On the Mathematical Modeling of Slender Biomedical Continuum Robots

Hunter B. Gilbert *

Department of Mechanical and Industrial Engineering, Louisiana State University, Baton Rouge, LA, United States

The passive, mechanical adaptation of slender, deformable robots to their environment, whether the robot be made of hard materials or soft ones, makes them desirable as tools for medical procedures. Their reduced physical compliance can provide a form of embodied intelligence that allows the natural dynamics of interaction between the robot and its environment to guide the evolution of the combined robot-environment system. To design these systems, the problems of analysis, design optimization, control, and motion planning remain of great importance because, in general, the advantages afforded by increased mechanical compliance must be balanced against penalties such as slower dynamics, increased difficulty in the design of control systems, and greater kinematic uncertainty. The models that form the basis of these problems should be reasonably accurate yet not prohibitively expensive to formulate and solve. In this article, the state-of-the-art modeling techniques for continuum robots are reviewed and cast in a common language. Classical theories of mechanics are used to outline formal guidelines for the selection of appropriate degrees of freedom in models of continuum robots, both in terms of number and of quality, for geometrically nonlinear models built from the general family of one-dimensional rod models of continuum mechanics. Consideration is also given to the variety of actuators found in existing designs, the types of interaction that occur between continuum robots and their biomedical environments, the imposition of constraints on degrees of freedom, and to the numerical solution of the family of models under study. Finally, some open problems of modeling are discussed and future challenges are identified.

Keywords: continuum robots, soft robots, dynamics, statics, mechanics, medical robotics

INTRODUCTION

Continuum robots use material deformation to move instead of joints. They may offer a technological solution to some of the difficult challenges of locomotion, perception, and manipulation found in a variety of unstructured and uncertain environments (Robinson and Davies, 1999). Biomedical applications have been a great motivator in the development of a wide variety of continuum and soft robots, ranging from surgery to therapy and other applications involving physical human-robot interaction. The great recent interest in these design paradigms stems from the observation that success in whatever form it is needed may be achieved without having complete control over the motion of a robot or its forces of interaction with the environment. In some cases, this is advantageous simply for reducing the complexity of
engineered systems, and in other cases, performance may be increased beyond what is possible with rigid machines. Several excellent examples of this general principle come from tools of modern medicine. A flexible endoscope can navigate the intestines without a great degree of control over its own shape. The same is true for an intravascular catheter. In these examples, it is the particular combination of geometry and just the right amount of mechanical “softness” that facilitates the completion of the task. Beyond this snake-in-a-pipe approach to navigation, recent research has argued that physical compliance is advantageous in grasping, underwater swimming, robustness to collision, and locomotion on soft terrains where low ground pressure is required. The interested reader is referred to several review articles for a survey of the benefits, applications, challenges, and history of soft and continuum robots (Kim et al., 2013; Burgner-Kahrs et al., 2015; Walker et al., 2016; Cianchetti et al., 2018).

Beyond being under-actuated, the modeler of a continuum manipulator also frequently faces other challenges. Designs are often difficult to separate into “components” since the structure and the actuator may be the same physical body. Actuators based on pneumatics, hydraulics, and composite structures may not be as easily characterized as electric motors. Friction and hysteresis models may be needed to explain observed mechanics, and environments rich with expected contacts may require the solution of contact models based on theories of nonlinear complementarity. Additionally, the standard kinematic descriptions based on the rigid transformations in the special Euclidean group $SE(3)$ are neither the most common approach to

![FIGURE 1](image)

**FIGURE 1** | (A) Concentric tube robots are comprised of hard (metallic) tubes which are precurved and nested inside one another. Rotating and translating the tubes results in motion. (B) Tendon-driven robots use one or more tendons or cables to provide internal actuation forces that bend a flexible, slender rod. (C) Pneumatic soft continuum robots use soft air muscles, which extend or contract with internal air pressure, to create bending in a composite structure. The supports could be hard or soft materials. (D) A fully soft pneumatic gripper uses asymmetry introduced by an inextensible fabric layer and an asymmetric air volume to create four slender fingers which bend to wrap around objects.
solid mechanics nor (necessarily) the most expedient approach to the description of solid continua undergoing deformation. With these considerations, one appreciates why the mathematical modeling of continuum and soft robots can be challenging.

This paper first reviews the state-of-the-art in the mathematical modeling of continuum manipulators having at least one “long” aspect in terms of its shape, which are termed slender in agreement with the mechanics literature. The goal of these models is to describe the dynamics (or statics) to relate actuator variables, other boundary conditions, and sensor measurements to the motion of the robot. The models are generally not concerned with other important aspects of robot design and analysis, such as repeatability, wear, safety, and other factors. For designs made of slender components, the motion of the robot is dominated by bending or beam-like deformations. This classification can be thought of as “arms,” “snakes,” or the individual “fingers” of a multi-fingered hand. Designs composed of individual components having this property are a natural extension, such as concentric tube robots (Mahoney et al., 2018) or multi-backbone continuum robots (Ding et al., 2013). For robots made of softer materials, such as the STIFF-FLOP designs, localized deformations may be complex, yet the dominant behavior is still beam-like (Cianchetti et al., 2014; Fraš et al., 2014). One of the goals of the work is to express the variety of methods encountered in the literature with a common notation. The review motivates a theoretical discussion rooted in the classical theories of solid mechanics. An analysis of the mechanics is used to support recommendations for future modeling efforts, with the conclusion that some choices for the model structure may result in better absolute model accuracy and efficiency (as quantified by the relationship between accuracy and dimensionality).

**REVIEW OF THE STATE OF THE ART**

Table 1 presents the unified nomenclature that will be used throughout this paper. In the discussion of other works, the original nomenclature has been changed to match what is shown. There are three primary considerations in any physics-based approach to modeling of solid continua: the adoption of kinematic hypotheses and coordinates describing the configuration of the body, the application of the laws of mechanics, and the selection of mathematical models that describe the behavior of materials (Sadati et al., 2019). Kinematic hypotheses alone allow the modeler to describe the geometry of the robot, but this alone is insufficient for most purposes because it does not reveal which configurations are possible or likely. The mechanics, which are formulated naturally as partial differential equations, provide the relationships between the kinematic degrees of freedom that indicate which path of configurations will be taken if particular conditions (actuation, environments, etc.) are imposed. Finally, the material models are needed to close the relationship between the kinematic degrees of freedom and the kinetic quantities related by the mechanics.

**Kinematic Descriptions**

The forebears of continuum manipulators are the hyper-redundant robots, defined as those having a large (or infinite in the case of continuum robots) relative degree of redundancy (Chirikjian and Burdick, 1994). In any robot with material deformation which is substantial with regard to the kinematics or dynamics, both the relative degree of redundancy and the degree of under-actuation are theoretically infinite since the configuration space is infinite-dimensional. Here the usual definition of a robot configuration is used: “a complete specification of the location of every point on the robot” (Spong et al., 2006). There have been two primary methods to date of describing the configuration of continuum and soft robots: the curve-based description and the general continuum description.

**The Curve-Based Description**

The state of the art curve-based description is that of the special Cosserat rod (Antman, 2005). Figure 2 depicts the curve, its

---

**TABLE 1 | Nomenclature used in this article.**

| Symbol | Meaning |
|--------|---------|
| $p$    | Position vector of a point with respect to an inertial frame of reference $F_0$ |
| $x$    | Vector $x$ resolved in Cartesian coordinates of frame $F$. The basis is held fixed if a derivative is taken, i.e. if $a = x \cdot d$, and $d$ are the unit vectors of $F$, then $\partial F_a^x = \partial_d x \cdot d$. |
| $F_i$  | Frame of reference $i$. $F_0$ is an inertial frame |
| $s$    | Arc length coordinate |
| $t$    | Time coordinate (may be real time or an arbitrary parameter describing changes in configuration depending on context) |
| $d_1$, $d_2$, $d_3$ | Director vectors of a framed curve |
| $g_0$, $R_0$, $p_0$ | Transformation in SE(3) consisting of rotation operator $R_0 \in SO(3)$ and displacement $p_0$, describing the transformation between $F_0$ and $F_0$ along a framed curve |
| $\gamma$ | A generalized coordinate for the $J^1$ degree of freedom |
| $\partial_\gamma (\cdot) | Partial derivative operator with respect to variable $\gamma$ |
| $u_\alpha$ | Strain variables in the special Cosserat rod description |
| $D$    | Subset of the real line, domain of the arc-length parameter for a rod |
| $\cdot \big|_{\text{bc}}$ | Canonical mapping $\mathbb{R}^3 \rightarrow \text{so}(3) \subset \mathbb{R}^{3 \times 3}$, $a, b = a \times b$ |
| $\alpha i \cdot d_i$ | Summation over repeated indices implied |
| $\tau_j$ | Actuator value $j$ |
| $\xi$, $\gamma$ | Task-space coordinate |
| $q$    | Time derivative of $q$ |
relationship to a solid body, and the quantities that are associated with the curve and the boundary conditions of a mechanical model. The elongated form of many continuum manipulators leads naturally to the concept of the “backbone curve,” which is typically defined to be a time-varying, piecewise differentiable curve in the standard three-dimensional affine Euclidean space equipped with associated vector space. A parametric representation gives the position of a point identified by a spatial parameter \( s \) at time \( t \in \mathbb{R} \) as a vector \( \mathbf{p}_b(s, t) \in \mathbb{R}^3 \) with respect to a specific frame of reference \( \mathcal{F}_0 \) in \( \mathbb{R}^3 \). The differentiability requirement on \( \mathbf{p}_b \) is always at least that the first derivative of \( \mathbf{p}_b \) with respect to \( s \) exists, is piecewise continuous, and is nowhere equal to zero. This condition guarantees that the curve is rectifiable, or in other words has a measurable arc length (Kreyszig, 1991). The curve changes over time, modeling the motion of the robot, and it is presumed to describe the dominant features of the motion of the robot. Since there is no finite set of coordinates that describes every possible curve meeting these requirements, the description of the shape is infinite-dimensional.

The usual type of modeling hypothesis for slender bodies is that other points, which are not located on the backbone, are described by some auxiliary relationship that describes their positions relative to the positions on the backbone. The standard theories from beam mechanics may be adopted for this purpose, in which case the backbone curve may be affixed to the body at the neutral axis of bending. One example is the Euler-Bernoulli hypothesis, which states that sections normal to the backbone remain normal for all deformations. Another is the hypothesis due to Timoshenko stating that normal sections rotate relative to the backbone but remain planar. Standard “warping” theories can be used to couple motion of the points normal to the sections with twisting about the backbone if the sections are not circular. Regardless of these additional hypotheses, the curve is of fundamental importance to the kinematic description.

Explicitly, the body of the robot is identified by the curve through the consideration of a reference configuration \( \mathcal{C}_0 \) of the robot. The backbone curve \( \mathbf{p}_b \) is placed on this reference configuration. The curve is then “framed” by a set of unit vectors \( \mathbf{d}_1(s, t), \mathbf{d}_2(s, t), \) and \( \mathbf{d}_3(s, t) \) termed the director vectors. The first two are chosen to be orthogonal and to span the section of the body at \( s \) which is normal to the curve. The third is taken to complete a right-handed, orthonormal coordinate frame as \( \mathbf{d}_3(s, t) = \mathbf{d}_1(s, t) \times \mathbf{d}_2(s, t) \). In terms of classical differential geometry, \( \mathbf{d}_3 \) is the tangent vector, and \( \mathbf{d}_1 \) and \( \mathbf{d}_2 \) could be selected as the normal and bi-normal vectors from Frenet’s formulas (Kreyszig, 1991). This procedure is problematic for general curves since torsion may be undefined, but many other alternative framings of the curve are possible which do not suffer this problem (Bishop, 1975). The backbone position and unit vectors together describe a local reference frame \( \mathcal{F}(s, t) \) for each point along the curve. The unit vectors equivalently define a spatiotemporal field of rotation operators \( \mathbf{R}_b(s, t) \in \mathcal{SO}(3) \). The rotation field can be represented by matrices (Rucker and Webster, 2011), quaternions (Boyer et al., 2020), or any other suitable representation. Together with the position vector, a spatiotemporal field of transformations \( \mathbf{g}_b(s, t) \in \mathcal{SE}(3) \) is defined by.

\[
\mathbf{g}_b(s, t) = \{\mathbf{R}_b(s, t), \mathbf{p}_b(s, t)\}
\]

The vectors \( \mathbf{u}(s, t) = \xi d_1 \) and \( \mathbf{v}(s, t) = \eta d_2 \) are termed the “strain variables.” They describe deformation of the body and are invariant under rigid transformations. The vector \( \mathbf{u} \) has been widely called the “curvature” vector in the robotics literature, but this may be misleading since it is not generally the curvature of the deformed backbone curve. The term “flexural strain” is preferred for \( \mathbf{u}_1 \) and \( \mathbf{u}_2 \), and “torsional strain” for \( \mathbf{u}_3 \). The variables \( \eta \) and \( \varepsilon \) are called the shear strains, and \( \gamma \) is the dilatation. The change in length or “extension” of the backbone curve is characterized by \( \partial_s \mathbf{p} = \mathbf{v} \cdot \mathbf{v} \). The strain variables are related to the framed curve by the following relationships.

\[
\partial_s \mathbf{p}_b(s, t) = \mathbf{v}(s, t), \quad \partial_s \mathbf{d}_k(s, t) = \mathbf{u}(s, t) \times \mathbf{d}_k(s, t).
\]

Finally, the vectors \( \partial_s \mathbf{p}_b(s, t) \) and \( \omega(s, t) \) represent the linear and angular velocity of the backbone curve and director vectors. The angular velocity satisfies the equation \( \partial_s \mathbf{d}_k(s, t) = \omega(s, t) \times \mathbf{d}_k(s, t) \). The four functions \( \mathbf{u}, \tilde{\mathbf{v}}, \partial_s \mathbf{p}_b, \) and \( \omega \) are not independent; they must satisfy \( \partial_s \mathbf{u} = \partial_s \mathbf{v} = \omega \times \mathbf{u} \). In the reference configuration, the flexure strains have non-zero values \( \mathbf{u}_0(\tilde{s}) \) if the backbone is not a straight. Generally, \( \mathbf{u}_0(\tilde{s}) = \mathbf{d}_3(\tilde{s}) \), but other choices are possible.

The General Continuum Description

The second approach to describing the configuration of continuum robots is to make as few prior kinematic hypotheses on the configuration as possible. The traditional description of a three-dimensional continuum in solid mechanics is used in this case. In this approach a reference configuration \( \mathcal{C}_0 \) is identified by their position vector relative to a frame of reference \( \mathcal{F}_0 \). Three coordinates \( \mathbf{X} \in \mathbb{R}^3 \) identify the position of each point in the body via a one-to-one, differentiable vector-valued function \( \mathbf{P}(\mathbf{X}) \). If \( \mathbf{X} \) is chosen as the Cartesian coordinates with respect to \( \mathcal{F}_0 \), then this function and its inverse are trivial. The final locations of the points are described by \( \mathbf{p}(\mathbf{X}, t) \). In some cases, it is useful to define a displacement field \( \mathbf{U} \) as follows.

\[
\mathbf{p}(\mathbf{X}, t) = \mathbf{P}(\mathbf{X}) + \mathbf{U}(\mathbf{X}, t).
\]

The amount of stretching can be quantified by the deformation gradient, defined by

---

1 There are additional considerations for this placement in the case of dynamical models, discussed below.
The deformation gradient straightforwardly describes the local changes in length (amount of stretching) and therefore plays a major role in the definition of strain measures. Note also that the curve-based description of the configuration, together with the classical Euler-Bernoulli hypothesis, can be placed into this more general framework using $X = (s, X_2, X_3)$ and $p(X, t) = p_b(s, t) + X_2d_2(s, t) + X_3d_3(s, t)$ (Antman, 2005).

**Perspective on Discretization and Configuration Spaces**

There are two perspectives that one might take when describing the kinematics or mechanics of continua. In the first perspective, the model consists of a (possibly nonlinear) PDE, a domain on which the PDE applies, and boundary conditions in the form of constraints or measurements. The robot’s state space consists of the dependent variables related by the PDE. The state space is therefore a particular Cartesian product space that might involve, in general, both finite-dimensional spaces and infinite-dimensional function spaces. In the process of computing a numerical solution to a model, any part of the state that belongs to an infinite-dimensional space must be approximated by a finite set of coordinates in $\mathbb{R}^n$, but the choice of coordinates does not need to be of great concern to the modeller. This perspective has been taken by numerous authors for general continuum manipulators (Trivedi et al., 2008; Till et al., 2019), concentric tube robots (Dupont et al., 2010; Rucker et al., 2010; Gilbert et al., 2016), parallel continuum robots (Black et al., 2018), and bioinspired locomotion by snakes and worms (Boyer et al., 2012). The modeller hopes that any approximation error is small enough to be ignored, and error-controlled numerical methods may provide some assurances. This first perspective is the natural one if, for example, the modeller selects an error-controlled, automatic step-size numerical integrator like the Dormand-Prince Runge Kutta pair to approximate the solution to a differential equation with a spatially distributed independent variable. The benefit to this perspective is that questions of convergence may generally be avoided. However, there are two main disadvantages: first, there is a relative paucity of tools available if the problem is not expressed with respect to a single independent spatial variable; second, the degrees of freedom chosen by automatic numerical methods may be unknowable in advance and may vary between model solutions, making it difficult to apply algorithms built on spaces like $\mathbb{R}^n$ or on manifolds where coordinate charts are available.

In a second perspective, the equations of an infinite-dimensional model are explicitly discretized through a suitable method such as the finite element method or a finite difference method (Renda et al., 2014; Back et al., 2015; Gilbert and Godage, 2019) or via a spectral method involving a “modal” decomposition (Chirikjian and Burdick, 1994; Godage et al., 2015; Chen et al., 2020). In this perspective, the modeller takes control over the discretization and fixes the dimensionality of the resulting model. One is free to take the perspective that a new model has been created that is not necessarily subordinate in any way to the infinite-dimensional model. In other words, the infinite dimensional dependent variables, ODEs, and/or PDEs, were only a steppingstone to the finite-dimensional model. The dimension may be varied according to a model hyper-parameter $N$, and often one wishes that as $N \to \infty$, the solutions to the sequence of fixed-dimensional models approach the solution to a corresponding infinite-dimensional model.

The second perspective is the standard one in generally accepted theories of robot kinematics and dynamics, in which the goal is to find a suitable coordinate set that describes the displacement field $u(X, t)$ that takes a material point located at initial position $P$ to its final position $p = p + u$. With rigid link manipulators, the space is partitioned into non-intersecting bodies indexed by number $i \in \mathbb{Z}^+$ and equipped with local coordinate frames, and then the machinery of $SE(3)$ is used to associate each body with its own displacement field expressed in terms of one of the coordinate transformations $^0T_i \in SE(3)$ representing the transformation between the base frame 0 and the frame of the $i^{th}$ body. For serial, rigid-link robots, the choice of finite dimensional coordinates parameterizing the displacement field is usually one of two conventions, the Denavit-Hartenberg convention (Denavit and Hartenberg, 1955) or the twist coordinate system and product-of-exponentials formula (Brockett, 1984).

For continuum and soft robots, neither the perspective (finite vs. infinite-dimensional) nor the approach to discretization (choice of coordinates) appears to be standardized. In some cases, restrictive assumptions do allow a set of finite coordinates that uniquely specify the configuration of a continuum robot. For example, Bretl and McCarthy showed that for the Kirchoff rod with no external loading, a configuration space isomorphic to $\mathbb{R}^6$ can be selected, corresponding physically to a known internal force and moment at the same location in space as the known orientation of the rod (Bretl and McCarthy, 2014). A similar result is known for coordinates of the configuration space of concentric tube robots without any external loads (Gilbert et al., 2016). The general principle is a basic result on initial value problems. If the mechanics of the system can be modeled by a system of $n$ first-order initial value problems, 

$$\frac{\partial y}{\partial t} = F(s, y), \quad y(s_0, t) = y_0(t),$$

with $F$ uniformly Lipschitz in $y$ and continuous in $s$, then the solutions are uniquely determined by $y_0$ (Schaeffer and Cain, 2016). Therefore, if all state information of the robot is contained in the functions $y_i(s, t)$, then it is clear that $y_0$ is a suitable set of coordinates for the configuration space of the robot. For curve-based models $y_0$ usually belongs to a space of the form $SE(3') \times \mathbb{R}^n$.

However, with less restrictive assumptions, low-dimensional configuration spaces are not generally found. Such is the case for parallel continuum robots (Black et al., 2018), for growing robots (Greer et al., 2019), or soft robotic hands (Schlagenhauf et al., 2018). It is in general impossible to find a “minimal” set of coordinates for the C-space of any continuum manipulator when
the locations and nature of external loads or contacts are a-priori unknown and when these loads cause substantial changes in the robot shape. The subsections that follow describe a variety of methods that have been used to mathematically represent the configurations of continuum robots.

Spectral Methods
Spectral methods were some of the earliest described methods for the kinematic modeling of backbone curves. In this method, the configuration is represented by a finite number of coordinates \( q(t) \in \mathbb{R}^N \) by assuming that some kinematic quantity is described by a truncated sum of “modal” shape functions \( \phi_i(s) \) in a manner analogous to a Fourier series. The general form is to have a scalar quantity \( S_j \) represented as

\[
S_j(s, t) = \sum_{i=1}^{N} a_{ij}(q(t)) \phi_i(s).
\]

The function \( a_{ij} \) may be simply an index into the vector \( q \) pulling out one of the components, or it may be a more complicated relationship. The mode shapes are generally selected among one of the standard families such as trigonometric functions \( \sin(k_i s \pi s) \) and \( \cos(k_i s \pi s) \) for a series of values \( k_i \in \mathbb{R} \) (directly analogous to a truncated Fourier series), the standard monomials \( \{1, s, s^2, \ldots\} \), the Legendre polynomials, Chebyshev polynomials, etc. (Chirikjian and Burdick, 1994; Zhang and Simaan, 2013; Chen et al., 2020). In general, to be classified as a spectral method, the mode functions should have global support rather than local support, which leads to the element-based methods described below.

There is a great deal of freedom within this approach. For example, the tangent vector \( d_1 \) can be expressed in spherical angles \( \theta(s, t) \) and \( \phi(s, t) \), and then \( S_1 = \theta \) and \( S_2 = \phi \), and \( v = (0, 0, 1) \) completes the kinematic description (Chirikjian and Burdick, 1994). \( S_j \) could also be chosen directly as a component of the displacement field of the backbone curve (Godage et al., 2015). These methods are extrinsic because they seek to approximate kinematic quantities as measured by the observer in the inertial frame \( \mathcal{F}_0 \). Parameterizations also possible which represent the strain variables \( u(s, t) \) and \( v(s, t) \) measured by an observer in the local frame \( \mathcal{F}_1(s, t) \) (Boyer et al., 2020). When coupled with a collocation method used to determine \( u \), it was shown that the Magnus expansion can be used to efficiently recover the position and orientation field (Orekhov and Simaan, 2020).

In the context of continuum robots, to the best of the author’s knowledge, the spectral methods have only been applied in conjunction with the curve-based descriptions discussed in The Curve-Based Description and not more for general continuum descriptions.

Element-Based Methods: PCC
The element-based methods, in contrast to the spectral methods, break up the problem spatially into adjacent sub-domains and attempt to model the kinematics on each sub-domain using a simpler hypothesis. This procedure can be carried out for both the curve-based description and the general continuum description. Many authors have adopted the kinematic hypothesis that the backbone curve is a sequence of circular arcs which are concatenated by imposing tangency conditions. There is a natural extension of this idea to piecewise helical curves. This approximation is termed the “piecewise constant curvature” (PCC) method, and many continuum robots have even been designed to exhibit deformation of this kind, at least in the absence of external loads (Webster and Jones, 2010). For example, multi-backbone robots and tendon-driven robots will adopt, with actuation, shapes very close to circular arcs with appropriate design decisions (Camarillo et al., 2008; Xu and Simaan 2008). On the other hand, even gravitational loading may cause more flexible robots to adopt shapes more complex than a single circular arc (Trivedi et al., 2008).

Within the curve-based framework described above, the “standard” PCC hypothesis including inextensibility and shearlessness is equivalent to a partitioning of the domain into \( m \) elements \( \Gamma_e = \{s_{i-1}, s_i\} \) with \( D = U_{i=1}^m \Gamma_e \) and an approximation of the flexural-torsional strain as the following sum:

\[
u(s, t) = \sum_{e=1}^{m} u_e(t) \chi_e(s), \quad \chi_e(s) = \begin{cases} 1 & s \in \Gamma_e, \\ 0 & s \notin \Gamma_e \end{cases}
\]

A similar approach is called the piecewise constant strain (PCS) method and extends the definition to include an approximating sum representing \( v \) (Renda et al., 2016). The vectors are resolved in Cartesian coordinates of the frame \( \mathcal{F}(s, t) \) and it is these body-frame coordinates that are presumed constant over the element.

The major advantage of the PCC method is that the extrinsic variables \( p_b \) and \( d_e \) (equivalently \( R_e \)) are easily computed using an explicit recursion presuming that a single boundary condition on the pose of the curve, \( g(s_0, t) = g_0(t) \), is known. On element \( \Gamma_e \), the transformation from \( \mathcal{F}_0 \) to \( \mathcal{F}(s, t) \) is given by

\[
g(s, t) = g(s_{s-1}, t) \exp((s - s_{s-1}) \xi_e(t)),
\]

\[
\xi_e(t) = \begin{bmatrix} \mathcal{F}(s) \chi_e(t) & \mathcal{E}(t) \end{bmatrix} = \begin{bmatrix} u_e(t) & v_e(t) \end{bmatrix}
\]

The mapping \( se(3) \rightarrow SE(3) \) is continuous and differentiable, including at the element \( 0 \in se(3) \), and it may be computed in closed form, below in vector form but originally discovered by Euler and Rodrigues (Rodrigues, 1840; Cheng and Gupta, 1989).

\[
\exp\begin{bmatrix} \omega \cdot v \\ 0 \end{bmatrix} = \begin{bmatrix} R(\omega) & p(\omega, v) \\ 0 & 1 \end{bmatrix}
\]

\[
R(\omega) = I + C(\omega)\omega + A(\omega)\omega^2 + B(\omega)(\omega \cdot v)\omega.
\]

The coefficient functions are as follows, with \( \text{sinc}(x) \) the unnormalized sinc function, which is available in the standard libraries of many programming languages.

\[
A(\omega) = \frac{1}{2} \text{sinc}^2\left(\frac{\|\omega\|}{2}\right), \quad B(\omega) = \frac{\|\omega\| - \sin(\|\omega\|)}{\|\omega\|^3}, \quad C(\omega) = \text{sinc}(\|\omega\|).
\]

In IEEE-754 double-precision arithmetic, the author has found that \( B(\omega) \) may be accurately computed by a truncated Taylor series \( B(\omega) \approx 1/6 - \|\omega\|^2/120 + \|\omega\|^4/5040 \) if \( \|\omega\| < 1 \times 10^{-4} \).
The PCC and/or PCS methods are the simplest explicit and consistent discretization methods for a framed curve which interpolate the intrinsic (strain) variables rather than the extrinsic (position and orientation) variables. In other words, given a curve with bounded flexural, torsional, shear, and extensional strains, the error between the curve and its PCC or PCS approximation (both in terms of $d_k$ and in terms of $p_b$) shrinks as the number of elements increases. This property is important because it means that the PCC/PCS framework can describe robots with practically any backbone curve if the domain is broken into enough elements. Figure 3 depicts a single isotropic rod under a combined twisting and bending moment, forming a helix.

The results of a PCC approximation under a linear elastic material law and subject to the virtual work principle discussed in Considerations for Kinematic Hypotheses below are shown in Figure 4. The flexural strain variables of the PCC model approach those of the exact solution as the number of elements increases. The example also demonstrates an important distinction between material flexural strain and the usual notion of curvature of a shape. The exact solution under the end load depicted in Figure 3 is indeed a helix, which is a “constant curvature” shape, yet the flexural strain variables, when resolved in components of the material frame $F(s, t)$ or any fixed frame of reference, are not constant functions of the arc length. For this reason and others discussed in Considerations for Kinematic Hypotheses, the simplicity of the PCC approach relative to others may not outweigh the drawbacks.

**Element-Based Methods: Higher-Order One-Dimensional**

The discretization of the curve into elements can also be accomplished with higher-order schemes than the piecewise constant strain approach. In general, as with the PCC approach if a boundary condition $g(s_0, t)$ is provided at a single location along the backbone curve, then the C-space can be defined by approximating the strain variables on each element with an expansion having more terms than the PCS approach (Boyer et al., 2020).
\[ \mathcal{F}(s,t) \begin{bmatrix} u(s,t) \\ v(s,t) \end{bmatrix} = \sum_{c=1}^{mNc} \phi_i(z_c)q_{ij}(t), \quad z_c = \frac{s - s_i}{s_k - s_{k-1}}. \]

The union operator is abused to mean here that the element-local terms contained within each argument of the union are "assigned" to provide the evaluation on that element. The coordinate \( z_c \) is a normalized element-local length coordinate identifying the cross section within the element. If the model is to contain flexure, torsion, shear, and extension, then a polynomial expansion of order \( p \) in each component entails \( N_G = 6(p+1) \) degrees of freedom within each element. Continuity conditions at the element boundaries may and/or boundary conditions may reduce the number of total degrees of freedom from \( mN_c \) as implied by the formula.

Other representations for the C-space based on segmentation into elements are possible but have not been widely pursued within the robotics community. Certainly, the more widely adopted approach within the finite element and computational mechanics communities has been to make the primary kinematic hypothesis at the level of \( p \) and/or \( d_i \) (Simo and Vu-Quoc, 1986), and this has also been studied in the context of continuum robots (Sadati et al., 2019). This is the approach taken, for example, by the ANSYS simulation software in handling nonlinear beam elements via the BEAM188 and BEAM189 elements. These methods require interpolation on the rotation group \( SO(3) \), and some care is required to ensure that the formulation is invariant to rigid-body transformations (Crisfield and Jelenic, 1999). One major advantage of strain interpolation is that it is frame-invariant directly by construction; the major disadvantage is that the calculation of inertial forces is greatly complicated by the spatial coupling of the degrees of freedom.

**Element-Based Methods: General Continuum**

More general finite-element descriptions have also been used to model soft and continuum robots. In this case, the degrees of freedom \( q_i(t) \) directly interpolate the position field \( p(X, t) \) over the three-dimensional domain of the body. Using typical first-order (linear) interpolants, the degrees of freedom are the Cartesian coordinates of the nodes of the mesh that breaks the body into discrete volumes. Direct nodal position discretization using finite elements can be used for closed-loop control using a dimensionality reduction scheme based on projection (Bieze et al., 2018). It has also been shown that high-order FEM models with an order reduction method involving fitting to PCC kinematics is effective (Runge et al., 2017). Finally, it has been demonstrated that general 3D FEM with model order reduction based on the Proper Orthogonal Decomposition can produce models amenable to dynamic closed-loop control (Katzschmann et al., 2019b).

**Direct Nodal Discretization**

Closely related to the element-based methods are those based on direct discretization of the variables. Differential operators in the mechanics can be replaced by their equivalent finite-difference operators to form algebraic equations directly, operating on the values of field variables specified at discrete spatial locations \( s_i \) for the curve-based approach. A finite difference scheme applied directly to the geometrically exact Cosserat equations has been demonstrated for both the planar case and the spatial case (Hasanzadeh and Janabi-Sharifi, 2014; Gilbert and Godage, 2019; Wang et al., 2021) and described for concentric tube robots (Webster and Rucker, 2009). Finite-difference methods were also used with direct spatial discretization to model a soft underwater arm driven by cables (Renda et al., 2014).

**Pseudo-Rigid Body Methods**

The pseudo-rigid body methods replace the continuum with an approximating rigid linkage. If the curve is broken into a sequence of chords with rotational joints at the nodes joining the chords, then this is equivalent to a spatial “lumping” of the flexural strains into a discrete point via the use of the Dirac delta distribution (Chirikjian and Burdick, 1991; Greigarn et al., 2019).

\[ \mathcal{F}(s,t) u(s,t) = \sum_{s,t}^m q_i(t) \delta(s-s_i) n_i. \]

A universal joint is the result if two orthogonal axes \( n_i \) and \( n_{i+1} \) are placed with \( s_i = s_{i+1} \) with both axes normal to the backbone curve. Three orthogonal axes create a spherical joint.

It has been shown that the kinematics of tip-loaded cantilever beams can be modeled adequately by a serial 3R mechanism (Su, 2009). Other PRB models have been created for modeling of catheters (Ganji and Janabi-Sharifi, 2009), tendon-driven continuum manipulators for minimally invasive surgery (Penning and Zinn, 2014), and MRI-actuated catheters (Greigarn et al., 2017). A 6-DOF PRB segment model has also been proposed (Venkiteswaran et al., 2019). An equivalence has also been shown between the coordinates of a PCC model and a suitably defined pseudo-rigid body model, indicating that PRB model segments with RPPR kinematics can be used to describe the same configuration space as PCC models (Katzschmann et al., 2019a).

**Initial Value Problem Concepts**

There are additionally a variety of other methods of analysis and computation which do not explicitly select the degrees of freedom in the kinematic description. In these methods, the unknowns are conceptually left as unknown functions, and numerical methods are used which automatically select the degrees of freedom used to represent the unknown functions, usually via an error estimation and control algorithm.

These methods have been used when the problem is re-cast as a one-dimensional boundary value problem with split boundary conditions.

\[ \partial_y f(s,y) + G_a(y(0,t), t_j) = 0, \quad G_b(y(L,t), t_j) = 0. \]

Solutions can then be provided by numerical codes which automatically determine the degrees of freedom used to approximate the function \( y(s, t_j) \) for each discrete value of \( t_j \). For continuum robots these methods have been demonstrated via collocation (Webster and Rucker, 2009) and shooting methods (Till et al., 2015; Mauzé et al., 2020) using numerical tools that approximate \( y(s, t_j) \) via piecewise polynomials. It has also been
shown recently that the dynamics problem for a wide variety of architectures based on single or multiple Cosserat rod sub-models can be cast as a shooting problem on an ODE once the time derivatives have been discretized using finite differences (Till et al., 2019).

**Differential Kinematics for Strain-Variable Hypotheses**

It is often necessary to calculate a manipulator “Jacobian field” based on the curve parametrization, and if the generalized coordinates are defined to interpolate the strain variables, this field is not trivial to calculate.

\[
\begin{bmatrix}
\frac{\partial p}{\partial t}
\end{bmatrix}
(s, t) = J_q(s, t) \frac{\partial q}{\partial t} = \begin{bmatrix}
J_p \\
J_w
\end{bmatrix} \frac{\partial q}{\partial t}.
\] (2)

Letting \( J_i \) be the column multiplied by \( \partial q_i / \partial t \), then the column can be calculated from the following differential relationships:

\[
\frac{\partial J_i}{\partial t} = \frac{\partial q}{\partial t} \\
\frac{\partial J_i}{\partial t} = \frac{\partial u}{\partial t} + u \times J_i.
\]

One must take care when the interpolation is carried out on the strain variables in coordinates of the local frame \( F_i(s, t) \). If desired, the coordinates in the body frame representations \( F_i(s, t) \) and \( F_i(s, t) \) may be calculated from the appropriate representation of these equations in the moving frame (Rucker and Webster, 2011).

\[
\frac{\partial}{\partial t} \left( F_i(s, t) \right) = -F_i(s, t) v + F_i(s, t) + \frac{\partial}{\partial q} \left( F_i(s, t) \right) + \frac{\partial}{\partial J} F_i(s, t) \times F_i(s, t) v
\]

From a known boundary condition where \( J_{ps}(0, t) = 0 \) and \( J_{ps}(0, t) = 0 \), the solution to these equations can be expressed in closed form as the solution to a time linearized system.

\[
\begin{bmatrix}
J_{ps} \\
J_{ps}
\end{bmatrix}
(s) = \int_0^s \exp \left( -\int_0^t \frac{\partial q}{\partial t} F_i(s, t) v \right) \frac{\partial q}{\partial t} F_i(s, t) u \, dr,
\]

\[
\frac{\partial}{\partial t} F_i(s, t) = F_i(s, t) v.
\]

**Mechanics**

Regardless of how the shape of a robot is described, the principles of classical mechanics are frequently used to describe the relationships between the model’s degrees of freedom, the internal stresses, and any imposed boundary conditions which may include external forces, imposed positions or orientations of parts of the robot, contact conditions. The robot’s actuators may generally be modeled in one of two ways: either they are described as constraints (a form of boundary condition) or as sources of internal stress.

**The Equations of Motion for the Special Theory of Cosserat Rods**

In the curve-based description, the equations of motion of the special theory of Cosserat rods serve as the strong form differential equations governing the mechanics (Antman, 2005).

\[
\begin{align*}
\frac{\partial n + f}{\partial t} = \rho A \frac{\partial n}{\partial t} + \rho I_k \frac{\partial n}{\partial t} d_k \\
\frac{\partial m + \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} + \frac{\partial r}{\partial z}}{\partial t} = \rho I_k \frac{\partial n}{\partial t} d_k + \frac{\partial (p \cdot \omega)}{\partial t}.
\end{align*}
\] (3)

The sum from \( k = 1 \) to 3 is implied over the terms involving \( I_k \) and \( d_k \). The variables \( n(s, t) \) and \( m(s, t) \) are the internal force and the internal moment, which are interpreted as the resultant force and resultant moment of the stress acting on section \( s \). In the case of a slowly accelerating body, which is typical in many biomedical applications, a quasistatic approximation may be used, in which all terms on the right-hand side are neglected (Burgner-Kahrs et al., 2015). \( f \) and \( \ell \) are externally applied forces and moments. Applied concentrated forces and moments require the Dirac \( \delta \) distribution to express in this formulation.

In the case of a model which allows freedom in all the strain variables, \( m \) and \( n \) are algebraically related to the kinematic variables through a suitable material constitutive law. On the other hand, in the shear-less and extension-less model, \( n \) is a basic unknown and is equivalent to a Lagrange multiplier which enforces the constraint \( v(s, t) = v_0(s) \). The parameter \( \rho A \) is the mass density (expressed per unit length) of the cross-section. \( \rho f \) is the mass moment of inertia (per unit length) of the section, which makes \( p \cdot \omega \) the angular momentum (per unit length) calculated about the mass center of the section. The three parameters \( \rho k \) account for linear momentum density of the cross section caused by angular velocity of the curve. The author is not aware of any works in the robotics literature for which this term has been nonzero; if the backbone curve is chosen to pass through the mass centers of the cross sections, then \( \rho f_k = 0 \) and the equations simplify considerably. However, it is noteworthy that this may in general result in the curve failing to pass through the cross-section centroids (if multiple materials are used) or it may be impossible to satisfy this requirement exactly if a single curve is used to model a body with complex geometric features.

**The Equations of Motion for Pseudo-Rigid Body Models**

With the PRB-type models, the equations of motion are exactly those of a classical multibody dynamical system with scleronomous, holonomic constraints. These equations are commonly given as follows (Murray et al., 2017).

\[
M(q) \ddot{q} + C(q, \dot{q})q + N(q, \dot{q}) = B(\tau).
\]

The right-hand side contains the non-conservative generalized forces associated with actuation and any other forces; since the robots are underactuated there are generally many more rows in this equation than actuator variables \( \tau_i \). Also, it is noteworthy that the inertial forces are not trivial to calculate since the motion of the continuum body is not the same as the motion of the rigid-link approximation. Some assumptions about how the continuum “tracks” the rigid-link approximation as it moves is needed. One approach is to match the centers of mass of chords along the curves of a PCC model with centers of mass of the links in the rigid link model (Della Santina et al., 2018).
The Equations of Motion for General Deformable Bodies

The dynamic equilibrium conditions of classical continuum mechanics serve as the defining relationship for general three-dimensional finite element models of soft and continuum robots. Rarely are these equations encountered explicitly in the literature on continuum robots, with most authors preferring to state the result after the strong form equations have been converted to the weak form and integrated. The resulting equations, incorporating constraint forces, are of the following form (Goury et al., 2021).

\[ M(q)\ddot{q} + \mathbf{F}(q, \dot{q}) + \mathbf{G}(q) = \mathbf{H}^T \lambda \]

The form of this equation is directly analogous to the classical form of the dynamical equations for rigid multibody systems. \( M(q) \ddot{q} \) accounts for the inertial forces, \( \mathbf{F}(q, \dot{q}) \) accounts for the internal forces produced by deformation of the material, and \( \mathbf{G}(q) \) accounts for gravitational forces. The matrix \( \mathbf{H} \) is associated with the constraints and boundary conditions and encodes the effect of the boundary and actuation forces contained in the vector \( \lambda \). The details of the construction procedure for this equation are out of the scope of this paper.

Projection via D’Alembert’s Principle

In the case of the curve-based models using either the PCS or higher-order models, the equations can be projected onto the degrees of freedom of the model using Galerkin’s principle, probably better known among mechanical engineers as the principle of virtual work (Greenwood, 1988). The method is also equivalent in results to Kane’s method of virtual power (Kane and Levinson, 1983; Rone and Ben-Tzvi, 2014). Because the backbone curve descriptions for the PCC, PCS, and higher order strain variable interpolants are described by independent degrees of freedom \( q \in \mathbb{R}^N \), a direct projection of the equilibrium equations is possible via D’Alembert’s principle, which amounts to an integration over the equations of motion.

\[
\int_0^L \left[ (\mathbf{F}(s,t) + \mathbf{F}'(s,t)) \cdot \dot{\mathbf{y}}(s,t) + (\mathbf{M}(s,t) + \mathbf{M}'(s,t)) \cdot \beta_j(s,t) \right] ds = \mathbf{Q}_{j,nc} \]

\[
\mathbf{F}(s,t) = -\partial_m \mathbf{n}(s,t), \quad \mathbf{F}'(s,t) = \rho \mathbf{A} \partial_\mathbf{p} \mathbf{p}(s,t) + \rho \mathbf{l}_m \partial_\mathbf{t} \mathbf{d}_m(s,t) \]

\[
\mathbf{M}(s,t) = -\partial_\mathbf{m} \mathbf{m} + \partial_\mathbf{p} \times \mathbf{n}, \quad \mathbf{M}'(s,t) = \rho \mathbf{l}_m \mathbf{d}_m \times \partial_\mathbf{p} + \partial_\mathbf{t} (\rho \mathbf{l}_m \mathbf{w}) \]

\[
\mathbf{Q}_{j,nc} = \int_0^L \mathbf{f}(s,t) \cdot \dot{\mathbf{y}}(s,t) + \ell(s,t) \cdot \beta_j(s,t) ds. \tag{5}
\]

The velocity coefficient function and angular velocity coefficient function are defined as

\[
\mathbf{y}(s,t) = \partial_\mathbf{q} \mathbf{p}(s,t) = \partial_\mathbf{p} \partial_\mathbf{q} \mathbf{p}, \quad \beta_j(s,t) = \partial_\mathbf{q} \mathbf{w}.
\]

The velocity coefficients are the “Jacobian field” satisfying the relation Eq. 2.

Since the time derivatives of the momentum density and angular momentum density, \( \partial_\mathbf{p} \partial_\mathbf{q} \mathbf{p} \) and \( \partial_\mathbf{t} (\rho \mathbf{l}_m \mathbf{w}) \), can be written as linear functions of the \( \partial_\mathbf{q} \mathbf{q} \), the equations of motion are linear in the accelerations of the generalized coordinates, as expected. In the case of the PCC/PCS kinematic description, the derivatives \( \partial_\mathbf{n} \) and \( \partial_\mathbf{m} \), if resolved in \( \mathcal{F}(s,t) \), are zero except at the element boundaries. The equations may be integrated by parts into a form which shows the conjugacy of \( \mathbf{n} \) and \( \mathbf{v} \) and the conjugacy of \( \mathbf{m} \) and \( \mathbf{u} \).

\[
\int_0^L \partial_\mathbf{n} \cdot \mathbf{v} \; ds = \left[ \mathbf{n} \cdot \mathbf{v} \right]_0^L - \int_0^L \mathbf{n} \cdot \partial_\mathbf{q} \mathbf{v} \; ds
\]

\[
\int_0^L \partial_\mathbf{m} \cdot \beta_j \; ds = \left[ \mathbf{m} \cdot \beta_j \right]_0^L - \int_0^L \mathbf{m} \cdot \left( \partial_\mathbf{q} \mathbf{u} + \mathbf{u} \times \beta_j \right) ds.
\]

In the local frame, the equations take the following forms.

\[
\int_0^L \partial_{\mathbf{n}} \cdot \mathbf{v} \; ds = \left[ \mathbf{n} \cdot \mathbf{v} \right]_0^L - \int_0^L \mathbf{n} \cdot \left( \partial_\mathbf{q} \mathbf{v} \right) ds + \int_0^L \mathbf{F}(s,t) \cdot \mathbf{v} ds
\]

\[
\int_0^L \partial_{\mathbf{m} \cdot \beta_j} \; ds = \left[ \mathbf{m} \cdot \beta_j \right]_0^L - \int_0^L \mathbf{m} \cdot \left( \partial_\mathbf{q} \mathbf{u} + \mathbf{u} \times \beta_j \right) ds.
\]

Note also that if \( \mathbf{n}(s,t) \) is constant over \( s \), the first integral is trivially zero.

Finally, note that if more than one rod-like body is present, then a sum over the bodies takes place in Eq. 5. Explicit constraints between the bodies may be handled via the method of Lagrange multipliers.

Learning-Based Approaches

Learning-based approaches, which are also sometimes referred to as “model-free” approaches, may be able to describe the relationships between the actuator inputs and observable outputs such as the end-effector motion without recourse to physical parameters and the laws of mechanics. These models usually serve a complementary purpose to those based on physical first principles. Since they require training data from a real robot or from another simulation model, they may be used for on-line control, inverse and forward kinematics, or for off-line analysis and testing of other algorithms such as for navigation and control. The a-priori prediction of behaviors from only design data is generally not possible to date using only learning-based methods.

A variety of purely kinematic approaches have been proposed. One learning approach uses an on-line estimation of the Jacobian matrix relating the time derivatives of the actuation variables \( \partial_\mathbf{q} \mathbf{r} \) to the time derivatives \( \partial_\mathbf{q} \mathbf{p}(L, t) \) and \( \partial_\mathbf{q} \mathbf{d}_m(L, t) \), and it has been shown that this approach works for both position control and hybrid position/force control when appropriate sensing is available in hardware (Yip and Camarillo, 2014; Yip and Camarillo, 2016). Since the method requires no information about the robot or the environment a-priori, it enables control in complex scenarios, where highly complex physics-based models may have poorly observable parameters or states. It has also been shown that inverse kinematics for continuum robots may be approximated by a multilayer perceptron network (George et al., 2017; Grassmann et al., 2018; Lai
et al., 2019), with multi-agent reinforcement learning (Ansari et al., 2016), with K-nearest neighbors and Gaussian mixture regression (Chen and Lau, 2016), and with deep reinforcement learning (Satheeshbabu et al., 2019). For reconfigurable robots subject to varying loads, it has been shown that classification of the load state using long short-term memory networks can substantially improve open-loop kinematic control (Nicolai et al., 2020). For flexible catheters, a combination of a support vector machine classifier and a fully connect neural network regressor were demonstrated achieving sub-millimeter trajectory tracking errors (Jolaei et al., 2020). Learning may also play an important role in proprioception for continuum robots with distributed sensing, where additional sensors beyond actuator-collocated ones are available for measuring the robot shape (Truby et al., 2020).

It has also been shown that dynamic models may be learned. Under a state observation of the form \( x = F(q) \) provided by sensors, where it is presumed that the dimension of \( x \) and \( q \) are the same and that \( F \) is invertible, the dynamics of the system can be posed as a one-to-one mapping \( (\tau, x, \partial x) \rightarrow \partial_t x \). On a real or simulated robot, this map representing the dynamics of the observables of the system can be approximated in discrete time via recurrent neural network (Thuruthel et al., 2017). Note that the kinematic relationships under static conditions are obviously also contained in this trained map for all points satisfying \((\tau, x, 0) \rightarrow 0\). A similar approach using deep neural networks has also been demonstrated (Gillespie et al., 2016). Data-driven system identification based on the Koopman operator theory has led to control-oriented dynamic models amenable to model-predictive control (Bruder et al., 2020). Autoregressive with exogenous input (ARX) and nonlinear autoregressive with exogenous input (NARX) models have been studied for a single-section tendon-driven continuum robot, with the conclusion that NARX provides advantages in modeled end-effector position accuracy (Parvaresh and Moosavian, 2019). For closed-loop dynamic control, the direct reinforcement learning of a control policy which learns the mapping from current robot states, previous robot states, and desired end-effector position to the appropriate control action is possible (Thuruthel et al., 2019).

There are also learning-based approaches to control which do not explicitly construct kinematic or dynamic models. One such approach is based on direct learning from demonstration in the actuator space, which was successfully demonstrated on a tendon-driven continuum manipulator (Xu et al., 2016). Learning can form a part of a.

**Actuator Models**

Actuators in continuum and soft robots have been classified as either extrinsic, in which case the actuators are not a part of the deformable body, or intrinsic, in which case the actuators are an integral part of the deformable body. Examples of the former include tendons, the boundary conditions placed on concentric tube robots. Examples of the latter include soft pneumatic muscles (Walker et al., 2016).

The actuators may be modeled (very generally) as relationships between the actuation variables, generalized forces, and the dynamic state of the robot consisting of \( q \) and \( \partial_t q \).

\[ G_i(\tau, q, \partial_t q, Q_{nc}) = 0. \]

However, the nature of the model may change depending on the exact form of \( G_i \). If \( G_i \) involves only \( \tau \) and \( q \), then it is exactly in the form of a holonomic constraint. In general, it may not be simple to find a reduced set of independent coordinates satisfying the constraint, and a Lagrange multiplier technique may be required to enforce it. On the other hand, if \( G_i \) can be inverted to find \( Q_{nc} = \mathcal{G}_i(\tau, q, \partial_t q) \), then the actuation can be directly coupled to the equations of motion. Which of these two views of actuation is the more natural one depends on the characteristics of the particular actuator(s) and sensor(s) chosen.

A first example is the model of a fiber-reinforced elastic actuator, in which \( V \) is the enclosed fluid volume of the actuator, \( \tau = P \) is the fluid pressure, and \( J_V \) is the Jacobian matrix relating the changes in the generalized coordinates to the change in volume of the fluid (Sedal et al., 2021). Then, the principle of virtual work indicates that

\[
\delta W = \tau \delta V = \tau J_V(q) \delta q = Q_{nc} \delta q \]

Another explicit example is found in the case of a tendon-driven robot. If enough support for the tendon is provided, a reasonable model for the points occupied by the tendon is a continuous curve described by \( p_i(s, t) = p(s, t) + a(s) \) with \( a(s) = a_1(s) d_1(s, t) + a_2(s) d_2(s, t) \) (Rucker and Webster, 2011). For the sake of simplicity, restrict the tendon to a planar path with \( a_2(s) = 0 \). The tendon length can then be calculated as an integral functional involving the deformation gradient evaluated along the tendon path using the curve-based kinematic hypotheses:

\[
\ell_i(q) = \int_0^1 d\ell_i \quad d\ell_i = ds^2(d_s + \partial_s a) F_i^T F_i (d_s + \partial_s a), \quad F_i = \frac{\partial p_i}{\partial P_i}.
\]

If the tendons are not fully constrained, other models for \( \ell_i(q) \) may be more appropriate (Rao et al., 2021). What is noteworthy about either length formulation is that the nonconservative generalized forces do not naturally appear. If the tendon lengths are a known quantity, the actuator model is a holonomic constraint on the generalized coordinates. The problem can be treated via the method of Lagrange multipliers. The Lagrange multiplier will be exactly the tendon tension, and the principle of virtual work can be used to reveal the exact form of the terms in \( Q_{nc} \) corresponding to the Lagrange multiplier.

\[
\delta W = \tau \delta \ell_i = \tau J_V(q) \delta q = Q_{nc} \delta q, \quad J_V(q) = \partial q \ell_i(q).
\]

Therefore, the effect of the tendon alone (not considering any frictional forces) must be

\[ Q_{nc} = J_V(q) \tau. \]

Note that the causal form in which the tendon tensions are known is “easier” to handle since no additional equations must be added. The causal form involving known tendon lengths requires
the addition of the nonlinear length constraints Eq. 6 to the
equation set and the tension becomes an algebraic unknown
along with the accelerations, forming a nonlinear differential-
algebraic system in the dynamic case or a nonlinear algebraic
system in the quasistatic case. The need to solve a DAE system
disappears if the tendon is considered a spring element, since then
the force is determined as a function of the difference between
$\ell(t, q)$ and the tendon displacement input $d$.

The resulting model form as a set of ordinary differential
equations or differential algebraic equations is shown for a variety
of common continuum robot actuators in Table 2.

### Materials

The kinematic hypotheses and mechanics models must be
augmented by constitutive laws (material models) to complete
the model of a continuum robot. For quasistatic models, the
choice is usually between linear elasticity and other hyperelastic
material models. For dynamic models, an additional choice of
damping or friction laws is generally required to produce realistic
responses.

#### Linear Elasticity

In the case of quasistatic models, a common assumption in the
literature has been to assume a Hookean (linear) material
response. In this case, if one assumes that the backbone curve
goes through the neutral axis of bending, the following constitutive
laws apply:

$$ F_{m} = \frac{1}{E I_{m}} (F_{t} u_{t} - F_{s} u_{s}) $$

$$ F_{n} = \frac{1}{G A} (F_{t} v_{t} - F_{s} v_{s}) $$

The matrices $K_{m}$ and $K_{n}$ may be calculated from the classical
Euler-Bernoulli or Timoshenko beam theories and the entries are the
flexural and torsional rigidities and shear and extension
rigidities, respectively. The explicit relationships follow below
(Antman, 2005).

$$ m_{s}(s, t) = (E I_{m}) (s) [u_{t}(s, t) - u_{t0}(s)] $$

Note that bending about $d_{1}$ and $d_{2}$ are characterized by the
second moments of area and the Young’s modulus $E$, while the
torsional moment is related to the torsional strain variable by a
coefficient $D$ solving the St. Venant torsion problem. Only in
the case of isotropic rods with circular cross section is this equal
to the usual shear modulus $G$ times the polar moment of area
$I_{33}$. Formulas for a wide variety of cross sections that are
uniform over $s$ have been tabulated (Roark et al., 2002). The
Timoshenko model for shear and elongation adds the following
relationships.

$$ n_{s} = (GA) (s) v_{sl} $$

$$ n_{3} = (EA) (s) [v_{s} - 1] $$

### Hyperelastic Material Models

Many other hyperelastic models are possible choices, such as
Yeoh, neo-Hookean, Gent, Ogden, and Mooney-Rivlin (He et al.,
2018; Shiva et al., 2019; Zhang et al., 2019; Antonelli et al., 2020;
Baccioti and Tarantino, 2021; Zhao et al., 2021). Although in
general one may expect that these more complex material models
should offer improved model accuracy, it has been shown
recently that, at least for some robot designs, a linear stress-
strain relationship may be more than adequate (Shiva et al.,
2019). Any hyperelastic law can be represented within the Cosserat rod
framework as a strain energy density function $W$.

$$ W = W(F_{s}(s, t) u_{s}, F_{t}(s, t) v_{t}) $$

$$ F_{s}(s, t) m = \partial_{s} W $$

$$ F_{t}(s, t) n = \partial_{v} W $$

The details of these calculations for each of the respective
hyperelastic models is omitted for the sake of brevity and can be
found in the cited references.

#### Damping and Friction

The introduction of dissipative mechanisms is generally
necessary to encourage numerical stability in dynamic models
and to produce realistic dynamic responses. Additionally, in some
cases static friction plays a significant role in determining the
quasistatic solutions, such as in tendon-driven catheters
(Jung et al., 2014). Viscous damping may be introduced via the Kelvin-
Voigt material model, which extends the linear elastic models to
include rate-dependence in the stress-strain relationship
(Gilbert and Godage, 2019; Mustaza et al., 2019).

In the curve-based framework, the Kelvin-Voigt law takes the
following form (Linn et al., 2013):

$$ F_{s}(s, t) m = F_{s}(s, t) K_{bt}(F_{s}(s, t) u - F_{s}(s) u_{s}) + F_{s}(s, t) B_{bt} \partial_{s}(F_{s}(s, t) u) $$

$$ F_{t}(s, t) n = F_{t}(s, t) K_{ae}(F_{t}(s, t) v - F_{t}(s) v_{s}) + F_{t}(s, t) B_{ae} \partial_{v}(F_{t}(s, t) v). $$

The matrices $K_{bt}$ and $B_{bt}$ are related by time constants referred
to as the extensional retardation time constant, $\tau_{e} = \eta E / E$, and the
shear retardation time constant, $\tau_{s} = \eta G$, with $\eta$ the
the “extensional viscosity” and $\eta$ the shear viscosity.

$$ F_{s}(s, t) B_{bt} = F_{t}(s, t) K_{bt} \cdot \diag(\tau_{e}, \tau_{e}, \tau_{e}) $$

$$ F_{t}(s, t) B_{ae} = F_{t}(s, t) K_{ae} \cdot \diag(\tau_{e}, \tau_{e}, \tau_{e}) $$

Static friction models have also been considered for concentric
tube robots (Lock and Dupont, 2011), tendon-driven continuum
robots (Li et al., 2020), and continuum robots having sheathed
tendons or multiple actuated backbones (Roy et al., 2017).

### DISCUSSION

The wide variety of modeling choices that have been described
offer the modeler an almost paralyzing array of choices. In the

---

**TABLE 2** Model form as an ODE or DAE system based on actuator type, assuming a single rod model architecture for the model.

| Actuator input                | Model form |
|------------------------------|------------|
| Inextensible tendon length   | DAE        |
| Extensible tendon length     | ODE        |
| Tendon force                 | ODE        |
| Pneumatic pressure           | ODE        |
| Hydraulic pressure (incompressible fluid) | ODE        |
| Hydraulic volume (incompressible fluid) | DAE        |

Frontiers in Robotics and AI | www.frontiersin.org  12   October 2021 | Volume 8 | Article 732643
subsections that follow, several questions are posed. The available evidence from the literature as well as analyses guided by classical theories of mechanics are used to discuss these questions and to provide guidance during the initial stages of selecting modeling approaches.

**Considerations for Kinematic Hypotheses**

The literature on modeling of continuum and soft robots suggests that errors in kinematic models, quantified by the absolute tip positioning error as a percentage of the overall root length, are typically on the order of a few percent. Therefore, there may be little benefit to increasing the order of a spectral method or to further subdividing the domain in an element-based method once the absolute accuracy with respect to the true solution reaches this point. In the sections that follow, analysis and recommendations for kinematic hypotheses which are derived from consideration of the mechanics of bending are offered. Table 3 provides a summary of the recommendations in terms of increasing either the number of elements or the order of the interpolation (assuming that $u$ is the interpolated variable). Figure 5 depicts the decisions leading to the various types of models that have been discussed.

**Considerations for Cantilevered Concentrated Loadings**

For continuum robots which are soft enough to exhibit substantial compliance to environmental loads (for example those that may be presented by contact with human anatomy), one of the first considerations for modeling should be consistency with the requirements for accurately modeling cantilevered, concentrated loads.

Let the Cosserat rod equations be recast in terms of the angle of the tangent vector and the load and deformation fixed to the plane defined by $d_1(0,t)$ and $d_1(0,t)$, let the boundary condition $g(0,t)$ be fixed, and let the load be concentrated at $s = L$ and modeled by $F = P\delta(s - L)d_1(0,t)$ for the scalar force magnitude $P$. Furthermore, assume the material response is linear and that the robot is inextensible. These restrictions simplify the problem and result in the following dimensionless boundary value problem governing the angle $\theta$, which represents the tangent vector:

$$d_2(s, t) = \cos(\theta)d_2(0, t) + \sin(\theta)d_1(0, t)$$

$$\partial_2 \theta + \lambda \cos(\theta) = 0, \quad \xi = \frac{x}{L}, \quad \lambda = \frac{PL^2}{EI} \theta(0, t) = 0, \quad \partial_2 \theta(1, t) = 0.$$

The boundary value problem has a known solution:

$$\theta(\xi, t) = 2 \sin^{-1}(k \sin(K(k) - (1 - \xi) \sqrt{\lambda}; k^2)) - \frac{\pi}{2}$$

The quantity $k$ is a scalar that may be found by Newton-Raphson iteration on the following equation, which is implied by the boundary condition $\theta(0) = 0$

$$k \sin(K(k) - \sqrt{\lambda}; k^2) = \sqrt{2}/2.$$

The PRB models have the attractive property that they map the problem back into the domain of traditional robotic manipulators, with the obvious advantage that all the tools and knowledge that have been developed in that context (in general, restricted to underactuated mechanisms) now apply to the continuum robot. In the traditional PRB models, the inertia properties are lumped into the links formed by the model, and the stiffness and damping properties are lumped into the joints between links. This lumping introduces error, but it has been shown that optimization of the parameters of the rigid body model can lead to accurate mechanical responses for both cantilevered transverse loads and for applied or internal moments (Chen et al., 2011). Given that the optimal 3R planar PRB model has three degrees of freedom, it is a fair comparison to place the model in competition with other three-DOF models. Here we consider the following three sets of potential kinematic hypotheses and matching constitutive laws and compare them with each other and with the exact solution. Without loss of generality, let $L = 1$ and $EI = 1$.

Piecewise constant curvature:

$$u_{\text{PCC}} = \sum_{i=1}^{3} u_i \mathcal{X}_i(s) d_2, \quad \mathcal{X}_i(s) = \begin{cases} 1 & s \in \Gamma_i, \\ 0 & s \not\in \Gamma_i. \end{cases}$$

$$D = \left[0, \frac{1}{3}\right] \cup \left[\frac{1}{3}, \frac{2}{3}\right] \cup \left[\frac{2}{3}, 1\right] m(s) = EI u(s).$$

**PRB:**

$$u_{\text{PRB}} = \sum_{i=1}^{3} q_i \delta(s - s_i) d_2, \quad s_1 = 0.125, \quad s_2 = 0.475, \quad s_3 = 0.863,$n

$$D = [0, 1], \quad m(s_1) = 3.25, \quad m(s_2) = 2.84, \quad m(s_3) = 2.95.$$

**Spectral:**

$$u_5 = \sum_{i=1}^{3} q_i s^{i-1} d_2, \quad D = [0, 1].$$

$$m(s) = EI u(s)$$

How well does each of the strategies perform when given 3-DOF to capture the deformation? The answer is depicted in...
Figure 6, showing tip position error in percent of robot length versus the dimensionless cantilevered load index $\lambda$. For cantilevered loads, a single spectral element which is quadratic in $u$ is a far better choice than either a 3-element PCC model or a 3R PRB model. If nonzero shear forces are expected, the PCC model seems to have little in its favor; the Jacobian for the PRB
model is simpler to calculate, meaning that the statics equations in Eq. 5 are easier to formulate, and the tip position is predicted more accurately, which also implies that the overall structural stiffness is more accurate for the PRB model than for the PCC model. The Jacobian for the spectral model, unlike the other two, cannot be obtained in an exact closed form.

The results imply that the typical piecewise constant-curvature assumption used in the development of geometrically nonlinear models for robots is a poor choice from the perspective of mechanics whenever a concentrated external load is present and is expected to produce internal shear forces which are transverse to \( \mathbf{v} \). In summary, if point loads are present on the robot, a linear interpolant of internal moment (equivalently curvature) is necessary to accurately capture the static equilibrium configurations for unrefined elements even in the small deflection case, and degrees of freedom are better spent on increasing the order of the interpolants than on increasing the number of elements.

Considerations for Non-Uniform Flexural Rigidity

In the case of non-uniform flexural rigidity, element refinement is more effective at reducing approximation error than increases in order. This conclusion is easily justified by the observation that if \( K(s) \) is a linear function, say for example \( 1 + as \), then in the simplest planar case with a constant internal moment one would be tasked to find another polynomial function \( k(s) \) such that \( k(s)K(s) = C \) for some constant \( C \). But this is obviously impossible, because \( k = C/K \) is a rational function, not a polynomial, and the Maclaurin series at \( s = 0 \) has a finite region of convergence. In the example case, the expansion is \( C \cdot \sum (-1)^{l}a^{l}s^{l} \). The series does not converge unless \( |s| < 1/|a| \) and as \( s \) approaches this upper bound, the number of terms in the series required to obtain convergence to a fixed tolerance increases without bound. Element refinement, on the other hand, has exactly the effect of reducing \( |s| \), ensuring convergence. For this reason, single-element, spectral-type methods are not recommended as a first choice if non-uniform flexural rigidity is present.

Considerations for Uniformly Distributed Loads

Distributed loads may act on biomedical continuum robots. The most obvious of these loads is a gravitational force distributed along the length of the robot. Other common forces may include buoyancy forces, electric forces, magnetic forces, and aerodynamic and hydrodynamic forces. The simplest possible model of a distributed load is a uniform one that is applied normal to the body of a robot which is initially in a straight configuration. In this case, the solution to the linearized Euler-Bernoulli equation is in general a fourth-order polynomial in position. The shear force is a linear function of arc length and the internal moment (and hence curvature in the linear elastic case) is quadratic. If the shape is discretized at the level of angle, the discretization should be cubic to accommodate a uniform load.

Considerations for Elastic Environmental Contact

For continuum robots in contact with soft bodies such as the soft tissues of the human anatomy, the contact might be well-described, at least in the region of contact, by a model like the linear elastic foundation model. For small deflections, the linearized Euler-Bernoulli model with a linear elastic foundation is modeled by the following differential equation (see Figure 7A).

\[
(EIy_{ss})_{ss} = -ky.
\]

The homogeneous solution to this equation has the following form.

\[
y_h(s) = \exp(-\beta s)\left[C_1 \sin(\beta s) + C_2 \cos(\beta s)\right] + \exp(\beta s)\left[C_3 \sin(\beta s) + C_4 \cos(\beta s)\right].
\]

The constant \( \beta \) depends only on the properties of the beam and the foundation.

\[
\beta^2 = \sqrt{\frac{k}{4EI}}.
\]

To what degree of accuracy does a polynomial shape function (assuming the small-deflection case) approximate \( y_h \)?

To answer this question, one should find the best approximation of \( y_h \) under a particular norm on \( L^2[0, \ell] \). Here we select the 2-norm and study the approximation error for 3rd, 4th, and 5th order polynomials. Since \( \beta \) is related to the ratio of stiffnesses \( k \) and \( EI \), and has dimension Length \(^{-2}\), we restrict the range of the dimensionless group \( \ell \beta \) to \( (0.1, 10) \). This range is consistent with the idea of compliance matching as a form of embodied intelligence in biomedical continuum robots. Note that as \( \ell \beta \to 0 \), the solution \( x_4 \) approaches a constant, which is easy to interpolate. As \( \ell \beta \to \infty \), the elastic foundation is becoming infinitely stiff relative to the body of the robot, modeling a hard contact. In this case, the internal forces and moments and the resulting deformations become strongly localized, and a point load may be a more appropriate model for the contact than an elastic foundation.
The physical solutions to the equation decay away from the application of a point load. Therefore, we restrict the approximation problem to the consideration of the two functional forms that follow on a domain \[0, 1\] for \(\beta \in (0.1, 10)\):

\[x_1 = \exp(-\beta s) \sin(\beta s), \quad x_2 = \exp(-\beta s) \cos(\beta s)\]

See Figure 7B for examples with \(\beta = 3\). Errors for polynomials \(p_1\) and \(p_2\) approximating \(x_1\) and \(x_2\), respectively, are shown in Figure 4 as the maximum absolute errors.

\[
e_i = \frac{\max_s |p_i(s) - x_i(s)|}{\max_s |x_i(s)|}
\]

For a single 5th order polynomial in shape, the maximum absolute error in approximating either \(x_1\) or \(x_2\) remains below 1% if \(\ell \beta < 4.3\) (Figure 8). As \(\ell \beta\) increases beyond this value, the polynomial approximations to \(x_1\) and \(x_2\) begin to oscillate with increasing maximum error.

To put this in a practical perspective, a typical colonoscope has a linearized flexural rigidity of \(EI \approx 0.02 \text{ N m}^2\) (Wehrmeyer et al., 1998). Soft tissues may have an elastic foundation stiffness of approximately 4 kN/m² (Asadian et al., 2011). This results in \(\beta = 15\) and therefore a hypothesis which is 5th order in position (3rd order in strain variables) should not have elements longer than approximately 0.28 m. Note that for a hypothesis that is linear in \(s\) for the strain variables, the length requirement would drop to approximately 96 mm, and for PCC elements, the length would drop to only 38 mm. For a spatial robot model that is inextensible and un-shearable and is 1 m long, this would result in a PCC model with approximately 66 degrees of freedom (27 elements at 3 DOF/element), a linear strain variable model with approximately 48 degrees of freedom (4 elements at 12 DOF/element). Therefore, if environmental contacts are soft and distributed over a long length, there is a strong incentive to develop models with higher-order strain variable hypotheses.

**Considerations for Numerical Methods Solution Multiplicity**

In general, the problem defined by Eq. 5 together with any constraints is a nonlinear algebraic problem, even if linear material models are used. This is either a consequence of the nonlinear geometry, which shows up in any finite-strain relationship between the strain variables and the position and orientation of the body, or a consequence of nonlinear material behavior, or both. In special cases, the problem may become linear; for example, if the actuators and generalized forces are related linearly, linear constitutive laws are used, and no external loads are present. For nonlinear static problems, the Newton-Raphson method and trust-region methods like the Levenberg-Marquardt method generally work well, but the modeler must be cautious of the possibility of solution multiplicity.

In other words, a function \(q = f(\mathbf{r})\) does not always exist because there may be two or more values of \(q\) which satisfy the equilibrium conditions given \(\mathbf{r}\). This solution multiplicity is accompanied by a singular tangent stiffness matrix for some value of \(q\) and possibly a tangent stiffness matrix with negative eigenvalues, as is the case for so-called “negative-stiffness mechanisms” (Platus, 1999). The coupling between kinematics and mechanics means that it is not always safe to assume the existence of a “forward kinematic mapping” which computes the
C-space coordinates from the actuator variables and then the task-space variables from the C-space coordinates. Consider the case in which Eq. 5 is of the form $F(q, \tau, Q_{nc}) = 0$ where $Q_{nc}$ includes only those generalized forces which are not algebraically related with $q$ and $\tau$. Then a perturbation analysis yields the C-space Jacobians with respect to $\tau$ and $Q_{nc}$

$$\delta F = \partial_q F \delta q + \partial_\tau F \delta \tau + \partial_{Q_{nc}} F \delta Q_{nc} = 0 \tag{5}$$

$$\delta q = -\left(\partial_q F\right)^{-1} \left(\partial_\tau F \delta \tau + \partial_{Q_{nc}} F \delta Q_{nc}\right).$$

It is evidently at configurations with singular $\partial_q F$ where multiple solutions may arise. This is one reason that quasistatic “resolved-rate” or continuation-type methods may fail to converge; dynamic models do not suffer this problem since the accelerations are resolved.

Time Stepping

For time stepping, explicit ode integrators can become prohibitively computationally expensive. This is a consequence of the fact that unresolved vibrational modes (as defined for linear test problems) become unstable using explicit methods. Implicit integrators and those designed for solving stiff ODEs and DAEs, such as the trapezoidal method or the backwards difference formulae, are preferable. Energy-preserving integrators have the benefit that the damping behavior is caused entirely by the material model, ensuring repeatable dynamic behavior with different time steps.

Current and Future Challenges in Modeling Generalizability and Re-Usability

Despite the growing body of evidence that models built on the foundation of the Cosserat rod equations are an adequate description of many continuum robots, one challenge that still faces practitioners is a lack of standardized tools to build new model simulation codes. For rigid robots, a wide variety of domain-specific modeling languages are available and permit concise descriptions within an easy-to-use interface to build new models. One example of this is the Universal Robot Description Format and Gazebo simulator within the Robot Operating System, but there are many others presently available including Simulink/Simscape, Dymola, and other Modelica-language based tools such as OpenModelica (Brück et al., 2002; Fritzon et al., 2006; Miller and Wendlandt, 2011;Sucan and Kay, 2019). To enable the widespread re-use of validated modeling components, a library of reusable “model building blocks” for continuum robots should be designed. Some important capabilities of such a library would be the following:

- Coupling of curve-based models to rigid multibody models.
- Coupling of curve-based models and general finite element models.
- Incorporation of common actuator models.
- Incorporation of common constraints (length, concentricity, no-penetration, selective inextensibility/strong anisotropy, revolute and prismatic joints, etc.)
- User-selected switching between dynamic and quasi-static model generation.

For biomedical continuum robots in particular, models which couple to mechanical models of human anatomy are needed. Coupling of state-of-the-art models for continuum robots or their direct incorporation with real-time finite element codes using GPU acceleration is a promising approach (Allard et al., 2007; Duriez and Bieze, 2017).

Novel Kinematic Hypotheses

There is a great deal of freedom in element-based kinematic hypotheses which has yet to be explored. One interesting avenue is the use of a shared or constrained DOF between elements. The motivation for this idea is that for dynamic models, time stepping is sometimes restricted or difficult for “stiff” problems having many eigenvalues. The equations of motion for solid continua are wave equations, which means that if many elements are stacked end-to-end, acoustic waves (axial compression and tension) and twist waves (torsional waves) through the structure may be resolved by the model. For most robotics applications, these modes are likely to be irrelevant, and constraining the problem so that they do not exist in the model may improve computational performance. The elimination of twist waves in elastic rod models was previously considered by an energy minimization argument (Bergou et al., 2008).

Furthermore, adaptive kinematic hypotheses based on pre-defined, switchable degrees of freedom that permit local, automatic refinement of the model may allow greatly improved computational efficiency in problems involving a-priori unknown environmental interactions or constraints. This will permit, for example, a single high-order element to describe the deformation in free-space, while local refinement can take place where a catheter contacts a vessel wall, a robotic endoscopic system contacts the colon, or where multi-fingered hands contact an object to manipulate it.

Learning

Within the context of continuum and soft robotics, data-driven methods have begun to demonstrate strong utility. For example, Long Short Term Memory networks can capture hysteresis in pneumatically actuated catheters (Wu et al., 2021), and offline simulation of first-principles models can be used to learn reduced-order models using the snapshot-based proper orthogonal decomposition, resulting in new models suitable for real-time control and other applications requiring fast computation (Goury and Duriez, 2018; Katzschmann et al., 2019b). The continued development of learning methods enabling low-DOF representations will be an important future area of research.

There are also interesting opportunities for learning that amalgamate first-principles models with data-driven model “correctors,” or which use constrained learning techniques to identify models which are topologically like a curve-based model. One possibility is to use a low-DOF curve-based model capturing some of the behavior and to introduce a nonconservative generalized force $Q_{nc}$ which is learned from observed data to close the gap between simulation and reality. Learning-based methods which are constrained to obey fundamental principles are another emerging area of research, such as learning the
Lagrangian or Hamiltonian function of systems directly from data (Ahmadi et al., 2018; Lutter et al., 2018).

Dynamic Model Validation

Although many dynamic models have been proposed, the validation of these models is currently lacking. There are many opportunities for rigorous evaluation and comparison of models with experimentally obtained data. The best and strongest form of model validation would be to instrument real robots with enough sensors to measure all the quantities appearing in Eq. 5 or the equivalent formulations for PRB and general continuum models, and to calculate the model residuals over conditions ranging over static, low-acceleration, and high-acceleration (e.g. sudden contact) regimes. This is clearly a challenging experimental task that may require state reconstruction and many sensors just to measure the configuration trajectory \( q(t) \). Other options for validation may include comparison of standard test signal response characteristics (e.g. rise time, percent overshoot, settling time, steady-state error, and oscillation period) in response to both actuator inputs and environmental perturbations.

There are also many other interesting questions that can be asked and answered which are quantitative in a different sense, but which may be even more aligned with the spirit of soft and continuum robotics theory. For example, a model and simulated controller could be used to predict the success or failure of the navigation of a robotic catheter through tortuous vasculature parameterized by some measure of “tortuosity,” and then the classification error could be assessed via experiment matching the simulations.

CONCLUSION

Continuum robots offer solutions to problems in biomedical applications which may not be solvable by traditional robotics technologies. With these new robots came the need for new models. A wide variety of physics-based and learning-based approaches to the modeling of continuum manipulators—both those made of hard materials and soft materials—are now available to the roboticist who needs them. This can lead to a dizzying array of choices for the uninitiated. This manuscript has reviewed the state-of-the-art approaches using a common language, discussed considerations which can guide the modeler when selecting which methods to use and some numerical difficulties to be aware of, and offered a view of the current and future challenges in the modeling of continuum robots. As modeling techniques continue to improve in terms of predictive power, as techniques begin to standardize, and as system identification techniques for soft and continuum robots mature, there is every reason to expect that the field will continue to expand, find new applications, and ultimately lead to transformative robotic solutions for human problems.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

HG is solely responsible for all aspects of this manuscript.

FUNDING

This material is based upon work supported by the National Science Foundation under Grant No. 2024795. Any opinions, findings, and conclusion or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Ahmadi, M., Topcu, U., and Rowley, C. (2018). “Control-Oriented Learning of Lagrangian and Hamiltonian Systems,” in 2018 Annual American Control Conference (ACC), Milwaukee, WI, June 27–29, 2018, 520–525. doi:10.23919/ACC.2018.8431726

Allard, J., Cotin, S., Faure, F., Bensoussan, P.-J., Poyer, F., Duriez, C., et al. (2007). “SOFA - an Open Source Framework for Medical Simulation,” in MMVR 15 - Medicine Meets Virtual Reality, Palm Beach, 13–18.

Ansari, Y., Faltotico, E., Mollard, Y., Busch, B., Cianchetti, M., and Laschi, C. (2016). “A Multiagent Reinforcement Learning Approach for Inverse Kinematics of High Dimensional Manipulators with Precision Positioning,” in 2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob), Singapore, June 26–29, 2016 (IEEE), 457–463. doi:10.1109/BIOROB.2016.7523669

Antman, S. S. (2005). Nonlinear Problems of Elasticity. 2nd ed. Applied Mathematical Sciences, (Springer).

Antonelli, M. G., Beomonte Zobel, P., D’Ambrogio, W., and Durante, F. (2020). Design Methodology for a Novel Bending Pneumatic Soft Actuator for Kinematically Mirroring the Shape of Objects. Actuators 9 (4), 113. doi:10.3390/act9040113

Asadian, A., Kermani, M. R., and Patel, R. V. (2011). “An Analytical Model for Deflection of Flexible Needles during Needle Insertion,” in 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, September 25–30, 2011, 2551–2556. doi:10.1109/IROS.2011.6094959

Bacciochi, M., and Tarantino, A. M. (2021). Finite Bending of Hyperelastic Beams with Transverse Isotropy Generated by Longitudinal Porosity. Eur. J. Mech. - A/Solids 85 (1), 104131. doi:10.1016/j.euromechsol.2020.104131

Back, J., Manwell, T., Karim, R., Rhode, K., Althoefer, K., and Liu, H. (2015). “Catheter Contact Force Estimation from Shape Detection Using a Real-Time Cosserat Rod Model,” in 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, September 28–October 2, 2015 (IEEE), 2037–2042. doi:10.1109/RSJ.2015.7353647

Bergou, M., Wardetzky, M., Robinson, S., Audoly, B., and Grinspun, E. (2008). Discrete Elastic Rods. ACM Trans. Graph. 27, 1–12. doi:10.1145/1360612.1360662

Biess, T. M., Largilliere, F., Kruzewski, A., Zhang, Z., Mrozowski, R., and Duriez, C. (2018). Finite Element Method-Based Kinematics and Closed-Loop Control of Soft, Continuum Manipulators. Soft Robotics 5 (3), 348–364. doi:10.1089/soro.2017.0079

Bishop, R. L. (1975). There Is More Than One Way to Frame a Curve. The Am. Math. Monthly 82 (3), 246–251. doi:10.2307/231946

Black, C. B., Till, J., and Rucker, D. C. (2018). Parallel Continuum Robots: Modeling, Analysis, and Actuation-Based Force Sensing. IEEE Trans. Robot. 34 (1), 29–47. doi:10.1109/TRO.2017.2753829

Boyer, F., Ali, S., and Porez, M. (2012). Macrocontinuous Dynamics for Hyperredundant Robots: Application to Kinematic Locomotion Bioinspired
