TECHNICAL RESEARCH REPORT

Computer-Aided Sketching of Epicycle-Type Automatic Transmission Gear Trains

by G. Chatterjee and L-W. Tsai

T.R. 95-92

INSTITUTE FOR SYSTEMS RESEARCH

Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry

COMPUTER-AIDED SKETCHING OF EPICYCLIC-TYPE AUTOMATIC TRANSMISSION GEAR TRAINS

Goutam Chatterjee
Department of Mechanical Engineering and
Institute for Systems Research
University of Maryland
College Park, Maryland

Lung-Wen Tsai
Department of Mechanical Engineering and
Institute for Systems Research
University of Maryland
College Park, Maryland

ABSTRACT

The enumeration of epicyclic gear mechanisms in the form of graphs gives rise to the need of a methodology for reverse transformation, that is, for constructing the mechanisms from graphs. This paper addresses the issue by discretizing an epicyclic gear mechanism into Fundamental Geared Entities. Further, these geared entities are shown to be a conglomeration of four primitives; namely, the carrier, sun, ring, and the planet gear. An algorithm is formulated to create the entities from a graph by using these primitives. The entities are then connected together to form a mechanism.

INTRODUCTION

A literature search reveals that little work has been done in the area of computer aided sketching of mechanisms (Chieng and Hoeltzel, 1990), particularly in the area of automotive transmission mechanisms. Most automotive automatic transmissions use epicyclic gear trains[1] (EGTs) to achieve the desired reduction ratios. Typically, an automatic transmission uses a one-dof (degree-of-freedom) EGT supported on one axis by bearings housed in the casing. This results in a two-dof fractionated mechanisms. Henceforth, we will call the mechanism formed by an EGT and the casing of an automatic transmission gear box an *Epicyclic Gear Mechanism* (EGM).

Graph Reprsentation

The graph representation of an EGT was described by Buchsbaum and Freudenstein (1970) in their pioneering paper. In a graph representation (Fig. 1(a) and (b)) links are represented by vertices, while joints are represented by edges. The gear joints are represented by thick edges while the revolute joints are represented by thin edges to distinguish one from the other.

The problem of isomorphism always arises in graph representation. Two graphs are isomorphic if there exists a one-to-one correspondence between their vertices and edges which preserves the incidence and labeling. For EGTs, this problem is further convoluted by the fact that two mathematically non-isomorphic graphs can represent mechanisms that are kinematically equivalent. For example, the mechanism shown in Fig. 1(a) can be reconfigured into the mechanism shown in Fig. 1(c) by rearranging the revolute joints among its coaxial links. Though these two mechanisms appear to be structurally non-isomorphic (Mruthyunjaya and Ravisankar, 1985), they are kinematically equivalent, and for the purpose of structural synthesis are considered the same. However, the graphs of the two mechanisms (Figs. 1(b) and (d) respectively) are mathematically non-isomorphic. Such graphs are called *pseudoisomorphic* graphs (Tsai and Lin, 1989).

Canonical Graph Representation

The problem of pseudoisomorphism can be averted by imposing some rules that result in unique arrangement of the edges of the same label. Such a graph is called a *canonical graph* (Tsai, 1988; Chatterjee and Tsai, 1994).

In the canonical graph representation of an EGM, the vertex representing the casing is marked as the root of the graph. Among various arrangements of the coaxial joints there exists a unique configuration such that all the thin edged paths originating from the root and ending at all the other vertices will have distinct edge labels. This unique graph representation is called the canonical graph represen-

[1] A few use countershaft type of arrangement that is typical of manual transmission. A hybrid type using a combination of both is also possible

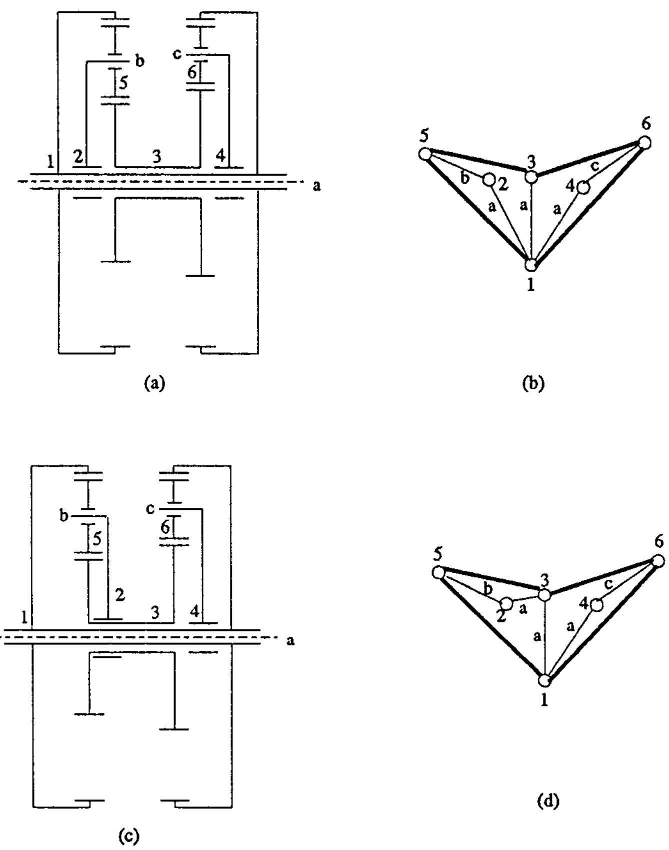

Figure 1. EGTs: (a,c) functional representations, (b,d) corresponding graph represntations.

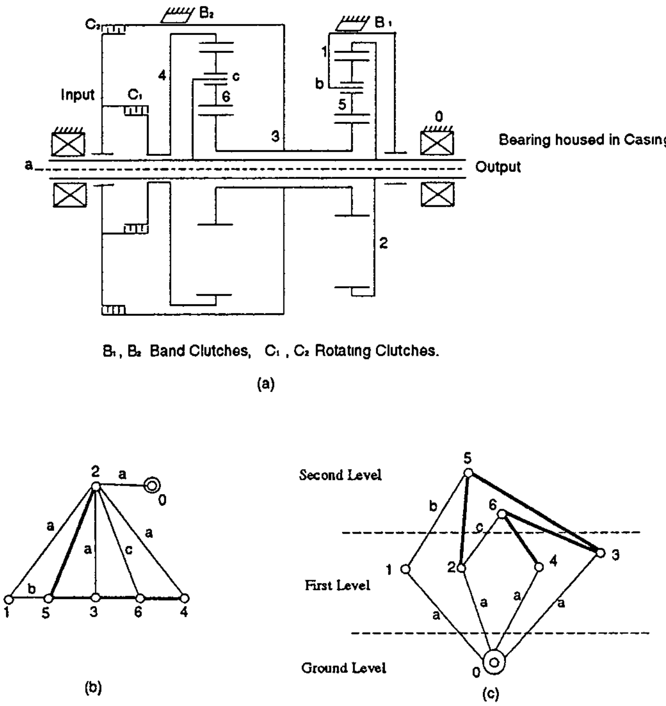

Figure 2. EGM employing Simpson gear set: (a) functional representation, (b) conventional graph represntation, (c) canonical graph representation.

tation. Using canonical graph representation, the vertices can be divided into several levels. The casing is denoted as the *ground* level vertex. A vertex that is connected to the root by only one thin edge is defined as a *first* level vertex. A vertex that is connected to the root by two thin edges is defined as the *second* level vertex and so on.

For example, Fig. 2(a) shows the Simpson gear set, Fig. 2(b) shows its conventional graph representation, while Fig. 2(c) shows the canonical graph representation. The first level vertices represent the sun gears, ring gears, and the carriers while the second level vertices represent the planet gears. A review of the work of various authors Gott (1991), Larew (1966), Levai (1964), Tsai (1988) had not revealed a single automatic transmission gear box having a link located on the third or higher levels. Therefore, in what follows, we will limit ourselves to canonical graph with vertex distribution only upto the second level.

Assignment of Gear Type

The canonical graph as described above does not distinguish between external and internal gear pairs. However, for the purpose of drawing the schematics, such a distinction is important. To do this we observe that a gear edge connecting a second-level vertex to a first-level vertex represents a planet-to-sun or planet-to-ring gear pair depending on whether the gear pair is external or internal. A heavy edge connecting two vertices of the second level represents a planet-to-planet interaction. Theoretically a gear pair between two planets can be either external or internal. However, for all the EGMs studied, only external gear meshes between planets have been found. Therefore to keep the sketching algorithm simple, we shall assume that a heavy edge connecting two vertices at the second level represents an external gear pair. In the adjacency matrix, an internal gear pair will be represented by a upper case G, an external gear pair will be represented by a lower case g, and a revolute joint will be represented by its edge label. Fig. 3(b)

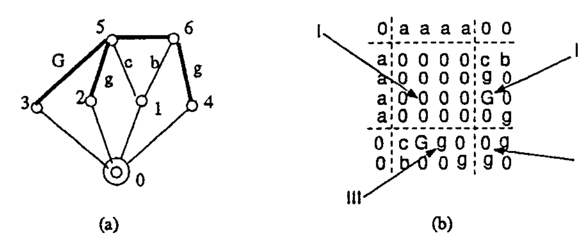

Figure 3. (a) Canonical Graph Representation of the Epicyclic Gear Mechanism shown in Fig. 3(a), (b) Link-to-Link Adjacency Matrix of (a).

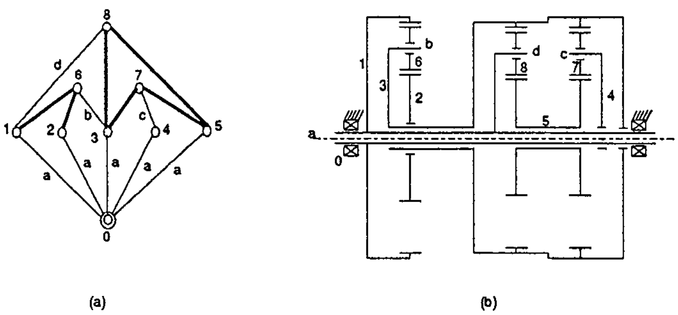

Figure 4. Epicyclic gear mechanism of 8 link: (a) graph representation, (b) functional representation showing coaxial shafts and overhead connection.

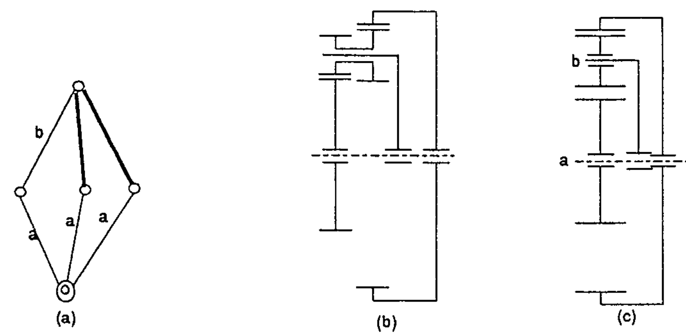

Figure 5. Simple Planetary Gear Set: (a) graph representation, (b) functional representation with three independent gear diameters, (c) functional representation with only two independent gear diameters.

shows the adjacency matrix for the canonical graph representation shown in Fig. 3(a). If we remove the first row and the first column corresponding to the root, we can divide the remaining matrix into four sub-matrices marked I, II, III, and IV in Fig. 3(b) according to the vertex level distributions. Since the adjacency matrix is symmetrical, the sub-matrices III and IV are transposes of each other. All elements in the sub-matrix I are zero, since there can be no connections among the first level vertices. The sub-matrix II gives the interaction among the second level vertices. The sub-matrix III represents the interaction between the first and the second level vertices. Therefore, only the elements of sub-matrix III and IV can have the edge label G.

Functional Representation

The graph representation of a mechanism helps in enumerating the basic configuration of a mechanism and stores all the relevant topological data. However, it does not aid the designer in visualizing how a mechanism looks and how it will work. Such knowledge is a prerequisite for the next phase of mechanism synthesis, i.e. the dimensional synthesis phase.

The functional representation of an EGT, unlike those of other planar kinematic chain, has traditionally been sketched by using a sidewise orthographic projection. Some of the desired features of a good functional schematics are:

No Interference - There should not be any crossing (or interference) between links in the functional schematic of an EGT. (Chieng and Hoeltzel, 1990)

Accessibility - The first level links of an EGM are used as input, output or fixed links. The link chosen as the output link is to be permanently connected to the final reduction unit of a transmission. The other links are connected either to the input shaft through rotating clutches or to the casing through band clutches, depending on the clutching condition. Therefore, it is necessary that, in the functional schematic of an EGM, these links be arranged in such a way that they remain accessible and hence can be connected to other elements of the transmission as, and when, required.

Few Coaxial Shafts - Link No. 1 of the functional schematic shown in Fig. 4(b) has a carrier and a ring gear connected[2] to each other by a shaft. This shaft is coaxial with another shaft, which connects the carrier and the ring gear of link 3. From a manufacturing point of view coaxial shafts are undesirable as tolerance limits on such shafts are too rigid. Therefore, their number should be kept as low as possible.

Low Inertia - Link No. 3 of the functional schematic shown in Fig. 4(b) has two ring gears attached edge to edge. Such connection will be referred to as *overhead* connection. This way one can reduce the number of coaxial shafts. However, this results in a high moment of inertia of the link. It also requires more material in its manufacture. Thus one has to strike a balance between the two desirable features. A method of maintaining a proper balance between coaxial shafts and overhead connection will be described in a later section.

It should be noted that the accessibility of a link is a more desirable feature than the last two features. Therefore, the latter two are sacrificed whenever the accessibility of a link is in question. Since the vertices of a graph do not carry information about the shape of links, it is possible to draw structures that are less general, and put additional constraint on the dimensions of certain links. For example, the functional schematic shown in Figs. 5(b) and (c) has the same graph representation as that shown in Fig. 5(a). However the EGT of Fig. 5(c) must have all its gears of the same module and the diameters of the gears are not independent.

FUNDAMENTAL GEARED ENTITIES (FGEs)

In this section, we shall divide the graph of an EGM into several FGEs by using the planet gears as the dividing ele-

[2] The difference between the words *connections* and *joints* as used here is: two or more members such as gears or carriers are said to be *connected* when they form a link with no relative motion between them, whereas two links are said to be *joined* when they can have specific relative motion with respect to each other.

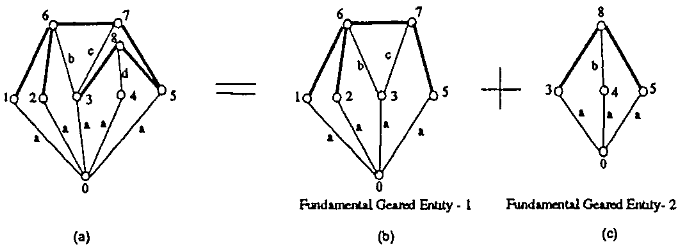

Figure 6. (a) Canonical graph of an EGM, (b,c) subgraphs representing fundamental geared entities.

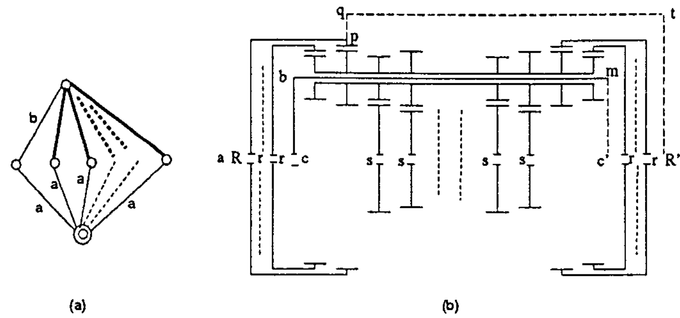

Figure 7. General Form of Single-Planet FGE.

ments. The planet gears can either be a rigid link by itself or several planets joined together as a planet train. From the graph point of view, the second-level vertices either become isolated or form several second-level vertex-chains after the removal of all the lower level vertices. For example, in the graph shown in Fig. 6(a) vertices 6 and 7 along with the geared edge connecting them forms a double planet train, while vertex 8 represents a single planet. The mechanism represented by a subgraph formed by an isolated second-level vertex or a chain of second-level vertices, and all the lower level vertices connected to them by geared or revolute edges, is called a *fundamental geared entity.* Figs. 6(b) and (c) show two subgraphs formed from the graph of Fig. 6(a), each of which represents an FGE. From the definition of FGE it follows that links corresponding to the vertices of a second level vertex chain of one FGE do not directly interact with those of the others. Therefore, to draw the functional representation of an EGM, *one should draw the FGEs first and then connect the common links of the FGEs by means of coaxial shafts or overhead connections.*

FGEs that have more than two planets in a chain are not very practical and will be excluded from further consideration. The FGEs can therefore be categorized into two general types, one containing only one planet and the other containing two meshing planets. Their graphs are shown in Fig. 7(a) and Fig. 8(a), respectively. The general form of a *single-planet FGE* as it appears in the functional schematic is shown in Fig. 7(b). The ring gears can occur only at either end of the FGE. The general form of the functional schematic of a *double-planet FGE* is shown in Fig. 8(b). It appears to be complicated because of the presence of a ring gear in the middle of the FGE. If there is no sun gear meshed with the planet at the higher label (axis c), then the ring gear can be drawn as shown in Fig. 8(c). Thus, there are two general forms of the second-type FGEs. One may note that the restriction we have imposed on the position of the ring gears (joined to the planet at its lower label) as shown in Fig. 8(b) is artificial. Clearly, one can also place the ring gears at the end by offsetting the lower label (axis *b*) of the carrier (see Fig. 8(d)). This, however, leads to too many hidden lines in a 2D representation and therefore is avoided. Such an artificial restriction however may not permit the achievement of the desired features (of functional representation) in some cases.

PRIMITIVES

The functional schematic of an FGE can be sketched by using four primitives depicting the sun, ring, planet and the carrier. The dimensions that characterize these primitives are decided by the program based on the arrangement of the primitives in the FGE. A methodology to decide the arrangement of the primitives is described in the next section. However, since the computer does not have any aesthetic sense, it will be given some proportions to begin with. These are described under specific primitives in the following subsections.

Carrier

The carrier primitive is shown in Fig. 9(a). All the planets in an FGE are to be supported on the higher level axes of a carrier. The height of the axis *b* from the centerline *a* of the carrier is introduced as a constant in the computer. We will call this constant the *reference value.* For the case of a double planet, the ratio of the distance between two higher level axes *b* and *c*, to the reference value is also defined. Whenever an axis corresponding to a new label is to be created, the program creates a variation about the reference value to define its height. The number of axes, the options of whether the carrier should face left or right, and the other parameters are decided by the program. The coordinates of the numbered points 1, 2, 3 etc. as shown in Fig. 9(a) are expressed as functions of the reference parameter with respect to the coordinates of the local origin of the primitive.

Sun Gear

The sun gear primitive is shown in Fig. 9(b). The parameters characterizing the primitive are shown in the same figure. Some of these parameters are defined as constants in the program as they don't affect any of the desirable features of the functional representation. They are defined in proportion to the above mentioned reference value. These parameters are the *face width* of the gear, the *gap between the teeth* of two meshing gear, the *hub* width and the *nominal hub diameter*. Another proportion that needs to be defined is the ratio of the radius of a sun gear to that of the meshing planet. The computer always uses a pre-specified ratio and creates a small variation about this ratio while

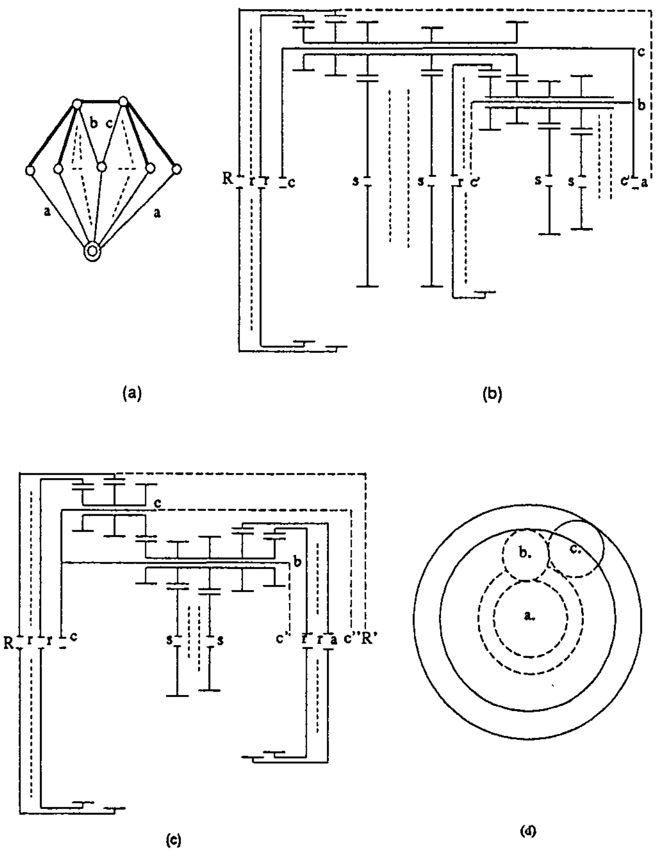

Figure 8. General Forms of Double-Planet FGE.

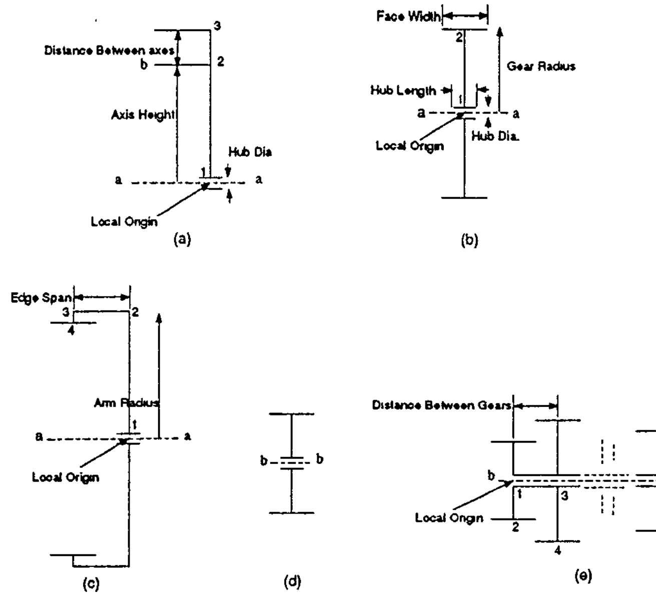

Figure 9. Primitives: (a)carrier, (b) sun gear, (c) ring gear, (d) single-planet gear, (e) multiple planet gear.

sketching the structure. In the graph of an FGE this primitive appears as a first level vertex incident by a gear edge labeled g.

Ring Gear

The ring gear primitive and the parameters characterizing it are shown in Fig. 9(c). Some parameters defined for the previous primitive also apply here. In addition, the *gear radius*, the *ratio of arm radius to gear radius*, and the *nominal edge span* as a proportion to the reference value are also defined. The program, however, may override these specifications under certain circumstances that will be described later. In the case of a ring-planet pair the program also has to decide the orientation of the primitive.

Planet Gear

The planet consists of one or more gears attached to each other side by side as shown in Fig. 9(d) and (e). In the graph representation of an FGE it appears as a second-level vertex or. The number of gears in the planet is determined by the number of geared edges incident on the vertex. In this case the distance between planet gears are defined in proportion to the reference value.

CONNECTING THE FGEs

The next task is to connect the common links of the FGEs by means of coaxial shafts or overhead connections. To reduce the number of choices we will forego the overhead connections in the beginning. Later, to reduce the number of coaxial shafts, we will present a method of converting some of the coaxial shafts to overhead connections.

Consider the general form of a single planet FGE as shown in Fig. 7(b). The points along the axis of the FGE that are labeled with letters can be connected by shafts to similar points of other FGEs. Such points will be called the *welding* points. The left most or the right most ring gear can have two welding points as shown in Fig. 7. Such a construction is permissible because it does not hamper the accessibility to any other points. In the same manner the carrier can also have two welding points: one on the left-hand-side and the other on the right-hand-side (line mc') of the shaft as shown in Fig. 7(b). Therefore, the welding points along the axis can be represented as R r r .. c .. s s .. c' .. r r R', where the r's represent the welding point of the ring gears, the c's that of the carrier and the s's that of the sun gears. The R represents the ring gear that has two welding points. Simalarly, the welding points of the two double planets FGEs shown in Figs. 8(b) and (c) can be represented as R r r .. c .. s s s ..r r .. c' .. s s c" R' and R r r .. c .. s s c' .. r r c" R', respectively.

The link numbers corresponding to sun gears joined to the same planet of an FGE can be permuted among each other because of symmetry. The same thing can be done for the ring gears that are joined to the same planet. Fig. 10(a) shows a canonical graph with geared edges properly labeled. It can be divided into three subgraphs, each of which rep-

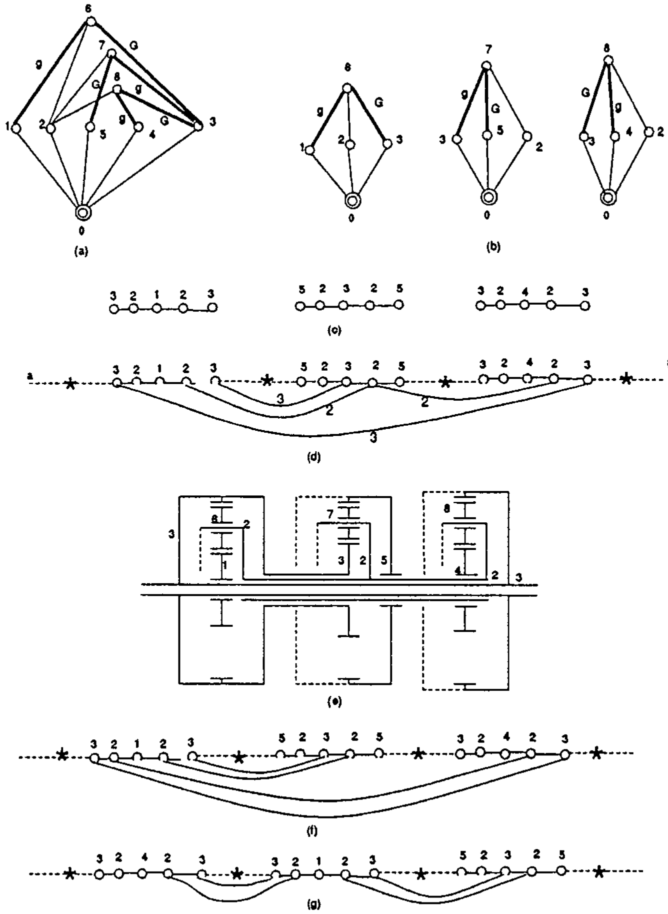

Figure 10. Constructing the Kinematic Structure from Primitives.

resents an FGE as shown in Fig. 10(b). Each FGE consists of only one carrier, one ring gear, and one sun gear. Each ring gear and carrier has two welding points. The welding points of the FGEs can be arranged graphically as shown in Fig. 10(c). The welding points are labeled by the number of the links. The welding points belonging to the same FGE are joined by edges, which are called *primary edges*. Edges connecting two welding points of two different FGEs are called *secondary edges*. A secondary edge can only connect two welding points of the same label. The label of a secondary edge is the same as those of its end points. A secondary edge between two points represents a shaft connection between the two links. The secondary edges should be drawn according to the following rules.

M1 The secondary edges should be contained in the half space below the line *ab* formed by the primary edges as shown in Fig. 10(d). This restriction is necessary because we are only considering shaft connections.

M2 Any two welding points of the same label should be connected by a path consisting of either secondary edges of the same label or several primary edges or a combination of both.

M3 There should not be a circuit formed by only secondary edges of the same label and primary edges.

M4 A secondary edge should not intersect or coincide with another secondary or primary edge except at its end points.

After all the above rules are satisfied, a test for accessibility of links should be carried out. A welding point in Fig. 10(d) is accessible if one can draw a secondary edge from that welding point to any one of the points marked by an asterisk without tracing on or crossing any edge. Otherwise, the point is inaccessible. A link is inaccessible if all its welding points are inaccessible. For example, the way the secondary edges has been drawn in Fig. 10(d), welding point 1 is inaccessible. Fig. 10(e) shows the corresponding connections in the mechanism. It is evident that link 1 is indeed inaccessible. However, if one connects the FGEs in the way shown in Fig. 10(f) instead, link 1 becomes accessible. Although this serves the purpose, a pair of connections between two non-adjacent FGEs means two long coaxial shafts are needed. This should be avoided, if possible. In this case, it can be avoided by arranging the FGEs as shown in Fig. 10(g).

Based on the above discussion the following algorithm is proposed to find a proper arrangement of the FGEs and that of the links within the FGEs. The first two steps describe how the first FGE should be constructed. The rest of the steps describe how the subsequent FGEs should be constructed.

Step 1 Draw the welding points of an FGE that has only two of its links connected to other units. There is at least one such unit in any EGM. If there is more than one such unit, choose any.

Step 2 Choose one welding point for each link in the first FGE, that is connected to links in other FGEs (to be added subsequently). Place these welding points adjacent to each other, otherwise the welding point that lies between them becomes inaccessible. The only welding point that can be trapped between is the one belonging to the carrier, provided the other one is accessible.

Step 3 Place another FGE to the right of the partially completed mechanism schematic. Select the one that has the maximum number of links to be connected to the preceding FGEs. If there is more than one choice, choose any. Henceforth, we will refer to an FGE that is being added as the current FGE and its welding points as current welding points. The FGEs that belong to a partially completed mechanism schematic will be referred to as the preceding FGEs.

Step 4 For every current welding point, choose a corresponding welding point in the preceding FGEs to which a connection can be made without crossing edges. If more than one such welding point is found choose the right-most among them. If no such welding point is found, then alter the arrangement of the welding points in the preceding FGEs. While altering the arrangement

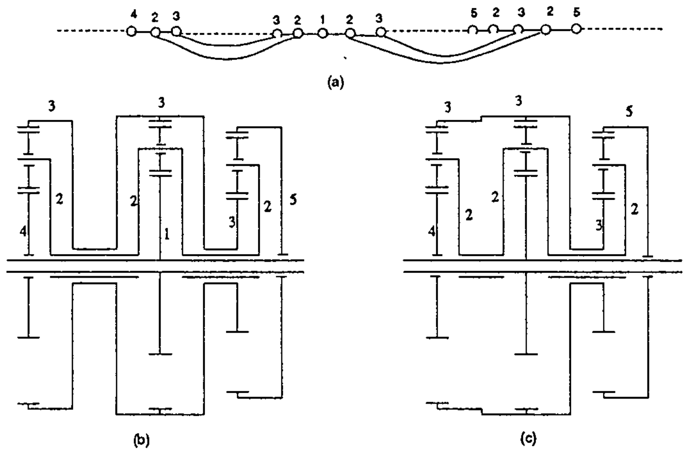

Figure 11. Constructing the Functional Schematic.

of the welding points in the preceding FGEs it should be kept in mind that the general form of the FGEs as described previously should not be changed. Also, permutations can be done only among those welding points that are either connected to the current FGE or not yet connected. If this does not work, make a different choice of the immediately preceding FGE and repeat Step 3.

Step 5 If a and b are two current welding points to be connected to the partially completed mechanism schematic, then a should be placed to the left of b if the corresponding welding point a in the preceding FGEs is located to the right of b (in the preceding FGEs). If this is not possible, alter the choice of the preceding welding points.

Step 6 The first connection from a current FGE to the preceding FGEs won't make any welding point inaccessible. However, any successive connections could. In that case check whether the corresponding link becomes inaccessible. If it is so, go to Step 4 and choose a different welding point. If this does not work, follow the permutation procedure given in Step 4. First permute the welding points of the preceding FGEs and then change the sequence of the FGEs themselves.

After the secondary edge connection is made so that rules M1 to M4 are satisfied, some of the welding points that are not connected by secondary edges can be removed if a corresponding welding point of the same label exists in the other FGEs. Once the above operations are complete the EGM can be sketched by mapping the welding points to the primitives of the links to which they belong. For example, functional schematic developed by mapping the welding points in Fig. 11(a) is shown in Fig. 11(b). Some of the shaft connections between adjacent units can be converted into overhead connections as shown in Fig. 11(c). This can be done only if all of the following conditions are satisfied.

(i) the welding points are adjacent to each other,

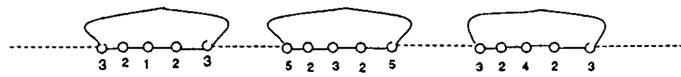

Figure 12. Arrangement of Connection Point to Sketch Both Overhead and Shaft Connections.

(ii) they belong to a ring or a carrier,

(iii) there is one more welding point of the same label in either of the FGEs.

Till now we have only discussed shaft connections and converting shaft connections of adjacent FGEs to overhead connections. However, the method discussed above can easily be generalized to include all possible overhead connections by arranging the welding points of the FGEs as shown in Fig. 12, and then making the secondary-edge connections so that they satisfy rules M2 to M4. This however, will result in complicated connection that won't be easy to sketch in practice. It is also not necessary if we limit ourselves to a small number, of FGEs, say four.

DISPLAY OF THE FUNCTIONAL SCHEMATIC

The functional schematic of an EGM is stored in the computer as a network of structures. For example, each of the primitives is stored as a structure. These structures contain not only the coordinates of the points that define the dimensions of the primitives, but also the pointers to other primitives to which they are connected. Each of the primitives is created in the local coordinate system of an FGE to which it belongs. Then the FGEs are translated to occupy their relative position with respect to each other in the global coordinate system. The lines corresponding to shaft connections are then created. The hub diameter of the gears and carriers are adjusted so that they don't interfere with the shafts. To display the structure on to the computer screen, a graphics package called PHIGS is used. The PHIGS window is a square window with its sides of unit length. The lower left corner has the coordinate of (0,0) and the upper right corner has the coordinate (1,1). Therefore, the structure before being displayed is first transformed from the World Coordinate System to the Screen Coordinate System and then scaled appropriately to fit the window.

CONCLUSIONS

A method to sketch the functional schematic of an EGM from its graph representation has been formulated. Using this method the functional schematic of a fairly large class of mechanisms can be sketched automatically. The method may be generalized to sketch EGMs having other type of fundamental entities. At present, a C-language program has been developed to sketch those EGMs in which each FGE contains only one planet gear. Fig. 13(b) shows an 11-link EGM that is sketched from the graph representations shown in Fig. 13(a) using this program.

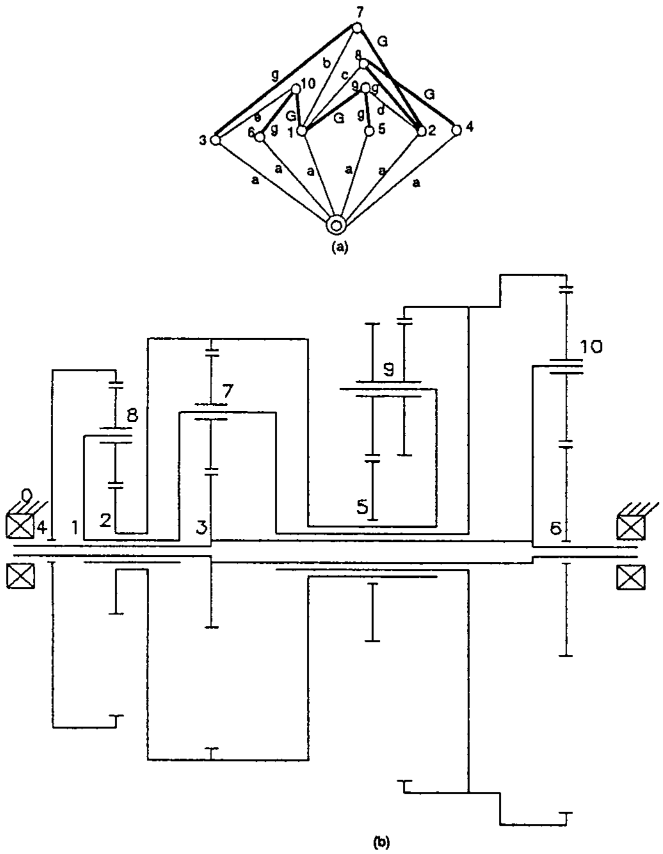

Figure 13. An 11-link EGM: (a) Graph Representation (b) Functional Representation.

REFERENCES

[1] Buchsbaum, F., and Freudenstein, F., 1970, "Synthesis of Kinematic Structure of Geared Kinematic Chains and Other Mechanisms," *Journal of Mechanisms and Machine Theory*, Vol. 5, pp. 357-392.

[2] Chatterjee, G., and Tsai, L.W., 1994, "Enumeration of Epicyclic Type Automatic Transmission Gear Trains," Accepted for publication at the SAE 1994 Congress and Exposition, Detroit, Michigan.

[3] Chieng, W.H., and Hoeltzel, D.A., 1990, "A Combinatorial Approach for the Automatic Sketching of Planar Kinematic Chains and Epicyclic Gear Trains," *ASME Journal of Mechanical Design*, Vol. 112, pp. 6-15.

[4] Gott, P.G., 1991, *Changing Gears*, Society of Automotive Engineeers, Inc., Warrendale, Pennsylvania.

[5] Larew, W.B., 1966, *Automatic Transmissions*, Chilton Book Company, Radnor, Pennsylvania.

[6] Levai, Z., 1964, *Theory of Epicyclic Gears and Epicyclic Change-Speed Gears*, Doctoral Dissertation, Technical University of Building, Civil and Transportation Engineering, Budapest.

[7] Mruthyunjaya, T.S., and Ravisankar, R., 1985, "Computerized Synthesis of the Structure of Geared Kinematic Chains," *Journal of Mechanisms and Machine Theory*, Vol. 20, pp. 367-387.

[8] Tsai, L.W., 1988, "The Kinematics of Spatial Robotic Bevel-Gear Trains," *IEEE Journal of Robotics and Automation*, Vol. 4, pp. 150-156.

[9] Tsai, L.W., and Lin, C.C., 1989, "The Creation of Non-Fractionated Two-Degree-of-Freedom Epicyclic Gear Trains," *ASME J. of Mechanisms, Transmissions, and Automation in Design*, Vol. 111, pp. 524-529.