
This version is available from Sussex Research Online: http://sro.sussex.ac.uk/54502/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
Dynamics and tribological analysis of a toroidal CVT

K. Philpot\textsuperscript{1}, R. Glovnea\textsuperscript{2*}
\textsuperscript{1,2}University of Sussex, Brighton BN19QT, United Kingdom

Abstract

The continuously variable transmission investigated in this paper works with contacts in the elastohydrodynamic regime of lubrication, thus the tangential forces are transmitted between elements through the shearing the lubricant film. The behavior of the lubricant film when subjected to shear depends of the nature of the lubricant and the relative motion between the contacting surfaces. In this paper a non-Newtonian behavior is assumed for the lubricant while the relative motion is determined for every point on the contact area by kinematic methods. The net tractive force in the sliding direction and the spin torque are evaluated and from these the power losses in the contacts are calculated. The dynamic behaviour of the device is evaluated taking into account the behavior of the lubricant as extracted from traction measurements. A simplified method for the evaluation of the dynamic response of the device to a rapid variation of the resistive torque is presented. The results show the importance of rheological behaviour of the lubricant, measured by the slope of the linear region of traction curves.

Keywords: CVT, elastohydrodynamic, traction, dynamics

\textsuperscript{*}Corresponding author: Romeo Glovnea  (r.p.glovnea@sussex.ac.uk).

1. INTRODUCTION

Continuously variable transmissions (CVT) are mechanical transmissions which unlike ordinary gearboxes offer seamless change of the transmission ratio through a theoretically infinite range. There are many design configurations of these transmissions however they all have in common the continuous change of a length or angular dimension which results in an adjustment of the velocity of the output element. Most CVT designs include tribological contacts which can transmit power while changing the geometrical configuration of the contacting elements. This requires that the motion and torque between elements is transmitted through friction. The type CVT analysed in this paper is a member of a subset of continuously variable transmissions known as traction drives, in which the tangential force between rotating elements is transmitted through the shearing of a lubricant film. Of these, the toroidal design is best known, because devices of this type have been used in automotive applications. There is an extensive literature on the design, operation and modelling of traction drives. According to Loewenthal et al [1] toroidal traction drives were first used in the transmission chain of machinery in factories or in wood processing equipment, in which the power was transmitted between toroidal metallic discs via wooden or leather bound discs. Towards the end of the nineteenth and the beginning of the twentieth century the spread of motor vehicles gave a great momentum to the design of many kinds of transmissions which allowed seamless variation of speed, based on friction. As the power requirements increased, after an initial period of competitiveness, CVT transmissions lost out to traditional geared drives, mostly due to high wear rates, comparatively low power to weight ratio and complexity.

Due to their relative recent use in automotive industry toroidal and half-toroidal CVTs have received much attention during the past few decades. The research, both theoretical and experimental, focused on kinematic and dynamic analysis, for example in Nagata et al [2], Zhang et al [3], Cretu and Glovnea [4], power loss and efficiency due to Yamamoto et al [5], Newall et al [6], Burke et al [7], fatigue life and durability Coy et al [8], Nikas [9], and lubricant Tevaarwerk and Kohnson [10], Anghel et al [11], Ohno [12], to mention only a few of the published papers.

In this paper the authors carry out a kinematic and dynamic analysis of an un-conventional type of toroidal CVT markedly different from those currently used predominantly in automotive applications [4].
2. BACKGROUND

Elastohydrodynamic traction drives transmit tangential forces between contacting elements through shearing of the lubricant film. While the shape and thickness of the film have been revealed by theory and experiment [4, 5] the exact nature and the behavior of the elastohydrodynamic films under shear was intensely debated over the years. Traction curves can be relatively easily obtained in disc or ball machines but finding a model which fits the entire curve and is valid in any conditions of load, sliding speed and temperature proved to be a difficult task. Research done by Hirst and Moore [14], Bair and Winer [15], as well as Evans and Johnson [16] among many others has converged towards the concept that the fluid behaves in a visco-elastic manner. From the many non-Newtonian fluid models the one proposed by Hirst and Moore and Johnson and Tevaarwerk [17] was adopted in this work. This model incorporates both a linear, visco-elastic term and a non-linear term based on the Eyring theory of viscous flow. For shear stresses much larger than Eyring stress, which is the case with traction drives, the relationship between shear stress and shear rate is express in equation (Eq1).

\[
\tau \cong \tau_0 \ln \dot{\gamma} + \tau_0 \ln \left( \frac{2\eta}{\tau_0} \right)
\]  

(Eq. 1)

In this relationship \( \tau \) is the shear stress, \( \dot{\gamma} \) is the shear rate, \( \eta \) is the viscosity of the lubricant and finally \( \tau_0 \) is a characteristic of the fluid called Eyring stress.

3. ANALYSIS OF THE CVT

3.1. CVT design features

The design and operation of the CVT analysed in this paper have been described more rigorously in previous publications [e.g. 4]. It comprises two input discs, one conical and the other toroidal, as seen in Figure 1. The conical disc is fixed to the input shaft while the toroidal disc can slide axially on the shaft, but is fixed to rotate with it. The output disc is also conical but the contacting surface with the intermediary elements is on the inner side of the cone. The intermediary elements are balls, without a materialized axis of rotation. They are prevented from touching by a fixed separator, not shown in the figure.

![Figure 1. Schematic of the CVT](image)

An axial force, generated by various means (e.g. elastic, hydraulic, electro-mechanic) acts upon the toroidal disc, ensuring that the discs and the ball make contact at any time.

The output disc is connected to the output shaft by a coupling which translates a difference in torque between input and output shafts into an axial force, which displaces the output disc. This displacement of the output disc causes a radial displacement of the balls and thus the transmission ratio is changed. Due to this correlation between the output torque and the transmission ratio this CVT was also called constant-power CVT [4]. Apart from this feature the design has the advantage over other toroidal-type CVTs that there is no need of complex synchronization mechanisms for the intermediary elements, which choose their axis of rotation function of the geometry of the disc and the kinematic conditions.
3.2. Contact forces

For this analysis the contact points between the discs and one ball are denoted by A – toroidal disc, B – input conical disc, and C – output disc. The angles between the radii corresponding to these points and the horizontal are denoted by $\alpha$, $\beta$ and $\gamma$ respectively. The forces in the contacts between the discs and the intermediary balls arise from the axial forces acting on the toroidal input and conical output discs. These forces have to generate tangential forces large enough to transmit the required output torque. Free – body – diagrams of one ball is shown in Figure 2. The plane of the cut through the ball contains the axes of rotation of the ball and discs. The forces normal to the surface of the ball lie in this plane, while the traction (friction) forces are all perpendicular to this plane.

![Free-body diagram of one ball](image)

For design purposes the traction coefficient of the lubricant is limited to 0.045 in this analysis, in order to account for overloading or unexpected temperature rises. This is a rather conservative value as the maximum traction coefficient for specialized traction fluids is about 0.1 [11], but it is adopted by other researchers for the reasons mentioned above [e.g. 3]. If the output torque (that is the torque required to be transmitted) is $T_o$ and there are $n$ ball elements then the normal force acting in the contact between one ball and the output disc is given by Eq. 2.

$$F_{nc} = \frac{T_o}{\mu n R_c}, \quad \text{(Eq. 2)}$$

where $\mu$ is the traction coefficient, and $R_c$ is the radial distance of the contact point of the ball and the output disc and the axis of the disc. The distance $R_c$ can be written as a function of the radial position of the centre of the ball, which in its turn is a function of the angle of contact $\alpha$:

$$R_c = r + R \sin \gamma = r_0 - \Delta R \sin \alpha + R \sin \gamma \quad \text{(Eq. 3)}$$

In this relationship $r_0$ is the radius of ball centre when $\alpha$ is zero and $\Delta R$ is the difference between the minor radius of the torus $R_1$ and of the ball $R$. The normal forces in the other two contacts are then found from balance of forces, neglecting the weight of the ball.

$$F_{na} = F_{nc} \frac{\sin (\gamma - \beta)}{\sin (\alpha + \beta)} = \frac{T_o}{\mu n R_c} \frac{\sin (\gamma - \beta)}{\sin (\alpha + \beta)} \quad \text{(Eq. 4)}$$

$$F_{nb} = \frac{T_o}{\mu n R_c} \frac{\cos \alpha \sin (\gamma - \beta) + \cos \gamma \sin (\alpha + \beta)}{\cos \beta \sin (\alpha + \beta)} \quad \text{(Eq. 5)}$$

These forces are further used to calculate the dimensions of the contact area at each of the discs, and contact pressures. From $F_{na}$ the value of the force required to act axially upon the toroidal disc can also be calculated.
\[
F_x = \frac{T_m \sin(\gamma - \beta)}{\mu R_c \cos \alpha \sin(\alpha + \beta)}
\]  
(Eq. 6)

Depending on the type of system used to generate it the expression of this force can take different forms, as for example was calculated in [4] for elastic force. To be noted that contact forces, output torque and transmission ratio depend in the end of the variable angle \( \alpha \) between balls and toroidal disc.

3.3. Relative kinematics on contact area

The traction force in the elastohydrodynamic contacts is generated by the shear of the lubricant film, which implies a difference in the velocities of the contacting surfaces. This speed difference called slip velocity [10] denoted by \( \Delta U \), changes whenever there is a need to adjust the torque required to be transmitted.

From the point of view of distribution over the contact area, the slip velocity is constant throughout this area. The slip velocity is defined by the sliding/rolling ratio \( \xi = \Delta U / U \), which for traction drives takes typical values between 0.01 and 0.03. In the above relationship \( U \) is the rolling velocity, that is half the sum of the surface velocities of the two bodies.

Apart from the slip velocity there are other components of the velocity that are a byproduct of the CVT geometry and kinematics and that vary over the contact area. Tevaarwerk and Johnson [10] define these as side slip and spin. Side slip occurs when the axes of rotation of the contacting elements are not contained in the same plane. Due to the arrangement of the discs and balls and the fact that the axes of rotation of the balls are not materialized, side slip does not occur in the current design.

There is a component of the velocity which is perpendicular to the slip direction, but this is generated by a change of the transmission ratio rather than the relative position of the axes of rotation. For this reason the side slip will be ignored in this study.

![Figure 3. The contact between one ball and the toroidal disc](image3.png)

![Figure 4. Geometric parameters and angular velocities used for kinematic analysis](image4.png)
The example below is given for the contact between one ball and the input toroidal disc, but the same analysis is valid for the other two contacts. As seen in figures 3 and 4 three Cartesian systems of coordinates are chosen for this analysis. One attached to the disc, Ox₁y₁z₁, set with axis Oy₁ coinciding with the axis of rotation of the disc; another attached to the ball, Ox₂y₂z₂ with axis Oy₂ coinciding with the axis of rotation of the ball; and the third attached to the contact ellipse Axyz with axis z perpendicular to the contact surface. All three systems of coordinates have axes x parallel to each other while axes y and z are contained in the same plane. The angle between the axis of rotation of ball and that of disc is denoted by λ.

The spin velocity is according to Johnson [13] the difference between the components of the angular velocities of the disc and ball on axis Az.

\[ \omega_{sp} = \omega_{1,z} - \omega_{2,z} = \omega_{1} \cos \alpha + \omega_{2} \cos (\alpha - \lambda) \]  
(Eq. 7)

The spin motion results in additional components of the relative velocity in directions x and y:

\[ U_x = -y \omega_{sp} \]  
(Eq. 8)

\[ U_y = x \omega_{sp} \]  
(Eq. 9)

The expression of the relative velocity resulting from spin shows that the spin pole is located at the centre of the contact. This is an important finding as in this case the net force in the direction of slip, generated by the shear due to spin is zero. On the other hand it is obvious that the shear stress due to spin adds to that due to slip over one half of the contact and subtract over the other. Although the shear of the lubricant film due to spin does not generate a useful tractive force in the direction of slip, it does generate power loss.

There is a fourth component of the relative motion between the contacting surfaces which results from the difference in the distance between various points on the contact area and the axis of rotation of a given element. This relative motion can be called micro-slip.

On the contact ellipse an arbitrary point \( M(x,y) \) is chosen as seen in Figure 3. The velocities of point \( M \) belonging to the disc and ball are:

\[ \vec{U}_{1,M} = \vec{OM} \times \vec{\omega}_1 \]  
\[ \vec{U}_{2,M} = \vec{OM} \times \vec{\omega}_2 \]  
(Eq. 10)

Performing the cross product and resolving the resulted vectors on axes Ax and Ay, the components of the velocity in the plane tangent to the contact are found to be:

\[ U_{1,Mx} = -\omega_{1} \left( -\frac{r}{\tan \lambda} + R \cos \alpha + y \sin \alpha - z \cos \alpha \right) \]  
(Eq. 11)

\[ U_{1,My} = \omega_{1} \left( -\frac{r}{\sin \lambda} + R \cos (\alpha - \lambda) + y \sin (\alpha - \lambda) - z \cos (\alpha - \lambda) \right) \]  
(Eq. 12)

\[ U_{1,My} = \omega_{1} x \sin \alpha \]  
(Eq. 13)

\[ U_{1,Mz} = -\omega_{2} x \sin (\alpha - \lambda) \]  
(Eq. 14)

From the point of view of the shearing the lubricant film, the difference of the components of these velocities on axes Ax and Ay are needed. The component containing coordinate z, which results from the fact that the contact ellipse is not flat, can be written, according to Johnson [13] in terms of x and y as follows:

\[ z = \frac{y^2}{2R} + \sqrt{\frac{x^2}{(R - y^2/2)^2 + y^2}} \]  
(Eq. 15)

The micro-slip velocity is small in comparison to the slip but it still contributes to the power loss of the contacts. Similar components of the relative velocity resulting from spin and micro-slip are derived for the other two contacts.
With a model for the shear behavior of the lubricant chosen, these allow the calculation of the shear stress distribution of the contact area and subsequently of the power loss and efficiency of the elastohydrodynamic contacts. Calculations of the efficiency of the contacts were carried out for the following numerical values of the geometrical and working parameters of the CVT.

Radius of balls (R): 25 mm
Major radius of torus (r₀): 50 mm
Minor radius of torus (R₁): 43 mm
Number of balls (n): 4
Angle β of input conical disc: 20°
Angle γ of output disc: 62°
Traction coefficient (μ): 0.045
Input torque (T₁): 100 Nm
Maximum output torque (Tₒ): 182 Nm
Maximum Hertzian pressure at input toroidal disc: 2.95 GPa
Maximum Hertzian pressure at input conical disc: 3.23 GPa
Maximum Hertzian pressure at output disc: 2.59 GPa
Input speed: 3000 rpm

To be noted that the Hertzian pressure is not constant for all values of the transmission ratio. The values shown were obtained for a transmission ratio of 1.82, which for this configuration corresponds to an angle α = 58°. These values may seem large, but in practical situations the CVT does not work continuously at the peak transmission ratio, thus for most of the operating time the contact pressure is lower. For lowest transmission ratio, which in this case is 0.38, the maximum Hertzian pressure is generated at the input conical disc and it has a value of 2.47 GPa. The Hertzian pressure is a decisive factor in the design of a toroidal CVT, for a given application, as it limits the normal and tangential forces in the contact and consequently maximum torque which can be transmitted.

![Figure 5](image)

**Figure 5. Efficiency of the elastohydrodynamic contacts function of angle α**

Figure 5 shows the efficiency for all EHD contacts of the CVT for the configuration specified. For reference, the variation of the transmission ratio and that of the Hertzian pressure at the input conical disc, are also shown. A marked decrease of the efficiency at low values of the angle is observed. This is mainly due to the large sliding and spin velocities, which in turn depend on the angular velocities of the discs and balls. To be noted that at small values of angle α, corresponding to lower transmission ratios the angular velocity of the balls reaches large values. As the angle α increases (this implies an increase of the transmission ratio) the efficiency shows values above 95 percent, due to small area of contact and relatively low spin velocity.

### 3.4. Dynamics of the CVT

Once the traction forces at each contact are evaluated, a simplified analysis of the response of the CVT components when subjected to a variation of the torque at the output element is carried out. The input torque T₁ is applied to the
input shaft. For the purpose of this analysis the input torque is considered constant. The principle of angular momentum for the input shaft/discs, neglecting the friction in the bearings, can be written as:

\[ T_1 - nF_{1A}R_A - nF_{1B}R_B = I_1\dot{\omega}_1 \]  \hspace{1cm} (Eq. 16)

\( F_{1A} \) and \( F_{1B} \) are the traction forces in the contacts between the ball and the input toroidal and conical discs respectively, while \( R_A \) and \( R_B \) are the corresponding contact radii on the two discs. \( \omega_1 \) is the angular velocity of the input shaft/discs.

For each ball the dynamic equation of motion is:

\[ F_{1A}r_A + F_{1B}r_B - F_{1C}r_C = I_2\dot{\omega}_2 \]  \hspace{1cm} (Eq. 17)

\( r_A, r_B, \) and \( r_C \) are the distances between the axis of rotation of the ball and the contact points with respectively the input toroidal, conical and output discs. Figure 6 shows one of the balls with the tangential forces acting upon it. These forces are perpendicular to the plane formed by the axes of rotation of the ball and discs. The radii of contact from Eq. 7 are also shown.

Finally the dynamic equation of motion for the output disc is:

\[ nF_{1C}R_C - T_3 = I_3\dot{\omega}_3 \]  \hspace{1cm} (Eq. 18)

where \( T_3 \) is the resistive torque applied to the output shaft. This torque is a function of time, deriving from the system the CVT is part of. In the above equations the radii of the contact points on the discs and ball are also functions of time, through angle \( \alpha \), which changes when the output torque varies. Also the terms \( I_1, I_2, \) and \( I_3 \) are the moments of inertia, about their axis of rotation of the input shaft/discs, ball and output shaft/disc respectively.

The tangential forces in equations (Eqs 16 -18) depend on the sliding velocity in each contact through the shear stresses. Using, for example equation (Eq 1) a relationship between the tangential force in the contacts and the sliding speed can be written, and subsequently employed in (Eq 18) to evaluate the dynamic response of the CVT. This however proved a difficult task as it leads to a system of non-linear differential equations, with variable coefficients, in the unknowns \( \omega_1, \omega_2 \) and \( \omega_3 \) that proved difficult to solve.

An alternative solution was found by observing that the traction drives are designed to work on the linear region of the traction curve of the lubricant. This means that the tangential force generated by the shear of the lubricant can be written as:
Depending on the contact the index \( i \) can take the values of 2 or 3.

The coefficient \( m \) is the slope of the linear region of the traction curve. This is extracted from traction experiments for a certain lubricant and given conditions of load and temperature. Introducing (Eq 20) into (Eqs 16 - 18) and grouping the like terms produces a non-linear system differential equations of first order, with the unknowns \( \omega_1 \), \( \omega_2 \) and \( \omega_3 \).

It is subsequently assumed that the resistive torque at the output disc varies between minimum and maximum values in a known time interval. This variation can be of any kind, but for exemplification a linear function was chosen. Further the time period of variation of the torque was divided in smaller intervals \( \Delta t \), over which the variation of the angular velocities can be considered also as linear. In this way the time derivative of the angular velocity in equations (Eqs 16 - 18) becomes the average angular acceleration over \( \Delta t \). Only equation (Eq 18) is re-written for exemplification.

\[
\left( \frac{\omega_2^i - \omega_2^{i-1}}{\omega_2^i - \omega_2^{i-1}} \right) \left( \frac{\omega_3^i - \omega_3^{i-1}}{\omega_3^i - \omega_3^{i-1}} \right) N_c R_c - \frac{T_2^i - T_2^{i-1}}{4m} = \frac{I_3}{2m} \left( \frac{\omega_2^i - \omega_2^{i-1}}{t^i - t^{i-1}} \right)
\]

(Eq. 20)

In this way the system of differential equations was transformed into an algebraic system of non-linear equations, in the unknowns \( \omega_1 \), \( \omega_2 \) and \( \omega_3 \).

4. RESULTS

Equations like (Eq 20) were written for the input discs/shaft as well as balls, and the system solved for the torque, initial angular velocity of the input shaft, and angles \( \beta \) and \( \gamma \) at a given variation of the resistive torque. A time step of one millisecond was chosen. Another important parameter in the analysis is the slope of the linear part of the traction curve, \( m \). Zhang and co-authors (3) use a value of \( m = 5 \), without stating the reason, so a first value chosen for these calculations is \( m = 5 \).

The transmission ratio \( i = \omega_2 / \omega_3 \), for both steady state and dynamic conditions are shown in Figure 7. The ratio of the actual to maximum output torque is also shown. As seen taking into account the inertia of the elements and the shearing of the lubricant make the transmission ratio change slower for the same variation of the resistive torque. Ideally the two transmission ratio curves should meet but in the time interval of the variation of the torque, angle \( \alpha \) exceeds its maximum value for current dimensions.

![Figure 7](https://example.com/figure7.png)

**Figure 7.** Comparison between steady state and dynamic transmission ratios, \( m = 5 \)

Results of tests presented in [11] showed that for temperatures up to 100°C the slope of the traction curves is 12, for pressure of about 1.5 GPa and a wide range of traction fluids tested.
Figure 8. Comparison between steady state and dynamic transmission ratios, m = 12

For m = 12 the variation of the dynamic transmission ratio is closer to the steady state curve but still lags behind it for the same value of the torque, as seen in Figure 8.

Figure 9. Comparison between steady state and dynamic transmission ratios, m = 100

Ideally for an infinite value of m (in practice calculations were carried out for m = 100) the same calculations show identical steady state and dynamic curves. This should be expected, as in this case the fluid responds instantaneously to a shear strain applied. This is shown in Figure 9. It should be mentioned that realistic simulations must use a value of this parameter obtained from traction measurements carried out for a given oil, in the same conditions of speed, pressure and temperature as used in the simulation. In this CVT the lubricant film thickness and consequently the shear rate, as well as the pressure vary with parameter α (that is with the transmission ratio) so it was not possible to have a single value for m. Instead, by going through a range of values, from rather modest to very large, the trend of the effect this parameter has upon the dynamic response of the device was possible to be evaluated.

5. CONCLUSIONS

This paper shows an analysis of the relative kinematics on the contacts of a toroidal-type CVT. This CVT is able to automatically change the transmission ratio as a function of the resistive torque at the output element. The total slip in the direction of rolling and the spin, on the contact areas are determined and are then used to calculate the power losses of the device. The dynamic response of the device to a rapid change of the output torque is evaluated by considering the shearing of the lubricant film and the inertia of the elements. It was found that the dynamic transmission ratio is slower to respond to the change of the output torque than the steady state transmission ratio. An important parameter which determines the response of the device is the slope of the traction curve of the lubricant.

Nomenclature

Fₙ – force normal to the surfaces
$F_t$ – force normal to the surfaces
$I$ – moment of inertia of balls or discs
$m$ – slope of linear region of lubricant’s traction curve
$r_0$ – major radius of torus
$r_{A,B,C}$ – distances between contact points and axis of rotation of ball
$R_{A,B,C}$ – distances between contact points and axis of rotation of discs
$T_{1,3}$ – torque at input, output shafts
$U_{1,2}$ – surface velocities of bodies 1 or 2
$U$ – rolling velocity $(U_1 + U_2)/2$
$\Delta U$ – sliding velocity $(U_1 - U_2)$
$U_{x,y}$ – components of surface velocities in directions $x$ or $y$
$x, y, z$ – coordinates in a system attached to contact ellipse
$\alpha$ – angle of the radius of the contact point on the toroidal disc
$\beta$ – angle of the radius of the contact point on the input, conical disc
$\gamma$ – angle of the radius of the contact point on the output disc
$\lambda$ – angle between axes of rotation of balls and discs
$\mu$ – traction coefficient
$\omega_{1,2,3}$ – angular velocity of balls or discs
$\xi$ – slide/roll ratio

Indexes A, B and C refer to contacts between balls and input toroidal, input conical and output discs respectively

6. REFERENCES