Engineering with Computers (2000) 16: 117–130

Engineering with Computers

Using the Skeleton Method to Define a Preliminary Geometrical Model for Three-Dimensional Speed Reducers

J. C. Fauroux[1], M. Sartor[2] and M. Paredes[2]

[1]Laboratoire de Recherche et Applications en Mécanique Avancée (LaRAMA), Institut Français de Mécanique Avancée (IFMA), Aubière, France; [2]Laboratoire de Génie Mécanique de Toulouse (LGMT), Institut National des Sciences Appliquées (INSA), Toulouse, France

Abstract. *This paper deals with the first design phase of three-dimensional speed reducers. Concepts to help designers during the entry-level model definition are presented and, more particularly, those related to the task of geometrical synthesis. From these principles, a specific CAD tool has been built and is also described. A rough model, called a skeleton, is introduced to represent each conventional reducing stage category, thus enabling the automatic formation of entire geometrical models of speed reducers. Shaft orientations and positions are calculated from products of transformation matrices, in the same way as for spatial kinematic chain closure conventional problems. Optimisation techniques are used to obtain a preliminary dimensioning of the structure. The numerical processing is achieved progressively in three steps in order to improve the final convergence and, if necessary, to enlighten further the designer on failure origins within specification data. Three examples are given to illustrate both the creation of the geometrical model and the way results are obtained.*

Keywords. 3D speed reducer design; Closed-loop chain; Geometrical model optimisation; Mechanism synthesis; Skeleton

1. Introduction

Most common power transmission lines encountered in machines are one-degree-of-freedom mechanisms built as a series of elementary stages, such as gear, belt or chain stages, connected to each other by intermediate shafts. Their main function is to transmit a rotative movement from an input shaft to a remote output shaft whose rotation speed is often quite small in comparison with that of the input. These mechanical devices can be designed without difficulty when input and output shafts are parallel or on the same planar surface. Attention will be focused here on cases where the output shaft must take a complex spatial position with regard to the input one, such devices being called three-dimensional speed reducers (*3D.S.R.*). At least two stages with non-parallel shafts, such as bevel or worm gearing, must generally be used to obtain the required output position. The design of these 3D mechanisms leads consistently to synthesis problems for the solution of which methods and assistance tools are welcome.

The purpose of (*3D.S.R.*) preliminary design is to establish an entry-level model of the device, which specifies the composition of the mechanism and gives starting values to positions and dimensions for all main components such as shaft axes and gearing wheels. This model, often represented by a structural scheme with some dimensional data, is built from the kinematic point of view, first considering the geometrical aspects of the problem. It constitutes the specification sheet of the next design phase, which will take into account many other aspects in a more detailed study of each device part.

The preliminary design process mentioned above may be considered as the three phase process presented in Fig. 1.

At the starting point, the main available specifications are the expected spatial positions of the input and output shafts, the required speed ratio and some other global characteristics. The first task necessarily consists in defining the nature of all the elementary reducing stages which will be serially connected to form the transmission line (Phase 1). A wide range of literature concerning the structural synthesis of mechanisms is available [1–6]. Gener-

Correspondence and offprint requests to: Professor J.-C. Fauroux, Institut Français de Mécanique Avancée (IFMA), Campus de Clermont-Ferrand, Les Cézeaux, BP 265, 63175 Aubière, France. Email: fauroux@ifma.fr, Web: www.ifma.fr

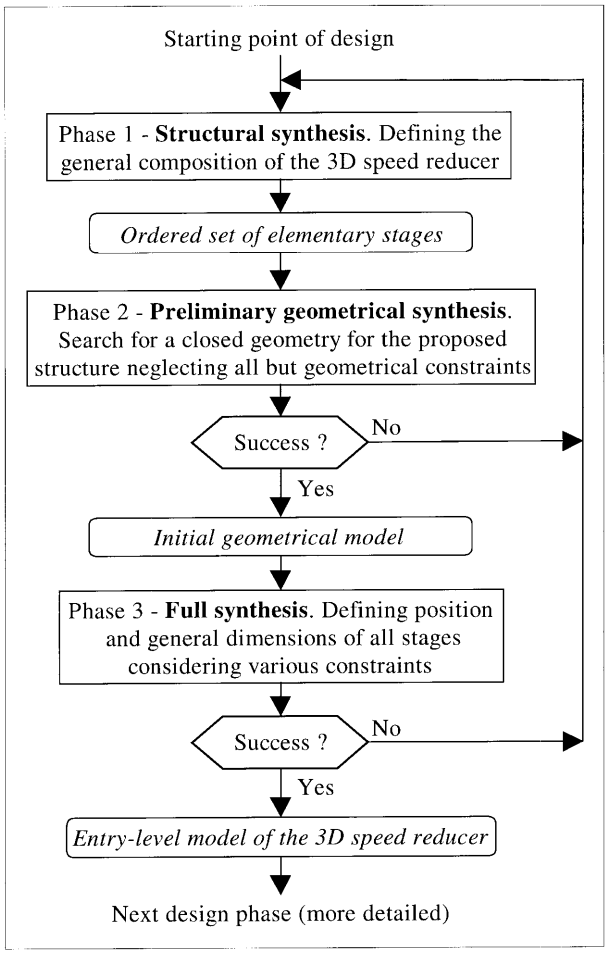


Fig. 1. Preliminary design process of 3D speed reducers.

ally based on graph representation, methods have been proposed essentially for the classification and enumeration of mechanisms according to kinematic structures. Buchsbaum and Freudenstein [2] and Ravisankar and Mruthyunjaya [5] applied such methods to gear transmission and differential drives. Unfortunately, these methods do not apply to practical cases where a *3D.S.R.*. architecture fulfilling given specifications is to be found. Usually, the task of structural synthesis is rather carried out using the designer's skills and knowledge. In any case, the final design choice takes the form of an ordered set of elementary stages, but without the guarantee that this choice will be appropriate nor optimal.

Afterwards, the definition of an entry-level model for a *3D.S.R.* consists of:

- searching for the main dimensions of structure components, i.e. searching for a closed spatial chain running from the input to the output position and respecting the structural characteristics of each selected elementary stage;
- under different constraints such as:
  - obtaining the required speed ratio;
  - avoiding part interference;
  - using realistic wheel dimensions (taking into consideration the main conventional design criteria: contact stress, fatigue life, proportion ratios …).

The whole problem is relatively cumbersome, and there is little chance of succeeding in attempts to solve it directly, especially as the existence of a closed geometry is uncertain. A more suitable way to solve this problem is to divide the resolution into two phases.

Thus, the design process first continues with a preliminary task of geometrical synthesis (Phase 2), intended merely to find a closed geometric chain. If this search fails, the previous choice (i.e. an ordered set of stages) has to be reconsidered. If it succeeds, an initial geometrical model is obtained, and the process may continue using this model as a starting point. The last step is a full synthesis dealing with the complete problem (Phase 3).

In this paper, attention will only be focused on the task of preliminary geometrical synthesis, and particularly on Phase 2. A software tool which has been developed in order to help the designer build *3D.S.R.* initial geometrical models is presented. The main difficulty has been to find a formulation which is general enough to allow the automatic processing of all power transmission lines, with any number and any nature of stages. This obstacle has been overcome by transforming the problem into a conventional closed spatial chain synthesis. This transformation is based on the use of filar representations of elementary stages, called stage *skeletons*, and which contain only the data necessary to build the model.

Many studies are available concerning closed spatial chain analysis or synthesis. Algebraic methods have been applied to various space bar-linkages made from rotational, spherical or prismatic joints, such as the RGGR mechanism studies presented by Shigley and Uicker [7] or by Söylemez and Freudenstein [8]. Analytical approaches, characterised by the search for mechanism closed positions from matrix loop equations, have been initiated by Denavit and Hartenberg [9], then widely used, especially in the robotics field. These latter appear to be perfectly suited to our problem, because they offer the means of constructing a model by standard unit association. The relevant literature provides several methods of formulating spatial loop equations [10]. In our case of a simple closed-loop chain,

Denavit and Hartenberg's notations [11] have seemed the most appropriate because of their simplicity.

Using transformation matrix association, it thus became possible to build automatically a geometrical model for any *3D.S.R.* Such a model being usually redundant, nonlinear optimisation techniques have been chosen to determine the geometrical parameters that minimise the device overall dimensions.

2. Classification of the Reducer Elementary Stages

Let us begin with the presentation of the most common elementary stages which can be encountered in power transmission lines. To prepare the following developments, stages are classified according to their general geometrical structure. Figures 2 and 3 emphasise two main stage categories:

- *Stages with non-concurrent shafts*. The most common values of the angle defined in Fig. 2 from $\vec{k}$ to $\vec{u}$ and measured in the $(\vec{i}, \vec{j}, \vec{k})$ direct ordinate system, are 0°, 90°, 180° and 270°. This category includes stages with parallel shafts either on opposite sides ($\varphi = 0°$) or on the same side ($\varphi = 180°$), the associated elementary mechanisms being: external/external and external/internal cylindrical gearing; belt-pulley transmission; chain transmission. Some stages with non-parallel, non-concurrent shafts, such as worm gearing, crossed-axis helical gearing and hypoïd gearing, are also part of this class, with generally $\varphi = 90°$ or $\varphi = 270°$.
- *Stages with concurrent shafts*. When the angle φ, defined in Fig. 3, equals 0° or 180°, input and output shafts are coaxial. Some epicyclic trains and specific devices such as the harmonic drive joint are concerned. Other possible devices are bevel gearing (the two usual configurations shown in Fig. 3 square with $\varphi = 90°$, but provide opposite senses of output rotation) and Cardan joint (under certain conditions).

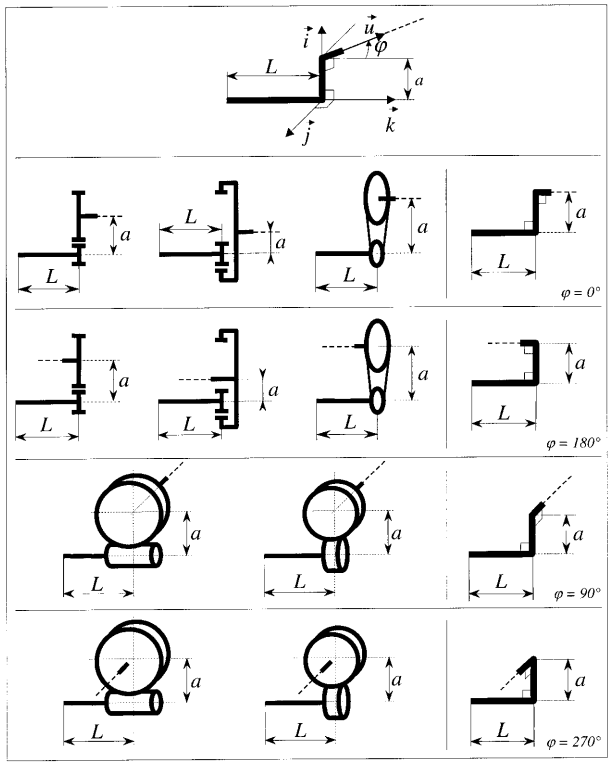


Fig. 2. Stages with non-concurrent shafts.

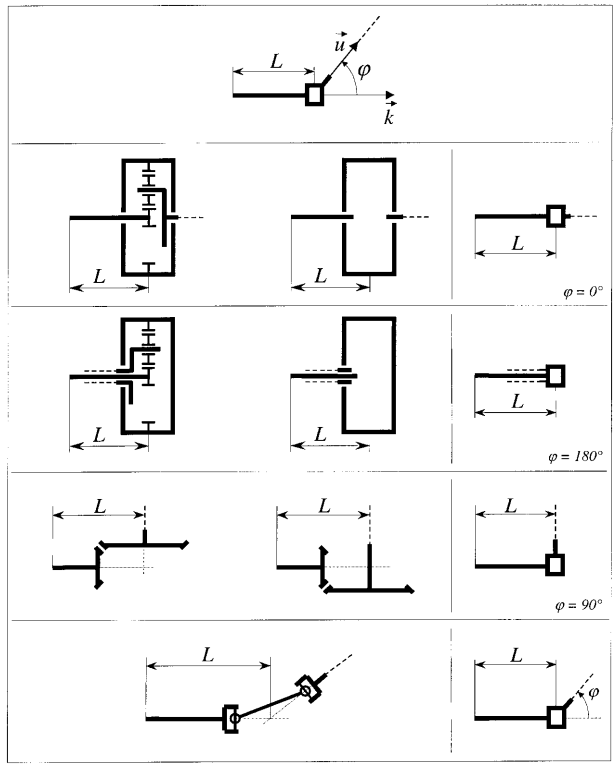


Fig. 3. Stages with concurrent shafts.

The diagrams on the right of Figs 2 and 3 are called the 'skeletons' of the stages. These linear structures will be used later to construct the geometrical model of 3D speed reducers. Their main role is to represent the stage spatial architecture. Dimensions used to mark these architectures are only general:

- the input shaft length *L*;
- the distance *a* between input and output axis;

- the angle φ between both axes.

There is no length associated with the output shafts in order to decrease the parameter number. At this wide abstraction level, this length will be included in the input shaft length of the next stage.

3. Problem Setting

The task of preliminary geometrical synthesis requires only the following initial specifications:

- the nature of the successive stages which will constitute the power transmission line. It is assumed this preliminary choice will not necessarily lead to a feasible solution;
- a coordinate system $R_0 = (O_0, \vec{X}_0, \vec{Y}_0, \vec{Z}_0)$ which defines the input shaft position. Actually, only O_0, the shaft starting point, and $\vec{Z}_0$, the shaft orientation, are related to the practical problem. $\vec{X}_0$ and $\vec{Y}_0$ are just required for the calculations;
- a coordinate system $R_s = (O_s, \vec{X}_s, \vec{Y}_s, \vec{Z}_s)$ which defines the output shaft position: O_s is the shaft terminal point and $\vec{Z}_s$ the axis orientation;
- the volume within which all the reducer components must be contained, the most simple case being a parallelepiped defined by two opposite points P_m and P_M, relatively to the global coordinate system $R_g = (O_g, \vec{X}_g, \vec{Y}_g, \vec{Z}_g)$;
- the objective function the geometrical structure has to optimise. Hereafter it is considered that the overall dimensions of the reducer must be minimal.

Geometrical specifications are illustrated in Fig. 4.

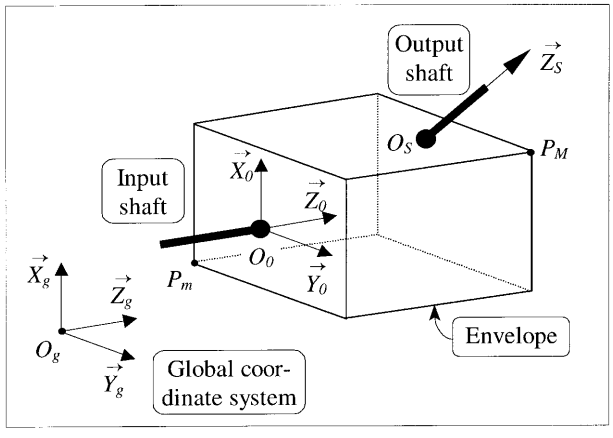


Fig. 4. Geometrical specifications.

4. Towards a Closed Chain Synthesis Problem

To illustrate our purpose in this section, the example of a three stage *3D.S.R.* made of two successive bevel gearing stages, followed by a cylindrical gearing, as shown in Fig. 5(a), is considered.

The first level dimensional synthesis problem to be solved consists in determining all the intermediate shaft spatial positions and all the stage general dimensions (i.e. those presented in stage skeletons). The nature of this problem is not the one traditionally encountered in the field of kinematics, as the motion transformation law is not the aim of our study. Each stage is homokinetic, so the overall linkage is also homokinetic.

All input, intermediate and output shafts are linked by a rotational joint to the housing, so their axes are fixed in space and there is no other motion than the rotations about these axes. The problem

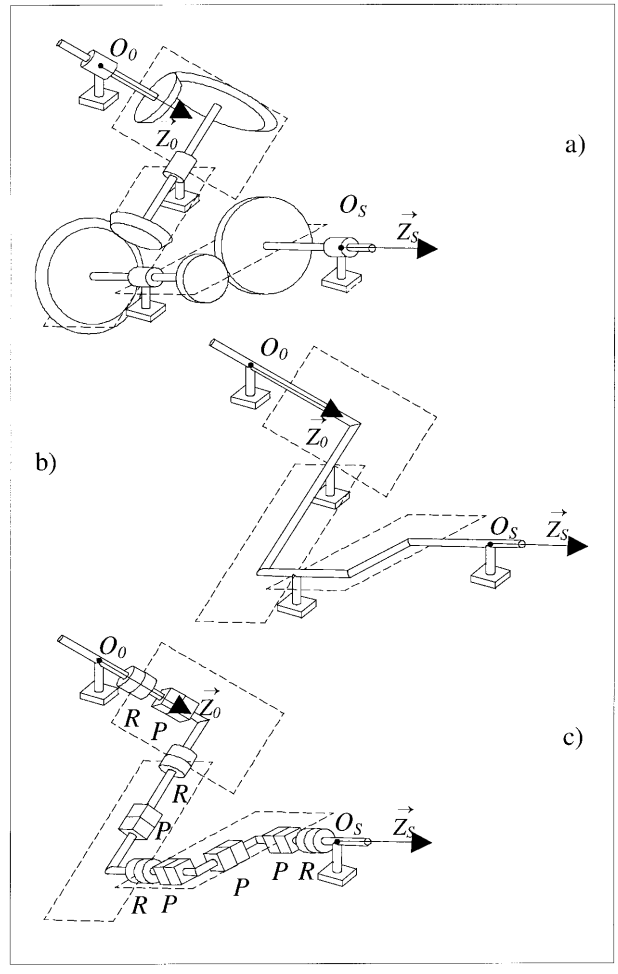


Fig. 5. Problem transformations.

under study is not influenced by these rotations, which may be ignored. In this case, all shafts may be considered as rigid bodies completely linked to the housing. Moreover, each stage may be replaced by its associated skeleton. This constitutes a first transformation of the problem leading to the overall skeleton of the *3D.S.R.* illustrated in Fig. 5(b).

In order to search for the best positions of all intermediate shafts and the best general dimensions of all stages, some structural constraints may be artificially relaxed. All intermediate shafts are liberated from the housing, and joints are added in such a way that the search becomes possible:

- a prismatic joint *P* is introduced wherever a variable length appears in the skeleton. It gives the possibility of simulating modifications in the reducer geometry by extension or compression of the associated dimension. A final prismatic joint is added onto the output axis to model the output shaft length which has not been represented within any stage skeleton;
- a rotational joint *R* is introduced at each extremity of the *3D.S.R.* and between two successive stages. It represents the possibility of modifying the reducer geometry by rotating one stage with respect to the previous one around their common linking shaft.

In such a way, the relative position of the different stage elements is preserved, so the geometrical architecture of each stage is not affected by the transformation. For instance, all angles φ remain unchanged. Figure 5(c) illustrates this second and last transformation of the problem. The initial problem has become the *RPRPRPPPR* study, which belongs to the well-known class of closed chain synthesis problems.

5. The mDH Notations

The system obtained from the transformation presented above is considered as being composed of *NJ* joints and (*NJ*+1) links, links (*O*) and (*NJ*) belonging to the same fixed base. The 'modified Denavit–Hartenberg's' notations (mDH) proposed by Khalil and Kleinfinger [12] are such that:

- Joint J_j connects link $(j-1)$ and link (j);
- A coordinate frame R_j is assigned fixed with respect to link (j);
- The axis of joint J_j is supposed to lie along $\vec{Z}_j$;
- The $\vec{X}_{j-1}$ axis is defined as $\vec{X}_{j-1} = \vec{Z}_{j-1} \wedge \vec{Z}_j$.

Four parameters define the relative positions of two successive co-ordinate systems (see Fig. 6):

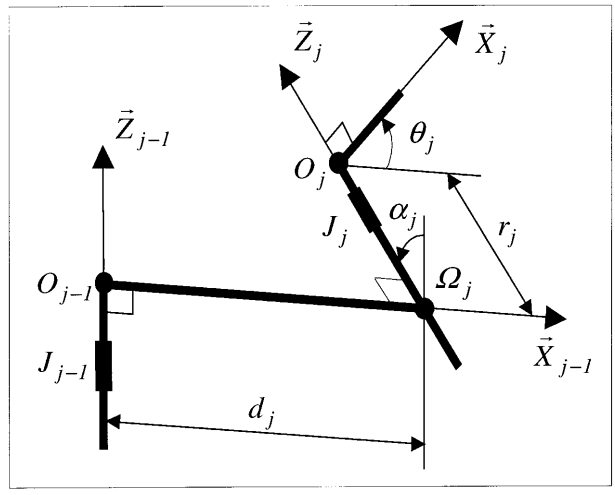


Fig. 6. mDH notations.

- α_j angle between $\vec{Z}_{j-1}$ and $\vec{Z}_j$ about $\vec{X}_{j-1}$;
- d_j distance between O_{j-1} and Ω_j;
- θ_j angle between $\vec{X}_{j-1}$ and $\vec{X}_j$ about $\vec{Z}_j$;
- r_j distance between Ω_j and O_j.

The variable associated with joint J_j, denoted by q_j, is θ_j if J_j is rotational or r_j if J_j is prismatic. Hence,

$$q_j = \theta_j \cdot (1 - \sigma_j) + r_j \cdot \sigma_j \tag{1}$$

where $\sigma_j = 0$ if J_j is rotational and $\sigma_j = 1$ if J_j is prismatic.

The transformation matrix which defines the frame R_j with respect to frame R_{j-1} is equal to

$$[T]_{j-1}^{j} = \begin{bmatrix} \cos\theta_j & -\sin\theta_j & 0 & d_j \\ \cos\alpha_j \sin\theta_j & \cos\alpha_j \cos\theta_j & -\sin\alpha_j & -r_j \sin\alpha_j \\ \sin\alpha_j \sin\theta_j & \sin\alpha_j \cos\theta_j & \cos\alpha_j & r_j \cos\alpha_j \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

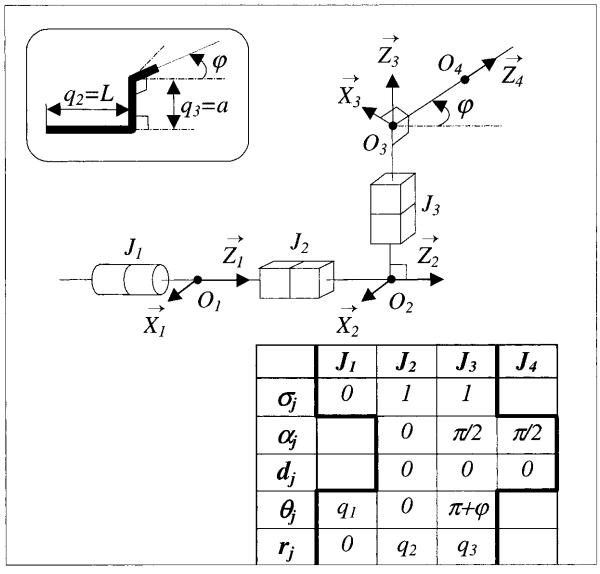
Simple-closed chains are such that

$$[T]_0^{NJ} = [T]_0^1 \,.\, [T]_1^2 \,\ldots\, .[T]_{NJ-1}^{NJ} \tag{3}$$

where the right member represents the successive changes among the links connected by joints, and the left member is the closure transformation between co-ordinate systems R_0 and R_{NJ} defined in the same fixed base.

6. Elementary Geometrical Models According to Stage Categories

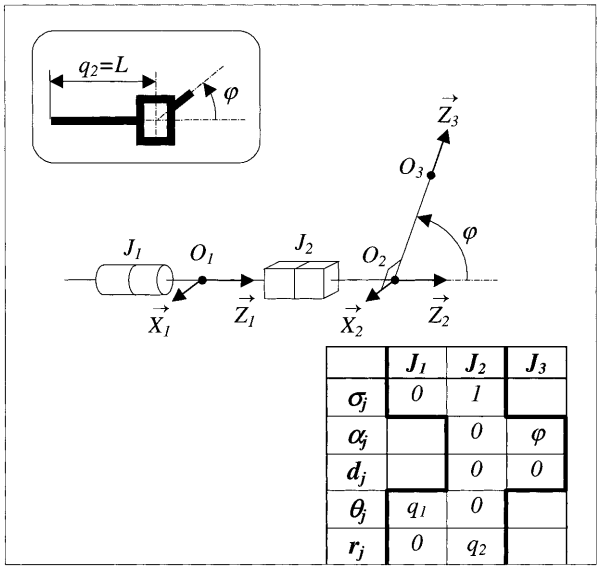
The aim being to build a tool based on standard unit manipulation, a pre-definite geometrical model



	J_1	J_2	J_3	J_4
σ_j	0	1	1	
α_j		0	π/2	π/2
d_j		0	0	0
θ_j	q_1	0	π+φ	
r_j	0	q_2	q_3	

Fig. 7. Geometrical model for non concurrent shaft stages.

is associated with each stage category. This definition is done in accordance with the principles of Section 4. A prismatic joint takes the place of each variable length. A rotational joint is added at the beginning of the chain. Figures 7 and 8 give the two elementary geometrical models, each one being assigned to its respective stage category. The last vector prefigures the first link direction of the next joint. Tables of parameters are drawn up using the mDH notations.



	J_1	J_2	J_3
σ_j	0	1	
α_j		0	φ
d_j		0	0
θ_j	q_1	0	
r_j	0	q_2	

Fig. 8. Geometrical model for concurrent shaft stages.

7. Automatic Construction of the Geometrical Model Associated with a *3DSR* Problem

It is easy to deal with the case of any reducer made of serially connected speed reducing stages. The associated geometrical model may be obtained in this way:

- The first elementary stage table is set and completed with parameters $\alpha_1 = d_1 = 0$ considering that the first joint J_1 is a rotational one.
- The next stage table is added like a puzzle piece at the end of the previous table, and so on up to the last stage.
- Two columns are added at the end in order to materialise a prismatic joint representing the possible length variation of the output shaft, and a rotational joint representing the possible rotation of the structure about the output axis.

Let us consider the example of a three stages *3D.S.R.* whose constitution is:

- two successive worm gearing stages with $\varphi = 90°$;
- an external/external cylindrical gearing stage with shafts on opposite sides ($\varphi = 0°$).

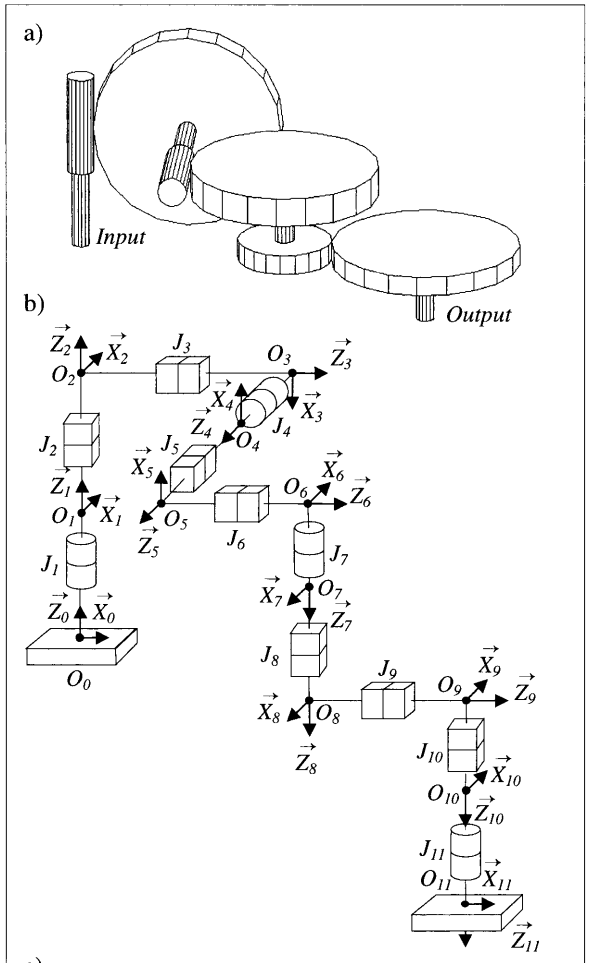
Figure 9(a) shows the diagram of the reducer given in an arbitrary starting position. Figure 9(b) illustrates the whole geometrical model which can be automatically generated, and Fig. 9(c) is the associated parameter table.

8. Expression of the Optimisation Problem

8.1. Problem Setting

Let us consider a *3D.S.R.* whose model is composed of *NJ* prismatic or rotational joints. These joints are characterised by variables q_j ($j = 1 \ldots NJ$) which must take values such as the output shaft satisfies the imposed position and orientation. This is the problem, often carried out in robotics [10], of solving the inverse geometrical model of a manipulating arm. Two main difficulties have been encountered.

First, the model is usually redundant: the variable number is higher or equal to the closure equation number for reducers of more than two stages. As a consequence, the number of solutions will often be infinite. To choose from among these solutions, it will be necessary to search for the one which optimises a certain criterion. Secondly, constraints



	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8	J_9	J_{10}	J_{11}
σ_j	0	1	1	0	1	1	0	1	1	1	0
α_j	0	0	$\pi/2$	$\pi/2$	0	$\pi/2$	$\pi/2$	0	$\pi/2$	$\pi/2$	0
d_j	0	0	0	0	0	0	0	0	0	0	0
θ_j	q_1	0	$3\pi/2$	q_4	0	$3\pi/2$	q_7	0	π	0	q_{11}
r_j	0	q_2	q_3	0	q_5	q_6	0	q_8	q_9	q_{10}	0

Fig. 9. *3D.S.R.* example and its whole geometrical model.

concerning the overall dimension must be respected: the entire skeleton must lay inside the envelope formerly specified. As the skeleton is a succession of segments connecting points O_j, all these points must be included within the envelope (this is true only for convex envelopes). Consequently, there are six inequality constraints for each point O_j, because each co-ordinate must be inside an interval.

So an optimising method has to be found to solve the inverse geometrical problem. The solution will have to minimise a criterion, and to respect constraints concerning the overall dimensions.

8.2. Choice of the Optimisation Criterion

The criterion must be geometrical because the problem is treated from a geometrical point of view. For instance, the overall dimensions of the whole reducer may be minimised. The following function F is proposed:

$$F(q_1, \cdots, q_{NJ}) = \sum_{j=1}^{NJ} \sigma_j \cdot q_j \tag{4}$$

This function is the sum of the lengths of the prismatic joints. Minimising F is equivalent to imagining a rubber band that would run along all the prismatic joints, and would try to compress them. This prevents the skeleton from deviating from a hypothetical axis joining O_0 to O_s.

8.3. Equality Constraints

Equality constraints are used to ensure that the output shaft respects the required orientation and position. According to the presence of NJ joints in the geometrical model, the closure equation takes the following form.

$$[T]_g^0 \,.\, [T]_0^1 \,.\, [T]_1^2 . \cdots [T]_{NJ-2}^{NJ-1} \,.\, [T]_{NJ-1}^{NJ} = [T]_g^s \tag{5}$$

where $[T]_g^0$ is the transformation matrix from R_g (global co-ordinate system) to R_0 (input co-ordinate system);
$[T]_{j-1}^{j}$ is the transformation matrix defined by Eq. (2) for each joint j;
$[T]_g^s$ is the transformation matrix from R_g to R_s (required output co-ordinate system).

The matrix equality of Eq. (5) has a geometrical signification, considering each column separately as follows.

$$\text{1st column} \rightarrow \vec{X}_{NJ} = \vec{X}_s \tag{6}$$

$$\text{2nd column} \rightarrow \vec{Y}_{NJ} = \vec{Y}_s \tag{7}$$

$$\text{3rd column} \rightarrow \vec{Z}_{NJ} = \vec{Z}_s \tag{8}$$

$$\text{4th column} \rightarrow \vec{O}_{NJ} = \vec{O}_s \tag{9}$$

This leads to 12 scalar relationships which however are not independent from each other. To decrease as far as possible the number of equality constraints to be processed, the closure chain expression will be reduced to only six independent scalar relationships by considering the last but one link system, instead of the last one. Only the following constraints will be used:

- Three orientation constraints:

$$\vec{Z}_{NJ-1} = \vec{Z}_s \rightarrow \begin{cases} Ge_1 = x_{Z_{NJ-1}} - x_{Z_s} = 0 \\ Ge_2 = y_{Z_{NJ-1}} - y_{Z_s} = 0 \\ Ge_3 = z_{Z_{NJ-1}} - z_{Z_s} = 0 \end{cases} \quad (10)$$

- Three position constraints:

$$O_{NJ-1} = O_s \rightarrow \begin{cases} Ge_4 = 1 - x_{O_{NJ-1}}/x_{O_s} = 0 \\ Ge_5 = 1 - y_{O_{NJ-1}}/y_{O_s} = 0 \\ Ge_6 = 1 - z_{O_{NJ-1}}/z_{O_s} = 0 \end{cases} \quad (11)$$

where the different components are considered in the co-ordinate system R_g. Constraints are expressed in a non-dimensional way for best optimisation efficiency.

The simplification of the closure expression is made possible thanks to the particular properties of the last rotational joint J_{NJ}:

$$\alpha_{NJ} = d_{NJ} = r_{NJ} = 0 \quad (12)$$

The associated transformation matrix is

$$[T]^{NJ}_{NJ-1} = \begin{bmatrix} \cos q_{NJ} & -\sin q_{NJ} & 0 & 0 \\ \sin q_{NJ} & \cos q_{NJ} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

from which it can be concluded that

$$\vec{Z}_{NJ-1} = Z_{NJ} \text{ and } O_{NJ-1} = O_{NJ} \quad (14)$$

Therefore, $\vec{Z}_{NJ-1}$ and O_{NJ-1} can replace $\vec{Z}_{NJ}$ and O_{NJ} in the initial closure Eqs (8) and (9). Moreover, there is no need to impose orientation conditions on vectors $\vec{X}_{NJ-1}$ and $\vec{Y}_{NJ-1}$, because it is certain that a q_{NJ} value (rotation about $\vec{Z}_{NJ-1} = \vec{Z}_{NJ} = \vec{Z}_S$) will always be available to transform these vectors into the required $\vec{X}_S$ and $\vec{Y}_S$ vectors. The only alteration this simplification introduces is that the last rotational joint variable q_{NJ} does not appear in the closure formulation any more, and its calculus is skipped. This can be considered as an advantage, since it decreases the size of the problem. It remains possible to deduce, if necessary, q_{NJ} value from the other variables after the problem has been solved.

In practice, $\vec{Z}_{NJ-1}$ and O_{NJ-1} components are calculated for current values of q_j ($j = 1 \ldots (NJ - 1)$) by processing numerically the product of the successive basic change matrices:

$$[T]^{NJ-1}_g = [T]^0_g \cdot \prod_{j=1}^{NJ-1} [T]^j_{j-1} \quad (15)$$

8.4. Inequality Constraints

Inequality constraints (16) force points O_j to stay within the envelope:

$$\begin{cases} x_m \le x_{O_j} \le x_M \\ y_m \le y_{O_j} \le y_M \\ z_m \le z_{O_j} \le z_M \end{cases} \Leftrightarrow \begin{cases} 1 - x_{O_j}/x_m \le 0 \text{ and } x_{O_j}/x_M - 1 \le 0 \\ 1 - y_{O_j}/y_m \le 0 \text{ and } y_{O_j}/y_M - 1 \le 0 \\ 1 - z_{O_j}/z_m \le 0 \text{ and } z_{O_j}/z_M - 1 \le 0 \end{cases} \quad (16)$$

where:

- $[x_m, x_M]$, $[y_m, y_M]$ and $[z_m, z_M]$ are the intervals which define the envelope in the coordinate system R_g. It must be carefully chosen for having non null values of x_m, x_M, y_m, y_M, z_m, z_M, thus avoiding divisions by zero in Eq. (16);
- x_{O_j}, y_{O_j}, z_{O_j} are the co-ordinates of point O_j in R_g. They are located on the fourth column of the intermediate change matrix $[T]^j_g = [T]^0_g \,.\, [T]^1_0 \ldots [T]^j_{j-1}$.

9. Solving the Optimisation Problem

9.1 Algorithm Refinements

The problem had to be enhanced to obtain valid initial conditions. In fact, since arbitrary initial values are given to variables q_j at the outset, equality constraints are false, hence a low convergence. The closure equation of the reducer must be satisfied above all. So, to find good initial values, two preliminary optimisation steps are achieved. Figure 10 sums up the process.

It can be easily shown that $\vec{Z}_{NJ-1}$ components are only dependent on the angular variables (not on those of length). For instance, in the example of Section 7, the literal development of orientation constraints yields

$$\begin{cases} -\sin q_1 \sin q_4 = x_{Z_S} \\ \cos q_1 \sin q_4 = y_{Z_S} \\ \cos q_4 = z_{Z_S} \end{cases} \quad (17)$$

where q_1 and q_4 are two angular variables.

So, a preliminary optimisation may be done by making only $(q_{a1}, \ldots q_{ak})$ angular variables vary and minimising a penalty function which represents the violation of the three orientation constraints related to Eq. (10):

$$F_{pa}(q_{a1}, \ldots q_{ak}) = \sum_{i=1}^{3} (Ge_i(q_1, \ldots q_{NJ-1}))^2 \quad (18)$$

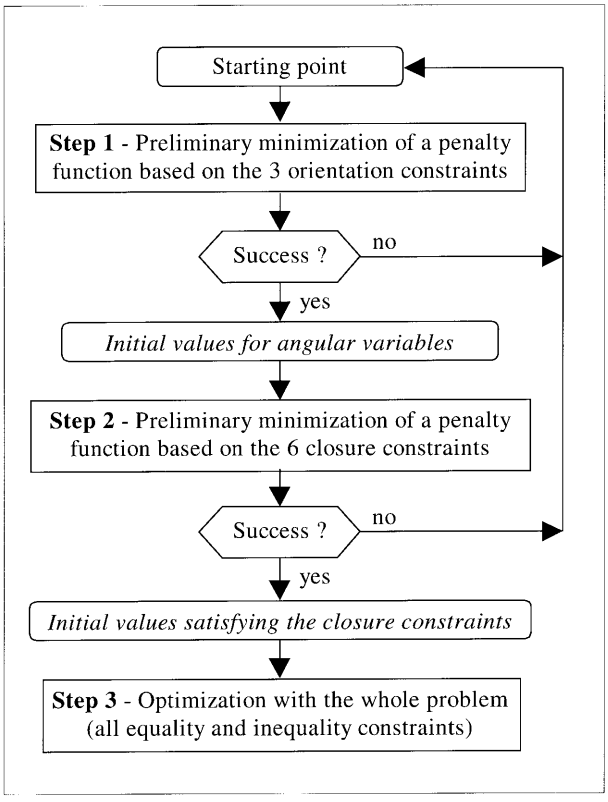


Fig. 10. The three-step solving process.

From this improved starting point, the closed position is calculated minimising the lack of closure. This approach is not far from that presented by Hall, Root and Sandgren [14]. The entire set of variables and all six closure constraints are considered now as

$$F_p(q_1, \cdots q_{NJ-1}) = \sum_{i=1}^{6} (Ge_i(q_1, \cdots q_{NJ-1}))^2 \qquad (19)$$

Then, optimising the whole problem can be processed. Starting from a configuration where the skeleton is already closed, this last step progressively alters the dimensions of this skeleton to inscribe it within the envelope, and to compact it as much as possible. All the constraints can be initially violated without compromising final convergence.

9.2. Computer Implementation

From the previously described concepts, software was implemented in plain ANSI C. The algorithm is based on the *B.F.G.S.* method (Broydon–Fletcher–Goldfarb–Shanno) from the Fortran optimisation library *D.O.T.* [13]. A Graphical User Interface (GUI) was added for convenience, and implemented in Tcl/Tk. From a practical point of view, the notion of steps is transparent for the user: he just has to fill the blank fields in the user interface, select other advanced options if necessary, and start the optimisation process (Fig. 11). Data to be input are the nature of the stages, initial dimensions of the skeleton, optimisation method and some other technical parameters.

Then, results can be obtained in two forms:

- an exhaustive file of numerical data (vector of solutions, values of constraints and optimum function, number of iterations . . .);
- a 3D model of the skeleton with an associated 3D animation file showing the skeleton retraction during its optimisation. User can replay each of the previously calculated iteration and, moreover, can rotate and zoom the model at a given iteration (cf. iteration 1431 in Fig. 11) for best viewing particular configurations.

10. First Example

The example already introduced in Section 7 is resumed below with the following complementary specifications:

- *Envelope* $x_m = y_m = z_m = 100$ mm
  $x_m = y_m = z_m = 1100$ mm
- *Input shaft* $O_0 = (600,600,100)$
  $\vec{X}_0 = (1, 0, 0)$
  $\vec{Z}_0 = (0, 0, 1)$
- *Output shaft* $O_s = (1100, 1100, 1100)$
  $\vec{Z}_s = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$

Co-ordinates are given in the general co-ordinate system R_g. Upper and lower bounds have to be set for length variables. The upper boundary is taken as equal to the envelope diagonal length. The lower boundary is arbitrarily set to $L_{\min} = 200$ mm hereafter. In a more detailed approach, each lower boundary could be chosen separately according to the nature of the variable (shaft length or distance between shafts) and to technological considerations such as minimum tooth number, minimum gearing modulus, etc.

The initial position, illustrated in Fig. 12(a), is deliberately taken a long away from the optimum to illustrate the algorithm efficiency. Initial values of q_j are provided in Table 1. The first step predetermines new values for angular variables:

$$\{q_1,q_4,q_7\} = \{2.36rad, 5.33rad, 1.57rad\} \qquad (20)$$

Note that q_7 remains unchanged. This is quite

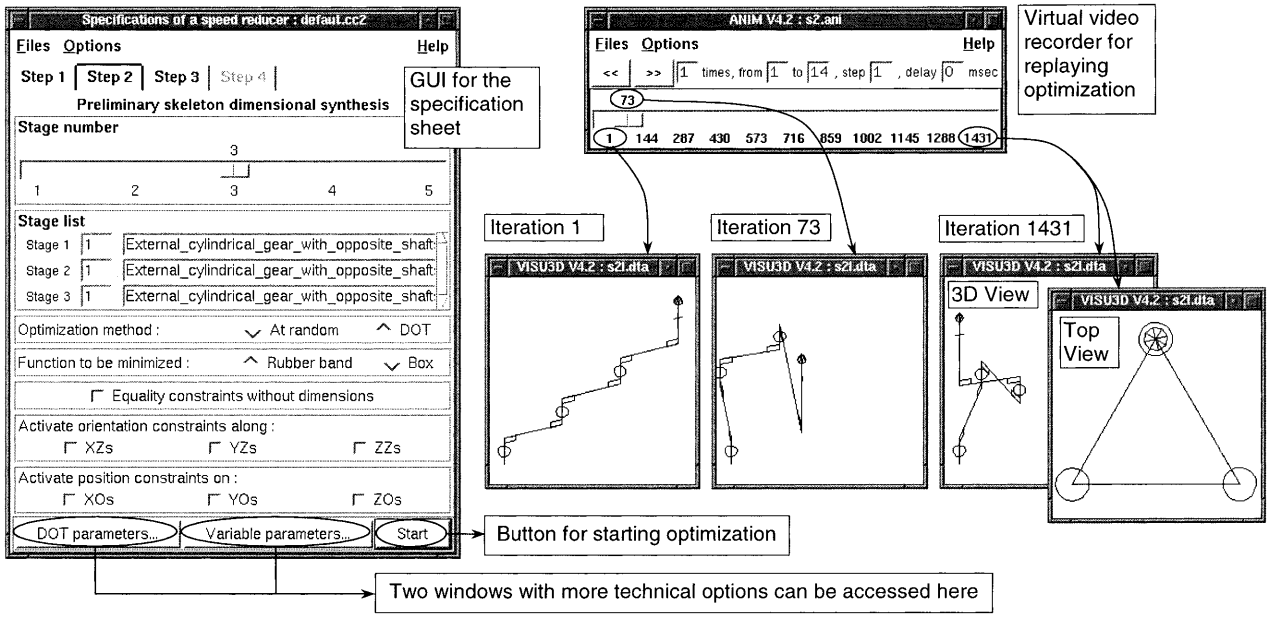


Fig. 11. The graphical user interface provided.

normal, since the last cylindrical gearing stage has parallel shafts and so does not introduce any variation of the terminal shaft direction. Equation (17) confirms this property, showing that q_7 does not appear in $\vec{Z}_{NJ-1}$ component expression.

The second step is intended to find values that verify the closure equation. Its duration is about 1 second with a Pentium Pro 233 Mhz. The penalty objective function is less than 10^{-13}, giving proof that the preliminary closure task is working well. The skeleton has the configuration shown on Fig. 12(b). The objective function is then $F = 2081$ mm.

The third step can be run now to compress the skeleton, and to obtain, in less than 1 second, what is represented in Fig. 12(c). The objective function has decreased to $F = 1930$ mm.

As for any numerical calculation method running from an arbitrary point, the influence of the starting point has to be analysed. By making the starting point vary, only one other solution can be found for q_1 and q_4 by the first preliminary step:

$$\{q_1, q_4, q_7\} = \{5.50rad,\ 0.95rad,\ 1.57rad\} \qquad (21)$$

This second solution in fact represents a dual geometrical disposition of the first one, as illustrated in Fig. 13. Note that it can be easily verified that q_1 and q_4 values given in Eqs (20) and (21) fit the whole solutions of Eq. (22) following from Eq. (17):

$$\begin{cases} -\sin q_1 \sin q_4 = 1/\sqrt{3} \\ \cos q_1 \sin q_4 = 1/\sqrt{3} \\ \cos q_4 = 1/\sqrt{3} \end{cases} \qquad (22)$$

The geometry obtained from this point is shown in Fig. 14, leading to $F = 2344$ mm. This second configuration is a local optimum, when the previous one is the global optimum.

11. Second Example

The three stage *3D.S.R.* defined below is considered as:

- a worm gearing stage with $\varphi = 90°$;
- an external/external cylindrical gearing stage with shafts on opposite sides ($\varphi = 0°$);
- a worm gearing stage with $\varphi = 90°$.

Figure 15 illustrates the general architecture of the reducer. Note that it represents a variant of the previous example, the two last stages having been exchanged. The geometrical specifications are taken as identical to those of Section 10.

Angular variable values satisfying the three orientation equality constraints are found without difficulty. When the starting point varies, q_1 takes either the value $3\pi/2$ or the value $5\pi/2$, when q_4

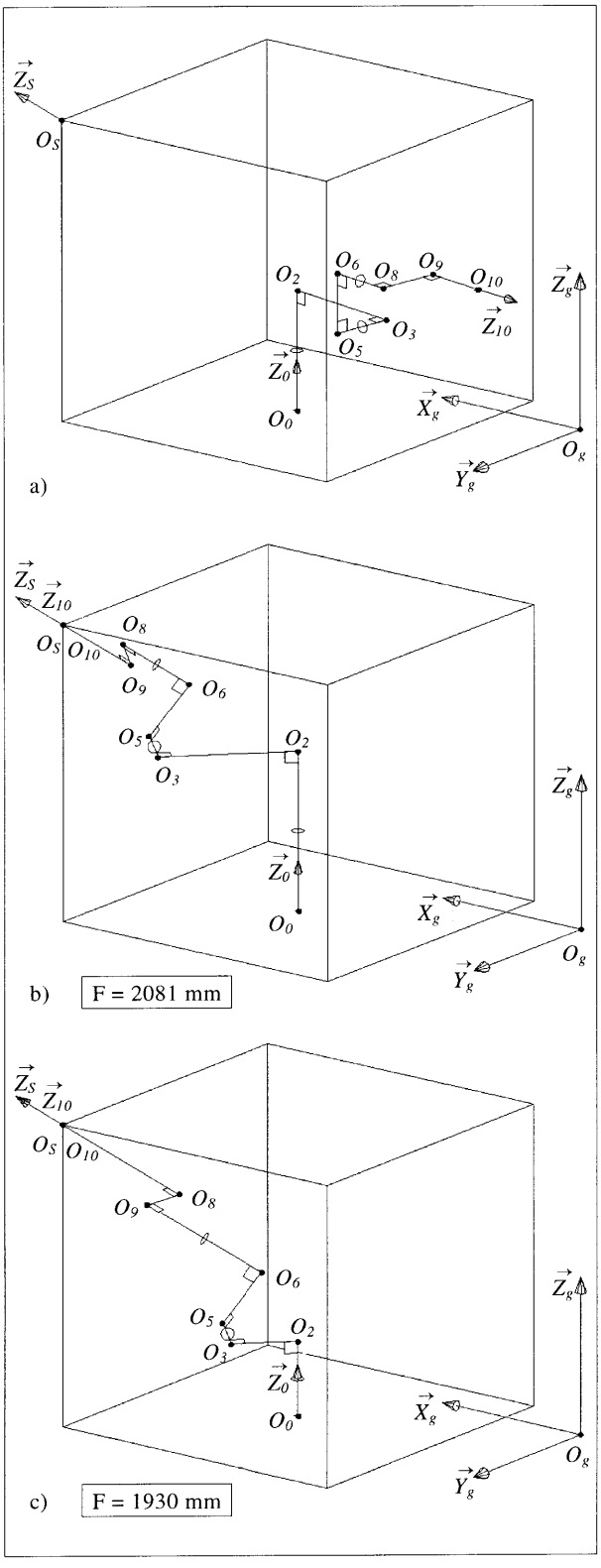


Fig. 12. Towards the first example global optimum.

Table 1. Numerical values associated with the 1st example

	Initial (a)	Step 1	Step 2 (b)	Step 3 (c)
q_1 (rad)	4.71	2.36	2.36	2.36
q_2 (mm)	400		541	248
q_3 (mm)	400		421	200
q_4 (rad)	4.71	5.33	5.33	5.33
q_5 (mm)	200		205	200
q_6 (mm)	200		209	200
q_7 (rad)	1.57	1.57	1.57	2.11
q_8 (mm)	200		247	420
q_9 (mm)	200		205	234
q_{10} (mm)	200		252	428
F (mm)			2081	1930

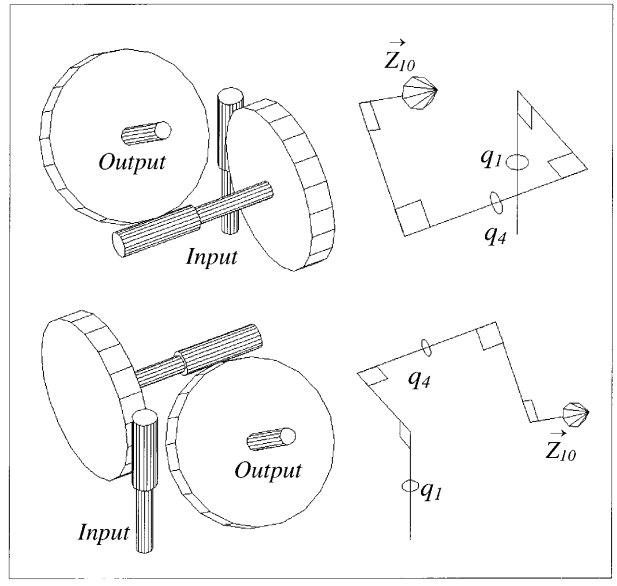


Fig. 13. Two dual geometrical configurations.

and q_7 always take different values. This typically characterises the existence of an infinity of solutions.

The second step is now to determine values which have to verify the six closure equality constraints. However, in this case, the optimisation algorithm fails in its search from various starting points. Figure 16 shows an example among the different results which could be obtained: output orientation is correct, but not output position. This seems to indicate the impossibility of fulfilling the specifications with the considered structure.

To confirm the validity of this conclusion, literal expressions of the six equality constraints associated with the geometrical model in use are developed below. Equations (23) and (24), respectively, represent orientation and position constraints:

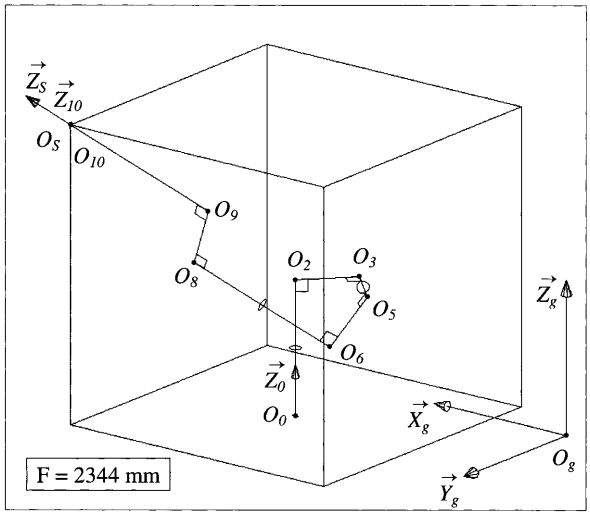


Fig. 14. First example local optimum.

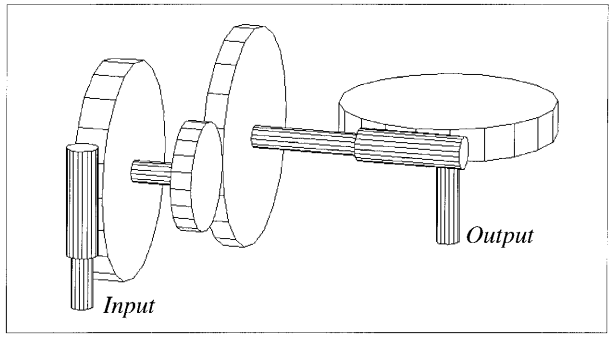


Fig. 15. Second example general architecture.

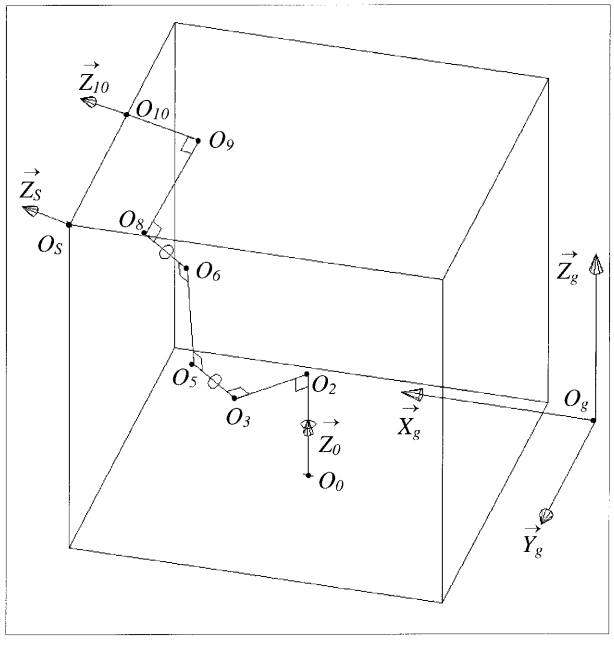


Fig. 16. Example of a closure failure.

$$\begin{cases} s_1 s_4 c_7 + s_1 c_4 s_7 = 1/\sqrt{3} \\ -c_1 s_4 c_7 - c_1 c_4 s_7 = 1/\sqrt{3} \\ -c_4 c_7 + s_4 s_7 = 1/\sqrt{3} \end{cases} \quad (23)$$

$$\begin{cases} (s_1 s_4 c_7 + s_1 c_4 s_7) q_{10} + (s_1 c_4 c_7 - s_1 s_4 s_7) q_9 \\ -c_1 q_8 - s_1 c_4 q_6 - c_1 q_5 + s_1 q_3 = 1100 \\ -(c_1 s_4 c_7 + c_1 c_4 s_7) q_{10} + (c_1 s_4 s_7 - c_1 c_4 c_7) q_9 \\ -s_1 q_8 + c_1 c_4 q_6 - s_1 q_5 - c_1 q_3 = 1100 \\ (s_4 s_7 - c_4 c_7) q_{10} + (c_4 s_7 + s_4 c_7) q_9 - s_4 q_6 + q_2 = 1100 \end{cases} \quad (24)$$

where s_i and c_i represents respectively sin q_i and cos q_i. Equation (23) may be simplified as follows:

$$\begin{cases} -\sin q_1 \sin(\pi + q_4 + q_7) = 1/\sqrt{3} \\ \cos q_1 \sin(\pi + q_4 + q_7) = 1/\sqrt{3} \\ \cos(\pi + q_4 + q_7) = 1/\sqrt{3} \end{cases} \quad (25)$$

Equation (25) is similar to Eq. (22). This proves the existence of an infinity of solutions for the angular variables. According to Eqs (20) and (21), they are such that

$$\{q_1, (\pi + q_4 + q_7)\} = \{2.36rad, 5.33rad\} \quad (26)$$

$$(q_1, (\pi + q_4 + q_7)\} = \{5.50rad, 0.95rad\} \quad (27)$$

Making some substitutions between Eqs (23) and (24), the following relation may be deduced:

$$q_8 + q_5 = 0 \Leftrightarrow O_3O_5 + O_6O_8 = 0 \text{ (Fig. 16)} \quad (28)$$

This relation cannot be satisfied, since both variables *q5* and *q8* are supposed to be strictly positive. This confirms the impossibility of closing the chain.

This example illustrates the interest of the preliminary optimisation steps: they achieve a better confinement of the failure location, and then they help the analyst who will have to imagine modifications for his specification data. In that case, if the output orientation must be preserved, the output point position has to be modified. The designer can try to find a more appropriate position for O_s before calling the current reducer architecture into question.

12. Third Example

A five stage *3D.S.R.* example is now presented to show the algorithm's behaviour with a larger number of variables. This example has deliberately been chosen so that the optimum can be verified from direct human reasoning. The *3D.S.R.* is made of:

- an external/external cylindrical gearing stage with shafts on opposite sides ($\varphi = 0°$);

- a bevel gearing stage with $\varphi = 45°$;
- an external/external cylindrical gearing stage with shafts on opposite sides ($\varphi = 0°$);
- two external/external cylindrical gearing stages with shafts on the same side ($\varphi = 180°$).

Figure 17 gives a rough sketch of the reducer under study. Other specifications are (in the general co-ordinate system R_g:

- *Envelope* $x_m = y_m = z_m = 100$ mm
  $x_M = 600$ mm
  $y_M = z_M = 1100$ mm
- *Input shaft* $O_0 = (350,300,100)$
  $\vec{X}_0 = (1,0,0)$
  $\vec{Z}_0 = (0,0,1)$
- *Output shaft* $O_s = (350,1100,1100)$
  $\vec{Z}_s = (0,1/\sqrt{2}, 1/\sqrt{2})$

The starting point is shown in Fig. 18(a). Figures 18(b) and 18(c) present the configuration obtained after a few seconds (about five seconds). The objective function is $F = 2448.5$ mm. Shafts of the three last stages are perfectly arranged in an equilateral triangle with sides of $L_{\min}$ (Fig. 18(c)). This disposition, which seems fairly obvious when proposed by the algorithm, was not so easy to imagine by direct thinking. A designer would be naturally inclined to dispose all shafts in a single planar surface.

Note that for this example, there is an infinity of optimal solutions:

- points O_7 to O_{14} can be translated along $\vec{Z}_s$ direction,
- points O_8 to O_{13} can be rotated about the axis $(O_s\vec{Z}_s)$.

without changing the optimum value of function F.

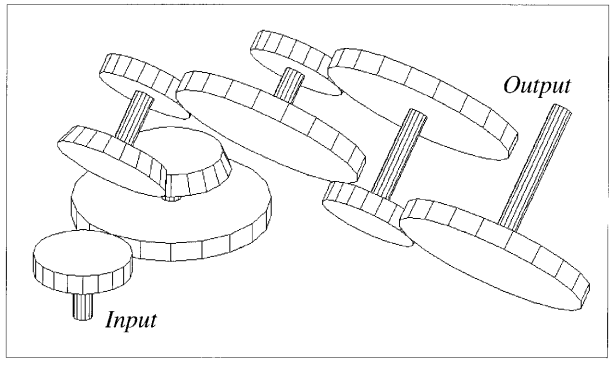


Fig. 17. Third example general architecture.

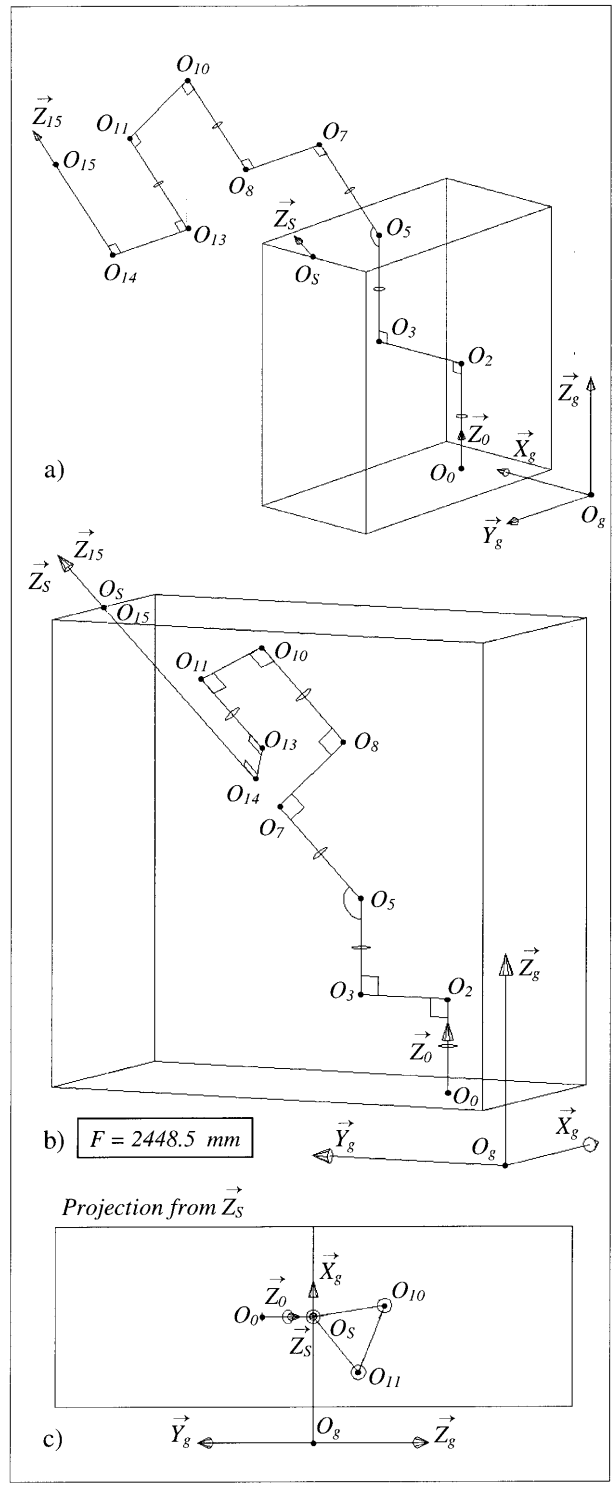


Fig. 18. Third example initial and final configurations.

This only appears from the numerical results when the analyst varies the starting point. The next design phase has to be able to model and exploit this vagueness.

13. Conclusion

The principles for a preliminary design CAD tool have been presented. This mechanism synthesis tool is specifically dedicated to help designers create 3D speed reducer (3DSR) structures.

The new concept of mechanism *skeleton* was introduced to represent and position in space the main geometrical characteristics of any 3D transmission mechanism. A formulation using the well-known mDH notation was then presented and applied to basic standard reducer stages to create a skeleton database. Finally, an assembly method was described for automatic construction of the skeleton associated with any complex 3D power transmission line.

These concepts were associated with matrix loop formulation, and proved to be extremely useful for solving the general speed reducer best 3D geometry problem. It was expressed as an optimisation problem where the overall dimensions of the reducer should be minimised ('rubber band' function) under two types of constraints: equality constraints for closure conditions; and inequality constraints for keeping the skeleton inside a given envelope.

Some particular dispositions have been taken to improve the performance of the solving process, and the help offered to the designer:

- expression of the closure chain equation has been reduced to only six scalar relations, decreasing the number of equality constraints to be processed, and thus facilitating the work of the optimizer;
- two preliminary optimisation phases have been added at the beginning to obtain progressively consistent initial values and, should the occasion arise, to enlighten the designer about the causes of failure when the structure under study appears to be inappropriate for the specification data.

Software was implemented to show the method advantages. Thanks to a powerful optimisation tool, 3D CAD software and a user-friendly graphical interface, the designer has the opportunity to perform a rough feasibility analysis on any type of 3DSR.

This approach is quite general; it may be extended either to tree-structured power transmission lines, such as drilling machines with multiple spindles, or to any device whose architecture can be represented by filar structure at the early stage of design.

References

1. Artobolevski, I. I. (1986) Mechanisms in Modern Engineering Design (Vols 1–5, 3440 p.), Moscow, MIR Publishers
2. Buchsbaum, F., Freudenstein, F. (1970) Synthesis of kinematic structure of geared kinematic chains and other mechanisms. Journal of Mechanisms, 5, 357–392
3. Freudenstein, F., Maki, E. R. (1979) The creation of mechanisms according to kinematic structure and function. Environment and Planning, 6, 375–391
4. Hoeltzel, D. A., Chieng, W. H. (1990) Knowledge-based approaches for the creative synthesis of mechanisms. Computer-Aided Design, 22, 57–67
5. Ravisankar, R., Mruthyunjaya, T. S. (1985) Computerized synthesis of the structure of geared kinematic chains. Mechanism and Machine Theory, 20, 367–387
6. Subramanian, D., Wang, C. S. (1993) Kinematic synthesis with configuration spaces. Proceedings of International Workshop on Qualitative Reasoning about physical systems (QR-93), Orcas Islands, USA, 228–239
7. Shigley, J. E., Uiker, J. J. (1980) Theory of Machines and Mechanisms, New York, McGraw-Hill
8. Söylemez, E., Freudenstein, F. (1982) Transmission optimization of spatial 4-link mechanisms. Mechanism and Machine Theory, 17(4), 263–283
9. Hartenberg, R. S., Denavit, J. (1964) Kinematic Synthesis of Linkages, New York, McGraw-Hill
10. Megahed, S. M. (1993) Principles of robot Modeling and Simulation, New York, Wiley
11. Denavit, J., Hartenberg, R. S. (1955) A kinematic notation for lower-pair mechanisms based on matrices. Journal of Applied Mechanics, 22, 215–221
12. Khalil, W., Kleinfinger, J. F. (1986) A new geometric notation for open and closed-loop robots. Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'86), San Francisco, CA, 1174–1180
13. Design Optimization Tools, Version 4.20 (1995) Vanderplaats R&D, Colorado Springs, USA
14. Hall, A. S., Root, R. R., Sandgren E. (1977) A dependable method for solving matrix loop equations for the general three-dimensional mechanism. Journal of Engineering for Industry, August, 77, 547–550