GEOMETRY-BASED SHAPE OPTIMIZATION TO IMPROVE DYNAMICAL BEHAVIOR OF PLANETARY GEAR BOXES IN AERO ENGINES

Aris Konstantinidis  
Chair of Aero Engine Design  
BTU Cottbus-Senftenberg  
03046 Cottbus, Germany  
Aris.Konstantinidis@b-tu.de

Klaus Höschler  
Chair of Aero Engine Design  
BTU Cottbus-Senftenberg  
03046 Cottbus, Germany  
hoeschler@b-tu.de

Patrick Hauk  
Rolls-Royce Deutschland Ltd & Co KG  
15827 Blankenfelde-Mahlow, Germany  
Patrick.Hauk@Rolls-Royce.com

1. INTRODUCTION

The constant pursuit for higher fuel efficiency has led the development of aero engines in the past decades. Meanwhile, thermodynamic efficiency has reached a very high level, due to increased combustor exit temperatures and engine pressure ratios. Therefore, the focus moved to increasing propulsion efficiency by larger bypass ratios, to further improve specific fuel consumption significantly. Engines with large bypass ratios demand a different architecture and increasing the bypass ratio above 10 makes it attractive to design an engine with a gearbox, separating the shaft of the low-pressure turbine system from the fan. The aerospace environment requires the gearbox to be applicable for high speeds and high torque while being lightweight, raising many design challenges. [1] [2] [3]

Challenges like these necessitate sophisticated design processes. Optimizing designs in the early development stages become increasingly important to minimize the number of necessary design iterations, resulting in time and cost savings. Common methods include topology, size and shape optimization. These optimization strategies typically discretize either the design space or the component, using a finite element approach. Interpreting the results and transferring them back into a functional CAD model is a difficult, often time-consuming and has been a focus of research in the past. [4] [5] [6]

Topology optimization has become a well-established structural optimization method in the past years and of its key advantages is the fast number of possible designs it can create. However, a lightweight gearbox in an aerospace application requires specific flexibilities and stiffness criteria for many components. This forms a challenging task for commercial topology optimization solvers, often leading to unwanted results as described by Deng [7] and De Lima [8], although this problem has been addressed in academia far earlier, for example by Kikuchi et al [9]. Along with flexibility requirements and linear elastic analysis, buckling under high torque needs to be considered as well. Besides improving structural properties on a component-based level, it is also important to optimize dynamic behavior on a
system level, as unwanted Eigen modes in the running range can lead to hazardous failures in the gear train during extreme events.

The objective of this paper is to present an alternative approach that allows the structural optimization of highly flexible structures. The component is optimized to improve dynamic behavior of the system, via an automated simulation process. Creating a bridge between component design parameters and the system's response improves the design process in the early development stages. While improving the system's dynamic behavior, flexibility, buckling and weight requirements on a component level are considered. The approach uses a CAD based geometry and parametrization, avoiding the issue of transferring the optimized results back into a functional CAD model. Three different parametrization strategies will be compared, aiming to provide a large design variety. Furthermore, the approach allows the use of multi-objective optimization algorithms, which produce many optimal designs on a pareto front. This provides more information to the design engineer and gives the ability to trade off different optimization goals.

PROBLEM DESCRIPTION

A planetary gearbox consists of a sun gear located in the center, surrounded by a number of planetary gears and a ring gear around it, as shown in figure 1. The input torque is applied to the sun gear, while the output torque is either extracted through the ring gear with the planetary gears fixed or through the planetary gears with the ring gear fixed. Holding the ring gear fixed leads to a greater gear ratio and is the setup for this example. With a fixed ring gear, the gearbox is attached to the surrounding engine architecture via the ring gear mount (RGM), a thin cylindrical component. The stiffness characteristics of the RGM impact the overall dynamic behavior of the system and also the load distribution within the system.

![Figure 1: Planetary Gear Box Example](image1)

Gear train analysis has identified the mode shown in figure 2 to be critical and within the running range. In this mode, the ring gear mount moves upwards and the sun gear moves down and across towards the gap in between two planet gears. Due to a helical gearing the planet gears rotate as the sun gear protrudes the gap. From a side view the top and bottom planets shift back and forth along with the ring gear mount as it tilts. The sun gear rotates around its center axis against the RGM. This mode in planetary gear trains has also been identified and analyzed in greater detail by [10].

![Figure 2: Critical Eigen Frequency Mode](image2)

The ring gear mount also needs to meet minimum flexibility criteria to protect the gears from unwanted load distributions and to decouple the gearbox from external disturbances. Therefore it is not possible to increase the system’s Eigen frequency by stiffening the mounting generally. In this case, the gear train analysis states maximum stiffness criteria for axial stiffness ($K_a$), radial stiffness ($K_r$), bending stiffness ($K_b$) and bending coupling stiffness ($K_{bc}$). The RGM must not buckle under load and needs to be lightweight. This leads to a tightly constraint optimization problem, making it difficult to improve the baseline design by engineering judgement alone.

In summary, the goal is to optimize the design parameters ($p$) that describe the ring gear mount’s shape to increase the bending mode frequency ($f$) while minimizing its mass ($m$) and staying within the maximum limit for stiffness in particular directions ($K_i$), buckling Eigen value ($b$) and design space ($d$) constraints:

$$\min_{p \in P} \{-f\}(p)$$

with $P := \{p \in \mathbb{R}^n | K_i(p) \leq 0, \ b(p) \leq 0 \leq b^u \}$ (1)

2. METHODS

Figure 3 visualizes the optimization process. The optimization algorithm generates the input parameters that create the CAD geometry based on the individual
optimization strategy. In this case, the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is used for single-objective optimizations and the Archived-based Micro Genetic Algorithm (AMGA) is used for multi-objective optimizations. Both algorithms are well suited for non-linear, non-convex optimization problems. They are less suited for long running simulation, since they require many design iterations but allow parallelization to reduce the required optimization time. More information about those algorithms can be found in [11] and [12] respectively. In principle, the approach shown in this paper is independent of the chosen algorithm.

The CAD geometry is altered based on the optimization parameters and delivers the resulting mass. It is also verified that the resulting geometry continues to stay within the boundaries of the design space. For each iteration the finite element (FE) model is automatically remeshed, while the boundary conditions remain. The FE model can now be solved for different structural problems. Problems like linear elastic, non-linear or Eigen frequency analysis of the component are implemented, but are outside of the scope of this example. In this case a reduction of the stiffness and mass matrix to the interface points and a linear buckling analysis is performed. After the component analysis is completed, the resulting stiffness properties and buckling Eigen values are passed back to the optimization algorithm. The reduced RGM Model is then implemented into a larger systems model, which is used for Eigen frequency analysis on a system level, using the FE solver NASTRAN. The CAD models of the optimal designs are stored and can easily be utilized to transfer the results into other CAD models or for further analysis.

**FIGURE 3: OPTIMIZATION PROCESS**

The automation of the analysis process and the integration of multiple models into a single analyses loop is one of the key challenges of this approach. The utilized workflow comprises the automatization of design and simulation tasks in Siemens NX 12 by using NXOpen, its Python API and the optimization tools within Isight 6. Any results from the simulation, such as displacements, stresses, reaction forces, average and maximum values in a specific area or the relation between different results of any complexity can be accessed and used. This makes this approach very flexible for a wide range of engineering tasks.

Since this approach is based on a CAD geometry, the modelling approach is a main focus. Splines are common to describe complex shapes with few parameters and are well suited for optimization tasks. In regards to spline definition and usage, Siemens NX allows multiple ways to create splines. Two common ways are described as “through points” and “by poles” in the user interface. Internally NX iteratively solves the equations (3) and (4) respectively for each definition to display the resulting spline. The B-splines \([x(u),y(u)]^T\) is described with the basis functions \(N_k(u)\) and \((n + 1)\) control points \((x_k,y_k)\).

\[
\begin{align*}
\begin{bmatrix} x(u) \\ y(u) \end{bmatrix} &= \sum_{k=0}^{n} N_k(u) \begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} X_i \\ Y_i \end{bmatrix} \quad (3) \\
\begin{bmatrix} x(u) \\ y(u) \end{bmatrix} &= \sum_{k=0}^{n} N_k(u) \begin{bmatrix} x_k \\ y_k \end{bmatrix} \quad (4)
\end{align*}
\]

From a user perspective, the “by poles” definition (4) allows the manipulation of the control points directly. The “through points” definition creates a set of \((n + 1)\) points \((X_i,Y_i)\) in the geometry. The curve must pass through. Then the corresponding control points are internally calculated by Siemens NX.

For the purpose of demonstration, two splines with a third degree base function are created with five defining points each, see figure 4. One uses the “through points” definition and the other uses the “by poles” definition. Each point is defined by two coordinates \([X_i,Y_i]\) and their values range between \(0 < X_{2,3,4}, Y_{1,2,3,4} < 100\), except for the horizontal position of the start and endpoint fixed at \(X_1 = 0, X_5 = 100\). Based on this design space 200 parameter combinations have been created using an optimal-latin-hypercube distribution. The set of parameter combinations is used to create 200 spline variations for each definition. In the context of a CAD geometry, a spline is typically used to extrude or revolve a surface or volume from it. Therefore, a self-intersecting spline will lead to a failure in the geometry generation and needs to be avoided. Since the parameter range for each point is very large in this example, many failed design can be expected.

Comparing the two possible spline definitions, “by poles” shows a lower failure rate, see Table 1. The “through points” definition is especially critical if the control points are allowed to move in a way that their position, counting from the beginning of the spline to the end, is interchangeable. Figure 4 illustrates that issue. A lower failure rate will improve optimization performance, as every
failed CAD geometry generated by the optimization algorithm is not returning any information. However, in cases with small spline shape variations and unchanged control points order, the “through points” definition still offers benefits and is applicable, since the maximum and minimum coordinates of the spline are close to the control point coordinates and no large overswings occur.

Performing a study, judging the parametrization quality and the chosen spline description, is recommended, since a failed geometry cannot provide any information to the optimization algorithm. Further, only the CAD geometry is build and no simulation is needed, therefore the computational effort is low.

**PARAMETRIZATION STRATEGY**

The parametrization strategy of the CAD geometry plays a decisive role in this approach. It forms the CAD model, on which the simulation is based, determines the design space for the optimization and has a significant impact on the results. Size-, shape- and topology optimizations are common approaches to optimize structural properties and the CAD parametrization can be chosen in a way that each of these optimization strategies can be conducted by the approach. The goal of the parametrization is to cover a wide range of different and feasible designs, while using as few parameters as possible. The parametrization and the resulting designs also need to be considered in the FE model, to ensure a stable remeshing and consistent boundary conditions. In regards to manufacturability, any self-intersecting designs are rejected by the process. Further, designs with a local curvature higher than the accepted threshold are also rejected. Since the component is milled, almost all non-self-intersecting designs with a low local curvature can be manufactured.

The focus of the optimization is on improving the component’s shape. A topological redesign of the ring gear mount is not within the scope of this example. However, potentially large design changes and exploring a vast variety of designs is an objective. Every parametrization in this paper displays a cross section of the RGM, which is revolved around the engines center axis. Three optimizations have been performed, one for each approach. The first and second parametrization vary the center line of the cross section, affecting the overall shape via straight lines and a spline, while the third parametrization varies the thickness distribution along a given center line.

**TABLE 1: SPLINE DEFINITION COMPARISON**

<table>
<thead>
<tr>
<th></th>
<th>designs created</th>
<th>successful designs</th>
<th>failure rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>through points</td>
<td>200</td>
<td>92</td>
<td>54%</td>
</tr>
<tr>
<td>by poles</td>
<td>200</td>
<td>142</td>
<td>29%</td>
</tr>
</tbody>
</table>

Figure 5 shows the first parametrization, which connects four points through straight lines. The points zero and three can move vertically with their horizontal positions fixed at 0 and l. The points one and two can move vertically and horizontally within the given design space. The lines are joined by two constant radii $R$. The minimum distance to the design space outline is measured for each design variation and has to be greater than zero. The goal is to understand the possible improvement regarding the increase of the critical Eigen frequency, while retaining a simple and easy to manufacture shape. Since the baseline design is robust against buckling and the wall thickness remains unchanged, the optimization target function is simplified to:

$$
\min_{p \in P} m(p)
$$

with $P := \left\{ p \in \mathbb{R}^n \mid \left[ K \right] \left( p \right) \leq 0, \quad p^l \leq p \leq p^u \right\}$ (2)
Figure 6 shows the second parametrization, using seven points and twelve design parameters. The seven points form a B-spline via the “by poles” definition in the CAD suite. The points zero and six can move vertically within the design space, with their horizontal position fixed at 0 and \( l \). The points one to five can move vertically and horizontally inside as well as outside the design space. The minimum distance to the design space outline is measured and has to be greater than zero again, ensuring the resulting B-spline stays within the given design space. The goal is to find a further improvement regarding the target function (2) compared to the first approach, by trading in some manufacturability due to a more complex shape.

Figure 7(A) and (B) show the third parametrization, which uses two B-splines to form the inner and outer contour of the RGM. The inner contour remains static and is derived from either the baseline design or the result of an earlier shape optimization. Points are created along the inner contour curve length \( u \) to the spline. This leads to an uneven distribution of support points along the spline, with more points in areas of greater curvature and fewer points in areas of less curvature. The number of support points created varies depending on the curvature and the predetermined distance \( s \). At each point \( u_i \) a support line of length \( t_i \) is created perpendicular to the inner contour. The endpoints of the support lines represent the defining points of the outer contour B-spline through (3). The length \( t_i \) is determined through a third, formative curve (5), shown in figure 7(B), which describes the thickness distribution along the component. The formative curve is a B-spline created with \( (m+1) \) evenly distributed defining points and the “through points” definition. The parameters \( p_k \) are the variable design parameters for the optimization.

\[
\left[ \begin{array}{c} v(r_f) \\ t(r_f) \end{array} \right] = \sum_{k=0}^{m} N_k(r_f) \left[ \begin{array}{c} v_k \\ t_k \end{array} \right] \frac{k/m}{p_k}, r \in [0,1] \subset \mathbb{R} \tag{5}
\]

The variable \( r \) runs along the spline and is defined to be between 0 and 1. With the inner counter curve length given as \( u_{\text{max}} \), the corresponding position of each support point \( u_i \) in the formative curve coordinate system is given by (6).

\[
v_i = \frac{u_i}{u_{\text{max}}} v_{\text{max}}
\tag{6}
\]

\[
v(r_f) = v_i \tag{7}
\]

The next step is to find the corresponding curve parameter \( r_f \) where the spline is located at \( v_i \) in the coordinate system. The first equation of (5) and a root search algorithm is used to find \( r_f \) where (7) is satisfied. The found \( r_f \) value can then be used in the second equation of (5) to find \( t_i = t(r_f) \) which represents the wall thickness at \( u_i \) of the ring gear mount. Alternatively, the distance from the v-axis to the spline at \( v_i \) can also be measured within the CAD suite to determine the length of \( t_i \) as an associative variable.

It should be noted that the control points of the formative spline are independent of the number and position of the contour splines support points. The variable spread of support points leads to a more accurate representation of the thickness distribution along the component, given by the formative spline. By using a formative spline to derive the thickness distribution from, instead of varying each length \( t_i \) individually, the number of design parameters is reduced, improving optimization performance. In comparison, the formative spline also leads to a smoother, more gradual change in thickness, which is favorable in many use cases.
This approach has also been used in a different context and described in more detail by Chao [13].

3. RESULTS AND DISCUSSION

The optimization target parameters will be referenced as factors of the baseline designs values, with a baseline mass \( m_\# \) and critical Eigen frequency \( f_\# \). The structural optimization of the RGM component has been performed as an iterative process, increasing design complexity from one step to the next. In general, this approach allows the design engineer to gather information about a particular problem and improve his parametrization strategy decisions for the next step.

As a starting point an optimization using the CMA-ES algorithm and the parametrization shown in figure 4 has been performed, resulting in a design that improved the critical Eigen frequency to \( 1.135f_\# \). Note that CMA-ES is a single objective algorithm and the components mass is not considered in the target function (2). However, it is possible to include the mass into the optimization through a weighted target function or as an optimization constraint. The optimized design shown in figure 8 increased the diameter of the upper and lower horizontal section and moved the vertical section right up to design space boundary. Further improvement of the design is limited by the design space and the radial stiffness constraint, indicated in figure 13.

Although, the top section of the second design appears similar to the first, it does provide more insight to the optimization problem. Concluding from the first optimization, the axial stiffness constraint can be assessed sufficiently by increasing the diametric separation between the upper and lower horizontal section and the radial constraint becomes the primary limiting factor to further improvement. Different design ideas arise for the top section, adding bellows and sweeps to potentially gain further improvement. Four control points for the B-spline in the top section increase the design flexibility in this area, accounting for those ideas. Multiple optimizations, varying the mutation parameters and the initial design, converge back to the design in figure 9. Therefore, the simple right-angle shape for the top section can be accepted as optimal.
The last optimization is based on the result in figure 9, as it aims to improve the thickness distribution along the spinal shape of the second optimization. 20 support points are used along the component’s length for an accurate representation of the thickness distribution provided by the formative spline, which is shaped by seven design parameters. This time, a multi-objective optimizing, targeting maximal critical Eigen frequency and minimal mass as objectives, has been performed, using the AMGA strategy. A multi-objective optimization does not provide a single solution as a result, but a pareto optimal front of many designs. The resulting front is shown in figure 11. A single objective optimization, only maximizing the critical Eigen frequency, will lead to the design at the top right corner of the pareto front and is shown in figure 10(A) with its thickness distribution in figure 10(B). The pareto front allows the design engineer to understand the tradeoff between mass and frequency for this particular design, providing a better understanding of the problem.

Figure 11 shows the pareto front and a linear relation between mass and Eigen frequency for lower frequencies and a drop off at the top for higher frequencies found. The drop off is a result of the radial stiffness constraint becoming active for the higher Eigen frequency designs as the optimization algorithm cannot increase the frequency as efficiently as for lower frequency designs, see also figure 12. Since the wall thickness is decreased in some areas of the component, buckling under maximum load needs to be considered. The buckling Eigen values for the pareto front are displayed by the coloring of each point in figure 11. Depending on the minimum safety regarding buckling, some lighter design may be excluded.

Summarizing the results, the performance improvements regarding (1) arise from an increase of radial and bending stiffness (\(K_r\) and \(K_b\)) and are limited by the maximum radial stiffness allowed, indicated by figure 13. Provided with this information, the engineer can choose a design with the most desirable property combination or choose to increase the maximum radial stiffness allowance to further increase the critical Eigen frequency. The first two optimizations show similar results, both creating a vertical section close to the design space boundary and increasing the diameter of the horizontal section which connects to the ring gear. With the increased design freedom of the second parametrization the optimization algorithm explored various design variations in this section and converged back to the simple shape. However, the lower horizontal section differed, creating a conical shape which led to an increase in radial stiffness, improving the target function. The third optimization focused on the thickness distribution. It increased the wall thickness along the conical shape and reduced the wall thickness along the vertical section towards the ring gear connection. This led to a slight further improve of the critical Eigen frequency and reduced the RGM’s mass by 5% compared to the baseline design, while being compliant with buckling requirements under torsional load. Since optimization process is based on a CAD geometry, a fully functional CAD model is directly available for each
design point. The results using the reduced model have later been confirmed in a higher fidelity, unreduced model.

4. CONCLUSION

The investigation in this paper demonstrates that the use of a CAD geometry-based optimization approach in conjunction with a reduced system model is suitable to improve dynamic behavior of a system by optimizing a component’s shape. Remeshing each design iteration allows for large deviations from the baseline design, while ensuring accurate analysis results. The low number of design variables, when compared to FE mesh-based shape optimizations, also allows for the utilization of various optimization algorithms, including evolutionary based algorithms. Those are well suited for non-linear, non-convex optimization problems. The ability to perform multi-objective optimizations provides the possibility to gather more information for the design engineer, improving the development process in the early stages. Using CAD model parameters as input values for a structural optimization allows a simple interpretation of the dependencies between input and target values. Extending the optimization process beyond the component level onto a systems level also extends this understanding of those dependencies. The investigation further showed that the rate of improvement heavily depends on the used parametrization.

Moreover, future investigations may include the optimization of multiple components and their interdependent behavior simultaneously, in order to improve the dynamic behavior of a system. With the increased number of optimization parameters, the use of response surfaces models to improve optimization time is an area to investigate. Structurally optimizing components that share the same or adjacent design spaces is called concurrent topology optimization and recently researched among others by Previati, Ballo and Gobbi [14]. Using CAD model based optimization strategies along with standard assembly design capabilities could provide an alternative approach to this problem.

ACKNOWLEDGEMENTS

This work has been carried out in collaboration with Rolls-Royce Deutschland Ltd. & Co.KG as part of the research project KEPLER (project number 80171294) funded by the State of Brandenburg and the European Regional Development Fund. Rolls-Royce Deutschland’s permission to publish this work is greatly acknowledged.

REFERENCES


