International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 3 Issue 3, March - 2014

# Spring Synthesis for Nonlinear Force-Displacement Function

Arefin Ahmed, Hong Zhou
Department of Mechanical Engineering
Texas A&M University-Kingsville
Kingsville, Texas, USA

***Abstract*—Nonlinear springs have nonlinear force-displacement relationships and wide applications in damping, suspension, vibration absorption and prosthetic systems. Conventional spring designs are mainly for linear springs and difficult to be used for nonlinear springs because of their customized nonlinear features. In this paper, a nonlinear spring is considered as a piece of elastic material and its configuration is defined by topology, shape and size. A systematic method is presented for synthesizing nonlinear springs. Topology of a synthesized nonlinear spring is controlled by the locations of input, output, intermediate and support nodes in the design domain. Shape of the synthesized nonlinear spring is decided by a set of spline curves that interpolate the input, output, intermediate and support nodes. Size of the nonlinear spring is described by in-plane widths that are perpendicular to the spline curves. The complete configuration of a synthesized nonlinear spring is defined by variable width spline curves which are controlled by their interpolation circles. The synthesis of a nonlinear spring is thus systematized as optimizing the control parameters of the interpolation circles. The presented method is demonstrated by the synthesis of a hardening nonlinear spring.**

***Keywords*—nonlinear spring; design; spline; variable width; control parameter; interpolation circle.**

## I. Introduction

Springs are mechanical parts made from elastic materials and in particular configurations to provide a range of force over a significant deflection and/or to store and release potential energy. Springs are designed to provide a push, a pull or a twist force (torque), or to primarily store and release energy [1]. The performance of a spring is characterized by the force ($F$) applied to it and the deflection or displacement ($D$) which the applied force results. The slope of the $F$-$D$ curve is the spring rate or stiffness denoted by $k$. If the slope of a spring is constant, the spring is linear. Otherwise, it is nonlinear. Linear springs obey the Hooke's Law, $F = k\ D$. Conventional spring designs are mainly for linear springs. The applied force in a nonlinear spring is not proportional to its deflection. Two typical $F$-$D$ curves of nonlinear springs are progressive (hardening) and degressive (softening) curves, which are curves 1 and 3 in Fig. 1, respectively. Curve 2 shows a linear $F$-$D$ relationship.

A progressive nonlinear spring is a spring that gradually increases its spring rate as the spring deflection progresses, which provides a progressively hardening reaction as the spring gets compressed or extended. In contrast, a degressive nonlinear spring gradually decreases its spring rate as the spring deflection increases and provides a softening reaction. The $F$-$D$ relationship of a nonlinear spring is usually based on its specific application and has its individual feature [2]. Because of the unique nonlinear characteristics, design methods for linear springs are difficult to apply, which makes designing nonlinear springs more challenging than linear springs. The motivation of this paper is to provide a systematic synthesis method for nonlinear springs.

Mechanisms are mechanical devices used to transfer or transform motion, force or energy [3]. Conventional mechanisms are composed of rigid links that are connected by kinematic joints. In a rigid mechanism, a desired output motion in the output link is generated by a simple input motion (typically a rotation from a motor or an engine) in the input link through relative motions of connected links. Mechanism functions are realized in compliant mechanisms [4-5] by elastic deformations. The configuration of a compliant mechanism is usually a piece of elastic material without any rigid joint. The jointless configurations of compliant mechanisms offer them benefits including the elimination of backlash, friction, wear and lubrication, and the reduction of vibration and noise, manufacturing and assembly cost. The single-piece configuration of a compliant mechanism is commonly described its topology, shape and size. Topology is the overarching material layout of a compliant mechanism, and is the connectivity relationship among nodes in the compliant mechanism. A compliant mechanism has input, output, support and intermediate nodes. Its topology is the topology of the network formed by these nodes [6]. Its shape is the shape of the network, which is defined by the locations of nodes and the connection curves of nodes. Its size is the cross-sectional sizes of node connections [7]. Thus, the synthesis of a compliant

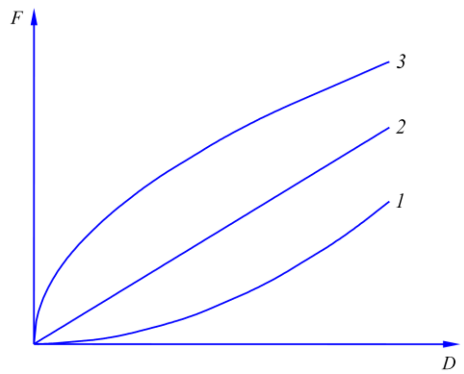


Fig. 1 Spring $F$-$D$ curves.

mechanism is to decide its topology, shape and size.

Springs usually have jointless configurations and utilize elastic deformations of materials to generate desired force-displacement relationships. They are a special class of compliant mechanisms. The input and output nodes of a spring usually coincide. The design of linear springs has been systematized. The spring rate of a linear helical coil spring can be calculated by the following equation [1].

$$k = \frac{d^4 G}{8D^3 N_a} \tag{1}$$

In (1), $d$ is the wire diameter, $D$ is the mean coil diameter, $N_a$ the number of active coils, and $G$ is the shear modulus of the spring material. The standard form of helical coil springs has constant coil diameter, constant pitch and constant spring rate. To make the spring rate of a coil spring variable, its coil diameter or pitch can be varied. A coil spring with varied coil diameter can have a conical, barrel or hourglass shape. However, the change of spring rate in such kinds of nonlinear springs is often limited. The springs synthesized in this paper are focused on translational springs in which desired nonlinear force and displacement relationships can be realized. A nonlinear spring is synthesized as a compliant mechanism by deciding its topology, shape and size.

The remainder of the paper is organized as follows. The synthesis strategies on topology, shape and size of nonlinear springs are provided in section II. The optimization approach of spring parameters is presented in section III. Section IV is on the synthesis of a hardening spring using the synthesis method introduced in this paper. Conclusions are finally drawn in section V.

## II. FORMULATION OF SPRING SYNTHESIS

The synthesis of a nonlinear spring is to decide its topology, shape and size. Topology is defined by the network of nodes in the spring, which are decided by nodes and node connections. Nodes in a spring can be input, output, support or intermediate nodes. Each connection has two end nodes plus one or more intermediate nodes. The more nodes a connection has, the more flexible the variation of its shape is. The shape of a connection is determined by the locations of its nodes and the connection curve of the nodes. Besides location parameters, each node can include one or more size parameters for the cross-section sizes of the connection. If a connection is regarded as a building block [8], a spring is then composed of one or more connections. If one connection can meet the needs, we use one connection. Otherwise, we gradually increase the number of connections. It is preferred for a spring to have fewer connections in order to simplify its structure and lower its cost.

Connections are used as building blocks for spring synthesis in this paper. The shape of a connection is the curve that interpolates all nodes of the connection. If a connection has only two nodes, i.e. two end nodes, the shape of the connection is just a straight line segment that connects its two ends. If a curved connection is needed, it takes at least three nodes. A popular node interpolation is to use Lagrange interpolation. Given $n$+1 nodes ($P_0$, $P_1$, $\cdots$ $P_n$) with pairwise parameter values ($t_0$, $t_1$, $t_2$, $\cdots$, $t_n$), its Lagrange interpolation is given as follows.

$$P(t) = \sum_{j=0}^{n} P_j L_j(t) \tag{2}$$

$$L_j(t) = \prod_{\substack{k=0 \\ k \neq j}}^{n} \frac{(t - t_k)}{(t_j - t_k)} \tag{3}$$

$$t_0 < t_1 < t_2 \cdots < t_n \tag{4}$$

$L_j(t)$ in (3) is of degree $n$ and has the following property.

$$L_j(t_k) = \delta_{jk} = \begin{cases} 1, & for\ j = k \\ 0, & for\ j \neq k \end{cases} \tag{5}$$

$\delta_{jk}$ is the Kronecker delta. The above property assures that $L_j$ has value of 1 at $t = t_j$ and vanishes at other parameter values. Although Lagrange interpolation is simple, it has the disadvantages that the individual Lagrange polynomials are complicated and depend on the location of the parameter values, and thus, all of them have to be recomputed whenever any one of the parameter values is modified. The degree of Lagrange polynomial is the total number of interpolation nodes minus 1, which becomes undesirably high when the number of interpolation nodes is not low. High degree polynomials have a strong tendency to oscillate [9]. Here is an example to show the undesirable oscillation from Lagrange interpolation. Function $y = 1/(1+x^2)$ is interpolated on $-5 \leq x \leq 5$ at 7 equally spaced nodes along $x$ axis. The 7 interpolation nodes are shown in Fig. 2. The solid curve is the desired curve from the given $y$ function while the dotted curve is the curve from Lagrange interpolation that pass through the 7 interpolation nodes and has degree of 6. The undesirable oscillation is obvious in the Lagrange interpolation curve, especially near the left and right ends. Fig. 3 shows the spline interpolation curve from the same interpolation nodes as Fig. 2. The dotted curve has no oscillation and is very close to the desired solid curve.

In spring synthesis, synthesizers are normally interested in smooth and tight interpolation curves that path through all interpolation nodes in order. Spline interpolation can meet the

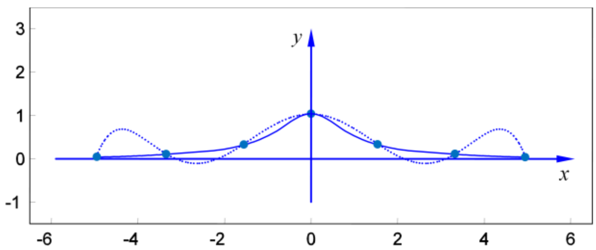


Fig. 2 Lagrange interpolation.

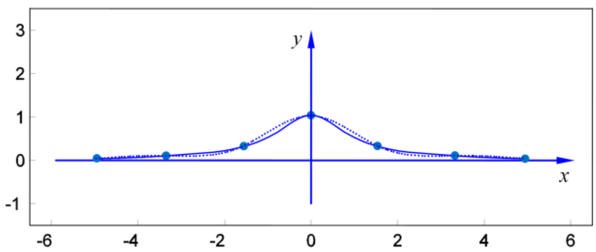


Fig. 3 Spline interpolation

needs. A cubic spline interpolation curve is a set of polynomials of degree 3 that are smoothly connected at given interpolation nodes. The slope and curvature at internal nodes are continuous between two adjacent polynomials. At the two end nodes, their conditions can be chosen differently which include natural end conditions (two end curvatures are set as zero), not-a-knot end conditions (the third derivative is continuous at both the first and last internal nodes) or clamped end conditions (two end slopes are specified).

Given $n$+1 nodes ($P_0$, $P_1$, $\cdots$ $P_n$) with pairwise parameter values $(t_0, t_1, t_2, \cdots, t_n)$, its piecewise cubic spline interpolation is given by $n$ cubic polynomials between each successive pair of nodes as follows [10].

$$S_j(t),\quad t_j \le t \le t_{j+1},\quad j=0,1,\cdots,n-1 \tag{6}$$

$$\begin{aligned} S_j(t) &= M_j\frac{(t_{j+1}-t)^3}{6h_j} + M_{j+1}\frac{(t-t_j)^3}{6h_j} + \\ &+\left(P_j - \frac{M_j h_j^2}{6}\right)\frac{t_{j+1}-t}{h_j} + \left(P_{j+1} - \frac{M_{j+1}h_j}{6}\right)\frac{t-t_j}{h_j} \end{aligned} \tag{7}$$

$$h_j = t_{j+1} - t_j \tag{8}$$

$P_j$ in (7) can be any coordinate of node $j$. (6) and (7) can be used for any of X, Y or Z coordinate. For different coordinates, coefficients in (7) are different. $M_j$ can be solved as follows.

$$AM = D \tag{9}$$

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \mu_1 & 2 & \lambda_1 & \cdots & 0 & 0 & 0 \\ 0 & \mu_2 & 2 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \mu_{n-1} & 2 & \lambda_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \tag{10}$$

$$M = [M_0 \quad M_1 \quad M_2 \quad \cdots \quad M_{n-1} \quad M_n]^T \tag{11}$$

$$D = [0 \quad d_1 \quad d_2 \quad \cdots \quad d_{n-1} \quad 0]^T \tag{12}$$

$$\lambda_j = \frac{h_j}{h_{j-1}+h_j},\quad j=1,\cdots,n-1 \tag{13}$$

$$\mu_j = \frac{h_{j-1}}{h_{j-1}+h_j},\quad j=1,\cdots,n-1 \tag{14}$$

$$d_j = 6\frac{(P_{j+1}-P_j)/h_j - (P_j - P_{j-1})/h_{j-1}}{h_{j-1}+h_j},\quad j=1,\cdots,n-1 \tag{15}$$

The above equations are for natural end conditions. Related equations for other end conditions can be similarly established.

The shape of a connection is defined by its cubic spline curve that passes through all interpolation nodes. The spline curve depends on the locations of the nodes. At each node, one or more size parameters can be added to define the cross-sectional sizes of the connection. (6) to (8) are also applicable to interpolate cross-section sizes. A spline curve with its perpendicular width is called a variable width spline curve in this paper. A connection in a nonlinear spring can then be described by a variable width spline curve. For a 2D variable width spline curve, each node has 3 interpolation parameters: 2 location parameters and 1 width parameter. The 3 parameters can be represented as a circle that is called interpolation circle in the paper. The center of the circle is the location of the node and its diameter represents the width. Then, a variable width spline curve is defined by a set of interpolation circles. Fig. 4 shows 5 ordered interpolation circles. The centers of every two successive circles are connected by a chord. The chord length is used as interpolation parameter $t$. The variable width spline curve shown in Fig. 5 is from the interpolation circles of Fig. 4. Both the center spline curve and the variable width spline curve are smooth in Fig. 5.

The locations of the 5 interpolation circles in Fig. 6 are exactly the same as those in Fig. 4, but two right interpolation circles have larger diameters than their corresponding ones in Fig. 4. The variable width spline curve from these 5 interpolation circles is shown in Fig. 7. Although the center spline curve is smooth and has no cusp, the variable width spline curve is unsmooth and has a cusp, which is undesirable since it is a source of stress concentration. To avoid cusp, the curvature radius of the center spline curve at any point has to be greater than half of the perpendicular width at that point. When the following constraint is satisfied, cusp will not happen.

$$R(t) > 0.5w(t) \tag{16}$$

$$R(t) = \frac{\left[\dot{x}^2(t) + \dot{y}^2(t)\right]^{3/2}}{\left|\dot{x}\ddot{y} - \ddot{x}\dot{y}\right|} \tag{17}$$

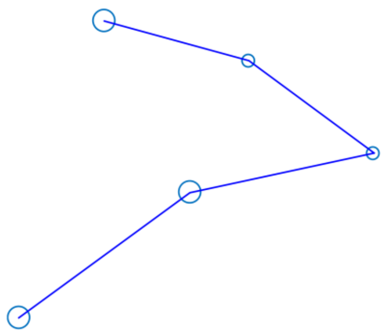

Fig. 4 The interpolation circles of a variable width spline curve.

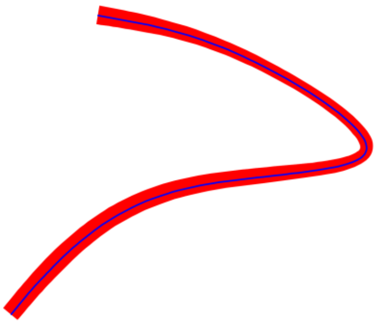

Fig. 5 The variable width spline curve defined by the five interpolation circles

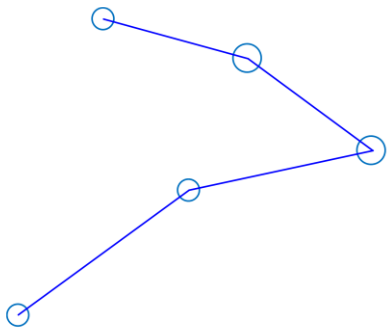

Fig. 6 The interpolation circles with two larger diameters.

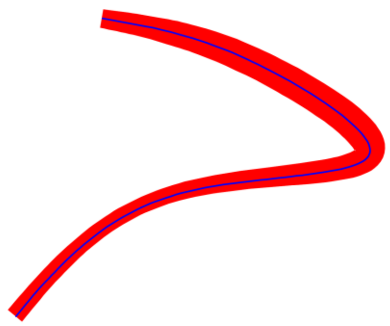

Fig. 7 The variable width spline curve with an unsmooth cusp

In (16), $R(t)$ is the curvature radius and $w(t)$ is the width at point $t$ of the center spline curve.

## III. Optimization of Spring Parameters

The configuration of a nonlinear spring is modeled as one or more variable width spline curves in this paper. Each variable width spline curve is defined by its interpolation circles. The synthesis of a nonlinear spring is thus systematized as optimizing the parameters of the interpolation circles that define the spring.

The control parameters of interpolation circles of a synthesized nonlinear spring are optimized in this paper by using Global Optimization Toolbox of MATLAB [11-12]. Global Optimization Toolbox gives methods that search for optimal solutions to problems that have multiple maxima or minima. Global Search Solver in MATLAB's Global Optimization Toolbox is used in the paper to optimize parameters. This solver is to find a global solution and has an efficient local solver, fmincon.

The performance of a synthesized nonlinear spring is evaluated by finite element analysis software ANSYS [13-14]. Providing the parameters of a set of interpolation circles, the force, displacement and stress of the related nonlinear spring are analyzed by ANSYS. An ANSYS batch file is created in MATLAB on elements, material properties, boundary conditions and input information. ANSYS is then called from MATLAB to execute the batch file. After preprocessor, processor and postprocessor stages, ANSYS generates an output file on the performance information of the synthesized spring. MATLAB reads the ANSYS output file and calculates the objective and constraint functions for optimization. ANSYS Parametric Design Language is used in the paper as a communication tool between MATLAB optimization and ANSYS finite element analysis.



## IV. Synthesis of a Hardening Spring

A hardening spring is a nonlinear spring that gradually increases its spring rate as the spring displacement progresses. The displacement and force ranges of the synthesized spring are from 0 to 20 mm and from 0 to 20 N, respectively. The *F-D* relationship is defined by 5 desired ($D$, $F$) target points: ($D_0$, $F_0$), ($D_1$, $F_1$), ($D_2$, $F_2$), ($D_3$, $F_3$) and ($D_4$, $F_4$). The displacement range of 20 mm is divided into 5 equal-distance points, i.e. $D_j = 5j,\ j = 0, 1, 2, 3, 4$. $F_0$ and $F_4$ correspond to the two end points of the force range, so that $F_0$ = 0 and $F_4$ = 20 N. The 3 internal force points are set to $F_1 = 0.04F_4$, $F_2 = 0.17F_4$ and $F_3 = 0.48F_4$ [2]. The spline curve that interpolates the 5 target points is shown in Fig. 8. The design domain is 100 mm x 100 mm which is shown Fig. 9. The input node is marked by a crossed circle and is at the middle point of the top edge of the design domain with input displacement upward. The potential support nodes are marked by filled circles and are on the bottom edge of the design domain. The material for the spring is engineering plastic with yield strength of 71 MPa and modulus of elasticity of 2200 MPa. The out-of-plane thickness is set at 4 mm. The in-place width is varied from 1.0 mm to 3.0 mm. The spring is symmetric to the vertical line that passes the input node, and is composed of two connections. Each connection is modelled as a variable width spline and is defined by 5 interpolation circles ($P_0$ to $P_4$ in Fig. 9). Only the right variable width spline curve is lettered because of the symmetry. The support node can be chosen from two different locations (corner and middle points), so there are two available solutions. Parameters to be optimized are the locations and diameters of the interpolation circles. The two end circles have fixed locations, but their diameters are variable. Thus, there are totally 11 parameters to be optimized, which are represented as a design variable vector $X$.

$$X = \lfloor w_0 \quad p_{1x} \quad p_{1y} \quad w_1 \quad p_{2x} \quad P_{2x} \quad w_2 \quad P_{3x} \quad P_{3y} \quad w_3 \quad w_4 \rfloor \tag{18}$$

The spring is synthesized to minimize the error between the actual force from the spring and the desired force when a certain displacement is input to the spring. This error is measured by the average deviation at the 4 target displacements ($D_1$ to $D_4$) as follows.

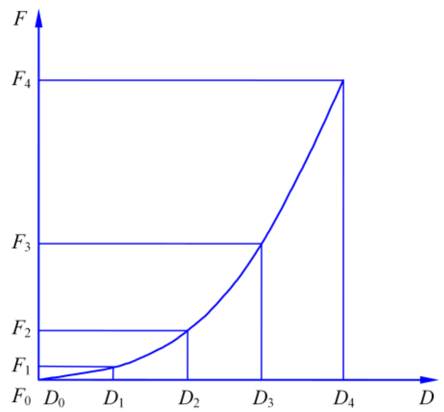


Fig. 8 The *F-D* curve of the synthesized spring.

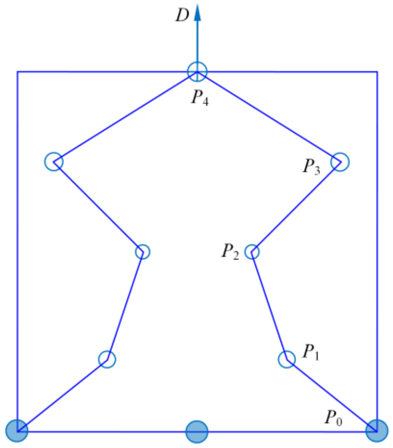


Fig. 9 The design domain of the synthesized spring.

$$FE = \frac{1}{n}\sum_{j=1}^{n}\left|F_{a,j} - F_{d,j}\right| \tag{19}$$

$FE$ is the average force error. $F_{a,j}$ is the actual force generated by the spring when target displacement $D_j$ is input to the spring while $F_{d,j}$ is the desired spring force and equals target force $F_j$.

The synthesis result is shown in Fig. 10 when the two support nodes are at the corners. The design variable vector $X$ for this solution is

$$X = [2.61 \quad 19.89 \quad 13.44 \quad 1.02 \quad 9.66 \quad 56.85 \quad 1.04 \quad 25.39 \quad 77.03 \quad 1.73 \quad 1.59] \tag{20}$$

In this solution, the desired and actual spring forces at 5 targets are: (0, 0), (0.80, 1.57), (3.40, 4.37), (9.60, 9.67) and (20, 20.00). The spline curves that interpolate the two sets of spring forces are shown in Fig. 11. The blue spline is from the desired target forces while the red one interpolates the actual spring forces. When the spring displacement is over 15 mm, the actual spring force is almost the same as the desired force. The maximum stress in the spring is 55.29 MPa, which happens when the spring has displacement of 20 mm.

Fig. 12 shows the undeformed and deformed beam elements of the spring, which is from ANSYS with the input displacement of 20 mm.

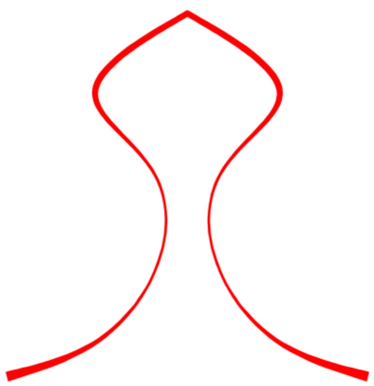

Fig. 10 The spring with corner supports.



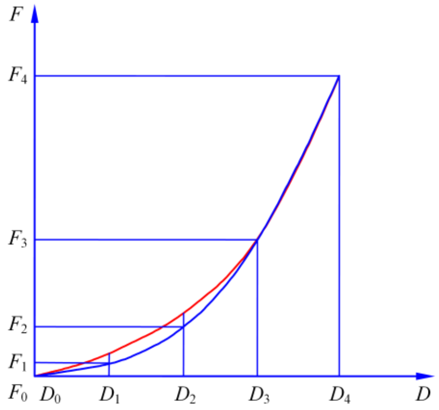


Fig. 11 The desired and actual *F-D* curves of the spring.

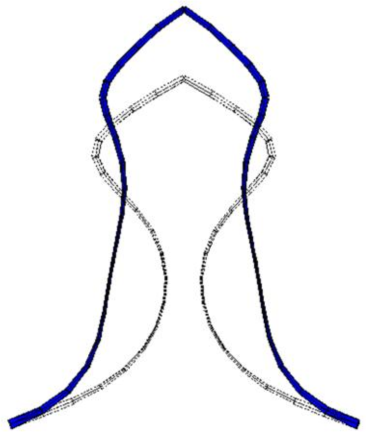

Fig. 12 The deformed spring with corner supports.

When spring force and stress are analyzed in ANSYS, the input displacement of 20 mm is divided into 4 even load steps and geometric nonlinearity command "NLGEOM" is turned on. The spring is discretized into beam elements and modeled by BEAM188 that allows tapered beam cross-sections.

When the two support nodes coincide at the middle bottom point in Fig. 9, the synthesis result is shown in Fig. 13. The design variable vector $X$ for this solution is

$$X = [2.13 \quad 27.20 \quad 32.13 \quad 1.13 \quad 40.55 \quad 50.86 \quad 2.34 \quad 21.60 \quad 71.76 \quad 1.47 \quad 1.48] \tag{21}$$

In this solution, the desired and actual spring forces at 5 targets are: (0, 0), (0.80, 1.98), (3.40, 4.85), (9.60, 9.67) and (20, 20.00). The actual *F-D* curve of this spring solution is close to that in Fig. 11 and is thus not included here. The maximum stress in the spring is 51.35 MPa when the spring has the displacement of 20 mm. The deformed spring of this solution is shown in Fig. 14.

Two alternative spring solutions are shown here that generate close *F-D* functions, but have different locations of support nodes in the design domain. One spring solution can be chosen from them based on the spring application.

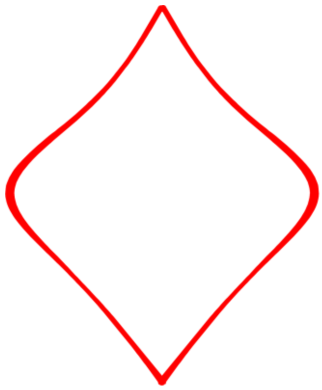

Fig. 13 The spring with middle support.

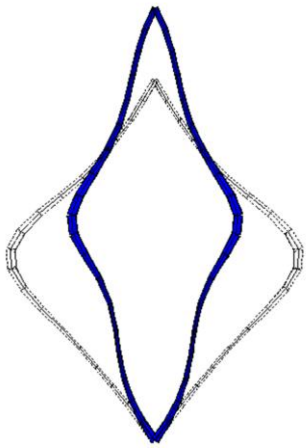

Fig. 14 The deformed spring with middle support.

## V. CONCLUSIONS

A systematic synthesis method of nonlinear springs is presented in the paper. A nonlinear spring is modeled as an elastic piece of material and synthesized as a special compliant mechanism. The topology of a synthesized nonlinear spring is on the connectivity of a network of nodes. The network is formed by input, output, support and intermediate nodes of the spring. The shape of the synthesized spring is on the shape of the node network, which is decided by the locations of the nodes and their connections. The size of the synthesized spring is on the cross-sectional sizes of the node connections. Each connection is modeled in the paper as a variable width spline curve and controlled by its interpolation circles. A spring is then composed of a set of variable width spline curves and controlled by the parameters of interpolation circles. The synthesis of a spring is thus systematized as the optimization of control parameters of interpolation circles.

MATLAB's Global Optimization Toolbox is used in the paper for the optimization of control parameters of synthesized nonlinear springs. The deviation between the desired force-displacement function and the actual force-displacement function is minimized. The deviation is measured by the differences between desired spring forces and actual spring forces under certain spring displacements. The maximum stress in the spring is constrained below the yield strength of the spring material. The spring force and stress of a synthesized spring is analyzed using ANSYS. The communication between MATLAB and ANSYS is based on ANSYS Parametric Design Language. A hardening nonlinear spring is synthesized in the paper to verify the effectiveness of the presented synthesis method.

## REFERENCES

[1] R.L., Norton, Machine Design, Fifth Edition. Upper Saddle River, NJ: Prentice Hall, 2013.

[2] C.V., Jutte, S., Kota, "Design of Nonlinear Springs for Prescribed Load-Displacement Functions," ASME Journal of Mechanical Design, 2008, 130, (081403):1-10.

[3] D.H., Myszka, Machines and Mechanisms, Third Edition. Upper Saddle River, NJ: Prentice Hall, 2005.

[4] L.L., Howell, Compliant Mechanisms. Hoboken, NJ: Wiley, 2001.

[5] L.L., Howell, S.P., Magleby, B.M., Olsen, Handbook of Compliant Mechanisms. Hoboken, NJ: Wiley, 2013.

[6] H., Zhou, K.L., Ting, "Topological Synthesis of Compliant Mechanisms Using Spanning Tree Theory,"ASME Journal of Mechanical Design, 2005, 127, pp:753-759.

[7] H., Zhou, K.L., Ting, "Shape and Size Synthesis of Compliant Mechanisms Using Wide Curve Theory," ASME Journal of Mechanical Design, 2006, 128, pp:551-558.

[8] C.J., Kim, Y.M., Moon, S., Kota, "A Building Block Approach to the Conceptual Synthesis of Compliant Mechanisms Utilizing Compliance and Stiffness Ellipsoids," ASME Journal of Mechanical Design, 2008, 130, (022308):1-11.

[9] J., Hoschek, D., Lasser, Fundamentals of Computer Aided Geometric Design. Wellesley, MA: A K Peters, 1993.

[10] E., Kreyszig, Advanced Engineering Mathematics, Tenth Edition. Hoboken, NJ: Wiley, 2011.

[11] D., Hanselman, B., Littlefield, Mastering MATLAB. Upper Saddle River, NJ: Prentice Hall, 2012.

[12] MathWorks, Global Optimization Toolbox User's Guide. Natick, MA: MathWorks, 2012.

[13] E.H., Dill, The Finite Element Method for Mechanics of Solids with ANSYS Applications. 6000 Broken Sound Parkway, NY: CRC Press, 2012.

[14] ANSYS, ANSYS Parametric Design Language Guide, Canonsburg, PA: ANSYS, 2012.