Optimal Design of Compact Spur Gear Reductions

M. Savage, S.B. Lattime, and J.A. Kimmel
The University of Akron
Akron, Ohio

and

H.H. Coe
Lewis Research Center
Cleveland, Ohio

Prepared for the
Sixth International Power Transmission and Gearing Conference
sponsored by the American Society of Mechanical Engineers
Scottsdale, Arizona, September 13–16, 1992
OPTIMAL DESIGN OF COMPACT SPUR GEAR REDUCTIONS

M. Savage, S.B. Lattime, and J.A. Kimmel
The University of Akron,
Akron, Ohio 44325

H.H. Coe
National Aeronautics and Space Administration
Lewis Research Center
Cleveland, Ohio 44135

ABSTRACT

The optimal design of compact spur gear reductions includes the selection of bearing and shaft proportions in addition to the gear mesh parameters. Designs for single mesh spur gear reductions are based on optimization of system life, system volume, and system weight including gears, support shafts, and the four bearings. The overall optimization allows component properties to interact, yielding the best composite design. A modified feasible directions search algorithm directs the optimization through a continuous design space. Interpolated polynomials expand the discrete bearing properties and proportions into continuous variables for optimization. After finding the continuous optimum, the designer can analyze near optimal designs for comparison and selection. Design examples show the influence of the bearings on the optimal configurations.

NOMENCLATURE

A distance from inboard roller bearing to gear or pinion, in.
a bearing life adjustment factor
B distance from outboard ball bearing to gear or pinion, in.
B gear tooth surface material constant, psi
b Weibull slope
C center distance, in.
C dynamic capacity, lb
D shaft outside diameter, in.
E elastic modulus, psi
c_t goodness of fit error limit
F force, lb
F_d dynamic tooth load, lb
f gear face width, in.
∇f unit gradient in the feasible direction
∇h unit gradient in the violated constraints
J AGMA bending strength tooth form factor
l service life, hr
Δ load cycle count
M merit function
∇m unit gradient in the merit function
N number of teeth
P_d diametral pitch, in.^{-1}
p load life factor
R reliability
ΔS optimization step size
V inequality constraint
v bearing load adjustment factor
X independent design parameter
Y scaled independent design parameter
Γ gamma function
θ characteristic life, hr
v Poisson’s ratio
ρ radius of curvature, in.
The design of compact gear sets is an optimization problem which has received considerable attention (Tucker, 1980; Savage et al., 1982; Carroll and Johnson, 1984; Savage et al., 1991; and Errichello, 1989). Tucker (1980) summarized the traditional gear design tradeoff between involute interference and tooth bending strength. The traditional method yielded many successful gear designs without the aid of a digital computer. Savage et al. (1982) applied computer optimization to find the compact gear mesh designs which balance resistance to gear tooth bending, gear surface pitting, and gear tip scoring. Carroll and Johnson (1984) extended this approach to search in the actual design space of whole tooth gears with the AGMA bending strength formulas. In a preliminary effort to this research, Savage et al. (1991) applied the modified gradient optimization procedure of this work to the compact gear mesh design problem. Similar results to earlier optimization studies were obtained. Errichello (1989) adapted the optimal gear design determination of minimum weight gears back to a hand calculation procedure. Once a design optimum has been identified, he rightly states that similar designs can be obtained without retracting the optimization path for each design.

However, optimization with the assistance of the modern digital computer offers a designer the opportunity to expand the scope of the gear selection process. Initial efforts in optimal gear design have focused on the gears due to the complexity of the gear mesh and its loading. One logical extension is to treat the entire transmission as a complete system.

In aircraft transmissions, service life between overhauls is affected mainly by the lives of the bearings which support the gears (Astridge and Savage, 1990; and Lewicki et al., 1986). Instead of designing optimal gears and then sizing bearings and support shafts to accommodate the gears, a better approach would be to design the transmission as an optimized system composed of gears, bearings, and shafts. The overall service life, weight, and size of the reduction are more important properties than the life, weight, and size of any component.

Lundberg and Palmgren (1952) developed a theory for the life and capacity of ball and roller bearings. This life model is based on the two-parameter Weibull statistical distribution (Weibull, 1951) and is the international standard for bearing life and capacity calculations (Harris, 1984). Coy et al. (1976) extended this theory to describe the pitting fatigue life of spur gears. In turbine powered transmissions which see little shock loading, the pitting fatigue of the gear teeth is the dominant life parameter for gears designed to withstand tooth breakage and scoring. The pitting mode of failure exhibits no threshold strength for infinite life.

Savage et al. (1989) combined these models into a system life model based on the two-parameter Weibull distribution for transmissions composed of bearings and gears. This system model enables the design of a transmission for long service life in combination with other transmission properties such as low weight and compact size.

Many optimization techniques are available from zero order random search procedures through first order gradient and modified gradient procedures to more complex higher order procedures (Vanderplaats, 1984). A modification of the feasible directions gradient method of Zoutendijk (1960) provides balance between algorithm complexity and efficiency for the desktop computing environment. All discrete variables and parameters are fit with continuous polynomials to allow gradient calculations. After finding a continuous optimum, one can select and analyze proximate realistic designs.

This paper describes the computerized optimal design of an enclosed parallel shaft spur gear reduction. The reduction should transmit an input torque and speed to an output shaft at a given reduction ratio. Standard ball or roller bearings may support the input and output shafts in one of several pre-selected configurations. Inequality constraints restrict the designs to have adequate tooth bending strength, tooth scoring resistance, involute contact geometry, and shaft strength and stiffness at the full AGMA estimated dynamic load (AGMA Standard, 1988). Independent design parameters include the gear mesh diametral pitch and face width, the number of teeth on the pinion, the axial locations of the four support bearings, and both shaft diameters.

The merit function for this study is the transmission life divided by the cubed product of the component weights and the transmission volume. This combined objective criterion yields compact designs with long lives. After finding a continuous optimum, the program displays the values and properties so the designer can modify the design by selecting nearby realistic parameter values. The program checks and displays the modified design and once again, gives the designer the opportunity for modification. Optimal designs are reported and compared.

TRANSMISSION MODEL

Configurations

The transmission configurations for design are variations of the single mesh spur gear reduction. An input shaft supports a pinion on two rolling element bearings, and an output shaft supports a larger gear on two more rolling element bearings. Standard 100-, 200-, or 300-series ball or roller bearings may support the input and output shafts. The bearing type and series are specified for a given design as are the bearing life and capacity adjustment factors. Designs with all ball bearing support, all roller bearing support, and a mixture of ball and roller bearing support on each shaft are possible.

Three bearing support configuration geometries, shown in Fig. 1, are treated as separate design problems. These configurations are:

(1) Overhung support for both gears with the pinion support bearings on one side and the gear support bearings on the other side

(2) Overhung support for both gears with all supporting bearings on the same side of the two gears

(3) Straddle support for both the pinion and gear.

For all configurations, the total gear mesh force acts along the line of action with radial and tangential components in the direction of the...
(a) Opposite overhung.
(b) Single sided overhung.
(c) Straddle.

Figure 1.—Bearing support configuration geometries.

Figure 1

line of centers and pitch circle tangent. The full load produces the resultant radial reactions on the bearings, while the radial component produces the separating deflection of the pinion and gear shafts and the tangential component produces the misaligning shaft slope along the gear teeth.

Gear Strength

The bending fatigue model uses the AGMA I factor (AGMA Standard, 1988) to estimate the bending stress with the load at the highest point of single tooth loading on the pinion. The load is estimated as the full dynamic load, $F_d$, using the AGMA velocity factor model (AGMA Standard, 1988). The formula for the bending stress is:

$$\sigma_b = \frac{F_d \cdot P_d}{f \cdot J}$$

The maximum contact stress and gear tip Hertzian pressure are calculated as:

$$\sigma_H = \left[ \frac{F_d}{\pi f \cos \phi} \left( \frac{1}{\rho_p} + \frac{1}{\rho_g} \right) \left( \frac{1 - \nu_p^2}{E_p} + \frac{1 - \nu_g^2}{E_g} \right) \right]^{1/2}$$

The maximum contact stress occurs at the lowest point of single tooth contact on the pinion tooth. The gear tip Hertzian pressure uses one half of the total dynamic load since the load is shared between two tooth pairs at this point.

The gear tip scoring model includes the pressure times velocity factor and the critical oil scoring temperature model from lubrication theory. The normal pressure times sliding velocity is proportional to the frictional power loss of the gear set. This factor is the highest for contact at the gear tip, where the normal pressure is the gear tip Hertzian pressure.

Life

The life model comes from rolling element bearings (Lundberg and Palmgren, 1952). Lundberg and Palmgren determined that the scatter in the life of a bearing can be modeled with a two-parameter Weibull distribution:

$$\ln \left( \frac{1}{R} \right) = \left( \frac{t}{\delta} \right)^b$$

In terms of a 90-percent probability of survival life, $\xi_{10}$ the two-parameter Weibull distribution is:

$$\ln \left( \frac{1}{R} \right) = \ln \left( \frac{1}{\alpha} \right) \cdot \left( \frac{t}{\xi_{10}} \right)^b$$

The life to reliability relationship of Eq. (4) is for a specific load which determines the $\xi_{10}$ life. This load, $F$, is related to the component dynamic capacity, $C$, as:

$$\xi_{10} = \left( \frac{C}{\sqrt{F}} \right)^p$$

Since the life at the dynamic capacity is one million load cycles, it does not appear as a variable in the equation.

The Rolling-Elements Committee of the Lubrication Division of the ASME modified Eq. (5) with adjustment factors (Bamberger et al., 1971). These factors extend Eq. (5) to cover many different end use situations so that designers can size bearings properly. The revised code equation for $\xi_{10,a}$, the adjusted 90-percent reliability life for the bearing, is:

$$\xi_{10,a} = a \left( \frac{C}{\sqrt{F}} \right)^p$$

In Eq. (4), the Weibull slope, $b$, is normally $10/9$ for ball and $9/8$ for straight roller bearings, and in Eqs. (5) and (6), the load-life exponent, $p$, is 3.0 for ball bearings and 3.33 for roller bearings.

Gear tooth pitting failures are similar to bearing failures, with the possible difference of surface initiation. So the two-parameter Weibull distribution also describes the scatter in gear life with a different Weibull slope. The life of a gear, $\xi_{10,a}$, is related to the life of a tooth on the gear, $\xi_{10,t}$, by:

$$\xi_{10,a} = \xi_{10,t}$$

A relationship for the dynamic capacity, $C_t$, of a spur gear tooth as a function of Buckingham's load-stress factor, $B$, which has the dimensional units of stress (AGMA Standard, 1988) is:

$$C_t = B \left( \frac{f}{\rho_g \cdot \rho_p} \right)$$

The dynamic capacity of the gear, $C_g$, in terms of the dynamic capacity of a single tooth, $C_t$, is:

$$C_g = \frac{C_t}{N_g \cdot (b_p) \cdot (b_p)}$$

3
Drive system reliability, $R_s$, is a strict series probability of all the component reliabilities (Savage, et al., 1989), which makes it the product of the reliabilities of all the components. The system life can also be expressed as a two-parameter Weibull distribution in terms of the system reliability parameters, $b_s$ and $\theta_{10,s}$:

$$\log \left( \frac{1}{R_s} \right) = \log \left( \frac{1}{0.9} \right) \cdot \left( \frac{t_{10,s}}{\theta_{10,s}} \right)^{b_s}$$

(10)

In terms of the $t_{10}$ life, the mean time between overhauls for full transmission replacement is the mean life, which is calculated using the gamma function:

$$t_{av} = t_{10} \cdot \Gamma (1 + 1/b) \left[ \frac{1}{\ln (1/0.9)} \right]^{1/b}$$

(11)

**Size**

Two measures of transmission size are incorporated in the merit function: component weights and transmission volume. The component weights include the weights of steel discs with the outlines of the components. All discs are hollow with the shaft outside diameter taken as the inside diameter of the gears and bearings and the shaft inside diameter being 1/2 in. less than the outside diameter. The second measure of transmission size is the volume of the smallest rectangular solid which includes the two gears and four bearings.

The bearing properties of the outside diameter, width, and dynamic capacity are expanded from catalog data (Harris, 1984) with low-order polynomial curves. The polynomials describe continuous properties which increase monotonically with bearing bore for the gradient calculations. On checking the designs, a table look up of properties for standard size bearings produces practical bearing characteristics.

**OPTIMIZATION METHOD**

As with most optimization techniques, the procedure begins with several vectors. These vectors are the independent design variables, X; the inequality constraints, V; the parameters of the merit function, Pr; and the constants which define the specific problem, C. An optimization solution is the design variable values, X, which minimize or maximize the merit function value while maintaining all constraint values, V, inside their specified limits. A procedure starts with a guess for the design variable, X, and iterates with some logic to find the optimal design variable.

To maintain balance among the independent design parameters, the design space is scaled into a continuous, dimensionless design space. The scaled design parameters, Y, vary from -1.0 to +1.0 as specified by upper and lower bounds on the independent design parameters, X.

**Gradient**

Central to the method is the gradient calculation. This is performed with small perturbations in the design variables from the nominal position. The gradient in the merit function, $\nabla M$, is the vector of partial derivatives in the merit function, M, with respect to the scaled design variables, Y:

The magnitude of the gradient vector is given by:

$$\|\nabla M\| = \left[ \sum_{j=1}^{n} \left( \frac{\partial M}{\partial Y_j} \right)^2 \right]^{1/2}$$

(12)

For minimization, the direction of change in $Y$ which reduces the merit function, $M$, at the greatest rate is determined by the unit vector, $\nabla m$:

$$\nabla m = -\frac{\nabla M}{\|\nabla M\|}$$

(13)

For maximization, the sign in Eq. (13) reverses.

In the simple gradient method, Eq. (13) defines the direction for the step change in the scaled design vector:

$$Y_{j+1} = Y_j + \Delta S \nabla m$$

(14)

where $\Delta S$ is the scalar magnitude of the step. If no constraints are violated, this will be the next value for $Y$ in the search.

**Step Size**

Step size, $\Delta S$, is a significant element of any optimization procedure (Vanderplaats, 1984). For stability and directness, the step size is nominally fixed. Initially, the step size is 5 percent of the range of a single design parameter. But the procedure halves the step whenever a local minimum is reached or the step is trapped in a constraint corner. On completion, the search declares a solution when the percent change in the merit function, M, is less than a pre-set limit.

$$\left| M_{j+1} - M_j \right| < \epsilon$$

(15)

**Initial Value**

The optimization procedure described above is scaled, fixed step, and steepest decent. When the initial guess is in the acceptable design space, and the optimum is a relative minimum, this method works quite well. However, placing the initial guess in the acceptable design region is often difficult. Also when the best design is determined by a "trade-off" among conflicting design constraints at the edge of the feasible design space, a direct merit function gradient iteration cannot slide along the constraint boundary toward the optimum.

These problems are addressed with a second gradient in a constraint variable:

$$\nabla V_k = -\frac{\nabla V_k}{\|\nabla V_k\|}$$

(16)

where $\nabla V_k$ is a unit vector in the direction of decreasing value in the constraint, $V_k$. For upper bound constraints, moving through the design space in the direction of $\nabla V$ reduces the constraint value $V_k$. For lower bound constraints, a sign reversal in Eq. (16) produces an increase in the constraint value, $V_k$, for motion in the gradient direction. The vector sum of the gradients in the violated constraints, $\nabla h$, is the second gradient of the algorithm:

$$\nabla h = \sum_k \nabla V_k$$

(17)

The gradient in the violated constraints, $\nabla h$, points towards the acceptable design space from the unacceptable design space. By itself, it enables the algorithm to turn an unacceptable initial guess into an acceptable trial design by a succession of steps:

$$Y_{j+1} = Y_j + \Delta S \nabla h$$

(18)

**Feasible Direction**

Once inside the acceptable design region, the algorithm proceeds along the steepest descent direction until the calculated step places the
next trial outside the acceptable design space. To avoid this condition, the algorithm selects a feasible direction for the next step. Figure 2 shows a constraint limit intersecting contour lines of improving merit function values. The figure shows gradients in the merit function, \( \nabla m \), and the impending constraint, \( \nabla h \). The feasible direction selected, \( \nabla f \), is the unit vector sum of these two gradients:

\[
\nabla f = \frac{\nabla m + \nabla h}{|\nabla m + \nabla h|}
\]

And the next design step becomes:

\[
Y_{j+1} = Y_j + \Delta S \nabla f
\]

![Merit contour](image)

Figure 2.—Feasible direction gradient vector.

Algorithm Use

By using subroutines to calculate the merit function and constraint values for each design trial, the procedure separates the logic of the algorithm from the analysis necessary to define the problem. This allows the design problem to be changed without concern for the optimization procedure. The directness of the procedure adds additional steps, but enables the program to run on a personal computer.

An additional benefit of separating the analysis routines from the optimization logic is the ability to modify the design at execution and verify the characteristics of similar, more practical designs with the same program. The optimization procedure works with a continuous design space, which includes gears with fractions of teeth and nonstandard bearings. By allowing the user to see the ideal continuous variable solution and to modify this to designs with whole numbers of teeth and standard sizes, the procedure enables a designer to determine a practical optimum design easily.

TRANSMISSION DESIGN

Consider the design of a 2:1 gear reduction to transmit an input torque of 600 lb/in. at 1000 rpm at a power level of 9.5 hp. Due to the higher loads near the gears, roller bearings are placed adjacent to the pinion and gear and lower capacity ball bearings are placed at the more lightly loaded outboard positions.

Designs are obtained with all bearings of the same series or with roller bearings of one series and lighter series ball bearings. Support configurations include opposite overhung support, single-sided overhung support and straddle support as shown in Fig. 1.

Nine independent design variables were used for the design problem:

1. Gear diametral pitch (in.), \( P_d \)
2. Gear mesh face width (in.), \( f \)
3. Number of pinion teeth, \( N_p \)
4. Distance from the pinion to its roller bearing (in.), \( A_p \)
5. Distance from the pinion to its ball bearing (in.), \( B_p \)
6. Distance from the gear to its roller bearing (in.), \( A_g \)
7. Distance from the gear to its ball bearing (in.), \( B_g \)
8. Pinion shaft diameter (in.), \( D_p \)
9. Gear shaft diameter (in.), \( D_g \)

The first three determine the gear design and the last six define the location and size of the support bearings. Figure 3 shows these location and size parameters. The number of teeth on the gear is twice that on the pinion and the gear center distance, \( C \), is three-halves of the number of teeth on the pinion divided by the diametral pitch.

The design constraints include:

1. Pinion tooth bending strength limit (Savage et al., 1991)
2. Gear tooth surface contact stress limits (Savage et al., 1991)
3. Gear scoring limits of contact pressure times sliding velocity and oil flash temperature (Savage et al., 1991)
4. Involute interference avoidance limit (Savage et al., 1982)
5. Minimum axial clearance between the pinion or gear and its roller bearing equal to one-half the combined width of the gear and bearing
6. Shaft outside diameter limit below the maximum gear rim diameter for a rim thickness equal to three-halves of the tooth height
7. Maximum pinion or gear slope limit (Young, 1989)
8. Maximum pinion or gear deflection limit (Young, 1989).

For the single-sided overhung support and straddle support cases, an additional constraint is added which limits the sum of the maximum pinion bearing outside diameter and the maximum gear bearing outside diameter to be less than twice the center distance by a specified clearance.

The merit function to be maximized in the design is the transmission life as measured by its mean time to failure divided by the cube of the product of the transmission volume and the weight of its components.

![Figure 3.—Location and size design variables.](image)
Examples

For a first try at this design problem, a different merit function of life divided by volume and weight was used. In this case, the bearing life dominated the designs. A trade-off design, for which the merit function is at a relative maximum, was reached, but the design was lightly loaded and large. Figure 4 is a schematic of the overhung bearing design for 300-series roller bearings close to the gears and 200-series ball bearings away from the gears. A 1-in. square box is included for scaling.

![Figure 4](image)

The gears are thin and of large diameter with a face width of 0.1875 in. and a center distance of 12.25 in. Table 1 lists the design parameter and merit function component values for this and all reported designs. The first column lists the initial guess values for all designs, and the second column lists this design's values. The roller bearings are placed as close to the gears as possible, considering the fifth design constraint of axial clearance, and the shaft sizes are large with a pinion shaft diameter of 90 mm and a gear shaft diameter of 85 mm. A long mean transmission life between overhauls of 9.75×10^6 hr or 1120 years results.

This transmission configuration favors the bearings. Large diameter gears with thin widths carry small loads. Large shaft diameters and small bearing overhangs place strong bearings under minimum load to maximize the life of the inboard roller bearings. Since these are the most heavily loaded bearings in the transmission, maximizing the life of the weakest component, maximizes the life of the transmission.

In obtaining this design, the pinion and gear shaft diameters were increased from the initial values to improve the bearing capacities. In the final constraint summary, the active constraints were the axial clearances between the roller bearings and the pinion and gear. Bringing these bearings as close as possible to the two gears reduced the loading on the bearings while enlarging the shaft diameters increased their capacities. The transmission has a volume of 6345 in.^3 and a component weight sum of 44.6 lb. The gears are not loaded to their capacity with a pinion bending stress of 85 percent allowable and a maximum surface contact pressure of two-thirds the limit.

When the optimizing program used the merit function of the transmission life divided by the cube of the product of the transmission volume and weight, the selected designs were smaller and had more balance between the gear and bearing capacities. Figure 5 is a schematic diagram of the design for the same conditions as that of Fig. 4 with only the merit function changed. A shorter life of 55 970 hr or 6.4 years of continuous service is available from a gearbox which is smaller at 710 in.^3 and lighter with 21.2 lb of components. The design also has changed shape with the gears having a greater face width of 0.75 in. and smaller center distance of 4.5 in. The parameters of this design are listed in the third column of Table 1.

![Figure 5](image)

**TABLE 1 - DESIGN PARAMETER VALUES**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Initial guess</th>
<th>Opposite overhung</th>
<th>Opposite overhung^2</th>
<th>Single overhung</th>
<th>Stacked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merit function</td>
<td></td>
<td></td>
<td>Figure 4</td>
<td>Figure 5</td>
<td>Figure 6</td>
<td>Figure 7</td>
</tr>
<tr>
<td>Diametral pitch</td>
<td>in.²</td>
<td>14</td>
<td>3</td>
<td>16</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>Mesh face width</td>
<td>in</td>
<td>1.0</td>
<td>0.1875</td>
<td>0.75</td>
<td>0.5</td>
<td>0.625</td>
</tr>
<tr>
<td>Pinion tooth number</td>
<td></td>
<td>53</td>
<td>68</td>
<td>48</td>
<td>46</td>
<td>50</td>
</tr>
<tr>
<td>A_l roller distance</td>
<td>in.²</td>
<td>3.0</td>
<td>1.875</td>
<td>2.375</td>
<td>1.75</td>
<td>2.75</td>
</tr>
<tr>
<td>B_g ball distance</td>
<td>in.²</td>
<td>5.0</td>
<td>6.875</td>
<td>5.25</td>
<td>5.25</td>
<td>3.875</td>
</tr>
<tr>
<td>A_r roller distance</td>
<td>in.²</td>
<td>3.0</td>
<td>1.875</td>
<td>2.5</td>
<td>2.0</td>
<td>2.75</td>
</tr>
<tr>
<td>B_g ball distance</td>
<td>in.²</td>
<td>5.0</td>
<td>6.25</td>
<td>5.0</td>
<td>5.25</td>
<td>3.875</td>
</tr>
<tr>
<td>Pinion shaft diameter</td>
<td>mm</td>
<td>30</td>
<td>90</td>
<td>65</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td>Gear shaft diameter</td>
<td>mm</td>
<td>50</td>
<td>55</td>
<td>70</td>
<td>60</td>
<td>55</td>
</tr>
<tr>
<td>Transmission life</td>
<td>hr</td>
<td>9 747 600</td>
<td>55 970</td>
<td>148 500</td>
<td>414 100</td>
<td></td>
</tr>
<tr>
<td>Transmission volume</td>
<td>in.³</td>
<td>6345</td>
<td>710</td>
<td>560</td>
<td>630</td>
<td></td>
</tr>
<tr>
<td>Transmission component weight</td>
<td>lb</td>
<td>44.6</td>
<td>21.2</td>
<td>18.8</td>
<td>21.1</td>
<td></td>
</tr>
</tbody>
</table>
Higher volume and weight penalties have led to a more heavily loaded transmission. Since the gear load life factor of 8.93 (AGMA Standard, 1988) is nearly three times the roller bearing load life factor of 10/3, gear life affects heavily loaded designs more than lightly loaded designs which are strongly influenced by bearing life. The selected bearings have shorter lives than those of the first design.

Maximizing gear life alone would make the face width larger and the gear diameters smaller to conserve space. The pinion diameter was at its minimum possible for the shaft size which the bearing capacities required. In this design, the roller bearings were not pulled in to the gears as tightly as they were in the first design. Pinion and gear slope limits contributed to the shaft sizes and the gear bending stress was at the allowable limit, showing the balance in the design.

Applying the same procedures to the other two cases of single-sided overhung support and straddle support configurations produced similar results. In both cases, the designs were more compact and had longer lives. An additional interference constraint was added to both cases to maintain a clearance separation between the bearings of the two shafts. Figures 6 and 7 show the design configurations for single-sided overhung transmissions and straddle support transmissions. Their design and merit function parameters are listed in the last two columns of Table I.

Both of these designs were obtained with the merit function which divides the transmission life by the cube of the product of transmission volume and component weight. And the active constraints for both designs were the gear and pinion shaft slope limits. As the designs become more heavily loaded and the shafts reduce in size, the ability of the shaft to maintain gear and roller bearing alignment decreases.

The single-sided overhung design of Fig. 6 has a center distance of 5.75 in. and an average life of 148 500 hr or 17 years of continuous service before pitting failures require maintenance. Its pinion tooth bending stress is also at the design limit.

The straddle design of Fig. 7 uses the same set of 300-series roller bearings and 200-series ball bearings for comparison with the other designs. The lighter capacity ball bearings are placed further from the gears to reduce their loading in optimizing the life of the transmission. The pinion tooth bending stress is 85 percent of the allowable stress, its center distance is 5.36 in., and the transmission life is 414 800 hr or 47.5 years — longer than the other compact designs.

Which merit function should be used to select a design is not obvious. It should be set by the relative importance of life, size, and weight in the application. To get a more direct comparison between designs, one could select one factor in the merit function such as life and make it a lower bound against which to minimize the weight and volume of the transmission. If this were done, all designs would have the same life and comparisons could be based on size and weight considerations.

CONCLUSIONS

Optimization has been applied to the design of a transmission in the form of a single gear mesh reduction with its two support shafts and four support bearings using a system description of the transmission. The object of the design was to determine a small, light-weight transmission with a long service life. System models of the transmission's mean service life, volume, and component weights were obtained.

Three configurations of the transmission were treated: double overhung, single sided overhung, and straddle support. Nine independent design parameters and 26 different design constraints were applied in the optimization. The merit function was the transmission life divided by the cube of the product of the volume and weight.

The optimal designs were determined by a combination of bearing, shaft, and gear characteristics with the bearings and support geometry influencing the designs more than the gear mesh parameters. For lightly loaded designs, the bearing properties dominated the optimal designs.

With the interactive nature of the optimization procedure, actual bearing and gear proportions can be entered in a checking stage to determine the performance of realistic designs. Graphical output aids the interpretation of different designs. Additional changes in the merit function and constraints of the optimizing procedure may tailor the design procedure to match different actual design situations easily.
REFERENCES


**Optimal Design of Compact Spur Gear Reductions**

**M. Savage, S.B. Lattime, J.A. Kimmel, and H.H. Coe**

The optimal design of compact spur gear reductions includes the selection of bearing and shaft proportions in addition to the gear mesh parameters. Designs for single mesh spur gear reductions are based on optimization of system life, system volume and system weight including gears, support shafts and the four bearings. The overall optimization allows component properties to interact, yielding the best composite design. A modified feasible directions search algorithm directs the optimization through a continuous design space. Interpolated polynomials expand the discrete bearing properties and proportions into continuous variables for optimization. After finding the continuous optimum, the designer can analyze near optimal designs for comparison and selection. Design examples show the influence of the bearings on the optimal configurations.