is a second independent variable. The thread surface is formed when the rotation of the thread is accompanied by a translation, this translation is controlled by the pitch, p, and therefore the surface of the thread can be given as

$$\{s_t(u,v)\} = (\frac{d_2}{2} + u)\cos v\{i\} + (\frac{d_2}{2} + u)\sin v\{j\} + (\frac{p}{2\pi}v - u\tan\frac{\alpha}{2})\{k\} \tag{6.13}$$

One of the features of having the thread surface in parametric form is that we can find the normal, $\{N\}$, to the surface in any point of the surface by the vectorial cross product

$$\{N(u,v)\} = \frac{d\{s_t\}}{du} \times \frac{d\{s_t\}}{dv} \tag{6.14}$$

We find

$$\frac{d\{s_t\}}{du} = \cos v\{i\} + \sin v\{j\} - \tan\frac{\alpha}{2}\{k\} \tag{6.15}$$

$$\frac{d\{s_t\}}{dv} = -(\frac{d_2}{2} + u)\sin v\{i\} + (\frac{d_2}{2} + u)\cos v\{j\} + \frac{p}{2\pi}\{k\} \tag{6.16}$$

For the special case $u = 0$ and $v = 0$ we have

$$\{N(0,0)\} = \frac{d_2}{2}\left\{\begin{array}{c} \tan\frac{\alpha}{2} \\ -\tan\beta \\ 1 \end{array}\right\} \tag{6.17}$$

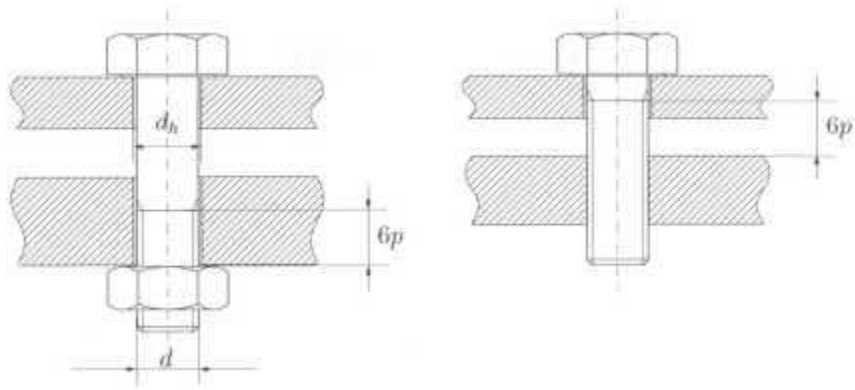
6.4 Types of bolts and nuts

Bolts. Figure 6.5 shows basic layouts of bolted connections and figure 6.6 shows three examples of layouts of bolted connections. The bolts in Figure 6.6 differ in the geometry of head, shank and end. The head form is defined by the method of driving, hexagonal head, hexagon socket head, slotted head and crosshead. The end shape is determined, among other things, by the method of manufacture or type of assembly.

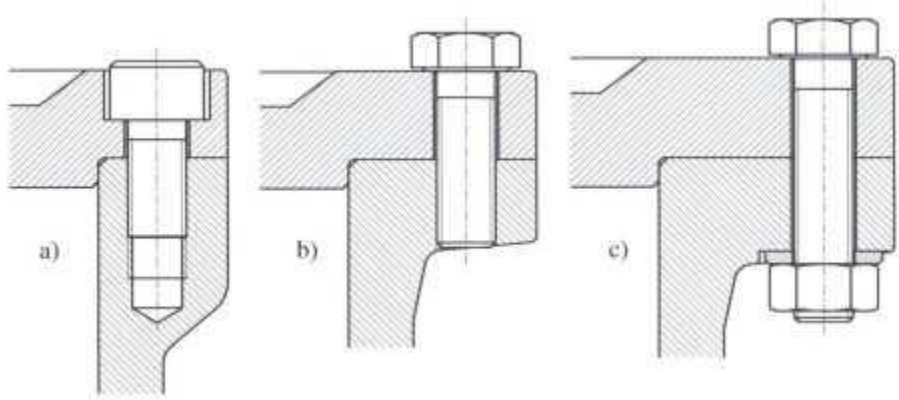
The shank form and diameter differ depending on application and/or manufacturing method. Special bolts have shanks with pilot surface for exact alignment of connected parts. On common mass produced bolts the thread is rolled into the shank which implies that the shank diameter is approximately equal to the thread pitch diameter. For classical bolts the shank diameter equals the nominal thread diameter.

Nuts. In mechanical engineering hexagon nuts are the most frequently used. Cap nuts, see Figure 6.7(a), provide protection of the thread. Special forms of nut, such as nuts with grooves, see Figure 6.7(b), or capstan nuts, see Figure 6.7(f), are used to provide axial location of hubs and rings on shafts or transmit axial force and must be tightened with a special spanner. Knurled nuts, slotted round nuts or ring nuts, see Figure 6.7(c), can be considered for low initial stresses, in sheet metal structures hexagon weld nuts, see Figure 6.7(e), may be fastened to the base material by spot welding. Hexagon nuts with centering shoulder, see Figure 6.7(g), and

capped nuts, see Figure 6.7(h), are used together with screws with reduced shanks.



[billedtekst start]**Figure 6.5:** Basic forms of bolted or screwed connections.[billedtekst slut]

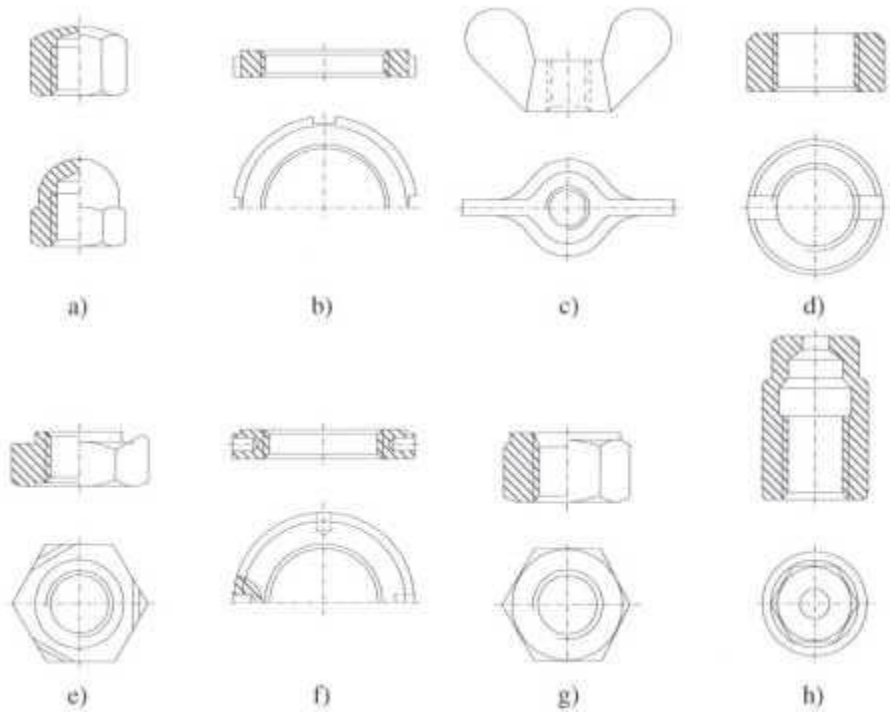


[billedtekst start]**Figure 6.6:** Various forms of bolted connections, a) Hexagon socket head screw in blind threaded hole, b) Hexagon head screw in through-threaded hole, c) Bolt and nut.[billedtekst slut]

Washers. These must be used under bolts and nuts if the base material has a tendency to set or would be over stressed. In the case of u-beams and I-beams special washers, with non-parallel surfaces must be used to allow for the slope of surfaces. Numerous washers with special surface geometries are claimed to secure the nut or protect the plate surfaces. In general these washers have no proven positive effect and they should be avoided.

6.5 Material specification for bolts and nuts

Bolt materials are identified according to property class. The designation for each property class is a real number, for example: 5.6, 6.8, 8.8, 9.8, 10.9 or 12.9. The digits before the decimal point is equal to a hundredth of the nominal tensile strength σ_{ut} in N/mm^2. The digits after the decimal point indicates 10 times the ratio of the nominal yield stress σ_y or $\sigma_{0.2}$ to the nominal tensile strength σ_{ut}, i.e. for a class 5.6 we have



[billedtekst start]**Figure 6.7:** Different types of standardized nuts.[billedtekst slut]

$$\sigma_{ut} = 5\ 100\text{N/mm}^2 = 500\text{MPa}$$

$$\sigma_y = \sigma_{ut} \cdot 6 \cdot \frac{8}{10} = 300\text{MPa}$$

Nuts with specified proof load stresses are designated by a property class between 4 and 12. The designation number is equal to a hundredth of the minimum tensile strength of a bolt in N/mm^2, which can be loaded up to the minimum yield strength, when used together with the nut. To fully utilize the strength of a 8.8 screw it should be paired with a nut of quality 8 or higher.

The length of the contact between the thread of the bolt/screw and the counter part being a nut or a blind hole is important. For standard nuts the height is given, if blind holes are used instead of a nut the minimum thread depths for selected cases can be found in Table 6.4.

6.6 Force and torque to preload a bolt

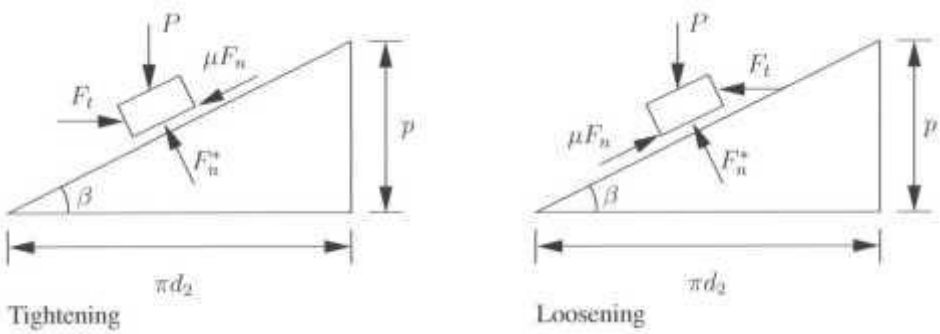
To determine the torque required to tighten or loose a bolted connection, it is required to study the forces acting on the thread. These forces depend on the friction coefficient, the thread geometry and the external load. The coefficient of friction is μ, if it is the static or the dynamic friction coefficient that should be used depends on the tightening process. In Figure 6.8 the forces on the thread is shown for the tightening and loosening case.

In Figure 6.8 we have used F_n to indicate the size of the normal load and F_n^* is the projection of this force onto the plane illustrated. The size of F_n^* depend on the flank angle α

and the pitch angle β. In

Table 6.4: Minimum thread depths in blind hole threads in nut material [13].

Bolt class	8.8		10.9
Thread fineness *d*/*p*	< 9	≥ 9	≥ 9
Al-alloy	1.1*d*	*1* 4*d*	
GG25	1.0 *d*	1.25 *d*	14 *d*
S235, St37, C15M	1.0*d*	1.25 *d*	14*d*
E295, St50, C35M	0.9 *d*	1.0*d*	1.2*d*
Steel with σ_{ut} > 800N/mm²	0.8 *d*	0.9d	1.0*d*

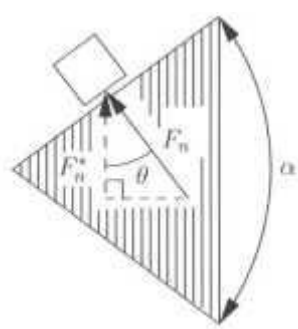


[billedtekst start]**Figure 6.8:** Forces acting on thread in a plane that has the normal of the plane perpendicular to the bolt axis.[billedtekst slut]

Figure 6.9 a normal cut trough Figure 6.8 is shown. We need to establish the angle θ, in Figure 6.10 the relationship among the angles are shown.

The angle is therefore given by

$$\tan\theta = \cos\beta \tan\frac{\alpha}{2} \tag{6.18}$$



[billedtekst start]**Figure 6.9:** A normal cut trough the free body diagram in Figure 6.8.[billedtekst slut]

We now have that

$$F_n^* = F_n \cos\theta \tag{6.19}$$

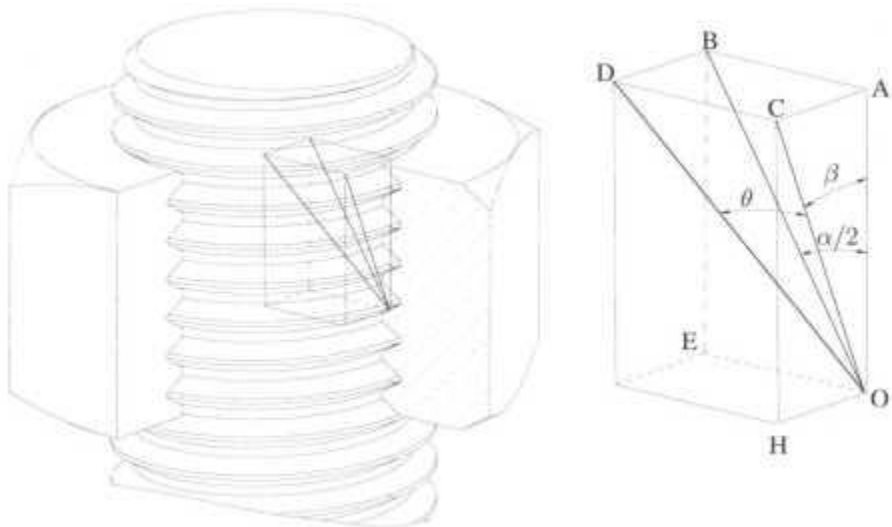


Figure 6.10: Forces acting on parallelepiped.

For the tightening case the two force equilibriums give

$$F_t - \mu F_n \cos\beta - F_n \cos\theta \sin\beta = 0 \tag{6.20}$$

$$-P - \mu F_n \sin\beta + F_n \cos\theta \cos\beta = 0 \tag{6.21}$$

which can be combined to

$$F_t = P\frac{\mu\cos\beta + \cos\theta\sin\beta}{\cos\theta\cos\beta - \mu\sin\beta} \tag{6.22}$$

For the loosening case we find from force equilibrium

$$F_l = P\frac{\mu\cos\beta - \cos\theta\sin\beta}{\cos\theta\cos\beta + \mu\sin\beta} \tag{6.23}$$

The torque necessary for tightening, T_t, and loosening, T_l is found from F_t since

$$T = F_t\frac{d_2}{2} \tag{6.24}$$

so therefore

$$T_t = P\frac{d_2}{2}\frac{\mu\cos\beta + \cos\theta\sin\beta}{\cos\theta\cos\beta - \mu\sin\beta} \tag{6.25}$$

$$T_l = P\frac{d_2}{2}\frac{\mu\cos\beta - \cos\theta\sin\beta}{\cos\theta\cos\beta + \mu\sin\beta} \tag{6.26}$$

In addition to the friction between the thread of the bolt and the nut there is also friction between the bearing surface of the nut and the assembled part. The frictional moment here can be given as $r_n \mu_n P$. The radius r_n is given as

$$r_n = \frac{d_w + d_h}{4} \tag{6.27}$$

where d_w is the diameter of the washer or bolt head, d_h, is the diameter of the hole and μ_n is the friction coefficient for the nut surface contact.

The general expressions for the torques are

The tightening torque The tightening torque T_t can be expressed as

$$T_t = P\left(\frac{d_2}{2}\left(\frac{\mu + \cos\theta \tan\beta}{\cos\theta - \mu \tan\beta}\right) + r_n \mu_n\right) \tag{6.28}$$

Loosening torque. The torque T_l in the thread required to loosen it is

$$T_l = P\left(\frac{d_2}{2}\left(\frac{\mu - \cos\theta \tan\beta}{\cos\theta + \mu \tan\beta}\right) + r_n \mu_n\right) \tag{6.29}$$

We talk about self-locking provided that the torque required for loosening is greater than zero, $T_l > 0$. Self-locking stops as soon as $T_l = 0$. The total torque T_l required for loosening is approximately equal to 0.7 to 0.9 times the tightening torque T_t for metric ISO V-threads so long as no vibrations reduce the effective coefficient of friction μ,.

Preload P and tightening torque T_t have an effect on tensile and torsional stresses in the bolt. The nominal tensile stress σ_z is calculated using either the stressed cross section of the thread A_s or the cross-section of the reduced shank if smaller. The von Mises equivalent tensile stress σ_{ref} then provides the actual stress in the material. If a 90% utilization of the material is considered permissible, then permissible assembly forces and the associated tightening torques for a specified coefficients of friction can be calculated.

An alternative derivation of the torque can be found using the parametric description of the thread surface. A unit vector in the normal direction can be given as (found from (6.17))

$$\{e_n\} = \frac{1}{G}\begin{Bmatrix} \tan\frac{\alpha}{2} \\ -\tan\beta \\ 1 \end{Bmatrix} \tag{6.30}$$

where $G = \sqrt{1 + \tan^2\frac{\alpha}{2} + \tan^2\beta}$. With this we directly have the components of the normal force

$$\frac{F_n}{G}\begin{Bmatrix} \tan\frac{\alpha}{2} \\ -\tan\beta \\ 1 \end{Bmatrix} \tag{6.31}$$

and the free body diagram for tightening and loosening are shown in Figure 6.11. The third component of the normal force is pointing towards the bolt axis and this can be neglected.

From the two diagrams in Figure 6.11 we find

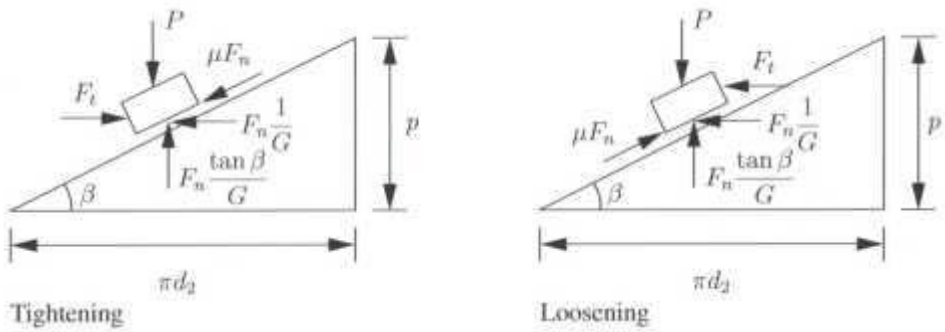


Figure 6.11 Forces acting on thread in a plane that has the normal of the plane perpendicular to the bolt axis.

The tightening torque

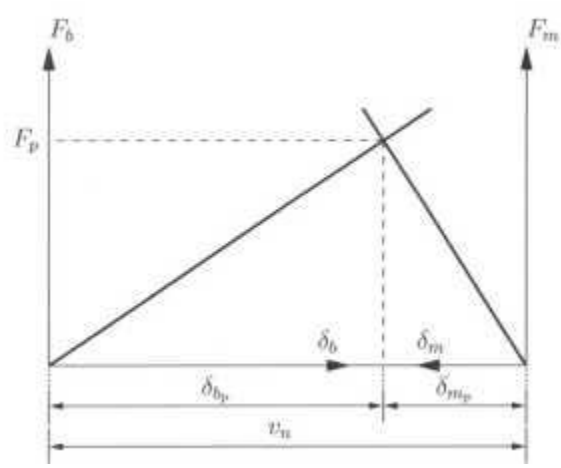
$$T_t = P\left(\frac{d_2}{2}\left(\frac{\mu G\cos\beta + \tan\beta}{1 - \mu G\tan\beta}\right) + r_n\mu_n\right) \tag{6.32}$$

Loosening torque.

$$T_l = P\left(\frac{d_2}{2}\left(\frac{\mu G\cos\beta - \tan\beta}{1 + \mu G\tan\beta}\right) + r_n\mu_n\right) \tag{6.33}$$

6.7 Deflection in joints due to preload

The forces and deflections arising in the bolt and plates after tightening are dependent on the effective assembly (preload) force F_p. If linear stiffness/deflection behaviour is assumed, the so called deflection triangle in Figure 6.12 can be drawn.



[billedtekst start]**Figure 6.12:** Deflection triangle. Left: The load (F_b,) displacement (δ_b,) coordinate system for the bolt. Right: The load (F_m) displacement (δ_m) coordinate system for the member plates.[billedtekst slut]

In this graphical representation the characteristic force-deflection lines for bolt and

plates are combined. In Figure 6.12 the preload F_p is defined together with the preload displacements of the bolt δ_{bp}

and the member plate preload displacements δ_{mp}. In the figure the movement of the nut along the thread v_n is also shown. Using the designations as indicated, the stiffness k_b of the bolts is expressed as

$$k_b = \frac{F_b}{\delta_b} \tag{6.34}$$

The inverse stiffness will here be termed flexibility, in the literature, see e.g. [1], this is also termed resilience. The flexibility f_b of the bolts is

$$f_b = \frac{1}{k_b} \tag{6.35}$$

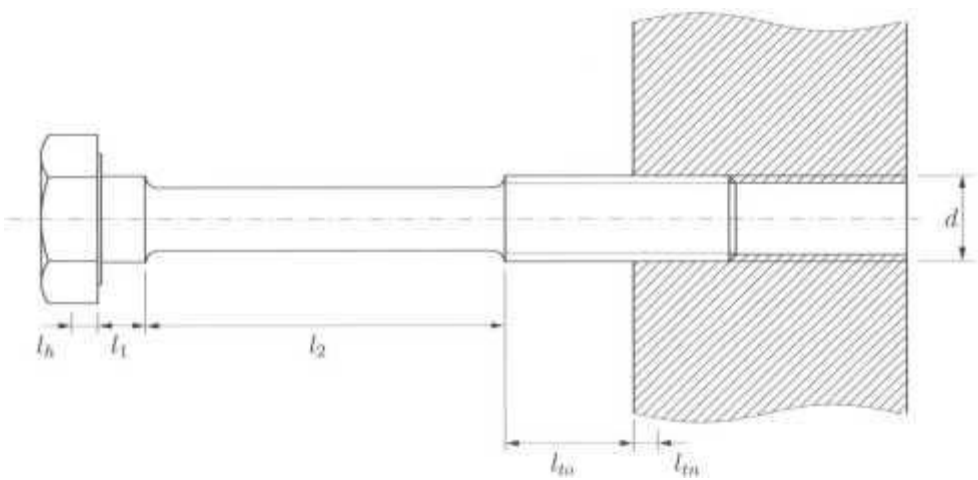
The stiffness of the member plates between the bearing surfaces of the bolt head and the nut is

$$k_m = \frac{F_m}{\delta_m} \tag{6.36}$$

and the flexibility

$$f_m = \frac{1}{k_m} \tag{6.37}$$

if retained in the center.



[billedtekst start]**Figure 6.13:** A bolt with different shaft diameters.[billedtekst slut]

Flexibility of bolts. Bolts consist of a number of individual elements which can be readily substituted by imaginary cylinders of varying lengths l_i and cross-sections A_i, see Figure 6.13. It follows that the flexibility of an individual cylindrical element is

$$f_i = \frac{l_i}{(E_b A_i)} \tag{6.38}$$

where E_b is the modulus of elasticity of the bolt material. The total flexibility of the bolt f_b becomes $f_b = \sum f_i$.

In the following an approximation for the individual flexibilities of parts of the bolt is

given. The elastic flexibility of the bolt head is estimated as

$$f_h = \frac{l_h}{E_b A_n} \tag{6.39}$$

where l_h, = 0.5d for hexagonal head bolts and l_h = 0.4d for socket head cap screws. The area A_n is given by

$$A_n = \frac{\pi d^2}{4} \tag{6.40}$$

The flexibility of the open portion of the thread is

$$f_{to} = \frac{l_{to}}{E_b A_3} \tag{6.41}$$

where l_{to} is the length and A_3 is the root cross-section.

$$A_3 = \frac{\pi d_3^2}{4} \tag{6.42}$$

The flexibility of the engaged part of the thread including the nut or the tapped thread region is given by

$$f_{tn} - f_{te} + f_n \tag{6.43}$$

i.e., the flexibility is the sum of the flexibility of the engaged thread f_{te} and that of the nut f_n. The flexibility of the engaged thread is estimated as

$$f_{te} = \frac{0.5d}{E_b A_3} \tag{6.44}$$

while the flexibility of the nut or tapped joint is estimated as

$$f_n = \frac{l_{tn}}{E_b A_n} \tag{6.45}$$

where l_{tn} = 0.4d for the nut and l_{tn} = 0.33d for the tapped thread joint. The total flexibility of the bolt in Figure 6.13 can therefore be expressed as

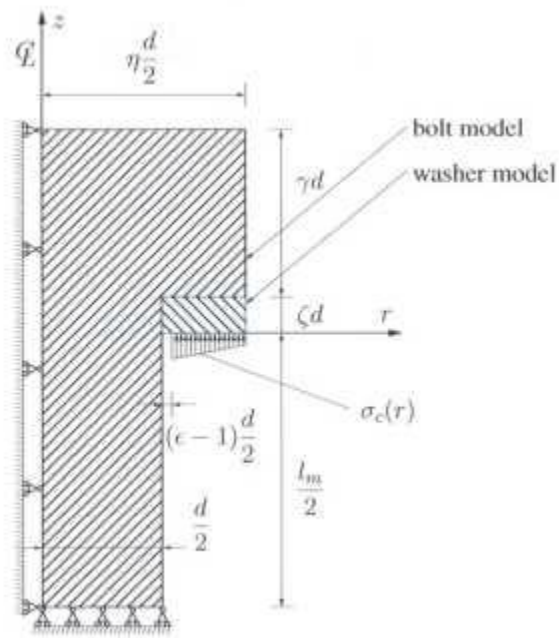
$$f_b - f_h + f_1 + f_2 + f_{to} + f_{tn} \tag{6.46}$$

the values for f_h, f_{to} and f_{tn} are all taken from [12].

Flexibility of member/retained plates. Exact determination of the flexibility of plates (or more general: compressed members) is rather complicated and requires in most cases a FE-analysis. A thorough discussion of different stiffness evaluations can be found in [7]. Good approximations can be obtained by the models described in the following, where a rather extensive review of the different methods proposed in the literature is presented. The primary idea behind is to illustrate the large variation in the found results using the different methods. The most used expression is probably the latest VD1 form 2003.

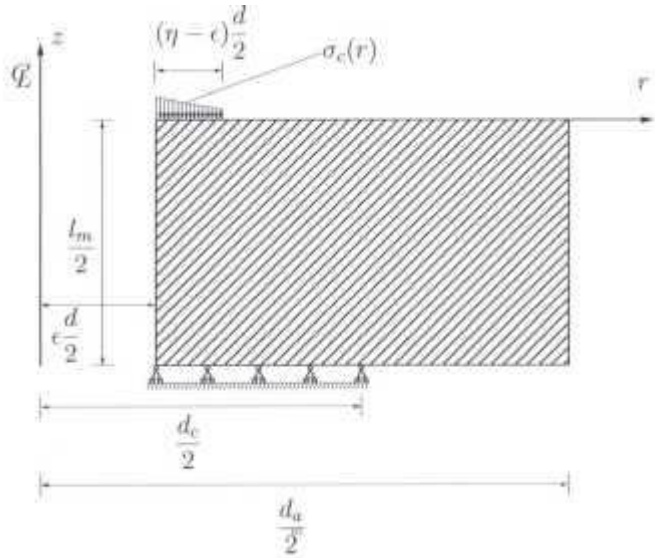
The geometry of the connection is presented in Figure 6.14 which shows a quarter of the bolt cross section, in the figure we show the dimensions of the bolt including a washer. In many analyses performed this part is neglected and further an assumption of axis symmetry is used. The washer can also be treated as an integrated part of the bolt.

The two plate members that are assembled by the bolt are assumed to be of equal thickness, and of the same material, so we may, in addition to the axis symmetry model, only model one of the plates. Figure 6.15 shows a quarter of the plate assembly cross section, were dimensions are also shown. In



[billedtekst start]**Figure 6.14:** The symmetric part of a bolt including a washer.[7][billedtekst slut]

Figures 6.14 and 6.15 the contact pressure distribution between the washer and the plate is indicated for illustrative purposes. If the FEM is used to calculate the stiffness then this contact pressure must be determined.



[billedtekst start]**Figure 6.15:** The member plate for the symmetric problem.[7][billedtekst slut]

The determination of the member stiffness, i.e., the assembled plates, has a long history that dates back to [9]. The assumption behind most formulas of the stiffness is that the width, d_a, of the member is infinite or so large that there is no elastic energy in the outer part of the member. In [9] it was proposed that the stress in the members is uniform in two frusta symmetric around the symmetry line, with a top angle of 2φ given by $\varphi = \pi/4$, see also Figure 6.16. This leads to member stiffness k_m

$$k_m = \frac{\pi E_m}{4l_m}((\eta d + \frac{l_m}{2})^2 - (\epsilon d)^2) \tag{6.47}$$

where E_m is Young's modulus, l_m is the combined length of the members, d is the nominal diameter of the bolt and ϵ and η are non-dimensional factors as indicated in Figures 6.14 and 6.15, i.e. $d_h = \epsilon d$ and $d_w = \eta d$. In many papers the clearance of the hole is set equal to zero and as such it is the nominal diameter of the bolt that is used in (6.47), i.e., $\epsilon = 1$ is used. It should however be the diameter of the hole that is used instead, this is the reason for including ϵ in (6.47) here. The assumption that leads to (6.47) is a very simple and straightforward assumption but not very accurate, also it is based on an assumption that is not in static equilibrium. Assuming that the total prestress contact load is P and integrating over the frusta we get (see e.g. [10])

$$k_m = \frac{\pi E_m \epsilon d \tan(\varphi)}{2\ln\left(\frac{(l_m \tan(\varphi) + (\eta - \epsilon)d)(\eta + \epsilon)}{(l_m \tan(\varphi) + (\eta + \epsilon)d)(\eta - \epsilon)}\right)} \tag{6.48}$$

which can be simplified in the case of $\varphi = \pi/4$.

The assumption of $\varphi = \pi/4$ is not appropriate as it was experimentally verified by [2] or [3], A more suitable assumption is to use the value $\varphi = \pi/6$. In [5] the stiffness is defined as (rearrangement to fit the definition used here)

$$k_m = \frac{\pi E_m}{4l_m}((\eta d + \frac{l_m}{2\sqrt{3}})^2 - (\epsilon d)^2) \tag{6.49}$$

which corresponds to the same assumption as [9] but with $\varphi = \pi/6$ instead. In [10] it is proposed to use (6.48) also with $\varphi = \pi/6$. As an alternative to these stiffness calculations both [5] and [10] suggest to use the results found by [14] which is a curve fit to FE results that is given by

$$k_m = E_m \epsilon d C_1 e^{C_2(\epsilon d/l_m)} \tag{6.50}$$

where C_1 and C_2 are constants that depend on the material data. The constants C_1, C_2 may be found in Table 6.5. The curve fit using an exponential function should be used with caution because the approximation fail drastically when $\epsilon d/l_m$ becomes large.

In [4] other earlier expressions for the member stiffness are given together with a new one. The assumption by [4] is that the stress in the member can be given as

$$\sigma(r, z) = a4r^4 + a_3r^3 + a_2r^2 + a_1r + a_0 \tag{6.51}$$

where the coefficients $a4 - a_0$ are functions of the z position (see Figure 6.15). It is assumed that

the stress will vanish at a distance $r_0(z)$ which also is a function of z. The assumed boundary conditions for finding these values are

Table 6.5: Constants A_i and B_i for the Wileman plate stiffness formula [14].

Material	**Modulus of elasticity** E_m [GPa]	**Poisson's ratio** υ_m [–]	C_1 [–]	C_2 [–]
Steel	206.8	0.291	0.78715	0.62873
Aluminium	71.0	0.334	0.79670	0.63816
Cast Iron	100	0.211	0.77871	0.61616

$$
\begin{aligned}
r &= \frac{\epsilon d}{2} && \Rightarrow && \frac{d\sigma}{dr} = 0 \\
r &= r_0 && \Rightarrow && \sigma = 0, \frac{d\sigma}{dr} = 0, \frac{d^2\sigma}{dr^2} = 0 \\
P &= \int_{\epsilon d/2}^{r_0} 2\pi r \sigma dr
\end{aligned}
\tag{6.52}
$$

Different definitions of $r_0(z)$ are given, one corresponds to a frustum

$$2r_0 = \eta d - 2z \tan(\varphi) \tag{6.53}$$

where the equation is rewritten to fit the definition of the z coordinate given here. The stiffness is then found using the unidirectional stress state, Hookes law and the displacement at a point $r = (\epsilon + \eta)d/A$.

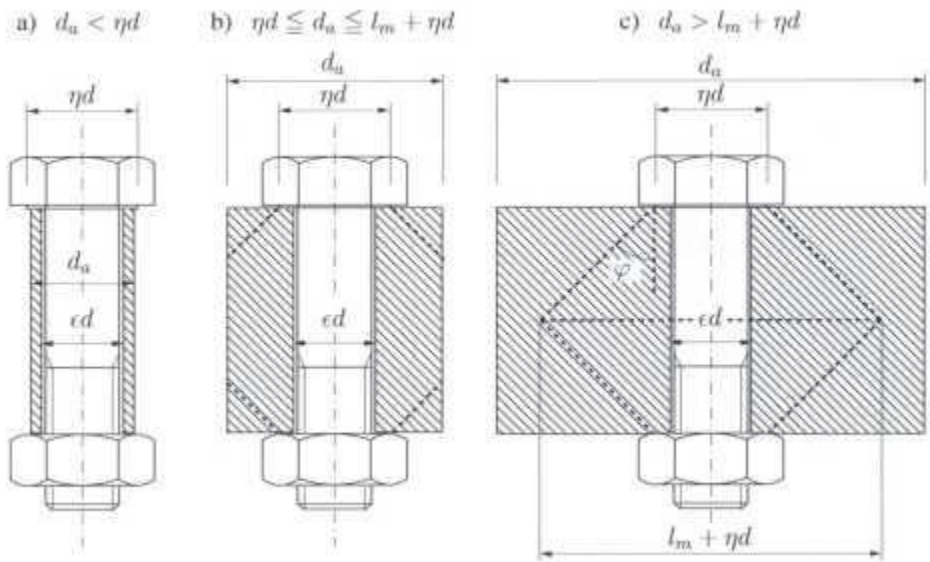
A more general expression for the stiffness is found in [13] which also specifies the stiffness in the case when the width of the member, d_a, has a limited size. The stiffness according to [13] is given by

$$
\begin{aligned}
d_a &\leqq \eta d && \Rightarrow && k_m = \frac{\pi E_m}{4l_m}(d_a^2 - (\epsilon d)^2) \\
\eta d < d_a &< 3\eta d && \Rightarrow && k_m = \frac{\pi E_m d^2}{4l_m}(\eta^2 - \epsilon^2) + \frac{\pi E_m}{8l_m}(\frac{d_a}{\eta d} - 1)(\frac{\eta d l_m}{5} + \frac{l_m^2}{100}) \\
d_a &\geqq 3\eta d && \Rightarrow && k_m = \frac{\pi E_m}{4l_m}((\eta b + \frac{l_m}{10})^2 - (\epsilon d)^2)
\end{aligned}
\tag{6.54}
$$

where the last formula for $d_a \geqq\geqq 3\eta d$ is the one that should be compared to the previously given expressions for the stiffness. The stiffness expressions in (6.54) are later changed as it may be found in e.g. [ll] into the expression

$$
\begin{aligned}
d_a < \eta d \quad &\Rightarrow \quad k_m = \frac{\pi E_m}{4l_m}(d_a^2 - (\epsilon d)^2) \\
\eta d \leqq d_a \leqq l_m + \eta d \quad &\Rightarrow \quad k_m = \frac{\pi E_m d^2}{4l_m}(\eta^2 - \epsilon^2) + \frac{\pi E_m}{8l_m}\eta d(d_a - \eta d)((\sqrt[3]{\frac{l_m \eta d}{d_a^2}} + 1)^2 - 1)) \\
d_a > l_m + \eta d \quad &\Rightarrow \quad k_m = \frac{\pi E_m d^2}{4l_m}(\eta^2 - \epsilon^2) + \frac{\pi E_m}{8}\eta d((\sqrt[3]{\frac{l_m \eta d}{(\eta b + l_m)^2}} + 1)^2 - 1)) \qquad (6.55)
\end{aligned}
$$

the three different ranges for d_a in (6.55) are illustrated in Figure 6.16.



[billedtekst start]**Figure 6.16**: The three different plate width situations for which separate definitions of the member stiffness are given.[billedtekst slut]

More recently the suggestions by VDI have been changed again. The new suggestions can be found in [12]. This expression is simplified to the case presented here given by

$$
\begin{aligned}
d_a < \eta d \quad &\Rightarrow \quad k_m = \frac{\pi E_m}{4 l_m}(d_a^2 - (\epsilon d)^2) \\
\eta d < d_a < d_{a,gr} \quad &\Rightarrow \quad k_m = \frac{\pi E_m}{\dfrac{2}{\epsilon d \tan(\varphi_d)} \ln\left(\dfrac{(\eta-\epsilon)(d_a-\epsilon d)}{(\eta+\epsilon)(d_a+\epsilon d)}\right) + \dfrac{4}{d_a^2-(\epsilon d)^2}\left(l_m - \dfrac{d_a - \eta d}{\tan(\varphi_d)}\right)} \\
d_a > d_{a,gr} \quad &\Rightarrow \quad k_m = \frac{\pi E_m \epsilon d \tan(\varphi_d)}{2 \ln\left(\dfrac{(l_m \tan(\varphi_d) + (\eta-\epsilon)d)(\eta+\epsilon)}{(l_m \tan(\varphi_d) + (\eta+\epsilon)d)(\eta-\epsilon)}\right)}
\end{aligned}
\tag{6.56}
$$

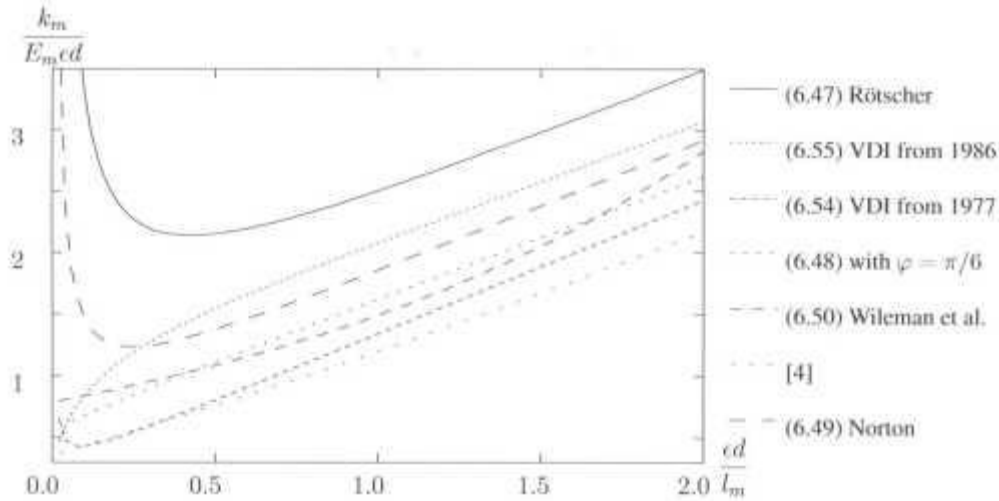
where $d_{a,gr}$ and $\tan(\varphi d)$ are given by

$$\tan(\varphi_d) = 0.362 + 0.032 \ln\left(\frac{l_m}{\eta d}\right) + 0.153 \ln\left(\frac{d_a}{\eta d}\right) \tag{6.57}$$

$$d_{a,gr} = \eta d + l_m \tan(\varphi_d) \tag{6.58}$$

From the equation it is seen that the [12] suggestions comply with (6.48) found in [10]

Other expressions for the stiffness can be found in the literature. However just from the many expressions for the stiffness presented here it is clear that there is a large variation in the suggestions. This is best illustrated graphically. In Figure 6.17 we have plotted, as also done in [14] and [10], the dimensionless stiffness, $k_m/(E_m \epsilon d)$, as a function of the ratio between the clearance of the hole and the length of the member, $\epsilon d/l_m$. The plot is done for the case where ϵ = 1.1 and η = 1.7. Unfortunately, such a plot will be highly depending on the value of η.



[billedtekst start]**Figure 6.17:** Dimensionless stiffness plots versus aspect ratio of clearance diameter to length of members.[7][billedtekst slut]

In [7] an extensive study using the finite element method and the elastic energy to determine the stiffness of the member have resulted in new simplified equations for determining the stiffness.

$$d_a \leqq \eta d \quad \Rightarrow \quad k_m = \frac{\pi E_m}{4L}(d_a^2 - (\epsilon d)^2) \tag{6.59}$$

$$\eta d < d_a < \eta d + L \quad \Rightarrow \quad k_m = \frac{k_0 - k_{max}}{e^{\left(\frac{\pi E_m \eta d (d_a - \eta d)}{2L(k_{max} - k_0)}\right)}} + k_{max} \tag{6.60}$$

$$d_a \geqq \eta d + L \quad \Rightarrow \quad k_m = k_{max} \tag{6.61}$$

where k_0 is the value of the stiffness for $d_a = \eta d$ and k_{max} is the asymptotic value of the stiffness (infinite d_a). These are given by

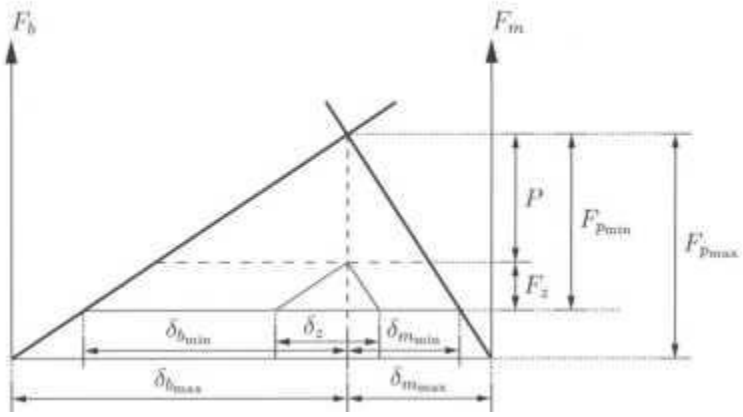
$$k_0 = \frac{\pi E_m}{4L}(\eta^2 - \epsilon^2)d^2 \tag{6.62}$$

$$k_{max} = E_m d(0.59(\eta^2 - \epsilon^2)\frac{d}{L} + 0.20(\eta + \epsilon)) \tag{6.63}$$

Dispersions during tightening. Depending on the tightening process of the bolt there will be dispersion in the assembly force. The dispersions of F_p between F_{Pmin} and F_{Pmax} that can occur during tightening is illustrated in Figure 6.18. The maximum preload F_{pmax} must remain smaller than the permissible bolt force, which is equal to 90% of yield stress for bolts up to *M*39.

Set/embedding. During the tightening process up to the assembly preload F_p the bearing surface under bolt head and nut and the contact planes between the plates are flattened. This will also take place when the external load is applied. Different estimates for this set or embedding can be found in the literature. The most simple expression (see [11]) is completely independent both of the number of junction lines and of the size of the irregularities of the mating surfaces. For solid joints the set δ_z (see Figure 6.18) is directly given by the relative

length (l_m/d).



[billedtekst start]**Figure 6.18:** Deflection diagram used to determine the influence of yielding and preload distribution.[billedtekst slut]

$$\delta_z = 3.29\left(\frac{l_m}{d}\right)^{0.34} 10^{-3}\,\text{mm} \tag{6.64}$$

Other expressions that also depend on the number of junction lines can be found. But here we limit the discussion to expression (6.64). Due to the setting of the joint by the amount δ_Z the assembly preload F_p is reduced by the amount F_z. The preload P $F_p - F_z$ is what remains. P must be greater than or equal to the required preload. The set causes a reduction in the elongation of the bolt by $F_z f_b$ and the compression of the plates by $F_z f_m$. Therefore

$$\delta_z = F_z f_b + F_z f_m \tag{6.65}$$

and

$$F_z = \frac{\delta_z}{(f_b + f_m)} \tag{6.66}$$

In order to avoid setting when using high preload bolts; no washers, locking plates or spring rings must be used under the bolt head or nut. It is important that the surfaces under bolt head and nut are properly finished and at right angles to the axis of the bolt.

6.8 Superposition of preload and working loads

If an axial tensile load F_a acts centrally on a symmetrically shaped and (centrally)preloaded bolted connection under the bolt head and nut of a through bolt, the bolt is elongated by a amount δ_{ba} and the compression of the plates is reduced by the equivalent amount δ_{ma}. The bolt and plate are engaged parallel to the tensile load F_a. The additional bolt force is expressed as

$$F_{ba} = k_b \delta_{ba} \tag{6.67}$$

and

$$F_a = (k_b + k_m)\,\delta_{ba} \tag{6.68}$$

The clamping force in the plates is reduced by

$$F_{ma} = F_a - F_{ba} \tag{6.69}$$

The forces are shown in the deflection diagram Figure 6.19.

Defining the force ratio ϕ by

$$F_{b_a} = \Phi F_a = \frac{k_b}{k_b + k_m} F_a \tag{6.70}$$

it follows that

$$\Phi = \frac{k_b}{k_b + k_m} = \frac{f_m}{f_b + f_m} \tag{6.71}$$

The residual clamping force in the junction line F_c after setting and loading is

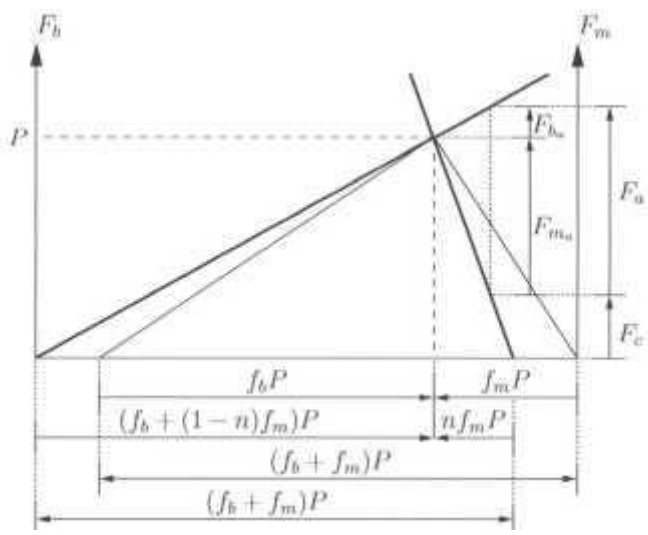
$$F_c = P - F_{ma} = P - (1 - \phi)F_a \tag{6.72}$$

it must be at least equal to the required clamping force $F_c \geq F_{Creq}$.

With α_a as the tightening factor the maximum assembly load becomes

$$F_{pmax} = \alpha_a F_{Pmin} = \alpha_a (F_{Creq} + (1 - \phi)F_a + F_z) \tag{6.73}$$

Distribution of forcc over retained members. In general the external axial force does not act directly below the bolt head and nut even when acting centrally, but within the retained members. If it is assumed that the points where the force is acting are not at a distance l_m between the bearing surfaces of the bolt head and nut, but only at the distance nl_m where $0 < n < 1$, then all the areas of the plate are no longer relaxed by the axial force F_a as the stiffness ratio of the loaded and relaxed areas of the bolted joint change. The relationships are shown in Figure 6.19.



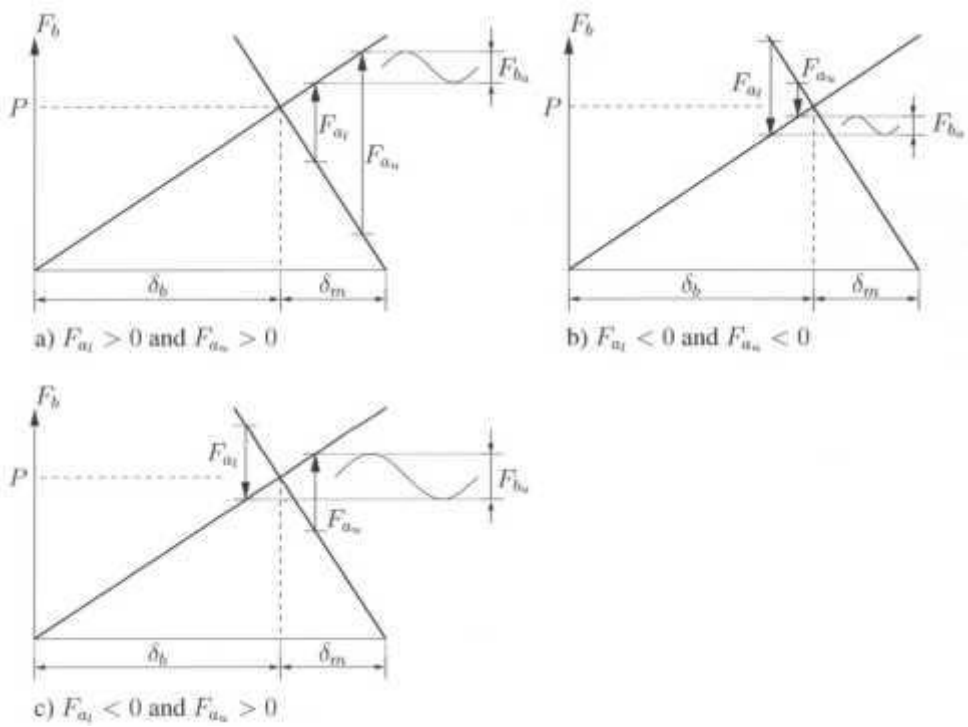
[billedtekst start]**Figure 6.19:** Deflection diagram for working load inside the restrained members.[billedtekst slut]

Dynamic loading. One of the important reasons for preloading bolts is the reduction in the amplitude stress the bolts "feel" from a dynamic load. This is can be illustrated graphically as done in Figure 6.20.

The externally applied load is assumed to vary between a lower value F_{at} and an upper value F_{au}. Both are regarded as being positive when applying tension to the bolt. In Figure 6.20 the three different possibilities for the sign of the two forces are shown. In Figure 6.20(a) both forces are positive, it is clear from the figure that the amplitude force the bolt "feels" are much smaller than the externally applied amplitude, i.e.

$$\frac{F_{b_a}}{2} = \Phi \frac{F_{a_u} - F_{a_l}}{2} \tag{6.74}$$

The same is true for the two other cases in Figures 6.20(b) and (c). The reduction in the amplitude is controlled by the size of the force ratio ϕ.

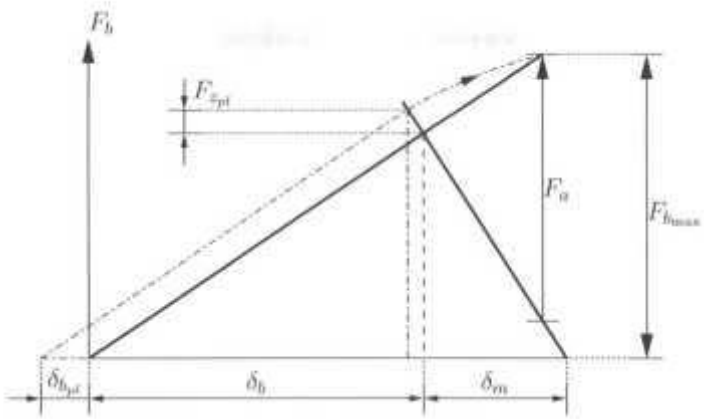


[billedtekst start]**Figure 6.20:** Dynamic loading. In the three figures the variation of the force in the bolt is illustrated together with the size (including sign) of the upper and lower size of an external loading.[billedtekst slut]

Loading up to the plastic region. If a bolt is stressed by a centrally acting external tensile load F_a into the plastic region, a change in the preload triangle (represented by a broken line) then follows, as shown in Figure 6.21. After the relaxation the removal of the external load F_a only the preload reduced by F_{Zpl} remains. F_{Zpl} is obtained with

$$F_{z_{pl}} = \frac{\delta_{b_{pl}}}{(f_b + f_m)} \tag{6.75}$$

where δ_{mpl} is the portion of plastic deflection under the total bolt force $F_{b\mathrm{max}}$ after F_a is applied.



[billedtekst start]**Figure 6.21** Plasticity in the bolted connections.[billedtekst slut]

6.9 Failure of bolted connections

In order to design bolted connections the external loads that occur during operation must be known as precisely as possible. It is useful for design purposes to make a distinction between frequently and rarely occurring working loads. The rare heavy extra loads are statistically assessed in terms of strength calculations, whereas for frequently occurring working loads it is usually an assessment of fatigue strength that is aimed for at least at the design stage. The ideal is a bolted connection that fully compensates the connecting sections of the members for any working extra loads that may arise. Bolted connections can therefore fail either by static failure or by dynamic failure (fatigue).

Static failure

1) If a minimum force, F_{creq}, is needed between the assembled parts then the maximum preload must have the value given by

$$F_{Pmax} = \alpha_a F_{Pmin} = \alpha_a (F_{Creq}, + (1 - \phi)F_a + F_z) \tag{6.76}$$

2) The maximum stress in the bolt shank should be less than the proof strength, σ_p, i.e.

$$\sigma = \frac{P + \Phi F_{a_u}}{A_s} \leq \sigma_p \tag{6.77}$$

The proof strength is the stress at which the bolt begins to take a permanent set. This stress is typically a little smaller than the yield stress. If the proof strength is not know the yield stress can be applied instead with some caution.

3) The stress in the thread should be checked. The stress level depends on the bolt being made from a ductile or brittle material. The difference is that with a ductile material all threads will carry the load and therefore the shear area of all the threads is used. In the stress level check for a brittle material only one thread will carry the load. Here it is worth mentioning that a good assumption is that the first turn of the thread carries half

the load ($P/2$) the next half of the remaining load ($P/4$) etc.

In order to find the stress level we need the shear area. If the nut is the strongest then the thread of the bolt will be stripped at the minor diameter $d_3$if the bolt is the strongest the thread of the nut is stripped

Table 6.6 Surface pressures σ_g..

Material	**Ultimate tensile stress** σ_{ut} N/mm²	**Surface pressure** σ_g N/mm²
S235, St37	370	260
E295, St50	500	420
C 45	800	700
42 CrMo 4	1000	850
30 CrNiMo 8	1200	750
X 5 CrNiMo 18 10	500-700	210
X 10 CrNiMo 18 9	500-750	220

at the nominal diameter d and if they are of equal strength the stripping will take place somewhere in between most probably at the pitch diameter d_2.

The shear area for one full revolution of the thread depends on the thread design but for the ISO thread we can approximate the three different situations by

$$A_{sh} = 0.87\pi d_3 p \tag{6.78}$$

$$A_{sh} = 0.8757\pi\, d_3 p \tag{6.79}$$

$$A_s h = 0\text{-}5\pi d_2 p \tag{6.80}$$

For a brittle material we have that the following must be fulfilled

$$\sqrt{3}\frac{P + \Phi F_{a_u}}{2A_{sh}} \leq \sigma_{ut} \tag{6.81}$$

while for a ductile material with nut height l_n we have

$$\sqrt{3}\frac{P + \Phi F_{a_u}}{\frac{l_n}{p}A_{sh}} \leq \sigma_y \tag{6.82}$$

4) Pressure under bolt head and nut must be checked whether the permissible pressure in the bearing surface of the bolt head and nut has been adhered to.

$$\frac{P + \Phi F_{a_u}}{\frac{\pi}{4}(d_w^2 - d_h^2)} \leq \sigma_g \tag{6.83}$$

Here elastic tightening is assumed. For permissible unit pressures σ_g see e.g. Table 6.6.

Dynamic failure

To verify that the bolt does not fail by fatigue we must find the mean and alternating stress in the bolt. The mean and amplitude value of an alternating force in the bolt are given by

$$F_{amp} = \Phi \frac{F_{a_u} - F_{a_l}}{2} \tag{6.84}$$

$$F_{mean} = P + \Phi \frac{F_{a_u} - F_{a_l}}{2} \tag{6.85}$$

The corresponding stresses are

$$\sigma_{amp} = K_f \frac{F_{amp}}{A_s} \tag{6.86}$$

$$\sigma_{mean} = K_{fm} \frac{F_{mean}}{A_s} \tag{6.87}$$

Estimate for the stress concentration factors for bolted connection are ([5])

ISO class ≤ 5.8 ⇒ Kf = 2.2(Rolled thread)

K_f = 2.8(Cut thread)

K_f = 2.1 (Fillet under head)

ISO class ≥ 6.6 ⇒ K_f = 3.0(Rolled thread)

K_f = 3.8(Cut thread)

K_f = 2.3(Fillet under head)

K_{fm} = 1.0

The mean and alternating stress can then be used in a modified Goodmann diagram, see Chapter 4.

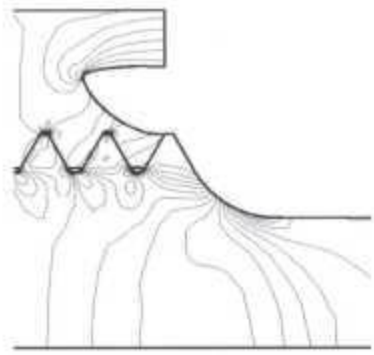
6.10 Design modification/optimization

Improvement in fatigue life of bolts can be achieved in three principally different ways

- Improving the joint stiffness factor by minimizing the bolt stiffness or/and maximizing the clamped material stiffness
- Improving the load distribution along the thread, by design changes made to the nut
- Minimizing the stress concentration factor in the bolt by applying shape optimization to the bolt design

The first bullet is a matter of practical design and selecting appropriate bolt nominal diameter relative to the thickness of the material clamped between the bolt head and the nut. Improvement of the joint stiffness factor is discussed in e.g. [8], The second bullet deals with the practical problem that for a traditional thread design the load is not evenly distributed along the thread, and the first turn of the thread can take up to 50% of the total load. To improve this point design changes must be made e.g. changing the design of the nut.

In [6] it is shown how the maximum stress of a M12 bolt/nut connection can be reduced by 34%. This reduction is achieved by design changes made both to the nut and the shank of the bolt, i.e. the third bullet. The optimized design is shown in Figure 6.22



[billedtekst start]**Figure 6.22:** Von Mises stress contour plot, zoom of the first 2 thread roots and the shank fillet of optimized design. [6][billedtekst slut]

6.11 Nomenclature

a_c	mm	Thread clearance between bolt and nut for trapezoidal thread
d	mm	Bolt nominal diameter
d_a	mm	Outer diameter of member plates
d_c	mm	Diameter contact area between member plates
d_h	mm	Hole diameter in plates
d_w	mm	Bolt head diameter or washer diameter
d_2	mm	Pitch diameter of bolt thread (mean flank diameter)
d_3	mm	Minor diameter of bolt thread
f_b	mm/N	Flexibility of bolt
F_h	mm/N	Flexibility of bolt head
f_m	mm/N	Flexibility of member plates
f_n	mm/N	Flexibility of nut
f_{te}	mm/N	Flexibility of engaged thread
f_{to}	mm/N	Flexibility of open part thread root
f_{tn}	mm/N	Flexibility of engaged thread and nut

$f_{1..n}$	mm/N	Flexibility of individual bolt shank segments
h	mm	Thread profile height
k_b	N/mm	Stiffness of the bolt
k_m	N/mm	Stiffness of the member plates
l_i	mm	Length of shank section i
l_m	mm	Thickness of member plates
l_t	mm	Length of open thread section
l_h	mm	Assumed length of bolt head
l_{tn}	mm	Assumed length of nut
p	mm	Pitch
p_h	mm	Lead
r	mm	Axis symmetric coordinate
r_n	mm	Mean radius of contact between nut and plate (head and plate)
z	mm	Axis symmetric coordinate
A_i	mm^2	Cross section area of bolt section

A_s	mm²	Stressed cross section of bolt thread
C_i	–	Stiffness constants
D	mm	Major diameter of nut thread
E_b	N/mm²	Modules of elasticity of bolt material
E_m	N/mm²	Modules of elasticity of plate material
F_a	N	External working load
F_{amp}	N	Amplitude of alternating force
F_{al}	N	Minimum working load
F_{au}	N	Maximum working load
F_b	N	Force in bolt
F_{ba}	N	Part of working load in the bolt
F_c	N	Clamping force
F_m	N	Force in the plates
F_{mean}	N	Mean value of alternating force
F_{ma}	N	Part of working load in the plates
F_n	N	Force acting normal to thread surface
F_p	N	Assembly force
G	–	Scaling factor
K_f	–	Fatigue stress concentration factor
K_{fm}	–	Fatigue stress concentration factor (on mean value)
P	N	Preload in the bolt
F_z	N	The setting force, the reduction in clamping force due to set between components

F_{zpl}	N	Preload reduction due to plastic deflection
T_l	Nm	Loosening torque
T_t	Nm	Tightening torque
α	rad	Flanks angle
α_a	rad	Assembly force ratio (tightening factor)
β	rad	The pitch angle at d_2
δ_b	mm	Elongation of bolt
δ_{bp}	mm	Assembly elongation of bolt
δ_m	mm	Compression of the member plates
δ_{mp}	mm	Assembly compression of the plates
ϵ	–	Non dimensional bolt head width parameter
η	–	Non dimensional bolt hole parameter
γ	–	Non dimensional bolt head height parameter
μ	–	The coefficient of friction in the thread
μ_n	–	The coefficient of friction between nut and plate (head and plate)
ϕ	–	Force ratio
$\sigma_c(r)$	N/mm^2	Surface pressure distribution between bolt head and member plates
σ_g	N/mm^2	Maximum surface pressure between bolt head and member plates
σ_p	N/mm^2	Proof strength of bolt
σ_{ut}	N/mm^2	Ultimate tensile stress
σ_y	N/mm^2	Yield stress
θ	rad	Angle between thread force components in space

φ	rad	Half of top angle of stress frustum
ζ	–	Non dimensional washer height parameter

6.12 References

[1] DUBBEL:. *Handbook of mechanical engineering, 24. Auft*. Springer Verlags, 1994.

[2] H. H. Gould and B. B. Mikic. Areas of contact and pressure distribution in bolted joints. *J Eng Ind Trans ASME*, 94 Ser B(312):864–870, 1972.

[3] Y. Ito, J. Toyoda, and S. Nagata. Interface pressure distribution in a bolt-flange assembly. *J Mech Des Trans ASME*, 101(2):330–337, 1979.

[4] N. Motosh. Determination of joint stiffness in bolted connections. *J Eng Ind Trans ASME*, 98 Ser B(3):858–861, 1976.

[5] R. L. Norton. *Machine design, an integrated approach, fifth edition.* Prentice-Hall Inc., Upper Saddle River, N.J. 07458, 2014.

[6] N. L. Pedersen. Overall bolt stress optimization. *Journal of Strain Analysis for Engineering Design*, 48(3): 155–165, 2013.

[7] N. L. Pedersen and P. Pedersen. On prestress stiffness analysis of bolt-plate contact assemblies. *Archive of Applied Mechanics*, 78(2):75–88, 2008.

[8] N. L. Pedersen and P. Pedersen. Stiffness analysis and improvement of bolt-plate contact assemblies. *Mechanics Based Design of Structures and Machines*, 36(1):47–66, 2008.

[9] F. Rötscher. *Maschienenelemente.* Springer, Berlin, 1927.

[10] J. E. Shigley and C. R. Michke. *Mechanical Engineering Design 7th ed.* McGraw Hill, Singapore, 2004.

[11] VDI 2230. *Systematische berechnung hochbeanspruchter schraubenverbindungen*. Berlin und Köln: Beuth-Verlag, 1986.

[12] VDI 2230 Blatt 1. *Systematische berechnung hochbeanspruchter schraubenverbindungen zylindrische einschraubenverbindungen, systematic calculation of high duty bolted joints. Joints with one cylindrical bolt.* Beuth-Verlag GmbH, 10772 Berlin, 2003.

[13] VDI 2230. *Systematische berechnung hochbeanspruchter schraubenverbindungen*. Berlin und Koln: Beuth-Verlag, 1977.

[14] J. Wileman, M. Choudhury, and I. Green. Computation of member stiffness in bolted connections. *Journal of Mechanical Design*, 113(4):432–t37, 1991.

Chapter 7 Couplings and universal joints

7.1 Introduction to couplings

Couplings (and clutches) are power transmitting machine elements. A torque is transmitted with a rotational speed from one shaft to another, placed along the same axis or nearly the same axis.

Couplings. are used to *transmit* torque permanently in aligned and non-aligned shafts and often in addition, in order to improve the *dynamic characteristics* of a drive mechanism and to *engage* the torque.

Clutches.[1] are classified into the two main groups related to the torque transmission principle: *positive* or *frictional.* Further on, clutches are divided into *closing* clutches and *opening* clutches. Finally, they can be described by the method of actuation (mechanical, pneumatic, hydraulic and electromagnetic). Special types of clutches are centrifugal clutches and overload clutches.

Brakes and one-way clutches can be seen as belonging to the second group.

7.2 Functional characteristics

As mentioned, couplings and clutches can be classified into the two main groups: couplings with constant connection between two shafts (permanent torsionally stiff couplings and permanent elastic couplings) and clutches where the shafts under certain conditions can rotate relatively to each other. To the last group belongs: frictional clutches, centrifugal clutches and hydrodynamic clutches. Another way of describing the functional characteristics is:

1. couplings for shaft elongation or shaft division
2. couplings for misaligned shafts or perception of angular deviation of shafts
3. clutches for man-operated engagement or disengagement
4. clutches for torque limitation
5. clutches activated by speed
6. clutches activated by differences in angular rotation

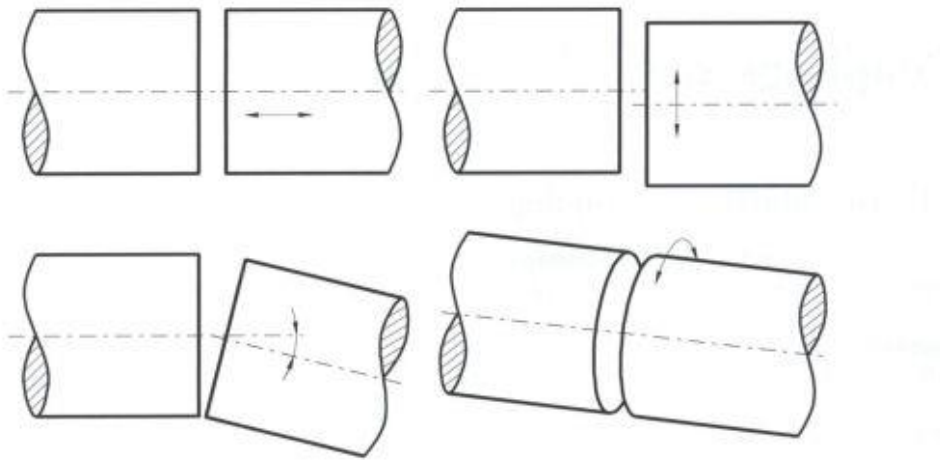
We will not be able to group all occurrences of couplings and clutches into these distinct functions as some occurrences will belong to more than one function group. Even though, when choosing a coupling, it is of great importance to analyze exactly which functional requirements are to be fulfilled.

---

[1] Only mechanical clutches are described here.

7.2.1 Shaft elongation or shaft division

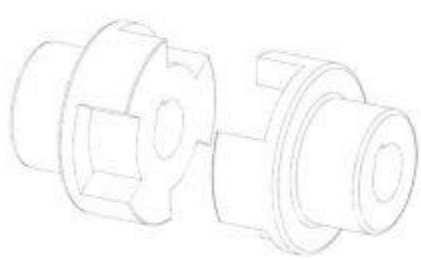
Long shafts may need to be divided into two or more parts due to assembly problems, e.g. if space is limited. Other reasons could be easier handling, transportation or manufacturing. Possible methods of assembling long shafts are by the use of *flange couplings,* Figure 7.12, *double tooth couplings,* Figure 7.3, *split muff couplings,* Figure 7.13 or special types of similar shape as the split muff coupling, but based on the principle of conical seat and hydraulic pressure fit.



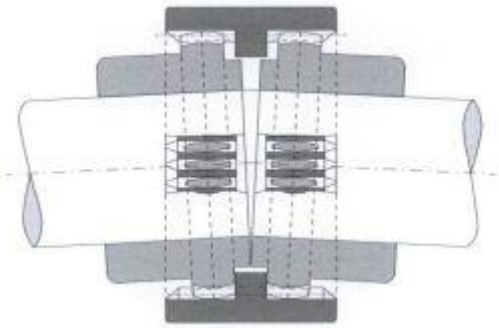
[billedtekst start]**Figure 7.1:** Possible functional demands for couplings.[billedtekst slut]

7.2.2 Misaligned shafts or angular deviation

Two shafts are not always in the ideal position relative to each other. To prevent axial forces or radially directed forces from one shaft to the other, special types of couplings can be used. Axial forces can be caused by thermal expansion. To prevent damaging axial forces the simple *positive contact coupling,* Figure 7.2 or a more advanced type, the *double tooth coupling,* Figure 7.3, can be used.



[billedtekst start]**Figure 7.2:** Positive contact coupling, only for axial displacement.[billedtekst slut]



[billedtekst start]**Figure 7.3:** Double tooth coupling. (Bovex). Suitable for small axial, radial and angular displacements.[billedtekst slut]

Radial forces can be caused by shaft deflection or misalignment of the shafts. Misalignment often results in large radial forces damaging the nearby bearings. Elastic couplings will normally not be able to solve problems related to misalignment.

Only a few coupling types can handle these problems. One is the *double tooth coupling,* eventually equipped with an elongation tube (spacer) bolted between the two outer parts of the coupling. (See Figure 7.3).

For even bigger angular deviation between the axes, a *Cardan shaft* (see Figure 7.22) has to be used. Examples are in rolling mills, milling machines and in the transmission to the wheels in automobiles.

Another type of angular deviation is the torsional deviation as obtained in elastic couplings used for reducing torque fluctuations or vibrations, thereby preventing damage to the machinery. For almost all elastic couplings (see page 162) the torsio-elastic function is the absolutely dominating and this type of couplings requires careful alignment of the shafts.

7.2.3 Man-operated engagement or disengagement

The simplest type is the *positive (interlocking) clutch.* Positive clutches can be engaged only when the shafts are at a standstill or running at the same speed. Disengagement can however take place when running even under torque, provided that the disconnecting force can be established. The most well-known type of man-operated clutch is the *friction clutch.* This type of clutch can be operated mechanically, electromechanically, or by the use of pneumatic or hydraulic force.

7.2.4 Torque-sensitive clutches

Torque-Sensitive Clutches, also called overload clutches, are safety clutches which protect machinery from damage as they do not exceed a predetermined torque load. The way of operation differs depending on the application. Examples are:

- *constant slip torque.* Typical a friction clutch that has to maintain a high torque during acceleration. Suitable as start-up clutches for accelerating large masses and for limiting short duration peak torques during operation. Constant surveillance is needed to prevent overheating. See Figure 7.4.
- *slip with pulsating torque.* See Figure 7.5. A radial pin clutch for PTO drive shaft (PTO: Power Take Off). These clutches are radially-acting ratchet clutches. Upon creation of an overload condition, torque is limited and transmitted in a pulsating manner during the slipping action that at the same time, provides audible warning.

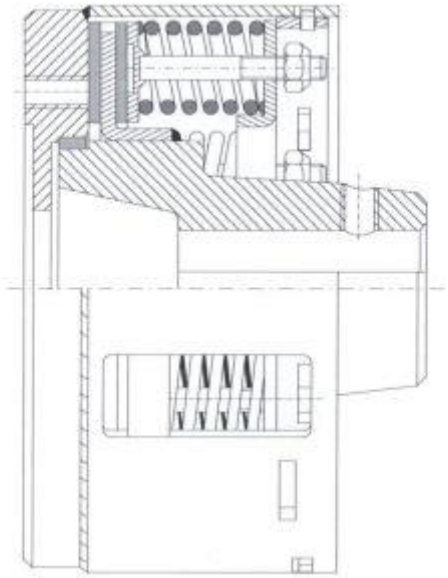
7.2.5 Speed-sensitive clutches

These are clutches that allow smooth starting for the driving machines, electric motors or

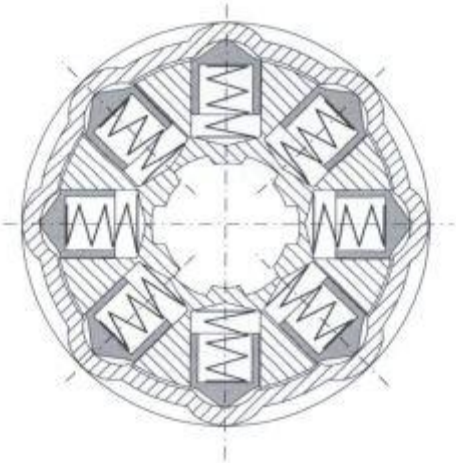
combustion engines, to accelerate first of all and drive the machine next. *A soft starting clutch* makes it possible to design for a reduced motor size or even reduced electricity supply for a machine that has to be accelerated to a high rotational speed with a high moment of inertia or load torque.

Different types are:

- Centrifugal clutches with segments.
- *Centrifugal clutches* with powder or granulate as torque transmitting "fluid".
- *Speed-Sensitive Clutches* based on the hydrodynamic principle.



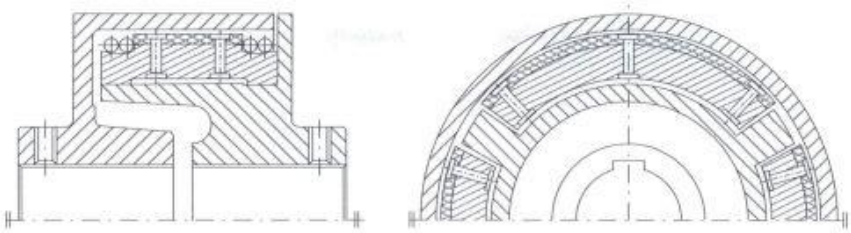
[billedtekst start]**Figure 7.4:** Overload clutch for PTO drive shaft.[billedtekst slut]



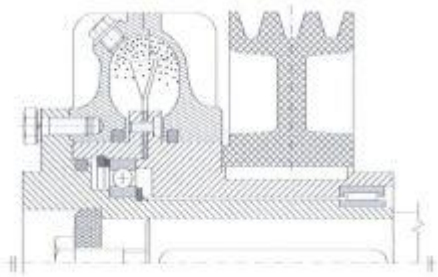
[billedtekst start]**Figure 7.5:** Radial pin clutch for PTO drive shaft.[billedtekst slut]

An additional function for the speed-sensitive clutches is of course that they are torque-sensitive.

Even though the centrifugal clutch is normally recognized as a soft starting clutch, it has been seen as a simple "automatic" clutch for small lawn mowers, where the reel function seems to be an on-off function caused by the small moments of inertia in the lawn mower.



[billedtekst start]**Figure 7.6:** Centrifugal clutch.[billedtekst slut]

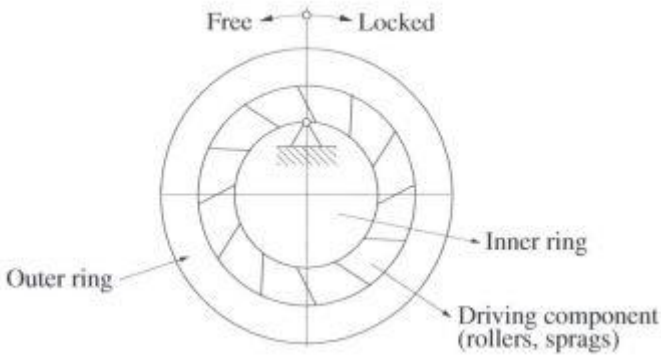


[billedtekst start]**Figure 7.7:** Centrifugal clutch.[billedtekst slut]

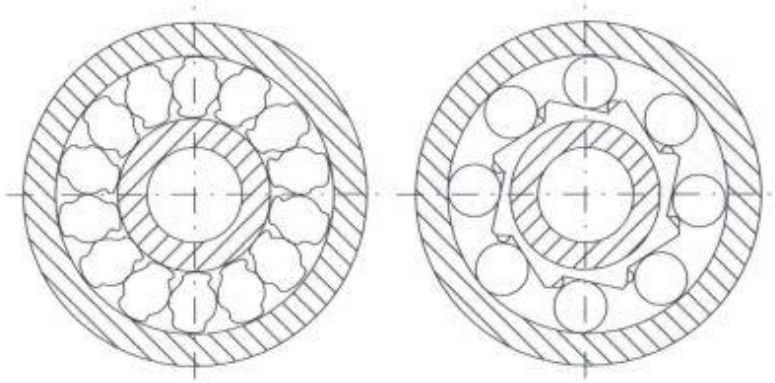
7.2.6 Directional (one-way) clutches, overrun clutches

For this type of clutches the torque can only be transmitted from one part of the clutch to the counter part by relative rotation in one direction (locked condition). In the other direction, there can be no torque transmission at all (freewheeling condition). Examples of applications are:

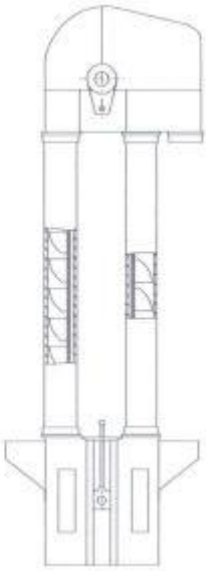
- *overrun clutch* (for bicycle hubs, starting motor drives).
- *return stop* (for conveyor belts, centrifugal pumps, automatic gearboxes for motor vehicles).
- *step-by-step freewheel* (for shaping machines, feed mechanisms, ratchet mechanisms).



[billedtekst start]**Figure 7.8:** The freewheeling principle.[billedtekst slut]



[billedtekst start]**Figure 7.9:** Free wheels: one type is the *sprag freewheel* which has outer and inner rings with cylindrical races. Arranged in between are the individually sprung sprags. The drive mode is free from slip. Various sprag shapes make it possible to design different types suitable for: high torque, high indexing accuracy, or non-constant overrunning operation. Another type is the *roller ramp freewheel*, where either the inner or the outer ring has roller ramps, the other ring has a cylindrical race. The individually sprung rollers are arranged in between.[billedtekst slut]

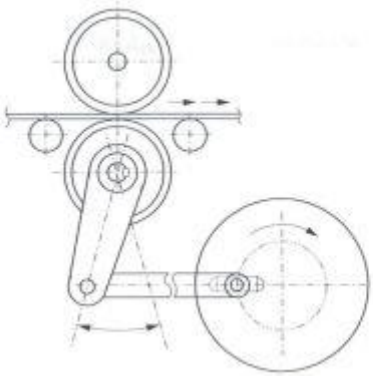


[billedtekst start]**Figure 7.10:** A freewheel clutch used as backstop in inclined conveyors or elevators to prevent the load from running back when the motor is switched off or in case of power failure. Other applications are in pumps, blowers and ventilators, to prevent reverse running due to the pressure from the flow medium after switching off.[billedtekst slut]

7.3 Permanent torsionally stiff couplings

7.3.1 Rigid couplings

Rigid couplings require the two shafts to be exactly aligned through the coupling. In order to account for small misalignments, the bearings on each side of the coupling should be located so far away from

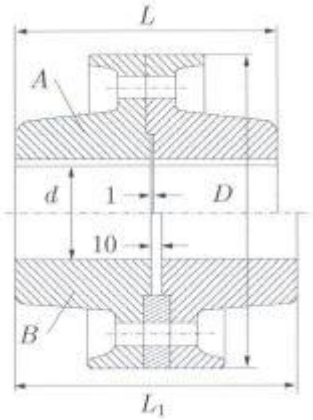


[billedtekst start]**Figure 7.11:** An indexing freewheel clutch used for producing fine feeds in packing-paper processing, printing or textile machines.[billedtekst slut]

the coupling that the shaft can deflect. This is in contrast to an elastic coupling, where it is required that both bearings should be placed near to the coupling.

Flange couplings are normally dimensioned to transmit the torque only with frictional locking provided by preloaded screws. The additional torque transmission capability due to shear stress in the bolts is neglected for safety reasons.

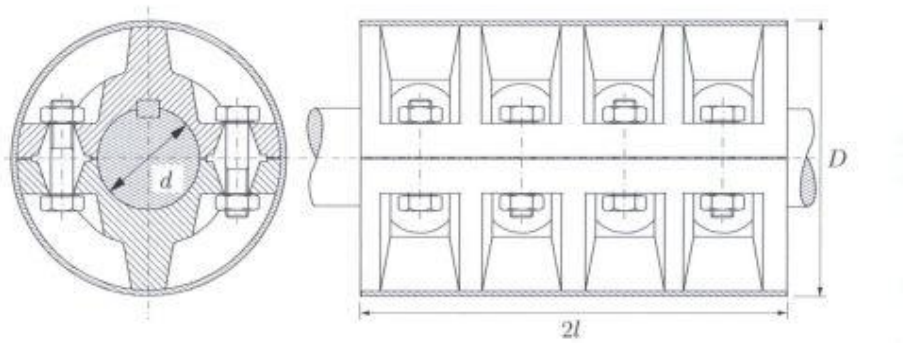
Alignment is secured by male and female part or by the use of a two-piece intermediate faceplate that allow for radial disconnection.



[billedtekst start]**Figure 7.12:** Flange coupling after DIN 116.[billedtekst slut]

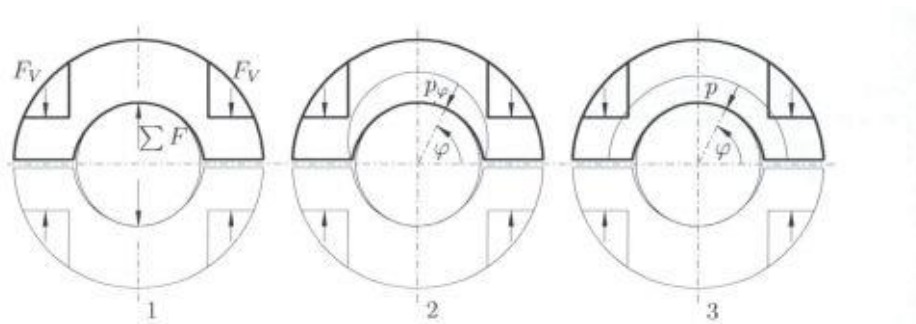
Split muff couplings (after DIN 115) may be used for small and medium range torques, but they are not suitable for alternating or intermittent loads. Before assembly the coupling must be bored to correct diameter usually with a thick piece of paper between the two halves to ensure sufficient pressure against the shafts after assembly.

For Split Muff Couplings transmitting the torque solely by friction, it is essential to use a proper calculation method. In the absence of exact information about manufacturing tolerances, which is the normal case, a safe calculation method has to be used. Add to this that the two shafts shall have exactly



[billedtekst start]**Figure 7.13:** Split Muff Coupling after DIN 115.[billedtekst slut]

the same diameter as a prerequisite for the shaft-hub calculation. Please note that the split muff coupling basically consists of two shaft-hub connections.



[billedtekst start]**Figure 7.14:** Three different premises for the calculation. 1: Line contact. 2: sine pressure contact. 3: constant pressure.[billedtekst slut]

1. Line contact between muff and shaft

The total force from the bolts on one shaft and one side of the coupling is F. According to the example in Figure 7.13 and Figure 7.14 the force is

$$\sum F = 4F_V \tag{7.1}$$

The permissible maximum torque is calculated to

$$T_s = 2\mu r \sum F = \mu d \sum F \tag{7.2}$$

2. Sine shaped pressure distribution

$$p_\varphi = p_0 \sin\varphi \tag{7.3}$$

$$dF = p_\varphi r d\varphi \cdot l \tag{7.4}$$

See detailed methods of calculation for permissible maximum torque T_s in Chapter 9, (Drum brakes).

$$T_s = \frac{4}{\pi}\mu d \sum F \tag{7.5}$$

3. Constant pressure between muff and shaft.

$$p = \text{constant} \tag{7.6}$$

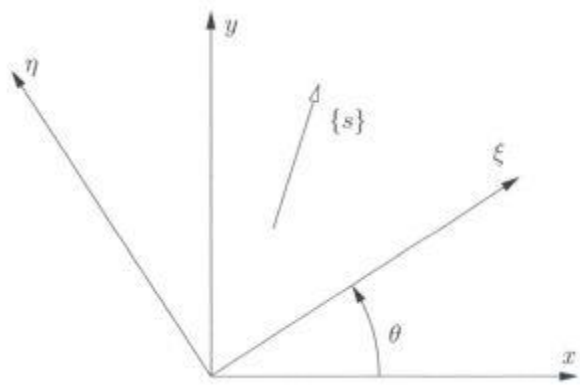
$$dF = prd\varphi \cdot l \tag{7.7}$$

$$\begin{aligned} \sum F &= 2\int_0^{\pi/2} dF \cdot \sin\varphi \\ &= pld \cdot [-\cos\varphi]_0^{\pi/2} \\ &= pld \end{aligned} \tag{7.8}$$

$$\begin{aligned} T_s &= \mu p(\pi dl)r \\ &= \frac{\pi}{2}\mu d \sum F \end{aligned} \tag{7.9}$$

7.3.2 Universal joints and other special joints

The universal joint is also termed a Cardan joint or a hook joint. In order to define a universal joint we use the methodology from multibody dynamics as it may be found in [7]. We define first the concept of transformation matrix in 2D.



[billedtekst start]**Figure 7.15:** Global and local coordinate systems, and a geometric vector.[billedtekst slut]

In Figure 7.15 the global coordinate system $x - y$ and a local coordinate system ξ–η are defined, the origo of the two coordinate systems coincides and the local coordinate system is rotated the angle θ counter clockwise. The angle is generally measured positive counter clockwise and by the arrow we define from which axis we measure the angle. The figure also show a geometric vector $\{s\}$. We assume that the geometric vector is fixed to the local coordinate system and as such will rotate with this coordinate system. Seen from the global coordinate system the vector is termed $\{s\}$ and from the local coordinate system $\{s'\}$ (a vector described in local coordinates is indicated by a prime). From the geometry we find

$$\{s\} = [A]\{s'\} \tag{7.10}$$

where $[A]$ is the transformation matrix that transform from local to global coordinates. In 2D this transformation matrix is given by

$$[A] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{7.11}$$

From (7.10) we directly find the time derivative

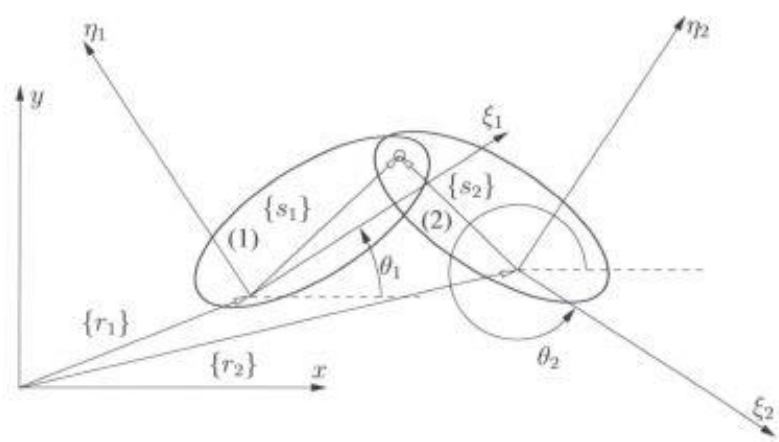
$$\frac{d\{s\}}{dt} = [\dot{A}]\{s'\} + [A]\{\dot{s}'\} = [\dot{A}]\{s'\} \tag{7.12}$$

where the dot indicates differentiation with respect to time as usual. The last simplification is due to the assumption that the vector is fixed in the local coordinate system $(\{\dot{s}'\} = \{0,0\}^T)$. Using the notation $w = d\theta/dt$ we find that

$$\frac{d\{s\}}{dt} = \omega \begin{bmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{bmatrix} \{s'\} \tag{7.13}$$

In 3D the transformation matrix becomes more involved, but for the case of a universal joint we may make some simplifications as shown later.

The next concept is constraints. A constraint is something that the system must fulfil at all times. The simplest form of joint is a revolute joint shown in Figure 7.16.



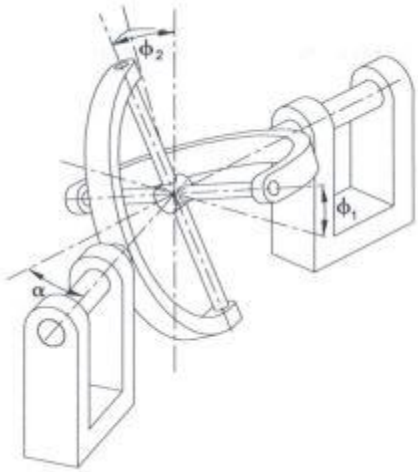
[billedtekst start]**Figure 7.16:** Two bodies constraint by a revolute joint.[billedtekst slut]

The constraints that specifies the connection of the two bodies is given by

$$\{r_1\} + \{s_1\} = \{r_2\} + \{s_2\} \tag{7.14}$$

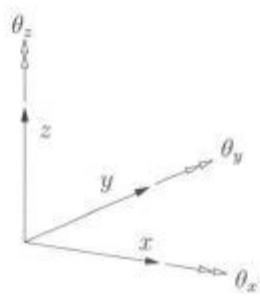
or by using the transformation matrix for each body

$$\{r_1\} + [A_1]\{s_1'\} - \{r_2\} - [A_2]\{s_2'\} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{7.15}$$



[billedtekst start]**Figure 7.17:** Universal joint.[billedtekst slut]

The purpose is to define the universal joint shown in Figure 7.17. The universal joint permits angles up to $2\pi/9$ between the two rotation axes, but transforms a uniform angular velocity $w_1 = d\phi_1/dt$ into an angular velocity $w_2 = d\phi_2/dt$ depending on the bend angle α and on the actual angle of rotation ϕ_1 of shaft 1 as we shall see later. As indicated the transformation matrix in 3D is more complicated, this is because three angles do not result in a unique position unless the order of rotation is given. In some cases simplifications are however possible.



[billedtekst start]**Figure 7.18:** 3D coordinate system and definition of rotation angles.[billedtekst slut]

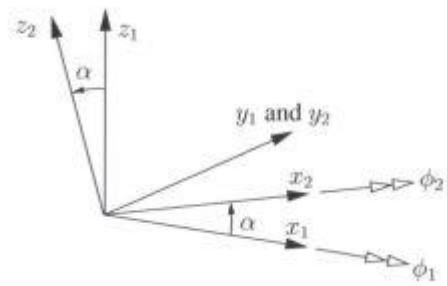
The transformation matrix in 3D when we only rotate around one axis as shown in Figure 7.18 is given respectively by

$$[A]_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{bmatrix} \qquad [A]_y = \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{bmatrix} \qquad (7.16)$$

$$[A]_z = \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A universal joint constraint four d.o.f. (degrees of freedom). The first three constraints is similar to a spherical joint in that the center of the cross (see Figure 7.17) is identical - seen from the local coordinate

system on each body. Mathematically the constraint of a spherical joint is identical to the definition of the revolute joint in 2D presented in (7.15). We aim here to find the connection between the rotation of body one ϕ_1 and the rotation of body two ϕ_2 (see Figure 7.17). To find this connection we first place the local coordinate systems as shown in Figure 7.19.



[billedtekst start]**Figure 7.19:** lacement of local coordinate systems in a universal joint.[billedtekst slut]

As shown the axis of rotation is x_1 for body one and x_2 for body two. Also shown is the angle α that defines the angle between the two rotation axes. The origo of the local coordinate systems is placed in the center of the cross of the universal joint. We may then choose to let the cross be defined by the y_1 and z_2 axes or the y_2 and z_1 axes. In any of the cases the transformation matrices are found directly from (7.16).

$$[A_1] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_1 & -\sin\phi_1 \\ 0 & \sin\phi_1 & \cos\phi_1 \end{bmatrix} \tag{7.17}$$

$$[A_2] = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi_2 & -\sin\phi_2 \\ 0 & \sin\phi_2 & \cos\phi_2 \end{bmatrix} \tag{7.18}$$

Now lets assume that the cross is defined by the y_1 and z_2 axes this gives us two local vectors aligned with these two directions

$$\{s_1'\} = \{0,1,0\}^T \qquad \{s_2'\} = \{0,0,1\}^T \tag{7.19}$$

The constraint that defines the relation between the angle ϕ_1 and ϕ_2 is then given by the orthogonality

$$\{s_1\}^T\{s_2\} = \{s_1'\}^T[A_1]^T[A_2]\{s_2'\} = 0 \tag{7.20}$$

By using $[A_1]$ and $[A_2]$ we can simplify (7.20) and find

$$-\cos\phi_1 \sin\phi_2 + \cos\alpha \sin\phi_1 \cos\phi_2 = 0 \Rightarrow$$

$$\tan\phi_2 - \cos\alpha \tan\phi_1 = 0 \tag{7.21}$$

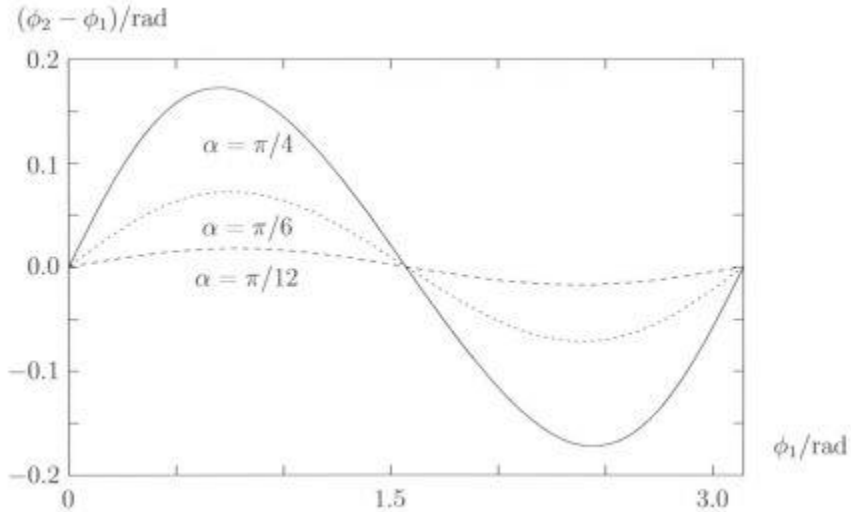
If we had made the other choice of axes to define the cross, i.e.,

$$\{s_1'\} = \{0, 0, 1\}^T \qquad \{s_2'\} = \{0, 1, 0\}^T \tag{7.22}$$

The resulting equation is

$$\tan \phi_1 - \cos \alpha \tan \phi_2 = 0 \tag{7.23}$$

from (7.21) and (7.23) we see that the position of the cross is important.



[billedtekst start]**Figure 7.20:** The lack of output angle ϕ_2 relative to input angle ϕ_1.[billedtekst slut]

Figure 7.20 shows a plot of

$$\phi_2 - \phi_1 = \arctan\left(\frac{\tan \phi_1}{\cos \alpha}\right) - \phi_1 \tag{7.24}$$

as a function of the input angle ϕ_1 based on (7.23). It is directly seen that the two angles ϕ_1 and ϕ_2 are not identical, except for each quarter of a full turn. Using this in (7.21) we find

$$\tan(\phi_2 + \pi/2) - \cos \alpha \tan(\phi_1 + \pi/2) = 0 \Rightarrow$$

$$\tan(\phi_1) - \cos \alpha \tan(\phi_2) = 0 \tag{7.25}$$

i.e. we see that (7.21) and (7.23) are principally identical and the difference expresses the position of the cross.

The next important question is the output angular speed w_2 as a function of the input angular speed w_1. This is found directly by differentiating (7.23) with respect to time. We find

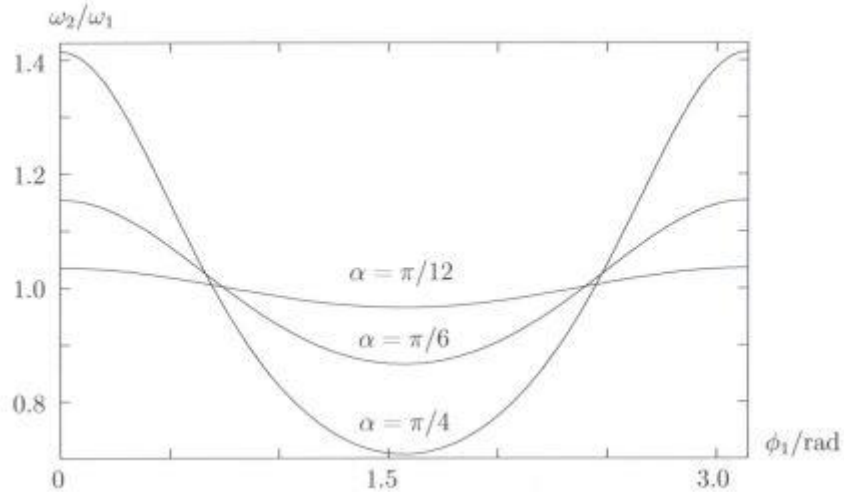
$$\omega_1 \frac{1}{\cos^2 \phi_1} - \omega_2 \frac{1}{\cos^2 \phi_2} \cos \alpha = 0 \tag{7.26}$$

which we may rewrite to

$$\frac{\omega_2}{\omega_1} = \frac{\cos^2 \phi_2}{\cos^2 \phi_1 \cos \alpha} = \frac{\cos \alpha}{1 - \cos^2 \phi_1 \sin^2 \alpha} \tag{7.27}$$

In Figure 7.21 the ratio w_2/w_1 is shown as a function of the input angle. From (7.27) we find that then maximum and minimum value of the output shaft angular speed is

$$(w_2)_{max} = w_1 \cos\alpha \qquad (w_2)_{min} = w_1/\cos\alpha \qquad (7.28)$$



[billedtekst start]**Figure 7.21:** The ratio between output angular speed and input angle.[billedtekst slut]

The next quantity we calculate is the angular acceleration of the output shaft under the assumption that the angular acceleration of the input shaft is zero, i.e., that $\dot{w}_1 = 0$. We find the acceleration by simple differentiation of (7.27) with respect to time

$$\dot{\omega}_2 = \frac{-\cos\alpha(2\cos\phi_1\sin^2\alpha\sin\phi_1)\omega_1}{(1-\cos^2\phi_1\sin^2\alpha)^2}\omega_1 \qquad (7.29)$$

and we may rewrite this

$$\frac{\dot{\omega}_2}{\omega_1^2} = -2\sin^2\alpha\cos\alpha\frac{\cos\phi_1\sin\phi_1}{(1-\cos^2\phi_1\sin^2\alpha)^2} = -2\sin^2\alpha\cos\alpha\frac{\tan\phi_1(1+\tan^2\phi_1)}{(\cos^2\alpha+\tan^2\phi_1)^2} \qquad (7.30)$$

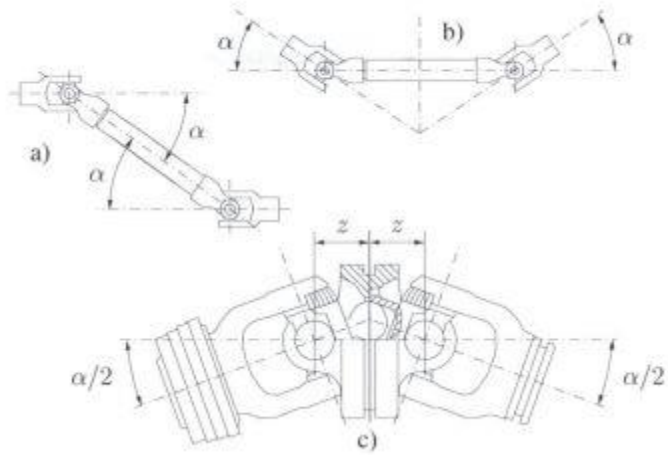
As it is seen from (7.30) and (7.27) we get fluctuations in both angular speed and angular acceleration of the output shaft although the input shaft have constant angular speed.

Due to the forces caused by acceleration and deceleration of the second shaft the designer has to keep the bend angle as low as possible. According to the company Walterscheid (See [5]) an increase in bend angle α from $\pi/36$ to $\pi/18$ will cause a decrease in lifetime to 50% of the original value.

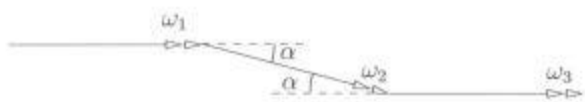
Two universal joints are normally (when possible) combined to what is called a *Cardan shaft*, see Figures 7.22 and 7.23 . In that way the fluctuation in speed is neutralized and the "only" problem is that the intermediate shaft is fluctuating in speed for which reason the moment of inertia has to be kept as low as possible. As seen in Figures 7.22 and 7.23 it is a precondition for proper operation that the two bend angles are of the same size.

It is also important that the two crosses are rotated $\pi/2$ compared to each other, this can

e.g. be seen in Figure 7.22(a), where we notice the different placement of the cross of the two universal joints. This



[billedtekst start]**Figure 7.22:** Cardan shaft.[billedtekst slut]



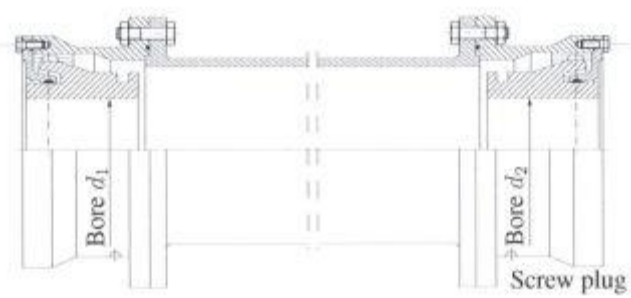
[billedtekst start]**Figure 7.23:** Definition of the angular speeds in a Cardan shaft.[billedtekst slut]

corresponds to the center shaft having the two axes in the to universal joints aligned. If this is not done you will not remove the fluctuations, but increase them further. That the fluctuations are removed follows directly from (7.21), (7.23) and Figure 7.23. We find that

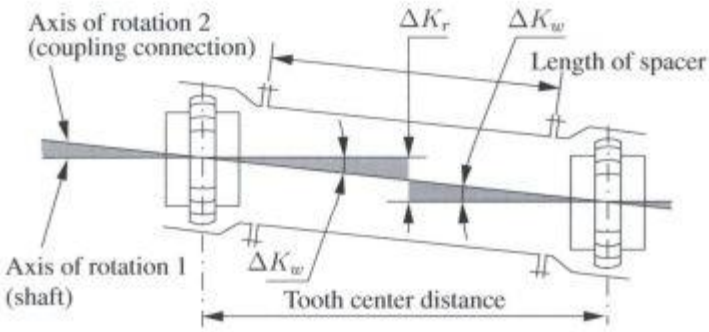
$\tan \phi_2 = \tan \phi_1 / \cos \alpha$

$\tan \phi_3 = \tan \phi_2 \cos \alpha \Rightarrow \tan \phi_3 = \tan \phi_1 \Rightarrow \phi_3 = \phi_1 \Rightarrow w_3 = w1 \Rightarrow \dot{w}_3 = \dot{w}_1$

Other torsionally stiff "self-aligning" couplings are the *curved tooth couplings,* which allows for some axial and angular deviations. They have to be lubricated with grease or oil to avoid wear.



[billedtekst start]**Figure 7.24:** Curved tooth coupling.[billedtekst slut]



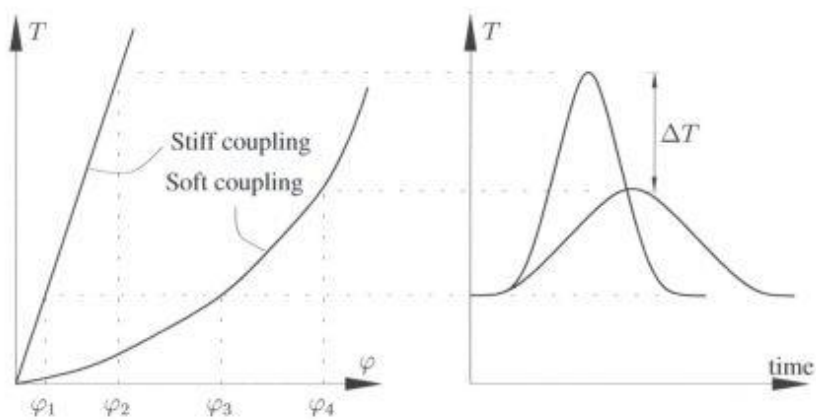
[billedtekst start]**Figure 7.25:** Schematic depiction of misalignment. Radial misalignment ΔK_r dependent on length of spacer. Angular misalignment ΔK_w up to $\pi/120$ per coupling half.[billedtekst slut]

7.4 Permanent elastic couplings

7.4.1 General purpose

Elastic couplings transmit rotary motion without slip and are primary used for reducing torque fluctuations or vibrations, thereby preventing damage to machinery. At least two goals are to be achieved:

- *Reducing peak loads* by the elastic accumulating effect of the transmitting devices. A large twist angle reduces peak torque T at the start. See Figure 7.26. The primary part twist the angle $\Delta\varphi = \varphi_2 - \varphi_1$ relative to the secondary part. By using a softer coupling this $\Delta\varphi$ increase. Stretching the impulse over a larger time span may reduce the peak torque to an acceptable level.
- *Avoiding torsional frequencies of resonance.* That is to choose or design the elastic coupling so that the operating speed range is outside the resonance area. (Normally comfortably higher than the resonance speed).

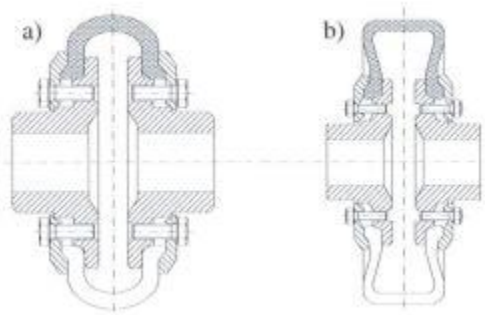


[billedtekst start]**Figure 7.26:** For the same amount of absorbed energy the soft coupling results in the lowest peak moment.[billedtekst slut]

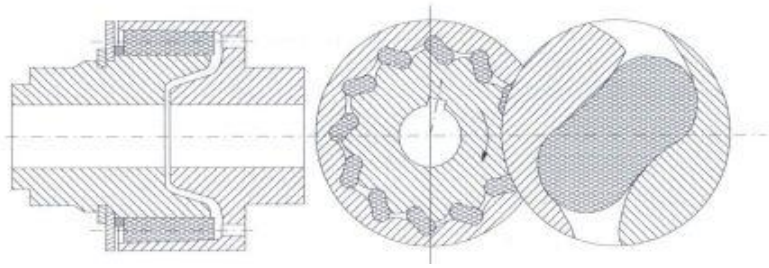
The reduction in torsional load ΔT for the coupling will be even bigger if a fraction of the energy is transformed into heat. This is called damping. Damping in the coupling for the most part relies on the material damping of the elastomers used and on the friction coefficient in the contact surfaces. The best damping effect is normally achieved by using a soft coupling.

By periodic pulsating torsional load (from combustion engines) it might be necessary to calculate the amplitude from the forced vibration. For a simple two-mass system this can be done when mass moment of inertia on both sides of the coupling, and characteristics for the coupling as the dynamic stiffness and the relative damping coefficient are available.

A coupling normally has to be chosen so that the critical rotational speed w_C is substantially lower than the working speed w as the amplification factor is strongly dependent on the ratio w/w_C. A soft coupling has a low critical speed. Often it is an advantage to choose a coupling with a progressive characteristic in order to limit the angular fluctuations, between the two coupling halves.



[billedtekst start]**Figure 7.27:** Permanent elastic coupling. The elastic element consists of a reinforced rubber "tire" with a radial cut. The b)-type allows bigger axial misalignment than the a)-type.[billedtekst slut]



[billedtekst start]**Figure 7.28:** Compact torsio-elastic coupling. (Up to ±Π/18). Type: Rollastic from the company: SEW-Eurodrive GmbH & Co KG. The shape of the elastic elements is cylindrical when the coupling is not loaded.[billedtekst slut]

7.4.2 Selection procedures

The selection of a suitable coupling is normally done in two steps, first the type of elastic coupling should be chosen and next the coupling size has to be determined. Figure 7.29 shows a

system with coupling schematically.

Type selection. Simple speed drive mechanisms (including electric motors, centrifugal pumps and fans) are coupled to elastomer couplings with medium elasticity ($\Delta\varphi < \pi/36$) in order to compensate for starting impulse and minor shaft misalignment. Highly variable-speed drive mechanisms (piston engines, presses etc.) or the transfer of resonance speed require highly elastic couplings ($\Delta\varphi < \pi/36$ to $\Pi/6$).

Size selection. This can partly be done on the basis of information from the manufacturers and partly on the basis of calculations described in DIN 740 Part 2 [4]. The permissible *nominal torque* (T_{CN}) of the coupling must be at least equal to the nominal torque (T_{LN}) on the driven machine (or to the nominal torque (T_{AN}) on the driving motor).

$$T_{CN} \geq T_{LN} S_\theta \tag{7.31}$$

Where

T_{LN} the nominal torque for the driven machine

S_θ coefficient that makes allowance for the decrease in strength of elastic rubber material when exposed to heat. $S_\theta = 1$ for steel

Table 7.1: Temperature coefficient S_θ.

Material for elastic elements	–20... + 30°C	+30... + 40° C	+40... + G0°C	+60... + 80°C
Natural rubber	1.0	1.1	1.4	1.6
Polyurethan	1.0	1.2	1.5	not allowed
Aerylnitril -Butadien	1.0	1.0	1.0	1.2

The permissible *maximum torque* (T_{Cmax}) for the coupling must be at least equal to the peak torque (T_{AS}) or (T_{LS}) that occur in operation as a result of torsional vibration on the drive side and load side taking into account the mass moments of inertia I_A and I_L, the impulse coefficients S_A, S_L, the frequencies of starts and the temperature coefficient S_θ.

$$T_{Cmax} \geq T_S S_z S_\theta \tag{7.32}$$

where

T_S maximum torque at coupling, $T_S = T_{AS}\frac{1}{m+1}S_A$ or $T_S = T_{LS}\frac{m}{m+1}S_L$

T_{AS} maximum input torque from motor

T_{LS} maximum load torque

m Ratio of drive end to driven end mass moment of inertia, $m = \frac{I_A}{I_L}$

S_A , S_L impulse coefficient (dynamic factor)

S_z starting coefficient, S_z = 1 for up to 120 starts per hour, S_z = 1.3 for 120 to 240 starts per hour

Table 7.2: Impulse coefficients S_A, S_L.

Impulse Coefficient	**Driving Machine**				
Driven Machine	E-motor	Combustion Engine, with more than 4 Cylinders	Combustion Engine		
			3 Cyl.	2 Cyl.	1 Cyl.
Generators	1.5 .. 1.7	1.7.. 1.9	2.0 .. 2.2	2.3 .. 2.5	2.7 .. 2.9
Elevators	1.6.. 1.8	1.9 .. 2.1	2.2 .. 2.4	2.5 .. 2.7	2.9 .. 3.1
Cranes	1.8 .. 2.0	2.1 .. 2.3	2.4.. 2.6	2.7 .. 2.9	3.1 .. 3.3
Piston pumps and compressors with flywheel Wheel	2.1 .. 2.4	2.4.. 2.7	2.7 .. 3.0	3.1 .. 3.4	3.5 .. 3.8

In case of a *continuous periodic torque fluctuation* from the drive system or from the driven system the fatigue torque limit T_{CW} for the coupling must not be exceeded. These (rough) calculations of coupling performance characteristics can be made under the condition that in terms of torsional vibrations, the machinery can be reduced to a linear two-mass vibration generating system.

$$T_{CW} \geq T_{Wi} S_\theta S_f \tag{7.33}$$

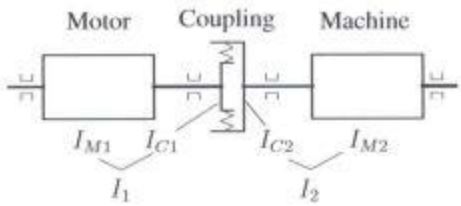
where

T_{Wi} amplitude of the *i*-th harmonic torque component, $T_{Wi} = T_{Ai}\frac{1}{m+1}V_{fi}$ or $T_{Wi} = T_{Li}\frac{m}{m+1}V_{fi}$

T_{Ai},T_{Li} amplitude of the external torque excitation of the *i*-th order acting on the drive end,. . . acting on the driven end

S_f frequency coefficient, $S_f = 1$ for $f \leq 10$Hz, $S_f = \sqrt{f/10}$ for $f > 10$Hz

V_{fi} amplification factor depending on the type of clutch $V_{fi} = \frac{1}{\sqrt{1-\frac{n}{n_R}}}$ see (7.37)



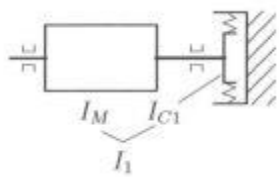
[billedtekst start]**Figure 7.29:** Schematic depiction of two-mass-system with elastic coupling.[billedtekst slut]

For an one-mass system the natural frequency is

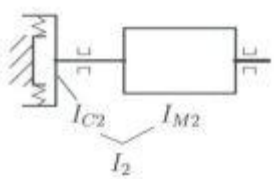
$$f_{e1} = \frac{1}{2\pi}\sqrt{\frac{R_{t,dyn}}{I_1}} \tag{7.34}$$

$$f_{e2} = \frac{1}{2\pi}\sqrt{\frac{R_{t,dyn}}{I_2}} \tag{7.35}$$

R_t,d_{yn} is the dynamic torsional stiffness of the coupling.



[billedtekst start]**Figure 7.30:** Schematic depiction of one-mass- system. Motor side.[billedtekst slut]



[billedtekst start]**Figure 7.31:** Schematic depiction of one-mass- system. Machine side.[billedtekst slut]

For a two-mass system, coupled together through an elastic coupling, the natural frequency is

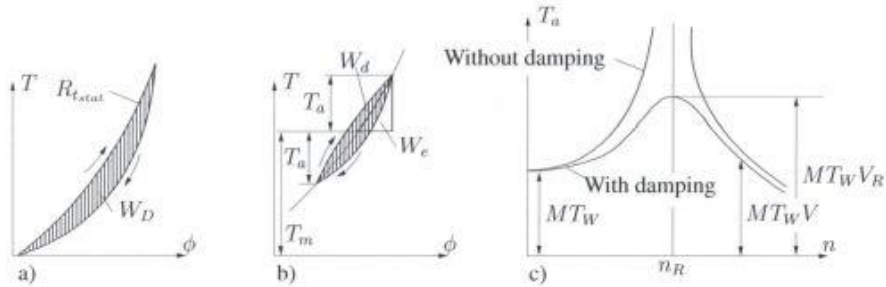
$$f_e = \frac{1}{2\pi}\sqrt{R_{t,dyn}\frac{I}{I_1 I_2}} \quad [\mathrm{Hz}] \tag{7.36}$$

where $I = I_1 + I_2$ is the total mass moment of inertia of the system. Exciting a two-mass system with impulses equal to the natural frequency will cause resonance and if the system is not sufficiently damped the amplitude will cause breakage. The critical rotational speed for a system is defined as

$$n_R = \frac{f_e}{i} \tag{7.37}$$

where i is the order number (= number of impulses per revolution) Ex. is $i = 2$ for a 4 cylinder 4 stroke motor, because of 2 impulses per revolution. (One impulse for every two revolution per cylinder).

7.4.3 Damping



[billedtekst start]**Figure 7.32:** Damped (torsional) vibrations in elastic couplings. The hatched area to the left represents the static work (W_D) absorbed in the coupling. The hatched area in the middle represents the dynamic heat loss (W_d) in the elastic elements. The dynamic load on the coupling can be regarded as a steady-state mean torque T_m superimposed with an alternating load with an amplitude T_a.[billedtekst slut]

A relative damping factor ψ is defined as the relation between W_d and W_e (W_e= the additional elastic energy from the + T_a-component).

$$\psi = \frac{W_d}{W_e} \tag{7.38}$$

At resonance speed (= the critical speed) the amplitude T_a will be increased with

$$V_R = \frac{2\pi}{\psi} \text{(= amplification factor at resonance speed)} \tag{7.39}$$

In case of an undamped torsional vibration, where $W_d = 0$ (no absorption of energy); $\psi = 0$ and $V_R \Rightarrow \infty$. Under normal running conditions the amplification factor can be calculated from

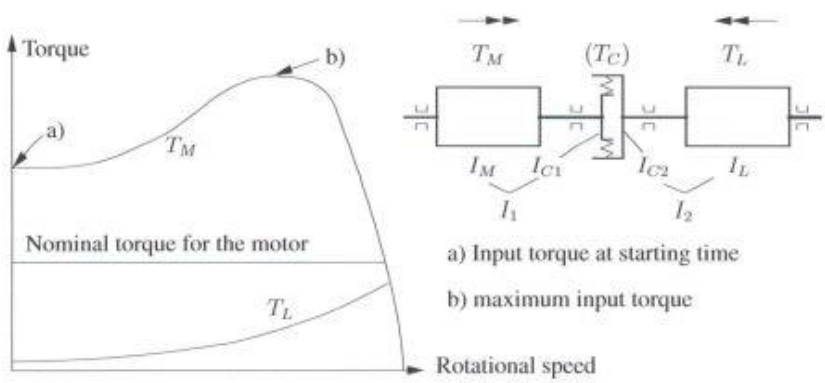
$$V_{fi} = \sqrt{\frac{1 + \frac{\psi^2}{4\pi^2}}{(1 - (\frac{f_i}{f_c})^2)^2 + \frac{\psi^2}{4\pi^2}}} \tag{7.40}$$

where f_i, [Hz] is the operating rotational speed.

In Figure 7.32c) it can be seen that the best running condition is achieved, when running with over- critical speed. Therefore, it is normally advantageous to choose a soft coupling with n_R substantially lower than the running speed.

7.4.4 Max coupling torque for squirrel-cage motor

Especially for the squirrel-cage motor it is interesting to calculate the maximum coupling torque based on information about the torque as a function of the rotational speed of the driving motor and the machine during a start-up. An example could be the starting of a ventilator or centrifugal pump by means of a squirrel-cage motor.



[billedtekst start]**Figure 7.33:** The torque function for a squirrel-cage motor and a centrifugal pump as function of the number of revolutions.[billedtekst slut]

Using a permanent elastic coupling the acceleration can be calculated

$$\frac{d\omega}{dt} = \frac{T_M - T_L}{I_1 + I_2} \tag{7.41}$$

The torque for the coupling can be calculated by either two equations

$$T_C = T_L + (I_{C2} + I_L)\frac{d\omega}{dt} \tag{7.42}$$

or

$$T_C = T_M - (I_{C1} + I_M)\frac{d\omega}{dt} \tag{7.43}$$

It is important to notice that the squirrel-cage motor passes through its torque curve (from left to right) every time it starts. It can be seen that the coupling load is nearly three times the nominal torque load for the motor during a starting period.

7.5 Overload couplings and safety couplings

Overload Couplings and Safety Couplings are used to protect machinery against a too high torque load. A typical scope of application is the Power Take Off drive shaft between a tractor and a machine. The significant difference in available power from the tractor and the power required for the machine, combined with a heavy flywheel in the tractor makes it necessary with an overload coupling as a protection during start-up as well as under working conditions.

7.6 Nomenclature

f_e	Hz	Natural frequency
f_{e1}, f_{e2}	Hz	Natural frequency for an one mass system
i	–	Order number (=number of impulses per revolution)
m	–	Ratio of drive end to driven end moment of inertia
N_R	Hz	Critical rotational speed
p_0	N/mm²	Pressure extreme between components
p_φ	N/mm²	Pressure value depending on angular position
$\{s\}$	mm	Geometric vector defined in global coordinate system
$\{s'\}$	mm	Geometric vector defined in local coordinate system
$[A]$	–	Transformation matrix
F_V	N	Bolt force
I_1, I_2	kgm²	Moment of inertia

R_t, *dyn*	kgm²/s²	Dynamic torsional stiffness of the coupling
S_f	–	Frequency coefficient
S_z	–	Starting coefficient
S_A	–	Impulse coefficient
S_L	–	Impulse coefficient
S_θ	–	Temperature coefficient
T_{Ai}, T_{Li}	Nm	Amplitude of the external torque excitation of the i-th order acting on the drive end, acting on the driven respectively end
T_{AS}	Nm	Maximum input torque
T_C	Nm	Torque transmitted by clutch
T_{Cmax}	Nm	The permissible maximum torque in the coupling
T_{CN}	Nm	The permissible nominal torque in the coupling
T_{CW}	Nm	Fatigue torque for the coupling
T_L	Nm	"Static" torque load to drive the machine
T_{LN}	Nm	The nominal torque for the driven machine

T_{LS}	Nm	Maximum load torque
T_M	Nm	Mean torque transmitted clutch
V_{fi}	–	Amplification factor at resonance speed
V_R	-	Amplification factor depending on the type of
W_d	Nm	The dynamic heat loss in the elastic elements
W_D	Nm	Static work absorbed in the coupling
W_e	Nm	Additional absorbed energy from the $+T_a$-component
α	rad	Angle between shaft in universal joint
μ	–	Coefficient of friction
w	rad/s	Angular speed
ϕ	rad	Angle
ψ	–	Relative damping factor
θ	rad	Angle

7.7 References

[1] *K. Decker. Maschinenelemente, Funktion, Gestaltung und Berechnung, 18. aktualisierta Auflage. Carl Hanser Verlag, München, Wien, 2011.*

[2] DIN 115. Ausgabe 1973-09. antriebselemente, schalenkupplungen.

[3] DIN 116. Ausgabe 1971-12. antriebselemente, scheibenkupplungen.

[4] DIN 740-2. Ausgabe 1986-08. antriebstechnik, nachgiebige wellenkupplungen. begriffe und berechnungsgrundlagen.

[5] GKN Walterscheid. Hot rolled steels for quenched and tempered springs, technical delivery conditions. Hauptstrasse 150, D-53797 Lohmar, Deutchland.

[6] KTR Kupplungstechnik GMBH. Rodder Damm 170, 48432 Rheine, Deutchland. Represented in Denmark by the firm: Lönne Scandinavia, Hjulmagervej 9D, 7100 Vejle.

[7] P. E. Nikravesh. *Computer-Aided Analysis of Mechanical Systems.* Prentice-Hall International, Inc., Englewood Cliffs, Nj 07632, 1988.

[8] Renk Aktiengesellschaft. Augsburg. - Represented in Denmark by the firm: Hans Manicus & Co. Aps.

[9] Ringspann GmbH. Bad Hamburg, Germany. - Represented in Denmark by the firm: Hans Manicus & Co. Aps.

[10] SEW-Eurodrive GmbH& Co KG. Represented in Denmark by the firm: SEW-Eurodrive A/S, Geminivej 28-30, 2670 Greve.[

11] Stromag. Hansastr. 120, 59425 Unna. Represented in Denmark by the firm: Stromag AB, Hvedemarken 11, 3550 Slangerup.

Chapter 8 Clutches

Clutches[1] are classified into the two main groups related to the torque transmission principle: *positive* or *frictional.* Furthermore clutches are divided into *closing* clutches and *opening* clutches and finally they can be described by the method of actuation (mechanical, pneumatic, hydraulic and electromagnetic). Special types of clutches are centrifugal clutches and overload clutches.

Positive (interlocking) clutches (direct engaged clutches)

To disengage one shaft from another, both running at low and high speed is relatively simple. It just results in instantaneous fall-out of the transmitted torque. In addition, an engagement between two shafts running at the same speed (or under stand-still condition) can take place without problems.

Instantaneous engagement of shafts running with different speeds is normally not possible as the difference in kinetic energy

$$\Delta E_{kin} = \frac{1}{2} I(\omega_M^2 - \omega_L^2) \tag{8.1}$$

has to be supplied or removed instantaneously. According to the principle of angular momentum ($T_{acc} = I(dw/dt)$) it is seen that the torque should be infinite if the change in speed (dw) was instantaneous. As this is not possible, two shafts running with different speed must be engaged smoothly over a suitable span of time.

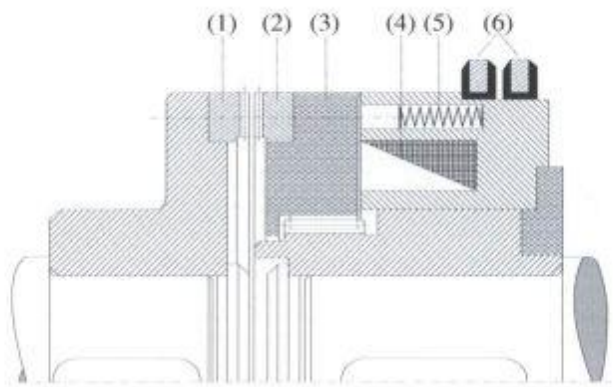
8.1 Friction clutches

Externally activated friction clutches are used in the transmission of torque from a driving to a driven shaft. The connection is provided by mechanical frictional engagement.

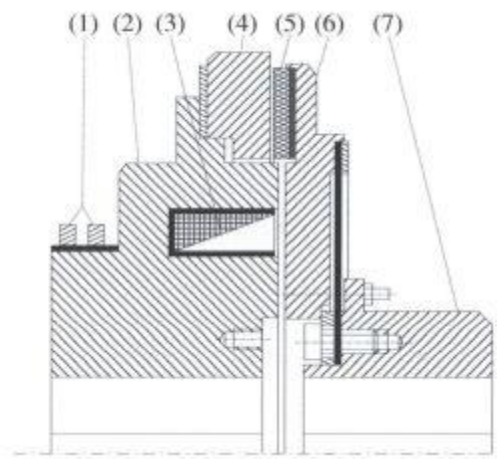
Frictional clutches can be classified according to number and arrangement of frictional surfaces in

- *single surface clutches. See Figure 8.2.*
- *dual surface (single-disc) clutches.*
- *multiple surface (multiple-disk) clutches. See Figure 8.4.*
- *cylindrical surface clutches*
- *cone (surface) clutches*

[1] Only mechanical clutches are described here.



[billedtekst start]**Figure 8.1:** Electromagnetic tooth clutch. Shaft clutch with sliprings, engaged by spring force (5) and released by magnetic force. The axially movable armature plate (3) engage by spring force through the radially toothed parts (1) and (2)[billedtekst slut]



[billedtekst start]**Figure 8.2:** Electromagnetic single-disc clutch. Many special designs exist for different applications. This one is with slipring body (1), coil body (2), coil (3), friction ring (4), carrier with friction lining (5), armature plate (6) and driving hub (7).[billedtekst slut]

Each type has different characteristics. Without going into details, the single disc type is the most robust one(heat absorption and heat transfer). Multiple disc types are, generally speaking cheaper and space saving. They may be either made dry or wet (oil lubricated).

8.1.1 Torque transmission (static)

For a small surface element $2\pi rdr$ on a disc the friction moment is

$$dT = 2\pi rdr \cdot p\mu r \tag{8.2}$$

This is integrated over the entire surface to

$$T_1 = \int_{r_i}^{r_o} 2\pi p\mu r^2 dr \tag{8.3}$$

1. *Assume uniform distribution of interface pressure.* This assumption is valid for an unworn (new), accurately manufactured clutch with rigid outer "disks". With reference to Figure 8.3 the total torque for one friction surface is

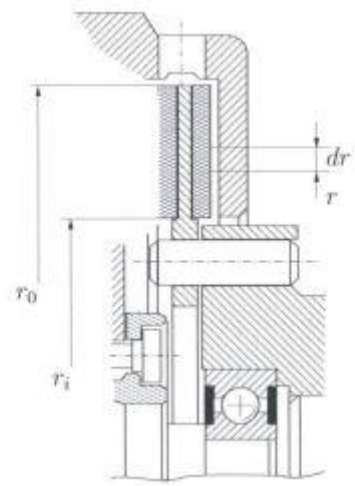
$$T_1 = \int_{r_i}^{r_o} 2\pi p \mu r^2 dr = \mu \frac{2\pi}{3}(r_o^3 - r_i^3)p \tag{8.4}$$

where p is the uniform level of interface pressure. The total normal force acting on the area is

$$F = \int_{r_i}^{r_o} 2\pi p r dr = \pi(r_o^2 - r_i^2)p \tag{8.5}$$

When eliminating the pressure p the total torque for one friction surface can be expressed by the normal acting force F

$$T_1 = \mu \frac{2}{3}\frac{r_o^3 - r_i^3}{r_o^2 - r_i^2}F \tag{8.6}$$



[billedtekst start]**Figure 8.3:** Cut-out from a clutch drawing.[billedtekst slut]

2. *Assume uniform rate of wear at interface.* It is a generally accepted assumption that wear rate is proportional to the rate of friction work which is friction force times rubbing speed. With an uniform coefficient of friction, the wear rate is proportional to the product of pressure times sliding speed. On the clutch face, speed is proportional to radius. On this basis, a new clutch (with uniform distribution of interface pressure) will have its greatest initial wear at the outer radius. After this initial running wear, the friction lining tends to wear at an uniform rate. Thus, assuming that

$$pr = C \quad \Rightarrow \quad p = \frac{C}{r} \tag{8.7}$$

where C is a constant.

Introducing this in (8.3) the torque can again be calculated

$$\begin{aligned} T_1 &= \int_{r_i}^{r_o} 2\pi p \mu r^2 dr \\ &= \int_{r_i}^{r_o} 2\pi C \mu r dr \\ &= \pi \mu C (r_o^2 - r_i^2) \end{aligned} \tag{8.8}$$

To eliminate the constant C the total axial force (normal force) on the friction surface can be calculated

$$\begin{aligned} F &= \int_{r_i}^{r_o} 2\pi r dr p \\ &= \int_{r_i}^{r_o} 2\pi C dr = 2\pi C (r_o - r_i) \end{aligned} \tag{8.9}$$

$$C = \frac{F}{2\pi(r_o - r_i)} \tag{8.10}$$

With C inserted in the torque (8.8) this becomes very simple

$$T_1 = \mu \frac{r_o + r_i}{2} F \tag{8.11}$$

For a disc clutch that has been running for some time the friction moment can be calculated as the friction force $F\mu$ acting on the mean radius r_m.

Friction radius

In the literature the friction torque is often written as:

$$T = \mu r_f F \tag{8.12}$$

where r_f is called friction radius. For the disc friction clutch the friction radius r_f is obviously equal to the mean radius r_m as seen in (8.11). For the new unworn disc clutch with uniform distributed interface pressure, as well as for a flange coupling the friction radius is

$$r_f = \frac{2}{3} \frac{r_o^3 - r_i^3}{r_o^2 - r_i^2} \tag{8.13}$$

The difference in value between the two expressions for the friction radius is normally negligible as the ratio r_i/r_0 normally is greater than or equal to 0.5.

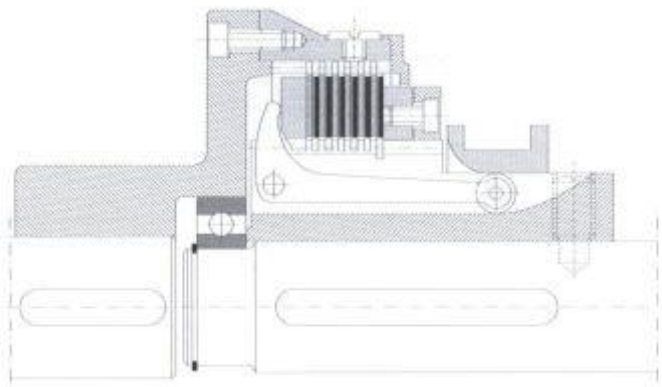
8.1.2 Transient slip in friction clutches during engagement

During the engagement of a friction clutch under running condition, friction energy will cause a raise in temperature that can damage the clutch, if there is a mismatch (disproportion) between the clutch size and the working condition.

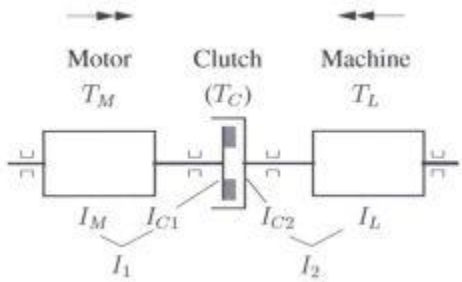
A case will be described as a background for an analysis of the transient slip in a friction clutch during engagement. A machine is assumed to be started under a constant load situation. The task for the motor is

- to accelerate the rotating masses in the machine (the inertia of the working

machine).



[billedtekst start]**Figure 8.4:** Example of mechanical operated multi-disc clutch, [billedtekst slut]engaged.



[billedtekst start]**Figure 8.5:** Principle arrangement of driving motor connected to machine through clutch.[billedtekst slut]

- to balance the work load T_L (coming from the process in the working machine).

The work load T_L is assumed to be constant and not depending on time or speed. The clutch torsional moment (torque transferred from the motor side of the clutch to the machine side of the clutch) is in Figure 8.6 shown to be a function of the time $T_C = f(t)$. The angular speed for the motor is supposed to be constant during the engagement period. Due to that, the mass moment of inertia for the motor I_1 is of no interest and the torque from the motor T_M will be equal to the clutch torque T_C. At the beginning of the engagement period the machine will not rotate $w2 = 0$ until the clutch torque is equal to the work load torque T_L. From that time ($t = t_1$) a torque is left for an acceleration $T_{acc} = T_C - T_L$ of the mass of inertia on the machine side. The equilibrium of angular moment gives

$$T_{acc} = T_C - T_L = I_2 \frac{d\omega_2}{dt} \tag{8.14}$$

and

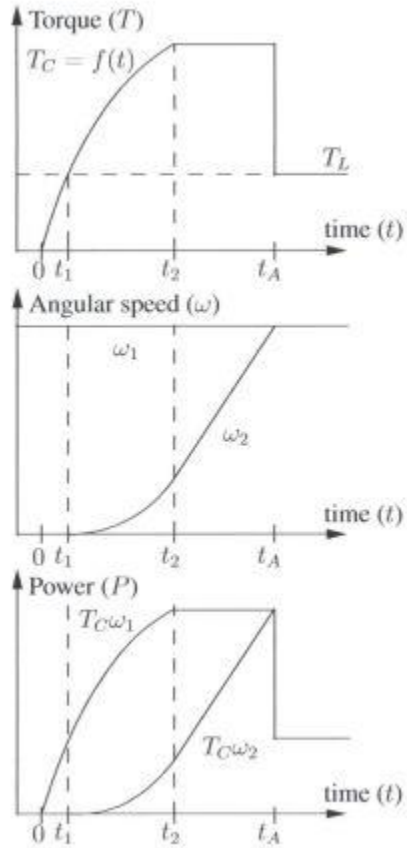
$$\omega_2 = \int_{t_1}^{t} \frac{T_C - T_L}{I_2} dt \tag{8.15}$$

when $t = t_A$ then $w_2 = w_1$ and the clutch locks up. Note that from t_2 to t_A we have that

$$\frac{T_C - T_L}{I_2} = \text{constant} \tag{8.16}$$

leading to a linear w_2-function.

The clutch will during the engagement period receive the energy



[billedtekst start]**Figure 8.6:** An engagement case in general.[billedtekst slut]

$$Q_M = \int_0^{t_A} T_C\omega_1 dt \tag{8.17}$$

A part from that is transferred directly to the machine as mechanical energy

$$Q_L = \int_0^{t_A} T_C\omega_2 dt \tag{8.18}$$

and the difference will remain in the clutch as "heat"

$$Q_C = \int_0^{t_A} T_C(\omega_1 - \omega_2) dt \tag{8.19}$$

In a regular case where $T_L \neq 0$, the dissipated energy in the clutch Q_C is

$$\begin{aligned} Q_C &= \int_0^{t_A} T_C(\omega_1 - \omega_2)dt \\ &= \int_0^{t_A} \left(I_2 \frac{d\omega_2}{dt} + T_L\right)(\omega_1 - \omega_2)dt \\ &= \frac{1}{2} I_2 \omega_1^2 + \int_0^{t_A} T_L(\omega_1 - \omega_2)dt \end{aligned} \tag{8.20}$$

In a case where the work load during the engagement period is negligible (T_L, = 0), the acceleration torque T_{acc} will be

$$T_{acc} = T_C = I_2 \frac{d\omega_2}{dt} \tag{8.21}$$

and the heat developed in the clutch Q_{CD}

$$\begin{aligned} Q_{C_0} &= \int_0^{t_A} I_2 \frac{d\omega_2}{dt}(\omega_1 - \omega_2)dt \\ &= I_2 \int_{\omega_2=0}^{\omega_2=\omega_1} (\omega_1 - \omega_2)d\omega_2 \\ &= I_2 \left[\omega_1\omega_2 - \frac{\omega_2^2}{2}\right]_{\omega_2=0}^{\omega_2=\omega_1} \\ &= \frac{1}{2} I_2 \omega_1^2 \end{aligned} \tag{8.22}$$

Often the clutch torque function $T_C = f(t)$ is unknown. In this case it is usually assumed that T_C momentarily is brought to its maximum value, which of course has to be larger than the load torque T_L.

In this simplified case the acceleration torque is

$$T_{acc} = T_C - T_L = \text{constant} \tag{8.23}$$

from which

$$T_C - T_L = I_2 \frac{d\omega_2}{dt} \tag{8.24}$$

and the w_2 function

$$\begin{aligned} \omega_2 &= \int_0^t \frac{T_C - T_L}{I_2} dt \\ &= \frac{T_C - T_L}{I_2} t \end{aligned} \tag{8.25}$$

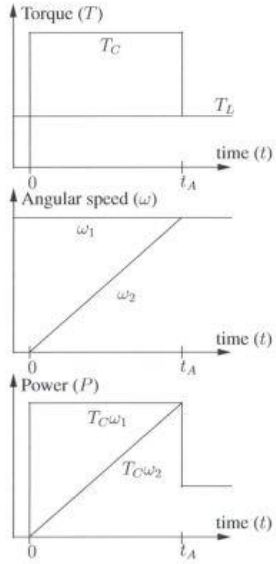
Now it is possible to calculate the engagement time t_A as

$$\frac{d\omega_2}{dt} = \frac{T_C - T_L}{I_2} \tag{8.26}$$

$$dt = \frac{I_2}{T_C - T_L} d\omega$$

$$t_A = \int_0^{\omega_1} \frac{I_2}{T_C - T_L} d\omega$$

$$t_A = \frac{I_2}{T_C - T_L} \omega_1 \tag{8.27}$$



[billedtekst.start]**Figure 8.7:** An engagement case, simplified.[billedteskst.slut]

The dissipated energy in the clutch is (in the simplified case)

$$Q_C = \int_0^{t_A} T_C(\omega_1 - \omega_2)dt \tag{8.28}$$

In this equation can be calculated from

$$\omega_2 = \frac{\omega_1}{t_A}t \tag{8.29}$$

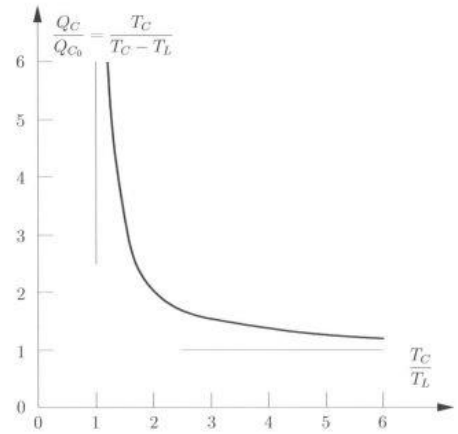
$$\begin{aligned} Q_C &= T_C\omega_1 \int_0^{t_A} \left(1 - \frac{t}{t_A}\right) dt \\ &= \frac{1}{2}T_C\omega_1 t_A \end{aligned} \tag{8.30}$$

which with t_A from (8.27) gives

$$Q_C = \frac{1}{2}I_2\omega_1^2\frac{T_C}{T_C - T_L} \tag{8.31}$$

Comparing the dissipated energy Q_C developed in the clutch during engagement under load condi tion with the dissipated energy under an unloaded engagement period $Q_{C_0} = 1/2 \cdot I_2\omega_1^2$ gives a good impression of the influence of the fraction T_C/T_L on the real heat load for a normal loaded start situation

$$\frac{Q_C}{Q_{C_0}} = \frac{T_C}{T_C - T_L} \tag{8.32}$$



[billedtekst.start]**Figure 8.8:** Dissipated energy as a function of T_C/T_L.[billedteskst.slut]

T_C has to be greater than about 2 times T_L, to limit the developed heat to an acceptable level. An increase to more than 3 times results on the other hand only in minimal reduction in Q_C/Q_{C_0}. See Figure 8.8.

8.1.3 Dissipated energy in the clutch

The dissipated energy from one engagement period for the clutch causes an increase in temperature in disc and neighbouring parts. Of course the highest raise in temperature will take place in the sliding surface and from there propagate to other parts of the clutch depending on heat conductivity resistance and heat transfer coefficient.

The cooling of the clutch will take place during and after the engagement, see Figure 8.9, illustrating the conditions for a multi disc clutch. The heat transfer is calculated by using the Fourier equation for in-stationary heat transmission. A one dimensional analysis perpendicular to the sliding surfaces is often sufficient.

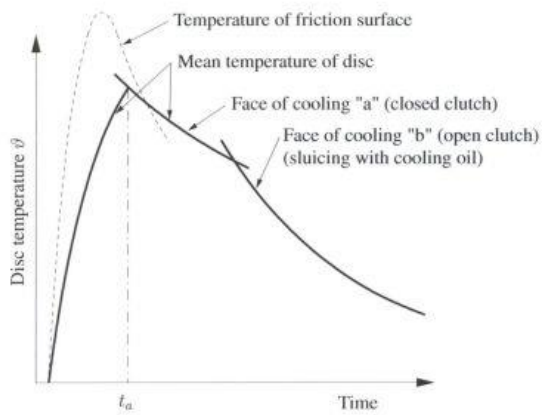
$$\frac{\partial \vartheta}{\partial t} = a \frac{\partial^2 \vartheta}{\partial x^2} \tag{8.33}$$

where

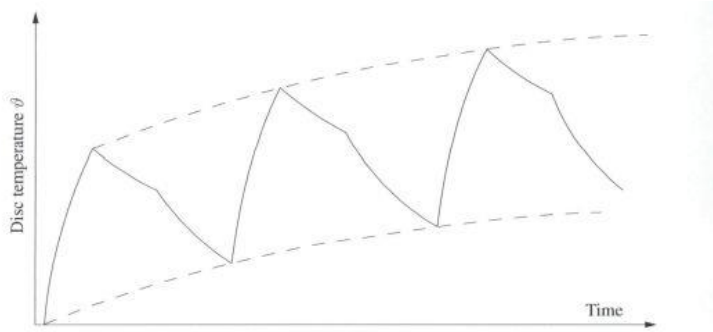
$$a = \frac{\lambda}{c\rho} \left[\frac{m^2}{s} \right] \quad \text{(thermal diffusivity)} \tag{8.34}$$

For further calculations it is advantageous to use this equation in a dimensionless form.

Often calculated is the mean temperature raise $\Delta\theta$ that would occur if the dissipated energy for one engagement would remain in the disks.



[billedtekst.start]**Figure 8.9:** Temperature in a clutch.[billedteskst.slut]



[billedtekst.start]**Figure 8.10:** Qualitative description of temperature with many engagements.[billedteskst.slut]

$$\Delta\vartheta = \frac{Q}{mc} \tag{8.35}$$

where m is mass of the heat absorbing disks and c is specific heat. (For steel: 0.46kJ/(kg°C), for cast iron: 0.55kJ/(kg°C)). With high frequency of engagement the above method is not sufficient.

Instead a friction face load is calculated. In one hour the dissipated energy Q_h is

$$Q_h = \frac{1}{2} I_2 \omega_1^2 \frac{T_C}{T_C - T_L} S_h \tag{8.36}$$

where S_h, is the number of engagements per hour. This gives a friction surface load:

$$\dot{Q}_h = \frac{1}{2} I_2 \omega_1^2 \frac{T_C}{T_C - T_L} \frac{S_h}{3600} \tag{8.37}$$

The admissible heat load value $\dot{q}_{adm}$ is defined as the acceptable friction surface load per unit area. On the basis hereof it is possible to calculate the required friction area A

$$A = \frac{\dot{Q}_h}{\dot{q}_{adm}} \tag{8.38}$$

Table 8.1: Example of admissible heat loadings [4].

Clutch type	**Admissible heat loading** $\dot{q}_{adm}[\mathrm{W/mm^2}]$	**Remarks**
Single disc clutches	1-2	Electro. clutch
	1-4	Pneum. clutch depending on speed and size
Multi-disc clutches	0.2-0.35	Steel/org. Lining, dry operation
	0.1-0.25	Wet operation (oil mist)
	0.2-0.45	Oil Splash
	1-3	Oil inner cooling

8.1.4 Layout design of friction clutches

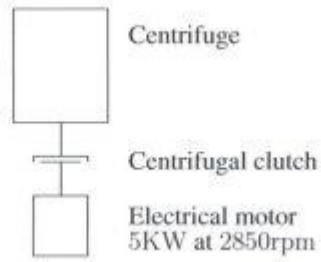
A clutch is basically designed according to the maximum torque to be transmitted and the engagement force that produces it, and so the thermal stress is usually the determining factor for selecting the appropriate size. The torque to be transmitted is governed by the nominal torque of the drive motor and operating machine, where allowance must be made for cyclic variations or the torque of tilt ($2T_{AN}$ to $3T_{AN}$) in squirrel-cage motors. The engaging and disengaging torque of a clutch is generally smaller than the transmittable torque, due to the difference in static and dynamic coefficient of friction. In particular, for wet clutches the sliding friction coefficient is smaller than the static friction coefficient.

8.2 Automatic clutches

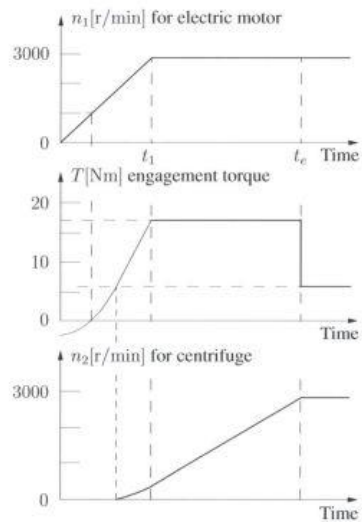
Automatic clutches can be divided into two groups depending on the type of activation: Speed-sensitive Clutches and Directional (One Way) Clutches. Sometimes Overload Clutches and Safety Clutches are also classified as automatic clutches.

8.2.1 Speed-sensitive clutches (centrifugal clutches)

Centrifugal clutches (and hydrodynamic clutches) are used as "soft-start" clutches especially where large masses of inertia have to be accelerated to high rotational speeds by relatively small driving motors. A centrifuge is a typical example. The characteristic of a centrifugal clutch is that the transmittable torque increases with the square of the rotational speed. To prevent the clutch from transferring torque already from stand-still, elastic springs are build in to delay the engagement. Some characteristics can be shown in the following example:



[billedtekst.start]**Figure 8.11:** The motor and the centrifuge are coupled via a centrifugal clutch. The clutch starts engaging at l000rpm (equilibrium between the centrifugal force on the shoes and the spring force) so that the transmittable power is exactly 5kW at 2850rpm. The work load for the centrifuge is constant = 6Nm (independent of the rotational speed).[billedteskst.slut]



[billedtekst.start]**Figure 8.12:** Solution to the example.[billedteskst.slut]

Example: centrifugal clutch for a centrifuge

Based on the information above, 3 diagrams can now be drawn. See Figure 8.12. In the first a rpm-curve for the motor (simplified) is already shown. The task is now, based on simple calculations, to draw

- a curve for transmitted torque to the centrifuge.
- a curve for the rotational speed of the centrifuge (The centrifuge is supposed to reach 2850rpm at the time t_e).

For $t = 0$ to $t = t_1$ the transmittable torque is a second degree curve. The torque has to equal 0 for $n = 1000$ rpm.

$$T_C = k(\omega^2 - (\frac{\pi}{30}1000\text{s}^{-1})^2) \tag{8.39}$$

For $w = (\pi/30)2850\text{s}^{-1}$ is

$$T = \frac{P}{\omega} = \frac{5000\text{W}}{\frac{\pi}{30}2850\text{s}^{-1}} = 16.75\text{Nm} \tag{8.40}$$

Now k can be calculated

$$\begin{aligned} k &= \frac{T}{\omega^2 - (\frac{\pi}{30}1000\text{s}^{-1})^2} \\ &= \frac{16.75\text{Nm}}{(\frac{\pi}{30}2850\text{s}^{-1})^2 - (\frac{\pi}{30}1000\text{s}^{-1})^2} \\ &= 2.144510^{-4}\text{Nms}^2 \end{aligned} \tag{8.41}$$

The second degree curve for the torque crosses the ordinate axis in

$$T_k = -k(\frac{\pi}{30}1000\text{s}^{-1})^2 = -2.3517\text{Nm} \tag{8.42}$$

The Centrifuge will start when the torque reach a level of 6Nm

$$T_k = k\omega_c^2 - 2.3517\text{Nm} = 6\text{Nm} \tag{8.43}$$

from which w_c can be calculated

$$w_c = 197.3\text{s}^{-1}$$
$$n_c = 1885\text{rpm}$$

8.2.2 Directional (one-way) clutches, overrun clutches

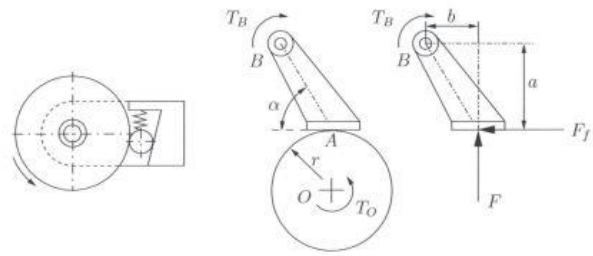
Directional clutches are designed to transmit torque in only one direction. However, torque is only transmitted between the two clutch-halves as long as the driving part makes an attempt to rotate faster than the driven one. This type of clutch is well known from the freewheel in a bicycle. As the function for a freewheel clutch is only dependent of the relative rotary motion, this type of clutch will also be able to function as a backstop. See description in Figure 7.10.

This chapter only deals with backstops based on friction elements. Typical types of backstops are based on mechanisms with latches or pawls and on mechanisms with friction elements.

For the backstop mechanism shown to the right in Figure 8.13, the braking torque T_O has to be calculated as a function of the external activation torque T_b and the geometry for the components. Balance of moments about the pivotal point B for the shoe and lever assembly

$$T_b + F_f a - Fb = 0 \tag{8.44}$$

At first, presume that the angle α is so small that the mechanism will act as a brake and sliding will take place so that $F_f = \mu F$ or $F = F_f / \mu$.



[billedtekst.start]**Figure 8.13:** Backstop mechanisms based on two different principles.[billedteskst.slut]

$$T_B + F_f a - F_f \frac{b}{\mu} = 0 \tag{8.45}$$

Now the "braking force" F_f can be calculated as a function of T_B

$$F_f = \frac{T_B}{\frac{b}{\mu} - a} \quad \Rightarrow \quad F_f = T_B \frac{\mu}{b} \frac{1}{1 - \mu \frac{a}{b}} \tag{8.46}$$

Finally the "braking torque" $T_O = F_f R$ can be found

$$T_O = T_B \mu \frac{r}{b} \frac{1}{1 - \mu \tan \alpha} \tag{8.47}$$

For $\mu \tan \alpha \rightarrow 1$ then $T_O/T_B \rightarrow \infty$, and for $\mu \tan \alpha \geq 1$ the "brake" acts as a backstop as long as $T_B \geq 0$.

When designing backstops the angle α has to be chosen as close down to the limit as a secure function allows. That is as close to the line arctan(l/μ) as possible in order to limit the contact pressure between the shoe and the drum.

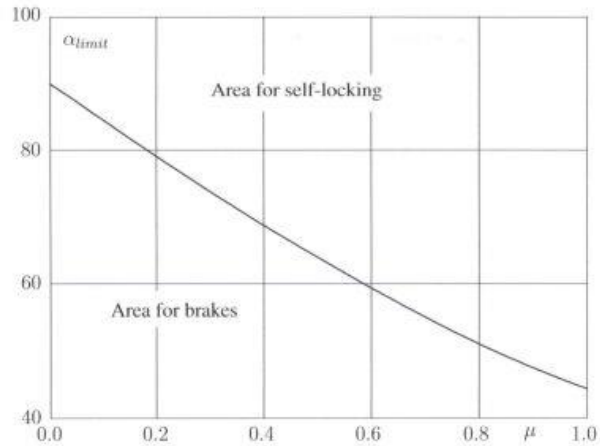
Paper-clip example. This is an example with two rollers, the big one with fixed pivot point. The paper is to be placed between the two rollers. This mechanism can be compared with the previously described shoe and drum, as the contact point B can be regarded as pivot point. The condition of correct function is that

$$\alpha \geq \arctan \frac{1}{\mu} \tag{8.48}$$

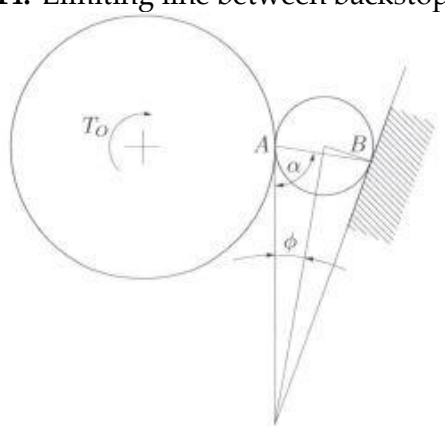
which is the same as

$$\varphi < \arctan \boldsymbol{\mu} \tag{8.49}$$

Please notice that it is the lower value of μ at A or B which is crucial for the proper function of the mechanism.



[billedtekst.start]**Figure 8.14:** Limiting line between backstops and brakes.[billedteskst.slut]



[billedtekst.start]**Figure 8.15:** "Backstop"-mechanism used as paper-clip.[billedteskst.slut]

8.3 Nomenclature

a	m²/s	Thermal diffusivity
c	J/kg°C	Heat capacity system
p_0	N/mm²	Pressure extreme between components
p_φ	N/mm²	Pressure distribution between components depending on angular position
F_V	N	Bolt force

I_1, I_2	kgm²	Moment of' inertia
P	Nm/s	Transmitted power
Q	Nm	Dissipated energy in the clutch
R_T,DYN	kgm²/s²	Dynamic torsional stiffness of the coupling
T_{ACC}	Nm	Accelerating torque
T_L	Nm	"Static" torque load to drive the machine clutch
λ	W/m°C	Heat conductivity
μ	–	Coefficient of friction
p	kg/m³	Density
W	rad/s	Angular speed

8.4 References

[1] Fichtel & Sachs. Today ZF Sachs. Friedrichshafen AG. Represented in Denmark by the company: ZF Danmark, Tåstrupgårdsvej 8-10, 2630 Tåstrup.

[2] Malmedie Antriebstechnik GMBH. Dycker Feld 28, D-42653 Solingen.

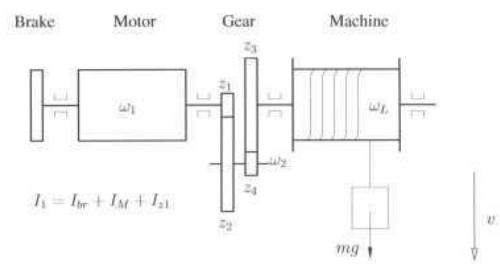
[3] Ortlinghaus Werke GmbH. Represented in Denmark by the firm: Bdr. Klee.

[4] Stromag. Hansastr. 120, 59425 Unna. Represented in Denmark by the firm: Stromag AB, Hvedemarken 11, 3550 Slangerup.

Chapter 9 Brakes

The basic function for a brake is to absorb kinetic and potential energy from moving parts (translational or rotational).Convert this energy to friction heat and dissipate the resulting heat without developing destructively high temperatures. As indicated, a brake can be linear as in mechanical lifts, where there has to be a linear safety brake to prevent the chair from falling down in case of broken wires, or it can be a brake for slowing down the rotational speed of e.g. a wind turbine, and bring it to standstill. Other examples are brakes in cars and in cranes and hoists.

Example: a brake for a hoist A hoist may consist of an electro-motor with integrated brake, a gear and a wire pulley. See a sketch in Figure 9.1. The brake can be an electrically activated brake automatically brought to function when the current to the motor is switched off. However, it is important to know that the brake needs a small amount of time to react (delay time), which means that when the stop-button is activated, there will be a short period, where the load will be in a free fall as neither the motor nor the brake will be able to prevent this.



[billedtekst start]**Figure 9.1:** Schematic depiction of hoist.[billedtekst slut]

In the present calculations we do not incorporate the delay time. The brake torque which is necessary to decelerate the rotating masses (when lowering the load) and the translational moving masses *m* have to be

$$T_{br} = \sum I_{Transformed}\dot{\omega}_1 + T_{Transformed} \tag{9.1}$$

In this equation inertia masses and torsional moments are transferred to the braking shaft. From the knowledge of the braking time, the braking torque can be calculated. The load torque can be transferred to the braking shaft by a simple calculation of moment equilibrium

$$T_{Transf.} = mgr\frac{\omega_L}{\omega_1} = mg\frac{v}{\omega_1} \quad (r = \text{radius}) \qquad (9.2)$$

$$(v = \text{the translational speed of the payload})$$

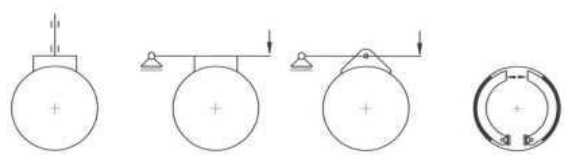
The transformation of the masses of inertia can be done by means of an energy consideration, as the kinetic energy for a rotating mass is: $E_{kin} = \frac{1}{2}I\omega^2$ we have to remove this completely when braking to stand-still. As seen in (9.1) all calculations are referred to the motor shaft including all masses of inertia

$$\sum I_{Transf.} = I_1 + I_2\left(\frac{\omega_2}{\omega_1}\right)^2 + I_L\left(\frac{\omega_L}{\omega_1}\right)^2 + m\left(\frac{v}{\omega_1}\right)^2 \qquad (9.3)$$

To complete the calculations for the hoist it is absolutely necessary to incorporate the delay time as you need to know how far the hook will move after the stop-button has been activated. It is possible to be informed about the delay time for braking motors, and as an example the delay time can vary from 15 milliseconds to 400 milliseconds, depending on the size of the motor and the methods of wiring the brake to the motor. For a particular type and size of an Asea breake motor the stopping length for the hook can be up to 300% larger for one method of connecting the brake than for another one. (There is of course a difference in price for the different methods of wiring the brake to the motor)

9.1 Drum brakes

Drum brakes, or more generally: radially activated brakes can be found in many variants. A few is shown here



[billedtekst start]**Figure 9.2:** Variants of drum brakes.[billedtekst slut]

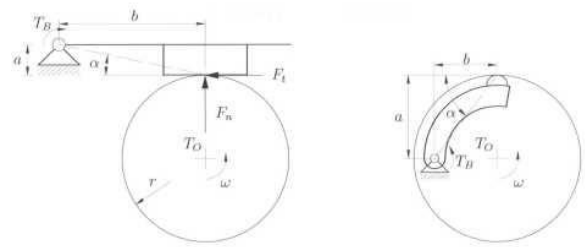
For radially actuated brakes, an often seen formation concept is the self-energizing brake. The level of self-energizing is expressed as an amplification factor λ.

9.1.1 Self-energizing

Continuing the calculations from the backstop section for the braking torque: T_O as a function of the external activation torque T_B and for the indicated direction of rotation of the drum, this could be written

$$T_O = T_B\mu\frac{r}{b}\frac{1}{1-\mu\tan\alpha} \qquad (9.4)$$

For the opposite direction of rotation, similarly



[billedtekst start]**Figure 9.3:** Schematic depiction of external and internal drum brake.[billedtekst slut]

$$T_O = T_B \mu \frac{r}{b} \frac{1}{1 + \mu \tan \alpha} \tag{9.5}$$

For $\alpha = 0$ that is $a = 0$ the two expressions for the braking torque are identical

$$T_O = T_B \mu \frac{r}{b} \tag{9.6}$$

For $\alpha \neq 0$we have a self-energizing or de-energizing effect characterized by the amplification factor λ

$$\lambda = \frac{1}{1 - \mu \tan \alpha} \tag{9.7}$$

Please notice that the amplification factor λ is very sensitive to changes in the friction coefficient μ if the angle α_{design} is chosen close to α_{limit}. An example will illustrate this: With a μ-value of 0.6 it

can be calculated that α_{limit} = 1.0304. If the designer choose to use an $\alpha_{design} = 5\pi/18$, it can be calculated from (9.7) that a variation in μ-value of 0.6 ±0.1 will cause a variation in the amplification factor λ between 2.47 and 6.03.

9.1.2 Braking torque and friction radius

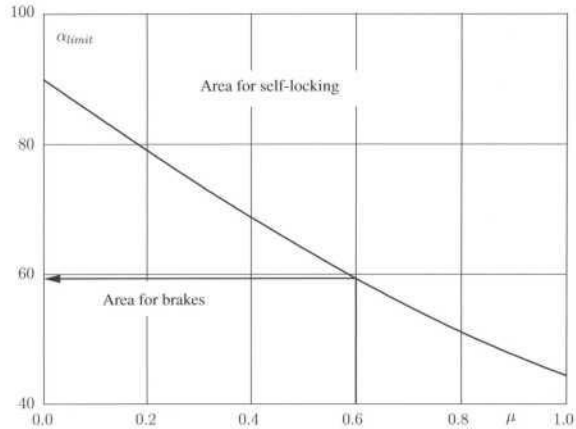
Assuming that the pressure distribution is a known function of φ,($P_\varphi = f(\varphi)$), e.g. as sketched in Figure 9.5, the friction torque can be calculated

$$T = \int_{-\theta/2}^{+\theta/2} p_\varphi \mu r d\varphi \cdot wr \tag{9.8}$$

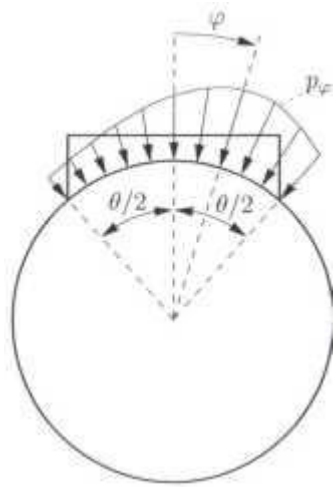
This torque can be written as

$$T = F_e \mu r_f \tag{9.9}$$

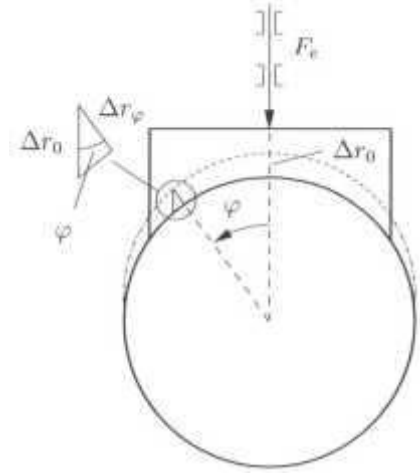
where F_e is the external activating force and r_f is determined such that (9.8) and (9.9) will give the same value for the torque), r_f is called the friction radius. In the next subsection some typical examples will be analyzed.



[billedtekst start]**Figure 9.4:** Selection of lead angle α_{design}.[billedtekst slut]



[billedtekst start]**Figure 9.5:** Friction torque and friction radius.[billedtekst slut]



[billedtekst start]**Figure 9.6:** Wear and surface pressure.[billedtekst slut]

9.1.3 Wear and normal pressure for parallel guided shoe

The assumption is that the shoe is guided in exact vertical direction and the external force is F_e. In Figure 9.6 a dotted line is showing the worn shoe as it looks after some time. The moved distance (during the

wear) in vertical direction is called Δr_0. As is seen in Figure 9.6 the wear in radial direction depends on φ in the following way

$$\Delta r_\varphi = \Delta r_0 \cos\varphi \qquad (9.10)$$

It is assumed that normal wear is proportional to friction work and friction work is proportional to surface pressure. That is

$$p_\varphi = const \cdot \Delta r_\varphi \qquad (9.11)$$

From this result follows that the pressure function can be written as

$$p_\varphi = p_0 \cos(\varphi - \psi) \qquad (9.12)$$

The p_0-value can now be calculated as a function of F_e and the geometry of the brake.

$$\begin{aligned} F_e &= \int_{-\theta/2}^{+\theta/2} p_\varphi w r d\varphi \cdot \cos\varphi \\ &= \int_{-\theta/2}^{+\theta/2} p_0 \cos(\varphi - \psi) \cos\varphi d\varphi \cdot wr \end{aligned} \qquad (9.13)$$

If the pressure distribution is symmetric, ψ = 0, the integration gives

$$F_e = p_0 w r \left(\frac{\theta}{2} + \cos\frac{\theta}{2} \sin\frac{\theta}{2} \right) \qquad (9.14)$$

The actual value of the brake torque is

$$T = \int_{-\theta/2}^{+\theta/2} p_\varphi w r d\varphi \cdot \mu r \qquad (9.15)$$

If ψ= 0 the integration gives

$$T = 2\mu w r^2 p_0 \sin\frac{\theta}{2} \qquad (9.16)$$

This torque can be written by the use of the friction radius as $T = F_e \mu r_f$. The purpose of using r_f is to simplify the calculations without reducing the accuracy.

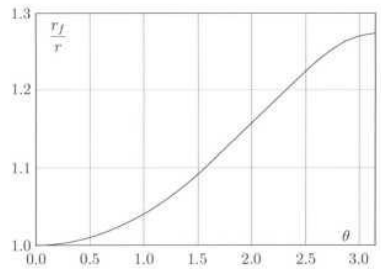
From the knowledge of the relations between F_e and p_0 in (9.13) and the two equations for the braking torque, the friction radius can now be calculated, ψ = 0

$$\frac{r_f}{r} = \frac{2\sin\frac{\theta}{2}}{\frac{\theta}{2} + \cos\frac{\theta}{2}\sin\frac{\theta}{2}} = \frac{4\sin\frac{\theta}{2}}{\theta + \sin\theta} \qquad (9.17)$$

The similarity with the calculation performed for the split muff coupling, where we assumed sine- formed pressure distribution, is obvious. With two half parts of the coupling and an angle of contact of , the friction torque is

$$T = 2\frac{r_f}{r} r F_e \mu = 2\frac{4}{\pi} r F_e \mu \qquad (9.18)$$

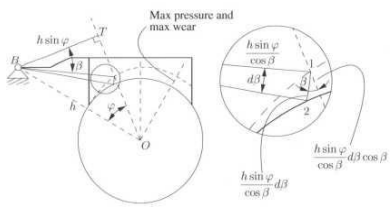
It can also be seen that r_f/r is the fraction between the arithmetical (real) sum of friction forces on radius r and the geometrical sum (vectorial) on the radius r_f.



[billedtekst start]**Figure 9.7:** r_f/r for symmetrical pressure distribution.[billedtekst slut]

9.1.4 Wear and normal pressure for non-pivoted long shoe

Figure 9.8 shows a brake shoe in contact with a drum. As the shoe wears, it pivots about *B.* For a rather extreme amount of wear an arbitrary point 1 moves to 2. The actual wear for the point 1 is represented by a line from point 1 to the drum surface in the direction of the drum center *O.* During the wear the shoe has rotated an angle $d\beta$, and this angle is of course the same for all points on the contact surface.



[billedtekst start]**Figure 9.8:** Brake arm with non-pivoted long shoe.[billedtekst slut]

From the figure it can be seen that the wear in radial direction is

$$\frac{h \sin\varphi}{\cos\beta} d\beta \cdot \cos\beta = h d\beta \cdot \sin\varphi \qquad (9.19)$$

As h and $d\beta$ both are constant over the wear surface the wear will always be shaped as a sine- function. According to this, the wear will obtain its maximum value for

$$\varphi = \frac{\pi}{2}$$

At the same angle we will find the maximum pressure and the pressure function can now be written

$$P_\varphi = P_{\varphi=\pi/2} \sin \varphi \qquad (9.20)$$

Design hint. By adjusting the placement for the pivot point, it is often possible to achieve a symmetrical pressure distribution as for the parallel guided shoe in Figure 9.6. This is advantageous because of the best utilization of the linings.

About calculation method. The analysis of the above shoe pressure could be done much easier by noticing that for a rotation (dβ) around a pivot point *(B),* the component of movement in an arbitrary chosen direction *(OT)* will be given by the distance from the pivot point to that direction *(h* sin *φ)* times the rotation angle *(dβ).* Consequently, the displacement in the *OT*-direction for an angular rotation *dβ* will be

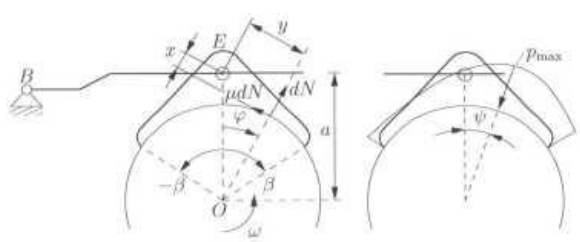
$$h d\beta \cdot \sin \varphi \qquad (9.21)$$

9.1.5 Wear and normal pressure for pivoted long shoe

The pivoted long shoe is often designed symmetrically around *OE* (see Figure 9.9), which results in the same braking torque for both directions of rotation. This type of brake will also normally be built with a shoe on both sides of the braking drum in order to reduce the side force on the shaft.

When designing the brake it is advantageous to place the pivot point as close to the drum as practical possible to reduce the tendency to "capsize".

During the braking process the sum of the moments attacking the shoe will balance. This is obtained automatically by the surface pressure distribution. The pressure distribution and the friction radius can now be calculated.



[billedtekst start]**Figure 9.9:** Brake arm with pivoted long shoe.[billedtekst slut]

At first the moments around pivot point *E* is calculated with the shown direction of revolution

$$\int_{-\beta}^{+\beta} dN \cdot y - \int_{-\beta}^{+\beta} \mu dN \cdot x = 0 \qquad (9.22)$$

or after insertion of *x* and *y*

$$\int_{-\beta}^{+\beta} dN \cdot a \sin \varphi - \int_{-\beta}^{+\beta} \mu dN \cdot (a \cos \varphi - r) = 0 \qquad (9.23)$$

In this equation dN can be calculated as a function of width of the lining w and the drum radius r.

$$dN = p_\varphi w r d_\varphi \tag{9.24}$$

With this inserted in (9.23) we have

$$\int_{-\beta}^{+\beta} p_\varphi w r d\varphi \cdot a \sin\varphi - \int_{-\beta}^{+\beta} \mu p_\varphi w r d\varphi \cdot (a\cos\varphi - r) = 0 \tag{9.25}$$

The pressure distribution as a function of φ is unknown, but (as earlier) supposed to be proportional with the wear and through that a function of cosines, p_0 is the symbol for the surface pressure, where this has its maximum, p is presumed to have its maximum value for the angle ψ.

$$P_\varphi = P_0 \cos(\varphi - \psi) \tag{9.26}$$

We first calculate the arithmetic sum (F_f) of the friction forces at radius r. From this can p_0 and p_φ be calculated.

$$\begin{aligned} F_f &= \int_{-\beta}^{+\beta} \mu p_\varphi w r d\varphi \\ &= \mu w r p_0 \int_{-\beta}^{+\beta} \cos(\varphi - \psi) d\varphi \\ &= \mu w r p_0 \left[\sin(\varphi - \psi)\right]_{-\beta}^{+\beta} \\ &= 2\mu w r p_0 \sin\beta \cos\psi \end{aligned} \tag{9.27}$$

$$p_0 = \frac{F_f}{2\mu w r \sin\beta \cos\psi} \tag{9.28}$$

$$p_\varphi = \frac{F_f \cos(\varphi - \psi)}{2\mu w r \sin\beta \cos\psi} \tag{9.29}$$

This expression for p_φ is inserted in (9.25) and the angle φ can now (after some rewriting) be calculated to

$$\tan\psi = \frac{\mu a(2\beta + \sin 2\beta) - 4\mu r \sin\beta}{a(2\beta - \sin 2\beta)} \tag{9.30}$$

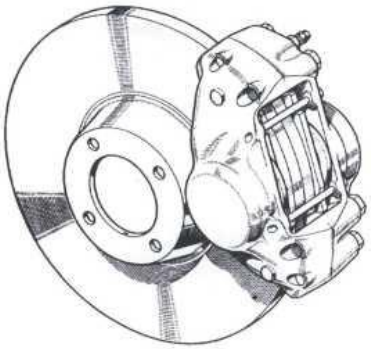
For the following data (a = 1.2r, β = π/2 and μ = 0.75) ψ can be calculated to approximately 4π/45.

If it is possible to bring the numerator in the fraction for ψ to zero, which will bring the angle ψ to zero, we would have obtained a symmetrical surface distribution. We also discover that the pivot point in that case is placed in the distance r_f from the center of the drum. From Figure 9.7 we can see that for practical reasons the angle θ has to be rather high. As an example an angle of θ = 7π/9 and a distance $a = r_f$ of 1.218r will cause the angle ψto be zero.

During the wear of the lining the distance a will decrease, which will cause a heavier wear of the lining at one of the shoe ends. Which one?

9.2 Disc brakes

Today hydraulic activated calliper disc brakes is the most commonly used type of brake for cars and motor cycles. Other applications where disc brakes are used almost exclusively are as parking and emergency brakes for wind turbines.

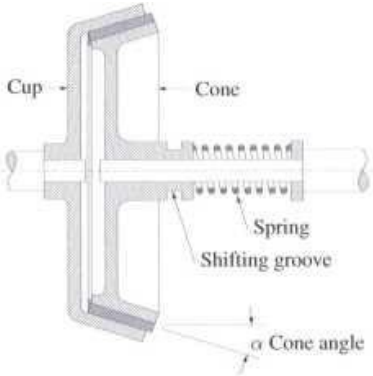


[billedtekst start]**Figure 9.10:** Calliper disc brake, hydraulic operated (Manufacture: Girling).[billedtekst slut]

The design principle is simple. The basic component is a cast ventilated disc, where air ventilation through the interior passages provides substantial additional cooling. The fixed calliper has hydraulic activated cylinders with brake pads on both sides of the disc. The activating force for a disc brake is substantially bigger than for a drum brake that make a brake amplifier necessary for cars.

9.3 Cone brakes

By using a cone brake the activating force will be less than for a disc brake for the same braking torque, but the cone angle must not be chosen too small to secure a safe disengagement.



[billedtekst start]**Figure 9.11**: Simplified principle drawing of a cone clutch, the same principle is used in the brake.[billedtekst slut]

A typical application is in hoists. For example in electrical chain hoists, where a helical compression spring will be activated and loaded in the same instant the motor is started. The magnetic field pulls the rotor axially into the stator by which the brake is released. On one hand it is required that the spring force is so large that the payload on the hook can be brought to stand-still quickly, and on the other hand the rotor should be able to overcome the spring force, when the motor is started.

Cone brakes are used in another well known application: the synchromesh in automobile gearboxes. Here it is used not as a cone brake, but as a cone clutch to synchronize the rotational speed for two shafts, before a claw coupling can engage and take over the torque transmission.

The cone brake uses wedging action to increase the normal force on the brake lining. The area of a surface element is

$$dA = rd\theta \frac{dr}{\sin\alpha} \tag{9.31}$$

where α is the half-cone angle. The normal force on the small area element is

$$dF_e = pdA \tag{9.32}$$

The actuating force is the thrust component dW of the normal force dF_e, so that

$$dW = dF_e \cdot \sin\alpha = pdA \cdot \sin\alpha = prdrd\theta \tag{9.33}$$

Now we can express the actuating force

$$W = \int_{d/2}^{D/2} \int_0^{2\pi} prdrd\theta = 2\pi \int_{d/2}^{D/2} prdr \tag{9.34}$$

and finally the braking torque is

$$T = \int_A \mu r dF_e = \frac{2\pi}{\sin\alpha} \int_{d/2}^{D/2} \mu pr^2 dr \tag{9.35}$$

9.3.1 Uniform pressure model

For uniform pressure distribution where $p = p_0$ the integration of (9.34) gives

$$W = \frac{\pi p_0}{4}(D^2 - d^2) \tag{9.36}$$

Similarly the torque is found as

$$T = \frac{2\pi p_0 \mu}{3\sin\alpha}\left(\frac{1}{8}\right)(D^3 - d^3) = \frac{\pi p_0 \mu}{12\sin\alpha}(D^3 - d^3) \tag{9.37}$$

or

$$T = \frac{\mu W(D^3 - d^3)}{3\sin\alpha(D^2 - d^2)} \tag{9.38}$$

9.3.2 Uniform wear model

Assuming uniform wear with $pr = C$ (C is a constant) gives the actuating force as

$$W = 2\pi c \int_{d/2}^{D/2} dr = \pi c(D - d) \tag{9.39}$$

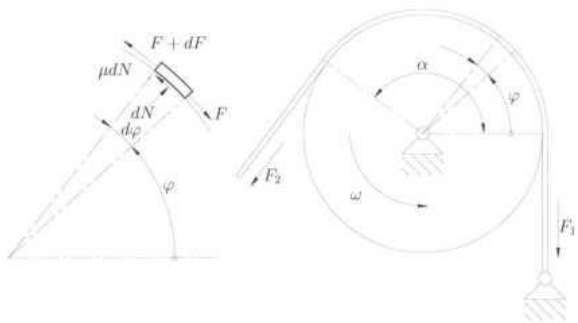
and the torque is

$$T = \frac{2\pi\mu c}{\sin\alpha}\int_{d/2}^{D/2} r dr = \frac{\pi\mu c}{4\sin\alpha}(D^2 - d^2) \tag{9.40}$$

$$T = \frac{\mu W}{4\sin\alpha}(D + d) \tag{9.41}$$

9.4 Band brakes

The band brake consists of a band pulled over a drum. By pulling in one end of the band with the force F_2, the brake is activated. When the drum rotates as shown in Figure 9.12, the force will increase over the contact angle from a value F_2 to a higher value F_1. The braking torque will consequently be $(F_1 - F_2)r$.



[billedtekst start]**Figure 9.12:** Band brake sketch and forces on an infinite (small) element.[billedtekst slut]

In order to calculate the force F_1, the forces on an infinite band element positioned an angle φ from the beginning of the contact angle α is analyzed. Forces are applied to a small cut-out element so that the increase in force (dF) is calculated in the same direction as the increase in angle ($d\varphi$).

Projection on the tangent direction

$$dF = -\mu dN \tag{9.42}$$

Projection on the radius direction

$$dN = 2F\sin\frac{d\varphi}{2} + dF \cdot \sin\frac{d\varphi}{2} \tag{9.43}$$

In this equation two simplifying assumptions are to be introduced

$$\sin\frac{d\varphi}{2} \text{ is approximated with } \frac{d\varphi}{2} \tag{9.44}$$

$$dF \cdot \sin\frac{d\varphi}{2} \approx 0 \tag{9.45}$$

In that way (9.43) is reduced to

$$dN \approx F d_{\varphi} \tag{9.46}$$

This expression for dN is inserted in (9.42)

$$dF = -\mu F d\varphi \tag{9.47}$$

or after separating the variables

$$\frac{dF}{F} = -\mu d\varphi \tag{9.48}$$

Integrated over the complete covering angle α we find

$$\int_{F_1}^{F_2} \frac{dF}{F} = \int_0^{\alpha} -\mu d\varphi \tag{9.49}$$

$$[\ln F]_{F_1}^{F_2} = [-\mu\varphi]_0^{\alpha} \tag{9.50}$$

$$\ln F_2 - \ln F_1 = -\mu\alpha \tag{9.51}$$

$$\ln \frac{F_1}{F_2} = \mu\alpha \tag{9.52}$$

$$\frac{F_1}{F_2} = e^{\mu\alpha} \tag{9.53}$$

This last equation is normally called Eytelwein's equation.

Typical industrial applications were previously big cranes. Today only few applications are known. One is for automatic gearboxes for cars.

On the other hand the principle of Eytelwein is known of all sailors. He knows how to sling a rope a couple of times around a bollard on the quay or around the bitt on the ship and by this he is able to control the pull in the cable and when necessary let go.

9.5 Nomenclature

a	mm	Linear dimension on brake arm
b	mm	Linear dimension on brake arm
d	mm	Inner diameter of disc
r	mm	Radius of brake drum
r_f	mm	Friction radius. (Theoretical value.)
P	N/mm^2	Pressure
w	mm	Width of brake

z	–	Number of teeth on gear
C	N/mm	Constant
D	mm	Outer diameter of disc
E_{kin}	Nm	Kinetic energy in a system
F	N	Force
F_e	N	External force
F_1	N	Large force in brake band (tight part)

F_2	N	Small force in brake band (slack part)
F_f	N	Friction force on brake drum
I_1,I_2	kgm²	Moment of inertia
N	N	Normal force on brake drum
T	Nm	Torque
T_B	Nm	Actuation torque on brake
W	N	Actuation force
α	rad	Angle
λ	–	Amplification factor
μ	–	Coefficient of friction
ψ	rad	Angular position of maximum pressure
w	rad/s	Angular speed
w_L	rad/s	Angular speed of load

Chapter 10 Belt Drives

10.1 Introduction

The two most common speed reduction mechanisms in industry are belts drives and gears. The efficiency of a belt transmission is generally less than that of a gear transmission.

Table 10.1: Comparison between power transmission systems (Rough approximations).

Drive type	**Gear**	**Chain**	**Timing belt**	**V-belt**	**Flat belt**
Power transmission through	contact	contact	contact	friction	friction
Max power transmission [kW]					
Normal	3000	200	100	100	150
Extreme	70000	4000	400	4000	4000
Max torque [Nm]	10^8	10^6	10^4	10^4	10^4
Max speed v [m/s]					
Normal	50	10	40	25	60
Extreme	200	40	70	40	120
Max ratio					
Normal	7:1	6:1	8:1	8:1	5:1
Extreme	1000:1	10:1	12:1	15:1	20:1
Efficiency *[%]*	93 - 99	94-98	93-98	92 - 94	94-98

Power transfer between gears is enabled by the normal action/reaction force at the tooth contact, and friction plays only a minor role. The transfer of power in a v-belt and a flat belt drive requires friction. The tensions F_{max} and F_{min} in the two strands cause a normal pressure over the belt-pulley contact, and the corresponding distributed friction force gives a moment about the pulley center, which equilibrates the shaft torque T.

For gears, the speed reduction ratio and the torque amplification ratio are equal to the radius ratio, so that the output power equals the input power and the efficiency is close to 100%.

The speed ratio across a pair of gears always equals the ideal ratio, because of the kinematic restrictions (a so-called positive drive), whereas sliding friction results in a torque ratio, which is less than ideal.

If the creep between belt and pulleys is neglected the speed reduction ratio and the torque amplification ratio are equal to the radius ratio for a belt transmission as well. For a practical belt drive the torque ratio equals the ideal ratio, but creep results in a larger speed reduction ratio than the ideal. The creep is

due to belt elements changing length as they travel between F_{max} and F_{min}, and since the pulley is rigid there must be relative motion between belt and pulley.

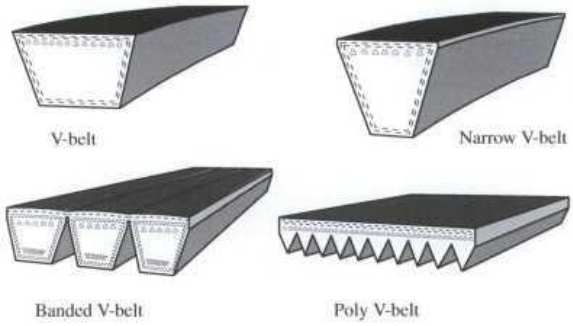
10.1.1 Reasons for choosing belt drives

Compared to other types of mechanical transmissions the following advantages can be mentioned:

- Low maintenance cost. (No lubrication needed)
- Within reasonable limits: Free choice of ratio and center distance.
- Easy to install. Low requirement to shafts line-up.
- High degree of reliability.
- Low noise level.
- Built in elasticity in the transmission.
- Easy handling of spare parts.
- Inexpensive.

10.2 The belts

Modern flat belts are of composite construction with cord reinforcement. They are particularly used for belt conveyors that are typically used to convey bulk material of all kinds. The smallest conveyors are transporting a few grams per hour whereas the biggest may transport more than 1000 tons per hour.



[billedtekst start]**Figure 10.1:** Schematic illustrations of different types of V-belts.[billedtekst slut]

Flat belt power transmission drives can be used up to 4000kW, but causes very high loads to the shaft and the bearings. The V-belt drives causes less shaft forces due to the V-grooved pulleys, and due to the newer wedged type of V-belts, the space requirements compared to a gear transmission are less dramatic and they are normally much cheaper than the toothed gear.

If a single V-belt is inadequate for power transmission then multiple belts and corresponding multigrooved pulleys are used.

V-belts normally comprise cord tensile members located at the pitch line, embedded in a softer matrix that is in turn embedded in a wear resistant cover.

The groove semi-angle lies usually in the range $4\pi/45 \leq \gamma \leq 19\pi/180$.

V-belts are available in a number of standard cross-sectional sizes, designated in order of increasing size A, B, etc, for classical V-belts and as SPA, SPB, etc. for the narrow type of belts. Each size is suitable for a particular power range.

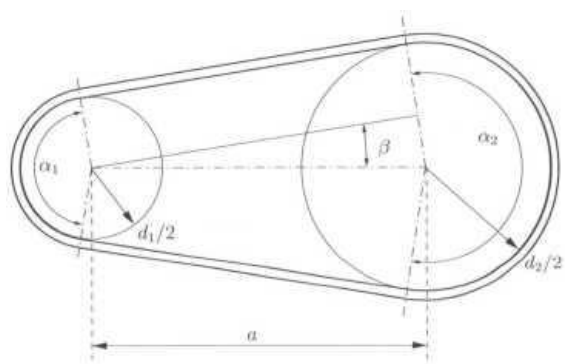
V-belts are manufactured in certain discrete standard pitch lengths. The power demand very often necessitates a number of belts on multi-grooved pulleys.

Before a belt drive can transmit a torque and thereby power, an initial tension F_0 must be applied to the belt by the shafts being pulled apart and then fixed in position.

Drive commences by applying a torque T_1 to the shaft of the small driving pulley, causing it to rotate at a steady speed n_1 [rev/s]. The tension in the 'tight' straight strand will exceed F_0, while the tension in the 'slack' strand will become less than F_0. This tension difference applies a torque T_2 to the driven pulley that rotates at uniform speed n_2.

10.3 Belt drive geometry (kinematics)

A typical belt drive is illustrated in Figure 10.2. The diameter of the small driving pulley is d_1 and that of the large driven pulley is d_2.



[billedtekst start]**Figure 10.2:** A typical belt drive.[billedtekst slut]

The pulley diameters and belt length are discrete variables giving a theoretical center distance a. The drive design must be capable to allow for belt installation and initial tightening.

If we express the speed of the belt on each pulley as function of the number of revolutions (n_1[rev/s] and n_2[rev/s]) we find

$$\upsilon_1 = \pi(d_1 + t)n_1 \tag{10.1}$$

$$\upsilon_2 = \pi(d_2 + t)n_2 \tag{10.2}$$

where t is the thickness of the belt. We may now define the creep ψ as

$$v_2 = v_{1(1-\psi)} \qquad (10.3)$$

which express the reduction in the speed of the belt on the driven pulley as compared on the driver pulley. By rearranging we find

$$\psi = \frac{v_1 - v_2}{v_1} = 1 - \frac{(d_2 + t)n_2}{(d_1 + t)n_1} \qquad (10.4)$$

The creep is not to be confused with sliding. The reason for the creep is that the belt tension has a different value on each side of the pulley and therefore different elongation. This transition of elongation takes place during the belts contact with the pulley. The size of the creep is in the order of a few percent (say 2%).

We may now express the speed ratio

$$i = \frac{n_2}{n_1} = \frac{(d_1 + t)v_2}{(d_2 + t)v_1} = \frac{(d_1 + t)(1 - \psi)}{d_2 + t} \qquad (10.5)$$

If we now make the assumption that the creep is negligible and the thickness of the belt is small as compared to the diameter of the pulleys, we can simplify and find

$$i \approx \frac{d_1}{d_2} \qquad (10.6)$$

It also follows from the assumption of zero creep that the belt speed can be approximated as

$$\upsilon \approx \pi d_1 n_1 \approx \pi d_2 n_2. \qquad (10.7)$$

Normally the belt speed v should be less than 30m/s for the usual cast iron pulley material. Most V-belts are designed to have optimum performance at speeds of around 20m/s.

The length of the belt can be found by using Figure 10.2. The angles $a\backslash$ and a_2 is

$$\alpha_1 = \pi - 2\beta \qquad \alpha_2 = \pi + 2\beta \qquad (10.8)$$

this leads to the length

$$L = \alpha_1 \frac{d_1}{2} + \alpha_2 \frac{d_2}{2} + 2a\cos\beta \qquad (10.9)$$

reformulation gives

$$L = \frac{\pi}{2}(d_1 + d_2) + \beta(d_2 - d_1) + 2a\cos\beta \qquad (10.10)$$

The angle β can be derived from

$$\frac{1}{2}(d_2 - d_1) = a\sin\beta \quad \Rightarrow \quad \sin\beta = \frac{d_2 - d_1}{2a} \qquad (10.11)$$

The presented equation are useful, when we have d_1, d_2 and a and want to find L. If instead L is known and we want to find a the same formulas could be used. However often you will find alternative formulas based on the assumption

$$d_2 - d_1 \ll a \quad \Rightarrow \quad \beta \ll 1 \quad \Rightarrow \quad \sin\beta \approx \beta \quad \text{and} \quad \cos\beta \approx 1 - \frac{\beta^2}{2} \tag{10.12}$$

Using (10.12) in (10.11) and (10.10) we have

$$\beta = \frac{d_2 - d_1}{2a}$$

$$L = \frac{\pi}{2}(d_1 + d_2) + \beta(d_2 - d_1) + 2a(1 - \frac{\beta^2}{2})$$

which lead to the following second order polynomial in a

$$\frac{1}{2}a^2 + (\frac{\pi}{8}(d_1 + d_2) - \frac{L}{4})a + \frac{(d_2 - d_1)^2}{16} = 0 \tag{10.13}$$

Solving (10.13) with respect to a gives

$$a = a_1 + \sqrt{a_1^2 - a_2^2} \tag{10.14}$$

where

$$a_1 = \frac{L}{4} - \frac{\pi}{8}(d_1 + d_2) \tag{10.15}$$

$$a_2 = \frac{d_2 - d_1}{2\sqrt{2}} \tag{10.16}$$

10.4 Belt forces

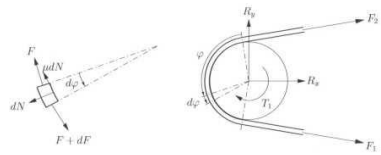
Following the geometry consideration we now use force equilibrium to find the maximum possible transmitted torque.

When a rope is wrapped around a stationary cylinder, it can remain in equilibrium although its ends are pulled with different forces, provided that one end is not pulled excessively. If the difference in the applied forces is sufficient to overcome the friction slip occurs and the rope slides around the cylinder. The belt behaves in a similar way and this can be used to find out the friction limited torque capacity.

10.4.1 Flat belt

Consider first a flat belt wrapped around a stationary pulley of radius $d_1/2$ as shown in Figure 10.3, the contact extending from $\varphi = 0$ on the slack side where the force is F_2 to $\varphi = \alpha_1$ on the tight side where the force is F_1, i.e. $F_1 > F_2$. We initially assume that the belt slide on the pulley.

Because the belt is sliding on the pulley, we know the friction force and may express the transmitted torque



[billedtekst start]**Figure 10.3:** Free body diagram of one pulley, and a infinitesimal piece of the belt. It is assumed here that the belt slip counter clockwise on the pulley, which result in the shown friction force μdN.[billedtekst slut]

$$T_1 = (F_1 - F_2)\frac{d_1}{2} \tag{10.17}$$

Initially, we assume that the belt speed is so low that acceleration can be neglected. From the small cut-out of the belt shown in Figure 10.3 the force equilibrium in tangential direction is

$$(F + dF)\cos(\frac{d\varphi}{2}) - F\cos(\frac{d\varphi}{2}) - \mu dN = 0 \tag{10.18}$$

This can be reduced to

$$dF = \mu dN \tag{10.19}$$

Equilibrium in the radial direction gives

$$(F + dF)\sin(\frac{d\varphi}{2}) + F\sin(\frac{d\varphi}{2}) - dN = 0 \tag{10.20}$$

This can be reduced to

$$Fd\varphi = dN \tag{10.21}$$

Combining (10.19) and (10.21) gives

$$\frac{dF}{F} = \mu d\varphi \tag{10.22}$$

integrating over the contact angle α_1 gives Eytelwein's equation

$$\ln\frac{F_1}{F_2} = \mu\alpha_1 \tag{10.23}$$

or

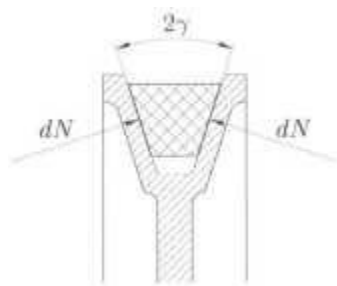
$$\frac{F_1}{F_2} = e^{\mu\alpha_1} \tag{10.24}$$

In the above deduction it is assumed that the driving pulley, with index 1, is the smallest one, giving the smallest wrap angle ($\alpha_1 < \alpha_2$). If this is not the case the angle α_1 should be substituted by α_2. Correspondingly, the coefficient of friction may be different for the two pulleys. In (10.24) we have the theoretical maximum difference between the force on the slack and tight side, i.e., if we express the maximum and minimum force as and F_{max} and F_{min} respectively we have

$$\frac{F_{max}}{F_{min}} = e^{\mu\alpha_1} \tag{10.25}$$

10.4.2 V-belt

If we now consider a V-belt the situation is slightly changed. The element contacts the two inclined sides of the groove, giving rise to normal reactions *dN* shown in Figure 10.4 together with a friction force μ*dN* at each contact. Since the two normal reaction components lying parallel to the pulley shaft in Figure 10.4 equilibrate one another, the resultant of the two contacts appear as in Figure 10.4. Force equilibrium gives



[billedtekst start]**Figure 10.4**: Belt forces.[billedtekst slut]

$$(F + dF)\cos\frac{d\varphi}{2} - F\cos\frac{d\varphi}{2} - 2\mu dN = 0 \tag{10.26}$$

and

$$(F + dF)\sin\frac{d\varphi}{2} + F\sin\frac{d\varphi}{2} - 2dN\sin\gamma = 0 \tag{10.27}$$

these two equations yield

$$\frac{dF}{F} = \frac{\mu}{\sin\gamma}d\varphi \tag{10.28}$$

where μ is the coefficient of friction between belt and groove.

Integrating over the contact angle α_1 gives

$$\ln\frac{F_1}{F_2} = \frac{\mu}{\sin\gamma}\alpha_1 \tag{10.29}$$

or

$$\frac{F_1}{F_2} = e^{(\mu/\sin\gamma)\alpha_1} \tag{10.30}$$

the same comment as under the flat belt concerning the validity of (10.30) applies, i.e. with respect to the wrap angle and the coefficient of friction.

10.4.3 Including inertia

Flat belts For belt transmissions operating at higher speeds the "useful belt tension" is influenced by the centrifugal forces acting on the belt, see Figure 10.5. Due to the centrifugal force the tension in the belt increases and the contact force between belt and pulleys decreases.

[billedtekst start]**Figure 10.5:** Free body diagram and kinetic diagram of small cut-out of belt.[billedtekst slut]

The force balance along a tangent to the pulley, including the centrifugal effect, gives

$$(F + dF)\cos\frac{d\varphi}{2} - F\cos\frac{d\varphi}{2} - \mu dN = 0 \tag{10.31}$$

which again is reduced to

$$dF = \mu dN \tag{10.32}$$

since there is no angular acceleration normally (assumed here).

The mass of the infinitesimal belt piece in Figure 10.5 is

$$dm = (w\,t\,\frac{d_1}{2}\,d\varphi)\rho = q\,\frac{d_1}{2}\,d\varphi \tag{10.33}$$

where w is the width of the belt, t is the thickness and ρ is the density of the belt material, $q = w\,t\,\varrho$ is defined as the mass pr. length of the belt.

The force equilibrium in radial direction is

$$(F + dF)\sin\frac{d\varphi}{2} + F\sin\frac{d\varphi}{2} - dN = dm\,\frac{v^2}{d_1/2} \quad \Rightarrow$$

$$(F + dF)\sin\frac{d\varphi}{2} + F\sin\frac{d\varphi}{2} - dN - qv^2 d\varphi = 0 \tag{10.34}$$

This can be reduced to

$$Fd\varphi - dN - qv^2 d\varphi = 0 \tag{10.35}$$

and

$$dN = Fd\varphi - qv^2 d\varphi \quad \Rightarrow$$
$$\frac{1}{\mu} dF = (F - qv^2) d\varphi \tag{10.36}$$

from which the variables can be separated and integrated

$$\int_{F_2}^{F_1} \frac{dF}{F - qv^2} = \int_0^{\alpha_1} \mu d\varphi \tag{10.37}$$
$$ln \frac{F_1 - qv^2}{F_2 - qv^2} = \mu\alpha_1 \tag{10.38}$$
$$\frac{F_1 - qv^2}{F_2 - qv^2} = e^{\mu\alpha_1} \tag{10.39}$$

this is called the extended Eytelwein's equation.

V-belts For V-belt transmissions the same equation can be used with implementation of a new artificial friction coefficient μ_e with

$$\mu_e = \frac{\mu}{\sin\gamma} \tag{10.40}$$

as the effective coefficient of friction which reflects the amplification of the actual coefficient μ by the wedging action in the groove of angle γ.

It is important to note that the tension in the belt is F_1 and F_2, but the force that acts on the pulley is reduced by the speed of the belt to $F_1 - qv^2$ and $F_2 - qv^2$ respectively.

Large pulleys and tension ratios are preferred, but tension ratios are limited by the friction coefficients and wrap angles that are encountered in practice.

Drive capacity may be increased by increasing the initial tension, but this will in turn reduce the belts fatigue live. Large pulleys reduce loads, but the cost of the pulleys themselves increases with size. The manufacturers of electric motors recommend minimum pulley diameters for acceptable motor bearing lives.

In the most simple setup a drive comprises two pulleys, which may have different values of friction coefficient μ and wrap angle α. In order to fully exploit the friction capability between the belt and the pulley, it is important that

$$F_2 \geq 0 \tag{10.41}$$

It is common practice to specify the pretension as

$$F_0 = \frac{F_{max} + F_{min}}{2} \tag{10.42}$$

If the pretension level is sufficient the maximum possible transmitted torque is given by

$$T_{max} = ((F_{max} - qv^2) - (F_{min} - qv^2))\frac{d_1}{2} = (F_{max} - F_{min})\frac{d_1}{2} \tag{10.43}$$

and the forces must fulfil

$$\frac{F_{\max} - qv^2}{F_{\min} - qv^2} = e^{(\mu_e \alpha)_{\min}} \tag{10.44}$$

Where

$$(\mu_e\alpha)\min = \min(\mu_e\alpha_1, \mu_e\alpha_2) \tag{10.45}$$

and in the case of v-belt drive we must use

$$\mu_e = \mu / \sin\gamma \tag{10.46}$$

If μ_e is the same for both pulleys, as is normal when both pulleys are grooved, then the smaller pulley will be limiting since $\alpha_1 \leq \alpha_2$.

The maximum power that can be transmitted by a single belt is

$$P_{\max} = T_{\max}\omega = T_{\max}\frac{v}{d_1/2} = (F_{\max} - F_{\min})v \tag{10.47}$$

If we alternatively assume a number z of equally size belts we have

$$P_{max} = z(F_{\max} - F_{\min})v \tag{10.48}$$

or

$$F_{\max} - F_{\min} = \frac{P_{\max}}{zv} \tag{10.49}$$

This relation is insufficient for determining $F_{\max}$, F_{min} individually for a given power per belt *(Pmax/z)* and velocity v. However if we apply the extended Eytelwein's equation we achieve

$$F_{\max} = \frac{P_{\max}}{zvk_\varphi} + qv^2 \tag{10.50}$$

$$F_{\min} = \frac{P_{\max}(1 - k_\varphi)}{zvk_\varphi} + qv^2 \tag{10.51}$$

where

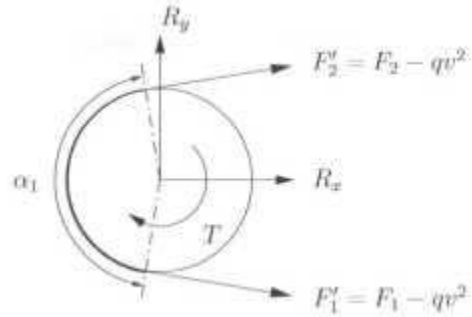
$$k_\varphi = 1 - \frac{1}{e^{(\mu_e\alpha)_{\min}}} = \frac{e^{(\mu_e\alpha)_{\min}} - 1}{e^{(\mu_e\alpha)_{\min}}} \tag{10.52}$$

it is seen that these expressions relate to the boundary of slip and therefore the applied real values should be less.

From a design point of view it is of interest to know the radial force acting on the pulley shafts. This is found by vectorial summation of the belt forces. In Figure 10.6 a free body diagram of a pulley is shown. The force from the belt acting on the pulley includes the contribution from the speed of the belt. The resulting forces are termed F'and F'.

The total load on the shaft is $F_t = \sqrt{R_x^2 + R_y^2}$ or expressed in the resulting forces from the belt on the pulley

$$F_t = \sqrt{F_1'^2 + F_2'^2 - 2F_1'F_2'\cos\alpha_1} \tag{10.53}$$



[billedtekst start]**Figure 10.6:** Free body diagram of pulley.[billedtekst slut]

or alternatively by defining the ratio between the loads $\epsilon = F_1'/F_2'$ and the difference $\kappa = F_1' - F_2'$ as

$$F_t = \kappa \frac{\sqrt{\epsilon^2 + 1 - 2\epsilon \cos\alpha_1}}{\epsilon - 1} = \kappa K \tag{10.54}$$

i.e., expressed as a number (K) times the load difference. We also note that the load on the shaft is largest at stand still.

10.5 Belt stress (flat belt)

The belt experience a variation in the stress as it goes through a full revolution. In this section the different contributions to the stress is discussed.

Pretension of the belt results in a stress

$$F_0 = \frac{F_1 + F_2}{2} \quad \Rightarrow \quad \sigma_0 = \frac{F_0}{A} \tag{10.55}$$

where $A = w\ t$ is the cross sectional area of the belt. We have previously shown that there is a contribution to the stress from the speed of the belt. This stress is

$$F_c = qv^2 \quad \Rightarrow \quad \sigma_c = \frac{F_c}{A} \tag{10.56}$$

The stress (10.56) can be considered to be constant in the belt. There must also be a stress difference that enables the torque transfer. This stress difference is

$$F_d = F_1 - F_2 \quad \Rightarrow \quad \sigma_d = \frac{F_d}{A} \tag{10.57}$$

The forces in the two free spans of the belt are F_1 and F_2. The forces on the pulleys are defined as $F_1' = F_1 - F_c$ and $F_2' = F_2 - F_c$ and the corresponding stresses

$$\sigma_1 = \frac{F_1'}{A} \qquad \sigma_2 = \frac{F_2'}{A} \tag{10.58}$$

it then follows that

$$\frac{F_1}{A} = \sigma_1 + \sigma_c = \sigma_0 + \frac{\sigma_d}{2} \tag{10.59}$$

$$\frac{F_2}{A} = \sigma_2 + \sigma_c = \sigma_0 - \frac{\sigma_d}{2} \tag{10.60}$$

From these equations we find that (σ_d can be given as

$$\sigma_d = \sigma_1 - \sigma_2 \tag{10.61}$$

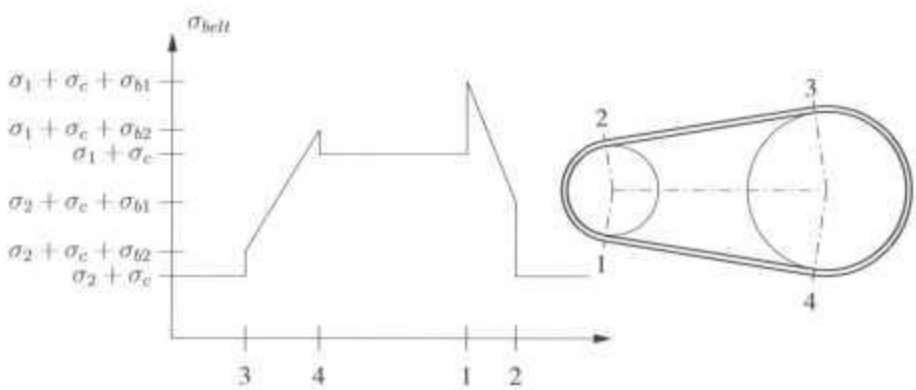
On the free slack part of the belt we have the stress $a_2 + {}^a c$ and on the tight free part have the stress $\sigma_1+\sigma_c = \sigma_2+\sigma_c+\sigma_d$, i-e. the stress must build up on one pulley by σ_d and reduce by the same amount on the other pulley. This is the cause of creep.

Finally, there is also bending stresses in the belt due to bending around the pulleys. This stress is directly related to the diameters of the pulley and therefore different on the two pulleys

$$\sigma_{b1} = \frac{Ey}{d_1/2} \qquad \sigma_{b2} = \frac{Ey}{d_2/2} \tag{10.62}$$

where E is the modulus of elasticity of the belt and y is the distance from the pulley, which leads to the maximum values in the outer fiber

$$\sigma_{b1\,\max} = \frac{Et}{d_1} \qquad \sigma_{b2\,\max} = \frac{Et}{d_2} \tag{10.63}$$



[billedtekst start]**Figure 10.7:** Principal graph of stress in belt, it is assumed that the transition of creep is linear.[billedtekst slut]

By adding up all the contributions we may show a principal graph of the stress in the belt. In principle the stress cycle shown in Figure 10.7 gives a double cycle of loads with each complete rotation that should be dealt with in fatigue calculations of the belt. The cycle is however, highly related to the individual design of the belt drive that may not be as simple as illustrated here. Specific calculations should therefore be done for actual designs.

10.6 Optimization of belt-drives

From $F_0 = (F_1 + F_2)/2$ and (10.39) together with the power given as

$$P = (F_1 - F_2)v \qquad (10.64)$$

we may express the limiting power in different ways.

$$P_{\max} = \frac{2(e^{(\mu_e\alpha)_{\min}} - 1)}{e^{(\mu_e\alpha)_{\min}} + 1}(F_0 - qv^2)v \qquad (10.65)$$

$$P_{\max} = \frac{e^{(\mu_e\alpha)_{\min}} - 1}{e^{(\mu_e\alpha)_{\min}}}(F_1 - qv^2)v \qquad (10.66)$$

$$P_{\max} = (e^{(\mu_e\alpha)_{\min}} - 1)(F_2 - qv^2)v \qquad (10.67)$$

Using (10.65) we may express the optimal belt speed for given pretension F_0 by using that

$$\frac{dP_{\max}}{dv} = \frac{2(e^{(\mu_e\alpha)_{\min}} - 1)}{e^{(\mu_e\alpha)_{\min}} + 1}(F_0 - 3qv^2) = 0 \qquad (10.68)$$

which gives

$$v_{opt} = \sqrt{\frac{F_0}{3q}} \qquad (10.69)$$

If instead the tension load F_1 is to be minimized we rewrite (10.66) to find

$$F_1 = \frac{P_{\max}}{v}\frac{e^{(\mu_e\alpha)_{\min}}}{e^{(\mu_e\alpha)_{\min}} - 1} + qv^2 \qquad (10.70)$$

which gives

$$\frac{dF_1}{dv} = 0 \quad \Rightarrow \quad v = \sqrt[3]{\frac{P_{\max}}{2q}\frac{e^{(\mu_e\alpha)_{\min}}}{e^{(\mu_e\alpha)_{\min}} - 1}} \qquad (10.71)$$

Finally, it is possible to minimize the bearing load. A crude assumption of the bearing load is

$$F_l = F_1' + F_2' \qquad (10.72)$$

Rewriting (10.65) we get

$$P_{\max} = \frac{e^{(\mu_e\alpha)_{\min}} - 1}{e^{(\mu_e\alpha)_{\min}} + 1}(F_1' + F_2')v \qquad (10.73)$$

or that

$$F_1' + F_2' = \frac{P_{\max}}{v}\frac{e^{(\mu_e\alpha)_{\min}} + 1}{e^{(\mu_e\alpha)_{\min}} - 1} \qquad (10.74)$$

From (10.74) we see that in order to minimize the bearing load we have to maximize $(\mu_e\alpha)_{\min}$ and maximize v.

10.7 Plot of the belt forces

In many cases a plot of the involved belt forces is very illustrative. Universally the following equilibrium holds for all configurations

$$P = (F_1' - F_2')v = (F_1 - F_2)v \tag{10.75}$$

Secondly we can express the force on the slack part of the belt as

$$F_2' = \phi\frac{P}{v} + F_0' \tag{10.76}$$

where ϕ is a givenconstant that depends on the configuration. F_0' is the belt preload. Finally the belt forces must follow Eytelwein's equation

$$\frac{F_1'}{F_2'} \leq e^{\mu\alpha} \tag{10.77}$$

Using (10.75) and (10.77) the limit for F_1' and F_2' are given by

$$F_1' \geq \frac{P}{v}\frac{e^{\mu\alpha}}{e^{\mu\alpha}-1} \tag{10.78}$$

$$F_2' \geq \frac{P}{v}\frac{1}{e^{\mu\alpha}-1} \tag{10.79}$$

Case I. In this case the mean value of the belt load is constant and equal to F_0', this corresponds to the case where there is no active tightening taking place. For this case the following apply

$$\phi = -\frac{1}{2} \Rightarrow$$

$$F_1' = \frac{1}{2}\frac{P}{v} + F_0' \tag{10.80}$$

$$F_2' = -\frac{1}{2}\frac{P}{v} + F_0' \tag{10.81}$$

In Figure 10.8 the resulting forces is given as a function of *P/v.*

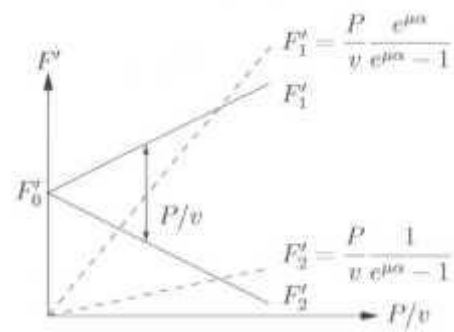
Case II. In this case there is constant force in the slack part of the belt, i.e. $F_2' = constant.$ For this configuration the following apply

$$\phi = 0 \Rightarrow$$

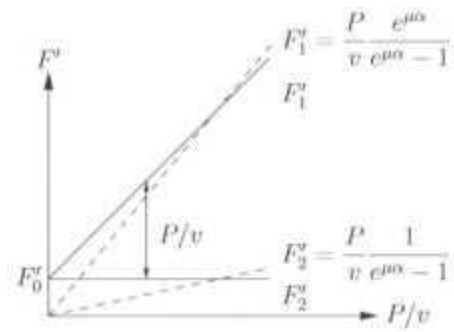
$$F_1' = \frac{P}{v} + F_0' \tag{10.82}$$

$$F_2' = F_0' \tag{10.83}$$

In Figure 10.9 the resulting forces is given as a function of *P/v.*



[billedtekst start]**Figure 10.8:** Belt forces as a function of P/v for the specific case $\phi = -1/2$.[billedtekst slut]



[billedtekst start]**Figure 10.9:**Belt forces as a function of P/v for the specific case $\phi = 0$.[billedtekst slut]

General case For the general case ϕ can have a positive or negative value. Using (10.79) in (10.76) we find

$$\frac{P}{v}\frac{1}{e^{\mu\alpha}-1} \leq \phi\frac{P}{v} + F_0^v \Rightarrow$$

$$\frac{P}{v}\left(\frac{1}{e^{\mu\alpha}-1} - \phi\right) \leq F_0^v \tag{10.84}$$

From this the limit to the power, P_{max}, assuming a given speed, v, can be found. Assuming that P is given result in a minimum speed, υ_{min}

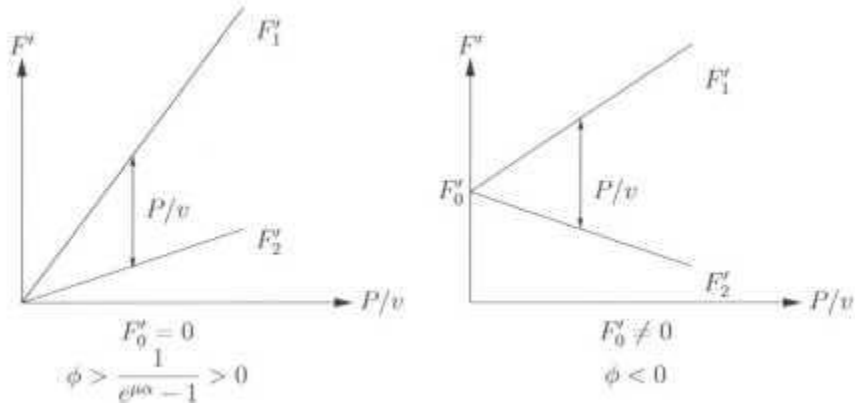
$$P_{max} = F_0^v v \frac{e^{\mu\alpha}-1}{1-\phi(e^{\mu\alpha}-1)} \tag{10.85}$$

$$v_{min} = \frac{P}{F_0^v}\frac{1-\phi(e^{\mu\alpha}-1)}{e^{\mu\alpha}-1} \tag{10.86}$$

In Figure 10.10 two examples of the general case is shown. If we select $\phi > I/(e^{\mu\alpha} - 1)$ then no preload is needed as seen in Figure 10.10.

Bearing loads At stand still we have the following two belt loads

$$F_2^v = F_0, \qquad F_1^v = F_0 \tag{10.87}$$



[billedtekst start]**Figure 10.10:** Two examples of belt forces for the general case.[billedtekst slut]

where

$$F_0 = F_0' + qv^2 \tag{10.88}$$

From the belt forces the bearing loads can be found. At idle, i.e. when $P = 0$ and $v \neq 0$the two belt loads are

$$F_2' = F_0', \qquad F_1' = F_0' \tag{10.89}$$

When the maximum power, P_{max}, is transmitted for a given speed, v, we have

$$F_2' = \frac{1}{e^{\mu\alpha}-1}\frac{P_{max}}{v}, \qquad F_1' = F_2' + \frac{P_{max}}{v} = \frac{e^{\mu\alpha}}{e^{\mu\alpha}-1}\frac{P_{max}}{v} \tag{10.90}$$

The minimum bearing loads are achieved for the specific case

$$\phi = \frac{1}{e^{\mu\alpha}-1} \quad \text{and} \quad F_0' = 0 \tag{10.91}$$

For Figures 10.8 to 10.10 it should be noted that P/v also indicates the transmitted torque since

$$\frac{P}{v} = \frac{T\omega}{v} = \frac{T\omega}{\omega r} = \frac{T}{r} \tag{10.92}$$

where r is the radius of the driving pulley.

10.8 Nomenclature

a	mm	Center distance between pulleys
d_1	mm	Diameter of driving (small) pulley
d_2	mm	Diameter of driven (large) pulley
i	–	Speed ratio for the pulleys

k_φ	–	The drive property
n_1	rev/s	Rotational speed of pulley 1
n_2	rev/s	Rotational speed of pulley 2
q	kg/m	The belt's mass density per unit length
t	m	The belt thickness.
υ	m/s	Belt speed
w	m	Belt width
z	–	Number of belts on a pulley
E	N/m^2	Modulus of elasticity
F_0	N	Initial tension in belt
F_1	N	Large belt force (tight part)
F_2	N	Small belt force (slack part)
F_1'	N	Large belt force on pulley (tight part)
F_2'	N	Small belt force on pulley (slack part)
Fmax	N	Maximal belt force (tight part)
F_{min}	N	Minimal belt force (slack part)
F_t	N	Total load on shaft
L	mm	Length of belt
N	N	Normal force on belt from pulley
P	W	Transmitted power
P_{max}	W	Maximum possible transmitted power
T_{max}	Nm	Maximal torque on driving pulley

T'_2	Nm	Torque on driven pulley
α_1	rad	Wrap angle on driving pulley
α_2	rad	Wrap angle on driven pulley
β	rad	Angle between belt and line between pulley centers
γ	rad	Groove semi-angle on pulley
μ	–	Coefficient of friction between belt and pulley
μ_e	–	Effective coefficient of friction between belt and pulley
w	rad/s	Angular speed
ρ	kg/m^3	Mass density of belt
ψ	–	Creep
σ_0	N/m^2	Belt stress due to pretension
σ_c	N/m^2	Belt stress due inertia
σ_d	N/m^2	Belt stress due to transmission of torque
σ_1	N/m^2	Belt stress in tight part
σ_2	N/m^2	Belt stress in slack part
σ_b	N/m^2	Bending stress in belt

10.9 References

[1] K. Casper. Kræfter i fladremtræk (in danish). DIAM nr. 645 (Technical University of Denmark), 1981.

[2] D. Muhs, H. Wittel, M. Becker, D. Jannasch, and J. Voβiek. *Roloff/Matek maschinenelemente, 16. Auslage.* Viewegs Fachbiivher der Technik, Wiesbaden, 2003.

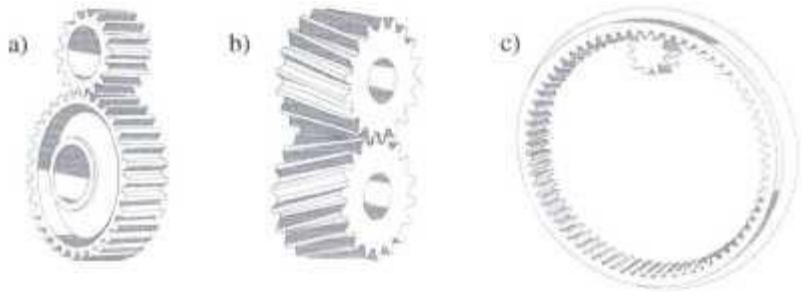
[3] G. Niemann and H. Winter. *Maschinenelemente Band III.* Springer-Verlag, Berlin, Heildelberg, New York, 1983.

[4] H. Palmgren. *Kileremshandboken.* Studentlitteratur, Lund, Sweden, 1985.

Chapter 11 The geometry of involute gears

11.1 Introduction

Of all the many types of machine elements that exist today, gears are among the most commonly used. The basic idea of a wheel with teeth is extremely simple and dates back several thousand years. It is obvious to any observer that one gear drives another by means of the meshing teeth, and to the person who has never studied gears, it might seem that no further explanation is required. It may therefore come as a surprise to discover the large quantity of geometric theory that exists on the subject of gears, and to find that there is probably no branch of mechanical engineering, where theory and practice are more closely linked. Enormous improvements have been made in the performance of gears during the last two hundred years or so, and this has been due principally to the careful attention given to the shape of the teeth. The theoretical shape of the tooth profile used in most modem gears is an involute. When precision gears are cut by modem gear-cutting machines, the accuracy with which the actual teeth conform to their theoretical shape is quite remarkable and far exceeds the accuracy that is attained in the manufacture of most other types of machine elements.

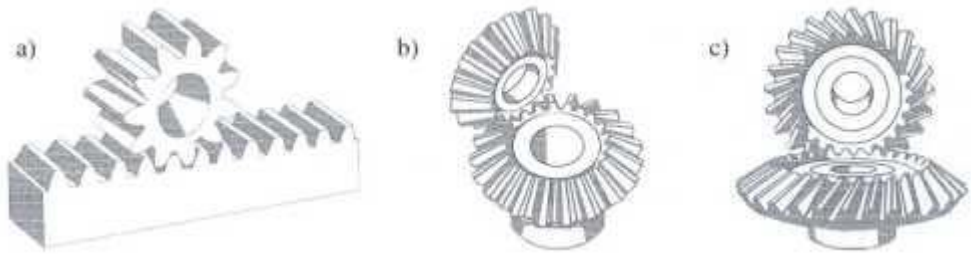


[billedtekst start]**Figure 11.1:** Gear types: a) Spur gear, b) Helix gear, c) Internal gear pair.[billedtekst slut]

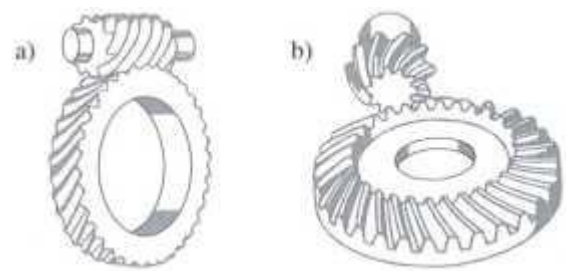
11.2 Internal and external gears

A spur gear is cut from a cylindrical blank, with teeth that are parallel to the gear axis. If the teeth face outwards, the gear is called an external gear, and if they face inwards, the gear is known as an internal gear. This chapter deals with the subject of external gears.

Gears are highly standardized products and this chapter mainly refers to the German standard DIN 3990 [2] and the corresponding ISO standard [3],



[billedtekst start]**Figure 11.2:** Gear types: a) Rack and pinion, b) and c) Bevel gear pair.[billedtekst slut]



[billedtekst start]**Figure 11.3:** Gear types: a) Worm gear, b) Hypoid gear.

11.3 Gear ratio

When describing gears, the term angular speed ratio is often used to describe the relation between the angular speed for the driving wheel, w_A, and the driven wheel, w_B,.

$$i = \frac{\omega_B}{\omega_A} = \pm\frac{r_A}{r_B} \tag{11.1}$$

where T_A is the pitch radius (see Figure 11.4) of the driving wheel and r_B is the radius of the driven wheel. In some textbooks one can also find the inverse definition of the angular speed ratio.

One can also define a mechanical advantage or torque ratio given by

$$t_r = \frac{T_B}{T_A} = \frac{1}{i} = \frac{\omega_A}{\omega_B} = \pm\frac{r_B}{r_A} \tag{11.2}$$

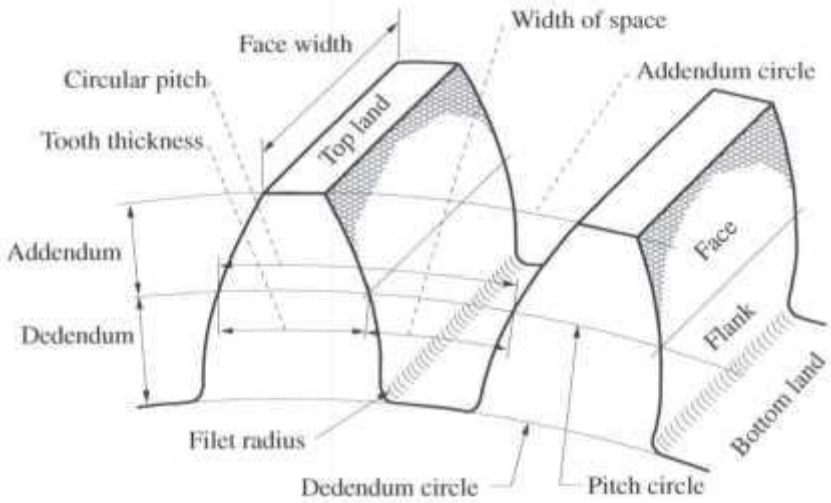
where T_A is the torque on the driving wheel, i.e. the input torque, and T_B is the torque on the driven wheel, i.e. the output torque. However, in conjunction with the geometrical calculations, it is today primarily the relation between the number of teeth on the larger gear wheel, z_2, and the smaller gear wheel, zi, that is used, this is defined as the gear ratio. The gear ratio is therefore given by

$$u = \frac{z_2}{z_1} \tag{11.3}$$

The smaller gear wheel is often called the pinion. It should be noted that the gear ratio u is equal to the magnitude of either i or t_r depending on which one is numerical larger than 1.

11.4 Gears in mesh

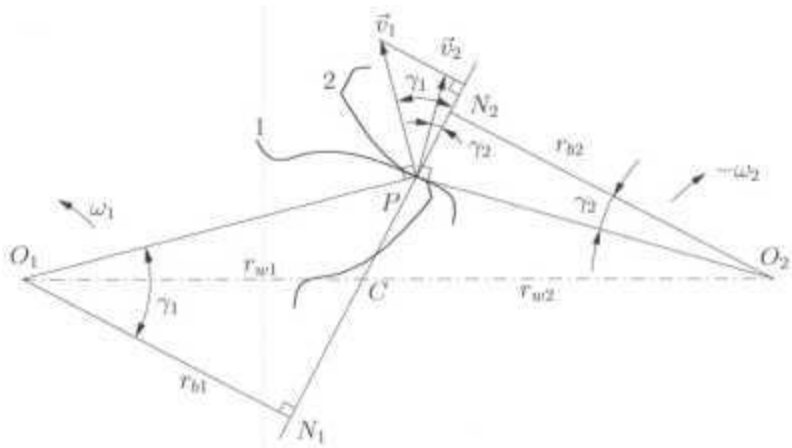
The pair of gears 1 and 2, turning around two fixed points 0_1 and 0_2 meet in point P. For the teeth is the current speed in point P



[billedtekst start]**Figure 11.4:** Gear nomenclature (Simplified geometry).[billedtekst slut]

$$|\vec{v}_1| = \overline{O_1P}\omega_1 \tag{11.4}$$

$$|\vec{v}_2| = -\overline{O_2P}\omega_2 \tag{11.5}$$



[billedtekst start]**Figure 11.5:** Meshing gears.[billedtekst slut]

In order for the gears not to move towards or away from each other, the velocity components in the direction normal to the teeth, along N_1 N_2, must be the same for the two gears. This means that

$$|\vec{v}_1| \cos\gamma_1 = |\vec{v}_2| \cos\gamma_2 \tag{11.6}$$

or

$$\overline{O_1P}\omega_1 \cos\gamma_1 = -\overline{O_2P}\omega_2 \cos\gamma_2 \tag{11.7}$$

At the same time we have

$$\overline{O_1P} = \frac{r_{b1}}{\cos\gamma_1} \tag{11.8}$$

and

$$\overline{O_2P} = \frac{r_{b2}}{\cos\gamma_2} \tag{11.9}$$

By inserting (11.8) and (11.9) in (11.7) we have

$$r_{b1}w1 = -r_{b2}w_2 \tag{11.10}$$

or

$$\frac{\omega_2}{\omega_1} = -\frac{r_{b1}}{r_{b2}} \tag{11.11}$$

Since O_1 N_1 and O_2N_2 are both perpendicular to the line of action, they each make the same angle with the line of centers, and we find that

$$\frac{r_{b1}}{r_{b2}} = \frac{r_{w1}}{r_{w2}} \tag{11.12}$$

We can therefore conclude that the condition for constant angular speed ratio w_1/w_2 is, that the gears common normal in the present point, constantly goes through the point *C*, which divides the center line O_1O_2 in the constant ratio rw_2/rw_1.

The angular speed ratio will be

$$i = \frac{\omega_2}{\omega_1} = -\frac{r_{w1}}{r_{w2}} \tag{11.13}$$

The two radii r_{w1} and r_{w2} are called the pitch radii. It should be noted that the formulas of the present section is for an external gear set, with an internal gear set the formulas are identical, but the sign of should be changed. This is the reason for the plus minus sign in (11.1) and (11.2), i.e., it relates to an internal or external gear set.

11.5 Tooth shapes

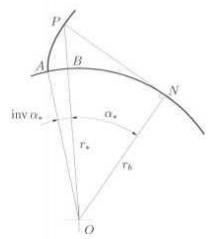
Usable gears must fulfil the following conditions in order to be of interest:

- The two gear wheels must have the same "distance" (pitch) between two following teeth.
- Furthermore the space between two teeth must be bigger than (or equal to) the tooth thickness on the other gear wheel.
- At last the conclusive condition mentioned in the previous section is to be fulfilled too.

These conditions can be fulfilled by several tooth shapes from which well-known types are: cycloid, Novikov and involute. Only the involute type will be explained here.

11.6 Involute tooth shape basics

For involute gears the tooth flank is shaped as an involute to a circle. As the circle radius increase the tooth shape tend to be straighter. A common reference for involute gears is the (basic) rack that can be described as a part of an involute gear with an infinite number of teeth. That is a gear wheel which radius is infinite. A reference of that type is relative easy to produce to a very high degree of precision as the tooth shape is absolutely straight.



[billedtekst start]**Figure 11.6:** Basic involute geometry.[billedtekst slut]

For the point P at the involute curve the curvature radius is given by the distance *PN* and the curvature center is in *N*.

Also to observe is that the length of the arc *AN* is equal to the length of *PN.*

$$\widehat{AN} = \overline{PN} = r_b \tan \alpha_* \qquad (11.14)$$

$$\widehat{BN} = r_b \alpha_* \qquad (11.15)$$

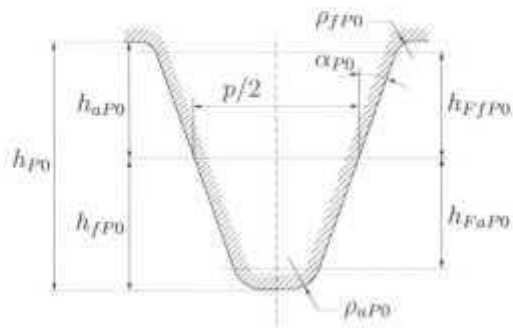
$$\widehat{AB} = \widehat{AN} - \widehat{BN} \qquad (11.16)$$

$$= r_b(\tan\alpha_* - \alpha_*) \text{ inv } \alpha_* = \tan \alpha_* - \alpha_* \qquad (11.17)$$

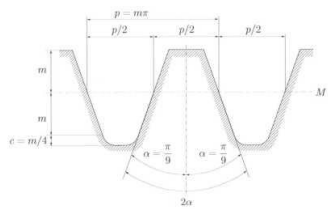
11.7 Basic rack

The basic rack, can be constructed as a gear wheel with an infinite number of teeth. Thus will all radii be infinite, and the circles will be straight lines. On the basic rack, the involute is a straight line, perpendicular to the line of action and the tooth flank will be a plane.

The basic rack is easy to manufacture to a high degree of precision. It is used as a reference profile for involute teeth, and gear wheels can be manufactured with a tool shaped as a slightly modified basic rack.



[billedtekst start]**Figure 11.7:** Tool shape for producing involute gears.[billedtekst slut]



[billedtekst start]**Figure 11.8:** Basic rack.[billedtekst slut]

11.8 Pitch and module

The standard pitch circle diameter d is given in terms of the basic rack pitch *p* (circular pitch)

$$d = \frac{p}{\pi} z \tag{11.18}$$

For a system of gears conjugate to a particular basic rack, it would therefore be necessary to specify only the value of the pitch *p*, which is the same for every gear in the system, and we would then use (11.18) to calculate the standard pitch circle diameter of each gear. This method of specification was in fact used in the past, and gears in which the circular pitch is specified as a convenient length are known as "circular pitch gears". However, they are seldom made today, as they have one slight disadvantage. If the value of the circular pitch is chosen as an integer number, the standard pitch circle diameter is always an inconvenient size, due to the presence of the factor π in (11.18). It has been found more practical to design gears in which the standard pitch circle diameter is an integer number. With this consideration in mind, we introduce a quantity called the module *m*, defined in terms of the basic rack pitch

$$m = \frac{p}{\pi} \tag{11.19}$$

We now combine (11.18) and (11.19), in order to express the standard pitch circle diameter in terms of the module

$$d = mz \tag{11.20}$$

and, since once again the circular pitch of the gear is equal to the basic rack pitch, a relation between the circular pitch and the module can be found immediately from (11.19)

$$p = \pi m \tag{11.21}$$

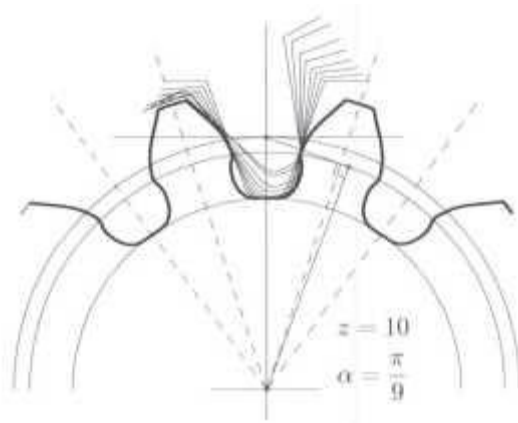
The module, which we have shown is proportional to the circular pitch, is used not only in the calculation of the standard pitch circle radius, but also as a measure of the tooth size. When two gears are meshed together, they must clearly have teeth of approximately the same size, and in practice they are designed with the same module m. In other words, the two gears are both conjugate to the same basic rack.

In countries where distances are measured in inches, not in mm, a system using the diametral-pitch (p_d) and circular-pitch (p) is used instead of the module.

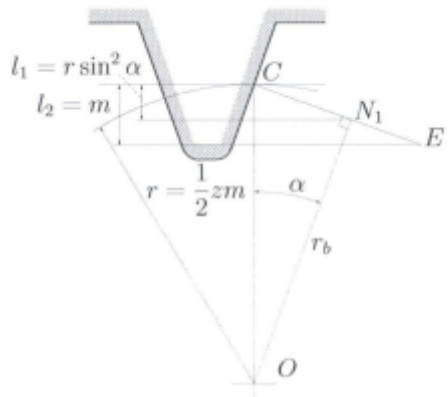
$$p_d = \frac{z}{d\,[\text{inches}]} = \frac{\pi}{p[\text{inches}]} \quad \left(= \frac{25.4}{m[\text{mm}]}\right) \tag{11.22}$$

11.9 Under-cutting

At small numbers of teeth, the rack will penetrate the base of the teeth, and under-cutting will occur.



[billedtekst start]**Figure 11.9:** Under cutting.[billedtekst slut]



[billedtekst start]**Figure 11.10:** Limiting the under-cut.[billedtekst slut]

Under-cutting will occur when the contact point is not between point C and point N_1 on the line of action, i.e., when there is contact further out than point N_1 on the line of action. So to avoid undercutting, l_2 must be smaller or equal to l_1, as shown 011 the Figure 11.10.

$$l_2 = m \tag{11.23}$$

$$l_1 = r\sin^2\alpha = \frac{1}{2}zm\sin^2\alpha \tag{11.24}$$

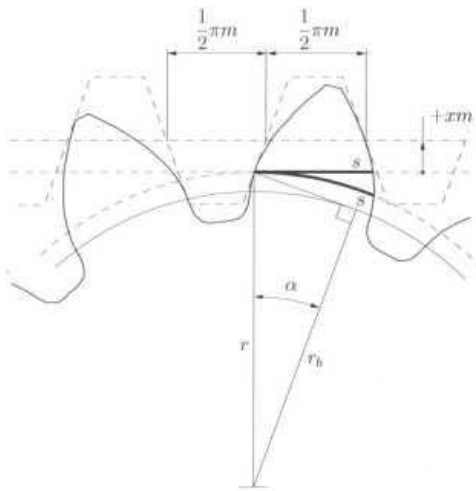
From the (11.23) and (11.24) we find that

$$z = \frac{2}{\sin^2 \alpha} \tag{11.25}$$

The minimum number of teeth before under-cutting will occur, z_{min}, is calculated as z rounded up, until the nearest integer. For $\alpha = \Pi/9$, $z_{min} = 17$. Under-cutting should be avoided, as the base of the teeth is weakened. How to do this, is explained below.

11.10 Addendum modification (profile shift)

Addendum modification is achieved by moving the rack the length *xm*, outwards from the wheel center, x is called the addendum modification factor. The sign for the addendum modification factor is defined so that positive addendum modification results in a thicker tooth. This should be understood as, the tooth base will be thicker, while the tooth tip will be thinner.



[billedtekst start]**Figure 11.11:** Addendum modification.[billedtekst slut]

At positive addendum modification, the minimum number of teeth to avoid under-cutting is changed so a smaller gear wheel without under-cutting can be obtained.

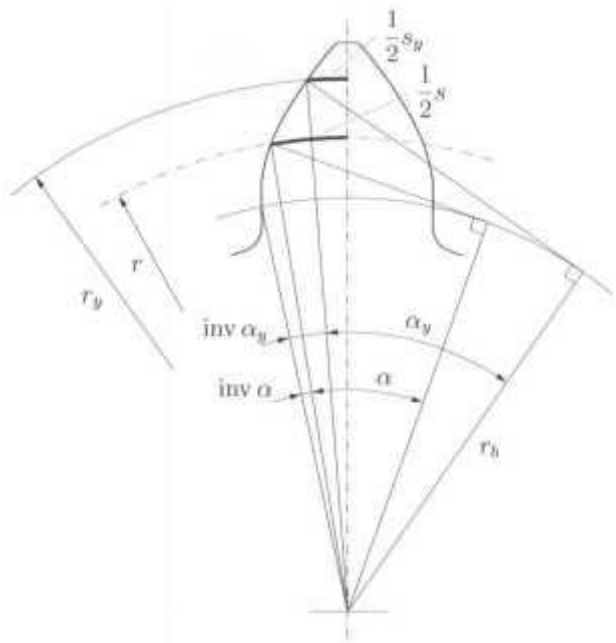
11.11 Tooth thickness

As a starting point, the tooth thickness, *s*, is calculated for the basic pitch circle.

$$s = \frac{1}{2}\pi m + 2xm \tan \alpha \tag{11.26}$$

The tooth thickness on radius r_y is calculated from the tooth thickness on the pitch circle *r*.

$$\frac{1}{2}s_y = \frac{1}{2}s\frac{r_y}{r} - r_y\left(\operatorname{inv}\alpha_y - \operatorname{inv}\alpha\right) \tag{11.27}$$



[billedtekst.start]**Figure 11.12:** Tooth thickness. (Basic pitch circle drawn with wrong radius for clearness in figure. Compare to Figure 11.11).[billedtekst.slut]

$$s_y = 2r_y\left(\frac{s}{2r} + \text{inv}\,\alpha - \text{inv}\,\alpha_y\right) \tag{11.28}$$

$$s_y = 2r_y\left(\frac{1}{z}\left(\frac{\pi}{2} + 2x\tan\alpha\right) + \text{inv}\,\alpha - \text{inv}\,\alpha_y\right) \tag{11.29}$$

The angle α_y is defined by the two expressions for the base circle radius r_b

$$r_b = r\cos\alpha = r_y\cos\alpha_y \tag{11.30}$$

$$\cos\alpha_y = \frac{r}{r_y}\cos\alpha \tag{11.31}$$

The tooth thickness at the tip is now found by inserting r_a in place of r_y. When designing gears, it is common to select the addendum modification, so that the tooth tip arc length $s_a \geq 0.25m$, however, for carburized gears $s_a \geq 0.4m$.

11.12 Calculating the addendum modification

For two gear wheels, defined by the module, m, the number of teeth, z_1 and z_2, and manufactured with a rack, in accordance with ISO/R 2334, the sum of the addendum modification factors, $x_1 + x_2$, for a given center distance, is calculated as follows: The sum of the thickness of the teeth on the pitch circle, s_{w1} and s_{w2}, equals the pitch on the same circle, p_w

$$s_{w1} + s_{w2} = Pw \tag{11.32}$$

furthermore, we have

$$s_{w1} = 2r_{w1}\left(\frac{1}{z_1}\left(\frac{\pi}{2} + 2x_1 \tan\alpha\right) + \operatorname{inv}\alpha - \operatorname{inv}\alpha_w\right) \tag{11.33}$$

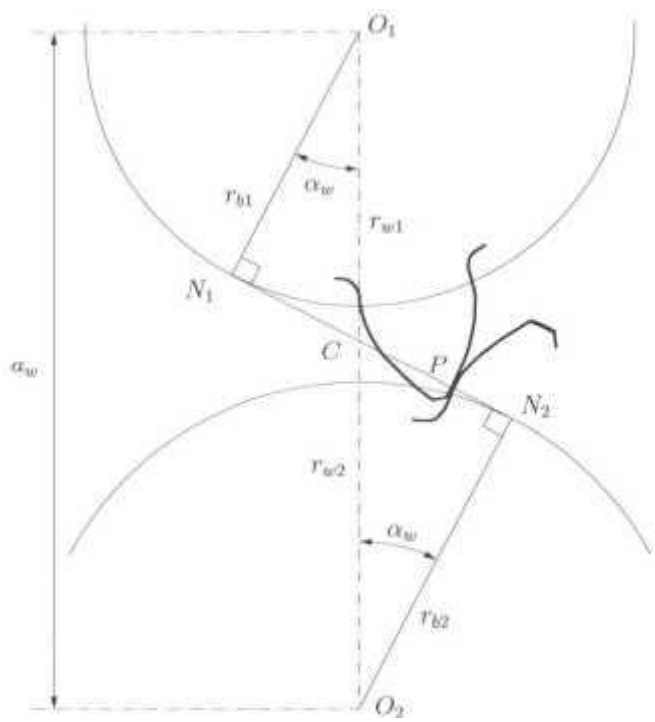
$$s_{w2} = 2r_{w2}\left(\frac{1}{z_2}\left(\frac{\pi}{2} + 2x_2 \tan\alpha\right) + \operatorname{inv}\alpha - \operatorname{inv}\alpha_w\right) \tag{11.34}$$

$$2r_{w1} = \frac{z_1 p_w}{\pi} \quad and \quad 2r_{w2} = \frac{z_2 p_w}{\pi} \tag{11.35}$$

By insertion in (11.32)

$$x_1 + x_2 = (z_1 + z_2)\frac{\operatorname{inv}\alpha_w - \operatorname{inv}\alpha}{2\tan\alpha} \tag{11.36}$$

please notice that the above calculation is based on the assumption that no clearance exist between the teeth. In practice, meshing gears must be calculated with clearance to allow oil or grease lubrication.



[billedtekst.start]**Figure 11.13:** Relations between center distance, radii and angles.[billedtekst.slut]

The transverse pressure angle α_w is found by

$$r_{w1} = \frac{r_{b1}}{\cos\alpha_w} = r_1\frac{\cos\alpha}{\cos\alpha_w} \tag{11.37}$$

$$r_{w2} = \frac{r_{b2}}{\cos\alpha_w} = r_2\frac{\cos\alpha}{\cos\alpha_w} \tag{11.38}$$

$$a_w = r_{w1} + r_{w2} \tag{11.39}$$

$$a_w = (r_1 + r_2)\frac{\cos\alpha}{\cos\alpha_w} \tag{11.40}$$

$$\cos\alpha_w = \frac{r_1 + r_2}{a_w}\cos\alpha \tag{11.41}$$

The angles inv α_w and inv α is calculated by the following formula

$$\text{inv}\ \alpha_y = \tan\alpha_y - \alpha_y \tag{11.42}$$

In the case where the center distance equals $r_1 + r_2$ it is found that α_w equals α, thus resulting in $x_1 + x_2 = 0$.This center distance, known as a (no indices), can be found from the following

$$a = r_1 + r_2 \quad or \quad a = \frac{1}{2}m(z_1 + z_2) \tag{11.43}$$

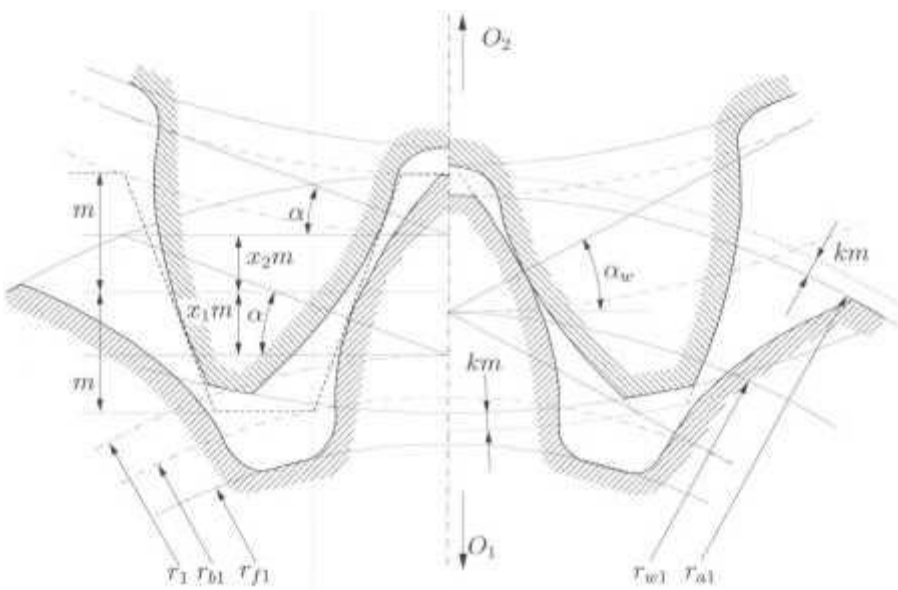
We then find that

$$a_w = a\frac{\cos\alpha}{\cos\alpha_w} \tag{11.44}$$

$$\cos\alpha_w = \frac{a}{a_w}\cos\alpha \tag{11.45}$$

11.13 Radial clearance

The radial clearance for a gear wheel in mesh is the distance between the tip circle of that gear and the tooth root circle of the meshing gear. In other words, the amount by which the dedendum of the gear exceeds the addendum of the meshing gear.



[billedtekst.start]**Figure 11.14:** Two gears with profile shift. Left: engaged with a rack Right: in clearance free engagement.[billedtekst.slut]

For a simple gear with no addendum modification $x_1 + x_2 = 0$, and manufactured with a basic rack in accordance with ISO/R 2334, the clearance is (see Figure 11.8)

$$c = 0.25m \tag{11.46}$$

With this clearance the addendum and dedendum height of the tooth is

$$h_{ad} = m_n \tag{11.47}$$

$$h_{dd} = m_n + 0.25_{mn} \tag{11.48}$$

where m_n is the normal module for helical gears (see later in this chapter), for spur gears $m_n = m$.

For a gear with addendum modification, the clearance might be unacceptable. As the center distance is less than $(x_1 + x_2)m+a$, the height of the teeth must be altered, if the clearance has to be fully retained ($c = 0.25m$). The tip relief factor, k, is calculated from (11.39) and (11.43) and Figure 11.14.

$$km = a_w - (a + x_1m + x_2m) \tag{11.49}$$

$$km = (a_w - a) - (x_1 + x_2)m \tag{11.50}$$

where k is, numerically, a small factor ≤ 0. For $k > -0.1$ the tip relief is not carried out, and for $k < -0.1$ an effective tip relief factor $k_w = k + 0.1$ is used.

$$k_w = \begin{cases} 0 & for \quad -0.1 < k \\ k + 0.1 & for \quad k < -0.1 \end{cases} \tag{11.51}$$

11.14 Gear radii

Based on the pitch radii and Figure 11.14, we find:

Base radius

$$r_{b1}=r_1 \cos \alpha \tag{11.52}$$

Root radius

$$r_{f1} = r_1 + x_1m - m - 0.25m \tag{11.53}$$

$$r_{f1} = r_1 - (1.25 - x_1)m \tag{11.54}$$

Tip radius

$$r_{a1} = r_1 + x_1m + m + k_wm \tag{11.55}$$

$$r_{a1} = r_1 + (1 + x_1 + k_w)m \tag{11.56}$$

The tip radius can, if the tip relief is taken fully into account, be calculated from

$$r_{a1} = a_w - (r_{f2} + 0.25m) \tag{11.57}$$

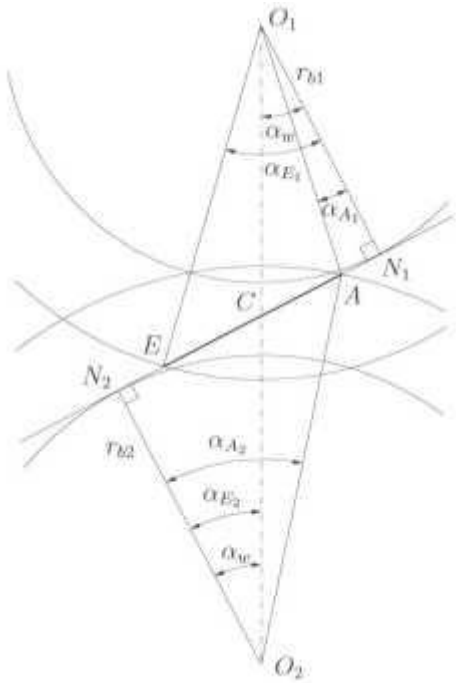
Pitch radii

$$r_{w1} = \frac{1}{1+u} a_w \tag{11.58}$$

$$r_{w2} = \frac{u}{1+u} a_w \tag{11.59}$$

11.15 Contact ratio

The transverse contact ratio, $\varepsilon\alpha$, is defined as the ratio between the length of path of contact, g_α, and the contact pitch, p_e.



[billedtekst.start]**Figure 11.15:** Length of path of contact.[billedtekst.slut]

Based on Figure 11.15, the length of the path of contact $g_\alpha = AE$ is calculated

$$g_\alpha = r_{b1} (\tan \alpha_{E1} - \tan \alpha_{A1}) \tag{11.60}$$

$$\tan \alpha_{E_1} = \frac{EN_1}{r_{b1}} \tag{11.61}$$

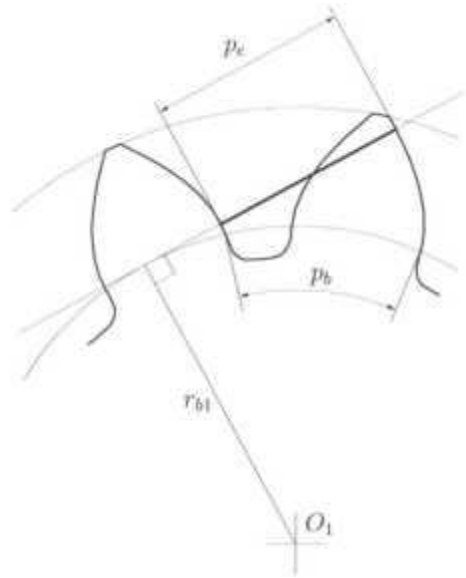
$$\tan \alpha_{E_1} = \sqrt{\left(\frac{d_{a1}}{d_{b1}}\right)^2 - 1} \tag{11.62}$$

$$\tan \alpha_{A1} = \frac{N_1N_2 - N_2A}{r_{b1}} \tag{11.63}$$

$$\tan \alpha A_1 = (1 + u) \tan \alpha_w - u \tan \alpha_{A2} \tag{11.64}$$

$$\tan \alpha_{A_2} = \sqrt{\left(\frac{d_{a2}}{d_{b2}}\right)^2 - 1} \tag{11.65}$$

The contact pitch, p_e, equals the base pitch, as the line of action is wounded off the base circle. The circumference of the base circle is



[billedtekst.start]**Figure 11.16:** Contact pitch.[billedtekst.slut]

$$2\pi r_b = z p_b \tag{11.66}$$

We can then calculate the base pitch p_b

$$p_b = \frac{2\pi r_b}{z} = \frac{\pi z m \cos\alpha}{z} = \pi m \cos\alpha \tag{11.67}$$

And as the contact pitch, $p_e = p_b$, we have

$$p_e = \pi m \cos\alpha \tag{11.68}$$

We defined the transverse contact ratio, ε_α, as the relation between the length of path of contact and the contact pitch.

$$\varepsilon_\alpha = \frac{g_\alpha}{p_e} \tag{11.69}$$

$$\varepsilon_\alpha = \frac{r_{b1}\left(\tan\alpha_{E_1} - \tan\alpha_{A_1}\right)}{\pi m \cos\alpha} \tag{11.70}$$

and as we have

$$\pi m \cos\alpha = \frac{2\pi r_{b1}}{z_1} \tag{11.71}$$

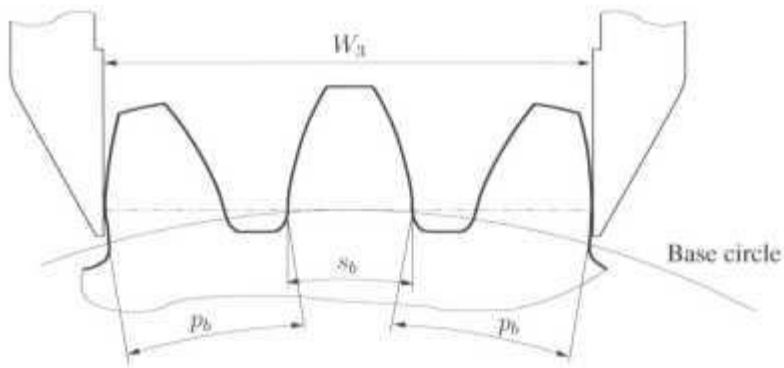
we can then calculate the transverse contact ratio as

$$\varepsilon_\alpha = \frac{z_1}{2\pi}\left(\tan\alpha_{E_1} - \tan\alpha_{A_1}\right) \tag{11.72}$$

The transverse contact ratio, ε_α, is used during the design phase, to evaluate a gears possibility for low noise level. A gear with a big contact ratio will normally give a lower noise level, than a gear with a small contact ratio, given the same quality of manufacturing and the same working conditions.

11.16 Base tangent length

A control of all teeth on a gear is time consuming and is seldom done. Normally, the control is a measurement of the base tangent length. The base tangent length is the distance between two parallel planes, which are tangents to teeth, on respectively the left flank and the right flank.



[billedtekst.start]**Figure 11.17:** Base tangent length.[billedtekst.slut]

The number of teeth to measure, must be selected so that the points of contact are as close to the base circle as possible. As the line of measure is normal to the tooth flanks at the points of contact, it is at the same time tangent to the base circle. Thus, it is possible to calculate the base tangent length as the corresponding curve length on the base line

$$W_{zw} = (z_w - 1)p_b + s_b \qquad (11.73)$$

$$W_{z_w} = (z_w - 1)\pi m \cos\alpha + 2r_b\left(\frac{1}{z}\left(\frac{\pi}{2} + 2x\tan\alpha\right) + \operatorname{inv}\alpha\right) \qquad (11.74)$$

$$W_{Zw} = m \cos\alpha\,(\pi\,(z_w - 0.5) + z \operatorname{inv}\alpha + 2x \tan\alpha) \qquad (11.75)$$

The tolerance for base tangent length can be found in DIN 3967 or DIN 3963. The number of teeth to measure is calculated as

$$z_w = z\frac{\alpha}{\pi} + 0.5 \qquad (11.76)$$

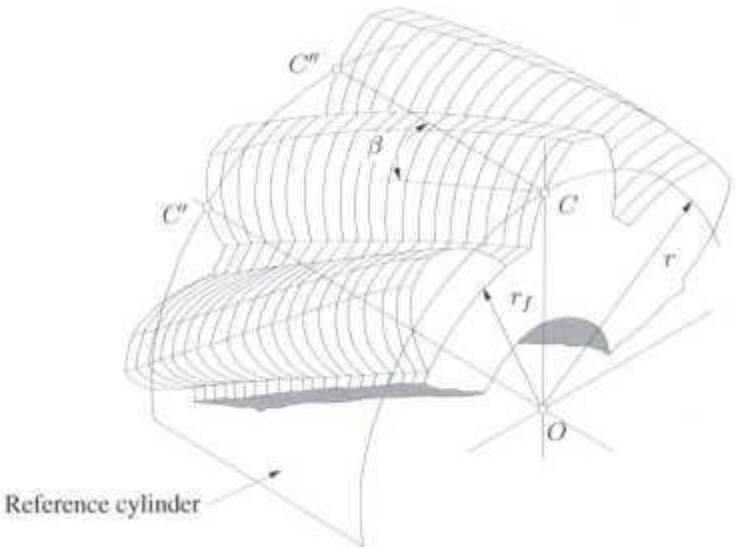
rounded to an integer.

11.17 Helical gears

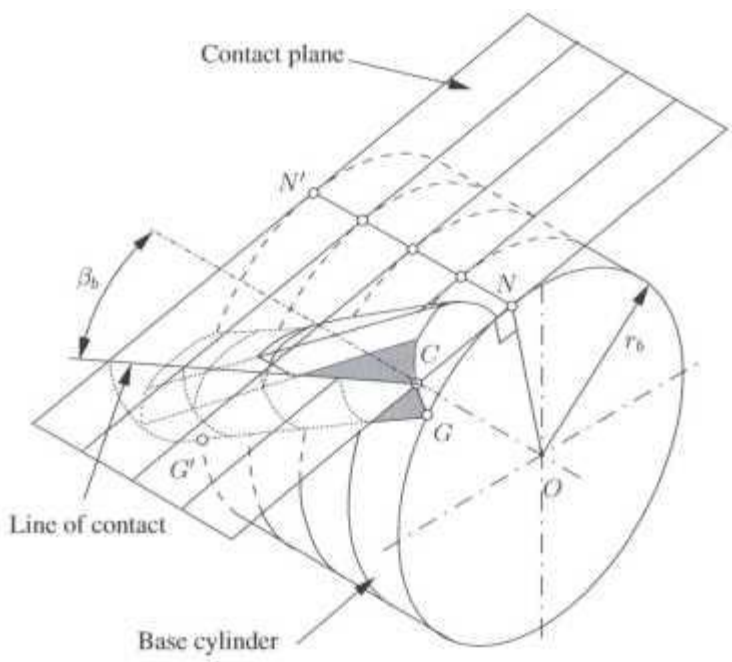
A helical gear can be thought of as created of many thin spur pieces, displaced slightly. As the displacement is the same for all pieces, the line of contact in the plane of contact will be a straight line, rotated the angle β_b from the line of contact for a spur gear.

For practical reasons are dimensions defined in the normal section given the index *n*, while dimensions defined in the cross section are given the index *t*. Even though the module m_n, for a helical gear is the same as the module *m*, for a spur gear, the index is used to avoid any mistakes. The helix angle for a gear, is defined as the angle the cutting rack is rotated, compared

to manufacturing of spur gear. This angle is called β (no indices), as it is in a tangent plane to the reference circle.



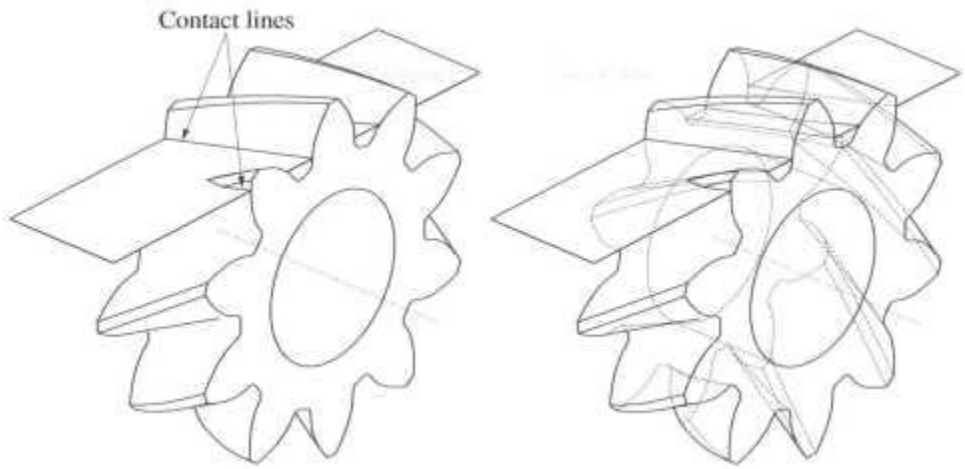
[billedtekst.start]**Figure 11.18:** Helical gear.[billedtekst.slut]



[billedtekst.start]**Figure 11.19:** Plane of contact for a helical gear.[billedtekst.slut]

The reference diameter is

$$d = \frac{p_t}{\pi} z = m_t z \tag{11.77}$$

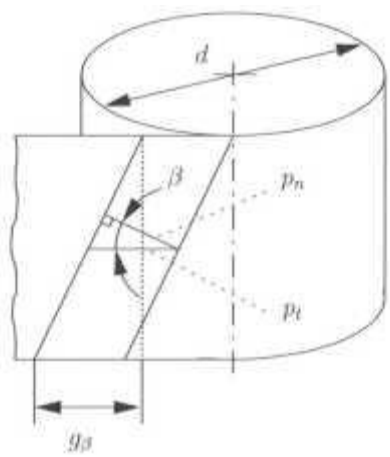


[billedtekst.start]**Figure 11.20** Plane of contact for a helical gear. The gear is shown with and without hidden lines.[billedtekst.slut]

$$p_t = \frac{p_n}{\cos\beta} \tag{11.78}$$

$$d = \frac{p_n}{\pi} z \frac{1}{\cos\beta} \tag{11.79}$$

$$d = m_n z \frac{1}{\cos\beta} \tag{11.80}$$

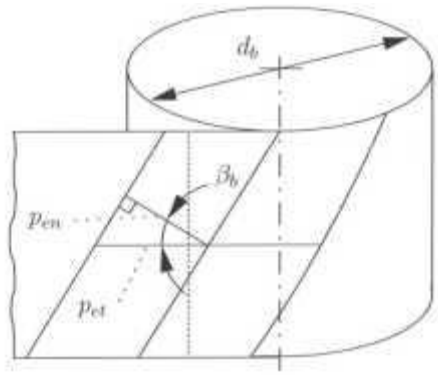


[billedtekst.start]**Figure 11.21:** Cutter reference plane - tangent plane to the reference circle.[billedtekst.slut]

The base circle diameter is

$$d_b = d \cos \alpha_t \tag{11.81}$$

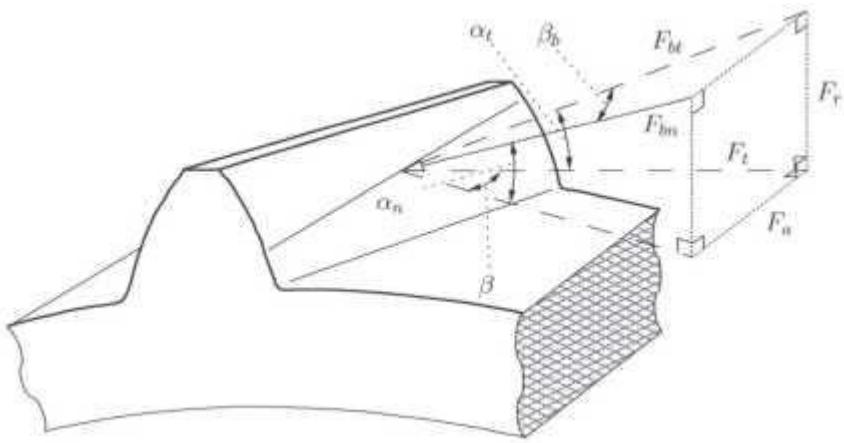
Contact pitch in the cross section is



[billedtekst.start]**Figure 11.22:** Contact plane - tangent plane to the base circle.[billedtekst.slut]

$$P_{et} = P_{bt} = P_t \cos \alpha_t \qquad (11.82)$$

In order to establish relations between the interesting angles in, respectively the normal section and the cross section, a point on the reference circle is considered in Figures 11.23 and 11.24. The only difference between the figures is that the gear in Figure 11.23 is a right hand helical gear and the gear in Figure 11.24 is a left hand helical gear.

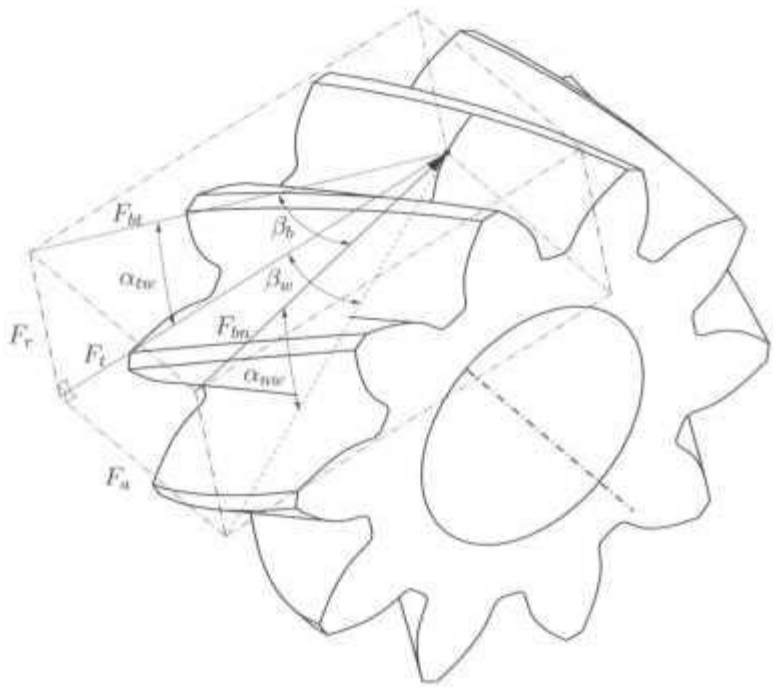


[billedtekst.start]**Figure 11.23**: Angular relations for a right hand helical gear, without profile shift.[billedtekst.slut]

From Figures 11.23 and 11.24 we see that β_b is not depending on a profile shift. With no profile shift we have the following geometric and force relations

$$\tan \alpha_t = \frac{\tan \alpha_n}{\cos \beta} \qquad (11.83)$$

$$\sin \beta_b = \sin \beta \cos \alpha_n \qquad (11.84)$$



[billedtekst.start]**Figure 11.24:** Angular relations for a left hand helical gear, with profile shift.[billedtekst.slut]

$$F_\alpha = F_t \tan \beta \tag{11.85}$$

$$F_r = F_t \tan \alpha_t \tag{11.86}$$

with profile shift we have

$$\tan \alpha_{tw} = \frac{\tan \alpha_{nw}}{\cos \beta_w} \tag{11.87}$$

$$\sin \beta_b = \sin \beta_w \cos a_{nw} \tag{11.88}$$

$$F\alpha = F_t \tan \beta_w \tag{11.89}$$

$$F_r = F_t \tan \alpha_{tw} \tag{11.90}$$

The contact ratio for helical gears, is calculated based on two contribution contact ratios, ε_α and ε_β, and the sum of these two contributions, $\varepsilon_\gamma = \varepsilon_\alpha + \varepsilon_\beta$. The transverse contact ratio, ε_α, is the same as for spur gears.

$$\varepsilon_\alpha = \frac{g_\alpha}{p_{et}} \tag{11.91}$$

$$\varepsilon_\alpha = \frac{r_{b1}\left(\tan\alpha_{E_1} - \tan\alpha_{A_1}\right)}{\frac{\pi m_n}{\cos\beta}\cos\alpha_t} \tag{11.92}$$

and as we have

$$\pi m_n \frac{\cos\alpha_t}{\cos\beta} = \frac{2\pi r_{b1}}{z_1} \tag{11.93}$$

we find

$$\varepsilon_\alpha = \frac{z_1}{2\pi}\left(\tan\alpha_{E_1} - \tan\alpha_{A_1}\right) \tag{11.94}$$

$$\tan\alpha_{E_1} = \sqrt{\left(\frac{d_{a1}}{d_{b1}}\right)^2 - 1} \tag{11.95}$$

$$\tan\alpha_{A1} = (1 + u)\tan\alpha_{tw} - u\tan\alpha_{A2} \tag{11.96}$$

$$\tan\alpha_{A_2} = \sqrt{\left(\frac{d_{a2}}{d_{b2}}\right)^2 - 1} \tag{11.97}$$

ε_β is called the contact ratio addendum and is a measure for how long one side of a tooth is in front of the other side, g_β, in relation to the pitch of the reference circle, p_t.

$$\varepsilon_\beta = \frac{g_\beta}{p_t} = \frac{b\tan\beta}{p_t} \tag{11.98}$$

If the helix angle, β, or the tooth width, b, is adjusted so that ε_β is an integer (1 - 3), the length of the line of action is at all times constant. With respect to the noise from the gear this should be an advantage. Other geometric dimensions are

$$d_{w1} = \frac{2}{1+u}a_w \tag{11.99}$$

$$\cos\alpha_{tw} = \frac{d_{b1}}{d_{w1}} \tag{11.100}$$

$$x_1 + x_2 = \frac{1}{2}(z_1 + z_2)\frac{\operatorname{inv}\alpha_{tw} - \operatorname{inv}\alpha_t}{\tan\alpha_n} \tag{11.101}$$

$$d\alpha_1 = d_1 + 2(1 + x_1 + k_w)m_n \tag{11.102}$$

$$d_{f1} = d_1 - 2(1.25 - x_1)m_n \tag{11.103}$$

$$a_w = \frac{1}{2}(z_1 + z_2)\frac{\cos\alpha_t}{\cos\alpha_{tw}}\frac{m_n}{\cos\beta} \tag{11.104}$$

$$\cos\alpha_{tw} = \frac{1}{2}(z_1 + z_2)\frac{\cos\alpha_t}{a_w}\frac{m_n}{\cos\beta} \tag{11.105}$$

$$\cos\alpha_{tw} = \frac{a}{a_w}\cos\alpha_t \tag{11.106}$$

11.18 Nomenclature

a	mm	Center distance, but (without suffix) only if $x_1 + x_2 = 0$
a_w	mm	Center distance
b	mm	Width of wheel (axial length)
c	mm	Clearance
d_1, d_2	mm	Reference diameter. 1: pinion, 2: wheel
d_{a1} , d_{a2}	mm	Tip diameter
d_{b1}, d_{b2}	mm	Base diameter
d_{f1},d_{f2}	mm	Root diameter
d_{w1}	mm	Diameter of working circle
f	μm	Individual error
g_α	mm	Length of path of contact
h_{ad}	mm	Addendum of basic rack
h_{dd}	mm	Dedendum of basic rack
i	–	Angular speed ratio/transmission ratio. $i = w_{driven}/w_{driving}$
k	–	Addendum modification factor
l	mm	Undercut length
m	mm	Module. (Standardized)
m_n	mm	Normal module. (Standardized modules. Suffix n used for helical gears)
m_t	mm	Transverse module
P	mm	Pitch (or circular pitch)
P_d	mm	Diametral pitch

r_1, r_2	mm	Reference radius. 1: pinion, 2: wheel
r_{a1}, r_{a2}	mm	Tip radius
r_{b1}, r_{b2}	mm	Base radius
r_{f1}, r_{f2}	mm	Root radius
r_{w1} r_{w2}	mm	Pitch radius
S	mm	Tooth thickness at reference diameter
s_w	mm	Tooth thickness at (working) pitch diameter
U	–	Gear ratio z_2/ z_1
X	–	Addendum modification coefficient
Z	–	Number of teeth
C	–	Pitch point
W	mm	Base tangent length
A	rad	Pressure angle for tool (basic rack)
α_w	rad	Pressure angle for two meshing gear wheels
β	rad	Helix angle (without suffix: at reference cylinder)
β_b	rad	Helix angle at base circle
Γ	rad	Auxiliary angle
E	–	Contact ratio for spur gear
ε_α	–	Transverse contact ratio
ε_β	–	Contact ratio addendum (for helical gear)
ε_γ	–	Contact ratio, total (for helical gear)
$\rho_a P_0$	mm	Tip Radius of tool
ω	rad/s	Angular speed

()		Reference circle (without suffix)

$()_a$	Axial direction
$()_b$	Base circle, base cylinder
$()_f$	Tooth root, dedendum
$()_n$	Normal section
$()_r$	Radial direction
$()_t$	Transverse section
$()_\alpha$	Transverse contact, profile
$()_\beta$	Helix
$()_\gamma$	Total, total value
$()_1$	Pinion
$()_2$	Wheel

11.19 References

[1] K. Decker. Maschinenelemente, Funktion, Gestaltung und Berechnung, 18. aktualisierta Auflage. Carl Hanser Verlag, München, Wien, 2011.

[2] DIN 3990. Basic principles for the calculations of load capacity of spur and helical gears.(in german).

[3] ISO 6336. Basic principles for the calculations of load capacity of spur and helical gears.

[4] J. E. Shigley and C. R. Michke. *Mechanical Engineering Design 7th ed.* McGraw Hill, Singapore, 2004.

[5] Tochtermann, W. and Bodenstein, W. *Konstruktionselemente des Maschinenbaues, Teil 2.* Springer- Verlag, 1979.

Chapter 12 The strength of involute gears

12.1 Introduction

The strength analysis in this chapter, is mainly based on the standard ISO 6336 [4], which provides a uniform mean of comparing and relating gear performances over a wide range of designs and applications. However, where appropriate simpler methods of evaluation is described in DIN 3990 [3] these methods are introduced.

This chapter covers calculation of load capacity, as limited by contact stress (pitting) and tooth root stress causing tooth breakage, for spur and helical gears. The nomenclature used is described in Chapter 11.

12.2 General influence factors

12.2.1 Nominal tangential load, *FNt*

The nominal tangential load, tangential to the reference cylinder, is calculated directly from the power, *P*, transmitted by the gear pair.

$$F_{Nt} = \frac{T_1}{r_1} = \frac{T_2}{r_2} \tag{12.1}$$

$$T_1 = \frac{P}{\omega_1}, \quad T_2 = \frac{P}{\omega_2} \tag{12.2}$$

$$\omega_1 = 2\pi \frac{n_1}{60\mathrm{s/min}}, \quad \omega_2 = 2\pi \frac{n_2}{60\mathrm{s/min}} \tag{12.3}$$

12.2.2 Application factor, *KA*

The application factor K_A accounts for dynamic overload from sources external to the gear. The factor depends on the physical characteristics of the driving and driven machine, on the couplings and on the operating conditions.

The application factor should be determined by precise measurement or by comprehensive system analysis. However, if this is not possible, a rough guidance is given intheTables 12.1, 12.2 and 12.3.

The values in the tables are only valid for gears not running in the resonance speed range. Experience show that K_A may be a little greater for a speed increasing transmission, than for a speed reducing transmission. Consequently, the data in Table 12.1 should be increased with a factor 1.1 for a speed increasing transmission. Tables 12.2 and 12.3 show examples of the characters on various machines.

Table 12.1: Application factor for speed reducing gears [7].

Driving-Driven	**Uniform**	**Moderate shocks**	**Heavy shocks**
Uniform	1	1.25	1.75
Light shock	1.25	1.50	2
Medium shock	1.50	1.75	2.00+

Table 12.2: Examples for driving machines [7].

Character	**Driving machine**
Uniform	electric motor, steam turbine, gas turbine
Light shock	multi cylinder combustion engine
Medium shock	single cylinder combustion engine

Table 12.3: Examples for driven machines [7].

Character	**Driven machine**
Uniform	generator, belt conveyer, platform conveyer, worm conveyer, light elevator, electric hoist, feed gears of machine tools, ventilator, turbo blower, turbo compressor
Light shock	main drive to machine tool, heavy elevator, turning gears of crane, mine ventilator, multi cylinder piston pump, feed pump
Heavy shock	press, shear, rubber dough mill, rolling mill drive, power shovel, heavy centrifuge, heavy feed pump, rotary drilling apparatus, pug mill

12.2.3 Dynamic factor, *KV*

The dynamic factor, K_V, accounts for internally generated loads due to vibrations of pinion and gear against each other. K_V is defined as the ratio between the maximum force which occurs at the mesh of an actual gear pair, and the corresponding load due to the externally applied load. The main influences are:

- Transmission errors.

- Masses of pinion and gear.
- Mesh stiffness.
- Transmission load, including application factor.

Further influences are

- Lubrication.
- Damping characteristics of the gear system.
- Shaft and bearing stiffness.
- Bearing pattern on loaded tooth flank.

$$K_V \approx 1 + f_F K z_1 v \sqrt{\frac{u^2}{1+u^2}} \cdot 10^{-5} \qquad (12.4)$$

where

f_F The load adjustment factor, see Table 12.4. K Tooth factor, [s/m], see Table 12.4

z_1 Number of teeth on pinion.

V Tangential speed in gear contact, [m/s].

u Teeth ratio, z_2 / z_1.

For helical gears with $\varepsilon_\beta < 1$ the values in Table 12.4 should be interpolated between the values for $\varepsilon_\beta = 0$ and $\varepsilon_\beta \geq 1$. The load per unit width(divided by the face width *b)* of the tooth is defined as

$$w_t = \frac{F_{Nt}}{b} K_A K_V \qquad (12.5)$$

Table 12.4: Load adjustment factor f_F and tooth factor K [3], [4] and [2].

	Quality class		
	5	6	7
w_t [N / mm]	Spur gear f_F		
200	1.34	1.43	1.52
350	1.00	1.00	1.00
500	0.86	0.83	0.79
K[s/m]	36	47	62
w_t [N / mm]	Helical gear f_F, $\varepsilon_\beta \geq 1$		
200	1.47	1.55	1.61
350	1.00	1.00	1.00

500	0.81	0.78	0.76
K[s/m]	23	32	46

12.3 Longitudinal (axial) load distribution factors, *KHβ, KFβ*

The load distribution factors account for the effects of non-uniform load distribution across the face width. These factors are also called width factors.

$K_{H\beta}$ accounts for the effect on the Hertzian pressure on the tooth flank

K_F accounts for the bending stress at the tooth root.

The main influence factors are:

- Cutting errors
- Errors in mounting due to bore errors
- Internal bearing clearance
- Wheel and pinion shaft alignment errors
- Tooth stiffness
- Shaft stiffness
- Housing stiffness
- Bearing deflections
- Thermal expansion and distortion
- Tangential and axial load
- Running-in effects

An approximate expression for width factor for tooth foot strength, $K_{F\beta}$, is

$$K_{F\beta} \approx 1 + (K_\beta - 1) f_w f_P \qquad (12.6)$$

Where

K_β Basic width factor, see Table 12.5.

f_w Line load adjustment factor, see Table 12.6.

f_p Materials factor, see Table 12.7.

an approximate expression for width factor for tooth surface strength, $K_{H\beta}$, is

$$K_{H\beta} \approx K_{F\beta}^{1.39} \qquad (12.7)$$

12.3.1 Principles of longitudinal load distributions

Figures 12.1 and 12.2, illustrate the effect of misalignment and tooth loading on load distribution before running-in.

Table 12.5: Basic width factor K_β [3], [4] and [2].

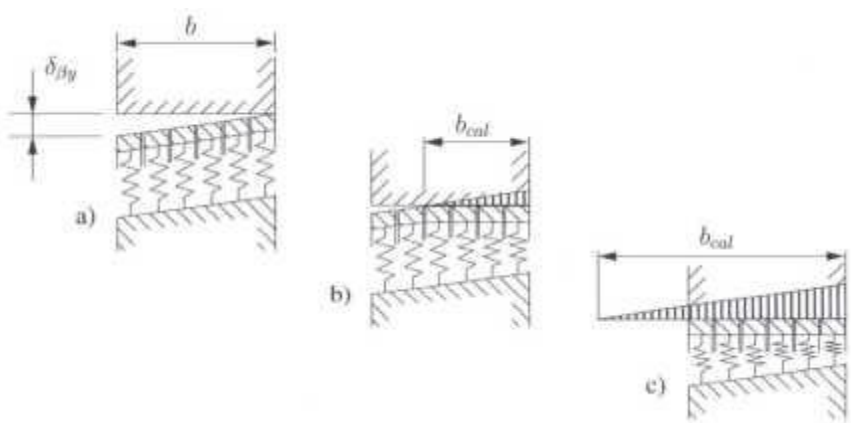
Tooth width [mm]		**Quality class**		
Over	Up to	5	6	7
	20	1.07	1.08	1.10
20	40	1.08	1.09	1.11
40	100	1.09	1.09	1.13
100	160	1.12	1.13	1.16
160	315	1.14	1.15	1.18

Table 12.6: Adjustment factor for line load f_w [3], [4] and [2].

Adjustment factor for line load f_w	1	1.15	1.3	1.45	1.6
Load per unit length w_t [N/mm]	> 350	≈ 300	≈ 250	≈ 200	≤ 100

Table 12.7: Materials factor f_p [2]

Materials factor f_p	= 1	≈0.7	≈0.5
Materials	Steel/Steel	Cast steel/Cast steel	Cast iron/Cast iron



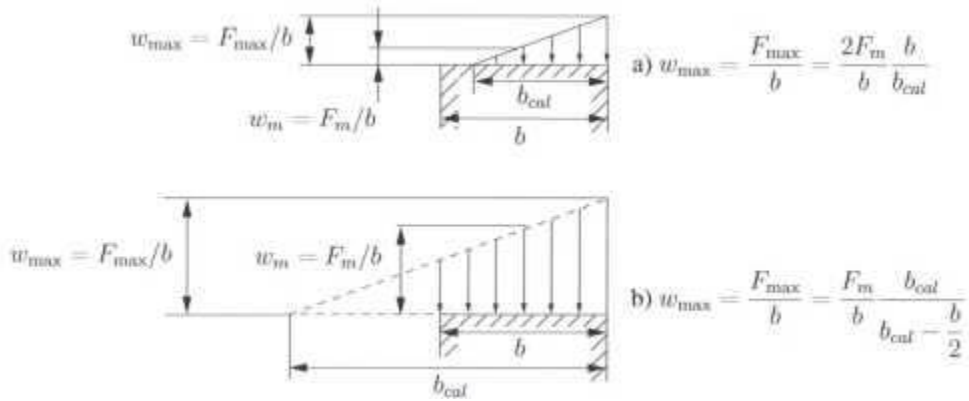
[billedtekst.start]**Figure 12.1:** Longitudinal load distribution. $\delta_{\beta y}$[μm] is effective equivalent misalignment (after running-in). b_{cal} [mm] is calculated facewidth. Principle: a) unloaded; b) light load and/or high helix error; c) heavy load and/or small helix error.[billedtekst.slut]

12.4 Transverse load distribution factors, $K_{H\alpha}$, $K_{F\alpha}$

The distribution of total tangential load over several pairs of meshing teeth depends, in the case of given gear dimensions, on gear accuracy and the value of the total tangential load. The factor $K_{H\alpha}$,, takes into account the effect of the load distribution on gear-tooth contact stresses and the factor $K_{F\alpha}$ takes the effect of the load distribution on tooth root strength into account.

The main influences are:

- Total mesh stiffness.



[billedtekst.start]**Figure 12.2:** Longitudinal load distribution. Maximum specified load F_{max}/b [N/mm]; a) light load and/or high helix error; b) heavy load and/or small helix error.[billedtekst.slut]

- Total tangential tooth load.
- Base pitch error.
- Tip relief.
- Face width.
- Running-in allowance.
- Gear tooth dimensions.

12.4.1 Formulas for determination of factors

Gears with total contact ratio $\varepsilon_\gamma \leq 2$

$$K_{Ha} = K_{Fa} = \frac{\varepsilon_\gamma}{2}\left(0.9 + 0.4\frac{c_\gamma(f_{pe} - y_p)}{w_t K_{F\beta}}\right) \tag{12.8}$$

Where

ε_γ Is the total contact ratio.

c_γ The mean total tooth stiffness in transverse plane, depends mainly on the material. For steel the value is ≈ 20N/(mm · μm). For cast iron the value is ≈ 14N/(mm · μm).

f_{pe} The maximum allowable base pitch error on pinion or wheel, see Tabl 12.8.

y_p Running in allowance. The base pitch error, f_{pe}, is typically reduced by y_p during the running in process.

w_t The tangential load per unit length, see (12.5).

$K_{F\beta}$ Width factor for tooth foot strength, see (12.6).

Gears with total contact ratio $\varepsilon_\gamma > 2$

$$K_{H\alpha} = K_{F\alpha} = 0.9 + 0.4\sqrt{\frac{2(\varepsilon_\gamma - 1)}{\varepsilon_\gamma}}\frac{c_\gamma(f_{pe} - y_p)}{w_t K_{F\beta}} \tag{12.9}$$

All parameters are previously described.

If the outcome of using (12.8) or (12.9) is that $K_{F\alpha} = K_{H\alpha} < 1$ then the values should be adjusted to $K_{F\alpha} = K_{H\alpha} = 1$. This happens when more than one pair of teeth transmits the load.

If (12.8) or (12.9) gives a value that is larger than obtained by (12.10), (12.11), (12.12) and (12.13) then only one pair of teeth transmits the load and the values for $K_{F\alpha}$ and $K_{H\alpha}$ from (12.10) and (12.11) should be used for further calculations.

$$K_{F\alpha} = \frac{1}{Y_\varepsilon} \tag{12.10}$$

$$K_{H\alpha} = \frac{1}{Z_\varepsilon^2} \tag{12.11}$$

$$Y_\varepsilon = 0.25 + \frac{0.75}{\varepsilon_\alpha} \tag{12.12}$$

$$Z_\varepsilon = \sqrt{\frac{4 - \varepsilon_\alpha}{3}(1 - \varepsilon_\beta) + \frac{\varepsilon_\beta}{\varepsilon_\alpha}} \tag{12.13}$$

If $\varepsilon_\beta > 1$ then use $\varepsilon_\beta = 1$ in (12.13).

Table 12.8: Accepted base pitch errors for industrial gears, when calculating gear strength [4], [3] and [2].

Quality class	**5**	**6**	**7**
Max allow. base pitch error	9	15	25
Running in allowance y_p[μm]	1.5	2.5	4

12.5 Calculation of surface durability (pitting)

In Chapter 11 it is analyzed how the contact point (contact line) between two gear teeth moves along the flanks as the gears rotate. In the contact a Hertzian pressure distribution is present due to the torque transmitted. If the Hertzian pressure (compressive stress) reaches a certain level a fatigue failure called pitting, may start to develop on the teeth flanks. Pitting causes small debris to leave the surface of the teeth flanks. When pitting failure evolves in a transmission, it becomes noisy and the vibration level increases. The maximum Hertzian pressure is the most significant influencing factor on the development of pitting failure, but also the materials, the hardening, the sliding speed in the contact, the teeth flank surface finish and the lubricant are important, when the pitting failure risk is evaluated. Allowance is made for these factors in the expressions developed in this section.

12.5.1 Fundamental formulas

The calculation of the surface load capacity is based on the Hertzian pressure on the operating pitch circle. The Hertzian stress, σ_H, at the pitch circle must be equal to or less than the allowable Hertzian stress, σ_{HP}

$$\sigma_H = \sigma_{H0}\sqrt{K_A K_V K_{H\beta} K_{H\alpha}} \leq \sigma_{HP} \tag{12.14}$$

where σ_{H0} is the basic value of contact stress

$$\sigma_{H0} = Z_H Z_E Z_\varepsilon Z_\beta \sqrt{\frac{F_{Nt}}{d_1 b}\frac{u+1}{u}} \tag{12.15}$$

Z_H Zone factor, see Subsection 12.5.4. Takes into account the flank curvatures of the pitch point and the relation of the tangential load at the pitch circle to that of the reference cylinder.

Z_E Elasticity factor, see Subsection 12.5.5. Takes into account the material properties.

Z_ε Contact ratio factor, see Subsection 12.5.6. Takes into account the influence of the effective length of the lines of contact.

Z_β Helix angle factor, see Subsection 12.5.7. Takes into account the influence of the helix angle.

d_1, b, u Other factors relate to the geometry.

12.5.2 Allowable contact stress

The allowable contact stress σ_{HP} is to be evaluated separately for pinion and wheel.

$$\sigma_{HP} = \frac{\sigma_{Hlim} Z_N}{S_{Hmin}} Z_L Z_R Z_V Z_W Z_X \tag{12.16}$$

where

σH_{lim} Endurance limit for contact stress (material related factor).

Z_N Life factor for contact stress, see Subsection 12.5.8. Permits higher load capacity for a limited number of cycles.

S_{Hmin} Minimum required safety factor for contact stress.

In general, the factors Z_L, Z_R, Z_V takes into account the influence of the oil film on surface fatigue.

Z_LLubrication factor, see Subsection 12.5.9.

Z_R Roughness factor, see Subsection 12.5.10.

Z_V Speed factor, see Subsection 12.5.11.

Z_W Work hardening factor, see Subsection 12.5.12. Accounts for the effect on load capacity of meshing with a surface hardened mating gear.

Z_X Size factor for contact stress. Normally Z_X has a value of 1.

12.5.3 Safety factor for contact stress (against pitting)

The calculated safety factor for contact stress must be checked separately for pinion and wheel.

$$S_H = \frac{\sigma_{Hlim} Z_N}{\sigma_{H0}} \frac{Z_L Z_R Z_V Z_W Z_X}{\sqrt{K_A K_V K_{H\beta} K_{H\alpha}}} \tag{12.17}$$

12.5.4 Zone factor

The zone factor accounts for the influence on the Hertzian pressure of tooth flank curvature at pitch point and converts the tangential force at the reference cylinder to the normal force at the pitch cylinder.

$$Z_H = \sqrt{\frac{2\cos\beta_b \cos\alpha_{wt}}{\cos^2\alpha_t \sin\alpha_{wt}}} = \sqrt{\frac{2\cos\beta_b}{\cos^2\alpha_t \tan\alpha_{wt}}} \tag{12.18}$$

12.5.5 Elasticity factor

The elasticity factor accounts for the influence of the material properties E (Young's modulus) and v (Poisson's ratio) on the Hertzian pressure.

$$\frac{1}{E_{12}} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \tag{12.19}$$

$$Z_E = \sqrt{\frac{1}{\pi \frac{1}{E_{12}}}} = \sqrt{\frac{E_{12}}{\pi}} \tag{12.20}$$

When Young's modulus and Poisson's ratio are the same for both pinion and wheel, (12.20) can be written as

$$Z_E = \sqrt{\frac{E}{2\pi(1-\nu^2)}} \tag{12.21}$$

For steel v = 0.3 and therefore

$$Z_E = \sqrt{0.175E} \tag{12.22}$$

For mating gears in materials having different Young's modulus but same Poisson ratio, the equivalent modulus can be found as

$$E = \frac{2E_1E_2}{E_1+E_2} \tag{12.23}$$

12.5.6 Contact ratio factor

The contact ratio factor accounts for the influence of the transverse contact ratio, and the overlap ratio on the specific surface load of gear teeth.

Spur gears

$$Z_\varepsilon = \sqrt{\frac{4-\varepsilon_\alpha}{3}} \tag{12.24}$$

Helical gears

$$Z_\varepsilon = \sqrt{\frac{4-\varepsilon_\alpha}{3}(1-\varepsilon_\beta)+\frac{\varepsilon_\beta}{\varepsilon_\alpha}} \qquad \varepsilon_\beta < 1 \qquad (12.25)$$

$$Z_\varepsilon = \sqrt{\frac{1}{\varepsilon_\alpha}} \qquad \varepsilon_\beta \geq 1 \qquad (12.26)$$

The transverse contact ratio, ε_α, and the overlap ratio, ε_β, can be found in Chapter 11.

12.5.7 Helix angle factor

Independently of the influence of helix angle on the length of line of contact, the helix angle factor, Z_β, accounts for the influence of helix angle on surface durability, allowing for such variables as the distribution of load along the lines of contact.

$$Z_\beta = \sqrt{\cos\beta} \qquad (12.27)$$

12.5.8 Life factor

The life factor takes account of a higher permissible Hertzian stress if only limited durability endurance (number of cycles) is demanded.

Based on the number of cycles, N_L, material and hardening, we can calculate Z_N as:

For through hardened steels or surface hardened steel when a certain amount of pitting is allowed

$$Z_N = 1.6 \quad N_L \leq 6 \cdot 10^5 \qquad (12.28)$$

$$Z_N = \left(\frac{3 \cdot 10^8}{N_L}\right)^{0.0756} \qquad 6 \cdot 10^5 < N_L \leq 10^7 \qquad (12.29)$$

$$Z_N = \left(\frac{10^9}{N_L}\right)^{0.057} \qquad 10^7 < N_L \leq 10^9 \qquad (12.30)$$

$$Z_N = 1 \quad 10^9 \leq N_L \qquad (12.31)$$

For through hardened steels or surface hardened steels

$$Z_N = 1.6 \qquad N_L \leq \quad 10^5 \qquad (12.32)$$

$$Z_N = \left(\frac{5 \cdot 10^7}{N_L}\right)^{0.0756} \qquad 6 \cdot 10^5 < N_L \leq 10^7 \qquad (12.33)$$

$$Z_N = 1 \quad 5\text{-}10^7 \leq N_L \qquad (12.34)$$

For through hardened or nitriding steels, gas nitridation cast iron

$$Z_N = 1.3 \quad N_L \leq 10^5 \qquad (12.35)$$

$$Z_N = \left(\frac{2 \cdot 10^6}{N_L}\right)^{0.0875} \qquad 10^5 < N_L \leq 2 \cdot 10^6 \qquad (12.36)$$

$$Z_N = 1 \qquad 2 \cdot 10^6 \leq N_L \qquad (12.37)$$

For through hardened steels, bath nitridation

$$Z_N = 1.1 \quad N_L \leq 10^5 \qquad (12.38)$$

$$Z_N = \left(\frac{2 \cdot 10^6}{N_L}\right)^{0.0318} \qquad 10^5 < N_L \leq 2 \cdot 10^6 \qquad (12.39)$$

$$Z_N = 1 \quad 2 \quad 10^6 \qquad \leq N_L \qquad (12.40)$$

12.5.9 Lubrication factor

The lubrication factor, Z_L, accounts for the influence of the type of lubricant and its viscosity on the surface load capacity.

$$Z_L = C_{ZL} + \frac{4(1 - C_{ZL})}{\left(1.2 + \frac{80\frac{mm^2}{s}}{\nu_{50}}\right)} = C_{ZL} + \frac{4(1 - C_{ZL})}{\left(1.2 + \frac{134\frac{mm^2}{s}}{\nu_{40}}\right)} \tag{12.41}$$

$$C_{ZL} = \frac{\sigma_{Hlim} - 850\frac{N}{mm^2}}{350\frac{N}{mm^2}}0.08 + 0.83 \tag{12.42}$$

If σ_{Hlim} is less than 850N/mm², use σ_{Hlim} = 850N/mm². If σ_{Hlim} is greater than 1200N/mm², use σ_{Hlim} = 1200N/mm².

12.5.10 Roughness factor

The roughness factor accounts for the influence of surface texture of tooth flanks on surface load capacity.

$$Z_R = \left(\frac{3}{\mathtt{Rz100}}\right)^{C_{ZR}} \tag{12.43}$$

where Rz100 is the mean relative roughness, relative to a center distance of 100mm.

$$\mathtt{Rz100} = \frac{\mathtt{Rz1} + \mathtt{Rz2}}{2}\sqrt[3]{\frac{100}{a}} \tag{12.44}$$

$$C_{ZR} = 0.12 + \frac{1000\frac{N}{mm^2} - \sigma_{Hlim}}{5000\frac{N}{mm^2}} \tag{12.45}$$

If σ_{Hlim} is less than 850N/mm² use σ_{Hlim} = 850N/mm²; if σ_{Hlim} has a higher value than 1200N/mm², use σ_{Hlim} = 1200N/mm². For estimates one may use Ra ≈ 0.lRz.

12.5.11 Speed factor

The speed factor accounts for the influence of the pitch line velocity on the surface load capacity.

$$Z_V = C_{ZV} + \frac{2(1 - C_{ZV})}{\sqrt{0.8 + \frac{32\frac{m}{s}}{v}}} \tag{12.46}$$

$$C_{ZV} = 0.85 + \frac{\sigma_{Hlim} - 850\frac{N}{mm^2}}{350\frac{N}{mm^2}}0.08 \tag{12.47}$$

If σ_{Hlim} is less than 850N/mm² use σ_{Hlim} = 850N/mm²; if σ_{Hlim} has a higher value than 1200N/mm², use σ_{Hlim} = 1200N/mm². v[m/s] is the tangential speed in the pitch point.

12.5.12 Work hardening factor

The work hardening factor accounts for the increase of surface durability due to meshing a steel

wheel, with a hardened pinion with smooth tooth surface.

$$Z_W = 1.2 - \frac{\mathrm{HB} - 130}{1700} \tag{12.48}$$

where HB is the Brinell hardness of the hardened pinion.

For hardness inside the range 130 < HB < 400, the value assigned to Z_W is 1.0.

12.6 Calculation of load capacity (tooth breakage)

The methods are based on the assumption that the highest tooth root tensile stress arises by application of the force at the outer point of single tooth pair contact (in the case of helical gears for the virtual teeth in the normal section).

12.6.1 Fundamental formulas

The tooth root stress, σ_F, must be equal to or less than the allowable tooth root pressure, σ_{FP}

$$\sigma_F = \sigma_{F0} K_A K_V K_{F\beta} K_{F\alpha} \leq \sigma_{FP} \qquad (12.49)$$

where σ_{F0} is the basic value of tooth root stress.

$$\sigma_{F0} = \frac{F_{Nt}}{bm_n} Y_{Fa} Y_{Sa} Y_\varepsilon Y_\beta = \frac{F_{Nt}}{bm_n} Y_{FS} Y_\varepsilon Y_\beta \qquad (12.50)$$

where

Y_{Fa} Tooth form factor, see Subsection 12.6.4. Takes the influence of the tooth form on the nominal bending stress into account.

Y_{Sa} Stress modification factor. Takes into account theconversion of the nominal bending stress to the local tooth root stress.

Y_{FS} Shape factor for tooth, see Table 12.9.

Y_ε Contact ratio factor, see (12.12).

Y_β Helix angle factor, see Subsection 12.6.5.Takes into account the influence of the helix angle.

12.6.2 Allowable tooth root stress

The allowable tooth root stress, σ_{FP}, is to be evaluated separately for pinion and wheel.

$$\sigma_{FP} = \frac{\sigma_{Flim} Y_{NT}}{S_{Fmin}} Y_\delta Y_R Y_X \qquad (12.51)$$

Where

σ_{Flim} Endurance limit for tooth root stress (material related factor).

Y_{NT} Life factor for tooth root stress, see Subsection 12.6.6. Permits higher load capacity for a limited number of cycles.

S_{Fmin} Minimum demanded safety factor for tooth root stress.

Y_δ Relative notch sensitivity factor, see Subsection 12.6.7. Indicates to what extent the theoretical stress concentration lies above the endurance limit in the case of fatigue breakage.

Y_R Relative surface condition factor, see Subsection 12.6.8. Takes into account the

dependence of the tooth root strength on the surface condition in the tooth root fillet.

Y_X Size factor for tooth root stress, see Subsection 12.6.9.

12.6.3 Safety factor for tooth root stress (against tooth breakage)

The calculated safety factor for tooth root stress must be checked separately for pinion and wheel

$$S_F = \frac{\sigma_{Flim} Y_{NT}}{\sigma_{F0}} \frac{Y_\delta Y_R Y_X}{K_A K_V K_{F\beta} K_{F\alpha}} \tag{12.52}$$

12.6.4 Tooth form factor

The tooth form factor, Y_{FS}, takes into account the influence of the tooth form on the nominal bending stress. See Tables 12.9 and 12.10.

Table 12.9: Tooth form factor Y_{FS} for positive profile shift factor [2].

Profile shift factor											
z_n	0	+0.1	+0.2	+0.3	+0.4	+0.5	+0.6	+0.7	+0.8	+0.9	+ 1.0
16	4.88	4.74	4.59	4.47	4.38	4.32	4.26	4.21	4.16	4.11	4.03
17	4.83	4.67	4.55	4.45	4.37	4.31	4.26	4.22	4.18	4.13	4.05
18	4.77	4.63	4.51	4.42	4.36	4.30	4.27	4.23	4.19	4.15	4.08
19	4.71	4.58	4.48	4.40	4.34	4.30	4.27	4.24	4.20	4.16	4.10
20	4.66	4.55	4.45	4.38	4.33	4.30	4.27	4.25	4.22	4.18	4.12
50	4.27	4.28	4.29	4.32	4.34	4.36	4.38	4.40	4.40	4.40	4.38
60	4.26	4.28	4.30	4.33	4.35	4.38	4.41	4.43	4.43	4.43	4.41
70	4.26	4.29	4.31	4.34	4.37	4.40	4.43	4.45	4.45	4.45	4.45

Table 12.10: Tooth form factor Y_{FS} for negative profile shift factor [2].

Profile shift factor					
z_n	-0.5	-0.4	-0.3	-0.2	-0.1
20					4.80
22			5.0	4.85	4.70

24		5.07	4.89	4.74	4.62
50	4.37	4.33	4.30	4.28	4.27
60	4.28	4.25	4.20	4.25	4.25
70	4.23	4.22	4.22	4.23	4.24

12.6.5 Helix angle factor

The helix angle factor, Y_β, takes into account the difference between the helical gear and the virtual spur gear in the normal section.

$$Y_\beta = 1 - \varepsilon_\beta \frac{\beta}{2\pi/3} \geq Y_{\beta min} \qquad (12.53)$$

If $\varepsilon_\beta > 1$ in (12.53) then ε_β 1 should be used. If $\beta > \pi/6$ in (12.53) then $\beta = \pi/6$ should be used.

$$Y_{\beta min} = 1 - 0.25\varepsilon_\beta \geq 0.75 \qquad (12.54)$$

12.6.6 Life factor

The life factor, Y_{NT}, takes into account that, in the case of limited life (number of cycles), a higher tooth root stress can be permitted. For a given number of load cycles, N_L, can the life factor be found, based on the material.

For through hardened and tempered steel

$$Y_{NT} = 2.5N_L \leq 10^4 \tag{12.55}$$

$$Y_{NT} = \left(\frac{3 \cdot 10^6}{N_L}\right)^{0.16} \qquad 10^4 \leq N_L \leq 3 \cdot 10^6 \tag{12.56}$$

$$Y_{NT} = 1\ 3{\cdot}10 < N_L \tag{12.57}$$

For surface hardened steels

$$Y_{NT} = 2.5\ N_L \leq 10^3 \tag{12.58}$$

$$Y_{NT} = \left(\frac{3 \cdot 10^6}{N_L}\right)^{0.115} \qquad 10^3 \leq N_L \leq 3 \cdot 10^6 \tag{12.59}$$

$$Y_{NT} = 1\ 3\ 10^6 < N_L \tag{12.60}$$

For through hardened or nitriding steels, gas nitridation

$$Y_{NT} = 1\text{-}6\ N_L \leq 10^3 \tag{12.61}$$

$$Y_{NT} = \left(\frac{3 \cdot 10^6}{N_L}\right)^{0.059} \qquad 10^3 \leq N_L \leq 3 \cdot 10^6 \tag{12.62}$$

$$Y_{NT} = 1\ 3{\cdot}10^6 < N_L \tag{12.63}$$

For through hardened steels, bath nitridation

$$Y_{NT} = 1.2\ N_L \leq 10^3 \tag{12.64}$$

$$Y_{NT} = \left(\frac{3 \cdot 10^6}{N_L}\right)^{0.012} \qquad 10^3 \leq N_L \leq 3 \cdot 10^6 \tag{12.65}$$

$$Y_{NT} = 1\ 3\ 10 < N \tag{12.66}$$

12.6.7 Relative notch sensitivity factor, Y_δ

The relative notch sensitivity factor, Y_δ, indicates the sensitivity to stress concentration in the case of fatigue breakage.

The factor is based on the actual stress concentration factor and the gear material. It is mostly found

that

$$Y_\delta = 1.0 \tag{12.67}$$

12.6.8 Relative surface condition factor

The relative surface condition factor, Y_R, takes into account the dependence of the tooth root strength on the surface condition in the root fillet, mainly the dependence on the peak to valley roughness.

The factor is based on the surface roughness and the material.

For Rz < 1μm. For steel

$$Y_R = 1.120 \tag{12.68}$$

For surface hardened steels and perlitic malleable iron

$$Y_R = 1.070 \tag{12.69}$$

For cast iron and nitridation steels

$$Y_R = 1.025 \tag{12.70}$$

For 1μm ≤ Rz ≤ 40μm. For steel

$$Y_R = 1.674 - 0.529(Rz + 1)^{0.1} \tag{12.71}$$

For surface hardened steels and perlitic malleable iron

$$Y_R = 5.306 - 4.203(Rz + I)^{0.01} \tag{12.72}$$

For cast iron and nitridation steels

$$Y_R = 4.299 - 3.259(Rz + I)^{0.005} \tag{12.73}$$

12.6.9 Size factor

The size factor, Y_X, takes into account the strength decrease with increasing size. The factor is based on the normal module and material for the gear.

For steel, perlitic malleable iron, spheroidal cast iron

$$Y_X = 1.03 - 0.006m_n \quad 5 < m_n < 30 \tag{12.74}$$

$$Y_X = 0.85 \quad 30 < m_n \tag{12.75}$$

For surface hardened steels

$$Y_X = 1.05 - 0.01mn \quad 5 < m_n < 30 \tag{12.76}$$

$$Y_X = 0.75 \quad 30 < m_n \tag{12.77}$$

For cast iron

$$Y_X = 1.075 - 0.01m_n \quad 5 < m_n < 30 \tag{12.78}$$

$$Y_X = 0.7 \quad 30 < m_n \tag{12.79}$$

For $m_n \leq 5$ use $Y_X = 1.0$.

12.7 Elastohydrodynamic lubrication in gears

Due to the very high pressure in the contact between the teeth elastohydrodynamic lubrication occurs. To determine an appropriate oil viscosity the "Stribeck reference contact pressure" is determined

$$k_s = \frac{3F_t}{bd_1}\frac{u+1}{u} \tag{12.80}$$

where

k_s [N/mm²] " Stribeck reference contact pressure".

F_t[N] Tangential force in contact.

b[mm] Width of pinion.

$d1$ [mm] Reference diameter of pinion.

u Gear ratio.

The "Stribeck reference contact pressure" is not the actual contact pressure in the contact, but rather a value used to determine the appropriate viscosity.

The lubrication coefficient k_s/v is used as parameter in Table 12.11 to find the viscosity

Table 12.11: The relation between k_s/v and the kinematic viscosity at 40°C for the lubricant recommended for the gear [2].

k_s/v [MPa · s/m]	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
V [mm²/s]	47	52	56	60	63	66	69	71	74	77
k_s/v [MPa · s/m]	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
v[mm²/s]	77	95	120	140	150	160	168	175	185	195
k_s/v [MPa · s/m]	2	3	4	5	6	7	8	9	10	20
v[mm2/s]	270	330	380	420	470	495	520	550	570	740

12.8 Design modification/optimization

The previous part of the chapter has described how the strength of gears are assessed according to the ISO 6336 standard. In gear design there are, as indicated, two primary limiting factors with respect to the strength, both are related to fatigue. The first is a fatigue fracture that leads to breakage of the whole tooth. This is due to the bending stress at the tooth root. The second limiting factor is surface fatigue, primarily pitting. This fatigue is due to the contact pressure between the teeth on contacting gears. The design modifications shown here are related to the

first limiting factor, the bending stress.

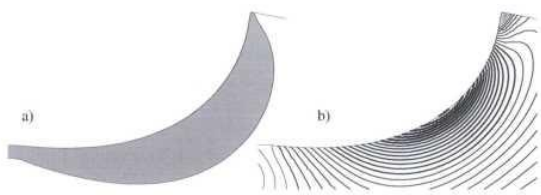
The estimation and evaluation of bending stress in gear teeth has a long history that goes back to the 19th century, see e.g. [5]. In this paper an earlier result from 1868 for evaluating the strength of gear is reported. The evaluation is given by

$$X = 2000 p_c f \tag{12.81}$$

where X is the tooth breaking load in pounds, p_c is the circular pitch in inches and f is the tooth face width in inches. The result in (12.81) has many deficiencies. First of all it does not account for the actual

tooth thickness at the root. The strength is therefore not calculated today using this formula but by the method presented in the previous parts of the chapter.

The bending stress can also be calculated using the Finite Element Method (FEM), see e.g. [6], This is in contrast to the older method of photo elasticity, see e.g. [1], In gear design the maximal stress at the tooth root is controlled both by the nominal bending stress and by the stress concentration due to the geometrical changes. Using FE and methods of shape optimization it is possible to reduce stress concentration. The design/shape of involute teeth is highly controlled by the ISO standard that specifies the cutting tool shape for the teeth. For the function of the transmission the most important parameter is the involute tooth shape. Design changes can however be made to the gear tooth root without affecting the involute shape and therefore the functional goodness of the transmission can remain the same. The maximum bending stress in gears is however not in the involute part of the tooth but in the root, the shape of which is controlled by the cutting tool tip design, see Figure 11.7.



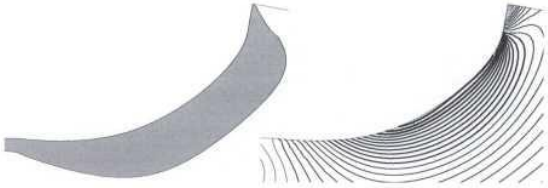
[billedtekst start]**Figure 12.3:** Close-up of the stress concentration zone for the ISO design with 17 teeth, a) Illustration of stress along the tooth root, b) Stress contour lines.[8][billedtekst slut]

In Figurel2.3 the bending stress for the ISO profile is shown for the case of a gear with 17 teeth. To illustrate the stress concentration, Figure 12.3a) shows the size of the largest principal stress along the part of the boundary where the stress concentration is present. The size of the stress is indicated by the gray area, the perpendicular thickness of the gray area corresponds to the stress level. Figure 12.3b) shows the contour lines of constant stress

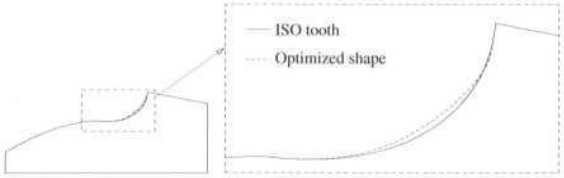
In doing shape optimization of gear teeth there are two possibilities. The most simple is to optimize the tooth root directly, but from a practical point-of-view it should instead be the tool that cuts the shape that should be designed as this shape indirectly controls the finished tooth shape. The cutting tool corner is a circular arc, this shape is the typical selected shape in many different standards related to machine elements, although it is seldom the optimal choice in relation to stress level minimization.

In [8] it is shown that it is possible to make improvements in the bending stress for involute spur teeth that can be meshed with standard ISO teeth. The improvement in the bending stress is achieved through a stress concentration lowering. In the case of a spur gear with 17 teeth the improvement in the bending stress reported as compared to the ISO profile is 11.9%, see Figures 12.4 and 12.5, only half a tooth is shown in Figure 12.5. For the case of 12 teeth the improvement is 14.3%, see Figures 12.6 and 12.7. In [8] it is demonstrated that significant improvement in the maximum bending stress is possible. The geometric changes to the cutting tool are simple and can therefore be used for practical purposes.

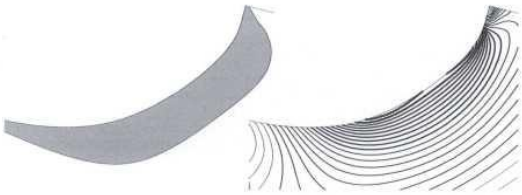
If the transmission is only running in one direction it is also possible to use asymmetric gear teeth, and the bending stress of these type of gears can also be optimized see e.g. [9].



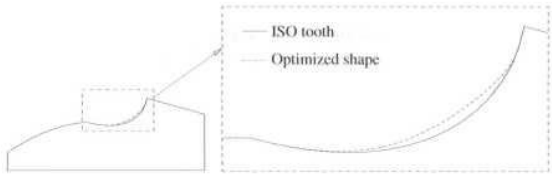
[billedtekst start]**Figure 12.4:** Close-up of the stress concentration zone for the optimal tooth for a gear with 17 teeth. The improvement in the maximum bending stress is 11.9% as compared to the ISO profile.[8][billedtekst slut]



[billedtekst start]**Figure 12.5:** Left contour of half the tooth of a gear with 17 teeth. Right: Zoom of the area where the design of the tooth is changed. The ISO profile is shown together with the optimal shape. [8][billedtekst slut]



[billedtekst start]**Figure 12.6:** Close-up of the stress concentration zone for the optimal tooth for a gear with 12teeth. The improvement in the maximum bending stress is 14.3% as compared to the ISO profile.[8][billedtekst slut]



[billedtekst start]**Figure 12.7:** Left contour of half the tooth of a gear with 12 teeth. Right: Zoom of the area where the tooth design is changed. The ISO profile is shown together with the optimal shape.[8][billedtekst slut]

12.9 Nomenclature

a	mm	Shaft center distance
b	mm	Face width
b_{cal}	mm	Width of loaded tooth face according to Figure 12.2
C_γ	N/(mm, μm)	Mean value of total tooth stiffness (or mesh stiffness) per unit face width
d_1,d_2	mm	Reference diameter of pinion, wheel
f	in	Tooth face width
f_p	–	Materials factor
f_{pe}	μm	Maximum allowable base pitch error
f_w	–	Line load adjustment factor
f_F	–	Load adjustment factor
k_s	N/mrn	Stribeck reference contact pressure
m_n	mm	Normal module
n_1,n_2	min^{-1}	Revolutions
Pc	in	Pitch (circular)
u	–	Gear ratio $zi/z\backslash$
v	m/s	Tangential speed
w_m	N/mm	Mean load per unit width according to Figure 12.2

w_{max}	N/mm	Maximum load per unit width according to Figure 12.2
W_t	N/mm	Tangential force including overload factors divided by tooth width
y_p	μm	Running in allowance
Z_1, Z_2	–	Number of teeth of pinion, wheel
C_{ZL}	–	Auxiliary lubrication factor dependent of material strength
C_{ZR}	–	Auxiliary roughness factor dependent of material strength
C_{ZV}	–	Auxiliary speed factor dependent of material strength
E_1, E_2	N/mm^2	Modulus of elasticity for pinion and wheel
E_{12}	N/mm^2	Elasticity factor
F_m	N	Mean load according to Figure 12.2
F_{max}	N	Maximum specified load according to Figure 12.2
Ft	N	Tangential force at reference circle
F_{Nt}	N	Nominal tangential load. (Tangential to the reference cylinder)
K	s/m	Tooth factor
K_V	–	Dynamic factor
K_A	–	Application factor

$K_{F\alpha}$	–	Transverse load distribution factor for bending stress
$K_{F\beta}$	–	Longitudinal load distribution factor for bending stress
$K_{H\alpha}$	–	Transverse load distribution factor for contact stress (Hertzian pressure)
$K_{H\beta}$	–	Longitudinal load distribution factor for contact stress (Hertzian pressure)
$K\beta$	–	Basic width factor
N_L	–	Number of cycles
P	Nm/s	Transmitted power
S_F	–	Safety factor for bending stress (against breakage)
S_{Fmin}	–	Minimum required safety factor for bending stress (against breakage)
S_H	–	Safety factor for contact stress (Hertzian pressure) (against pitting)
S_{Hmin}	–	Minimum required safety factor for contact stress (Hertzian pressure)
T_1, T_2	Nm	Nominal torque; pinion and wheel respectively
X	lb	Tooth breaking load
$Y_{F\alpha}$	–	Tooth form factor
Y_{FS}	–	Shape factor for tooth form factor
Y_{NT}	–	Life factor for tooth root stress
Y_R	–	Surface condition factor
$YS\alpha$	–	Stress modification factor
Y_X	–	Size factor for bending stress
Y_β	–	Helix angle factor for bending stress
$Y_{\beta min}$	–	Minimum helix angle factor for bending stress
Y_δ	–	The relative notch sensitivity factor

Ye	–	Contact ratio factor for bending stress
Z_V	–	Speed factor
Z_E	$\sqrt{N/mm^2}$	Elasticity factor
Z_H	–	Zone factor for Hertzian pressure at pitch point
Z_L	–	Lubricant factor
Z_N	–	Life factor for contact stress
Z_R	–	Roughness factor for contact stress
Z_w	–	Work hardening factor
Z_X	–	Size factor for contact stress
Z_β	–	Helix angle factor for contact stress
Z_ε	–	Contact ratio factor for contact stress
Ra	μm	Surface roughness parameter
Rz	μm	Surface roughness parameter
$\delta_{\beta y}$ εy	μm	Effective equivalent misalignment (after running-in)
ε_α	–	Transverse contact ratio
ε_β	–	Overlap ratio
ε_γ	–	Total contact ratio
υ_{50}	mm^2/s	Kinematic viscosity at 50°C
σF	N/mm^2	Tooth root stress
$\sigma Flim$	N/mm^2	Endurance limit for tooth root stress
σFp	N/mm^2	Allowable tooth root stress
ΣH	N/mm^2	Contact stress (Hertzian pressure)

$\Sigma H0$	N/mm^2	Basic value of contact stress
$\sigma Hlim$	N/mm^2	Endurance limit for contact stress
σHP	N/ mm^2	Allowable contact stress (allowable Hertzian pressure)
w_1, w_2	rad/s	Angular velocity of pinion, wheel

12.10 References

[1] T. Aida and Y. Terauchi. On bending stress of spur gear –1,3. *Japan Society of Mechanical Engineers – Bulletin,* 5(17): 161–183, 1962.

[2] *K. Decker. Maschinenelemente, Funktion, Gestaltung und Berechnung, 18. aktualisierta Auflage. Carl Hanser Verlag, Mimchen, Wien, 2011.*

[3] DIN 3990. Basic principles for the calculations of load capacity of spur and helical gears.(in german).

[4] ISO 6336. Basic principles for the calculations of load capacity of spur and helical gears.

[5] W. Lewis. Investigation of the strength of gear teeth. *Proceedings of the engineers' club of Philadelphia,* X: 16-24, October 1892.

[6] S. Li. Finite element analysis for contact strength and bending strength of a pair of spur gears with machining errors, assembly errors and tooth modifications. *Mechanism and Machine Theory,* 42:88–114, 2007.

[7] G. Niemann, Winter H., and Höhn G. *Maschinenelemente Band I.* Springer-Verlag, Berlin, Heildelberg, New York, 2005.

[8] N. L. Pedersen. Reducing bending stress in external spur gears by redesign of the standard cutting tool. *Structural and Multidisciplinary Optimization,* 38(3):215–227, 2009.

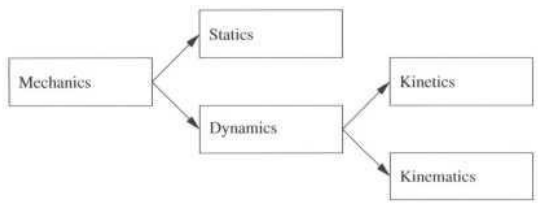
[9] N. L. Pedersen. Improving bending stress in spur gears using asymmetric gears and shape optimization. *Mechanism and Machine Theory,* 45(11):1707–1720, 2010.

Chapter 13 2D Joint Kinematics

13.1 Introduction

In Chapter 7 we have already covered couplings and universal joints between shafts. The present chapter is devoted to practical aspects in relation to 2D kinematic joints. As such the chapter can be seen as an introduction into the subject of mechanisms or multibody dynamics. For textbooks describing multibody analysis in more details see e.g. [1] or [2], or specifically in relation to 2D [3].

The subject is closely related to mechanics in general. We can in principle split the subject of mechanics into two parts; statics and dynamics. The dynamics subject is then typically again split also into two parts; kinetics and kinematics. Graphically it can be interpreted as in Figure 13.1



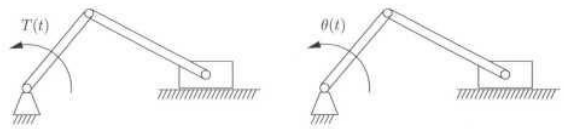
[billedtekst start]**Figure 13.1:** Relationship between mechanics and kinematics[billedtekst slut]

Basically we may say that in kinetics we look at the forces and derive the motion which is the result of these forces, in kinematics we start from the motion without any necessary knowledge of the forces. A simple example of this conceptual difference is illustrated by the slider-crank mechanism in Figure 13.2.

If we want to apply kinematics to find the motion, it is necessary that we have as many driving constraints as there are degrees of freedom (d.o.f.) in the system. The driving constraint should be given as a function of time and preferably in an analytical form as e.g.

$$\theta(t) = K_1 \cdot \frac{t}{s} + K_2 \tag{13.1}$$

where K_1 and K_2 are dimensionless constants, t is time here measured in s.



[billedtekst start]**Figure 13.2:** Conceptual difference between kinetics and kinematics. In kinetics the motion is controlled by the forces/moments, here the torque $T(t)$. In kinematics the motion is controlled directly, here by the rotation angle $\theta(t)$ of the crank.[billedtekst slut]

13.2 Joints in 2D

In the simplest form multibody dynamics deals with rigid bodies connected through joints. In 2D the primary joints are (lower order pairs)

- Revolute joint
- Translational joint

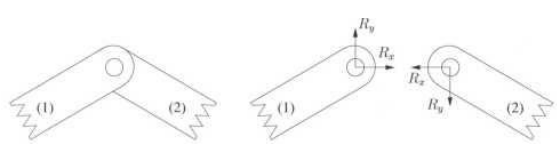
The secondary and more involved joints (higher order pairs) are

- Point contact (sliding) joint
- Rolling joint

In the following the joints are described mathematical and described by constraints. The concept of constraints and the mathematical description of the revolute joint are presented in Section 7.3.2. In relation to the specific joints we also add some practical aspects.

Revolute joint

Figure 13.3 shows a simple schematic revolute joint, in the exploded view the reaction forces are added. The contact geometry of a revolute joint is a cylinder, i.e. the contact is over a surface. If the contact geometry of a joint if a surface we have a lower order pair, we classify a joint as a higher order pair if the contact is a line or a point, therefore the revolute joint is a lower order pair. Lower order pairs are generally simpler than higher order pairs. The constraints of lower order pairs can be defined without specific knowledge of the physical layout of the joint, contrary to higher order pairs.



[billedtekst start]**Figure 13.3:** Schematically representation of a revolute joint and the corresponding reaction forces.[billedtekst slut]

The number of d.o.f. that is removed by adding a revolute joint between two bodies is two. For any joint the number of d.o.f. that is removed by the joint corresponds to the number of independent reaction forces/moments. The mathematical constraints of the revolute joint are given in Section 7.3.2.

The general positive points related to a revolute joint are

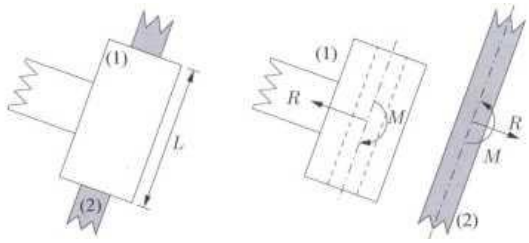
- Simple (inexpensive)
- Constraint forces are transmitted evenly over the contact geometry
- Easy to lubricate
- Rather insensitive to dirt

Due to these points the revolute joint is the most used joint in mechanisms. The reason for using other types of joints is

- Some motions are difficult to produce exclusively by the use of revolute joint in a limited space, other motions cannot be achieved, e.g. indexed motion or gearing.

Translational joint

In Figure 13.4 a simple schematic translational joint is shown, in the exploded view the reaction force and moment are added.



[billedtekst start]**Figure 13.4:** Schematically representation of a translational joint and the corresponding reaction force and moment.[billedtekst slut]

For the translational joint the contact geometry is a surface and therefore the translational joint is a lower order pair. The number of independent reactions is two so the joint removes two d.o.f. corresponding to the reaction force and the reaction moment. Compared to the revolute joint, a translational joint is more difficult to produce and therefore more expensive. This is due to the needed accuracy of the contact geometry. The reaction force, *R*, is usually evenly distributed over the full length of the bushing (the length *L* in Figure 13.4). But the moment *M* must be added to this reaction force, the moment and the force will therefore in a real application act as edge forces (force couple) as seen in Figure 13.5.

This concentration of the contacting force gives rise to an altogether rather uneven force distribution, and implies high demands on the surface finish and on the length *L*. The joint will have a tendency to press the lubricant away and in front of the joint, this also implies a higher sensitivity to dirt.

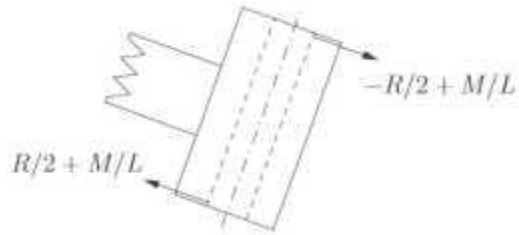
The general positive point related to a translational joint is

- Possibility of precise linear motion within a compact space

Other points in relation to a translational joint are

- More expensive than the revolute joint

- Constraint forces are unevenly distributed

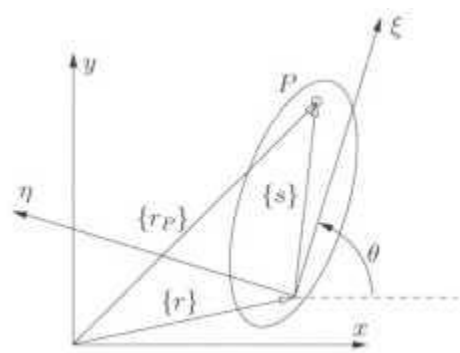


[billedtekst start]**Figure 13.5:** Reaction forces concentrated at the bushing edge.[billedtekst slut]

- Difficult to lubricate
- Sensitive to dirt

To describe the constraints of the translation joint mathematically we first extend the notation used in Section 7.3.2.

In Figure 13.6 a rigid body is shown, a local coordinate system ξ, η has been rigidly attached to the body. That a rigid body in 2D has three d.o.f. (relative to a global/inertial frame) can be seen directly in figure, since the body is fully constraint if the position $\{r\}$ and the orientation θ of the local coordinate system is prescribed.



[billedtekst start]**Figure 13.6:** A Point P attached to rigid body.[billedtekst slut]

The position of a point P rigidly attached to the body can be given as

$$\{r_P\} = \{r\} + \{s\} = \{r\} + [A]\,\{s'\} \tag{13.2}$$

where $\{s'\}$ is the vector $\{s\}$ expressed in the local coordinates and $[A]$ is the transformation matrix that transform from local to global coordinates

$$[A] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{13.3}$$

Utilizing that the body is rigid we have that the vector $\{s'\}$ is constant and that the velocity of point P is given by

$$\{\dot{r}p\} = \{\dot{r}\} + [\mathring{A}]\{s'\} = \{\dot{r}\} + w[B]\{s'\} \tag{13.4}$$

Where $\omega = \dot{\theta}$ and

$$[B] = \begin{bmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{bmatrix} \tag{13.5}$$

If we define a 2D hat vector $\{\hat{s}\}$ as

$$\{s\} = \begin{Bmatrix} s_x \\ s_y \end{Bmatrix} \Rightarrow \{\hat{s}\} = \begin{Bmatrix} -s_y \\ s_x \end{Bmatrix} \tag{13.6}$$

we find that

$$[B]\{s'\} = [A]\{\hat{s}'\} = \{\hat{s}\} \tag{13.7}$$

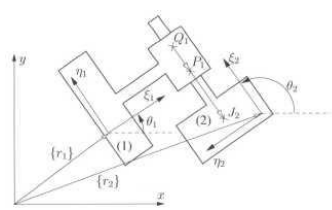
and therefore that

$$\{r_F\} = \{r\} + CJ\{S\} \tag{13.8}$$

The acceleration of point P can be given as

$$\{\ddot{r}_P\} = \{\ddot{r}\} + \omega[\dot{A}]\{\hat{s}'\} + \dot{\omega}[A]\{\hat{s}'\} = \{\ddot{r}\} - \omega^2\{s\} + \dot{\omega}\{\hat{s}\} \tag{13.9}$$

For a translational joint the number of constraints needed is two, i.e. the number of constraint equals the number of independent reaction forces and moments. To formulate the constraints for a translational joint we use Figure 13.7.



[billedtekst start]**Figure 13.7:** Two bodies constraint by a translational joint[billedtekst slut]

The first constraints is rather simple in that the two bodies cannot rotate relative to each other so

$$\phi_1 = (\theta_1 - \theta_1^0) - (\theta_2 - \theta_2^0) = 0 \tag{13.10}$$

where θ_1 and θ_2 are the rotational coordinate for the two bodies and θ_1^0 and θ_2^0 are initial values. The second constraint needed is that the bodies cannot translate normal to the bushing. To facilitate this three point on the symmetry line is defined (Q_1, P_1 and J_2). Two of the points are attached to body 1 and the last point is attached to body 2. Two global vectors can be defined

$$\{s_1\} = \begin{Bmatrix} x_1^P - x_1^Q \\ y_1^P - y_1^Q \end{Bmatrix} \tag{13.11}$$

$$\{d\} = \begin{Bmatrix} x_2^J - x_1^P \\ y_2^J - y_1^P \end{Bmatrix} \tag{13.12}$$

please notice that the coordinates are given as found in (13.2), i.e. for the point Q_1 we find

$$\begin{Bmatrix} x_1^Q \\ y_1^Q \end{Bmatrix} = \{r_1\} + [A_1]\{s_1'^Q\} = \begin{Bmatrix} x_1 \\ y_1 \end{Bmatrix} + \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{Bmatrix} \xi_1^Q \\ \eta_1^Q \end{Bmatrix} \tag{13.13}$$

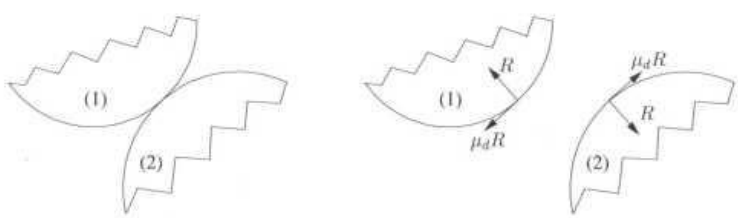
Where $\{\xi_1^Q, \eta_1^Q\}^T$ is the coordinates for point Q_1 in the local coordinate system of body 1.

The vector $\{s_1\}$ is a constant vector as seen from body 1 whereas $\{d\}$ is not a constant because it is defined between two points on two different bodies. If these two vectors stay aligned then the two bodies cannot translate relative to each other perpendicular to this line. Using the hat notation this can be given as

$$\Phi_2 = \{\hat{s}_1\}^T\{d\} = 0 \tag{13.14}$$

The constraints of a translational joint can therefore be given by (13.10) and (13.14). As it was the case for the revolute joint we find that the mathematical constraints can be defined without any specific information about the physical layout of the joint.

Point contact joint



[billedtekst start]**Figure 13.8:** Schematically representation of a point contact joint and the corresponding reaction forces.[billedtekst slut]

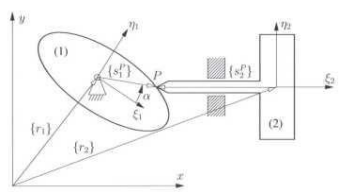
In a point contact as seen in Figure 13.8 the contact geometry is a line (a point in 2D), this joint is therefore a higher order joint. The number of independent reaction forces is one (R) and therefore this type of joint removes one d.o.f., the size of the other force is directly related to the normal force through the dynamic coefficient of friction Compared to the revolute and translational joint the point contact joint will not keep the two bodies together so other joints/forces must be added to achieve this. A common practical application of this type of joint is the valve-cam connection in an engine. The normal

force R will in a real application be of the Hertzian type so that the force is distributed over an area. Compared to a revolute and translational joint the size of this area is very limited, so the joint has higher potential difficulties with regards to lubrication, dirt and wear. Usually the surfaces need to be hardened. The most common application of this type of joint is the gears where we know that excessive lubrication is needed. Typically the gear-box is therefore sealed which also to some extend solves the sensitivity to dirt problem.

The general positive point related to a point contact joint is

- Many possibilities for special motions, i.e. constant gearing or indexed motion.

In formulating the constraints of a point contact joint we need to include the physical layout of the joint and different simplifications can be made depending on the joint at hand. The mathematical formulation is for this reason only exemplified by the configuration in Figure 13.9.



[billedtekst start]**Figure 13.9:** Two bodies constraint by a point contact joint.[billedtekst slut]

The constraints can be formulated as

$$\{\phi\} = \{r_2\} + \{s_2^P\} - \{r_1\} - \{s_1^P\} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{13.15}$$

The problem is that we only need one constraint not two as given here. The vector $\{s_2^P\}$ is a constant vector in body 2, i.e. $\{s_2^P\} = [A_2]\{s'^P_2\}$, but $\{s_1^P\}$ is not constant and depends on the angle α. The simple way of solving this problem is to keep the constraint as formulated here and then add the α variable as an "artificial" coordinate to the coordinates i.e. the coordinates of body 1 become

$$\{x_1, y_1, \theta_1, \alpha\}^T \tag{13.16}$$

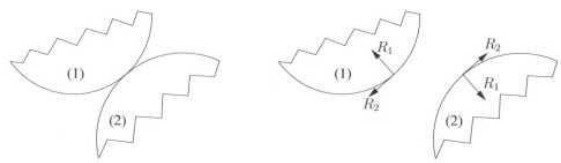
The vector $\{s_1^P\}$ dependency on α can either be given in an analytical form or in numerical form where interpolation then is needed for intermediate points. This shows that the physical layout of the joint is needed for the definition of the point contact joint constraints, the layout is given by the vector $\{s_1^P(\alpha)\}$.

Rolling joint

For the rolling joint seen in Figure 13.10 the contact geometry is the same as for the point contact joint, so this is also a higher order joint. The difference is that there are two independent reaction forces if

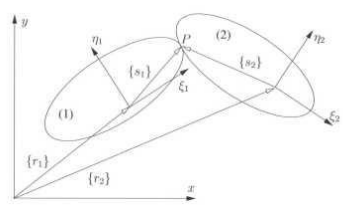
$$R_2 < \mu_s R_1 \tag{13.17}$$

where μ_s is the static coefficient of friction. So this type of joint removes two d.o.f. This type of joint also needs external forces to keep the two bodies together. If the reaction force R_2 exceeds the limiting size then the rolling joint is transformed into a point contact joint. The difference between the two similar joints is also that typically we would like the friction force to be different. The friction should be low for the point contact joint since this gives rise to losses while we for the rolling joint would like the friction force limit to be high in order to broaden the application range, a typical application of this type of joint is the tire/road contact.



[billedtekst start]**Figure 13.10:** Schematically representation of a rolling joint and the corresponding reaction forces.[billedtekst slut]

In formulating the constraints of a rolling joint we also need to include the physical layout of the joint and many specific designs exist. We exemplify by the configuration in Figure 13.11. Compared to the previous joint the formulation is more involved and most easily expressed in velocities. The constraint can be formulated as; the contact point of the two bodies must have the same velocity.



[billedtekst start]**Figure 13.11**: Two bodies constraint by a rolling joint.[billedtekst slut]

It should be noted that the vectors $\{s'_1\}$ and $\{s'_2\}$ are not constant vectors since they are defined from the origo of the local coordinate system to the contact point, and during the rolling motion this point will move. The two constraints can mathematically be expressed as

$$\{\phi\} = \{\dot{r}_1\} + \omega_1\{\dot{s}_1\} - \{\dot{r}_2\} - \omega_2\{\dot{s}_2\} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{13.18}$$

If needed these equations can be integrated in order to be formulated in position coordinates. The layout of the joint enters the mathematical formulation of the joint through the vectors $\{s_1\}$ and $\{s_2\}$.

13.3 Degrees of freedom

An important property of any mechanism is the overall number of degrees of freedom, $N^{d.o.f.}$, this number corresponds directly to the number of independent inputs which are usually controlled by an actuator to fully control the mechanism motion. The overall degrees of freedom is also termed mobility. A free body in 2D has three d.o.f., so if we have N_{bo} free bodies we have $3N_{bo}$ overall d.o.f. for the system.

We may find $N^{d.o.f.}$ from

$$N^{d.o.f.} = 3N_{bo} - 2N_{re} - 2N_{tr} - N_{po} - 2N_{ro} \tag{13.19}$$

where

N_{bo} number of bodies

N_{re} number of revolute joints

N_{tr} number of translational joints

N_{po} number of point contact joints

N_{ro} number of rolling joints

please notice that (13.19) can not be used in all configurations and might lead to inconsistent results due to the geometry of the mechanism. The reason is that the constraints must be linearly independent, i.e. no redundant constraints.

13.4 Position, velocity and acceleration analysis

If a mechanism with N_{bo} bodies is fully constrainted, i.e., we have a driver constraint $\Phi_d = 0$ for each of the independent inputs, it is possible to find the position, velocity and acceleration by kinematic analysis. The number of driver constraints N_{dr} is

$$N_{dr} = N^{d.o.f.} \tag{13.20}$$

Position analysis

Overall we have a non-linear set of equations

$$\{\Phi\} = \begin{Bmatrix} \{\phi(\{q\})\} \\ \{\phi_d(\{q\}, t)\} \end{Bmatrix} = \{0\} \tag{13.21}$$

where $\{\Phi(\{q\})\}$ contains all kinematic constraints of the mechanism, $\{\Phi_d(\{q\}, t)\}$ contains all driver constraints, t is time and $\{q\}$ is the Cartesian coordinates of the mechanism

$$\{q\} = \{x_1, y_1, \theta_1, x_2, y_2, \theta_2, \ldots, x_{N_{bo}}, y_{N_{bo}}, \theta_{N_{bo}}\}^T \tag{13.22}$$

any "artificial" coordinates can be added to this coordinate vector.

Assuming that (13.21) constitutes N linearly independent equations in N unknowns, this may be solved using Newton-Raphson iteration. This is an iterative numerical procedure that involves the calculation in each iteration step of the Jacobian matrix $[\Phi_q]$ given by

$$[\Phi_q] = \frac{\partial\{\Phi\}}{\partial\{q\}} \tag{13.23}$$

The iteration of the Newton-Raphson iteration involves the following steps

$$[\Phi_q(\{q\}_i)]\{\Delta q\}_i = -\{\Phi(\{q\}_i)\} \tag{13.24}$$

$$\{q\}_{i+1} = \{q\}_i + \{\Delta q\}_i \tag{13.25}$$

where i is the iteration index. The iterations are continued until convergence, i.e., until (13.21) is fulfilled to desired numerical precision. Newton-Raphson iteration does not converge in all cases and usually a suitable initial guess of the position vector is needed. If the time steps are reasonably small the previous position can be used. Alternatively one might use

$$\{q(t+\Delta t)\} = \{q(t)\} + \{\dot{q}(t)\}\Delta t + \frac{1}{2}\{\ddot{q}(t)\}\Delta t^2 \tag{13.26}$$

Velocity analysis

With the position $\{q\}$ found we can find the corresponding velocity $\{\dot{q}\}$ of the mechanism by differentiating the constraints with respect to time

$$\{\dot{\Phi}\} = [\Phi_q]\{\dot{q}\} + \frac{\partial\{\Phi\}}{\partial t} = \{0\} \Rightarrow \tag{13.27}$$

$$[\Phi_q]\{\dot{q}\} = -\frac{\partial\{\Phi\}}{\partial t} \tag{13.28}$$

i.e. no iteration is involved in the velocity analysis and we can reuse the final inverted version of the Jacobian matrix, the numerical cost of finding the velocities is therefore small compared to the position analysis.

Acceleration analysis

With the position $\{q\}$ and velocity $\{\dot{q}\}$ found we can find the accelerations $\{\ddot{q}\}$ by differentiation once more with respect to time.

$$\{\ddot{\Phi}\} = \frac{d([\Phi_q]\{\dot{q}\})}{dt} + \frac{d(\partial\{\Phi\}/\partial t)}{dt} = \{0\} \Rightarrow \tag{13.29}$$

$$[\Phi_q]\{\ddot{q}\} = -\frac{\partial([\Phi_q]\{\dot{q}\})}{\partial\{q\}}\{\dot{q}\} - 2\frac{\partial[\Phi_q]}{\partial t}\{\dot{q}\} - \frac{\partial^2\{\Phi\}}{\partial t^2} \tag{13.30}$$

i.e. no iteration is involved in the acceleration analysis, and the acceleration analysis reuse the final inverted version of the Jacobian matrix.

13.5 Mechanism design

In designing a mechanism there are some overall useful point to follow, these are

- Reuse existing designs (be inspired by already existing design)
- The fewer actuators needed for the motion of the mechanism the better is the design
- The fewer bodies in the mechanism the better is the design
- Use revolute joints instead of translational joints if possible
- Use lower order pairs instead of higher order pairs if possible

13.6 Nomenclature

$\{d\}$	mm	Geometric vector defined in global coordinate system
i	–	Iteration index
$\{q\}$	Mixed	Cartesian coordinate vector of mechanism
$\{s\}$	mm	Geometric vector defined in global coordinate system
$\{s'\}$	mm	Geometric vector defined in local coordinate system
t	s	Time
$\{r\}$	mm	Geometric vector defining centre of local coordinate system in global coordinate system
$\{rp\}$	mm	Geometric vector defining position of point attached to body in the global coordinate system
$[A]$	–	Transformation matrix
K_1, K_2	–	Constants
L	mm	Length of bushing
M	Nm	Moment
$N^{d.o.f.}$	–	Number of degrees of freedom (mobility)
N_{dr}	–	Number of driver constraints

N_{bo}	–	Number of bodies
N_{po}	–	Number of point contact joints
N_{re}	–	Number of revolute joints
N_{ro}	–	Number of rolling joints
N_{tr}	–	Number of translational joints
Q,P,J	–	Points
R	N	Reaction force
α	rad	Angle
μ_d	–	Dynamic coefficient of friction
μ_s	-	Static coefficient of friction
Φ	–	Constraint
θ	rad	Angle
$\{\Phi\}$	–	Constraint vector
$[\Phi_q]$	–	Jacobian matrix

13.7 References

[1] E. J. Haug. *Computer-Aided Kinematics and Dynamics of Mechanical Systams, volume I; Basic Methods.* Allyn and Bacon, Boston, 1987.

[2] P. E. Nikravesh. Computer-Aided Analysis of Mechanical Systems. Prentice-Hall International, Inc., Englewood Cliffs, Nj 07632, 1988.

[3] P. E. Nikravesh. Planar Multibody Dynamics, Formulation Programminf and Applications. Taylor & Francis Group, LLC, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL., 2008.

Appendix A Tables with ISO-tolerances and fits

Table A.l: Selected standard tolerance grades IT for basic sizes up to 500mm.

Tolerance grade		IT5	IT6	IT7	IT8	IT9	IT10	IT11	IT12
Diameter									
>	≤								
mm	mm	μm	μm	μm	μm	μm	μm	μm	μm
1	3	4	6	10	14	25	40	60	100
3	6	5	8	12	18	30	48	75	120
6	10	6	9	15	22	36	58	90	150
10	18	8	11	18	27	43	70	110	180
18	30	9	13	21	33	52	84	130	210
30	50	11	16	25	39	62	100	160	250
50	80	13	19	30	46	74	120	190	300
80	120	15	22	35	54	87	140	220	350
120	180	18	25	40	63	100	160	250	400
180	250	20	29	46	72	115	185	290	460
250	315	23	32	52	81	130	210	320	520
315	400	25	36	57	89	140	230	360	570
400	500	27	40	63	97	155	250	400	630

Table A.2: Upper deviation values for shaft sizes up to 500mm. For j s the deviation is ±IT.

Tol. pos.	a	b	c	cd	d	e	ef	f	fg	9	h	js	
Diameter													
>	≦												
1	3	-270	-140	-60	-34	-20	-14	-10	-6	-4	-2	0	
3	6	-270	-140	-70	-46	-30	-20	-14	-10	-6	-4	0	
6	10	-280	-150	-80	-56	-40	-25	-18	-13	-8	-5	0	
10	18	-290	-150	-95	-	-50	-32	-	-16	-	-6	0	
18	30	-300	-160	-110	-	-65	-40	-	-20	-	-7	0	
30	40	-310	-170	-120	-	-80	-50	-	-25	-	-9	0	
40	50	-320	-180	-130	-	-80	-50	-	-25	-	-9	0	
50	65	-340	-190	-140	-	-100	-60	-	-30	-	-10	0	
65	80	-360	-200	-150	-	-100	-60	-	-30	-	-10	0	
80	100	-380	-220	-170	-	-120	-72	-	-36	-	-12	0	
100	120	-410	-240	-180	-	-120	-72	-	-36	-	-12	0	
120	140	-460	-260	-200	-	-145	-85	-	-43	-	-14	0	
140	160	-520	-280	-210	-	-145	-85	-	-43	-	-14	0	
160	180	-580	-310	-230	-	-145	-85	-	-43	-	-14	0	
180	200	-660	-340	-240	-	-170	-100	-	-50	-	-15	0	
200	225	-740	-380	-260	-	-170	-100	-	-50	-	-15	0	
225	250	-820	-420	-280	-	-170	-100	-	-50	-	-15	0	
250	280	-920	-480	-300	-	-190	-110	-	-56	-	-17	0	

280	315	-1050	-540	-330	-	-190	-110	-	-56	-	-17	0	
315	355	-1200	-600	-360	-	-210	-125	-	-62	-	-18	0	
355	400	-1350	-680	-400	-	-210	-125	-	-62	-	-18	0	
400	450	-1500	-760	-440	-	-230	-135	-	-68	-	-20	0	
450	500	-1650	-840	-480	-	-230	-135	-	-68	-	-20	0	

Table A.3: Lower deviation values for basic shaft sizes up to 500mm.

Tol. pos.		j	j	j	k	k	m	n		P
Tol. grade		5 and 6	7	8	4 to 7	from 8	all	all		all
Diameter										
>	≤									
1	3	-2	-4	-6	0	0	+2	+4		+6
3	6	-2	-4	-	+1	0	+4	+8		+12
6	10	-2	-5	-	+1	0	+6	+10		+15
10	18	-3	-6	-	+ 1	0	+7	+ 12		+18
18	30	-4	-8	-	+2	0	+8	+ 15		+22
30	50	-5	-10	-	+2	0	+9	+ 17		+26
50	80	-7	-12	-	+2	0	+11	+20		+32
80	120	-9	-15	-	+3	0	+13	+23		+37
120	180	-11	-18	-	+3	0	+ 15	+27		+43
180	250	-13	-21	-	+4	0	+ 17	+31		+50
250	315	-16	-26	-	+4	0	+20	+34		+56
315	400	-18	-28	-	+4	0	+21	+37		+62
400	500	-20	-32	-	+5	0	+23	+40		+68

Table A.4: Lower deviation values for basic shaft sizes up to 500mm.

Tol.	pos.	r	s	t	u	V	X	y	z	za	zb	ZC
All tolerance grades												
Diameter												
>	≦											
1	3	+ 10	+ 14	-	+ 18	-	+20	-	+26	+32	+40	+60
3	6	+ 15	+ 19	-	+23	-	+28	-	+35	+42	+50	+80
6	10	+ 19	+23	-	+28	-	+34	-	+42	+52	+67	+97
10	14	+23	+28	-	+33	-	+40	-	+50	+64	+90	+130
14	18	+23	+28	-	+33	+39	+45	-	+60	+77	+108	+150
18	24	+28	+35	-	+41	+47	+54	+63	+73	+98	+136	+188
24	30	+28	+35	+41	+48	+55	+64	+75	+88	+118	+160	+218
30	40	+34	+43	+48	+60	+68	+80	+94	+112	+148	+200	+274
40	50	+34	+43	+54	+70	+81	+97	+114	+136	+180	+242	+325
50	65	+41	+53	+66	+87	+102	+122	+144	+172	+226	+300	+405
65	80	+43	+59	+75	+102	+120	+146	+174	+210	+274	+360	+480
80	100	+51	+71	+91	+ 124	+146	+178	+214	+258	+335	+445	+585
100	120	+54	+79	+104	+144	+172	+210	+254	+310	+400	+525	+690
120	140	+63	+92	+122	+170	+202	+248	+300	+365	+470	+620	+800
140	160	+65	+100	+134	+190	+228	+280	+340	+415	+535	+700	+900
160	180	+68	+108	+146	+210	+252	+310	+380	+465	+600	+780	+1000
180	200	+77	+ 122	+ 166	+236	+284	+350	+425	+520	+670	+880	+1150
200	225	+80	+ 130	+ 180	+258	+310	+385	+470	+575	+740	+960	+1250

225	250	+84	+ 140	+196	+284	+340	+425	+520	+640	+820	+ 1050	+ 1350
250	280	+94	+ 158	+218	+315	+385	+475	+580	+710	+920	+1200	+ 550
280	315	+98	+ 70	+240	+350	+425	+525	+650	+790	+1000	+1300	+700
315	355	+108	+90	+268	+390	+475	+590	+730	+900	+1150	+1500	+1900
355	400	+114	+208	+294	+435	+530	+660	+820	+1000	+1300	+1650	+2100
400	450	+126	+232	+330	+490	+595	+740	+920	+1100	+1450	+1850	+2400
450	500	+132	+252	+360	+540	+660	+820	+1000	+1250	+1600	+2100	+2600

Table A.5: Lower deviation values for basic hole sizes up to 500mm. For Js the deviation is ±IT.

Tol.	pos.	A	B	C	CD	D	E	EF	F	FG	G	H	Js
Diameter													
>	≦												
1	3	+270	+140	+60	+34	+20	+14	+10	+6	+4	+2	0	
3	6	+270	+140	+70	+46	+30	+20	+14	+10	+6	+4	0	
6	10	+280	+150	+80	+56	+40	+25	+18	+13	+8	+5	0	
10	18	+290	+150	+95	-	+50	+32	-	+16	-	+6	0	
18	30	+300	+160	+110	-	+65	+40	-	+20	-	+7	0	
30	40	+310	+170	+120	-	+80	+50	-	+25	-	+9	0	
40	50	+320	+180	+130	-	+80	+50	-	+25	-	+9	0	
50	65	+340	+190	+140	-	+100	+60	-	+30	-	+10	0	
65	80	+360	+200	+150	-	+100	+60	-	+30	-	+10	0	
80	100	+380	+220	+170	-	+120	+72	-	+36	-	+12	0	
100	120	+410	+240	+180	-	+120	+72	-	+36	-	+12	0	
120	140	+460	+260	+200	-	+145	+85	-	+43	-	+14	0	
140	160	+520	+280	+210	-	+145	+85	-	+43	-	+14	0	
160	180	+580	+310	+230	-	+145	+85	-	+43	-	+14	0	
180	200	+660	+340	+240	-	+170	+100	-	+50	-	+15	0	
200	225	+740	+380	+260	-	+170	+100	-	+50	-	+15	0	
225	250	+820	+420	+280	-	+170	+100	-	+50	-	+15	0	
250	280	+920	+480	+300	-	+190	+110	-	+56	-	+17	0	

280	315	+1050	+540	+330	-	+190	+110	-	+56	-	+17	0	
315	355	+1200	+600	+360	-	+210	+125	-	+62	-	+18	0	
355	400	+1350	+680	+400	-	+210	+125	-	+62	-	+18	0	
400	450	+1500	+760	+440	-	+230	+135	-	+68	-	+20	0	
450	500	+ 1650	+840	+480	-	+230	+135	-	+68	-	+20	0	

Table A.6: Upper deviation values for basic hole sizes up to 500mm.

Tol.	pos.	J	J	J	K	K	M	M	N	N	A	A	A	A
Tol.	grade	6	7	8	to 8	fr. 9	to 8	fr. 9	to 8	fr. 9	5	6	7	8
Diameter														
>	≦													
1	3	+2	+4	+6	0	0	-2	-2	-4	-4	0	0	0	0
3	6	+5	+6	+ 10	–1 + Δ	-	–4 +Δ	–4	–8 + Δ	0	1	3	4	6
6	10	+5	+8	+ 12	–1 + Δ	–	–6 + Δ	–6	–10 + Δ	0	2	3	6	7
10	18	+6	+ 10	+ 15	–1 + Δ	–	–7 + Δ	–7	–12 + Δ	0	3	3	7	9
18	30	+8	+12	+20	–2 + Δ	–	–8 + Δ	–8	–15 + Δ	0	3	4	8	12
30	50	+ 10	+ 14	+24	–2 + Δ	–	–9 + Δ	–9	–17 + Δ	0	4	5	9	14
50	80	+ 13	+18	+28	–2 + Δ	–	–11 + Δ	–11	–20 + Δ	0	5	6	11	16
80	120	+ 16	+22	+34	–3 + Δ	–	–13 + Δ	–13	–23 + Δ	0	5	7	13	19
120	180	+ 18	+26	+41	–3 + Δ	–	–15 + Δ	–15	–27 + Δ	0	6	7	15	23
180	250	+22	+30	+47	–4 + Δ	–	–17 + Δ	–17	–31 + Δ	0	6	9	17	26
250	315	+25	+36	+55	–4 + Δ	–	–20 + Δ	–20	–34 + Δ	0	7	9	20	29
315	400	+29	+39	+60	–4 + Δ	–	–21 + Δ	–21	–37 + Δ	0	7	11	21	32
400	500	+33	+43	+66	–5 + Δ	–	–23 + Δ	–23	–40 + Δ	0	7	13	23	34

Table A.7: Upper deviation values for basic hole sizes up to 500mm. For tolerance positions from P to ZC with tolerance grades up to (and including) 7 the same tolerance grade as from tolerance grade 8 should be used but with the appropriate Δ value added.

Tol.	pos.	P	R	S	T	u	V	X	Y	z	ZA	ZB	ZC
All tolerance grades													
Diameter													
>	≦												
1	3	-6	-10	-14	-	-18	-	-20	-	-26	-32	-40	-60
3	6	-12	-15	-19	-	-23	-	-28	-	-35	-42	-50	-80
6	10	-15	-19	-23	-	-28	-	-34	-	-42	-52	-67	-97
10	14	-18	-23	-28	-	-33	-	-40	-	-50	-64	-90	-130
14	18	-18	-23	-28	-	-33	-39	-45	-	-60	-77	-108	-150
18	24	-22	-28	-35	-	-41	-47	-54	-63	-73	-98	-136	-188
24	30	-22	-28	-35	-41	-48	-55	-64	-75	-88	-118	-160	-218
30	40	-26	-34	-43	-48	-60	-68	-80	-94	-112	-148	-200	-274
40	50	-26	-34	-43	-54	-70	-81	-97	-114	-136	-180	-242	-325
50	65	-32	-41	-53	-66	-87	-102	-122	-144	-172	-226	-300	-405
65	80	-32	-43	-59	-75	-102	-120	-146	-174	-210	-274	-360	-480
80	100	-37	-51	-71	-91	-124	-146	-178	-214	-258	-335	-445	-585
100	120	-37	-54	-79	-104	-144	-172	-210	-254	-310	-400	-525	-690
120	140	-43	-63	-92	-122	-170	-202	-248	-300	-365	-470	-620	-800
140	160	-43	-65	-100	-134	-190	-228	-280	-340	-415	-535	-700	-900
160	180	-43	-68	-108	-146	-210	-252	-310	-380	-465	-600	-780	-1000

180	200	-50	-77	-122	-166	-236	-284	-350	-425	-520	-670	-880	-1150
200	225	-50	-80	-130	-180	-258	-310	-385	-470	-575	-740	-960	-1250
225	250	-50	-84	-140	-196	-284	-340	-425	-520	-640	-820	-1050	-1350
250	280	-56	-94	-158	-218	-315	-385	-475	-580	-710	-920	-1200	-1550
280	315	-56	-98	-170	-240	-350	-425	-525	-650	-790	-1000	-1300	-1700
315	355	-62	-108	-190	-268	-390	-475	-590	-730	-900	-1150	-1500	-1900
355	400	-62	-114	-208	-294	-435	-530	-660	-820	-1000	-1300	-1650	-2100
400	450	-68	-126	-232	-330	-490	-595	-740	-920	-1100	-1450	-1850	-2400
450	500	-68	-132	-252	-360	-540	-660	-820	-1000	-1250	-1600	-2100	-2600

Table A.8: Interference fits with hole and shaft base.

Fit		H7/r6 or R7/h6		H7/s6 or S7/h6		H7/t6 or T7/h6		H7/u6 or U7/h6		H7/v6 or V7/h6	
Diameter											
>	≦	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**
3	6	3	23	7	27			11	31		
6	10	4	28	8	32			13	37		
10	14	5	34	10	39			15	44		
14	18	5	34	10	39			15	44		
18	24	7	41	14	48			20	53	26	60
24	30	7	41	14	48	20	54	27	61	34	68
30	40	9	50	18	59	23	64	35	76	43	84
40	50	9	50	18	59	29	70	45	86	56	97
50	65	11	60	23	72	36	85	57	106	72	121
65	80	13	62	29	78	45	94	72	121	90	139
80	100	16	73	36	93	55	112	89	146	111	168
100	120	19	76	44	101	69	126	109	166	137	194
120	140	23	88	52	117	82	147	130	195	162	227
140	160	25	90	60	125	94	159	150	215	188	253
160	180	28	93	68	133	106	171	170	235	212	277
180	200	31	106	76	151	120	185	190	265	238	313
200	225	34	109	84	159	134	209	212	287	264	339

225	250	38	113	94	169	150	225	238	313	294	369
250	280	42	126	106	190	166	250	263	347	333	417
280	315	46	130	118	202	188	272	298	382	373	457
315	355	51	144	133	226	211	304	333	426	418	511
355	400	57	150	151	244	237	330	378	471	473	566
400	450	63	166	169	272	267	370	427	530	532	635
450	500	69	172	189	292	297	400	477	580	597	700

Table A.9: Interference fits with hole and shaft base.

Fit		H7/x6 or X7/h6		H7/z6 or z7/h6		H7/za6 or ZA7/h6		H7/zb6 or ZB7/h6		H7/zc6 or ZC7/h6	
Diameter											
>	≦	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**
3	6	16	36	23	43	30	50	38	58	68	88
6	10	19	43	27	51	37	61	52	76	82	106
10	14	22	51	32	61	46	75	72	101	112	141
14	18	27	56	42	71	59	88	90	119	132	161
18	24	33	67	52	86	77	111	115	149	167	201
24	30	43	77	67	101	97	131	139	173	197	231
30	40	55	96	87	128	123	164	175	216	249	290
40	50	72	113	111	152	155	196	217	258	310	341
50	65	92	141	142	191	196	245	270	319	370	424
65	80	116	165	180	229	244	293	330	379	450	499
80	100	143	200	223	280	300	357	410	467	550	607
100	120	175	232	275	332	365	422	490	547	655	712
120	140	208	273	325	390	430	495	580	645	760	825
140	160	240	305	375	440	495	560	660	725	860	925
160	180	270	335	425	490	560	625	740	805	960	1025
180	200	304	379	474	549	624	699	834	909	1104	1179
200	225	339	414	529	604	694	769	914	989	1204	1279

225	250	379	454	594	669	774	849	1004	1079	1304	1379
250	280	423	507	658	742	868	952	1148	1232	1498	1582
280	315	473	557	738	822	942	1052	1248	1332	1648	1732
315	355	533	626	843	936	1093	1186	1443	1536	1843	1936
355	400	603	696	943	1036	1243	1336	1593	1686	2043	2136
400	450	677	780	1037	1140	1387	1490	1787	1890	2337	2440
450	500	757	860	1187	1290	1537	1640	2037	2140	2537	2640

Appendix B Stress concentration factors

The graphs of stress concentration factors on the next pages are taken from [2], [3] and [1], Please note that the curves are plots of estimated curve-fits given in the graphs. The curves are also shown outside the validity domain given on the graphs, this is done for illustration purposes.

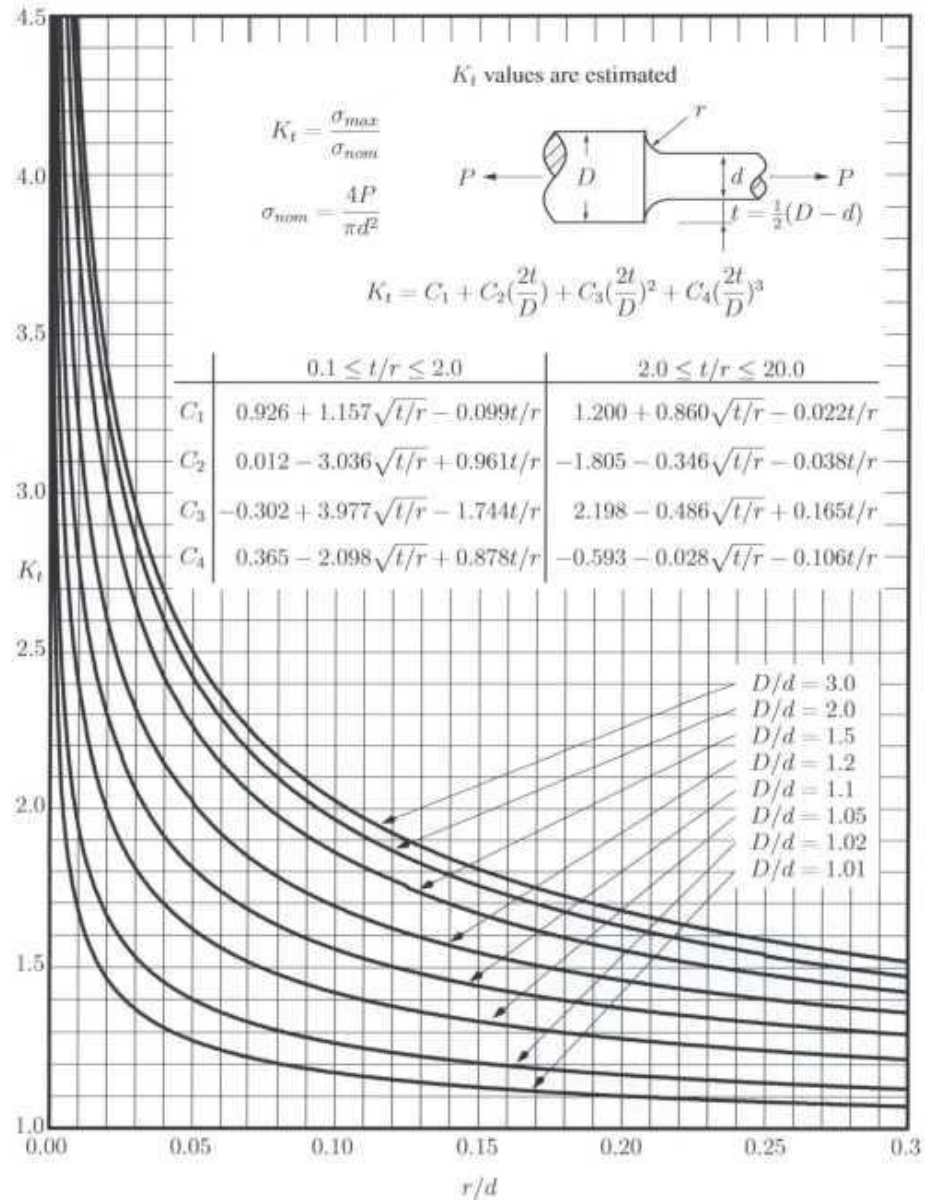
B.1 References

[1] R. L. Norton. *Machine design, an integrated approach, fifth edition*. Prentice-Hall Inc., Upper Saddle River, N.J. 07458, 2014.

[2] R. E. Peterson. *Stress concentration design factors.* John Wiley & Son, Inc., New York, USA, 1953.

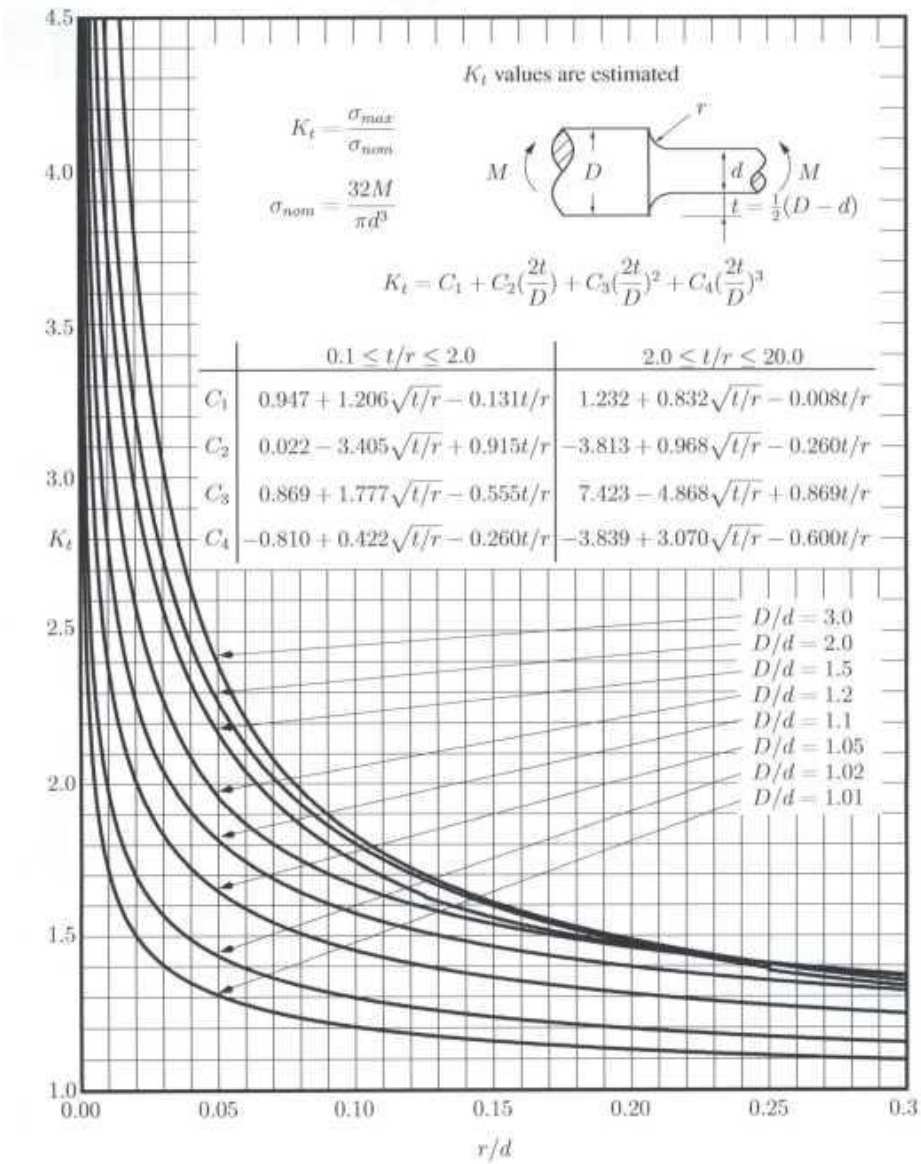
[3] W. D. Pilkey. *Peterson 's Stress Concentration Factors.* John Wiley and Sons, New York, USA, 2nd edition, 1997.

Side 284



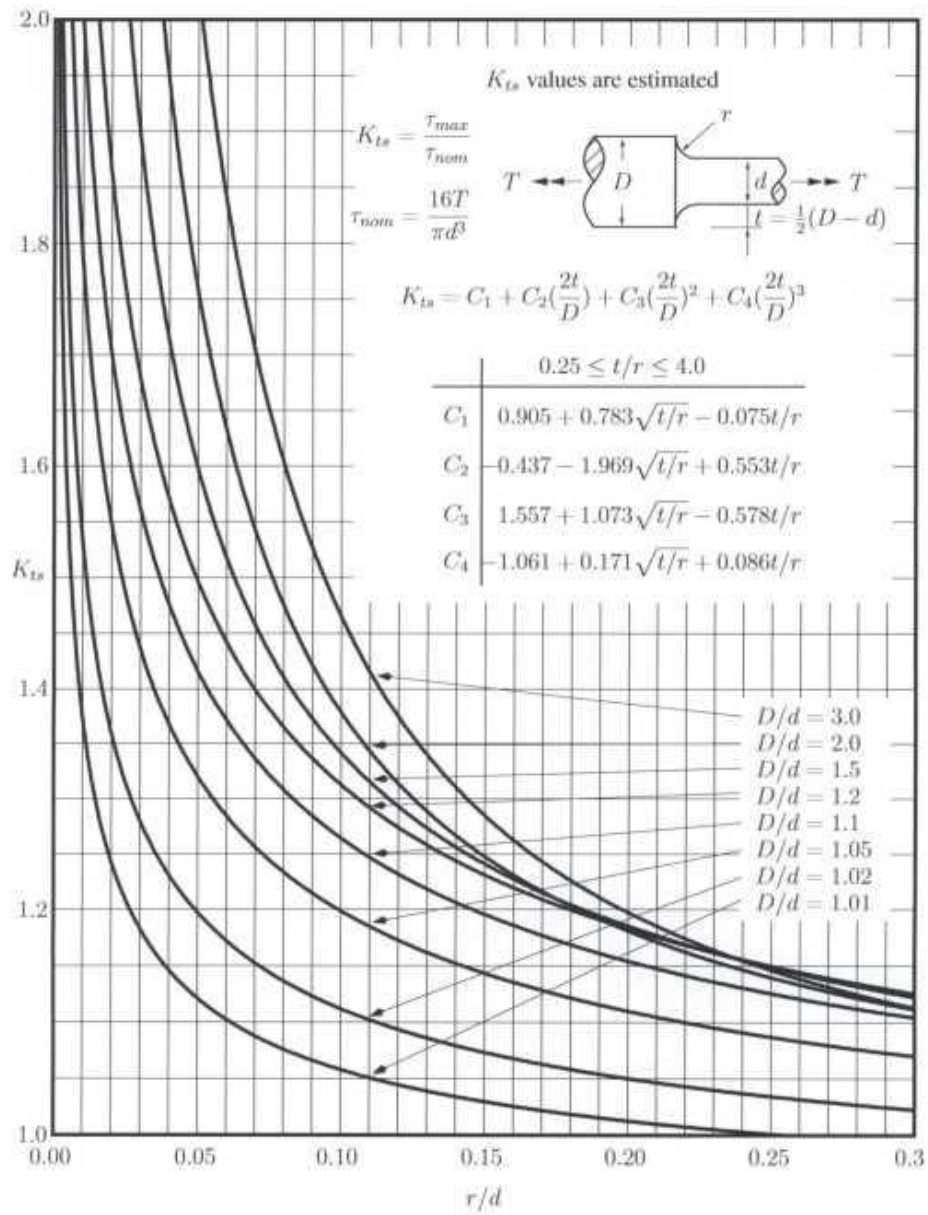
[billedtekst start]**Figure B.1:** Stress concentration factor K_t for a shaft in tension with a fillet, curve-fits taken from [3].[billedtekst slut]

Side 285



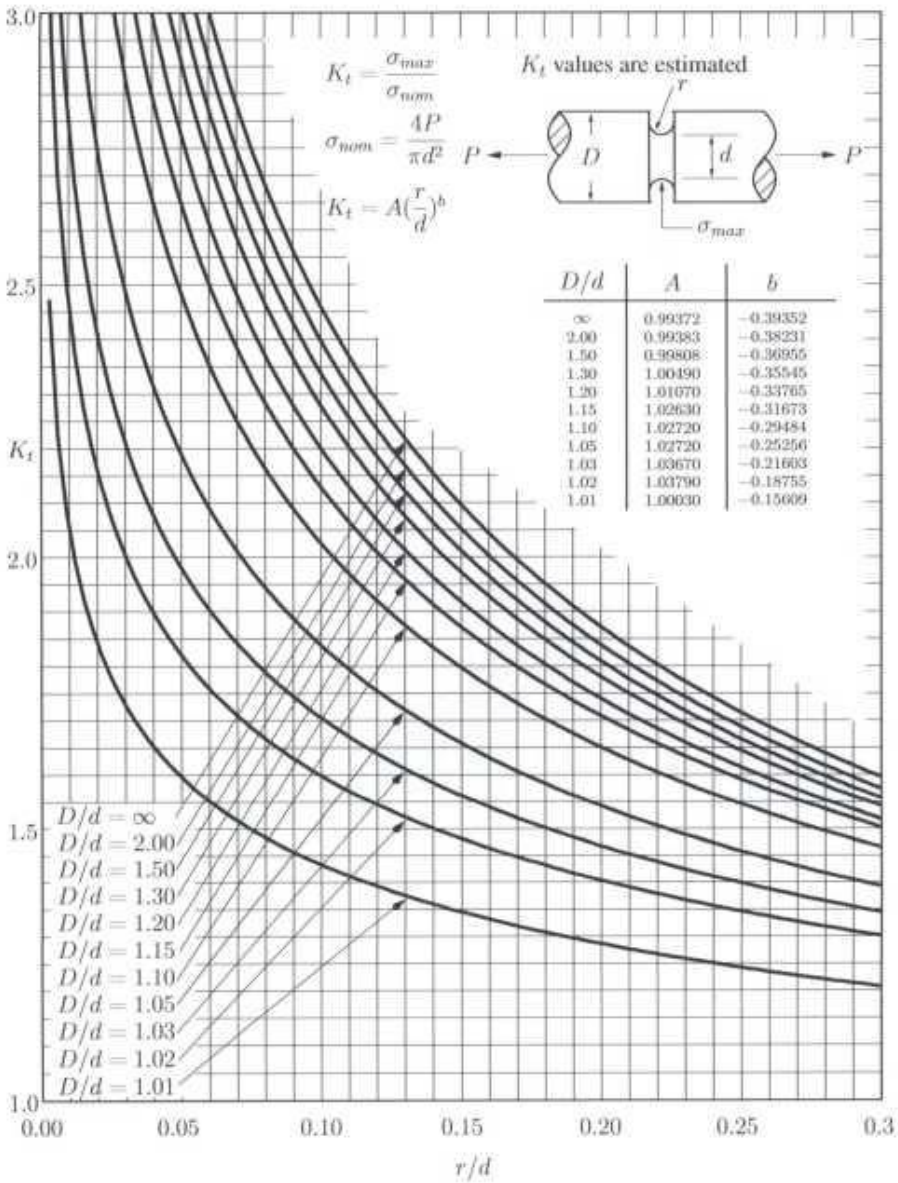
[billedtekst start]**Figure B.2:** Stress concentration factor K_t for a shaft in bending with a fillet, curve-fits taken from [3]. Note that some of the curve-fits cross each other, this is not physical correct but due to the given curve-fits.[billedtekst slut]

Side 286



[billedtekst start]**Figure B.3:** Stress concentration factor K_{ts} for a shaft in torsion with a fillet, curve-fits taken from [3], Note that some of the curve-fits cross each other, this is not physical correct but due to the given curve-fits.[billedtekst slut]

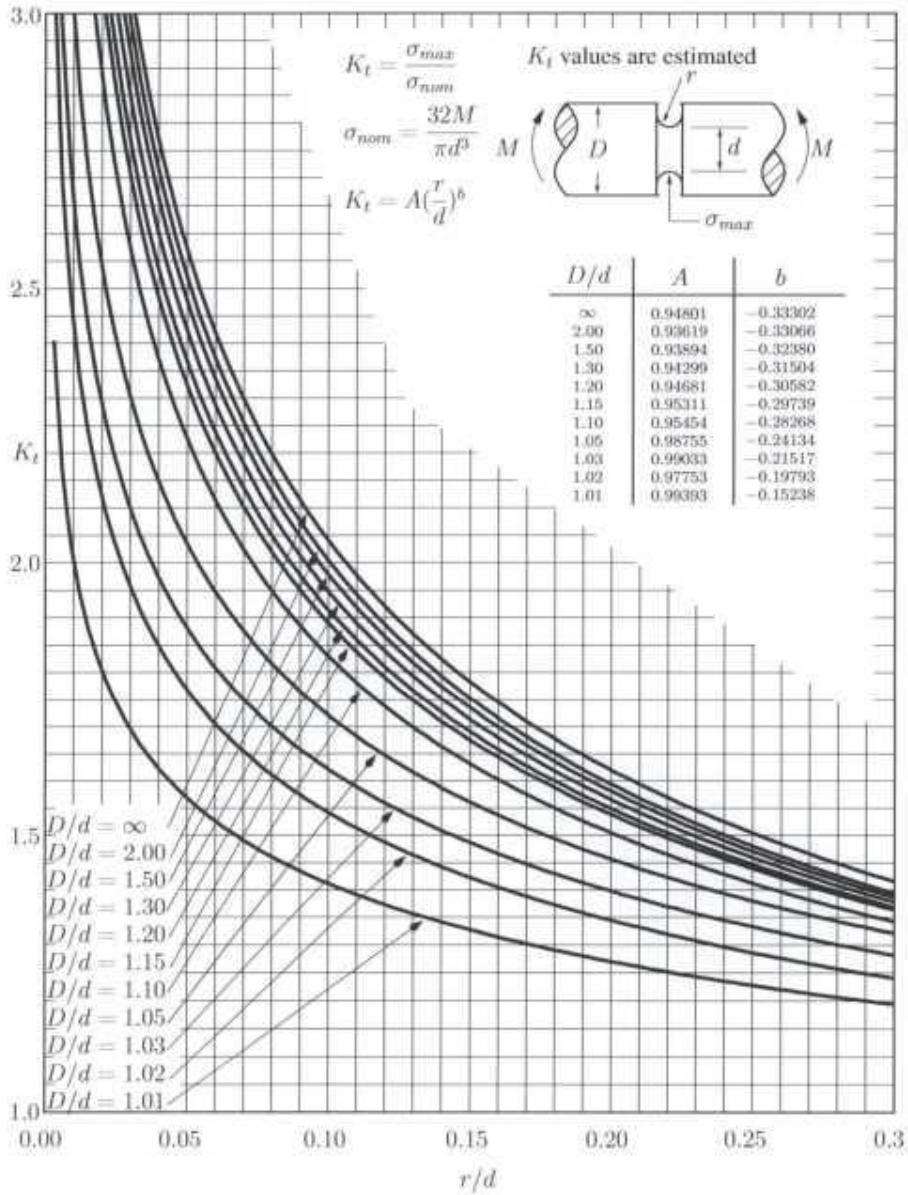
Side 287



D/d	A	b
∞	0.99372	−0.39352
2.00	0.99383	−0.38231
1.50	0.99808	−0.36955
1.30	1.00490	−0.35545
1.20	1.01070	−0.33765
1.15	1.02630	−0.31673
1.10	1.02720	−0.29484
1.05	1.02720	−0.25256
1.03	1.03670	−0.21603
1.02	1.03790	−0.18755
1.01	1.00030	−0.15609

[billedtekst start]**Figure B.4:** Stress concentration factor K_t for a shaft in tension with a U-shaped groove, curve-fits taken from [1].[billedtekst slut]

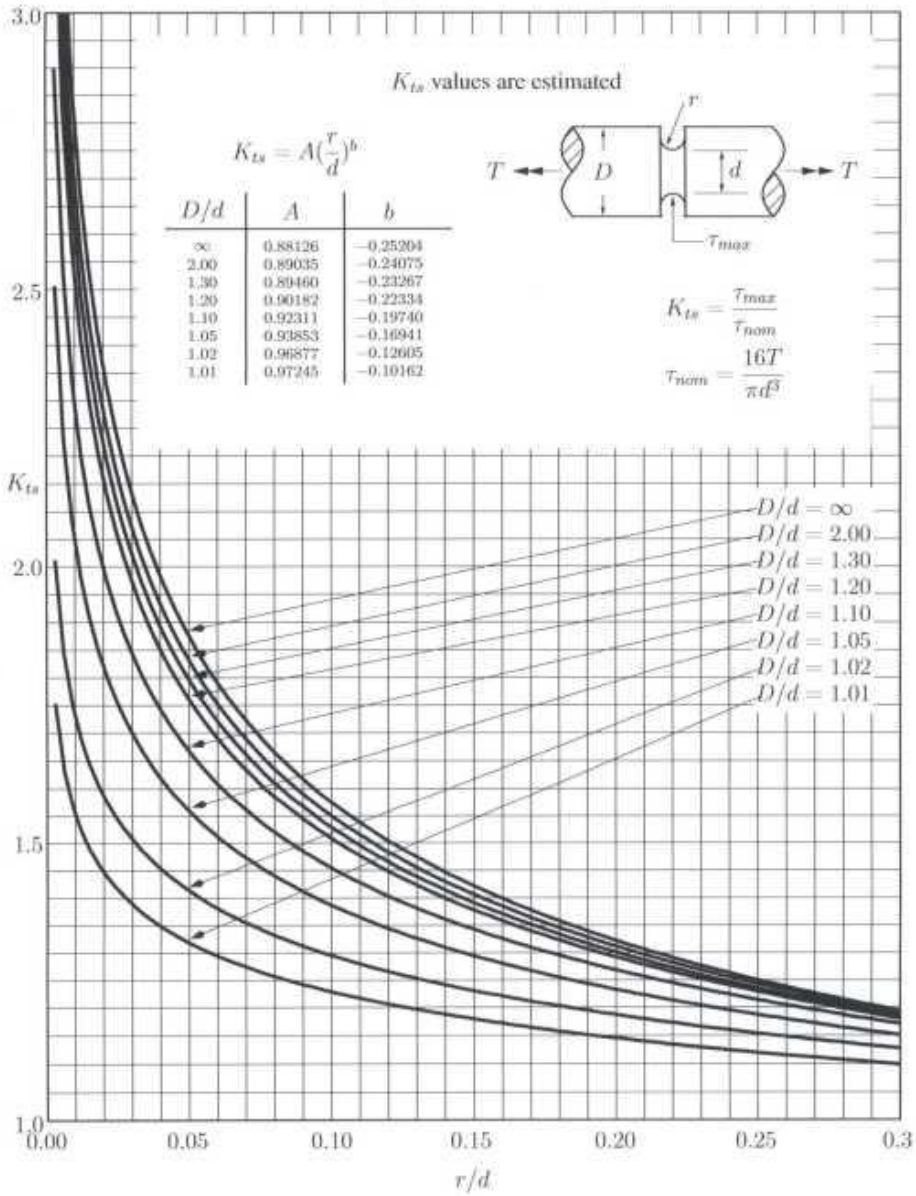
Side 288



D/d	A	b
∞	0.94801	−0.33302
2.00	0.93619	−0.33066
1.50	0.93894	−0.32380
1.30	0.94299	−0.31504
1.20	0.94681	−0.30582
1.15	0.95311	−0.29739
1.10	0.95454	−0.28268
1.05	0.98755	−0.24134
1.03	0.99033	−0.21517
1.02	0.97753	−0.19793
1.01	0.99393	−0.15238

[billedtekst start]**Figure B.5:** Stress concentration factor K_t for a shaft in bending with a U-shaped groove, curve-fits taken from [1].[billedtekst slut]

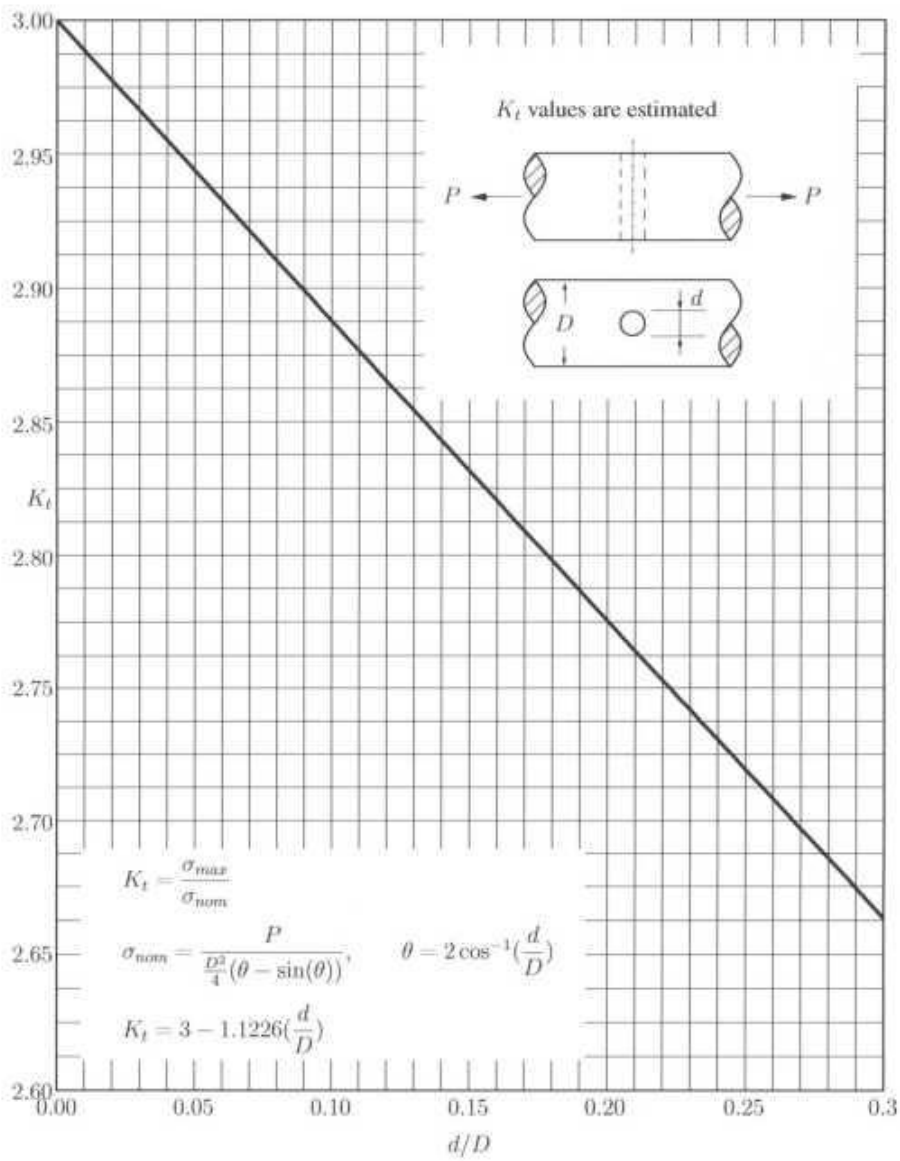
Side 289



D/d	A	b
∞	0.88126	−0.25204
2.00	0.89035	−0.24075
1.30	0.89460	−0.23267
1.20	0.90182	−0.22334
1.10	0.92311	−0.19740
1.05	0.93853	−0.16941
1.02	0.96877	−0.12605
1.01	0.97245	−0.10162

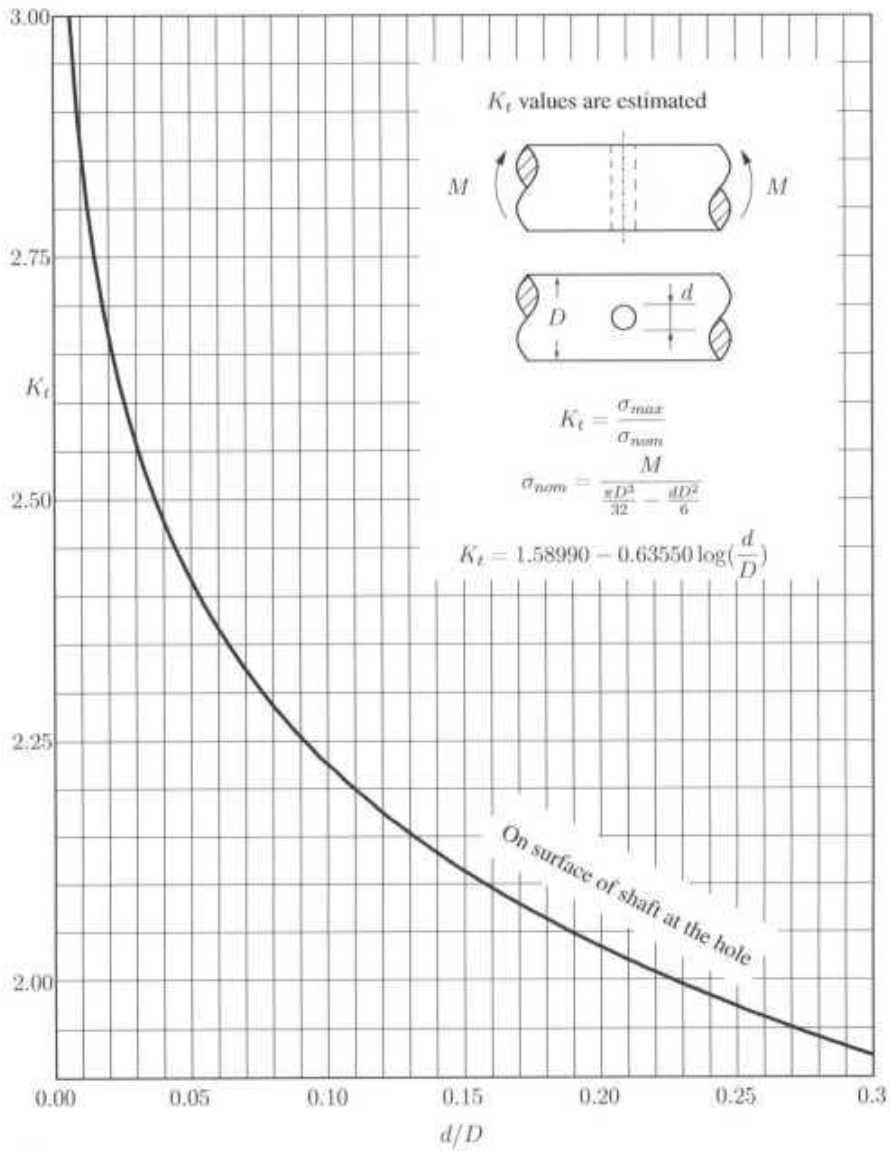
[billedtekst start]**Figure B.6:** Stress concentration factor K_{ts} for a shaft in torsion with a U-shaped groove, curve-fits taken from [1].[billedtekst slut]

Side 290



[billedtekst start]**Figure B.7:** Stress concentration factor K_t for a shaft in tension with a transverse hole. The shown curve is a curve-fit of the curve found in [3].[billedtekst slut]

Side 291

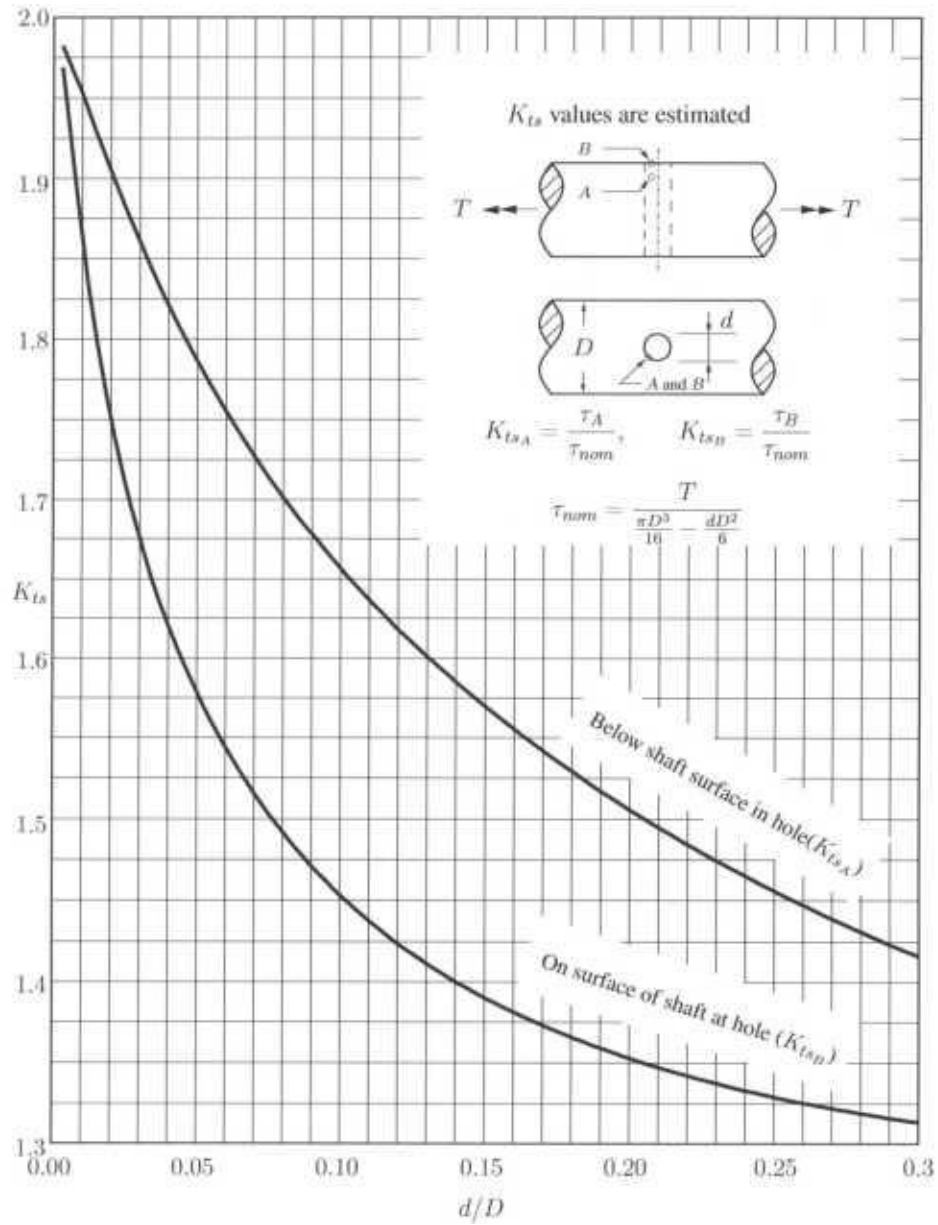


[billedtekst start]**Figure B.8:** Stress concentration factor K_t for a shaft in bending with a transverse hole, curve-fits taken from [1].[billedtekst slut]

Side 292

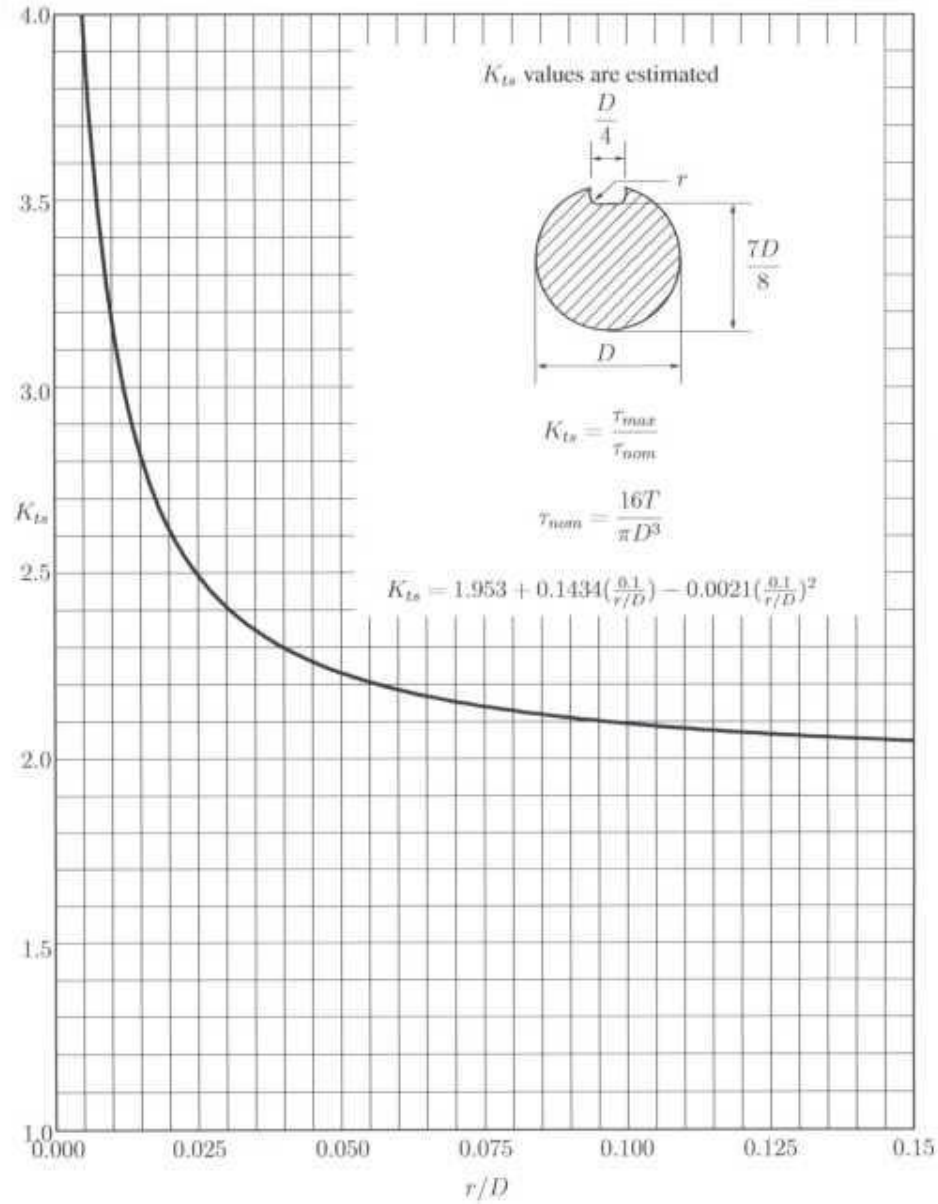
$$K_{ts_A} = 0.0475\log(\frac{d}{D})^4 + 0.2949\log(\frac{d}{D})^3 + 0.4776\log(\frac{d}{D})^2 - 0.2149\log(\frac{d}{D}) + 1.2115$$

$$K_{ts_B} = 0.0038\log(\frac{d}{D})^4 + 0.1305\log(\frac{d}{D})^3 + 05053\log(\frac{d}{D})^2 + 0.2496\log(\frac{d}{D}) + 1.3242$$



[billedtekst start]**Figure B.9:** Stress concentration factor K_{ts} for a shaft in torsion with a transverse hole. The shown curves are curve-fits of the curves found in [2].[billedtekst slut]

Side 293



[billedtekst start]**Figure B.10:** Stress concentration factor K_{ts} for a shaft in torsion with a keyseat, curve-fit taken from [3].[billedtekst slut]

Index

C

D

E

F

G

H

I

J

K

L

M

N

P

R

S

T

U

V

W