ROBOT MOTION PLANNING WITH NONHOLONOMIC CONSTRAINTS

by

Zexiang Li and J. F. Canny

Memorandum No. UCB/ERL M89/13

9 February 1989

ROBOT MOTION PLANNING WITH NONHOLONOMIC CONSTRAINTS

by

Zexiang Li and J. F. Canny

Memorandum No. UCB/ERL M89/13

9 February 1989

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ROBOT MOTION PLANNING WITH NONHOLONOMIC CONSTRAINTS

by

Zexiang Li and J. F. Canny

Memorandum No. UCB/ERL M89/13

9 February 1989

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Robot Motion Planning with Nonholonomic Constraints

Zexiang Li* and J.F. Canny *
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley

January 1, 1989

Abstract

Rolling constraint is a classical example of a class of constraints knwon as non-holonomic constraints. Such a constraint is usually difficult to be dealt with. In this paper, we study motion of an object relative to another under rolling constraints. In particular, we address the following two problems: *(1) Given the geometries of two contacting bodies, assert if a motion exists between two contact configurations without violating the constraints, and (2) if a motion exists, plan a path that links the two contact configurations.* Using the kinematic equations of contact, we first transform contact constraints in the configuration manifold to a system of differential equations in the parameter space. Then, we apply a generalized version of the Frobenius's theorem, called Chow's Theorem, to assert the existence of motion. The smallest involutive distribution, ∇, generated by the constrained vector fields is computed using Macsyma. If two contact configurations are within the same maximal integral manifold of ∇, then a motion exists. To plan a path that links two contact configurations, given that it exists, we first identify the contact angle as the holonomy angle of a closed path in the object surface. Then, we use Lie bracket motion and the Gauss-Bonnet Theorem to generate a desired path. The motion planning algorithm is geometric, and is easy to visualize. Potential applications of this study include (1) rolling a *magnetically levitated mobile robot* in a crowded environment, (2) adjusting contact configurations of a multifingered robot hand without slipping, and (3) following a workpiece by a robot end effector without dissipation or wearing.

*This research is supported in part by NSF PYI Grant DMC 8451129 and by the Defense Advanced Research Projects Agency (DoD), monitored by Space and Naval Warfare Systems Command under Contract N00039-88-C-0292.

1 Introduction

In this paper, we study motion of an object relative to another under *rolling constraints.* For this, consider the two objects shown in Figure 1, which are labelled by *obj1* and obj2, respectively. Here, *obj1* may represent the end-effector of a robot manipulator, the fingertip of a multifingered robot hand, or a *magnetically levitated mobile robot*, and *obj2* represents the contacting environment (or the workpiece) of the manipulator, the object being grasped by the robot hand, or the terrain to be traveled by the mobile robot.

There are at least three advantages in executing rolling motion over sliding motion, which is known to be holonomic. First, *because constraint forces for non-holonomic constraints are workless, wearing problems associated with the contacting bodies are absent.* Second, *the associated control problems can be more easily dealt with..* In order to control sliding motion, for example, the coefficient of friction has to be known exactly, which is in general difficult. Even the world's best figure skater has trouble in managing controlled sliding. On the other hand, rolling motion can be made safe by staying sufficiently close to the center of the friction cone ([LHS88], [CHS88]). Finally, as we will see in this paper, *the set of reachable configurations for rolling is much larger than that for sliding.* This is due to the non-involutivity property of the corresponding constrained vector fields.

Suppose that, say because of the apparent advantages of rolling motion, we have decided to move an object from one contact configuration to another by rolling, then the following two problems are natural and basic to a robotics engineer.

Problem 1 *Assert, using possibly the geometric data of the objects, if a motion that lead from one contact configuration to another without violating the constraints exists.*

Problem 2 *If a motion exists, plan a path that not only satisfies the constraints but also connects the two contact configurations.*

We call the first problem *the existence problem*, the second problem the *planning problem.* The objectives of this paper are to provide solutions to these two

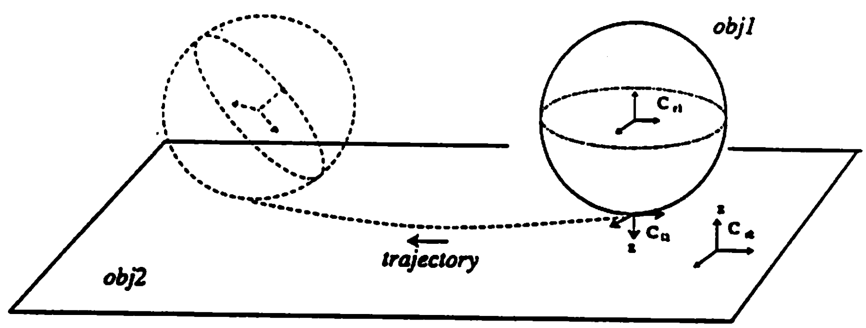

Figure 1: Motion of an object with rolling constraints

problems.

This paper is organized as follows. In Section 2, we review the geometries of a surface and the kinematics of contact. In Section 3, we define the (contact) configuration manifold and describe the contact constraints. Then, we transform the contact constraints in the configuration manifold to a system of differential equations in the parameter space. The *existence problem* is reduced to a problem of characterizing all possible flows generated by the system of differential equations. But, the later problem can be dealt with using existence techniques in nonlinear control theory. In Section 4, using geometric techniques we present a geometric motion planning algorithm.

2 Preliminaries

In this section, we review briefly the geometries of a surface and the kinematics of contact. See [MP78], [Kli78] and [Spi74] for a further treatment on the geometries of a surface and [Mon86], [LHS88], [CR87] and [Ker85] for the kinematics of contact. A short introduction to local and global surface theory is also provided in Appendix A, where the notions of covariant differentiation and holonomy angles and Gauss-Bonnet Theorem are discussed. These issues are key to the development of the *geometrical motion planning algorithm* in Section 4.

Notation 2.1 Let C_i and C_j be two coordinate frames of $\mathbb{R}^3$, where i and j are arbitrary subscripts. Let $r_{i,j} \in \mathbb{R}^3$ and $R_{i,j} \in SO(3)$ denote the position and orientation of C_i relative to C_j. For simplicity, we will write $g_{i,j} \triangleq (r_{i,j}, R_{i,j}) \in SE(3)$ the configuration of C_i relative to C_j.

Definition 2.1 The velocity of C_i relative to C_j is defined by

$$\begin{bmatrix} v_{i,j} \\ w_{i,j} \end{bmatrix} = \begin{bmatrix} R_{i,j}^t \dot{r}_{i,j} \\ \mathcal{S}^{-1}(R_{i,j}^t \dot{R}_{i,j}) \end{bmatrix}, \text{ or } \xi_{i,j} = \begin{bmatrix} \mathcal{S}(w_{i,j}) & v_{i,j} \\ 0 & 0 \end{bmatrix}$$

where $\mathcal{S} : \mathbb{R} \longrightarrow so(3)$ identifies $\mathbb{R}^3$ with the space of 3×3 skew-symmetric matrices.

Note that every $\xi \in se(3)$ defines a left invariant vector field on $SE(3)$, where the corresponding flow $g(t) \in SE(3), t \in I$, satisfies the following differential equation.

$$\dot{g}(t) = g(t)\xi. \tag{1}$$

Thus, a trajectory of an object is given by specifying a left invariant vector field, which in turn is generated by an element $\xi \in se(3)$. On SE(3), which is a Lie group, the solution to (1) for any $\xi(t) \in se(3)$ is defined globally.

Proposition 2.1 *Consider three coordinate frames C_1, C_2 and C_3. The following relation exists between their relative velocities.*

$$\begin{bmatrix} v_{3,1} \\ w_{3,1} \end{bmatrix} = Ad_{g_{3,2}^{-1}} \begin{bmatrix} v_{2,1} \\ w_{2,1} \end{bmatrix} + \begin{bmatrix} v_{3,2} \\ w_{3,2} \end{bmatrix} \tag{2}$$

where $Ad_{g_{3,2}^{-1}}$ is a similarity transformation, given by

$$Ad_{g_{3,2}^{-1}} = \begin{bmatrix} R_{3,2}^t & -R_{3,2}^t S(r_{3,2}) \\ 0 & R_{3,2}^t \end{bmatrix}.$$

Remark 2.1 Suppose that C_3 is fixed relative to C_2, then the velocity of C_3 relative to C_1 is related to that of C_2 by a constant transformation, namely the $Ad_{g_{3,2}^{-1}}$ map.

Definition 2.2 A space curve is the image of a C^2 map $c : I \longrightarrow \mathbb{R}^3$, where I is an interval. The pair (c, I) is called a parameterization of the space curve. c is regular if $\dot{c}(t) \neq 0, \forall t \in I$. A space curve is regular if it has a regular parameterization.

Notation 2.2 U will always denote an open subset of $\mathbb{R}^2$. A point of U will be denoted by $\mathbf{u} \in \mathbb{R}^2$, or by $(u_1, u_2) \in \mathbb{R} \times \mathbb{R}$, or $(u, v) \in \mathbb{R} \times \mathbb{R}$. Let $f : U \longrightarrow \mathbb{R}^3$ be a differentiable map, $df_u : T_u\mathbb{R}^2 \longrightarrow T_{f(u)}\mathbb{R}^3$ denotes the tangent map of f, and f_u, f_v denote the partial derivatives of f with respect to u and v, respectively.

Definition 2.3 A surface (or an embedded 2 manifold) in $\mathbb{R}^3$ is a subset $S \subset \mathbb{R}^3$ such that for every point $s \in S$, there exists an open subset S_s of S (in the induced topology of $\mathbb{R}^3$) with the property (1) $s \in S_s$, (2) S_s is the image of a C^3 map $f : U \longrightarrow \mathbb{R}^3$, where $f_u \times f_v \neq 0, \forall (u, v) \in U$, and (3) $f : U \longrightarrow S_s \subset \mathbb{R}^3$ is a diffeomorphism.

S_s is called a coordinate patch and the pair (f, U) is called a (local) coordinate system of S. The coordinates of a point $s \in S_s$ are given by $(u, v) = f^{-1}(s)$. From now on, if the coordinate system is clear from the context, we shall not distinguish between a point $s \in S_s$ and its coordinates. The collection of coordinate patches $\{S_s\}$ which covers S, i.e., $S = \cup S_s$, is called an atlas of S. By a curve in S we mean a curve $c : I \longrightarrow \mathbb{R}^3$, which can be expressed as $f \circ \mathbf{u}(t)$ for some curve $\mathbf{u} : I \longrightarrow U$ in U.

Remark 2.2 The surface of a smooth object is a compact embedded surface, and thus with a finite number of atlases.

Example 2.1 The sphere S of radius ρ is an embedded surface. To prove this, let $U = \{(u, v) \in \mathbb{R}^2, -\frac{\pi}{2} < u < \frac{\pi}{2}, -\pi < v < \pi\}$ and consider the following coordinate systems

$$f : U \longrightarrow \mathbb{R}^3 : (u, v) \longmapsto (\rho \cos u \cos v, -\rho \cos u \sin v, \rho \sin u)$$

and

$$\hat{f} : U \longrightarrow \mathbb{R}^3 : (u, v) \longmapsto (-\rho \cos u \cos v, \rho \sin u, \rho \cos u \sin v).$$

The partial derivatives of f and $\hat{f}$ are

$$\begin{aligned} f_u &= (-\rho \sin u \cos v, \rho \sin u \sin v, \rho \cos u) \\ f_v &= (-\rho \cos u \sin v, -\rho \cos u \sin v, 0) \end{aligned}$$

and

$$\begin{aligned} \hat{f}_u &= (\rho \sin u \cos v, \rho \cos u, -\rho \sin u \sin v) \\ \hat{f}_v &= (\rho \cos u \sin v, 0, \rho \cos u \cos v) \end{aligned}$$

Clearly, $f_u \times f_v \neq 0$ and $\hat{f}_u \times \hat{f}_v \neq 0$, $\forall (u, v) \in U$. Moreover, $S_1 = f(U)$ and $S_2 = \hat{f}(U)$ covers S. Thus, S is an embedded 2 manifold.

The unit sphere (i.e., $\rho = 1$) of $\mathbb{R}^3$ is denoted by S^2.

Example 2.2 The ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ can be parametrized by the following coordinate system

$$f : U \longrightarrow \mathbb{R}^3 : (u, v) \longmapsto (a \cos u \cos v, -b \cos u \sin v, c \sin u)$$

and

$$\hat{f} : U \longrightarrow \mathbb{R}^3 : (u, v) \longmapsto (-a \cos u \cos v, b \sin u, c \cos u \sin v)$$

where U is given by the previous example.

Definition 2.4 The Gauss map of a surface S is a continuous map $n : S \longrightarrow S^2$ such that $n(s)$ is normal to S. Note that n induces an orientation on S. We will also use n to denote the map $n \circ f : U \longrightarrow S^2$, and $T_u n : T_u\mathbb{R}^2 \longrightarrow T_{f(u)}\mathbb{R}^3$ denote the tangent map of n.

Definition 2.5 A coordinate system (f, U) is called orthogonal if $f_u \cdot f_v = 0,\ \forall(u, v) \in U$, and right-handed if $f_u \times f_v / |f_u \times f_v| = n \circ f(\mathbf{u})$. (f, U) is called a geodesic coordinate system if it is orthogonal and $f_u \cdot f_u = 1$ and $f_v \cdot f_v = q^2$, for some $q > 0$ which is a function of (u, v). Let (f, U) be an orthogonal right-handed coordinate system for a surface patch $S_0 \subset S$. We define the Gaussian frame at a point $s \in S_0$ as the coordinate frame with origin at $f(\mathbf{u})$ and coordinate axes

$$\mathbf{x}(\mathbf{u}) = f_u/|f_u|, \quad \mathbf{y}(\mathbf{u}) = f_v/|f_v|, \text{ and } \mathbf{z}(\mathbf{u}) = n \circ f(\mathbf{u}).$$

Definition 2.6 Let S_0 be a coordinate patch of S, with an orthogonal coordinate system (f, U). At a point $s \in S_0$, the curvature form K is defined as the 2×2 matrix

$$K = [\mathbf{x}(\mathbf{u}), \mathbf{y}(\mathbf{u})]^t [\mathbf{z}_u(\mathbf{u})/|f_u|, \mathbf{z}_v(\mathbf{u})/|f_v|],$$

where $\mathbf{u} = f^{-1}(s)$. The torsion form T is the 1×2 matrix

$$T = \mathbf{y}(\mathbf{u})^t [\mathbf{x}_u(\mathbf{u})/|f_u|, \mathbf{x}_v(\mathbf{u})/|f_v|],$$

and the metric tensor M is the 2×2 matrix

$$M = \begin{bmatrix} |f_u| & 0 \\ 0 & |f_v| \end{bmatrix}.$$

Note that M is the square root of the first fundamental form.

Example 2.3 Consider the sphere S of radius ρ. Let $S_1 = f(U)$ be the coordinate patch of S studied in Example 2.1. The Gauss frame at a point $s \in S_1$ is given by

$$\mathbf{x}(\mathbf{u}) = \begin{bmatrix} -\sin u \cos v \\ \sin u \sin v \\ \cos u \end{bmatrix}, \quad \mathbf{y}(\mathbf{u}) = \begin{bmatrix} -\sin v \\ -\cos v \\ 0 \end{bmatrix} \text{ and } \mathbf{z}(\mathbf{u}) = \begin{bmatrix} \cos u \cos v \\ -\cos u \sin v \\ \sin u \end{bmatrix}.$$

The curvature form, torsion form and metric tensor are given by

$$K = \begin{bmatrix} 1/\rho & 0 \\ 0 & 1/\rho \end{bmatrix}, \ T = [0 - \tan u/\rho], \text{ and } M = \begin{bmatrix} \rho & 0 \\ 0 & \rho \cos u \end{bmatrix}.$$

We now consider the two objects that move while maintaining contact with each other (see Figure 1). Choose reference frames C_{r1} and C_{r2} fixed relative to *obj1* and *obj2*, respectively. Let $S_1 \subset \mathbb{R}^3$ and $S_2 \subset \mathbb{R}^3$ be the embeddings of the surfaces of *obj1* and *obj2* relative to C_{r1} and C_{r2}, respectively. Let n_1 and n_2 be the Gauss maps (outward normal) for S_1 and S_2. Choose atlases $\{S_{1,i}\}_{i=1}^{m_1}$ and $\{S_{2,i}\}_{i=1}^{m_2}$ for S_1 and S_2. Let $(f_{1,i}, U_{1,i})$ be an orthogonal right handed coordinate system for $S_{1,i}$ with Gauss map n_1. Similarly, let $(f_{2,i}, U_{2,i})$ be an orthogonal, right-handed coordinate system for $S_{2,i}$ with n_2.

Let $c_1(t) \in S_1$ and $c_2(t) \in S_2$ be the positions at time t of the point of contact relative to C_{r1} and C_{r2}, respectively. We will restrict attentions to an interval I such that $c_1(t) \in S_{1,i}$ and $c_2(t) \in S_{2,j}$ for all $t \in I$ and some i and some j. The coordinate systems $(f_{1,i}, U_{1,i})$ and $(f_{2,j}, U_{2,j})$ induce a normalized Gauss frame at all points in $S_{1,i}$ and $S_{2,j}$. We define a continuous family of coordinate frames, two for each $t \in I$, as follows. Let the local frames at time t, C_{l1} and C_{l2}, be the coordinate frames that fixed relative to C_{r1} and C_{r2}, respectively, that coincide at time t with the normalized Gauss frames at $c_1(t)$ and $c_2(t)$ (see Figure 1).

We now define the parameters that describe the 5 degrees of freedom for the motion of the point of contact. The coordinates of the point of contact relative to the coordinate system $(f_{1,i}, U_{1,i})$ and $(f_{2,j}, U_{2,j})$ are given by $\mathrm{u}_1(t) = f_{1,i}^{-1}(c_1(t)) \in U_{1,i}$ and $\mathrm{u}_2(t) = f_{2,j}^{-1}(c_2(t)) \in U_{2,j}$. These account for 4 degrees of freedom. The final parameter is the angle of contact $\psi(t)$, which is defined as the angle between the x−axis of C_{l1} and C_{l2}. We choose the sign of ψ so that a rotation of C_{l1} through $-\psi$ around its z axis aligns the x−axis.

We describe the motion of *obj1* relative to *obj2* at time t, using the local coordinate frames frames C_{l1} and C_{l2}. Let v_x, v_y and v_z be the components of translation velocity of C_{l1} relative to C_{l2} at time t. Similarly, let w_x, w_y and w_z be the components of rotational velocity.

The symbols K_1, T_1 and M_1 represent, respectively, the curvature form, torsion form and metric at time t at the point $c_1(t)$ relative to the coordinate system $(f_{1,i}, U_{1,i})$. We can analogously define K_2, T_2, and M_2. We also let

$$R_\psi = \begin{bmatrix} \cos\psi & -\sin\psi \\ -\sin\psi & -\cos\psi \end{bmatrix}, \ \tilde{K}_2 = R_\psi K_2 R_\psi.$$

Note that $\tilde{K}_2$ is the curvature of *obj2* at the point of contact relative to the x- and y -axes of C_{l1}. Call $K_1 + \tilde{K}_2$ the relative curvature form.

The following kinematic equations that describe motion of the point of contact over the surface of *obj1* and *obj2* in response to a relative motion between these objects are due to Montana ([Mon86]).

Theorem 2.1 *(Kinematic equations of contact) At a point of contact, if the relative curvature form is invertible, then the point of contact and angle of contact evolve according to*

$$\dot{\mathbf{u}}_1 = M_1^{-1}(K_1 + \tilde{K}_2)^{-1}\left(\begin{bmatrix} -w_y \\ w_x \end{bmatrix} - \tilde{K}_2 \begin{bmatrix} v_x \\ v_y \end{bmatrix}\right), \tag{3}$$

$$\dot{\mathbf{u}}_2 = M_2^{-1}R_\psi(K_1 + \tilde{K}_2)^{-1}\left(\begin{bmatrix} -w_y \\ w_x \end{bmatrix} + K_1 \begin{bmatrix} v_x \\ v_y \end{bmatrix}\right), \tag{4}$$

$$\dot{\psi} = w_z + T_1 M_1 \dot{\mathbf{u}}_1 + T_2 M_2 \dot{\mathbf{u}}_2, \tag{5}$$

$$0 = v_z. \tag{6}$$

The last equation is called the constraint equation.

Example 2.4 (*The classical example re-visited*) Let's consider the classical example of a unit disk rolling on the plane, as shown in Figure 2 (See [Gol80], and [Gre77]). The coordinates of the plane are given by $(u_2, v_2) \in \mathbf{R}^2$, and the coordinate of the disk is $u_1 \in \mathbf{R}$. Embed the disk in $\mathbf{R}^3$ with the following parametrization

$$f : U_1 \longrightarrow \mathbf{R}^3 : u_1 \longmapsto (\cos u_1, \sin u_1, 0).$$

We define the Gauss frame of the disk by the frame with origin at $f(u_1)$ and coordinate axes

$$\mathbf{x}(u_1) = f', \ \mathbf{z}(u_1) = f'', \text{ and } \ \mathbf{y}(u_1) = \mathbf{z} \times \mathbf{x}.$$

Let ψ be the angle of contact. Let (v_x, v_y, v_z) be the components of translational velocity of C_{l1} relative to C_{l2}, and $(0, w_y, w_z)$ be the components of rotational velocity. *Note that the disk has only two degrees of rotational freedom.* Following a procedure outlined in [Mon86], we derive the following kinematic equations of contact for the

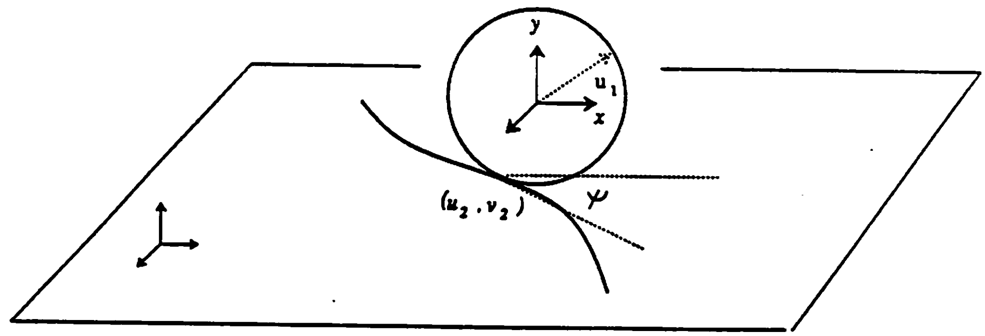

Figure 2: A unit disk roll in the plane

moving disk.

$$\begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \\ \dot{v}_2 \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} -1 \\ -\cos\psi \\ \sin\psi \\ 0 \end{bmatrix} w_y + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} w_z + \begin{bmatrix} 0 \\ \cos\psi \\ \sin\psi \\ 0 \end{bmatrix} v_x + \begin{bmatrix} 0 \\ -\sin\psi \\ \cos\psi \\ 0 \end{bmatrix} v_y, \quad (7)$$

$$v_z = 0.$$

From Greenwood ([Gre77]), the constraints of no slippage are described in differential forms by

$$du_2 - \cos\psi du_1 = 0, \quad (8)$$

$$dv_2 - \sin\psi du_1 = 0. \quad (9)$$

It is important to note that the two differential forms (or 1-forms) given by Equations (8) and (9) annihilate exactly the first two vector fields of Equation (7) . For example, we have that

$$(-1, -\cos\psi, \sin\psi, 0) \cdot (-\cos\psi, 1, 0, 0)^t = 0.$$

In other words, rolling constraints can be described by one of the following two equivalent conditions: *(1) the differential forms given by (8) and (9) be zero, or (2) the vector fields which do not annihilate these one-forms vanish from the kinematic equations of contact, which in turn is equivalent to* $v_x = v_y = 0$. This duality should always be kept clear in mind.

Definition 2.7 We define three special modes of contact (or contact constraints) in terms of the relative velocity components by

(1) Fixed point of contact:

$$\begin{bmatrix} v_x \\ v_z \end{bmatrix} = 0, \text{ and } \begin{bmatrix} w_x \\ w_y \end{bmatrix} = 0; \tag{10}$$

(2) Rolling contact:

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = 0, \quad \text{and } \ w_z = 0; \tag{11}$$

(3) Sliding contact:

$$\begin{bmatrix} w_x \\ w_y \\ w_z \end{bmatrix} = 0. \tag{12}$$

The following results follow immediately from Theorem 2.1.

Corollary 2.1 *The kinematic equations of contact correspond to each of the contact modes are*

$$\begin{cases} \dot{\mathbf{u}}_1 = 0, \\ \dot{\mathbf{u}}_2 = 0, \\ \dot{\phi} = w_z, \end{cases} \tag{13}$$

for fixed point of contact,

$$\begin{cases} \dot{\mathbf{u}}_1 = M_1^{-1}(K_1 + \tilde{K}_2)^{-1} \begin{bmatrix} -w_y \\ w_x \end{bmatrix}, \\ \dot{\mathbf{u}}_2 = M_2^{-1} \hat{R}_\phi (K_1 + \tilde{K}_2)^{-1} \begin{bmatrix} -w_y \\ w_x \end{bmatrix}, \\ \dot{\phi} = T_1 M_1 \dot{\mathbf{u}}_1 + T_2 M_2 \dot{\mathbf{u}}_2. \end{cases} \tag{14}$$

for rolling contact, and

$$\begin{cases} \dot{\mathbf{u}}_1 = -M_1^{-1}(K_1 + \tilde{K}_2)^{-1} \tilde{K}_2 \begin{bmatrix} v_x \\ v_y \end{bmatrix}, \\ \dot{\mathbf{u}}_2 = M_2^{-1} \hat{R}_\phi (K_1 + \tilde{K}_2)^{-1} K_1 \begin{bmatrix} v_x \\ v_y \end{bmatrix}, \\ \dot{\phi} = T_1 M_1 \dot{\mathbf{u}}_1 + T_2 M_2 \dot{\mathbf{u}}_2. \end{cases} \tag{15}$$

for sliding contact.

3 Existence of Motion

In this section, we use the kinematic equations of contact and a generalized version of the Frobenius Theorem to verify the existence of motion between two contact configurations under rolling. For convenience, we restate the problem here, with a slight modification.

Problem 1' *Consider motion of obj1 relative to obj2, as shown in Figure 1. Let* $g^0_{r1,r2} = (r^0_{r1,r2}, R^0_{r1,r2}) \in SE(3)$ *be an initial contact configuration of obj1 relative to obj2, and* $g_{r1,r2}(t) \in SE(3), t \in [0, t_f]$*, be a trajectory of obj1 that satisfies the rolling constraints and* $g_{r1,r2}(0) = g^0_{r1,r2}$*. Characterize the set of reachable configurations of obj1, that is, a configuration* $g^f_{r1,r2} \in SE(3)$ *such that there exists a* $t_f \in [0, \infty)$ *with* $g^f_{r1,r2} = g_{r1,r2}(t_f)$*.*

In Section 2, contact constraints have been described in terms of the local coordinate frames. We need to transform these constraints to the reference frame C_{r1}. By Proposition 2.1, the velocity of C_{r1} is related to that of C_{l1} by a similarity transformation

$$\begin{bmatrix} v_{r1,r2} \\ w_{r1,r2} \end{bmatrix} = Ad_{g_{l1,r1}} \cdot \begin{bmatrix} v_{l1,r2} \\ w_{l1,r2} \end{bmatrix} \tag{16}$$

where $g_{l1,r1}$ is the configuration variable of the local frame of *obj1* at the point of contact relative to C_{r1}. On the other hand, since C_{l2} is fixed relative to C_{r2}, we conclude that

$$v_{l1,r2} = v_{l1,l2} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} \text{ and } w_{l1,r2} = w_{l1,l2} = \begin{bmatrix} w_x \\ w_y \\ w_z \end{bmatrix}.$$

Thus, rolling constraints require that the velocity field of *obj1* be expressible in the form

$$\begin{bmatrix} v_{r1,r2} \\ w_{r1,r2} \end{bmatrix} = Ad_{g_{l1,r1}} \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ w_x \\ w_y \\ 0 \end{bmatrix} \tag{17}$$

for some $(w_x, w_y) \in \mathbb{R}^2$. By Equation (1) the trajectory of *obj1* is uniquely determined by specifying the rolling velocity (w_x, w_y).

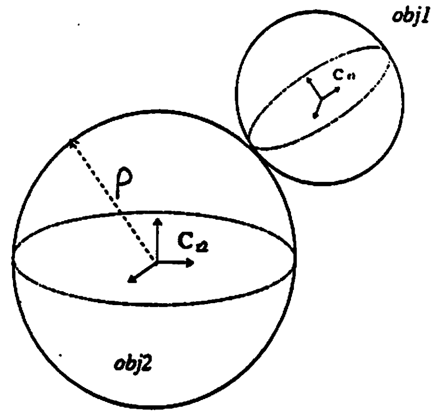

Figure 3: Motion of a unit ball over another ball

By maintaining contact with *obj2*, the set of reachable configurations (of *obj1*) must be within a five dimensional manifold M of SE(3). M is called the contact configuration manifold, and locally M can be described as the zero set of a function $h : SE(3) \longrightarrow \mathbf{R}$, i.e., $M = h^{-1}(0)$. Here, h is called a height function.

Example 3.1 Let *obj1* be a unit ball and *obj2* be the plane, as shown in Figure 1. Let the reference frame C_{r1} be fixed to the center of the ball and the reference frame C_{r2} be arbitrary at the plane, except that its z-axis points up. Consider the following function

$$h : SE(3) \longrightarrow \mathbf{R} : (r, R) \longmapsto r_z - 1,$$

where $r = (r_x, r_y, r_z)^t$. Clearly, a reachable configuration must be an element of $M = h^{-1}(0)$, which is a 5 dimensional submanifold of $SE(3)$.

Example 3.2 Let *obj1* be a unit ball and *obj2* be a ball of radius ρ, as shown in Figure 3. Let C_{r1} and C_{r2} be fixed to the center of *obj1* and *obj2*, respectively. Then, M is the zero set of the following function

$$h : SE(3) \longrightarrow \mathbf{R} : (r, R) \longmapsto |r| - (1 + \rho).$$

From now on, the 5 dimensional manifold M will be called the configuration space (or manifold) of *obj1*. In principle, in order to determine the set of reachable points within M under rolling constraints, we need to know M. But, as we see from the above examples that an analytic description of M depends on the geometric of the objects and becomes complicated when these objects have less no symmetries at all. Fortunately, we will need to deal with another space, called the parameter space, which is much easier to characterize than M. Moreover, the parameter space is diffeomorphic to M. Thus, if a motion exists between two contact configurations within M then a flow exists in the parameter space such that the images of the two contact configurations are linked by the flow, and conversely. We now proceed to make these statements more precise.

Definition 3.1 Let $S_1 \subset \mathbb{R}^3$ and $S_2 \subset \mathbb{R}^3$ be the embeddings of the surfaces of *obj1* and *obj2* relative to C_{r1} and C_{r2}, respectively. Let S^1 be the unit sphere of $\mathbb{R}^2$ representing the space of contact angles. Then, the parameter space P is the following product space

$$P = S_1 \times S_2 \times S^1.$$

P is a five dimensional manifold. The topology of P is given by the product topology. In other words, an atlas of P is given by $\{S_{1,i}\}_{i=1}^{m_1} \times \{S_{2,i}\}_{i=1}^{m_2} \times \{S_i^1\}_{i=1}^{m_3}$, where $\{S_{1,i}\}_{i=1}^{m_1}$ is an atlas of S_1, and etc. Consequently, a coordinate system of P is given by $(f_{1,i}, U_{1,i}) \times (f_{2,j}, U_{2,j}) \times \psi$. The coordinate system for the contact angle is induced by a coordinate system of S_1 and a coordinate system of S_2.

Proposition 3.1 *If both objects are convex and at least one of them is strictly convex, then the configuration manifold M is diffeomorphic to the parameter space P.*

Proof. This is essentially a restatement of the kinematic equations of contact. Since both objects are convex and at least one of them is strictly convex, the relative curvature form is invertible. This implies that contact must occur over isolated points. Let $f : M \subset SE(3) \longrightarrow P$ be the map that takes a contact configuration to the corresponding contact parameters. Clearly, f is one-to-one and onto, and the tangent map of f is just the kinematic equations of contact. Thus, f is a diffeomorphism. □

From Section 2, rolling constraints in the parameter space are described by a system of differential equations given in Theorem 2.1. Rearrange the kinematic equations of contact, we have

$$\begin{bmatrix} \dot{u}_1 \\ \dot{v}_1 \\ \dot{u}_2 \\ \dot{v}_2 \\ \dot{\psi} \end{bmatrix} = X_1(\mathbf{u}_1, \mathbf{u}_2, \psi)w_x + X_2(\mathbf{u}_1, \mathbf{u}_2, \psi)w_y, \tag{18}$$

where $\mathbf{u}_1 = (u_1, v_1)$ and $\mathbf{u}_2 = (u_2, v_2)$. $X_1(\mathbf{u}_1, \mathbf{u}_2, \psi)$ and $X_2(\mathbf{u}_1, \mathbf{u}_2, \psi)$ are the (constrained) vector fields on P, which correspond to rolling motion. Equation (18) is a system of differential equations in local coordinates on the manifold P.

Definition 3.2 Let $p_0 = (u_1^0, v_1^0, u_2^0, v_2^0, \psi^0)^t$ be an initial point in P. A point $p_f \in P$ is said to be *reachable from* p_0 *if there exists a choice of* $(w_x, w_y) \in \mathbb{R}^2$ *such that the flow* $p(t)$ *of (18) reaches* p_f *after some finite* $t_f \in [0, \infty)$*, i.e.,* $p(0) = p_0$ *and* $p(t_f) = p_f$*. Such a* $p_f \in P$ *is said to be in the reachable space of* p_0.

We can now restate the reachability (or existence) problem in terms of the system (18).

Problem 1" *Given an initial point* p_0 *(*p_0 *corresponds to the initial contact configuration* $g^0_{r1,r2}$ *with the diffeomorphism of Proposition 3.1). Characterize the set of reachable points from* p_0 *by the system (18).*

We apply Chow's theorem to solve *Problem 1"*. Chow's theorem is a generalization of the Frobenius's Theorem and has has been widely used in nonlinear control theory ([Isi85], [HK77], [Spi74]).

Theorem 3.1 *(Chow's Theorem). Consider the following system of differential equations on a n-dimensional manifold* N.

$$\dot{x} = f_1(x)u_1 + ... f_m(x)u_m, \quad m \leq n \tag{19}$$

where $x \in N$ *is the states in local coordinates,* $f_i(x) \in \chi^\infty(N), i = 1, ...m$*, is a* C^∞ *vector field on* N *and* $(u_1, ...u_m) \in \mathbb{R}^m$ *are the control inputs.*

Let $\nabla = \{f_1, ...f_m\}_{LA}$ *denote the smallest involutive distribution containing* $\{f_1, ...f_m\}$ *(or the the smallest Lie algebra of vector fields generated by* $\{f_1, ...f_m\}$ *) and* $\tilde{N}_{x_0}$ *the maximal integral manifold of* ∇ *through* $x_0 \in N$ *(* $\tilde{N}_{x_0}$ *exists and is unique by Frobenius's Theorem). Then,*

1. *A point $x \in N$ is reachable from x_0 if and only if $x \in \tilde{N}_{x_0}$, i.e., (x, x_0) belong to the same maximum integral manifold of ∇.*

2. *Every point in N is reachable if and only if $\tilde{N}_{x_0} = N$ if and only if* $\dim \nabla = n$.

The following algorithm computes ∇.

Algorithm 3.1 Input: *A collection of vector fields $\{f_1, ...f_m\} \in \chi^\infty(N)$*

Output: $\nabla = \{f_1, ...f_m\}_{LA}$.

Step 1: *Set*

$$\nabla_0 = \{f_1, f_m\};$$

Step 2: *Compute*

$$\nabla_k = \nabla_{k-1} + \sum_{i=1}^{m} [f_i, \nabla_{k-1}];$$

until an integer k^* such that $\nabla_{k^*} = \nabla_{k^*+1}$, then $\nabla = \nabla_{k^*}$ and return.

Remark 3.1 1. $k^* \leq n$.

2. $[f_1, f_2]$ denotes the Lie bracket vector field of $f_1, f_2 \in \chi^\infty(N)$. In local coordinates $(x_1, ...x_n) \in \mathbf{R}^n$ of N, the Lie bracket vector field is computed by the formula

$$[f_1, f_2] = \frac{\partial f_2}{\partial x_i} f_1 - \frac{\partial f_1}{\partial x_i} f_2,$$

where $\frac{\partial f_1}{\partial x_i}$ is the ordinary Jacobian matrix of f_1.

Using Chow's Theorem and Algorithm 3.1 we have the following algorithm for verifying the existence of motion.

Algorithm 3.2 *(Existence of Motion)*

Input: *1. The coordinates of two points $p_0, p_f \in P$.*

2. Constrained vector fields $X_1, X_2 \in \chi^\infty(P)$ that describe the rolling motion.

Output: *A binary answer on if p_f can be reached from p_0.*

Step 1: *Compute the coordinate expressions of the constrained vector fields* $X_1(u_1, u_2, \psi)$ *and* $X_2(u_1, u_2, \psi)$ *for* $(u_1, u_2) \in \left(\{U_{1,i}\}_{i=1}^{m_1} \times \{U_{2,j}\}_{j=1}^{m_2}\right)$. *(Assume that* S_1 *is covered by* $\{S_{1,i}\}_{i=1}^{m_1}$ *and* S_2 *is covered by* $\{S_{2,j}\}_{j=1}^{m_2}$ *).*

Step 2: *Compute the following Lie bracket vector fields*

$$\begin{aligned} X_3 &= [X_1, X_2] = \frac{\partial X_2}{\partial p_i} X_1 - \frac{\partial X_1}{\partial p_i} X_2, \\ X_4 &= [X_1, X_3], \\ X_5 &= [X_2, X_3], \end{aligned} \tag{20}$$

where $p = (u_1, v_1, u_2, v_2, \psi)^t$. *Set* $\nabla' = \{X_1, X_2, X_3, X_4, X_5\}$, *which is an involutive distribution containing* (X_1, X_2).

Step 3:

- *If* $dim(\nabla') = 5$, $\forall p \in P$, *then* $\nabla = \nabla'$ *and* $\tilde{N}_{p_0} = P$. *Return true for any* p_0 *and* p_f *in* P. *(Thus, every contact configuration in* M *is reachable by rolling.)*
- *If* $dim(\nabla') = n < 5$. *Let* ∇ *be the smallest involutive distribution contained in* ∇' *with rank* n *and* $\tilde{N}_{p_0}$ *the maximum integral manifold of* ∇ *through* p_0. *If* $p_f \in \tilde{N}_{p_0}$ *then return true, otherwise return false.*

Remark 3.2 The above algorithm can be computed symbolically using Macsyma.

Example 3.3 Consider the example of a unit ball moving on the plane, as shown in Figure 1. From Example 2.1, the ball can be covered by two coordinate systems, and the plane by a sinle coordinate system in the obvious way. The curvature form, metric and torsion form of the unit ball are computed in Example 2.3. The geometric data of the plane can be easily computed using the following coordinate system.

$$f_{2,1} : U_{2,1} \longrightarrow \mathbb{R}^3 : (u_2, v_2) \longmapsto (u_2, v_2, 0),$$

where $U_{2,1} = \mathbb{R}^2$. Clearly, the metric is the identity and the torsion form as well as the curvature form are zero.

Following Algorithm 3.2, we have

Step 1: On the first coordinate system of P, the kinematic equations of contact are

$$\begin{bmatrix} \dot{u}_1 \\ \dot{v}_1 \\ \dot{u}_2 \\ \dot{v}_2 \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 0 \\ \sec u_1 \\ -\sin\psi \\ -\cos\psi \\ -\tan u_1 \end{bmatrix} w_x + \begin{bmatrix} -1 \\ 0 \\ -\cos\psi \\ \sin\psi \\ 0 \end{bmatrix} w_y \quad (21)$$
$$\triangleq X_1 w_x + X_2 w_y.$$

Step 2: The successive Lie brackets of X_1 and X_2 are

$$X_3 = [X_1, X_2] = \begin{bmatrix} 0 \\ -\sec u_1 \tan u_1 \\ -\sin\psi \tan u_1 \\ -\cos u_1 \tan u_1 \\ -\sec^2 u_1 \end{bmatrix},$$

$$X_4 = [X_1, X_3] = \begin{bmatrix} 0 \\ 0 \\ -\cos\psi \\ \sin\psi \\ 0 \end{bmatrix},$$

and

$$X_5 = [X_2, X_3] = \begin{bmatrix} 0 \\ (1+\sin^2 u_1)\sec^3 u_1 \\ 2\sin\psi \sec^2 u_1 \\ 2\cos\psi \sec^2 u_1 \\ 2\sec^2 u_1 \tan u_1 \end{bmatrix}.$$

Step 3: Compute the rank of

$$\nabla' = \{X_1, X_2, X_3, X_4, X_5\}$$

It is easy to verify that, through elementary row and column operations, the determinant of ∇' is identically 1.

Steps 1 through 3 are repeated for the second coordinate system of P and ∇' is again nonsingular.

Output: *It is true that a unit ball can reach any contact configuration on the plane by rolling!*

♠

Example 3.4 Our second example consists of a unit ball moving relative to another ball of radius ρ (See Example 3.2). Since each ball has two coordinate systems, P has a total of four coordinate systems.

Step 1: In the first coordinate system the kinematic equations of contact are

$$\begin{bmatrix} \dot{u}_1 \\ \dot{v}_1 \\ \dot{u}_2 \\ \dot{v}_2 \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 0 \\ (1-\beta)\sec u_1 \\ -\beta\sin\psi \\ -\beta\cos\psi\sec u_2 \\ \beta\tan u_2\cos\psi - (1-\beta)tanu_1 \end{bmatrix} w_x + \begin{bmatrix} -(1-\beta) \\ 0 \\ -\beta\cos\psi \\ \beta\sin\psi\sec u_2 \\ -\beta\tan u_2\sin\psi \end{bmatrix} w_y$$

$$\triangleq X_1 w_x + X_2 w_y,$$

where $\beta = \frac{1}{1+\rho}$.

Step 2: Using Macsyma, the successive Lie brackets of X_1 and X_2 are computed.

$$X_3 = [X_1, X_2] = \begin{bmatrix} 0 \\ (1-\beta)^2\sec^2 u_1 \\ \beta(1-\beta)\sin\psi\sin u_1\sec u_1 \\ \beta(1-\beta)\cos\psi\sin u_1\sec u_1\sec u_2 \\ X_{3,5} \end{bmatrix},$$

where

$$X_{3,5} = -\frac{\beta(1-\beta)\cos\psi\cos u_1\sin u_1\sin u_2 + \{-\beta^2\cos^2 u_1 + (\beta-1)^2\}\cos u_2}{\cos^2 u_1\cos^2 u_2};$$

$$X_4 = [X_1, X_3] = \begin{bmatrix} 0 \\ 0 \\ \beta(2\beta-1)\cos\psi \\ -\beta(2\beta-1)\sin\psi\sin u_2\sec u_2 \\ \beta(2\beta-1)\sin\psi\sin u_2\sec u_2 \end{bmatrix};$$

$$X_5 = [X_2, X_3] = \begin{bmatrix} 0 \\ -\{-(1-\beta)^3\cos^2 u_1 + 2(1-\beta)^3\}\sec^3 u_1 \\ -\{\beta^3\sin\psi\cos^2 u_1 - 2\beta(1-\beta)^2\sin\psi\}\sec^2 u_1 \\ -\{\beta^3\cos\psi\cos^2 u_1 - 2\beta(1-\beta)^2\cos\psi\}\sec^2 u_1\sec u_2 \\ X_{5,5} \end{bmatrix};$$

where

$$X_{5,5} = \frac{\{\beta^3\cos\psi\cos^3 u_1 - 2\beta(1-\beta)^2\cos\psi\cos u_1\}\sin u_2 + \alpha}{\cos^3 u_1\cos u_2}$$

and

$$\alpha = \{\beta^2(1-\beta)\cos^2 u_1 - 2(1-\beta)^3\}\sin u_1\cos u_2.$$

Step 3: Computing the determinant of

$$\nabla' = \{X_1, X_2, X_3, X_4, X_5\}$$

gives

$$\det \nabla' = -\frac{(\beta-1)^2\beta^2(2\beta-1)^3}{\cos u_1 \cos u_2}, \quad \beta = \frac{1}{1+\rho}.$$

∇' is singular for the following cases

- $\beta = 1 \rightarrow \rho = 0$: This corresponds to *obj2* being a single point. Note that the rank of ∇' is 3 (not 2!). This can also be seen from the multiplicity of the zeros in the determinant.
- $\beta = \frac{1}{2} \rightarrow \rho = 1$: This corresponds to the case when both objects are balls of identical radius. In fact, counting the multiplicity of the zeros at $\beta = \frac{1}{2}$, or computing the rank of ∇', the reachable space has dimension 2! This fact can be interpreted using the notion of holonomy angles (See Section 4).
- $\beta = 0 \rightarrow \rho = \infty$. The result is degenerate because from the previous example we know that a unit ball can reach any contact configuration on the plane by rolling.

Steps 1 through 3 are repeated for the other three coordinate systems and the results are consistent.

Output: *It is true that a unit ball can reach any contact configuration by rolling relative to another ball of radius ρ if and only if ρ is not zero and $\rho \neq 1$.*

Example 3.5 (*The classic example re-visited*). Consider again the classic example of a unit disk on the plane. Note that the two rotations are different here as from Example 3.3. According to [Gre77] and [Gol80], the disk can reach any contact configurations by rolling, but as far as we are aware of no proof has been given in any mechanics textbook. The constrained vector fields from Example 2.4 are

$$X_1 = \begin{bmatrix} -1 \\ -\cos\psi \\ \sin\psi \\ 0 \end{bmatrix} \text{ and } X_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

The Lie algebra of vector fields generated by $\{X_1, X_2\}$ consists of, in addition to $\{X_1, X_2\}$, the following vectors

$$X_3 = [X_1, X_2] = \begin{bmatrix} 0 \\ -\sin\psi \\ -\cos\psi \\ 0 \end{bmatrix},$$

and

$$X_4 = [X_2, X_3] = \begin{bmatrix} 0 \\ -\cos\psi \\ \sin\psi \\ 0 \end{bmatrix}.$$

Note that $[X_1, X_3] = 0$. It is simple to verify that

$$\nabla = \{X_1, X_2, X_3, X_4\}$$

has rank 4, for all points in the parameter space. This shows that any contact configuration is indeed reachable by rolling.

4 The Motion Planning Algorithm

In this section, we solve the following planning problem.

Problem 2' *Suppose that a motion exists between two contact configurations* $(g^0_{r1,r2}, g^f_{r1,r2}) \in M$*. Find a trajectory of* obj1 *that satisfies the contact constraints and links* $g^f_{r1,r2}$ *to* $g^0_{r1,r2}$.

Let $p_0 = (u^0_1, v^0_1, u^0_2, v^0_2, \psi^0)^t$ and $p_f = (u^f_1, v^f_1, u^f_2, v^f_2, \psi^f)^t$ be the point in P that corresponds to the contact configuration $g^0_{r1,r2}$ and $g^f_{r1,r2}$, respectively. Uniqueness of p_0 and p_f are guaranteeed by Proposition 3.1. We may assume without loss of generality that p_0 and p_f belong to the same coordinate system. The objective of the planning problem is to construct a trajectory $(\mathbf{u}_1(t), \mathbf{u}_2(t), \psi(t)) \in P$ (or $(c_1(t), c_2(t), \psi(t)) \in P$ for a coordinate invariant description) that satisfies the rolling constraints and links p_f to p_0. We remark that as a generic property of robot motion planning, the paths are not unique, unless additional constraints such as minimal distance, maximum safety margin ([Kod87], [Can88]) and the grasp condition ([LCS89]) are imposed.

First, let's characterize relations between contact trajectories $c_1(t) \in S_1$ of *obj1* and $c_2(t) \in S_2$ of *obj2*. For this, let TS_1 denote the tangent bundle of S_1, TS_2 the tangent bundle of S_2 and TS^1 the tangent bundle of S^1. We claim that there exist bundle maps

$$\varphi_1 : TS_1 \longrightarrow TS_2 \times TS^1, \tag{22}$$

and

$$\varphi_2 : TS_2 \longrightarrow TS_1 \times TS^1 \tag{23}$$

that takes the contact trajectory of one object to the contact trajectory of the other and the trajectory of the contact angle. The coordinate expression of the map ψ_2 is computed as follows. Let $c_2(t), t \in I$, be a trajectory of contact for *obj2*, or equivalently, $(\mathbf{u}_2, \dot{\mathbf{u}}_2) \in TS_2$. By the second equation of (14), the components of rolling velocity can be expressed in terms of $\dot{\mathbf{u}}_2$ as

$$\begin{bmatrix} -w_y \\ w_x \end{bmatrix} = (K_1 + \tilde{K}_2)\tilde{R}_\psi M_2 \dot{\mathbf{u}}_2, \quad \tilde{R}_\psi^{-1} = \tilde{R}_\psi. \tag{24}$$

Substitute this into the remaining equations of (14), yield

$$\begin{bmatrix} \dot{\mathrm{u}}_1 \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} M_1^{-1}\tilde{R}_\psi \\ T_1\tilde{R}_\psi + T_2 \end{bmatrix} M_2\dot{\mathrm{u}}_2. \tag{25}$$

Equation (25), the coordinate expression of the map ψ_2, defines the evolution of the contact coordinates of *obj1* and the contact angle. Let $\mathrm{u}_1(t), t \in I$, be the solution of (25) and $c_1(t) = f_1^{-1}(\mathrm{u}_1(t)) \in S_1, t \in I$, for (f_1, U_1) a coordinate system of S_1. $c_1(t)$ is called *the contact trajectory of obj1 induced by the contact trajectory $c_2(t)$ of obj2.* Conversely, interchange the role between *obj2* and *obj1* , we have

$$\begin{bmatrix} \dot{\mathrm{u}}_2 \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} M_2^{-1}\tilde{R}_\psi \\ T_1 + T_2\tilde{R}_\psi \end{bmatrix} M_1\dot{\mathrm{u}}_1. \tag{26}$$

which is the coordinate expression of the map ψ_1.

The angle of contact, ψ, whose evolution is defined by (26) has a useful geometric interpretation when *obj2* is torsion free, i.e., $T_2 = 0$. Let $c_1(t), t \in [t_0, t_1]$, be a piecewise regular simple closed curve in S_1 representing the contact trajectory of *obj1*, and $\delta\psi = \psi(t_1) - \psi(t_0)$ denote the net change of contact angle induced by c_1. We have

Proposition 4.1 *$-\delta\psi$ is equal to the holonomy angle of the loop c_1 (See Appendix A for the definition of holonomy angle). In other words, $-\delta\psi = \iint_R k\,dA$, where k is the Gaussian curvature of S_1 and R is the region bounded by c_1.*

Remark 4.1 This is a key result leading to the development of the motion planning algorithm. In order to realize a desired change of contact angle without altering the point of contact relative to S_1, we may plan a closed curve in S_1 such that the Gaussian curvature integral over the region bounded by the loop is equal to the net angle change.

Proof. By appendix A, we may assume that $c_1(t) \in S_1$ is contained within a geodesic coordinate system. Thus, the metric tensor takes the form

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & q \end{bmatrix}, \text{ i.e., } |f_{u_1}| = 1, \text{ and } |f_{v_1}| = q.$$

and the Christoff symbols are given by

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{q_1}{q},\ \Gamma^1_{22} = -qq_1,\ \Gamma^2_{22} = \frac{q_2}{q}$$

$$\Gamma^1_{11} = \Gamma^1_{22} = \Gamma^1_{12} = \Gamma^2_{11} = 0;\ \text{and}\ \ k = -\frac{q_{11}}{q}.$$

where $q_1 = \frac{\partial q}{\partial u_1}$ and $q_2 = \frac{\partial q}{\partial v_1}$. $\dot{\psi}$ from (26) is given by (when $T_2 = 0$)

$$\dot{\psi} = T_1 M_1 \dot{\mathbf{u}}_1.$$

But,

$$T_1 M_1 = \frac{f^t_{v_1}}{q}\left[\frac{f_{u_1u_1}}{|f_{u_1}|}, \frac{f_{u_1v_1}}{|f_{v_1}|}\right]\begin{bmatrix} 1 & 0 \\ 0 & q \end{bmatrix} = \left[\frac{f_{v_1}\cdot f_{u_1u_1}}{q}, \frac{f_{v_1}\cdot f_{u_1v_1}}{q^2}\right] \tag{27}$$

Using Gauss's formula and from the special forms of the Christoff's symbols, we have

$$f_{u_1u_1} = \Gamma^1_{11} f_{u_1} + \Gamma^2_{11} f_{v_1} + h_{11} n = 0 + 0 + h_{11} n,$$

and

$$f_{u_1v_1} = \Gamma^1_{12} f_{u_1} + \Gamma^2_{12} f_{v_1} + h_{12} n = \frac{q_1}{q} f_{v_1} + h_{12} n.$$

Thus,

$$f_{v_1}\cdot f_{u_1u_1} = 0 \text{ and } f_{v_1}\cdot f_{u_1v_1} = q_1/q.$$

Finally,

$$\dot{\psi} = q_1 \dot{v}_1 \tag{28}$$

which is precisely the expression (differing by a sign to account for the reversed orientation) for the derivative of the holonomy angle (Appendix A). □

Using (25), (26) and Proposition 4.1, we have the following algorithm that generates a desired path for the planning problem. The example of a unit ball on the plane is used for illustration.

Algorithm 4.1 *(The Motion Planning Algorithm)*

Input: *1. Initial and final contact configurations* $p_0 = (u^0_1, v^0_1, u^0_2, v^0_2, \psi^0)$ *and* $p_f = (u^f_1, v^f_1, u^f_2, v^f_2, \psi^f)$.

2. *Geometric data of* obj1 *and* obj2: *curvature forms* (K_1, K_2), *metric tensors* (M_1, M_2) *and torsion forms* (T_1, T_2). *(We assume that* obj2 *is torsion free).*

Output: *A curve that links* p_0 *to* p_f *and satisfies the rolling constraints.*

Step 1: *Construct a curve* $c_2(t) \in S_2, t \in [t_0, t_1]$, *such that*

$$\mathbf{u}_2(t_0) = \begin{bmatrix} u_2^0 \\ v_2^0 \end{bmatrix} \text{ and } \mathbf{u}_2(t_f) = \begin{bmatrix} u_2^f \\ v_2^f \end{bmatrix}. \tag{29}$$

Let $c_1(t) \in S_1$ *and* $\psi(t) \in S^1, t \in [t_0, t_1]$, *denote the induced contact trajectory of* obj1 *and the trajectory of the contact angle, respectively. At* $t = t_1$, *the contact point of* obj1 *and the contact angle reach some intermediate values, denoted by*

$$\begin{bmatrix} \hat{u}_1 \\ \hat{v}_1 \end{bmatrix} = \mathbf{u}_1(t_1) \text{ and } \hat{\psi} = \psi(t_1).$$

Step 2: *Construct a closed curve* $c_2(t) \in S_2, t \in [t_1, t_2]$, *such that the induced contact trajectory of* obj1 *has the property*

$$\mathbf{u}_1(t_1) = \begin{bmatrix} \hat{u}_1 \\ \hat{v}_1 \end{bmatrix} \text{ and } \mathbf{u}_1(t_2) = \begin{bmatrix} u_1^f \\ v_1^f \end{bmatrix}.$$

Let $\psi(t) \in S^1, t \in [t_1, t_2]$, *denote the induced trajectory of the contact angle. At* $t = t_2$, *the angle of contact reaches some intermediate value denoted by*

$$\tilde{\psi} = \psi(t_2), \text{ where } \psi(t_1) = \hat{\psi}.$$

Step 3: *Let* $\delta\psi = \psi^f - \tilde{\psi}$ *be the desired holonomy angle. Construct a closed curve* $c_1(t) \in S_1, t \in [t_2, t_f]$, *such that (1) the induced trajectory* $c_2(t) \in S_2, t \in [t_2, t_f]$, *is also* <u>*closed*</u> *and (2) the Gaussian curvature integral over the region bounded by* c_1 *is equal to the desired holonomy angle.*

Output: *Return the curve* $(\mathbf{u}_1(t), \mathbf{u}_2(t), \psi(t)) \in P$, $t \in [t_0, t_1,] \cup [t_1, t_2] \cup [t_2, t_f]$, *which is the union of the curves constructed in Step 1, 2 and 3.*

Remark 4.2 The desired contact point $\begin{bmatrix} u_2^f \\ v_2^f \end{bmatrix}$ of *obj2* is achieved in Step 1. Then, using a closed curve relative to *obj 2* in Step 2 the desired contact point $\begin{bmatrix} u_1^f \\ v_1^f \end{bmatrix}$ of

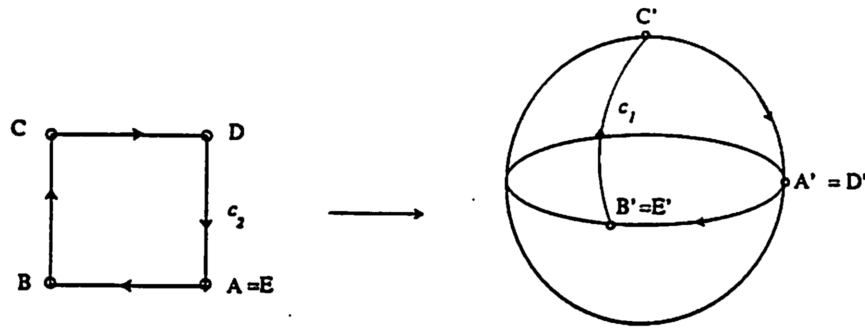

Figure 4: A Lie bracket motion

obj1 is realized without sacrificing the desired contact point of *obj 2*. Finally in Step 3, using a closed curve relative to *obj1*, which also induces a closed curve relative to *obj2*, the desired holonomy angle is realized.

We now use the example of a unit ball on the plane to illustrate the algorithm. Clearly, Step 1 can be easily done using existing techniques in robot motion planning ([Kod87], [Can88]). Step 2 and Step 3 are carried out as follows.

Step 2A: Without loss of generality, we may assume that the initial anf final points of contact in the unit sphere S^2 are within the same coordinate system. Let $\hat{\mathbf{u}}_1 = \begin{bmatrix} \hat{u}_1 \\ \hat{v}_1 \end{bmatrix}$ and $\mathbf{u}_1^f = \begin{bmatrix} u_1^f \\ v_1^f \end{bmatrix}$ be the coordinates of these contact points. We wish to construct a *closed curve* $c_2(t), t \in [t_1, t_2]$, in the plane so that the induced contact trajectory $c_1(t), t \in [t_1, t_2]$, of S^2 links $\hat{\mathbf{u}}_1$ to $\mathbf{u}_1^f$, i.e., $f^{-1}(c_1(t_1)) = \hat{\mathbf{u}}_1$ and $f^{-1}(c_1(t_2)) = \mathbf{u}_1^f$.

Lemma 4.1 *Let* $c_1^0 = f(\hat{\mathbf{u}}_1)$ *and* $c_1^f = f(\mathbf{u}_1^f)$ *be exactly* $\pi/2$ *distance apart in the unit sphere* S^2. *Then, the square of side length* $\pi/2$, *shown in Figure 4 will induce a contact trajectory* c_1 *which links* c_1^0 *to* c_1^f.

Proof. We need to demonstrate that the square has the desired features. Label the point c_1^0 and c_1^f in the sphere by A' and B', respectively, as shown in the figure. $d(A', B') = \pi/2$. There exists a unique geodesic, i.e., an arc of the great circle, that connects A' to B'. The great circle will be called the equator. Let A denote the initial point of contact in the plane. Thus, tracing the geodesic from A' to B'

induces a straight line in the plane with end point B, and $d(B, A) = \pi/2$ (by arc length constraint). Going from the point B to the point C in the plane is equivalent to going from the point B' to the north pole, C', in the sphere. Note that $\measuredangle(ABC)$ and $\measuredangle(A'B'C')$ are both right angles. Now, tracing the straight line from C to D in the plane induces a curve in the sphere which ends at the starting point A'. Consequently, by closing the curve in the plane with a straight line joining D to A, we have arrived at the point B' in the sphere. This shows that the square indeed induces a curve in the sphere which has a net incremental distance $\pi/2$. This is called a Lie bracket motion. □

We now return to the more general case.

Step 2B: By Lemma 4.1, we may assume that $d(c_1^0, c_1^f) < \pi/2$. Otherwise, Lemma 4.1 can be applied repeatedly until some intermediate point which is less than $\pi/2$ distance away from c_1^f is reached. Let $l = d(c_1^0, c_1^f) < \pi/2$ be the distance of these two points. We wish to construct a closed curve $c_2(t), t \in [t_1, t_2]$, in the plane such that the induced contact trajectory $c_1(t), t \in [t_1, t_2]$, has an incremental distance l along the geodesic connecting c_1^0 to c_1^f.

We propose to use for c_2 the closed curve $ABCDE$ shown in Figure 5, where $x = d(A, B)$ is to be determined, $d(B, C) = d(C, D) = \pi/2$, and

$$\theta = 2\tan^{-1}\frac{x}{\pi/2}.$$

We like to show that for some choice of x, the closed curve $ABCDE$ will induce a curve $c_1(t), t \in [t_1, t_2]$, in the sphere that links c_1^0 to c_1^f. First, by tracing the straight line from A to B and then to C induces a curve in the sphere which starts at A', passes through B' and then comes to the north pole, C'. Note that $d(B', A') = x$ and $\measuredangle(A'B'C') = 90^o$. Going down from C to D with an angle θ and by a distance $\pi/2$ is equivalent to going down in the sphere from C' to some point D' at the equator. Clearly, $d(B', D') = \theta$. Now, Connect D to A by a straight line, and we claim that *(1)* $\measuredangle CDA = 90^o$ *and (2)* $d(A, D) = x$. To see this, note that by definition $\measuredangle ACD = \theta/2$ and AC is common to both the triangles $\triangle ABC$ and $\triangle ACD$. Thus, they must

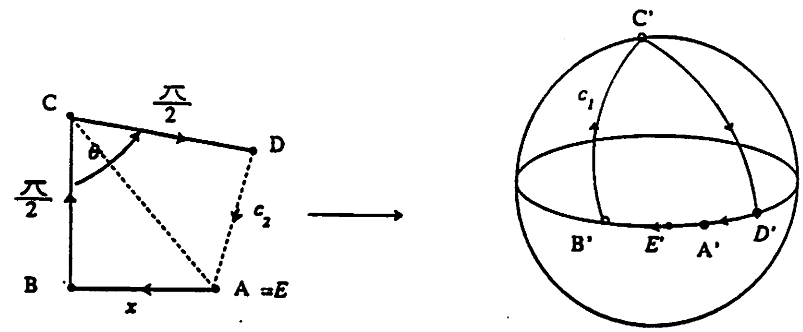

Figure 5: A (general) Lie bracket motion

be congruent triangles and the claim follows.

Thus, by tracing the straight line from D back to A in the plane, we have followed the equator from D' to some point E', and $d(E', D') = x$. With c_2 being the closed curve $ABCDE$ for some choice of x, the induced curve c_1 in the sphere has its starting point A' and its ending point E', where $d(E', A')$, the net incremental distance, is a function of x. Let $f(x) = d(E', A')$. It is not hard to see that

$$f(x) = 2x - \theta = 2x - 2\tan^{-1}\frac{x}{\pi/2}.$$

The hope is to find an x, if possible, that solves the equation

$$f(x) \stackrel{?}{=} l. \tag{30}$$

We claim that there exists a unique x that solves (30). To show this, note that $f(0) = 0$ and $f(\pi/2) = \pi/2 > l$. Thus, solutions exist. For the uniqueness part, we compute the derivative of $f(x)$, which is given by

$$f'(x) = 2 - 2\frac{2/\pi}{1+\frac{4x^2}{\pi^2}} = \frac{2 - 2/\pi + 4x^2/\pi^2}{1 + 4x^2/\pi^2} > 0.$$

Thus, $f(x)$ is a monotone function and the solution to (30), denoted by $x*$, is unique!

Consequently, the curve $ABCDE$, with $d(B, A) = x^*$, has all the desired features.

Step 3': We wish to construct a *closed curve* $c_1(t), t \in [t_2, t_f]$, in S^2 such that *(1) the induced curve* $c_2(t), t \in [t_2, t_f]$, *in the plane is also closed and (2)* c_1 *has a desired holonomy angle* $\delta\psi$.

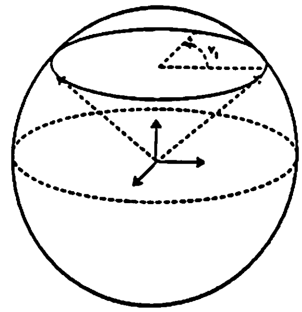

Figure 6: Another Lie bracket motion

We may assume that $0 < -\delta\psi < 2\pi$. Consider the latitude circle with $u_1(t) = u_1(0)$,and $v_1(t) = v_1(0) + t$, $t \in [t_2, t_2 + 2\pi]$. We claim that *(1) the induced trajectory c_2 is also a circle and (2) the holonomy angle of c_1 ranges from 0 to 2π for $0 < u_1(0) < \pi/2$* . To see this, substitute the expression of $\begin{bmatrix} u_1(t) \\ v_1(t) \end{bmatrix}$ into (26) and after some algebra, we get

$$\psi(t) - \psi(0) = -\sin u_1(0)t \triangleq \alpha t, \quad \alpha = -\sin u_1(0),$$

and

$$u_2(t) = \beta\cos(\alpha t + \psi_0) + \gamma_0, \quad \gamma_0 = u_2(0) - \cos\psi_0 \cos u_1(0)/\alpha,$$

$$v_2(t) = -\beta\sin(\alpha t + \psi_0) + \delta_0, \quad \delta_0 = v_2(0) + \sin\psi_0 \cos u_1(0)/\alpha.$$

Thus, we have

$$(u_2(t) - \gamma_0)^2 + (v_2(t) - \delta_0)^2 = \beta^2.$$

This shows the claim.

5 Conclusion

In this paper, we have studied robot motion planning with nonholonomic constraints. First, using the kinematic equations of contact developed by Montana, we have transformed contact constraints in the configuration manifold to a system of differential equations in the parameter space. To verify the existence of motion between two contact configurations, we have established an algorithm that compute the smallest involutive distribution generated by the constrained vector fields. If the distribution has full rank, then any two contact configurations can be reached from each other by rolling. Otherwise, a motion exists if and only if the two contact configurations belong to the same maximum integral manifold of the distribution.

As we have shown by examples that, it is precisely due to the nonholonomicity of the rolling constraints that an object can reach an arbitrary contact configuration by rolling. We conjecture that this is a generic property of any two objects. We have also given an example, namely two balls of equal radius, where this property fails.

For many interesting applications, one object is approximately torsion free. In these cases, we have also given an algorithm that generates a desired path when it exists. The general case is currently under investigation with the framework developed in this paper.

A more difficult problem which we are also very interested in is, *finding a curve in S_2 of shortest distance such that the desired contact point of* obj1 *and the desired contact angle can be realized.* This problem has a close relation with another class of problems recently studied by others ([Mon88]).

Acknowledgement

The author would like to thank professors Shankar Sastry, John Canny and JiangHua Lu for the support, encouragement and valuable discussions while this research was conducted, John Hauser for the help in using Macsyma to generate the ∇ for example 3.4.

References

[Can88] J.F. Canny. *The complexity of robot motion planning.* MIT Press, 1988.

[CHS88] A. Cole, J. Hauser, and S. Sastry. Kinematics and control of a multifingered robot hand with rolling contact. In *IEEE International Conference on Robotics and Automation*, 1988.

[CR87] C. Cai and B. Roth. On the spatial motion of rigid bodies with point contact. In *IEEE International Conference on Robotics and Automation*, pages 686–695, 1987.

[Gol80] H. Goldstein. *Classic Mechanics.* Addison-Wesley, 2nd edition edition, 1980.

[Gre77] D. Greenwood. *Classical Dynamics.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1977.

[HK77] R. Hermann and A.J. Krener. Nonlinear controllability and observability. *IEEE Transaction on Automatic Control*, 22:728–740, 1977.

[Isi85] A. Isidori. *Nonlinear control systems: An introduction. Lecture notes in control and information sciences*, Springer-Verlag, 1985.

[Ker85] J. Kerr. *An analysis of multifingered hand.* PhD thesis, Department of Mechanical Engineering, Stanford University, 1985.

[Kli78] W. Klingenberg. *A course in differential geometry.* Springer-Verlag, 1978.

[Kod87] D.E. Koditschek. Exact robot navigation by means of potential functions: some topological considerations. In *IEEE International Conference on Robotics and Automation*, pages 1–6, 1987.

[LCS89] Z.X. Li, J. F. Canny, and S.S. Sastry. On motion planning for dextrous manipulation. In *IEEE International conference on robotics and automation (submitted)*, 1989.

[LHS88] Z.X. Li, P. Hsu, and S. Sastry. On grasping and coordinated manipulation by a multifingered robot hand. *International Journal of Robotics Research (to appear)*, 1988.

[Mon86] D. Montana. *Tactile sensing and kinematics of contact.* PhD thesis, Division of Applied Sciences, Harvard University, 1986.

[Mon88] R. Montgomery. *Shortest loops with a fixed holonomy.* Technical Report, Mathematical Science Research Institute, University of California, Berkeley, 1988.

[MP78] R.S. Millman and G.D. Parker. *Elements of differential geometry.* Prentice-Hall, Inc., Englewood Cliffs, N.J. 07632, 1978.

[Spi74] M. Spivak. *A Comprehensive Introduction to Differential Geometry.* Volume I and II, Publish or Perish, Boston, Mass., 1974.

Appendix A

In this appendix, we provide a short introduction to some local and global surface theory. The concepts to be discussed here include the first, second fundamental forms in a more general setting, the Christoff symbols, covariant differentiations and geodesics, holonomy angles and the Gauss-Bonnet Formula for computing the holonomy angle. An in-depth treatment of these subjects can be found in ([MP78], [Kli78] and [Spi74]).

Notation 5.1 As in Section 2, U will always denote an open subset of $\mathbf{R}^2$. A point of U will be denoted by $\mathrm{u} \in \mathbf{R}^2$, or by $(u_1, u_2) \in \mathbf{R} \times \mathbf{R}$, or $(u, v) \in \mathbf{R} \times \mathbf{R}$. Let $f : U \longrightarrow \mathbf{R}^3$ be a differentiable map, $df_u : T_u\mathbf{R}^2 \longrightarrow T_{f(u)}\mathbf{R}^3$ denotes the tangent map of f, and f_u, f_v denote the partial derivatives of f with respect to u and v, respectively. Also, f_{ij} denotes $\frac{\partial^2 f}{\partial u_i \partial u_j}, i, j = 1, 2$.

Notation 5.2 Let S denote an embedded surface, with a coordinate system (f, U). Let $n : S \longrightarrow S^2$ denote the Gauss map of S. In a coordinate system (f, U) we will also use n to denote the map $n o f : U \longrightarrow S^2$ and $dn_u : T_u\mathbf{R}^2 \longrightarrow T_{f(u)}\mathbf{R}^3$ the corresponding tangent map. The first fundamental form I and the second fundamental form II are denoted by

$$I_u = \begin{bmatrix} f_1 \cdot f_1 & f_1 \cdot f_2 \\ f_2 \cdot f_1 & f_2 \cdot f_2 \end{bmatrix} \triangleq \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix},$$

and

$$II_u = -dn_u \cdot df_u = -\begin{bmatrix} n_u \cdot f_u & n_u \cdot f_v \\ n_v \cdot f_u & n_v \cdot f_v \end{bmatrix} \triangleq \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}.$$

Remark 5.1

1. When (f, U) is orthogonal, the metric M defined in Section 2 is the square root of the first fundamental form.
2. When (f, U) is orthogonal, the curvature form is given by $K = M^{-t} II M^{-1}$, i.e., a change of basis.

The Gaussian curvature k is given by $k = \frac{\det(II_u)}{\det(I_u)}$ and the Christoff symbols of the second kind are defined by

$$\Gamma_{ij}^k = \frac{1}{2}\sum_{l=1}^{2} g^{lk}\left(\frac{\partial g_{il}}{\partial u_j} - \frac{\partial g_{ij}}{\partial u_l} + \frac{\partial g_{lj}}{\partial u_i}\right), \tag{31}$$

where g^{lk} is the $l-k$ entry of the inverse matrix of I_u. The Gauss's formula expresses f_{ij} in terms of the basis (f_1, f_2, n)

$$f_{ij} = \sum_k \Gamma^k_{ij} f_k + h_{ij} n, \quad i,j = 1,2. \tag{32}$$

Definition 5.1 Let X be a tangent vector field along a curve $c(t)$ in S, which in local coordinates can be written as $X(t) = \sum_i x_i(t) f_i \circ \mathbf{u}(t)$. The covariant derivative of X along c is defined by

$$\frac{\nabla X}{dt}(t) = \sum_k \left(\dot{x}_k + \sum_{i,j} x_i \dot{u}_j \Gamma^k_{ij} \right) f_k \circ \mathbf{u}(t) \tag{33}$$

and X is called parallel along c if $\frac{\nabla X}{dt}(t) = 0$

A curve $c(t)$ on S is called a geodesic if and only if $\frac{\nabla \dot{c}(t)}{dt} = 0$. By (33) this is equivalent to the second derivative of $c(t)$ being normal to the surface. Thus, all geodesics to the unit sphere S^2 are arcs of the great circles.

Proposition 5.1 *Let $c(t)$ be a curve on a surface S. Let $\hat{X}$ be a tangent vector to S at $c(t_0)$. Then, there exists a unique vector field $X(t)$ that is parallel along $c(t)$ with $X(t_0) = \hat{X}$.*

Proof. Consider the following initial value problem

$$\tfrac{dx_k}{dt} = -\textstyle\sum_{i,j} \Gamma^k_{ij} x_i \dot{u}_j,$$

$$x_k(t_0) = \hat{x}_k, \; k = 1,2.$$

By Picard's theorem (i.e., existence and uniqueness of O.D.E), this has a unique solution for values of t near t_0. The solution clearly is parallel along $c(t)$, where it is defined. Repeated application of this gives a unique X defined along all of $c(t)$. □

The unique vector field $X(t)$ parallel along $c(t)$ such that $X(t) = \hat{X}$ is called the parallel translate of $\hat{X}$ along $c(t)$. Note that if two vector fields $X(t)$ and $Y(t)$ are both parallel along a curve c in S. Then, $|X(t)|$ is constant and so is the angle between $X(t)$ and $Y(t)$. To see this, let $g(t) = \langle X(t), Y(t) \rangle$. $dg/dt = \langle dX/dt, Y \rangle + \langle X, dY/dt \rangle = 0 + 0 = 0$, so g is constant. If $Y = X$, this implies $|X|$

is constant. The cosine of the angle between X and Y is $g(t)/|X||Y|$, which is then constant and so is the angle.

We now introduce the notion of holonomy angle.

Let c be a piecewise regular simple closed curve in a surface S with period L and is contained within a simply connected geodesic coordinate system, bounding a region R. Let X be a unit vector field parallel along c. (Such an X exists by Proposition 5.1.) In general, we have $X(0) \neq X(L)$. We shall be interested in the angle between $X(0)$ and $X(L)$, which we denote it by $\delta\theta = \measuredangle(X(0), X(L))$. $\delta\theta$ is called the holonomy angle of c. Since an orientation of S is given by the Gauss map, the angle between $X(0)$ and $X(L)$ is well defined (i.e., counter clockwise rotation about the normal).

Theorem 5.1 *(Gauss-Bonnet Formula) Let $c(t)$ be the curve on S as described above. Let k_g be the geodesic curvature of c and $\alpha_1, ...\alpha_n$ be the jump angles at the junctions. Then,*

1.

$$\delta\theta = \measuredangle(X(0), X(L)) = \iint_R k dA, \tag{34}$$

where k is the Gaussian curvature of c and

2.

$$\iint_R k dA + \int_c k_g ds + \sum \alpha_i = 2\pi. \tag{35}$$

Remark 5.2 1. Let c be a unit speed curve. The geodesic curvature k_g of c is defined by the following formula

$$k_g(n \times \dot{c}(t)) = \sum_k \left(\ddot{u}_k + \sum_{i,j} \Gamma_{ij}^k \dot{u}_i \dot{u}_j \right) f_k. \tag{36}$$

Since (f_1, f_2) are linearly independent, k_g is uniquely defined by (36). A curve c is a geodesic if and only if $k_g = 0$.

2. $f : U \longrightarrow \mathbf{R}^3$ is a geodesic coordinate system, if the first fundamental form takes the form

$$I_u = \begin{bmatrix} 1 & 0 \\ 0 & g^2 \end{bmatrix}, \tag{37}$$

for some $q > 0$, which is a function of (u_1, u_2). Every surface can be covered by geodesic coordinate systems.

Proof (of Theorem 5.1). We shall only prove (1). The proof of (2) can be found in any standard text book on differential geometry (e.g. [MP78], [Kli78] and [Spi74]).

In a geodesic coordinate system, the first fundamental form takes the form of (37). By (31), the Christoff symbols are

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{q_1}{q},\ \Gamma^1_{22} = -qq_1,\ \Gamma^2_{22} = \frac{q_2}{q}$$

$$\Gamma^1_{11} = \Gamma^1_{22} = \Gamma^1_{12} = \Gamma^2_{11} = 0, \text{ and } k = -\frac{q_{11}}{q},$$

where $q_i = \partial q/\partial u_i$.

Since all the quantities in the conclusion are independent of parametrization, we may assume that c is parametrized by arc length. Let $\theta = \measuredangle(f_1, X)$. Note that $\cos\theta = \langle f_1, X\rangle$ so that $-(\sin\theta)\dot{\theta} = \langle \dot{f}_1, X\rangle + \langle f_1, \dot{X}\rangle = \langle \dot{f}_1, X\rangle$ since X is parallel along c. However,

$$\dot{f}_1 = \dot{u}_1 f_{11} + \dot{u}_2 f_{12}$$

so that Gauss's formula (32) yield

$$\begin{aligned} -(\sin\theta)\dot{\theta} &= \langle \dot{u}_1 f_{11} + \dot{u}_2 f_{12}, X\rangle \\ &= \langle (\dot{u}_1\Gamma^1_{11} + \dot{u}_2\Gamma^1_{12}) f_1 + (\dot{u}_1\Gamma^2_{11} + \dot{u}_2\Gamma^2_{12}) f_2,\ X\rangle. \end{aligned}$$

The specific form of Γ^k_{ij} allows us to conclude

$$-(\sin\theta)\dot{\theta} = \dot{u}_2 \frac{q_1 \langle f_2, X\rangle}{q}. \tag{38}$$

Because $g_{12} = 0, \{f_1, f_2/|f_2|\}$ is an orthonormal basis of $T_{f(u)}S$, for each $u \in U$. Hence,

$$X = \langle X, f_1\rangle f_1 + \frac{\langle X, f_2/|f_2|\rangle}{|f_2|} f_2.$$

Since X is a unit vector and $\langle X, f_1\rangle = \cos\theta$, we have $\langle X, f_2\rangle = \sin\theta$. Hence (38) becomes $\dot{\theta} = -q_1\dot{u}_2$ so that

$$\delta\theta = -\int_c q_1 \dot{u}_2 ds = -\int_c q_1 du_2 \tag{39}$$

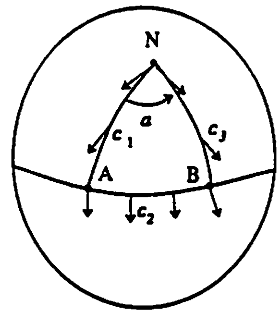

Figure 7: A geodesic triangle in S^2

By Green's Theorem,

$$-\int_c q_1 du_2 = -\iint_R q_{11} du_1 du_2 = -\iint_R \frac{q_{11}}{q} q du_1 du_2 = \iint_R k dA, \tag{40}$$

where $\iint_R dA = \iint_R q du_1 du_2$ is the area integral.

On the other hand, we can write

$$\begin{aligned} \measuredangle(X(0), X(L)) &= \measuredangle(f_1 \circ u(L), X(L)) - \measuredangle(f_1 \circ u(0), X(0)) \\ &= \theta(L) - \theta(0) = \int_c \frac{d\theta}{ds} ds = \delta\theta. \end{aligned} \tag{41}$$

□

Example 5.1 Let A, B be two points on the equator of the unit sphere S^2, separated by a distance $\alpha < \pi$, and N the north pole (Figure 7) . Clearly, the three points can be connected by unique geodesics (i.e., arcs of the great circles) to form a triangle. This is called a geodesic triangle. The three edges are labeled by c_1, c_2 and c_3. Consider the closed curve c which consists of the union of c_1, c_2 and c_3. Let $X(0)$ be the unit tangent vector to c_1 at $c_1(0)$, as shown in the figure, and $X(t)$ the vector field which is the parallel translate of $X(0)$. The holonomy angle $\delta\theta = \measuredangle(X(0), X(L))$ can be computed using either (34) or (35).

By (34), $k = 1$ for the unit sphere. Thus,

$$\delta\theta = \iint_R k dA = \text{area of } R = \frac{1}{2} \cdot 4\pi \cdot \frac{\alpha}{2\pi} = \alpha.$$

On the other hand, by (35) $k_g = 0$ since c is a geodesic triangle and the jump angles are $\alpha_1 = \pi/2, \alpha_2 = \pi/2$, and $\alpha_3 = \pi - \alpha$. Therefore,

$$\delta\theta = 2\pi - \in k_g ds - \sum \alpha_i = \alpha.$$

Remark 5.3 Given two points A and B in M, if B is not a conjugate point of A, then there exists a unique geodesic connecting B to A. In other words, there exists no Jacobi field along the geodesic connecting B to A ([Kli78]).