Diffusion-Based Motion Planning for a Nonholonomic Flexible Needle Model

Wooram Park, Jin Seob Kim, Yu Zhou
Noah J. Cowan, Allison M. Okamura, Gregory S. Chirikjian
Department of Mechanical Engineering
Johns Hopkins University
Baltimore, MD 21218
{wpark7,ncowan,aokamura,gregc}@jhu.edu

Abstract—Fine needles facilitate diagnosis and therapy because they enable minimally invasive surgical interventions. This paper formulates the problem of steering a very flexible needle through firm tissue as a nonholonomic kinematics problem, and demonstrates how planning can be accomplished using diffusion-based motion planning on the Euclidean group, SE(3). In the present formulation, the tissue is treated as isotropic and no obstacles are present. The bevel tip of the needle is treated as a nonholonomic constraint that can be viewed as a 3D extension of the standard kinematic cart or unicycle. A deterministic model is used as the starting point, and reachability criteria are established. A stochastic differential equation and its corresponding Fokker-Planck equation are derived. The Euler-Maruyama method is used to generate the ensemble of reachable states of the needle tip. Inverse kinematics methods developed previously for hyper-redundant and binary manipulators that use this probability density information are applied to generate needle tip paths that reach the desired targets.

Index Terms—needle steering, nonholonomic path planning, probability density function, Euler-Maruyama method, medical robotics

I. INTRODUCTION

Needle insertion is a critical aspect of many medical diagnoses, treatments and scientific studies, including percutaneous procedures requiring therapy delivery to, or sample removal from, a specific location. However, errors in needle targeting can cause unnecessary discomfort for the patient and mitigate the effectiveness of diagnosis or therapy. In addition, many procedures are currently not amenable to needles because of obstacles, such as bone or sensitive tissues, which lie between feasible entry points and potential targets. Thus, there is a clear motivation for needle steering in order to provide accurate and dexterous targeting. It is interesting to note that in clinical practice, some surgeons do make use of needle steering through a combination of lateral, twisting, and inserting motions under visual feedback from imaging systems such as ultrasound [1]. Surgeons accomplish this from experience, making it difficult to teach and limiting the accuracy to that of human hand/eye coordination.

In this work, we formulate the problem of steering a very flexible needle through firm tissue as a nonholonomic kinematics problem. The asymmetry of a bevel tip serves as the source of the nonholonomic constraint. Using this model, we analyze reachability and demonstrate how planning can be accomplished using diffusion-based motion planning on the Euclidean group, SE(3).

Recent work in needle/tissue interaction modeling has been used for planning and simulation of medical procedures [2], [3], [4]. The effects of needle bending have been explored by several groups. O’Leary, et al. [5] demonstrated experimentally that needle bending forces are significantly affected by the presence of a bevel tip. Others have generated needle bending using different strategies such as incorporating a pre-bent stylus inside a straight canula [7], or a telescoping double canula where the internal canula is pre-bent [8]. Kataoka, et al. [9] proposed a model for needle deflection neglecting the effect of a bevel tip, although they acknowledge that the bevel is likely the main source of deflection. Two previous studies [10], [11] have analyzed needle paths for steering and obstacle avoidance, but neither explore the effect of tip asymmetry on steering. This paper builds on our previous experimental work [6], in which we validated a nonholonomic model for needle steering.

Many approaches exist for motion planning of robots both with and without nonholonomic constraints [20], [21]. A planning method based on maximizing probability densities was developed by Ebert-Uphoff and Chirikjian for manipulator inverse kinematics in [15]. Mason and Burdick extended the approach in [15] to the context of mobile robots operating in the plane in [19]. The density function in this approach was generated by convolution on SE(2) as in [16]. The generation of densities in planar problems by solving a diffusion equation on SE(2) was introduced in [18], [24]. These methods are extended to the three-dimensional case of needle steering in this paper.

II. NEEDLE STEERING MODEL

A very flexible needle is treated here as a nonholonomic robot whose end follows a path in three-dimensional space that is prescribed by the properties of the ambient material, the geometry of the needle’s bevel tip and the time-varying translational and rotational insertion speeds of the needle at the entry point. This is depicted in Fig. 1, which also shows the space frame fixed at the insertion point and the body frame fixed on the needle tip. As the needle is inserted...
without twisting, it will have a natural tendency to bend due to the bevel tip. The amount of bending is described by \( \kappa \). Due to its high flexibility, the rest of the needle will follow the path prescribed by the tip. Twisting of the needle at the insertion point is transmitted to the needle tip due to its high torsional stiffness. This allows the direction of bending to be controlled. However, unlike the kinematic cart, the trajectory that the needle follows must remain at its initial orientation to prevent the needle from twisting. At the insertion point, the needle tip frame aligns with the world frame.

The vector \( \xi \) contains both the angular and translational velocity of the motion \( g(t) \) as seen in the body-fixed frame of reference.

Let \( e_i, i = 1, \ldots, 6 \) denote the standard basis for \( \mathbb{R}^6 \). Associated with the Lie algebra \( se(3) \) is a basis given by a set of matrices \( E_i = \hat{e}_i, i = 1, \ldots, 6 \) which, when linearly combined and exponentiated, produce elements of \( SE(3) \). The Lie Algebra has a Lie bracket given by the matrix commutator \( [A, B] := AB - BA \), and an inner product given by \( \langle E_i, E_j \rangle = \delta_{ij} \). The commutator relations are written all together as

\[
[E_1, E_j] = \sum_{k=1}^{6} C_{ij}^k E_k
\]

where \( C_{ij}^k \) are called the structure constants of the Lie algebra. Using this notation we can rewrite (2) as

\[
\xi = \sum_{i=1}^{6} (g^{-1} \dot{g}, E_i) e_i. \tag{5}
\]

### B. A Simple Kinematic Needle Steering Model

Using this notation, our kinematic model is as follows:

\[
\xi = [\kappa v, 0, \omega, 0, 0, v]^T = (\kappa e_1 + e_6) v + (e_3) \omega, \tag{6}
\]

where \( v, \omega \in \mathbb{R}^3 \) are respectively the translational and rotational insertion speeds along and about the tangent to the needle at the point of insertion. The model assumes that when the needle is inserted with speed, \( v \) and twist rate, \( \omega \), they are both transmitted to the tip. The term \( \kappa v \) reflects the fact that the needle will bend with a certain curvature, \( \kappa \), that depends on the bevel angle and material properties of the tissue, and that the rate of bending is proportional to the speed of insertion. When \( \omega = 0 \), this needle model makes a circular arc of curvature, \( \kappa \), as shown in Fig. 1. We constrain \( v \geq 0 \), since “pulling” on the needle does not behave in a symmetric fashion to insertion.

### III. REACHABILITY

We examine the reachability of (6) by computing the control Lie algebra. Since the configuration space, \( SE(3) \), is a Lie group, and the needle steering vector field (6) is left-invariant, computing Lie brackets of the vector fields reduces to computing the Lie brackets in the Lie algebra, \( se(3) \). Computationally, we take successive Lie brackets of \( \mathcal{S}_1 = \hat{s}_1 \) and \( \mathcal{S}_2 = \hat{s}_2 \) and verify that the result spans \( se(3) \). Specifically, let \( \mathcal{S}_3 = [\mathcal{S}_1, \mathcal{S}_2], \mathcal{S}_4 = [\mathcal{S}_1, \mathcal{S}_3], \mathcal{S}_5 = [\mathcal{S}_2, \mathcal{S}_3] \) and \( \mathcal{S}_6 = [\mathcal{S}_1, \mathcal{S}_5] \), and let \( \mathcal{s}_i = (\mathcal{S}_i)^\vee \in \mathbb{R}^6 \), where \( \vee \) is defined in (2). Checking if span\{\( \mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_6 \)\} = \( se(3) \), is equivalent to verifying that span\{\( s_1, s_2, \ldots, s_6 \)\} = \( \mathbb{R}^6 \). Let

\[
M = \begin{bmatrix}
 0 & \kappa & 0 & 0 & 0 & 0 \\
 0 & 0 & -\kappa & 0 & 0 & 0 \\
 0 & 1 & 0 & -\kappa^2 & 0 & 0 \\
 0 & 0 & 0 & \kappa & 0 & 0 \\
 0 & 0 & 0 & 0 & \kappa & 0 \\
 1 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]
and note that
\[
\det (M) = \kappa^4 \neq 0
\]
and hence \(\text{span}\{ s_1, s_2, \ldots, s_6 \} = \mathbb{R}^6\). Thus, from every initial condition, the system (6) can reach an open set of final configurations.

It is worth noting that while the system is “reachable”, it is not necessarily small-time locally controllable. This is due to the fact that “small Lie bracket motions” are not allowed, since we constrain \(v \geq 0\).

IV. GENERATING THE DENSITY OF REACHABLE END POSITIONS

Knowing that any position and orientation in an open set can be reached is a starting point for constructing path planning algorithms. We now take the second step of examining how the distribution of reachable end states is related to the inputs at the base of the needle. Knowing this distribution will help us in the next section to select insertion parameters to reach a desired target.

Starting with the deterministic kinematic model in (6), let us ask what the distribution of end positions and orientations would look like if we could generate the needle conformations corresponding to all possible reasonable orientations would look like if we could generate the needle trajectories, as described in Section V. The approach we take is to view the input parameter \(\omega\) as a random variable. This artificial construction enables us to sample the space of needle trajectories. Since in practice the needle is pushed forward and cannot perform car-like maneuvers such as a “K-turn,” we will assume that
\[
v(t) = 1
\]
and
\[
\omega(t) = \lambda w(t),
\]
where \(w(t)\) is a unit Gaussian white noise with the property that \(\langle \omega(t)w(t + \tau) \rangle = \delta(\tau)\), and \(\lambda\) is a parameter that allows us to choose how much artificial noise to inject for the purpose of sampling trajectories. This means that (6) becomes the following stochastic differential equation (SDE):
\[
(g^{-1}\dot{g})^\gamma dt = \begin{bmatrix} \kappa \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \lambda \\ 0 \\ 0 \\ 0 \end{bmatrix} dW
\]
where \(dW = W(t + dt) - W(t) = w(t)dt\) are the non-differentiable increments of a Wiener process \(W(t)\) and \(dt\) is taken as a finite but small value rather than the classical infinitesimal differential.

This equation, which is of the more general form
\[
(g^{-1}\dot{g})^\gamma dt = cd\gamma + bdW
\]
can be solved to generate an ensemble of sample paths from which the desired distribution information can be extracted.

In contrast, one can consider the evolution of \(f = f(g(t))\), the probability density function (pdf) describing the relative frequency of occurrence of the needle end pose at a particular time. Computing this pdf is useful for generating needle trajectories, as described in Section V.

The Fokker-Planck equation corresponding to the SDE (10) describes how \(f(g, t)\) evolves, and takes the form
\[
\frac{\partial f}{\partial t} = \left( \sum_{i=1}^{6} c_i E_i^R + \frac{1}{2} \sum_{j=1}^{6} \sum_{k=1}^{6} b_{jk} E_j^R E_k^R \right) f
\]
where \(E_i^R\) is a random variable, and \(\lambda\) is a parameter that allows us to choose how much artificial noise to inject for the purpose of sampling trajectories. This means that (6) becomes the following stochastic differential equation (SDE):

\[
\int \mathcal{F}(f(g)) = \int_{\text{SE}(3)} f(g) U(g, p) dg
\]
where \(U(g, p)\) is one of an infinite number of irreducible unitary representation matrices with the property \(U(g_1 \circ g_2, p) = U(g_1, p)U(g_2, p)\). The set of all such matrices is parameterized by the nonnegative real numbers \(p\). The particular property that is useful in the current context is that
\[
\mathcal{F}(E_i^R f) = u(E_i, p) \hat{f}(p)
\]
where
\[
u(E_i, p) = \frac{d}{dt} U(\exp(tE_i))|_{t=0}
\]
This means that the Fokker-Planck equation on (11) can be reduced to a system of linear ordinary differential equations with constant coefficients in this generalized Fourier space.
For a general SDE of the form
$$dx = c(x, t)dt + B(x, t)dW,$$
the discretized version is expressed as
$$x_{i+1} = x_i + c(x_i, t_i)\Delta t + B(x_i, t_i)\Delta W_i,$$
where
$$\Delta t = t_{i+1} - t_i,$$
$$\Delta W_i = W(t_{i+1}) - W(t_i).$$

The simple update procedure represented by these equations is the Euler-Maruyama method. Since $W(t)$ is the Wiener process, we can conclude that $\Delta W_i$ has a Gaussian distribution with zero mean and variance $\Delta t$, which means
$$\Delta W_i \sim \sqrt{\Delta t}N(0, 1)$$

In order to modify (12) for use with (10), we use the following equation, which is suitable for numerical integration of SE(3),
$$g_{i+1} = g_i \exp(\xi),$$
where $g_i = g(t_i)$, $\xi = c(x_i, t_i)\Delta t + B(x_i, t_i)\Delta W_i$ and $\exp(\cdot)$ is the matrix exponential.

We can generate an ensemble of random paths by numerical integration of (10) using (13). Fig. (2) shows the resulting random paths. The ensemble represents the density of the variable, $g$, for a certain time, $t$. Therefore we can obtain a time-evolving PDF of the system variable by the Euler-Maruyama method.

We can see the probability density functions at the final time (10 sec) in Fig. 3. From the Euler-Maruyama method, we can easily construct the function in Fig. 3(a) by counting the number of $(r, \theta)$ pairs in each small bin. However, the probability density in certain bins might be zero, even if the probability density in adjacent bins is large, since this probability density function is obtained from a finite number of random paths. Here we generate the modified probability density function by using Gaussian functions. Each $(r, \theta)$ pair is replaced with a Gaussian distribution.

As an alternative to the Fourier approach, the Euler-Maruyama method can be used to generate the density of the system variable. In other words, the probability density function, $f(g; t)$ can be obtained by using the Euler-Maruyama method to generate an ensemble of sample paths.

As an alternative to the Fourier approach, the Euler-Maruyama method can be used to generate the density of the system variable. In other words, the probability density function, $f(g; t)$ can be obtained by using the Euler-Maruyama method to generate an ensemble of sample paths.

In the next section, numerical examples of this equation are presented, and a method for solving the inverse kinematics problem using this information is demonstrated.

V. INVERSE KINEMATICS USING THE DENSITY OF REACHABLE END POSITIONS

Both the Fourier-based solution method for the needle Fokker-Planck equation and numerical methods for enumerating discrete sample paths are useful for the needle path planning problem. The Fourier method is good for generating results for relatively large values of time (or
needle insertion length). In this way, relatively few terms in the Fourier inversion formula can be used at low computational cost to approximate the solution very well. On the other hand, numerical integration of the SDE, like brute force sampling of discrete conformations, can work very well and be relatively efficient for small values of time, where the needle pdf is highly concentrated.

Let $L$ denote the length of the needle with which we desire to reach a given target. Suppose we know $f(g; t)$ at times $t = kL/N$ for $k = 1, ..., N$ (solved in principle using the SE(3)-Fourier method or the Euler-Maruyama method) and that we also know $\{g_i\}$, a discrete set consisting of $K$ end poses for a needle segment of length $L/N$ generated by integrating the needle SDE $K$ times. Then our approach adapted from [15], [17], [23], evaluates $f(g^{-1}\circ g_{\text{goal}}; L(N-1)/N)$ for each of the $K$ values of $i$, and selects the value of $i$ for which this probability of hitting is maximized. Then, with this value, $i_1$, fixed, we seek $g_{i_1}$ such that $f((g_{i_1}\circ g_t)^{-1}\circ g_{\text{goal}}; L(N-2)/N)$ is maximized. We call this value $g_{i_2}$. At the $k^{th}$ step we seek $g_k = \arg\max f((g_{i_{k-1}}\circ \cdots \circ g_{i_2}\circ g_t)^{-1}\circ g_{\text{goal}}; L(N-k)/N)$.

This process is repeated until $k = N - 1$. At the last step, the value $g_{i_N}$, is chosen from $\{g_i\}$ in order to minimize the distance to the target.

While in principle one would ideally use $f(g; t)$ in this procedure, the amount of time required to compute and store all of these values for $g$ in a fine grid in SE(3) and $N$ values of $t$ can be significant. Since in needle steering we often do not care about the orientation with which a site (such as a lymph node) is approached, the orientational dependence of the pdf can sometimes be ignored. Likewise, if position is written in spherical coordinates as $a = [r \sin \phi \sin \theta, r \cos \phi \sin \theta, r \cos \theta]^T$, and if we obtain a ‘solution’ for which $r$ and $\theta$ are correct but the value of $\phi$ is not, then this ‘solution’ can be used by simply changing the initial insertion roll of the needle. Therefore, the marginal probability density

$$f(r, \theta; t) = \int_0^{2\pi} \int_{SO(3)} f(A, a(r, \theta, \phi); t)dAd\phi$$

can be used for the purpose of planning. This can be efficiently generated using the SE(3)-Fourier method directly without first computing $f(g, t)$ and integrating, or by binning Euler-Maruyama trajectories on a 2D grid (as in Fig. 3) rather than a 6D grid.

We can see some practical results in Fig. 4. This inverse kinematics approach gives us paths reaching various points. Note that using this method, we can find a path going to a point on the $z$-axis (Fig. 4(b)), even though our initial conditions are always along the $z$-axis and the deterministic version of (9) would follow a circular arc. In this case, we confirmed that the needle should follow a helix, which is reasonable because the needle cannot follow a straight line and must make an arc due to the effect of the bevel tip. Fig. 4(c) shows that the needle can reach the points whose distances from the insertion point are approximately the same.

Fig. 5 shows the reachable points in the y-z plane for the given values of the parameters $\kappa$ and $\lambda$. Since $\lambda$ governs the strength of the effect of a unit Gaussian white noise in (8) and (9), it is reasonable that the tip of the needle can reach a wider area with larger $\lambda$ value. If we take too small $\lambda$, the workspace will be too small and it will be hard for the tip to hit the goal point. In other words, $\lambda$ should be large enough. If the distance between the surface of the tissue and a goal point is not within a certain range, then the point is not reachable. This can be overcome by changing $\lambda$ or the insertion speed of the needle, $v$. For the other unreachable goal points, we can relocate the insertion point and the initial insertion orientation.

VI. CONCLUSIONS

A simple nonholonomic kinematic model is presented to describe how an ideal needle with bevel tip moves through firm tissue. The reachability criteria are proven for this model, and an inverse kinematics method based on the propagation of needle-tip pose probabilities is presented. These probabilities are generated as a diffusion process on the Euclidean motion group using methods from stochastic differential equations, Fokker-Planck theory, and harmonic analysis on the Euclidean group. A numerical example demonstrates the method. Future work will address the application and refinement of both this model and the
solution methodology presented here to solve the planning problem for real needles in real tissues, including the case when obstacles are present.

ACKNOWLEDGMENTS

The authors thank Kevin Lynch for his insight on the topic of reachability. This material is based upon work supported by the National Science Foundation under grant No. 0098382, “Diffusion Processes in Motion Planning and Control” (G. Chirikjian, PI), and the National Institutes of Health grant No. R21-EB003452, “Biomechanical Modeling for Steerable Needles” (A. Okamura, PI).

REFERENCES


Fig. 5. Reachable points in the y-z Plane.