Naval Research Laboratory
Washington, DC 20375-5320

NRL/8230/FR--2022/3

Fundamentals of External Gear Pump Design

LOGAN T. WILLIAMS

Dynamics and Control Systems Branch
Naval Center for Space Technology

January 20, 2022

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
		START DATE	END DATE
January 20, 2022	Formal Report	October 1, 2013	January 20, 2022

4. TITLE AND SUBTITLE
Fundamentals of External Gear Pump Design

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**
		6897

6. AUTHOR(S)
Logan T. Williams

7. PERFORMING ORGANIZATION / AFFILIATION NAME(S) AND ADDRESS(ES)
U. S. Naval Research Laboratory
4555 Overlook Avenue SW
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER
NRL/8230/FR--2022/3

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR / MONITOR'S ACRONYM(S) NUMBER	11. SPONSOR / MONITOR'S REPORT NUMBER(S)
Office of Naval Research One Liberty Center 875 N. Randolph Street, Suite 1425 Arlington, VA 22203-1995	ONR	NRL/8232/FR- -2022/xx

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Statement A, Approved for public release. Distribution is unlimited.

13. SUPPLEMENTAL NOTES
supplemental notes

14. ABSTRACT
This report seeks to provide a foundational level of understanding of the primary mechanisms of operation and loss sources that provides designers a starting point for a gear pump design. The scope of this report includes the quantitative and qualitative parameters and metrics to define pump performance, design of individual pump components, some of the critical mutual dependencies of those components, pump fabrication, and the integration and testing of a gear pump prototype.

15. SUBJECT TERMS
External gear pump, design methodology, hydraulics

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	c. THIS PAGE		
U	U	U	UU	128

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
Logan T. Williams	202-767-9166

This page intentionally left blank

CONTENTS

FIGURES

TABLES

This page intentionally left blank

EXECUTIVE SUMMARY

This report presents a consolidated summary of the fundamental concepts and lessons learned in the design of external gear pumps that were accumulated in the course of research at the US Naval Research Laboratory. External gear pumps are a ubiquitous mechanical device, valued for their robustness and simplicity. However, this apparent simplicity can belie the sensitivity of the performance of a gear pump to numerous design parameters. As such, while the basic composition and operation of the external gear pump are considered common knowledge, the analytical formulations of the critical mechanisms in play are not. This leads to the problem that while information on what an external gear pump is, or how to make one, is readily available, information about how to make an efficient pump is not. This is particularly the case in high-power and/or miniature gear pumps. This report seeks to provide a foundational level of understanding of the primary mechanisms of operation and loss sources that provides designers a starting point for a gear pump design. The scope of this report includes the quantitative and qualitative parameters and metrics to define pump performance, the design of individual pump components, some of the critical mutual dependencies of those components, pump fabrication, and the integration and testing of a gear pump prototype.

This page intentionally left blank

FUNDAMENTALS OF EXTERNAL GEAR PUMP DESIGN

1. INTRODUCTION TO EXTERNAL GEAR PUMPS

The external gear pump is a ubiquitous device in fluidic systems, commonly used for driving a variety of fluid flows such as oil, chemicals, or polymers. A simple, compact, and rugged device, external gear pumps are effective in supplying specific flow rates at a range of pressures; this combination of reliability and accuracy makes external gear pumps appealing in industrial and scientific applications. Despite the widespread use of external gear pumps, it can be difficult to find consolidated information as to their nature of operation in sufficient detail to design high-performance pumps. While such knowledge is common within industry, it is harder for acquire for an academic or someone outside the industry. In contrast, information about other, more technically and conceptually challenging pump types, such as the centrifugal pump, can be acquired much more readily. The purposes of this report are to establish the fundamental mechanisms that drive gear pump performance and to provide a framework for designers to design, build, and evaluate their own gear pumps. We begin with this section, in which we discuss the background of external gear pumps, where they fit in comparison to other pumps, and a basic overview of how they work. In order to provide an effective means of evaluation, we also define the key performance metrics and establish a methodology for comparing pump performance between design iterations.

1.1 Background

The external gear pump is one type of pump amongst a family tree of possible pump configurations. The various pump designs that exist can be grouped according to their functional similarities. The highest-level distinction is the overall manner in which the fluid is pressurized. Dynamic pumps impart energy into a flow and leverage the conversion of fluid kinetic energy into static pressure. Displacement pumps draw fluid into an enclosed pumping chamber and then compress the volume to expel the fluid; the mechanical force required to compress the pumping chamber then pressurizes the fluid to ensure fluid expulsion at the rate of volume contraction. Dynamic pumps can in turn be sorted by the mechanism of energy transfer to the fluid, such as rotary impellers, electromagnetic effects, or fluid hammer, while displacement pumps are grouped by whether the pumping cycle is reciprocal or rotary in nature. Further categorizations can be made within each group in successive detail, for example by the mechanism of expanding and contracting the fluid-pumping chamber. Gear pumps are a type of rotary displacement pump in which the pumping chamber is the space between gear teeth and the meshing action of two gears provides the change in volume of the pumping chamber to drive the fluid. Gear pumps can be divided further by the type of gears used, such as external gears, internal gears, and gerotors. A representative configuration breakdown is shown in Fig. 1.

Manuscript approved May 19, 2022.

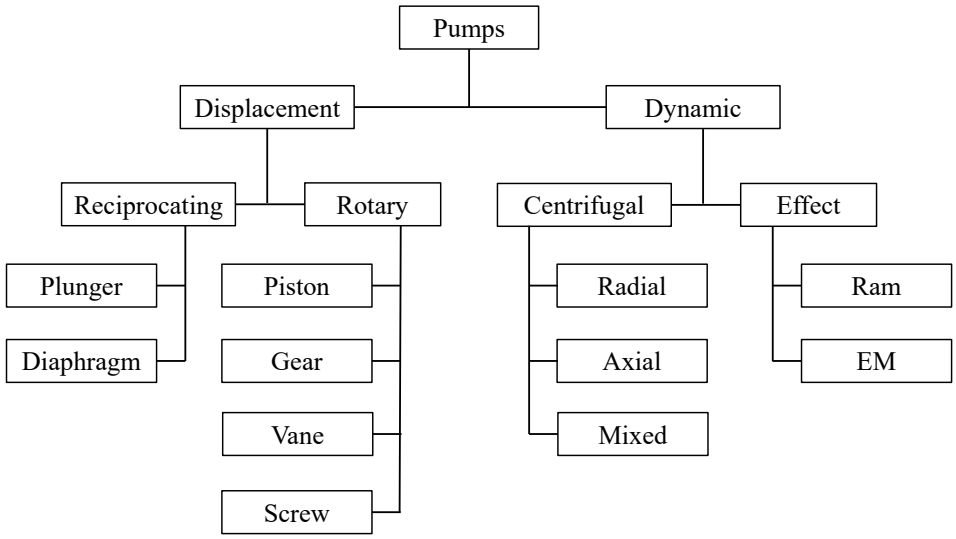

Fig. 1—Pump configuration breakdown

Due to the high number of possible pump configurations for consideration, a quantitative parameter is useful during preliminary selection. A commonly used parameter in pump design is the pump specific speed, N_s, which is a nondimensional parameter that captures the pump's operating conditions. The specific speed is defined differently in the US and SI systems,

$$N_s = \frac{N[rpm]Q^{1/2}[gpm]}{H^{3/4}[ft]}(US)$$
$$\Omega_s = \frac{\omega[rad/s]Q^{1/2}[m^3/s]}{(gH)^{3/4}[m^2/s^2]}\ (SI), \tag{1.1}$$

where N is the pump shaft speed, ω is the angular shaft speed, Q is the volumetric flow rate, H is the pump head, and g is acceleration due to gravity. The units for these terms are given in brackets. Empirically it has been shown that a pump configuration will have a range of specific speeds for which it is optimal, summarized in Fig. 2 [1]. From the figure, it can be seen that gear pumps are ideal for applications in which moderate flow rates are required at midrange to high pressures (200–3,000 psi). While piston pumps can offer higher pressures and vane pumps can offer higher flow rates, gear pumps excel at providing a combination of both.

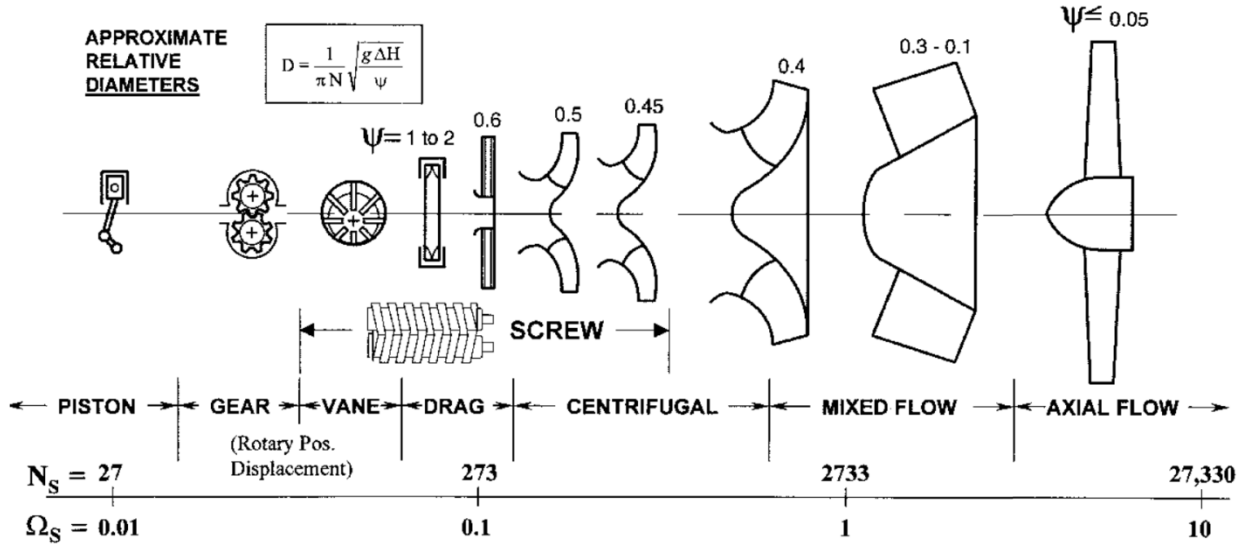

Fig. 2—Optimal pump configuration as a function of specific speed [1]

1.2 Overview of External Gear Pumps

The external gear pump operates through the action of two meshed external gears. As the gears rotate, the teeth unmesh at the suction side, creating a void and drawing in fluid. The gears then carry the fluid along the wall of the pumping chamber as they continue to rotate. The fluid trapped between the gear teeth is then forced into the discharge region as the gear teeth mesh. The clearances between the gears and the adjacent components create the seal that restricts backflow through the pump. An illustration in Fig. 3 gives the cross section of an external gear pump with a color gradient to represent fluid pressure within the pump — red denoting high pressure and blue denoting low pressure.

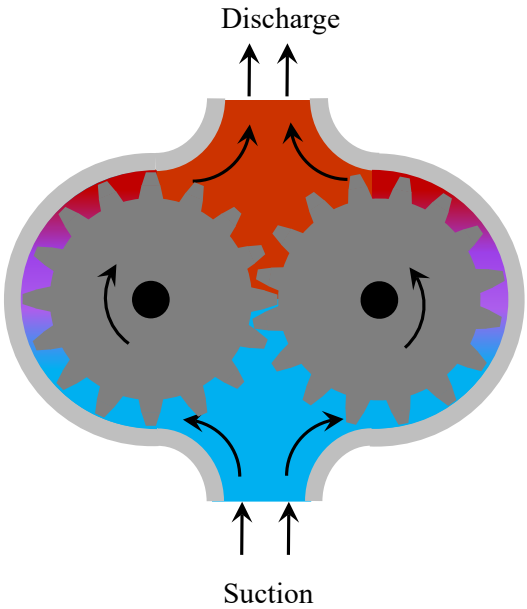

Fig. 3—Illustration of the cross section of an external gear pump

While there is no single external gear pump design, a common configuration is shown in Fig. 4, with an exploded rendering shown in Fig. 5. The core design concept of the pump is a pumping chamber that is formed inside the pump casing between two wear plates. These plates are pressed into the gears and provide axial containment, while the pump casing provides radial confinement of the fluid. Internal sealing is provided by o-rings for static seals and a lip seal for the rotary seal at the drive shaft. The two shafts are supported by journal bearings in the wear plate that counteract the radial load on the gears. There are three fluid connections, with two used for the pump inlet/outlet, and the third used as a bearing drain. This allows fluid in the pump to leak through the journal bearings and out through the drain, creating a fluid circuit that continuously lubricates the bearings.

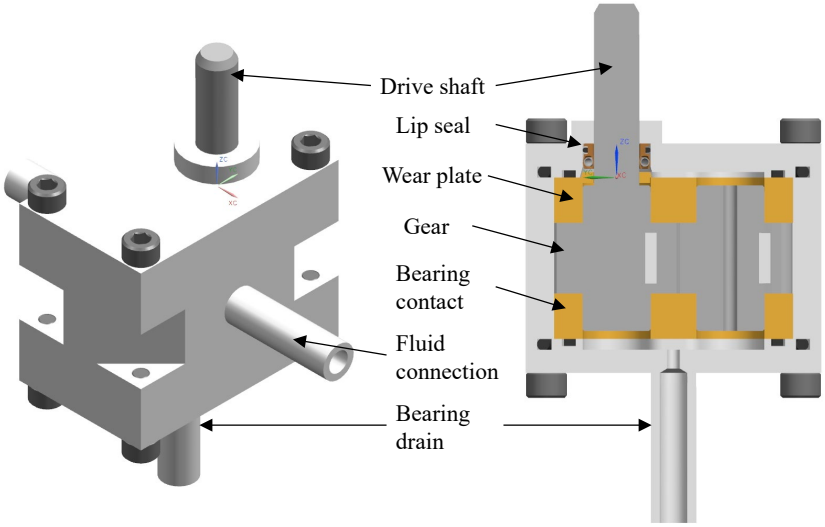

Fig. 4—CAD model of example gear pump. Exterior view (*left*) and cross-section view (*right*)

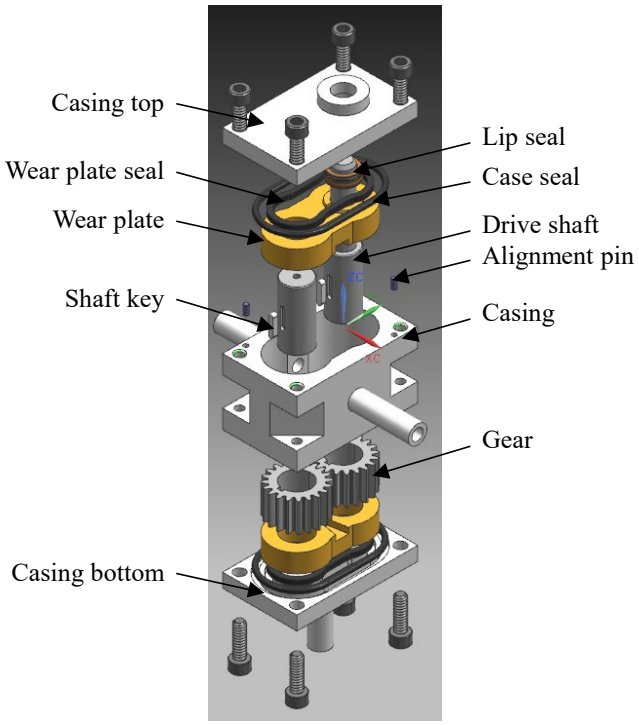

Fig. 5—Exploded view of gear pump CAD model

While gear pumps are simple in concept with few moving pieces, the design and fabrication of a highly efficient gear pump is more complicated in practice, particularly for high-pressure pumps. Fluid within the discharge region is driven back through the pump into the suction region due to the pressure gradient between discharge and suction. The clearances between the gears and the pump casing as well as between the gears and the wear plates dictate this recirculation flow. These interfaces between components also affect viscous forces on the gears, increasing the torque load of the pump. Additional concerns include shaft support, casing strength, static and dynamic seals, and component wear. Properly addressing these design considerations is what separates a compact, efficient pump driving a 2,000-psi hydraulic system and a pump that ineffectually recirculates fluid internally at 600 psi.

1.3 **Performance Metrics**

If the goal of a gear pump designer is to provide a good pump, then we need to define parameters to describe the performance of a pump and to quantify what constitutes "good." There can be many system considerations for a pump, from mass limitations to material restrictions to cost constraints. Many of these are application specific — what is optimal in the hydraulic industry would be a poor fit in the food-processing industry. Setting aside such specific requirements, there are several common parameters that capture operating conditions and overall performance and can be used to quantify the "goodness" of a pump. These parameters describe the mechanical input to the pump, the hydraulic output from the pump, and the efficiency of the pump.

1.3.1 *Shaft Input*

A pump does not operate in isolation, but must be driven by some prime mover, be it an electric motor, a belt drive, magnetic coupling, or some other mechanical means. Most gear pumps feature a drive shaft that couples the pump to the prime mover. The speed at which this shaft is driven and the torque applied define the input mechanical power to the pump. This input shaft power, P_s, can be written as

$$P_s = \omega T_s, \tag{1.2}$$

where T_s is the shaft torque. The shaft speed sets the pump speed, and the torque required to drive the pump at this speed reflects the force required to pressurize the fluid and to overcome frictional forces within the pump. Pump speed is often used in either the Imperial system, denoted by N and given in revolutions per minute, or in the SI, denoted by ω and given in radians per second. The two are related by the conversion below.

$$\frac{2\pi}{60} N = \omega \tag{1.3}$$

1.3.2 *Hydraulic Output*

In basic terms, the output flow from the pump can be characterized by the volumetric flow rate out the discharge of the pump, Q, and the pressure of the discharge flow. It is more accurate to say that the pump provides a pressure differential relative to suction pressure; this is particularly relevant for closed pumping systems in which fluid is drawn from a reservoir, driven through some load, and returned to the reservoir. In such systems, the reservoir pressure sets the baseline pressure to which all other system pressures are relevant. An electrical analogy is that of an isolated circuit that is biased above ground, and the pump is like a voltage supply referenced to that electrical common. The pressure differential of the pump can be defined as

$$\Delta p = p_d - p_s\,, \tag{1.4}$$

where p_d is the pump discharge pressure and p_s is the suction pressure. The product of the volumetric flow rate and the pressure differential is the hydraulic power output of the pump:

$$P_{out} = Q\Delta p\,. \tag{1.5}$$

A common method of presenting pump performance is to plot one parameter as a function of another, the most common of which is output flow rate as a function of pump speed. An example plot is shown in Fig. 6.

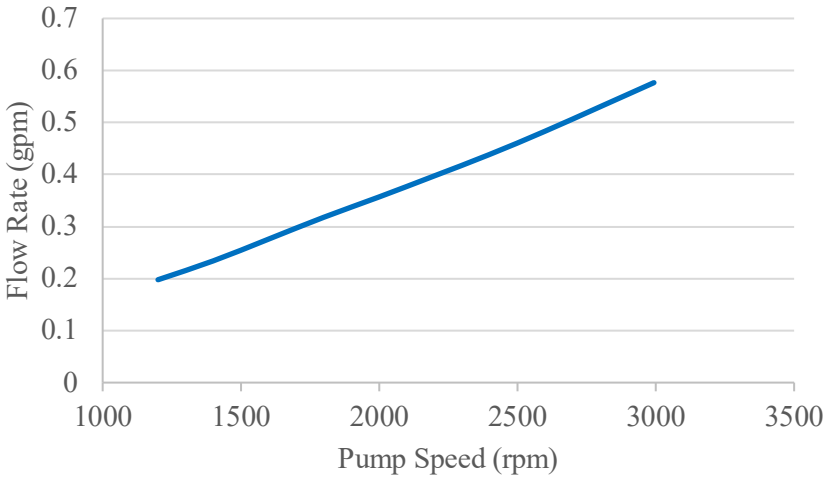

Fig. 6—Example performance curve of output flow rate as a function of pump speed

1.3.3 *Pump Efficiency*

The primary goal of a pump is to convert the input mechanical power into output hydraulic power. The ratio of the output power to the input shaft power is the overall efficiency of the pump, and by utilizing Eqs. (1.2) and (1.5), we can quantify this parameter:

$$\eta = \frac{P_{out}}{P_s} = \frac{Q\Delta p}{\omega T_s}\,. \tag{1.6}$$

Any real pump will have losses that waste power, and these losses can come in different forms, each quantified by their individual efficiencies. In a complex system, the product of all the constituent efficiencies yields the total efficiency for the device. For a hydraulic pump, there are three efficiencies of interest: volumetric, or the loss of delivered flow rate due to internal leakage; mechanical, or the loss of applied power due to friction; and hydraulic, or the loss of flow energy due to fluid losses at the discharge. Hydraulic efficiency is generally considered for centrifugal pumps, where the stagnation of the fluid in the volute passage and subsequent flow through the discharge can dissipate a non-negligible fraction of the delivered power. For gear pumps and other designs that have no such obstacle to the discharge port, hydraulic inefficiencies can be discounted. Thus, the two primary forms of loss that must be considered are volumetric efficiency losses due to internal leakage of the discharge flow around the gears back to the suction side, and mechanical efficiency losses due to friction and viscous losses.

Volumetric efficiency is a measure of how much of the nominal volume of pumped fluid is lost to internal leakage between the discharge and suction ends of the pump. As with any device with moving

components, there are finite clearances that separate dynamic objects, in this case the gears from the pump casing or wear plates. The pressure difference between discharge and suction thus will drive a leakage flow through these clearances. There are two main pathways for fluid to leak around the gears: radial leakage between the gear teeth and the side walls of the pump casing, and axial leakage between the axial face of the gears and the wear plates. An illustration of the two leakage pathways is shown in Fig. 7. A third leakage pathway is through the meshing zone of the gears, where fluid is driven through any finite gaps between mating teeth surfaces or where any fluid not expelled into the discharge region is returned to the suction region. This last leakage pathway is determined primarily by the design and fabrication quality of the gears used in the pump.

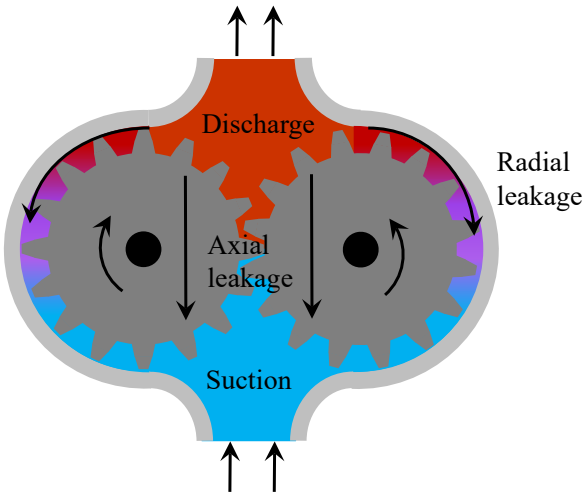

Fig. 7—Internal leakage pathways around the gears

The volumetric efficiency of a pump can be defined as the ratio of the actual flow rate supplied to the ideal flow rate expected.

$$\eta_v = \frac{Q_{actual}}{Q_{ideal}} \tag{1.7}$$

Fixed-displacement pumps have a nominal displacement, *D*, which is the ideal amount of fluid pumped per cycle. The ideal flow rate thus can be defined as the product of the displacement and the pump speed.

$$Q_{ideal} = \frac{ND}{60} \tag{1.8}$$

If you look closely at Fig. 6, you can see that the flow rate curve is not perfectly linear; this reflects the impact of volumetric efficiency on the output flow of a pump. By plotting the flow rate and the volumetric efficiency together, shown in Fig. 8, we can see that efficiency increased as pump speed increased, which in turn increased the slope of the flow rate curve.

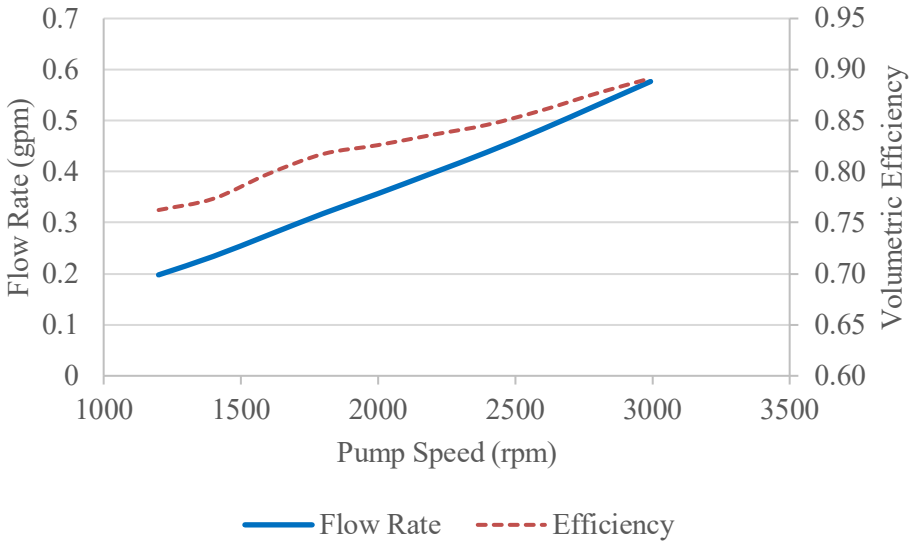

Fig. 8—Example plot of flow rate and volumetric efficiency as a function of pump speed

Whereas volumetric efficiency is concerned with the loss of fluid flow, mechanical efficiency relates to the loss of input mechanical power due to friction and viscous losses. This takes the form of dynamic seal friction (such as a shaft seal on the drive shaft), friction between moving components (a gear and the wear plate), and viscous dissipation within fluid films. For an external gear pump, there are three primary areas where friction occurs: between the gear teeth and the radial wall, between the axial faces of the gears and the wear plates, and within the bearings that support the pump shafts. There are other places where friction can occur, such as at the shaft seal or the gear teeth meshing. However, a single-stage gear train generally has very high efficiency, and likewise, a properly selected shaft seal imparts only minor torque losses; both can be neglected for now. Thus, the design focus for maximizing mechanical efficiency lies in minimizing friction between the gears and the surrounding components while also providing a low-friction bearing option for shaft support.

The primary consideration for limiting friction within the pump is to ensure the generation of a full hydrodynamic film between components. There are three lubrication regimes: boundary lubrication, elastohydrodynamic lubrication, and full-film lubrication (in this application, hydrodynamic lubrication). Due to the nature of the lubrication interface being driven by Couette flow, the creation of the fluid film is based on the availability of fluid and the shaft speed of the pump. The first, and obvious, point is that the pump should never be run dry; a less obvious consideration is that the pump should be given a wet assembly with all interfaces pre-lubricated to ensure fluid engagement at start-up. A second consideration is that the fluid film will better reduce the coefficient of friction at higher speeds when more fluid can be entrained; thus, low-speed, high-pressure operation is not as preferable to high-speed operation. This is especially important for journal bearing operation, which is discussed in greater detail later on. After a certain point, further increases in speed will increase the coefficient of friction, but this is a more gradual response.

Assuming a hydrodynamic film has formed, the friction between two surfaces at differing velocities is dependent on the viscosity of the fluid, the thickness of the film, and the difference in velocity. This creates a trade between mechanical efficiency and volumetric efficiency, as a close-fit assembly minimizes leakage at a cost of higher friction. This can be mitigated somewhat by operating at a high pressure, as the frictional power lost is independent of the pressure of the fluid; thus, the higher the fluid pressure pumped, the greater fluid power delivered for the same friction power spent. This is an oversimplification, as there are mechanisms for pressure to impact the lubrication boundaries that may need to be considered, such as wear

plate interaction or deflection of the shafts leading to a decrease in radial clearance. However, it is useful as a rule of thumb that mechanical efficiency for a well-sealed pump improves at high pressures.

Mechanical efficiency can be quantified as the ratio of the ideal hydraulic power output of the pump to the actual shaft power required to drive the pump.

$$\eta_m = \frac{Q_{ideal}\Delta p}{\omega T_s} \tag{1.9}$$

It can be shown that the product of Eqs. (1.7) and (1.9) yields Eq. (1.6).

$$\eta = \eta_v \eta_m$$

$$= \frac{Q_{actual}}{Q_{ideal}} \frac{Q_{ideal}\Delta p}{\omega T_s}$$

$$\eta = \frac{Q\Delta p}{\omega T_s}$$

With these parameters, the "goodness" of a pump can be evaluated in terms of performative capability in its output flow rate and pressure differential as well as its performative quality in how efficiently it achieves these outputs. Pump sizing is the process in pump design (or pump selection, if purchasing from a commercial supplier) where the nominal performative capability of the pump is chosen, and it is primarily decided by the gear size and the nominal operating speed. On the other hand, the pump efficiency is determined by the entire pump design and the interfacing of the different components. Thus, the most complex task for a gear pump designer is the manner in which to achieve high pump efficiency for the nominal operating conditions and any other external constraints.

1.4 **Performance Analysis Methodology**

If you were to ask ten engineers for their recommendations of the best pump design, you likely would get ten different answers. The fact of the matter is that for systems beyond some minimum complexity, the various performance parameters are in tension, where improving one aspect of a design requires sacrificing another. For example, increasing the pressure capability of a pump inevitably requires increasing the mass of the pump. The increased pressure will decrease the volumetric efficiency of the pump, requiring additional design elements to compensate and ultimately increasing fabrication cost. The permutations of all these design trade-offs create a trade space within which all pump designs exist. The specific application in mind for a given design defines the region within the trade space that is optimal; a pump designed for high-viscosity fluids is a poor choice for low-viscosity fluids. Many of these considerations are based on the nominal use-case for the pump, and still others can be dependent on external factors such as cost or limitations on fabrication methods. Thus, defining the "best" design is highly subjective and impossible to define. However, a key element in the design process is the ability to define and quantify how one pump design might be "better" than another, particularly within the scope of objective performance metrics. This is particularly useful while iterating on a single element within the design in order to optimize that component or subassembly.

1.4.1 *Parameter Interdependency*

Yet even this sort of design comparison within a restricted focus can be difficult, as pump performance metrics are functions of multiple operating parameters. Thus, comparing one design iteration to another requires ensuring constant operating parameters, but this can be complicated further when these parameters

are mutually dependent. As an example, consider volumetric efficiency as a function of pump pressure differential. For a pump test setup using a relief valve as a variable pressure load, the discharge pressure is a function of the cracking pressure, p_c, of the relief valve and the output flow rate, and thus pump speed. The pump performance can be thought of as falling on lines of either constant speed or constant cracking pressure. Higher speeds increase volumetric efficiency at constant pressure, but increasing the speed inevitably will increase the pump pressure (due to nonzero system resistance), which lowers efficiency. A conceptual drawing of volumetric efficiency along lines of constant speed and cracking pressure is shown in Fig. 9.

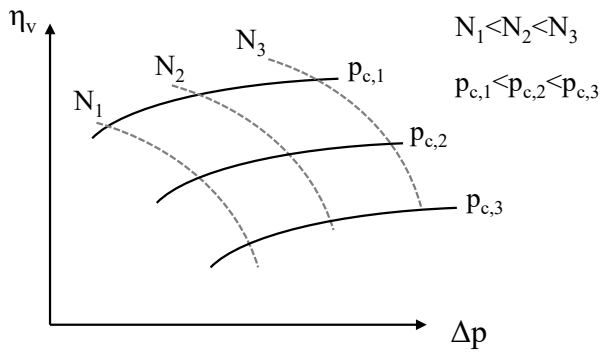

Fig. 9—Conceptual illustration of pump volumetric efficiency as a function of speed and pressure isolines

Extending the example, if one pump design has a nominally higher volumetric efficiency than another design, then the output flow rate of the two pumps at a given pump speed will be different, and will diverge as a function of increasing pressure. In such a scenario, it is nearly impossible for the two pump designs to operate at identical conditions for a direct comparison. Instead, it would be more useful if pump performance metrics could be given as a function of a single operating parameter, thus making direct comparisons more feasible. This is difficult for most performance metrics, but there is one for which this can be achieved: volumetric efficiency.

The classical approach to gear pump efficiency is simply to examine each component efficiency as a function of each parameter in turn. While this method allows easy conceptualization of the relationship between individual parameters and efficiency, it results in a large amount of performance curves and is highly susceptible to the interdependencies of the operational parameters. In such circumstances, comparisons between two pump designs require collection of performance data over the same operational space, which is not always available. Thus, there is a need for a methodology that readily facilitates comparisons between similar gear pumps without having to do comprehensive performance evaluations every time.

There are several existing frameworks for analyzing gear pump efficiency. Paul and Wanke have examined using the dimensionless Stribeck number as a single parameter to define pump efficiency [2]. This approach combines volumetric and mechanical efficiencies to create a function for the total efficiency utilizing analytical approximations for leakage and frictional torque. While the approach is reasonably able to capture the shape of the volumetric efficiency data, it is unable to consolidate the mechanical efficiency data accurately. A study by Schänzle, Störmer, and Pelz modeled gear pump efficiency using a nondimensional approach based on the similarity of gear pump architectures [3]. Their work aggregated the various parameters that define gear pump performance (pressure, speed, fluid properties, and pump geometry) into four dimensionless variables to capture total pump efficiency. The goal of this approach is to use a limited set of data on a given pump to extrapolate performance to different operating conditions or even different pumps within the same model family. Comparison of the model to empirical data showed

that the similarity method could reasonably estimate volumetric efficiency within 3% and mechanical efficiency within 10%, but was unable to reliably estimate fluid leakage or frictional torque losses directly.

One key problem with both approaches is that they are dependent on simplified loss models, which cannot fully capture the problem, or on multiple dimensionless loss coefficients that must be empirically determined from pump performance data. If the goal is to accurately capture pump performance to enable ready comparisons, then a methodology must have a reasonable expectation of accuracy while requiring limited curve fitting to measured data. The solution is to assume a simple and robust model for loss mechanisms that can be easily fit to a small sample of measured pump metrics. To begin with, a change of perspective of the problem is required.

1.4.2 *Hydraulic Resistance*

The volumetric efficiency term captures the internal leakages within the gear pump that reduce the delivered volumetric flow rate from the ideal flow rate possible. One inescapable fact of hydraulics is that a dynamic interface between two surfaces will always leak, and the gears within the pump are no different. The interface between the radial tips of the gear teeth and the pump casing, as well as between the axial faces of the gears and the casing (or wear plates, depending on design), allow some amount of backflow from the discharge side of the pump to the suction side. Additionally, some pump designs feature a return pathway from the shaft bearings to pump suction to ensure lubrication. Static clearances between components, seal damage, etc., all create pathways that pressurized fluid can take to leak back to suction. When evaluating pump performance, it is helpful to consider all of these pathways as a whole, rather than simply to focus on the pump discharge to the load in isolation. Figure 10 shows a block diagram of the flow pathways in a common gear pump architecture.

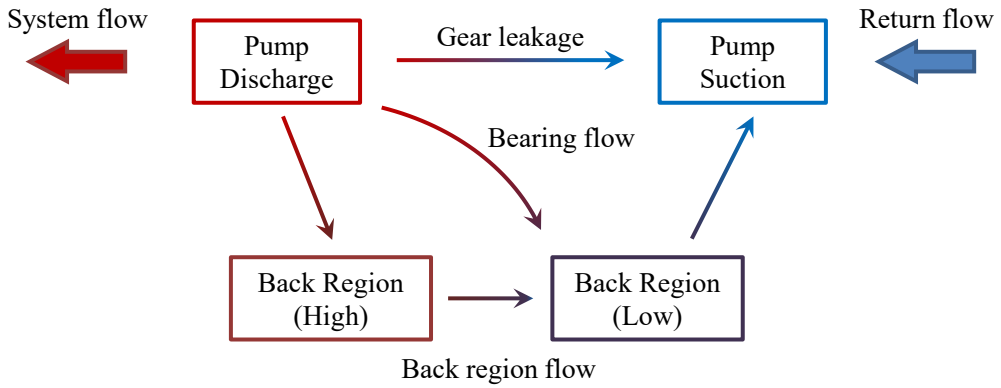

Fig. 10—Block diagram example of flow pathways in a gear pump

One method to evaluate all these potential fluid pathways is to consider the pump as a hydraulic circuit. In this analogy, volumetric flow rate is current, pressure is voltage, and flow restrictions are resistances. As such, the volumetric flow rate can be given in terms of the hydraulic resistance, R_h, of the flow pathway and the pressure drop across that pathway.

$$Q = \frac{\Delta p}{R_h}. \tag{1.10}$$

As a fixed-displacement pump, the pressure rise across a gear pump acts as a voltage source with a current output of the ideal flow rate, and each leakage pathway acts as a resistance in parallel. Figure 11

shows a simplified circuit for the gear pump schematic from Fig. 10, where the subscript *G* denotes the gears, *B* denotes the back region, and *b* denotes the bearings. Ideally, a well-designed bearing should not allow a significant amount of fluid through in comparison to the deliver flow rate, so this pathway can sometimes be neglected. One advantage to using the circuit analogy is that offers an easy way to conceptualize the different leakage pathways and how they are interrelated. It also provides an easy way to represent the pump efficiency as the ratio of the load system resistance to the overall equivalent resistance:

$$\eta_v = \frac{R_{eq}}{R_{sys}}. \tag{1.11}$$

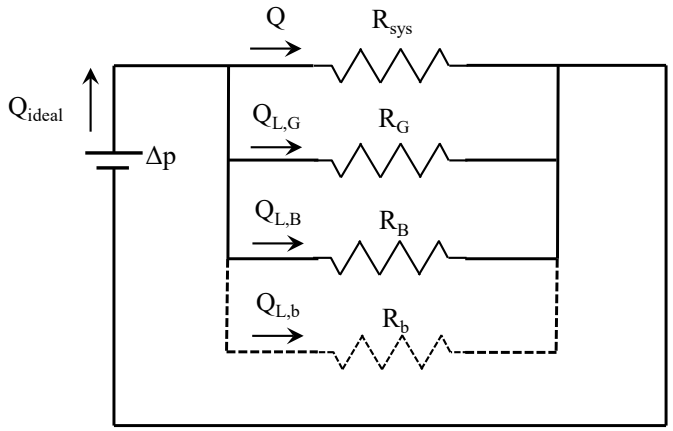

Fig. 11—Simplified hydraulic circuit of the gear pump

The main disadvantage with the circuit analogy is that in practice, fluid leakage past the gears is not solely a function of the pressure rise, but is also a function of shaft speed. Similarly, the system load may not necessarily have a linear relation between flow rate and pressure. As an example, a relief valve does not fit this simplified model of system load resistance, as the pressure drop across the valve is a function of the cracking pressure, p_c, in addition to the flow rate and some nominal load resistance, R_L.

$$\Delta p = p_c + QR_L \tag{1.12}$$

The observed system pressure drop is shown in Fig. 12 for a relief valve, a filter, and associated hydraulic lines as a function of flow rate. The data closely fits the form of Eq. (1.12), though the slope of a linear regression line for a data set decreases as the cracking pressure increases. Therefore, while the circuit analogy can be a useful tool for conceptualizing the problem, it is not a complete analytical tool.

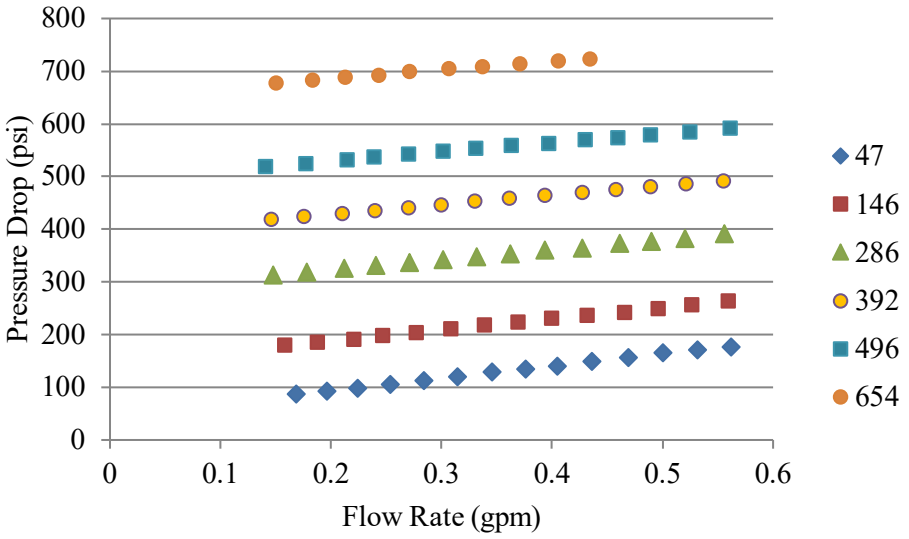

Fig. 12—System load pressure drop as a function of flow rate and relief valve cracking pressure in psi

1.4.3 *Hydraulic Resistance as an Operating Parameter*

Two of the most important operating conditions that dictate pump performance are the pressure differential across the pump and the pump speed. Since the pump speed is related to volumetric flow rate, the two parameters can be linked by the hydraulic resistance in Eq. (1.10). In this convention, we do not use the hydraulic resistance as a description of the load itself, but instead create a new pump operating parameter from the ratio of the pump pressure to the flow rate. As an example, suppose we had a gear pump where the only leakage pathway was through the gears, and the leakage flow rate, Q_L, was linearly related to the pump speed and pressure.

$$Q_L = Q_G = \frac{\Delta p}{R_G} + \frac{N}{60\mathbb{R}_G}. \tag{1.13}$$

In Eq. (1.13), we define the leakage flow as the sum of a pressure term and a speed term. Continuing and expanding on the hydraulic circuit analogy, the pressure-dependent leakage is a function of the gear fluid resistance, R_G, and speed-dependent leakage is a function of the gear reluctance, $\mathbb{R}_G$. The term reluctance is used for convenience and does not imply any behavioral similarity to magnetics. We can then define the volumetric efficiency in terms of ideal and leakage flow rates.

$$\eta_v = \frac{Q_{actual}}{Q_{ideal}} = \frac{Q_{ideal} - Q_L}{Q_{ideal}}$$

$$\eta_v = 1 - \frac{Q_L}{Q_{ideal}}$$

$$\eta_v = 1 - \frac{\frac{\Delta p}{R_G} + \frac{N}{60\mathbb{R}_G}}{ND/60}$$

$$\eta_v = 1 - \frac{60\Delta p}{NDR_G} - \frac{1}{D\mathbb{R}_G}$$

Substituting in Eqs. (1.7) and (1.10),

$$\eta_v = 1 - \eta_v \frac{R_h}{R_G} - \frac{1}{D\mathbb{R}_G}$$
$$\eta_v = \frac{1 - \frac{1}{D\mathbb{R}_G}}{1 + \frac{R_h}{R_G}}. \tag{1.14}$$

Equation (1.14) thus reduces the volumetric efficiency to a function of a single operating parameter and three pump configuration parameters. The gear resistance, gear reluctance, and pump displacement are all determined by the pump assembly geometry and the choice of fluid. One caveat is that, as will be demonstrated later, the fluid viscosity is captured by the gear resistance term; this means that the gear resistance is a function of the fluid used and the temperature of the fluid at any given point in time.

The utility of using hydraulic resistance as an operating parameter is best demonstrated through an example case study. Suppose we take a small gear pump in the fixed wear plate configuration (i.e. the axial and radial clearance around the gears are set at pump assembly and are constant) and test the pump at various speeds and pressures. Each test duration is limited to a few minutes so that heating of the working fluid is minimal. The measurement of the output flow compared to the pump's displacement using Eqs. (1.7) and (1.8) yields the volumetric efficiency of the pump at each point. If we plot the volumetric efficiency against either pump speed or pressure differential, as shown in Fig. 13, the result is of only marginal utility. This is due to the fact that neither pump speed nor pressure differential, alone, is sufficient to describe the operating state of the pump. The plots in Fig. 13 are only useful to establish rough trends that volumetric efficiency improves with increasing pump speed and decreasing pressure differential.

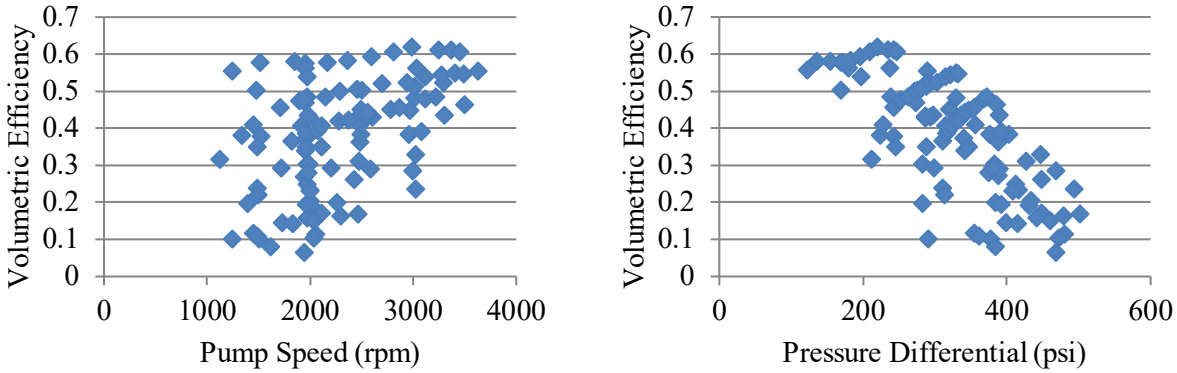

Fig. 13—Volumetric efficiency of a fixed wear plate configuration pump as a function of either pump speed or pump pressure differential. Data collected across multiple tests over multiple days at near constant fluid temperature.

In contrast, the same data is plotted in Fig. 14 as a function of the hydraulic resistance. Here we can see that the data lies along a single curve defined by two constants and the hydraulic resistance. By comparing the curve fit equation in Fig. 14 to Eq. (1.14), we can link these two constants to the pump displacement, the gear resistance, and the gear reluctance. This demonstrates that for fixed fluid viscosity, pump displacement, and internal clearances, the volumetric efficiency is defined solely as a function of the hydraulic resistance, with two key consequences. The first is that the comparison of the volumetric efficiency of two pump designs with a similar displacement is greatly facilitated, as it is simply a comparison of two curves defined by a single parameter, rather than a broad collection of data points. The second consequence of importance is that the volumetric efficiency curve can be determined from a curve fit of empirical data as part of a pump performance evaluation. This curve fit can then be used to calculate

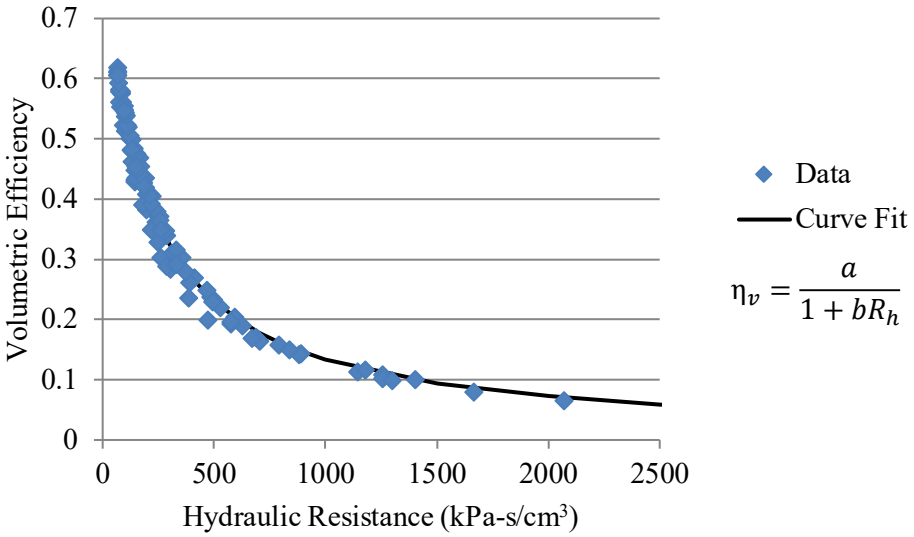

Fig. 14—Volumetric efficiency of a fixed wear plate configuration pump as a function of hydraulic resistance. Data corresponds to the data plotted in Fig. 13, and the curve fit matches the form of Eq. (1.14).

a full curve across a range of hydraulic resistances. Thus, a comparison of two pump designs does not require the two pumps to be tested at the same operating conditions.

The use of the hydraulic resistance is similar to the method by Michael and Wanke that utilizes the Stribeck number, as both approaches seek to consolidate the multiple operating conditions down to a single parameter. The Stribeck number more conveniently captures viscosity, whereas the hydraulic resistance parameter cannot and must rely on either a variable gear resistance, pulling viscosity out of the resistance term and accounting for it as well (resulting in two parameters to track), or some combined parameter instead of the hydraulic resistance. However, the hydraulic resistance approach does not require a preset leakage or loss model and instead can empirically determine the gear resistance and reluctance using a limited set of data. Higher-fidelity modeling of these terms can be used for predictive purposes, but the method is not constrained by them. The Stribeck number method also does not allow for the uniqueness of the speed and pressure effects; for example, analytical modeling suggests that speed-dependent leakage flow is not a function of viscosity, but this is not captured by the Stribeck method.

Conversely, the similarity method is more robust in capturing empirical trends, as the loss models have more parameters that can capture the complexity of the system. However, that complexity, in turn, requires more empirical data and analysis in order to determine five nondimensional coefficients. Both the similarity method and the hydraulic resistance method capture the speed and pressure effects, yet the hydraulic resistance method requires only two empirical parameters to define the performance curve, which can be solved with a single curve fit.

1.4.4 *The Hydraulic Parameter*

One critical assumption when using hydraulic resistance as an operating parameter is that the fluid viscosity is constant, both within a given pump performance evaluation and between two different pump performance evaluations when comparing designs. This is due to the fact that, as mentioned earlier, fluid viscosity is captured by the gear resistance term, and in order to assume a fixed gear resistance, we must assume viscosity is likewise fixed. However, in practice, this is rarely the case, as viscosity is a function of fluid temperature, which, in turn, is a function of any heat loads, ambient conditions, mechanical efficiency of the pump, and the power throughput of the pump over time. Additionally, in order to make an accurate

comparison between two pump designs, the corresponding performance evaluations would have to take place with identical fluid temperatures. Such onerous conditions defeat the original intent of the approach, which is to enable design comparisons without requiring such rigorous condition matching. Therefore, the hydraulic resistance method can be improved by removing the assumption that includes viscosity within the gear resistance term.

To begin, let us define an inviscid gear resistance, R_g, that can be related to the viscous gear resistance term, R_G, that we have used thus far:

$$R_g = \frac{R_G}{\mu}, \tag{1.15}$$

where μ is the dynamic viscosity of the fluid. Substituting this into Eq. (1.14),

$$\eta_v = \frac{1 - \frac{1}{D\mathbb{R}_G}}{1 + \frac{R_\mu}{R_g}}, \tag{1.16}$$

where R_μ is the hydraulic parameter, defined as

$$R_\mu = \frac{R_h}{\mu} = \frac{\Delta p}{\mu Q}. \tag{1.17}$$

Much like the hydraulic resistance, the hydraulic parameter captures the impact of pump speed (in the form of flow rate) and pressure, but it additionally includes the fluid viscosity. This removes the requirement of constant fluid temperature, or even of similar fluids, when making a performance comparison between pumps. This also applies when determining the gear resistance and the gear reluctance through curve fitting to empirical data, as individual data points no longer need to be taken at the same temperature. This is of particular use when testing high-pressure pumps, as the high fluid power results in rapid onset of fluid heating, which requires either an active system cooling or a time-consuming wait for thermal equilibrium.

To demonstrate the advantages of the choice of hydraulic parameter over hydraulic resistance as the operating parameter, we again can examine the performance of a case study pump. Once again, the pump is of a fixed wear plate configuration, but rather than enforcing constant fluid temperature by cooling the pump after each data point, data is collected in four sets, with minimal delay between sets. During each set, the action of the pump heats the fluid, leading to a gradual increase in fluid temperature of approximately 10 °C. A comparison of two plots is shown in Fig. 15, one with the volumetric efficiency as a function of the hydraulic resistance and the other as a function of the hydraulic parameter. The plot with the volumetric efficiency as a function of the hydraulic resistance shows four distinct curves, which demonstrates the impact of variable fluid viscosity on apparent gear resistance. In contrast, plotting the volumetric efficiency against the hydraulic parameter shows a single curve across all four data sets.

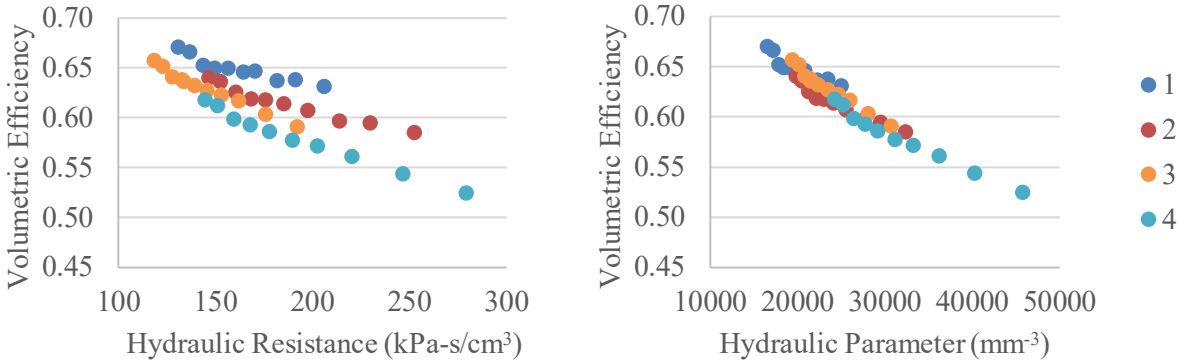

Fig. 15—Comparison of volumetric efficiency plotted as either a function of hydraulic resistance (*left*) or the hydraulic parameter (*right*). Fluid temperatures for each sample set are: Set 1: 28.5 °C to 30 °C, Set 2: 30 °C to 32 °C, Set 3: 37 °C to 38 °C, and Set 4: 38 °C to 39 °C.

1.4.5 *Viscosity Calculations*

While useful for facilitating performance comparisons, use of the hydraulic parameter does require knowledge of the viscosity over time. It is not generally feasible to measure fluid viscosity itself, but rather to calculate it based on the measured fluid temperature. Most fluid specifications report viscosity in terms of the kinematic viscosity, so calculating dynamic viscosity will require knowledge of both the kinematic viscosity and the fluid density as a function of temperature. One of the more robust models for kinematic viscosity is the Walther formula,

$$log[log(\nu + \lambda)] = A - Blog(T)\,, \tag{1.18}$$

where the viscosity, ν, is given in cSt, T is the absolute temperature of the fluid, and A, B, and λ are empirical constants. For hydrocarbons, λ is generally given a value of 0.7, while A and B can be calculated from two known viscosities (ν_1 and ν_2) at two known temperatures (T_1 and T_2).

$$A = log[log(\nu_1 + \lambda)] + Blog(T_1) \tag{1.19}$$

$$B = \frac{log[log(\nu_1 + \lambda)] - log[log(\nu_2 + \lambda)]}{log(T_2) - log(T_1)} \tag{1.20}$$

Commercial fluids often give multiple viscosities at different temperatures within the specifications list, which is sufficient to solve for A and B. The kinematic viscosity in cSt at any temperature is then given by the following:

$$\nu = 10^{10^{A - Blog(T)}} - 0.7\,. \tag{1.21}$$

The density for hydrocarbon fluids generally varies linearly with temperature and is modeled from density plots as

$$\rho = \rho_0 - mT\,, \tag{1.22}$$

where ρ_0 is the density of the fluid at 0 °C, T is the fluid temperature in °C, and m is the temperature sensitivity that varies with ρ_0.

$$m = m_0 - n\rho_0 \tag{1.23}$$

The values of m_0 and n can be determined using a linear regression of oil density curve slopes. Hydraulic fluid densities are not necessarily specified at 0 °C, therefore ρ_0 must be calculated from the given reference density, ρ_{ref}, and temperature, T_{ref}.

$$\rho_0 = \frac{\rho_{ref} + m_0 T_{ref}}{1 + nT_{ref}} \tag{1.24}$$

The dynamic viscosity can then be calculated from the kinematic viscosity and the density:

$$\mu = \rho\nu\,. \tag{1.25}$$

1.5 Introduction to Gears

Thus far, we have not yet discussed the eponymous gears, themselves, as we have been building a framework to discuss gear pump design. While an in-depth discussion on gear design occurs in Section 3, it is worth giving a brief overview of the most common gear architecture used in external gear pumps to establish a foundation of nomenclature. This will allow for an examination of several key phenomena within gear pumps with some context of the structure of the gears used and the relevant geometric parameters.

The type of gear most commonly used in external gear pumps is the involute spur gear; "spur" denotes a gear with a constant axial profile (i.e., no twist), and "involute" denotes the defining curve for the cross-sectional tooth profile. An involute curve is one defined in reference to another curve, in this case, a circle concentric with the axis of the gear, called the base circle. The solid core of the gear extends outward to the root circle, beyond which a set number of teeth, n_t, extend. The pitch circle is a reference circle that is used to standardize and define the tooth parameters, and the pitch circles of two meshing gears should intersect only at one point. The addendum circle defines the outer diameter of the gear. The outer surface at the edge of the gear teeth is called the top land, and the surface along the root circle between teeth is the bottom land. The side of of each tooth where meshing contact occurs is called the tooth face. Gears often will have a fillet between the bottom land and the tooth face. The "height" of the cylindrical body that the gear creates is known as the land width, or simply the width. Gears used in pumps generally will have a central bore for shaft mounting and potentially a shaft keyway to prevent slippage. The defining features of an involute spur gear are shown in Fig. 16.

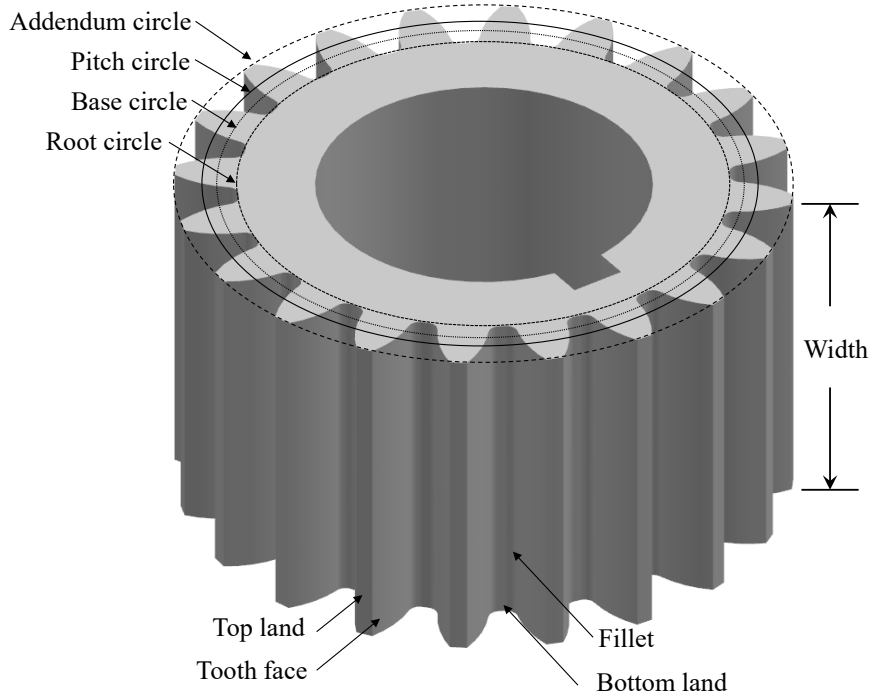

Fig. 16—Defining features of an involute spur gear

2. DESIGN CONSIDERATIONS

With a basic understanding of the makeup of an external gear pump established, we now can begin to address the design of the pump itself. There are two overall questions that an engineer must ask when starting the design process for a device:

1. What objective am I trying to achieve with this device, i.e., what metrics and values thereof define success?

2. What are the phenomena of interest that will inhibit the performance of this device, and how can I account for them within the design?

The first question was addressed in Section 0, where we discussed performance metrics and a methodology for comparison. The specific application in mind for the pump also will define the absolute performance requirements in terms of flow rate and pressure. In this chapter, we will delve into the second question of key phenomena during pump operation that will impact performance and drive the design process. These include leakage, meshing dynamics, pressure spikes, cavitation, and friction.

2.1 Leakage

One of the most critical elements of a gear pump is the degree to which it is able to form static and dynamic seals between interfaces and to limit leakage of the working fluid. Leakage can be categorized into two types: external leakage, where fluid within the pump leaks to outside the casing, and internal leakage, where fluid leaks from regions of higher pressure in the pump to regions of lower pressure. External leakage is addressed primarily through seal design between casing components and at the drive shaft, and is more a task of good design practice and knowledge of the standards for seal gland design rather than a design challenge specific to gear pumps. In contrast, addressing internal leakage requires a thorough

understanding of the pump dynamics in play. The most common form of internal leakage in gear pumps is fluid from the discharge region leaking around the gears to the suction region. This leakage can be divided further between radial leakage around the tips of the gears and axial leakage across the axial face of the gears. The total leakage around the gears, Q_L, can be written as the sum of the radial and axial leakages, Q_r and Q_z, respectively.

$$Q_L = Q_r + Q_z \tag{2.1}$$

The volumetric efficiency of a pump can be written in terms of the ideal flow rate and the leakage flow rate.

$$\eta_v = 1 - \frac{Q_L}{Q_{ideal}} \tag{2.2}$$

Thus, limiting these leakages through the design of the pump components and assembly tolerances while tailoring these to the nominal operating conditions is the biggest challenge the designer faces.

2.1.1 *Radial Leakage*

The radial leakage consists of successive fluid leakage from one intertooth volume to the next from the discharge region to the suction region. This can be approximated as 2D Couette flow across each tooth tip driven by the pressure differential between adjacent intertooth volumes, as has been done in other research [2,4,5]. Since the thin gap between the tooth tip and the casing wall is the site of the primary pressure drop, the radial leakage across each gear can be modeled as a single, thin-film flow with length of the summation of the tooth tip thickness of half of the gear teeth. While the geometry is represented more accurately in cylindrical polar coordinates, for the sake of simplified mathematics, it is formulated in Cartesian coordinates. Considering that the gap thickness is much smaller than the radius of the gear, this is a reasonable simplification. The approximated geometry used is illustrated in Fig. 17.

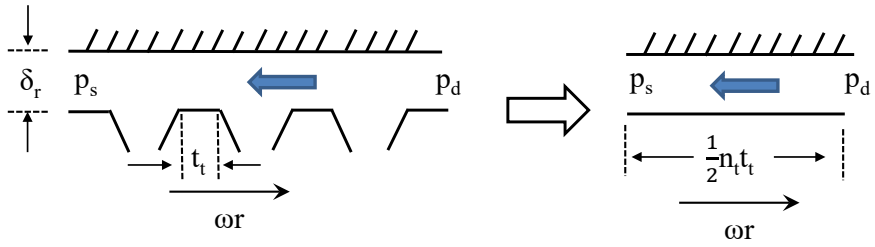

Fig. 17—Approximation of radial gear leakage across multiple teeth as Couette flow

The velocity profile of Couette flow for such a simplified system can be derived readily from the Navier-Stokes equations and integrated along the radial gap to calculate the radial leakage flow. Note that the factor of two corresponds to two gears with radial leakage.

$$u(y) = \frac{\Delta p y^2}{\mu n_t t_t} - \left(\frac{\Delta p \delta_r}{\mu n_t t_t} + \frac{r_t \omega}{\delta_r}\right) y + r_t \omega$$

$$Q_r = 2w \int_0^{\delta_r} u(y) dy$$

$$Q_r = \frac{\Delta p \delta_r^3 w}{3\mu n_t t_t} - w \delta_r r_t \omega \tag{2.3}$$

In Eq. (2.3), Δp is the pressure drop from discharge to suction, δ_r is the radial clearance between the gear and the casing, w is the width of the gear (i.e., the length of the gear along the axis of rotation), μ is the fluid viscosity, n_t is the number of gear teeth, t_t is the thickness of the gear teeth tip, r_t is the radius of the gear at the tip, and ω is the angular speed of the gear. It should be noted that mathematically, the integration results in a negative Q_r, but this denotes flow from discharge to suction, which, by convention, is chosen to be a positive leakage, thus Eq. (2.3) is the negative of the actual integration result.

An interesting result of this basic model is that the radial gear leakage is the superposition of a pressure term and a speed term. In particular, it is worth noting that the motion of the gears acts to suppress the leakage flow. Comparing Eq. (2.3) to Eq. (1.13), we can write the radial gear resistance and reluctance as

$$R_{G,r} = \frac{3\mu n_t t_t}{\delta_r^3 w} \tag{2.4}$$

$$\mathbb{R}_{G,r} = -\frac{1}{2\pi\delta_r w r_t}. \tag{2.5}$$

Note that Eqs. (2.4) and (2.5) are consistent with the hydraulic resistance method formulation in Section 1.4, namely that the gear resistance is a viscous term that is linear with viscosity, and the gear reluctance is inviscid. The inviscid gear resistance for use with the hydraulic parameter can be found readily by dividing out the viscosity.

2.1.2 *Axial Leakage*

Analytically modeling axial leakage is more complex, as the flow through the gap between the face of the gear and the casing is not two-dimensional due to the rotating gear boundary condition. Such a system requires numerical computation to solve the Reynolds equation for the pressure distribution, and thus does not fit as a convenient analytical tool for initial design or performance analysis. However, an initial approximation can be made that the axial leakage will be a superposition of a pressure-driven flow and a gear speed-driven flow akin to the radial leakage.

$$Q_z = f(\Delta p) + g(\omega)$$

For the pressure term, the removal of the speed dependency returns the problem to a 2D Reynolds equation across two semicircles (the half of each gear along the center of the pump). The leakage then can be integrated along the width of the semicircle:

$$f(\Delta p) = 2\int_0^{r_p} \frac{\Delta p \delta_z^3}{12\mu}\frac{w}{L}dx$$

$$f(\Delta p) = 2\frac{\Delta p \delta_z^3}{12\mu}\int_0^{r_p} \frac{dx}{2\sqrt{r_p^2 - x^2}}$$

$$f(\Delta p) = 2\frac{\Delta p \delta_z^3}{12\mu}\arcsin\left(\frac{x}{r}\right)\Big|_0^r$$

$$f(\Delta p) = \frac{\Delta p \delta_z^3 \pi}{12\mu}$$

$$R_{G,z} = \frac{12\mu}{\delta_z^3 \pi}. \tag{2.6}$$

Regardless of the specifics of the model for gear resistance, it is reasonable to assume that the pressure term always will have a linear dependency with the fluid viscosity, while the speed term will be independent of viscosity, similar to radial leakage.

2.1.3 *External Leakage*

External leakage is flow from within the pump through some finite gap to the ambient conditions. We can conceptualize a simple external leakage flow as planar Poiseuille flow, illustrated in Fig. 18, where there is a gap between two surfaces of constant height h, and length L. On one side of this interface is ambient pressure, p_a, and on the other side is the fluid pressurized to the discharge pressure, p_d. The interface of the two surfaces has some perimeter P that can be approximated as a linear width out of the plane of Fig. 18.

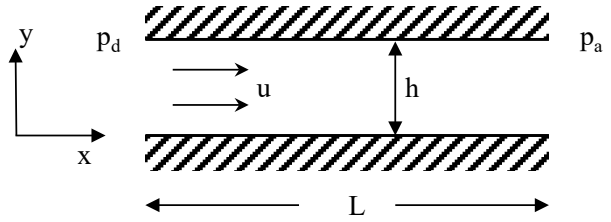

Fig. 18—Leakage problem formulation as planar Poiseuille flow

Solving the 2D Cartesian formulation of the Navier-Stokes equations assuming steady state, fully developed incompressible flow in x yields the solution for the velocity, u.

$$u(y) = \frac{1}{2\mu}\frac{\partial p}{\partial x} y(y-h) \tag{2.7}$$

Integrating the velocity profile across the gap and multiplying by the interface perimeter yields the leakage flow rate through the gap:

$$\begin{aligned} Q_L &= P\int_0^h u(y)dy \\ &= \frac{1}{2\mu}\frac{\partial p}{\partial x} P\int_0^h y(y-h)dy \\ &= \frac{1}{2\mu}\frac{(p_a - p_d)}{L} P\left[\frac{y^3}{3} - \frac{y^2 h}{2}\right]_0^h \\ Q_L &= \frac{P(p_d - p_a)h^3}{12\mu L}. \end{aligned} \tag{2.8}$$

Similar to internal leakage, the leakage flow rate is a function of the cubic power of the gap height, and thus even the channels created by surface roughness can allow for a non-negligible leakage flow for high-

pressure pumps. This high sensitivity to the gap height drives the design goal for minimizing the gap size between components and establishing minimum surface finish requirements between interfaces.

Compared to the situation with internal leakage, resolving external leakage is relatively straightforward, as it arises at either static interfaces along the casing or at dynamic interfaces such as the drive shaft, where elastomeric seals can be used to minimize the effective gap height. For casing seals, this becomes a static seal design issue whereby usually an elastomer seal is compressed into a gland and grips the two sealing surfaces to restrict fluid flow. For shaft seals, a two-part dynamic seal is statically sealed against a casing gland and dynamically seals against the shaft itself. In both cases, the usual dominant factor is the discharge pressure within the pump driving fluid through a finite gap that is opposed by the contact pressure between the seal and the sealing surface. It is this pressure balance that dictates the efficacy of a seal and the degree of leakage.

The basic premise for static seals is the presence of an elastomer material placed at a junction between the two surfaces to be sealed, usually within some groove or housing called a gland. The compression of the elastomer within the gland during assembly, referred to as the squeeze, imparts an initial contact stress between the seal and the sealing surfaces. During operation, pressurized fluid presses the seal against the wall of the gland, further deforming the seal and providing additional contact stress against the sealing surfaces. If the contact stress between the seal and the sealed surfaces exceeds the fluid pressure on the seal, a static seal is formed and leakage is restricted. However, excessive stress on the seal can induce the elastomer to extrude into the gap between the surfaces, leading to plastic deformation and potential failure of the seal. Illustrated in Fig. 19 is how a simple static seal transitions from assembly to operational conditions, and finally to extrusion under excessive pressure loads. The design of these glands and the selection of the corresponding elastomer seal are well established and readily available through seal vendor guides [6,7].

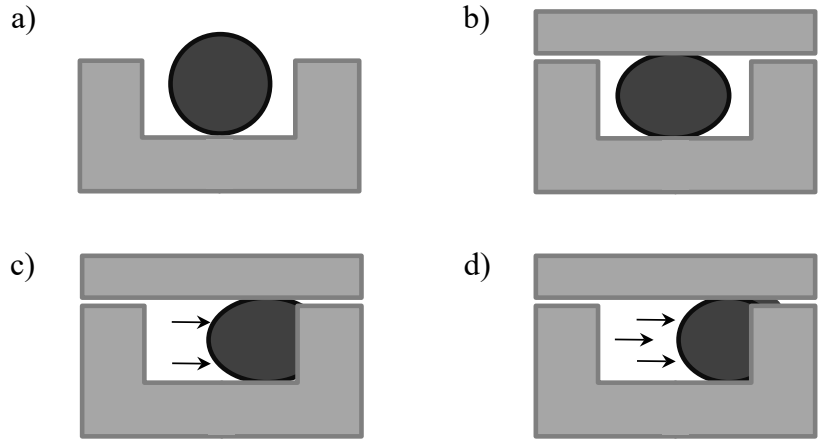

Fig. 19—Cross section of a simple static seal: a) uncompressed elastomer prior to assembly, b) elastomer with initial squeeze from assembly, c) seal under further compression from pressurized fluid, and d) overcompressed seal undergoing extrusion

Dynamic external seals are complicated by constant or intermittent movement of one of the sealing surfaces due to operation of the device, in this case, usually the drive shaft. A gland design similar to what is used in static seals anchors the seal and usually involves the compression of a relatively broad elastomer. The part of the seal in contact with the shaft is often a harder material focused to a sharp edge or a rounded lip to improve resistance to abrasion while maintaining a high contact pressure on the shaft over a limited area. For high-pressure applications, the seal may be designed to utilize fluid pressure or an internal spring

to increase compression of the seal against the shaft. Selection of dynamic seals and the design of the mating gland vary greatly by application and are set by seal vendor guidelines [8].

2.2 Meshing Dynamics: Output Ripple

The beating heart of the gear pump is the pair of external gears; the unmeshing of the teeth creates the vacuum that draws in fluid from the suction side of the pump and the meshing of the teeth compresses the fluid and drives it into the discharge region. As such, most of the dynamic phenomena within gear pumps occur within or nearby the meshing zone, such as flow and pressure ripple in the pump output or pressure spikes within the mesh. This aspect of gear pumps is still an active field of research [9-12], and it is important that a designer understands how these phenomena arise in order to properly account for them in a pump design.

One thing to remember when considering meshing dynamics is that gear pumps are primarily used with incompressible fluids that are characterized by a large bulk modulus, or resistance to compression. As a volume of fluid is compressed, the pressure within the volume is impacted by the net flow of fluid into the volume and the rate of change of that volume. The rate of change of pressure is related to the bulk modulus, β, the volume of the fluid, V, and the net flow into the volume, Q.

$$\frac{dp}{dt} = \frac{\beta}{V}\left(Q - \frac{dV}{dt}\right) \tag{2.9}$$

2.2.1 *Model Formulation*

Equation (2.9) shows that the output flow rate, and by extension, the discharge pressure, is a product of the time rate of change of the volume available for the fluid within the discharge region. Considering the standard spur involute gear, which is commonly used in gear pumps, the time rate of change in intertooth volume is not constant throughout the mesh at constant pump speed. Additionally, the mesh of one set of teeth may not yet be completed when the next set of teeth begins to mesh, particularly for gears with constant contact (e.g. low backlash involute spur gears). The superposition of multiple intertooth volumes undergoing the mesh at different phases results in a ripple in the output flow rate as a function of angular position of the gears, and thus time. A basic analytical model can be made in order to better understand the impact of the gear mesh on output ripple, based on work done by Manring [9]. Taking the conceptualized pump geometry from Fig. 3, we can define a control volume for the discharge region of the pump that is bounded by the pump casing and the root circles of the gears, shown in Fig. 20. For the sake of simplicity, the boundary of the control volume at the mesh is chosen to be along the line connecting the axes of rotation.

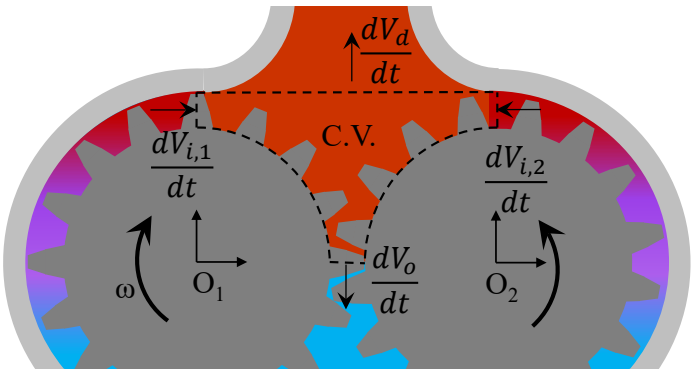

Fig. 20—Discharge region control volume

The drive gear brings into the control volume a combination of solid gear material and fluid, where the composition of this net volume addition varies with the angular position of the gear. The output of the control volume is either through the mesh, where a combination of gear material and fluid exits, or through the pump discharge. We can then rewrite Eq. (2.9) in terms of the flow rate out the discharge, Q_d, and the rate of volume change of the discharge region, V_d.

$$\frac{dp_d}{dt} = \frac{\beta}{V}\left(-Q_d + \frac{dV_d}{dt}\right) \tag{2.10}$$

Integration of Eq. (2.10) over time and concurrently calculating the resulting discharge flow rate as a function of discharge pressure and system flow resistance will yield both pressure and flow ripple, but this is a nontrivial task. A simplifying approximation is to assume that the discharge flow rate is equal to the rate of change of the discharge region, which is reasonable for high-bulk-modulus fluids. This essentially treats the discharge region as constant pressure averaged over time. The time rate of volume of material into the control volume is therefore balanced by the time rate of volume out of the control volume, which determines flow rate out of the discharge as a function of time.

$$\frac{dV_d}{dt} = \frac{dV_{i,1}}{dt} + \frac{dV_{i,2}}{dt} - \frac{dV_o}{dt} \tag{2.11}$$

It is important here to establish the parameters to define the problem, as this is a rotational system with cyclical behavior. Thus, we can utilize two polar coordinate reference frames centered on each gear axis, called O_1 and O_2. The origins of these two reference frames are separated by the sum of the pitch radii of the gears. Any point within the control volume can be defined with respect to either polar reference frame by a radius and an angular position, (r_1, θ_1) or (r_2, θ_2). The rotation of the gears superimposes an angular progression of the gears on these reference frames such that the position of a fixed point on a gear tooth within O_1 or O_2 is a function of the angular progression of the corresponding gear, ψ_i. In essence, this is taking a gear profile that is a function of θ and rotating it by ψ. The two reference frames and how a point on a gear profile can be defined in either are illustrated in Fig. 21.

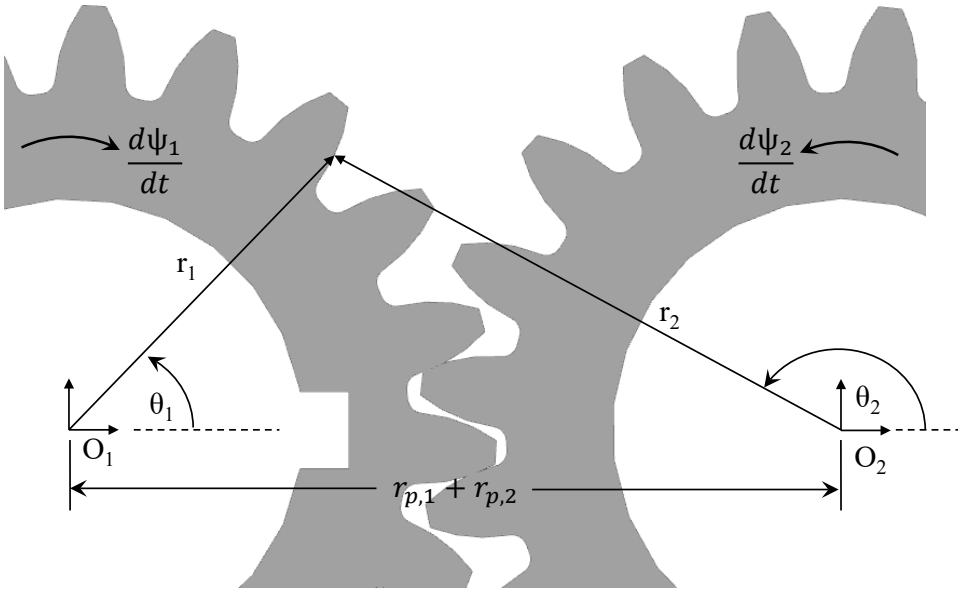

Fig. 21—Reference frames for defining a point on the gear profile

The volume of material of entering the control volume from the drive gear, designated by the subscript one, is a combination of the gear material and fluid. Assuming a spur gear where the tooth profile is constant along the axis of the gear, the differential volume added to the control volume can be estimated as a differential wedge.

$$\frac{dV_{i,1}}{dt} = \frac{dV_{iG,1}}{dt} + \frac{dV_{if,1}}{dt}$$

$$= \frac{1}{2} w\left(r_1^2 - r_{r,1}^2\right) \frac{d\psi_1}{dt} + \frac{1}{2} w\left(r_{a,1}^2 - r_1^2\right) \frac{d\psi_1}{dt}$$

$$\frac{dV_{i,1}}{dt} = \frac{1}{2} w\left(r_{a,1}^2 - r_{r,1}^2\right) \frac{d\psi_1}{dt} \tag{2.12}$$

In the above equation, r_a is the radius of the addendum circle, r_r is the radius of the root circle, and w is the width of the gear. By assuming an incompressible fluid, we can treat this input volume rate as constant. Similarly, the rate of volume addition by the idle gear is constant with respect to angular position regardless of the makeup of that volume addition.

$$\frac{dV_{i,2}}{dt} = \frac{1}{2} w\left(r_{a,2}^2 - r_{r,2}^2\right) \frac{d\psi_2}{dt} \tag{2.13}$$

The volume of material exiting the control volume is the combination of the gear material from both gears and any remaining fluid between the gear teeth.

$$\frac{dV_o}{dt} = \frac{dV_{oG,1}}{dt} + \frac{dV_{oG,2}}{dt} + \frac{dV_{of}}{dt} \tag{2.14}$$

The volume rate of each gear exiting the control volume is calculated in the same manner as Eq. (2.12).

$$\frac{dV_{oG,1}}{dt} = \frac{1}{2} w\left(r_1^2 - r_{r,1}^2\right) \frac{d\psi_1}{dt} \tag{2.15}$$

$$\frac{dV_{oG,2}}{dt} = \frac{1}{2} w\left(r_2^2 - r_{r,2}^2\right) \frac{d\psi_2}{dt} \tag{2.16}$$

From the fundamental law of gearing, we know that the two gears must have the same tangential velocity at the point of contact; from this, we can relate the two gear angular speeds.

$$r_{p1} \frac{d\psi_1}{dt} = r_{p2} \frac{d\psi_2}{dt} \tag{2.17}$$

The volumetric rate of liquid leaving the control volume through the mesh is somewhat more complicated, as the velocity of the fluid orthogonal to the boundary of the control volume is dependent on position relative to the two gears. To simplify, we can assume a linear relationship between the normal exit velocity out of the control volume (u_o) and the radial position relative to the closest gear (based on the position of the pitch point).

$$u_o(r) = \begin{cases} r\dfrac{d\psi_1}{dt} & if\ r \le r_{p1} \\ (r_{p1} + r_{p2} - r)\dfrac{d\psi_2}{dt} & if\ r > r_{p1} \end{cases} \tag{2.18}$$

The assumed form of the velocity profile is illustrated in Fig. 22.

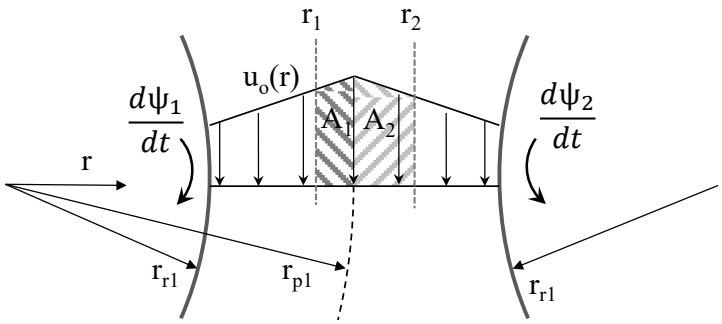

Fig. 22—Fluid velocity profile out of the control volume at the mesh

If we define the velocity profile as a function of radial position from the axis of rotation of the drive gear (r) and the radial position of the edge of the drive and idle gear profiles (r_1 and r_2, respectively), we can formulate the fluid volumetric flow rate out of the mesh as a definite integral of the velocity profile:

$$\frac{dV_{of}}{dt} = w\int_{r_1}^{r_2} u_o dr\,. \tag{2.19}$$

Given the shape of the velocity profile, we can think of the definite integral as the sum of the areas of two trapezoids, A_1 and A_2. If $r_2 \le r_{p1}$, then $A_2 = 0$ and

$$\begin{aligned} \frac{dV_{of}}{dt} &= w\int_{r_1}^{r_2} r\frac{d\psi_1}{dt}dr \\ &= \frac{1}{2}w\,r^2\frac{d\psi_1}{dt}\bigg|_{r_1}^{r_2} \\ \frac{dV_{of}}{dt} &= \frac{1}{2}w(r_2^2 - r_1^2)\frac{d\psi_1}{dt}\,. \end{aligned} \tag{2.20}$$

If $r_1 > r_{p1}$, then $A_1 = 0$ and

$$\begin{aligned} \frac{dV_{of}}{dt} &= w\int_{r_1}^{r_2}(r_{p1} + r_{p2} - r)\frac{d\psi_2}{dt}dr \\ &= w\left[(r_{p1} + r_{p2})r\frac{d\psi_2}{dt} - \frac{1}{2}r^2\frac{d\psi_2}{dt}\right]_{r_1}^{r_2} \\ \frac{dV_{of}}{dt} &= w\left[(r_{p1} + r_{p2})(r_2 - r_1) - \frac{1}{2}(r_2^2 - r_1^2)\right]\frac{d\psi_2}{dt}\,. \end{aligned} \tag{2.21}$$

Finally, if $r_2 > r_p$ and $r_1 < r_p$, then the integral is the sum of A_1 and A_2.

$$\frac{dV_{of}}{dt} = A_1 + A_2$$
$$\frac{dV_{of}}{dt} = \frac{1}{2}w\left(r_{p1}^2 - r_1^2\right)\frac{d\psi_1}{dt} + w\left[\left(r_{p1} + r_{p2}\right)\left(r_2 - r_{p1}\right) - \frac{1}{2}\left(r_2^2 - r_{p1}^2\right)\right]\frac{d\psi_2}{dt} \qquad (2.22)$$

2.2.2 *Gear Alignment Determination*

Solving for the discharge flow rate in Eq. (2.11) therefore requires the gear profile of each gear as a function of angular position $r_i(\theta)$ and a reference point to establish the relative angular orientation of the two gears. As will be discussed in Section 3, the involute curve is defined parametrically, which makes it difficult to define the radial position of the curve as a function of angular position. For the purposes of this exercise, the involute curve can be approximated by using a second-order polynomial curve fit. As for determining the relative orientation of the gears, we can use the locus of all contact points during the mesh, called the line of action. A close-up of the gear mesh and the associated geometry at the start of a meshing cycle is shown in Fig. 23. The subscript "s" denotes a parameter value at the start of the mesh.

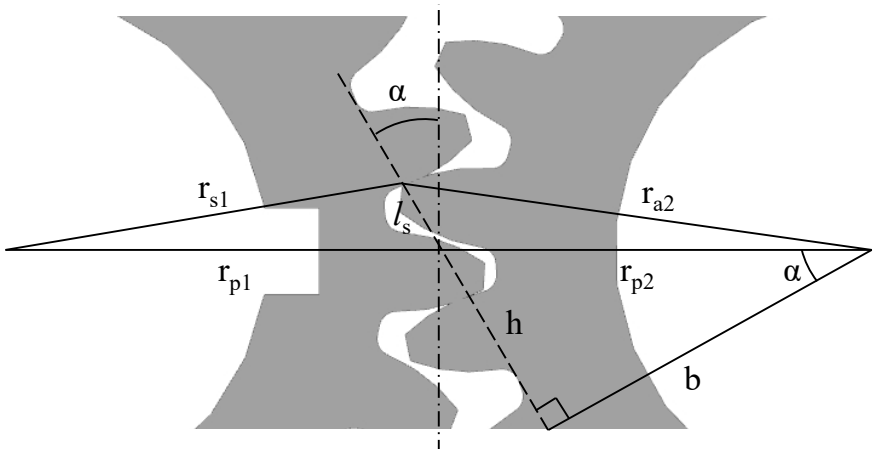

Fig. 23—Gear geometry at the start of the mesh

For an involute gear, the line of action is at an angle equal to the pressure angle, with the starting point of the mesh coincident with the addendum circle of the idle gear. The length of action is the line segment connecting the point of contact along the line of action to the midpoint between the two gears; at the start of the mesh, this value is at its maximum, l_s, which can be determined from geometry.

$$r_{a2}^2 = b^2 + (h + l_s)^2$$
$$l_s = \sqrt{r_{a2}^2 - b^2} - h$$
$$l_s = \sqrt{r_{a2}^2 - r_{p2}^2\cos^2\alpha} - r_{p2}\sin\alpha \qquad (2.23)$$

The radial position that a tooth from the driving gear first contacts a mating tooth then can be calculated using the Law of Cosines.

$$r_{s1}^2 = r_{p1}^2 + l_s^2 - 2r_{p1}l_s\cos\left(\frac{\pi}{2} - \alpha\right)$$
$$r_{s1} = \sqrt{r_{p1}^2 + l_s^2 - 2r_{p1}l_s\sin\alpha} \tag{2.24}$$

The values of r_{s1}, r_{a2}, and the position of the start of the mesh then can be used to determine relative orientation of the two gears. However, we first must consider how the profile of the gear is defined in polar coordinates.

Ideally, the tooth profile is defined as a radial position, r, as a function of angular position, θ. Most gears used will have a multipart profile, with different curves defining the tooth face, the top land, and the bottom land. As an example, an involute tooth profile with no fillet between the tooth flank and the bottom land can be defined by four curves with junctions at four defined angular positions, θ_{j1} through θ_{j4}.

$$r(\theta) = \begin{cases} r_{inv,f}(\theta) & \text{for } \theta \in [0, \theta_{j1}] \\ r_a & \text{for } \theta \in (\theta_{j1}, \theta_{j2}) \\ r_{inv,r}(\theta) & \text{for } \theta \in [\theta_{j2}, \theta_{j3}] \\ r_r & \text{for } \theta \in (\theta_{j3}, \theta_{j4}) \end{cases} \tag{2.25}$$

In the above equation, $r_{inv,f}$ and $r_{inv,r}$ define the involute curve radial position of the forward and reverse tooth faces, respectively. The cyclical nature of the gears means that this profile is repeated a number of times equal to the number of teeth on the gear, n_t. The entire gear profile across a full 2π radian sweep then can be written in terms of a cyclical angular position, φ.

$$r(\varphi) = \begin{cases} r_{inv,f}(\varphi) & \text{for } \varphi \in [0, \varphi_{j1}] \\ r_a & \text{for } \varphi \in (\varphi_{j1}, \varphi_{j2}) \\ r_{inv,r}(\varphi) & \text{for } \varphi \in [\varphi_{j2}, \varphi_{j3}] \\ r_r & \text{for } \varphi \in (\varphi_{j3}, \varphi_{j4}) \end{cases} \tag{2.26}$$
$$\varphi = \mathrm{mod}\left(\theta, \frac{2\pi}{n_t}\right) \tag{2.27}$$

However, the formulation of the cyclical angular position in Eq. (2.27) neither accounts for gear orientation required to fulfill meshing requirements nor accounts for gear rotation relative to the absolute reference frame, *O1*. We can account for both by including the angular progression of the gear introduced earlier, as well as a reference angle, θ_{ref}. Equation (2.27) then can be rewritten as

$$\varphi_1 = \mathrm{mod}\left(\theta + \theta_{ref,1} + \psi_1, \frac{2\pi}{n_t}\right) \tag{2.28}$$
$$\varphi_2 = \mathrm{mod}\left(\theta + \theta_{ref,2} + \psi_2, \frac{2\pi}{n_t}\right). \tag{2.29}$$

The angular progression of the two gears is related through the fundamental law of gearing.

$$\psi_2 = -\frac{r_{p1}}{r_{p2}}\psi_1 \tag{2.30}$$

For the drive gear, the value of $\theta_{ref,1}$ is chosen such that the value of $r_1(\varphi)$ is equal to r_{1s} at $\theta = \theta_{SOA,1}$,

$$\theta_{SOA,1} = \cos^{-1}\left(\frac{r_{s1}^2 + r_{p1}^2 - l_s^2}{2r_{s1}r_{p1}}\right), \tag{2.31}$$

and for the idle gear, the value of $\theta_{ref,2}$ is chosen such that $r_2(\varphi)$ is equal to r_{a2} for the first time at $\theta = \theta_{SOA,2}$.

$$\theta_{SOA,2} = \pi - \cos^{-1}\left(\frac{r_{a2}^2 + r_{p2}^2 - l_s^2}{2r_{a2}r_{p2}}\right) \tag{2.32}$$

2.2.3 *Case Study*

To make a simple case study, we can examine the flow ripple of a simple gear pump utilizing two identical involute spur gears. The gear geometry used is given in Table 1, with the gear profiles plotted in Fig. 24 oriented at the start of the mesh.

Table 1—Assumed Gear Geometry

Parameter	Value
r_p	7.94 mm
r_r	6.97 mm
r_a	8.68 mm
r_b	7.46 mm
w	5.46 mm
α	20°

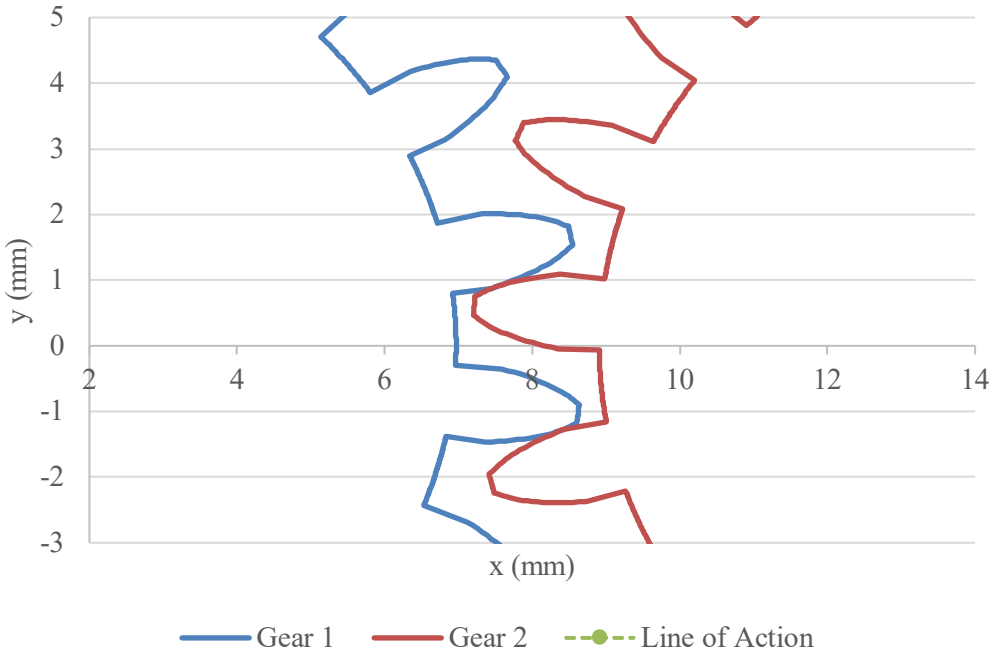

Fig. 24—Plot of both gear profiles at the meshing region and the line of action

If we now assume that the pump is operated at 2,000 rpm, we can calculate the discharge flow rate as a function of time, shown in Fig. 25. The variable widths of the pulsations reflect the noticeable backlash of the gears. Key findings from Manring's work are that the amplitude of these pulsations is driven primarily by the number of teeth within the drive gear, and, to a lesser extent, the number of teeth on the idle gear [9]. This reflects the overall paradigm of the gear pump, in which the gear teeth entering the discharge region create a cyclic compression of volume available for fluid to occupy, analogous to the compression of a piston in a cylinder. As the number of teeth increases for a given gear pitch diameter, the size of the teeth decreases (as covered in Section 3.1), which results in higher-frequency, lower-amplitude compression of the discharge region. The higher the frequency of the pulsation, the more the relaxation from one compression cycle blends into the initiation of the next compression cycle. It is worth noting these pulsation frequencies will be high; as an example, at 1,800 rpm, a gear with 20 teeth will have a pulsation frequency of 600 Hz.

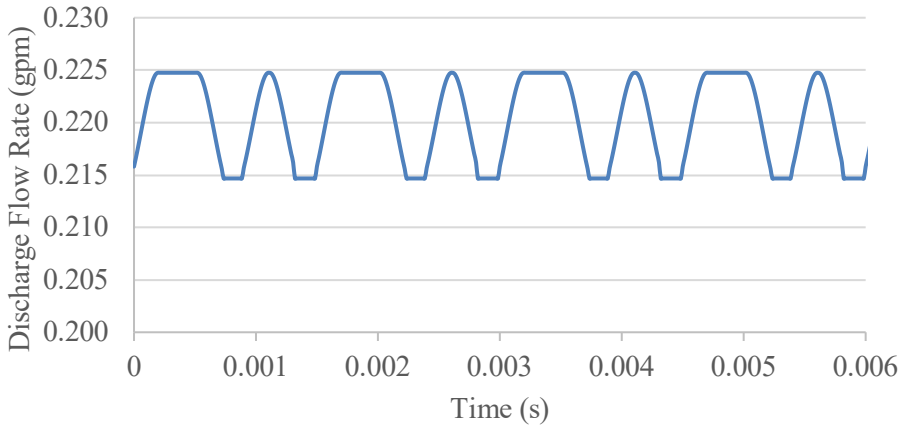

Fig. 25—Discharge flow ripple as a function of time

2.3 **Meshing Dynamics: Pressure Spikes**

As mentioned already, the primary driver of fluid pumping in a gear pump is the creation and elimination of fluid volumes between gear teeth. The creation of these pockets at the unmeshing of the gears creates the suction that draws fluid into the gear teeth, and the meshing of the teeth compacts the fluid and expels it into the discharge. In an idealized pump, the meshing process fully expels all fluid into the discharge; however, for a real involute gear pump, the addendum of one tooth does not perfectly mate against the dedendum of the other, pictured in Fig. 26. This is intended to provide sufficient clearance to allow some variation in the offset between the gears; while optimally, they are separated by the pitch diameter, the backlash allows some error without damaging the gears. While useful for power transmission, in pumping applications, this results in some fluid remaining, creating some inherent leakage back to suction. A more important impact is that when the meshing teeth come into contact along the line of force, the remaining fluid is effectively trapped during the compression process.

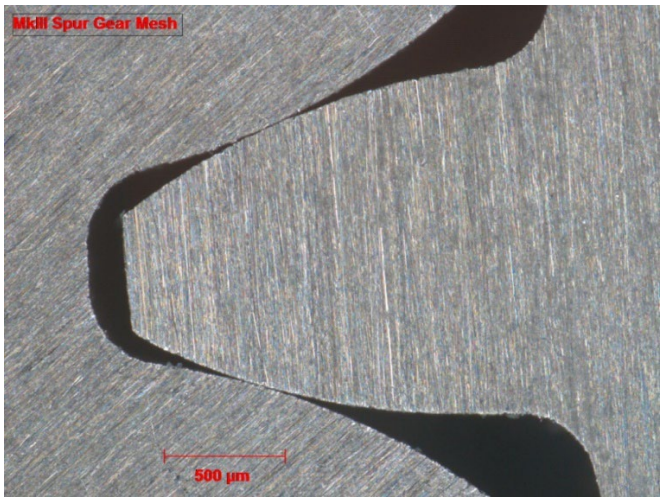

Fig. 26—Microscope imaging of a pocket formed during meshing of two involute gears

As the teeth continue to mesh, the pressure of the trapped fluid must rise according to Eq. (2.9). The flow rate out of the pocket through an orifice is

$$Q = C_d A \sqrt{\frac{2\Delta p}{\rho}}, \tag{2.33}$$

where C_d is the discharge coefficient for the orifice, A is the area of the orifice, Δp is the pressure drop across the orifice, and ρ is the density of the fluid.

Conceptually, as the volume of the pocket decreases, the pressure of the fluid will increase until it is sufficient to force a flow rate through all available orifices to match the volume rate of change of the meshing action. For a pump with very tight clearances between components, as is desired in high-pressure pumps, this can lead to considerable pressures in the meshing action. As an example of the impacts of such pressure spikes, Fig. 27 shows microscope imaging of gear teeth from a 500-psi gear pump. The highlighted indentations in the teeth are consistent with a radially inward force acting on the tooth tips during operation. These pressure spikes also contribute to the running noise of the pump and can increase flow ripples propagated into the hydraulic system being supplied [10]. For these reasons, it is worthwhile to reduce the pressure spikes encountered during operation.

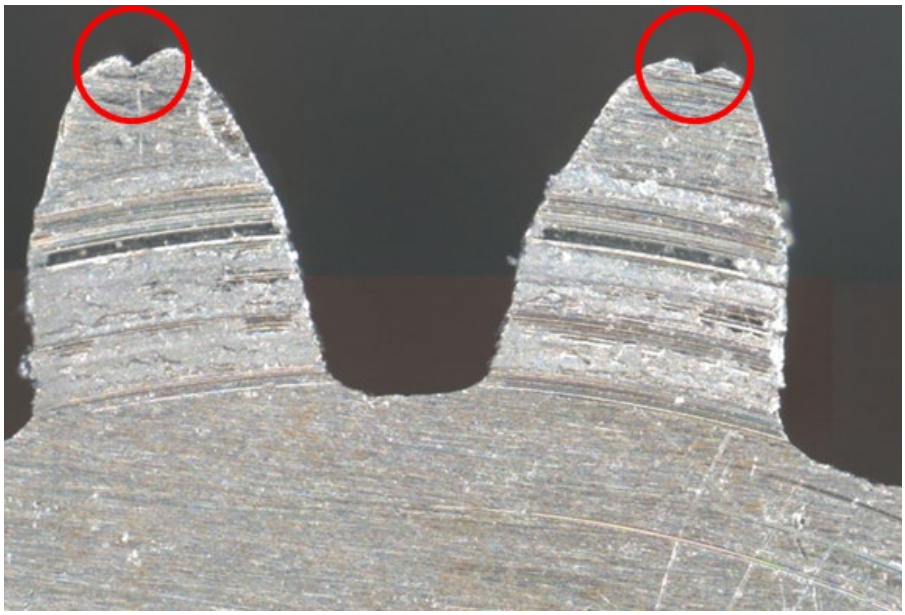

Fig. 27—Tip damage in tooth top lands due to pressure spikes in gear meshing

There are a few options for reducing fluid entrapment and the resulting pressure spikes, and both involve increasing the flow area to the trapped fluid, which would reduce the required pressure rise to drive the necessary flow rate. The first is to use a gear type other than spur gears, such as a helical or herringbone gear. The gradual tooth engagement would mean that the fluid begins to be expelled before the tooth is fully engaged, increasing the exit area of the pocket and allowing the majority of the pocket volume to be expelled before the remaining fluid is trapped. While this adds to the trade for gear type selection, the other factors at hand (namely, cost and difficulty in acquiring high-precision helical or herringbone gears of the required size) make this option potentially difficult. The other option is to provide a larger exit area for in the surrounding components. This increased orifice area commonly takes the form of a groove machined into the surface of the wear plate in contact with the gears. This relief groove extends a little less than halfway along the width of the wear plate; the exact amount is a trade between volumetric efficiency (due to the reduced leakage pathway) and mechanical efficiency and noise (reducing pressure pulsations). Another option is to increase the volume of the cutaway at the root of the gear teeth, thereby increasing the overall volume of the trapped fluid pocket and reducing the amplitude of the pressure spike. However, this also carries a penalty to volumetric efficiency, as this represents extra fluid that is pressurized by the pump that is carried back to the suction region.

3. GEAR DESIGN

Selection and design of the eponymous gears is one of the most critical decisions to make in the design of the pump. The type of gear, the choice of gear tooth, the number of teeth, the gear size, and the aspect ratio all have an impact on the performance of the pump. Additionally, the design of the gears will have a strong influence on the design choices of the other pump elements. As such, the first step in pump development is usually the gears. When selecting a gear, the first detail to be decided is the type of gear used; while we already have restricted the scope of this work to external-tooth gears, there are still further subdivisions to consider. External gears can be categorized by two main features: the tooth profile and the axial profile. The tooth profile is the shape of the cross-section of the tooth in the plane orthogonal to the axis of rotation, and the axial profile is the line or curve created by projecting the same point of the tooth profile along the axis of rotation. A CAD rendering of a spur involute gear is presented in Fig. 28 with the

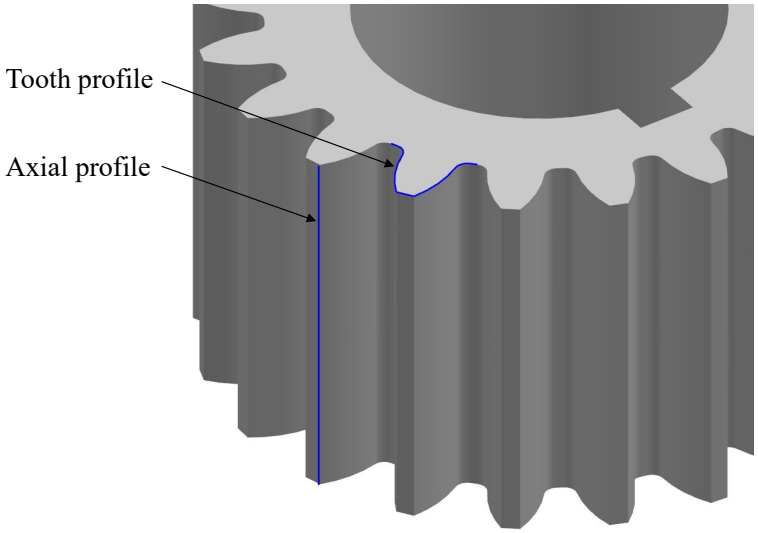

Fig. 28—Highlight of the two critical profiles in an external gear

two profiles highlighted. Beyond the form and structure of the tooth, gears can be defined by an assortment of geometric parameters. While at first glance, these parameters may not seem as impactful as the tooth structure, these parameters are actually the defining foundation for the tooth design and are worth examining first.

3.1 **Gear Parameters & Geometry**

Before a full discussion on gear design and fabrication can be made, we first must review the standard gear nomenclature used for defining gear geometry. These parameters and their definitions provide a designer with a set of variable and defining expressions that can provide a starting point for gear design, which then can be utilized in initial pump sizing estimates. It can be useful to establish the overall geometry of the gear first to guide the design of the tooth profile based on these sizing estimates. In turn, these geometric estimates require a familiarity with the defining gear parameters.

3.1.1 *Overview of Gear Parameters*

The combination of a complex tooth structure and a multitude of gear-design permutations results in a large number of parameters, some of which can be known by multiple names. Below is a list of the key gear parameters that are useful in the design process, an illustration of which is shown in Fig. 29. It can be useful to separate parameters into those that are independent, arbitrarily set by the designer, and those that are dependent, calculated in terms of other parameters.

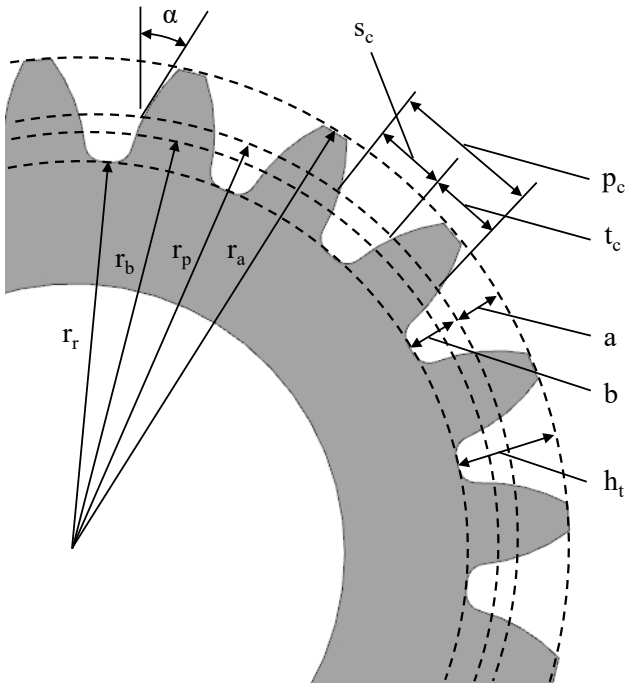

Fig. 29—Primary gear parameters

Independent Parameters

- Pitch diameter, d_p, or pitch radius, r_p, defines the reference circle that many of the gear parameters are measured against. Establishes the nominal point of contact between meshing gears.
- Addendum circle radius, r_a, defines the outermost extension of the gear teeth. Also can be called the outer radius or tip radius.
- Root circle radius, r_r, defines the minimum radius of the gear edge. Also can be thought of as the radius of the cylindrical hub from which the gear teeth extend.
- Profile angle, α, is the angle between the outer surface of the gear tooth and a radial vector. Note that in general, the profile angle is equal to the pressure angle, the angle at which a mated gear applies a force on the gear tooth. Usually between 14.5° and 20°.
- Base circle radius, r_b, defines the base circle that involute or cycloid curves are defined from that constitute the tooth profile. Note that Fig. 29 depicts an involute tooth gear, while for a cycloid gear tooth profile, multiple base circles may be used. Generally, the base circle radius is defined in terms of the pitch radius,
$$r_b = r_p \cos\alpha \, ,$$
but this is not always the case and the base radius may be used as an independent parameter.
- Tooth number, n_t, defines the number of teeth on the gear.
- Gear width, w, is the dimension of the gear along the axis of rotation; also called the land width.

Dependent Parameters

- Circular pitch, p_c, is the arc length along the pitch circle per gear tooth.
$$p_c = \frac{\pi d_p}{n_t}$$
- Diametral pitch, P, is the ratio of the number of teeth to the pitch diameter

$$P = \frac{n_t}{d_p} = \frac{\pi}{p_c}$$

- Addendum, a, is the distance that the gear tooth extends beyond the pitch circle.

$$a = r_a - r_p$$

- Dedendum, b, is the distance between the pitch circle and the root circle.

$$b = r_p - r_r$$

- Circular thickness, t_c, is the thickness of the gear tooth along the pitch circle.

$$t_c = \frac{p_c}{2}$$

- Circular backlash, B_c, is the gap between mated gear teeth along the pitch circle

$$B_c = s_c - t_c$$

- Radial backlash, B_r, is the gap between the outer radius of one gear and the root radius of the other mated gear.

$$B_r = b_1 - a_2$$

- Tooth spacing, s_c, is the arc length along the pitch circle between two adjacent teeth.

$$s_c = \frac{p_c + B_c}{2}$$

- Whole depth, h_t, is the distance from the tooth tip to the root circle and is generally a function of the diametral pitch.

$$h_t = \begin{cases} \dfrac{2.157}{P} & P < 20 \\ \dfrac{2.2}{P} + 0.002 & P \geq 20 \end{cases}$$

- Tip thickness, t_t, is the thickness of the gear tooth at the tip.

$$t_t = t_c - 2a \tan \alpha$$

Note that some dependent and independent parameters can be switched, e.g., either the addendum or the addendum circle radius can be independent with the other the dependent parameter.

3.1.2 *Gear Sizing*

With the basic terminology of gears defined, we now can proceed to initial sizing estimates. Going back to our two key design questions from the beginning of Section 2, we must have some metric of performance that a pump design must achieve, and one of these metrics is usually output flow rate. From Eqs. (1.7) and (1.8), we know that the output flow rate of the pump is a function of the pump displacement, the pump speed, and the volumetric efficiency at the current operating conditions.

$$Q_{actual} = D\eta_v \frac{N}{60} \tag{3.1}$$

The volumetric efficiency is dependent on multiple factors, such as pump speed, pump pressure differential, and the assembly clearances, which makes it difficult to predict efficiency this early in the design process. Instead, we can select some estimated volumetric efficiency based on an assessment of the nominal operating conditions and the precision of fabrication and assembly. Volumetric efficiency increases with higher pump speeds, lower pump pressure differentials, and tighter tolerances in the assembly. High-pressure, high-precision pumps can achieve volumetric efficiencies of 80–95%, while high-pressure pumps with lesser-quality but still adequate fabrication and assembly can reach 50–75%. Low-pressure pumps can be highly variable in acceptable efficiencies, depending on the application.

In a similar fashion, the nominal pump speed can be difficult to determine at this stage of the design process. In some cases, the speed may be dictated by external factors, such as the prime mover being a

fixed-speed motor, or some seal having a restricting slip speed. In most cases, the pump speed will not be so constrained and instead will be an independent design parameter throughout the pump design process. As such, the nominal pump speed will be a compromise between competing factors within the trade space. In these situations, we can use an initial estimate within the range of 1,500–4,000 rpm.

Given these initial estimates, we now can establish some scaling relations to aid in defining some of the geometric parameters of the gear. The first step is to parameterize the pump displacement in terms of the gear geometry. At this stage of the design process, we have not yet identified a tooth profile, but the most common is the involute tooth, though the following analysis also should hold for most tooth profiles. We can calculate the displacement of the pump as the sum of the volumes of the intertooth gaps of both gears; assuming constant gap volume, V_{gap}, around the gear and identical gears, we can calculate the displacement as

$$D = 2n_t V_{gap} \, . \tag{3.2}$$

The gap volume can be estimated as a wedge segment with an average width equal to the tooth spacing along the pitch circle and a height equal to the gear width, *w*. An illustration of the tooth gap geometry is shown in Fig. 30. This method is not exact, as it relies on the assumption that the volume of gear material within the wedge is approximately equal to the volume of the intertooth gap not included in the wedge. However, it is a useful approximation for early gear sizing.

$$V_{gap} = h_t s_c w \tag{3.3}$$

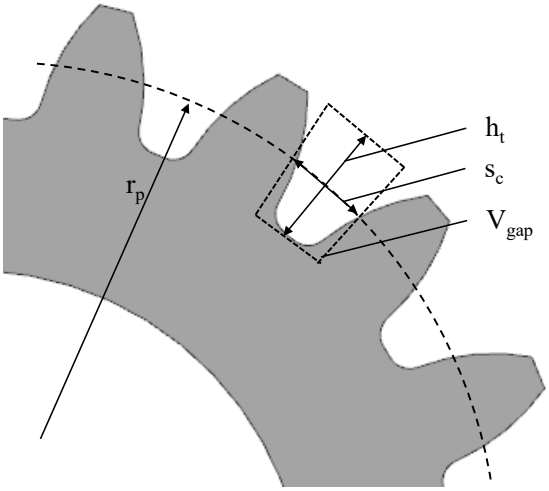

Fig. 30—Intertooth gap geometry

The displacement of the pump then can be written by combining Eqs. (3.2) and (3.3).

$$D = 2n_t h_t s_c w \tag{3.4}$$

A more thorough calculation of the displacement can be performed once a gear geometry has been established, either through numerical integration of the tooth profile or calculation in a CAD program. By substituting several definitions of several gear parameters from Section 3.1.1 into Eq. (3.4), we can create a scaling relation for the pump displacement. For the sake of simplicity, we will assume negligible backlash and a diametral pitch of 20 or greater, and will drop the 0.002 from the whole depth relation.

$$D = 2n_t\left(\frac{2.2}{P}\right)\left(\frac{p_c}{2}\right)w$$

$$= 2n_t\left(\frac{2.2d_p}{n_t}\right)\left(\frac{\pi d_p}{2n_t}\right)w$$

$$D = \frac{2.2\pi w d_p^2}{n_t} \tag{3.5}$$

Substituting Eq. (3.5) into Eq. (3.1) and solving for the pitch diameter yields a relation for the pitch diameter as a function of design requirements and independent parameters.

$$Q_{req} = \left(\frac{2.2\pi w d_p^2}{n_t}\right)\eta_v \frac{N}{60}$$

$$d_p^2 = \frac{60 n_t Q_{req}}{2.2\pi w N \eta_v}$$

$$d_p = 2.95\sqrt{\frac{n_t Q_{req}}{wN\eta_{v,est}}} \tag{3.6}$$

This initial gear sizing is based on parameters that most likely are not set at this stage, nor even concretely identified, such as the volumetric efficiency and the pump speed. Additionally, there are two other geometric parameters, the gear width and the tooth number, that create a trade space around the gear geometry. The width of the gear primarily affects the form factor of the pump and the relative prevalence between radial and axial leakage. In contrast, the tooth number has a pronounced impact on pump performance and capability in terms of pump ripple and maximum gear load. Therefore, the tooth number and its trade space should be examined next.

3.1.3 *Gear Tooth Number & Bending Stress*

Selecting the gear tooth number is primarily a trade between flow ripple, pump displacement, and tooth stress. Assuming a fixed pitch diameter, as the tooth number increases, the pump displacement decreases, as per Eq. (3.5). At the same time, this increases the frequency and decreases the amplitude of pump ripple, as discussed earlier in Section 2.2. The final component of the trade space is the bending stress on the gear teeth. The Lewis equation for the bending stress, σ_b, is [13]

$$\sigma_b = \frac{F_t P}{wY}, \tag{3.7}$$

where Y is the Lewis form factor, a dimensionless geometric factor. The original formulation of the Lewis form factor has been modified multiple times to account for additional factors that were unknown at the time of Lewis's formulation, such as stress concentration. An updated geometry factor, J, can be used in the AGMA bending stress equation [14],

$$\sigma_b = \frac{F_t P}{wJ}\frac{K_a K_m}{K_v} K_s K_B K_I, \tag{3.8}$$

where the K terms are empirical factors that account for additional effects. The geometric factor J can be found in tables or calculated using AGMA standard 908-B89. The dynamic factor, K_v, accounts for vibration loads due to tooth impacts from imperfect meshing, also known as transmission error.

$$K_v = \left(\frac{A}{A + \sqrt{V_t}}\right)^B \quad [US]$$
$$K_v = \left(\frac{A}{A + \sqrt{200V_t}}\right)^B \quad [SI] \tag{3.9}$$

$$A = 50 + 56(1 - B) \tag{3.10}$$

$$B = \frac{(12 - Q_v)^{2/3}}{4} \quad Q \in I, [6,11] \tag{3.11}$$

In the above equations, V_t is the tangential velocity of the gear at the pitch circle, also known as the pitch-line velocity, given in either ft/min for the US system formulation or m/s for the SI formulation. The dynamic factor is dependent on the quality and fabrication method of the gear, abstracted as the quality index, Q_v. It is recommended practice to use the quality index of the lower-quality gear to be conservative.

The load distribution factor, K_m, accounts for axial misalignment or deviation that causes the transmitted load between teeth to be unevenly distributed. This is a particular problem for gears with larger widths. Suggested values for K_m are given in Table 2. A suggested standard is to limit the gear width to $8/P < w < 16/P$.

Table 2—Suggested Load Distribution Factor

w (in)	w (mm)	K_m
< 2	< 50	1.6
6	150	1.7
9	250	1.8
> 20	> 500	2.0

The application factor, K_a, accounts for time-varying loads on the gear, as the original Lewis equation assumed constant tooth loading. This can be particularly important for a gear pump, as pressure changes can substantially alter loading on the gears. The application factor is dependent on a qualitative assessment of how uniform the prime mover and the driven load are, with suggested values given in Table 3.

Table 3—Suggested Load Distribution Factor

	Expected Load		
Prime Mover	**Uniform**	**Moderate Shock**	**Heavy Shock**
Uniform (Electric motor, turbine)	1.00	1.25	> 1.75
Light Shock (Multi-cylinder engine)	1.25	1.50	> 2.00
Medium Shock (Single-cylinder engine)	1.50	1.75	> 2.25

The size factor, K_s, comes from fatigue loading analysis and represents how empirical data on fatigue failure is based on relatively small test specimens. However, most gear-strength data have been collected from actual gear teeth and not facsimiles, so the use of the size factor is not necessarily recommended. AGMA standard is to use a value of K_s of 1.0 unless a larger value is desired as a factor of safety for particularly large gear teeth. In such instances, a value of K_s between 1.25–1.50 is conservative.

The rim thickness factor, K_B, is a relatively new factor to account for gears with a rim-and-spoke design, rather than a solid disc. For gear pumps, this would arise when the gear's bore diameter, d_b, is close to its root diameter, d_r. AGMA standards have defined the rim thickness factor in terms of a backup ratio, m_B. Should the gear not have a bore, the rim thickness factor is 1.

$$K_B = \begin{cases} -2m_B + 3.4 & 0.5 \leq m_B \leq 1.2 \\ 1.0 & m_B > 1.2 \end{cases} \tag{3.12}$$

$$m_B = \frac{(d_r - d_b)}{2h_t} \tag{3.13}$$

The idler factor, K_I, accounts for the fact that the idle gear is subjected to more stress cycles per unit time and to higher-magnitude loads than the drive gear. For this reason, the idler factor is 1.42 for the idle gear and 1.0 for the drive gear. Note that AGMA standards typically apply the reciprocal of this factor as a reduction of the gear material strength rather than as an amplification of the applied load.

The final element required for estimating the tooth-bending stress is an estimate of the applied load. The conservative approach is to assume only one pair of teeth transfers the full load at any point in time, even if in practice there are multiple pairs in contact during the mesh, and that the load is applied the tip of the tooth. In practice, the point of contact when only a single pair of teeth are in contact is usually closer to the pitch circle, but assuming the maximum moment arm is the more conservative approach.

$$F_t = r_p \tau \tag{3.14}$$

The torque on the gear due to the required pump pressure differential, Δp_{req}, can be estimated by comparing the nominal input shaft power to the required output flow power.

$$\begin{aligned} \tau \omega_{nom} &= \frac{1}{\eta_v \eta_m} \Delta p_{req} Q_{req} \\ \tau \left(\frac{2\pi N_{nom}}{60} \right) &= \frac{1}{\eta_v \eta_m} \Delta p_{req} \left(\eta_v \frac{N_{nom} D}{60} \right) \\ \tau &= \frac{D \Delta p_{req}}{2\pi \eta_m} \end{aligned} \tag{3.15}$$

Equation (3.15) requires an estimate of the mechanical efficiency of the pump across the range of pressures of interest. Well-designed pumps can achieve mechanical efficiencies in the range of 70–90%, depending on the operating pressures. Gear pumps with floating wear plates tend to have increased mechanical efficiency at higher pressures and decreased mechanical efficiency at higher pump speeds. In practice, the mechanical efficiency of a pump is not constant across its range of operating conditions, but Eq. (3.15) provides a reasonable initial estimate. As an example, Fig. 31 shows a comparison of modeled and measured pump torque load for small gear pump (D = 0.82 cm^3). An average mechanical efficiency of 82% was used as the estimated efficiency, which demonstrates broad applicability of Eq. (3.15).

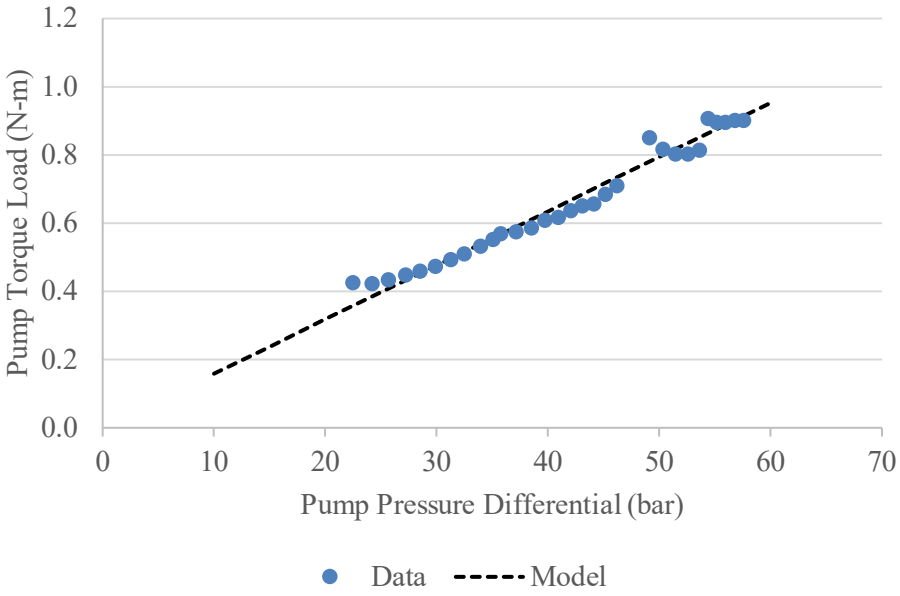

Fig. 31—Comparison of empirical and modeled pump torque load as a function of pump pressure differential. Model corresponds to Equation (3.15).

To conduct a basic scaling analysis, let us go back to the original Lewis equation, neglecting for now the additional factors of the AGMA equation. If we substitute Eqs. (3.14) and (3.15) along with the definition of the diametral pitch into Eq. (3.7), we can solve for the tooth-bending stress based on our initial requirements and performance estimates.

$$\sigma_b = \frac{1}{wY}\left(\frac{d_p}{2}\frac{D\Delta p_{req}}{2\pi\eta_m}\right)\left(\frac{n_t}{d_p}\right)$$
$$\sigma_b = \frac{1}{4\pi}\frac{Dn_t\Delta p_{req}}{wY\eta_m} \tag{3.16}$$

Thus, for a given displacement (set by the nominal pump speed and required flow rate), the tooth-bending stress increases as the tooth number increases. The Lewis form factor increases with tooth number, but of order less than 1, which results in a stress dependency on tooth number of order between 0 and 1.

From this, we can see that the trade on tooth number is between reduced flow ripple on one side and increased flow rate for a given pump size and lower tooth stress. The specific application of the pump and any specific circumstances or additional criteria come into play. For sensitive or precise applications, the need for reduced pump ripple may drive the design toward a greater tooth number, where the increased tooth stress must be compensated by high-strength gear material. Conversely, designs that feature low-strength materials, such as plastics, will be inherently limited in possible tooth numbers, depending on the expected pressure loads.

3.1.4 *Gear Aspect Ratio*

The keen reader might have noticed that up to this point, while we have defined initial performance estimates, nominal operating conditions, and required outputs, there is still one underconstrained element of the gear-sizing process — the displacement. The displacement itself is set by the nominal pump speed, the required flow rate, and the estimated volumetric efficiency, but two key independent gear parameters that define the pump displacement are not yet established: the pitch diameter and the gear width. There is

a continuum of possible combinations of pitch diameters and gear widths that will provide the required displacement. This brings the design to a question of the aspect ratio of the gear, or the ratio of the width to the pitch diameter. As with most elements in an interconnected system, this creates a series of trades between competing forces that must be considered.

The first consideration to make is that the pitch diameter of the gears sets the shaft separation distance, which, in turn, imposes hard limits on the space available for shaft support. Thus, the pitch diameter must be greater than the diameter of the chosen bearing element, whether it be a rolling-element bearing, a journal bearing, etc. Associated with this is that the shaft diameter is constrained by the pitch diameter, as a bore diameter too close to the root diameter increases the rim thickness factor and weakens the gear. This is particularly important for high-pressure pumps, where thicker shafts are needed to transmit the higher torque loads.

The second consideration is the relative balance between radial and axial leakage. For large aspect ratios, radial leakage dominates, while for small aspect ratios, the axial leakage is the primary contributor. This can be important when attempting to reduce fabrication costs by minimizing the number of precision components: a gear with a high aspect ratio will require greater emphasis on the pump casing, while a low-aspect-ratio gear will place greater importance on the wear plates (or the axial clearance, if no wear plates are used). Another related concern is the balance between volumetric and mechanical efficiency. The larger the diameter of the gear, the greater the mechanical losses due to viscous dissipation during operation, but also the longer the leakage pathway across the face of the gear. Thus, increasing the diameter for a given displacement increases volumetric efficiency and decreases mechanical efficiency.

There are several other impacts that, though of smaller importance, should be considered as well. Increasing the width of the gear also increases the volume of the fluid pocket between the gear teeth that must escape through the relief grooves during compression, which increases the pressure pulsations during operation. There are also form-factor considerations, such as the overall pump size envelope, whether it has to conform to any nearby subsystems as part of a larger assembly, etc. There also can be manufacturing considerations to the aspect ratio. A common practice in the gear pump industry is to design a family of pumps with a similar overall design subdivided by gear pitch diameters. Within these subdivisions, the gear width can be varied to create a range of pump displacements that share multiple components, such as wear plates, casing end plates, seals, etc.

In summary, a designer inevitably will confront an underconstrained system that will require some measure of qualitative assessment to determine what traits are needed by a specific application. Returning to these assumptions is a fundamental element of the design iteration process and a designer always should be willing to revisit the specifications of the gear itself as the pump design evolves.

3.2 **Tooth Profiles**

The next aspect of the gear is the gear tooth shape, or the cross-sectional profile. The tooth profile generally consists of a combination of several curves to define the tooth flank, tip, and root. There have been many gear tooth profiles developed utilizing various curves, but there are three curves in particular worth discussing: the involute curve, the cycloidal curve, and the circular arc. In addition, hybrid profiles are possible that combine elements of different curves to create a compound curve that provides a mix of benefits and features not possible with any single curve.

3.2.1 *Involute Profile*

The most common type of gear tooth is the involute tooth, where the edges of the tooth are defined using involute curves. Involutes are created by “unwinding” a line from a reference curve; for gears, the

reference curve is a circle and is referred to as the base circle. There are three primary advantages to involute gear teeth. The first is that involute teeth offer the desirable feature that at every angle of rotation during the mesh, the point of contact between two gears exists along a single line, called the line of action. This also ensures a consistent load force vector during the mesh. A second advantage of the involute profile is that there is greater tolerance for lateral displacement between the two axes of rotation while maintaining a smooth mesh. In gear pumps, this can arise either through errors in tolerancing (such as in the shaft diameter or the bearing bore diameter, or in the wear plate), or through deflection of the shaft at high pressures. The third advantage to involute teeth is the relative ease of fabrication and wider availability. The primary disadvantage of involute teeth is that the tip of one tooth does not conform to the bottom land of the mating gear, which results in trapped fluid pockets during the mesh. These fluid pockets are passed through the mesh back to the suction region, reducing volumetric efficiency. These pockets also can be compressed during the mesh, leading to pressure spikes that can damage pump components and generate undesirable noise.

The profile begins with the root circle, which is the minimum radius that the tooth profile reaches; referencing the gear parameter relations from Section 3.1.1, the root radius, r_r, is defined as

$$r_r = r_p - b\,, \tag{3.17}$$

where r_p is the pitch radius and b is the dedendum. The tooth profile extends radially outward from the root circle to the base circle; the base circle, r_b, is generally defined as

$$r_b = r_p \cos\alpha\,, \tag{3.18}$$

where α is the pressure angle of the gear. It is important to note that the base circle is generally distinct from the pitch circle. Some pressure angles commonly used in standard spur gears are14.5°, 20°, and 25°. Mathematically, the Cartesian coordinates of the involute curve can be expressed parametrically as

$$x = r_b(\cos t + t\sin t) \tag{3.19}$$
$$y = r_b(\sin t - t\cos t)\,, \tag{3.20}$$

where t is the parametric variable. An alternate definition in polar coordinates is shown below.

$$r = r_b\sqrt{1+t^2} \tag{3.21}$$
$$\theta = \theta_{ref} + t - \tan^{-1} t \tag{3.22}$$

The reference angle, θ_{ref}, sets the angular position where the curve begins. The involute curve extends from the base circle to the maximum radius of the gear at the addendum circle.

$$r_a = r_p + b \tag{3.23}$$

The opposing flank of the tooth is defined similarly, starting with a radial component from the root circle. The angular position of the second flank is offset by the angular thickness of the tooth, which is the ratio of the circular thickness to the pitch radius. The second flank curve is an involute of the opposite direction, so the parametric definition in polar coordinates is as follows:

$$r = r_b\sqrt{1+t^2} \tag{3.24}$$
$$\theta = \left(\theta_{ref} + \frac{t_c}{r_p}\right) - t + \tan^{-1} t\,. \tag{3.25}$$

The very tip of the gear is an arc section of the addendum circle between the termination of the two involute curves. An example of an involute tooth profile is plotted in Fig. 32. Additional features can be added, such as a fillet on the tooth flank, to improve tooth strength or to compensate for other factors. However, some of these modifications, such as undercutting at the root, can increase the volume of fluid pockets formed and can reduce volumetric efficiency further.

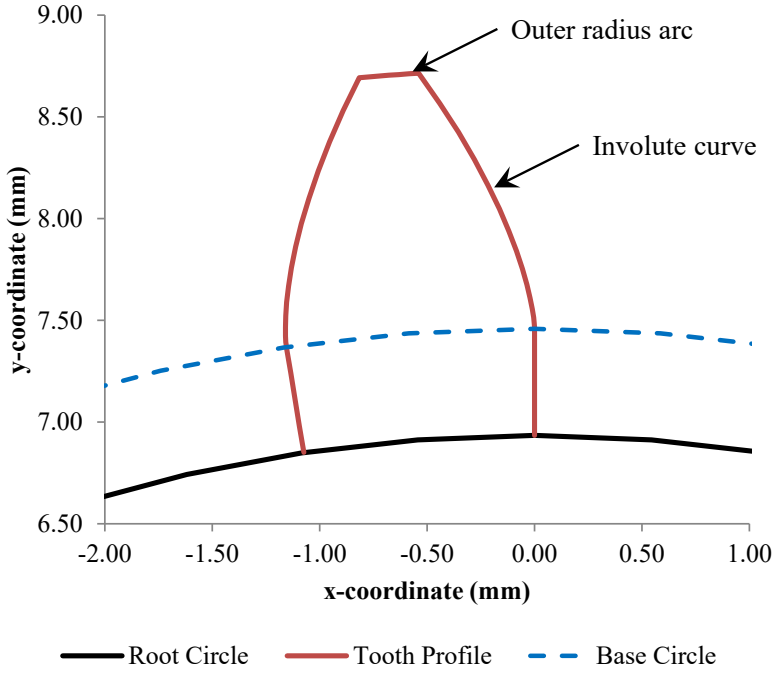

Fig. 32—Involute tooth profile

3.2.2 *Cycloidal Profile*

Cycloid tooth gears traditionally have been used for mechanical clocks, and rarely have been used for pumps or power transmission. However, recent interest in tooth profiles that fully sweep the mating gear, thus eliminating the trapped fluid pockets that can cause pressure spikes, has renewed interest in cycloid tooth profiles. The cycloid curve is generated as the pathline of a point on a generating circle that is rolled along a reference circle. For gear teeth, this reference circle is the pitch circle, rather than a separate reference circle as is the case with involute gear teeth. The addendum curve is produced by a generating circle that rolls on the outside of the reference circle, while the dedendum curve is produced by the generating circle rolling on the inside of the reference circle. These two curves are called the epicycloid and hypocycloid curves, respectively. Mathematically, the Cartesian coordinates of the epicycloid curve as a function of the angular position is defined as

$$x(\theta) = (r_p + R_G)\cos\theta - R_G \cos\left(\frac{r_p + R_G}{R_G}\theta\right) \tag{3.26}$$

$$y(\theta) = (r_p + R_G)\sin\theta - R_G \sin\left(\frac{r_p + R_G}{R_G}\theta\right), \tag{3.27}$$

where R_G is the radius of the generating circle. The definition of the hypocycloid curve can be defined in a similar manner.

$$x(\theta) = (r_p - R_G)\cos\theta + R_G \cos\left(\frac{r_p - R_G}{R_G}\theta\right) \tag{3.28}$$

$$y(\theta) = (r_p - R_G)\sin\theta - R_G \sin\left(\frac{r_p - R_G}{R_G}\theta\right) \tag{3.29}$$

The most basic cycloidal profile is to combine alternating epicycloid and hypocycloid curves where they meet at the pitch circle. This creates a lobed gear structure, shown in Fig. 33. However, a wide variety of tooth profiles can be created by combining more than two cycloid curves to create a single tooth curve, similarly to how the involute tooth profile consists of two involute curves with interconnecting circular and radial curves. Additionally, the cycloid curves can be truncated or fileted to modify the final tooth profile.

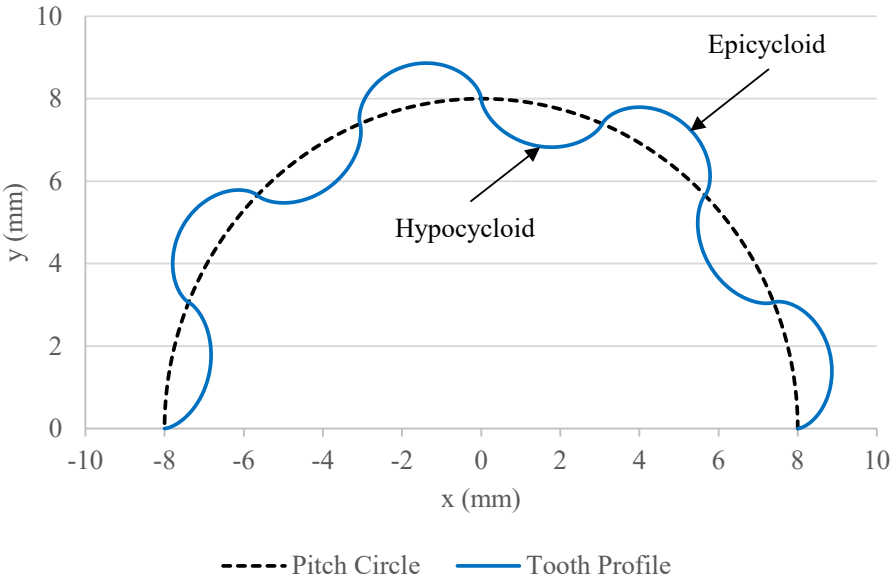

Fig. 33—Compound tooth profile combining epicycloid and hypocycloid curves

Historically, cycloid gears were advantageous due to the relative ease of fabrication at size scales relevant to watch making, along with additional factors such as higher tooth strength and lower friction. For pump applications, the primary advantage of a cycloid tooth profile is that it can readily be tailored to provide conforming tooth profiles through the mesh that eliminate trapped fluid pockets. Frictional concerns are generally less relevant in fluid pumps compared to timepieces, since the gears in pumps generally are lubricated by the working fluid. The main disadvantage of the cycloid profile is that it is much more sensitive to the distance between axes of rotation. Thus, high-pressure pumps that seek to use cycloid profiles must pay considerable attention to shaft deflection and tolerancing.

3.2.3 *Circular Arc Profile*

Another method to provide a tooth profile that conforms to the bottom land of the mating gear is to use a circular arc for both the addendum and dedendum curves. Similar to the cycloidal tooth profile, where epicycloid and hypocycloid curves are joined to define the tooth, the circular arc tooth profile consists of two circular arcs joined at the pitch circle to define the top and bottom lands. These arcs are centered on the pitch circle, offset by an angle, θ_D, which is half of the angle of the circular pitch.

$$\begin{aligned}\theta_D &= \frac{1}{2}\frac{p_c}{r_p} \\ &= \frac{1}{2}\frac{2\pi r_p}{r_p n_t} \\ \theta_D &= \frac{\pi}{n_t}\end{aligned} \tag{3.30}$$

The addendum arc is of radius a and extends beyond the pitch circle to both points where the arc would contact the pitch circle. The dedendum arc, with radius b, likewise proceeds within the pitch circle across both points of intersection with the pitch circle. The angular progression through each circular arc can be defined by the angle γ. An illustration of the profile geometry is shown in Fig. 34.

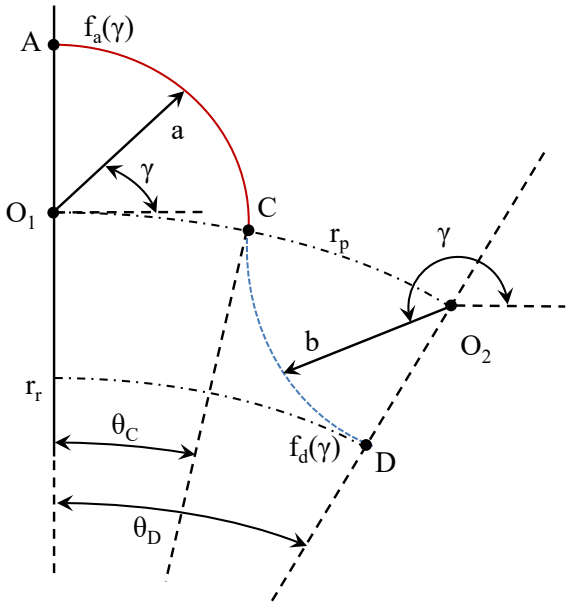

Fig. 34—Circular tooth profile reference geometry

For a tooth profile where full arcs are used, that is, no truncation of the curve is employed, the addendum and the dedendum are constrained by the need to ensure an integer number of teeth. If we assume a selected pitch radius and tooth number, the addendum and dedendum are constrained to one possible solution for a given value of θ_C that allows both curves to meet at point C. Since we know that the line segment O_1C has a length of a, the value of a is related to the value of θ_C by the Law of Cosines.

$$\begin{aligned}a^2 &= r_p^2 + r_p^2 - 2r_p^2\cos\theta_C \\ &= 2r_p^2(1-\cos\theta_C) \\ a &= r_p\sqrt{2(1-\cos\theta_C)}\end{aligned} \tag{3.31}$$

The dedendum is similarly uniquely defined.

$$b = r_p\sqrt{2(1-\cos(\theta_D-\theta_C))} \tag{3.32}$$

With the addendum and the dedendum defined, the circular curves for the tooth can be defined in Cartesian coordinates as a function of the reference angle γ.

$$x_a(\theta) = x_{O_1} + a\cos\gamma \quad (3.33)$$
$$y_a(\theta) = y_{O_1} + a\sin\gamma \quad (3.34)$$

$$x_d(\theta) = x_{O_2} + b\cos\gamma \quad (3.35)$$
$$y_d(\theta) = y_{O_2} + b\sin\gamma \quad (3.36)$$

The bounds of the value of the reference angle, γ, are chosen such that the curves start and terminate on the pitch circle. The coordinates of the curve origin points of the ith tooth are defined by a reference angle for the position of the first tooth, θ_{ref}, and the angle θ_D.

$$x_{O_{1,i}}(\theta) = r_p\cos\left(\theta_{ref} + 2\theta_D(i-1)\right) \quad (3.37)$$
$$y_{O_{1,i}}(\theta) = r_p\sin\left(\theta_{ref} + 2\theta_D(i-1)\right) \quad (3.38)$$

$$x_{O_{2,i}}(\theta) = r_p\cos\left(\theta_{ref} + \theta_D(2i-1)\right) \quad (3.39)$$
$$y_{O_{2,i}}(\theta) = r_p\sin\left(\theta_{ref} + \theta_D(2i-1)\right) \quad (3.40)$$

An example of a set of addendum and dedendum curves for a circular arc tooth are shown in Fig. 35. Much like cycloidal tooth profiles, the circular arc can be truncated to produce a narrower tooth, though this comes at the cost of the mutually conforming addendum and dedendum profiles.

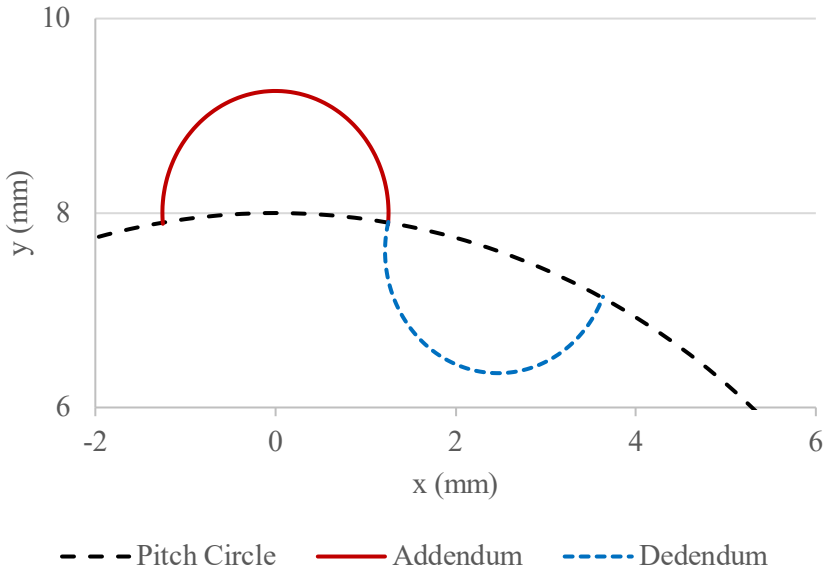

Fig. 35—Example circular arc tooth curves

3.2.4 *Hybrid Profiles*

We already have seen that there is no inherent restriction that the tooth profile must be a single curve. The standard involute tooth profile is a combination of two involute curves for the flanks and circular arcs along the addendum and root circles for the top and bottom lands, respectively. Likewise, tooth profiles may include truncated cycloidal or circular curves connected by arcs along the addendum or root circle. This approach yields a compound curve for the tooth profile with a greater degree of flexibility compared

to a full cycloidal or circular curve that is uniquely defined based on the tooth number. Breaking down the tooth profile into multiple curves can be extended such that the tooth flank itself consists of more than one curve joined together to be smooth and continuous. These hybrid profiles offer a designer nearly limitless possibilities for customization.

As an example, suppose a tooth profile that combines three curves to create half of a symmetric tooth profile. The hybrid tooth profile consists of three sections: the addendum arc, the flank curve, and the dedendum arc. The two arcs have their separate origins located on the pitch circle. The beginning and end points of the profile are labeled A and D, respectively, while the transition points between addendum or dedendum arcs and the flank curve are labeled B and C, respectively. An illustration of the geometry for a three-curve hybrid tooth profile is shown in Fig. 36.

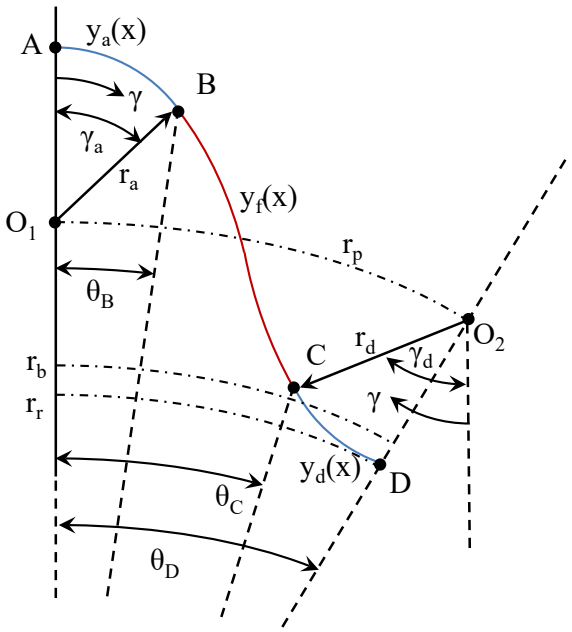

Fig. 36—Geometry of one half of a hybrid tooth profile

In order to create a complete hybrid profile, the three component curves must be combined and then reflected across the y-axis. The criteria for merging the three curves are that the created curve is continuous and differentiable. The curves defined thus far satisfy the criteria in isolation, which means that the solution criteria reduces solely to where the curves meet — points B and C. Mathematically, this means that at point B,

$$y_a(x_B) = y_f(x_B) \tag{3.41}$$

$$\left.\frac{dy_a}{dx}\right|_{x_B} = \left.\frac{dy_f}{dx}\right|_{x_B}, \tag{3.42}$$

and at point C,

$$y_d(x_C) = y_f(x_C) \tag{3.43}$$

$$\left.\frac{dy_d}{dx}\right|_{x_C} = \left.\frac{dy_f}{dx}\right|_{x_C}. \tag{3.44}$$

Given the use of parametric equations, it is worth noting that

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt}.$$

Solving Eqs. (3.41) through (3.44) for the three curve formulations is the crux of the generation of the hybrid curve. This is complicated by the fact that generally, the transition between curves, points B and C, is not uniquely defined. This in addition to the fact that in such hybrid profiles, the base circle can become an independent variable, resulting in a problem that is underconstrained. As such, there is no unique solution, so additional criteria will be required from the designer to refine the profile design. Numerical calculations of the curve-splicing criteria and iteration of the design are highly likely for hybrid profiles, making for a much more involved design process compared to standard tooth profiles. However, hybrid tooth profiles offer the potential for capturing the benefits of multiple profiles while minimizing the flaws.

3.3 **Axial Profiles**

3.3.1 *Design Considerations*

With a tooth profile established, the next step in the design of the gear is to define the axial profile, which, as a reminder, is the pathline of the same point on the tooth profile along the width of the gear. The three primary categories of axial profiles for pump application are spur, helical, and symmetric. A spur gear has an axial profile that is a straight line parallel to the axis of rotation and represents a simple extrusion of the gear cross section. A helical gear has an axial profile that is a straight line at an angle to the axis of rotation, resulting in a tooth that wraps around the gear hub like a helix. In theory, the axial profile of the helical gear does not have to be a straight line and could be any monotonic curve. In practice, this complicates fabrication for minimal performance gain, so straight-line profiles are the standard form for helical gears. The final category consists of axial profiles that are symmetric with respect to the midplane of the gear. This group includes curvilinear and herringbone gears, the latter of which is pictured in Fig. 37.

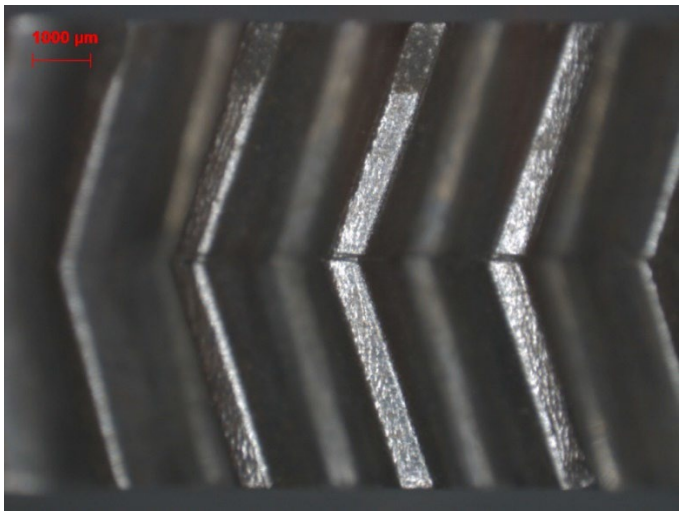

Fig. 37—Microscope imaging of the side of a herringbone gear

The primary design consideration for the axial profile is the engagement of the teeth during meshing. Spur gears engage along the entire tooth length simultaneously, which offers a larger contact area and thus a higher load capacity. However, this results in a more sudden application of force, which generates noise.

Helical gears offer a more gradual tooth engagement, which reduces noise at the cost of reducing load capacity. However, the angled contact surface between the two teeth imparts an axial force on the gear during meshing, illustrated in Fig. 38. When gears mesh, all force exchange is normal to the surface of the tooth at the point of contact. For spur gears, this results in primarily tangential loads with some small radial loads, as the contact surface is orthogonal the axial direction. For helical gears, contact force vector contains a component parallel to the axis of rotation, resulting in large tangential and axial loads in addition to minor radial loads. This axial load must then be compensated, either by the wear plate or some thrust bearing. The extra load often can result in increased wear on components and frictional losses. Symmetric gears provide the gradual tooth engagement of helical gears but without the axial load on the gears, as each half of the gear has opposing force vectors. The primary drawback with symmetric gears is the complexity involved with fabrication and the resulting expense.

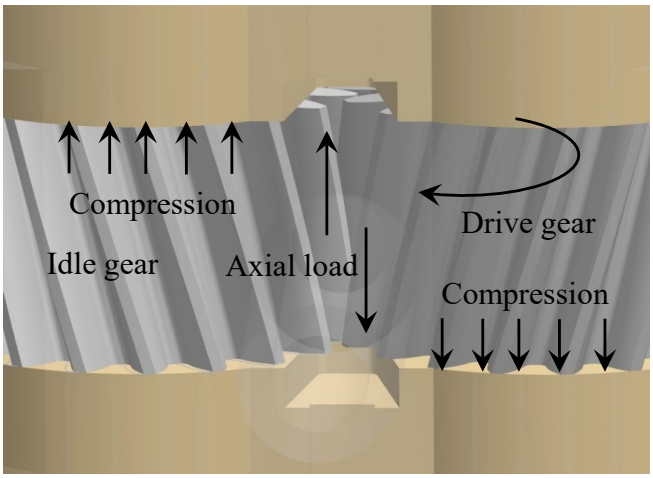

Fig. 38—CAD rendering of the meshing of helical gears and the induced axial loading

An additional consideration in the selection of the axial profile is pressure spikes within the mesh. Recalling Section 2.3, pressure spikes occur when fluid pockets are compressed within the meshing region with insufficient exit area to escape. The use of helical or symmetric gears reduces the magnitude of the pressure spikes, as the gradual tooth engagement allows for compression of the fluid pocket from one or both ends of the gear before full tooth engagement. This results in a larger exit area for the fluid, reducing the pressure required to drive the fluid from the pocket. Similarly, helical and symmetric gears offer reduced pump ripple, as the relative rate of change of volume into or out of the discharge region of any given tooth is smaller.

Therefore, the selection of the axial profile is primarily a trade between fabrication complexity (and thereby cost) and pump performance, with a consideration for tooth strength. Spur gears are the simplest and easiest to design and fabricate, and are the most robust against bending stress. However, this comes at the cost of the highest operating noise level and, depending on wear plate design, higher pressure spikes within the mesh. Helical gears are marginally more complex to design and fabricate, but offer reduced noise, pump ripple, and pressure spikes. However, the induced axial loading on the gears imparts additional design constraints on other components within the pump, namely the wear plate and the shaft supports. Symmetric gears capture most of the benefits of helical gears with the pump design simplicity of spur gears, having no net axial loading, but are the most complex to design and fabricate, and thus the most expensive to include in a pump. These are the considerations that a designer must consider when selecting an axial profile.

3.3.2 *Profile Formulation*

Once an axial profile has been selected, the next challenge is to formulate it in whatever design platform is used. Often this may require representing that which exists on a cylindrical surface in Cartesian coordinates, depending on the capabilities of the software package. For spur gears, the axial profile is trivial, as the gear can be defined as a simple one-dimensional extrusion of the gear cross section. For helical and symmetric gears, the task is more involved, as it requires projecting a two-dimensional curve onto a three-dimensional surface. We begin with defining both a Cartesian and a cylindrical coordinate system origin at the intersection between the axis of rotation and the top axial surface of the gear, shown in Fig. 39. The axial profile is a curve in the x-z plane that is projected onto the cylindrical surface; the angular position of the axial profile at z = 0 is at some initial angular position θ_0. This value acts as a reference angle that positions the axial profile.

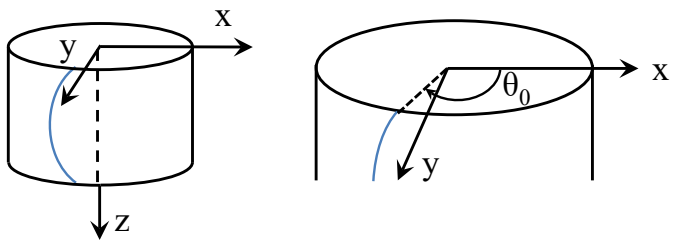

Fig. 39—Coordinate system for tooth axial profile generation. Dashed, black line denotes the axis of rotation, and solid, blue line denotes the axial profile.

Suppose that the axial profile is defined as some function f in the x-z plane that is projected onto the surface of the cylinder, where

$$x = f(z)\,. \tag{3.45}$$

The corresponding angular position on the cylindrical surface is then set by the radius of the cylinder, in this case, the root circle.

$$\theta = \cos^{-1}\left(\frac{f(z)}{r_r}\right) \tag{3.46}$$

The y-coordinate of this point can then be solved as a function of the axial position, z.

$$\begin{aligned} y &= r_r \sin\left[\cos^{-1}\left(\frac{f(z)}{r_r}\right)\right] \\ &= r_r\sqrt{1-\left(\frac{f(z)}{r_r}\right)^2} \\ y &= \sqrt{r_r^2 - \big(f(z)\big)^2} \end{aligned} \tag{3.47}$$

Equations (3.45) and (3.47) together define the projection of the axial profile onto the root cylinder as a function of the axial position in Cartesian coordinates.

One potential challenge in the practice of gear design is dependent on the CAD software used to design the gear, particularly for complex tooth profiles. In some cases, when a tooth profile is swept along an axial profile, the CAD program will require that the axial profile and the tooth profile intersect at an explicit reference point. The formulation of the axial profile above is defined along the root circle, which only intersects with the tooth profile at the bottom land of the gear. However, it is easy to conceive of the axial profile as defining the top land, which, if not corrected for, can lead to errors. Therefore, it is useful to define the axial profile starting at some initial angle θ_0 that corresponds to the angular position of the first point where the tooth profile intersects the root circle (e.g., θ_D from Fig. 34). This then can be used to define the starting point of the axial profile.

$$f(0) = r_r \cos\theta_0 \tag{3.48}$$

3.4 Gear Case Study

As an example, let us examine the design of a curvilinear circular-arc tooth (CCAT) gear. Functionally, we desire that the gear must display the following characteristics:

- a circular profile of the addendum to match the dedendum, so that one tooth completely sweeps the volume between tooth flanks of the mating gear,
- gradual engagement of the teeth during meshing, enabling smoother operation and lower-pressure pulsations, and
- balanced axial force transfer during engagement to eliminate net axial loads on the gears.

In order to achieve these goals, the tooth profile must hybridize the profiles of the involute curve for smooth force transfer with the void-clearing capability of a circular-arc tip. Likewise, the gradual engagement of a helical gear must be made symmetrical to balance the axial loading.

3.4.1 *Axial Profile*

To begin, we will choose a half-period sine wave for the axial profile. The surface of a cylinder of radius r is defined by

$$\begin{aligned} x &= r\cos t \\ y &= r\sin t \\ z &\in [0, w], \end{aligned} \tag{3.49}$$

where t is the parametric parameter and w is the width of the gear. The standard form of the equation for a sine wave in 2D Cartesian coordinates is

$$y = A\sin\left(\frac{2\pi}{T}x\right), \tag{3.50}$$

where A is the amplitude of the sine wave and T is the period of the wave. In this application, the sine wave is wrapped around the cylinder such that

$$\begin{aligned} y &= r(\theta - \theta_0) \\ x &= z\ , \end{aligned}$$

where θ is the angular position of an arbitrary point on the axial profile, and θ_0 is the angular position of the start of the curve when z = 0 (see Fig. 39). The period of the sine wave is arbitrarily chosen to be equal

to $2w$ so that the axial profile is one-half of a sine wave. Substituting this into Eq. (3.50) yields the polar equation for the sine wave.

$$r(\theta - \theta_0) = A\sin\left(\frac{\pi}{w}z\right) \tag{3.51}$$

However, the sine wave needs to be expressed in 3D Cartesian parametric form. To start, the parametric parameter is defined as

$$t \in [0,1], \tag{3.52}$$

which then allows Eq. (3.49) to define the z parameter.

$$z(t) = tw \tag{3.53}$$

Substituting Eq. (3.53) into Eq. (3.51),

$$\theta = \frac{A}{r}\sin(\pi t) + \theta_0 \ , \tag{3.54}$$

which can be substituted into Eq. (3.49). For the sake of convenience, the axial profile is chosen arbitrarily such that it runs along the dedendum of the tooth profile (point D in the next section). Thus, the cylinder of interest is the root cylinder where $r = r_r$.

$$x(t) = r_r \cos\left[\frac{A}{r_r}\sin(\pi t) + \theta_0\right] \tag{3.55}$$

$$y(t) = r_r \sin\left[\frac{A}{r_r}\sin(\pi t) + \theta_0\right] \tag{3.56}$$

The two remaining parameters that must be defined are the angular position of the start of the axial profile, θ_0, and the amplitude, A. With regard to θ_0, the midpoint of the tooth profile is defined arbitrarily to be at $\theta = 90°$. Since θ_0 is chosen to be coincident with the dedendum of the tooth profile, its value is dependent on the arc of the tooth profile,

$$\theta_0 = \frac{\pi}{2} + \theta_D \ \ \text{[radians]} \ , \tag{3.57}$$

where θ_D is the half-angle of the tooth profile, defined in the next section. The final parameter, the amplitude of the axial profile, is used to tailor the initial angle of the teeth — analogous to the helix angle for helical gears. Revisiting Eq. (3.51) and treating $r(\theta - \theta_0)$ as y again, the angle at the edge of the gear, β, can be found by

$$\left.\frac{dy}{dz}\right|_{z=0} = A\frac{\pi}{w}\cos\left(\frac{\pi}{w}z\right)\Bigg|_{z=0}$$

$$\tan\beta = A\frac{\pi}{w},$$

which can be rewritten to give the amplitude as a function of a desired initial gear tooth angle.

$$A = \tan\beta \frac{w}{\pi} \tag{3.58}$$

For this design, an initial angle of 25° is selected, as it is similar to standard helix angles and it provides sufficient twist such that during meshing, no single tooth fully disengages before the next set of teeth engages.

3.4.2 *Tooth Profile*

There are two goals that the tooth profile must achieve: teeth that fully sweep the flanks to avoid fluid pocket formation, and smooth meshing that allows quiet and efficient power transmission without creating a leakage pathway. While previous sections have discussed the limitations of pure involute toothed gears, one of the chief advantages of the involute curve is the smooth transfer of force during meshing. Additionally, how a tooth sweeps the flank of a mating tooth is primarily a function of the design of the addendum and the dedendum of the tooth. Therefore, rather than do away with involute curves entirely, the CCAT profile is a hybrid of circular and involute curves.

The hybrid tooth profile consists of three sections: the addendum circular arc, the involute flank curve, and the dedendum circular arc. The two circular arcs have their separate origins located on the pitch circle. The addendum arc origin, O_1, is located at the 12 o'clock position and is where the azimuthal position, θ, is defined to be zero. The dedendum arc origin, O_2, is defined to be at the angle that defines the end of the half-tooth profile. The beginning and end points of the profile are labeled A and D, respectively, while the transition points between circular arcs and the involute curve are labeled B and C. An illustration of the key parameters that define the CCAT tooth profile is shown in Fig. 40.

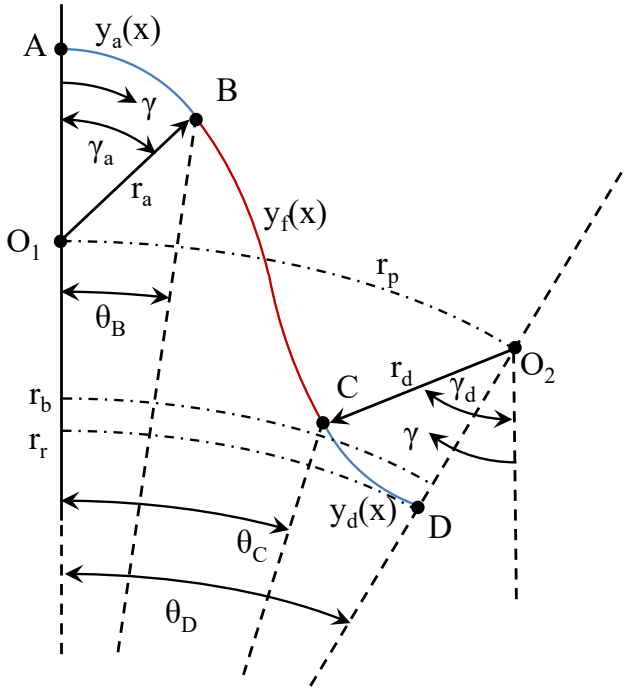

Fig. 40—Geometry of one-half of a tooth profile. Blue lines denote circular curves, red lines denote the involute curve. Curve end points are identified A through D. Pitch, base, and root circles are denoted by the dot-dash lines and are labeled as r_p, r_b, and r_r, respectively.

The arc length of the tooth in radians is determined by the number of teeth, n_t,

$$\theta_D = \frac{\pi}{n_t} \tag{3.59}$$

which also defines the angular position of O_2. Due to the mathematics involved, it is easier to work with parametric equations. The equations for the addendum circular arc are given by

$$x_a = r_a \sin\gamma \tag{3.60}$$

$$y_a = r_p - r_a \cos\gamma \tag{3.61}$$

and the dedendum arc by

$$x_d = x_{O2} - r_d \sin\gamma \tag{3.62}$$

$$y_d = y_{O2} - r_d \cos\gamma, \tag{3.63}$$

where γ is the angular position of the point, (x_a, y_a) or (x_d, y_d), depending on the curve in question, with respect to the vertical axis, as outlined in Fig. 40, where

$$\gamma \in \begin{cases} [0, \gamma_a] & y_a(x) \\ [\theta_D, \gamma_d] & y_d(x) \end{cases}. \tag{3.64}$$

The coordinates of the origin for the dedendum curve are

$$x_{O2} = r_p \sin\theta_D \tag{3.65}$$

$$y_{O2} = r_p \cos\theta_D, \tag{3.66}$$

while the coordinates for point D are

$$x_D = (r_p - r_d)\sin\theta_D \tag{3.67}$$

$$y_D = (r_p - r_d)\cos\theta_D. \tag{3.68}$$

The involute curve of the flank already has been defined by Eqs. (3.19)—(3.22) for both polar and Cartesian coordinate systems. However, Eqs. (3.19) and (3.20) are the general forms of the equations for an involute curve and correspond to a curve originating off the x-axis. In this case, the involute curve $y_f(x)$ originates at a point, C_0, which lies on the base circle, passes through point C, and terminates at point B. Therefore, Eqs. (3.19) and (3.20) need to be rotated by an angle, θ_R, so that the involute curve begins at point C_0, as shown in Fig. 41.

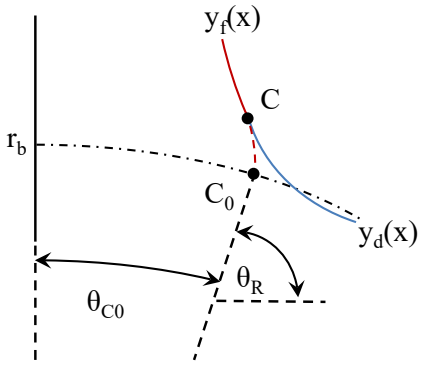

Fig. 41—Origin of involute curve $y_f(x)$. Red, dashed line denotes the part of the involute curve that is not used in the actual tooth profile.

Multiplying Eqs. (3.19) and (3.20) by a rotation matrix yields

$$x = r_b(\cos t + t\sin t)\cos\theta_R - r_b(\sin t - t\cos t)\sin\theta_R \tag{3.69}$$

$$y = r_b(\cos t + t\sin t)\sin\theta_R + r_b(\sin t - t\cos t)\cos\theta_R . \tag{3.70}$$

For the convenience of having all angles be in reference to the y-axis, Eqs. (3.69) and (3.70) are given in terms of the complement of θ_R, called θ_{C0}.

$$x_f(t) = r_b(\cos t + t\sin t)\sin\theta_{C0} - r_b(\sin t - t\cos t)\cos\theta_{C0} \tag{3.71}$$

$$y_f(t) = r_b(\cos t + t\sin t)\cos\theta_{C0} + r_b(\sin t - t\cos t)\sin\theta_{C0} \tag{3.72}$$

3.4.3 *Curve Splicing*

In order to create a complete hybrid profile, the three component curves must be combined and then reflected across the y-axis. The criteria for merging the three curves are that the created curve is continuous and differentiable. The curves defined thus far satisfy the criteria alone, which means that the solution criteria reduces solely to where the curves meet — points B and C. Mathematically, this means that at point B,

$$y_a(x_B) = y_f(x_B) \tag{3.73}$$

$$\left.\frac{dy_a}{dx}\right|_{xB} = \left.\frac{dy_f}{dx}\right|_{xB}, \tag{3.74}$$

and at point C,

$$y_d(x_D) = y_f(x_D) \tag{3.75}$$

$$\left.\frac{dy_d}{dx}\right|_{xD} = \left.\frac{dy_f}{dx}\right|_{xD} . \tag{3.76}$$

Given the use of parametric equations, it is worth noting that

$$\frac{dy}{dx} = \frac{\dfrac{dy}{dt}}{\dfrac{dx}{dt}}.$$

Substitution of Eqs. (3.60), (3.61), (3.71), and (3.72) into Eqs. (3.73) and (3.74) yields the curve splice criteria for point B.

	Circular arc		Involute	
x_B	$r_a \sin\gamma_a$	=	$r_b(\cos t + t\sin t)\sin\theta_{C0} - r_b(\sin t - t\cos t)\cos\theta_{C0}\big\vert_{t=tB}$	(3.77)
y_B	$r_p + r_a\cos\gamma_a$	=	$r_b(\cos t + t\sin t)\cos\theta_{C0} + r_b(\sin t - t\cos t)\sin\theta_{C0}\big\vert_{t=tB}$	(3.78)
$\left.\dfrac{dy}{dx}\right\vert_B$	$-\tan\gamma_a$	=	$\left.\dfrac{1+\tan t\tan\theta_{C0}}{\tan\theta_{C0}-\tan t}\right\vert_{t=tB}$	(3.79)

Similarly, the curve splice criteria at point C are below.

	Circular arc		Involute	
x_C	$x_{O2} - r_d\sin\gamma$	=	$r_b(\cos t + t\sin t)\sin\theta_{C0} - r_b(\sin t - t\cos t)\cos\theta_{C0}\big\vert_{t=tC}$	(3.80)
y_C	$y_{O2} - r_d\cos\gamma$	=	$r_b(\cos t + t\sin t)\cos\theta_{C0} + r_b(\sin t - t\cos t)\sin\theta_{C0}\big\vert_{t=tC}$	(3.81)
$\left.\dfrac{dy}{dx}\right\vert_C$	$-\tan\gamma_d$	=	$\left.\dfrac{1+\tan t\tan\theta_{C0}}{\tan\theta_{C0}-\tan t}\right\vert_{t=tC}$	(3.82)

3.4.4 *Numerical Solution*

While Eqs. (3.77)–(3.82) provide the criteria for splicing the three constituent curves of the tooth profile, they cannot be solved analytically. Instead, a numerical approach is used where the termination parameters for all three curves are calculated using the above equations. The squared differences between the corresponding values are calculated and summed, and the remaining unknown parameters are passed through a numerical solving algorithm to minimize the sum error.

$$E_{SS} = \left(x_{B,circle} - x_{B,involute}\right)^2 + \left(y_{B,circle} - y_{B,involute}\right)^2 + 0.1\left(\left.\frac{dy}{dx}\right|_{B,circle} - \left.\frac{dy}{dx}\right|_{B,involute}\right)^2 \ldots$$
$$\ldots + \left(x_{C,circle} - x_{C,involute}\right)^2 + \left(y_{C,circle} - y_{C,involute}\right)^2 + 0.1\left(\left.\frac{dy}{dx}\right|_{C,circle} - \left.\frac{dy}{dx}\right|_{C,involute}\right)^2 \quad (3.83)$$

It should be noted that due to the sensitivity of derivative calculations in Eqs. (3.79) and (3.82) to small changes in values, the errors in the derivative are weighted less than the errors in position. This is to prioritize the solver to make a continuous curve, as otherwise, it prioritizes matching the derivative at the expense of a single, continuous curve. The Excel Solver is used to minimize sum of the squared errors, which makes use of the generalized reduced gradient code. The variable parameters in the curve fitting and their restrictions are listed in Table 4.

Table 4—Curve-Splicing Parameters and Their Range Restrictions

Parameter	Range Restriction
γ_a	$0 \leq \gamma_a \leq \pi/4$
γ_d	$\theta_D \leq \gamma_d \leq \pi/4$
t_B	$t_C \leq t_B$
t_C	$0 \leq t_C$
θ_{C0}	$0 \leq \theta_{C0} \leq \theta_D$
r_b	$r_r \leq r_b \leq r_p$
b	$a \leq b$

The solver then iterates on the parameters until an optimal solution is found. Decreasing the constraint precision and the convergence values helps ensure a smooth solution (e.g., both were set to $1x10^{-10}$). The resulting curve is shown in Fig. 42. This curve is then reflected across the y-axis to create the full tooth profile.

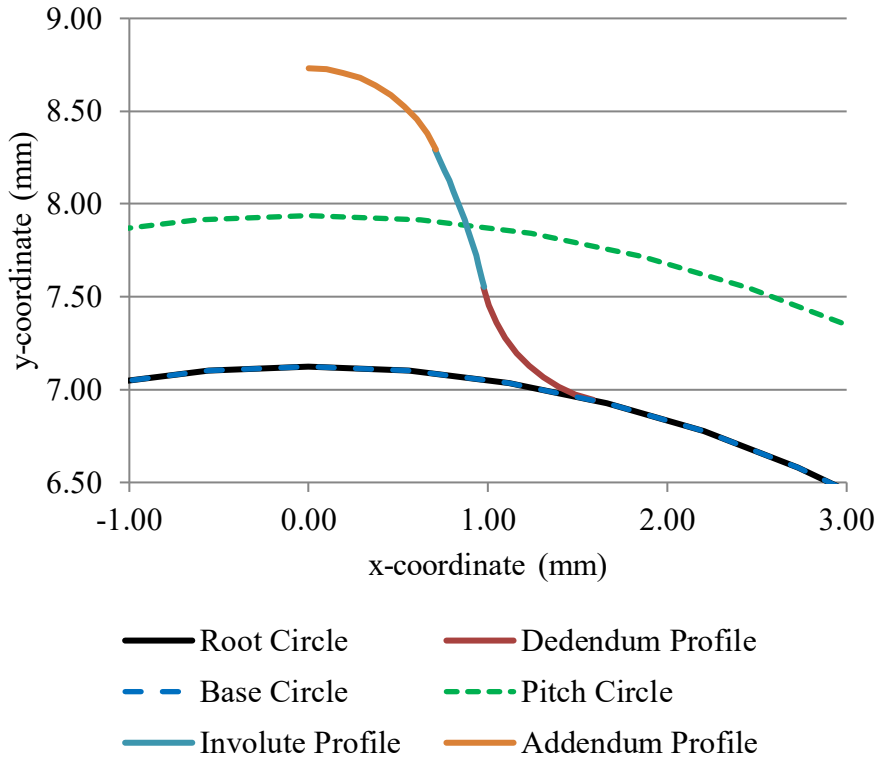

Fig. 42—Post-splice tooth profile

Once the tooth profile is complete, the gear is generated in a CAD software package; for this design, NX 10 was used. First the root cylinder is created, of radius r_r and height w. Next, the bore for the shaft is cut, along with the keyway. Then the tooth axial profile is created as a 3D parametric curve. The tooth profile is then created using 67 data points calculated using the tooth profile parameters; these data points are imported into NX and are used to generate the profile. The circular arc sections are created using the arc command, while the involute section is created using a fit spline to the corresponding data points. The tooth profile is then closed with a circular arc of radius r_r and then swept along the axial profile, generating

the first tooth. The component features of this tooth are then grouped and propagated in a circular pattern to complete tooth formation. The process is illustrated in Fig. 43.

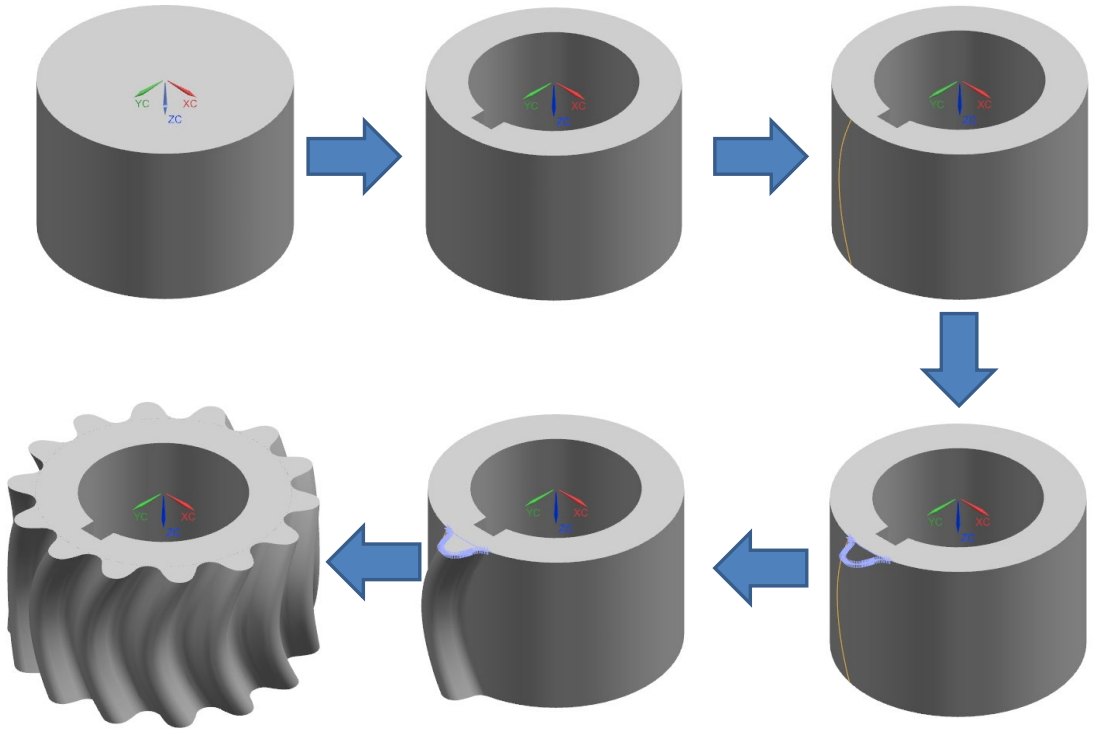

Fig. 43—Graphical outline of the gear design process in CAD

3.5 **Material Considerations**

Many of the factors discussed thus far will have a strong influence on the choice of materials from which to fabricate the gears. The first point of consideration is the tooth-bending stress, which will be a factor of the tooth profile, number of teeth, and the nominal pressure differential in the pump. While it is feasible for low-pressure pumps to utilize plastics or softer metals like aluminum for the gears, high-pressure pumps usually will require high-strength steel gears. A second consideration is the choice of material for the casing. A general rule for gear pumps is for the gear to be made of a harder material than the casing. Paired with a slightly undersized casing, this allows for the gear to run-in the casing to provide a tighter radial seal to reduce leakage. The higher material hardness also pairs well with ensuring sufficient tooth strength and reducing wear.

Another restriction that is relevant in the design process is material compatibility and the various manifestations that can take. When designing a pump, it is important to consider the working fluid and whether that introduces any restrictions on material selection. For most hydraulics applications, this is rarely an issue, as petroleum-based hydraulic fluids (e.g. mineral oil) are nonreactive with metals and most plastics. For more chemically active working fluids, such as salt water or certain chemicals, care should be taken to examine material reactivity. Additionally, pumps designed for the food or health care industries will have stricter requirements on what materials can come into contact with the working fluid.

In addition to the choice of raw material for a gear, a designer also can consider material treatments to achieve the desired characteristics. Heat treatment can provide a harder surface that is more wear-resistant,

either by through-hardening or by case-hardening. It is important to consider any possible distortion from the process that may change final gear dimensions, especially for smaller gears. Other options include anodizing, carburizing, or nitriding to increase surface hardness and to reduce corrosion susceptibility while preserving a softer and more ductile core. High-pressure pumps often will feature some form of treated steel, such as heat-treated 4000 series steel, or some other metal to resist wear and to ensure longevity. It is important to remember that the selection of gear material does not occur in a vacuum and is subject to material and design choices made elsewhere in the pump. As such, though the gears are the first components to be designed for the pump, a designer should not fear revisiting these choices during the design iteration process.

3.6 **Fabrication Methods**

The final stage of the gear design process is to consider the source and the manufacturing technique used to fabricate the gear. There are multiple options available, ranging from simple, extruded-bar gear stock to hobbing and grinding. For general power-transmission purposes, the distinction often does not matter, but for high-efficiency fluid pumping at small scales, having a regular and sharply defined gear tooth profile is important for limiting leakage. Extruded gear stock offers a cost-effective source for standard gear sizes of either spur or helical gears, but it restricts the designer to standard gear sizes, pitches, etc. Additionally, extruded gear stock is prone to profile irregularities that increase leakage through the mesh and decrease volumetric efficiency. Gears produced through hobbing and grinding offer a cleaner profile, and for standard sizes can be obtained affordably. However, nonstandard gears may require additional cost for relatively small quantities. A compromise approach for a nonstandard spur gear design is to use wire electrical discharge machining (EDM). This method can produce a sharp tooth profile with a relatively smooth surface finish, potentially as low as 16 μ-in Ra, depending on the process. Microscope imaging of the tooth profiles of involute spur gears fabricated from extruded bar stock and wire EDM are presented in Fig. 44; note the difference in the tooth profile between the two methods. Some multi-axis wire EDM platforms offer the ability to fabricate helical gears using this method, though this may depend on the tooth profile and the gear size.

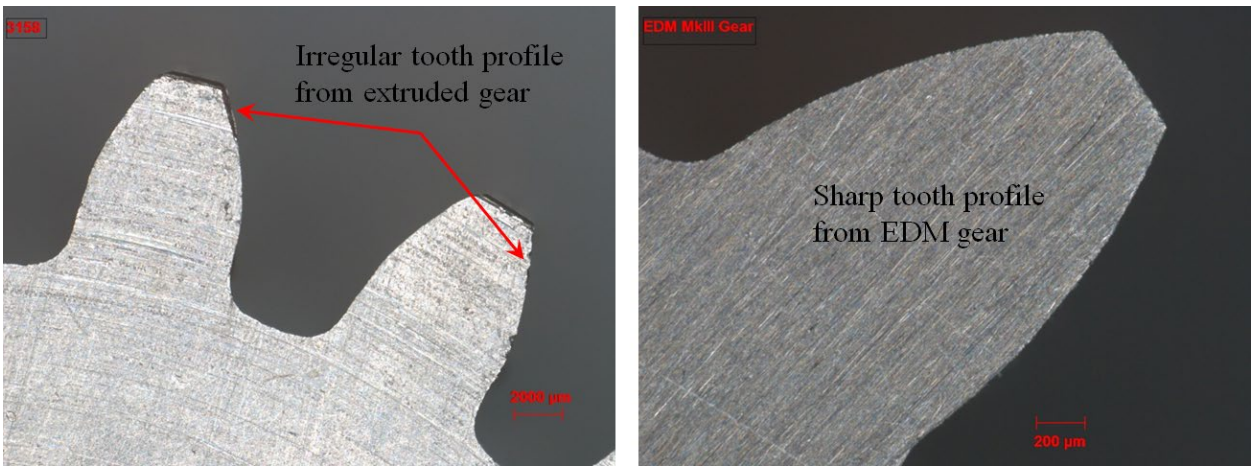

Fig. 44—Microscope images of tooth profiles from spur gears fabricated from extrusion (*left*), and wire EDM (*right*)

The manufacturing process is more complicated for gears with symmetric axial profiles. Some techniques, like wire EDM, are impractical due to line-of-sight issues, particularly for small gears. An additional challenge is that some axial profiles, like the herringbone, create sharp angles that are difficult to machine. A general rule of thumb is that an axial profile that is not continuously differentiable (i.e.,

smooth) likely will pose fabrication challenges. Even a rounded herringbone can be problematic depending on the size of the gear. The herringbone gear pictured in Fig. 37 was actually fabricated by riveting together two helical gears with the same pitch diameter (0.625 in.) and tooth profile but opposite helicity, and then grinding the two axial faces of the gear to smooth out the rivets. However, during the riveting process, the two halves came slightly out of alignment, which means that during the mesh, one-half of the tooth will be less engaged than the other, creating a leak within the mesh. Thus, while some unorthodox processes can produce some gears at a pitch diameter less than what standard gear-fabrication techniques could offer, it can come at the cost of increased leakage and reduced volumetric efficiency.

Another option for gear fabrication that has become more applicable in recent years is the use of computer numerical control (CNC) milling machines to cut the gear from stock material. The precision and resolution of CNC mills have increased sufficiently in the past 10 to 20 years that even complex, small-diameter gears can be fabricated with a fine surface finish. As an example, Fig. 45 shows a photograph of two gears from the case study in Section 3.4 fabricated using a Kitamura Mytrunnion 5-axis CNC mill. Additive manufacturing also offers some potential for fabricating complex gears, but the resulting surface irregularities likely will require post-processing to ensure a sufficient surface finish, particularly for small gears.

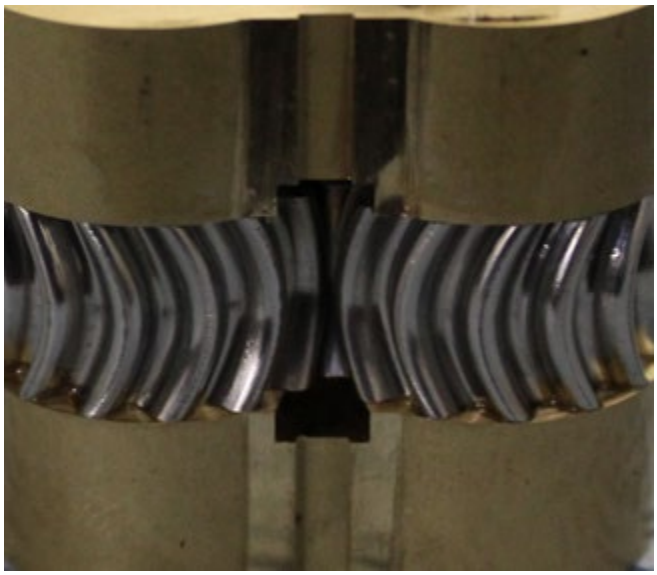

Fig. 45—CCAT gear machined using a 5-axis CNC mill

4. WEAR PLATES

A primary design challenge for gear pumps is to maintain volumetric efficiency by limiting internal fluid leakage. It was discussed in Section 2.1 that the two primary forms of leakage are radial leakage around the teeth of the gears and axial leakage across the axial face of the gears. Radial leakage is determined by pump speed, the pressure differential across the pump, and the clearance between the gear teeth and the casing. As discussed in Section 3.5, a common approach is to design the pump with a very close radial tolerance between the gears and the casing, allowing the gears to wear the casing in and ensuring a tight seal. Related to this is the task of minimizing the axial leakage. The most basic approach is to house the gears within a pumping chamber with minimal axial clearance. While simple and relatively cheap to fabricate, this method is highly dependent on assembly clearances and can lead to excessive wear on the casing. As such, this approach is generally only used for simple, low-pressure pumps.

The more thorough approach is to use wear plates, which fit against the axial faces of the gears to provide a seal against the gears with a surface that is more readily replaceable than the pump casing itself. There are two overall design philosophies for wear plates: a fixed-plate approach in which the axial boundaries are fixed by the pump assembly and the gear is either constrained within by shaft placement or allowed to float between the plates; the other method is to use floating plates that are hydrostatically loaded against the gears [15]. Fixed-plate designs are highly dependent on placement of the gear/shaft assemblies for maximum efficiency and often are optimized for a specific set of operating conditions. Floating-plate designs, in contrast, are capable of maintaining high volumetric and mechanical efficiency across a wider range of operating pressures but require a careful balancing of pressures on the wear plates that necessitates a more involved design.

Beyond this choice of configuration, the design choices that must be addressed include the features that characterize the functionality of the wear plates, such as the relief groove, shaft bearing elements, and any pressure equalization channels. As with the gears, a designer also must take careful consideration of how the wear plates fit into the system as a whole, as well as the fabrication of the wear plates, specifically the choice of material and the fabrication method.

4.1 Types of Wear Plates

The core design function of wear plates is to suppress axial leakage, which is often the primary component to internal pump leakage. To understand why, recall in Section 2.1 the analytical formulation for the two forms of leakage. In radial leakage, the rotation of the gears imparts a shear stress on the fluid in opposition to the pressure gradient out of the discharge region. This superposition of a Couette flow and an adverse pressure gradient is what creates the opposing pressure and pump speed terms in Eq. (2.3). Thus, radial leakage can be suppressed by a combination of tight clearances and increased pump speed. In contrast, there is no such clear tension between opposing factors to suppress axial leakage. A formulation for the added momentum to the fluid film between the axial face of the gear and the mating surface is nontrivial to derive, but overall the rotation of the gears imparts a shear stress on the fluid that has a component along the direction from discharge to suction. Thus, any increase in pump speed will serve only to increase axial leakage, not to reduce it. Therefore, the only mechanism to reduce axial leakage is either to reduce the operating speed and pressure (which usually, for a defined application, is not possible), or to reduce the clearance between the gear and the neighboring surface.

This requirement for a tight clearance between the axial face of the gear and the bounding surface imposes several challenges. The close proximity of a moving surface to a stationary surface will place a shear load on both surfaces even if no contact is made, and in practice, the finite surface roughness of one or both components ensures at least occasional contact. This can be exacerbated by the presence of small particulates in the pumped fluid. Over time, these factors lead to wear of one or both of the surfaces, which will enlarge the gap and will increase leakage. A pump designer essentially has two choices: either some finite gap at the axial face of the gears can be accommodated (set by design or by eventual wear), or the wear process must be accounted for in the design. The former is a viable option for low-pressure pumps where the resulting leakage rate is small, or for low-cost pumps with a limited nominal life span. These pumps may have the gears in contact with the pump casing on all sides and the pump's lifespan is set by wear of the pump components. In some cases, a pump of this design may use a gear material that is softer than the casing to allow replacement of the gears to extend pump life, as gear replacement may be more cost-efficient than replacing the casing.

The other approach is to accept that the cost of close axial clearances is the occurrence of wear and to design the pump accordingly. Well-made gears and precisely machined casings are usually too expensive to replace than simply to allow them to wear continuously (this does not include the initial wearing-in of a

pump assembly). Instead, an expendable component can be added to the pump assembly between the axial face of the gears and the casing that is made of a softer material than the gears. These wear plates then can be replaced over time more cost-effectively than replacing the gears or the casing. The two design practices with wear plates is either to have the wear plates fixed in space relative to the gears or to allow a single degree of freedom such that the gap between the gear and the wear plate is variable. These floating wear plates then rely on fluid pressure to press the plates against the axial faces of the gears. Both approaches have their advantages and disadvantages, and should be considered with the requirements and design goals of the pump in mind.

4.1.1 *Fixed Wear Plates*

Pump designs featuring fixed wear plates generally come in two varieties. The first fixes the placement of the wear plates within the assembly and allows the gear to float along the shaft within some allowable range. The average axial clearance gap thickness is then half of the difference between the separation of the plates and the gear width. The second variety features a constrained "stack" consisting of the gears, the wear plates, and the shaft support such that there is a predefined gap between the gears and the wear plates. This stack is then secured into the pump casing. Both approaches are highly dependent on placement of the components within the assemblies. This can take the form of either fine-tolerance components to secure the wear plates or very precise shimming within the stack to axially center the gears within the pumping chamber.

One of the main design challenges with a fixed-wear-plate pump is ensuring a fixed, nonzero clearance between the gears and the wear plates. Recall in Section 2.1.2 our limited model that showed that the pressure term for axial leakage scales with the third power of the axial gap. As such, even small variations in the axial gap can have large impacts on the leakage flow rate. However, a designer also must keep the size of the gear in consideration. As gear size, and thus pump displacement, increases for a fixed leakage flow rate, the volumetric efficiency of the pump increases. Therefore, gear size should be considered when deciding on the required precision of assembly for fixed-wear-plate pumps.

An empirical example of the impact of the axial clearance is shown in Fig. 46. The data was collected from a small gear pump with a displacement of 0.8 cm^3/rev and a gear pitch diameter of 8 mm. The pump featured no wear plates, and instead allowed the gears to float between the top and bottom of the pump casing. Multiple versions of the gears were fabricated from the same extruded bar stock to create sets of gears with slightly different widths. Combined with a fixed pumping-cavity depth, the different gear widths effectively changed the average axial clearance in an otherwise identical pump configuration. Performance

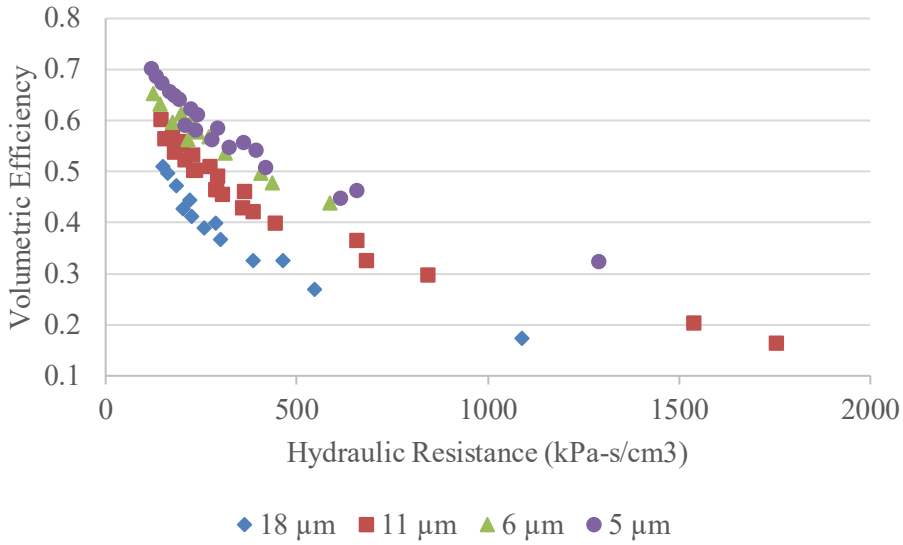

Fig. 46—Pump volumetric efficiency as a function of hydraulic resistance and axial clearance

testing of the pump revealed that the volumetric efficiency of the pump could change by as much as 20% for a change in axial clearance of 13 µm. While larger pumps will have slightly more leeway in internal tolerancing, it is an important demonstration of how sensitive axial leakage is to the gap between the gears and the wear plates.

The primary advantage of a fixed-wear-plate pump is the simplified design process. Floating wear plates, as will be discussed next, are generally more complicated to design for maximum efficiency. In contrast, the challenges of a fixed-wear-plate approach focus more on assembly and component tolerancing rather than the design of said components. There are also generally fewer components in designs featuring fixed wear plates, such as complex seals, that can add to fabrication costs. The main disadvantage to this approach is a narrower performance envelope. Leakage losses and friction losses generally have opposite trends with clearances — the larger the clearance, the greater the leakage flow rate, but the lesser the frictional losses. Thus, selecting the nominal axial clearance of the pump design is choosing a point within a trade space between volumetric and mechanical efficiency. If the gap is too small for a given operating pressure, frictional losses dominate, and if the gap is too large, leakage losses dominate. Another factor is that as operating pressure increases, the fluid power generated by the pump increases compared to the frictional losses, but at the same time, this increased pressure increases leakage. This results in a relatively narrow window of operating conditions for a given axial clearance that maximizes efficiency. Therefore, in practice, fixed-wear-plate pumps tend to be used for pressures below 70 bar.

4.1.2 *Floating Wear Plates*

The main detraction from fixed-wear-plate designs is that the axial clearance is fixed at assembly and cannot be adjusted to meet changes in operating conditions. Floating-wear-plate designs operate on the concept of using axially unconstrained wear plates that are hydrostatically loaded against the gears by the discharge pressure. The axial clearance between the gears and the wear plates is set by a balance of forces. On one side is this hydrostatic load and on the other side is the combination of the discharge pressure in the pumping chamber and the pressure within a hydrodynamic fluid film between the gears and wear plates. Thus, as the pump pressure increases, the force on the wear plates increases, further compressing the plates, decreasing the fluid film thickness, and decreasing leakage. This comes at a cost of higher friction, but at

higher pressures, the delivered fluid power increases at a faster rate than the frictional power loss, maintaining a reasonable mechanical efficiency.

The primary advantage of this configuration is that when properly designed, it allows for a variable axial clearance that automatically adjusts to changes in pump discharge pressure. As pressure increases, the clearance decreases to suppress leakage, but the additional fluid power output by the pump offsets the increased frictional losses, maintaining mechanical efficiency. As a result, a well-designed pump with floating wear plates can operate at a wide range of pressures while maintaining high efficiency. This also aids in pump startup, as it enables a slightly larger axial clearance to facilitate formation of a hydrodynamic lubrication film before reducing the clearance at higher pressures. For these reasons, most high-pressure pumps utilize floating wear plates.

The primary disadvantage of floating wear plates is the added complexity. The addition of floating plates introduces additional concerns that the design must address, such as ensuring a pressure balance between the top and bottom of the pumping chamber or the tilting of the wear plates. Another matter to consider is that the back region of the wear plates that are placed under hydrostatic load must be sealed to prevent fluid leakage around the wear plate and into the pumping chamber. Since the pressurized fluid in the back region is supplied from the pump discharge, this essentially would create another leakage pathway that contributes to volumetric inefficiency. Each of these factors must be considered carefully to ensure maximum pump efficiency.

4.2 Design Features

The features and characteristics of a set of wear plates in a given pump design can vary based on the configuration of wear plate used and the specific needs of the pump design. This stems from the fact that for some pumps, wear plates may provide multiple functions; as such, wear plates may contain additional features to fulfill these functions. One reason for this is the desire to minimize the number of high-precision components. If the wear plates already require the machining of a precise outer profile, there may be only an incremental cost to add a seal gland or a bearing element, rather than to impose that constraint on an otherwise low-cost casing element. Therefore, wear plates can offer a high degree of variability and complexity that can make or break a design. The three main elements of a wear plate are the relief groove, the shaft bearing element, and the pressure channels, shown in Fig. 47. Depending on the pump, some wear plate designs may feature only some or none of these features, but high-performance pump designs generally will feature all three. Therefore, it is important for a designer to be able to understand the function of each feature and the best practices for their incorporation.

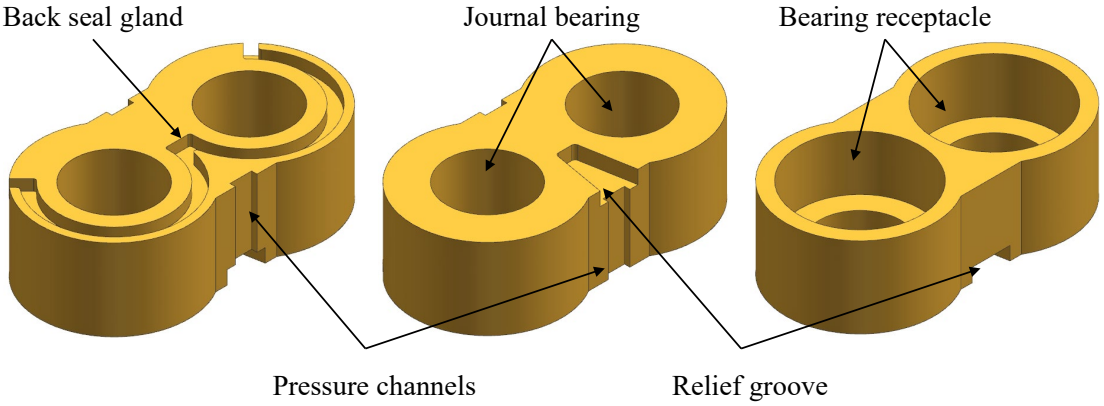

Fig. 47—Design features of multiple variants of wear plates

4.2.1 *Relief Groove*

As discussed in Section 2.3, one of the phenomena that a designer must consider when designing a pump is the compression of fluid pockets within the gear mesh. This can cause pressure spikes within the pump, leading to increased noise and wear of components. One method to minimize these pressure spikes is to utilize gears with an axial profile that allows gradual meshing, which was covered in Section 3.3. Another option is to increase the exit area from the trapped fluid pocket, thereby reducing the pressure differential needed to provide the requisite expulsion flow rate. This increased orifice area commonly takes the form of a groove machined into the surface of the wear plate in contact with the gears. There are three parameters to choose when designing the relief groove: placement, depth, and shape. The placement of the relief groove is dictated by the geometry of the gears used, while the shape and depth of the groove are constrained primarily by the wear plate and any shaft bearing elements.

Placement of the relief groove is a balance between competing factors. The relief groove must extend into the meshing zone in order to be of any use in attenuating the worst of the pressure spikes; however, doing so reduces the leakage pathway through the mesh from the discharge to suction, increasing leakage. At a minimum, the relief groove must extend to the point where the trapped fluid pocket is formed — when the meshing teeth make contact. This point is dependent on the exact tooth profile used in the gears, but we can estimate it by reexamining the gear-alignment geometry of involute gears from Fig. 23. If we focus on the angle of orientation of the gear at which contact occurs, θ_s, we can redraw the relevant geometry, shown in Fig. 48.

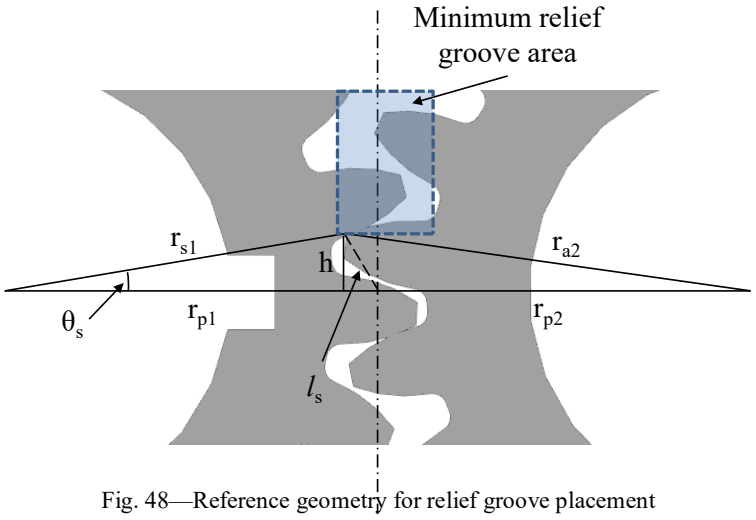

Fig. 48—Reference geometry for relief groove placement

The angle θ_s can be solved for using the Law of Cosines,

$$\theta_s = \cos^{-1}\left(\frac{r_{s1}^2 + r_{p1}^2 - l_s^2}{2r_{s1}r_{p1}}\right), \tag{4.1}$$

which then can be used to solve for the distance from the midpoint of the mesh, h.

$$h = r_{s1}\sin\theta_s \tag{4.2}$$

The value of r_{s1} can be found using Eq. (2.24).

In practice, the edge of the relief groove should extend beyond this amount and into the meshing zone, i.e., the value of h above should be taken as a noninclusive maximum. The actual extent is a trade between pressure pulsation attenuation, volumetric efficiency, and gear design. As an example, pump designs that utilize spur gears benefit more from extensive relief grooves than do pumps that use helical gears, as there is no attenuation of the pressure spikes due to gradual meshing. High-pressure pumps will favor less extensive relief grooves to limit leakage, which can place additional demands on the gear design. Another factor to consider is the pocket geometry as a function of gear engagement through the mesh. The utility of the relief groove diminishes as the cross section of the pocket is reduced below the exit area of the relief groove. While the relief groove in that case would still provide a reduction in the pressure differential between the pocket and the discharge, it may not be worth the reduction in volumetric efficiency. As such, relief grooves rarely extend to the midpoint of the mesh.

The shape of the relief groove also must be considered alongside placement. The geometry of the relief groove must maximize fluid transport within the given constraints of the wear plate design and taking into account the gear geometry. Ideally, the relief groove should cover as much of the meshing zone as possible while not extending below the gear root circle. This is because even before a set of teeth comes into contact during the mesh, the volume between adjacent teeth is compressed by the movement of the mating tooth into the space. As the mating teeth come closer, the effective exit area of the space is reduced, raising the pressure within the space. While this rise in pressure is of a lower magnitude than what is seen in meshing zone, in can contribute to pressure pulsations and noise. In practice, however, machining costs tend to limit the complexity of the relief groove shape. Rectangular channels are the most common, though trapezoidal channels can offer greater flow area for marginal increase in effort. Some rectangular grooves feature circular extensions to maximize coverage at the start of the mesh [4]. Placement of the relief groove always should be on the discharge side of the pump; if the pump is expected to be run reversibly, then relief grooves should be on both the suction and discharge regions.

Once the positioning and shape of the relief groove has been established, the next factor to set is the depth. The core function of the relief groove is to increase the effective exit area from the trapped fluid pocket from a thin band with the thickness of the fluid film, δ_z, to a rectangular orifice orders of magnitude taller. The depth of the groove should be selected based on analytical estimates of the pressure rise using Eqs. (2.9) and (2.33) while not compromising the structural integrity of the wear plate. For this estimation, the exit area should be estimated as the area of the relief groove at the boundary of the meshing zone, illustrated in Fig. 49. The depth of the relief groove will be limited primarily by the geometry of the wear plate, as it must fit within the thickness of the wear plate plus some safety margin. This can be complicated by the presence of any bearing elements in the wear plate. For high-pressure pumps, these bearings are subjected to large stresses, which may impose some minimum material thickness to support. This could result in potential trades between groove depth and groove shape that must be evaluated by the designer.

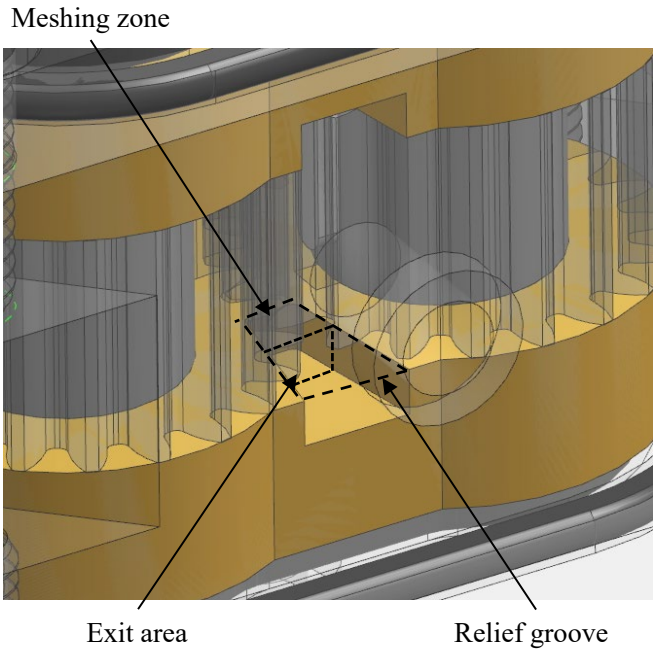

Fig. 49—CAD representation of relief groove in relation to the meshing zone and the effective exit area for the trapped fluid

4.2.2 *Bearing Elements*

The primary function of the wear plates is to ensure a tight seal against the axial face of the gears in a form that protects the gears from wear. However, efficiency of system design often imposes additional functionality to a component to reduce part counts and to simplify designs. Thus, while some wear plates are designed as thin plates solely to provide a tight seal, there is increasing use of wear plates for shaft support. This can come in the form of either journal bearings directly machined into the wear plates, or as recesses to house commercial rolling-element bearings. The topic of shaft support is discussed in detail in Section 5, but it is important to remember that if embedded shaft support is implemented in a wear plate design, it will strongly dictate the wear plate thickness and bore diameter. Thus, wear plate design is often in parallel with bearing design, both initially and throughout the design iteration process.

4.2.3 *Pressure Channels*

One of the key features of floating wear plate designs is that pressurized fluid is supplied to a region behind the wear plates to compress the plates against the gears. One method to provide that fluid is to have small channels or ports machined into the wear plate that connect these back regions to the discharge region. Such ports and other channels also have been used to promote lubrication of bearings or to reduce cavitation [16]. In most cases, these fluid passages are relatively small, as these applications require minimal fluid flow and this helps to minimize the impacts of incidental leakage.

4.3 **Pressure Balance**

The primary purpose of the wear plates is to provide a seal against the axial faces of the gears through a tightly controlled axial clearance between the gears and the wear plate. For fixed-wear-plate designs, this clearance is set by the assembly and the tolerances of the constituent components, while for floating-wear-plate designs, this clearance is determined by the balance of forces on the wear plates. However, during operation, the high-pressure fluid in the discharge region of the pump applies a force on the wear plates

away from the gears. Pumps utilizing fixed wear plates must be designed so that the wear plates are supported within the assembly against this outward force by some hard backstop. In contrast, floating-wear-plate designs must utilize pressurized fluid from the pump discharge to apply a countering force on the opposing sides of the wear plates. This can be done with one or more enclosed volumes on the opposite side of the wear plates supplied with fluid, either at discharge or suction pressure, and separated by static seals. The contact pressure of the seal against the wear plate provides additional compressive force on the wear plates and gears that must be considered. A conceptual illustration of this force balance is shown in Fig. 50.

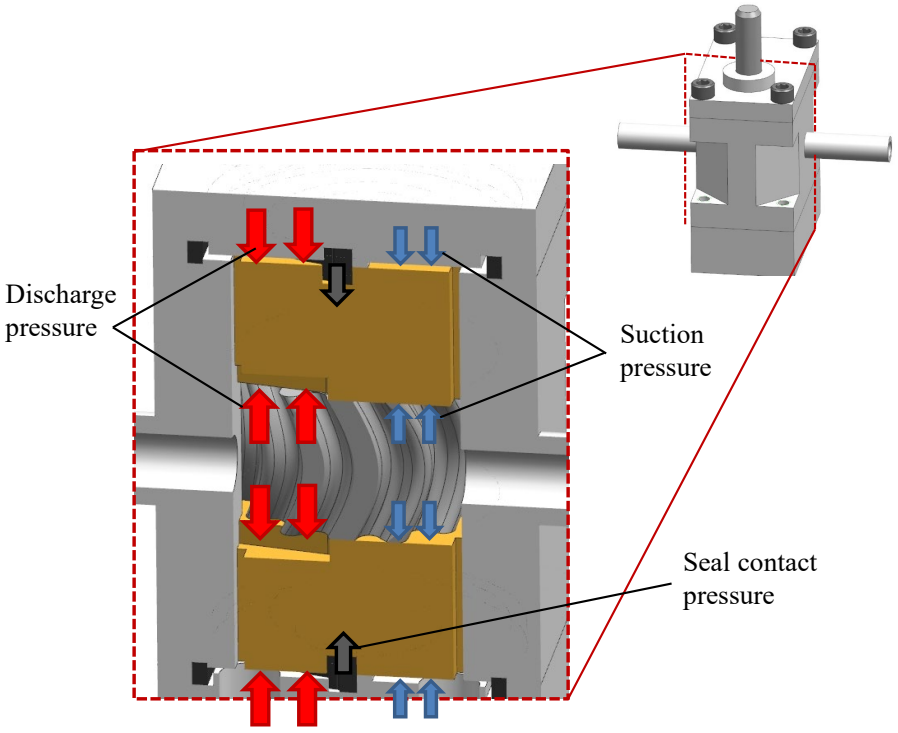

Fig. 50—Conceptual representation of floating-wear-plate force balance in a cross section of a gear pump

One of the primary challenges with floating-wear-plate designs is the proper balance of these countering forces within the stack of wear plates and gears. If there is too much pressure in the back regions, the stack is overly compressed, leading to excessive friction and loss of mechanical efficiency. In contrast, insufficient compression allows the fluid within the discharge region to separate the stack, leading to increased axial clearances and reduced volumetric efficiency. It is important for a designer to understand the trade space between these two competing effects.

4.3.1 *Force Balance Case Study*

As a demonstration of this, consider a gear pump with floating wear plates with a single back region behind each wear plate, similar to the pump design in Fig. 5. These two back regions are connected via a pressure equalization bore within the idle gear shaft and are fed from a single external port. A three-way valve allows this port to be closed, open to suction, or open to discharge through a variable relief valve to allow the setting of the back region pressure. Pressures and flow rates are measured throughout the setup, along with the shaft speed and torque. A schematic of the circuit is shown in Fig. 51. From this data, the mechanical and volumetric efficiencies of the pump then can be calculated.

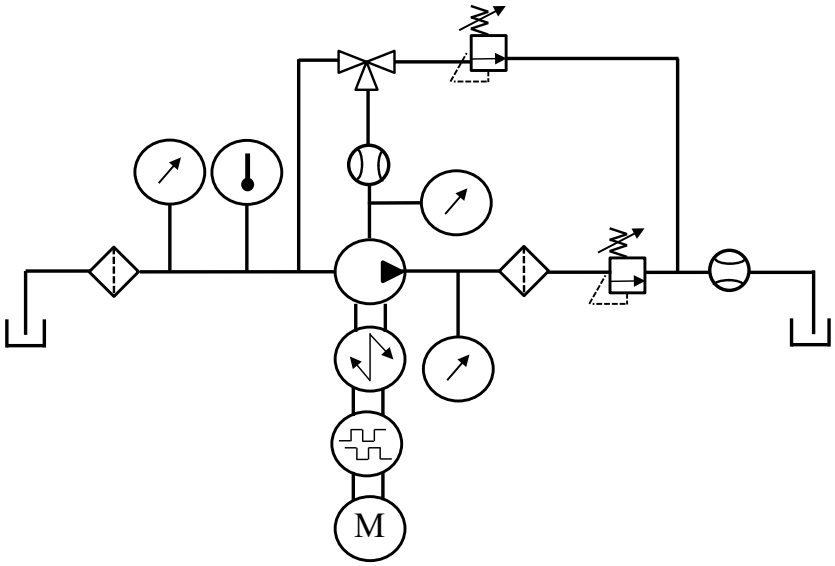

Fig. 51—Schematic of the pump testbed for the force balance case study

The volumetric, mechanical, and total efficiency are presented in Fig. 52 at the same nominal operating conditions as a function of the ratio of the back region pressure to the pump discharge pressure. As this pressure ratio increases, mechanical efficiency decreases while volumetric efficiency increases, as expected. However, this trade-off is not a zero-sum exercise, as the total efficiency increases to a pressure ratio around 0.8, indicating that there exists some optimal force balance that provides the maximum total efficiency.

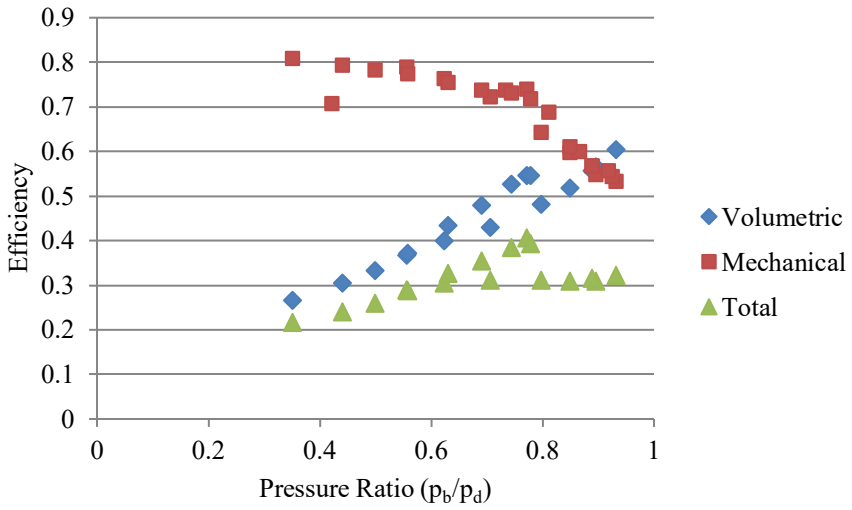

Fig. 52—Pump efficiencies as a function of the ratio of back region pressure to discharge pressure

The problem with the idea of "selecting" a back region pressure is that dropping fluid pressure requires some fluid flow across a restriction, be it a relief valve or some calibrated orifice. Thus, dropping the back region pressure inherently requires some fluid flow — an additional leakage flow path. Since the goal is to balance the forces on the wear plate, the pressure of the back region itself is not important, but rather the

net force it imparts. Therefore, an alternative force balance could separate the back region into two separate volumes: one at discharge pressure and one at suction pressure. The two regions would be distinct and separate to reduce leakage flow between them. This shifts the design goal from setting a pressure to setting the respective areas of the high- and low-pressure regions to give the proper force balance.

4.3.2 *Design Guidelines for Pressure Balance*

The testbed for the above case study was designed such that the bearing drain connection could be varied between the pump discharge or the pump suction, or closed off from the system. Testing the pump with each of the three configurations revealed several design guidelines that are worth noting. The first stems from the assumption that a connection between the bearing drain and the pump output will lead to fluid flowing from the output through the bearing drain into the back region and up through the bearings. Instead, it was found that fluid is driven through the bearings, through the back region, and out the bearing drain as a parallel pathway to the pump discharge. In general, fluid flows in a gear pump originate in the discharge region of the pump and flow outwards from there to sources of lower pressure, be it through the pump discharge to tank, across the gears back to suction, or through the bearings to tank.

The second guideline is that the fluid pressure in the bearings sets the fluid pressure against which the shaft seal on the drive shaft must contain. If the bearings are opened to discharge pressure, either there will be an increased leak at the drive shaft, or the shaft seal must be rated for the resulting pressure with the inevitable increase in frictional losses. For that reason, the bearings should be at suction pressure; an additional benefit is that the pressure differential between discharge pressure in the pumping chamber and suction pressure at the bearings should ensure sufficient bearing lubrication should journal bearings be used. This does come with the consideration that close diametral clearances between the shafts and the bearings as well as the shafts and the wear plate bores are necessary to avoid excessive leakage flow through the bearings.

The third guideline is that a closed-bearing-drain configuration is generally only appropriate for pump designs that utilize rolling-element bearings, such as needle bearings. Journal bearings require some finite fluid flow through the bearing assembly due to the nature of the azimuthal flow in the bearing as well as to flush out debris and to provide convective cooling. A closed-bearing configuration offers the highest volumetric efficiency, as it eliminates a finite fluid loss, but doing so would risk bearing failure in journal bearings.

4.4 **Implementation**

There have been multiple studies to develop methodologies for designing an optimal wear plate force balance, through both experimental [5,15,17] and computational [4,18,19] means. The best approach for a designer will depend on their individual skills and the resources available. However, there are some general guidelines that apply across most pump designs involving material selection, integration into the pump assembly, and fabrication methods.

4.4.1 *Material Selection*

The central function of the wear plate, to provide a seal against the axial face of the gears, requires that the wear plate material is softer than the gears and provides a low-friction interface during operation. The lower relative hardness is important, as it preserves the gears against wear and enables the gears to smooth the wear plate surface during the initial wearing-in process. However, an overly large disparity in hardness can lead to excessive erosion of the wear plate and particulate formation within the pump. For that reason, wear plates made from soft plastics should not be paired with hard, metallic gears, though there can be feasible use of such materials when paired with gears of a similar material in low-pressure pumps. For

moderate-to-high-pressure pumps with steel gears, copper alloys are usually favored for the wear plates, with aluminum bronze as a particularly good choice for its high strength and corrosion resistance. Navy bronze is another option that combines strength with corrosion resistance, though it is somewhat softer and denser. An additional factor for pumps with journal bearings is the metallurgical compatibility between the wear plate and the pump shafts, discussed in more detail in Section 5.1.4.

4.4.2 *Component Integration*

Integrating the wear plate into the pump assembly is an exercise in attention to detail, the importance of which is dependent on the pump displacement, the pressure differential, and the desired efficiency. As discussed in Section 2, internal leakage is a strong function of the height of the flow passage, such as the gap between the edge of the wear plate and the pump casing, and, to a lesser extent, the pressure differential across the gap. At the same time, the volumetric efficiency is dependent on the sum total of internal leakage relative to the pump displacement and the speed. In practical terms, this means that as pump displacement decreases, the impact of internal leakage on volumetric efficiency increases. This can be ameliorated somewhat by increasing pump speed, yet with a potential cost of increased frictional losses and decreased mechanical efficiency. The alternative is to ensure that the wear plates, gears, and casing all have tight tolerances that minimize clearances to what is necessary for assembly. This requirement inversely scales with pump displacement, as a larger pump is more tolerant of internal leakage, but is of particular importance for high-pressure, small-scale pumps.

A related design concern for floating wear plates is the possibility for the plates to rotate around the long axis of the wear plate. This can arise when the pressure loads on the wear plates are not properly balanced. Take as an example the core pump subassembly (gears, shafts, wear plates, and bearings) in Fig. 53. This subassembly corresponds to a pump design with a single-chamber back region at discharge pressure. When the pump is running, the interior pressure is divided between suction and discharge pressures, but the entire back region will rise to discharge pressure. While the net force on the plate acts to compress the subassembly, it also creates a net torque on the plates. As a result of this torque, there is greater contact pressure on the gears in the suction region than in the discharge region. This effect is magnified in the case of no back pressure on the wear plate and can lead to the wear plate's losing contact with the gears in the discharge region. A comparison of an uncompensated wear plate with no back pressure, and a pressure-compensated wear plate after use is shown in Fig. 54. The uneven wear pattern is consistent with the wear plate's rotating away from the gears in the discharge region.

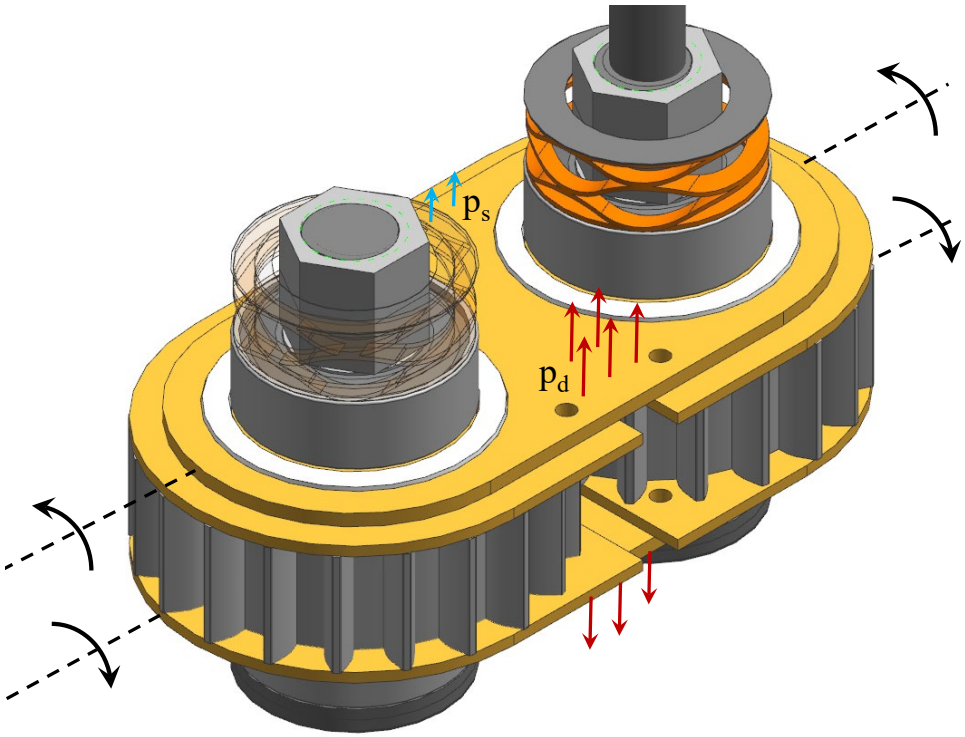

Fig. 53—CAD rendering of a thin wear plate inner pump assembly with a single-chamber back region at discharge pressure. Arrows denote interior pressure on the wear plates. The pressure imbalance creates a torque on the wear plates that compresses the wear plates into the gears at the suction end of the pump.

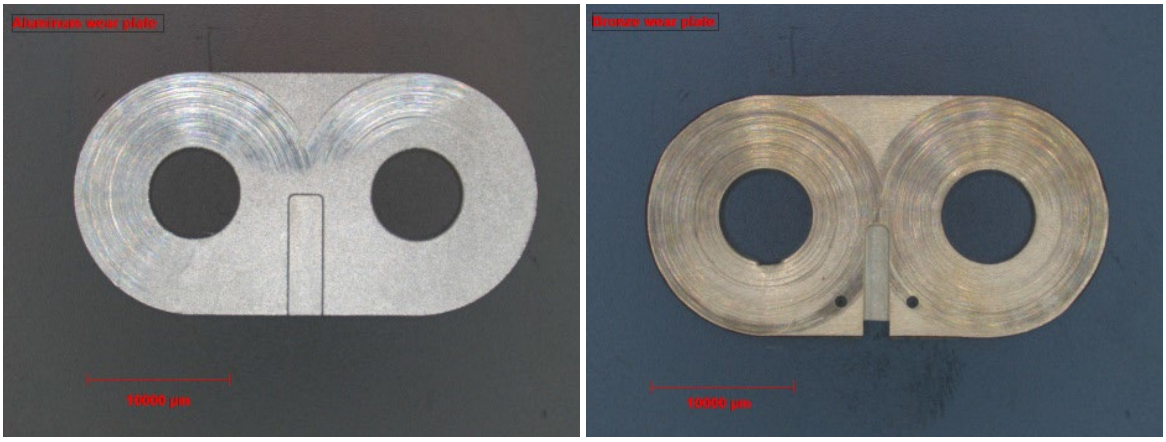

Fig. 54—Comparison of regions of wear for uncompensated (*left*) and pressure-compensated wear plates (*right*). Unworn region of the uncompensated wear plate corresponds to pump discharge region.

The design concern here is that such rotation reduces the hydraulic resistance across the leakage flow path along the axial face of the gear, reducing volumetric efficiency. This has been found consistently in empirical investigations [5,15,17], and can also be the result of imperfections within the assembly or the components. Some examples include bearing misalignment, the axial face of the gear not being square against the axis of the shaft, a nonflat wear plate surface, a taper to the pump casing, or distortions of the pump casing under pressure [15]. Rotation of the wear plates poses an additional problem as the variance in the gap between the gear and the wear plate creates a hydrodynamic bearing that imparts an additional

force separating the wear plates from the gears. Empirical examination has found that this tilt of the wear plate generally extends to the maximum angle allowable by the clearances between the casing and the wear plate.

To quantify this, let us suppose an idealized wear plate and casing geometry where the cross section of the wear plate is rectangular. The wear plate with a width w_p and thickness t is housed within the casing with a bore width of w_b. The width of the wear plate is less than the bore width to allow a radial clearance of δ. The maximum dimension within the wear plate cross section is the diagonal, d, the alignment of which can be defined as an angle with the horizontal, α. The degree of tilt of the wear plate, θ, is defined as the angle between the wear plate surface and the horizontal. The geometric parameters are illustrated in Fig. 55.

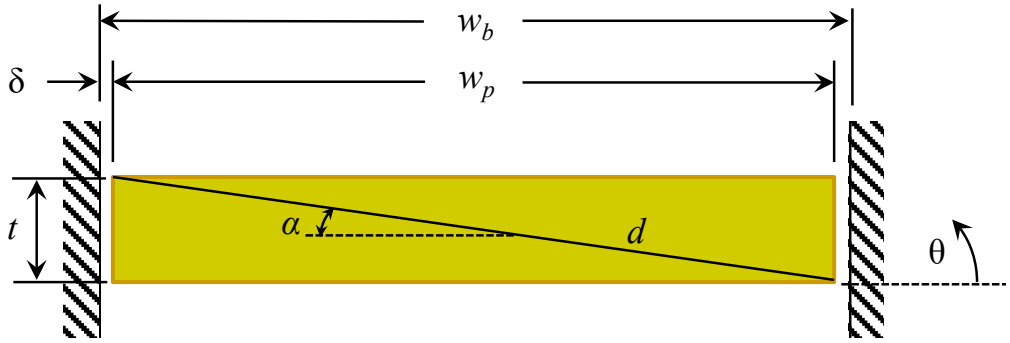

Fig. 55—Idealized geometry of wear plate tilt; long axis of the wear plate is into the plane of the page

The apparent width of the wear plate, w, within the casing is a function of the diagonal and the orientation of the wear plate.

$$w = d\cos\alpha \tag{4.3}$$

We can rewrite the angle of the diagonal to the horizontal in terms of the tilt of the wear plate in reference to the nominal angle of the diagonal when the wear plate is horizontal, α_0.

$$w = d\cos(\alpha_0 - \theta) \tag{4.4}$$

The values of d and α_0 are determined from the wear plate dimensions.

$$d = \sqrt{w_p^2 + t^2} \tag{4.5}$$

$$\alpha_0 = \tan^{-1}\left(\frac{t}{w_p}\right) \tag{4.6}$$

In order for the casing to restrict the tilting of the wear plate at some angle, the diagonal must be greater than or equal to the bore width.

$$d \geq w_b$$
$$d \geq w_p + 2\delta$$
$$\sqrt{w_p^2 + t^2} \geq w_p + 2\delta$$

$$\delta \leq \frac{1}{2}\left(\sqrt{w_p^2 + t^2} - w_p\right) \tag{4.7}$$

If this criterion is met, then the maximum tilt angle occurs when the apparent width is equal to the bore width.

$$\begin{aligned} w &= w_b \\ d\cos(\alpha_0 - \theta) &= w_p + 2\delta \\ \sqrt{w_p^2 + t^2}\cos(\alpha_0 - \theta) &= w_p + 2\delta \\ \cos(\alpha_0 - \theta) &= \frac{w_p + 2\delta}{\sqrt{w_p^2 + t^2}} \\ \theta &= \tan^{-1}\left(\frac{t}{w_p}\right) - \cos^{-1}\left(\frac{w_p + 2\delta}{\sqrt{w_p^2 + t^2}}\right) \end{aligned} \tag{4.8}$$

In pump designs that feature embedded bearings within the wear plate, either rolling-element or journal bearings, the needs of the bearings often require some minimum wear plate thickness that usually ensures Eq. (4.7) is satisfied. Thus, wear plate tilt is generally a concern specific to thin wear plates and ensuring proper fabrication practices to ensure square surfaces and properly aligned bearings. Should a pump design feature thin wear plates, extra care should be taken to ensure adequate force balance and small clearances to prevent tilting and the accompanying loss of volumetric efficiency.

4.4.3 *Back Region Sealing*

One of the challenges with a single-chamber back region design is that the pressurized back region will usually overcompensate for the internal pressure in the discharge region of the pump. As a result, as the pump discharge pressure increases, the net compressive force on the wear plates will increase, leading to reduced mechanical efficiency. To address this, the pump design can utilize a compound seal that not only serves as a static seal for the pump casing, but also divides the back region into high-pressure and low-pressure regions. The two regions then can be connected to the appropriate pressure through a small pressure channel cut into each side of the wear plate. The compressive force on the wear plate is thus the sum of the two pressures applied across each area. The force balance then can be tailored by altering the shape of the seal, thereby altering the total area of the wear plate within each chamber. A CAD rendering of an example two-chamber configuration is shown in Fig. 56.

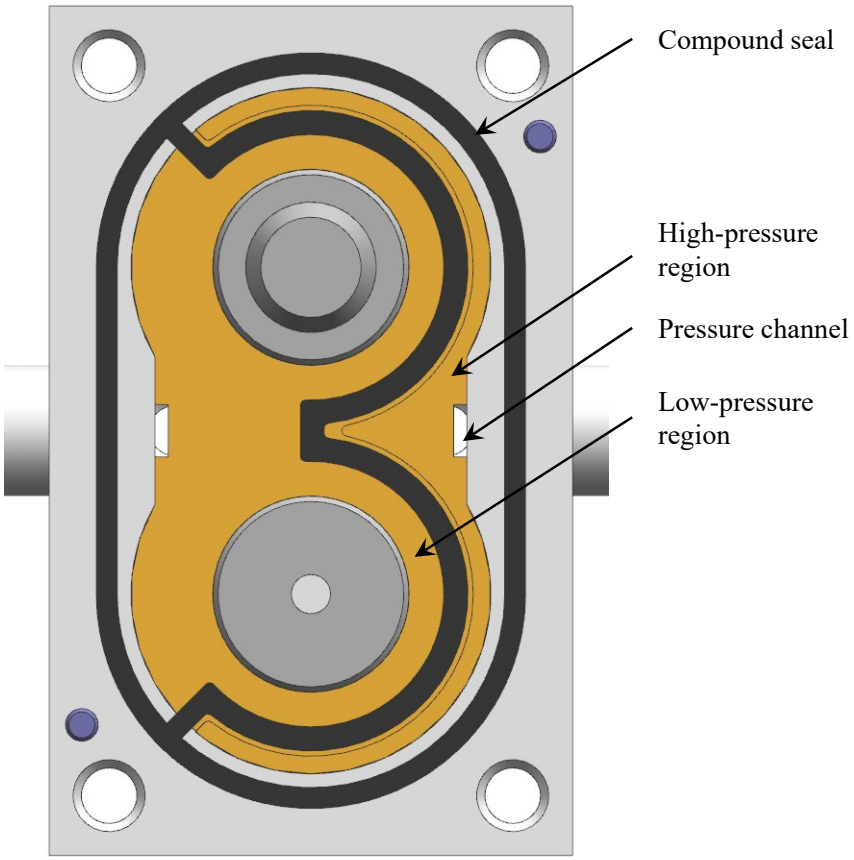

Fig. 56—CAD rendering of an example two-chamber back region using a compound seal. View is from the top with the top casing removed.

As a design rule, the journal bearings should be included in the low-pressure region for two reasons. The first is that establishing a clear pressure differential between inside the pump discharge and the low-pressure region facilitates formation of stable hydrodynamic bearings to support the shafts and to provide convective cooling of the bearings. The second reason is that venting the bearings to the low-pressure region mitigates the potential for high-pressure surges along the bearings and potential leakage from the shaft seal.

The top and bottom case elements are generally the optimal part to house the gland for the compound seal, as these components generally should not experience wear and should not require replacement. A good rule of thumb is to minimize the amount of precision machining in the wear plates, as they are designed to be the first expendable part. The gland should be designed so that when the pump is assembled, the seal is compressed to give a 20% squeeze on the compound seal. The squeeze on a seal is defined as

$$Sq = \frac{d_0 - d_c}{d_0}, \tag{4.9}$$

where d_0 is the initial dimension of interest of the seal (diameter or height) prior to assembly and d_c is the dimension of the seal after compression. The surface of the casing that corresponds to the high and low-pressure regions then can be milled down to create a recessed pocket that, upon assembly, forms the two respective back regions. A CAD rendering of the top casing element and the key design features is shown in Fig. 57.

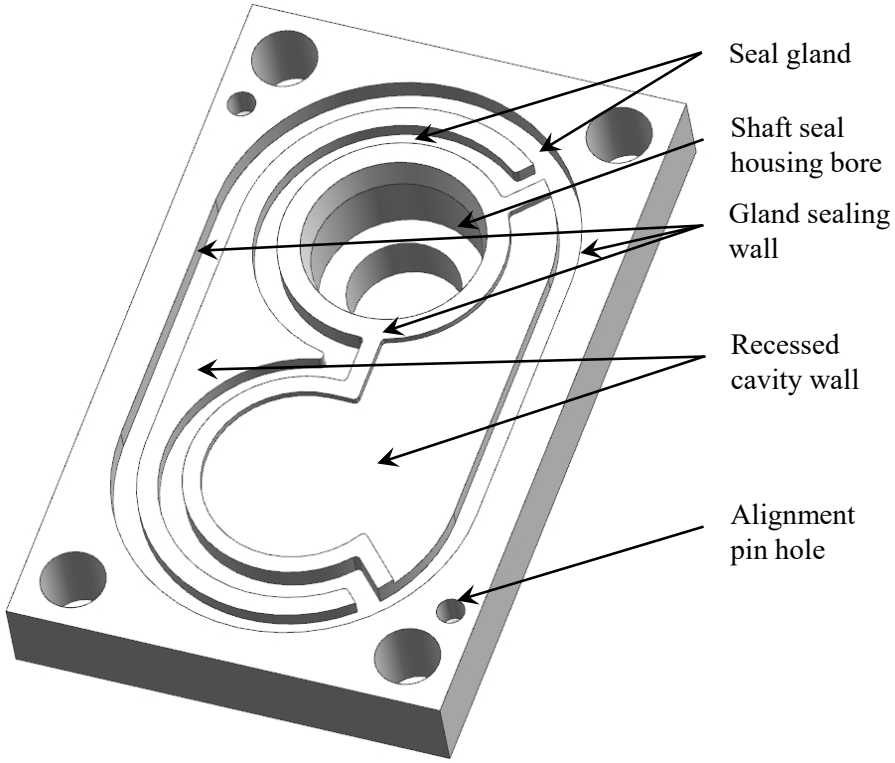

Fig. 57—Top casing design features

There are two main seal-and-gland designs that can be used, one in which the gland was entirely within the casing and one in which there was a mating gland in the wear plate with an extended inner seal, illustrated in Fig. 58. The two concepts are distinguished by the alignment of the sealing surface, either parallel to the wear plate or orthogonal to the wear plate. The key design question with the inner seal is whether fluid pressure on the seal deforms the material sufficiently to generate enough contact stress to seal. At the same time, the mobility of the floating wear plates introduces the potential for the wear plate to move away from the inner seal to reduce the contact stress. Thus, if axial displacement of the wear plates is anticipated to be of sufficient magnitude to reduce the squeeze on the seal, then the orthogonal sealing surface is recommended, though at the cost of increased machining of the wear plate. However, for most pump assemblies, this is generally not required, and the parallel sealing surface is sufficient.

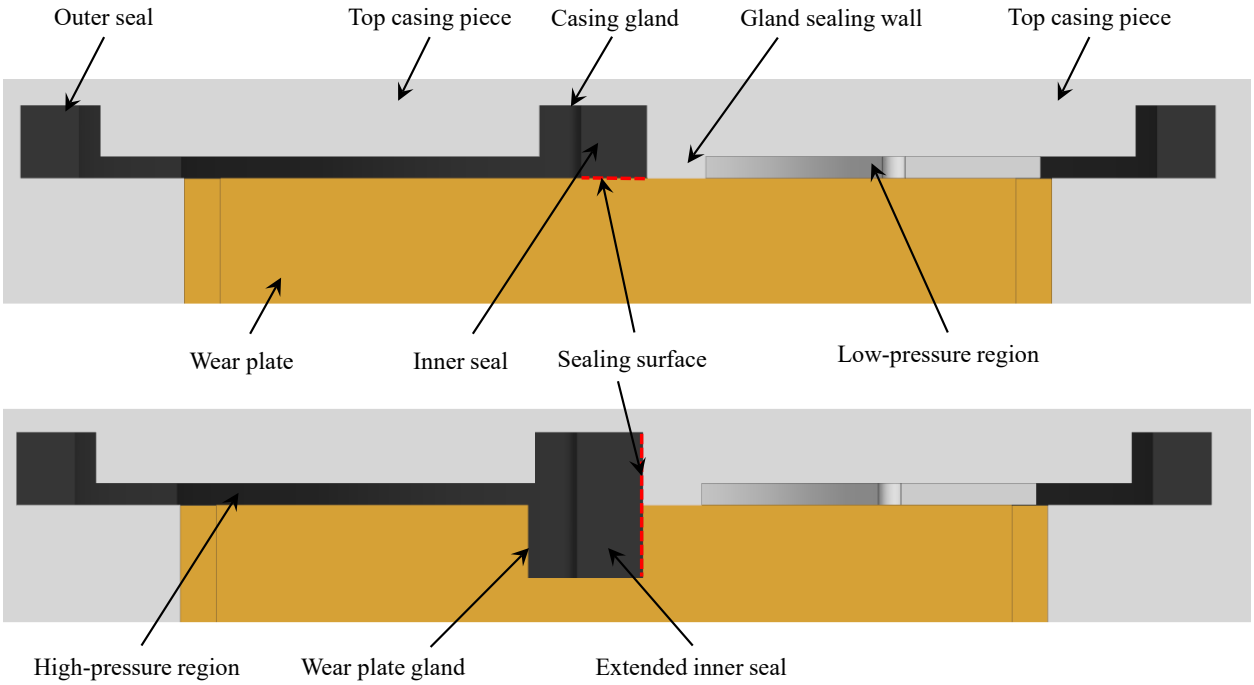

Fig. 58—Cross section of the top casing-wear plate interface and two gland-and-seal design variants: parallel sealing surface with casing gland only (*top*), and orthogonal sealing surface using a mating wear plate gland with extended inner seal (*bottom*)

Iteration is an element within the design process, and as such, it can be useful to quantify the efficacy of the seal. Going back to Section 1.4, we can consider the pump as a circuit with multiple leakage paths in parallel and quantify the leakages by comparing the pressures in each region. Recall from Eq. (1.10) that a flow rate can be calculated as the ratio of a pressure differential to the hydraulic resistance between that differential. For a two-chambered back region of the type we have discussed, we can use the block diagram in Fig. 10 as a reference. The high-pressure region is separated from the discharge region of the pump through the pressure channel in the wear plate, with an effective hydraulic resistance R_c. The low-pressure and high-pressure back regions are separated by the compound seal with an effective hydraulic resistance R_s. If we assume that the leakage flow across the back seal, Q_b, passes first through the high-pressure chamber, we can say that

$$\frac{p_d - p_{b,H}}{R_c} = Q_b = \frac{p_{b,H} - p_{b,L}}{R_s}$$
$$\frac{R_s}{R_c} = \frac{p_{b,H} - p_{b,L}}{p_d - p_{b,H}}. \tag{4.10}$$

This quantifies the effective resistance of the seal in reference to another hydraulic resistance that can be determined either empirically or analytically. The wide range of pressure transducers available enable measurement of the back region pressures directly, even for small-scale pumps, as illustrated in Fig. 59. Depending on the geometry of the specific transducer to be embedded, some modifications to the casing may be required for the duration of the evaluation, such as increasing the casing thickness to accommodate the transducer length. We can assume equal pressure between the top and bottom back regions, as they are connected through their respective pressure channels. The low-pressure region also usually is connected through an inner bore in the idle gear shaft.

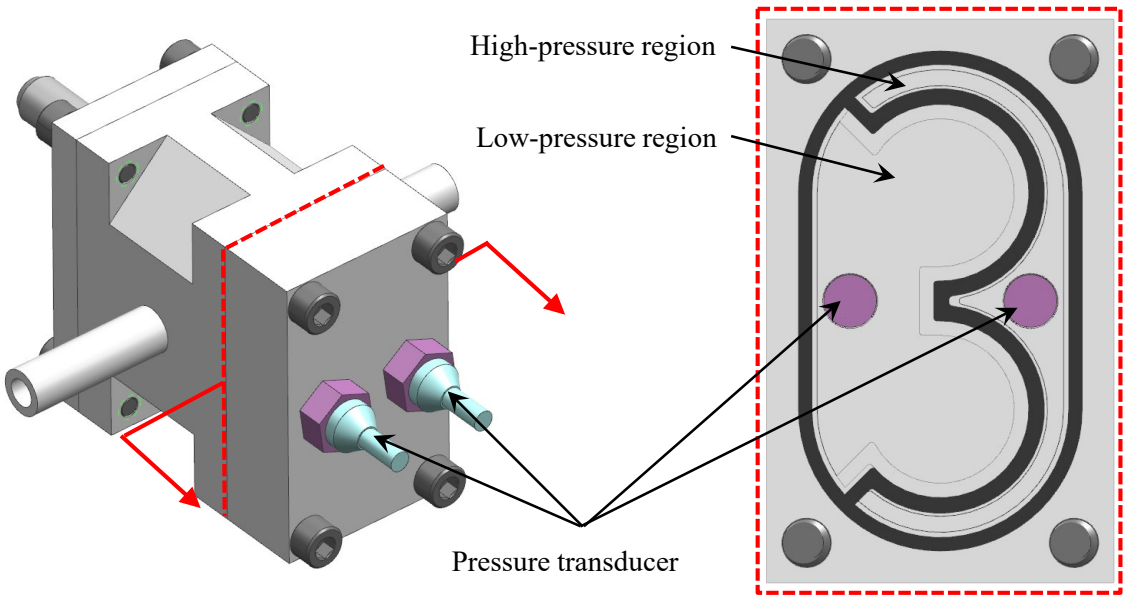

Fig. 59—Modified bottom casing and installed pressure transducers

An example data set is shown in Fig. 60 that, when evaluated using Eq. (4.10), yields an average value for R_s/R_c of 1.89. This indicates that seal provides a resistance to leakage less than a factor of two higher than that of the pressure channel. Considering that the pressure channel is relatively wide, as seen in Fig. 47, and should not be considered a high-resistance orifice, this demonstrates the relative poor performance of this particular seal. In practice, the high-pressure and low-pressure regions should have pressures closer to their corresponding sources, as the majority of the pressure drop in the leakage path should correspond to the largest hydraulic resistance — the seal.

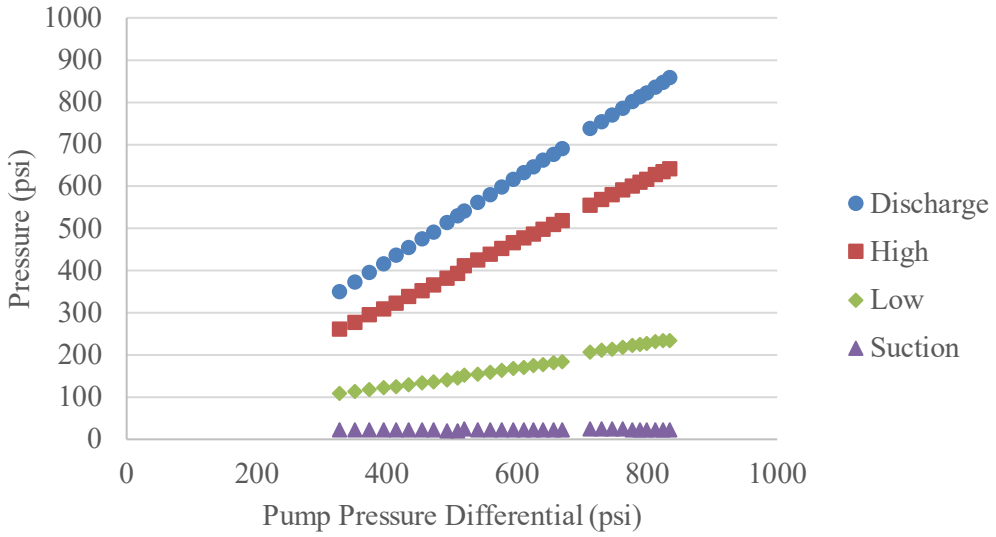

Fig. 60—Example plot of the two back chamber pressures and the pump operating pressures

The data in Fig. 60 is a good example of what to look for when iterating on the pump design and it is important to consider such issues as problems within a pump design, not just a component failure. It is worth revisiting Fig. 56 and noticing that the compound seal is compressed against both the wear plate and the casing. This introduces an additional leakage pathway, as there is the possibility fluid can pass between the wear plate and the casing, but it also creates an underlying assumption that the back face of the wear plate is coplanar with the end face of the central casing piece. If the faces of these two parts are not coplanar, there is a discontinuity in the gland structure at the juncture between the two parts. If the discontinuity is too large for the seal hardness, a gap can occur, which reduces the hydraulic resistance of the seal. As such, it is ideal if the two back regions can be isolated with seals that contact only the wear plate, though this can be difficult for small-scale pumps. This cautionary tale is a reminder for a designer that some performance losses can be the result of the interplay of components without an obvious failure that would present itself upon examination of each component.

4.4.4 *Fabrication Methods*

The wear plate is a difficult balance between high-precision requirements and the fact that it is the expendable component within the assembly. As such, the design must balance tolerancing requirements (and thereby cost) with performance. For this reason, wire-EDM is the preferred machining option, as it can meet strict tolerancing requirements with a satisfactory surface finish, though additional polishing may be required for a journal bearing. The primary limitation with wire-EDM is that the resultant part is an extrusion of a planar shape, and thus any feature not shared throughout the depth of the part, such as a relief groove, must be added in post-machining. A primary concern during machining is the tolerancing on the profile of the wear plate, as it must both fit within the casing and allow two shafts to pass through. An example set of draft dimensions is shown in Fig. 61, which highlights the key tolerances and features.

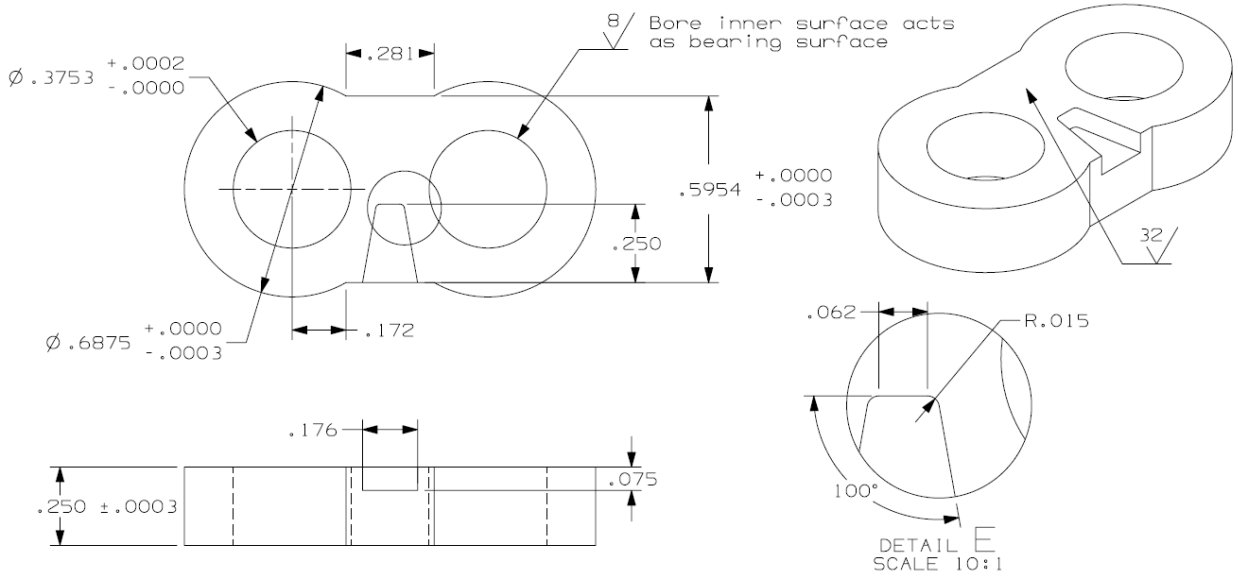

Fig. 61—Example draft of a wear plate and key tolerances

Basic seals generally can be sourced commercially, but something more complex like a compound seal will usually require custom fabrication, such as die cutting or additive manufacturing. The increasing availability of additive manufacturing of polymers provides an easy pathway for complex seal geometries with greater control over material properties. The seal in Fig. 56 was manufactured by Protolabs using an

Objet 260 Connex3 Polyjet machine. The Polyjet utilizes a photopolymer droplet spray that is UV-cured to build the part in 30-micron layers. The process results in a polymeric material that can formed into any seal shape with high resolution as long as there are no voids. The material hardness can be varied between durometer A 30 and 95. For seal applications, testing has found that a hardness between 70 and 95 is effective.

5. SHAFTS AND BEARINGS

The functional purpose of a gear pump is to convert mechanical power from the prime mover into hydraulic power in the form of a pressurized fluid flow. Aside from pumps that utilize magnetic coupling or other non-contacting methods of power transmission, the usual mechanism of coupling to the prime mover is a drive shaft that is connected to one of the gears and extends outside the pump casing. Most pump designs also feature an idle shaft to which the other gear is mounted; this idle shaft is entirely within the pump casing. The secondary purpose of the drive shaft, and the primary purpose of the idle shaft, is to support and spatially constrain the gears during operation. This prevents the gears from contacting the pump housing during operation and the resulting wear, as well as ensuring an optimal mesh of the gears.

During operation of the pump, the generated pressure differential between discharge and suction regions imparts a radial load on the gears. In the absence of any constraints, this radial load would displace the gears into the pump casing, leading to wear and efficiency loss. Depending on application and material selection, such losses may be acceptable, but most pump designs will need to provide some form of compensation to prevent this from occurring. In most cases, this takes the form of radial bearings that support the shafts against this pressure-based force and restrict shaft placement within a narrow tolerance. This combination of shaft and bearings ensures smooth operation of the pump and a steady transfer of input mechanical power to output hydraulic power.

Complicating the task of designing shaft support is the myriad options available, and the challenge of how to discern which ones are most suitable. Some of the most common methods include bushings, rolling element bearings, and hydrodynamic journal bearings, with each offering its own set of advantages and disadvantages. Bushings are simple to design and compact, but can erode, create particulates, and have limited lifespans. Rolling-element bearings are robust and provide low-friction support, but are difficult to design and usually are utilized as stock components that may limit design flexibility. Journal bearings are also compact and capable of supporting large loads but require detailed design and precise fabrication. In pump applications, the chief characteristics of interest for bearings are the load capacity, durability, and the friction of the interface between shaft and bearing. As will be discussed this chapter, the load capacity required will depend on the nominal pump operating conditions and the gear geometry. However, we can make some general assessments of the friction of each bearing type by comparing the shaft torque required to drive the pump as a function of the pump pressure differential. Higher shaft torques for the same pressure load denote increased friction at the bearing. An example set of such data for the pump from Fig. 5 is shown in Fig. 62, and from this we can see clearly that bushings, while simple to design and integrate, operate with a higher frictional load. Combined with the likelihood for increased wear and particulates, we generally can assume that bushings are not suitable except for low-cost, low-pressure pumps. As such, we will focus our discussion on rolling-element and journal bearings.

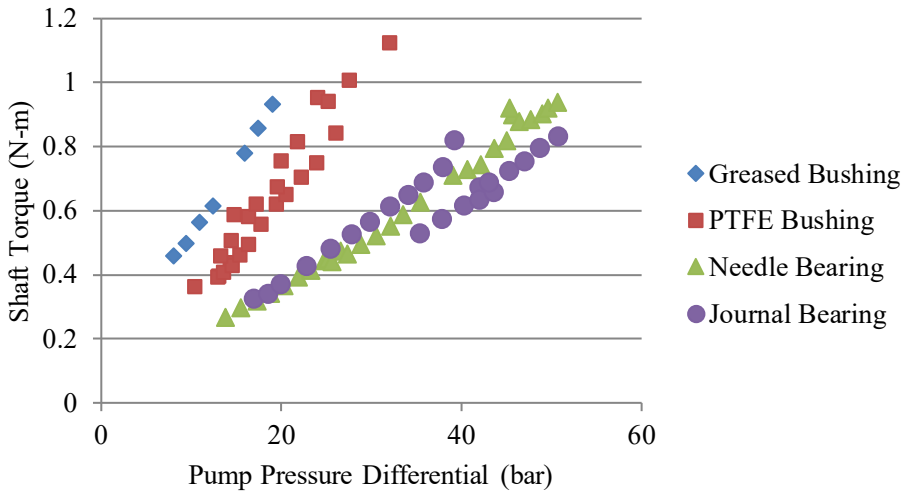

Fig. 62—Comparison of shaft torque as a function of pump pressure differential for four bearing options

5.1 **Shaft Design**

Before discussion can begin in detail on the bearings, we first must start with the shafts themselves, as they drive the design of the bearings in terms of geometry and material selection. The main considerations that a designer should make at this stage include the torsion load carried by the drive shaft, material selection, design features, and fabrication methods. Design of the shaft should be considered in parallel with the bearings (and wear plates, if applicable) and the gears, as these components are interdependent.

5.1.1 *Torsion Load and Minimum Diameter*

The first design requirement of the drive shaft is that it is capable of transmitting the input torque from the prime mover without failing. In this case, failure is defined as the stress on the shaft exceeding the material yield stress, σ_y. The shear stress on a cylindrical rod with an applied torque τ is

$$\sigma_\varphi = \frac{\tau r}{J_T}, \tag{5.1}$$

where J_T is the torsion constant and r is the radial position within the shaft. For a cylindrical rod, the torsion constant is the polar moment of inertia of the cross section.

$$J_T = \frac{\pi r_s^4}{2} \tag{5.2}$$

From Eq. (5.1), we know that the maximum shear stress occurs at the outer surface of the shaft, where $r = r_s$. By taking this value and substituting Eq. (5.2) into Eq. (5.1), we can calculate the maximum shear stress on the shaft.

$$\sigma_\varphi = \frac{2\tau}{\pi r_s^3} \tag{5.3}$$

Recall that in Section 3, we established that the torque load of a gear pump could be estimated by the pump displacement and the pressure differential. Substituting Eq. (3.15) in for the torque in Eq. (5.3), we can quantify the highest shear stress on the shaft in terms of pump specifications and operating conditions.

$$\sigma_{\varphi} = \frac{D\Delta p}{\pi^2 \eta_m r_s^3} \tag{5.4}$$

From the Von Mises criterion for pure shear, onset of yielding occurs at a shear stress of $\sigma_y/\sqrt{3}$, where σ_y is the yield strength of the chosen shaft material. If we add in a factor of safety, X_s, the maximum possible shear stress allowable is

$$\sigma_{max} = \frac{\sigma_y}{X_s\sqrt{3}} \tag{5.5}$$

If we then substitute Eq. (5.5) for the shear stress term in Eq. (5.4), we can solve for the minimum allowable shaft radius as a function of the nominal pump operating conditions.

$$r_{s,min} = \left(\frac{\sqrt{3}X_s D\Delta p_{nom}}{\pi^2 \eta_m \sigma_y}\right)^{1/3} \tag{5.6}$$

5.1.2 *Design Features*

There are several other design features relevant to the pump shafts that are relevant beyond torque transmission capacity. A secondary consideration to transferring the input mechanical power to the fluid is the connection between the shafts and the gears, particularly for the drive shaft. In some pump assemblies, the gear and the shaft are relatively fixed, either machined together as a single piece, or two separate pieces permanently joined. Another common design practice is to mount the gear onto the shaft in a reversible process to facilitate component replacement. With the exception of the case in which the shaft and gear are machined as a single part, the pump design will feature some method of securing the gear to the shaft that is capable of transferring the input shaft torque to the gear. Two common methods to achieve this are to use an interference fit or a shaft key, or a combination of both. Interference fit specifications vary based on the material used and the shaft geometry and offer a simple and cost-effective mounting method. The risk is that for a high torque-to-shaft-radius ratio, the gear could slip on the shaft and decrease the transmitted torque. In small-scale or high-pressure pumps, shaft keys offer a more robust means of transferring torque at the cost of additional complexity.

One design trade for the drive shaft is the competing effects of the shaft diameter on the bearings and the shaft seal. As will be discussed further this chapter, the load force that a bearing can sustain increases as the shaft diameter increases. In contrast, a larger shaft diameter increases the surface speed at which the shaft seal contacts and provides a larger leakage area, placing a larger burden on the shaft seal design. A common method to resolve this design tension is to step down the drive shaft diameter as it nears the casing wall. This reduces the demands on the shaft seal while allowing a larger diameter bearing, but the minimum shaft radius from Eq. (5.6) still applies.

Another element to consider when designing the shafts is the assembly process. Gear pumps in general rely on tight tolerancing to enable minimum clearances to maintain internal sealing and to ensure high volumetric efficiency, particularly in high-pressure pumps. Most of the assembly clearances will be very small, and as such, strategic placement of chamfers at sharp edges can facilitate assembly. For this reason, it is recommended to put 45° chamfers at the top and bottom edges of the shaft, as well as at any change in shaft diameter in the drive shaft. These chamfers can be as small as 0.01 inches, but can ease installation of

press-fit components. Some commercial shaft seals may have additional recommendations on shaft chamfers, so be sure to consult all requisite guidelines when designing the shafts.

Finally, one minor feature of the idle shaft that is easy to overlook is an internal bore in pump designs featuring floating wear plates. This channel provides a means to ensure pressure equalization between the two back regions behind the wear plates and ensures a balanced load on the gears. Due to the limited flow required to achieve this pressure balance, the diameter of this passage can be very small.

5.1.3 *Fabrication Guidelines*

When planning shaft fabrication, there are three factors to consider: shaft material, surface finish, and tolerancing. Selection of the material for the shaft generally is driven by a combination of factors, including size constraints, torque load, and seal requirements. In general, the shaft should be made of a hard material to resist wear from the shaft seal and to reduce the risk of stripping at the coupler to the prime mover. A common approach is to use a hardened steel, though this can be relaxed for some low-pressure pumps. A harder shaft material is also better suited to maintaining its surface finish. A low surface roughness is imperative for a shaft for multiple reason. Surface defects can increase stress along the shaft due to stress concentration, thereby decreasing a shaft's overall torque capacity. Additionally, a low surface roughness is required for many shaft seals to minimize leakage and wear of the seal material. Note that many seals require some minimum surface roughness to allow containment of a thin layer of fluid to lubricate the seal. It is worth repeating that the design of the shafts is done in conjunction with the selection or design of the shaft seals and bearings. An example of some of the machining specifications for a drive shaft are shown in Fig. 63.

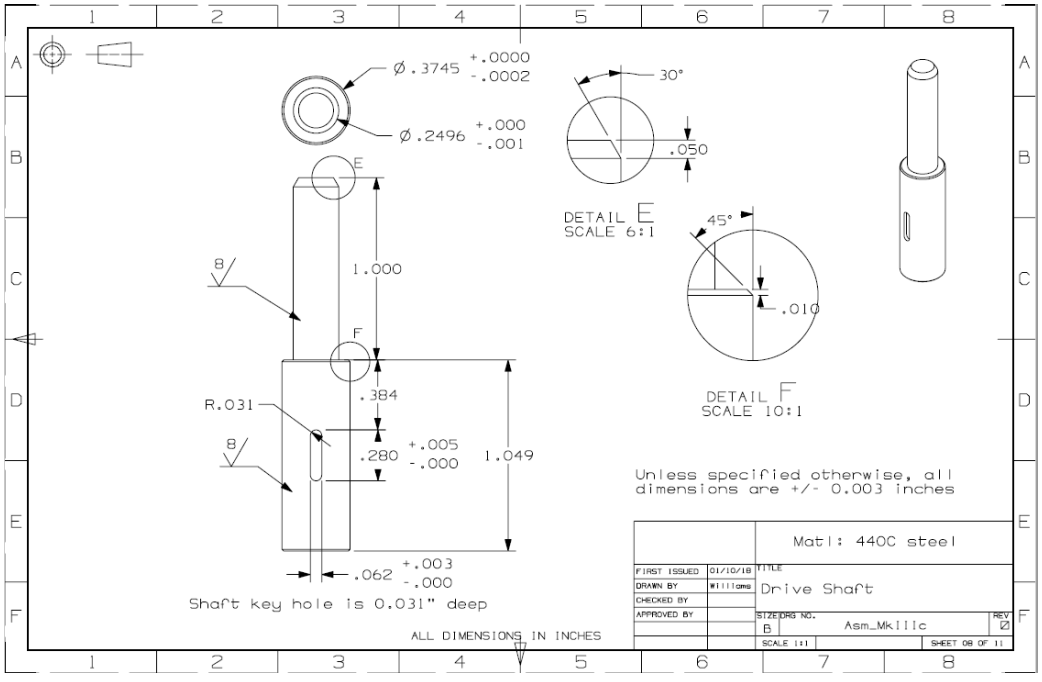

Fig. 63—Example drive shaft fabrication drawings for miniature gear pump

The final factor to consider with the shaft design is the tolerancing required. Each of the functional elements of the drive shaft discussed this far, from shaft seals to bearings, will have specific tolerancing guidelines. For the drive shaft in particular, this can lead to each design feature having a unique set of

tolerances. The effects of poor tolerancing with the bearing will be discussed in more detail further on, but it is important to remember that no component is fabricated perfectly or exactly. Instead, parts are specified within a range of allowable dimensions. While considering tolerances is standard design practice, it is important to keep in mind that the shafts, and the drive shaft in particular, interface with almost every major component in the pump assembly, from gears to casing. Therefore, the drive shaft will have the most opportunity and the least capacity for tolerancing error. A designer should take special care to keep all of the individual tolerancing requirements in mind during design iteration.

5.1.4 *Wear and Material Considerations*

One important wear mechanism to consider when designing the shafts is galling. Galling is a form of adhesive wear in which the asperities of two surfaces begin to contact and deform. During such contact, material can be transferred from one surface to the other, and asperities can break off and contribute to debris within the fluid bearing. This debris then can be embedded into the surfaces, creating larger asperities and further adding to debris formation. As the oxide layers at the surface of the metals are worn off and at high enough contact pressures, the asperities undergo partial cold-welding of the two surfaces and lead to material transfers large enough to see with the naked eye. This is primarily a concern for journal bearings at high loads, often seen in high-pressure pumps. An example of galling on the drive shaft represented in Fig. 63 was photographed using an optical microscope and is shown in Fig. 64.

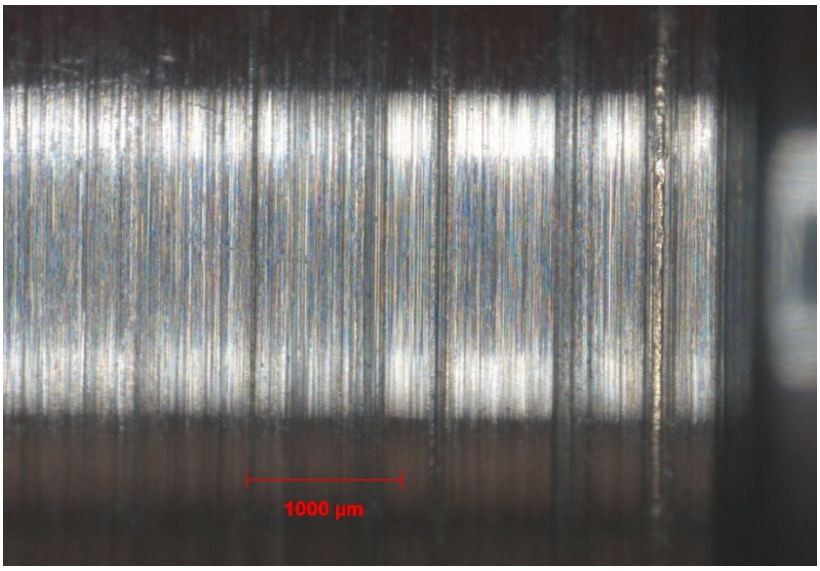

Fig. 64—Microscope image of galling on a drive shaft

Galling and wear impose additional concerns on material selection. As an example, despite 17-4 PH and 300 series stainless steels' having high hardness, they have poor wear resistance [20]. As such, materials such as hardened 440C stainless steel are recommended to combine hardness and wear resistance. Another option is to incorporate shaft coatings to improve wear performance, particularly if journal bearings are used. Some of the factors influencing adhesive wear are mutual solubility or the ability to form compounds from the two materials; these factors are called metallurgical compatibility. For journal bearings, the goal is to avoid material interactions, and as such, we desire metallurgical *incompatibility*. For journal bearings in wear plates, the iron in the steel shaft is compatible with the copper in the bronze wear plate, reducing wear resistance. To counter this, the shafts can to be coated with a material that is incompatible with the bearing material. Metallurgical incompatibility is defined by two conditions [14]:

1. The metals are mutually insoluble and do not form an alloy
2. At least one of the materials must be from the B-subgroup (elements to the right of the nickel group)

Since the desired bearing materials are usually bronze alloys, shaft coatings should incorporate materials that are incompatible with copper at a minimum. Copper is incompatible with molybdenum and chromium, and is partially incompatible with tungsten and niobium; coating options featuring these elements are thus highly advantageous. One simple and widely available option would be to chrome plate the shafts. A more advanced technique is to combine metallurgical incompatibility with other wear-resistant traits, such as lubricity and beneficial mutual material transfer (as opposed to the destructive material transfer in galling). One such process is carbon-based coatings applied by chemical or plasma vapor deposition. Some examples of this include BALINIT CNI and BALINIT C by Oerlikon Balzers. BALINIT CNI uses a sputtering process to deposit a smooth, dense, but thin chromium nitride coating to increase the hardness, wear resistance, and corrosion resistance. SEM imaging of the CNI coating is shown in Fig. 65. Such coatings are suggested for applications where sufficient lubrication is generally available (hydrodynamic or mixed lubrication) [21].

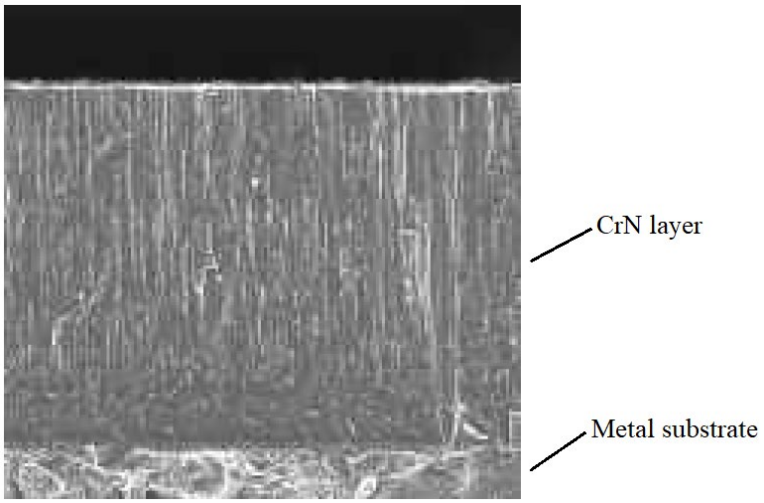

Fig. 65—Coating structure of BALINIT CNI [21]

BALINIT C creates a more complex multilayer structure of many thin layers of tungsten carbide and amorphous carbon matrix on a chromium interface to bond the layers to the substrate, shown in Fig. 66. Both the tungsten and the carbon fulfill the second condition of metallurgical incompatibility, and the successive layers of each yield a low coefficient of friction even in a dry running condition. One of the beneficial features of multilayer coatings is that during wear-in, some of the material will be embedded between asperities in the mating surface, leading to smoother interface between the two surfaces, as shown in Fig. 67. This material transfer is designed primarily for dry-run applications, but it also will occur to a lesser extent in lubricated conditions.

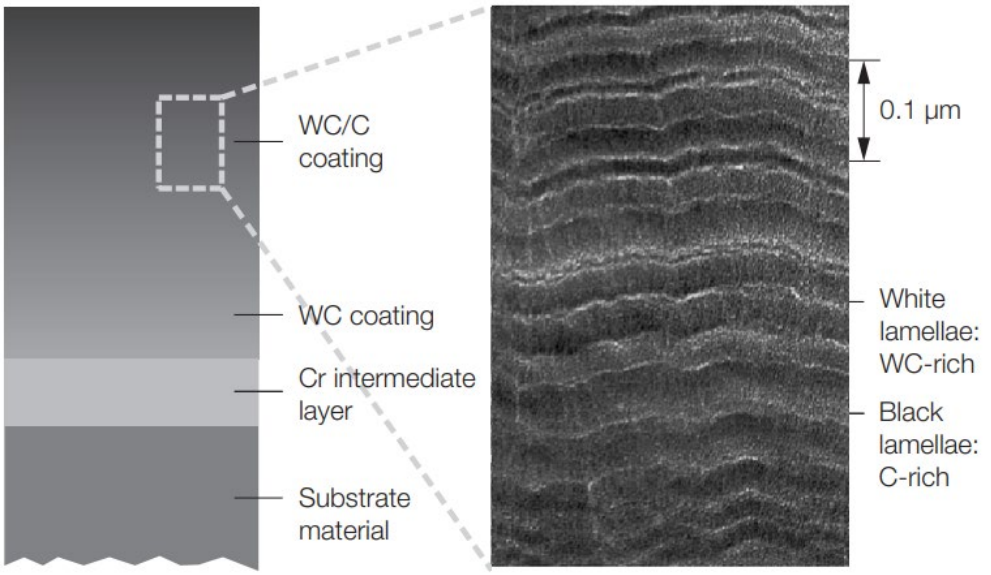

Fig. 66—Coating structure of BALINIT C [21]

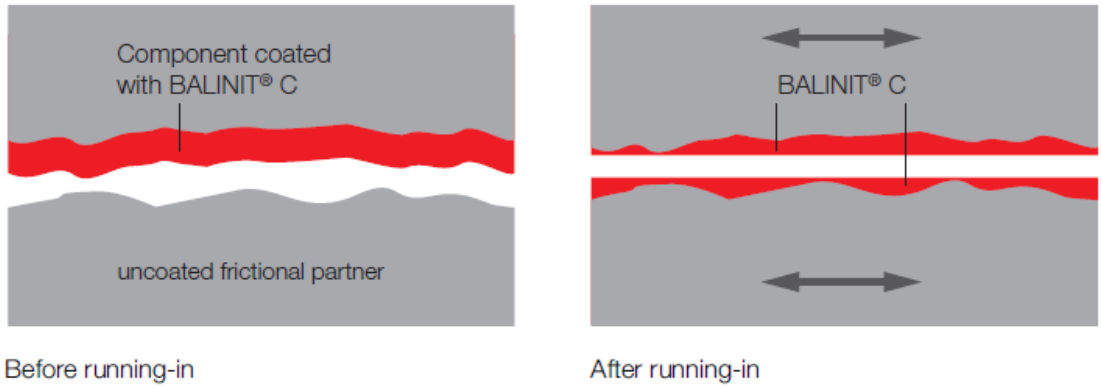

Fig. 67—Run-in material transfer with BALINIT C coating [21]

Different coating options will have their respective advantages and the selection will depend on the specific application in mind. For example, BALINIT C has a lower hardness (1,000–1,500 HK) than BALINIT CNI, but offers superior resistance against abrasive wear due to its running-in and burnishing properties. BALINIT CNI has microcracks within the chromium nitride layer that aid in oil retention for lubricated operation, while BALINIT C has similar characteristics due to pinhole defects, but to a lesser extent than BALINIT CNI. The two surfaces also are designed for separate lubrication regimes. BALINIT CNI is better suited to mixed lubrication, where contact is periodic but not excessive, thus making the high-hardness chromium nitride layer advantageous for resisting surface deformation. This most likely would be the case during pump operation at moderate pressures, where the shafts may run close to the bearing surface. BALINIT C is better in boundary lubrication, and thus would be beneficial during startup of the pump, when the shaft does not have a fully formed hydrodynamic fluid film, or at high radial loads, when the pump operates at higher pressures. Should shaft coatings be a priority for a designer, these kinds of trades will need to be evaluated across the range of possible vendors and coatings.

5.2 Rolling-Element Bearings

Due to the interdependency between components, it is prudent to design the shafts concurrently with the bearings to ensure mutual compatibility and optimal performance. The two most common options for shaft support are rolling-element bearings and journal bearings. Rolling-element bearings, such as ball bearings or needle bearings, feature load-bearing elements that freely rotate between the supported body and the surrounding structure. These bearings can be categorized by structure, such as the shape of the rolling element, whether cages are used to constrain the rolling elements, and whether the bearing has one, two, or zero races. Such bearings usually are sourced commercially rather than designed as custom components; as such, it is more of a bearing-selection process rather than a design process. However, the form factor and specifications of commercial bearings are important to consider in conjunction with the shaft, the wear plate, and the casing design, as each component places constraints on the others.

5.2.1 *Bearing Load*

The first consideration for bearings is the load capacity. Needle bearings offer greater surface contact compared to ball bearings of the same dimensions, and thus can support a larger radial load. An additional benefit for gear pump design is that needle bearings usually have smaller annular thickness compared to ball bearings. This smaller form factor facilitates embedding the bearings into the wear plate or the casing. However, needle bearings are less tolerant to angular misalignment compared to ball bearings. A general rule is that, form factor aside, ball bearings tend to perform better in high-speed, low-load applications.

The net radial load on the bearing can be estimated as the summation of the fluid pressure applied over the area of the gear. The exact pressure distribution is dependent on the gear and casing design, as in practice, the pressure around the gear drops from discharge pressure to suction pressure across the primary hydraulic resistance — the narrow gap between the gear teeth and the casing wall. However, we can approximate this as a continuous function in terms of the angular position around the gear, θ, and simplify the gear as a solid cylinder with radius r_p, and height w. The radial load then can be calculated by integrating the vector pressure force around the circumference of the cylinder.

$$\vec{F}_r = \int_0^{2\pi} -p(\theta)\vec{r}(\theta)w\,d\theta \tag{5.7}$$

As an example, suppose we have a casing design in which half of the gear is exposed to the pumping chamber and half of the gear is against the casing wall, illustrated in Fig. 68. Furthermore, let us assume that of the half exposed to the pumping chamber, half is in fluid at discharge pressure and half is at suction pressure. The remaining half of the gear is exposed to fluid that varies linearly with angular position from suction pressure to discharge pressure.

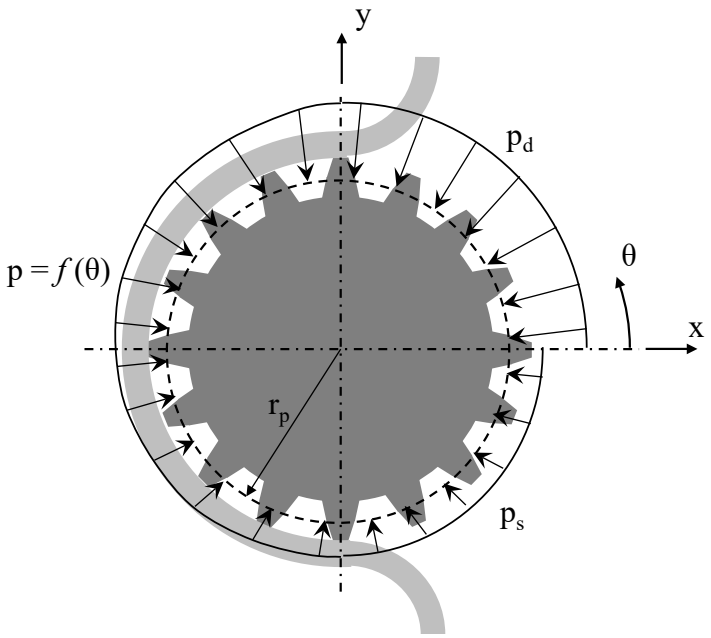

Fig. 68—Illustration of example gear pressure load profile

Next, we can separate the vector integral into components along the *x*- and *y*-axes.

$$F_x = -\int_0^{2\pi} p(\theta) r_p w \cos\theta \, d\theta \tag{5.8}$$

$$F_y = -\int_0^{2\pi} p(\theta) r_p w \sin\theta \, d\theta \tag{5.9}$$

The integral in Eq. (5.7) then can be broken up into three segments by angular position, one each for the discharge region, the suction region, and the transition between the two along the casing wall. Taking the *x*-component first,

$$F_x = -\int_0^{\pi/2} p_d r_p w \cos\theta \, d\theta - \int_{\pi/2}^{3\pi/2} f(\theta) r_p w \cos\theta \, d\theta - \int_{3\pi/2}^{2\pi} p_s r_p w \cos\theta \, d\theta \,, \tag{5.10}$$

where the pressure along the casing wall can be approximated as a linear function in θ:

$$f(\theta) = p_d - \frac{p_d - p_s}{\pi}\left(\theta - \frac{\pi}{2}\right). \tag{5.11}$$

Substituting Eq. (5.11) into Eq. (5.10) and integrating yields the following:

$$F_x = -\int_0^{\pi/2} p_d r_p w \cos\theta \, d\theta - \int_{\pi/2}^{3\pi/2} \left(p_d - \frac{p_d - p_s}{\pi}\left(\theta - \frac{\pi}{2}\right)\right) r_p w \cos\theta \, d\theta - \int_{3\pi/2}^{2\pi} p_s r_p w \cos\theta \, d\theta$$

$$= -\left[p_d r_p w \sin\theta\right]_0^{\pi/2} - \left[\left(p_d + \frac{\Delta p}{2}\right) r_p w \sin\theta - \frac{\Delta p}{\pi} r_p w (\cos\theta + \theta \sin\theta)\right]_{\pi/2}^{3\pi/2} - \left[p_s r_p w \sin\theta\right]_{3\pi/2}^{2\pi}$$

$$= -p_d r_p w + 2\left(p_d + \frac{\Delta p}{2}\right) r_p w - 2\Delta p r_p w - p_s r_p w$$
$$= [p_d - p_s - \Delta p] r_p w$$
$$F_x = 0\,. \tag{5.12}$$

The y-component of the load can be calculated in a similar fashion.

$$F_y = -\Delta p w r_p \left(1 - \frac{2}{\pi}\right) \tag{5.13}$$

As a reminder, in the above equations, Δp is the pressure differential of the pump, defined as

$$\Delta p = p_d - p_s\,. \tag{5.14}$$

The magnitude and direction of the radial load then can be calculated as follows:

$$\left|\vec{F}_r\right| = \sqrt{F_x^2 + F_y^2} \tag{5.15}$$

$$\angle \vec{F}_r = \tan^{-1}\left(\frac{F_y}{F_x}\right). \tag{5.16}$$

The key takeaway from this example is that while the direction of the radial load is dependent on the casing design, the magnitude is proportional to the pressure differential across the pump and the dimensions of the gear. This load is transferred to the shaft, which, in turn, must be supported by the shaft bearings. In most pump designs, there usually are two bearings per shaft, which means the rated load capacity of the bearing must equal half of the radial load. Thus, we can formulate a conservative estimate for the required load capacity for the bearing as follows:

$$F_R = \frac{1}{2}\Delta p w r_p\,. \tag{5.17}$$

5.2.2 *Bearing Features*

Another factor for consideration when selecting bearings is what design features are desired and/or available, such as the number of races and cages for the rolling elements. While of less immediate functional importance compared to the load capacity or the rated speed, these sorts of details are still relevant to the design process. As an example, as bearing size decreases, it becomes increasingly difficult to produce models with cages, and for miniature pumps, such bearings may not be an option. While cageless bearings still offer robust support of the shaft, the unconstrained rolling elements can make disassembly much more tedious. Usually, such bearings are sold with packing grease to secure the bearings during assembly, but the grease will be washed out quickly during operation. Thus, cageless bearings are more suitable for pump designs that require less maintenance. When selecting bearings, these qualitative factors should be considered in any trade study alongside the quantitative specifications, such as load capacity and form factor.

5.2.3 *Lubrication*

While rolling-element bearings are not as reliant on lubrication as journal bearings, some form of lubrication is required to ensure optimal bearing life. The two primary forms of bearing lubrication are either a liquid, such as oil or hydraulic fluid, or a grease. Most commercial bearings will specify any lubrication requirements in the case of liquid lubrication, or will come prepackaged with internal lubrication

in the case of grease. A designer, therefore, should ensure that the pump design provides bearing lubrication to match the bearings selected. Both forms of lubrication have their respective advantages and disadvantages, but there are some considerations for each when applied specifically to gear pumps.

Liquid lubrication for the bearings is often the first option considered for hydraulic pumps, as most hydraulic fluids also serve as excellent lubricants for bearings. This cross-compatibility greatly simplifies pump design, as contamination control is not required. For pumps designed for fluids that either are poor lubricants, such as fluids with high viscosity or reactivity concerns, or that must meet food safety standards, separating and sealing the bearings from the main pumping chamber may be required to allow for a suitable reservoir of lubrication fluid. A designer also must consider thermal factors, either heat generation within the bearings due to high loads or heat transfer from the pumping chamber. Due to the narrow clearances in bearing assemblies and between the bearings and the shaft, excessive heating of the bearings can lead to thermal expansion of components beyond their tolerances, leading to seizure of the bearings. Therefore, it is recommended to design a lubrication pathway that provides some small circulation of fluid through the bearings for cooling. Low-power pumps or pumps that will operate at low duty cycles may be able to dispense with lubrication circuits, but it is not ideal.

Another option for lubricating the bearings that can be more robust against wear is to use grease. Greases usually consist of two components: an oil and a thickening agent. Various oils or hydraulic fluids can be used for this purpose, while the thickening agents can vary from soap to clay. The various combinations of these different oils and thickening agents are what distinguish different greases. The purposes of the thickening agent are to bind the oil and to increase viscosity, as the higher viscosity aids in confinement of the grease and limits infiltration of particulates. As with liquid lubrication, there is a design choice that must be made as to whether the bearings will be isolated from the pumping chamber. For designs that allow mixing of the pumped fluid with the bearings, the mutual compatibility of the grease and the working fluid should be considered. A greased bearing assembly can leech the base oil from the grease into the working fluid if they are dissimilar. In general practice, this should not be of grave concern, as the relative amount of grease in most bearings is small compared to the volume of fluid in the system and thus should not greatly affect the fluid properties. However, care should be taken to ensure that there are no compatibility issues elsewhere in the system, or if the pump is used for food-rated or medical applications. Likewise, the solubility of the grease in the working fluid should be considered. One of the advantages of the use of grease is particulate control, but if the grease is broken down over the span of several days and no longer serves as a barrier, the purpose of the grease is undermined. A comparison of two pump shafts coated in two different greases and left in the same hydraulic fluid over time is shown in Fig. 69. Note the difference in the appearance of the grease film on the shafts within the region applied. Due to the number of permutations of grease and working fluid, it may not be possible to know which combinations are optimal without testing.

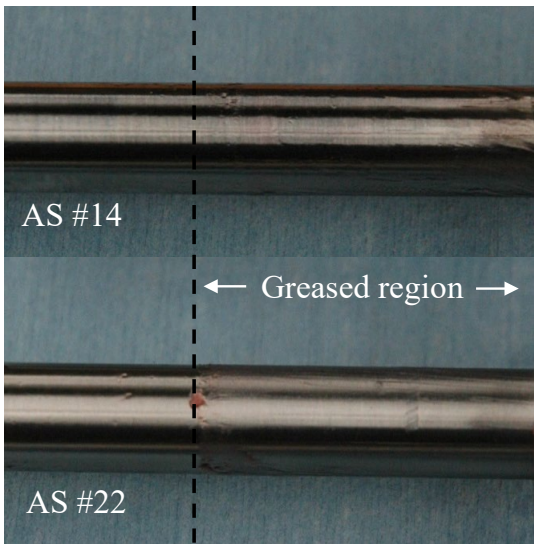

Fig. 69—Comparison of greased shaft exposed to Aeroshell fluid #31 for six days. The top shaft is coated with Aeroshell grease #14, while the bottom shaft is coated with Aeroshell grease #22. The grease layer on the bottom shaft is more prevalent and demonstrates increased stability in the fluid.

5.2.4 *Case Study of Bearing Tolerances*

As has been mentioned already, careful tolerancing is crucial and can mean the difference between a smoothly operating pump and one with excessive wear that underperforms design expectations. One of the foundational elements of a gear pump design is the spatial relationship between the gears and the casing, which determines internal leakage and wear. The placement of the gears is set by the shaft, which, in turn, is set by the placement of the bearings and the wear plates (if applicable). In each instance, if each component is not sufficiently constrained, the narrow clearances between the gears and the casing will be compromised. Recall that the pressure differential within the pump imparts a radial load on the gears, and should underconstrained gears contact the casing while under this load, excessive wear of the casing can result.

Examining a case study of this interaction in practice is a useful way to illustrate the processes. A gear pump that utilized rolling-element bearings housed within the wear plates was observed to require greater input torque to operate at a given pressure than expected. In addition, the torque profile was highly irregular with large-amplitude, single-sided variation compared to operation at a lower pressure. A plot of the two torque profiles is shown in Fig. 70.

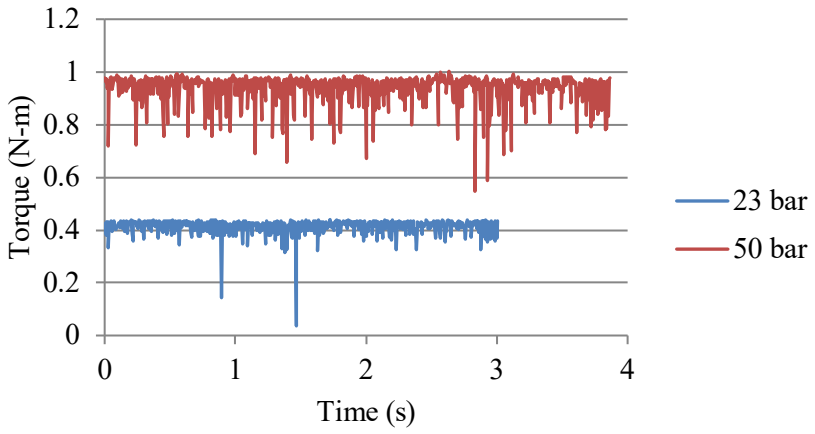

Fig. 70—Torque profiles of steady-state pump operation at 2,600 rpm at two discharge pressures. Note the higher pressure profile with a larger amount of single-sided oscillations. Expected torque load at 50 bar is 0.8 N-m.

Optical inspection of the pump casing revealed that erosion of the casing wall had occurred during operation. As Fig. 71 shows, there is a distinct region where the wall profile shifts, as well as a clear change from the vertical striations caused by EDM fabrication to lateral striations. This is indicative of the gears' pressing into the casing wall as pressure increased. This wearing-in also served to create a raised, ridgelike structure that created a slight contraction within the pumping chamber. One consequence of this was that after a sufficient time span of pump operation, the gears became fixed axially while still maintaining the ability to rotate, and could be removed from the casing only with considerable force.

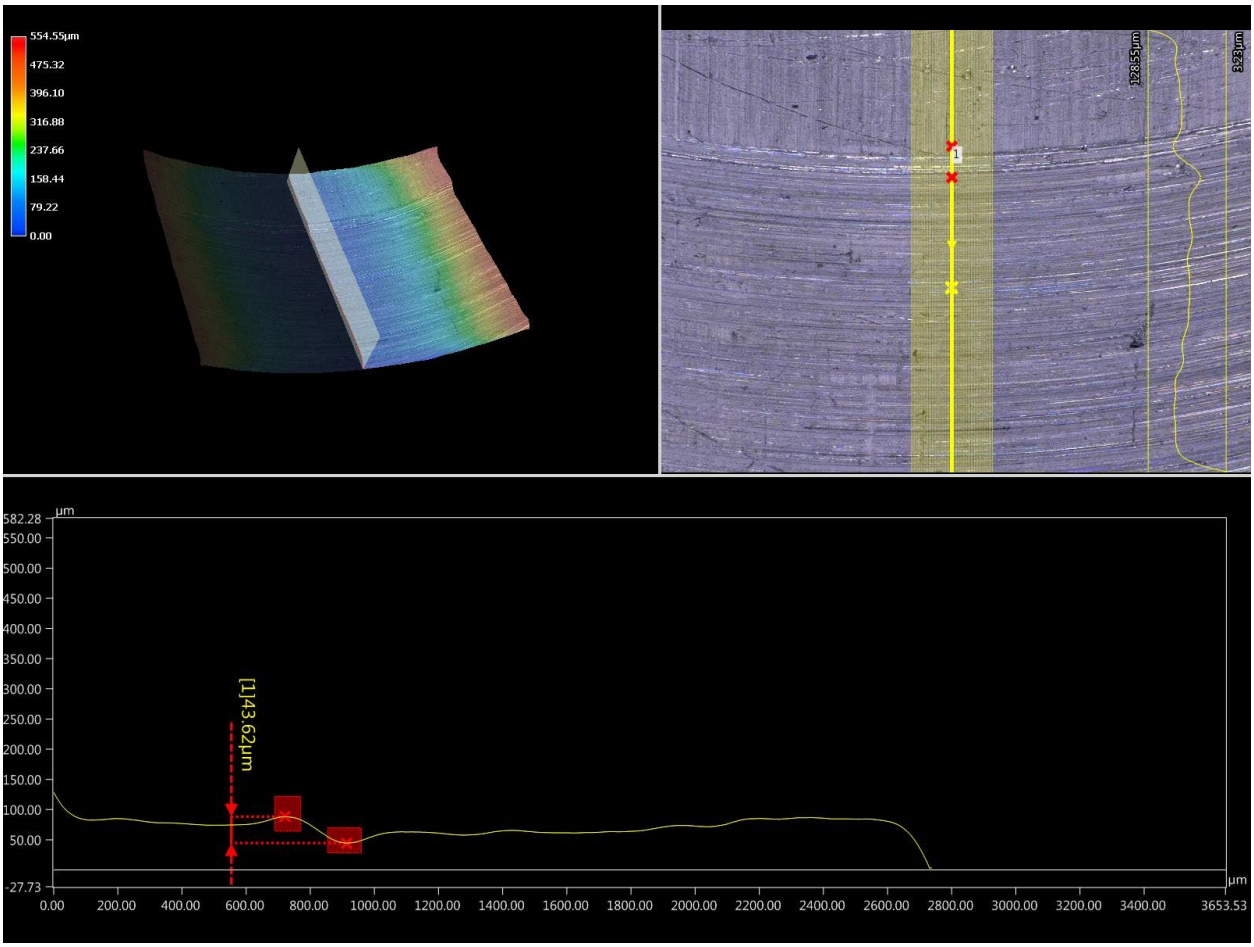

Fig. 71—Optical inspection of casing wall wear and wall profile. Resulting images are a topographic contour of a section of a curved section of the casing wall (*top left*), a line profile of the relative surface height along the same section (*top right*), and an expanded plot of the relative profile height (*bottom*).

From the combination of optical measurements and pump diagnostic data, we can conclude that the cause is lateral movement of the shafts and gears into the pump casing wall. Inspection of the needle bearings used in the pump revealed that they were out of specification, with an average bore diameter of 0.379 inches rather than the designated 0.375 inches. Thus, for a shaft with an actual diameter of 0.374 inches, the diametral clearance was 0.005 inches, much larger than the designed diametral clearance of 0.001 inches between the gear and the casing wall. As the pressure in the discharge section of the pumping chamber increases, the radial load on the gears increases and pushes the gears into casing wall and erodes the wall. The 43-micron (0.0017-inch) change in wall depth measured in Fig. 71 fits within the diametral clearance between the needle bearings and the shaft. In essence, this oversized clearance meant that for sizeable radial loads, the gear-wall interface becomes the effective bearing support, rather than the actual bearings. Therefore, when using commercial bearings, a designer should verify bore sizing when designing the mating shaft. This is one of the many examples in which pump component development is best conducted concurrently, rather than in series.

5.3 **Journal Bearings**

Unlike rolling-element bearings, which support the load through the conjugate action of individual rollers, journal bearings rely on a hydrodynamic fluid film between the bearing surface and the shaft. Journal bearings take of the form of a polished bore in a supporting material into which the shaft is inserted. During operation, the rotation of the shaft induces Couette flow within the film. As a radial load is applied, displacement of the shaft from concentricity with the bore creates an asymmetric flow structure that increases fluid pressure. At a sufficient shaft displacement, this increased fluid pressure provides sufficient force to counter the applied radial load. Due to the marginal thickness of the fluid film, journal bearings must be machined with very strict tolerancing based on the dimensions of the shaft to be supported. The use of journal bearings can simplify pump designs by reducing the number of components at the cost of imposing additional machining requirements. Additionally, journal bearings require a constant supply of lubricant to ensure stability of the hydrodynamic film. The immediate distinction for a pump designer is that while most rolling-element bearings generally are selected from commercial stock, journal bearings usually are machined into existing components, such as the wear plate. Thus, the designer must be familiar with the concept fundamentals of the journal bearing.

5.3.1 *Journal Bearing Design*

The hydrodynamic journal bearing (HJB) is essentially a rotating shaft placed inside a bearing bore, shown in Fig. 72. The difference in shaft and bore radii defines the nominal radial clearance of the bearing, r_c, and the length of the bearing, l, is set by design. The shaft rotates at a defined speed N with a radial load F_r. In most bearing applications, the ends of the bearing are exposed to atmosphere, but in this case, the ends of the bearing are at either the pump chamber pressure, p_c, or a bearing sink pressure, p_s. The chamber pressure corresponds to the back region behind the wear plate, which could be at discharge pressure or suction pressure, depending on the pump design. The bearing sink is a return line to the hydraulic system tank that allows a small circulation of fluid through the bearing to ensure sufficient lubrication. Alternatively, the sink region can be pressurized and can act as a fluid source that pumps fluid into the pump chamber through the bearing, which has the added benefit of applying a compressive force on the wear plates and increases contact between the wear plates and the gears.

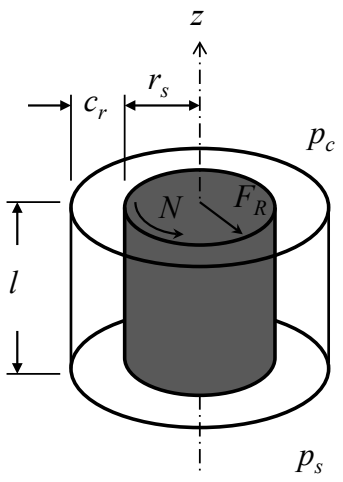

Fig. 72—Hydrodynamic journal bearing geometry

The HJB functions by creating an azimuthal flow of fluid around the annulus between the shaft and the bore. As the shaft rotates, viscous shear imparts a tangential velocity on the fluid, creating a velocity gradient across the clearance. If the shaft is perfectly concentric with the bore, the fluid flow created is a standard Couette flow in a cylindrical configuration, which is readily understood; however, such a flow is azimuthally symmetric in terms of velocity and pressure. As such, there is no net radial force on the shaft by the fluid, and there is no bearing support. When a radial force is applied to the shaft, the shaft experiences a displacement from centerline, which creates an eccentricity, e, between the shaft and the bore. The critical result of this eccentricity is that the fluid film thickness, h, now varies with the angular position, θ, changing from a maximum at one end to a minimum at the other. The rotation of the shaft now serves to force fluid into this reduced flow area, which increases the fluid pressure. Due to the asymmetry of the film pressure, the fluid film now imparts a net radial force on the shaft to balance the original load that created the eccentricity. The pressure distribution within the film is nonlinear and peaks at an angle to the applied load, α. In order for the net radial load on the shaft from the film balance the external load, the shaft leads the bearing, resulting in the line of eccentricity between the centers of the shaft and bore existing at an angle to the external load, φ. A diagram of the film pressure distribution and resulting angles is shown in Fig. 73.

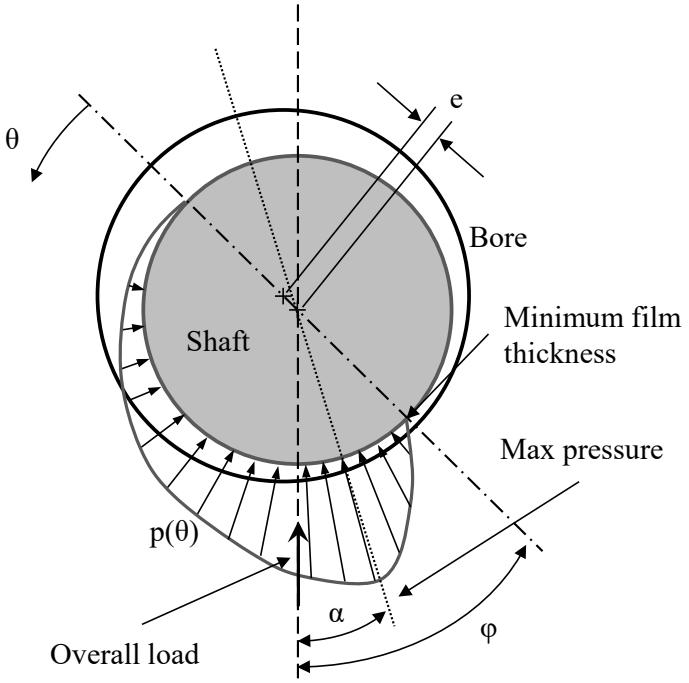

Fig. 73-Diagram of bearing fluid film pressure distribution

Bearings can be analyzed either as long bearings, where l/d_s is very large, or as short bearings. For this analysis, short bearings are assumed, where the pressure as a function of the azimuthal and axial coordinates is based on the theoretical equation for short bearings [14],

$$p(\theta, z) = \frac{\mu U}{r_b c_r^2}\left(\frac{l^2}{4} - z^2\right)\frac{3\varepsilon \sin\theta}{(1+\varepsilon\cos\theta)^3} + \frac{p_c - p_s}{l}z + \frac{p_c - p_s}{2}, \tag{5.18}$$

where U is the surface speed of the shaft, μ is the viscosity of the fluid, and ε is the eccentricity ratio, defined as e/c_r. The radial load supported by the bearing, which must be equal to the applied load, F_r, is

$$F_r = K_\varepsilon \frac{\mu U l^3}{c_r^2}, \tag{5.19}$$

where K_ε is the eccentricity factor, defined in terms of the eccentricity ratio:

$$K_\varepsilon = \frac{\varepsilon\sqrt{\pi^2(1-\varepsilon^2)+16\varepsilon^2}}{4(1-\varepsilon^2)^2}. \tag{5.20}$$

With the radial load, F_r, defined by the pump operating conditions and the other terms set by geometry or operating conditions, Eq. (5.19) uniquely defines the eccentricity factor required to support the load, which, in turn, uniquely defines the eccentricity ratio. Due to the form of Eq. (5.20), the resulting eccentricity ratio cannot be solved analytically, but must be solved numerically. Determining the eccentricity ratio then allows for the calculation of the pressure distribution within the bearing using Eq. (5.18). It should be noted that Eq. (5.18) predicts negative film pressures for θ in the range of 180 to 360 degrees, with a minimum pressure equal in magnitude to the maximum film pressure found in 0 to 180 degrees. In practice, this does not occur except for with very low bearing loads, thus the pressure distribution for bearings is set to zero after the point of minimum film thickness [22]. This region of low pressure allows for an axial flow of fluid to balance the flow of fluid induced by the pressurization of the film.

The film thickness, assuming no voids in the flow, is set by the gap between the shaft and the bearing, which, in turn, is set by the eccentricity.

$$h(\theta) = c_r(1 + \varepsilon\cos\theta) \tag{5.21}$$

From this, the maximum and minimum film thicknesses can be calculated easily.

$$h_{min} = c_r(1-\varepsilon) \tag{5.22}$$
$$h_{max} = c_r(1+\varepsilon) \tag{5.23}$$

The torque on the rotating member (the shaft), T_s, is a function of the torque on the stationary member (the bearing) T_b and the radial load.

$$T_s = T_b + F_r e \sin\varphi \tag{5.24}$$

$$T_b = \mu \frac{d_b^2 l U}{2c_r} \frac{\pi}{\sqrt{1-\varepsilon^2}} \tag{5.25}$$

$$\varphi = \tan^{-1}\left(\frac{\pi\sqrt{1-\varepsilon^2}}{4\varepsilon}\right) \tag{5.26}$$

The effective coefficient of friction, μ_{eff}, for the bearing then can be estimated from the shaft torque and the radial load.

$$\mu_{eff} = \frac{T_s}{r_s F_r} \tag{5.27}$$

A key design consideration is to maintain a low coefficient of friction to ensure that power dissipation in the bearing is kept low, both to limit heat generation and to maximize the mechanical efficiency of the pump.

5.3.2 *Journal Bearing Case Study*

Journal bearings share with rolling-element bearings a high priority on precision, specifically precise design and machining. It is vital that the surface roughness of both the shaft and the bearing should not meet or exceed the minimum film thickness, otherwise full hydrodynamic flow will not be achieved, but instead, the bearing will be in elastohydrodynamic flow, where surface asperities of the two components will be in contact, increasing wear and friction. As a case study, example journal bearing design parameters are given in Table 5. This design results in a minimum film thickness of approximately 27 µin; thus, an average surface roughness of 10 µin on both the shaft and the bearing should be sufficient to ensure smooth operation.

Table 5—Example Journal Bearing Design Parameters

Parameter	Value
r_s	4.8 mm
c_r	6.6 µm
l	9.5 mm
p_s	1 bar
p_c	100 bar
N	3,000 rpm
F_r	965 N
μ	0.0062 Pa-s

Next, the pressure in the bearing is calculated using Eq. (5.18) and is plotted in Fig. 74 as a function of azimuthal position for multiple axial positions in mm, where $z = 0$ corresponds to the center of the bearing. It should be noted that z = 4.76 mm corresponds to the end of the bearing at the average pump chamber pressure, which, for the given operating conditions, is 100 bar. The peak pressure within the film is much higher than the pump chamber pressure; however, this pressure is still much lower than the material strength of most materials that would be used for the wear plate.

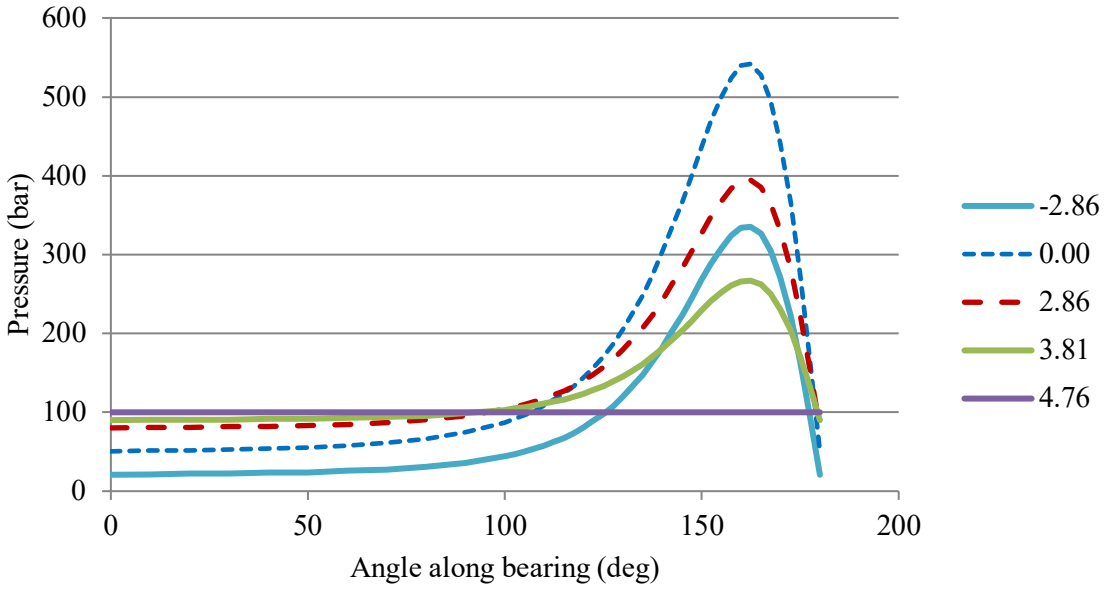

Fig. 74—Bearing film pressure as a function of angular and axial position

In order to estimate the hoop stress on the bearing, we need to calculate the average fluid pressure.

$$p_{avg} = \frac{F_r}{ld_b} + \frac{p_c + p_s}{2} \tag{5.28}$$

The hoop stress is calculated as

$$\sigma_h = \frac{r_b p_{avg}}{t_b}, \tag{5.29}$$

where t_b is the radial thickness of the bearing. The minimum thickness of the bearing then can be defined in terms of a nominal factor of safety for the bearing, X_s, and the yield strength of the bearing material, σ_Y.

$$t_b = X_s \frac{r_b p_{avg}}{\sigma_Y} \tag{5.30}$$

Continuing the example, with a factor of safety of four, an average bearing pressure of 47.4 bar, and a yield strength for aluminum bronze of 221 MPa, the thickness of the bearing needs to be at least 1.3 mm. This is readily achievable with most wear plate designs, which is one of the advantages of integrating the journal bearing into the wear plate.

The next step is to evaluate the tolerancing of the shaft and the bore. A defined dimension can be written as the sum of a nominal value and an allowable variance, δ.

$$x = x^0 + \begin{cases} -\delta^- \\ \delta^+ \end{cases} \tag{5.31}$$

The allowable variance can be either symmetric, in which $\delta^- = \delta^+$, or asymmetric. Tolerancing for journal bearings is often asymmetric to provide a hard upper limit to the shaft diameter and a lower limit to the bore diameter, as this prevents part interference. The critical quantity of interest in this case is the radial clearance, c_r, as this determines the allowable range of motion of the shaft. An overly large radial clearance can result in larger-than-expected displacements of the gears, similar to what is discussed in Section 5.2.4. Alternatively, a radial clearance that is too small can result in contact between the two surfaces and increased wear. Thus, the combined variances of all tolerances in the bearing assembly should be considered.

Continuing our case study example, suppose a surface coating is applied to the shaft with a thickness of t_c = 3.5 µm ± 1.5 µm. If we set the uncoated shaft and bore dimensions as

$$d_s = 9.5123 + \begin{cases} -0.0051 \\ 0 \end{cases} [mm] \tag{5.32}$$

$$d_b = 9.5363 + \begin{cases} 0 \\ 0.0051 \end{cases} [mm], \tag{5.33}$$

then we can calculate the bounds of the radial clearance.

$$c_r = \frac{1}{2}\left[\left[d_b^0 - (d_s^0 + 2t_c^0)\right] + \begin{cases} \delta_b^- - \delta_s^+ - 2\delta_t^+ \\ \delta_b^+ - \delta_s^- - 2\delta_t^- \end{cases}\right] \tag{5.34}$$

For our example here, this results in the radial clearance range below:

$$c_r = 6.67 + \begin{cases} -1.50 \\ 6.58 \end{cases} \mu m \,. \tag{5.35}$$

The bearing calculations then can be repeated with the nominal, minimum, and maximum radial clearances substituted for the radial clearance in Table 5. The goal is to ensure that in all three cases, the resulting shaft eccentricity is less than the sum of the radial clearance and the surface roughness of the shaft and the bearing.

5.3.3 *Experimental Characterization of Journal Bearings*

It can be useful to experimentally characterize a journal bearing design to verify modeled expectations ensure sufficient load capacity for the associated pump. A test rig can be used to measure the associated performance conditions, such as shaft speed and torque, while applying a known radial load on a bearing element. Much of the hardware required is the same used in characterizing the gear pump: a shaft encoder, a torque transducer, and a motor. CAD models are shown in Fig. 75 and Fig. 76 of an example bearing test rig installed in a pump test bed, and a cross section of the bearing test rig itself, respectively. The bearing test rig consists of a 0.3741-in-diameter shaft suspended between two fixed ball bearings. A test journal bearing with a bore diameter of 0.3752 in and 0.975 in long was suspended along the shaft to measure the applied radial load on the bearing. A turnbuckle placed along the connection between the bearing and the load cell in combination with multiple Bellville washers in series allowed for setting a variable load on the bearing. The bearing then was immersed in a bath of hydraulic fluid to provide sufficient lubrication.

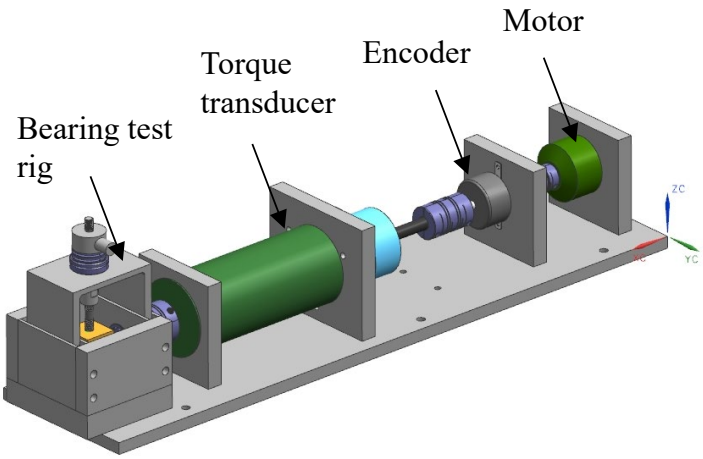

Fig. 75—Bearing test rig setup on the pump testbed

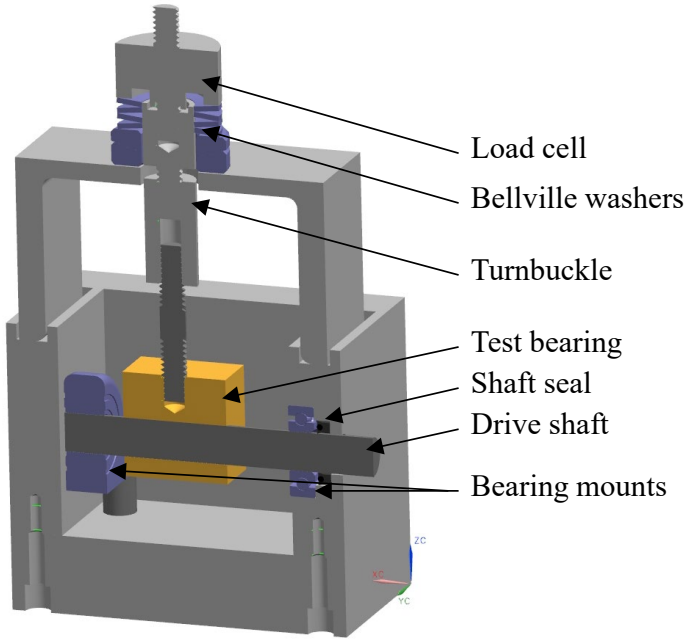

Fig. 76—Cross section of bearing test rig

When building such a test apparatus, it is important to remember that some elements may require wearing in, such as shaft seals. As Fig. 77 shows, when the rig was run for the first time without the test bearing at constant motor voltage, the measured shaft speed and torque drifted over time. Without an initial wearing-in process, these transients can introduce errors in the measured performance of the bearing. Additionally, a test rig should be run at various speeds to measure the required torque to turn the shaft at speed against the shaft seal-and-bearing mount friction. The resultant data can be used to create a linear approximation for the system frictional torque as a function of shaft speed, shown in Fig. 78.

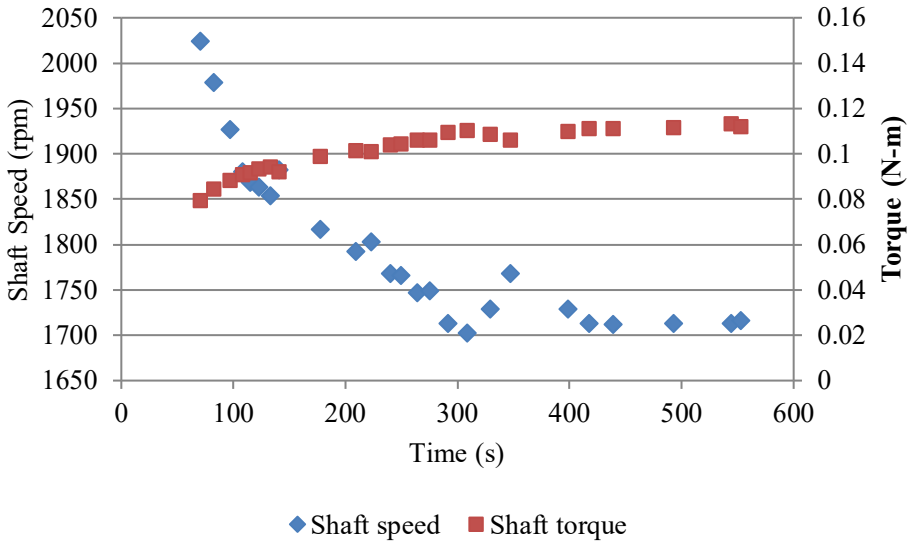

Fig. 77—Bearing test rig wear-in

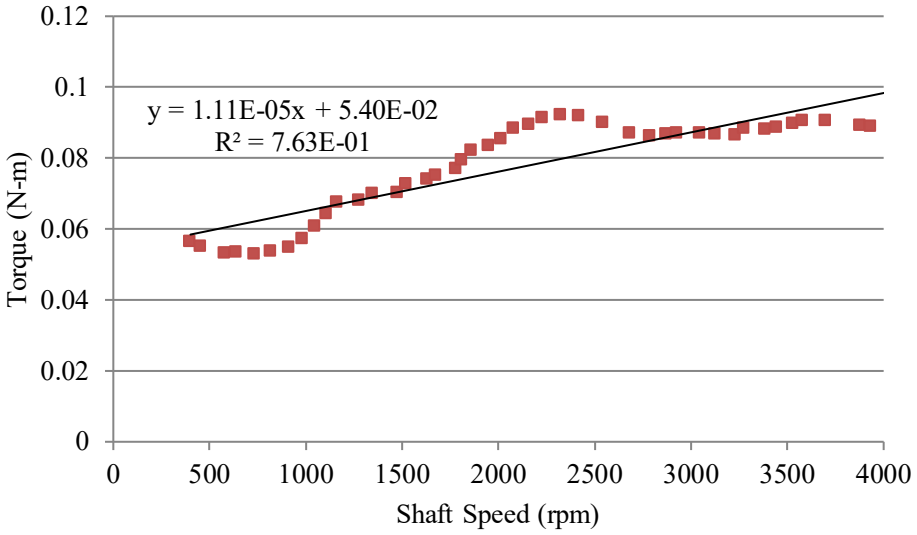

Fig. 78—Tare of bearing test rig seal-and-bearing mount frictional torque

To evaluate the bearing, the test rig must be run across a range of shaft speeds at a multiple radial loads while measuring the required shaft torque. The effective coefficient of friction for the bearing then is calculated from Eq. (5.27), an example of which is shown in Fig. 79 for five radial loads tested.

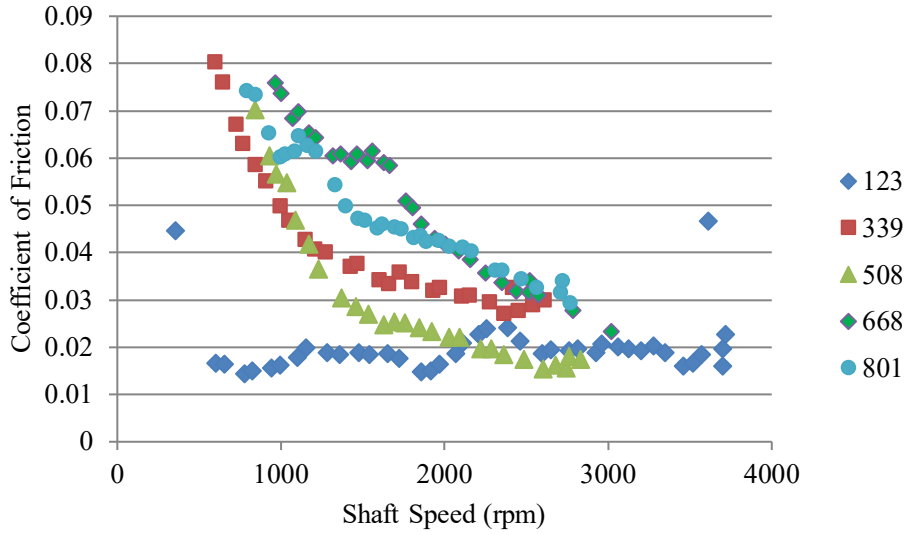

Fig. 79—Bearing friction as a function of shaft speed and radial load. Radial load given in N.

Due to the effective coefficient of friction's dependence on two variables, it is difficult to gain a comprehensive understanding of the bearing performance from Fig. 79. To assist in the analysis, the Sommerfeld number is used, which can be used to describe the bearing operating conditions within a single parameter. The Sommerfeld number, S, is defined as

$$S = \left(\frac{r_s}{c_r}\right)^2 \frac{\mu l r_s}{F_r} \omega \,. \tag{5.36}$$

The Sommerfeld number also can be calculated as a function of the eccentricity ratio [14],

$$S = \frac{(2+\varepsilon^2)\sqrt{1+\varepsilon^2}}{12\pi\varepsilon}, \tag{5.37}$$

which can be useful for connecting the physical operating parameters of Eq. (5.36) to the displacement of the shaft.

Plotting the effective coefficient of friction of the bearing against the Sommerfeld number results in a plot that closely follows expectations found in the literature, shown in Fig. 80. For low Sommerfeld numbers, the lubrication film can be characterized as "thin" and results in a relatively high coefficient of friction. As the Sommerfeld number increases, the film becomes "thick" and friction decreases. Due to the inverse relation of the Sommerfeld number to the radial load, the Sommerfeld number decreases overall for each successive data set where the radial load is increased, which thus increases the coefficient of friction. Similarly, as the shaft speed increases in each data set, the Sommerfeld number increases and the coefficient of friction decreases. After some minimum, the coefficient of friction increases with increasing Sommerfeld number, representing increased viscous shear losses within the fluid film.

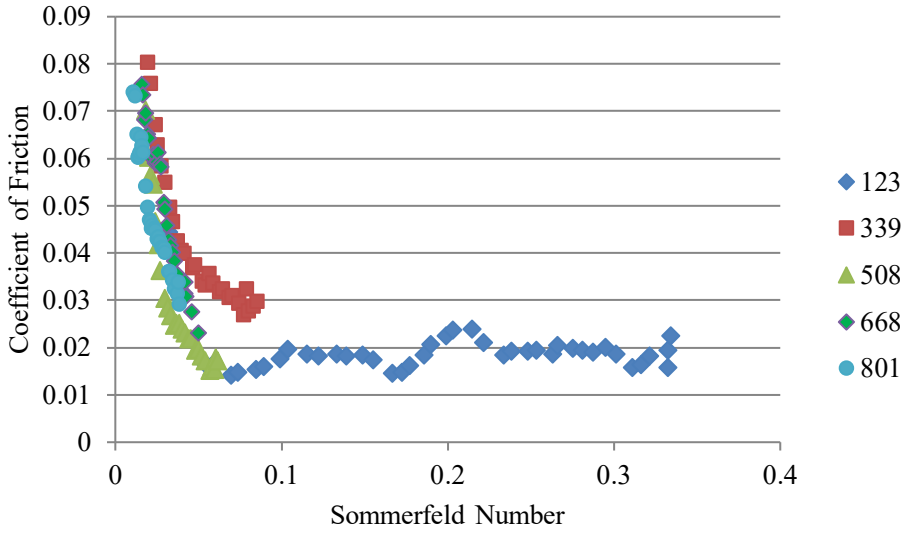

Fig. 80—Bearing friction as a function of Sommerfeld number for different radial loads. Radial load given in N.

An additional consideration when considering such data is the stability of the thick- and thin-film regimes with respect to temperature and viscosity. In the thick-film regime, as temperature increases (e.g., from frictional heating) the viscosity decreases, which decreases the Sommerfeld number, which, in turn, decreases the coefficient of friction. This decreases the heat generation in the bearing, creating a negative feedback to stabilize bearing performance. In contrast, for the thin-film regime, the coefficient of friction increases with decreasing Sommerfeld number, creating more frictional heat, which creates a positive feedback loop. Thus, operation within the thick-film regime is considered more stable.

From the results of this example bearing test case study, there are several design recommendations that can be made:

- The ideal operating regime for the bearing is to have a Sommerfeld number of at least 0.1, as this puts the bearing in the thick-film regime, which promotes stability against thermal feedback.
- The minimum Sommerfeld number acceptable is approximately 0.05, as while this is likely within the thin-film regime, it still offers a low coefficient of friction to ensure limited bearing friction, though it is susceptible to loss of performance as temperature increases.
- The goals of a stable bearing with a high factor of safety and a low size and mass pump are at odds, as a small bearing length, while beneficial to pump form factor, decreases the Sommerfeld number.
- A large shaft diameter is generally beneficial, as it serves to increase projected bearing area (and thus decreases fluid pressure required), increases the Sommerfeld number, and limits axial leakage in the pump.
- Small radial clearance in the bearing is necessary for at least three reasons:
    - It increases the Sommerfeld number, which reduces bearing friction.
    - It reduces peak film pressure.
    - It reduces the required eccentricity of the shaft to generate the film pressure needed, which relaxes the tolerancing required in the rest of the pump design related to drive shaft placement.

6. CASING, SEALS, AND ASSEMBLY

The combination of the gears, wear plates, shafts, and bearings creates the core pumping mechanism, the beating heart of the pump. The final element to a gear pump is to house this stack subassembly in a tightly sealed enclosure so that the meshing action of the gears is able to draw in and pressurize fluid. The primary design task for the casing is to ensure that this enclosure provides a closely fit boundary around the core stack to minimize internal leakage. Since the casing will consist of two or more parts, static seals, along with a dynamic shaft seal, are needed to prevent external leakage. This chapter also presents some guidance on seal design practices, pump assembly, and initial performance evaluation.

6.1 Pump Casing

The primary purpose of the casing is to provide a rigid confinement of the fluid and an envelopment of the core pump stack. A tight envelopment of the gears establishes the suction and discharge regions within the pump. The primary design goals of the casing are:

1. to confine the pressurized fluid,
2. to minimize radial leakage from discharge to suction, and
3. to provide a structural reference point for the assembly, particularly the drive shaft.

These goals are often pursued while optimizing for a specific parameter, such as cost or mass. This can result in the casing having an additional design goal, such as minimal wall thickness, or minimizing the number of precision-machined features. These are secondary design goals that represent the specific application of the pump or the priorities of the designer, while the above three primary goals are near universal across all pumps.

6.1.1 *Strength Requirements*

The first objective of the casing is to confine the pressurized fluid, and to do so, the casing must be capable of withstanding the interior pressure of the fluid on the casing. For multiple-part casings, as is almost universally the case, this gives two main strength criteria. The first is that in the main body of the casing, the component that constitutes the majority of the envelopment of the gears, the pressure-induced stress on the casing does not exceed the yield stress within some factor of safety. The second criterion is

that the fasteners of the additional casing components can maintain the assembly under pressure. For the former, we treat the casing as a pressure vessel and restrict the loads to that component. For a three-part casing design of the type shown in Fig. 5, where a central housing houses the gear stack with two end plates, we can treat the problem as a two-dimensional pressure vessel. The net force along each axis on the casing from the fluid pressure results in two critical wall-thickness dimensions: the circumferential thickness, t_c, and the port wall thickness, t_p. We can simplify the pressure distribution within the pumping chamber by approximating it as half at suction pressure, p_s, and half at discharge pressure, p_d. A representation of the assumed geometry and conditions for the casing is shown in Fig. 81.

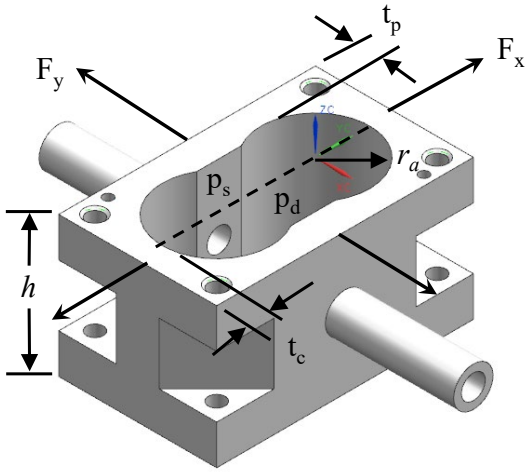

Fig. 81—Casing load geometry and assumed pressure distribution

The net force along each axis that imparts tension on the casing can be solved by integrating the pressure along the radial surface elements on the interior surface of the case. For casings that are symmetric along the *x*- and *y*-axes, the tensile force along the *x*-axis is

$$\begin{aligned} F_x &= 2\int_{-\pi/2}^{\pi/2} hr_a p\cos\theta\, d\theta \\ F_x &= 2\int_{-\pi/2}^{0} hr_a p_d\cos\theta\, d\theta + 2\int_{0}^{\pi/2} hr_a p_s\cos\theta\, d\theta \\ F_x &= 2hr_a p_d\sin\theta\big|_{-\pi/2}^{0} + 2hr_a p_s\sin\theta\big|_{0}^{\pi/2} \\ F_x &= 2hr_a(p_s+p_d)\,, \end{aligned} \tag{6.1}$$

where h is the height of the casing and r_a is the addendum radius of the gear to be housed. The force along the *y*-axis is similarly

$$\begin{aligned} F_y &= 2\int_{0}^{\pi/2} hr_a p_s\sin\theta\, d\theta + 2r_p hp_s + 2\int_{0}^{-\pi/2} hr_a p_d\sin\theta\, d\theta + 2r_p hp_d \\ F_y &= 2hr_a p_s\cos\theta\big|_{0}^{\pi/2} + 2r_p hp_s + 2hr_a p_d\cos\theta\big|_{0}^{-\pi/2} + 2r_p hp_d \\ F_y &= 2h\big(r_a+r_p\big)(p_s+p_d)\,, \end{aligned} \tag{6.2}$$

where r_p is the pitch radius of the gears used. Assuming this tensile load is shared equally between each of the two symmetric sides of the casing, the normal stress on the casing is

$$\sigma_i = \frac{F_i}{2th}, \tag{6.3}$$

where t is the minimum wall thickness orthogonal to the force vector. If we use a factor of safety, X_s, then the minimum wall thickness can be determined from the yield strength of the casing material, σ_y.

$$t_{min} = \frac{X_s F_i}{2h\sigma_y} \tag{6.4}$$

Substituting Eqs. (6.1) and (6.2) into Eq. (6.4) in turn yields the minimum circumferential and port wall thicknesses.

$$t_{c,min} = X_s \frac{(r_a + r_p)(p_s + p_d)}{\sigma_y} \tag{6.5}$$

$$t_{p,min} = X_s \frac{r_a(p_s + p_d)}{\sigma_y} \tag{6.6}$$

In general, the pitch radius is similar to the addendum radius, which means that the minimum circumferential wall thickness is close to twice the minimum port wall thickness. Another consideration is any mounting hole or cutaway for mass reduction, as this can cause stress concentration and can be another point of potential failure.

The second strength criterion for the casing is that the attachment of the secondary casing plates can withstand the pressure of the interior fluid. In this case, the simplified pressure distribution assumed in Fig. 81 does not apply, as that corresponds to the pressure along the radial boundary of the gears and the wear plates. Instead, the pressure distribution against the casing end plates follows the shape of the back region that was discussed in Section 4.4.3. Likewise, the area over which the fluid pressure acts is dependent on wear plate shape and the size of the static seal gland in the casing end plate. As a conservative estimate, we can treat the entire fluid pressure at discharge pressure over a simplified area.

$$F_z = p_d(\pi r_a^2 + 4 r_p r_a) \tag{6.7}$$

This load is distributed between n_b number of bolts used to secure the casing end plate. The axial stress on each bolt from this load is

$$\sigma = \frac{F_z}{n_b A_t}, \tag{6.8}$$

where the effective area of a threaded bolt is given in terms of the bolt pitch diameter, d_p, and the minor diameter, d_r [14].

$$A_t = \frac{\pi}{4}\left(\frac{d_p + d_r}{2}\right)^2 \tag{6.9}$$

The positioning of these bolts should be distributed as evenly as possible to ensure that local deflection of the casing end plates is minimized.

Preloading the bolts can be done to ensure that the bolts are capable of supporting the load prior to operation. For relatively static operation, the preload can be up to 90% of the proof load, while for more

cyclically used pumps, a preload up to 75% of the proof load can be used. The proof load is the maximum load the bolt can take before the onset of plastic deformation. Thus, for a given bolt size, the proof load is

$$F_{proof} = A_t \sigma_y \tag{6.10}$$

The torque, T, required to apply the desired preload, F_p, for lubricated threads can be approximated in terms of the bolt diameter, d.

$$T \cong 0.21 F_p d \tag{6.11}$$

6.1.2 *Design Guidelines*

When designing the casing, one of the first decisions to make is the number of casing elements, which has a strong impact on the fabrication process to be used. The two main approaches are casings that are separated into two or three parts. In two-part casings, the core pump stack is housed in a blind hole and is sealed with a single end plate. Three-part casings utilize two end plates to seal the central casing element, which allows machining of a through hole for the core pump stack. While two-part casings have fewer components to machine and one less static seal to design, there are some disadvantages to consider. The first is that it can be difficult to extract components from a blind hole with tight clearances, particularly the bottom wear plate. This makes maintenance of the pump more labor-intensive or requires additional design elements to facilitate component removal.

The second disadvantage to two-part casings is that it can be difficult to machine a blind hole into a complex shape with the tight tolerances needed for high volumetric efficiency. Techniques like wire EDM are incompatible with blind holes, and milling can be time-consuming and has difficulty ensuring a sharp corner at the edge of the hole. In the case where the pump design utilizes wear plates, a slight rounding of the corner at the end of the blind hole can prevent the stack from being fully inserted, effectively decreasing the overall depth of the hole. For pumps without wear plates, this creates a region of interference between the casing and the gears. If the casing is made of a sufficiently hard material compared to the gears, this can lead to erosion of the corners of the gear teeth. Microscope imaging and a height profile of a gear tooth that suffered wear in a blind bore are shown in Fig. 82.

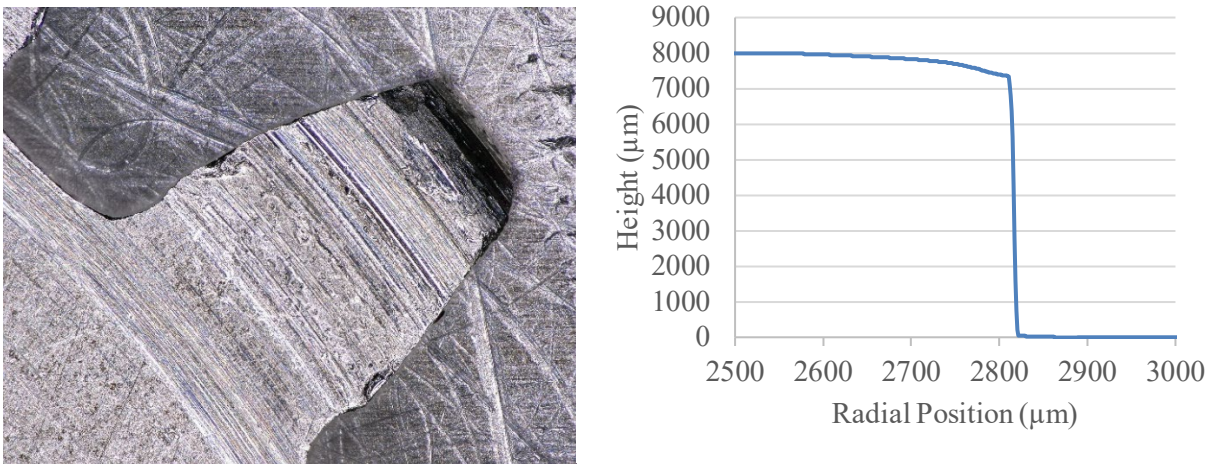

Fig. 82—Microscope image of gear tooth wear (*left*) and radial profile of tooth height at the tip (*right*)

An additional problem with this kind of wear is that if the material hardness is similar between the casing and the gear, the wear may not be fully one-sided, where one material wears away the other to create complementary surfaces. Instead, this can result in both materials' wearing such that a gap is created, greater internal leakage after wear-in and decreased volumetric efficiency. A comparison of pump volumetric efficiency is shown in Fig. 83 of a basic pump without wear plates that utilized the gear shown in Fig. 82 before and after the wear occurred. Therefore, should a blind hole be used for fabricating the pump casing, the gear should be made of a considerably harder material to prevent gear degradation during the initial wear-in process. This is yet another example of the general rule that the gears should be of a harder material than the casing. For these reasons, two-part casings are generally feasible for low-pressure pumps, for which tight clearances are not as high a priority and the reduction in part count can provide cost savings. It is recommended that high-pressure pumps utilize a three-part casing design, as it is easier to ensure precise pumping volume dimensions.

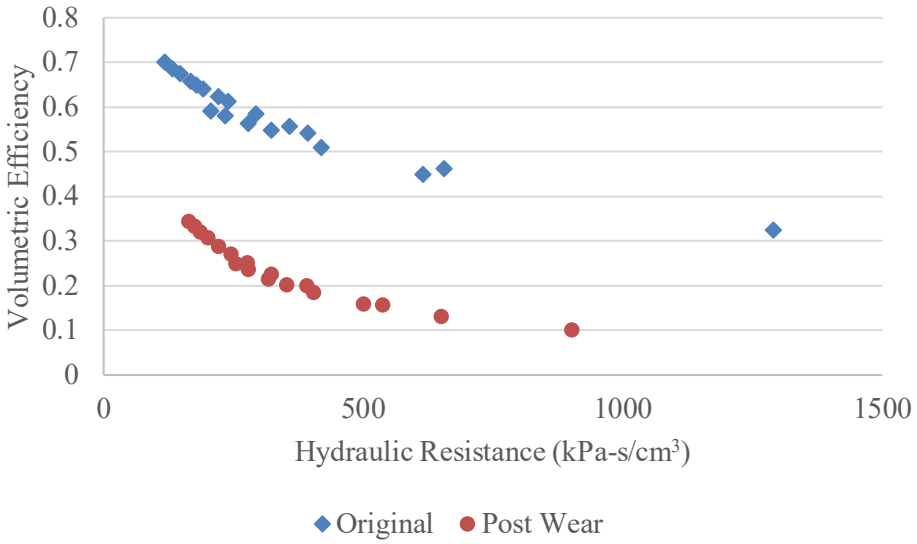

Fig. 83—Comparison of pump performance before and after gear wear

Another design element of the casing is one shared with the wear plates, but has not been discussed to this point, and that is the shape of the pumping chamber. Due to the nature of the central casing element, the pumping chamber is essentially the vertical extrusion of the cross section of the wear plate. Thus, the shape of the interior of the casing must match the planar profile of the wear plate. The simplest form of this profile is the "pill" shape of two connected semicircles, shown in Fig. 53 and Fig. 54. However, as discussed in Section 2.1.1, the radial leakage is a function of the number of gear teeth enveloped by the casing at any given point in time. Extending the arc of the circular walls that surround the gears potentially can increase the number of teeth enveloped, depending on the total number of gear teeth and the degree of envelopment of the gears. This can improve the volumetric efficiency of the pump for gears for the cost of a more complex casing and wear plate geometry. Excessive envelopment of the gears can be detrimental, as it could impede fluid flow into or out of the meshing region. This is particularly important in the suction region, as it can contribute to the onset of cavitation.

The final consideration for the casing design is the tolerancing of the pumping chamber in relation to the gear addendum diameter. As has already been mentioned, standard practice is to choose a gear material that has a greater hardness than the casing material, as this facilitates initial run-in and can provide a better seal against radial leakage. For this reason, the pumping chamber can have the same nominal diameter as the addendum diameter. However, for ease of initial assembly and to reduce particulate formation, it is

recommended that the casing bore be toleranced slightly larger, and the gear slightly smaller. The magnitude of this tolerance depends on the diameter of the gear used; small-scale pumps may call for tolerances less than 0.001 inches.

6.1.3 *Component Alignment*

One of the critical functions of the casing that can be missed easily is that the central casing piece is the point of reference for the entire pump assembly. The primary design goal is to ensure that the drive shaft is properly aligned with the bore in the casing end plate and the associated shaft seal. As has been discussed in Section 4.4.2, shaft misalignment can contribute to reduced volumetric efficiency. Some additional factors to consider with drive shaft misalignment are increased stress on the shaft seal or impingement of the drive shaft on the casing bore. These can lead to premature shaft seal failure, increased frictional losses and wear on the shaft, or external leakage of the working fluid. The need to maintain shaft alignment extends to the bearings supporting the shafts and, by extension, the wear plates in the case of embedded bearings. This is of particular importance for journal bearings, as the pressure distribution within the fluid film that supports the shaft is dependent on the position of the shaft relative to the central axis of the bearing. Thus, the ability to accurately fix these elements within a tolerable range is critical to efficient operation.

A common method to ensure alignment between the various components is to use precision-ground alignment pins. These pins are then press-fit into blind holes in the central casing piece and mate to corresponding holes in the casing end plate. An example of this method of dimensioning is given in Fig. 84. By dimensioning the key features in relation to these alignment holes, we can enforce a spatial alignment tolerance between features in mating parts, such as the drive shaft bore in the casing end plate to the wear plate bore. This design practice reinforces the need for minimal clearance between the wear plate and the casing, as in floating wear plate designs; this is usually the only method to constrain the wear plate in the relevant dimensions. It should be mentioned that the use of press-fit alignment pins could increase the difficulty in separating the casing elements, so a designer should consider including features to facilitate disassembly. One example would be to add a threaded hole through the end plate that meets the blank face of the central casing element outside of the static seal. This allows a user to use a bolt threaded into the feature to apply a force between the two casing elements.

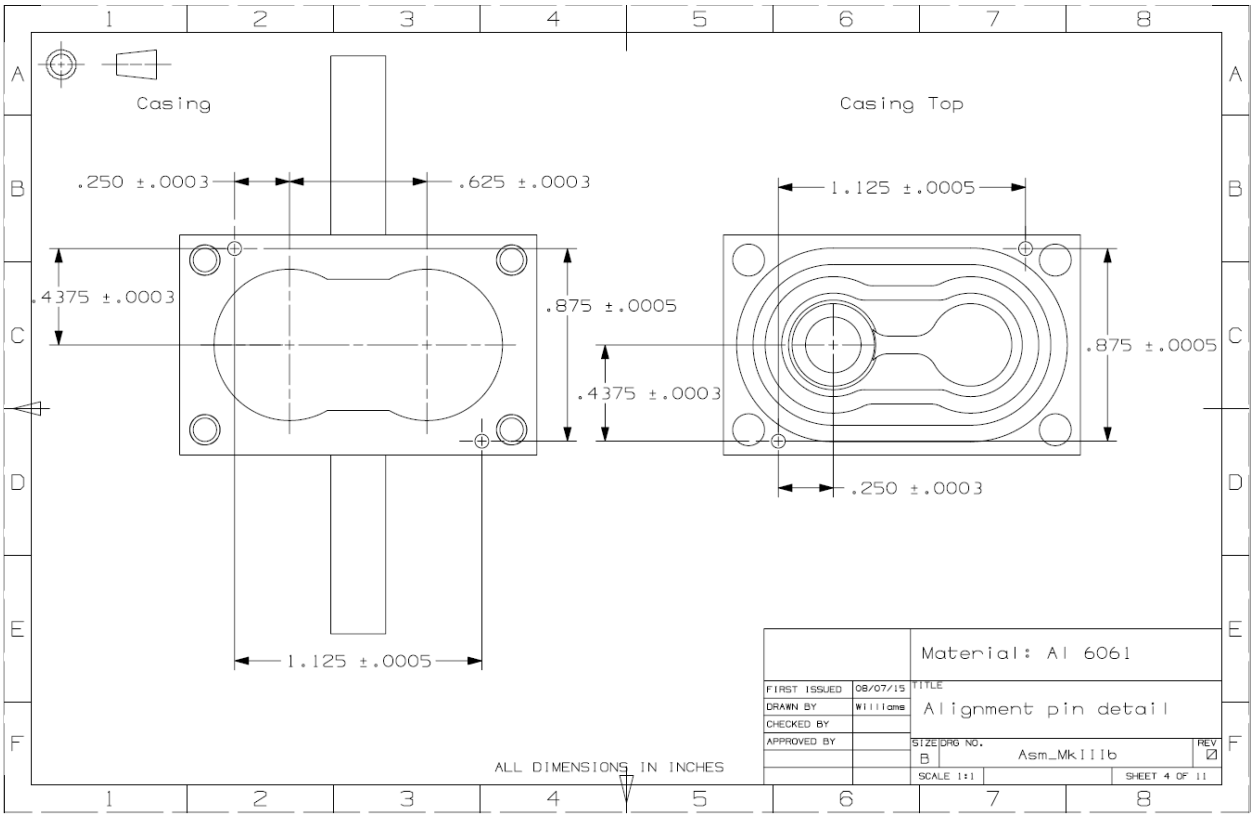

Fig. 84—Example fabrication drawing showing feature placement in relation to the alignment holes for the central casing element (*left*) and the top end plate (*right*)

6.2 **Seals**

Given the breadth of options for commercial seals available, a designer is usually better served selecting an off-the-shelf model, making this more a task of selection rather than design. However, certain seal models may have different requirements on the gland design. Therefore, when designing the casing, it is important to conduct seal selection for both static and dynamic seals in parallel to ensure that the seals are sufficient for the expected operating conditions while also having sufficient space in the casing for the corresponding gland. Seals also will have specific requirements for gland surface finish for optimum performance. Choice of material also should be made with an eye to material compatibility with the working fluid, both for longevity of the seals as well as for any potential fluid contamination concerns.

6.2.1 *Static Seals*

The static seal between casing elements is a relatively simple design challenge that usually can be addressed with standard o-rings. The theory behind leakage in this application and the idea behind o-ring seals already has been discussed in Section 2.1.3. For convenience and reduced cost, it is usually recommended to place the gland so that the resulting perimeter matches a standard o-ring diameter. Custom-diameter o-rings can be produced from o-ring stock and rubber adhesive, or by additive manufacturing, but rarely will such a custom configuration merit the added cost.

Each vendor of o-rings usually will have guidelines for designing the gland. These usually are presented in the form shown in Fig. 85, with the relevant gland dimensions and parameters. The main gland dimensions are the gland depth, L, and the gland width, G. O-ring catalogues often will present gland

dimensions in tabular form, with gland depth and width varying with each o-ring size. The guidelines also will present recommended feature dimensions, such as the fillet radius, R, the maximum gland wall angle, φ, and the corner radius, r_c. Recommended wall surface finishes, R_a, also will be given, which usually will not be uniform between all surfaces. While static seals are a relatively simple design feature, it is important to keep these manufacturer guidelines in mind to ensure optimal performance of the seals to prevent leakage.

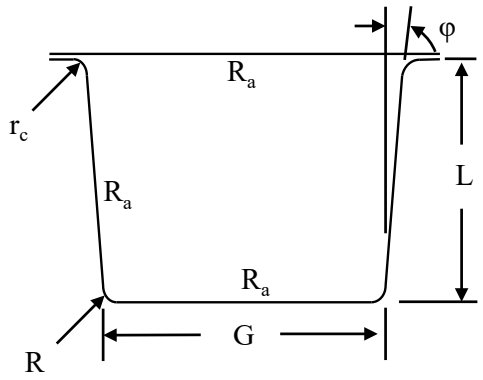

Fig. 85—O-ring gland design parameters

6.2.2 *Shaft Seals*

Sealing the junction between the drive shaft and the casing is a more complex and sensitive challenge than the junction between two casing elements, as this is a dynamic interface with the drive shaft in motion in close proximity to the stationary casing. While a basic o-ring seal can be used for low-speed, low-pressure pumps, more complex, compound seals generally are used that may consist of multiple elastomers or backing springs. The various configurations of seals generally fit on a spectrum between lip seals and squeeze seals, generally determined by the shape and structure of the seal and the magnitude with which it grips the shaft. Lip seals have smaller radial stiffness and are characterized by low friction and low wear at the cost of reduced sealing capability with pressure. Squeeze seals are designed to have greater stiffness, which generates a greater contact stress on the shaft over a larger area; this reduces leakage at higher pressures at the cost of greater frictional losses [8].

The two primary factors driving shaft seal selection are the anticipated fluid pressure along the drive shaft and the surface speed of the shaft. The fluid pressure along the drive shaft may not necessarily be at discharge pressure, depending on the design of the pump. It is usually advantageous to design the bearings to drain to suction pressure to ensure sufficient fluid flow and to reduce fluid pressure at the shaft seal. However, it is important to remember that fluid pressure almost certainly will be higher that suction pressure due to nonzero hydraulic resistance between the bearings and pump suction. It also can be useful to consider scenarios in which pump pressure may surge above nominal levels, such as relief valve failure or system blockage, and to design to seal against pressures higher than expected. The second factor, surface speed, is defined as

$$u_s = r_s \omega \,, \tag{6.12}$$

where r_s is the radius of the shaft and ω is the angular speed of the shaft. At sufficiently high surface speeds, the fluid entrained on the surface of the shaft can lift the seal off the surface of the shaft, referred to as hydroplaning, resulting in leakage. Some seal configurations are better suited to inhibit hydroplaning than others, though each configuration comes with its own set of trade-offs. High surface speeds also can result

in frictional heating of the seal, particularly for squeeze seals. Therefore, it is important to weigh the nominal pump speed, the drive shaft diameter, and the expected duty cycle of the pump and to estimate potential thermal effects. Vendors often will list what surface speeds and pressures a given seal is rated for.

Once a shaft seal has been selected, the corresponding gland in the casing can be designed. Vendors generally will list the gland specifications, similar to what is discussed in Section 6.2.1. One important distinction is the need to ensure that the shaft seal remains stationary relative to the casing. Shaft seals generally are designed so that the dynamic interface is between the seal and the shaft, not the seal and the casing. For this reason, glands for shaft seals are designed particularly to grip the seal with a greater overall force than the shear force between the seal and the shaft. Tolerancing on the gland is likely to be stricter than with a basic o-ring gland, and there is usually a minimum surface roughness needed on one of the gland walls to ensure a sufficient coefficient of friction. Similarly, some seals may impose a maximum shaft surface roughness to minimize wear on the seal, as well as a minimum hardness on the shaft material to prevent erosion of the shaft. These requirements are listed by the seal vendor, and it is important for a designer to keep these seal requirements in mind when designing the shaft.

6.3 **Assembly**

Assembly practices will vary with the specific design and application of the pump, tempered by cost and workspace considerations. Ideally, all pumps should be assembled in a clean environment free of particulates to minimize contamination. Larger pumps and low-pressure pumps, which are less reliant on very tight clearances, are usually more tolerant of assembly conditions that are less than ideal. However, where possible, it is recommended to use these safe practices to ensure maximum performance and pump lifespan.

The first step to pump assembly is to process and clean the individual components. It is conservative to assume that all parts not specifically listed as cleaned or assembly-ready should be cleaned to remove any machining fluid, particulates, or other residue. Immersed cleaning in an alcohol bath in an ultrasonic cleaner is ideal, with hand cleaning with alcohol at a minimum. Handling of the parts should take place in a clean workspace that is not downwind of any source of particulates. Use of lint-free or reduced-particulate cloths is recommended to prevent introduction of new contaminants. Once each part has been cleaned and dried, it should be placed in a clean, sealed container until assembly begins.

Once all components are cleaned and prepared, arrange them on a clean surface along with a sample of the working fluid. In general, components in any dynamic interface or with tight clearances should be assembled with surfaces wetted by the working fluid. In practice, that means almost every surface except the exterior of the casing. Begin with installing wetted seals into their glands and staging all relevant fasteners, shims, and other related components. Next, press fit any shaft keys and attach the gears to the corresponding shafts, making sure to wet every surface of the gear. For smaller gears, a hypodermic syringe can be a useful means of distribution of the fluid. If the shaft bearings are embedded in the wear plates, insert them now, and then affix the wear plates to the shafts. For non-spur gears, such as helical or herringbone gears, both gears must be meshed first and then inserted together into the wear plates, as mesh can be achieved only radially, not axially. Once the core pump stack has been assembled, it can be inserted into the main casing element. Due to the tight clearances between the pump stack and the casing, errors in tolerancing can make this a very tight fit. Make sure that the wear plates are able to move along the axis of the shaft relatively freely, as excessive friction can inhibit performance of floating wear plates. Gradual polishing of the inner wall can be performed, though it is not ideal. The final step is to mount the casing end plates and to secure the fasteners to the established preload torque. A photographic flow chart of this process with a miniature gear pump is presented in Fig. 86.

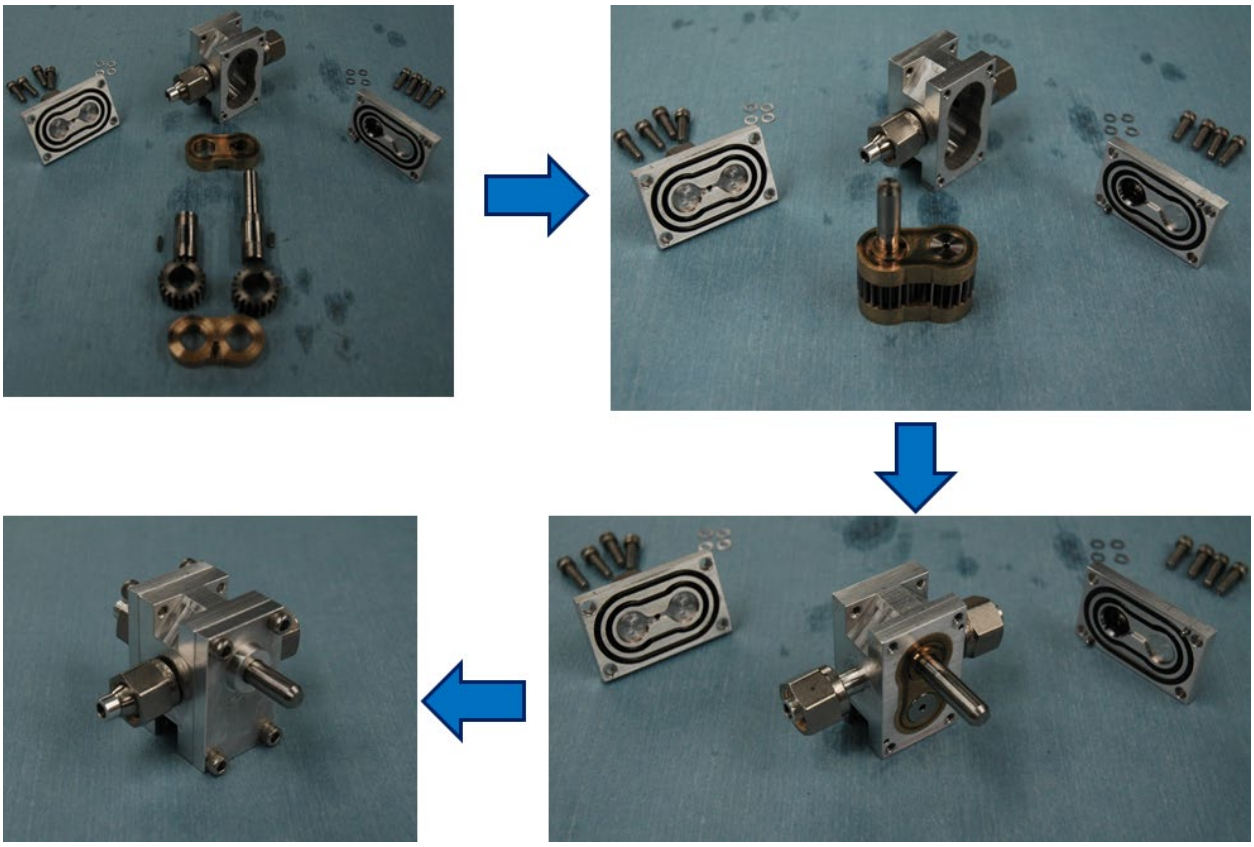

Fig. 86—Example assembly of a miniature gear pump

6.4 **Performance Evaluation**

Once a pump model has been designed, fabricated, and assembled, the next step in the process is to test the pump to measure the performance. A designer should never assume that a pump design will function as intended on the first attempt, but rather should expect some amount of iteration. Evaluation of gear pump designs has two main facets: the ability to test the pump in controlled conditions with all operating conditions monitored, and a framework to allow meaningful comparisons between design iterations across the range of possible operating conditions. The former can be achieved with a modular pump testbed that allows full control over all operating conditions of the pump while measuring fluid pressure, flow rate, and temperature while also measuring input mechanical power. The latter is obtained from the performance evaluation methodology presented in Section 1.4 that allows direct comparison of pump design iterations across disparate operating conditions, reducing experimental constraints.

6.4.1 *Pump Testbed*

At a basic level, the pump testbed consists of a modular mounting system (e.g., an optical table) to which system elements can be attached. These elements include the pump, diagnostics, the prime mover, fluid plumbing, filters, and a variable hydraulic load. Typical diagnostics include a pressure transducer on the suction and discharge sides of the pump, an in-line fluid thermocouple to measure fluid temperature at suction, and a flow meter downstream of the pump discharge to measure the output flow rate. It is a safe practice to place inline filters at the intake from the reservoir and between pump discharge and the test hydraulic load to prevent particulate contamination of either the pump or the load. The pore size of the filter depends on the internal clearances within the pump, but 7 μm is a conservative starting point. It is

recommended that the fluid reservoir of the system be pressurized above atmospheric pressure (1–2 bar or more, depending on the intake filter and pump speed) to avoid cavitation within the pump. The powertrain of the prime mover usually will include some electric motor, an encoder to measure shaft speed, and a torque transducer to measure delivered shaft torque. A schematic of an example pump testbed is shown in Fig. 87.

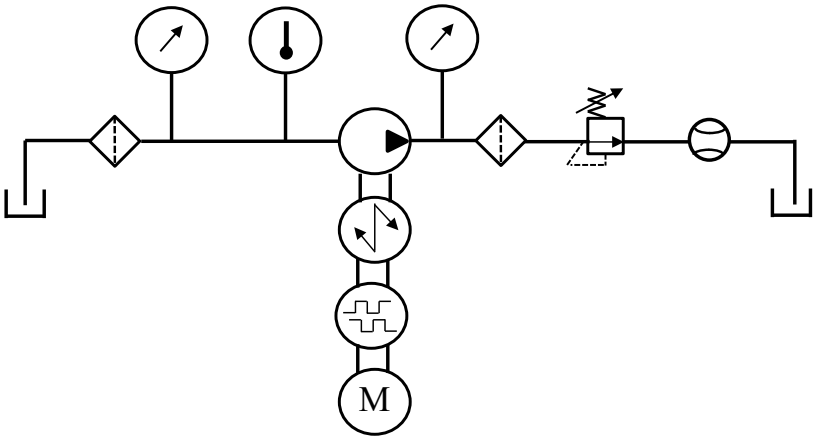

Fig. 87—Pump testbed hydraulic schematic

6.4.2 *Wear-In*

The first hour or so of operation of a pump is when a majority of the dynamic interfaces wear in. Two primary examples are the gear teeth against the casing wall and the axial faces of the gears against the wear plates. It is important to note that a noticeable change in performance may be observed as surfaces are worn smooth and harder components impose a complementary conformity on softer parts. Typically, frictional contributions to shaft torque should decrease over time, improving mechanical efficiency. Some of these effects can be seen as quickly as after a few seconds of operation, as shown in Fig. 88. Due to this erosion of material surfaces, increased particulate formation from the pump should be expected during wear-in. This may necessitate cleaning or replacing the discharge filter once wear-in has been completed.

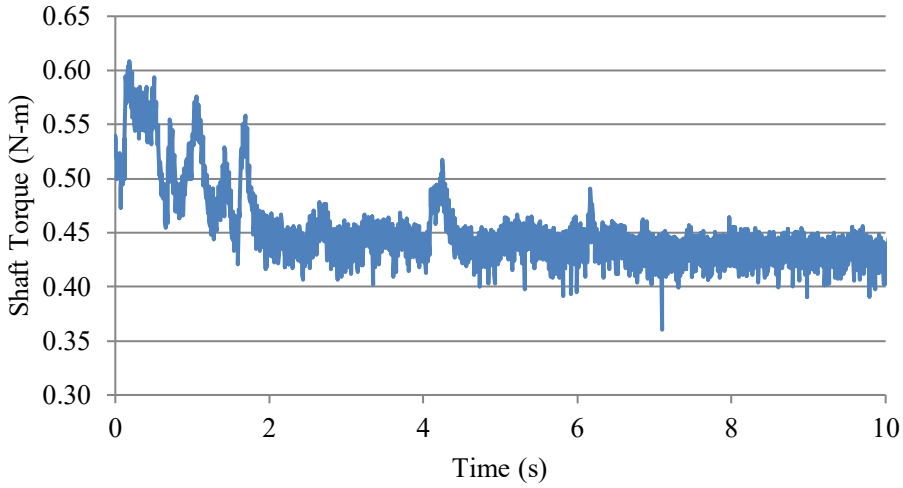

Fig. 88—Measured shaft torque as a function of time for initial wear-in of a miniature gear pump

It is important to note that the wear-in process creates a unique conformation between the individual components within the pump, particularly between the gears and the wear plate or the casing. Despite a common manufacturing process, different samples of the same part may have unique features or defects, or may sit in different points within the same dimensional tolerancing guidelines. As such, replacing components after wear-in may not necessarily result in the exact same performance as measured before. One particular example is swapping gears with the same nominal diameter and width while maintaining the other components. Should the new set of gears have a marginally smaller addendum diameter, the radial clearance around the gears will be larger, leading to increased leakage. Conversely, using a set of gears with a slightly larger diameter within the tolerance range may result in additional wearing in. This would provide similar performance with the new set of gears, but attempting to reuse the old set of gears later may result in a marked decrease in pump performance. A comparison of pump volumetric efficiency is shown in Fig. 89 with the same set of gears but before and after running the pump with an alternative set of gears with the same nominal size. The fact that the alternative set of gears had an addendum diameter that was on the higher end of the tolerance band meant additional wear of the casing occurred, reducing subsequent volumetric efficiency when the original gears were restored.

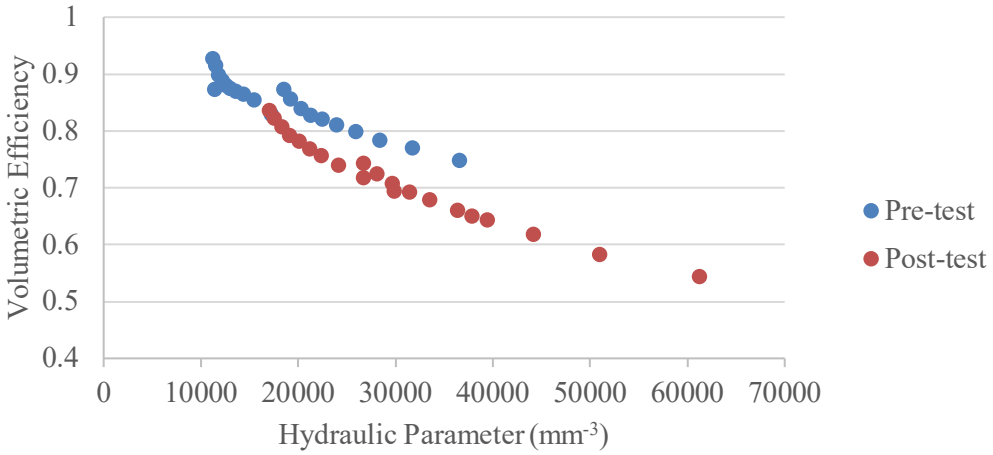

Fig. 89—Volumetric efficiency as a function of the hydraulic parameter of a pump with a set of gears before and after testing the pump with an alternative set of gears. Both performance curves correspond to pump performance with the original gear set.

7. CONCLUSION

The information presented here establishes a foundation of the key concepts and some guidelines for the design, fabrication, and testing of external gear pumps. With this, a designer should be able to develop a functional gear pump design from initial estimates of desired performance and to conduct proof-of-concept performance evaluation. From this beginning, iteration and optimization of the design can be done, as well as further research into the multiple specialized topics of gear pump design and operation, such as cavitation. A designer should not consider this work to be the totality of the knowledge on external gear pumps, but rather the starting point. The hope is that information presented here will provide designers with enough understanding to frame the questions that must be answered along their own unique design processes and a starting point from which to begin.

8. REFERENCES

[1] I. J. Karassik, J. P. Messina, P. Cooper, and C. C. Heald, *Pump Handbook* (Fourth Edition, McGraw-Hill, New York 2008).

[2] P. Michael, H. Khalid, and T. Wanke, "An Investigation of External Gear Pump Efficiency and Stribeck Values," SAE Technical Papers, 8, (2012).

[3] C. Schänzle, N. Störmer, and P. Pelz, "Modeling the Efficiency of External Gear Pumps Based on Similarity Considerations," Proceedings of the BATH/ASME 2018 Symposium on Fluid Power and Motion Control, (University of Bath, UK 2018).

[4] M. Borghi, F. Paltrinieri, B. Zardin, and M. Milani, "External Gear Pumps and Motors Bearing Block Design: Influence on the Volumetric Efficiency," Proceedings of the 51st National Conference on Fluid Power, pp 557-571, (2008).

[5] C. Hooke and E. Koç, "End Plate Balance in Gear Pumps," Proceedings of the Institution of Mechanical Engineers. Part B, Journal of Engineering Manufacture, 198, pp 55-60, (1984).

[6] Parker Hannifin Corporation, Parker O-ring Handbook, ORD 5700, (2007).

[7] Apple Rubber Products Inc., Seal Design Guide, (2009).

[8] Parker Hannifin Corporation, Fluid Power Seal Design Guide, EPS 5370, (2007).

[9] N. D. Manring and S. B. Kasaragadda, "The Theoretical Flow Ripple of an External Gear Pump," ASME Journal of Dynamic Systems, Measurement, and Control, 125, pp. 396-404, (2003).

[10] S. Wang, H. Sakurai, and A. Kasarekar, "The Optimal Design in External Gear Pumps and Motors," IEEE/ASME Transactions on Mechatronics, 16, pp. 945-952, (2011).

[11] P. Casoli, A. Vacca, and G. L. Berta, "A Numerical Model for the Simulation of Flow in Hydraulic External Gear Machines," Proceedings of the Conference on Power Transmission and Motion Control, (University of Bath, UK, 2006**),** pp. 147-168.

[12] A. Vacca and M. Guidetta, "Modeling and Experimental Validation of External Spur Gear Machines for Fluid Power Applications," Simulation Modeling Practice and Theory, 19, pp. 2007-2031, (2011).

[13] J. E. Shigley and L. D. Mitchell, *Mechanical Engineering Design*, (Fourth Edition, McGraw-Hill, New York, 1983).

[14] R. L. Norton, *Machine Design: An Integrated Approach*, (Second Edition, Prentice-Hall, Upper Saddle River, NJ, 2000).

[15] E. Koç, "Analytical and Experimental Investigation into the Sealing and Lubrication Mechanisms of the Gear Ends in Pumps," Wear, 135, pp. 79-94 (1989).

[16] E. Koç, A. O. Kurban, and C. J. Hooke, "An Analysis of the Lubrication Mechanisms of the Bush-Type Bearings in High Pressure Pumps," Tribology International, 30, pp. 553-560 (1997).

[17] E. Koç, and C. J. Hooke, "An Experimental Investigation into the Design and Performance of Hydrostatically Loaded Floating Wear Plates in Gear Pumps," Wear, 209, pp. 184-192 (1997).

[18] M. Borghi, M. Milani, F. Paltrinieri, and B. Zardin, "Studying the Axial Balance of External Gear Pumps," Commercial Vehicle Engineering Congress and Exhibition, 2005-01-3634 (2005).

[19] S. Dhar and A. Vacca, "A Novel CFD – Axial Motion Coupled Model for the Axial Balance of Lateral Bushings in External Gear Machines," Simulation Modelling Practice and Theory, 26, pp. 60-76 (2012).

[20] Committee of Stainless Steel Producers, "Review of the Wear and Galling Characteristics of Stainless Steels," A Designers' Handbook Series, (1978).

[21] OC Oerlikon Balzers AG, "Coated Components: Greater Performance and Reliability," (2010).

[22] G. B. Dubois and F. W. Ocvirk, "Analytical Derivation and Experimental Evaluation of Short-Bearing Approximation for Full Journal Bearings," National Advisory Committee for Aeronautics, Report 1157, (1953).