A REVIEW OF GROOVED DYNAMIC GAS BEARINGS

Lili Gu*, Elliott Guenat, Jürg Schiffmann
Laboratory for Applied Mechanical Design
École polytechnique fédérale de Lausanne
Neuchâtel 2000, Switzerland

Abstract

This paper offers an extensive review of publications dealing with the modeling, the design and the experimental investigation of grooved dynamic gas lubricated bearings. Recent years have witnessed a rise in small-scale and high-speed turbomachinery applications. Besides the well-known gas foil bearings, grooved bearings offer attractive advantages, which unveil their potential in particular at small scale due to the structural simplicity as well as satisfying predictability. This paper starts with a general background of the application of gas lubricated bearings and introduces and compares the different gas bearing topologies. Further, the state-of-the-art modeling of grooved gas lubricated bearings is introduced in a systematic way assessing the advantages and inconveniences of two major approaches, the Narrow Groove Theory (NGT) and direct discretization method. Since NGT method is an elegant and efficient approach to model the complex effects of periodic grooves, a critical section is dedicated to the Narrow Groove Theory. In a second phase different models to include additional physical phenomena such as real gas lubrication, rarefaction or turbulence effects are reviewed. The paper concludes with a critical assessment of the state-of-the-art and indicates potential fields of research that would allow to shed more light into the understanding of these bearings, as well as with some thoughts on the integrated design methodologies of gas bearing supported rotors.

1 Introduction

1.1 Typical Gas Bearing Applications
The four main bearing technologies encountered in high-speed turbomachinery applications [1] are (1) rolling-element bearings, (2) active magnetic bearings, (3) liquid fluid film bearings and (4) gas-lubricated bearings (GBs). Rolling-element bearings are standardized machine elements. They are well suited to withstand shocks and high loads even at rest. However, the mechanical contact between the rolling elements and the inner and outer races and the inertial forces limit its lifetime at high speeds and requires well-controlled lubrication [2], which calls for auxiliaries such as oil management systems, seals and pumps.

Active magnetic bearings generate load capacity through controlled magnetic fields. As a consequence, they do not require lubricants and can sustain a load under vacuum conditions, which makes it an appealing solution for oil-free turbomachinery applications. However, it requires

* Corresponding author. Lili Gu, Scientist, Laboratory for Applied Mechanical Design, École polytechnique fédérale de Lausanne, Neuchâtel 2000, Switzerland, lili.gu@epfl.ch
electric control auxiliaries and catcher bearings and thus needs more space compared to other bearing technologies.

Fluid film bearings with liquid lubricants, most often oil, can sustain high loads compared to GBs but require auxiliaries for providing the lubricant. By nature, they exhibit a cross-coupled behavior, which affects the rotordynamics. Moreover, cavitation of the liquid lubricant prevents their applicability at high speeds [3].

GBs often utilize gas lubricant that originates directly from the working gas processed by the rotating machine. The gaseous lubricant can either be operated under ideal gas conditions such as in air cycle machines [4] or gas processing [5] or close to the saturation line such as in heat pumps or chillers [6]. In dynamic bearings, no auxiliaries are needed, and the machines can easily be conceived as hermetic units, thus avoiding the need for complicated seals to prevent leakage. Further, the low viscosity of gas [7,8] reduces drag losses and improves the overall efficiency. Another advantage of gaseous lubricants is associated with its relatively stable physical and chemical properties since there is no vaporization, cavitation, decomposition or solidification over a wide range of temperatures. The low viscosity, on the other hand, limits both the load capacity and the damping [9], which together with the cross-coupled behavior may cause rotodynamic instabilities [10,11]. These effects are typically balanced by smaller bearing clearance to diameter ratios compared to oil-film bearings, usually by one order of magnitude [2]. As a result, GBs are very sensitive to radial expansion for journal bearings resulting from thermal distortion or centrifugal growth [4]. In addition, external damping through flexible support, e.g., with rubber-‘O’-ring-support for journal bearings [1] and viscoelastic dampers behind tilting pads for thrust bearings [12], can be introduced to increase the stability threshold [13–15]. However, it is difficult to predict the dynamic behavior of such assemblies due to the complex behavior of the viscoelastic support material [11].

With the increasing demand for lighter, more compact, economical, and environmentally-friendly energy conversion systems, gas bearings play a crucial role in oil-free turbomachinery [16–18]. Since the first gas-bearing supported turbomachinery delivered by UK Atomic Energy Establishment in 1958, gas bearings have demonstrated their advantages in various applications such as in turbo-compressors, gas turbines and turbochargers/expanders [4,19–21]. As shown in Figure 1, typical applications scale from MEMS-sized gas turbines [20,22], to 10-20 mm tip diameter fuel-cell (FC) application [23], 10-30 mm tip diameter turbomachinery applied for heating, ventilation or air conditioning (HVAC) [1,24,25], and to large size systems such as turbochargers and air cycle machines [26]. The power of known gas bearing supported turbomachinery ranges from an order of magnitude of 10W to 100kW, while the rotor speeds range from 10 krpm up to more than 1000 krpm.
Between the 1950’s and 1970’s interesting reviews emerged on the design of GBs by Gross [9] and Castelli and Pirvics [27]. Powell [8] provided an extension to manufacturing concerns and Fuller [7] pointed out potential industrial applications for the GB technology, whereas Sternlight and Arwas [19,28] focused on the GB-supported turbomachinery. Limited manufacturing accuracy, the lack of available high-performance materials as well as computational capabilities before 1970 restricted the market penetration of GBs, as suggested by Pan and Sternlicht [29]. Yet, the progress in computer power and manufacturing technology over the past decades has promoted growing interests in the study of GBs as implied by the rising number of publications over the past years on GB in Fig. 2.

As a consequence of the renewed and rising interest, additional reviews have been published since the end of the last century. In their reviews, Agrawal [4] and Walton and Heshmat [26] focused on

---

† Note: This statistical result is merely a rough estimation of the trend of the publications related to gas bearings based on the database search algorithms provided by Web of Science. The numbers of related publications might deviate slightly from the actual number.
compliant GB type limited to structures with journal diameters larger than 20 mm used in oil-free air cycle machines [4] and cryogenic devices. In 2009, Spakovszky [20] gave an overview on the fundamentals, theoretical and experimental concerns of MEMS-based aerodynamic GBs, with bearing length to diameter ratios smaller than 0.1 and Reynold number smaller than 500.

1.2 Classification of gas-lubricated bearings
Gas lubricated bearings can be classified both according to their pressurization mechanism and to the nature of the bearing surface. Depending on the mechanism of pressure generation within the fluid film, GB can be categorized into four distinct classes:

- Aerodynamic gas bearings (HDGB) [28,30,31] generate load capacity through the relative motion between the rotor and bearing sleeves. The viscous shear within the fluid film drags the lubricant into an aerodynamic wedge, which then generates a pressure gradient. The pressure gradient develops load capacity, stiffness, and damping. An increase in rotor speed generally improves the load capacity [8].

- In aerostatic gas bearings (HSGB) externally pressurized gas is fed directly into the fluid film either through orifices [32,33] or porous materials [34–36]. The external pressure source makes the load capacity insensitive to the rotational speed to a large extent, and the bearings offer load capacities and stiffness even at zero rotor speed. Generally, the load capacity is enhanced with the rotor speed due to the additional pressure generated by the viscous shear effect.

- The external pressurization combined with the self-acting principle in hybrid gas-bearings may postpone the self-excited whirl instability in addition to eliminating the dry friction during the start-stop phases [37–40].

- Squeeze film gas bearings (SFGB) [41–43] generate pressure via squeezing-film action, i.e., a lateral motion between the rotor and the bearing surface. Therefore, they are independent of external gas supply or of the presence of an aerodynamic wedge associated with HDGBs. The squeezing action can be imposed by an external motion exciter, e.g., piezoelectric or electromagnetic devices [44]. SFGBs also provide viscous damping, which makes them very interesting for MEMS or micro flexible rotors [45,46]. According to Blech[47], Ausman [48], and Mohite et. al. [46], the excitation frequencies and the gap size determine whether an SFGB is offering rather a high stiffness or a high damping ratio.

In addition to the pressure generation mechanism GBs can be further categorized into rigid geometry and compliant bearings. Fig. 3 represents an overview of typical types of GBs in each category.
• **Rigid Gas Bearings.** Rigid aerostatic bearings can attain high stiffness and load capacities and stability thresholds if designed properly [8]. A well-known disadvantage for rigid aerostatic bearings is that they are prone to pneumatic hammer, which is caused by a sudden loss of damping and can limit the working range [45, 46]. Rigid aerodynamic gas bearings yield limited load capacity [51] and low stability thresholds [52] [53]. Numerous authors have investigated the issue of the self-excited instability of gas lubricated bearings, which is a consequence of both the low damping and the cross-coupled behavior [54,52,10,55]. Significant efforts have been made to enhance the lubrication performance and to postpone the instability onset. The results of these efforts target the geometry of the lubricant surfaces, which could break the symmetry of pressure and therefore reduce the cross-coupled behavior or increasing the damping coefficients. This can be achieved by texturing the surfaces in contact with the gas lubricant, such as Rayleigh steps [56,57] pockets [58] and by applying series of grooves either on the rotor or on the bearing bushing. Grooved bearings have proven to be the most promising solution in terms of stability among self-acting gas bearings [59]. Malanoski experimentally showed that the herringbone-grooved journal bearings (HGJB) were less likely to suffer from the whirl instability than other types of dynamic journal bearings [60].

• **Compliant-Surface GBs.** Compliant surface gas bearings use compliant structures to overcome the drawbacks of rigid bearings. They consist either of flexible foils or of divided bearing sleeves (into separate pads such as tilting-pad GBs) [39,61,62], which deform and tilt in response to the local pressure field within the fluid film. The most significant advantages of compliant GBs are that they generate high load capacity and are tolerant to misalignment, manufacturing deviation, and thermal gradients. Foil bearings, a particular class of compliant GBs, have been extensively studied by Heshmat, San Andres, Kim, et al. [17,63–69]. They can be classified according to the nature of the compliant structure upon which rests the top foil. Classical foil bearings use corrugated bump foils or use interleaved foil. Alternatives to the corrugated bump foils such as metal mesh or
viscoelastic material have been investigated as well [12,67]. Foil GBs [11, 52] generally require special coatings on the surfaces of foils and rotors to minimize wear at the start and stop. High standards are imposed to the tolerances on the foil thickness as well as on the forming of bumps to avoid effects of manufacturing deviation on the bearing performance [70]. Compliant GBs are often encountered in air cycle compressors for aircraft environmental control, turbocompressor and turboexpanders in cryogenic industry, turbochargers and small scale gas turbine engines and for air management of fuel cells [71].

1.3 Scope of the Review
In general, gas lubricated bearings require very small fluid film clearances in order to overcome the low stability thresholds and to generate sufficient load capacity. These tight clearances, approximately 3 orders of magnitude lower than the bearing diameter or length, require particular attention in the manufacturing and in the material selection process. The recent evolution with regards to both computational and manufacturing techniques have led to a rise of the interests in gas bearings, as implied by the rising number of articles published over the last years exhibited in Fig. 2. Among those GBs investigated in the literature, both the hybrid and the compliant GBs tend to achieve higher load capacities compared to other topologies. Nonetheless, grooved gas bearings offer an exciting alternative for small-scale turbomachinery due to their compactness and simplicity in manufacturing and operation as well as in their high stability thresholds. As a consequence, there is a growing interest in grooved dynamic gas bearings (GDGBs), especially in oil-free small-scale turbomachinery.

In view of this, the objective of the paper is to offer an overview of the state-of-the-art of this promising technology, providing insights into the understanding and modeling of the underlying mechanisms governing these bearings and offering design guidelines.

2 The Fundamentals of Grooved Dynamic Gas Bearings

2.1 General Background
GDGBs belong to the rigid bearing category and generally feature a grooved surface pattern that forms an angle with the principal tangential surface speed. These angled grooves act as a viscous pump and drag gas from the bearing edge into the fluid film domain. Figure 4 illustrates two typical GDGBs utilized as journal bearings and thrust bearings. The design parameters indicated in Figure 4 are summarized in Table 1.

| Table 1 Design parameters for grooved dynamic gas bearings given in Fig. 4 |
|-----------------|-----------------|-----------------|
| Nominal clearance | $h_r$          |                 |
| Bearing size     | $L/D$           | $R_o/R_I$       |
| Groove depth ratio | $h_g/h_r$      |                 |
| Groove region length ratio | $\gamma = (L - L_{land})/L$ | $\gamma = (R_o - R_G)/(R - R_G)$ |
Groove width ratio \[ \alpha = \frac{a}{a + b} \]

Groove angle \[ \beta \]

Grooved journal bearings are generally composed of two countervacting pumping groove patterns (herringbone grooves) [72], which increase the pressure within the fluid film through the rotation of the shaft. This pressure increase has been shown to delay the onset of self-excited instability. The grooves contribute to controlling the attitude angle of the self-acting bearing which contributes to the improvement of dynamic stability [73]. An example of a herringbone groove pattern applied in journal GBs is shown in Fig. 4(a). Herringbone GBs have found their applications in gyroscopes [8], automotive turbochargers [74] and in turbocompressor for refrigeration [72]. Groove patterns also have been applied to thrust GBs, with inward, outward and combined effect configurations [75,76]. Spiral grooved thrust GBs offer the highest load capacity among all no-compliant aerodynamic GBs as suggested by Wong et al. [77]. An example of inward grooves is shown in Fig. 4(b). The clockwise rotation of the grooved member drags gas from the outer diameter inwards, which results in a net pressure rise and thus a net load capacity.

The main challenges of grooved dynamic bearings are the stringent manufacturing tolerances and the required surface roughness [78,79] that go hand in hand with the tight clearances and the shallow groove depths. From an engineering perspective, an associated drawback is the thermal distortion due to thermal dissipation, which could result in a seizure or a significant reduction of the thrust bearing load capacity.

From a modeling perspective, grooved dynamic gas bearings are challenging since the grooves introduce periodic geometrical discontinuities of the fluid film. Several approaches have been introduced in order to overcome these modeling difficulties. Fig. 5 summarizes the different methodologies applied to model grooved bearings. Starting from the Navier-Stokes equations several sets of assumptions allow simplifying the description of the evolution of the fluid film pressure in gas lubricated bearings. In general, there are two branches in the theoretical analysis of grooved GBs. The first branch is based on the narrow groove theory (NGT) [73,80], which assumes an infinite number of grooves. This assumption leads to a modified Reynolds equation for the resulting “smooth pressure”. The second branch models groove and ridge regions directly
by the corresponding Reynolds equation [81,82], which is then solved with numerical discretization methods. Fig. 5 gives a detailed overview of these methodologies.

![Diagram showing the modeling tree of grooved GBs](image)

Figure 5: Modeling tree of grooved GBs

### 2.2 Early Form of Reynolds-Equation-Based Narrow Groove Theory

Starting from the Navier-Stokes equations, Reynolds presented the classic lubrication theory for incompressible lubricants in the late 19th century [83]. The Reynolds equation was extended to compressible lubricants by Harrison in 1913 [84]. His theory assumes (1) Newtonian fluid, (2) domination of viscous forces over inertial and body forces, (3) no-slip boundaries, (4) constant density, viscosity and pressure across the fluid film and (5) a film thickness much smaller than the other dimensions, thus corresponding to the set of assumptions A1 [65]. In order to capture the periodical nature of the groove patterns of spiral groove thrust bearings (SGTBs), Whipple [59,85] introduced the Narrow Groove Theory (NGT) in an attempt to avoid the numerical complexity of grooved patterns on a thrust-GB. The NGT theory refers to a groove pattern with an infinite number of grooves. This assumption leads to infinitesimal pressure fluctuations across a groove-ridge pair in view of smoothening the overall pressure distribution within the fluid film. This elegant approach yields a modified Reynolds equation, which takes the pumping action of the grooves into account yet results in a continuous and smooth pressure evolution. The set of assumptions A2 (Fig. 5) that lead to the modified Reynolds equation for grooved patterns are given as follows:

1. The total pressure variation over a ridge-groove pair is equal to the sum of the change in pressure over the ridge and over the groove.
2. The derivative of pressure over each of the two regions (ridge and groove) is constant. This implies the local incompressibility of the fluid for an infinite number of grooves.
3. The pressure gradient parallel to the ridge-groove interface is constant on both sides of the interface in order to avoid pressure discontinuity across it.
4. The mass flow normal to the interface respects the continuity principle across it.
5. The pressure is periodic along the circumferential direction over a pair of groove-ridge pair.
6. The groove curvature is negligible.

Although these assumptions seem very far-reaching, this theory can provide a “fairly rough but useful estimate of the operating characteristics and the optimal proportions of such bearings” as commented by Boeker et al. [86] in his survey of Whipple’s theory in 1958. Whitley and Williams [87] applied the same modeling approach as Whipple, but included effects due to surface distortion and plate vibration to thrust GB. In 1967, Whitley[88] extended Whipple’s work by introducing the end effects, the influence of the groove number and curvature. It is important to note that these early studies treat the film as isothermal and incompressible for practical purpose. Moreover, only straight grooves with the constant rectangular cross section extending outwardly either in an inclined pattern [59] or along a circular arc [73,89] are considered, which were also referred to as parallel groove patterns [90].

2.3 Muijderman’s NGT
Muijderman [90,91] assessed the applicability of Whipple’s incompressible NGT theory in depth combining the findings of other researchers, e.g., Ford, Boeker and Wordsworth et al. [30,86,92]. In order to overcome the difficulties in matching Whipple’s theory and the experiment conducted by Wordsworth [92], Muijderman modified Whipple’s NGT method by considering grooves along a spiral curvature including end effects for thrust GBs. Moreover, to obtain a more accurate pressure build-up along the logarithmic spiral grooves, he applied a conformal mapping method to map a spiral groove pattern into a parallel one to simplify the numerical calculations. It is important to point out that the lubricant studied in Ref. [90] is incompressible, which is less time-consuming to obtain solutions with acceptable accuracy.

2.4 Vohr and Chow’s NGT
At the same time as Muijderman [90,91], Vohr and Pan published an NGT-based model [93] for grooved journal GBs in 1964. Their model assumes an infinite number of grooves and is formulated for parameterized groove shapes (width, angle, depth, and length). In 1965 Vohr and Pan’s theory was extended by Malanoski and Pan [89] to study the static and dynamic performance of spiral-grooved thrust GBs. In comparison with experiments, their theory showed a deviation on the load capacity within an error band of 3%, thus validating their NGT-based model.

Vohr and Chow used the NGT to formulate a model for herringbone-grooved journal bearings (HGJBs) in 1965 [73]. Their model uses the assumption sets A1 and A2 to obtain a modified Reynolds equation that allows predicting the evolution of the pressure within the fluid film of a journal bearing with herringbone shaped grooves.

Moreover, two additional assumptions (A3 in Fig. 5) are usually adopted to further simplify both the Reynolds equation and its modified form by providing a relationship between the density and the pressure:

1. Isothermal fluid
2. Ideal gas

These assumptions allow transforming the original Reynolds equation from an expression in pressure and density into an equation expressed in pressure only.

In order to predict the stiffness and damping properties of herringbone grooved journal bearings, Vohr and Chow [73] applied the perturbation analysis to linearize the modified Reynolds equation around a centered position. The limitation to a centered rotor position allows to introduce a separation of variables (circumferential and axial directions), thus transforming the modified 2D non-linear Reynolds equation into a 1D linear differential equation that can be integrated very easily and rapidly. A particular feature of herringbone grooved journal bearings is that they can offer stable operation even at concentric rotor positions. This has been shown to be theoretically impossible for plain journal bearings.

In order to solve the modified Reynolds-equation obtained by applying the NGT, finite difference (FD), finite element (FE) and finite volume FV based solving schemes have been applied successfully [94–96] yielding a satisfactory accuracy. Of course, the solutions are valid only as long as the number of grooves is sufficiently large.

2.5 The Experimental Validation of the Classical NGT

The NGT was experimentally validated first in its incompressible form by Hirs [97], who ran a floating HGJB with a load applied on the non-rotating bushing. Malanoski [60] ran a 38 mm shaft supported by shallow-grooved HGJBs without external load up to 60 kRPM with no sign of half-whirl instability. Measurements of radial stiffness were in good agreement with the 1D centered NGT, with most of the measurement points within 10% deviation from the concentric NGT predictions.

Castelli and Vohr [112] renounced to the simplification of the separation of variables introduced by Vohr and Chow, thus obtaining bearing properties valid also at large eccentricities. He solved the NGT equation (in circumferential and axial directions) at eccentric rotor positions and compared his computations with experimental load-displacement curves by Malanoski at different speeds, with a good agreement, even at eccentricity ratios above 0.5.

Cunningham and Fleming [98] ran a 38mm HGJB-supported rotor with an externally-pressurized load shoe acting on the rotating shaft to investigate the load capacity. Although the theory from Vohr and Cow [73] and experiments matched well at low eccentricity ratio, the total static stiffness deviated significantly from linearity at high eccentricity ratio. Stanev et al. [99] experimentally validated a design of hybrid HGJB theoretically investigated by Fleming in 1970, finding a relative deviation of less than 25% for the load capacity. More recently, Ognjanović [2] experimentally confronted the NGT with the experimental results obtained from a floating 60mm HGJB supported on rolling elements bearings, suggesting deviations up to 15% between the centered 1D NGT model and experimental data. The deviation from the theory was observed to increase when the number of grooves was decreased, with a further depreciation of 10% of the load capacity when the number of grooves was lowered from 36 to 12.
Although the literature is insufficient in validating cases for the NGT, the theory is supported by numerous successful implementations of NGT-designed HGJBs and SGTBs for high-speed rotors of small-scale turbomachinery [1,100,101] and laser scanners [102].

### 2.6 Direct discretization of the Reynolds Equation for Grooved Bearings

Instead of using the NGT to simplify the analysis of grooved bearings researchers also looked into the discretization of the fluid film while directly solving the Reynolds equation. Castelli and Pirvics [27] overviewed the discretization methods of Reynolds equation and the associated solution techniques for both time-dependent and time-independent considerations. In 1997, van der Stegen [103] applied finite difference methods to compute the load capacity for gas-lubricated HGJBs operated at large compressibility numbers. When using the finite difference method (FDM) to solve the non-linear Reynolds equation, the discontinuity of the local film thickness due to the presence of grooves are suggested to yield numerical challenges. One well-recognized challenge is associated with numerical oscillations of the solution, which can be attenuated only with very fine meshes [103]. Similar problems may also appear in Galerkin weighted residual approaches, especially at high compressibility numbers where convective flow components become significant [104].

Appropriate weight functions associated with the finite element approach are suggested to avoid the oscillations exhibited in FDM. Reddi and Chu [105] initiated the FEM-based analysis of compressible fluids within a grooved gas bearing applied for video and audio tapes in 1970. In an analysis of HGJBs with a small number of grooves, Bonneau and Absi [82] adopted self-adaptive upwind schemes, which used Petrov-Galerkin weighted functions varying with surface velocity, mesh quality, viscosity, pressure, and thickness. They applied a Newton-Raphson approach to deal with the nonlinear nature in the discretized equations for grooved gas bearings. Bonneau and Absi’s analysis successfully eliminated the barrier imposed by the NGT which is limited to a sufficiently high number of grooves. In order to overcome the numerical diffusion related to these formulations, Faria and San Andrés [81,104] introduced high-order shape functions for the analysis of high-speed thin gas films.

Besides the finite difference and the finite element formulations, finite volume methods (FVM) have been employed as well in the analysis of discontinuous gas films [106,107]. Arghir [107] discretized the Reynolds equation with an unstructured grid in the irregular domains to obtain similar geometric flexibility as for the FEM-based methodologies.

Both the FEM and FVM have demonstrated their capability in dealing with discontinuities of the fluid film thickness more efficiently than the FDM [108–111]. Miller and Green [109] compared FEM with FVM for spiral-grooved gas face seals, and found that FVM needs significantly less time (approximate 50% lower for their application) than FEM for their analysis which considers rotor runout and static stator misalignment.

In conclusion, for grooved bearings, the mesh needs to coincide with the fluid film discontinuities, which can be challenging for the discretization [27,107]. Moreover, models based on the discretization of the fluid film are suggested to be inefficient and inaccurate if the grooved member is rotating [82,104]. Therefore, despite the flexibility of direct fluid film discretization, it would
be very time consuming to embed such methods into integrated design optimizations because their accuracy is greatly dependent on the mesh size and the computational time [1].

3 Limitations and Enhancement of the NGT

The assumption set A2 in Fig. 5 associated with the classic NGT models limit its use and validity. Therefore, numerous efforts were made in order to extend its range by assessing additional effects [60,112–117] such as the influence of (1) the local compressibility across a groove-ridge pair, (2) the end effect at the bearing edges, (3) journals operated at large eccentricities, (4) the effect of a finite number of grooves and (5) the effect of leakage flow. The investigations and the results of these additional effects are discussed in the sections below.

3.1 Local Compressibility Effect

When deriving the classical narrow groove theory (NGT) one of the assumptions is a linear pressure profile (constant pressure gradient) across a groove-ridge pair, which imposes a locally incompressible flow. This assumption is in conflict with the nature of a compressible gaseous lubricant. Based on the classical NGT Vohr and Chow [73] to conclude that the dimensionless load capacity for HGJB increased almost linearly as the $\Lambda$ grows over 30. However, in 1969, Constantinescu and Castelli [118] modified Vohr and Chow’s NGT equations with locally compressible relationships of the pressure and density. Their modified NGT analysis yields an asymptotic pattern of the non-dimensional load as a function of the compressibility number $\Lambda$. In the same work, Constantinescu and Castelli [118] also investigated the conditions under which the local incompressibility could be used without incurring significant errors. For their analysis, they defined a local compressibility number as follows:

$$\Lambda_{\text{local}} = \frac{\mu V_{\text{local}} L_{\text{local, char}}}{h_{\text{local}, g} P_a} = \frac{\mu (V \sin \beta) \left( \frac{2\pi R}{N_g \sin \beta \alpha} \right)}{h_{\text{char}}^2 P_a} = \frac{\mu L_{\text{char}}}{h_{\text{local}, g}^2 N_g P_a} = \frac{h_{\text{local}, g}^2 \alpha}{h_{\text{char}, g}^2 N_g} \frac{1}{\sin^2 \beta \Lambda}$$

(1)

where $h$ and $h_g$ correspond to the ridge and groove film thickness, respectively, $\alpha$ the groove width ratio, and $\beta$ the groove angle as indicated in Table 1 and Figure 4. Their analysis suggests that for an infinitely long bearing with a groove depth ratio of 2-4, the local compressibility effects can be neglected as long as the local compressibility number is below the value of $\Lambda_{\text{local}} = [0.1 – 0.15]$. Through Eq. (1) it can be deduced that there is a number of grooves above which the assumption of local incompressibility yields only minimal errors.

Besides the load capacity, the local compressibility also has an impact on other factors as well. In 1971 Emel’yanov and Emel’yanova [116] derived fundamental equations that considered both the global and local compressibility as well as the rarefaction effect for logarithmic groove patterns of SGTBs. Their results suggested that the local incompressibility effect overestimated the pressure rise within the fluid film due to both the compressibility and the rarefaction effects. Comparison with experimental data showed that their model performed better than the classical NGT model such as applied by Hsing and Malanoski [119].
3.2 Edge Effect (End Effect)
The edge effect (end effect) refers to the phenomenon that the linear pressure rise and drop across a groove-ridge pair in the classic NGT analysis, leads to a sub-ambient pressure at the bearing edge. This effect is in conflict with the ambient-pressure boundary condition. Further, this effect is more and more pronounced if the number of grooves decreases, i.e. when the width of a groove-ridge pair is no longer narrower in comparison to the bearing length [126]. Muijderman [90,91], Elrod [114], Pan [120] and Gupta [117] proposed corrections to account for this contradiction. According to Muijderman the edge effect leads to an effective radius reduction of a thrust bearing, which depends on the operating conditions, the number of grooves and their geometry. A similar correction of the edge effect is suggested by Elrod [114] and by Gupta [117] for SGTB. All these end effect corrections lead to a reduction of the load capacity compared to the uncorrected theory. Later, Pan [121] proposed a compressible NGT with a two-scale analysis, one at the groove scale and the other at the bearing scale, getting rid of the edge and compressibility problems of the classical NGT at the price, however, of an increased mathematical complexity.

3.3 Eccentricity Effect
The introduction of infinitesimal perturbations around a static eccentricity allows to obtain linearized equations for the zeroth and the first order pressures, ultimately yielding static load capacity as well as the stiffness and damping properties. In addition, in some cases, a separation of variables (theta and z) is suggested for journal bearings operated in concentric positions, which yields only the dynamic bearing coefficients [60,98,122,123]. A comparison with the FEM analysis by Bonneau and Absi [82] suggests that the separation-of-variable approaches remain valid as long as the static eccentricity remains below 20%. In order to overcome the limitation of lightly loaded journal bearings direct numerical methods, such as the finite difference [112,118] and the finite element [81,82,104] methods were implemented for solving the modified Reynolds equation resulting from the NGT. Using the finite difference method, Castelli and Vohr [112] identified that the optimum groove depth ratios are dependent of the eccentricity, which seems intuitively correct since the effective film thickness varies with eccentricity.

3.4 Number of Grooves
When applying the NGT method, it is essential to keep the pressure fluctuations across a groove-ridge pair small. Fleming and Hamrock [122] suggested an analytical relationship limiting the ratio of compressibility number and the groove number to a threshold value. The limiting values correspond to the condition in which local compressibility effects should be considered as discussed by Constantinescu and Castelli [118]. It is noted that the effect of the number of grooves on the bearing performance can only be assessed by using a discretized solving scheme of the Reynolds equation [1,82]. In order to expand the NGT to a finite number of grooves, Pan and San Andrés [124] applied a two-scale formulation, which adopted the NGT for the macro-level analysis and the rarefied and a locally compressible gas film lubrication in the micro level (across a groove-ridge pair).
4 Further Enhancements of Grooved Gas Bearing Models

The applicability of the classical Reynolds equation has its limitations due to the thin-film assumptions made for its derivation from the Navier-Stokes equations (A1 in Fig. 5). Limitations associated with these assumptions include the effects of rarefaction [113,125–127], the transition from laminar to turbulent flows [115,128] as well as the neglection of inertial effects [129–131]. In addition, the assumption of an ideal-gas and isothermal fluid film (A3) allows obtaining a simplified equation in pressure only. However, this simplification is incompatible with the considerations of real gas effects [132,133] and thermal effects within the fluid film [29,134,135]. The following discusses these limitations in detail.

4.1 Rarefaction Effects and Slip-Flow Boundaries

Gas rarefaction effect may become critical in gas lubricated bearings in particular at low pressures and for gases with small molecules. The effect of rarefaction has been investigated thoroughly for slider bearings in hard disk drives. The effect scales with the mean free molecular path. If the ratio between the clearance and the molecular mean free path drops below a threshold value, rarefaction effects start manifesting [1,119,136,137] resulting in a departure from the zero-velocity flow condition at the boundaries. The Knudsen number $Kn$, defined as the ratio between the molecular mean free path and the fluid film thickness, indicates the severity of rarefaction effect. At $Kn < 0.01$, the fluid is considered as a continuum and the classical equations for compressible fluids can be applied. In general, for $Kn > 0.01$ the gas film is regarded as rarefied [138–140].

The rarefaction effect can be considered either by correcting the classical continuous fluid equations by including a slip flow boundary conditions at the wall surfaces [138,141] or by deriving the governing equation from the Boltzmann equation (accounting for molecular behavior) [137,142]. Hsing and Malanoski [119] found deviations up to 20% between stiffness obtained from models with and without the rarefaction effect at $Kn=0.1$ for a SGTB, thus suggesting a significant impact on the bearing performance. Shukla et al. [141] suggested that the gas-bearing load capacity decreased with increasing $Kn$. To account for gas rarefaction effects, Fukui and Kaneko [137] provided a database of the slip flow correction factors and Pan [127] deduced an empirical correlation for the slip correction factor based on their work.

Fukui and Kaneko [140] suggested that Burgdorfer’s model [138] with the 1st-order velocity slip boundary condition overestimated the load capacity while Hsia’s model [136] with the 2nd-order slip boundary condition underestimated it. Hwang et al. [139] introduced an empirical correlation based on the Boltzmann model to correct the Poiseuille mass flow, which allows for a more rapid computation of the slip-flow compared to the method by Fukui and Kaneko [137]. Based on this result Schiffmann et al. [1,133] applied Hwang’s model for the integrated optimization of gas bearing supported rotor systems. The difference of the load capacity between models with and without rarefaction effects decreases with increasing compressibility numbers since the Couette flow is independent of the degree of rarefaction and becomes dominant compared with Poiseuille flow [140]. Investigations by Schiffmann and Favrat [133] suggest a significant effect of the rarefaction on the dynamic properties of herringbone grooved journal bearings.
4.2 Turbulence Effects
Laminar conditions in a thin fluid film is usually valid for Reynolds numbers up to 2000 with the local clearance as the characteristic length. The Reynolds equation can be corrected to account for the turbulence using a correction factor in the two Poiseuille flow components (in circumferential and axial directions) to account for the bulk velocities in the film [143–145]. The tangential and perpendicular directions to the bearing surface speed are usually corrected differently [143]. Several models are provided to compute these correction factors [143–145] as a function of the local Reynolds number, although they are all derived from incompressible flows. Some theories, such as Hirs’ study [145], distinguish between Poiseuille and Couette-dominated flows to account for the turbulence correction factors. In all cases, the effect of correction factors increase with the Reynolds numbers and, therefore, reduce the influence of the diffusion in the Reynolds equation. Hsing [120] derived the NGT with correction factors accounting for the turbulence, distinguishing the correction factors for grooves and ridges. Constantinescu [146] proposed turbulence parameters based on a theory valid for non-smooth (i.e., grooved) surfaces. Therefore, the combination of the two works can provide a tool for the performance prediction of grooves bearings operating in the turbulent regime, as employed by Guenat and Schiffmann [147] for the simulation of HGJB in high-density refrigerants with real-gas effects.

4.3 Inertia Effect
The groove patterns introduce a discontinuity in the film thickness, which may introduce a variation of flow field, also known as concentrated inertia effect [108]. A significant deviation of the pressure field has been suggested to exist between the Navier-Stokes and Reynolds solution for gas lubricants under certain circumstances as explained in by Qiu et al. [106] and Arghir et al. [130]. Through the study of gas-lubricated textured slider bearings, Qiu et al. [106] suggest that deviation is due to the fact that velocity gradients in the flow direction and radial direction become dominant at an increased aspect ratio (larger relative dimple depth). The influence of relative texture depth on the validity of Reynolds equation has also been confirmed by Feldman et al. [130] for gas-lubricated parallel surfaces. Inertia effects around groove-like film discontinuities should also be taken into account as demonstrated by Guardino et al. [148] and Jarray et al. [149]. Through the investigation of the surface roughness effect on air-riding seals, Guardino et al. [148] suggested that the difference between the Reynolds equation and full Navier-Stokes equation is less significant for compressible fluids than incompressible conditions. Nonetheless, with the growing interest of hermetic and compact turbomachinery working with dense fluids, such as supercritical CO2 [150,151] whose density and viscosity are comparable to a liquid, the inertia effects will manifest for such applications, particularly for thrust bearings at high-speed operations.

4.4 Real Gas Effect
Usually, the gas lubricant is treated as perfect/ideal gas in many applications [7]. Nonetheless, in applications where the lubricant is operating close to the saturation line such as for refrigeration compressors [147] or fuel-cell air supplies operated with humid air [23] this assumption is far from valid. The density terms in the Reynolds equation become a complex, non-linear function of both pressure and temperature and is, therefore, difficult to be simplified. Fairuz and Jahn [152] found that the real-gas effect is particularly significant for the supercritical-CO2-lubricated dry gas seals (with groove textures) running close to the critical point. Schiffmann and Favrat suggested [133]
that both the synchronous direct stiffness and the load capacity decrease if the lubricant is operated increasingly close to the saturation line. This trend is enhanced at large compressibility numbers. Since the damping coefficients are affected by real gas effects as well, the stability of the bearings is modified under real gas effects. In recent work by Guenat and Schiffmann [147], the real gas effect is studied for three types of gas bearing geometries, i.e., Rayleigh-step-slider, plain and herringbone-grooved journal gas bearings. Their numerical results based on modified NGT suggest a pronounced degeneration of the bearing performance, reaching 50% for the load capacity and almost 100% for the dynamic stability at certain configurations. However, in some cases, the dynamical stability could be increased by several orders of magnitude. Furthermore, their study also indicated the influence of the real gas effect on the applicability of the quasi-incompressible NGT theory due to the enhanced compressibility in the real-gas situation for the HGJBs. Another effect associated with non-ideal behavior of the gas lubricant is the partial condensation of water in saturated humid air, as investigated by Guenat and Schiffmann [96]. Although humid air effect has a negligible effect on the load capacity of a journal bearing, the dynamic stability can be reduced by 25% in realistic atmospheric conditions, which suggests that a stability margin should be considered in applications with warm and humid ambient conditions.

4.5 Thermal Effects
Hughes and Osterle [134] suggested the applicability of adiabatic and isothermal operating conditions within the radial clearance of journal bearings. Their analysis shows that the fluid film behaves isothermally due to the small fluid film volume compared to the wetted surface. As a consequence, an isothermal approximation is usually valid and applied for gas film lubrication [59,153,154]. From the modeling point of view, the isothermal assumption allows not only the spatial-invariant gas viscosity but also the density to be expressed in terms of the pressure only. Castelli and Vohr [112] validated this assumption for spiral groove journal bearings by comparing their numerical results with experimental data obtained by Malanoski [60]. However, Kao [135] pointed out that the error due to such an assumption could influence the prediction of the load capacity because the heat dissipation is neglected through the surfaces.

In spite of the validity of isothermal boundaries in most gas-lubricated bearings, particular caution is required at high-temperature gradients that could lead to the structural distortion of the fluid film surfaces. Such a distortion could lead to a significant drop in load capacity and even to premature failure. Wilson [155] suggested that a non-isothermal model achieved a better accuracy in the prediction of the bearing performance for a wide range of bearing temperature, up to 1310.9 K. Moreover, utilizing non-isothermal boundary conditions is suggested to help increase the prediction accuracy of load capacity. Pan and Sternlicht [156] theoretically studied the influence of thermal distortion on the SGTB load capacity, which showed that the load capacity loss for typical metallic alloys could be non-negligible at high loads.

5 Effects of the Groove Design on the Bearing Performance
The purpose of introducing grooved patterns on the gas bearing surfaces is to improve its static and dynamic performance. Vohr and Chow [73] found that the attitude angle of a grooved bearing can be reduced significantly compared to a plain bearing and that a grooved gas bearing may not
suffer from the half-frequency whirl instability that is associated with plain bearings under unloaded conditions. As a consequence, the grooved bearing performance is influenced by the groove geometry such as size, profile shape, number and length. The following section summarizes how the groove designs influence the bearing performance.

5.1 Groove Shape
A large number of research groups [59,98,112,122,123,157] have oriented their research to maximize both the stability and the load capacity by optimizing the groove designs. Initially Chow and Vohr [115] concluded that the advantages in adopting helical grooves for journal bearings with incompressible lubricant include (1) the greater stability under lightly loaded conditions, (2) the self-pumping effect of grooves eliminating the need of a pressurized supply, and (3) the fact that the performance can be enhanced even further by an external supply of pressurized lubricant.

Hashimoto and Ochiai [158] tested four sets of thrust gas bearings, i.e., (1) stepped (2) pocketed (3) spiral grooved and (4) herringbone grooved, with the same depth of grooves/steps/pockets. The results suggested that the spiral groove thrust bearing (SGTB) presents the smallest friction torque while maintaining the highest stiffness coefficients for high speed operating conditions. Later, Hashimoto and Namba [76] suggested that a spiral groove pattern with bends in the vicinity of the outer circumference of the thrust bearing was preferred if (1) the minimum ratio of friction torque to the dynamic stiffness or (2) the maximum product of the flying height and the dynamic stiffness is sought. In 2011, the concept of varying depth along the groove’s length was proposed by Fesanghary and Khonsari [159] to enhance the load capacity of a micro fluid film between two relatively sliding plates. The results suggest that an optimized spline-shaped groove profile can improve the load capacity of a thrust bearing by 45% compared to the original design, which implies that curved grooves are preferred over a tapered or flat groove depth.

Originally, the grooves of a herringbone-grooved journal bearing (HGJB) are of helical shape whereas the grooves on a thrust bearing follow a logarithmic spiral to keep the angle between the groove and the radii constant. As a consequence, for classical grooved journal/thrust bearings the groove geometry is determined by the groove angle, the relative groove width, the groove depth, and the relative groove length. For the determination of the best combination of the groove shape Fleming and Hamrock [122] optimized the HGJB groove geometry for maximum stability and load capacity in 1974 and summarized the resulting groove geometries for various compressibility number and length to diameter ratios.

The concept of varying groove geometry is also effective to improve the bearing performance for herringbone-grooved journal GB. Schiffmann [160] applied polynomial functions to describe the axial evolution of groove width, depth, and the angle, and optimized the axial groove shape evolution based on the modified NGT theory (limited to small journal orbits) for maximum stability and minimum power losses. The theoretical comparison of the enhanced groove patterns with the classic herringbone grooves suggests that the clearance can be increased by up to 1.8 times for the same stability requirement, which allows to reduce the windage loss and to decrease the manufacturing tolerance.
5.2 Effect of Pumping Direction on SGTBs

The grooves in dynamic thrust bearings can be shaped so that the pumping action is towards the ID (inward), the OD (outwards), or a combination of both. In general, inward pumping grooves are applied in SGTBs to achieve a higher stiffness and load capacity, as a result of transporting the lubricant inward towards the smaller peripheral areas close to the inner radius. Malanoski and Pan [89] found that the inward pumping configuration yields a higher load capacity than the other two configurations. In addition, the inward pumping configuration provides stronger sealing functions because of its higher dam effect. Shahin et al. [161] compared the inward- and outward-pumping designs for spiral-grooved dry gas seals. They found that the outward-pumping action could induce negative stiffness and a tendency for instability.

5.3 Rotation of the Smooth or Grooved Surface

The groove patterns can be applied to both the rotating and the stationary members. Work by Malanoski and Pan [89] on spiral-grooved thrust GBs and by Fleming and Hamrock [122] on herringbone-grooved journal GBs suggested that the stability increased significantly if the grooved member rotates, in particular at high compressibility numbers. Pan and Kim [162] plotted a stability map for a conical herringbone-groove gas bearing, which shows a much higher critical mass, even unbounded at a compressibility number of 25 when the grooved member rotates. Similar findings are obtained by Fleming and Hamrock [122] who suggested significantly higher critical masses for grooved member rotation compared to smooth member rotation for HGJBs.

Bonneau and Absi [82] suggested that it is important to consider the eccentricity when comparing the cases with the grooved or with the smooth member rotating. Nonetheless, it is worthy to recall that the numerical computation tends to suffer from numerical oscillations when applying direct discretization approaches to the case with a rotating grooved member [81,154].

5.4 Number of Grooves

Zirkelback and San Andrés [154,163] applied a FE based model to SGTB and spiral groove face seals (SGFS) and found that both the axial stiffness and the mechanical losses grow with the number of grooves. The axial damping, on the other hand, dropped to negative values for groove numbers above 10. Nonetheless, at extremely high excitation frequencies, the influence of the number of grooves on the dynamic coefficients become negligible. For both SGTBs and HGJBs with fixed bearing geometries (except the groove number) and a constant speed, Hashimoto et al. [164] showed that the ratio of the power loss over stiffness decreased with the number of grooves. For journal GBs, Faria [104] puts forward that the direct stiffness reduces with the number of grooves at a low compressibility number ($\Lambda = 15.3$).

Van der Stegen [103] revealed that the number of grooves affected the load capacity differently depending on the range of compressibility number, as summarized in Table 2. It is noted that the load capacity increases with the number of grooves (in the range of 4-16) for relatively high compressibility number. There is a limit where the bearing performance is invariant with the increase in the number of grooves [154]. This is when the NGT theory becomes valid as discussed earlier.

Table 2 The influence of the groove number on the bearing performance
<table>
<thead>
<tr>
<th>Studied Cases</th>
<th>Variation trend with groove number</th>
<th>Groove Number</th>
<th>Compressibility number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zirkelback[154,163] and San Andrés on SGTBs and SGFS</td>
<td>Power loss ↑</td>
<td>[6→18]</td>
<td>3.95</td>
</tr>
<tr>
<td></td>
<td>Seal opening force ↓</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Axial stiffness ↑ Negligible influence at high excitation frequency</td>
<td>[6→48]</td>
<td>160.9 &amp; 321.8</td>
</tr>
<tr>
<td></td>
<td>Axial damping ↓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hashimoto [164] on SGTBs and HGJBs</td>
<td>Power loss/Axial Stiffness ↓</td>
<td>[8→24]</td>
<td>Around 4.35</td>
</tr>
<tr>
<td>Faria [104] on HGJB (stationary groove surface)</td>
<td>Axial stiffness ↓</td>
<td>[4-20]</td>
<td>15.3</td>
</tr>
<tr>
<td></td>
<td>Axial damping ↓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van der Stegen [103] on HGJB</td>
<td>Load capacity↑</td>
<td>[4-16]</td>
<td>≈ [0-1.6]</td>
</tr>
<tr>
<td></td>
<td>↓</td>
<td></td>
<td>≈ [1.6-16]</td>
</tr>
<tr>
<td></td>
<td>↑</td>
<td></td>
<td>≈ [16-160]</td>
</tr>
</tbody>
</table>

* Unless indicated specifically, the results were calculated for cases with the groove member rotating.

5.5 Groove Depths
In thrust bearings a deviation from the optimum groove depth may decrease the pumping effect [163] and depreciate the damping coefficients. The axial stiffness, however, varies non-monotonically with the groove depth and yields a maximum value at a groove depth ratio of around 1.8 for SGTBs [163].

Experimental results obtained by Cunningham et al. [98] on HGJB agreed with the theoretical results predicted by Vohr and Chow [73] suggesting that the load capacity decreases with the groove-to-ridge-clearance ratio at low compressibility numbers. They pointed out that groove-to-ridge-clearance ratios around 2.0 to 2.4 yield the highest load capacity for HGJB. In general, the optimum groove depth is almost constant at compressibility numbers below 10 and gradually rises with increasing compressibility, as shown by van der Stegen [103].

5.6 Groove Regions
The choice of partially- or fully- grooved bearing has a significant influence on the bearing performance both in terms of the stability and load capacity. For a journal bearing, the typical herringbone groove pattern gradually increases the pressure within the fluid film towards the center of the bearing and improves the stability compared to a plain journal bearing. Partially-grooved HGJBs leave a plane land region in the middle, as shown in Fig 3(a). The un-grooved region results in a smaller effective fluid film clearance compared to a fully-grooved design. This is beneficial

---

† The plots were overlapping in Fig. 7.16 of Ref. [103], making the exact trend difficult to be asserted for the load capacity varying with the number of grooves at the compressibility number smaller than 16. However, the results became clear after compressibility number reaching 16.
in terms of bearing load capacity and stiffness. However, the smaller clearance for a partially grooved HGJB leads to a higher power loss than a fully-grooved HGJB as reported by Cunningham et al. [98,123]. Hence, there is a trade-off between competing objectives with regards to the optimal groove design.

For thrust bearings, the most common patterns are inward-pumping spiral grooves. Due to the one-sided pattern, such an inward pumping thrust bearing yields a net flow of lubricant from the OD towards the ID. The ungrooved region increases the flow resistance and therefore raises the pressure within the fluid film while impacting the load capacity of the bearing. Therefore, most SGTBs studied in the literature, either inward-pumping or outward-pumping [89,165], are partially-grooved. The groove patterns can be optimized for texturing other than herringbone and spiral profiles, which provides the extended flexibility for the best combination of the load capacity and the friction coefficient such as the work done by Gropper et al. [166], Cupillard et al. [166,167], and Kango et al. [168].

Two-sided SGTBs, with grooves on both the smooth and rotating surfaces, proposed by A. Yemelyanov and I. Yemelyanov [169] in 1999 was compared to the classic single sided spiral-grooved GB. The results suggest that the two-sided thrust bearing outperforms the classic spiral-grooved GB by 60% regarding the load capacity using a model based on the NGT. However, the study carried out by Lehn and Schweizer [170] suggests that two-sided SGTBs are inferior to the classic one-sided version. Lehn and Schweizer attribute this to the fact that the one-sided SGTB have small ridge clearance, which plays the role of seals and therefore helps increase the local pressure. However, for the binary SGTBs, the nominal clearance for the ridges can be increased by the magnitude of groove depth at locations where part of a ridge on one side is facing the grooves on the other side.

5.7 The Influence of Groove Geometry on Friction Loss
In general, the clearance-to-radius-ratio for a herringbone-grooved journal GB is in the order of $0.5 \times 10^{-3}$ to assure stable operation [98]. The small clearance to diameter ratio inevitably leads to power losses. Cunningham et al. [98] found that the power loss increases slightly with the groove angle and decreases with the groove width ratio. Moreover, Hashimoto et al. [164] reported that there were decreased windage losses with an increasing number of grooves.

The groove pattern itself, applied to the rotor other than the bearing parts, can be explored as a viscous vacuum pump to decrease the windage losses. The concept of viscous pump gas bearing was published initially by James et al. in 1967 [171], proving its feasibility but showing a low compression efficiency. Other examples are found in work by Sato et al. [172] (using SGTBs) and Yoshimoto et al. [102] (employing HGJBs) making use of the pumping functions of grooves to decrease the pressure along the rotor in view of decreasing the windage losses.

6 Critical Assessment of State-of-the-Art
As suggested by this review, the literature is very rich in modeling methodologies for grooved journal and thrust bearings. Several ways of approaching the complex modeling of grooved bearings have been developed and implemented, ranging from the direct discretization of the fluid film and solving the Reynolds equation to applying the narrow groove theory, which assumes an
infinite number of grooves. As usual, the simplification of complex equations yields some limitations but also significant gains in computational time.

So far, the NGT has proven its reliability in design, high-level modeling, and analysis capability. Since the solving of the simplified equations is computationally very fast, NGT based models are particularly beneficial for integrated optimizations of systems working with grooved gas bearings. In addition, physical phenomena such as real-gas, rarefied gas and turbulence effects can be incorporated in the NGT. In our opinion, further theoretical improvement of NGT should also include the consideration of local inertia effects, which may become important in novel applications where the density of the lubricant is very high (heat pump compressor or super-critical Brayton cycles).

The NGT is inappropriate to analyze complex non-conventional bearing texturing such as non-periodic patterns and multiple pockets. Direct numerical methods applied to the Reynolds equation are more promising to explore exotic texturing patterns. However, the NGT is suitable for the modeling of continuous and periodic groove patterns beyond straight or spiral grooves and should be further exploited in this direction.

Since the real-gas effects can be significant, in particular when the ambient conditions approach the critical point, additional research should be carried out with regards to dense gas lubrication. An additional phenomenon, which has not been analyzed yet is the potential occurrence of condensation within a grooved bearing that is operated close to the saturation line. As matter of fact, the isothermal compression that occurs in gas lubricated bearings, in particular in herringbone grooved journal bearings, could easily lead to localized condensation. As the pressure in two phase fluids is governed by the temperature, the occurrence of condensation could influence the bearing performance considerably.

Although a significant number of physical phenomena have been included in the bearing models, the literature lacks systematic experimental validation data to support these theoretical developments. The only indirect validation of these models is the large number of grooved bearing supported rotors, which have been developed and successfully tested. Therefore, research efforts should be oriented toward the completion of the experimental benchmark database for grooved bearings.

Besides the theoretical developments that aim for more accurate models, one of the key limitations of grooved bearings are the tight bearing clearances and groove depth. These make the manufacturing of such bearings very challenging and expensive due to the tight tolerance fields. So far, grooved bearings have only been optimized for stability and/or static load capacity. Yet, it is likely that if these bearings are optimized for manufacturing robustness, i.e. low impact of wide tolerance fields on stability and/or load capacity different optimal geometries are obtained. Shape tolerances are also relevant since they can also affect the bearing performance. As their effect on the bearing performance is unexplored designed often indicate the tightest tolerance fields that the production process can reach. More insights into the effect of manufacturing deviation on grooved bearings would certainly improve their appeal.
7 Integrated Design and Optimization of GDGB Supported Rotors

With modern design tools and the improved knowledge of gas bearings, this technology is becoming more and more accessible. However, the design of gas bearing supported rotors remains a challenging task since such systems are complex and of multi-physical nature. Further, each application has different specifications and design objectives, thus leading to different bearing designs. This makes the generalization of bearing design very challenging.

In light of the discussion above, it becomes clear that the optimal bearing design is strongly dependent on the operational conditions and the characteristics of the supported rotor. Therefore, integrated design and optimization methodologies are strongly recommended in order to achieve the best design of gas bearings [1,157]. Often multi-objective optimizations are performed on standalone GDGB in order to investigate and understand the design tradeoffs between two or more competing design objectives. However, integrated approaches are suggested as well for the specific applications of a given GDGB supported rotor.

7.1 Optimization of Standalone Gas Bearings

Despite the continuous effort invested in the development of reliable and efficient gas bearings throughout the past decades, there is a lack of guidance to design gas lubricated bearings for different applications. The main challenge is that the design objectives are dependent on the specific applications and that the optimal bearing design is relying on the selected design objectives.

However, for standalone gas bearing, the design often aims to achieve the maximization of stability and load capacity and/or the mitigation of the mechanical losses. Ideally, the results are expressed with non-dimensional numbers so that the results can be transferred to a wide variety of applications.

Early optimizations of groove geometries for incompressible fluid were based on simplified analytical solutions, and targeted the highest load, such as proposed by Whitley [88] for straight-grooved thrust bearings and by Muijderman for SGTBs [90]. However, the design of high-performance gas bearings in high-speed applications call for modern optimization methods that take advantage of high-performance computers. Advanced optimization algorithms, such as gradient-based [159,173], particle swarm [174], or evolutionary algorithms [133] can be used to search for the optimal designs based on the numerical solution of reduced-order physical models, such as the modified Reynolds equation for gas lubricated bearings [175].

In general, gradient-based algorithms are efficient. However, they can easily remain trapped in a local minimum. This drawback manifests especially in cases where the response surface is seeded with local minima and discontinuities. This is the reason why gradient based optimizers often get combined with other, more robust algorithms. The sequential (or successive) quadratic programming (SQP) is a widely-used gradient-based method for nonlinear and constrained optimization problems [176], which uses the quasi-Newton updating method to approximate the Hessian matrix. Hashimoto and Ochiai [75,76] applied a direct search (namely grid search within all feasible areas of design variables) to roughly locate multiple optimum candidates and then employ the SQP to find the local optimal results. Eventually, the multiple local optima are compared to select the global optimization solution. Harmony search (HS), initially introduced in
2001 by Geem et al. [177], is an efficient evolution-based algorithm, which can be utilized for a preliminary global search before the predetermined iteration limit is reached. Fesanghary and Khonsari [159] employ the HS method to identify the near-global-optimum regions. Subsequently, they apply the gradient-based SQP approach to identify and refine the optima.

Fesanghary and Khonsari [159] implied that the deformable groove top surface makes the optimization more complicated due to the coupling between fluid film behavior and the structural deformation of the grooved surface. In 2013 the same authors applied SQP to optimize the grooves for the incompressible lubrication between two parallel plates and achieved an increase in load capacity by 14-36% compared to a classical design.

The particle-swarm-optimization method can search very large spaces, however, it does not necessarily guarantee the identification of the truly optimal solution. Lehn and Schweizer [170] adopted this method to optimize the binary (two-sided) groove geometries for a spiral SGTB with respect to the load capacity and friction coefficient, which suggested that a two-sided spiral SGTB is not necessarily superior to the one-sided spiral SGTB.

The advantage of evolutionary optimization approaches is that the search is very wide and offers a stochastic behavior (mutation), which allows such algorithms to identify the globally optimal solution. However, compared to other algorithms they are less efficient [178,179]. An additional advantage of evolutionary algorithms is that they can handle multi-objective optimizations without having to define weights. Schiffmann [160] successfully applied evolutionary algorithms to identify the ideal groove geometries of herringbone grooved journal bearings.

Hashimoto and Ochiai [164] conducted optimizations of a thrust bearing including the groove geometries, the outer radius, and the groove radius ratio, with the objective to minimize the ratio of friction torque to stiffness. The optimization procedure sets off with a search of the optimum number of grooves for both the spiral and herringbone thrust bearings at a fixed speed. Then the fixed groove number of 24 was used for the optimization of the four design variables related to the groove geometry. For both types of thrust bearings, the optimal values for the depth, and groove width ratio reach the upper limits for all rotational speeds. On contrary, the optimum groove angle favored the bottom limit, below 20°. Experimental validation was conducted for the optimized thrust bearings, which suggested that the optimized designs achieved better performance than the conventional designs.

In 2008, the same authors [75] continued their optimization effort to the modification of the groove shapes for both thrust bearing configurations. Starting from the well-accepted logarithmic spiral grooves as a base geometry, the optimization partitioned the groove in the radial direction and then obtained the spline interpolation functions for the subsequent search. The constraints were extended to limit the damping to the objective function was to maximize the axial stiffness. The optimizer found a new groove pattern that originates from the spiral grooves but resembles asymmetric and distorted herringbone contours with different orientations. The results showed that the optimized groove pattern produces an axial stiffness around four times higher than the original SGTB under light load condition. However, along with the increase in stiffness, there is a drop in the optimal film thickness, which generates a friction torque approximately twice of that in the
original SGTB. Nonetheless, this improvement due to the newly-optimized groove pattern decreases with axial load.

Subsequently, Hashimoto and Namba [76] added objective functions, such as the maximization of the air film thickness and the minimization of the bearing torque and investigated combinations of the three objectives. The comparison of the optimal solutions obtained by applying different objectives revealed different optimum groove profiles. A spiral groove pattern is preferred by the target is the maximization of film thickness or the minimization of friction torque. A modified spiral pattern with a distortion towards the outer radial part is yielded under the objectives of maximized dynamic stiffness under constant load, maximal stiffness and film thickness and the minimization of the friction torque along with the maximizing stiffness.

There are a handful of studies [180–183] on the optimization of general textures on fluid-film bearings. A good review about surface texturing of general fluid-film bearings is available by Gropper et al. [184]. The optimizations of complex texturing patterns involving multiscale analysis are introduced by Podsidiado and Stachowiak[180] using fractal signature methods and a unified computational approach. Moreover, a combination of genetic optimization and CFD could be used for three-dimensional designs of micro-thrust bearings [182].

These examples introduced above clearly highlight the benefit of using optimization algorithms to identify optimal bearing solutions. However, these applications of optimizers to the bearings design reveals that the optimal bearing solution is sensitive to the selected objectives, which makes the generalization of optimization results challenging.

### 7.2 Optimization of Gas Bearing Supported Rotors

Normally, a gas bearing is supposed to support a rotor. Hence, a good design of gas bearings not only needs to achieve a good load capacity but also needs to be able to support the specified rotor up to its maximum rotor speed in a stable manner. Since the optimal bearing geometry is strongly dependent on the optimization objectives applying the results of bearings optimized as stand-alone elements does not necessarily lead to an optimal solution. As an example, operating conditions of a gas bearing supported compressor can easily influence the thermal management of the gas bearing supported rotor, which could lead to a bearing clearance distortion. Since the bearing clearance is often very small, even a limited clearance distortion due to thermal gradients can affect the rotor dynamic behavior. As a consequence, it is advised to apply state-of-the-art integrated design and optimization approaches that include the rotor and the bearing design.

First steps towards this direction have been undertaken by Schiffmann and Favrat [72,133]. They have formulated the design of a gas bearing supported rotor and in a second step the design of a gas bearing supported turbocompressor for heat pump applications as a multi-objective optimization problem. The comparison between a turbocompressor design with individually optimized components and an integrally optimized machine suggests a seasonal compressor performance increase by 12 points as a result of the integrated approach. It is suggested that this improvement is due to the fact that the interactions between the various components are considered with an integrated approach. Besides identifying optimal solutions an integrated and automated design approach offers the benefit that the identified solution is truly optimal. This cannot be guaranteed with a manually designed machine.
The downside of integrated approaches is that high-fidelity and reduced-order models are required to reflect the governing physical phenomena of gas bearing supported rotors. This includes models for the gas lubricated bearings and models for the rotordynamics that are able to capture the gas bearing properties, which are dependent both on the rotational and an excitation frequency. Further, an integrated multi-objective optimization often uses evolutionary algorithms, which requires a large amount of evaluations in order to converge towards a Pareto curve. Hence, the models not only need to be of high-fidelity but also need to be time-efficient.

8 Summary and Concluding Remarks
This paper offers an overview of the state-of-the-art of grooved dynamic gas bearings, echoing the increased interest in oil-free, small-scale turbomachinery.

The review starts with a classification of gas lubricated bearings and then introduces different modeling methodologies with a focus on grooved dynamic gas lubricated bearings. A particular emphasis is given to the Narrow Groove Theory, which through the assumption of an infinite number of grooves allows to obtain a modified Reynolds equation that includes the pumping effect of the groove patterns. This approach offers an elegant workaround to the challenging discretization of the thin fluid film if direct numerical approaches are implemented.

The paper is then extended to give an overview of efforts to include physical phenomena that are often neglected in thin film lubrications such effects of real gas lubrication, rarefaction, local compressibility, turbulence, inertia, edge influence, and thermal boundaries. Some of these effects are of particular importance for more recent applications such as in turbocompressors for heat pumps, expanders for Organic Rankine Cycles or in supercritical CO2 Brayton cycles.

Due to the critical role of the groove design, the details of how different groove parameters influence the bearing performance are discussed. There is clearly a trend away from classical groove patterns towards more advanced geometries that improve the bearing performance but are also more robust with regards to manufacturing deviation.

Furthermore, this paper gives a critical assessment of the state-of-the-art and offers the authors’ insights on how to shed more light into the complex mechanisms governing grooved gas lubricated bearings. The main challenge in gas bearing research lies in the lack of thorough and extensive experimental data. To complete, the review extends to the optimization of grooved gas bearings as well as to the integrated design of gas bearing supported rotor systems, which links the scope of this review to practical applications.

ACKNOWLEDGMENTS
The authors acknowledge the funding by the Swiss National Science Foundation grant PYAPP2 _154278/1.

NOMENCLATURE

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
<th>Unit</th>
</tr>
</thead>
</table>
$C_{xy}, C_{yx}$ Direct stiffness in the two orthogonal directions $[N \cdot \text{s/m}]$

$h$ Film thickness $[\text{m}]$

$K_{xy}, K_{yx}$ Direct stiffness in the two orthogonal directions $[N/\text{m}]$

$N_g$ Number of grooves $[\text{kg}]$

$P$ Pressure $[\text{Pa}]$

$R$ Radius $[\text{m}]$

**Greek symbols**

$\alpha$ Groove width ratio

$\beta$ Groove angle

$\lambda$ Molecular mean free path $[\text{m}]$

$\Lambda$ Compressibility number / Bearing number $[-]$

$\mu$ Lubricant viscosity $[\text{Pa.s}]$

$\tau$ Pressure ratio $P/P_a$ $[-]$

$\Omega$ Rotating speed $[\text{rad/s}]$

**Subscripts**

$a$ Atmospheric

$g$ Groove

$i$ In/Inlet

$o$ Out/Outlet

$r$ Ridge

**Acronyms**

FC Fuel cell

FDM Finite difference method

FNG Finite number of grooves

FEM Finite element method

FVM Finite volume method

GB Gas (or Gas-Lubricated) bearing

GDGB Grooved dynamic (hydrodynamic) gas bearing

GDJGB Grooved dynamic(hydrodynamic) journal gas bearing

GDTGJB Grooved Dynamic Thrust GBs
REFERENCES


